



Università degli Studi di Napoli Federico II
Ph.D. Program in
Information **T**echnology and **E**lectrical **E**ngineering
XXXV Cycle

THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Antenna-Based Techniques for Modern Radars

by

MASSIMO ROSAMILIA

Supervisor: Prof. Antonio De Maio



SCUOLA POLITECNICA E DELLE SCIENZE DI BASE

DIPARTIMENTO DI INGEGNERIA **E**LETRICA E DELLE **T**ECNOLOGIE DELL'**I**NFORMAZIONE

*"There are two types of people in the world: those
who can extrapolate from incomplete data."*

ANTENNA-BASED TECHNIQUES FOR MODERN RADARS

Ph.D. Thesis presented
for the fulfillment of the Degree of Doctor of Philosophy
in Information Technology and Electrical Engineering
by

MASSIMO ROSAMILIA

December 2022



Approved as to style and content by

A handwritten signature in black ink, appearing to read 'Antonio De Maio', written over a horizontal line.

Prof. Antonio De Maio, Advisor

Università degli Studi di Napoli Federico II

Ph.D. Program in Information Technology and Electrical Engineering
XXXV cycle - Chairman: Prof. Stefano Russo



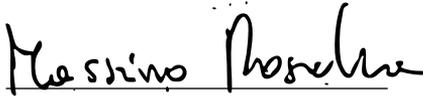
<http://itee.dieti.unina.it>

Candidate's declaration

I hereby declare that this thesis submitted to obtain the academic degree of Philosophiæ Doctor (Ph.D.) in Information Technology and Electrical Engineering is my own unaided work, that I have not used other than the sources indicated, and that all direct and indirect sources are acknowledged as references.

Parts of this dissertation have been published in international journals and/or conference articles (see list of the author's publications at the end of the thesis).

Napoli, December 12, 2022

A handwritten signature in black ink that reads "Massimo Rosamilia". The signature is written in a cursive style and is underlined.

Massimo Rosamilia

Abstract

Multichannel radar systems must comply with very stringent operational requirements, involving adaptive detection and estimation tasks, as well as the ability to handle many tracks simultaneously without degrading search performance when operating in interference environments. An essential role is hence played by the detection of the target and its accurate localization. Therefore, this PhD Thesis deals with the development of multichannel detection and estimation algorithms for radar signal processing, presenting innovative architectures for the joint target detection and angle estimation, tailored strategies to recover performance in missing-data scenarios, as well as novel signal processing methods developed for Frequency Diverse-Array Multiple Input Multiple Output (FDA-MIMO) radars (also exploiting the polarimetric domain). After an overview of the state of the art of modern multichannel radar systems, where some relevant issues of practical interest are highlighted, the problem of joint radar target detection and angle estimation is addressed in the context of a multichannel phased array radar. Moreover, the estimation of a structured covariance matrix under the missing-data context is tackled by means of the Expectation-Maximization (EM) algorithm. Within the same setting, the problem of detecting a prospective target embedded in Gaussian interference with an unknown (but possibly structured) covariance matrix is studied. After that, assuming a FDA-MIMO radar operating in a background of Gaussian interference with known spectral properties, a joint target angle and incremental range estimation architecture is derived. Then, the target detection problem with a polarimetric FDA-MIMO radar is investigated. The performance analysis of the proposed architectures and the corresponding discussions are developed in terms of computational complexity and detection performance and/or estimation accuracy. Moreover, for comparison purposes, some counterparts available in the open literature as well as theoretical benchmark limits are also considered, corroborating the effectiveness of the considered strategies in some diverse simulated scenarios. Finally, the advances described in this thesis work are summarized and discussed along with emerging trends and issues to be addressed in the future.

Keywords: modern radar system, adaptive detection, ML estimation, missing-data, FDA-MIMO, Polarimetric radar.

Sintesi in lingua italiana

I sistemi radar multicanale devono soddisfare requisiti operativi molto stringenti, che prevedono compiti di rivelazione adattiva e di stima, nonché la capacità di gestire molteplici tracce senza compromettere le prestazioni di ricerca quando si opera in presenza di interferenze. Un ruolo essenziale è quindi svolto dal processo di rivelazione del bersaglio e dalla sua accurata localizzazione. Per questo motivo, l'obiettivo di questa tesi di dottorato è la progettazione di algoritmi innovativi per l'elaborazione del segnale radar, sviluppando tecniche per rivelare il bersaglio e fornire congiuntamente la relativa stima dell'angolo, strategie ad-hoc per ristabilire le prestazioni in scenari con dati mancanti, nonché nuovi metodi di elaborazione del segnale sviluppati per radar Frequency Diverse-Array Multiple Input Multiple Output (FDA-MIMO), sfruttando anche il dominio polarimetrico. Dopo una panoramica sullo stato dell'arte dei moderni sistemi radar multicanale, in cui vengono evidenziate diverse problematiche di interesse pratico, viene studiato il problema di rivelare il bersaglio e fornire contemporaneamente la sua stima della direzione di arrivo utilizzando un radar multicanale phased array. Inoltre, è affrontato il problema della stima di una matrice di covarianza strutturata in uno scenario con dati mancanti, impiegando l'algoritmo Expectation-Maximization (EM). Nello stesso contesto, è esaminato il problema della rivelazione di un bersaglio immerso in un'interferenza Gaussiana con una matrice di covarianza non nota, ma eventualmente strutturata. Successivamente, ipotizzando un radar FDA-MIMO operante in un contesto di interferenza Gaussiana con proprietà spettrali note, viene proposto un algoritmo per la stima dell'angolo e della distanza incrementale del bersaglio. Inoltre, viene analizzato il problema della rivelazione adattiva del bersaglio con un radar FDA-MIMO polarimetrico assumendo disturbo additivo Gaussiano con matrice di covarianza non nota. La progettazione degli algoritmi proposti è rigorosamente descritta e corredata dall'analisi della relativa complessità computazionale nonché dallo studio delle proprietà di convergenza delle procedure di ottimizzazione impiegate. Infine, l'efficacia dei metodi elaborati in questo lavoro di tesi è analizzata mediante simulazione numerica al fine di studiare le prestazioni di rivelazione e/o l'accuratezza di stima. Completano l'analisi il confronto con alcune controparti disponibili nella letteratura aperta nonché con benchmark teorici di riferimento. Conclude questo lavoro di

tesi un riepilogo dei problemi affrontati e delle strategie di elaborazione sviluppate evidenziando possibili sviluppi futuri che potrebbero essere perseguiti.

Parole chiave: sistema radar moderno, rivelazione adattiva, stima ML, dati mancanti, FDA-MIMO, radar polarimetrico.

