Partial Relaxation: A Computationally Efficient Direction-of-Arrival Estimation Framework

Dem Fachbereich 18
Elektrotechnik und Informationstechnik
der Technischen Universität Darmstadt
zur Erlangung der Würde eines
Doktor-Ingenieurs (Dr.-Ing.)
vorgelegte Dissertation

von
M.Sc. Minh Hoang Trinh
geboren am 11.11.1991 in Hanoi, Vietnam

Referent:	Prof. Dr.-Ing. Marius Pesavento
Korreferent:	Prof. Dr. Mats Viberg
Tag der Einreichung:	16.03.2020
Tag der mündlichen Prüfung:	30.04.2020

D 17
Darmstädter Dissertation
Darmstadt, 2020

Trinh Hoang Minh: *Partial Relaxation: A Computationally Efficient Direction-of-Arrival Estimation Framework*

Darmstadt, Technische Universität Darmstadt

Jahr der Veröffentlichung der Dissertation auf TUprints: 2020

Tag der mündlichen Prüfung: 30.04.2020

Erklärungen laut Promotionsordnungen

§ 8 Abs. 1 lit. c PromO
Ich versichere hiermit, dass die elektronische Version meiner Dissertation mit der schriftlichen Version übereinstimmt.

§ 8 Abs. 1 lit. d PromO
Ich versichere hiermit, dass zu einem vorherigen Zeitpunkt noch keine Promotion versucht wurde. In diesem Fall sind nähere Angaben über Zeitpunkt, Hochschule, Dissertationsthema und Ergebnis dieses Versuchs mitzuteilen.

§ 9 Abs. 1 PromO
Ich versichere hiermit, dass die vorliegende Dissertation selbstständig und nur unter Verwendung der angegebenen Quellen verfasst wurde.

§ 9 Abs. 2 PromO
Die Arbeit hat bisher noch nicht zu Prüfungszwecken gedient.

Datum und Unterschrift

Acknowledgments

First and foremost, I would like to express my deepest appreciation to my supervisor Prof. Dr.-Ing. Marius Pesavento for his guidance during my doctoral studies. His constant support, suggestion and encouragement made this work possible.

I am also grateful to my co-supervisor Prof. Dr. Mats Viberg for giving me invaluable insight into this research topic. Our fruitful discussions immensely influenced this thesis and made my doctoral studies a remarkable experience.

I would like to thank Prof. Dr. mont. Mario Kupnik, Prof. Dr.-Ing. Klaus Hofmann and Prof. Dr.-Ing. Gerd Griepentrog for their work as committee members of my doctoral defense.

For valuable discussions and support, I would like to thank my current and former colleagues in the Communication Systems Group: Florian Bahlke, Yufan Fan, Ganapati Hegde, Gerta Kushe, Tianyi Liu, Fabio Nikolay, Pouyan Parvazi, Oscar Dario Ramos Cantor, David Schenck, Christian Steffens, Wassim Suleiman, Dima Taleb and Yang Yang. Their useful discussions and advice greatly improved my technical skills as a researcher. Special thanks go to Marlis Gorecki for her help in administrative and teaching-related issues.

Finally, I would like to thank my parents and my brother for their love, patience and encouragement.

Kurzfassung

Einfallsrichtungsschätzung (englisch: Direction-of-Arrival, kurz DOA, estimation) aus rauschbehafteten Sensordaten ist ein fundamentales Forschungsgebiet in der Sensordatenverarbeitung. Die Anwendung der DOA-Schätzverfahren ist nicht nur auf Radarsysteme beschränkt, sondern umfasst zusätzlich verschiedene Forschungsfelder, z.B. die Radioastronomie, biomedizinische Bildgebung, Reflexionsseismik, drahtlose Kommunikation und vieles mehr.

Aufgrund des breiten Anwendungsbereiches der Einfallsrichtungsschätzung wurden unterschiedliche DOA-Schätzverfahren erforscht, um das Auflösungsvermögen, die Recheneffizienz sowie die Robustheit der Schätzverfahren zu erhöhen. Ein Kompromiss zwischen der Schätzgüte und dem Rechenaufwand ist jedoch im Allgemeinen unvermeidlich. Unter dieser Prämisse werden in der vorliegenden Arbeit neue recheneffiziente Schätzverfahren entwickelt und untersucht. Das Ziel der neuen Schätzverfahren ist es, ein hohes Auflösungsvermögen im Schwellenbereich, d.h. bei einer geringen Anzahl zeitlicher Messungen oder geringem Signal-zu-Rausch-Verhältnis, zu erreichen.

Auf Grundlage von konventionellen DOA-Schätzverfahren aus der Literatur und den zugrundeliegenden Signalmodellen führen wir in dieser Dissertation das Konzept der partiellen Relaxierung (englisch: Partial Relaxation) ein. Im Partiellen Relaxierungskonzept werden die DOA Parameter durch die Eigenwertberechnung einer modifizierten Kovarianzmatrix in jedem Testwinkel des Sichtfeldes geschätzt. Numerische Simulationen zeigen, dass die entwickelten DOA-Schätzverfahren, unabhängig von bestimmten Strukturen der Sensorgruppe, eine sehr gute Schätzgüte im Schwellenbereich erzielen.

In der vorliegenden Arbeit werden sowohl theoretische als auch praktische Aspekte der partiell relaxierten DOA-Schätzer untersucht. Unter anderem wird für den praktischen Einsatz eine generelle und effiziente Methode zur Berechnung der relevanten Eigenwerte vorgestellt, die für alle entwickelten partiell relaxierten DOA-Schätzer von großer Bedeutung ist. Zur Beschleunigung der Eigenwertberechnungen nutzt die eingeführte Methode die gemeinsame Struktur in den Kostenfunktionen der Schätzer aus. Dadurch reduziert sich der Rechenaufwand der Schätzer mit der vorgeschlagenen Methode im Vergleich zu einer direkten Implementierung signifikant. Des Weiteren wird ein geschlossener Ausdruck für die theoretische untere Grenze der Varianz des Schätzfehlers für *alle* erwartungstreuen partiell relaxierten DOA-Schätzern hergeleitet. Darüber hinaus wird die ermittelte untere Grenze der partiell relaxierten Schätzer mit den entsprechenden theoretischen Ergebnissen in der Literatur verglichen und die sich daraus ergebenenen Schlussfolgerungen diskutiert.

Abstract

Direction-of-Arrival (DOA) estimation from data collected at a sensor array in the presence of noise has been a fundamental and long-established research topic of interest in sensor array processing. The application of DOA estimation does not only restrict to radar but also spans multiple additional fields of research, including radio astronomy, biomedical imaging, seismic exploration, wireless communication, among others.

Due to the wide applications of DOA estimation, various methods have been developed in the literature to increase the resolution capability, computational efficiency, and robustness of the algorithms. However, a trade-off between the estimation performance and the computational complexity is generally inevitable. This thesis addresses the challenge of developing low-complexity DOA estimators with the ability to resolve closely spaced source signals in the threshold region, i.e., low sample size or low Signal-to-Noise ratio.

Motivated by various interpretations of the conventional DOA estimators in the literature and their implied signal models, in this thesis, we introduce a novel class of DOA estimators which is referred to as the *Partial Relaxation* framework. In the Partial Relaxation framework, the DOA parameters are estimated from the eigenvalues of a particular modified covariance matrix at each look-direction in the Field-of-View. Simulations show that the proposed DOA estimators achieve excellent performance in the threshold region without exploiting any particular structure of the sensor array.

Theoretical and practical aspects of the DOA estimators under the proposed framework are investigated in this thesis. From the practical perspective, as the computation of selected eigenvalues is crucial for the proposed estimators, we introduce a general and efficient implementation which exploits the underlying structure of the matrix argument induced by the Partial Relaxation framework. Compared with the naive implementation, the execution time of the proposed estimators with the efficient implementation is reduced by multiple orders of magnitude. From the theoretical aspect, a closed-form expression of the lower bound for the estimation error of *all* unbiased estimators under the Partial Relaxation framework is derived. Consequently, implications and comparisons of the proposed lower bound with theoretical results in the literature are studied and discussed.

Contents

Chapter 1
Introduction

The history of direction finding in general and DOA estimation in particular is traced back to the beginning of the 20th century when the first passive radio frequency-based direction finder was patented [Sto02]. Thereafter, DOA estimation from data collected at a sensor array has become a classical and fundamental research topic in the field of sensor array signal processing. Furthermore, in other research fields, many important parameter estimation problems, for example, harmonic retrieval [Chi78, Kes86], time delay estimation [Car87,Kay93], sparse regression [CRT06,Bar07], to name a few, share a common underlying mathematical model with the DOA estimation problem. Therefore, algorithms developed for the task of DOA estimation are generally applicable to other research fields without considerable modifications [SJGL14], and developing novel DOA estimation algorithms has been a topic of interest. Nowadays, the application of the DOA estimation problem has spanned over multiple research areas, including wireless communication, radio astronomy, biomedical imaging, automotive radar, sonar, among others [RAKS09, VT04, KV96, ZB11].

Motivated by the wide application of DOA estimation, numerous methods have been developed to increase the resolution capability, computational efficiency and robustness of the DOA estimation. By exploiting the full statistical information of the signals impinging on the sensor array, Maximum Likelihood (ML) estimators formulate the DOA estimation problem as an optimization problem for the DOA parameters. It has been observed from numerical experiments as well as confirmed by comprehensive theoretical investigations that the error performance of ML estimators is excellent in the asymptotic region [SN89, SS90, SN90b]. In addition, ML estimators outperform subspace-based DOA estimators in the non-asymptotic region [FLB04, Ath05, AJM07, JAM08], e.g., low number of snapshots or low Signal-to-Noise Ratio (SNR). Such remarkable error performance of ML estimators in the non-asymptotic region is of great importance in practical applications, where the number of snapshots and the SNR are both limited due to non-stationary propagation environments. Nevertheless, the application of ML estimators in real-time scenarios is generally impractical as they require the optimization of multi-modal functions and the computational cost is prohibitive.

In order to avoid the high computational complexity of the ML estimators in estimating the DOAs of the source signals, various estimators have been proposed in the literature which provide different trade-offs between the error performance and the computational complexity. Among the spectral-based DOA algorithms, the conventional

beamformer [BCD+84] computes the spectrum efficiently by match-filtering the receive signal with the array response of all candidate directions in the Field-of-View. However, regardless of the number of snapshots or SNR, the conventional beamformer is subject to limited resolution capability indicated by the classical Rayleigh limit. In the family of subspace-based algorithms, Multiple Signal Classification [Sch86], abbreviated as MUSIC, relies on the signal subspace calculated from the spatial sample covariance matrix and performs a spectral search for the estimated DOAs. Unlike the conventional beamformer, the MUSIC estimator [Sch86] does not suffer from the Rayleigh resolution limit provided that the number of snapshots or the SNR is sufficient. In contrast to MUSIC in which the DOAs are estimated from a spectral search, root-MUSIC [Bar83], ESPRIT [RK89] and their variants [PGH00], [QHSS16], [HN95], [RG09], [PPG12] provide search-free DOA estimates by exploiting uniform linear and shift-invariant array structures, respectively. The exploitation of the sensor array structures results in a considerable reduction in the computational time and enhancement in the estimation performance [RH89,KFP92,SN91,HPRK14]. However, without exploitation of specific structure of the sensor array, the trade-off between the computational complexity and the error performance is generally inevitable.

From the modeling perspective, when formulated as non-linear least squares problems, conventional spectral-based algorithms ignore the existence of multiple sources in the snapshots and therefore can be regarded as single-source approximation of multi-source criteria [SS90], [POR+93]. As a consequence, if the interference power from other sources is high, the performance of conventional spectral-based algorithms strongly degrades [SN89], [VB95]. This scenario occurs, e.g., when two or multiple sources are closely spaced. In contrast, ML estimators consider the influence of multiple source signals in the non-linear cost function. Therefore, the error performances of ML estimators do not degrade as critically as spectral-based algorithms in the closely spaced source signals case, however, at the cost of high computational complexity.

Based on aforementioned discussions, the research objective of this dissertation is to develop novel low-complexity DOA estimators that provide improved estimation performance compared with conventional spectral-based DOA estimators. More precisely, due to the fluctuation in the received signal strength in practical scenarios, the proposed DOA estimators should possess favorable resolution capability in the threshold regimes, e.g., low SNR or a limited number of measurements at the sensor array. Compared with ML estimators, the computational complexity of the proposed methods are required to be significantly lower, preferably at a complexity comparable to that of conventional spectral-based DOA estimators. In addition, the proposed DOA estimators should be applicable to sensor arrays of arbitrary geometry.

In order to achieve the research objective with specific requirements as mentioned, in this dissertation, the Partial Relaxation (PR) framework [THVP17, THVP18a, THVP18b] is presented. Taking a fundamentally different perspective from the conventional spectral-based algorithms, the PR-based DOA estimators take signals from both "desired" and "interfering" directions into account. However, while the manifold structure of the desired direction is unaltered, the manifold structure of the interfering directions is relaxed to make the problem computationally tractable, hence the name Partial Relaxation. Based on this concept, closed-form expressions of the optimal relaxed interference parameters are first determined and then substituted back into the multi-source criteria, resulting in simple spectral search procedures. In contrast to MUSIC, in which the eigenvectors spanning the noise subspace play an essential role in the calculation of the null-spectrum, it turns out that DOA estimators under the PR framework rely only on the eigenvalues of a particular modified covariance matrix at each candidate direction. Furthermore, the modified covariance matrices in the null-spectra of all proposed DOA estimators share a common structure induced from the PR framework. By exploiting this underlying structure, a general efficient implementation procedure of PR estimators based on the rational function approximation is proposed. In comparison with the corresponding conventional multi-source fitting methods, the PR-based DOA estimators admit significantly simpler solutions while obtaining superior error performance to the conventional spectral-based algorithms.

Concerning the statistical error performance of the PR-based estimators, due to the relaxation of the structure of the array response matrix, a performance loss of the PR-based DOA estimators compared with that of the conventional multi-source counterparts is in general inevitable. In order to precisely characterize the performance loss of the aforementioned PR estimators under the relaxed model as compared to that under the full-information model, one approach is to formulate the relaxed signal model that contains only those model assumptions exploited by the PR algorithms and then to derive the corresponding Cramér-Rao Bound (CRB) from it. The CRB under the relaxed signal model, which is referred to as PR-CRB, represents the lowest bound for the mean-square error (MSE) of all unbiased estimators under the PR framework. In order to guarantee the existence of the PR-CRB, it is essential to introduce a parameterization of the induced PR signal model that ensures the invertibility of the Fisher Information Matrix (FIM). This aspect is non-trivial, as it turn out that the relaxed signal model induced by the PR concept, on the one hand, simplifies the minimization of the estimation criteria, but on the other hand, introduces an over-parameterization which renders the FIM matrix singular if the parameters are redundant. Subsequently, the asymptotic MSEs of two DOA estimators under the PR framework based on the conventional Deterministic Maximum Likelihood (DML) and the Weighted Subspace

Fitting (WSF) are derived and compared with the proposed PR-CRB.

To summarize, the original contributions of this thesis are:

- We categorize conventional spectral-based and parametric DOA estimators in the literature into two groups, namely methods employing high-complexity multi-dimensional search and their low-complexity counterparts that employ single-source approximation.
- We introduce the concept of the PR framework for the DOA estimation problem based on the principle of non-complete relaxation of the array manifold while performing the multi-dimensional search. The adaptation of the PR concept to the conventional DOA estimation methods in the literature results in four corresponding estimators under the PR framework. From the simulation results, the proposed PR estimators exhibit excellent threshold performance without requiring any particular structure of the sensor array.
- In order to reduce the overall computational complexity, we propose an iterative efficient procedure for computing the required null-spectra of the proposed PR estimators.
- We derive a closed-form expression of the CRB of the DOA parameter under the PR signal model and Gaussian assumptions of the source signals and the noise. This compact expression, which is addressed as the partially relaxed CRB (PR-CRB), represents a lower bound for the achievable MSE for all unbiased estimators developed under the PR framework. In addition, further implications on the asymptotic performance of PR estimators and comparison with known results in the literature are discussed.

The detailed outline of this dissertation is as follows:

In **Chapter 2**, the signal model of the DOA estimation problem is introduced. Specific assumptions on the sensor array together with statistical properties of the source signals and the sensor noise are provided. At the end of this chapter, the general DOA estimation problem using data collected at the sensor array is formulated.

The aforementioned contributions are presented in details in the following Chapters 3-6. More precisely, in **Chapter 3**, existing DOA estimation methods in the literature are revised. By reinterpreting the conventional DOA estimators as solutions of non-linear least squares problems, a unified view on the conventional DOA estimators in the literature, which is the motivating background of the proposed work, is obtained.

In **Chapter 4**, the general concept of the proposed PR framework and the corresponding underlying induced signal model are introduced. Under the PR framework, we propose four spectral-based DOA estimators based on the classical Deterministic Maximum Likelihood, Weighted Subspace Fitting, constrained and unconstrained covariance fitting estimator. Unlike subspace-based spectral methods, the expressions of the null-spectra of the proposed PR estimators involve specific eigenvalues of some direction-dependent matrix arguments.

The computational aspect of the PR framework is the topic of discussion in **Chapter 5**. By exploiting the general structure in the matrix arguments in the PR-based DOA estimators, the eigenvalues are efficiently computed in parallel, and the computations of the null-spectra associated with the proposed PR estimators are accelerated. The content of Chapters 3-5 has been published in the following papers

- Minh Trinh-Hoang, Mats Viberg, and Marius Pesavento, "Improved DOA Estimators Using Partial Relaxation Approach," in IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP), December 2017
- Minh Trinh-Hoang, Mats Viberg, and Marius Pesavento, "An Improved DOA Estimator Based on Partial Relaxation Approach," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), April 2018
- Minh Trinh-Hoang, Mats Viberg, and Marius Pesavento, "Partial Relaxation Framework: An Eigenvalue-Based DOA Estimator Framework," IEEE Transactions on Signal Processing, Volume 66, Issue 23, Pages 6190 - 6203, December 2018

In **Chapter 6**, the asymptotic performance bound of the proposed PR estimators is investigated. The expression of the CRB for the DOA parameter under the PR signal model is derived in closed form and compared with known results in the literature. Asymptotic error behaviors of two PR estimators are studied in more details. This chapter is based on the following publications

- Minh Trinh-Hoang, Mats Viberg, and Marius Pesavento, "Cramér-Rao Bound for DOA Estimators under the Partial Relaxation Framework," in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), April 2019

- Minh Trinh-Hoang, Mats Viberg, and Marius Pesavento, "Cramér-Rao Bound for DOA Estimators under the Partial Relaxation Framework: Derivation and Comparison," submitted for IEEE Transactions on Signal Processing, Major Revision, October 2019

In order to illustrate the performance gain in terms of estimation errors and execution time of the proposed methods, in **Chapter 7**, simulation results based on synthetic data are presented. Additionally, the numerical simulation results are compared with the asymptotic statistical results developed in the previous chapter.

In **Chapter 8**, concluding remarks and possible extensions to further research are discussed.

Chapter 2
Signal Model and Problem Formulation

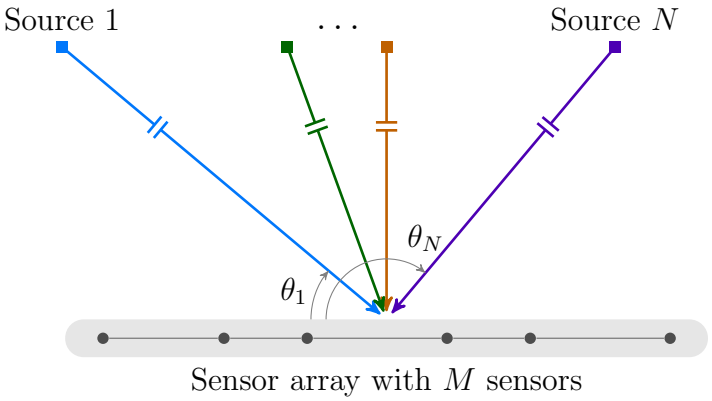

Figure 2.1. Setup of DOA estimation using sensor array

Consider an array of M omnidirectional sensors receiving N narrowband signals emitted from farfield point sources with the corresponding unknown DOAs $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_N]^{\mathrm{T}}$. The DOAs are assumed to be elements of the Field-of-View (FOV) set Θ, i.e., $\theta_n \in \Theta$ for $n = 1, \ldots, N$. The baseband receive signal $\boldsymbol{x}(t) = [x_1(t), \ldots, x_M(t)]^{\mathrm{T}} \in \mathbb{C}^M$ at the time instant t is modeled as

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

where $\boldsymbol{s}(t) = [s_1(t), \ldots, s_N(t)]^{\mathrm{T}} \in \mathbb{C}^N$ denotes the baseband source signal vector from N sources and $\boldsymbol{n}(t) \in \mathbb{C}^M$ represents the noise vector at the sensor array with the noise covariance matrix $\mathbb{E}\left\{\boldsymbol{n}(t)\boldsymbol{n}(t)^{\mathrm{H}}\right\} = \nu \boldsymbol{I}_M$. Regarding the property of the source signal $\boldsymbol{s}(t)$ and the noise vector $\boldsymbol{n}(t)$, the following assumptions are made throughout the thesis:

(As.1) *The number of sources N is known, and the number of sensors M is larger than the number of sources N.*

(As.2) *The source signals $\boldsymbol{s}(t)$ are samples of a stationary process with zero mean and a source covariance matrix $\boldsymbol{P} = \mathbb{E}\left\{\boldsymbol{s}(t)\boldsymbol{s}(t)^{\mathrm{H}}\right\}$. We further assume that the samples in different time instants are statistically independent.*

(As.3) *The noise vector $\boldsymbol{n}(t)$ represents spatially and temporally white, circularly symmetric Gaussian sensor noise with zero mean and variance ν.*

The steering matrix $\boldsymbol{A}(\boldsymbol{\theta}) \in \mathbb{C}^{M \times N}$ in (2.1), which is assumed to have full column rank, is given by

$$\boldsymbol{A}(\boldsymbol{\theta}) = [\boldsymbol{a}(\theta_1), \ldots, \boldsymbol{a}(\theta_N)], \tag{2.2}$$

where $\boldsymbol{a}(\theta_n)$ denotes the sensor array response for the DOA θ_n. We remark that the information on the geometry of the sensor array is captured in the structure of the steering vector $\boldsymbol{\theta}$. In order to estimate the DOA parameter $\boldsymbol{\theta}$ uniquely from the noiseless baseband signal, we make the following assumption on the geometry of the sensor array

(As.4) *For any collection of M distinct values of ϑ_m in the Field-of-View, i.e., $\vartheta_m \in \Theta$ for $m = 1, \ldots, M$, the matrix $[\boldsymbol{a}(\vartheta_1), \ldots, \boldsymbol{a}(\vartheta_M)]$ is full rank.*

We note that the assumption (As.4) is array-dependent and therefore requires further investigation for the specific sensor array and the FOV. In the special case of a Uniform Linear Array (ULA) with the sensor spacing $d \leq \frac{\lambda}{2}$, the steering vector has Vandermonde structure, and therefore the assumption (As.4) is satisfied (see [HJ12, Misc. 9.11.2]) for any M DOAs $\boldsymbol{\vartheta} = [\vartheta_1, \ldots, \vartheta_M]^{\mathrm{T}}$ in the FOV $\Theta = [0°, 180°)$. The N-source array manifold $\mathcal{A}_N$, which contains all possible steering matrices $\boldsymbol{A}(\boldsymbol{\vartheta})$ from N source signals in the at DOAs $\boldsymbol{\vartheta} = [\vartheta_1, \ldots, \vartheta_N]$ in the FOV, is parameterized as follows:

$$\mathcal{A}_N = \{\boldsymbol{A} | \boldsymbol{A} = [\boldsymbol{a}(\vartheta_1), \ldots, \boldsymbol{a}(\vartheta_N)], \vartheta_1 < \ldots < \vartheta_N\}. \tag{2.3}$$

The signal model in (2.1) is rewritten for multiple snapshots $t = 1, \ldots, T$ in a compact notation as

$$\boldsymbol{X} = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{S} + \boldsymbol{N}, \tag{2.4}$$

where $\boldsymbol{X} = [\boldsymbol{x}(1), \ldots, \boldsymbol{x}(T)] \in \mathbb{C}^{M \times T}$ is the baseband receive signal matrix. In a similar manner, we define the source signal matrix $\boldsymbol{S} \in \mathbb{C}^{N \times T}$ and the sensor noise matrix $\boldsymbol{N} \in \mathbb{C}^{M \times T}$ as $\boldsymbol{S} = [\boldsymbol{s}(1), \ldots, \boldsymbol{s}(T)]$ and $\boldsymbol{N} = [\boldsymbol{n}(1), \ldots, \boldsymbol{n}(T)]$, respectively.

From assumptions (As.1)-(As.3), the covariance matrix $\boldsymbol{R}$ of the receive signal $\boldsymbol{x}(t)$ in (2.1) is given by

$$\boldsymbol{R} = \mathbb{E}\left\{\boldsymbol{x}(t)\boldsymbol{x}(t)^{\mathrm{H}}\right\} = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{P}\boldsymbol{A}(\boldsymbol{\theta})^{\mathrm{H}} + \nu\boldsymbol{I}_M. \tag{2.5}$$

If the source signals are non-coherent, i.e., the source covariance matrix $\boldsymbol{P}$ is full rank, we define the eigenvalue decomposition of $\boldsymbol{R}$ into the signal subspace with dimension N and the noise subspace with dimension $(M - N)$ as follows:

$$\boldsymbol{R} = \boldsymbol{U}_{\mathrm{s}}\boldsymbol{\Lambda}_{\mathrm{s}}\boldsymbol{U}_{\mathrm{s}}^{\mathrm{H}} + \boldsymbol{U}_{\mathrm{n}}\boldsymbol{\Lambda}_{\mathrm{n}}\boldsymbol{U}_{\mathrm{n}}^{\mathrm{H}}, \tag{2.6}$$

where $\boldsymbol{\Lambda}_{\mathrm{s}} = \mathrm{diag}(\lambda_1, \ldots, \lambda_N)$ contains the N-largest eigenvalues of $\boldsymbol{R}$. We assume that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_N$. The corresponding principal eigenvectors, which are the columns of $\boldsymbol{U}_{\mathrm{s}}$, span the signal subspace. The remaining eigenvalues, which are all equal to the noise power ν, and the associated eigenvectors spanning the noise subspace are collected in $\boldsymbol{\Lambda}_{\mathrm{n}} = \nu\boldsymbol{I}_N$ and $\boldsymbol{U}_{\mathrm{n}}$, respectively.

In practice, the true covariance matrix $\boldsymbol{R}$ is not available and the sample covariance matrix $\hat{\boldsymbol{R}}$ is used instead:

$$\hat{\boldsymbol{R}} = \frac{1}{T}\boldsymbol{X}\boldsymbol{X}^{\mathrm{H}}. \tag{2.7}$$

Subspace techniques rely on the properties of the eigenspaces of the sample covariance matrix $\hat{\boldsymbol{R}}$, which is decomposed as

$$\hat{\boldsymbol{R}} = \hat{\boldsymbol{U}}\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{U}}^{H} \tag{2.8}$$

$$= \hat{\boldsymbol{U}}_{\mathrm{s}}\hat{\boldsymbol{\Lambda}}_{\mathrm{s}}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}} + \hat{\boldsymbol{U}}_{\mathrm{n}}\hat{\boldsymbol{\Lambda}}_{\mathrm{n}}\hat{\boldsymbol{U}}_{\mathrm{n}}^{\mathrm{H}}. \tag{2.9}$$

In (2.8), the eigenvalues $\hat{\lambda}_m$ for $m = 1, \ldots, M$ of the sample covariance matrix $\hat{\boldsymbol{R}}$, which are referred to as the sample eigenvalues, are collected in the diagonal matrix $\hat{\boldsymbol{\Lambda}} = \operatorname{diag}\left(\hat{\lambda}_1, \ldots, \hat{\lambda}_M\right) \in \mathbb{R}^{M\times M}$. The sample eigenvalues are sorted non-ascendingly, i.e., $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \ldots \geq \hat{\lambda}_M$. The m-th column of the orthonormal matrix $\hat{\boldsymbol{U}} = [\hat{\boldsymbol{u}}_1, \ldots, \hat{\boldsymbol{u}}_M] \in \mathbb{C}^{M\times M}$, which is denoted as $\hat{\boldsymbol{u}}_m$, represents one eigenvector corresponding to the sample eigenvalue $\hat{\lambda}_m$. In accordance with (2.6), the matrix $\hat{\boldsymbol{\Lambda}}_{\mathrm{s}} \in \mathbb{C}^{N\times N}$ in (2.9) is a diagonal matrix, containing the N-largest eigenvalues $\{\hat{\lambda}_1, \ldots, \hat{\lambda}_N\}$, and the matrix $\hat{\boldsymbol{U}}_{\mathrm{s}} \in \mathbb{C}^{M\times N}$ contains the corresponding N-principal eigenvectors of the sample covariance matrix $\hat{\boldsymbol{R}}$. Similarly, $\hat{\boldsymbol{\Lambda}}_{\mathrm{n}} \in \mathbb{C}^{(M-N)\times(M-N)}$ and $\hat{\boldsymbol{U}}_{\mathrm{n}} \in \mathbb{C}^{M\times(M-N)}$ contain the $(M-N)$-noise eigenvalues $\{\hat{\lambda}_{N+1}, \ldots, \hat{\lambda}_M\}$ and the associated noise eigenvectors, respectively.

The general DOA estimation problem based on the receive data at the sensor array is formulated as follows: given the receive baseband signal $\boldsymbol{X}$ from the parametric model in (2.4) and the mapping $\boldsymbol{\vartheta} \mapsto \boldsymbol{A}(\boldsymbol{\vartheta})$ for any $\boldsymbol{\vartheta} \in \Theta^N$, estimate the DOAs $\boldsymbol{\theta}$.

Chapter 3

State-Of-The-Art

In this chapter, we provide a short review of conventional DOA estimators in the literature. As numerous methods have been developed for the task of estimating the DOAs, the list of DOA estimators considered in this chapter is not exhaustive. Nevertheless, a unified perspective and algorithmic links between the considered DOA estimators are presented at the end of this chapter. Such insights play a critical role in the development of the proposed PR framework in the following chapters.

3.1 Spectral-based Algorithms

3.1.1 Beamforming Techniques

Initially developed to filter the desired signal from the interferers in the spatial domain, receive beamforming techniques are also applicable to estimate the DOAs of the source signals in the FOV [VT04,SM05]. By applying proper angle-dependent weights $\boldsymbol{w}(\theta) = [w_1(\theta), \ldots, w_M(\theta)]^{\mathrm{T}}$ on the receive signal $\boldsymbol{X} = [\boldsymbol{x}(1), \ldots, \boldsymbol{x}(T)]$ in a linear fashion, the average output signal power after beamforming is given by

$$P(\theta) = \frac{1}{T} \sum_{t=1}^{T} \left| \boldsymbol{w}(\theta)^{\mathrm{H}} \boldsymbol{x}(t) \right|^2 = \boldsymbol{w}(\theta)^{\mathrm{H}} \hat{\boldsymbol{R}} \boldsymbol{w}(\theta), \tag{3.1}$$

where $\hat{\boldsymbol{R}}$ is the sample covariance matrix of the receive signal defined in (2.7). The output signal power $P(\theta)$ in (3.1) is measured for all candidate directions $\theta \in \Theta$ to obtain the power spectrum. The locations of the separated peaks in $P(\theta)$ are considered as the estimated DOAs. We remark that the choice on the weight vector $\boldsymbol{w}(\theta)$ depends on specific applications and requirements [VB88]. In the literature, the two common choices on the weighting vector for the task of DOA estimation are:

- **Conventional beamformer**: Also known as the delay-and-sum beamformer, the weight vector $\boldsymbol{w}(\theta)$ is chosen such that the phase difference between sensors caused by the source signal impinging from the direction θ is compensated. By enforcing the power constraint on the weight vector $||\boldsymbol{w}(\theta)|| = 1$, the weight

vector of the conventional beamformer for the direction θ is given by [KV96, CVY14, SM05]

$$\boldsymbol{w}_{\text{Conv.BF}}(\theta) = \frac{\boldsymbol{a}(\theta)}{\sqrt{\boldsymbol{a}(\theta)^{\text{H}}\boldsymbol{a}(\theta)}}. \tag{3.2}$$

Substituting the weight vector in (3.2) back to (3.1), the power spectrum of the conventional beamformer is given by

$$P_{\text{Conv.BF}}(\theta) = \frac{\boldsymbol{a}(\theta)^{\text{H}}\hat{\boldsymbol{R}}\boldsymbol{a}(\theta)}{\boldsymbol{a}(\theta)^{\text{H}}\boldsymbol{a}(\theta)}. \tag{3.3}$$

One disadvantage of the conventional beamformer is the limited resolution capability. Using the conventional beamformer, two sources located inside the Rayleigh beamwidth cannot be resolved regardless of the number of snapshots T and the SNR [CVY14, VT04].

- **Capon beamformer**: Unlike the conventional beamformer where the phase difference is compensated, the Capon beamformer [Cap69] choose the weight vector $\boldsymbol{w}(\theta)$ such that the power of interfering directions and noise are minimized while maintaining a fixed gain for the desired direction θ after beamforming. Mathematically, the weight vector of the Capon beamformer $\boldsymbol{w}_{\text{Capon}}(\theta)$ is the solution of the following optimization problem:

$$\min_{\boldsymbol{w}(\theta)} \boldsymbol{w}(\theta)^{\text{H}}\hat{\boldsymbol{R}}\boldsymbol{w}(\theta) \quad \text{subject to} \quad \boldsymbol{w}(\theta)^{\text{H}}\boldsymbol{a}(\theta) = 1. \tag{3.4}$$

The minimizer of (3.4) is given by the following closed-form expression, which is also known as the Wiener-Hopf equation [VT04, SM05]:

$$\boldsymbol{w}_{\text{Capon}}(\theta) = \frac{\hat{\boldsymbol{R}}^{-1}\boldsymbol{a}(\theta)}{\boldsymbol{a}(\theta)\hat{\boldsymbol{R}}^{-1}\boldsymbol{a}(\theta)}. \tag{3.5}$$

Substituting (3.5) to (3.1), the Capon spectrum is given by

$$P_{\text{Capon}}(\theta) = \frac{1}{\boldsymbol{a}(\theta)\hat{\boldsymbol{R}}^{-1}\boldsymbol{a}(\theta)}. \tag{3.6}$$

Compared with the conventional beamformer, the Capon beamformer possesses improved spatial resolution capability [Lac71, HSS93]. For a sufficient number of snapshots T, multiple closely spaced source signals can be resolved by the Capon beamformer. However, the Capon beamformer requires the sample covariance matrix $\hat{\boldsymbol{R}}$ to be invertible, which implies that the number of snapshots T must be at least as large as the number of sensors M. In addition, the DOA estimation performance of the Capon beamformer degrades in the presence of correlated source signals [VB95, Ric05].

3.1.2 MUSIC and its Variants

In contrast to aforementioned beamformer methods which exploits the power information after spatial filtering to estimate the DOAs, the MUSIC estimator exploits the eigenstructure of the covariance matrix in (2.6). Under the assumptions (As.1)-(As.4) and uncorrelated source signals, the following identity holds:

$$\boldsymbol{U}_{\mathrm{n}}^{\mathrm{H}}\boldsymbol{a}(\theta) = \boldsymbol{0}, \tag{3.7}$$

with $\boldsymbol{U}_{\mathrm{n}}$ defined in (2.6) if and only if the candidate direction θ coincides with one of the true DOA, i.e., $\theta \in \{\theta_1, \ldots, \theta_N\}$. The identity in (3.7) is equivalent to

$$\boldsymbol{a}(\theta)^{\mathrm{H}}\boldsymbol{U}_{\mathrm{n}}\boldsymbol{U}_{\mathrm{n}}^{\mathrm{H}}\boldsymbol{a}(\theta) = 0, \tag{3.8}$$

which represents the core principle of the MUSIC algorithm. In practice, the noise subspace $\boldsymbol{U}_{\mathrm{n}}$ is not available, and the sample noise subspace $\hat{\boldsymbol{U}}_{\mathrm{n}}$ in (2.9) is used instead. The MUSIC estimates for the DOA parameters are obtained from the locations of the N-highest separated peaks of the following spectrum:

$$P_{\mathrm{MUSIC}}(\theta) = \frac{1}{\boldsymbol{a}(\theta)^{\mathrm{H}}\hat{\boldsymbol{U}}_{\mathrm{n}}\hat{\boldsymbol{U}}_{\mathrm{n}}^{\mathrm{H}}\boldsymbol{a}(\theta)}. \tag{3.9}$$

A simple generalization of the MUSIC algorithm is the weighted MUSIC (WMUSIC). The spectrum of the WMUSIC estimator is given by

$$P_{\mathrm{WMUSIC}}(\theta) = \frac{1}{\boldsymbol{a}(\theta)^{\mathrm{H}}\hat{\boldsymbol{U}}_{\mathrm{n}}\boldsymbol{W}\hat{\boldsymbol{U}}_{\mathrm{n}}^{\mathrm{H}}\boldsymbol{a}(\theta)}. \tag{3.10}$$

where $\boldsymbol{W} \in \mathbb{C}^{(M-N)\times(M-N)}$ is a positive semidefinite matrix. If $\boldsymbol{W} = \boldsymbol{I}_{M-N}$, i.e., the weighting matrix is the identity matrix, the WMUSIC estimator reduces to the conventional MUSIC estimator in (3.9). In [KB86], the following choice on the weighting matrix $\boldsymbol{W}$ results in the Min-Norm algorithm [Red79, KT83]:

$$\boldsymbol{W}_{\text{Min-Norm}} = \hat{\boldsymbol{U}}_{\mathrm{n}}^{\mathrm{H}}\boldsymbol{e}_1\boldsymbol{e}_1^{\mathrm{T}}\hat{\boldsymbol{U}}_{\mathrm{n}}, \tag{3.11}$$

where $\boldsymbol{e}_1$ is the vector corresponding to the first column of the identity matrix $\boldsymbol{I}_M$. The advantage of introducing a non-identity weighting matrix $\boldsymbol{W}$ in the spectrum is the improved performance in difficult scenarios, e.g., low SNR or low number of snapshots T [Red79, VT04]. Nevertheless, in the large sample case, i.e., as the number of snapshots T tends to infinity, a non-identity weighting matrix leads to degradation of the estimation performance [SN90a]. Among all WMUSIC estimators with different weighting matrices, only the MUSIC estimator is asymptotically efficient in the high SNR regime and uncorrelated source signals. However, similar to the Capon beamformer, the performance of MUSIC and WMUSIC strongly degrades in the case of correlated source signals [SN89, SN90a].

3.2 Parametric Methods

In contrast to the aforementioned spectral-based DOA estimators, the parametric methods jointly estimate the DOA parameters from the minimizer of some specific multidimensional optimization problems. As the information of the signal model in Chapter 2 is captured in the formulation of the optimization problems, the parametric methods generally outperforms the spectral-based DOA estimators at the cost of increased computational complexity. More details on the parametric methods and the corresponding optimization problems are provided in the following.

3.2.1 Deterministic Maximum Likelihood

Under the deterministic signal model, the source signal $\boldsymbol{s}(t)$ is assumed to be deterministic and unknown, while the noise $\boldsymbol{n}(t)$ is complex circular Gaussian distributed with the covariance matrix $\nu\boldsymbol{I}$. The corresponding joint negative log-likelihood function of the receive signal $\boldsymbol{x}(t)$ for $t = 1, \ldots, T$ w.r.t. the parameter vector $\boldsymbol{\alpha} = \left[\boldsymbol{\theta}^{\mathrm{T}}, \boldsymbol{s}(1),^{\mathrm{T}} \ldots, \boldsymbol{s}(T)^{\mathrm{T}}, \nu\right]^{\mathrm{T}}$ is given by [Boh84, Wax85]

$$\ell(\boldsymbol{\alpha}) = MT \log(\pi\nu) + \frac{1}{\nu} \sum_{t=1}^{T} ||\boldsymbol{x}(t) - \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{s}(t)||^2 . \tag{3.12}$$

In the Deterministic ML (DML) method, the DOAs are determined by searching for the optimal parameter vector $\hat{\boldsymbol{\alpha}}$ that minimizes the joint negative log-likelihood function in (3.12). For fixed $\boldsymbol{\theta}$, the ML estimates for the source signal waveform $\boldsymbol{s}(t)$ and the noise power ν are given by [Wax85]

$$\hat{\boldsymbol{s}}(t) = \boldsymbol{A}(\boldsymbol{\theta})^{\dagger}\boldsymbol{x}(t) \quad \text{for } t = 1, \ldots, T, \tag{3.13}$$

$$\hat{\nu} = \frac{1}{M}\mathrm{tr}\left(\boldsymbol{\Pi}^{\perp}_{\boldsymbol{A}(\boldsymbol{\theta})}\hat{\boldsymbol{R}}\right), \tag{3.14}$$

respectively. In (3.13), $\boldsymbol{A}(\boldsymbol{\theta})^{\dagger} = \left(\boldsymbol{A}(\boldsymbol{\theta})^{\mathrm{H}}\boldsymbol{A}(\boldsymbol{\theta})\right)^{-1}\boldsymbol{A}(\boldsymbol{\theta})^{\mathrm{H}}$ is the Moore-Penrose pseudo-inverse of the matrix $\boldsymbol{A}(\boldsymbol{\theta})$. In addition, we define $\boldsymbol{\Pi}_{\boldsymbol{A}(\boldsymbol{\theta})} = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{A}(\boldsymbol{\theta})^{\dagger}$ to be the projection matrix onto the subspace spanned by the columns of the matrix $\boldsymbol{A}(\boldsymbol{\theta})$. Similarly, in (3.14), $\boldsymbol{\Pi}^{\perp}_{\boldsymbol{A}(\boldsymbol{\theta})} = \boldsymbol{I}_M - \boldsymbol{\Pi}_{\boldsymbol{A}(\boldsymbol{\theta})}$ is the projection matrix onto the subspace which is the orthogonal complement to the subspace spanned by $\boldsymbol{A}(\boldsymbol{\theta})$. Substituting the ML estimates in (3.13) and (3.14) back to (3.14) and further simplifying, the optimization problem of the DML estimator w.r.t. the DOA parameters $\boldsymbol{\theta}$ is given by [Boh84, BM86]

$$\left\{\hat{\boldsymbol{\theta}}_{\mathrm{DML}}\right\} = \arg\min_{\boldsymbol{\theta}} \mathrm{tr}\left(\boldsymbol{\Pi}^{\perp}_{\boldsymbol{A}(\boldsymbol{\theta})}\hat{\boldsymbol{R}}\right). \tag{3.15}$$

Although it has been shown in [SN90b, VO91] that the DML estimates are statistically inefficient as the number of snapshots T tends to infinity, the estimation error of the DML method is generally favorable in the finite sample case [Ath05, RFCL06].