Contents

Abstract	i
Sintesi in lingua italiana	iii
Acknowledgements	ix
List of Acronyms	xiv
List of Figures	xxi
List of Tables	xxiii
List of Symbols	xxv
1 Introduction	1
2 Single-Pulse Simultaneous Target Detection and Angle Estimation in a Multichannel Phased Array Radar	9
2.1 Problem Formulation	10
2.1.1 Pointing Errors	12
2.2 System Design: Joint Detection and Angle Estimation	14
2.2.1 Constrained GLRT Detector for 1-D Scenario	17
2.2.2 Constrained GLRT Detector for 2-D Scenario	19
2.3 Performance Analysis	25
2.3.1 Linearized Array Manifold Signal Model	27
2.3.2 Actual Array Manifold	36

2.3.3	Performance Analysis Employing Two-stage Processors	40
2.4	Conclusion	42
3	Structured Covariance Matrix Estimation with Missing-Data via Expectation-Maximization	43
3.1	Problem Formulation	44
3.2	Covariance Matrix Estimation Procedure	47
3.2.1	EM Algorithm	48
3.2.2	Adaptive Beamforming	54
3.2.3	Detection of Number of Sources	57
3.3	Performance Analysis	62
3.3.1	Adaptive Beamforming	63
3.3.2	Detection of Number of Sources	66
3.4	Conclusion	70
4	Adaptive Radar Detection in the Presence of Missing-Data	73
4.1	Problem Formulation	74
4.2	Design of Decision Rules	76
4.2.1	Parameters estimation under H_1	77
4.2.2	Decision rules	82
4.3	Performance Analysis	83
4.3.1	Unconstrained estimation	85
4.3.2	Centro-Hermitianity constraint	85
4.3.3	Lower bound of the white noise power level constraint	87
4.3.4	Analysis on measured data	88
4.4	Conclusion	91
5	Single-Snapshot Angle and Incremental Range Estimation for FDA-MIMO Radar	93
5.1	Signal Model for FDA-MIMO Radar	94

5.1.1	Transmitted Signal Model	94
5.1.2	Received Signal Model	95
5.2	Joint Angle-Range Estimation for FDA-MIMO Radar	97
5.2.1	ML Estimation of u and δ	97
5.2.2	Approximated Methods for Range and Angle Estimation	98
5.2.3	Discussion on the Computational Complexity	107
5.3	CRLB for FDA-MIMO radar	109
5.4	Performance Analysis	111
5.4.1	White Noise Interference Scenario	112
5.4.2	Coherent Repeaters In The Interference Scenario	117
5.5	Conclusion	118
6	Adaptive Target Detection with Polarimetric FDA-MIMO Radar	119
6.1	Signal Model For Polarimetric FDA-MIMO Radar	120
6.1.1	Received Signal for Polarimetric FDA-MIMO radar	121
6.2	Formulation of the Detection Problem for Polarimetric FDA-MIMO Radar	124
6.2.1	Detection with GLRT	125
6.2.2	Detection with 2SGLRT	126
6.3	Optimization Problem (6.20)	127
6.3.1	LAM Method	128
6.3.2	GPM	131
6.3.3	CD Method	134
6.3.4	Bounded CFARness Analysis of the Derived Detectors and Computational Complexities	135
6.4	Simulation Results	137
6.5	Conclusion	144
7	Conclusions	147

A	Appendix to Chapter 2	151
A.1	Adaptive Monopulse Estimation Algorithm	151
A.2	Proof of Proposition 1	152
A.3	Solution to Problem (2.41)	152
A.4	CRLB Derivation	154
B	Appendix to Chapter 3	157
B.1	Proof of (3.22)	157
B.2	Proof of Proposition 2	158
B.2.1	Proof of the first item	159
B.2.2	Proof of the second item	160
B.3	Derivation of (3.28) and (3.29)	163
C	Appendix to Chapter 4	167
C.1	Closed-form expression of the score function (3.18)	167
C.2	Detailed expressions for LI-based detectors	169
D	Appendix to Chapter 5	171
D.1	Expressions of the \mathbf{s}_u and \mathbf{s}_δ	171
D.2	Computations of D_u and D_δ	172
E	Appendix to Chapter 6	175
E.1	Expressions for \mathbf{H}_u and \mathbf{H}_δ	175
E.2	Proof of Proposition 3	176
E.3	Computation of $\mathbf{g}_\Delta(u, \delta)$	178
E.4	Proof of $g(u, \delta) \in C_L^{1,1}$	178
E.5	Proof that the second derivatives of $g(u, \delta)$ are periodic functions	180
	Bibliography	183
	Author's Publications	195

Acknowledgements

This work has been partially developed during the visit of Massimo Rosamilia with:

- the Cranfield University, Shrivenham, U.K., from September to November 2021, under the scientific supervision of Prof. A. Balleri;
- the University of Luxembourg, Luxembourg, from September to December 2022, under the scientific supervision of Profs. Bhavani Shankar and Mohammad Alae-Kerahroodi.

Moreover, special thanks to Vincenzo, who has helped me countless times with all kinds of issues: hardware, software, and administrative! A particular acknowledgement goes to Augusto, for his willingness to solve the most complex mathematical quibbles. Special appreciation for my advisor, Prof. Antonio De Maio, who is a constant source of inspiration for me. Finally, a sincere gratitude to Prof. Stefano Marano, who has guided and helped me pursue this path.

List of Acronyms

The following acronyms are used throughout the thesis.

1-D	One-Dimensional
2-D	Two-Dimensional
3-D	Three-Dimensional
2SGLRT	Two-Step GLRT
AESA	Active Electronically Scanned Array
AGMP-CC	Adaptive Generalized Monopulse Procedure with Complex Correction
AIC	Akaike Information Criterion
AMF	Adaptive Matched Filter
AMP	Adaptive Monopulse Procedure
ASNR	Array Signal-to-Noise Ratio
CD	Coordinate Descent
CDO	CD Optimization

CFAR	Constant False Alarm Rate
CH	Centro-Hermitianity
CM	Conditional Maximizations
CNR	Clutter to Noise Ratio
CRLB	Cramér-Rao Lower Bound
CUT	Cell Under Test
DBF	Digital Beamformer
DO	Dinkelbach's Optimization
DOA	Direction of Arrival
DOF	Degree of Freedom
DSC	Double-Step Corrected
ECM	Expectation-Conditional Maximization
EM	Expectation-Maximization
ESM	Electronic Support Measurement
EVD	EigenValue Decomposition
FB	Forward-Backward
FDA	Frequency Diverse-Array
FIM	Fisher Information Matrix
FML	Fast Maximum Likelihood
GEM	Generalized EM

GLR	Generalized Likelihood Ratio
GLRT	GLR Test
GPM	Gradient Projection Method
HQC	Hannan–Quinn Information Criterion
IID	Independent and Identically Distributed
JNR	Jammer to Noise Ratio
LAM	Linearized Array Manifold
LI	Linear Interpolation
LRT	Likelihood Ratio Test
M-EM	Multi-cycle EM
MBI	Maximum Block Improvement
MDL	Minimum Description Length
MDGS	Multi-Dimensional Grid Search
MIMO	Multiple Input Multiple Output
ML	Maximum Likelihood
MLE	ML Estimate
MSE	Mean Square Error
MVDR	Minimum Variance Distortionless Response
P_{fa}	False Alarm Probability
P_D	Probability of Detection

PDF	Probability Density Function
RCS	Radar Cross-Section
RMSE	Root Mean Square Error
SCM	Sample Covariance Matrix
SD	Subspace Detector
SDR	Software Defined Radio
S/I	Signal to Interference Power Ratio
SINR	Signal to Interference plus Noise Ratio
SLL	SideLobe Level
SP	Single Polarization
SSBW	Single-Side Beam-Width
STAP	Space-Time Adaptive Processing
ULA	Uniform Linear Array
UMP	Uniformly Most Powerful
URA	Uniform Rectangular Array