3.2.2 Weighted Subspace Fitting

From the identities in (2.5) and (2.6), we observe that the subspace spanned by the columns of $\boldsymbol{U}_{\mathrm{s}}$ is identical to that of $\boldsymbol{A}(\boldsymbol{\theta})$. This observation implies that there exists an invertible matrix $\boldsymbol{V} \in \mathbb{C}^{N\times N}$ that satisfies

$$\boldsymbol{U}_{\mathrm{s}} = \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{V}. \tag{3.16}$$

Motivated by the identity in (3.16), the Weighted Subspace Fitting method estimates the DOAs by solving the following optimization:

$$\left\{\hat{\boldsymbol{\theta}}_{\mathrm{WSF}}, \hat{\boldsymbol{V}}\right\} = \underset{\boldsymbol{\theta},\boldsymbol{V}}{\arg\min} \left|\left|\boldsymbol{W}^{1/2}\left(\hat{\boldsymbol{U}}_{\mathrm{s}} - \boldsymbol{A}\boldsymbol{V}\right)\right|\right|_{\mathrm{F}}^{2}. \tag{3.17}$$

where we restrict the weighting matrix $\boldsymbol{W} \in \mathbb{C}^{N\times N}$ to be positive definite [VO91]. In accordance with the formulation of the DML method in (3.15), the concentrated optimization problem w.r.t. the DOA parameters $\boldsymbol{\theta}$ of the Weighted Subspace Fitting method is formulated in [VO91] as

$$\left\{\hat{\boldsymbol{\theta}}_{\mathrm{WSF}}\right\} = \underset{\boldsymbol{\theta}}{\arg\min}\ \mathrm{tr}\left(\boldsymbol{\Pi}^{\perp}_{\boldsymbol{A}(\boldsymbol{\theta})}\hat{\boldsymbol{U}}_{\mathrm{s}}\boldsymbol{W}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}}\right). \tag{3.18}$$

Similar to the weighted MUSIC, the performance of the WSF estimator depends on the choice of the weighting matrix $\boldsymbol{W}$. In [VO91], the authors showed that by choosing the weighting matrix as

$$\boldsymbol{W} = \hat{\tilde{\boldsymbol{\Lambda}}}^{2}\hat{\boldsymbol{\Lambda}}_{\mathrm{s}}^{-1} \tag{3.19}$$

with $\hat{\tilde{\boldsymbol{\Lambda}}} = \hat{\boldsymbol{\Lambda}}_{\mathrm{s}} - \hat{\nu}\boldsymbol{I}_N$ and $\hat{\nu} = \dfrac{1}{M-N}\sum\limits_{k=N+1}^{M}\hat{\lambda}_k$, the statistical asymptotic estimation error of the WSF method is minimized. In addition, the MSE of the WSF method with the weighting matrix $\boldsymbol{W}$ computed in (3.19) achieves the Cramér-Rao Bound as the number of snapshots T tends to infinity [OVSN93].

3.2.3 Covariance Fitting

Starting from the identity in (2.5) and applying the weighted least squares, we obtain the following formulation for the weighted covariance fitting problem:

$$\left\{\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{R}}_{\mathrm{s}}, \hat{\nu}\right\} = \underset{\boldsymbol{\theta},\boldsymbol{P}\succeq\boldsymbol{0},\nu\geq 0}{\arg\min} \left|\left|\boldsymbol{W}\left(\hat{\boldsymbol{R}} - \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^{\mathrm{H}} - \nu\boldsymbol{I}\right)\right|\right|_{\mathrm{F}}^{2}. \tag{3.20}$$

where $\boldsymbol{W} \in \mathbb{C}^{M \times M}$ is a suitable positive definite weighting matrix. It has been shown in [OSR98] that by choosing $\boldsymbol{W} = \hat{\boldsymbol{R}}^{-1}$, the estimates are asymptotically efficient, and the method is referred to Covariance Matching Estimation Technique, abbreviated as COMET. Similar to the Capon beamformer, COMET is only applicable if the number of snapshots T is sufficiently large so that the sample covariance matrix $\hat{\boldsymbol{R}}$ is invertible.

3.3 Discussions

As observed in the previous section, the DOA estimation problem of the parametric methods can be formulated in a general form as

$$\left\{\hat{\boldsymbol{A}}\right\} = \underset{\boldsymbol{A} \in \mathcal{A}_N}{\arg\min}\ f\left(\boldsymbol{A}, \boldsymbol{Y}\right), \tag{3.21}$$

where $f(\cdot)$ denotes a general cost function, and $\boldsymbol{Y}$ is a data matrix which is obtained from the received baseband signal matrix $\boldsymbol{X}$ through suitable mapping. We remark that different choices of the cost function $f(\cdot)$, the parameterization of $\boldsymbol{A}$ and the data matrix $\boldsymbol{Y}$ are associated with different error performance of the DOA estimators. Since the cost function $f(\cdot)$ is generally non-convex and multi-modal w.r.t. the steering matrix $\boldsymbol{A}$ and the N-source array manifold $\mathcal{A}_N$ is highly structured and non-convex, the optimization problem in (3.21) is generally challenging [VO91, OVSN93, ZW88]. If no particular structure of the array manifold $\mathcal{A}_N$ is exploited, solving for the global minimizer of the optimization problem in (3.21) involves multidimensional exhaustive search. As a consequence, the computational complexity of the parametric methods is in general prohibitively high.

In contrast to the parametric methods which require multidimensional search algorithms, the spectral-based algorithms possess lower computational complexity thanks to the usage of only one dimensional search. Therefore, the DOA estimates of the spectral-based algorithms are generally suboptimal to the estimates obtained from the multidimensional search algorithms. In fact, the methods requiring the spectral search in the aforementioned sections can be interpreted as a relaxed version of the corresponding parametric algorithms. The relaxation in this case is that, instead of considering the full parameters, only a single parameter is assumed in the model while minimizing the objective function, therefore the name single-source approximation [POR+93], [SM05]. In this approach, the single-source array manifold is the feasible set of the optimization problem in (3.21), i.e., $\boldsymbol{A} = \boldsymbol{a} \in \mathcal{A}_1$. While the number of considered parameters is reduced in the single-source approximation approach, the data matrix $\boldsymbol{Y}$ remains unchanged. The locations $\hat{\boldsymbol{a}}$ of N-deepest minima of the null-spectrum $f\left(\boldsymbol{a}, \boldsymbol{Y}\right)$, which

are obtained by performing a spectral sweep search on $\boldsymbol{a} \in \mathcal{A}_1$, correspond to the steering vectors of the estimated DOAs. The above mentioned steps are compactly expressed by the following notation:

$$\{\hat{\boldsymbol{a}}\} = {}^{N}\underset{\boldsymbol{a}\in\mathcal{A}_1}{\arg\min}\ f\left(\boldsymbol{a}, \boldsymbol{Y}\right). \tag{3.22}$$

From now on, unless the dependence of the steering vector $\boldsymbol{a}(\theta)$ on the direction is emphasized, the argument θ will be omitted. Under the single-source approximation approach, conventional spectral-based DOA estimators in Section 3.1 are retrieved by considering the optimizing criteria in Section 3.2 and different data matrices [POR+93]:

- **Measurement Fitting**: Using the cost function of the DML in (3.15) and the data matrix $\boldsymbol{Y} = \hat{\boldsymbol{R}}$ in (2.7) under the single-source approximation, the following optimization problem is obtained:

$$\{\hat{\boldsymbol{a}}\} = {}^{N}\underset{\boldsymbol{a}\in\mathcal{A}_1}{\arg\min}\ \mathrm{tr}\left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{R}}\right). \tag{3.23}$$

  Note that the objective function in (3.23) is the null-spectrum of the conventional beamformer [VT04].

- **Weighted Subspace Fitting (WSF)**: When the single-source approximation is adopted to the cost function in (3.18) with the data matrix $\boldsymbol{Y} = \hat{\boldsymbol{U}}_{\mathrm{s}}\boldsymbol{W}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}}$, where $\hat{\boldsymbol{U}}_{\mathrm{s}}$ is defined in (2.9), and $\boldsymbol{W} \in \mathbb{C}^{N\times N}$ is a positive definite weighting matrix, the following optimization problem is obtained:

$$\{\hat{\boldsymbol{a}}\} = {}^{N}\underset{\boldsymbol{a}\in\mathcal{A}_1}{\arg\min}\ \mathrm{tr}\left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{U}}_{\mathrm{s}}\boldsymbol{W}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}}\right). \tag{3.24}$$

  In a special case when $\boldsymbol{W} = \boldsymbol{I}_N$, the formulation in (3.24) can be shown to be equivalent to the MUSIC estimator [POR+93].

- **Covariance Fitting**: If the single-source approximation is adopted on (3.20) with the data matrix $\boldsymbol{Y} = \hat{\boldsymbol{R}}$ in (2.7) and the identity weighting matrix $\boldsymbol{W} = \boldsymbol{I}$ while neglecting the noise term, we obtain:

$$\{\hat{\boldsymbol{a}}\} = {}^{N}\underset{\boldsymbol{a}\in\mathcal{A}_1}{\arg\min}\ \underset{\sigma_{\mathrm{s}}^2\geq 0}{\min}\left|\left|\hat{\boldsymbol{R}} - \sigma_{\mathrm{s}}^2\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\right|\right|_{\mathrm{F}}^2. \tag{3.25}$$

  The inner optimization problem in (3.25) exhibits a closed-form minimizer $\hat{\sigma}_{\mathrm{s}}^2$ given by [SM05]

$$\hat{\sigma}_{\mathrm{s}}^2 = \underset{\sigma_{\mathrm{s}}^2\geq 0}{\arg\min}\left|\left|\hat{\boldsymbol{R}} - \sigma_{\mathrm{s}}^2\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\right|\right|_{\mathrm{F}}^2 = \frac{\boldsymbol{a}^{\mathrm{H}}\hat{\boldsymbol{R}}\boldsymbol{a}}{(\boldsymbol{a}^{\mathrm{H}}\boldsymbol{a})^2}, \tag{3.26}$$

which is the spectrum of the conventional beamformer in (3.3). If, for simplicity of the solution, a positive semidefinite constraint is enforced to the inner optimization problem in (3.26) as follows:

$$\begin{aligned} \hat{\sigma}_{\mathrm{s}}^2 = \underset{\sigma_{\mathrm{s}}^2 \geq 0}{\arg\min} \left|\left| \hat{\boldsymbol{R}} - \sigma_{\mathrm{s}}^2 \boldsymbol{a}\boldsymbol{a}^{\mathrm{H}} \right|\right|_{\mathrm{F}}^2 \\ \text{subject to } \hat{\boldsymbol{R}} - \sigma_{\mathrm{s}}^2 \boldsymbol{a}\boldsymbol{a}^{\mathrm{H}} \succeq \boldsymbol{0}, \end{aligned} \tag{3.27}$$

then the minimizer $\hat{\sigma}_{\mathrm{s}}^2$ is given by the Capon spectrum [SM05, p. 293], [LSW03]:

$$\hat{\sigma}_{\mathrm{s}}^2 = \frac{1}{\boldsymbol{a}^{\mathrm{H}} \hat{\boldsymbol{R}}^{-1} \boldsymbol{a}}. \tag{3.28}$$

To summarize, the following table provides the overview and links between the aforementioned DOA estimators under the general DOA estimation problem in (3.21):

Methods	**Multi-dimensional Search**	**Single-source Approximation**
Signal Fitting	Deterministic Maximum Likelihood	Conventional beamformer
Subspace Fitting	Weighted Subspace Fitting	Weighted MUSIC
Covariance Fitting	Unweighted COMET	Conventional beamformer (Capon beamformer)

Table 3.1. Overview of state-of-the-art DOA estimators

We remark that the optimization problems using the single-source approximation approach in (3.23), (3.24) and (3.26) are equivalent to the multi-source criteria counterparts under the assumptions of orthogonal steering vectors, i.e., $\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A} = M\boldsymbol{I}_N$, and uncorrelated source signals. In this case, the effects of interfering signals on the desired direction vanish. As a result, the steering vectors are decoupled, and the multi-source optimization problems can be decomposed into multiple single-source estimation problems. Conversely, if the steering vectors are not orthogonal, e.g., in the case of closely spaced sources, the conventional spectral-based methods deviate from the corresponding multi-dimensional search criteria due to the interference of neighboring source signals, resulting in considerable performance reduction.

Chapter 4

Partial Relaxation Framework

In order to relieve the aforementioned drawbacks of the conventional spectral-based algorithms, in this chapter, the general concept for the Partial Relaxation (PR) approach is introduced. Adopting the principle of the PR framework, we propose novel DOA estimators which are based on the multi-dimensional parametric classical least squares problems in (3.15), (3.18), and (3.20).

4.1 General Concept and Underlying Signal Model

Unlike the single-source approximation, our proposed PR framework considers the signals from both the "desired" and "interfering" directions. However, to make the problem more tractable than the multi-dimensional search methods, the array structures of the interfering signals are relaxed. This is the core principle of the PR framework, which is illustrated in Figure 4.1. More precisely, the general multi-dimensional optimization problem in (3.21) is relaxed by replacing the highly structured array manifold $\mathcal{A}_N$ with the relaxed array manifold $\bar{\mathcal{A}}_N$, which is parameterized as

$$\bar{\mathcal{A}}_N = \left\{ \boldsymbol{A} | \boldsymbol{A} = [\boldsymbol{a}(\theta), \boldsymbol{B}], \boldsymbol{a}(\theta) \in \mathcal{A}_1, \boldsymbol{B} \in \mathbb{C}^{M \times (N-1)} \text{ and } \operatorname{rank}(\boldsymbol{A}) = N \right\}. \tag{4.1}$$

Note that $\bar{\mathcal{A}}_N$ retains some structure depending on the geometry of the sensor array, hence the name Partial Relaxation. Due to the relaxation of the steering matrix as described in (4.1), if the cost function of (3.21) is minimized on the relaxed array manifold $\bar{\mathcal{A}}_N$, then the resulting minimizer contains the information of one DOA. Consequently, only one DOA is estimated. Therefore, the grid search is applied similarly to the single-source approximation in Section 3 as follows: for each given direction $\theta \in \Theta$, the objective function in (3.21) is minimized with respect to the unstructured matrix $\boldsymbol{B}$. The optimal value for each direction is essentially the null-spectrum of the PR estimator. The DOAs are then estimated by choosing the N-deepest local minima of the null-spectrum. The rationale for the PR framework is that, each time a candidate DOA θ coincides with one of the true DOAs θ_n for $n = 1, \ldots, N$, then with $\boldsymbol{B}$ modeling all other steering vectors, a perfect fit to the data is attained at high SNR or large number of snapshots T. When θ is different from all true DOAs, the number of degrees-of-freedom in $\boldsymbol{B}$ is not sufficiently large to match to the data perfectly.

Figure 4.1. Concept of the Partial Relaxation Framework

From the modeling perspective, by assuming the steering matrix $\boldsymbol{A}$ to be an element on the relaxed steering manifold $\bar{\mathcal{A}}_N$ instead of the true steering manifold $\mathcal{A}_N$, we imply that the receive signal at the sensor array is generated by a modified signal model that is different from the conventional signal model in Chapter 2. With a slight diversion from the defined notation in (2.4), the modified signal model induced by the PR framework, which is referred to as PR signal model, reads for the multiple snapshots case as

$$\boldsymbol{X} = \boldsymbol{A}(\vartheta)\boldsymbol{S} + \boldsymbol{N}. \tag{4.2}$$

In (4.2), $\boldsymbol{X}$, $\boldsymbol{S}$ and $\boldsymbol{N}$ are the receive signal, the source signal and the receive noise matrix defined in the similar manner as in Chapter 2. The relaxed steering matrix $\boldsymbol{A}(\vartheta) \in \bar{\mathcal{A}}_N$ is defined by the partition

$$\boldsymbol{A}(\vartheta) = [\boldsymbol{a}(\vartheta), \boldsymbol{B}] \tag{4.3}$$

with $\vartheta \in \{\theta_1, \ldots, \theta_N\}$ and $\boldsymbol{B} \in \mathbb{C}^{M \times (N-1)}$. The PR signal model in (4.2) essentially represents the scenario consisting of a structured signal under corruption of low-rank unstructured interference and noise. By applying the multidimensional optimizing criteria in Chapter 3 to the PR signal model, the corresponding PR-based DOA estimators are obtained. More details regarding the derivation of the proposed PR estimators are provided in the following sections.

4.2 Proposed Estimators

4.2.1 Partially Relaxed Deterministic Maximum Likelihood (PR-DML)

Adopting the concept of the PR framework on the objective function in (3.15) yields the following optimization problem:

$$\{\hat{\boldsymbol{a}}_{\text{PR-DML}}\} = {}^{N}\!\arg\min_{\boldsymbol{a} \in \mathcal{A}_1} \min_{\boldsymbol{B}} \operatorname{tr}\left(\boldsymbol{\Pi}^{\perp}_{[\boldsymbol{a},\boldsymbol{B}]}\hat{\boldsymbol{R}}\right). \tag{4.4}$$

Using the identity in (A.1), the objective function in (4.4) is rewritten in the following manner to decouple $\boldsymbol{a}$ and $\boldsymbol{B}$ partially:

$$\operatorname{tr}\left(\boldsymbol{\Pi}_{[\boldsymbol{a},\boldsymbol{B}]}^{\perp}\hat{\boldsymbol{R}}\right) = \operatorname{tr}\left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{R}}\right) - \operatorname{tr}\left(\boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}}\hat{\boldsymbol{R}}\right), \tag{4.5}$$

where we use the convention that $\operatorname{tr}\left(\boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}}\hat{\boldsymbol{R}}\right) = 0$ if $\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B} = \boldsymbol{0}$. Since the first term on the right hand side of (4.5) does not depend on $\boldsymbol{B}$, the inner optimization problem in (4.4) is equivalent to

$$\max_{\boldsymbol{B}\in\mathbb{C}^{M\times(N-1)}} \operatorname{tr}\left(\boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}}\hat{\boldsymbol{R}}\right). \tag{4.6}$$

A closed-form expression of the global maximum of the optimization problem (4.6) is obtained from the following theorem.

Theorem 4.1. *For any positive semidefinite matrix $\boldsymbol{Q} \in \mathbb{C}^{M\times M}$ with rank$(\boldsymbol{Q}) \geq N$ and any given non-zero vector $\boldsymbol{a} \in \mathbb{C}^{M}$, we have*

$$\max_{\boldsymbol{B}\in\mathbb{C}^{M\times(N-1)}} \operatorname{tr}\left(\boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}}\boldsymbol{Q}\right) = \sum_{k=1}^{N-1} \lambda_k\left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{Q}\right), \tag{4.7}$$

where $\lambda_k(\cdot)$ denotes the k-th largest eigenvalue of the matrix argument

Proof of Theorem 4.1. See Appendix B □

Applying Theorem 4.1 with $\boldsymbol{Q} = \hat{\boldsymbol{R}}$ and substituting (4.5) and (4.7) back into the objective function of (4.4), the null-spectrum of the PR-DML estimator is obtained as

$$\begin{aligned} f_{\text{PR-DML}}(\theta) &= \min_{\boldsymbol{B}} \operatorname{tr}\left(\boldsymbol{\Pi}_{[\boldsymbol{a}(\theta),\boldsymbol{B}]}^{\perp}\hat{\boldsymbol{R}}\right) \\ &= \sum_{k=N}^{M} \lambda_k(\boldsymbol{\Pi}_{\boldsymbol{a}(\theta)}^{\perp}\hat{\boldsymbol{R}}). \end{aligned} \tag{4.8}$$

The estimated DOAs $\hat{\boldsymbol{\theta}} = \left[\hat{\theta}_1, \ldots, \hat{\theta}_N\right]^{\mathrm{T}}$ are then determined by choosing the locations of the N-deepest minima of the null-spectrum $f_{\text{PR-DML}}(\theta)$. In Appendix C, an alternative derivation of the null-spectrum of the PR-DML estimator in (4.8) based on the best low-rank approximation problem [EY36] is provided. As further elaborated in Chapter 5, the null-spectrum $f_{\text{PR-DML}}(\theta)$ in (4.8) can be efficiently computed without necessarily performing a full eigenvalue decomposition at each direction, i.e., computing the corresponding set of eigenvectors of the matrix $\boldsymbol{\Pi}_{\boldsymbol{a}(\theta)}^{\perp}\hat{\boldsymbol{R}}$ is not required.

4.2.2 Partially Relaxed Weighted Subspace Fitting (PR-WSF)

From the mathematical formulation of the WSF estimator in (3.24), the optimization problem corresponding to the PR-WSF estimator reads

$$\{\hat{\boldsymbol{a}}_{\text{PR-WSF}}\} = {}^{N}\underset{\boldsymbol{a}\in\mathcal{A}_1}{\arg\min} \min_{\boldsymbol{B}} \operatorname{tr}\left(\boldsymbol{\Pi}^{\perp}_{[\boldsymbol{a},\boldsymbol{B}]}\hat{\boldsymbol{U}}_{\text{s}}\boldsymbol{W}\hat{\boldsymbol{U}}_{\text{s}}^{\text{H}}\right). \tag{4.9}$$

Comparing with (4.4), the formulation of the PR-WSF estimator is similar to the that of the PR-DML estimator with the sample covariance matrix $\hat{\boldsymbol{R}}$ replaced by $\hat{\boldsymbol{U}}_{\text{s}}\boldsymbol{W}\hat{\boldsymbol{U}}_{\text{s}}^{\text{H}}$. Therefore, by following the derivation of the PR-DML estimator and applying Theorem 4.1 with $\boldsymbol{Q} = \hat{\boldsymbol{U}}_{\text{s}}\boldsymbol{W}\hat{\boldsymbol{U}}_{\text{s}}^{\text{H}}$, the null-spectrum of the PR-WSF is calculated as

$$f_{\text{PR-WSF}}(\theta) = \sum_{k=N}^{M} \lambda_k(\boldsymbol{\Pi}^{\perp}_{\boldsymbol{a}(\theta)}\hat{\boldsymbol{U}}_{\text{s}}\boldsymbol{W}\hat{\boldsymbol{U}}_{\text{s}}^{\text{H}}). \tag{4.10}$$

Note that in a special case when $\boldsymbol{W} = \boldsymbol{I}_N$, the proposed estimator in (4.10) is equivalent to the MUSIC estimator, which is proven in Appendix D. From now on, if not further specified, the weighting matrix $\boldsymbol{W}$ is chosen as in (3.19).

4.2.3 Partially Relaxed Constrained Covariance Fitting (PR-CCF)

To derive novel estimators based on the covariance fitting problems in (3.26) and (3.27), we follow the principle of the PR framework for the array steering matrix by relaxing $\boldsymbol{A} = [\boldsymbol{a}, \boldsymbol{B}]$ with an arbitrary matrix $\boldsymbol{B} \in \mathbb{C}^{M\times(N-1)}$. Similarly, we partition the waveform matrix $\boldsymbol{S} = \left[\boldsymbol{s}, \boldsymbol{H}^{\text{T}}\right]^{\text{T}}$ in the signal model in (2.4) with $\boldsymbol{s} \in \mathbb{C}^{T\times 1}$ and $\boldsymbol{H} \in \mathbb{C}^{(N-1)\times T}$ to obtain

$$\boldsymbol{X} = \boldsymbol{a}\boldsymbol{s}^{\text{T}} + \boldsymbol{E} + \boldsymbol{N}, \tag{4.11}$$

where $\boldsymbol{E} = \boldsymbol{B}\boldsymbol{H} \in \mathbb{C}^{M\times T}$ models the received signal of the remaining $(N-1)$ sources with the relaxed array manifold structure and therefore $\operatorname{rank}(\boldsymbol{E}) \leq N-1$. Furthermore, we assume that the sample covariance matrix $\hat{\boldsymbol{R}}$ is positive definite, and the signals from other sources are uncorrelated with the signals from the direction $\boldsymbol{a}$. The justification for the positive definiteness of the sample covariance matrix $\hat{\boldsymbol{R}}$ will be discussed later. Similar to the Capon beamformer in (3.27), the partially relaxed constrained covariance fitting (PR-CCF) problem is formulated as follows:

$$\begin{aligned} \{\hat{\boldsymbol{a}}_{\text{PR-CCF}}\} = {}^{N}\underset{\boldsymbol{a}\in\mathcal{A}_1}{\arg\min} \min_{\sigma_{\text{s}}^2\geq 0, \boldsymbol{E}} & \left|\left|\hat{\boldsymbol{R}} - \sigma_{\text{s}}^2\boldsymbol{a}\boldsymbol{a}^{\text{H}} - \boldsymbol{E}\boldsymbol{E}^{\text{H}}\right|\right|_{\text{F}}^2 \\ \text{subject to } & \hat{\boldsymbol{R}} - \sigma_{\text{s}}^2\boldsymbol{a}\boldsymbol{a}^{\text{H}} - \boldsymbol{E}\boldsymbol{E}^{\text{H}} \succeq \boldsymbol{0} \\ & \operatorname{rank}(\boldsymbol{E}) \leq N-1. \end{aligned} \tag{4.12}$$

Keeping $\{\sigma_{\mathrm{s}}^2, \boldsymbol{a}\}$ fixed and minimizing the objective function of (4.12) with respect to $\boldsymbol{E}$, a low-rank approximation problem is obtained as [HW88]

$$\sum_{k=N}^{M} \lambda_k^2(\hat{\boldsymbol{R}} - \sigma_{\mathrm{s}}^2 \boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}) = \min_{\boldsymbol{E}} \left|\left| \hat{\boldsymbol{R}} - \sigma_{\mathrm{s}}^2 \boldsymbol{a}\boldsymbol{a}^{\mathrm{H}} - \boldsymbol{E}\boldsymbol{E}^{\mathrm{H}} \right|\right|_{\mathrm{F}}^2 \quad \text{subject to } \mathrm{rank}(\boldsymbol{E}) \leq N-1. \tag{4.13}$$

By performing the following eigenvalue decomposition:

$$\hat{\boldsymbol{R}} - \sigma_{\mathrm{s}}^2 \boldsymbol{a}\boldsymbol{a}^{\mathrm{H}} = \tilde{\boldsymbol{U}}_{\mathrm{s}} \tilde{\boldsymbol{\Lambda}}_{\mathrm{s}} \tilde{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}} + \tilde{\boldsymbol{U}}_{\mathrm{n}} \tilde{\boldsymbol{\Lambda}}_{\mathrm{n}} \tilde{\boldsymbol{U}}_{\mathrm{n}}^{\mathrm{H}}, \tag{4.14}$$

where $\tilde{\boldsymbol{\Lambda}}_{\mathrm{s}}$ and $\tilde{\boldsymbol{U}}_{\mathrm{s}}$ contain the $(N-1)$ - largest eigenvalues and the corresponding principal eigenvectors of $\hat{\boldsymbol{R}} - \sigma_{\mathrm{s}}^2 \boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}$, respectively, a minimizer $\hat{\boldsymbol{E}}$ of (4.13) satisfies

$$\hat{\boldsymbol{E}}\hat{\boldsymbol{E}}^{\mathrm{H}} = \tilde{\boldsymbol{U}}_{\mathrm{s}} \tilde{\boldsymbol{\Lambda}}_{\mathrm{s}} \tilde{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}}. \tag{4.15}$$

Substituting the minimizer $\hat{\boldsymbol{E}}$ in (4.15) back to the inner optimization problem in (4.12), we observe that, the constraint $\hat{\boldsymbol{R}} - \sigma_{\mathrm{s}}^2 \boldsymbol{a}\boldsymbol{a}^{\mathrm{H}} - \hat{\boldsymbol{E}}\hat{\boldsymbol{E}}^{\mathrm{H}} \succeq \boldsymbol{0}$ implies that $\hat{\boldsymbol{R}} - \sigma_{\mathrm{s}}^2 \boldsymbol{a}\boldsymbol{a}^{\mathrm{H}} \succeq \boldsymbol{0}$. Conversely, for each $\sigma_{\mathrm{s}}^2 \geq 0$, if $\hat{\boldsymbol{R}} - \sigma_{\mathrm{s}}^2 \boldsymbol{a}\boldsymbol{a}^{\mathrm{H}} \succeq \boldsymbol{0}$, from (4.14) and (4.15), we conclude that $\hat{\boldsymbol{R}} - \sigma_{\mathrm{s}}^2 \boldsymbol{a}\boldsymbol{a}^{\mathrm{H}} - \hat{\boldsymbol{E}}\hat{\boldsymbol{E}}^{\mathrm{H}} \succeq \boldsymbol{0}$. As a consequence, an equivalent formulation of the inner problem in (4.12) is obtained as follows:

$$\begin{aligned} &\min_{\sigma_{\mathrm{s}}^2 \geq 0} \sum_{k=N}^{M} \lambda_k^2(\hat{\boldsymbol{R}} - \sigma_{\mathrm{s}}^2 \boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}) \\ &\text{subject to } \hat{\boldsymbol{R}} - \sigma_{\mathrm{s}}^2 \boldsymbol{a}\boldsymbol{a}^{\mathrm{H}} \succeq \boldsymbol{0}. \end{aligned} \tag{4.16}$$

It can be easily shown from the Weyl's inequality regarding the eigenvalues of the modified Hermitian matrix [HRW98] and the positive semidefiniteness of $\hat{\boldsymbol{R}} - \sigma_{\mathrm{s}}^2 \boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}$ that, as long as the constraint in (4.16) is not violated, the objective function in (4.16) is strictly decreasing as σ_{s}^2 increases. Therefore, $\hat{\sigma}_{\mathrm{s,\,C}}^2$ is a minimizer of (4.16) if and only if the matrix $\hat{\boldsymbol{R}} - \hat{\sigma}_{\mathrm{s,\,C}}^2 \boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}$ possesses at least one eigenvalue equal to zero. Consequently, the minimizer $\hat{\sigma}_{\mathrm{s,\,C}}^2$ is obtained from the Capon spectrum [SM05, Eq. (6.5.33)]:

$$\hat{\sigma}_{\mathrm{s,\,C}}^2 = \frac{1}{\boldsymbol{a}^{\mathrm{H}} \hat{\boldsymbol{R}}^{-1} \boldsymbol{a}}. \tag{4.17}$$

Substitute (4.16) and (4.17) back into (4.12), the PR-CCF estimator returns the estimated DOAs by determining the N-deepest minima of the following null-spectrum:

$$f_{\text{PR-CCF}}(\theta) = \sum_{k=N}^{M} \lambda_k^2 \left(\hat{\boldsymbol{R}} - \frac{1}{\boldsymbol{a}(\theta)^{\mathrm{H}} \hat{\boldsymbol{R}}^{-1} \boldsymbol{a}(\theta)} \boldsymbol{a}(\theta)\boldsymbol{a}(\theta)^{\mathrm{H}} \right). \tag{4.18}$$

We remark that the expression of the PR-CCF estimator in (4.18) involves the inverse of the sample covariance matrix $\hat{\boldsymbol{R}}$. If the sample covariance matrix $\hat{\boldsymbol{R}}$ is singular,

e.g., in the case that the number of snapshots T is smaller than the number of sensor M, then the null-spectrum of the PR-CCF estimator cannot be computed using (4.18). More precisely, due to the positive semidefinite constraint in (4.12), the minimizer $\hat{\sigma}^2_{\mathrm{s,\,C}}$ of the optimization problem in (4.16) is zero regardless of the direction θ. As a result, the null-spectrum of the PR-CCF is constant, and no DOA can be estimated by the PR-CCF method.

From the aforementioned discussions, the positive semidefinite constraint in (4.12), on the one hand, simplifies the minimization of the cost function in the subsequent derivation by restricting the feasible region of σ_{s}^2. On the other hand, introducing the positive semidefinite constraint prevents the application of the PR-CCF estimator in the singular sample covariance matrix case. In order to alleviate this disadvantage of the PR-CCF estimator, the diagonal loading technique [LSW03], [Vor14] can be applied on the sample covariance matrix $\hat{\boldsymbol{R}}$. In this case, we replace the sample covariance matrix $\hat{\boldsymbol{R}}$ in (4.18) with $\hat{\boldsymbol{R}}-\gamma\boldsymbol{I}$, where γ is an angle-dependent loading factor. The choice on the loading factor γ and its influence on the estimation performance are subject of future research. Another alternative method is to neglect the positive semidefinite constraint while formulating the covariance fitting problem under consideration of the PR principle. More details on this approach are provided in the following section.

4.2.4 Partially Relaxed Unconstrained Covariance Fitting (PR-UCF)

Comparing with the constrained version presented in Section 4.2.3, the formulation of the PR-UCF omits the positive semidefiniteness constraint to yield the following optimization problem:

$$\begin{aligned}\{\hat{\boldsymbol{a}}_{\text{PR-UCF}}\} = \; {}^{N}\!\underset{\boldsymbol{a}\in\mathcal{A}_1}{\arg\min} \min_{\sigma_{\mathrm{s}}^2\geq 0,\boldsymbol{E}} \left|\left|\hat{\boldsymbol{R}} - \sigma_{\mathrm{s}}^2\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}} - \boldsymbol{E}\boldsymbol{E}^{\mathrm{H}}\right|\right|_{\mathrm{F}}^2 \\ \text{subject to } \operatorname{rank}(\boldsymbol{E}) \leq N-1.\end{aligned} \tag{4.19}$$

When minimizing with respect to $\boldsymbol{E}$, the minimizer $\hat{\boldsymbol{E}}$ of the optimization problem in (4.19) is obtained from the best rank-$(N-1)$ approximation of $\hat{\boldsymbol{R}} - \sigma_{\mathrm{s}}^2\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}$ as described in (4.14) and (4.15). Hence, the inner optimization of the PR-UCF estimator at each direction $\boldsymbol{a} = \boldsymbol{a}(\theta)$ is

$$\min_{\sigma_{\mathrm{s}}^2\geq 0} \sum_{k=N}^{M} \lambda_k^2\left(\hat{\boldsymbol{R}} - \sigma_{\mathrm{s}}^2\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\right). \tag{4.20}$$

Unlike the constrained variant in (4.16), the minimizer of $g(\sigma_{\mathrm{s}}^2) = \sum_{k=N}^{M} \lambda_k^2 \left(\hat{\boldsymbol{R}} - \sigma_{\mathrm{s}}^2 \boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\right)$ in (4.20) with respect to σ_{s}^2, denoted as $\hat{\sigma}_{\mathrm{s,\,U}}^2$, does not have a closed-form solution. However, a numerical solution of $\hat{\sigma}_{\mathrm{s,\,U}}^2$ can be determined by noting that the function $g(\sigma_{\mathrm{s}}^2)$ is continuously differentiable, and the derivative $\dfrac{\mathrm{d}g}{\mathrm{d}\sigma_{\mathrm{s}}^2}\left(\sigma_{\mathrm{s}}^2\right)$ is given by

$$\frac{\mathrm{d}g}{\mathrm{d}\sigma_{\mathrm{s}}^2}\left(\sigma_{\mathrm{s}}^2\right) = -\sum_{k=N}^{M} \frac{2\bar{\lambda}_k(\sigma_{\mathrm{s}}^2)}{\sigma_{\mathrm{s}}^4 \boldsymbol{a}^{\mathrm{H}} \left(\hat{\boldsymbol{R}} - \bar{\lambda}_k(\sigma_{\mathrm{s}}^2)\boldsymbol{I}_M\right)^{-2} \boldsymbol{a}}, \tag{4.21}$$

where we introduce the following shorthand notation:

$$\bar{\lambda}_k(\sigma_{\mathrm{s}}^2) = \lambda_k\left(\hat{\boldsymbol{R}} - \sigma_{\mathrm{s}}^2 \boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\right). \tag{4.22}$$

The proof for the identity in (4.21) is given in Appendix E. Note that the denominator in each summand of the expression in (4.21) is positive, we consider two extreme cases:

- If $\sigma_{\mathrm{s}}^2 \to 0$, then $\bar{\lambda}_k(\sigma_{\mathrm{s}}^2) \geq 0$ with $k = N, \ldots, M$ and therefore

$$\lim_{\sigma_{\mathrm{s}}^2 \to 0} \frac{\mathrm{d}g}{\mathrm{d}\sigma_{\mathrm{s}}^2}\left(\sigma_{\mathrm{s}}^2\right) < 0. \tag{4.23}$$

- If $\sigma_{\mathrm{s}}^2 \to \infty$, the rank-one component $-\sigma_{\mathrm{s}}^2 \boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}$ is dominant to $\hat{\boldsymbol{R}}$ and thus we obtain an asymptotic result for the smallest eigenvalues $\bar{\lambda}_M\left(\sigma_{\mathrm{s}}^2\right)$ as follows:

$$\lim_{\sigma_{\mathrm{s}}^2 \to \infty} \frac{\bar{\lambda}_M\left(\sigma_{\mathrm{s}}^2\right)}{\sigma_{\mathrm{s}}^2 \left|\left|\boldsymbol{a}\right|\right|_2^2} = -1. \tag{4.24}$$

  In addition, the remaining eigenvalues $\bar{\lambda}_k\left(\sigma_{\mathrm{s}}^2\right)$ with $k = N, \ldots, (M-1)$ are always bounded above and below thanks to the Weyl's inequality [HRW98]. Applying this remark and the identity in (4.24) to (4.21) leads to

$$\lim_{\sigma_{\mathrm{s}}^2 \to \infty} \frac{\mathrm{d}g}{\mathrm{d}\sigma_{\mathrm{s}}^2}\left(\sigma_{\mathrm{s}}^2\right) = \infty. \tag{4.25}$$

From (4.21), (4.23) and (4.25), there exists a sufficiently small $\sigma_{\mathrm{s,\,left}}^2$ and a sufficiently large $\sigma_{\mathrm{s,\,right}}^2$ so that the sign of the derivative $\dfrac{\mathrm{d}g}{\mathrm{d}\sigma_{\mathrm{s}}^2}\left(\sigma_{\mathrm{s}}^2\right)$ changes from negative to positive in the interval $\left[\sigma_{\mathrm{s,\,left}}^2, \sigma_{\mathrm{s,\,right}}^2\right]$. Therefore, the minimizer of the optimization problem in (4.20) exists. A simple bisection search [Bur10] can applied to compute a local minimizer $\hat{\sigma}_{\mathrm{s,\,U}}^2$ of (4.20). We remark that more advanced iterative methods can be applied for faster convergence to the local minimizer. However, computing the global minimizer of (4.20) is challenging as the cost function is non-convex (see the expression

of the second derivative in (E.8)). From the numerical results, which are illustrated in Chapter 7, the bisection search is sufficient for the task of DOA estimation. The steps to determine a search interval for the bisection search and the computation of the null-spectrum of the PR-UCF estimator at each direction $\boldsymbol{a}(\theta)$ are summarized in Algorithm 1.

Algorithm 1 Calculating the null-spectrum of PR-UCF at a given direction $\boldsymbol{a} = \boldsymbol{a}(\theta)$

Initialization: $\sigma^2_{\mathrm{s,\ left}} = \sigma^2_{\mathrm{s,\ right}} > 0$, tolerance ϵ, the derivative $\dfrac{\mathrm{d}g}{\mathrm{d}\sigma^2_{\mathrm{s}}}\left(\sigma^2_{\mathrm{s}}\right)$ defined in (4.21)
if $\dfrac{\mathrm{d}g}{\mathrm{d}\sigma^2_{\mathrm{s}}}\left(\sigma^2_{\mathrm{s,\ left}}\right) < 0$ **then**
  repeat
    $\sigma^2_{\mathrm{s,\ right}} \leftarrow 2\sigma^2_{\mathrm{s,\ right}}$
  until $\dfrac{\mathrm{d}g}{\mathrm{d}\sigma^2_{\mathrm{s}}}\left(\sigma^2_{\mathrm{s,\ right}}\right) > 0$
else
  repeat
    $\sigma^2_{\mathrm{s,\ left}} \leftarrow \sigma^2_{\mathrm{s,\ left}}/2$
  until $\dfrac{\mathrm{d}g}{\mathrm{d}\sigma^2_{\mathrm{s}}}\left(\sigma^2_{\mathrm{s,\ left}}\right) < 0$
end if
while $|\sigma^2_{\mathrm{s,\ right}} - \sigma^2_{\mathrm{s,\ left}}| > \epsilon$ **do**
  $\sigma^2_{\mathrm{s,\ mid}} \leftarrow \dfrac{\sigma^2_{\mathrm{s,\ left}} + \sigma^2_{\mathrm{s,\ right}}}{2}$
  if $\dfrac{\mathrm{d}g}{\mathrm{d}\sigma^2_{\mathrm{s}}}\left(\sigma^2_{\mathrm{s,\ mid}}\right) > 0$ **then**
    $\sigma^2_{\mathrm{s,\ right}} \leftarrow \sigma^2_{\mathrm{s,\ mid}}$
  else
    $\sigma^2_{\mathrm{s,\ left}} \leftarrow \sigma^2_{\mathrm{s,\ mid}}$
  end if
end while
$\sigma^2_{\mathrm{s,\ U}} \leftarrow \dfrac{\sigma^2_{\mathrm{s,\ left}} + \sigma^2_{\mathrm{s,\ right}}}{2}$
return $f_{\text{PR-UCF}}(\theta) = \sum\limits_{k=N}^{M} \lambda_k^2 \left(\hat{\boldsymbol{R}} - \hat{\sigma}^2_{\mathrm{s,\ U}} \boldsymbol{a}(\theta)\boldsymbol{a}(\theta)^{\mathrm{H}}\right)$

4.3 Summary

In this chapter, the concept of the PR framework has been introduced. At each candidate direction, an optimization problem which considers the existence of multiple

source signals impinging on the sensor array is formulated. To make the optimization problem more tractable, the manifold structure of the "interfering" signals is relaxed, which results in closed-form estimates for the interference parameters. Thanks to this relaxation, the conventional multi-dimensional optimization problem reduces to a simple spectral search in the FOV. Following this principle, four estimators under the PR framework are derived. The relation between the proposed methods and the state-of-the-art algorithms reviewed in Chapter 3 is summarized in Table 4.1.