List of Figures

2.1	Two dimensional symmetric planar array geometry with equi-spaced antennas in the x and y directions. (R, θ, ϕ) are the polar coordinates of target.	11
2.2	Normalized correlation versus direction cosine offset, assuming symmetric array: (a) 1-D case with $N = 9$ and $\bar{u} = 0$; (b) 2-D case with $N = M = 5$ and $\bar{u} = \bar{v} = 0$. Dashed red lines delimitate the 3 dB beamwidth.	15
2.3	Detection performance for a ULA with 16 antennas in Scenario 1 for different mismatches, assuming $K = 32$ and $\alpha = 0.5$. The actual DOA displacement is drawn from a uniform distribution over $[-\Delta, \Delta]$: (a) $\Delta = 0$, (b) $\Delta = 0.1$, (c) $\Delta = 0.25$, and (d) $\Delta = 0.5$	28
2.4	Estimation performance for a ULA with 16 antennas in Scenario 1 assuming $\alpha = 0.2$ and considering different sample support sizes, i.e., $K = [16, 32, 48, \infty]$, and several target locations: (a) $\Delta u = 0$, (b) $\Delta u = 0.05$, (c) $\Delta u = 0.1$, and (d) $\Delta u = 0.5$	29

- 2.5 Detection performance for a URA with 5×5 antennas, $K = 50$ secondary data, and different mismatch conditions, in Scenario 2. The constraint parameters α and β are set at 0.5. The target's location $(\Delta u, \Delta v)$ is modeled as a pair of statistically independent uniform random variables in the region $[-\Delta_1, \Delta_1] \times [-\Delta_2, \Delta_2]$: (a) $(\Delta_1, \Delta_2) = (0, 0)$, (b) $(\Delta_1, \Delta_2) = (0, 0.05)$, (c) $(\Delta_1, \Delta_2) = (0.05, 0.05)$, and (d) $(\Delta_1, \Delta_2) = (0.5, 0.5)$ 30
- 2.6 Estimation performance for a URA with 5×5 antennas in Scenario 2 for different sample support sizes, i.e., $K = [30, 50, 70, \infty]$ and several target locations, i.e., (a) $(\Delta u, \Delta v) = (0, 0)$, (b) $(\Delta u, \Delta v) = (0, 0.05)$, (c) $(\Delta u, \Delta v) = (0.05, 0.05)$, and (d) $(\Delta u, \Delta v) = (0.5, 0.5)$. Therein $\alpha = \beta = 0.2$. In Fig. (b) are also reported two box-and-whisker plots at 13 dB and 20 dB, respectively. 31
- 2.7 Detection performance for the actual array manifold assuming the interfering environment of Scenario 3. Figs. (a) and (b) report detection performance for a symmetric ULA with 9 antennas with $K = 18$, target location uniformly distributed over $[-\Delta, \Delta]$, and $\alpha = \Delta$: (a) $\Delta = 0.0523$ and (b) $\Delta = 0.891/N$. Figs. (c) and (d) correspond to a symmetric URA with 5×5 antennas, with $K = 75$, target location offset Δu and Δv modeled as a pair of statistically independent uniformly distributed random variables over $[-\Delta_1, \Delta_1]$ and $[-\Delta_2, \Delta_2]$, respectively, and $\alpha = \Delta_1$, $\beta = \Delta_2$: (c) $\Delta_1 = \Delta_2 = 0.1$ and (d) $\Delta_1 = \Delta_2 = 0.891/5 = 0.1782$. . . 32
- 2.8 Estimation performance for a symmetric ULA with 9 antennas in Scenario 3 assuming $K = \infty$ and $\alpha = 0.0523$. The target is located at Δu with (a) $\Delta u = 0$, (b) $\Delta u = 0.01745$, (c) $\Delta u = -0.01745$, and (d) $\Delta u = 0.02618$ 33

2.9	Estimation performance for a symmetric URA with 5×5 antennas in Scenario 3 assuming $K = \infty$, $\alpha = 0.1$, $\beta = 0.1$, and target located at $(\Delta u, \Delta v)$, with (a) $(\Delta u, \Delta v) = (0, 0)$, (b) $(\Delta u, \Delta v) = (0.0349, 0)$, (c) $(\Delta u, \Delta v) = (-0.01745, 0)$, and (d) $(\Delta u, \Delta v) = (0.0349, -0.001218)$	37
2.10	Bias and variance for a URA with 5×5 elements in Scenario 3 assuming $K = \infty$ and $\alpha = \beta = 0.1$. Dashed black ellipses refer to CRLB limit, variances are shown with solid red ellipses whereas the bias vectors are represented by solid blue arrows.	39
2.11	Detection performance for the actual array manifold assuming the interfering environment of Scenario 3. Figs. (a) and (b) consider a symmetric ULA with 9 antennas, $K = 18$ and target location uniformly distributed over $[-\Delta, \Delta]$ with $\alpha = \Delta$: (a) $\Delta = 0.0523$ and (b) $\Delta = 0.891/N$. Figs. (c) and (d) correspond to a symmetric URA with 5×5 antennas, assuming $K = 75$ and target location offset Δu and Δv modeled as a pair of statistically independent uniformly distributed random variables over $[-\Delta_1, \Delta_1]$ and $[-\Delta_2, \Delta_2]$, respectively. Besides, $\alpha = \Delta_1$, $\beta = \Delta_2$: (c) $\Delta_1 = \Delta_2 = 0.1$ and (d) $\Delta_1 = \Delta_2 = 0.891/5 = 0.1782$	41
3.1	Convergence rate analysis for the case study discussed in the main text, with $N = 10$. Fig. (a) displays the average rate of convergence versus the number of snapshots, while Fig. (b) displays the average number of iterations versus the number of snapshots. The norm difference $\ \boldsymbol{\theta}^{(h)} - \hat{\boldsymbol{\theta}}_{ML}\ $ in dB versus the number of iterations for Algorithm 4 is reported in Fig. (c), assuming $K = 40, 60, 80, 100$	53

3.2	Adaptive beamformer performance for a ULA with 20 antennas in Scenario 1. Figs. (a) and (c) consider $p_m = 0.1$ while Figs. (b) and (d) consider $p_m = 0.3$. Figs. (a) and (b) display the normalized average S/I versus number of snapshots, while Figs. (c) and (d) display the resulting beam-pattern with $K = 60$ (therein, the red-Xs along the θ -axis denote the sources directions).	64
3.3	Adaptive beamformer performance for a ULA with 20 antennas in Scenario 2. Figs. (a) and (c) consider $p_m = 0.1$ while Figs. (b) and (d) consider $p_m = 0.3$. Figs. (a) and (b) display the normalized average S/I versus number of snapshots, while Figs. (c) and (d) display the resulting beam-pattern with $K = 60$ (therein, the red-Xs along the θ -axis denote the sources directions).	65
3.4	Normalized average S/I versus number of snapshots for a ULA with 20 antennas. Fig. (a) considers Scenario 1 while Fig. (b) Scenario 2.	68
3.5	Detection performance for a ULA with 20 antennas assuming $K = 100$ and $p_m \in \{0.1, 0.3\}$. Figs. (a), (b), and (c) assume $d = 2$, Figs. (d), (e), and (f) assume $d = 3$, whereas Figs. (g), (h), and (i) assume $d = 4$ equal-power signals impinging the array, respectively, with signal separation corresponding to $0.891/N$. Moreover, Figs. (a), (d), and (g) consider AIC, Figs. (b), (e), and (h) consider MDL, whereas Figs. (c), (f), and (i) consider HQC.	69

3.6	Comparison of the P_D using EM and EM-FB estimation strategies for a ULA with 20 antennas assuming 3 equal-power signals impinging the array with signal separation corresponding to $0.891/N$. Figs. (a), (b), (c) consider $p_m = 0.1$, whereas $p_m = 0.3$ is assumed in Figs. (d), (e), (f). Besides, Figs. (a) and (d), Figs. (b) and (e), and Figs. (c) and (f) consider AIC, MDL, and HQC, respectively.	70
4.1	Detection performance for an ULA with $N = 16$ antennas and unconstrained estimation. Different sample support sizes are considered, i.e., (a) $K = 48$ and (b) $K = 64$	86
4.2	Detection performance for a symmetric ULA with $N = 15$ antennas assuming the covariance matrix with a centro-Hermitian structure, see (3.6). Different sample support sizes are considered, i.e., (a) $K = 30$ and (b) $K = 45$	87
4.3	Detection performance for a ULA with $N = 16$ antennas assuming the uncertainty set in (3.2). Different sample support sizes are considered, i.e., (a) $K = 24$ and (b) $K = 48$	88
4.4	Detection performance for a ULA with $N = 4$ antennas and unconstrained estimation assuming $K = 40$ secondary data. Target locations are indicated by black dotted lines.	90
4.5	Detection performance for a ULA with $N = 4$ antennas assuming the uncertainty set in (3.2) and $K = 24$ secondary data. Target locations are indicated by black dotted lines.	91
5.1	Signal transmission and reception in FDA-MIMO radar.	95
5.2	Signal processing at the receiver with multiple match-filtered waveforms.	96

5.3	Comparison of RMSE (dB) assuming white noise for some u and δ : (a) and (d) $u = -0.891/(2(N+M))$, $\delta = -\Delta f/(4B)$, (b) and (e) $u = 0$, $\delta = \Delta f/(4B)$, (c) and (f) $u = 0.891/(2(N+M))$, $\delta = \Delta f/(2B)$. The RMSE analysis w.r.t. u is reported in (a), (b), and (c), whereas that w.r.t. δ is reported in (d), (e), and (f).	113
5.4	Bias and variance analysis assuming white noise for 15 different pairs (u, δ) and SINR = 20dB.	114
5.5	Comparison of RMSE (dB) for some u and δ : (a) and (d) $u = 0.891/(2(N+M))$, $\delta = -\Delta f/(2B)$, (b) and (e) $u = -0.891/(2(N+M))$, $\delta = -\Delta f/(4B)$, (c) and (f) $u = 0$, $\delta = 0$, assuming two coherent repeaters, with SINR = 30dB, located at $u_1 = 5(0.891/(N+M))$ and $u_2 = -6(0.891/(N+M))$ with incremental range of $\delta_1 = \Delta f/(3B)$ and $\delta_2 = -\Delta f/(8B)$, respectively. The RMSE analysis w.r.t. u is reported in (a), (b), and (c), whereas that w.r.t. δ is reported in (d), (e), and (f).	115
5.6	Bias and variance analysis for 15 different pairs (u, δ) assuming SINR = 20dB and two coherent repeaters, with SINR = 30dB, located at $u_1 = 5(0.891/(N+M))$ and $u_2 = -6(0.891/(N+M))$ with incremental range of $\delta_1 = \Delta f/(3B)$ and $\delta_2 = -\Delta f/(8B)$, respectively.	116
6.1	Illustration of the polarimetric FDA-MIMO radar system.	121
6.2	Signal processing chain at the n -th polarimetric receiver with multiple match-filters.	122
6.3	Detection performance for white noise, assuming $K = 320$ and the parameters of Table 6.1. Fig. (a) considers the GLRT-based detectors, whereas Fig. (b) reports the 2SGLRT-based counterparts.	138

6.4	Detection performance for a mixed clutter environment (see Table 6.2) with $\rho_c = 0.4$, assuming $K = 320$ and the parameters of Table 6.1. Fig. (a) considers the GLRT-based detectors, whereas Fig. (b) reports the 2SGLRT counterparts.	141
6.5	Detection performance for a mixed clutter environment (see Table 6.2) with $\rho_c = 0.9$, assuming $K = 320$ and the parameters of Table 6.1. Fig. (a) considers the GLRT-based detectors, whereas Fig. (b) reports the 2SGLRT counterparts.	141
6.6	Detection performance for a clutter trees environment (see Table 6.2) with $\rho_c = 0.9$, assuming $K = 320$ and the parameters of Table 6.1. Fig. (a) considers the GLRT-based detectors, whereas Fig. (b) reports the 2SGLRT counterparts.	142
6.7	Detection performance of the LAM, GPM, and CD methods for different sample support sizes, i.e., $K = 320$, $K = 240$, $K = 120$, assuming a mixed clutter environment (see Table 6.2) with $\rho_c = 0.9$ and the parameters of Table 6.1. Figs. (a), (c), and (d) report the performance of the GLRT-based decision statistics; Figs. (b), (d), and (f) that of the 2SGLRT rules.	146

List of Tables

- 2.1 Computational Complexity of the considered processors. . . 35
- 3.1 Simulation Parameters 67
- 4.1 Simulation Parameters. 84
- 4.2 Targets Range and Angle Locations. 89
- 5.1 Computational complexity of the considered estimators. . . 109
- 5.2 Simulation Parameters of FDA-MIMO Radar 111
- 6.1 Simulation Parameters of the FDA-MIMO Radar 138
- 6.2 Typical parameters for polarimetric clutter characterization [37]144

List of Symbols

The following symbols are used within the thesis

$(\cdot)^\dagger$	conjugate transpose operator;
$(\cdot)^T$	transpose operator;
$(\cdot)^*$	complex conjugate operator (component-wise complex conjugate if the argument is a matrix or a vector);
$\mathbf{0}$	matrix with zero entries;
$\mathbf{1}$	all-ones column vector;
$\mathbf{1}_N$	all-ones column vector of size N ;
\mathbf{A}	matrix;
\mathbf{a}	column vector;
$\mathbf{A}(k, l)$	(k, l) -entry of a generic matrix \mathbf{A} (or vector \mathbf{a});
\mathbf{e}_k	k -th column vector of \mathbf{I} , whose size is determined from the context;
\mathbf{I}	identity matrix;
$\det(\mathbf{A})$	determinant of the matrix $\mathbf{A} \in \mathbb{C}^{N \times N}$;
$\lambda_{\max}(\mathbf{A})$	maximum eigenvalue of $\mathbf{A} \in \mathbb{H}^N$;
$\lambda_{\min}(\mathbf{A})$	minimum eigenvalue of $\mathbf{A} \in \mathbb{H}^N$;
$[a, b]$	a closed interval of \mathbb{R} ;
$\mathbb{C}^{N \times M}$	set of $N \times M$ complex matrices;
\mathbb{R}^{++}	set of real numbers greater than zero;