Methods	**Multi-dimensional Search**	**Partial Relaxation**	**Single-source Approximation**
Signal Fitting	DML	PR-DML	Conventional beamformer
Subspace Fitting	WSF	PR-WSF	Weighted MUSIC
Covariance Fitting	Unweighted COMET	PR-UCF PR-CCF	Conventional beamformer Capon beamformer

Table 4.1. Overview of PR estimators in relation to conventional DOA estimators

Comparing with the COMET estimator in (3.20), the weighting matrix and the noise covariance matching term are neglected in the formulations of the covariance fitting-based PR estimators in (4.12) and (4.19). By considering the additional impact of the sensor noise in the unconstrained covariance fitting problem, the partially relaxed full covariance fitting (PR-FCF) estimator has been proposed in [STM+19]. Possible extensions of the PR-FCF estimator by incorporating the weighting matrix are subjects to further investigation.

From the expression of the null-spectra in (4.8), (4.10), (4.18) and Algorithm 1, we note that the computation of specific eigenvalues of a matrix that depends on the desired direction θ is a crucial task for all estimators under the PR framework. Compared with the corresponding single-source approximation methods, whose expressions of the spectra only requires matrix-vector multiplications, the computational complexities of the proposed PR methods using off-the-shelf eigenvalue decomposition algorithms are significantly higher. In the next chapter, a general numerical procedure which allows more efficient computation of the null-spectra of the proposed estimators under PR framework is developed.

Chapter 5

Efficient Implementation of PR Estimators

5.1 Motivation

As introduced in Chapter 4, the proposed PR approach involves the estimation procedures which require extensive eigenvalue computation for the evaluation of the null-spectrum over the entire FOV. In fact, the null-spectra in (4.8), (4.10) and (4.18) and Algorithm 1 depend only on the $(M - N + 1)$-smallest eigenvalues of the matrix arguments, and therefore the explicit computation of the eigenvectors can be avoided. Generally, if no particular structure of the matrix is exploited, the eigenvalue decomposition requires $O(K^2L)$ operations where K is the dimension of the matrix and L is the number of required eigenvalues [GVL13, Ch.8]. This computational complexity may be prohibitive for specific practical applications and limit the usage of the proposed PR-based DOA estimators in practice if no acceleration procedure is considered. Therefore, the development of computationally efficient numerical methods to compute the eigenvalues associated with the null-spectra of PR estimators is of great interest.

From an algorithmic perspective, the mathematical expressions of the null-spectra of the proposed PR estimators in Chapter 4 share a common underlying problem structure in the sense that they all require, as a main computational task, the computation of the eigenvalues of a generic matrix form as follows:

$$\boldsymbol{D} - \bar{\rho}\boldsymbol{z}\boldsymbol{z}^{\mathrm{H}} = \bar{\boldsymbol{U}}\bar{\boldsymbol{D}}\bar{\boldsymbol{U}}^{\mathrm{H}}. \tag{5.1}$$

In (5.1), $\boldsymbol{D} = \mathrm{diag}\,(d_1, \ldots, d_K) \in \mathbb{R}^{K\times K}$ is a constant real diagonal matrix, $\bar{\rho} \in \mathbb{R}$ is an arbitrary positive real scalar and $\boldsymbol{z} = [z_1, \ldots, z_K]^{\mathrm{T}} \in \mathbb{C}^{K\times 1}$ is a direction-dependent complex-valued vector. The relationship between the generic form in (5.1) and the null-spectra or the PR estimators in (4.8), (4.10), (4.18) and Algorithm 1 is further detailed in the sections below. Since the expression on the left hand side of (5.1) denotes a Hermitian matrix obtained by subtracting a rank-one matrix from a constant diagonal matrix, the term *rank-one modified Hermitian matrix* is adopted. As presented in the following sections, this particular structure allows a faster implementation of the eigenvalue decomposition, and thus accelerates the computation of the null-spectrum of the proposed estimators in Chapter 4.

5.2 Eigenvalue Decomposition of a Rank-One Modified Hermitian Matrix

Initially proposed by Bunch, Nielsen and Sørensen in [BNS78] as a support for calculating the eigenvalues and eigenvectors of symmetric tridiagonal matrices in parallel, the procedure of determining a complex-valued eigensystem of a rank-one modified Hermitian matrix is based on the interlacing theorem as follows:

Theorem 5.1. *Let* $\{d_1, \ldots, d_K\}$ *be the elements on the diagonal of the matrix* $\boldsymbol{D} \in \mathbb{R}^{K\times K}$ *where* $\{d_1, \ldots, d_K\}$ *are distinct and sorted in descending order. Further assume that* $\bar{\rho} > 0$ *and* $\boldsymbol{z} \in \mathbb{C}^{K\times 1}$ *contains only non-zero entries. If the eigenvalues* $\{\bar{d}_1, \ldots, \bar{d}_K\}$ *of the matrix* $\boldsymbol{D} - \bar{\rho}\boldsymbol{z}\boldsymbol{z}^H$ *are also sorted in descending order, then:*

- $\{\bar{d}_1, \ldots, \bar{d}_K\}$ *are the solutions of the equation* $R(x) = 0$*, where the rational function* $R(x)$ *is given by:*

$$\begin{aligned} R(x) &= 1 - \bar{\rho}\boldsymbol{z}^{\mathrm{H}} \left(\boldsymbol{D} - x\boldsymbol{I}_K\right)^{-1} \boldsymbol{z} \\ &= 1 - \bar{\rho} \sum_{k=1}^{K} \frac{|z_k|^2}{d_k - x}. \end{aligned} \tag{5.2}$$

- $\{\bar{d}_1, \ldots, \bar{d}_K\}$ *satisfy the interlacing property, i.e.,*

$$d_1 > \bar{d}_1 > d_2 > \bar{d}_2 > \ldots > d_K > \bar{d}_K. \tag{5.3}$$

- *The eigenvector* $\bar{\boldsymbol{u}}_k$ *associated with the eigenvalue* $\bar{d}_k$ *is a multiple of* $\left(\boldsymbol{D} - \bar{d}_k\boldsymbol{I}_K\right)^{-1}\boldsymbol{z}$

Proof of Theorem 5.1. The proof follows from [GVL13, Th.8.4.3.] with trivial modification to handle the complex-valued vector case. □

We remark that the assumptions in Theorem 5.1 do not include the special cases of repeated elements on the diagonal matrix $\boldsymbol{D}$ or zero-valued entries in the vector $\boldsymbol{z}$. Nevertheless, under such specific circumstances, some eigenvalues of the matrix $\boldsymbol{D} - \bar{\rho}\boldsymbol{z}\boldsymbol{z}^{\mathrm{H}}$ are determined directly without any additional computation. More details on these particular cases are provided in Appendix F. Based on Theorem 5.1, rooting the rational function $R(x)$ in (5.2) is of great importance for the acceleration of our proposed partial relaxation approach. Due to the structure of the rational function $R(x)$ in (5.2) and the interlacing property in (5.3), the roots $\bar{d}_k$ for $k = 1, \ldots, K$ of the

rational function $R(x)$ can be determined independently of each other, thus allowing further improvement in the execution time through parallel computing. Without loss of generality, consider the k-th root $\bar{d}_k$ which lies inside the interval (d_{k+1}, d_k) with $k = 1, \ldots, K$ and $d_{K+1} = -\infty$. By defining the two auxiliary rational functions:

$$\psi_k(x) = -\bar{\rho} \sum_{l=1}^{k} \frac{|z_l|^2}{d_l - x} \tag{5.4}$$

$$\phi_k(x) = \begin{cases} -\bar{\rho} \sum\limits_{l=k+1}^{K} \dfrac{|z_l|^2}{d_l - x} & \text{if } 1 \leq k \leq K-1 \\ 0 & \text{if } k = K, \end{cases} \tag{5.5}$$

the equation for determining the root $\bar{d}_k$ in (5.2) is rewritten as

$$-\psi_k(x) = 1 + \phi_k(x). \tag{5.6}$$

Since both $\psi_k(x)$ and $\phi_k(x)$ are defined as the sum of multiple rational functions, a straightforward approach to solve (5.6) iteratively from a given point $x^{(\tau)}$ is using rational functions of first degree $\bar{\psi}_k(x)$ and $\bar{\phi}_k(x)$, respectively, as approximants. The author in [Li94, Subsec. 2.2.3] suggests the approximant of type

$$\bar{R}_{k;p,q}(x) = \begin{cases} p + \dfrac{q}{d_k - x} & \text{if } 1 \leq k \leq K \\ 0 & \text{if } k = K+1, \end{cases} \tag{5.7}$$

and choosing the parameters p and q such that the approximants coincide at a given point $x^{(\tau)}$ with the corresponding exact functions in (5.4) and (5.5), respectively, up to the first-order derivative. Different choices of the approximants were also introduced and discussed in [Li94], [Mel97]. For convenience purposes, the steps for determining the roots of the rational function $R(x)$ in (5.2) are summarized in Algorithm 2. Since the approximant in (5.7) is a rational function of degree one, Steps 3-5 in Algorithm 2 are solved in closed-form as given in Appendix G. As a consequence, the complexity of each iteration τ is $O(K)$, and the overall complexity of the eigenvalue decomposition procedure is $O(KLI)$, where L is the number of required eigenvalues, and I is the number of iterations required for the convergence of Algorithm 2. Note that Algorithm 2 converges quadratically [Li94], and when applied to the PR methods on a fine grid, the eigenvalues computed at one direction θ can be used as an initialization point for the eigenvalues at the next direction, which further reduces the number of iterations I required for convergence. From the simulation results, the number of iterations I required for the tolerance $\epsilon = 10^{-9}$ is less than 4 iterations. Therefore, we can assume that the complexity of the computational eigenvalue decomposition using Algorithm 2 is of order $O(KL)$.

Algorithm 2 Determining the k-th root of the rational function $R(x)$ in (5.2)

1: **Initialization**: Iteration index $\tau = 0$, arbitrary starting value $x^{(0)} \in (d_{k+1}, d_k)$, tolerance ϵ
2: **repeat**
3: Find the parameters p and q such that

$$\bar{R}_{k;p,q}(x^{(\tau)}) = \psi_k(x^{(\tau)}) \tag{5.8}$$

$$\frac{\mathrm{d}\bar{R}_{k;p,q}}{\mathrm{d}x}(x^{(\tau)}) = \frac{\mathrm{d}\psi_k}{\mathrm{d}x}(x^{(\tau)}) \tag{5.9}$$

4: Find the parameters r and s such that

$$\bar{R}_{k+1;r,s}(x^{(\tau)}) = \phi_k(x^{(\tau)}) \tag{5.10}$$

$$\frac{\mathrm{d}\bar{R}_{k+1;r,s}}{\mathrm{d}x}(x^{(\tau)}) = \frac{\mathrm{d}\phi_k}{\mathrm{d}x}(x^{(\tau)}) \tag{5.11}$$

5: Find $x^{(\tau+1)} \in (d_{k+1}, d_k)$ which satisfies

$$-\bar{R}_{k;p,q}(x^{(\tau+1)}) = 1 + \bar{R}_{k+1;r,s}(x^{(\tau+1)}) \tag{5.12}$$

6: $\tau \leftarrow \tau + 1$
7: **until** $\left|x^{(\tau+1)} - x^{(\tau)}\right| < \epsilon$
8: **return** $\bar{d}_k = x^{(\tau+1)}$

5.3 Application of the Acceleration Technique to PR Estimators

The procedure in Section 5.2 which allows specific eigenvalues to be computed in parallel is the starting point to accelerate the computation of the null-spectra of all proposed PR estimators in Chapter 4. In the following sections, more details on the integration of the efficient implementation in Algorithm 2 to the computation of the null-spectra expressions in (4.8), (4.10), (4.18) and Algorithm 1 are presented.

5.3.1 Efficient Implementation of PR-DML

As mentioned in Section 5.2, the complexity of evaluating the null-spectrum is proportional to the number of required eigenvalues L. Therefore, to accelerate the computation of the null-spectrum of the PR-DML method, the number of computed eigenvalues

should be as low as possible. If the number of source signals N satisfies $N \leq M/2+1$, similarly to the reformulation of the MUSIC estimator in [SN89, Eq.(3.7b)], the expression in (4.8) is rewritten as follows:

$$\begin{aligned}\sum_{k=N}^{M} \lambda_k\left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{R}}\right) &= \operatorname{tr}\left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{R}}\right) - \sum_{k=1}^{N-1} \lambda_k\left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{R}}\right) \\ &= \operatorname{tr}\left(\hat{\boldsymbol{R}}\right) - \frac{\boldsymbol{a}^{\mathrm{H}}\hat{\boldsymbol{R}}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{H}}\boldsymbol{a}} - \sum_{k=1}^{N-1} \lambda_k\left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{R}}\right).\end{aligned} \tag{5.13}$$

Using the reformulation in (5.13), only $(N-1)$ - eigenvalues out of M eigenvalues of $\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{R}}$ are computed, and therefore the computational complexity is reduced. In order to apply the eigenvalue decomposition procedure presented in Section 5.2, the term $\lambda_k\left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{R}}\right)$ is further rewritten as follows:

$$\begin{aligned}\lambda_k\left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{R}}\right) &= \lambda_k\left(\hat{\boldsymbol{R}}^{1/2}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{R}}^{1/2}\right) \\ &= \lambda_k\left(\hat{\boldsymbol{R}} - \frac{1}{||\boldsymbol{a}||_2^2}\hat{\boldsymbol{R}}^{1/2}\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\hat{\boldsymbol{R}}^{1/2}\right) \\ &= \lambda_k\left(\hat{\boldsymbol{\Lambda}} - \frac{1}{||\boldsymbol{a}||_2^2}\hat{\boldsymbol{\Lambda}}^{1/2}\hat{\boldsymbol{U}}^{\mathrm{H}}\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\hat{\boldsymbol{U}}\hat{\boldsymbol{\Lambda}}^{1/2}\right),\end{aligned} \tag{5.14}$$

with $\hat{\boldsymbol{\Lambda}}$ and $\hat{\boldsymbol{U}}$ defined in (2.8). From the expression in (5.14), the eigenvalue decomposition procedure introduced in Section 5.2 is applied with $\boldsymbol{D} = \hat{\boldsymbol{\Lambda}}$, $\bar{\rho} = \dfrac{1}{||\boldsymbol{a}||_2^2}$ and $\boldsymbol{z} = \hat{\boldsymbol{\Lambda}}^{1/2}\hat{\boldsymbol{U}}^{\mathrm{H}}\boldsymbol{a}$. From the computational perspective, except for the initial full eigenvalue decomposition in (2.9), the overall complexity of the calculation of the null-spectrum for the complete Field-of-View Θ with N_G directions is $O\left(\left(M^2 + M\left(N-1\right)\right)N_G\right) = O(M^2 N_G)$. This complexity is higher than the complexity required for computing the MUSIC null-spectrum, which is $O(MNN_G)$.

5.3.2 Efficient Implementation of PR-WSF

A similar iterative procedure for computing the eigenvalue decomposition as proposed in Section 5.2 and Section 5.3.1 can be applied directly to the PR-WSF method presented in Section 4.2.2. However, the computational complexity of the PR-WSF method is further reduced by exploiting the fact that all eigenvalues $\lambda_k(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{U}}_{\mathrm{s}}\boldsymbol{W}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}})$ for $k = N+1, \ldots, M$ are equal to zero since $\operatorname{rank}(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{U}}_{\mathrm{s}}\boldsymbol{W}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}}) \leq N$. Therefore, only the N-th eigenvalue $\lambda_N(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{U}}_{\mathrm{s}}\boldsymbol{W}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}})$ needs to be calculated. Furthermore, the

dimension of the matrix $\boldsymbol{D}$ in (5.1) can also be reduced. In fact, in accordance with (5.14), it can be shown that

$$\lambda_N\left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{U}}_{\mathrm{s}}\boldsymbol{W}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}}\right)=\lambda_N\left(\boldsymbol{W}-\frac{1}{||\boldsymbol{a}||_2^2}\boldsymbol{W}^{1/2}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}}\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\hat{\boldsymbol{U}}_{\mathrm{s}}\boldsymbol{W}^{1/2}\right). \tag{5.15}$$

Using the identity in (5.15), the procedure for computing the eigenvalue decomposition introduced in Section 5.2 is applied with $\boldsymbol{D}=\boldsymbol{W}$, $\bar{\rho}=\dfrac{1}{||\boldsymbol{a}||_2^2}$ and $\boldsymbol{z}=\boldsymbol{W}^{1/2}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}}\boldsymbol{a}$. Since the dimension of the matrix is reduced from $M\times M$ to $N\times N$, and only a single eigenvalue is required, the complexity reduces to $O\left((NM+N-1)\,N_G\right)=O(MNN_G)$, which is identical to the order of the computational complexity of the MUSIC algorithm. However, the computational overhead associated with PR-WSF is still higher than MUSIC since in the preprocessing step, additional calculations for determining the weighting matrix $\boldsymbol{D}=\boldsymbol{W}$ and the vector $\boldsymbol{z}=\boldsymbol{W}^{1/2}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}}\boldsymbol{a}$ are required.

5.3.3 Efficient Implementation of PR-CCF

The expression of the PR-CCF null-spectrum in (4.18) resembles the generic formulation of the rank-one modified Hermitian matrix in (5.1), except for the fact that the matrix $\hat{\boldsymbol{R}}$ is generally not diagonal. Therefore, the application of the eigenvalue decomposition in Section 5.2 is straightforward, in which we perform an orthogonal transformation on $\hat{\boldsymbol{R}}$ and $\boldsymbol{a}$ to diagonalize $\hat{\boldsymbol{R}}$. However, the number of the eigenvalues required for the computation of the null-spectrum is $(M-N+1)$, which is typically larger than the number of sources N. By rewriting the expression in (4.18) using the trace operator in accordance with (5.13), only the $(N-1)$-largest eigenvalues are calculated and the null-spectrum of the PR-CCF method is rewritten in the following form to utilize the principal eigenvalues:

$$\begin{aligned}\sum_{k=N}^{M}\lambda_k^2\left(\hat{\boldsymbol{R}}-\hat{\sigma}_{\mathrm{s,\,C}}^2\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\right)&=\sum_{k=N}^{M}\lambda_k\left(\left(\hat{\boldsymbol{R}}-\hat{\sigma}_{\mathrm{s,\,C}}^2\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\right)^2\right)\\&=\mathrm{tr}\left(\left(\hat{\boldsymbol{R}}-\hat{\sigma}_{\mathrm{s,\,C}}^2\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\right)^2\right)-\sum_{k=1}^{N-1}\lambda_k^2\left(\hat{\boldsymbol{R}}-\hat{\sigma}_{\mathrm{s,\,C}}^2\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\right)\\&=\mathrm{tr}\left(\hat{\boldsymbol{R}}^2\right)-2\hat{\sigma}_{\mathrm{s,\,C}}^2\boldsymbol{a}^{\mathrm{H}}\hat{\boldsymbol{R}}\boldsymbol{a}+\hat{\sigma}_{\mathrm{s,\,C}}^4\,||\boldsymbol{a}||_2^4-\sum_{k=1}^{N-1}\lambda_k^2\left(\hat{\boldsymbol{R}}-\hat{\sigma}_{\mathrm{s,\,C}}^2\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\right)\end{aligned} \tag{5.16}$$

Considering the formulation in (5.16), we observe that the PR-CCF method involves both the conventional beamformer in the form of (3.3) and Capon beamformer in the form of (3.6) for the evaluation of the null-spectrum. Similarly to the PR-DML

method, following the derivation in (5.14) for any eigenvalue $\lambda_k\left(\hat{\boldsymbol{R}}-\hat{\sigma}^2_{\mathrm{s,\,C}}\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\right)$, it can be shown that

$$\lambda_k\left(\hat{\boldsymbol{R}}-\hat{\sigma}^2_{\mathrm{s,\,C}}\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\right)=\lambda_k\left(\hat{\boldsymbol{\Lambda}}-\hat{\sigma}^2_{\mathrm{s,\,C}}\hat{\boldsymbol{U}}^{\mathrm{H}}\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\hat{\boldsymbol{U}}\right). \tag{5.17}$$

From (5.16) and (5.17), we apply the eigenvalue decomposition procedure presented in Section 5.2 with $\boldsymbol{D}=\hat{\boldsymbol{\Lambda}}$, $\bar{\rho}=\hat{\sigma}^2_{\mathrm{s,\,C}}$ and $\boldsymbol{z}=\hat{\boldsymbol{U}}^{\mathrm{H}}\boldsymbol{a}$. Thus, the overall computational complexity of the PR-CCF algorithm is $O\left(\left(M^2+M\left(N-1\right)\right)N_G\right)=O(M^2N_G)$.

5.3.4 Efficient Implementation of PR-UCF

Unlike the PR-DML, PR-WSF and PR-CCF estimators, the PR-UCF estimator requires additional steps of calculating the derivative $g'(\sigma^2_{\mathrm{s}})$ in (4.21) to obtain the minimizer $\sigma^2_{\mathrm{s,\,U}}$ of (4.20). To reduce the number of required eigenvalues for computing the derivative and the null-spectrum, the function $g(\sigma^2_{\mathrm{s}})=\sum\limits_{k=N}^{M}\lambda^2_k\left(\hat{\boldsymbol{R}}-\sigma^2_{\mathrm{s}}\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\right)$ is rewritten similarly to (5.16) as follows:

$$g(\sigma^2_{\mathrm{s}})=\mathrm{tr}\left(\hat{\boldsymbol{R}}^2\right)-2\sigma^2_{\mathrm{s}}\boldsymbol{a}^{\mathrm{H}}\hat{\boldsymbol{R}}\boldsymbol{a}+\sigma^4_{\mathrm{s}}\left|\left|\boldsymbol{a}\right|\right|^4_2-\sum_{k=1}^{N-1}\lambda^2_k\left(\hat{\boldsymbol{R}}-\sigma^2_{\mathrm{s}}\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\right). \tag{5.18}$$

The derivative $\dfrac{\mathrm{d}g}{\mathrm{d}\sigma^2_{\mathrm{s}}}\left(\sigma^2_{\mathrm{s}}\right)$ is calculated as

$$\frac{\mathrm{d}g}{\mathrm{d}\sigma^2_{\mathrm{s}}}\left(\sigma^2_{\mathrm{s}}\right)=-2\boldsymbol{a}^{\mathrm{H}}\hat{\boldsymbol{R}}\boldsymbol{a}+2\sigma^2_{\mathrm{s}}\left|\left|\boldsymbol{a}\right|\right|^4_2+\sum_{k=1}^{N-1}\frac{2\bar{\lambda}_k(\sigma^2_{\mathrm{s}})}{\sigma^4_{\mathrm{s}}\boldsymbol{a}^{\mathrm{H}}\left(\hat{\boldsymbol{R}}-\bar{\lambda}_k(\sigma^2_{\mathrm{s}})\boldsymbol{I}_M\right)^{-2}\boldsymbol{a}}, \tag{5.19}$$

where $\bar{\lambda}_k(\sigma^2_{\mathrm{s}})=\lambda_k\left(\hat{\boldsymbol{R}}-\sigma^2_{\mathrm{s}}\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\right)$. By substituting $\boldsymbol{z}=\hat{\boldsymbol{U}}^{\mathrm{H}}\boldsymbol{a}$, we obtain

$$\bar{\lambda}_k(\sigma^2_{\mathrm{s}})=\lambda_k\left(\hat{\boldsymbol{\Lambda}}-\sigma^2_{\mathrm{s}}\boldsymbol{z}\boldsymbol{z}^{\mathrm{H}}\right), \tag{5.20}$$

$$\frac{\mathrm{d}g}{\mathrm{d}\sigma^2_{\mathrm{s}}}\left(\sigma^2_{\mathrm{s}}\right)=-2\boldsymbol{z}^{\mathrm{H}}\hat{\boldsymbol{\Lambda}}\boldsymbol{z}+2\sigma^2_{\mathrm{s}}\left|\left|\boldsymbol{z}\right|\right|^4_2+\sum_{k=1}^{N-1}\frac{2\bar{\lambda}_k\left(\sigma^2_{\mathrm{s}}\right)}{\sigma^4_{\mathrm{s}}\sum\limits_{j=1}^{M}\dfrac{\left|z_j\right|^2}{\left(\hat{\lambda}_j-\bar{\lambda}_k\left(\sigma^2_{\mathrm{s}}\right)\right)^2}}. \tag{5.21}$$

Based on the expressions in (5.20) and (5.21), the null-spectrum of PR-UCF from Algorithm 1 is calculated by applying the procedure in Section 5.2 with $\boldsymbol{D}=\hat{\boldsymbol{\Lambda}}$, $\bar{\rho}=\sigma^2_{\mathrm{s},0}$ and $\boldsymbol{z}=\hat{\boldsymbol{U}}^{\mathrm{H}}\boldsymbol{a}$. The computational complexity of the PR-UCF method is therefore of order $O(M^2N_GN_I)$ where N_I is the number of bisection steps conducted in Algorithm 1.

5.4 Summary

In this chapter, a numerical procedure for efficiently computing the null-spectra of the proposed PR methods in Chapter 4 efficiently has been introduced. Instead of performing the full eigenvalue decomposition in each look-direction in the spectral search, the common underlying structure in the matrix argument from which the eigenvalues are computed is exploited. We show that the eigenvalues of the direction-depending matrix argument are identical to that of a constant diagonal matrix modified by a rank-one component. This particular structure is suitable for iteratively computing the selected eigenvalues in parallel. Furthermore, the convergence rate of the iterative algorithm is quadratic, which further reduces the computational workload in computing the null-spectra of the proposed methods. By adopting the efficient implementation procedure to the estimators under the PR framework, the PR-WSF estimator possesses the lowest computational complexity among the proposed PR methods, which is in the same order of magnitude as that of the MUSIC estimator. PR-UCF has the highest computational complexity among the proposed PR methods due to the additional iterative search to compute the signal power in each look-direction. Later in Chapter 7, the aforementioned improvement in the computational complexity of the proposed implementation is verified by numerical experiments.

Chapter 6

Asymptotic Performance Bound for PR Estimators

6.1 Motivation

Since the mean-square error (MSE) is the most common performance indicator in DOA estimation, the statistical analysis of parametric estimators has been the subject of investigation in the literature [GVL13]. Generally, the statistical distributions of the DOA estimates can only be obtained or approximated with high accuracy in specific cases [LLV93]. In the literature, the asymptotic case, in which the number of snapshots T tends to infinity, is the most investigated regime in which the performance of parametric estimators is studied. Based on the Central Limit Theorem, as the number of snapshots T tends to infinity, the estimates are asymptotically Gaussian distributed [SM14], and the error characteristics of the DOA estimators are fully characterized by the mean and the covariance matrix.

In order to gain further insights into the estimation error behavior of unbiased estimators, the CRB [Cra45, Rao45] has been widely accepted as a benchmark. However, the CRB is model-specific in the sense that it represents a lower bound for the achievable MSE for all unbiased estimators developed under the specific model [Kay93, SM14]. Nevertheless, different estimators exploit particular model properties as prior knowledge while deliberately ignoring others. For example, the relaxation of the steering matrix in the PR framework essentially implies a particular model of the steering matrix according to (4.3) which is different from the true steering matrix in (2.2). Compared with the conventional signal model in Chapter 2, the PR-induced signal model in Section 4.1, from which the PR estimators are derived, contains less information on the angular directions of the source signals. Due to such information loss, the error performance of the PR estimators degrades compared with the multi-dimensional counterparts under the full-information model.

In this chapter, we investigate the degradation in the DOA estimation performance of PR estimators by deriving the asymptotic error performance bound under the PR signal model. First, we revise the conventional stochastic CRB for the DOA parameters developed under the conventional signal model in Chapter 2. Then, the CRB for the

induced PR signal model, which is referred to as PR-CRB, is derived. From the closed-form expression of the PR-CRB, we quantify the theoretical performance loss induced by the relaxation of the array manifold. Furthermore, the theoretical asymptotic mean-square errors of two PR estimators, namely PR-DML and PR-WSF, are derived and compared with results in the literature.

6.2 Conventional Stochastic Cramér-Rao Bound for DOA parameters $\boldsymbol{\theta}$

Under the assumptions (As.1)-(As.4), the DOA estimation problem formulated in Chapter 2 is a special case of a generic parametric estimation problem with Gaussian distributed observations. In order to derive the conventional stochastic CRB for the DOA parameters $\boldsymbol{\theta}$, a parameterization of the distribution of the receive signal $\boldsymbol{x}(t)$ in (2.1) is required. As the source signal $\boldsymbol{s}(t)$ and the noise signal $\boldsymbol{n}(t)$ are zero mean and Gaussian distributed, the distribution of the receive signal $\boldsymbol{x}(t)$ is fully characterized by the receive covariance matrix $\boldsymbol{R}$ in (2.5). In the most commonly used model in the literature, the receive covariance matrix $\boldsymbol{R}$ is parameterized as

$$\boldsymbol{\alpha} = \left[\boldsymbol{\theta}^{\mathrm{T}}, \boldsymbol{p}^{\mathrm{T}}, \nu\right]^{\mathrm{T}}, \tag{6.1}$$

where $\boldsymbol{p} \in \mathbb{R}^{N^2}$ contains the N elements on the diagonal of the matrix $\boldsymbol{P}$ and the $(N^2 - N)$ elements characterizing the real and imaginary part of the respective entries on the upper-triangular part of the source covariance matrix $\boldsymbol{P}$ in the following order

$$\begin{aligned}\boldsymbol{p} = \Big[p_{11}, \ldots, p_{NN},\ &\mathrm{Re}\left\{p_{12}\right\}, \ldots, \mathrm{Re}\left\{p_{1N}\right\}, \ldots\ldots, \mathrm{Re}\left\{p_{N-1,N}\right\}, \\ &\mathrm{Im}\left\{p_{12}\right\}, \ldots, \mathrm{Im}\left\{p_{1N}\right\}, \ldots\ldots, \mathrm{Im}\left\{p_{N-1,N}\right\}\Big]^{\mathrm{T}}.\end{aligned} \tag{6.2}$$

The CRB matrix $\boldsymbol{C}(\boldsymbol{\alpha})$ for the parameter vector $\boldsymbol{\alpha}$ is computed as

$$\boldsymbol{C}(\boldsymbol{\alpha}) = \boldsymbol{\mathcal{I}}(\boldsymbol{\alpha})^{-1}. \tag{6.3}$$

In (6.3), the Fisher Information Matrix (FIM) $\boldsymbol{\mathcal{I}}(\boldsymbol{\alpha})$ is computed elementwise using the Slepian-Bangs formula [Sle54, Ban71] as follows:

$$\mathcal{I}_{uv}(\boldsymbol{\alpha}) = T\ \mathrm{tr}\left(\frac{\mathrm{d}\boldsymbol{R}}{\mathrm{d}\alpha_u}\boldsymbol{R}^{-1}\frac{\mathrm{d}\boldsymbol{R}}{\mathrm{d}\alpha_v}\boldsymbol{R}^{-1}\right), \tag{6.4}$$

with α_u denoting the u-th entry in the parameter vector $\boldsymbol{\alpha}$. The CRB matrix in (6.3) exists under certain regularity conditions [LC98, Ch. 6]. In [WZ89, OVSN93, SN90b], the assumptions (As.1)-(As.4) are sufficient conditions to guarantee the existence of

the CRB for the parameter $\boldsymbol{\alpha}$ in (6.1) under the signal model in Chapter 2. We further remark that in the context of DOA estimation, only the DOA parameters $\boldsymbol{\theta}$ in the parameter vector $\boldsymbol{\alpha}$ are of interest while the remaining parameters characterizing the source and noise covariance matrix are considered as nuisance parameters. By partitioning the FIM matrix and applying the block matrix inversion lemma in Appendix A, Equation (A.9), the submatrix corresponding to the DOA parameters $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_N]$, denoted by $\boldsymbol{C}_{\mathrm{CRB}}(\boldsymbol{\theta})$, is computed directly without inverting the full matrix $\boldsymbol{\mathcal{I}}(\boldsymbol{\alpha})$ as in (6.3). The closed-form expression of the block $\boldsymbol{C}_{\mathrm{CRB}}(\boldsymbol{\theta})$ is given by [SN90b, SLG01]

$$\boldsymbol{C}_{\mathrm{CRB}}(\boldsymbol{\theta}) = \frac{\nu}{2T}\mathrm{Re}\left\{\boldsymbol{M} \odot \left(\boldsymbol{D}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}\boldsymbol{D}\right)\right\}^{-1}, \tag{6.5}$$

with $\boldsymbol{M} = \left(\boldsymbol{P}\boldsymbol{A}(\boldsymbol{\theta})^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{P}\right)^{\mathrm{T}}$ and $\boldsymbol{D} = \left[\left.\frac{\mathrm{d}\boldsymbol{a}(\theta)}{\mathrm{d}\theta}\right|_{\theta=\theta_1}, \ldots, \left.\frac{\mathrm{d}\boldsymbol{a}(\theta)}{\mathrm{d}\theta}\right|_{\theta=\theta_N}\right]$. The matrix $\boldsymbol{C}_{\mathrm{CRB}}(\boldsymbol{\theta})$ in (6.5) is referred to in the following as the conventional CRB. It has been shown in [SN90b, SLG01] that the conventional CRB in (6.5) is the lower bound in the positive semidefinite sense for the MSE matrix of any unbiased DOA estimators $\hat{\boldsymbol{\theta}}$ developed under the conventional signal model:

$$\mathbb{E}\left\{\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)^{\mathrm{T}}\right\} \succeq \boldsymbol{C}_{\mathrm{CRB}}(\boldsymbol{\theta}). \tag{6.6}$$

As the DOA estimators under the PR framework are developed from the PR signal model in Section 4.1, their asymptotic error performance may not be correctly characterized by the conventional stochastic CRB in (6.5) due to different prior knowledge of the models. In the next sections, by following the derivation of the conventional CRB, the PR-CRB, which is the CRB for the DOA parameter ϑ under the PR signal model, is derived.

6.3 Parameterization of the PR Signal Model

Similar to the conventional signal model in (2.4), the receive signal $\boldsymbol{X}$ in the PR signal model in (4.2) are samples of a Gaussian distribution with zero mean and covariance matrix $\boldsymbol{R}$ given by

$$\boldsymbol{R} = \begin{bmatrix}\boldsymbol{a}(\vartheta) & \boldsymbol{B}\end{bmatrix}\boldsymbol{P}\begin{bmatrix}\boldsymbol{a}(\vartheta)^{\mathrm{H}} \\ \boldsymbol{B}^{\mathrm{H}}\end{bmatrix} + \nu\boldsymbol{I}. \tag{6.7}$$

As a result, the distribution of the receive signal $\boldsymbol{X}$ can be parameterized by the following vector

$$\boldsymbol{\alpha} = \left[\vartheta, \mathrm{Re}\left\{\mathrm{vec}\left(\boldsymbol{B}\right)\right\}^{\mathrm{T}}, \mathrm{Im}\left\{\mathrm{vec}\left(\boldsymbol{B}\right)\right\}^{\mathrm{T}}, \boldsymbol{p}^{\mathrm{T}}, \nu\right]^{\mathrm{T}}, \tag{6.8}$$

where $\boldsymbol{p}$ is defined as in (6.1). With respect to the parameterization in (6.8), the FIM $\boldsymbol{\mathcal{I}}(\boldsymbol{\alpha})$ for the PR signal model is computed elementwise as

$$\mathcal{I}_{uv}(\boldsymbol{\alpha}) = T \ \mathrm{tr}\left(\frac{\mathrm{d}\boldsymbol{R}}{\mathrm{d}\alpha_u}\boldsymbol{R}^{-1}\frac{\mathrm{d}\boldsymbol{R}}{\mathrm{d}\alpha_v}\boldsymbol{R}^{-1}\right). \tag{6.9}$$

However, the corresponding CRB matrix cannot be obtained directly by inverting the FIM in (6.9) as stated in the following theorem.

Theorem 6.1. *The FIM in* (6.9) *under the PR signal model with respect to the parameterization in* (6.8) *is singular.*

Proof of Theorem 6.1. In order to show that the FIM is singular, it is sufficient to show that the parameterization in (6.8) is not locally identifiable [Rot71, Th. 1]. In other words, the parameter vector $\boldsymbol{\alpha}$ in (6.8) cannot be uniquely determined in any sufficiently small neighborhood of the true parameter vector, even if the true covariance matrix $\boldsymbol{R}$ in (6.7) is available. For the notational simplicity, we assume that the covariance matrix $\boldsymbol{R}$ is generated according to (6.7) by the true generator $\{\boldsymbol{A}(\vartheta), \boldsymbol{P}, \nu\}$ with a relaxed steering matrix $\boldsymbol{A}(\vartheta) = [\boldsymbol{a}(\vartheta), \boldsymbol{B}] \in \bar{\mathcal{A}}_N$ and a positive semidefinite source covariance matrix $\boldsymbol{P}$. Define the scaling matrix $\tilde{\boldsymbol{T}}$ as

$$\tilde{\boldsymbol{T}} = \begin{bmatrix} 1 & \boldsymbol{0}^{\mathrm{T}} \\ \boldsymbol{0} & (1+\delta)\boldsymbol{I}_{N-1} \end{bmatrix} \tag{6.10}$$

where δ is an arbitrary positive scalar. Clearly, by choosing a sufficiently small value for δ, the covariance matrix $\boldsymbol{R}$ admits another representation by the generator $\left\{\boldsymbol{A}(\vartheta)\tilde{\boldsymbol{T}}, \tilde{\boldsymbol{T}}^{-1}\boldsymbol{P}\tilde{\boldsymbol{T}}^{-\mathrm{H}}, \nu\right\}$ in every arbitrarily small neighborhood around the true generator $\{\boldsymbol{A}(\vartheta),\ \boldsymbol{P},\ \nu\}$. From the aforementioned observation and according to [Rot71, Th. 1], the parameterization in (6.8) is not locally identifiable. Thus, the FIM is singular, and we conclude the proof. □

From the proof of Theorem 6.1, we observe that modeling the unstructured $\boldsymbol{B}$ matrix by $M \times (N-1)$ entries over-parameterizes the model in (6.7) and results in a singular FIM. In order to derive the CRB corresponding to the DOA parameter in the case of non-invertible FIM, one common approach is to replace the inverse with the Moore-Penrose pseudo-inverse [SM01, LY12]. Hence, the CRB of the desired DOA parameter ϑ under the PR signal model, which is denoted by $C_{\text{PR-CRB}}(\vartheta)$, is computed from (6.9) as

$$C_{\text{PR-CRB}}(\vartheta) = \left[\boldsymbol{\mathcal{I}}(\boldsymbol{\alpha})^{\dagger}\right]_{11}, \tag{6.11}$$

where the notation $[\cdot]_{11}$ refers to the upper-left entry of the matrix argument. Although the CRB in (6.11) is well-defined, the derivation of the analytic closed-form expression from the pseudo-inverse appears to be difficult.

Another common approach is to resolve the ambiguity induced by the relaxation by introducing a non-redundant parameterization of the PR signal model in (4.2). The proposed parameterization essentially removes the ambiguity by incorporating additional constraints [SB98, ZJLJ11] on the relaxed steering matrix $\boldsymbol{A}(\vartheta) \in \bar{\mathcal{A}}_N$. We emphasize that the non-redundant parameterization introduced in this section is not practically useful for estimating the DOA parameter as the parameterization unnecessary complicates the optimization procedure. However, it is crucial in the CRB analysis to ensure the invertibility of the FIM. Furthermore, it can be shown that the CRB in (6.11) can be retrieved from the inverse of the FIM corresponding to the non-redundant parameterization. The key observation which results in the derivation of the non-redundant parameterization is as follows: the condition $\operatorname{rank}(\boldsymbol{A}(\vartheta)) = N$ with $\boldsymbol{A}(\vartheta) \in \bar{\mathcal{A}}_N$ implies that there exists an invertible minor square matrix of $\boldsymbol{A}(\vartheta)$ with dimension $N \times N$. By re-indexing the sensors if necessary, we assume that, for the following matrix partition:

$$\boldsymbol{A}(\vartheta) = \begin{bmatrix} a_1(\vartheta) & \boldsymbol{b}_1^{\mathrm{T}} \\ \boldsymbol{a}_2(\vartheta) & \boldsymbol{B}_2 \\ \boldsymbol{a}_3(\vartheta) & \boldsymbol{B}_3 \end{bmatrix} \tag{6.12}$$

with $\boldsymbol{a}_2(\vartheta) \in \mathbb{C}^{M-N}$, $\boldsymbol{a}_3(\vartheta) \in \mathbb{C}^{(N-1)}$, $\boldsymbol{b}_1 \in \mathbb{C}^{N-1}$, $\boldsymbol{B}_2 \in \mathbb{C}^{(M-N)\times(N-1)}$ and $\boldsymbol{B}_3 \in \mathbb{C}^{(N-1)\times(N-1)}$, the minor block matrix $\begin{bmatrix} a_1(\vartheta) & \boldsymbol{b}_1^{\mathrm{T}} \\ \boldsymbol{a}_3(\vartheta) & \boldsymbol{B}_3 \end{bmatrix}$ is invertible with $a_1(\vartheta) \neq 0$. With the shorthand notation $\boldsymbol{a} = \boldsymbol{a}(\vartheta) = \left[a_1, \boldsymbol{a}_2^{\mathrm{T}}, \boldsymbol{a}_3^{\mathrm{T}}\right]^{\mathrm{T}}$, it follows from the property (A.11) of the Schur complement in Appendix A that the inverse $\left(\boldsymbol{B}_3 - \boldsymbol{a}_3 a_1^{-1} \boldsymbol{b}_1^{\mathrm{T}}\right)^{-1}$ exists. This remark in turn ensures the existence of the transformation matrix $\boldsymbol{T}$ and its inverse $\boldsymbol{T}^{-1}$, which are defined as

$$\boldsymbol{T} = \begin{bmatrix} 1 & -a_1^{-1}\boldsymbol{b}_1^{\mathrm{T}} \left(\boldsymbol{B}_3 - \boldsymbol{a}_3 a_1^{-1}\boldsymbol{b}_1^{\mathrm{T}}\right)^{-1} \\ \boldsymbol{0} & \left(\boldsymbol{B}_3 - \boldsymbol{a}_3 a_1^{-1}\boldsymbol{b}_1^{\mathrm{T}}\right)^{-1} \end{bmatrix}, \tag{6.13}$$

$$\boldsymbol{T}^{-1} = \begin{bmatrix} 1 & a_1^{-1}\boldsymbol{b}_1^{\mathrm{T}} \\ \boldsymbol{0} & \boldsymbol{B}_3 - \boldsymbol{a}_3 a_1^{-1}\boldsymbol{b}_1^{\mathrm{T}} \end{bmatrix}, \tag{6.14}$$

respectively. The signal model in (4.2) is rewritten as

$$\boldsymbol{X} = \bar{\boldsymbol{A}}\bar{\boldsymbol{S}} + \boldsymbol{N}, \tag{6.15}$$

where the transformed steering matrix $\bar{\boldsymbol{A}}$ and the transformed source signal $\bar{\boldsymbol{S}}$ are respectively given by

$$\bar{\boldsymbol{A}} = \boldsymbol{A}(\vartheta)\boldsymbol{T} = \begin{bmatrix} a_1 & \boldsymbol{0}^{\mathrm{T}} \\ \boldsymbol{a}_2 & \bar{\boldsymbol{B}} \\ \boldsymbol{a}_3 & \boldsymbol{I}_{N-1} \end{bmatrix} \tag{6.16}$$

$$\bar{\boldsymbol{S}} = \boldsymbol{T}^{-1}\boldsymbol{S}, \tag{6.17}$$

and the minor matrix $\bar{\boldsymbol{B}} \in \mathbb{C}^{(M-N)\times(N-1)}$ is given by

$$\bar{\boldsymbol{B}} = \left(\boldsymbol{B}_2 - \boldsymbol{a}_2 a_1^{-1} \boldsymbol{b}_1^{\mathrm{T}}\right)\left(\boldsymbol{B}_3 - \boldsymbol{a}_3 a_1^{-1} \boldsymbol{b}_1^{\mathrm{T}}\right)^{-1}. \tag{6.18}$$

As a result, any relaxed steering matrix $\boldsymbol{A}(\vartheta) \in \bar{\mathcal{A}}_N$ in (4.2) is uniquely characterized by the desired DOA ϑ, the minor matrix $\bar{\boldsymbol{B}}$ and the transformation matrix $\boldsymbol{T}$. Hence, the covariance matrix $\boldsymbol{R}$ of the receive signal $\boldsymbol{x}(t)$ according to the PR signal model is given by

$$\boldsymbol{R} = \bar{\boldsymbol{A}} \bar{\boldsymbol{P}} \bar{\boldsymbol{A}}^{\mathrm{H}} + \nu \boldsymbol{I}_M, \tag{6.19}$$

where $\bar{\boldsymbol{A}}$ is defined in (6.16) and

$$\bar{\boldsymbol{P}} = \boldsymbol{T}^{-1} \boldsymbol{P} \boldsymbol{T}^{-\mathrm{H}} \tag{6.20}$$

is the transformed source covariance matrix. From (6.16) and (6.20), we introduce the following parameterization for the covariance matrix $\boldsymbol{R}$ for the transformed PR model in (6.19):

$$\bar{\boldsymbol{\alpha}} = \left[\vartheta, \check{\boldsymbol{b}}^{\mathrm{T}}, \tilde{\boldsymbol{b}}^{\mathrm{T}}, \bar{\boldsymbol{p}}^{\mathrm{T}}, \nu\right]^{\mathrm{T}} \in \mathbb{R}^{\left(2+2(M-N)(N-1)+N^2\right)} \tag{6.21}$$

where $\bar{\boldsymbol{\alpha}}$ is the parameter vector characterizing the receive covariance matrix under the PR signal model. The vector $\check{\boldsymbol{b}}$, which contains the real part of the minor matrix $\bar{\boldsymbol{B}}$ in (6.16), is given by

$$\begin{aligned} \check{\boldsymbol{b}} &= \mathrm{Re}\left\{\mathrm{vec}\left(\bar{\boldsymbol{B}}\right)\right\} \in \mathbb{R}^{(M-N)(N-1)} \\ &= \left[\check{b}_{(1,1)}, \ldots, \check{b}_{(M-N,1)}, \ldots, \check{b}_{(1,N-1)}, \ldots, \check{b}_{(M-N,N-1)}\right]^{\mathrm{T}} \end{aligned} \tag{6.22}$$

with $\check{b}_{(k,l)} = \mathrm{Re}\left\{\bar{B}_{kl}\right\}$. In the same manner, we define $\tilde{b}_{(k,l)} = \mathrm{Im}\left\{\bar{B}_{kl}\right\}$ and

$$\begin{aligned} \tilde{\boldsymbol{b}} &= \mathrm{Im}\left\{\mathrm{vec}\left(\bar{\boldsymbol{B}}\right)\right\} \in \mathbb{R}^{(M-N)(N-1)} \\ &= \left[\tilde{b}_{(1,1)}, \ldots, \tilde{b}_{(M-N,1)}, \ldots, \tilde{b}_{(1,N-1)}, \ldots, \tilde{b}_{(M-N,N-1)}\right]^{\mathrm{T}} \end{aligned} \tag{6.23}$$

contains the imaginary part of the entries in the minor matrix $\bar{\boldsymbol{B}}$. Likewise, $\bar{\boldsymbol{p}} \in \mathbb{R}^{N^2}$ contains the N-elements $\left\{\bar{P}_{ii}\right\}$ for $1 \leq i \leq N$ and $N(N-1)$ elements $\left\{\mathrm{Re}\left\{\bar{P}_{iq}\right\}, \mathrm{Im}\left\{\bar{P}_{iq}\right\}\right\}$ for $1 \leq i < q \leq N$, which characterize the transformed source covariance matrix $\bar{\boldsymbol{P}}$. The ordering of the parameters in $\bar{\boldsymbol{p}}$ is similar to the ordering in (6.2). The element in the u-th row and v-th column of the FIM $\boldsymbol{\mathcal{I}}(\bar{\boldsymbol{\alpha}})$ for $u, v = 1, \ldots, 2 + 2(M-N)(N-1) + N^2$ is given by [SLG01, Eq. (4)]

$$\mathcal{I}_{uv}(\bar{\boldsymbol{\alpha}}) = T \ \mathrm{tr}\left(\frac{\mathrm{d}\boldsymbol{R}}{\mathrm{d}\bar{\alpha}_u} \boldsymbol{R}^{-1} \frac{\mathrm{d}\boldsymbol{R}}{\mathrm{d}\bar{\alpha}_v} \boldsymbol{R}^{-1}\right). \tag{6.24}$$

Note that in the parameterization (6.21), the parameter of interest is the DOA ϑ. As a consequence, the CRB of the parameter ϑ under the transformed signal model in (6.15) is given by the upper-left entry of the matrix $\boldsymbol{C} = \boldsymbol{\mathcal{I}}(\bar{\boldsymbol{\alpha}})^{-1}$. Nevertheless, the

FIM $\boldsymbol{\mathcal{I}}$ in (6.24) must be invertible for the existence of the CRB matrix. Furthermore, we must ensure that the parameterization in (6.21) does not alter the CRB of the desired parameter ϑ. In order to resolve the aforementioned concerns, we introduce in the following three additional assumptions:

(As.5) *The FIM defined in* (6.24) *has constant rank in a neighborhood of the true parameter* $\bar{\boldsymbol{\alpha}}$.

(As.6) *The source covariance matrix* $\boldsymbol{P} = \mathbb{E}\left\{\boldsymbol{s}(t)\boldsymbol{s}(t)^{\mathrm{H}}\right\}$ *is positive definite, i.e.,* $\operatorname{rank}(\boldsymbol{P}) = N$.

(As.7) *The set* $\mathbb{S} = \{\theta \in \Theta \,|\boldsymbol{a}\,(\theta) \in \operatorname{span}\,(\boldsymbol{U}_{\mathrm{s}})\}$ *with* $\boldsymbol{U}_s$ *defined in* (2.6) *contains only a finite number of elements.*

Then, the following theorem provides sufficient conditions for the invertibility of the FIM in (6.24).

Theorem 6.2. *Assume the PR signal model in* (4.2). *If (As.1)-(As.3) and (As.5)-(As.7) hold, then the FIM* $\boldsymbol{\mathcal{I}}(\bar{\boldsymbol{\alpha}})$ *in* (6.24) *is invertible, and therefore the CRB for the parameter vector* $\bar{\boldsymbol{\alpha}}$ *in* (6.21) *exists.*

Proof of Theorem 6.2. See Appendix H. □

From the sufficient conditions for the existence of the CRB as stated in Theorem 6.2, an alternative method to compute the PR-CRB in (6.11) based on the parameterization in (6.21) is provided in the following theorem.

Theorem 6.3. *Assume that the FIM matrix in* (6.21) *is invertible. Then the CRB of the parameter* ϑ *under the original PR signal model in* (4.2) *and that under the transformed PR signal model in* (6.15) *are identical, i.e.,*

$$C_{\text{PR-CRB}}(\vartheta) = \left[\boldsymbol{\mathcal{I}}(\boldsymbol{\alpha})^{\dagger}\right]_{11} = \left[\boldsymbol{\mathcal{I}}(\bar{\boldsymbol{\alpha}})^{-1}\right]_{11}, \tag{6.25}$$

where the parameter vectors $\boldsymbol{\alpha}$ *and* $\bar{\boldsymbol{\alpha}}$ *are defined in* (6.8) *and* (6.21), *respectively.*

Proof of Theorem 6.3. In order to show that the parameterization in (6.21) does not alter the CRB of the desired DOA parameter ϑ under the PR signal model, using the formula of the CRB for the parameter transformations [SM01, Eq. (16)] [Kay93], we obtain the following identity:

$$\boldsymbol{\mathcal{I}}(\boldsymbol{\alpha})^{\dagger} = \left(\frac{\partial \boldsymbol{\alpha}^{\mathrm{T}}}{\partial \bar{\boldsymbol{\alpha}}}\right) \boldsymbol{\mathcal{I}}(\bar{\boldsymbol{\alpha}})^{-1} \left(\frac{\partial \boldsymbol{\alpha}^{\mathrm{T}}}{\partial \bar{\boldsymbol{\alpha}}}\right)^{\mathrm{T}}, \tag{6.26}$$

where the entries of the FIMs $\boldsymbol{\mathcal{I}}(\boldsymbol{\alpha})$ and $\boldsymbol{\mathcal{I}}(\bar{\boldsymbol{\alpha}})$ are defined in (6.9) and (6.24), respectively. The entry on the u-th row and v-th column of the Jacobian matrix $\dfrac{\partial \boldsymbol{\alpha}^{\mathrm{T}}}{\partial \bar{\boldsymbol{\alpha}}}$ is given by $\left[\dfrac{\partial \boldsymbol{\alpha}^{\mathrm{T}}}{\partial \bar{\boldsymbol{\alpha}}}\right]_{uv} = \dfrac{\partial \alpha_v}{\partial \bar{\alpha}_u}$. Since the desired parameter $\bar{\alpha}_1 = \vartheta$, which is also the first entry in the parameter vector $\boldsymbol{\alpha}$, is independent of the remaining entries in $\boldsymbol{\alpha}$, we obtain

$$\frac{\partial \bar{\boldsymbol{\alpha}}^{\mathrm{T}}}{\partial \boldsymbol{\alpha}} = \begin{bmatrix} \boldsymbol{e}_1^{\mathrm{T}} \\ * \end{bmatrix}, \tag{6.27}$$

where the irrelevant entries are denoted by the placeholder symbol $*$. Substituting (6.27) into (6.26), the following identity holds:

$$\boldsymbol{\mathcal{I}}(\boldsymbol{\alpha})^{\dagger} = \begin{bmatrix} \boldsymbol{e}_1^{\mathrm{T}} \boldsymbol{\mathcal{I}}(\bar{\boldsymbol{\alpha}})^{-1} \boldsymbol{e_1} & * \\ * & * \end{bmatrix} = \begin{bmatrix} [\boldsymbol{\mathcal{I}}(\bar{\boldsymbol{\alpha}})^{-1}]_{11} & * \\ * & * \end{bmatrix}, \tag{6.28}$$

and therefore $\left[\boldsymbol{\mathcal{I}}(\boldsymbol{\alpha})^{\dagger}\right]_{11} = [\boldsymbol{\mathcal{I}}(\bar{\boldsymbol{\alpha}})^{-1}]_{11}$. As a result, the identity in (6.25) holds true, and we conclude the proof. □

As mentioned in Section 6.2 and Theorem 6.2, we emphasize that the sufficient conditions for the existence of the CRB under the conventional signal model in (2.4) and under the PR signal model in (4.2) are generally different: the source covariance matrix $\boldsymbol{P}$ must be full rank for the existence of the CRB under the PR signal model in (4.2). On the contrary, this condition is not required under the conventional signal model. Nevertheless, the relationship between the sufficient conditions for the existence of the CRB under the conventional signal model in Section 6.2 and the PR signal model in Section 4.1 is characterized by the following lemma.

Lemma 6.3.1. *Assume the PR signal model in* (4.2)*. If (As.1)-(As.6) hold, then the assumption (As.7) is also satisfied, and therefore the PR-CRB in* (6.25) *exists.*

Proof of Lemma 6.3.1. We prove by contradiction: Assume that the set $\mathbb{S}$ in (As.7) contains an infinite number of elements. Thus we can pick from the set $\mathbb{S}$ at least $(N+1)$ elements $\{\vartheta_1), \ldots, \vartheta_{N+1}\}$ whose steering vector $\boldsymbol{a}(\vartheta_n)$ for $n = 1, \ldots, N+1$ lie in the space spanned by $\boldsymbol{U}_{\mathrm{s}}$. This implies that $\mathrm{span}\left(\{\boldsymbol{a}(\vartheta_1), \ldots, \boldsymbol{a}(\vartheta_{N+1})\}\right) \subseteq \mathrm{span}\left(\boldsymbol{U}_{\mathrm{s}}\right)$. However, from (As.4) and noting that $N+1 \leq M$, we obtain the following inequality

$$\begin{aligned} N+1 &= \dim\left(\mathrm{span}\left(\{\boldsymbol{a}(\vartheta_1), \ldots, \boldsymbol{a}(\vartheta_{N+1})\}\right)\right) \\ &\leq \dim\left(\mathrm{span}\left(\boldsymbol{U}_{\mathrm{s}}\right)\right) = \mathrm{rank}\left(\boldsymbol{U}_{\mathrm{s}}\right) = N, \end{aligned} \tag{6.29}$$

which is a contradiction. This implies that the set $\mathbb{S}$ must contain only a finite number of elements, and in fact the number of elements in $\mathbb{S}$ must be at most N. As a result, the assumption (As.7) holds true. According to Theorem 6.2, the PR-CRB exists, and we conclude the proof. □

A direct consequence of Lemma 6.3.1 is that the assumptions (As.1)-(As.6) are sufficient conditions for the existence of the CRBs under both the conventional signal model and the PR signal model. Furthermore, any baseband receive matrix $\boldsymbol{X}$ generated from the conventional signal model in (2.4) can be likewise generated from the PR signal model in (4.2) with $\vartheta \in \{\theta_1, \ldots, \theta_N\}$ and $\boldsymbol{B}$ containing the steering vectors of the remaining DOAs. Therefore, in order to facilitate the comparison between the two CRBs, which is the main topic of the discussion in Section 6.5, we assume in the following that (As.1)-(As.6) hold true.

6.4 Derivation of the Cramér-Rao Bound for the DOA Parameter ϑ

In this section, inspired by the work in [SLG01], we derive the CRB for the parameter ϑ under the PR signal model in (4.2) directly from the FIM using the parameterization of the covariance matrix in (6.19) and (6.21). We partition the parameter vector $\bar{\boldsymbol{\alpha}}$ in (6.21) as follows:

$$\bar{\boldsymbol{\alpha}} = \left[\boldsymbol{\beta}^{\mathrm{T}} \;\middle|\; \boldsymbol{\gamma}^{\mathrm{T}}\right]^{\mathrm{T}} = \left[\vartheta, \check{\boldsymbol{b}}^{\mathrm{T}}, \tilde{\boldsymbol{b}}^{\mathrm{T}} \;\middle|\; \bar{\boldsymbol{p}}^{\mathrm{T}}, \nu\right]^{\mathrm{T}}, \tag{6.30}$$

where $\boldsymbol{\beta} = \left[\vartheta, \check{\boldsymbol{b}}^{\mathrm{T}}, \tilde{\boldsymbol{b}}^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{R}^{1+2(M-N)(N-1)}$ contains the parameters characterizing the transformed steering matrix $\bar{\boldsymbol{A}}$ in (6.16), and $\boldsymbol{\gamma} = \left[\bar{\boldsymbol{p}}^{\mathrm{T}}, \nu\right]^{\mathrm{T}} \in \mathbb{R}^{N^2+1}$ contains the remaining nuisance parameters. The FIM $\boldsymbol{\mathcal{I}} = \boldsymbol{\mathcal{I}}(\bar{\boldsymbol{\alpha}})$ in (6.24) and the CRB matrix $\boldsymbol{C} = \boldsymbol{\mathcal{I}}(\bar{\boldsymbol{\alpha}})^{-1}$ is partitioned conformably with (6.30) as follows:

$$\boldsymbol{\mathcal{I}}^{-1} = \begin{bmatrix} \boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}\boldsymbol{\beta}} & \boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}\boldsymbol{\gamma}} \\ \boldsymbol{\mathcal{I}}_{\boldsymbol{\gamma}\boldsymbol{\beta}} & \boldsymbol{\mathcal{I}}_{\boldsymbol{\gamma}\boldsymbol{\gamma}} \end{bmatrix}^{-1} = \begin{bmatrix} \boldsymbol{C}_{\boldsymbol{\beta}\boldsymbol{\beta}} & \boldsymbol{C}_{\boldsymbol{\beta}\boldsymbol{\gamma}} \\ \boldsymbol{C}_{\boldsymbol{\gamma}\boldsymbol{\beta}} & \boldsymbol{C}_{\boldsymbol{\gamma}\boldsymbol{\gamma}} \end{bmatrix} = \boldsymbol{C}. \tag{6.31}$$

We further define $\boldsymbol{\Sigma} = \boldsymbol{C}_{\boldsymbol{\beta}\boldsymbol{\beta}}^{-1}$ and partition two matrices $\boldsymbol{\Sigma}$ and $\boldsymbol{C}_{\boldsymbol{\beta}\boldsymbol{\beta}}$ corresponding to (6.30) as follows:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \Sigma_{\vartheta\vartheta} & \boldsymbol{\Sigma}_{\vartheta\check{\boldsymbol{b}}} & \boldsymbol{\Sigma}_{\vartheta\tilde{\boldsymbol{b}}} \\ \boldsymbol{\Sigma}_{\check{\boldsymbol{b}}\vartheta} & \boldsymbol{\Sigma}_{\check{\boldsymbol{b}}\check{\boldsymbol{b}}} & \boldsymbol{\Sigma}_{\check{\boldsymbol{b}}\tilde{\boldsymbol{b}}} \\ \boldsymbol{\Sigma}_{\tilde{\boldsymbol{b}}\vartheta} & \boldsymbol{\Sigma}_{\tilde{\boldsymbol{b}}\check{\boldsymbol{b}}} & \boldsymbol{\Sigma}_{\tilde{\boldsymbol{b}}\tilde{\boldsymbol{b}}} \end{bmatrix} = \boldsymbol{C}_{\boldsymbol{\beta}\boldsymbol{\beta}}^{-1} = \begin{bmatrix} C_{\vartheta\vartheta} & \boldsymbol{C}_{\vartheta\check{\boldsymbol{b}}} & \boldsymbol{C}_{\vartheta\tilde{\boldsymbol{b}}} \\ \boldsymbol{C}_{\check{\boldsymbol{b}}\vartheta} & \boldsymbol{C}_{\check{\boldsymbol{b}}\check{\boldsymbol{b}}} & \boldsymbol{C}_{\check{\boldsymbol{b}}\tilde{\boldsymbol{b}}} \\ \boldsymbol{C}_{\tilde{\boldsymbol{b}}\vartheta} & \boldsymbol{C}_{\tilde{\boldsymbol{b}}\check{\boldsymbol{b}}} & \boldsymbol{C}_{\tilde{\boldsymbol{b}}\tilde{\boldsymbol{b}}} \end{bmatrix}^{-1} \tag{6.32}$$

The objective is to compute the upper-left entry $C_{\vartheta\vartheta}$ of the matrix $\boldsymbol{C}_{\boldsymbol{\beta}\boldsymbol{\beta}}$, which is the CRB of the DOA parameter ϑ, under the knowledge of the entries $\mathcal{I}_{uv}$ in (6.24). Following the footsteps in [SLG01] to the PR signal model with the parameterization in (6.21), all entries of the matrix $\boldsymbol{\Sigma}$ in (6.32) are computed thanks to the following theorem.

Theorem 6.4. *The entry in the p-th row and q-th column of the matrix* $\boldsymbol{\Sigma}$ *in* (6.32) *is given by*

$$\Sigma_{pq} = \frac{2T}{\nu}\mathrm{Re}\left\{\mathrm{tr}\left(\bar{\boldsymbol{P}}\bar{\boldsymbol{A}}^{\mathrm{H}}\boldsymbol{R}^{-1}\bar{\boldsymbol{A}}\bar{\boldsymbol{P}}\frac{\mathrm{d}\bar{\boldsymbol{A}}}{\mathrm{d}\beta_p}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\frac{\mathrm{d}\bar{\boldsymbol{A}}}{\mathrm{d}\beta_q}\right)\right\}. \tag{6.33}$$

Proof of Theorem 6.4. See Appendix I □

From Theorem 6.4, we first compute the matrix $\boldsymbol{\Sigma}$ in (6.32) elementwise, then apply the matrix inversion lemma in Appendix A, Equation (A.9) to obtain the closed-form expression of the CRB corresponding to the parameter of interest ϑ. In order to compute each block of the matrix $\boldsymbol{\Sigma}$ in (6.32) using (6.33), first we note that

$$\frac{\mathrm{d}\bar{\boldsymbol{A}}}{\mathrm{d}\beta_p} = \begin{cases} \boldsymbol{d}\boldsymbol{e}_1^{\mathrm{T}} & \text{if } \beta_p = \vartheta, \\ \boldsymbol{e}_{k+1}\boldsymbol{e}_{l+1}^{\mathrm{T}} & \text{if } \beta_p = \check{b}_{(k,l)}, \\ j\boldsymbol{e}_{k+1}\boldsymbol{e}_{l+1}^{\mathrm{T}} & \text{if } \beta_p = \tilde{b}_{(k,l)}, \end{cases} \tag{6.34}$$

where $\boldsymbol{d} = \left.\dfrac{\mathrm{d}\boldsymbol{a}(\theta)}{\mathrm{d}\theta}\right|_{\theta=\vartheta}$ and $\boldsymbol{e}_k$ is an elementary vector of suitable dimension whose element on the k-th row is one and the remaining elements are zeros. Applying (6.34) to (6.33), the upper-left entry $\Sigma_{\vartheta\vartheta}$ is computed as follows:

$$\begin{aligned}\Sigma_{\vartheta\vartheta} &= \frac{2T}{\nu}\mathrm{Re}\left\{\mathrm{tr}\left(\bar{\boldsymbol{P}}\bar{\boldsymbol{A}}^{\mathrm{H}}\boldsymbol{R}^{-1}\bar{\boldsymbol{A}}\bar{\boldsymbol{P}}\frac{\mathrm{d}\bar{\boldsymbol{A}}}{\mathrm{d}\vartheta}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\frac{\mathrm{d}\bar{\boldsymbol{A}}}{\mathrm{d}\vartheta}\right)\right\} \\ &= \frac{2T}{\nu}\mathrm{Re}\left\{\left(\boldsymbol{e}_1^{\mathrm{T}}\bar{\boldsymbol{P}}\bar{\boldsymbol{A}}^{\mathrm{H}}\boldsymbol{R}^{-1}\bar{\boldsymbol{A}}\bar{\boldsymbol{P}}\boldsymbol{e}_1\right)\left(\boldsymbol{d}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{d}\right)\right\}. \end{aligned} \tag{6.35}$$

Defining

$$\bar{\boldsymbol{M}} = \left(\bar{\boldsymbol{P}}\bar{\boldsymbol{A}}^{\mathrm{H}}\boldsymbol{R}^{-1}\bar{\boldsymbol{A}}\bar{\boldsymbol{P}}\right)^{\mathrm{T}} = \begin{bmatrix}\bar{M}_{11} & \bar{\boldsymbol{M}}_{21}^{\mathrm{H}} \\ \bar{\boldsymbol{M}}_{21} & \bar{\boldsymbol{M}}_{22}\end{bmatrix}, \tag{6.36}$$

the element $\Sigma_{\vartheta\vartheta}$ is compactly given by

$$\Sigma_{\vartheta\vartheta} = \frac{2T}{\nu}\mathrm{Re}\left\{\bar{M}_{11}\boldsymbol{d}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{d}\right\}. \tag{6.37}$$

The entries of the block $\boldsymbol{\Sigma}_{\vartheta\check{\boldsymbol{b}}} = \left[\Sigma_{\vartheta\check{b}_{(1,1)}}, \ldots, \Sigma_{\vartheta\check{b}_{(M-N,N-1)}}\right]$ are computed from (6.33) and (6.34) as

$$\begin{aligned}\Sigma_{\vartheta\check{b}_{(k,l)}} &= \frac{2T}{\nu}\mathrm{Re}\left\{\mathrm{tr}\left(\bar{\boldsymbol{P}}\bar{\boldsymbol{A}}^{\mathrm{H}}\boldsymbol{R}^{-1}\bar{\boldsymbol{A}}\bar{\boldsymbol{P}}\frac{\mathrm{d}\bar{\boldsymbol{A}}}{\mathrm{d}\vartheta}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\frac{\mathrm{d}\bar{\boldsymbol{A}}}{\mathrm{d}\check{b}_{(k,l)}}\right)\right\} \\ &= \frac{2T}{\nu}\mathrm{Re}\left\{\left(\boldsymbol{e}_{l+1}^{\mathrm{T}}\bar{\boldsymbol{P}}\bar{\boldsymbol{A}}^{\mathrm{H}}\boldsymbol{R}^{-1}\bar{\boldsymbol{A}}\bar{\boldsymbol{P}}\boldsymbol{e}_1\right)\left(\boldsymbol{d}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{e}_{k+1}\right)\right\}, \end{aligned} \tag{6.38}$$

for $k = 1, \ldots, M-N$ and $l = 1, \ldots, N-1$. From (6.38), the compact expression of the block $\boldsymbol{\Sigma}_{\vartheta\check{\boldsymbol{b}}}$ is given by

$$\boldsymbol{\Sigma}_{\vartheta\check{\boldsymbol{b}}} = \boldsymbol{\Sigma}_{\check{\boldsymbol{b}}\vartheta}^{\mathrm{T}} = \frac{2}{\nu}\mathrm{Re}\left\{\bar{\boldsymbol{M}}_{21}^{\mathrm{H}} \otimes \left(\boldsymbol{d}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{J}\right)\right\}, \tag{6.39}$$

where

$$\boldsymbol{J} = [\boldsymbol{e}_2, \ldots, \boldsymbol{e}_{M-N+1}] \in \mathbb{R}^{M\times(M-N)} \tag{6.40}$$

is the selection matrix that extracts the second to the $(M-N+1)$-th column when right-multiplying with a matrix, and $\bar{\boldsymbol{M}}_{12}$ is defined as in (6.36). In the same manner, the compact expression of the block $\boldsymbol{\Sigma}_{\vartheta\tilde{\boldsymbol{b}}}$ is given by

$$\boldsymbol{\Sigma}_{\vartheta\tilde{\boldsymbol{b}}} = \boldsymbol{\Sigma}_{\tilde{\boldsymbol{b}}\vartheta}^{\mathrm{T}} = -\frac{2T}{\nu}\mathrm{Im}\left\{\bar{\boldsymbol{M}}_{21}^{\mathrm{H}} \otimes \left(\boldsymbol{d}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{J}\right)\right\}. \tag{6.41}$$

The element of the block $\boldsymbol{\Sigma}_{\check{\boldsymbol{b}}\check{\boldsymbol{b}}}$ corresponding to the entries $\check{b}_{(k,l)}$ and $\check{b}_{(r,s)}$, which is defined in (6.22) and (6.23), respectively, is calculated as

$$\begin{aligned}\Sigma_{\check{b}_{(k,l)}\check{b}_{(r,s)}} &= \frac{2T}{\nu}\mathrm{Re}\left\{\mathrm{tr}\left(\bar{\boldsymbol{P}}\bar{\boldsymbol{A}}^{\mathrm{H}}\boldsymbol{R}^{-1}\bar{\boldsymbol{A}}\bar{\boldsymbol{P}}\frac{\mathrm{d}\bar{\boldsymbol{A}}}{\mathrm{d}\check{b}_{(k,l)}}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\frac{\mathrm{d}\bar{\boldsymbol{A}}}{\mathrm{d}\check{b}_{(r,s)}}\right)\right\} \\ &= \frac{2T}{\nu}\mathrm{Re}\left\{\left(\boldsymbol{e}_{s+1}^{\mathrm{T}}\bar{\boldsymbol{P}}\bar{\boldsymbol{A}}^{\mathrm{H}}\boldsymbol{R}^{-1}\bar{\boldsymbol{A}}\bar{\boldsymbol{P}}\boldsymbol{e}_{k+1}\right)\left(\boldsymbol{e}_{l+1}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{e}_{r+1}\right)\right\},\end{aligned} \tag{6.42}$$

for $k, r = 1, \ldots, M-N$ and $l, s = 1, \ldots, N-1$. The corresponding matrix form is obtained as follows:

$$\boldsymbol{\Sigma}_{\check{\boldsymbol{b}}\check{\boldsymbol{b}}} = \frac{2T}{\nu}\mathrm{Re}\left\{\bar{\boldsymbol{M}}_{22} \otimes \left(\boldsymbol{J}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{J}\right)\right\}. \tag{6.43}$$

Performing the same procedure for the remaining blocks, we obtain

$$\boldsymbol{\Sigma}_{\tilde{\boldsymbol{b}}\check{\boldsymbol{b}}} = \frac{2T}{\nu}\mathrm{Re}\left\{\bar{\boldsymbol{M}}_{22} \otimes \left(\boldsymbol{J}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{J}\right)\right\}, \tag{6.44}$$

$$\boldsymbol{\Sigma}_{\check{\boldsymbol{b}}\tilde{\boldsymbol{b}}} = \boldsymbol{\Sigma}_{\tilde{\boldsymbol{b}}\check{\boldsymbol{b}}}^{\mathrm{T}} = -\frac{2T}{\nu}\mathrm{Im}\left\{\bar{\boldsymbol{M}}_{22} \otimes \left(\boldsymbol{J}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{J}\right)\right\}. \tag{6.45}$$

Therefore, the inverse CRB matrix for the block corresponding to the parameterization of the transformed matrix $\bar{\boldsymbol{A}}$, hence the matrix $\boldsymbol{C}_{\boldsymbol{\beta}\boldsymbol{\beta}}^{-1}$ defined in (6.32), is computed as follows:

$$\begin{aligned}&\begin{bmatrix} C_{\vartheta\vartheta} & \boldsymbol{C}_{\vartheta\check{\boldsymbol{b}}} & \boldsymbol{C}_{\vartheta\tilde{\boldsymbol{b}}} \\ \boldsymbol{C}_{\check{\boldsymbol{b}}\vartheta} & \boldsymbol{C}_{\check{\boldsymbol{b}}\check{\boldsymbol{b}}} & \boldsymbol{C}_{\check{\boldsymbol{b}}\tilde{\boldsymbol{b}}} \\ \boldsymbol{C}_{\tilde{\boldsymbol{b}}\vartheta} & \boldsymbol{C}_{\tilde{\boldsymbol{b}}\check{\boldsymbol{b}}} & \boldsymbol{C}_{\tilde{\boldsymbol{b}}\tilde{\boldsymbol{b}}} \end{bmatrix}^{-1} \\ =&\frac{2T}{\nu}\begin{bmatrix} \bar{\boldsymbol{M}}_{11}\boldsymbol{d}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{d} & \mathrm{Re}\left\{\bar{\boldsymbol{M}}_{21}^{\mathrm{H}} \otimes \left(\boldsymbol{d}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{J}\right)\right\} & -\mathrm{Im}\left\{\bar{\boldsymbol{M}}_{21}^{\mathrm{H}} \otimes \left(\boldsymbol{d}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{J}\right)\right\} \\ \mathrm{Re}\left\{\bar{\boldsymbol{M}}_{21} \otimes \left(\boldsymbol{J}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{d}\right)\right\} & \mathrm{Re}\left\{\bar{\boldsymbol{M}}_{22} \otimes \left(\boldsymbol{J}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{J}\right)\right\} & -\mathrm{Im}\left\{\bar{\boldsymbol{M}}_{22} \otimes \left(\boldsymbol{J}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{J}\right)\right\} \\ \mathrm{Im}\left\{\bar{\boldsymbol{M}}_{21} \otimes \left(\boldsymbol{J}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{d}\right)\right\} & \mathrm{Im}\left\{\bar{\boldsymbol{M}}_{22} \otimes \left(\boldsymbol{J}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{J}\right)\right\} & \mathrm{Re}\left\{\bar{\boldsymbol{M}}_{22} \otimes \left(\boldsymbol{J}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{J}\right)\right\} \end{bmatrix}.\end{aligned} \tag{6.46}$$

The inverse of the CRB of the desired direction ϑ, which corresponds to the entry $C_{\vartheta\vartheta}$, is then calculated using the block matrix inversion lemma (A.9) and the identities (A.2), (A.3), (A.4) and (A.8) in Appendix A as follows:

$$\begin{aligned} C_{\text{PR-CRB}}(\vartheta)^{-1} = C_{\vartheta\vartheta}^{-1} &= \frac{2T}{\nu}\Big(\bar{M}_{11}\boldsymbol{d}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{d} - \operatorname{Re}\Big\{\Big(\bar{\boldsymbol{M}}_{21}^{\mathrm{H}} \otimes \big(\boldsymbol{d}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{J}\big)\Big) \\ &\qquad \Big(\bar{\boldsymbol{M}}_{22}^{-1} \otimes \big(\boldsymbol{J}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{J}\big)^{-1}\Big)\big(\bar{\boldsymbol{M}}_{21} \otimes \big(\boldsymbol{J}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{d}\big)\big)\Big\}\Big) \\ &= \frac{2T}{\nu}\Big(\bar{M}_{11}\boldsymbol{d}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{d} - \Big(\bar{\boldsymbol{M}}_{21}^{\mathrm{H}}\bar{\boldsymbol{M}}_{22}^{-1}\bar{\boldsymbol{M}}_{21}\Big)\boldsymbol{d}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{J}}\boldsymbol{d}\Big). \end{aligned} \tag{6.47}$$

The expression in (6.47) is further simplified by the following identity.

Theorem 6.5. *Define the selection matrix $\boldsymbol{J}$ as in* (6.40)*. For any matrix $\boldsymbol{K} \in \mathbb{C}^{M\times N}$ such that the matrix $\boldsymbol{J}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{K}}^{\perp}\boldsymbol{J}$ is invertible, we have*

$$\boldsymbol{\Pi}_{\boldsymbol{K}}^{\perp} = \boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{K}}^{\perp}\boldsymbol{J}}. \tag{6.48}$$

Proof of Lemma 6.5. See Appendix J. □

Using Theorem 6.5 with $\boldsymbol{K} = \bar{\boldsymbol{A}}$ and noting that the invertibility of $\boldsymbol{J}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{J}$ is ensured due to the existence of the CRB with respect to the parameterization (6.21) as proven in Section 6.3, the CRB for the desired parameter ϑ under the PR signal model simplifies to

$$C_{\text{PR-CRB}}(\vartheta) = \frac{\nu}{2T}\Big(\Big(\bar{M}_{11} - \bar{\boldsymbol{M}}_{21}^{\mathrm{H}}\bar{\boldsymbol{M}}_{22}^{-1}\bar{\boldsymbol{M}}_{21}\Big)\boldsymbol{d}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\boldsymbol{d}\Big)^{-1}. \tag{6.49}$$

For easier comparison with the conventional stochastic CRB in the literature, the expression of the PR-CRB in (6.49) is rewritten in the following manner. First we note that, the transformation matrix $\boldsymbol{T}$ defined in (6.13) is invertible, and therefore it follows from (6.16) that

$$\begin{aligned} \boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp} &= \boldsymbol{I} - \boldsymbol{A}\boldsymbol{T}\big(\boldsymbol{T}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\boldsymbol{T}\big)^{-1}\boldsymbol{T}^{\mathrm{H}}\boldsymbol{A} \\ &= \boldsymbol{I} - \boldsymbol{A}\boldsymbol{T}\boldsymbol{T}^{-1}\big(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\big)^{-1}\boldsymbol{T}^{-\mathrm{H}}\boldsymbol{T}^{\mathrm{H}}\boldsymbol{A} \\ &= \boldsymbol{I} - \boldsymbol{A}\big(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\big)^{-1}\boldsymbol{A} = \boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}. \end{aligned} \tag{6.50}$$

In addition, if we partition

$$\boldsymbol{M} = \big(\boldsymbol{P}\boldsymbol{A}^{\mathrm{H}}\boldsymbol{R}^{-1}\boldsymbol{A}\boldsymbol{P}\big)^{\mathrm{T}} = \begin{bmatrix} M_{11} & \boldsymbol{M}_{21}^{\mathrm{H}} \\ \boldsymbol{M}_{21} & \boldsymbol{M}_{22} \end{bmatrix}, \tag{6.51}$$

and note that, the matrix $\bar{\boldsymbol{M}}$ in (6.36) is computed as $\bar{\boldsymbol{M}} = \big(\boldsymbol{T}^{-1}\big)^{*}\boldsymbol{M}\boldsymbol{T}^{-\mathrm{T}}$ where the inverse transformation matrix $\boldsymbol{T}^{-1}$ is defined in (6.14). From this observation, we

obtain

$$\bar{M}_{11} = M_{11} + 2\mathrm{Re}\left\{\boldsymbol{M}_{21}^{\mathrm{H}}\boldsymbol{b}_1 a_1^{-1}\right\} + |a_1|^{-2}\,\boldsymbol{b}_1^{\mathrm{H}}\boldsymbol{M}_{22}\boldsymbol{b}_1, \tag{6.52}$$

$$\bar{\boldsymbol{M}}_{21} = \left(\boldsymbol{B}_3 - \boldsymbol{a}_3 a_1^{-1}\boldsymbol{b}_1^{\mathrm{T}}\right)^* \left(\boldsymbol{M}_{21} + \boldsymbol{M}_{22}\boldsymbol{b}_1^* a_1^*\right), \tag{6.53}$$

$$\bar{\boldsymbol{M}}_{22} = \left(\boldsymbol{B}_3 - \boldsymbol{a}_3 a_1^{-1}\boldsymbol{b}_1^{\mathrm{T}}\right)^* \boldsymbol{M}_{22} \left(\boldsymbol{B}_3 - \boldsymbol{a}_3 a_1^{-1}\boldsymbol{b}_1^{\mathrm{T}}\right)^{\mathrm{T}}, \tag{6.54}$$

Substituting (6.52)-(6.54) into (6.49) and further simplifying, we summarize our result into the following theorem, which provides a closed-form expression of the CRB for the DOA parameter ϑ under the PR signal model.

Theorem 6.6. *Assume the PR signal model in* (4.2)*. Under the assumptions (As.1)-(As.6), the CRB of the parameter ϑ is given by the following expression:*

$$C_{\text{PR-CRB}}(\vartheta) = \frac{\nu}{2T}\left(\left(M_{11} - \boldsymbol{M}_{21}^{\mathrm{H}}\boldsymbol{M}_{22}^{-1}\boldsymbol{M}_{21}\right)\boldsymbol{d}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}\boldsymbol{d}\right)^{-1}, \tag{6.55}$$

with $\boldsymbol{A} = \boldsymbol{A}(\vartheta)$, $\boldsymbol{d} = \left.\dfrac{\mathrm{d}\boldsymbol{a}(\theta)}{\mathrm{d}\theta}\right|_{\theta=\vartheta}$, *and* $\boldsymbol{M}$ *defined in* (6.51).

Thus, we have proved one of our main results of this chapter, which is a closed-form expression of the CRB for the DOA parameter ϑ under the PR signal model.

6.5 Discussion

As the relaxed steering matrix $\boldsymbol{A}(\vartheta)$ in the PR signal model in (4.2) contains less structural information on the DOA than the true steering matrix $\boldsymbol{A}(\boldsymbol{\theta})$ in the conventional signal model in (2.4), the PR-CRB in (6.55) provides some insights regarding the impact of relaxing the array manifold on the DOA estimation performance. In order to investigate the impact of the relaxation, in this section, we consider the special case where the relaxed steering matrix $\boldsymbol{A}(\vartheta)$ is identical to the true steering matrix $\boldsymbol{A}(\boldsymbol{\theta})$ up to a column permutation. Under this special case, we obtain N different PR-CRBs corresponding to N distinct directions $\vartheta \in \{\theta_1, \ldots, \theta_N\}$ considered as the source of interest. We further remark that, although the true and the relaxed steering matrix are identical, the CRBs for the same DOA parameter are not identical due to different *a priori* knowledge on the structure of the steering matrix. Interestingly, it turns out that some established results in the literature can be re-derived indirectly by investigating the PR-CRB under this special case.

6.5.1 Relationship between Conventional CRB and PR-CRB

In the following, we provide an analytical comparison between the conventional CRB and the PR-CRB as stated in the following theorem.

Theorem 6.7. *The PR-CRB of the desired DOA parameter is lower-bounded by the conventional CRB corresponding to the same DOA parameter:*

$$C_{\text{PR-CRB}}(\theta_n) \geq [\boldsymbol{C}_{\text{CRB}}(\boldsymbol{\theta})]_{nn} \text{ for all } n = 1, \ldots, N. \tag{6.56}$$

Proof of Theorem 6.7. See Appendix K. □

The intuition behind Theorem 6.7 is that, the relaxation of the steering manifold removes information from the model and generally degrades the DOA estimation performance. However, Theorem 6.7 does not exclude special cases under which the PR-CRB can be equal to the conventional CRB. In the following, we provide one sufficient condition under which the PR-CRB coincides with the conventional CRB.