\mathbb{R}^N	set of N -dimensional column vectors of real numbers;
\mathbb{T}^N	set of $N \times N$ Hermitian Toeplitz matrices;
$\max(x, y)$	maximum between $x, y \in \mathbb{R}$;
$\text{Rank}(\mathbf{A})$	rank of the matrix $\mathbf{A} \in \mathbb{C}^{N \times N}$;
$\nabla_{\mathbf{x}} f(\mathbf{x})$	gradient of $f(\mathbf{x}) \in \mathbb{R}$ with respect to \mathbf{x} , with the partial derivatives arranged in a column vector;
\odot	Hadamard (i.e., elementwise) product;
$\text{Im}\{x\}$	imaginary part of a complex number x ;
$\text{Re}\{x\}$	real part of a complex number x ;
\otimes	Kronecker product;
$\rho(\mathbf{A})$	spectral radius of \mathbf{A} , i.e., $\sqrt{\lambda_{\max}(\mathbf{A} \mathbf{A}^\dagger)}$;
\succeq	generalized matrix inequality: for any $\mathbf{A} \in \mathbb{H}^N$, $\mathbf{A} \succeq \mathbf{0}$ means that \mathbf{A} is a positive semi-definite matrix ($\mathbf{A} \succ \mathbf{0}$ for positive definiteness);
$\text{diag}(\mathbf{x})$	diagonal matrix whose i -th diagonal element is $\mathbf{x}(i)$;
$\text{tr}\{\mathbf{A}\}$	trace of the matrix $\mathbf{A} \in \mathbb{C}^{N \times N}$;
$v \sim U(-\Delta, \Delta)$	a random variable v uniformly distributed over $(-\Delta, \Delta)$, $\Delta > 0$;
C^2	class of functions with continuous second-order derivatives;
$C_L^{1,1}$	class of functions with Lipschitz gradient with constant L ;
$E[\cdot]$	statistical expectation;
j	imaginary unit (i.e., $j = \sqrt{-1}$);
\mathbb{C}^N	set of N -dimensional column vectors of complex numbers;
\mathbb{H}^N	set of $N \times N$ Hermitian matrices;
\mathbb{H}_{++}^N	set of $N \times N$ Hermitian positive definite matrices;
$\text{sign}(y)$	sign of $y \in \mathbb{R}$, i.e., -1 if $y < 0$ and 1 if $y \geq 0$;
H_0	target absence hypothesis;
H_1	target presence hypothesis;

Chapter 1

Introduction

Modern radar research and development efforts are de facto characterized by a number of paramount keystones: namely, the evolution of subsystems hardware capabilities, the development of complex algorithms for specific radar signal processing tasks, as well as the design of novel phased array geometries and architectures. In particular, a major milestone within this technological trend is represented by Active Electronically Scanned Array ([AESA](#)) antennas which eventually enabled the use of advanced techniques such as Digital Beamformer ([DBF](#)) and Space-Time Adaptive Processing ([STAP](#)) for radar operations in hostile environments, clutter mitigation, and the like.

Considering the current state of the art, multichannel phased array radar systems must comply with very stressing operational requirements demanding surveillance at specific ranges (sometimes with different update rates depending on the different elevation sectors); tracking with variable update rates (which can be different from those adopted for search tasks), Three-Dimensional ([3-D](#)) target data measurements; the capability to manage many tracks simultaneously without decreasing the search performance, and the need to operate in clutter and jamming environments [[95](#), [89](#), [100](#), [23](#)]. To this end, agility in beam steering and adaptive [DBF](#) are key ingredients to cope with such challenges jointly with advanced signal and data processing algorithms aimed at boosting radar performance [[45](#)], [[62](#)]. At the root of all the aforementioned processes, there is target detection and localization measurements within a [3-D](#) co-

ordinate system.

Target detection has been the subject of plenty of articles in the open literature, mainly devoted to the development of adaptive detectors (as well as to their analysis) capable of operating in the presence of undesired disturbance, hostile interference, and clutter [2, 60, 75, 32, 21, 93, 87, 92, 15, 90, 39, 41]. Generally, to accomplish the detection task, at the design stage the received signal (under the assumption of target presence and taking into account receiver linearity) is represented via the superposition of the target echo and an interference-plus-noise contribution, which is usually modeled at baseband as the realization of a zero-mean Gaussian process with an unknown and possibly structured covariance matrix. Besides, the standard homogeneous radar environment assumption is often invoked whereby a set of secondary (training) data, free of target contamination, enables the estimation of the unknown interference covariance matrix and the derivation of adaptive architectures [60, 93, 39, 75, 21, 87, 92].

Within this framework, the target detection problem is formulated in terms of a binary statistical hypothesis test, whose optimal solution (in the Neyman-Pearson sense) is given by the Likelihood Ratio Test (LRT) [105, 39, 59, 90]. However, the LRT requires perfect knowledge of the parameters characterizing the likelihood function under the H_0 (target absent) and H_1 (target present) hypotheses, including their parameters related to either the target characteristics or the interference covariance matrix. In practical situations, such parameters are unknown and need to be estimated. This results in the development of implementable receivers based on sub-optimal approaches, such as the Generalized Likelihood Ratio (GLR), where the unknowns are replaced by their Maximum Likelihood (ML) estimates [104, 59].

Once the presence of a target has been established within a Cell Under Test (CUT), the estimation process of the target bearing could be accomplished by means of monopulse [79], generalized monopulse [81], or other bespoke techniques, thus implementing detection and estimation as two different signal processing tasks [120, 119]. Target parameter estimation is an enduring signal processing problem that has always raised persistent attention within the radar scientific community. Indeed, new threats demand strict radar performance requirements as well as the development of advanced algorithms capable of providing reliable estimates of the target

position, even at the expense of an increased computational complexity. With reference to the Direction of Arrival (DOA), it is well known that phased array radars can achieve highly accurate angle estimation by precisely forming a beam in the desired direction [39, 105]. Nevertheless, it is worth stressing that in a classic phased array radar system the beampattern is dependent only on the angular direction [118, 76, 28, 98] whereas it is not selective in the range domain. As a consequence, the range information is not directly embedded in the beamforming process. To overcome this drawback, a novel beam scanning array, referred to as Frequency Diverse-Array (FDA), has been proposed [5, 108, 17]. Specifically, employing a small frequency increment across adjacent array elements, the FDA induces discrimination in both the angle and range domains. In other words, FDA allows to glean jointly range and angle information. Moreover, the mentioned extra flexibility of the FDA radar paves the way for its capitalization in several applications, such as cognitive target tracking [48] and target localization [34]. Furthermore, FDA can be successfully applied to a Multiple Input Multiple Output (MIMO) architecture leading to an FDA-MIMO radar [96], where additional degrees-of-freedom become available by separating the different transmitted waveforms with appropriate matched filtering. Therefore, a range-angle-dependent transmit-receive beampattern can be realized via the FDA-MIMO manifold structure. That is, capitalizing the additional Degree of Freedom (DOF)s of an FDA-MIMO radar in the range domain, the target angle and range can be simultaneously estimated due to the range-angle-dependent characteristic of the FDA-MIMO steering vector.

To this end, it is worth mentioning that in multi channel radar architectures most procedures have been designed under the ideal conditions that all data at the output of the array are available. However, regardless of the type of system being used, practical radar architectures are not devoid of flaws and issues. In particular, measurement errors due to acquisition equipment, random sensor failure [111] caused by impulsive noise [123], range ambiguous echo returns affecting useful signal samples [91], as well as reception failures (e.g., in distributed radar architecture [51, 46]), can determine the lack of some observations. Therefore, missing sensor measurements can arise in a variety of radar signal processing problems, for instance, DBF, DOA estimation, interference cancellation, covariance esti-

mation, and target detection. This implies developing specific procedures to limit such a loss with respect to an ideal scenario where all measurements are available.

Within this framework, this thesis presents contributions pertaining to joint target detection and estimation architectures, innovative strategies to recover from missing-data scenarios, and novel signal processing methods developed for **FDA-MIMO** radars (also exploiting the polarimetric domain). Accordingly, the rest of the dissertation is organized into several chapters as follows.