Theorem 6.8. *If the matrix* $\boldsymbol{M}$ *in* (6.51) *is diagonal, then the PR-CRB and the conventional CRB are equal:*

$$C_{\text{PR-CRB}}(\theta_n) = [\boldsymbol{C}_{\text{CRB}}(\boldsymbol{\theta})]_{nn} \text{ for all } n = 1, \ldots, N. \tag{6.57}$$

Proof of Theorem 6.8. Without loss of generality, we assume that the desired DOA parameter under the PR signal model coincides with the first DOA, i.e., $\vartheta = \theta_1$. Since the matrix $\boldsymbol{M}$ is diagonal, it follows that the matrix $\boldsymbol{M} \odot \left(\boldsymbol{D}^{\text{H}} \boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp} \boldsymbol{D}\right)$ is also diagonal. As a result, the conventional CRB in (6.5) for the desired DOA parameter $\vartheta = \theta_1$ is given by

$$[\boldsymbol{C}_{\text{CRB}}(\boldsymbol{\theta})]_{11} = \frac{\nu}{2T} (M_{11} \boldsymbol{d}^{\text{H}} \boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp} \boldsymbol{d})^{-1}. \tag{6.58}$$

On the other hand, due to the diagonal structure of the matrix $\boldsymbol{M}$, the matrices $\boldsymbol{M}_{21}$ and $\boldsymbol{M}_{12} = \boldsymbol{M}_{21}^{\text{H}}$ in (6.51) vanish. Exploiting this fact and comparing (6.55) with (6.58), the result in (6.57) follows. □

Although theoretically proven, it may seem difficult to deduce implications from the sufficient condition in Theorem 6.8 as the expression of the matrix $\boldsymbol{M}$ in (6.51) lacks physical interpretation. In order to provide more insights on this condition, in the following we consider the high SNR regime. As the noise power ν is small, the matrix $\boldsymbol{M}$ in (6.51) is approximated by [VT04, Eq. (8.110)]:

$$\boldsymbol{M} = \left(\boldsymbol{P} \boldsymbol{A}^{\text{H}} \boldsymbol{R}^{-1} \boldsymbol{A} \boldsymbol{P}\right)^{\text{T}} \approx \boldsymbol{P}. \tag{6.59}$$

From the approximation in (6.59), we obtain a weaker but more useful corollary than Theorem 6.8: The PR-CRB in (6.55) is approximately equal to the conventional CRB in (6.5) if the SNR is sufficiently high and the source signals are uncorrelated. Under this particular scenario, the PR estimators are expected to achieve near-optimal asymptotic error performance of the conventional ML estimators with substantially reduced computational complexity. However, this argument is valid only if some PR estimators can achieve the PR-CRB asymptotically. For that reason, the asymptotic error performance of the PR estimators and their relation to the PR-CRB are the topics of discussion in the following section.

6.5.2 Asymptotic Error Performance of PR Estimators

In this section, the expressions of the asymptotic MSE of two PR estimators, namely the PR-WSF in (4.10) and the PR-DML in (4.8), are studied. As the estimators under the PR framework are derived from the PR signal model instead of the conventional signal model, we also study the statistical efficiency of the aforementioned estimators with respect to the PR-CRB in (6.55). First, the influence of the weighting matrix $\boldsymbol{W}$ on the asymptotic estimation performance of the PR-WSF estimator is investigated. We write $\hat{\theta}_n(\boldsymbol{W})$ to be the PR-WSF estimate of the true DOA parameter θ_n. The dependence of the statistical distribution of the estimate $\hat{\theta}_n(\boldsymbol{W})$ on the number of snapshots T is suppressed for notational convenience. Using this notation, the asymptotic statistical efficiency of the PR-WSF estimator is addressed in the following theorem.

Theorem 6.9. *The asymptotic MSE of the PR-WSF estimator for any positive definite weighting matrix $\boldsymbol{W} \succ \boldsymbol{0}$ is equal to the PR-CRB:*

$$\lim_{T\to\infty} \frac{C_{\text{PR-CRB}}(\theta_n)}{\mathbb{E}\left\{\left|\hat{\theta}_n(\boldsymbol{W}) - \theta_n\right|^2\right\}} = 1 \text{ for all } n = 1, \ldots, N. \tag{6.60}$$

Proof of Theorem 6.9. See Appendix L. □

Theorem 6.9 shows that, the asymptotic error performance of the PR-WSF is independent of the weighting matrix $\boldsymbol{W}$. This is indeed different from the WSF estimator developed for the conventional signal model. In fact, the asymptotic DOA estimation error of the conventional WSF with $\boldsymbol{W} = \boldsymbol{I}_N$ is strictly larger than that using the asymptotically optimal weighting matrix $\boldsymbol{W} = \hat{\boldsymbol{W}}_{\text{opt}}$ in (3.19) [VO91]. Furthermore, an interesting consequence of Theorem 6.9 regarding the asymptotic error performance of the PR-DML and MUSIC estimators is stated by the following theorem.

Theorem 6.10. *The asymptotic MSEs of the PR-WSF, PR-DML and MUSIC estimator are identical, and they are equal to the PR-CRB.*

Proof of Theorem 6.10. The MUSIC estimator can be considered as a special case of PR-WSF with $\boldsymbol{W} = \boldsymbol{I}_N$ as proven in Appendix D. Similarly, from [VO91, Lem. 2], which holds true for any parameterization of the steering matrix $\boldsymbol{A}$, the PR-DML estimator has the same asymptotic error performance as the PR-WSF estimator with the weighting matrix $\boldsymbol{W} = \hat{\boldsymbol{\Lambda}}_{\mathrm{s}} - \hat{\nu}\boldsymbol{I}_N$ with $\hat{\boldsymbol{\Lambda}}_{\mathrm{s}}$ defined in (2.9) and $\hat{\nu} = \frac{1}{M-N}\sum_{k=N+1}^{M}\hat{\lambda}_k$. From Theorem 6.9, we conclude our proof. □

As a consequence of Theorem 6.10, two corollaries are deduced: First, under the PR signal model, the PR-DML estimator is asymptotically efficient. This is again different from the results under the conventional signal model, where the DML estimator is inefficient under the conventional signal model for a finite number of sensors M [SN90b, Sec. IV.]. The second corollary from Theorem 6.10 is even more remarkable: In [SN90a, Sec. VI.], under the case with a particular source covariance matrix $\boldsymbol{P} = \left(\boldsymbol{A}^{\mathrm{H}}\boldsymbol{A}\right)^{-1/2}$ and a sufficiently large noise power ν, the asymptotic MSE of the MUSIC estimator is lower than that of the DML estimator. From a different perspective, from Theorem 6.10, the asymptotic MSEs of the MUSIC and PR-DML estimator are identical. Consequently, in the aforementioned special case, the asymptotic MSE of the PR-DML estimator is lower than the asymptotic MSE of the conventional DML estimator. In other words, the relaxation of the steering matrix $\boldsymbol{A}$ can indeed improve the DOA estimation performance for a certain setup. However, the performance improvement of the PR-DML estimator compared to the DML estimator in the aforementioned case does not contradict with Theorem 6.7 as the DML estimator is not an efficient estimator under the assumptions (As.1)-(As.4).

6.6 Summary

In this chapter, the CRB for the PR signal model, which is the lower bound for the MSE of all estimators under the PR framework, has been derived. The performance loss induced by the relaxation of the array manifold is quantified by comparing the closed-form expression of the PR-CRB with the conventional CRB in the literature. Concerning the asymptotic error performance of the proposed DOA estimators, we show that the MSEs of two PR estimators, namely the PR-DML and PR-WSF estimators, are equivalent to that of the MUSIC estimator as the number of snapshots T

tends to infinity. Furthermore, the asymptotic MSEs of the PR-DML, PR-WSF, and MUSIC estimators are identical to the PR-CRB.

We remark that the discussions on the asymptotic MSE behaviors of the proposed DOA estimators are valid only if the number of snapshots T tends to infinity. In the case of a finite number of snapshots T, the statement regarding the equivalent asymptotic error performance of PR-DML, PR-WSF, and MUSIC may not hold. This remark is illustrated by numerical simulations in the next chapter, where the error performance of the proposed estimators under the PR framework is investigated by extensive numerical simulations.

Chapter 7

Numerical Results

7.1 Simulation Setup

In this section, numerical simulation results concerning the error performance of DOA estimators under the PR framework in Chapter 4 and the theoretical lower bound PR-CRB in Chapter 6 are presented. In addition, the error performance of MUSIC, root-MUSIC [Bar83] and DML estimator as well as the stochastic CRB are used as benchmark.

In our simulations, if not further specified, we assume two uncorrelated but closely spaced source signals at $\boldsymbol{\theta} = [45°, 50°]^T$ which impinge on a ULA of $M = 10$ antennas with the spacing equal to half of the wavelength. We stress that in contrast to root-MUSIC, all PR methods are applicable to any array geometry. The source signals have the mean value of zero and unit power. The SNR is computed as $\text{SNR} = \frac{1}{\nu}$. Regarding the PR-WSF method, we choose the weighting as in (3.19).

The numerical experiments are conducted in MATLAB 2016b on a PC equipped with an OS of Arch Linux with a processor 8 x Intel Core i7-6700K 4.00GHz CPU and 16GB RAM. The iterative eigenvalue decomposition introduced in Chapter 5, which plays a key role in accelerating the execution of all PR methods, is implemented in C and imported in MATLAB through a MEX interface. We remark that the function `dlaed4()` in the LAPACK library [ABB+99, BCG+07] can be applied to compute the efficient eigenvalue decomposition in Chapter 5. Nevertheless, the function `dlaed4()` lacks the ability to reuse the eigenvalues of the previous direction as initialization for the next direction during the spectral search. Therefore, we do not employ this function in the implementation of the PR estimators. In order to avoid obtaining local minima while solving the optimization in (3.15) with iterative methods, the DOA estimates from the DML estimator are computed from a brute-force search over a dense grid on $\mathcal{A}_2$. However, we stress that the brute force grid-search implementation is not well-scalable with the increase in the number of sources N as the computation complexity increases exponentially with N. Such high computational complexity prevents the application of the DML estimator in practical cases. Nevertheless, the performance of the DML estimator is included as a benchmark for the threshold performance [Ath05].

The key performance indicators are the Root-Mean-Square Error (RMSE) and the execution time. The RMSE is calculated as

$$\mathrm{RMSE} = \sqrt{\frac{1}{N_R N} \sum_{r=1}^{N_R} \sum_{n=1}^{N} \left(\hat{\theta}_n^{(r)} - \theta_{\mathrm{n}} \right)^2}. \tag{7.1}$$

In (7.1), the estimated DOAs in the r-th Monte-Carlo run $\hat{\boldsymbol{\theta}}^{(r)} = [\hat{\theta}_1^{(r)}, \ldots, \hat{\theta}_N^{(r)}]^{\mathrm{T}}$ and the true DOAs $\boldsymbol{\theta} = [\theta_1, \ldots, \theta_N]^{\mathrm{T}}$ in (7.1) are sorted in ascending order. The number of Monte-Carlo trials is $N_R = 1000$. The execution time of investigated estimators is measured in each Monte-Carlo trial and then averaged.

7.2 Influence of SNR

First we investigate the error performance of the PR estimators compared with conventional methods when the SNR is varied. In the first experiment, we assume the number of snapshots $T = 40$, which is larger than the number of sensors $M = 10$. In this investigated experiment, the sample covariance matrix $\hat{\boldsymbol{R}}$ is positive definite, and thus PR-CCF is applicable. As depicted in Figure 7.1, the difference between the PR-CRB and the conventional CRB is negligible, which is predicted from Theorem 6.8. This observation implies that the relaxation of the array manifold in this scenario does not degrade the estimation performance of the DOA parameters much. The RMSE of both PR-CCF and PR-UCF does not approach the CRB. However, the difference in RMSE is insignificant. PR-DML and PR-WSF have similar performance behaviors, achieving the CRB at a lower SNR than MUSIC but much higher than root-MUSIC. In terms of non-asymptotic performance, the partial relaxation methods exhibit superior SNR threshold performance in comparison to the MUSIC algorithm. The PR-CCF and PR-UCF estimator possess almost identical estimation error performance in the inspected SNR region, where their thresholds occur at an even lower SNR than that of root-MUSIC. The PR-CCF and PR-UCF are outperformed by the brute-force DML in both the asymptotic and the non-asymptotic regions, although the difference in RMSE is small. This remark suggests that PR-CCF is more favorable than PR-UCF, since the computational complexity of PR-CCF is lower than that of PR-UCF while the error performances are comparable.

Figure 7.2 depicts a different setup where the number of snapshots $T = 8$ is smaller than the number of antennas $M = 10$. In this case, the sample covariance matrix calculated in (2.7) is singular, and therefore the PR-CCF is not applicable. Nevertheless, an

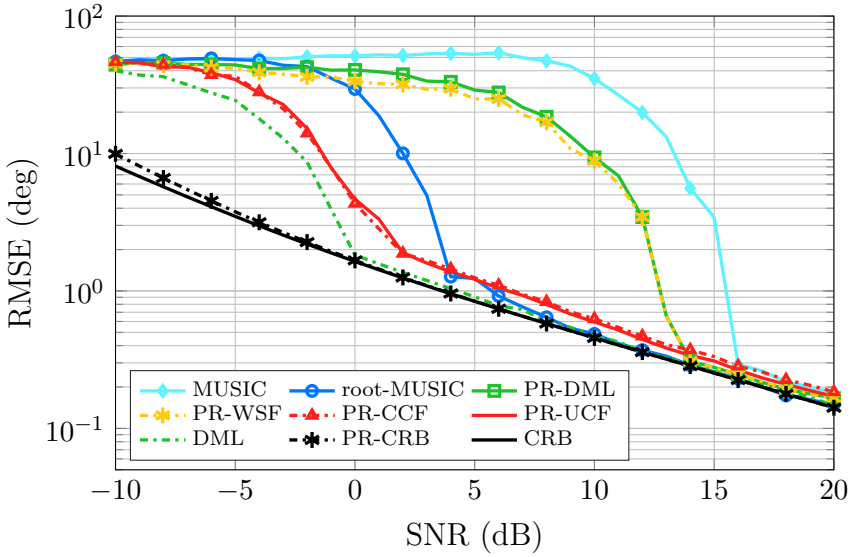

Figure 7.1. Uncorrelated source signals, number of snapshots $T = 40$

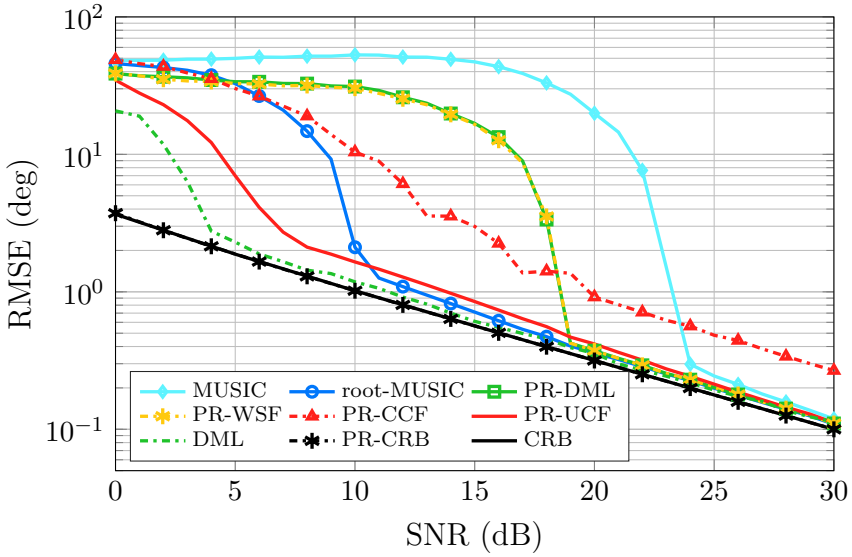

Figure 7.2. Uncorrelated source signals, number of snapshots $T = 8$

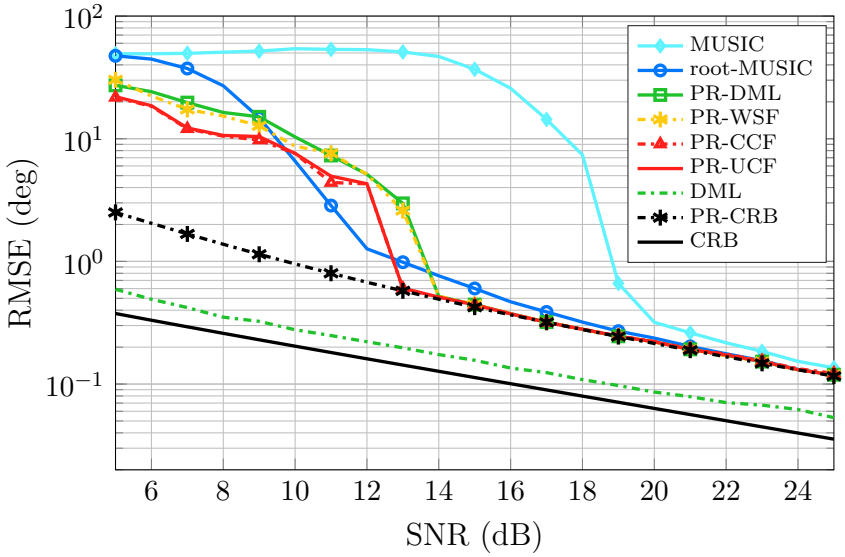

Figure 7.3. Correlated source signals with $\rho = 0.95$, number of snapshots $T = 200$

alternative approach is to apply the diagonal loading technique with the loading factor $\gamma = 10^{-4}$ on the sample covariance matrix before computing the PR-CCF spectrum. We further stress that the diagonal loading technique is applied only to the PR-CCF estimator. Concerning the PR-UCF estimator which does not impose such constraint on the positive definiteness of the covariance matrix, the initialization of $\sigma^2_{\text{s, left}}$ in Algorithm 1 is set at 10^{-6}. To avoid outliers in the RMSE caused by misdetection and to simulate the DOA tracking process [Hu14], 1% of the estimates with the largest error for all investigated algorithms are removed before calculating the RMSE. We observe that, even in the case of an extremely low number of snapshots, PR-UCF obtains a remarkable threshold behavior, outperforming other methods except for the brute-force DML. The RMSE of PR-UCF only slowly approaches the Cramer-Rao Bound as the SNR increases. In the high SNR regime, however, the RMSE of PR-UCF is very close to the Cramer-Rao Bound. The performance of PR-CCF is highly degraded due to the diagonal loading. Further research may be carried out regarding the optimal adaptive choice of the diagonal loading factor γ to achieve an improved performance using a direction-dependent factor. Similar to the above-investigated scenarios, the difference in RMSE between PR-CRB and the conventional CRB is negligible. PR-DML and PR-WSF outperform MUSIC consistently. In the next simulation, we consider the correlated signals of two sources while varying the SNR. The correlation coefficient ρ is defined as:

$$\rho = \frac{\mathbb{E}\left\{s_1(t)^{\mathrm{H}} s_2(t)\right\}}{\sqrt{\mathbb{E}\left\{|s_1(t)|^2\right\}\mathbb{E}\left\{|s_2(t)|^2\right\}}}. \tag{7.2}$$

In Figure 7.3, the correlation coefficient is set to $\rho = 0.95$. The number of snapshots is increased to $T = 200$. Note that spatial smoothing [SWK85, EEFSRJ82] or forward-backward averaging [PK89] is not applied. DML consistently outperforms other considered estimators in the inspected SNR region. The threshold of root-MUSIC occurs at a slightly lower SNR than that of the PR methods. On the other hand, in the high SNR region, all estimators under the PR framework have a lower RMSE than root-MUSIC. However, the improvement in RMSE is negligible. Concerning the two performance bounds, we observe a substantial gap between PR-CRB and the conventional CRB. This observation suggests one drawback of all PR methods, namely degraded estimation performance in the case of correlated source signals. In the high SNR regime where two sources are resolved in all Monte-Carlo trials, the MSE performance of PR estimators is accurately described by the PR-CRB. A more comprehensive investigation of the correlation factor ρ on the error performance is illustrated in the next experiment.

7.3 Influence of Correlation Factor ρ

In this experiment, we vary the correlation factor ρ defined in (7.2) between 0 and 0.99, which represents the uncorrelated case and the highly correlated cases, respectively. The number of snapshots T is set to $T = 100$, and the SNR is set to SNR = 10 dB. As observed in Figure 7.4, as the correlation factor increases, the performance of the proposed PR estimators degrades, which is partly predicted by the expression of the PR-CRB. The difference in RMSE between the PR-CRB and the conventional CRB increases as the correlation factor ρ increases. In contrast to other investigated DOA estimators, the MSE of the MUSIC estimator does not reach the PR-CRB even if the source signals are uncorrelated. This phenomenon is explained from the radial error components of the root associated with the MUSIC null-spectrum [GVL13] in the closely spaced source signals case. PR-CCF and PR-UCF exhibit similar error performances where the RMSE of the estimates departs from the PR-CRB only if the correlation factor is larger than 0.8. Such error behaviors of PR-CCF and PR-UCF indicate a higher resolution capability in the presence of correlated source signals compared to PR-DML and PR-WSF. Nevertheless, the performance of all investigated DOA estimators except for DML highly degrades as the correlation factor increases. This observation again confirms the disadvantages of the PR methods in estimating the DOAs of the correlated source signals.

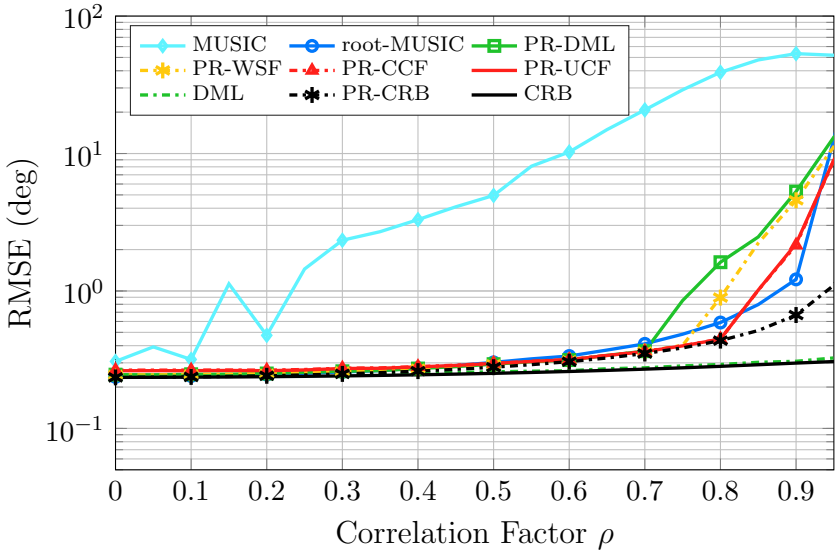

Figure 7.4. SNR = 10 dB, number of snapshots $T = 100$

7.4 Influence of Number of Snapshots T

In the next experiment depicted in Figure 7.5, the SNR is fixed at 3 dB, and the number of snapshots T is varied between 10 and 10000. Similar to previous setups concerning uncorrelated souce signals, no visible difference in RMSE between PR-CRB and the conventional CRB is observed from the simulation results in the investigated region. The RMSE performance of PR-UCF/PR-CCF resembles that of DML, achieving the asymptotic region at $T = 30$ samples, which is approximately an order of magnitude lower in the required snapshots than that for PR-DML/PR-WSF to reliably detect two sources. However, in the asymptotic region, the RMSE of PR-CCF and PR-UCF is not as close to the CRB as that of root-MUSIC, PR-DML or PR-WSF. PR-WSF outperforms PR-DML consistently in this simulation. In general, the proposed PR methods outperform MUSIC in both threshold and asymptotic regimes. This numerical result suggests that the choice of the weighting matrix $\boldsymbol{W}$ plays a crucial role in the estimation performance of the PR-WSF estimator in the non-asymptotic regime. In other words, by choosing a suitable weighting matrix $\boldsymbol{W}$, the asymptotic performance of the PR-WSF estimator remains unchanged as proven by Theorem 6.9, while the error performance of the PR-WSF estimator in the threshold region may be improved. The choice of the weighting matrix $\boldsymbol{W}$ for the optimal threshold performance is the subject of future research.

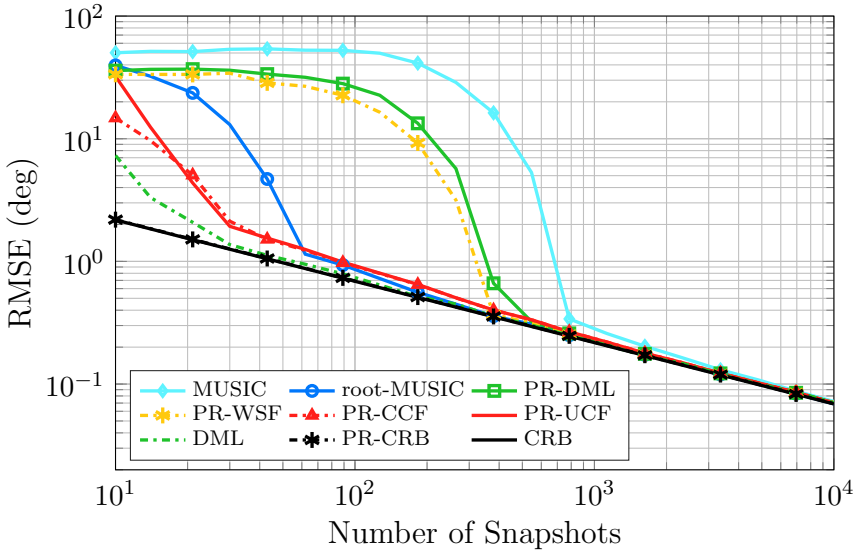

Figure 7.5. Uncorrelated source signals, SNR = 3 dB

7.5 Influence of Angular Separation $\Delta\theta$

In the next experiment, we investigate the estimation performance depending on the angular separation between two source signals. The DOA of the first source signal is fixed at $\theta_1 = 45°$ and the angular separation between two sources $\Delta\theta$ is varied from 0.5° to 6°. The number of available snapshots is set to $T = 100$. In Figure 7.6, similar to DML, the RMSE of PR-CCF/PR-UCF is close to the Cramer-Rao Bound even when the angular separation $\Delta\theta$ is as small as 1.25°, which is significantly smaller than the angular separation required for PR-DML/PR-WSF to resolve two sources. However, the RMSE of PR-CCF/PR-UCF slowly achieves the CRB only when $\Delta\theta > 5°$. PR-WSF slightly outperforms PR-DML, and both algorithms outperform MUSIC consistently in the investigated scenarios. Similar with the results in Section 7.4, this numerical result signifies the importance of the weighting matrix $\boldsymbol{W}$ in the error performance of the PR-WSF estimator in the threshold region.

7.6 Execution time

In Figure 7.7, the execution times in seconds of the DOA estimation algorithms with respect to the number of antennas M are depicted. We do not include the execution time of the brute-force DML method due to the extremely high execution time. The

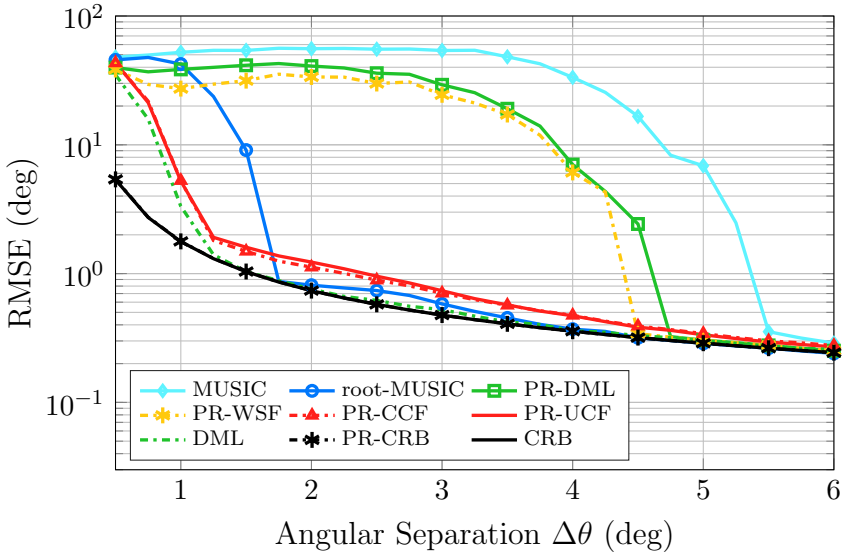

Figure 7.6. Uncorrelated source signals, SNR = 10 dB, number of snapshots $T = 100$

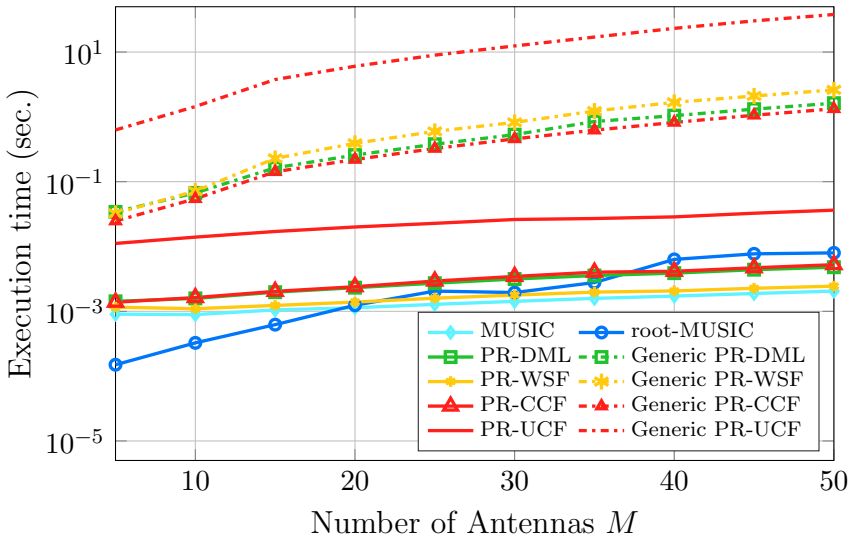

Figure 7.7. SNR = 10 dB, number of snapshots $T = 100$

FOV is partitioned uniformly into $N_G = 1800$ directions. The term *Generic* in Figure 7.7 refers to the naive implementation using the MATLAB command `eig` for the eigenvalue decomposition, which does not exploit any underlying structure of the matrix argument. The rooting process applied to root-MUSIC relies on determining the eigenvalues of the companion matrix associated with the polynomial, and therefore the execution time increases drastically with respect to the number of antennas M. All proposed PR methods, except for the PR-UCF estimator, follow similar trends as MUSIC, where the execution time is in the same order of magnitude. The execution time of the PR-UCF estimator is approximately ten times larger than other PR methods due to the number of bisection steps involved in each angular direction. Nevertheless, the PR-UCF estimator using the efficient eigenvalue decomposition based on the successive rational function approximation in Chapter 5 requires less execution time than the direct implementation with the MATLAB command. More precisely, thanks to the quadratic convergence behavior of Algorithm 2, the execution time is reduced by a factor of 20 to 1000 in comparison with the direct implementation using the generic command `eig()` in MATLAB. PR-WSF exhibits almost identical execution time behavior as MUSIC, indicating the possibility of applying the PR methods in practical cases.

7.7 Summary

In this chapter, the performance of the proposed estimators under the PR framework through extensive numerical simulations have been studied. The MSE and the execution time of the proposed estimators are computed and compared with that of state-of-the-art algorithms in the literature. Concerning the lower bound for the MSE for all PR estimators, we observe that the PR-CRB is approximately equal to the conventional CRB if the source signals are uncorrelated, which is in accordance with Theorem 6.8 in Chapter 6. Although the proposed lower bound PR-CRB is derived under the assumption that the number of snapshots T tends to infinity, from the simulation results, the expression of the PR-CRB can also accurately describe the MSE performance of the proposed PR estimators in the case of a finite number of snapshots as well. In the non-asymptotic regime, it is observed that the threshold performance of the proposed PR estimators outperforms that of the corresponding single-source approximation methods. Furthermore, given that the source signals are uncorrelated, the threshold performance of the low-complexity PR-CCF and PR-UCF estimators resembles that of the conventional high-complexity brute-force DML estimator. Such attractive performance properties of the proposed PR estimators are of great interest in practical scenarios where both the number of snapshots and the SNR are limited.

Chapter 8

Conclusions and Outlook

In this thesis, we introduce a novel class of low-complexity DOA estimators referred to as the Partial Relaxation (PR) framework. Under the proposed framework, four estimators are derived from different optimizing criteria. In addition, theoretical and practical aspects associated with the PR framework are investigated.

In Chapter 3, after revising state-of-the-art DOA estimation algorithms in the literature, we categorize conventional DOA estimators into two groups, namely methods employing high-complexity multi-dimensional search and low-complexity single-source approximation counterparts. Motivated by two aforementioned main groups of DOA estimators, in Chapter 4, we propose the PR framework, which is based on the principle of relaxing the sensor array manifold. Instead of enforcing the full structure on the steering matrix when formulating the DOA estimation problem as in the multi-dimensional search estimators, in the PR framework, only the structure in the steering vector of one source of interest is preserved, while the structure of the remaining interfering sources is relaxed. The solutions of the aforementioned optimizing problems, which are the null-spectra of the PR-based DOA estimators, involve only particular eigenvalues of modified covariance matrices. The eigenvalue-depending characteristic of the estimators under the PR framework is different from that of the conventional subspace-based DOA estimators, where the eigenvectors are considered and the eigenvalues are in general neglected.

Since particular eigenvalues are extensively required in the spectral search, developing an efficient method for computing the null-spectra of the proposed estimators is of great interest. In Chapter 5, we present a scalable numerical procedure which reduces the computational complexity of all estimators under the PR framework. By reformulating the matrix argument in the null-spectra of PR estimators as a rank-one modification of a diagonal matrix, the eigenvalues involved in the null-spectra of PR estimators are efficiently computed. Furthermore, the proposed approach allows any particular subset of eigenvalues to be computed in parallel. As a result, the execution time of the PR methods adopting the efficient implementation is significantly reduced when compared with the full eigenvalue decomposition approach.

In Chapter 6, the theoretical asymptotic performance of the estimators under the PR framework is studied. Under the induced PR signal model, the PR-CRB, which is the

theoretical asymptotic performance bound for the MSE of all unbiased PR estimators, is derived in closed-form. From the expression of the proposed PR-CRB, multiple conclusions are drawn and compared with results in the literature. Using a similar property of the corresponding WSF estimator under the conventional signal model, it is proven that the PR-WSF estimator achieves the PR-CRB asymptotically as the number of snapshots tends to infinity. However, different from the conventional WSF estimator, the asymptotic MSE of the PR-WSF estimator is independent of the weighting matrix. Furthermore, we prove that asymptotically, the error performance of the PR-WSF estimator is identical to that of the PR-DML and the MUSIC estimator. In addition, thanks to the implications obtained from the proposed PR-CRB, the asymptotic error performance of the MUSIC estimator is rediscovered indirectly through the PR-CRB.

In Chapter 7, the error performance of the proposed PR estimators are investigated through comprehensive numerical experiments. In addition, the DOA estimation performance of the PR estimators is compared with the state-of-the-art algorithms and the PR-CRB developed in Chapter 6. Simulation results show that the theoretical bound PR-CRB accurately characterizes the error performance of the proposed PR estimators if the number of snapshots or the SNR is sufficiently high. Furthermore, our four proposed low-complexity PR estimators outperform MUSIC in all investigated scenarios without exploiting any particular structure of the sensor array, e.g., Vandermonde structure from a uniform linear array. The PR estimators based on the covariance fitting problems, namely PR-CCF and PR-UCF, exhibit comparable performance as the computationally expensive DML estimator in the threshold region, i.e., in the case of a low number of snapshots, low SNR or closely spaced sources. Compared with the PR-UCF estimator, the inner approximation in the PR-CCF estimator helps to reduce the computational complexity without notably sacrificing the error performance. Also from the simulation results, we observe that the performance of the PR-WSF estimator in the non-asymptotic region is not as remarkable as the PR-UCF or PR-CCF estimator. Nevertheless, the PR-WSF estimator is still favorable in specific circumstances due to the excellent asymptotic behavior and the lower computational complexity, especially if the number of sensors is high.

From the experimental results, a drawback of the covariance fitting variants is the slight deviation from the PR-CRB in the asymptotic region. However, such slight deviation in the asymptotic region does not limit the potential of applying the PR-UCF and PR-CCF estimator for DOA estimation in practical applications. In fact, the corresponding asymptotic assumptions may not be fulfilled in real scenarios. Due to the changes in the propagating environment, the data collection time is generally limited, or the receive signal power can be low. In the aforementioned challenging scenarios, the covariance fitting-based PR estimators are the methods of choice thanks

to the excellent estimation performance in the threshold region and low computational complexity.

One aspect that has not yet been covered in this thesis is the theoretical asymptotic error performance for PR estimators based on the covariance matching criteria. In addition, the error behavior of the proposed PR estimators in the non-asymptotic region is another research topic that requires further investigation. In this dissertation, the improved threshold performance of the PR estimators compared with conventional spectral-based algorithms is revealed from the numerical simulations. However, to date, no theoretical explanation of the excellent threshold behavior of the PR estimators is available. Therefore, for future work, an analytical investigation of the error performance of PR-based DOA estimators in the threshold region with a limited number of snapshots or low SNR is of great interest. Extensions of the PR framework to different applications with more general setups, for example, sparse reconstruction [SPP18], non-coherent processing [SPPZ18] and so on, are topics of future research.

Appendix

A Mathematical Identities

We provide in the following a list of standard results that are used in this dissertation.