- **Chapter 2** deals with the problem of joint radar target detection and angle estimation. To this end, after a tailored linearization procedure, the target steering vector is represented as the superposition of the pointing direction term plus two contributions that account for the directional cosines offsets with respect to the nominal array steering. Based on this representation, target detection in the presence of interference is cast as a binary composite hypothesis testing problem with a different number of unknowns under the two alternatives. In particular, under the target presence hypothesis, the likelihood function exhibits dependence on the complex target amplitude, the interference covariance matrix, and the two-direction cosine displacement parameters (accounting for the unknown target location within the array mainbeam). This formulation paves the way for the development of signal processing architectures that are able to detect the target and, at the same time, determine its angular estimate. In this respect, the GLR Test (**GLRT**) criterion [39], [59] is considered, which under the target presence hypothesis, after concentration over the unknown interference covariance and target complex amplitude, demands the solution of an optimization problem over the unknown direction offsets. For the special case of a One-Dimensional (**1-D**) linear array, the problem is solved in closed form by simply computing the roots of a second-order equation. For the Two-Dimensional (**2-D**) planar array, the optimization becomes more challenging and falls within the class of box-constrained fractional quadratic problems. In this regard, two new solution methods are proposed. The former provides the optimal solution and relies on the use of the Dinkelbach's algorithm [99, 33, 16]. The latter
-

is an iterative fast-converging procedure based on a Coordinate Descent (CD) optimization [29]. However, in this last case, only convergence to a stationary point can be claimed.

- **Chapter 3** presents the problem of estimating a structured covariance matrix under the missing-data context. First of all, the homogeneous Gaussian environment observation model with missing-data is introduced capitalizing on a-priori information about the covariance matrix structure and/or specific array configurations. Then, the covariance matrix estimation process is formulated as an appropriately constrained optimization problem that is in general difficult to solve. Hence, an effective iterative solution technique, based on the Expectation-Maximization (EM) algorithm [42, 105, 56, 114], is introduced together with some convergence properties and rate of convergence results. Each iteration of the algorithm (for very common covariance structures of practical interest in radar signal processing applications) involves only closed-form solutions for the unknowns. As per the radar context, the developed theory is contextualized for a beamforming application and for the problem of detecting the number of sources. The study leads to some efficient methods capable of operating in the presence of missing-data with satisfactory performance.
 - In **Chapter 4**, leveraging the results proposed in **Chapter 3**, the problem of target detection in Gaussian interference with an unknown (but possibly structured) covariance matrix is addressed for a missing-data context. The detection problem is formulated as a composite hypothesis test characterized by different unknowns under the two (i.e., H_0 , H_1) hypotheses. In particular, a sub-optimal GLRT architecture is designed as a feasible detector. Besides, a variation of the conventional GLRT, i.e., the Adaptive Matched Filter (AMF) [93] test (also known as the Two-Step GLRT (2SGLRT)), is derived as well. The 2SGLRT requires the ML Estimate (MLE) of the covariance matrix under the H_0 hypothesis (to be computed from secondary data according to the method derived in **Chapter 3**) and the MLE of the complex target echo amplitude (assuming that the interference covariance matrix is known). However, the alterna-
-

tive hypothesis (with reference to the **GLRT** receiver) involves the joint **MLE** of the complex target echo amplitude and the covariance matrix. To handle this challenging task, an **EM**-based framework is proposed to determine optimized solutions to the maximization problem at hand with some quality guarantees. The resulting procedure involves only closed-form expressions for the cases of unstructured estimation and covariance matrix with a centro-Hermitian structure. For the case of the covariance matrix with a more general structure, the plain **EM** leads to an optimization problem where closed-form solutions are not available. Therefore, this case is addressed using two **EM** variations, i.e., the Expectation-Conditional Maximization (**ECM**) and Multi-cycle EM (**M-EM**) techniques, requiring the optimization of individual subsets of the unknowns at a given iteration [77, 73, 97], each of them updated in closed-form for some covariance structures of practical interest.

- **Chapter 5** investigates the problem of joint target angle and incremental range estimation using an **FDA-MIMO** radar in a background of Gaussian interference with known spectral properties. At the design stage, the target parameters, i.e., angle, incremental range, and echo-amplitude, are assumed unknown. Under the aforementioned setup, the estimation problem is formulated starting from a single data snapshot. Then, the **ML** estimator is derived maximizing the likelihood function with respect to the unknown parameters. Furthermore, to reduce the computational cost connected with the **2-D** grid search required by the implementation of the **ML** rule, three approximated methods are considered:
 - an iterative procedure based on the **CD** algorithm leveraging a sequence of **1-D** searches which arise alternating optimizations over one specific variable while keeping the other one fixed;
 - an adaptive monopulse approach that approximates the optimal search exploiting real bias/slope correction values;
 - a generalized monopulse procedure employing a complex slope and bias correction aimed at minimizing the mean square value of the noise error term.
-

At the analysis stage, Cramér-Rao Lower Bound (CRLB) for the angle and incremental range estimation via an FDA-MIMO radar is derived.

- Starting from the framework proposed in **Chapter 5**, **Chapter 6** deals with the target detection problem with a polarimetric FDA-MIMO radar. In this context, the synthesis of adaptive detectors is investigated taking into account a disturbance covariance matrix with unknown spectral characteristics. At the design stage, the target detection problem is formulated as a binary hypothesis test where the target parameters (i.e., angle, incremental range within the radar cell [66, 67], and scattering matrix) along with the interference covariance matrix are assumed unknown. Considering feasible detectors (derived resorting to the GLRT and 2SGLRT criteria), the problems at hand are tackled in a sub-optimal but effective way by means of three specific techniques:
 - the Linearized Array Manifold (LAM) method (proposed in **Chapter 2**), which solves an equivalent optimization problem leveraging a tailored linearization procedure. This entails representing the target steering vector as the superposition of the pointing direction term plus two contributions that account for the actual angle and incremental range offsets with respect to the nominal array steering;
 - the Gradient Projection Method (GPM) [18], which iteratively updates the unknown target location parameters, in the concentrated log-likelihood function under the alternative hypothesis, along the current ascent direction. This is accomplished by using either a constant stepsize or an adaptive one via the backtracking procedure from an initial guess until a stationary point is reached;
 - the CD procedure [19], composed of an alternating sequence of 1-D searches (conducted over discretized intervals) with respect to one target location parameter (either angle or incremental range) while keeping the other fixed.

Finally, the bounded Constant False Alarm Rate (CFAR) property as

well as the computational complexity of the aforementioned [GLRT](#)-based detectors are discussed.

- **Chapter 7** summarizes the advancements described in this thesis work and highlights emerging trends and issues to be addressed in the immediate future.
-

Chapter 2

Single-Pulse Simultaneous Target Detection and Angle Estimation in a Multichannel Phased Array Radar

This chapter¹ is focused on simultaneous target detection and angle estimation with a multichannel phased array radar. Resorting to a linearized expression for the array steering vector around the beam pointing direction, the problem is formulated as a composite binary hypothesis test where the unknowns, under the alternative hypothesis, include the target directional cosines displacements with respect to the array nominal coarse pointing direction. The problem is handled via the [GLR](#) criterion (both one-step and two-step) where decision statistics leveraging the [MLEs](#) of the parameters are compared with a detection threshold. If crossed, target presence is declared and the [MLEs](#) of the aforementioned displacements directly provide target angular position with respect to the pointing direction. From the analytic point of view, [ML](#) estimation involves a constrained fractional quadratic optimization problem whose optimal solution

¹©2020 IEEE. Reprinted, with permission, from A. Aubry, A. De Maio, S. Marano and M. Rosamilia, "Single-Pulse Simultaneous Target Detection and Angle Estimation in a Multichannel Phased Array Radar," *IEEE Transactions on Signal Processing*, vol. 68, pp. 6649-6664, 2020.

can be found via the Dinkelbach's algorithm or approximated through a fast-converging procedure based on a CD optimization. The performance analysis of the proposed architectures as well as the corresponding discussion is developed in terms of computational complexity, CFAR behavior, detection performance, and angular estimation accuracy, also in comparison with some counterparts available in the open literature and theoretical benchmark limits.

This chapter is organized as follows. Section 2.1 deals with problem formulation. In Section 2.2, three procedures to solve the constrained optimization problem are developed. Section 2.3 addresses the performance analysis and outlines comparisons for both detection and estimation tasks. Finally, Section 2.4, summarizes conclusions and outlines possible future research avenues.

2.1 Problem Formulation

Let us consider a radar system that collects spatial data via a narrow-band planar array (see Fig. 2.1) composed of NM antennas. After down-conversion, pulse compression, and fast-time sampling, the echo signal induced by a prospective target located at range R , azimuth θ_0 , and elevation ϕ_0 with respect to array boresight, is given by

$$a\mathbf{p}(u_0, v_0), \quad (2.1)$$

where a is an unknown complex parameter accounting for target backscattering and channel propagation effects, (u_0, v_0) indicates the target angular location in the space of directional cosines [105], i.e.,

$$u_0 = \sin(\theta_0) \cos(\phi_0), \quad v_0 = \sin(\theta_0) \sin(\phi_0), \quad (2.2)$$

and $\mathbf{p}(u_0, v_0)$ denotes the spatial steering vector $\mathbf{p}(u, v)$ evaluated at (u_0, v_0) . For a Uniform Rectangular Array (URA)

$$\mathbf{p}(u, v) = \mathbf{p}_u(u) \otimes \mathbf{p}_v(v), \quad (2.3)$$

with

$$\mathbf{p}_u(u) = [e^{j\frac{2\pi}{\lambda_0}x_0u}, e^{j\frac{2\pi}{\lambda_0}x_1u}, \dots, e^{j\frac{2\pi}{\lambda_0}x_{N-1}u}]^T \in \mathbb{C}^N, \quad (2.4)$$

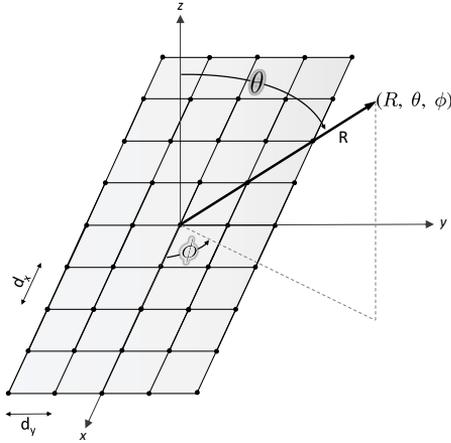


Figure 2.1. Two dimensional symmetric planar array geometry with equi-spaced antennas in the x and y directions. (R, θ, ϕ) are the polar coordinates of target.

$$\mathbf{p}_v(v) = [e^{j\frac{2\pi}{\lambda_0}y_0v}, e^{j\frac{2\pi}{\lambda_0}y_1v}, \dots, e^{j\frac{2\pi}{\lambda_0}y_{M-1}v}]^T \in \mathbb{C}^M, \quad (2.5)$$

where (x_i, y_h) , $i = 0, \dots, N-1$, $h = 0, \dots, M-1$, are the positions² of the array elements and λ_0 is the radar operating wavelength.

Assuming the potential useful echo signal buried in Gaussian interference with unknown spectral characteristics and supposing the availability of $K \geq N$ homogeneous secondary data (i.e., data vectors, free of useful target returns, exhibiting the same spectral property as that from the cell under test), the problem of detecting a target located at (θ_0, ϕ_0) , i.e., at

²For a **URA**, denoting by d_x and d_y the interelement spacing (usually given by $\lambda_0/2$) along the x and y axes, respectively, if the reference system center is located in the bottom-left corner

$$x_i = d_x i, \quad i = 0, 1, \dots, N-1, \quad y_h = d_y h, \quad h = 0, 1, \dots, M-1.$$

If instead the reference system center coincides with the array center, then

$$x_i = d_x \left(i - \left(\frac{N-1}{2} \right) \right), \quad i = 0, 1, \dots, N-1,$$

$$y_h = d_y \left(h - \left(\frac{M-1}{2} \right) \right), \quad h = 0, 1, \dots, M-1.$$

(u_0, v_0) in $u-v$ space, can be formulated as the following composite binary hypothesis testing problem

$$\begin{cases} H_0 : \begin{cases} \mathbf{r} = \mathbf{n} \\ \mathbf{r}_k = \mathbf{n}_k \quad k = 1, \dots, K \end{cases} \\ H_1 : \begin{cases} \mathbf{r} = a\mathbf{p}(u_0, v_0) + \mathbf{n} \\ \mathbf{r}_k = \mathbf{n}_k \quad k = 1, \dots, K \end{cases} \end{cases}, \quad (2.6)$$

where the interference plus noise components \mathbf{n} and \mathbf{n}_k , $k = 1, \dots, K$, are modeled as statistically independent, complex, zero-mean, circularly symmetric Gaussian random vectors with unknown positive definite covariance matrix

$$\mathbf{M} = E[\mathbf{n}\mathbf{n}^\dagger] = E[\mathbf{n}_k\mathbf{n}_k^\dagger], \quad k = 1, \dots, K. \quad (2.7)$$

Two important remarks are now necessary.

1. The presented framework is developed using a [URA](#), but it can be easily generalized to deal with different types of array. Besides, for the special case of a linear array, i.e., $M = 1$, the steering vector [\(2.3\)](#) degenerates into [\(2.4\)](#).
2. The interference scenario with the assumptions in [\(2.7\)](#) defines the so called “homogeneous environment”, well established and accepted in radar detection-estimation literature being theoretically justified and representative of many practical operative contexts [[60](#), [93](#), [14](#), [101](#)] and also explained in many technical books [[112](#), [45](#), [39](#)].

2.1.1 Pointing Errors

An array steering direction is not usually aligned with the target [DOA](#), especially when the radar is in search mode. In order to account for this mismatch, a specific model of the array steering vector is now developed, leveraging a linearization of the array manifold around the transmit look-direction (\bar{u}, \bar{v}) . Specifically, denoting by $(\Delta u, \Delta v)$ the directional cosine offset, i.e., $\Delta u = u_0 - \bar{u}$, $\Delta v = v_0 - \bar{v}$, the target steering vector is

approximated as

$$\mathbf{p}_a(\Delta u, \Delta v) = \mathbf{p}(\bar{u}, \bar{v}) + \frac{\partial \mathbf{p}(\bar{u}, \bar{v})}{\partial u} \Delta u + \frac{\partial \mathbf{p}(\bar{u}, \bar{v})}{\partial v} \Delta v, \quad (2.8)$$

where the explicit dependence of the approximated steering vector on the pointing direction is omitted to avoid unnecessary notational complications. Note that

$$\frac{\partial \mathbf{p}(\bar{u}, \bar{v})}{\partial u} = \frac{\partial \mathbf{p}(u, v)}{\partial u} \Bigg|_{\substack{u=\bar{u} \\ v=\bar{v}}} = j \frac{2\pi}{\lambda_0} (\mathbf{p}_u(\bar{u}) \odot [x_0, x_1, \dots, x_{N-1}]^T) \otimes \mathbf{p}_v(\bar{v}) \quad (2.9)$$

and

$$\frac{\partial \mathbf{p}(\bar{u}, \bar{v})}{\partial v} = \frac{\partial \mathbf{p}(u, v)}{\partial v} \Bigg|_{\substack{u=\bar{u} \\ v=\bar{v}}} = j \frac{2\pi}{\lambda_0} \mathbf{p}_u(\bar{u}) \otimes (\mathbf{p}_v(\bar{v}) \odot [y_0, y_1, \dots, y_{M-1}]^T) \quad (2.10)$$

represent the partial derivatives with respect to u and v .