1. For any matrices $\boldsymbol{\Lambda}$, $\boldsymbol{\Gamma}$ and $\boldsymbol{\Xi}$ of suitable dimensions and assuming that the inverses below exist, following identities hold

$$\boldsymbol{\Pi}^{\perp}_{[\boldsymbol{\Lambda}\ \boldsymbol{\Gamma}]} = \boldsymbol{\Pi}^{\perp}_{\boldsymbol{\Lambda}} - \boldsymbol{\Pi}_{\boldsymbol{\Pi}^{\perp}_{\boldsymbol{\Lambda}}\boldsymbol{\Gamma}} \tag{A.1}$$

$$\mathrm{Re}\left\{\boldsymbol{\Lambda}\boldsymbol{\Gamma}\right\} = \mathrm{Re}\left\{\boldsymbol{\Lambda}\right\}\mathrm{Re}\left\{\boldsymbol{\Gamma}\right\} - \mathrm{Im}\left\{\boldsymbol{\Lambda}\right\}\mathrm{Im}\left\{\boldsymbol{\Gamma}\right\} \tag{A.2}$$

$$\mathrm{Im}\left\{\boldsymbol{\Lambda}\boldsymbol{\Gamma}\right\} = \mathrm{Re}\left\{\boldsymbol{\Lambda}\right\}\mathrm{Im}\left\{\boldsymbol{\Gamma}\right\} + \mathrm{Im}\left\{\boldsymbol{\Lambda}\right\}\mathrm{Re}\left\{\boldsymbol{\Gamma}\right\} \tag{A.3}$$

$$\left(\boldsymbol{\Lambda}\otimes\boldsymbol{\Gamma}\right)^{-1} = \boldsymbol{\Lambda}^{-1}\otimes\boldsymbol{\Gamma}^{-1} \tag{A.4}$$

$$\mathrm{tr}\left(\boldsymbol{\Lambda}^{\mathrm{H}}\boldsymbol{\Gamma}\right) = \mathrm{vec}\left(\boldsymbol{\Lambda}\right)^{\mathrm{H}}\mathrm{vec}\left(\boldsymbol{\Gamma}\right) \tag{A.5}$$

$$\mathrm{vec}\left(\boldsymbol{\Gamma}\boldsymbol{\Xi}\boldsymbol{\Lambda}\right) = \left(\boldsymbol{\Lambda}^{\mathrm{T}}\otimes\boldsymbol{\Gamma}\right)\mathrm{vec}\left(\boldsymbol{\Xi}\right) \tag{A.6}$$

$$\det\left(\boldsymbol{I}_M + \boldsymbol{\Lambda}\boldsymbol{\Gamma}\right) = \det\left(\boldsymbol{I}_N + \boldsymbol{\Gamma}\boldsymbol{\Lambda}\right) \tag{A.7}$$

$$\begin{bmatrix}\mathrm{Re}\left\{\boldsymbol{\Lambda}\right\} & -\mathrm{Im}\left\{\boldsymbol{\Lambda}\right\}\\ \mathrm{Im}\left\{\boldsymbol{\Lambda}\right\} & \mathrm{Re}\left\{\boldsymbol{\Lambda}\right\}\end{bmatrix}^{-1} = \begin{bmatrix}\mathrm{Re}\left\{\boldsymbol{\Lambda}^{-1}\right\} & -\mathrm{Im}\left\{\boldsymbol{\Lambda}^{-1}\right\}\\ \mathrm{Im}\left\{\boldsymbol{\Lambda}^{-1}\right\} & \mathrm{Re}\left\{\boldsymbol{\Lambda}^{-1}\right\}\end{bmatrix}. \tag{A.8}$$

2. Assuming that an invertible matrix $\boldsymbol{\Lambda}$ with the partition $\boldsymbol{\Lambda} = \begin{bmatrix}\boldsymbol{\Lambda}_{11} & \boldsymbol{\Lambda}_{12}\\ \boldsymbol{\Lambda}_{21} & \boldsymbol{\Lambda}_{22}\end{bmatrix}$ where the submatrices $\boldsymbol{\Lambda}_{11}$ and $\boldsymbol{\Lambda}_{22}$ are invertible. The inverse of the matrix $\boldsymbol{\Lambda}$ is computed from the block matrix inversion lemma as follows:

$$\begin{aligned}
&\begin{bmatrix}\boldsymbol{\Lambda}_{11} & \boldsymbol{\Lambda}_{12}\\ \boldsymbol{\Lambda}_{21} & \boldsymbol{\Lambda}_{22}\end{bmatrix}^{-1}\\
&= \begin{bmatrix}\boldsymbol{\Lambda}_{11}^{-1} + \boldsymbol{\Lambda}_{11}^{-1}\boldsymbol{\Lambda}_{12}\left(\boldsymbol{\Lambda}_{22} - \boldsymbol{\Lambda}_{21}\boldsymbol{\Lambda}_{11}^{-1}\boldsymbol{\Lambda}_{12}\right)^{-1}\boldsymbol{\Lambda}_{21}\boldsymbol{\Lambda}_{11}^{-1} & -\boldsymbol{\Lambda}_{11}^{-1}\boldsymbol{\Lambda}_{12}\left(\boldsymbol{\Lambda}_{11} - \boldsymbol{\Lambda}_{12}\boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Lambda}_{21}\right)^{-1}\\ -\left(\boldsymbol{\Lambda}_{22} - \boldsymbol{\Lambda}_{21}\boldsymbol{\Lambda}_{11}^{-1}\boldsymbol{\Lambda}_{12}\right)^{-1}\boldsymbol{\Lambda}_{21}\boldsymbol{\Lambda}_{11}^{-1} & \left(\boldsymbol{\Lambda}_{22} - \boldsymbol{\Lambda}_{21}\boldsymbol{\Lambda}_{11}^{-1}\boldsymbol{\Lambda}_{12}\right)^{-1}\end{bmatrix}\\
&= \begin{bmatrix}\left(\boldsymbol{\Lambda}_{11} - \boldsymbol{\Lambda}_{12}\boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Lambda}_{21}\right)^{-1} & -\left(\boldsymbol{\Lambda}_{11} - \boldsymbol{\Lambda}_{12}\boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Lambda}_{21}\right)^{-1}\boldsymbol{\Lambda}_{12}\boldsymbol{\Lambda}_{22}^{-1}\\ -\boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Lambda}_{21}\left(\boldsymbol{\Lambda}_{11} - \boldsymbol{\Lambda}_{12}\boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Lambda}_{21}\right)^{-1} & \boldsymbol{\Lambda}_{22}^{-1} + \boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Lambda}_{21}\left(\boldsymbol{\Lambda}_{11} - \boldsymbol{\Lambda}_{12}\boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Lambda}_{21}\right)^{-1}\boldsymbol{\Lambda}_{12}\boldsymbol{\Lambda}_{22}^{-1}\end{bmatrix}
\end{aligned} \tag{A.9}$$

   In the special case that the upper-block matrix reduces to a scalar Λ_{11} and $\det\left(\boldsymbol{\Lambda}_{22}\right) \neq 0$, we define the Schur complement to the first entry of the matrix $\boldsymbol{\Lambda}$, denoted by $\mathcal{S}\left(\boldsymbol{\Lambda}\right)$, as follows

$$\mathcal{S}\left(\boldsymbol{\Lambda}\right) = \Lambda_{11} - \boldsymbol{\Lambda}_{21}^{\mathrm{H}}\boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Lambda}_{21}. \tag{A.10}$$

If the matrix $\boldsymbol{\Lambda}$ is invertible, then the upper-left entry of the inverse of the matrix $\boldsymbol{\Lambda}$, which is denoted by $\left[\boldsymbol{\Lambda}^{-1}\right]_{11}$, is computed as follows:

$$\left[\boldsymbol{\Lambda}^{-1}\right]_{11} = \boldsymbol{e}_1^{\mathrm{T}} \boldsymbol{\Lambda}^{-1} \boldsymbol{e}_1 = \mathcal{S}\left(\boldsymbol{\Lambda}\right)^{-1}. \tag{A.11}$$

In addition, the first column of the inverse $\boldsymbol{\Lambda}^{-1}$ is computed using the block matrix inversion lemma as

$$\boldsymbol{\Lambda}^{-1} \boldsymbol{e}_1 = \mathcal{S}\left(\boldsymbol{\Lambda}\right)^{-1} \begin{bmatrix} 1 \\ -\boldsymbol{\Lambda}_{22}^{-1} \boldsymbol{\Lambda}_{21} \end{bmatrix}. \tag{A.12}$$

B Proof of Theorem 4.1

The proof of Theorem 4.1, which plays a crucial role in the derivation of the null-spectra of PR-DML and PR-WSF, is provided in this section. In order to prove the identity in (4.7), we first introduce two important lemmas:

Lemma B.1. *Let $\boldsymbol{B} \in \mathbb{C}^{M\times(N-1)}$ be a non-zero matrix and $\boldsymbol{a} \in \mathbb{C}^{M\times 1}$ be a non-zero vector such that $\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}$ is a non-zero matrix. Then there exists a matrix $\boldsymbol{Z} \in \mathbb{C}^{M\times N'}$ with $1 \leq N' \leq N-1$ such that the following conditions are satisfied:*

$$\boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}} = \boldsymbol{Z}\boldsymbol{Z}^{\mathrm{H}} \tag{B.1a}$$

$$\boldsymbol{Z}^{\mathrm{H}}\boldsymbol{Z} = \boldsymbol{I}_{N'} \tag{B.1b}$$

$$\boldsymbol{Z}^{\mathrm{H}}\boldsymbol{a} = \boldsymbol{0}. \tag{B.1c}$$

Proof of Lemma B.1. First expand $\boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}}$ in the following manner:

$$\begin{aligned} \boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}} &= \boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}\left(\left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}\right)^{\mathrm{H}}\left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}\right)\right)^{-1}\left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}\right)^{\mathrm{H}} \\ &= \boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}\left(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}. \end{aligned} \tag{B.2}$$

Therefore, the rank of the matrix $\boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}}$ is bounded by:

$$\begin{aligned} 1 \leq N' &= \operatorname{rank}\left(\boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}}\right) \\ &\leq \min\left\{\operatorname{rank}\left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\right), \operatorname{rank}\left(\boldsymbol{B}\right), \operatorname{rank}\left(\left(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}\right)\right)\right\} \leq N-1. \end{aligned} \tag{B.3}$$

Furthermore, since $\boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}}$ contains only the eigenvalues 0 and 1, taking the eigenvalue decomposition of $\boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}}$ leads to:

$$\boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}} = \boldsymbol{Z}\boldsymbol{Z}^{\mathrm{H}}, \tag{B.4}$$

and the number of columns of $\boldsymbol{Z}$ is equal to N'. Clearly the matrix $\boldsymbol{Z} \in \mathbb{C}^{M \times N'}$ satisfies $\boldsymbol{Z}^{\mathrm{H}}\boldsymbol{Z} = \boldsymbol{I}_{N'}$. Therefore, the conditions in (B.1a) and (B.1b) are satisfied if $\boldsymbol{Z}$ is chosen as in (B.4). Finally, we observe that:

$$\begin{aligned} \boldsymbol{a}^{\mathrm{H}}\boldsymbol{Z}\boldsymbol{Z}^{\mathrm{H}}\boldsymbol{a} &= \boldsymbol{a}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}}\boldsymbol{a} \\ &= \boldsymbol{a}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}\left(\boldsymbol{B}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}\right)^{-1}\boldsymbol{B}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{a} = 0. \end{aligned} \tag{B.5}$$

The identity in (B.1c) follows immediately from (B.5). □

Lemma B.2. *Let $\boldsymbol{a} \in \mathbb{C}^{M\times 1}$ be a non-zero vector and $\boldsymbol{Q} \in \mathbb{C}^{M\times M}$ be a non-zero Hermitian positive semidefinite matrix. Then the eigenvectors which correspond to non-zeros eigenvalues of $\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{Q}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}$ are orthogonal to $\boldsymbol{a}$.*

Proof of Lemma B.2. Let $\boldsymbol{x}$ be an eigenvector corresponding to an eigenvalue $\lambda \neq 0$ of $\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{Q}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}$, then by definition:

$$\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{Q}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{x} = \lambda\boldsymbol{x}. \tag{B.6}$$

Taking the conjugate transpose of (B.6) and multiplying with $\boldsymbol{a}$ on the right, we obtain:

$$0 = \boldsymbol{x}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{Q}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{a} = \left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{Q}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{x}\right)^{\mathrm{H}}\boldsymbol{a} = \lambda^{*}\boldsymbol{x}^{\mathrm{H}}\boldsymbol{a}. \tag{B.7}$$

From (B.7) and the assumption that λ is non-zero, we can conclude that $\boldsymbol{x}$ is orthogonal to $\boldsymbol{a}$. □

Now we return to the main proof of (4.7). In the case that $\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B} = \boldsymbol{0}$, then by convention in Section 4.2.1, we obtain that $\mathrm{tr}\left(\boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}}\boldsymbol{Q}\right) = 0$. On the other hand, if the matrix $\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}$ is a non-zero matrix, applying the decomposition (B.1) in Lemma B.1 and noting that $\boldsymbol{Z}^{\mathrm{H}}\boldsymbol{a} = \boldsymbol{0}$, the objective function in (4.7) is rewritten as follows:

$$\begin{aligned} \mathrm{tr}\left(\boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}}\boldsymbol{Q}\right) &= \mathrm{tr}\left(\boldsymbol{Z}\boldsymbol{Z}^{\mathrm{H}}\boldsymbol{Q}\right) = \mathrm{tr}\left(\boldsymbol{Z}^{\mathrm{H}}\boldsymbol{Q}\boldsymbol{Z}\right) \\ &= \mathrm{tr}\left(\boldsymbol{Z}^{\mathrm{H}}\left(\boldsymbol{\Pi}_{\boldsymbol{a}} + \boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\right)\boldsymbol{Q}\left(\boldsymbol{\Pi}_{\boldsymbol{a}} + \boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\right)\boldsymbol{Z}\right) \\ &= \mathrm{tr}\left(\boldsymbol{Z}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{Q}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{Z}\right). \end{aligned} \tag{B.8}$$

Therefore, the optimization problem in (4.7) is reformulated as:

$$\underset{\boldsymbol{Z}\in\mathbb{C}^{M\times N'}}{\text{maximize}}\ \mathrm{tr}\left(\boldsymbol{Z}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{Q}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{Z}\right) \tag{B.9a}$$

$$\text{subject to } \boldsymbol{Z}^{\mathrm{H}}\boldsymbol{Z} = \boldsymbol{I}_{N'} \tag{B.9b}$$

$$\boldsymbol{Z}^{\mathrm{H}}\boldsymbol{a} = \boldsymbol{0}, \tag{B.9c}$$

with $1 \leq N' \leq N-1$. Dropping the constraint $\boldsymbol{Z}^{\mathrm{H}}\boldsymbol{a} = \boldsymbol{0}$ in (B.9c), we obtain the relaxed optimization problem:

$$\underset{\boldsymbol{Z}\in\mathbb{C}^{M\times N'}}{\text{maximize}} \ \mathrm{tr}\left(\boldsymbol{Z}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{Q}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{Z}\right) \tag{B.10a}$$

$$\text{subject to } \boldsymbol{Z}^{\mathrm{H}}\boldsymbol{Z} = \boldsymbol{I}_{N'}. \tag{B.10b}$$

From the Ky-Fan inequality in [Fan49], the optimization in the relaxed problem in (B.10) admits a maximizer $\hat{\boldsymbol{Z}}$ whose columns form an orthonormal basis of the eigenspace associated with the N'-largest eigenvalues of $\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{Q}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}$. However, Lemma B.2 implies that any maximizer $\hat{\boldsymbol{Z}}$ of (B.10) also satisfies (B.9c), i.e., $\hat{\boldsymbol{Z}}^{\mathrm{H}}\boldsymbol{a} = \boldsymbol{0}$. Therefore, any maximizer $\hat{\boldsymbol{Z}}$ of the optimization problem in (B.10) is also a maximizer of (B.9). As a consequence, we obtain the following result:

$$\begin{aligned}\sum_{k=1}^{N'} \lambda_k\left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{Q}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\right) = \max \ & \mathrm{tr}\left(\boldsymbol{Z}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{Q}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{Z}\right) \\ \text{subject to } & \boldsymbol{Z} \in \mathbb{C}^{M\times N'} \\ & \boldsymbol{Z}^{\mathrm{H}}\boldsymbol{Z} = \boldsymbol{I}_{N'}, \\ & \boldsymbol{Z}^{\mathrm{H}}\boldsymbol{a} = \boldsymbol{0}.\end{aligned} \tag{B.11}$$

Combining (B.3), (B.8) and (B.11), the following identity is obtained:

$$\begin{aligned}\max_{\boldsymbol{B}\in\mathbb{C}^{M\times(N-1)}} \mathrm{tr}\left(\boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}}\boldsymbol{Q}\right) &= \sum_{k=1}^{N-1} \lambda_k\left(\boldsymbol{\Pi}_{a}^{\perp}\boldsymbol{Q}\boldsymbol{\Pi}_{a}^{\perp}\right) \\ &= \sum_{k=1}^{N-1} \lambda_k\left(\boldsymbol{\Pi}_{a}^{\perp}\boldsymbol{Q}\right).\end{aligned} \tag{B.12}$$

The optimum in (B.12) is achieved if we choose one matrix $\boldsymbol{B} \in \mathbb{C}^{M\times(N-1)}$ such that $\boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}} = \boldsymbol{Z}\boldsymbol{Z}^{\mathrm{H}}$ and the columns of $\boldsymbol{Z}$ form an orthonormal basis of the eigenspace associated with $(N-1)$-principal eigenvalues of $\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{Q}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}$.

C Alternative Derivation of the PR-DML Estimator

In this section, we provide an alternative derivation of the PR-DML estimator based on the best rank approximation problem [EY36]. For notational convenience, we drop the dependence of the steering vector $\boldsymbol{a}(\theta)$ on the direction θ, i.e., $\boldsymbol{a} = \boldsymbol{a}(\theta)$. The inner optimization of the PR-DML estimator corresponding to (4.4) is reformulated as follows:

$$\min_{\boldsymbol{s}^{\mathrm{T}},\boldsymbol{B},\boldsymbol{J}} \left|\left|\boldsymbol{X} - \boldsymbol{a}\boldsymbol{s}^{\mathrm{T}} - \boldsymbol{B}\boldsymbol{H}\right|\right|_{\mathrm{F}}^{2}, \tag{C.1}$$

where $\boldsymbol{B} \in \mathbb{C}^{M\times(N-1)}$ represents the relaxed steering matrix of $(N-1)$ undesired source signals. The newly introduced variables $\boldsymbol{s}^{\mathrm{T}} \in \mathbb{C}^{1\times T}$ and $\boldsymbol{H} \in \mathbb{C}^{(N-1)\times T}$ represent the signals from the desired direction ϑ, and the remaining undesired directions, respectively. In order to solve the optimization problem in (C.1), we fix $\boldsymbol{B}$ and $\boldsymbol{H}$, optimize with respect to $\boldsymbol{s}^{\mathrm{T}}$ and then substitute the solution back to objective function in (C.1) to obtain:

$$\begin{aligned}\min_{\boldsymbol{s}^{\mathrm{T}}\in\mathbb{C}^{1\times T}} \left|\left|\boldsymbol{X}-\boldsymbol{B}\boldsymbol{H}-\boldsymbol{a}\boldsymbol{s}^{\mathrm{T}}\right|\right|_{\mathrm{F}}^2 &= \left|\left|\boldsymbol{X}-\boldsymbol{B}\boldsymbol{H}-\boldsymbol{a}\boldsymbol{a}^{\dagger}\left(\boldsymbol{X}-\boldsymbol{B}\boldsymbol{H}\right)\right|\right|_{\mathrm{F}}^2 \\ &= \left|\left|\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{X}-\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}\boldsymbol{H}\right|\right|_{\mathrm{F}}^2.\end{aligned} \tag{C.2}$$

By substituting $\bar{\boldsymbol{E}} = \boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{B}\boldsymbol{H}$ and note that $\operatorname{rank}\left(\bar{\boldsymbol{E}}\right) \leq N-1$, the inner optimization problem in (C.2) is equivalent to the following optimization problem:

$$\min_{\bar{\boldsymbol{E}}\in\mathbb{C}^{M\times T}} \quad \left|\left|\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{X}-\bar{\boldsymbol{E}}\right|\right|_{\mathrm{F}}^2 \tag{C.3}$$

$$\text{subject to} \quad \operatorname{rank}\left(\bar{\boldsymbol{E}}\right) \leq N-1 \tag{C.4}$$

$$\bar{\boldsymbol{E}}^{\mathrm{H}}\boldsymbol{a} = \boldsymbol{0}. \tag{C.5}$$

We note that apart from the constraint (C.5), the optimization problem in (C.3)-(C.5) is identical to the following best rank approximation problem

$$\begin{aligned}\min_{\bar{\boldsymbol{E}}\in\mathbb{C}^{M\times T}} \quad & \left|\left|\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{X}-\bar{\boldsymbol{E}}\right|\right|_{\mathrm{F}}^2 \\ \text{subject to} \quad & \operatorname{rank}\left(\bar{\boldsymbol{E}}\right) \leq N-1.\end{aligned} \tag{C.6}$$

We further remark that the minimizer of the optimization problem (C.6) satisfies the constraint (C.5) thanks to the Lemma B.2. This observation implies that any minimizer of the optimization problem (C.6) is also a minimizer of the optimization problem in (C.3)-(C.5). As a result, the PR-DML inner optimization problem in (C.1) is reformulated as follows:

$$\begin{aligned}\min_{\boldsymbol{s}^{\mathrm{T}},\boldsymbol{B},\boldsymbol{J}} \left|\left|\boldsymbol{X}-\boldsymbol{a}\boldsymbol{s}^{\mathrm{T}}-\boldsymbol{B}\boldsymbol{J}\right|\right|_{\mathrm{F}}^2 &= \sum_{k=N}^{M}\sigma_k^2(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{X}) \\ &= \sum_{k=N}^{M}\lambda_k(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\boldsymbol{X}\boldsymbol{X}\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}) = T\sum_{k=N}^{M}\lambda_k(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{R}}),\end{aligned} \tag{C.7}$$

where the last equality holds as a consequence of the identity (A.7). In (C.7), $\sigma_k(\cdot)$ denoted the k-th largest singular value of the matrix argument. Dropping the constant factor T, the result in (C.7) is identical to the null-spectrum of the PR-DML estimator in (4.8), and we conclude the proof.

D Equivalence between MUSIC and PR-WSF with Identity Weighting Matrix

In this section, we prove that MUSIC can be considered as a special case of the proposed PR-WSF estimator in (4.10) with $\boldsymbol{W} = \boldsymbol{I}_N$. Considering the expression of the null-spectrum in (4.10) for the steering vector $\boldsymbol{a} = \boldsymbol{a}(\theta)$, we note that the rank of $\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{U}}_{\mathrm{s}}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}}$ is at most N. Hence, $\lambda_k\left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{U}}_{\mathrm{s}}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}}\right) = 0$ for $k = N+1, \ldots, M$. Therefore, when calculating the null-spectrum in (4.10), only $\lambda_N\left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{U}}_{\mathrm{s}}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}}\right)$ is considered. The expression of $\lambda_N\left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{U}}_{\mathrm{s}}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}}\right)$ can be further rewritten as follows:

$$\begin{aligned}\lambda_N\left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{U}}_{\mathrm{s}}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}}\right) &= \lambda_N\left(\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}}\left(\boldsymbol{I}_M - \frac{1}{||\boldsymbol{a}||_2^2}\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\right)\hat{\boldsymbol{U}}_{\mathrm{s}}\right) \\ &= \lambda_N\left(\boldsymbol{I}_M - \frac{1}{||\boldsymbol{a}||^2}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}}\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\hat{\boldsymbol{U}}_{\mathrm{s}}\right) \\ &= 1 + \lambda_N\left(-\frac{1}{||\boldsymbol{a}||^2}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}}\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\hat{\boldsymbol{U}}_{\mathrm{s}}\right),\end{aligned} \tag{D.1}$$

where the first equality is a direct consequence of (A.7). Since $-\frac{1}{||\boldsymbol{a}||^2}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}}\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\hat{\boldsymbol{U}}_{\mathrm{s}}$ is a negative semidefinite rank-one matrix of size $N \times N$, it can be easily shown that:

$$\lambda_N\left(-\frac{1}{||\boldsymbol{a}||^2}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}}\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}}\hat{\boldsymbol{U}}_{\mathrm{s}}\right) = -\frac{1}{||\boldsymbol{a}||_2^2}\boldsymbol{a}^{\mathrm{H}}\hat{\boldsymbol{U}}_{\mathrm{s}}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}}\boldsymbol{a}. \tag{D.2}$$

Substituting (D.2) into (D.1) and using the orthogonality property between the signal and the noise subspace, we obtain:

$$\begin{aligned}\lambda_N\left(\boldsymbol{\Pi}_{\boldsymbol{a}}^{\perp}\hat{\boldsymbol{U}}_{\mathrm{s}}\hat{\boldsymbol{U}}_{\mathrm{s}}^{\mathrm{H}}\right) &= 1 - \frac{1}{||\boldsymbol{a}||_2^2}\boldsymbol{a}^{\mathrm{H}}\left(\boldsymbol{I}_M - \hat{\boldsymbol{U}}_{\mathrm{n}}\hat{\boldsymbol{U}}_{\mathrm{n}}^{\mathrm{H}}\right)\boldsymbol{a} \\ &= \frac{\boldsymbol{a}^{\mathrm{H}}\hat{\boldsymbol{U}}_{\mathrm{n}}\hat{\boldsymbol{U}}_{\mathrm{n}}^{\mathrm{H}}\boldsymbol{a}}{\boldsymbol{a}^{\mathrm{H}}\boldsymbol{a}}.\end{aligned} \tag{D.3}$$

The expression in (D.3) is identical to the null-spectrum of MUSIC. Therefore, with $\boldsymbol{W} = \boldsymbol{I}_N$, the expression in (4.9) is another equivalent formulation of the MUSIC estimator.

E On the Derivative of the PR-UCF Objective Function with Respect to σ_{s}^2

In this section, we derive the expression of the derivative of the PR-UCF objective function w.r.t. σ_{s}^2 in (4.21). By taking the eigenvalue decomposition as in (2.9) and

substituting $\boldsymbol{z} = \hat{\boldsymbol{U}}^\mathrm{H}\boldsymbol{a}$, the inner objective function of the PR-UCF in (4.20) is rewritten as:

$$g(\sigma_\mathrm{s}^2) = \sum_{k=N}^{M} \lambda_k^2 \left(\hat{\boldsymbol{\Lambda}} - \sigma_\mathrm{s}^2 \boldsymbol{z}\boldsymbol{z}^\mathrm{H}\right). \tag{E.1}$$

In the following steps, we calculate the derivative $\dfrac{\mathrm{d}\lambda_k\left(\hat{\boldsymbol{\Lambda}} - \sigma_\mathrm{s}^2 \boldsymbol{z}\boldsymbol{z}^\mathrm{H}\right)}{\mathrm{d}\sigma_\mathrm{s}^2}$. Applying the results from [Mag85, Th. 2.] leads to the following expression:

$$\frac{\mathrm{d}\lambda_k\left(\hat{\boldsymbol{\Lambda}} - \sigma_\mathrm{s}^2 \boldsymbol{z}\boldsymbol{z}^\mathrm{H}\right)}{\mathrm{d}\sigma_\mathrm{s}^2} = \frac{\bar{\boldsymbol{u}}_k^\mathrm{H} \dfrac{\mathrm{d}\left(\hat{\boldsymbol{\Lambda}} - \sigma_\mathrm{s}^2 \boldsymbol{z}\boldsymbol{z}^\mathrm{H}\right)}{\mathrm{d}\sigma_\mathrm{s}^2} \bar{\boldsymbol{u}}_k}{\bar{\boldsymbol{u}}_k^\mathrm{H}\bar{\boldsymbol{u}}_k} \tag{E.2a}$$

$$= -\frac{\bar{\boldsymbol{u}}_k^\mathrm{H}\boldsymbol{z}\boldsymbol{z}^\mathrm{H}\bar{\boldsymbol{u}}_k}{\bar{\boldsymbol{u}}_k^\mathrm{H}\bar{\boldsymbol{u}}_k}, \tag{E.2b}$$

where $\bar{\boldsymbol{u}}_k$ is an eigenvector corresponding to the eigenvalue $\lambda_k\left(\hat{\boldsymbol{\Lambda}} - \sigma_s^2 \boldsymbol{z}\boldsymbol{z}^\mathrm{H}\right)$. Interestingly, the expression on the numerator of (E.2b) can be shown to be independent of the eigenvectors. In fact, by using the shorthand notation $\bar{\lambda}_k(\sigma_\mathrm{s}^2) = \lambda_k\left(\hat{\boldsymbol{\Lambda}} - \sigma_\mathrm{s}^2 \boldsymbol{z}\boldsymbol{z}^\mathrm{H}\right)$ as in (4.22), and applying Identity 1 and 3 from Theorem 5.1 to the matrix $\hat{\boldsymbol{\Lambda}} - \sigma_\mathrm{s}^2 \boldsymbol{z}\boldsymbol{z}^\mathrm{H}$, we obtain:

$$0 = 1 - \sigma_\mathrm{s}^2 \boldsymbol{z}^\mathrm{H}\left(\hat{\boldsymbol{\Lambda}} - \bar{\lambda}\left(\sigma_\mathrm{s}^2\right)\boldsymbol{I}_M\right)^{-1}\boldsymbol{z}, \tag{E.3}$$

$$\bar{\boldsymbol{u}}_k = \left(\hat{\boldsymbol{\Lambda}} - \bar{\lambda}_k\left(\sigma_\mathrm{s}^2\right)\boldsymbol{I}_M\right)^{-1}\boldsymbol{z}. \tag{E.4}$$

Substituting (E.3) and (E.4) into (E.2b), the derivative of $\lambda_k\left(\hat{\boldsymbol{\Lambda}} - \sigma_\mathrm{s}^2 \boldsymbol{z}\boldsymbol{z}^\mathrm{H}\right)$ with respect to σ_s^2 is given by:

$$\begin{aligned}
\frac{\mathrm{d}\lambda_k\left(\hat{\boldsymbol{\Lambda}} - \sigma_\mathrm{s}^2 \boldsymbol{z}\boldsymbol{z}^\mathrm{H}\right)}{\mathrm{d}\sigma_\mathrm{s}^2} &= -\frac{\bar{\boldsymbol{u}}_k^\mathrm{H}\boldsymbol{z}\boldsymbol{z}^\mathrm{H}\bar{\boldsymbol{u}}_k}{\bar{\boldsymbol{u}}_k^\mathrm{H}\bar{\boldsymbol{u}}_k} \\
&= -\frac{\boldsymbol{z}^\mathrm{H}\left(\hat{\boldsymbol{\Lambda}} - \bar{\lambda}_k\left(\sigma_\mathrm{s}^2\right)\boldsymbol{I}_M\right)^{-1}\boldsymbol{z}\boldsymbol{z}^\mathrm{H}\left(\hat{\boldsymbol{\Lambda}} - \bar{\lambda}_k\left(\sigma_\mathrm{s}^2\right)\boldsymbol{I}_M\right)^{-1}\boldsymbol{z}}{\boldsymbol{z}^\mathrm{H}\left(\hat{\boldsymbol{\Lambda}} - \bar{\lambda}_k\left(\sigma_\mathrm{s}^2\right)\boldsymbol{I}_M\right)^{-2}\boldsymbol{z}} \\
&= -\frac{1}{\sigma_\mathrm{s}^4\boldsymbol{z}^\mathrm{H}\left(\hat{\boldsymbol{\Lambda}} - \bar{\lambda}_k\left(\sigma_\mathrm{s}^2\right)\boldsymbol{I}_M\right)^{-2}\boldsymbol{z}} \\
&= -\frac{1}{\sigma_\mathrm{s}^4\boldsymbol{a}^\mathrm{H}\left(\hat{\boldsymbol{R}} - \bar{\lambda}_k\left(\sigma_\mathrm{s}^2\right)\boldsymbol{I}_M\right)^{-2}\boldsymbol{a}}.
\end{aligned} \tag{E.5}$$

Taking the derivative of (E.1) by applying the identity in (E.5) concludes our proof of (4.21).

Remark 1: The expression of the derivative of $\bar{\lambda}\left(\sigma_{\mathrm{s}}^2\right) = \lambda_k\left(\hat{\boldsymbol{\Lambda}} - \sigma_{\mathrm{s}}^2 \boldsymbol{z}\boldsymbol{z}^{\mathrm{H}}\right)$ w.r.t. σ_{s}^2 can be alternatively derived by taking the total derivative of the rational equation in (E.3) w.r.t. σ_{s}^2 as follows:

$$\begin{aligned} 0 &= \frac{\mathrm{d}\left(1 - \sigma_{\mathrm{s}}^2 \boldsymbol{z}^{\mathrm{H}}\left(\hat{\boldsymbol{\Lambda}} - \bar{\lambda}\left(\sigma_{\mathrm{s}}^2\right)\boldsymbol{I}_M\right)^{-1}\boldsymbol{z}\right)}{\mathrm{d}\sigma_{\mathrm{s}}^2} \\ &= \frac{\mathrm{d}\left(1 - \sigma_{\mathrm{s}}^2 \sum\limits_{m=1}^{M} \frac{|z_m|^2}{\hat{\lambda}_m - \bar{\lambda}_k\left(\sigma_{\mathrm{s}}^2\right)}\right)}{\mathrm{d}\sigma_{\mathrm{s}}^2} \\ &= -\sum_{m=1}^{M} \frac{|z_m|^2}{\hat{\lambda}_m - \bar{\lambda}_k\left(\sigma_{\mathrm{s}}^2\right)} - \sigma_{\mathrm{s}}^2 \sum_{m=1}^{M} \frac{|z_m|^2}{\left(\hat{\lambda}_m - \bar{\lambda}_k\left(\sigma_{\mathrm{s}}^2\right)\right)^2} \frac{\mathrm{d}\bar{\lambda}_k\left(\sigma_{\mathrm{s}}^2\right)}{\mathrm{d}\sigma_{\mathrm{s}}^2} \\ &= -\frac{1}{\sigma_{\mathrm{s}}^2} - \sigma_{\mathrm{s}}^2 \frac{\mathrm{d}\bar{\lambda}_k\left(\sigma_{\mathrm{s}}^2\right)}{\mathrm{d}\sigma_{\mathrm{s}}^2} \sum_{m=1}^{M} \frac{|z_m|^2}{\left(\hat{\lambda}_m - \bar{\lambda}_k\left(\sigma_{\mathrm{s}}^2\right)\right)^2}. \end{aligned} \tag{E.6}$$

Rearranging the last expression in (E.6) yields

$$\begin{aligned} \frac{\mathrm{d}\bar{\lambda}_k\left(\sigma_{\mathrm{s}}^2\right)}{\mathrm{d}\sigma_{\mathrm{s}}^2} &= -\frac{1}{\sigma_{\mathrm{s}}^4 \sum\limits_{m=1}^{M} \frac{|z_m|^2}{\left(\hat{\lambda}_m - \bar{\lambda}_k\left(\sigma_{\mathrm{s}}^2\right)\right)^2}} \\ &= -\frac{1}{\sigma_{\mathrm{s}}^4 \boldsymbol{z}^{\mathrm{H}}\left(\hat{\boldsymbol{\Lambda}} - \bar{\lambda}\left(\sigma_{\mathrm{s}}^2\right)\boldsymbol{I}_M\right)^{-2}\boldsymbol{z}} = -\frac{1}{\sigma_{\mathrm{s}}^4 \boldsymbol{a}^{\mathrm{H}}\left(\hat{\boldsymbol{R}} - \bar{\lambda}_k\left(\sigma_{\mathrm{s}}^2\right)\boldsymbol{I}_M\right)^{-2}\boldsymbol{a}}, \end{aligned} \tag{E.7}$$

which is identical to the expression in (E.5).

Remark 2: If the Newton's method is employed to compute the minimum of the PR-UCF function instead of the bisection method, the second derivative of the modified eigenvalue $\bar{\lambda}_k\left(\sigma_{\mathrm{s}}^2\right)$ w.r.t. the scalar σ_{s}^2 is required. By taking the derivative of the expression in (E.5), we obtain:

$$\begin{aligned} \frac{\mathrm{d}^2\bar{\lambda}_k}{\mathrm{d}\sigma_{\mathrm{s}}^4} &= \frac{2}{\left(\sigma_{\mathrm{s}}^4 \boldsymbol{a}^{\mathrm{H}}\left(\hat{\boldsymbol{R}} - \bar{\lambda}_k\left(\sigma_{\mathrm{s}}^2\right)\boldsymbol{I}_M\right)^{-2}\boldsymbol{a}\right)^2} \times \\ &\quad \left(\sigma_{\mathrm{s}}^2 \boldsymbol{a}^{\mathrm{H}}\left(\hat{\boldsymbol{R}} - \bar{\lambda}_k\left(\sigma_{\mathrm{s}}^2\right)\boldsymbol{I}_M\right)^{-2}\boldsymbol{a} - \frac{\sigma_{\mathrm{s}}^4 \boldsymbol{a}^{\mathrm{H}}\left(\hat{\boldsymbol{R}} - \bar{\lambda}_k\left(\sigma_{\mathrm{s}}^2\right)\boldsymbol{I}_M\right)^{-3}\boldsymbol{a}}{\sigma_{\mathrm{s}}^4 \boldsymbol{a}^{\mathrm{H}}\left(\hat{\boldsymbol{R}} - \bar{\lambda}_k\left(\sigma_{\mathrm{s}}^2\right)\boldsymbol{I}_M\right)^{-2}\boldsymbol{a}}\right) \\ &= \frac{2}{\sigma_{\mathrm{s}}^8 \left(\sum\limits_{j=1}^{K} \frac{|z_j|^2}{\left(d_j - \bar{\lambda}_k\left(\sigma_{\mathrm{s}}^2\right)\right)^2}\right)^2} \left(\sigma_{\mathrm{s}}^2 \sum_{j=1}^{K} \frac{|z_j|^2}{\left(d_j - \bar{\lambda}_k\left(\sigma_{\mathrm{s}}^2\right)\right)^2} - \frac{\sum\limits_{j=1}^{K} \frac{|z_j|^2}{\left(d_j - \bar{\lambda}_k\left(\sigma_{\mathrm{s}}^2\right)\right)^3}}{\sum\limits_{j=1}^{K} \frac{|z_j|^2}{\left(d_j - \bar{\lambda}_k\left(\sigma_{\mathrm{s}}^2\right)\right)^2}}\right). \end{aligned} \tag{E.8}$$

F Deflation Process

In this section, we describe the deflation process [GVL13, p. 471], [BNS78, Sec. 2] to simplify the eigenvalue decomposition in (5.1) where the initial diagonal matrix $\boldsymbol{D} = \mathrm{diag}(d_1, \ldots, d_K)$ contains repeated eigenvalues, and there are one or multiple zero-valued entries in $\boldsymbol{z} = [z_1, \ldots, z_K]^T$.

(a) If there exists an index k such that $z_k = 0$, then the eigenvalue $\bar{d}_k$ in (5.1) is equal to d_k, since the k-th row and column of the diagonal matrix $\boldsymbol{D}$ are unperturbed by the rank-one matrix $\rho \boldsymbol{z}\boldsymbol{z}^{\mathrm{H}}$. The remaining eigenvalues $\bar{d}_l$ with $l \neq k$ are the eigenvalues of $\boldsymbol{D}_{\mathrm{r}} - \rho \boldsymbol{z}_{\mathrm{r}}\boldsymbol{z}_{\mathrm{r}}^{\mathrm{H}}$ where the diagonal matrix $\boldsymbol{D}_{\mathrm{r}}$ and the vector $\boldsymbol{z}_{\mathrm{r}}$ are obtained by removing the k-th entry from the diagonal matrix $\boldsymbol{D}$ and the vector $\boldsymbol{z}$, respectively.

(b) If there are two identical eigenvalues $d_k = d_i$ with $k \neq i$, we choose a Givens rotation matrix $\boldsymbol{G} = [\boldsymbol{g}_1, \ldots, \boldsymbol{g}_K]$ such that

$$\boldsymbol{g}_i^{\mathrm{H}}\boldsymbol{z} = \sqrt{|z_i|^2 + |z_k|^2} \tag{F.1a}$$

$$\boldsymbol{g}_k^{\mathrm{H}}\boldsymbol{z} = 0 \tag{F.1b}$$

$$\boldsymbol{g}_l^{\mathrm{H}}\boldsymbol{z} = z_l \text{ with } l \neq i, l \neq k. \tag{F.1c}$$

Since $\boldsymbol{G}$ is unitary and $\boldsymbol{G}^{\mathrm{H}}\boldsymbol{D}\boldsymbol{G} = \boldsymbol{D}$, the eigenvalues $\left\{\bar{d}_1, \ldots, \bar{d}_K\right\}$ of the original problem in (5.1) are identical to the eigenvalues of the matrix $\boldsymbol{G}^{\mathrm{H}}\left(\boldsymbol{D} - \rho \boldsymbol{z}\boldsymbol{z}^{\mathrm{H}}\right)\boldsymbol{G} = \boldsymbol{D} - \rho \tilde{\boldsymbol{z}}\tilde{\boldsymbol{z}}^{\mathrm{H}}$ with $\tilde{\boldsymbol{z}} = \boldsymbol{G}^{\mathrm{H}}\boldsymbol{z}$. However, the identity in (F.1b) implies $\tilde{z}_k = 0$, and therefore we can reduce this case to the case in (a).

In the deflation process, the two above mentioned steps are applied iteratively to determine all the eigenvalues which remain unchanged due to the Hermitian rank-one modification. Consequently, if the initial diagonal matrix $\boldsymbol{D}$ has an eigenvalue d_k of multiplicity $m \geq 2$, then d_k is also an eigenvalue of the rank-one modified matrix $\boldsymbol{D} - \rho \boldsymbol{z}\boldsymbol{z}^{\mathrm{H}}$ with the multiplicity of at least $(m-1)$. Furthermore, a deflated diagonal matrix $\hat{\boldsymbol{D}}$ with distinct eigenvalues and a deflated vector $\hat{\boldsymbol{z}}$ with non-zero entries are generated, which are of smaller dimensions than the original diagonal matrix $\boldsymbol{D}$ and the rank-one component vector $\boldsymbol{z}$, respectively. The remaining eigenvalues which are different from the initial eigenvalues of the matrix $\boldsymbol{D}$ are then determined by applying Algorithm 2 to the deflated matrix $\hat{\boldsymbol{D}} - \rho \hat{\boldsymbol{z}}\hat{\boldsymbol{z}}^{\mathrm{H}}$.

G Closed-form Expressions for Rational Approximation

In this section, based on [Li94], we provide the closed-form expressions for computing the required parameters in Steps 3-5 of Algorithm 2. For easier notation, we recall the definition of the rational function $R(x)$ in (5.2) as follows

$$R(x) = 1 - \bar{\rho} \sum_{l=1}^{K} \frac{|z_l|^2}{d_l - x} \quad = 1 + \psi_k(x) + \phi_k(x) \tag{G.1}$$

where the functions $\psi_l(x)$ and $\phi_l(x)$ are defined in (5.4) and (5.5), respectively. At this point, we consider two separate cases:

- **Case 1:** The k-th root of $R(x)$ is computed with $1 \leq k < K$. For notational convenience, we define

$$\psi_k = \psi_k\left(x^{(\tau)}\right) = -\bar{\rho} \sum_{l=1}^{k} \frac{|z_l|^2}{d_l - x^{(\tau)}}, \tag{G.2}$$

$$\psi_k' = \frac{\mathrm{d}\psi_k}{\mathrm{d}x}\left(x^{(\tau)}\right) = -\bar{\rho} \sum_{l=1}^{k} \frac{|z_l|^2}{\left(d_l - x^{(\tau)}\right)^2}, \tag{G.3}$$

$$\phi_k = \phi_k\left(x^{(\tau)}\right) = -\bar{\rho} \sum_{l=k+1}^{K} \frac{|z_l|^2}{d_l - x^{(\tau)}}, \tag{G.4}$$

$$\phi' = \frac{\mathrm{d}\phi_k}{\mathrm{d}x}\left(x^{(\tau)}\right) = -\bar{\rho} \sum_{l=k+1}^{K} \frac{|z_l|^2}{\left(d_l - x^{(\tau)}\right)^2}, \tag{G.5}$$

  where $x^{(\tau)}$ is the approximated root of the rational function $R(x)$ in (G.1) at the iteration τ. The closed-form expressions of the parameters for the rational approximation update in Steps 3 and 4 in Algorithm 2 are respectively computed as follows:

$$p = \psi_k - \psi_k' \delta_k, \qquad q = \psi_k' \delta_k^2, \tag{G.6}$$

$$r = \phi_k - \phi_k' \delta_{k+1}, \qquad s = \phi_k' \delta_{k+1}^2, \tag{G.7}$$

  with

$$\delta_k = d_k - x^{(\tau)} \text{ for all } k = 1, \ldots, K. \tag{G.8}$$

  Corresponding to Step 5 in Algorithm 2 and after some reformulations, the next iteration value $x^{(\tau+1)}$ which lies in the interval (d_{k+1}, d_k) is computed from the previous iteration $x^{(\tau)}$ as follows

$$x^{(\tau+1)} = x^{(\tau)} + \eta^{(\tau)}, \tag{G.9}$$

where the update $\eta^{(\tau)}$ is one solution of the quadratic equation

$$ax^2 + bx + c = 0. \tag{G.10}$$

In (G.10), the coefficients a, b and c are given by

$$\begin{aligned} a &= 1 + p + r \\ &= R(x^{(\tau)}) - \delta_k \psi' - \delta_{k+1}\phi', \end{aligned} \tag{G.11}$$

$$\begin{aligned} b &= -(1 + p + r)(\delta_k + \delta_{k+1}) - q - s \\ &= \delta_k \delta_{k+1} \frac{\mathrm{d}R}{\mathrm{d}x}\left(x^{(\tau)}\right) - (\delta_k + \delta_{k+1})\, R\left(x^{(\tau)}\right), \end{aligned} \tag{G.12}$$

$$\begin{aligned} c &= (1 + p + r)\delta_k \delta_{k+1} + q\delta_{k+1} + s\delta_k \\ &= \delta_k \delta_{k+1} R\left(x^{(\tau)}\right). \end{aligned} \tag{G.13}$$

In order to ensure that the next iteration point $x^{(\tau+1)} \in (d_{k+1}, d_k)$ and to avoid loss of significance due to subtraction, the update $\eta^{(\tau)}$ in (G.9) is computed as follows

$$\eta^{(\tau)} = \begin{cases} \dfrac{-b + \sqrt{b^2 - 4ac}}{2a} & \text{if } b \leq 0, \\ \dfrac{2c}{b + \sqrt{b^2 - 4ac}} & \text{if } b > 0. \end{cases} \tag{G.14}$$

- **Case 2:** The K-th root of $R(x)$ is computed. Using notations similar to the previous case, we have $\phi_k(x) = 0$ and therefore $r = s = 0$. The expressions for computing the parameters p and q remain unchanged as in (G.6) with $k = K$. The next iteration point $x^{(\tau+1)}$ is computed similarly to (G.9) as

$$x^{(\tau+1)} = x^{(\tau)} + \eta^{(\tau)}. \tag{G.15}$$

However, different from the expression in (G.14), the update $\eta^{(\tau)}$ in (G.15) is computed by

$$\begin{aligned} \eta^{(\tau)} &= \frac{q}{1+p} + \delta_K \\ &= \frac{R\left(x^{(\tau)}\right)\left(d_K - x^{(\tau)}\right)}{R\left(x^{(\tau)}\right) - \left(d_K - x^{(\tau)}\right)\dfrac{\mathrm{d}R}{\mathrm{d}x}\left(x^{(\tau)}\right)}. \end{aligned} \tag{G.16}$$

H On the Invertibility of the Fisher Information Matrix

In this section, the proof of the invertibility of the FIM in (6.24) is provided. A critical aspect in the proof of the invertibile FIM is the concept of local identifiability, which is

repeated in the following. Assume that the probability density function $f(\boldsymbol{y};\boldsymbol{\alpha})$ of the observation vector $\boldsymbol{y} \in \mathbb{R}^n$ is parameterized by the parameter vector $\boldsymbol{\alpha}$ in the parameter space $\mathbb{A}$. The definition of the local identifiability of a given parameter vector $\boldsymbol{\alpha}_0 \in \mathbb{A}$ is defined as follows [Rot71, Def.3]:

Definition H.1. *A parameter vector $\boldsymbol{\alpha}_0$ is said to be locally identifiable if there exists an open neighborhood of $\boldsymbol{\alpha}_0$ containing no other $\boldsymbol{\alpha} \in \mathbb{A}$ which holds*

$$f(\boldsymbol{y};\boldsymbol{\alpha}_0) = f(\boldsymbol{y};\boldsymbol{\alpha}) \text{ for all } \boldsymbol{y} \in \mathbb{R}^n. \tag{H.1}$$

Using Definition G.1, the proof of the invertible FIM in (6.24) is provided in the following. From the assumptions (As.2)-(As.3), the receive signal $\boldsymbol{X}$ in (4.2) is Gaussian distributed with zero mean. As a result, the probability density function of the receive signal in each snapshot $\boldsymbol{x}(t)$ is fully characterized by the covariance matrix $\boldsymbol{R}$ in (6.19). Write $\boldsymbol{R} = \boldsymbol{R}(\bar{\boldsymbol{\alpha}})$ to emphasize the dependence of the receive covariance matrix $\boldsymbol{R}$ in (6.19) on the parameter vector $\bar{\boldsymbol{\alpha}}$ according to the parameterization given in (6.21). We further define the ϵ-neighborhood of a vector $\boldsymbol{\alpha}_0$ as

$$\mathcal{B}_\epsilon(\boldsymbol{\alpha}_0) = \left\{ \boldsymbol{\alpha} \;\middle|\; ||\boldsymbol{\alpha} - \boldsymbol{\alpha}_0||_2 < \epsilon \right\}. \tag{H.2}$$

Under the Gaussian assumption of the source signal $\boldsymbol{s}(t)$ and the noise $\boldsymbol{n}(t)$, due to Assumption (As.5) and according to [Rot71, Th. 1], the FIM in (6.24) is invertible if and only the parameterization is locally identifiable, i.e., if there exists $\epsilon > 0$ such that

$$\boldsymbol{R}(\bar{\boldsymbol{\alpha}}_1) = \boldsymbol{R}(\bar{\boldsymbol{\alpha}}_2) \text{ with } \bar{\boldsymbol{\alpha}}_2 \in \mathcal{B}_\epsilon(\bar{\boldsymbol{\alpha}}_1) \text{ implies } \bar{\boldsymbol{\alpha}}_1 = \bar{\boldsymbol{\alpha}}_2, \tag{H.3}$$

where the parameter vectors $\bar{\boldsymbol{\alpha}}_1$ and $\bar{\boldsymbol{\alpha}}_2$ are defined conformably with (6.21) as $\bar{\boldsymbol{\alpha}}_1 = \left[\vartheta_1, \check{\boldsymbol{b}}_1^{\mathrm{T}}, \tilde{\boldsymbol{b}}_1^{\mathrm{T}}, \bar{\boldsymbol{p}}_1^{\mathrm{T}}, \nu_1\right]^T$ and $\bar{\boldsymbol{\alpha}}_2 = \left[\vartheta_2, \check{\boldsymbol{b}}_2^{\mathrm{T}}, \tilde{\boldsymbol{b}}_2^{\mathrm{T}}, \bar{\boldsymbol{p}}_2^{\mathrm{T}}, \nu_2\right]^T$, respectively. By applying the eigenvalue decomposition similarly to (2.6)

$$\boldsymbol{R}(\bar{\boldsymbol{\alpha}}_1) = \boldsymbol{R}(\bar{\boldsymbol{\alpha}}_2) = \boldsymbol{U}_{\mathrm{s}}\boldsymbol{\Lambda}_{\mathrm{s}}\boldsymbol{U}_{\mathrm{s}}^{\mathrm{H}} + \boldsymbol{U}_{\mathrm{n}}\boldsymbol{\Lambda}_{\mathrm{n}}\boldsymbol{U}_{\mathrm{n}}^{\mathrm{H}}, \tag{H.4}$$

the condition $\boldsymbol{R} = \boldsymbol{R}(\bar{\boldsymbol{\alpha}}_1) = \boldsymbol{R}(\bar{\boldsymbol{\alpha}}_2)$ implies that $\nu_1 = \nu_2$ since both values are equal to the $(M-N)$-smallest eigenvalue of $\boldsymbol{R}$. Furthermore, with ϑ_1 and ϑ_2 denoting the first element in $\bar{\boldsymbol{\alpha}}_1$ and $\bar{\boldsymbol{\alpha}}_2$, respectively, both ϑ_1 and ϑ_2 are members of the set $\mathbb{S} = \{\theta \in \Theta \,|\boldsymbol{a}(\theta) \in \operatorname{span}(\boldsymbol{U}_{\mathrm{s}})\}$ defined in (As.7). We will prove in the following that, if we choose a suitable radius $\epsilon > 0$, then ϑ_1 and ϑ_2 are identical. Clearly in the case that the set $\mathbb{S}$ contains only one element, then $\vartheta_1 = \vartheta_2$ for every value of $\epsilon > 0$. In the case that the set $\mathbb{S}$ contains more than one element, we define the scalar $\bar{\epsilon}$ as follows:

$$\bar{\epsilon} = \inf_{\substack{\theta_k, \theta_l \in \mathbb{S} \\ k \neq l}} |\theta_k - \theta_l|. \tag{H.5}$$

Intuitively, the scalar $\bar{\epsilon}$ in (H.5) represents the smallest possible distance between two arbitrary distinct elements in $\mathbb{S}$. As the set $\mathbb{S}$ is finite, $\bar{\epsilon}$ is strictly greater than zero. Choosing the radius $\epsilon = \frac{\bar{\epsilon}}{2} > 0$, we prove that $\vartheta_1 = \vartheta_2$ by contradiction: assume that $\vartheta_1 \neq \vartheta_2$. From the definition of $\bar{\epsilon}$ in (H.5), we obtain

$$|\vartheta_2 - \vartheta_1| \geq \bar{\epsilon}. \tag{H.6}$$

On the other hand, from the definition in (H.2) for $\bar{\boldsymbol{\alpha}}_2 \in \mathcal{B}_\epsilon(\bar{\boldsymbol{\alpha}}_1)$, the following inequality holds:

$$|\vartheta_2 - \vartheta_1| \leq ||\bar{\boldsymbol{\alpha}}_2 - \bar{\boldsymbol{\alpha}}_1||_2 < \epsilon = \frac{\bar{\epsilon}}{2}, \tag{H.7}$$

where the first inequality is a direct consequence of the definition of $||\cdot||_2$. As (H.7) contradicts to (H.6), the assumption $\vartheta_1 \neq \vartheta_2$ is false. As a result, ϑ_1 and ϑ_2 must be identical.

In order to prove that the remaining entries of the vector parameters $\bar{\boldsymbol{\alpha}}_1$ and $\bar{\boldsymbol{\alpha}}_2$ are equal, it is sufficient to prove that for any positive definite Hermitian matrices $\bar{\boldsymbol{P}}_1$ and $\bar{\boldsymbol{P}}_2$ and for any matrices $\bar{\boldsymbol{B}}_1$, $\bar{\boldsymbol{B}}_2$, the following statement must hold:

$$\text{if } \bar{\boldsymbol{A}}_1\bar{\boldsymbol{P}}_1\bar{\boldsymbol{A}}_1^{\mathrm{H}} = \bar{\boldsymbol{A}}_2\bar{\boldsymbol{P}}_2\bar{\boldsymbol{A}}_2^{\mathrm{H}}, \text{ then } \begin{cases} \bar{\boldsymbol{B}}_1 = \bar{\boldsymbol{B}}_2 \\ \bar{\boldsymbol{P}}_1 = \bar{\boldsymbol{P}}_2 \end{cases} \tag{H.8}$$

where the full rank matrices $\bar{\boldsymbol{A}}_1$ and $\bar{\boldsymbol{A}}_2$ are defined conformably with (6.16) as

$$\bar{\boldsymbol{A}}_1 = \begin{bmatrix} a_1 & \boldsymbol{0}^{\mathrm{T}} \\ \boldsymbol{a}_2 & \bar{\boldsymbol{B}}_1 \\ \boldsymbol{a}_3 & \boldsymbol{I}_{N-1} \end{bmatrix} \text{ and } \bar{\boldsymbol{A}}_2 = \begin{bmatrix} a_1 & \boldsymbol{0}^{\mathrm{T}} \\ \boldsymbol{a}_2 & \bar{\boldsymbol{B}}_2 \\ \boldsymbol{a}_3 & \boldsymbol{I}_{N-1} \end{bmatrix}, \tag{H.9}$$

with $\boldsymbol{a}(\vartheta_1) = \boldsymbol{a}(\vartheta_2) = \left[a_1, \boldsymbol{a}_2^{\mathrm{T}}, \boldsymbol{a}_3^{\mathrm{T}}\right]^{\mathrm{T}}$. In order to prove (H.8), we partition the source covariance matrices $\bar{\boldsymbol{P}}_1$ and $\bar{\boldsymbol{P}}_2$ as

$$\bar{\boldsymbol{P}}_1 = \begin{bmatrix} \sigma_1 & \boldsymbol{\rho}_1^{\mathrm{H}} \\ \boldsymbol{\rho}_1 & \boldsymbol{Q}_1 \end{bmatrix} \text{ and } \bar{\boldsymbol{P}}_2 = \begin{bmatrix} \sigma_2 & \boldsymbol{\rho}_2^{\mathrm{H}} \\ \boldsymbol{\rho}_2 & \boldsymbol{Q}_2 \end{bmatrix}. \tag{H.10}$$

Due to the positive definiteness of the matrices $\bar{\boldsymbol{P}}_1$ and $\bar{\boldsymbol{P}}_2$, the two submatrices $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$ are invertible. By substituting (H.9) and (H.10) back to (H.8), expanding and comparing each block matrices on the left and the right hand side, the equation (H.8) is equivalent to the following set of conditions:

$$|a_1|^2 \sigma_1 = |a_1|^2 \sigma_2 \tag{H.11}$$

$$\boldsymbol{a}_2\sigma_1 a_1^* + \bar{\boldsymbol{B}}_1\boldsymbol{\rho}_1 a_1^* = \boldsymbol{a}_2\sigma_2 a_1^* + \bar{\boldsymbol{B}}_2\boldsymbol{\rho}_2 a_1^* \tag{H.12}$$

$$\boldsymbol{a}_3\sigma_1 a_1^* + \boldsymbol{\rho}_1 a_1^* = \boldsymbol{a}_3\sigma_2 a_1^* + \boldsymbol{\rho}_2 a_1^* \tag{H.13}$$

$$\boldsymbol{a}_2\sigma_1\boldsymbol{a}_2^{\mathrm{H}} + 2\mathrm{Re}\left\{\bar{\boldsymbol{B}}_1\boldsymbol{\rho}_1\boldsymbol{a}_2^{\mathrm{H}}\right\} + \bar{\boldsymbol{B}}_1\boldsymbol{Q}_1\bar{\boldsymbol{B}}_1^{\mathrm{H}} = \boldsymbol{a}_2\sigma_2\boldsymbol{a}_2^{\mathrm{H}} + 2\mathrm{Re}\left\{\bar{\boldsymbol{B}}_2\boldsymbol{\rho}_2\boldsymbol{a}_2^{\mathrm{H}}\right\} + \bar{\boldsymbol{B}}_2\boldsymbol{Q}_2\bar{\boldsymbol{B}}_2^{\mathrm{H}} \tag{H.14}$$

$$\boldsymbol{a}_3\sigma_1\boldsymbol{a}_2^{\mathrm{H}} + \boldsymbol{\rho}_1\boldsymbol{a}_2^{\mathrm{H}} + \boldsymbol{a}_3\boldsymbol{\rho}_1^{\mathrm{H}}\bar{\boldsymbol{B}}_1 + \boldsymbol{Q}_1\bar{\boldsymbol{B}}_1^{\mathrm{H}} = \boldsymbol{a}_3\sigma_2\boldsymbol{a}_2^{\mathrm{H}} + \boldsymbol{\rho}_2\boldsymbol{a}_2^{\mathrm{H}} + \boldsymbol{a}_3\boldsymbol{\rho}_2^{\mathrm{H}}\bar{\boldsymbol{B}}_2 + \boldsymbol{Q}_2\bar{\boldsymbol{B}}_2^{\mathrm{H}} \tag{H.15}$$

$$\boldsymbol{a}_2\sigma_1\boldsymbol{a}_3^{\mathrm{H}} + 2\mathrm{Re}\left\{\boldsymbol{\rho}_1\boldsymbol{a}_3^{\mathrm{H}}\right\} + \boldsymbol{Q}_1 = \boldsymbol{a}_2\sigma_2\boldsymbol{a}_3^{\mathrm{H}} + 2\mathrm{Re}\left\{\boldsymbol{\rho}_2\boldsymbol{a}_3^{\mathrm{H}}\right\} + \boldsymbol{Q}_2. \tag{H.16}$$

From the partition in (6.12) which enforces $a_1 \neq 0$, the condition in (H.11) implies

$$\sigma_1 = \sigma_2. \tag{H.17}$$

As σ_1 and σ_2 are identical, the condition in (H.13) is equivalent to

$$\boldsymbol{\rho}_1 = \boldsymbol{\rho}_2. \tag{H.18}$$

Substituting (H.18) in (H.16), we obtain

$$\boldsymbol{Q}_1 = \boldsymbol{Q}_2. \tag{H.19}$$

From (H.17), (H.18) and (H.19), we conclude that

$$\bar{\boldsymbol{P}}_1 = \bar{\boldsymbol{P}}_2. \tag{H.20}$$

Next we prove that $\bar{\boldsymbol{B}}_1 = \bar{\boldsymbol{B}}_2$. The condition in (H.12) implies

$$\bar{\boldsymbol{B}}_1\boldsymbol{\rho}_1 = \bar{\boldsymbol{B}}_2\boldsymbol{\rho}_2. \tag{H.21}$$

Substituting (H.21) and (H.19) to (H.15) results in

$$\boldsymbol{Q}_1\bar{\boldsymbol{B}}_1^{\mathrm{H}} = \boldsymbol{Q}_2\bar{\boldsymbol{B}}_2^{\mathrm{H}} = \boldsymbol{Q}_1\bar{\boldsymbol{B}}_2^{\mathrm{H}}. \tag{H.22}$$

Since $\bar{\boldsymbol{P}}_1$ is a full column rank matrix, the submatrix $\boldsymbol{Q}_1$ is invertible. Multiplying both sides of the Equation (H.22) with the inverse of $\boldsymbol{Q}_1$, we obtain

$$\bar{\boldsymbol{B}}_1 = \bar{\boldsymbol{B}}_2. \tag{H.23}$$

From the identities in (H.20) and (H.23), the statement in (H.3) is true, and the parameterization in (6.21) is locally identifiable. Consequently, the FIM given in (6.24) is invertible. The existence of the CRB matrix follows immediately from the invertibility of the FIM, and we conclude the proof.

I Proof of Theorem 6.4

The proof of the identity in (6.33), which is an important intermediate result for the derivation of the closed-form expression of the PR-CRB, is provided in this section. The proof will follow the footsteps of [SLG01] with slight modification due to a different parameterization. The main steps are outlined below. We assume that, all inverses occurring in the following derivation exist, which is guaranteed by the invertibility of

the FIM. In accordance with the parameterization in (6.21), the FIM with respect to the parameterization in (6.21) is given by [SLG01, Eq. (8)]:

$$\boldsymbol{\mathcal{I}} = T \left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\bar{\boldsymbol{\alpha}}^{\mathrm{T}}} \right)^{\mathrm{H}} \left(\boldsymbol{R}^{-\mathrm{T}} \otimes \boldsymbol{R}^{-1} \right) \left(\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\bar{\boldsymbol{\alpha}}^{\mathrm{T}}} \right), \tag{I.1}$$

with

$$\boldsymbol{r} = \mathrm{vec}\left(\boldsymbol{R}\right) = \left(\bar{\boldsymbol{A}}^{*} \otimes \bar{\boldsymbol{A}} \right) \mathrm{vec}(\bar{\boldsymbol{P}}) + \nu \boldsymbol{I}. \tag{I.2}$$

By defining

$$[\boldsymbol{F} \mid \boldsymbol{\Delta}] = \left(\boldsymbol{R}^{-\mathrm{T}/2} \otimes \boldsymbol{R}^{-1/2} \right) \left[\frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\boldsymbol{\beta}^{\mathrm{T}}} \;\middle|\; \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\boldsymbol{\gamma}^{\mathrm{T}}} \right], \tag{I.3}$$

where the vectors $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are defined in (6.30). The partitioned FIM is given by:

$$\boldsymbol{\mathcal{I}} = \begin{bmatrix} \boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}\boldsymbol{\beta}} & \boldsymbol{\mathcal{I}}_{\boldsymbol{\beta}\boldsymbol{\gamma}} \\ \boldsymbol{\mathcal{I}}_{\boldsymbol{\gamma}\boldsymbol{\beta}} & \boldsymbol{\mathcal{I}}_{\boldsymbol{\gamma}\boldsymbol{\gamma}} \end{bmatrix} = T \begin{bmatrix} \boldsymbol{F}^{\mathrm{H}}\boldsymbol{F} & \boldsymbol{F}^{\mathrm{H}}\boldsymbol{\Delta} \\ \boldsymbol{\Delta}^{\mathrm{H}}\boldsymbol{F} & \boldsymbol{\Delta}^{\mathrm{H}}\boldsymbol{\Delta} \end{bmatrix}. \tag{I.4}$$

If we partition the CRB matrix $\boldsymbol{C}$ similarly to the FIM $\boldsymbol{\mathcal{I}}$, i.e.,

$$\boldsymbol{C} = \begin{bmatrix} \boldsymbol{C}_{\boldsymbol{\beta}\boldsymbol{\beta}} & \boldsymbol{C}_{\boldsymbol{\beta}\boldsymbol{\gamma}} \\ \boldsymbol{C}_{\boldsymbol{\gamma}\boldsymbol{\beta}} & \boldsymbol{C}_{\boldsymbol{\gamma}\boldsymbol{\gamma}} \end{bmatrix}, \tag{I.5}$$

then, using the block matrix inversion lemma, the matrix $\boldsymbol{\Sigma} = \boldsymbol{C}_{\boldsymbol{\beta}\boldsymbol{\beta}}^{-1}$ is obtained as [SLG01, Eqs. (14), (20) and (28)]:

$$\begin{aligned} \boldsymbol{\Sigma} = \boldsymbol{C}_{\boldsymbol{\beta}\boldsymbol{\beta}}^{-1} &= T\boldsymbol{F}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{\Delta}}^{\perp}\boldsymbol{F} \\ &= T\boldsymbol{F}^{\mathrm{H}}\boldsymbol{\Pi}^{\perp}_{\left(\boldsymbol{R}^{-1/2}\bar{\boldsymbol{A}}\right)^{*}\otimes\left(\boldsymbol{R}^{-1/2}\bar{\boldsymbol{A}}\right)}\boldsymbol{F}. \end{aligned} \tag{I.6}$$

In the following steps, we calculate all entries of the matrix $\boldsymbol{\Sigma}$ in (I.6). Each column $\boldsymbol{f}_p$ of the matrix $\boldsymbol{F}$ is calculated as

$$\begin{aligned} \boldsymbol{f}_p &= T \left(\boldsymbol{R}^{-\mathrm{T}/2} \otimes \boldsymbol{R}^{-1/2} \right) \frac{\mathrm{d}\boldsymbol{r}}{\mathrm{d}\beta_p} \\ &= T \left(\boldsymbol{R}^{-\mathrm{T}/2} \otimes \boldsymbol{R}^{-1/2} \right) \mathrm{vec}\left(\frac{\mathrm{d}\boldsymbol{R}}{\mathrm{d}\beta_p} \right) \\ &= \mathrm{vec}\left(\boldsymbol{Z}_p + \boldsymbol{Z}_p^{\mathrm{H}} \right), \end{aligned} \tag{I.7}$$

with $\boldsymbol{Z}_p = T\boldsymbol{R}^{-1/2}\bar{\boldsymbol{A}}\bar{\boldsymbol{P}}\dfrac{\mathrm{d}\bar{\boldsymbol{A}}}{\mathrm{d}\beta_p}^{\mathrm{H}}\boldsymbol{R}^{-1/2}$. Using the shorthand notation $\underline{\boldsymbol{A}} = \boldsymbol{R}^{-1/2}\bar{\boldsymbol{A}}$, the projection matrix $\boldsymbol{\Pi}^{\perp}_{\left(\boldsymbol{R}^{-1/2}\bar{\boldsymbol{A}}\right)^{*}\otimes\left(\boldsymbol{R}^{-1/2}\bar{\boldsymbol{A}}\right)} = \boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}^{*}\otimes\underline{\boldsymbol{A}}}$ is rewritten as [SLG01, Eq. (21)]:

$$\boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}^{*}\otimes\underline{\boldsymbol{A}}} = \boldsymbol{I}_M \otimes \boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}} + \boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}^{*}} \otimes \boldsymbol{I}_M - \boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}^{*}} \otimes \boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}}. \tag{I.8}$$

From the identity in (I.8) and the property

$$\boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}}\boldsymbol{Z}_p = \boldsymbol{0} = \boldsymbol{Z}_p^{\mathrm{H}}\boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}}, \tag{I.9}$$

we obtain:

$$\begin{aligned}\boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}^*\otimes\underline{\boldsymbol{A}}}\boldsymbol{f}_p &= \operatorname{vec}\left(\boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}}\left(\boldsymbol{Z}_p+\boldsymbol{Z}_p^{\mathrm{H}}\right)+\left(\boldsymbol{Z}_p+\boldsymbol{Z}_p^{\mathrm{H}}\right)\boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}}-\boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}}\left(\boldsymbol{Z}_p+\boldsymbol{Z}_p^{\mathrm{H}}\right)\boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}}\right)\\ &= \operatorname{vec}\left(\boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}}\boldsymbol{Z}_p^{\mathrm{H}}+\boldsymbol{Z}_p\boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}}\right).\end{aligned} \tag{I.10}$$

Using [SLG01, Eq. (31)], the entry Σ_{pq} on the p-th row and q-th column of the matrix $\boldsymbol{\Sigma}$ is computed as follows:

$$\begin{aligned}\Sigma_{pq} &= T\boldsymbol{f}_p^{\mathrm{H}}\boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}^*\otimes\underline{\boldsymbol{A}}}\boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}^*\otimes\underline{\boldsymbol{A}}}\boldsymbol{f}_q\\ &= T\operatorname{vec}\left(\boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}}\boldsymbol{Z}_p^{\mathrm{H}}+\boldsymbol{Z}_p\boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}}\right)^{\mathrm{H}}\operatorname{vec}\left(\boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}}\boldsymbol{Z}_q^{\mathrm{H}}+\boldsymbol{Z}_q\boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}}\right)\\ &= 2T\mathrm{Re}\left\{\operatorname{tr}\left(\boldsymbol{Z}_p\boldsymbol{\Pi}^{\perp}_{\underline{\boldsymbol{A}}}\boldsymbol{Z}_q^{\mathrm{H}}\right)\right\}\\ &= \frac{2T}{\nu}\mathrm{Re}\left\{\operatorname{tr}\left(\boldsymbol{R}^{-1/2}\bar{\boldsymbol{A}}\bar{\boldsymbol{P}}\frac{\mathrm{d}\bar{\boldsymbol{A}}}{\mathrm{d}\beta_p}^{\mathrm{H}}\boldsymbol{\Pi}^{\perp}_{\bar{\boldsymbol{A}}}\frac{\mathrm{d}\bar{\boldsymbol{A}}}{\mathrm{d}\beta_q}\bar{\boldsymbol{P}}\bar{\boldsymbol{A}}^{\mathrm{H}}\boldsymbol{R}^{-1/2}\right)\right\}\\ &= \frac{2T}{\nu}\mathrm{Re}\left\{\operatorname{tr}\left(\bar{\boldsymbol{P}}\bar{\boldsymbol{A}}^{\mathrm{H}}\boldsymbol{R}^{-1}\bar{\boldsymbol{A}}\bar{\boldsymbol{P}}\frac{\mathrm{d}\bar{\boldsymbol{A}}}{\mathrm{d}\beta_p}^{\mathrm{H}}\boldsymbol{\Pi}^{\perp}_{\bar{\boldsymbol{A}}}\frac{\mathrm{d}\bar{\boldsymbol{A}}}{\mathrm{d}\beta_q}\right)\right\},\end{aligned} \tag{I.11}$$

which concludes the proof.

J Proof of Theorem 6.5

In this section, we provide the proof of the identity (6.48) in Theorem 6.5. Under the assumption that the matrix $\boldsymbol{J}^{\mathrm{H}}\boldsymbol{\Pi}^{\perp}_{\boldsymbol{K}}\boldsymbol{J}$ is invertible with the $M\times(M-N)$ selection matrix $\boldsymbol{J}$ defined in (6.40), we have the following inequalities:

$$\begin{aligned}M-N &\geq \operatorname{rank}\left(\boldsymbol{\Pi}_{\boldsymbol{\Pi}^{\perp}_{\boldsymbol{K}}\boldsymbol{J}}\right)\\ &= \operatorname{rank}\left(\boldsymbol{\Pi}^{\perp}_{\boldsymbol{K}}\boldsymbol{J}\right)\geq\operatorname{rank}\left(\boldsymbol{J}^{\mathrm{H}}\boldsymbol{\Pi}^{\perp}_{\boldsymbol{K}}\boldsymbol{J}\right)=M-N,\end{aligned} \tag{J.1}$$

and therefore

$$\operatorname{rank}\left(\boldsymbol{\Pi}_{\boldsymbol{\Pi}^{\perp}_{\boldsymbol{K}}\boldsymbol{J}}\right)=M-N=\operatorname{rank}\left(\boldsymbol{\Pi}^{\perp}_{\boldsymbol{K}}\right). \tag{J.2}$$

As the two projection matrices $\boldsymbol{\Pi}_{\boldsymbol{\Pi}^{\perp}_{\boldsymbol{K}}\boldsymbol{J}}$ and $\boldsymbol{\Pi}^{\perp}_{\boldsymbol{K}}$ are idempotent, it follows that

$$\boldsymbol{\Pi}_{\boldsymbol{\Pi}^{\perp}_{\boldsymbol{K}}\boldsymbol{J}}\boldsymbol{\Pi}^{\perp}_{\boldsymbol{K}}=\boldsymbol{\Pi}^{\perp}_{\boldsymbol{K}}\boldsymbol{J}\left(\boldsymbol{J}^{\mathrm{H}}\boldsymbol{\Pi}^{\perp}_{\boldsymbol{K}}\boldsymbol{J}\right)^{-1}\boldsymbol{J}^{\mathrm{H}}\boldsymbol{\Pi}^{\perp}_{\boldsymbol{K}}=\boldsymbol{\Pi}^{\perp}_{\boldsymbol{K}}\boldsymbol{\Pi}_{\boldsymbol{\Pi}^{\perp}_{\boldsymbol{K}}\boldsymbol{J}}. \tag{J.3}$$

As a consequence, the projection matrices $\boldsymbol{\Pi}_{\boldsymbol{\Pi}^{\perp}_{\boldsymbol{K}}\boldsymbol{J}}$ and $\boldsymbol{\Pi}^{\perp}_{\boldsymbol{K}}$ commute, and therefore they are simultaneously diagonalizable [HJ12, Th.1.3.12]. This implies that there exists a unitary matrix $\boldsymbol{U}=[\boldsymbol{u}_1,\ldots,\boldsymbol{u}_M]$ and two diagonal matrices $\boldsymbol{\Gamma}=\operatorname{diag}(\gamma_1,\ldots,\gamma_M)$ and $\boldsymbol{\Upsilon}=\operatorname{diag}(\upsilon_1,\ldots,\upsilon_M)$ such that

$$\boldsymbol{\Pi}_{\boldsymbol{\Pi}^{\perp}_{\boldsymbol{K}}\boldsymbol{J}}=\boldsymbol{U}\boldsymbol{\Gamma}\boldsymbol{U}^{\mathrm{H}}\text{ and }\boldsymbol{\Pi}^{\perp}_{\boldsymbol{K}}=\boldsymbol{U}\boldsymbol{\Upsilon}\boldsymbol{U}^{\mathrm{H}}. \tag{J.4}$$

Furthermore, the eigenvalues γ_m and υ_m for $m = 1, \ldots, M$ are either 0 or 1. Now we prove (6.48) by contradiction: assume that $\boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{K}}^{\perp}\boldsymbol{J}} \neq \boldsymbol{\Pi}_{\boldsymbol{K}}^{\perp}$. This implies that there exists at least one index k such that $\gamma_k \neq \upsilon_k$. However, for all $\upsilon_m = 0$, then $\boldsymbol{\Pi}_{\boldsymbol{K}}^{\perp}\boldsymbol{u}_m = \boldsymbol{0}$, and therefore,

$$\boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{K}}^{\perp}\boldsymbol{J}}\boldsymbol{u}_m = \boldsymbol{\Pi}_{\boldsymbol{K}}^{\perp}\boldsymbol{J}\left(\boldsymbol{J}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{K}}^{\perp}\boldsymbol{J}\right)^{-1}\boldsymbol{J}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{K}}^{\perp}\boldsymbol{u}_m = \boldsymbol{0}. \tag{J.5}$$

Thus, we conclude that for all indices m with $\upsilon_m = 0$, the corresponding eigenvalues γ_m of the matrix $\boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{K}}^{\perp}\boldsymbol{J}}$ are also zero. Therefore, the assumption $\gamma_k \neq \upsilon_k$ holds if and only if $\upsilon_k = 1$ and $\gamma_k = 0$. However, this in turn implies that $\operatorname{rank}\left(\boldsymbol{\Pi}_{\boldsymbol{K}}^{\perp}\right) > \operatorname{rank}\left(\boldsymbol{\Pi}_{\boldsymbol{\Pi}_{\boldsymbol{K}}^{\perp}\boldsymbol{J}}\right)$, which contradicts with (J.2). As a result, we have $\upsilon_m = \gamma_m$ for all $m = 1, \ldots, M$ and therefore (6.48) holds.

K On the Comparison between PR-CRB and the Conventional CRB

In this section, an analytical proof of Theorem 6.7 is provided. Without loss of generality, we assume that $\vartheta = \theta_1$. Under this assumption, it is sufficient to prove that

$$C_{\text{PR-CRB}}(\theta_1) \geq \left[\boldsymbol{C}_{\text{CRB}}(\boldsymbol{\theta})\right]_{11}. \tag{K.1}$$

In order to prove (K.1), first we partition $\boldsymbol{D} = [\boldsymbol{d}, \boldsymbol{D_2}]$ and correspondingly

$$\boldsymbol{D}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}\boldsymbol{D} = \begin{bmatrix} \boldsymbol{d}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}\boldsymbol{d} & \boldsymbol{d}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}\boldsymbol{D}_2 \\ \boldsymbol{D}_2^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}\boldsymbol{d} & \boldsymbol{D}_2^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}\boldsymbol{D}_2 \end{bmatrix} = \begin{bmatrix} \bar{D}_{11} & \bar{\boldsymbol{D}}_{21}^{\mathrm{H}} \\ \bar{\boldsymbol{D}}_{21} & \bar{\boldsymbol{D}}_{22} \end{bmatrix}. \tag{K.2}$$

Further define $\tilde{\boldsymbol{M}} = \begin{bmatrix} \boldsymbol{M}_{21}^{\mathrm{H}}\boldsymbol{M}_{22}^{-1/2} \\ \boldsymbol{M}_{22}^{1/2} \end{bmatrix}\begin{bmatrix} \boldsymbol{M}_{22}^{-1/2}\boldsymbol{M}_{21} & \boldsymbol{M}_{22}^{1/2} \end{bmatrix}$ where $\boldsymbol{M}_{21}$ and $\boldsymbol{M}_{22}$ are defined in (6.51). Clearly both $\boldsymbol{D}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}\boldsymbol{D}$ and $\tilde{\boldsymbol{M}}$ are positive semidefinite, from [SN90a, Lem. A.1], the following matrix

$$\operatorname{Re}\left\{\tilde{\boldsymbol{M}} \odot \left(\boldsymbol{D}^{\mathrm{H}}\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}\boldsymbol{D}\right)\right\} = \operatorname{Re}\left\{\begin{bmatrix} \left(\boldsymbol{M}_{21}^{\mathrm{H}}\boldsymbol{M}_{22}^{-1}\boldsymbol{M}_{21}\right)\bar{D}_{11} & \boldsymbol{M}_{21}^{\mathrm{H}} \odot \bar{\boldsymbol{D}}_{21}^{\mathrm{H}} \\ \boldsymbol{M}_{21} \odot \bar{\boldsymbol{D}}_{21} & \boldsymbol{M}_{22} \odot \bar{\boldsymbol{D}}_{22} \end{bmatrix}\right\} \tag{K.3}$$

is also positive semidefinite. Using this property and the Schur complement of the most upper-left entry of the block matrix on the right hand side of (K.3), we obtain

$$\begin{aligned} &\operatorname{Re}\left\{\left(\boldsymbol{M}_{21}^{\mathrm{H}}\boldsymbol{M}_{22}^{-1}\boldsymbol{M}_{21}\right)\bar{D}_{11}\right\} - \\ &\qquad \operatorname{Re}\left\{\boldsymbol{M}_{21}^{\mathrm{H}} \odot \bar{\boldsymbol{D}}_{21}^{\mathrm{H}}\right\}\operatorname{Re}\left\{\boldsymbol{M}_{22} \odot \bar{\boldsymbol{D}}_{22}\right\}^{-1}\operatorname{Re}\left\{\boldsymbol{M}_{21} \odot \bar{\boldsymbol{D}}_{21}\right\} \geq 0, \end{aligned} \tag{K.4}$$

or equivalently

$$\begin{aligned} \operatorname{Re}\left\{\left(M_{11} - \boldsymbol{M}_{21}^{\mathrm{H}}\boldsymbol{M}_{22}^{-1}\boldsymbol{M}_{21}\right)\bar{D}_{11}\right\} \leq \operatorname{Re}\left\{M_{11}\bar{D}_{11}\right\} - \\ \operatorname{Re}\left\{\boldsymbol{M}_{21}^{\mathrm{H}} \odot \bar{\boldsymbol{D}}_{21}^{\mathrm{H}}\right\}\operatorname{Re}\left\{\boldsymbol{M}_{22} \odot \bar{\boldsymbol{D}}_{22}\right\}^{-1}\operatorname{Re}\left\{\boldsymbol{M}_{21} \odot \bar{\boldsymbol{D}}_{21}\right\}. \end{aligned} \tag{K.5}$$

Note that both expressions on the left and the right hand side of (K.5) are non-negative, taking the inverse of both sides and applying the block matrix inversion lemma for the most upper-left entry of the matrix on the right hand side again results in

$$\mathrm{Re}\left\{\left(M_{11}-\boldsymbol{M}_{21}^{\mathrm{H}}\boldsymbol{M}_{22}^{-1}\boldsymbol{M}_{21}\right)\bar{D}_{11}\right\}^{-1}\geq\left[\mathrm{Re}\left\{\begin{bmatrix}M_{11}&\boldsymbol{M}_{21}^{\mathrm{H}}\\ \boldsymbol{M}_{21}&\boldsymbol{M}_{22}\end{bmatrix}\odot\begin{bmatrix}\bar{D}_{11}&\bar{\boldsymbol{D}}_{21}^{\mathrm{H}}\\ \bar{\boldsymbol{D}}_{21}&\bar{\boldsymbol{D}}_{22}\end{bmatrix}\right\}\right]_{11}^{-1}. \tag{K.6}$$

From (6.5), (6.55), (K.2) and (K.6), we obtain the result in (K.1).

L On the Influence of Weighting Matrix on PR-WSF

We first prove that the asymptotic MSE of PR-WSF estimate for the DOA parameter $\vartheta=\theta_1$ is independent of the weighting matrix $\boldsymbol{W}$. The generalization of the proof for the remaining case $\vartheta=\theta_n$ with $n=2,\ldots,N$ follows immediately. From [VO91, Th. 2], which holds true for any parameterization and any positive definite weighting matrix $\boldsymbol{W}$, the distribution of estimates obtained by the PR-WSF estimator is given by

$$\hat{\boldsymbol{\beta}}-\boldsymbol{\beta}\sim\mathrm{As}N\left(\boldsymbol{0},\frac{1}{T}\boldsymbol{C}_{\text{PR-WSF}}\left(\boldsymbol{W}\right)\right), \tag{L.1}$$

where $\boldsymbol{\beta}$ is defined in (6.30), $\hat{\boldsymbol{\beta}}$ contains the corresponding parameters estimated by the PR-WSF estimator, and AsN denotes the asymptotic normal distribution as the number of snapshots T tends to infinity. The error covariance matrix $\boldsymbol{C}_{\text{PR-WSF}}\left(\boldsymbol{W}\right)$ in (L.1) is given by [VO91, Eq. (49)]

$$\boldsymbol{C}_{\text{PR-WSF}}\left(\boldsymbol{W}\right)=\left(\boldsymbol{V}\left(\boldsymbol{W}\right)\right)^{-1}\boldsymbol{Q}\left(\boldsymbol{W}\right)\left(\boldsymbol{V}\left(\boldsymbol{W}\right)\right)^{-1}, \tag{L.2}$$

where the elements on the p-th row and q-th column of the matrices $\boldsymbol{V}\left(\boldsymbol{W}\right)$ and $\boldsymbol{Q}\left(\boldsymbol{W}\right)$ are given by [VO91, Eqs. (50)-(51)]

$$V_{pq}\left(\boldsymbol{W}\right)=2\mathrm{Re}\left\{\mathrm{tr}\left(\frac{\partial\bar{\boldsymbol{A}}}{\partial\beta_p}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\frac{\partial\bar{\boldsymbol{A}}}{\partial\beta_q}\bar{\boldsymbol{A}}^{\dagger}\boldsymbol{U}_{\mathrm{s}}\boldsymbol{W}\boldsymbol{U}_{\mathrm{s}}^{\mathrm{H}}\bar{\boldsymbol{A}}^{\dagger\mathrm{H}}\right)\right\}, \tag{L.3}$$

$$Q_{pq}\left(\boldsymbol{W}\right)=2\nu\mathrm{Re}\left\{\mathrm{tr}\left(\frac{\partial\bar{\boldsymbol{A}}}{\partial\beta_p}^{\mathrm{H}}\boldsymbol{\Pi}_{\bar{\boldsymbol{A}}}^{\perp}\frac{\partial\bar{\boldsymbol{A}}}{\partial\beta_q}\bar{\boldsymbol{A}}^{\dagger}\boldsymbol{U}_{\mathrm{s}}\boldsymbol{W}\boldsymbol{\Lambda}_{\mathrm{s}}\tilde{\boldsymbol{\Lambda}}^{-2}\boldsymbol{W}\boldsymbol{U}_{\mathrm{s}}^{\mathrm{H}}\bar{\boldsymbol{A}}^{\dagger\mathrm{H}}\right)\right\}, \tag{L.4}$$

respectively. In (L.4), the matrix $\tilde{\boldsymbol{\Lambda}}$ is defined as $\tilde{\boldsymbol{\Lambda}}=\boldsymbol{\Lambda}_{\mathrm{s}}-\nu\boldsymbol{I}_N$. Since the interest is the asymptotic MSE in estimating the DOA parameter ϑ, we prove the following identity which holds true for any positive definite weighting matrix $\boldsymbol{W}$.

$$\left[\boldsymbol{C}_{\text{PR-WSF}}\left(\boldsymbol{W}\right)\right]_{11}=\left[\boldsymbol{C}_{\text{PR-WSF}}\left(\boldsymbol{W}_{\text{opt}}\right)\right]_{11}, \tag{L.5}$$

where $\boldsymbol{C}_{\text{PR-WSF}}(\boldsymbol{W})$ is defined in (L.2), and $\boldsymbol{W}_{\text{opt}} = \tilde{\boldsymbol{\Lambda}}^2 \boldsymbol{\Lambda}_{\text{s}}^{-1}$. In order to prove (L.5), we introduce the following matrix operator based on the Kronecker product and two properties related to this operator:

Definition L.1. *Let* $\boldsymbol{\Lambda} = \begin{bmatrix} \Lambda_{11} & \boldsymbol{\Lambda}_{21}^{\text{H}} \\ \boldsymbol{\Lambda}_{21} & \boldsymbol{\Lambda}_{22} \end{bmatrix} \in \mathbb{C}^{M \times M}$ *and* $\boldsymbol{\Gamma} = \begin{bmatrix} \Gamma_{11} & \boldsymbol{\Gamma}_{21}^{\text{H}} \\ \boldsymbol{\Gamma}_{21} & \boldsymbol{\Gamma}_{22} \end{bmatrix} \in \mathbb{C}^{M \times M}$ *be two Hermitian matrices with conformable partition where the upper-left block is a scalar. The extended blockwise Kronecker product between two matrices* $\boldsymbol{\Lambda}$ *and* $\boldsymbol{\Gamma}$ *is defined as follows:*

$$\boldsymbol{\Lambda} \boxtimes \boldsymbol{\Gamma} = \begin{bmatrix} \Lambda_{11} \otimes \Gamma_{11} & \boldsymbol{\Lambda}_{21}^{\text{H}} \otimes \boldsymbol{\Gamma}_{21}^{\text{H}} & j\boldsymbol{\Lambda}_{21}^{\text{H}} \otimes \boldsymbol{\Gamma}_{21}^{\text{H}} \\ \boldsymbol{\Lambda}_{21} \otimes \boldsymbol{\Gamma}_{21} & \boldsymbol{\Lambda}_{22} \otimes \boldsymbol{\Gamma}_{22} & j\boldsymbol{\Lambda}_{22} \otimes \boldsymbol{\Gamma}_{22} \\ -j\boldsymbol{\Lambda}_{21} \otimes \boldsymbol{\Gamma}_{21} & -j\boldsymbol{\Lambda}_{22} \otimes \boldsymbol{\Gamma}_{22} & \boldsymbol{\Lambda}_{22} \otimes \boldsymbol{\Gamma}_{22} \end{bmatrix}. \tag{L.6}$$

From this definition, we present in the following two simple but important properties of the proposed matrix product, which are easily proven by the mathematical identities in Appendix A:

Property 1. *The Schur complement to the upper-left entry of the matrix* $\text{Re}\{\boldsymbol{\Lambda} \boxtimes \boldsymbol{\Gamma}\}$ *is given by*

$$\mathcal{S}(\text{Re}\{\boldsymbol{\Lambda} \boxtimes \boldsymbol{\Gamma}\}) = \Lambda_{11}\Gamma_{11} - \boldsymbol{\Lambda}_{21}^{\text{H}} \boldsymbol{\Lambda}_{22}^{-1} \boldsymbol{\Lambda}_{21} \boldsymbol{\Gamma}_{21}^{\text{H}} \boldsymbol{\Gamma}_{22}^{-1} \boldsymbol{\Gamma}_{21}. \tag{L.7}$$

If the Schur complement to the upper-left entry of the matrix $\boldsymbol{\Gamma}$ *is zero, i.e.,* $\mathcal{S}(\boldsymbol{\Gamma}) = 0$*, then the result in* (L.7) *reduces to*

$$\mathcal{S}(\text{Re}\{\boldsymbol{\Lambda} \boxtimes \boldsymbol{\Gamma}\}) = \mathcal{S}(\boldsymbol{\Lambda})\,\Gamma_{11}. \tag{L.8}$$

Property 2. *Assuming that the matrices* $Re\{\boldsymbol{\Lambda} \boxtimes \boldsymbol{\Gamma}\}, \boldsymbol{\Lambda}_{22}$ *and* $\boldsymbol{\Gamma}_{22}$ *are invertible, from the simplification in* (6.47)*, the first column of the matrix* $\boldsymbol{K} = \text{Re}\{\boldsymbol{\Lambda} \boxtimes \boldsymbol{\Gamma}\}^{-1}$ *is given by*

$$\boldsymbol{K}\boldsymbol{e}_1 = \mathcal{S}(\text{Re}\{\boldsymbol{\Lambda} \boxtimes \boldsymbol{\Gamma}\})^{-1} \begin{bmatrix} 1 \\ \text{Re}\{\boldsymbol{K}_{21}\} \\ \text{Im}\{\boldsymbol{K}_{21}\} \end{bmatrix}, \tag{L.9}$$

where the matrix $\boldsymbol{K}_{21} \in \mathbb{C}^{(M-1)^2}$ *is given by*

$$\boldsymbol{K}_{21} = -\left(\boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Lambda}_{21}\right) \otimes \left(\boldsymbol{\Gamma}_{22}^{-1}\boldsymbol{\Gamma}_{21}\right). \tag{L.10}$$

Now we return to the proof of the identity (L.5). By using the extended blockwise Kronecker product with

$$\boldsymbol{\Gamma} = \begin{bmatrix} \boldsymbol{d}^{\text{H}}\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}\boldsymbol{d} & \boldsymbol{d}^{\text{H}}\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}\boldsymbol{J} \\ \boldsymbol{J}^{\text{H}}\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}\boldsymbol{d} & \boldsymbol{J}^{\text{H}}\boldsymbol{\Pi}_{\boldsymbol{A}}^{\perp}\boldsymbol{J} \end{bmatrix} \tag{L.11}$$

$$\boldsymbol{\Lambda}(\boldsymbol{W}) = \left(\bar{\boldsymbol{A}}^{\dagger}\boldsymbol{U}_{\text{s}}\boldsymbol{W}\boldsymbol{U}_{\text{s}}^{\text{H}}\bar{\boldsymbol{A}}^{\dagger\text{H}}\right)^{\text{T}} \tag{L.12}$$

$$\boldsymbol{\Xi}(\boldsymbol{W}) = \left(\bar{\boldsymbol{A}}^{\dagger}\boldsymbol{U}_{\text{s}}\boldsymbol{W}\boldsymbol{\Lambda}_{\text{s}}\tilde{\boldsymbol{\Lambda}}^{-2}\boldsymbol{W}\boldsymbol{U}_{\text{s}}^{\text{H}}\bar{\boldsymbol{A}}^{\dagger\text{H}}\right)^{\text{T}}, \tag{L.13}$$

and following the footsteps in (6.33)-(6.46), the matrices $\boldsymbol{V}(\boldsymbol{W})$ and $\boldsymbol{Q}(\boldsymbol{W})$ in (L.3)-(L.4) are compactly expressed as

$$\boldsymbol{V}(\boldsymbol{W}) = 2\mathrm{Re}\left\{\boldsymbol{\Lambda}(\boldsymbol{W}) \boxtimes \boldsymbol{\Gamma}\right\}, \tag{L.14}$$

$$\boldsymbol{Q}(\boldsymbol{W}) = 2\nu\mathrm{Re}\left\{\boldsymbol{\Xi}(\boldsymbol{W}) \boxtimes \boldsymbol{\Gamma}\right\}. \tag{L.15}$$

Furthermore, by defining $\boldsymbol{\Lambda}_0 = \boldsymbol{\Lambda}(\boldsymbol{W}_{\mathrm{opt}})$ and $\boldsymbol{\Xi}_0 = \boldsymbol{\Xi}(\boldsymbol{W}_{\mathrm{opt}})$ and noting that $\bar{\boldsymbol{A}}^{\dagger}\boldsymbol{U}_{\mathrm{s}}$ is invertible, the following identity is easily verified:

$$\boldsymbol{\Lambda}_0 = \boldsymbol{\Xi}_0 = \boldsymbol{\Lambda}\boldsymbol{\Xi}^{-1}\boldsymbol{\Lambda}, \tag{L.16}$$

where the dependence of the matrices $\boldsymbol{\Lambda}$ and $\boldsymbol{\Xi}$ on the weighting matrix $\boldsymbol{W}$ is suppressed for notational convenience. Also from (L.2), (L.14)-(L.16), we obtain

$$\boldsymbol{C}_{\text{PR-WSF}}(\boldsymbol{W}_{\mathrm{opt}}) = \frac{\nu}{2}\mathrm{Re}\left\{\left(\boldsymbol{\Lambda}\boldsymbol{\Xi}^{-1}\boldsymbol{\Lambda}\right) \boxtimes \boldsymbol{\Gamma}\right\}^{-1}. \tag{L.17}$$

As a consequence, the identity in (L.5) holds if and only if

$$\boldsymbol{e}_1^{\mathrm{T}}\mathrm{Re}\left\{\boldsymbol{\Lambda} \boxtimes \boldsymbol{\Gamma}\right\}^{-1}\mathrm{Re}\left\{\boldsymbol{\Xi} \boxtimes \boldsymbol{\Gamma}\right\}\mathrm{Re}\left\{\boldsymbol{\Lambda} \boxtimes \boldsymbol{\Gamma}\right\}^{-1}\boldsymbol{e}_1 = \left[\mathrm{Re}\left\{\left(\boldsymbol{\Lambda}\boldsymbol{\Xi}^{-1}\boldsymbol{\Lambda}\right) \boxtimes \boldsymbol{\Gamma}\right\}^{-1}\right]_{11}. \tag{L.18}$$

In order to prove (L.18), first we note that $\mathcal{S}(\boldsymbol{\Gamma}) = 0$, which is a direct consequence of Theorem 6.5 in Section 6.4. The expression on the left hand side of (L.18) is simplified using Property 2 and (A.2)-(A.3) as follows:

$$\begin{aligned}
&\boldsymbol{e}_1^{\mathrm{T}}\mathrm{Re}\left\{\boldsymbol{\Lambda} \boxtimes \boldsymbol{\Gamma}\right\}^{-1}\mathrm{Re}\left\{\boldsymbol{\Xi} \boxtimes \boldsymbol{\Gamma}\right\}\mathrm{Re}\left\{\boldsymbol{\Lambda} \boxtimes \boldsymbol{\Gamma}\right\}^{-1}\boldsymbol{e}_1 \\
&= \mathcal{S}\left(\mathrm{Re}\left\{\boldsymbol{\Lambda} \boxtimes \boldsymbol{\Gamma}\right\}\right)^{-2}\Big(\Xi_{11}\Gamma_{11} - \mathrm{Re}\left\{\boldsymbol{\Xi}_{21}^{\mathrm{H}}\boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Lambda}_{21}\Gamma_{11}\right\} - \mathrm{Re}\left\{\boldsymbol{\Lambda}_{21}^{\mathrm{H}}\boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Xi}_{21}\Gamma_{11}\right\} \\
&\qquad\qquad + \mathrm{Re}\left\{\boldsymbol{\Lambda}_{21}^{\mathrm{H}}\boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Xi}_{22}\boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Lambda}_{21}\Gamma_{11}\right\}\Big) \\
&= \mathcal{S}(\boldsymbol{\Lambda})^{-2}\Gamma_{11}^{-1}\left(\Xi_{11} - \boldsymbol{\Xi}_{21}^{\mathrm{H}}\boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Lambda}_{21} - \boldsymbol{\Lambda}_{21}^{\mathrm{H}}\boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Xi}_{21} + \boldsymbol{\Lambda}_{21}^{\mathrm{H}}\boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Xi}_{22}\boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Lambda}_{21}\right).
\end{aligned} \tag{L.19}$$

The right hand side of (L.18) is reformulated using Property 1, Property 2 and (A.11) as follows:

$$\begin{aligned}
\left[\mathrm{Re}\left\{\left(\boldsymbol{\Lambda}\boldsymbol{\Xi}^{-1}\boldsymbol{\Lambda}\right) \boxtimes \boldsymbol{\Gamma}\right\}^{-1}\right]_{11} &= \mathcal{S}\left(\boldsymbol{\Lambda}\boldsymbol{\Xi}^{-1}\boldsymbol{\Lambda}\right)^{-1}\Gamma_{11}^{-1} \\
&= \mathcal{S}(\boldsymbol{\Lambda})^{-2}\Gamma_{11}^{-1}\begin{bmatrix}1 & -\boldsymbol{\Lambda}_{21}^{\mathrm{H}}\boldsymbol{\Lambda}_{22}^{-1}\end{bmatrix}\begin{bmatrix}\Xi_{11} & \boldsymbol{\Xi}_{21}^{\mathrm{H}} \\ \boldsymbol{\Xi}_{21} & \boldsymbol{\Xi}_{22}\end{bmatrix}\begin{bmatrix}1 \\ -\boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Lambda}_{21}\end{bmatrix} \\
&= \mathcal{S}(\boldsymbol{\Lambda})^{-2}\Gamma_{11}^{-1}\left(\Xi_{11} - \boldsymbol{\Xi}_{21}^{\mathrm{H}}\boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Lambda}_{21} - \boldsymbol{\Lambda}_{21}^{\mathrm{H}}\boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Xi}_{21} + \boldsymbol{\Lambda}_{21}^{\mathrm{H}}\boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Xi}_{22}\boldsymbol{\Lambda}_{22}^{-1}\boldsymbol{\Lambda}_{21}\right).
\end{aligned} \tag{L.20}$$

Comparing (L.19) with (L.20), we obtain the identity in (L.18) and conclude the first part of the theorem, which states that the MSE of the PR-WSF estimator for the DOA parameter ϑ is independent of the weighting matrix $\boldsymbol{W}$.

Next we show that the expression in (L.18) is equal to the PR-CRB given in (6.55). A sufficient condition for such equality is that

$$\boldsymbol{C}_{\text{PR-WSF}}(\boldsymbol{W}_{\mathrm{opt}}) = T\boldsymbol{\Sigma}^{-1}, \tag{L.21}$$

where the matrix $\mathbf{\Sigma}$ is defined in (6.32)-(6.33). We remark that the left hand side of (L.21) is the asymptotic MSE of the PR-WSF estimator with the weighting matrix $\boldsymbol{W} = \boldsymbol{W}_{\text{opt}} = \tilde{\mathbf{\Lambda}}^2 \mathbf{\Lambda}_{\text{s}}^{-1}$. In order to prove (L.21), we use the results in (6.32), (6.46) and the definition of the extended blockwise Kronecker product to compactly express the matrix $\mathbf{\Sigma}$ as

$$\mathbf{\Sigma} = \frac{2T}{\nu} \text{Re} \left\{ \left(\bar{\boldsymbol{P}} \bar{\boldsymbol{A}} \boldsymbol{R}^{-1} \bar{\boldsymbol{A}}^{\text{H}} \bar{\boldsymbol{P}} \right)^{\text{T}} \boxtimes \mathbf{\Gamma} \right\}. \tag{L.22}$$

We further note that, by a similar argument to [VO91, Eq. (75)], the following identity for the PR signal model holds:

$$\bar{\boldsymbol{P}} \bar{\boldsymbol{A}} \boldsymbol{R}^{-1} \bar{\boldsymbol{A}}^{\text{H}} \bar{\boldsymbol{P}} = \bar{\boldsymbol{A}}^{\dagger} \boldsymbol{U}_{\text{s}} \boldsymbol{W}_{\text{opt}} \boldsymbol{U}_{\text{s}}^{\text{H}} \bar{\boldsymbol{A}}^{\dagger \text{H}} = \mathbf{\Lambda}_0. \tag{L.23}$$

From (L.16), (L.17), (L.22) and (L.23), the identity in (L.21) follows, and thus we conclude the proof.

List of Acronyms

CRB	Cramér-Rao Bound
COMET	Covariance Matching Estimation Techniques
DML	Deterministic Maximum Likelihood
DOA	Direction-of-Arrival
ESPRIT	Estimation of Signal Parameters via Rotational Invariance Techniques
FIM	Fisher Information Matrix
FOV	Field-of-View
ML	Maximum Likelihood
MSE	Mean-Square Error
MUSIC	Multiple Signal Classification
PR	Partial Relaxation
PR-CRB	Partially Relaxed Cramér-Rao Bound
PR-DML	Partially Relaxed Deterministic Maximum Likelihood
PR-WSF	Partially Relaxed Weighted Subspace Fitting
PR-CCF	Partially Relaxed Constrained Covariance Fitting
PR-UCF	Partially Relaxed Unconstrained Covariance Fitting
RMSE	Root-Mean-Squared Error
SNR	Signal-to-Noise Ratio
ULA	Uniform Linear Array
WMUSIC	Weighted MUSIC
w.r.t.	With respect to
WSF	Weighted Subspace Fitting

List of Symbols

Latin letters

$\boldsymbol{A}$	Steering matrix
$\bar{\boldsymbol{A}}$	Transformed steering matrix under the PR model
$\mathcal{A}_N$	N-dimensional steering manifold
$\bar{\mathcal{A}}_N$	N-dimensional relaxed steering manifold
$\boldsymbol{a}(\theta_n)$	Steering vector of the DOA θ_n
$\boldsymbol{B}$	Relaxed steering matrix of undesired source signals under the PR model
$\bar{\boldsymbol{B}}$	Minor matrix of the transformed steering matrix under the PR model
$\mathcal{B}_\epsilon(\boldsymbol{p})$	An ϵ-neighborhood of a point $\boldsymbol{p}$
$\check{\boldsymbol{b}}$	Vector containing the real part of all entries in matrix $\bar{\boldsymbol{B}}$
$\tilde{\boldsymbol{b}}$	Vector containing the imaginary part of all entries in matrix $\bar{\boldsymbol{B}}$
$\boldsymbol{C}$	Cramér-Rao bound matrix
$\boldsymbol{C}_{\text{CRB}}(\boldsymbol{\theta})$	Stochastic Cramér-Rao bound matrix for the DOA parameters $\boldsymbol{\theta}$
$C_{\text{PR-CRB}}(\vartheta)$	Cramér-Rao bound for the desired DOA parameter ϑ under the PR signal model
$\mathbb{C}$	Set of complex numbers
$\mathbb{C}^M$	Set of complex-valued vectors of length M
$\mathbb{C}^{M\times N}$	Set of complex-valued matrices of dimension $M \times N$
$\boldsymbol{D}$	Diagonal matrix
d	Sensor spacing in a ULA
d_k	The k-th entry on the diagonal of the diagonal matrix $\boldsymbol{D}$
$\boldsymbol{E}$	Sensor measurement matrix from the $(N-1)$-undesired source signals
$\boldsymbol{G}$	Givens rotation matrix
$\boldsymbol{H}$	Signals of $(N-1)$-interfering sources under the PR model
$\boldsymbol{\mathcal{I}}$	Fisher Information matrix
$\boldsymbol{J}$	Selection matrix
j	$\sqrt{-1}$
M	Number of sensors
N	Number of sources
N_G	Number of directions in the discretized FOV
N_I	Number of iterations
N_R	Number of Monte-Carlo runs
$\boldsymbol{N}$	Sensor noise matrix
$\boldsymbol{n}(t)$	Noise measurement vector at time instant t
O	Order of complexity
$P(\theta)$	Power of the beamformer output signal at the direction θ
$\boldsymbol{P}$	Source covariance matrix

$\bar{\boldsymbol{P}}$	Transformed source covariance matrix
$\bar{\boldsymbol{p}}$	Vector containing non-redundant elements of $\bar{\boldsymbol{P}}$
$R(x)$	Rational function of the rank-one modification problem (see (5.2))
$\bar{R}_{k;p,q}(x)$	First-order rational approximant with pole d_k and parameters p and q
$\boldsymbol{R}$	Sensor covariance matrix
$\hat{\boldsymbol{R}}$	Sample covariance matrix
$\mathbb{R}$	Set of real numbers
$\mathbb{R}^M$	Set of real-valued vectors of length M
$\mathbb{R}^{M\times N}$	Set of real-valued matrices of dimension $M \times N$
$\boldsymbol{S}$	Source signal matrix
$\mathbb{S}$	Set of DOAs whose steering vector lies on the subspace spanned by $\boldsymbol{U}_\mathrm{s}$
$\boldsymbol{s}(t)$	Source signal vector at time instant t
T	Number of snapshots
$\boldsymbol{T}$	Transformation matrix (see (6.13))
t	Time instant
$\boldsymbol{U}_\mathrm{s}$	Signal subspace of the true covariance matrix
$\boldsymbol{U}_\mathrm{n}$	Noise subspace of the true covariance matrix
$\hat{\boldsymbol{U}}_\mathrm{s}$	Signal subspace of the sample covariance matrix
$\hat{\boldsymbol{U}}_\mathrm{n}$	Noise subspace of the sample covariance matrix
$\boldsymbol{w}(\theta)$	Beamformer vector at the direction θ
$\boldsymbol{W}$	Weighting matrix
$\boldsymbol{X}$	Sensor measurement matrix
$\boldsymbol{x}(t)$	Sensor measurement vector at time instant t
$\boldsymbol{Y}$	Generic data matrix
$\boldsymbol{Z}$	Matrix whose columns form an orthonormal basis of the subspace spanned by a projection matrix (see (B.1))
$\boldsymbol{z}$	Vector characterizing the rank-one modification component

Greek letters

$\boldsymbol{\alpha}$	Parameter vector
$\bar{\boldsymbol{\alpha}}$	Parameter vector of the transformed PR signal model
$\boldsymbol{\beta}$	Parameter vector characterizing the transformed steering matrix $\bar{\boldsymbol{A}}$
γ	Diagonal loading factor
$\Delta\theta$	Angular separation between sources
ϵ	Tolerance
Θ	Set of DOAs in the FOV
θ	Auxiliary DOA in the FOV
θ_n	DOA of the n-th source
$\hat{\theta}_n$	Estimated DOA of the n-th source

$\hat{\theta}_n^{(r)}$	Estimated DOA of the n-th source in the r-th Monte Carlo run
ϑ	The DOA whose steering structure is maintained in the PR signal model
$\mathbf{\Lambda}_{\mathrm{s}}$	Diagonal matrix containing the true signal eigenvalues
$\mathbf{\Lambda}_{\mathrm{n}}$	Diagonal matrix containing the true noise eigenvalues
$\hat{\mathbf{\Lambda}}_{\mathrm{s}}$	Diagonal matrix containing the sample signal eigenvalues
$\hat{\mathbf{\Lambda}}_{\mathrm{n}}$	Diagonal matrix containing the sample noise eigenvalues
ν	Noise power
$\hat{\nu}$	Estimated noise power
$\lambda(\cdot)$	Eigenvalue of the matrix argument
$\hat{\lambda}_m$	The m-th largest eigenvalue of the sample covariance matrix
$\bar{\lambda}_m$	The m-th largest eigenvalue of a rank-one modified matrix
ρ	Correlation factor
$\bar{\rho}$	Scaling factor of the rank-one modification component
σ_{s}^2	Signal power
τ	Iteration index
$\phi_k(x)$	Sum of all rational functions with poles smaller than d_k
$\psi_k(x)$	Sum of all rational functions with poles larger than or equal to d_k

Vectors and matrices

$\mathbf{0}$	Matrix of zeros with conformable dimensions
$\boldsymbol{e}_m$	The m-th column of an identity matrix
$\boldsymbol{I}_M$	Identity matrix of dimension $M \times M$
$\succeq$	Matrix inequality between Hermitian matrices
$\otimes$	Kronecker product
$\odot$	Hadamard product, i.e., element-wise product
$\boxtimes$	Extended blockwise Kronecker product
$\mathbf{\Pi}_{\boldsymbol{A}}$	Projection matrix onto the column space of $\boldsymbol{A}$
$\mathbf{\Pi}_{\boldsymbol{A}}^{\perp}$	Orthogonal projection matrix onto the column space of $\boldsymbol{A}$
$[\cdot]_{ij}$	The entry on the i-th row and j-column of the matrix argument
$\mathcal{S}(\cdot)$	Schur complement of the upper-left entry of the matrix argument
$\lvert\cdot\rvert$	Absolute value of the scalar argument
$\lvert\lvert\cdot\rvert\rvert$	Norm of the vector or matrix argument
$\lvert\lvert\cdot\rvert\rvert_{\mathrm{F}}$	Frobenius norm of the matrix argument
$(\cdot)^{\mathrm{T}}$	Transpose operator
$(\cdot)^{*}$	Conjugate operator
$(\cdot)^{\dagger}$	Moore-Penrose pseudo-inverse operator
$(\cdot)^{-1}$	Inverse operator
$(\cdot)^{1/2}$	Principal square root
$(\cdot)^{\mathrm{H}}$	Hermitian operator (conjugate transpose)

$\det(\cdot)$	Determinant of the matrix argument
$\operatorname{diag}(\cdot)$	Diagonal matrix whose elements are specified by the arguments
$\operatorname{rank}(\cdot)$	Rank of the matrix argument
$\operatorname{tr}(\cdot)$	Trace of the matrix argument
$\operatorname{vec}(\cdot)$	Vectorization operator

Miscellaneous

$\operatorname{Re}\{\cdot\}$	Real part
$\operatorname{Im}\{\cdot\}$	Imaginary part
$\mathbb{E}\{\cdot\}$	Expectation operator
${}^{N}\arg\min f$	N-local minima of the function f
$\frac{\mathrm{d}f}{\mathrm{d}x}$	First-order total derivative of the function f with respect to x
$\frac{\partial f}{\partial x}$	First-order partial derivative of the function f with respect to x
$\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}$	Second-order total derivative of the function f with respect to x
$\mathrm{As}N$	Asymptotic normal distribution
$\dim(\cdot)$	Dimension of a subspace
$\operatorname{span}(\cdot)$	Subspace spanned by vector arguments or columns of the matrix argument

Bibliography

[ABB+99] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen, *LAPACK Users' Guide*, 3rd ed. Philadelphia, PA: Society for Industrial and Applied Mathematics, 1999.

[AJM07] Y. I. Abramovich, B. A. Johnson, and X. Mestre, "Performance Breakdown in MUSIC, G-MUSIC and Maximum Likelihood Estimation," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 2, Apr. 2007, pp. II–1049–II–1052.

[Ath05] F. Athley, "Threshold Region Performance of Maximum Likelihood Direction of Arrival Estimators," *IEEE Transactions on Signal Processing*, vol. 53, no. 4, pp. 1359–1373, Apr. 2005.

[Ban71] W. J. Bangs, "Array Processing with Generalized Beamformers," Ph.D. dissertation, 1971.

[Bar83] A. Barabell, "Improving the Resolution Performance of Eigenstructure-based Direction-Finding Algorithms," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 8, Apr. 1983, pp. 336–339.

[Bar07] R. G. Baraniuk, "Compressive Sensing [Lecture Notes]," *IEEE Signal Processing Magazine*, vol. 24, no. 4, pp. 118–121, Jul. 2007.

[BCD+84] A. J. Barabell, J. B. L. Capon, D. F. Delong, J. R. Johnson, and K. D. Senne, "Performance Comparison of Superresolution Array Processing Algorithms," in *Technical Report Project Rep. TST-72*. MIT Lincoln Laboratory, Lexington, Massachusetts, May 1984.

[BCG+07] I. Burylov, M. Chuvelev, B. Greer, G. Henry, S. Kuznetsov, and B. Sabanin, "Intel Performance Libraries: Multi-Core-Ready Software for Numeric-Intensive Computation." *Intel Technology Journal*, vol. 11, no. 4, pp. 299 – 308, 2007.

[BM86] Y. Bresler and A. Macovski, "Exact Maximum Likelihood Parameter Estimation of Superimposed Exponential Signals in Noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, no. 5, pp. 1081–1089, Oct. 1986.

[BNS78] J. R. Bunch, C. P. Nielsen, and D. C. Sorensen, "Rank-One Modification of the Symmetric Eigenproblem," *Numerische Mathematik*, vol. 31, no. 1, pp. 31–48, 1978.

[Boh84] J. Bohme, "Estimation of Source Parameters by Maximum Likelihood and Nonlinear Regression," in *ICASSP '84. IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 9, Mar. 1984, pp. 271–274.

[Bur10] Burden, R.L. and Faires, J.D., "Numerical Analysis." Cengage Learning, 2010, pp. 48–49.

[Cap69] J. Capon, "High-resolution Frequency-wavenumber Spectrum Analysis," *Proceedings of the IEEE*, vol. 57, no. 8, pp. 1408–1418, Aug. 1969.

[Car87] G. C. Carter, "Coherence and time delay estimation," *Proceedings of the IEEE*, vol. 75, no. 2, pp. 236–255, Feb. 1987.

[Chi78] D. Childers, *Modern Spectrum Analysis*, ser. IEEE Press Selected Reprint Series. IEEE Press, 1978.

[Cra45] H. Cramér, *Mathematical Methods of Statistics*. Princeton University Press, 1945.

[CRT06] E. J. Candès, J. K. Romberg, and T. Tao, "Stable Signal Recovery from Incomplete and Inaccurate Measurements," *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.

[CVY14] P.-J. Chung, M. Viberg, and J. Yu, "Chapter 14 - DOA Estimation Methods and Algorithms," in *Volume 3: Array and Statistical Signal Processing*, ser. Academic Press Library in Signal Processing, R. C. Abdelhak M. Zoubir, Mats Viberg and S. Theodoridis, Eds. Elsevier, 2014, vol. 3, pp. 599 – 650.

[EEFSRJ82] J. E. Evans, D. F. Sun, and J. R. Johnson, "Application of Advanced Signal Processing Techniques to Angle of Arrival Estimation in ATC Navigation and Surveillance Systems," *Technical report, MIT Lincoln Laboratory*, Jun. 1982.

[EY36] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, vol. 1, pp. 211–218, 1936.

[Fan49] K. Fan, "On a Theorem of Weyl Concerning Eigenvalues of Linear Transformations I," *Proceedings of the National Academy of Sciences*, vol. 35, no. 11, pp. 652–655, 1949.

[FLB04] P. Forster, P. Larzabal, and E. Boyer, "Threshold Performance Analysis of Maximum Likelihood DOA Estimation," *IEEE Transactions on Signal Processing*, vol. 52, no. 11, pp. 3183–3191, Nov. 2004.

[GVL13] G. Golub and C. Van Loan, *Matrix Computations*, 4th ed., ser. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, 2013.

[HJ12] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. New York, NY, USA: Cambridge University Press, 2012.

[HN95] M. Haardt and J. A. Nossek, "Unitary ESPRIT: How to Obtain Increased Estimation Accuracy With a Reduced Computational Burden," *IEEE Transactions on Signal Processing*, vol. 43, no. 5, pp. 1232–1242, May 1995.

[HPRK14] M. Haardt, M. Pesavento, F. Roemer, and M. N. E. Korso, "Chapter 15 - Subspace Methods and Exploitation of Special Array Structures," in *Volume 3: Array and Statistical Signal Processing*, ser. Academic Press Library in Signal Processing, R. C. Abdelhak M. Zoubir, Mats Viberg and S. Theodoridis, Eds. Elsevier, 2014, vol. 3, pp. 651 – 717.

[HRW98] R. A. Horn, N. H. Rhee, and S. Wasin, "Eigenvalue Inequalities and Equalities," *Linear Algebra and its Applications*, vol. 270, no. 1, pp. 29 – 44, 1998.

[HSS93] P. Handel, P. Stoica, and T. Soderstrom, "Capon Method for DOA Estimation: Accuracy and Robustness Aspects," in *Nonlinear Digital Signal Processing, 1993. IEEE Winter Workshop on*, 1993, pp. 71–75.

[Hu14] Y. H. Hu, "Chapter 18 - Source Localization and Tracking," in *Volume 3: Array and Statistical Signal Processing*, ser. Academic Press Library in Signal Processing, R. C. Abdelhak M. Zoubir, Mats Viberg and S. Theodoridis, Eds. Elsevier, 2014, vol. 3, pp. 799 – 817.

[HW88] T. Hayden and J. Wells, "Approximation by Matrices Positive Semidefinite on a Subspace," *Linear Algebra and its Applications*, vol. 109, pp. 115 – 130, 1988.

[JAM08] B. A. Johnson, Y. I. Abramovich, and X. Mestre, "MUSIC, G-MUSIC, and Maximum-Likelihood Performance Breakdown," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3944–3958, Aug. 2008.

[Kay93] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[KB86] M. Kaveh and A. Barabell, "The Statistical Performance of the MUSIC and the Minimum-norm Algorithms in Resolving Plane Waves in Noise," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 34, no. 2, pp. 331–341, Apr. 1986.

[Kes86] S. Kesler, *Modern Spectrum Analysis, II*, ser. IEEE Press Selected Reprint Series. IEEE Press, 1986.

[KFP92] H. Krim, P. Forster, and J. G. Proakis, "Operator Approach to Performance Analysis of Root-MUSIC and Root-Min-Norm," *IEEE Transactions on Signal Processing*, vol. 40, no. 7, pp. 1687–1696, Jul. 1992.

[KT83] R. Kumaresan and D. W. Tufts, "Estimating the Angles of Arrival of Multiple Plane Waves," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-19, no. 1, pp. 134–139, Jan. 1983.

[KV96] H. Krim and M. Viberg, "Two Decades of Array Signal Processing Research: The Parametric Approach," *IEEE Signal Processing Magazine*, vol. 13, no. 4, pp. 67–94, Jul. 1996.

[Lac71] R. T. Lacoss, "Data Adaptive Spectral Analysis Methods," *Geophysics*, vol. 36, no. 4, pp. 661–675, 1971. [Online]. Available: https://doi.org/10.1190/1.1440203

[LC98] E. Lehmann and G. Casella, *Theory of Point Estimation*. Springer, 1998.

[Li94] R.-C. Li, "Solving Secular Equations Stably and Efficiently," EECS Department, University of California, Berkeley, Tech. Rep. UCB/CSD-94-851, Dec. 1994.

[LLV93] F. Li, H. Liu, and R. J. Vaccaro, "Performance Analysis for DOA Estimation Algorithms: Unification, Simplification, and Observations," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 29, no. 4, pp. 1170–1184, Oct. 1993.

[LSW03] J. Li, P. Stoica, and Z. Wang, "On Robust Capon Beamforming and Diagonal Loading," *IEEE Transactions on Signal Processing*, vol. 51, no. 7, pp. 1702–1715, Jul. 2003.

[LY12] Y. Li and P. Yeh, "An Interpretation of the Moore-Penrose Generalized Inverse of a Singular Fisher Information Matrix," *IEEE Transactions on Signal Processing*, vol. 60, no. 10, pp. 5532–5536, Oct. 2012.

[Mag85] J. R. Magnus, "On Differentiating Eigenvalues and Eigenvectors," *Econometric Theory*, vol. 1, no. 2, pp. 179–191, 1985.

[Mel97] A. Melman, "A Numerical Comparison of Methods for Solving Secular Equations," *Journal of Computational and Applied Mathematics*, vol. 86, no. 1, pp. 237 – 249, 1997.

[OSR98] B. Ottersten, P. Stoica, and R. Roy, "Covariance Matching Estimation Techniques for Array Signal Processing Applications," *Digital Signal Processing*, vol. 8, no. 3, pp. 185–210, Jul. 1998.

[OVSN93] B. Ottersten, M. Viberg, P. Stoica, and A. Nehorai, *Exact and Large Sample Maximum Likelihood Techniques for Parameter Estimation and Detection in Array Processing*, ser. Springer Series in Information Sciences. Springer, Berlin, Heidelberg, 1993, vol. 25, pp. 99–151.

[PGH00] M. Pesavento, A. B. Gershman, and M. Haardt, "Unitary Root-MUSIC With a Real-Valued Eigendecomposition: A Theoretical and Experimental Performance Study," *IEEE Transactions on Signal Processing*, vol. 48, no. 5, pp. 1306–1314, May 2000.

[PK89] S. U. Pillai and B. H. Kwon, "Forward/backward Spatial Smoothing Techniques For Coherent Signal Identification," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 1, pp. 8–15, Jan. 1989.

[POR+93] A. Paulraj, B. Ottersten, R. Roy, A. Swindlehurst, G. Xu, and T. Kailath, "Subspace Methods for Directions-Of-Arrival Estimation," *Handbook of Statistics*, vol. 10, pp. 693 – 739, 1993.

[PPG12] P. Parvazi, M. Pesavento, and A. B. Gershman, "Rooting-Based Harmonic Retrieval Using Multiple Shift-Invariances: The Complete and the Incomplete Sample Cases," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1556–1570, 2012.

[QHSS16] C. Qian, L. Huang, N. D. Sidiropoulos, and H. C. So, "Enhanced PUMA for Direction-of-Arrival Estimation and Its Performance Analysis," *IEEE Transactions on Signal Processing*, vol. 64, no. 16, pp. 4127–4137, Aug. 2016.

[RAKS09] A. Rembovsky, A. Ashikhmin, V. Kozmin, and S. Smolskiy, *Radio Monitoring: Problems, Methods and Equipment*, ser. Lecture Notes in Electrical Engineering. Springer, 2009.

[Rao45] C. R. Rao, "Information and Accuracy Attainable in the Estimation of Statistical Parameters," *Bulletin of Calcutta Mathematical Society*, vol. 37, pp. 81–91, 1945.

[Red79] S. S. Reddi, "Multiple Source Location-A Digital Approach," *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-15, no. 1, pp. 95–105, Jan. 1979.

[RFCL06] A. Renaux, P. Forster, E. Chaumette, and P. Larzabal, "On the High-SNR Conditional Maximum-Likelihood Estimator Full Statistical Characterization," *IEEE Transactions on Signal Processing*, vol. 54, no. 12, pp. 4840–4843, Dec. 2006.

[RG09] M. Rubsamen and A. B. Gershman, "Direction-of-Arrival Estimation for Nonuniform Sensor Arrays: From Manifold Separation to Fourier Domain MUSIC Methods," *IEEE Transactions on Signal Processing*, vol. 57, no. 2, pp. 588–599, 2009.

[RH89] B. D. Rao and K. V. S. Hari, "Performance Analysis of Root-MUSIC," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 12, pp. 1939–1949, Dec. 1989.

[Ric05] C. D. Richmond, "Capon Algorithm Mean-squared Error Threshold SNR Prediction And Probability of Resolution," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2748–2764, Aug. 2005.

[RK89] R. Roy and T. Kailath, "ESPRIT-Estimation of Signal Parameters via Rotational Invariance Techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 984–995, Jul. 1989.

[Rot71] T. J. Rothenberg, "Identification in Parametric Models," *Econometrica*, vol. 39, no. 3, pp. 577–591, 1971. [Online]. Available: http://www.jstor.org/stable/1913267

[SB98] P. Stoica and Boon Chong Ng, "On the Cramér-Rao Bound under Parametric Constraints," *IEEE Signal Processing Letters*, vol. 5, no. 7, pp. 177–179, Jul. 1998.

[Sch86] R. Schmidt, "Multiple Emitter Location and Signal Parameter Estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[SJGL14] A. L. Swindlehurst, B. D. Jeffs, G. S. Granados, and J. Li, "Chapter 20 - Applications of Array Signal Processing," in *Academic Press Library in Signal Processing: Volume 3*, ser. Academic Press Library in Signal Processing, A. M. Zoubir, M. Viberg, R. Chellappa, and S. Theodoridis, Eds. Elsevier, 2014, vol. 3, pp. 859 – 953.

[Sle54] D. Slepian, "Estimation of Signal Parameters in the Presence of Noise," *Transactions of the IRE Professional Group on Information Theory*, vol. 3, no. 3, pp. 68–89, Mar. 1954.

[SLG01] P. Stoica, E. G. Larsson, and A. B. Gershman, "The Stochastic CRB for Array Processing: A Textbook Derivation," *IEEE Signal Processing Letters*, vol. 8, no. 5, pp. 148–150, May 2001.

[SM01] P. Stoica and T. L. Marzetta, "Parameter Estimation Problems with Singular Information Matrices," *IEEE Transactions on Signal Processing*, vol. 49, no. 1, pp. 87–90, Jan. 2001.

[SM05] P. Stoica and R. Moses, *Spectral Analysis of Signals*. Pearson Prentice Hall, 2005.

[SM14] B. M. Sadler and T. J. Moore, "Chapter 8 - Performance Analysis and Bounds," in *Academic Press Library in Signal Processing: Volume 3*, ser. Academic Press Library in Signal Processing, A. M. Zoubir, M. Viberg, R. Chellappa, and S. Theodoridis, Eds. Elsevier, 2014, vol. 3, pp. 297 – 322.

[SN89] P. Stoica and A. Nehorai, "MUSIC, Maximum Likelihood, and Cramer-Rao bound," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 5, pp. 720–741, May 1989.

[SN90a] ——, "MUSIC, Maximum Likelihood, and Cramér-Rao Bound: Further Results and Comparisons," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 12, pp. 2140–2150, Dec. 1990.

[SN90b] ——, "Performance Study of Conditional and Unconditional Direction-of-Arrival Estimation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 10, pp. 1783–1795, Oct. 1990.

[SN91] P. Stoica and A. Nehorai, "Performance Comparison of Subspace Rotation and MUSIC Methods for Direction Estimation," *IEEE Transactions on Signal Processing*, vol. 39, no. 2, pp. 446–453, Feb. 1991.

[SPP18] C. Steffens, M. Pesavento, and M. E. Pfetsch, "A Compact Formulation for the $\ell_{2,1}$ Mixed-Norm Minimization Problem," *IEEE Transactions on Signal Processing*, vol. 66, no. 6, pp. 1483–1497, 2018.

[SPPZ18] W. Suleiman, P. Parvazi, M. Pesavento, and A. M. Zoubir, "Non-Coherent Direction-of-Arrival Estimation Using Partly Calibrated Arrays," *IEEE Transactions on Signal Processing*, vol. 66, no. 21, pp. 5776–5788, 2018.

[SS90] P. Stoica and K. C. Sharman, "Maximum Likelihood Methods for Direction-of-Arrival Estimation," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 38, no. 7, pp. 1132–1143, Jul. 1990.

[STM+19] D. Schenck, M. Trinh-Hoang, X. Mestre, M. Viberg, and M. Pesavento, "Full Covariance Fitting DoA Estimation Using Partial Relaxation Framework," in *2019 27th European Signal Processing Conference (EUSIPCO)*, Sep. 2019, pp. 1–5.

[Sto02] J. S. Stone, "Method of Determining the Direction of Space-Telegraph Signals," US Patent US716 134A, 1902.

[SWK85] T.-J. Shan, M. Wax, and T. Kailath, "On Spatial Smoothing for Direction-of-Arrival Estimation of Coherent Signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 4, pp. 806–811, Aug. 1985.

[THVP17] M. Trinh-Hoang, M. Viberg, and M. Pesavento, "Improved DOA Estimators Using Partial Relaxation Approach," in *IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, Dec. 2017, pp. 1–5.

[THVP18a] ——, "An Improved DOA Estimator Based on Partial Relaxation Approach," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2018, pp. 3246–3250.

[THVP18b] ——, "Partial Relaxation Approach: An Eigenvalue-Based DOA Estimator Framework," *IEEE Transactions on Signal Processing*, 2018.

[VB88] B. D. Van Veen and K. M. Buckley, "Beamforming: A Versatile Approach to Spatial Filtering," *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4–24, Apr. 1988.

[VB95] C. Vaidyanathan and K. Buckley, "Performance Analysis of the MVDR Spatial Spectrum Estimator," *IEEE Transactions on Signal Processing*, vol. 43, no. 6, pp. 1427–1437, Jun. 1995.

[VO91] M. Viberg and B. Ottersten, "Sensor Array Processing Based on Subspace Fitting," *IEEE Transactions on Signal Processing*, vol. 39, no. 5, pp. 1110–1121, May 1991.

[Vor14] S. A. Vorobyov, "Chapter 12 - Adaptive and Robust Beamforming," in *Volume 3: Array and Statistical Signal Processing*, ser. Academic Press Library in Signal Processing, R. C. Abdelhak M. Zoubir, Mats Viberg and S. Theodoridis, Eds. Elsevier, 2014, vol. 3, pp. 503 – 552.

[VT04] H. Van Trees, *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*, ser. Detection, Estimation, and Modulation Theory. Wiley, 2004.

[Wax85] M. Wax, "Detection and estimation of superimposed signals," Ph.D. dissertation, Stanford University, Stanford, CA, Mar. 1985.

[WZ89] M. Wax and I. Ziskind, "On Unique Localization of Multiple Sources by Passive Sensor Arrays," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 996–1000, Jul. 1989.

[ZB11] R. Zekavat and R. Buehrer, *Handbook of Position Location: Theory, Practice and Advances*, ser. IEEE Series on Digital & Mobile Communication. Wiley, 2011.

[ZJLJ11] S. Zhiguang, Z. Jianxiong, H. Lei, and L. Jicheng, "A New Derivation of Constrained Cramér-Rao Bound Via Norm Minimization," *IEEE Transactions on Signal Processing*, vol. 59, no. 4, pp. 1879–1882, Apr. 2011.

[ZW88] I. Ziskind and M. Wax, "Maximum Likelihood Localization of Multiple Sources by Alternating Projection," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 10, pp. 1553–1560, Oct. 1988.

Lebenslauf

Name:	Minh Trinh Hoang
Geburtsdatum:	11. November 1991
Geburtsort:	Hanoi

Schulausbildung

2002-2009	Gymnasium Hanoi-Amsterdam in Hanoi, Vietnam

Studium

seit 2017	Promotion in Elektrotechnik und Informationstechnik, Fachgebiet Nachrichtentechnische Systeme, Institut für Nachrichtentechnik, Technische Universität Darmstadt
2014-2017	Studium der Elektrotechnik und Informationstechnik, Technische Universität Darmstadt, Studienabschluss: Master of Science
2011-2014	Studium der Elektrotechnik und Informationstechnik, Technische Universität Darmstadt, Studienabschluss: Bachelor of Science

Berufstätigkeit

seit Februar 2017	Wissenschaftlicher Mitarbeiter, Fachgebiet Nachrichtentechnische Systeme, Institut für Nachrichtentechnik, Technische Universität Darmstadt