To assess the accuracy of the approximation in (2.8), Fig. 2.2 reports the magnitude of the normalized correlation (mismatched angle cosine) between the actual steering vector and the approximated one, i.e.,

$$\frac{|\mathbf{p}_a(\Delta u, \Delta v)^\dagger \mathbf{p}(\bar{u} + \Delta u, \bar{v} + \Delta v)|}{\|\mathbf{p}_a(\Delta u, \Delta v)\| \|\mathbf{p}(\bar{u} + \Delta u, \bar{v} + \Delta v)\|}, \quad (2.11)$$

versus the directional cosines offsets. Specifically, assuming symmetric array configurations, in Fig. 2.2(a) the 1-D case is analyzed, with $N = 9$ and $\bar{u} = 0$, whereas Fig. 2.2(b) refers to the 2-D scenario with $M = 5, N = 5, \bar{u} = \bar{v} = 0$. The results clearly highlight the ability of $\mathbf{p}_a(\Delta u, \Delta v)$ to describe accurately the actual steering vector as long as the target DOA lies within the 3 dB beamwidth. Indeed, normalized correlation values higher than 0.83 are achieved³ if $|\Delta u| \leq 0.891/N$ and $|\Delta v| \leq 0.891/M$ where $0.891/N$ and $0.891/M$ represent the 3 dB single-side beamwidth of a planar array pointing at the boresight direction.

Hereafter, to simplify notation, the nominal steering vector $\mathbf{p}(\bar{u}, \bar{v})$ is indicated as \mathbf{p} whereas the steering derivatives (at the pointing directions)

³Normalized correlation values larger than or equal to 0.95 are achieved, in 1-D case provided that $|\Delta u| \leq 0.891/N$.

$\partial \mathbf{p}(\bar{u}, \bar{v})/\partial u$ and $\partial \mathbf{p}(\bar{u}, \bar{v})/\partial v$ are denoted by $\dot{\mathbf{p}}_u$ and $\dot{\mathbf{p}}_v$, respectively. As a result, equation (2.8) can be re-written as

$$\mathbf{p}_a(\Delta u, \Delta v) = \mathbf{p} + \mathbf{H} \Delta \boldsymbol{\theta}, \quad (2.12)$$

with $\mathbf{H} = [\dot{\mathbf{p}}_u, \dot{\mathbf{p}}_v]$ the Jacobian matrix, and $\Delta \boldsymbol{\theta} = [\Delta u, \Delta v]^T \in \mathbb{R}^2$. Note that, for the special case of a linear array, the steering model boils down to

$$\mathbf{p}_a(\Delta u) = \mathbf{p}_u(\bar{u}) + \frac{\partial \mathbf{p}(\bar{u})}{\partial u} \Delta u = \mathbf{p} + \dot{\mathbf{p}}_u \Delta u. \quad (2.13)$$

Now, leveraging the useful signal model (2.12), the target detection problem in the presence of pointing errors can be cast as

$$\begin{cases} H_0 : \begin{cases} \mathbf{r} = \mathbf{n} \\ \mathbf{r}_k = \mathbf{n}_k \quad k = 1, \dots, K \end{cases} \\ H_1 : \begin{cases} \mathbf{r} = a \mathbf{p}_a(\Delta u, \Delta v) + \mathbf{n} \\ \mathbf{r}_k = \mathbf{n}_k \quad k = 1, \dots, K \end{cases} \end{cases}, \quad (2.14)$$

where $\mathbf{p}_a(\Delta u, \Delta v)$ represents the approximated steering vector and the entries of $(\Delta u, \Delta v)$ indicate the unknown target direction cosines, with $|\Delta u| \leq \alpha$, $|\Delta v| \leq \beta$. The choice of the constraint levels α and β must reflect a compromise between DOA uncertainty and quality of the linear approximation. A reasonable choice is the 3 dB single-side beamwidth.

Notably, by capitalizing the a-priori information on the sensing system, namely the transmit beamwidth size, the feasible values of the unknowns DOA displacements can be appropriately constrained, laying the ground for an improved angular estimation.

2.2 System Design: Joint Detection and Angle Estimation

The optimum solution to the hypotheses testing problem (2.14) (in the Neyman-Pearson sense) is the LRT. However, its practical implementation is precluded as the knowledge of the parameters Δu , Δv , a and \mathbf{M} is required. In detection theory jargon this means that a Uniformly Most

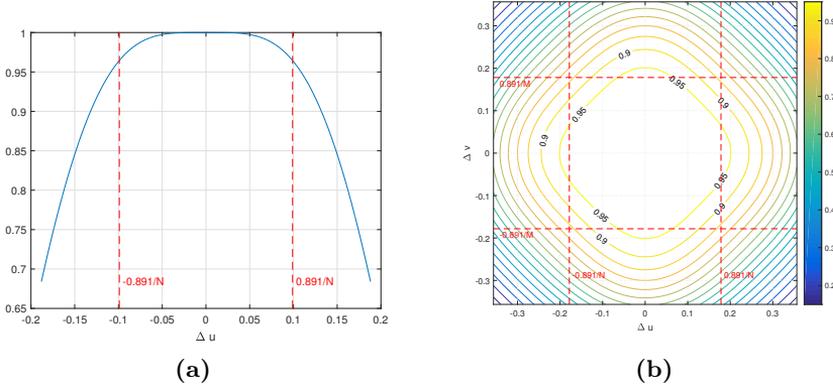


Figure 2.2. Normalized correlation versus direction cosine offset, assuming symmetric array: (a) 1-D case with $N = 9$ and $\bar{u} = 0$; (b) 2-D case with $N = M = 5$ and $\bar{u} = \bar{v} = 0$. Dashed red lines delimitate the 3 dB beamwidth.

Powerful (UMP) test does not exist. Thus, to come up with a practically implementable receiver, the GLR criterion is exploited, where the unknown parameters are replaced by their ML estimates under each hypothesis. Specifically, the following constrained GLRT decision rule is proposed

$$\frac{\max_{\substack{M, a, \\ |\Delta u| \leq \alpha, |\Delta v| \leq \beta}} f_{H_1}(\mathbf{r}, \mathbf{r}_1, \dots, \mathbf{r}_k | M, a, \Delta \theta)}{\max_M f_{H_0}(\mathbf{r}, \mathbf{r}_1, \dots, \mathbf{r}_k | M)} \underset{H_0}{\overset{H_1}{\geq}} T, \quad (2.15)$$

where $f_{H_1}(\cdot)$ and $f_{H_0}(\cdot)$ represent the probability density functions of the observations under the two hypotheses, and T is the detection threshold⁴ set to ensure a desired False Alarm Probability (P_{fa}). Besides, the AMF version of (2.15) (also known as two-step GLRT because it can be obtained computing the GLRT from the primary data vector (step 1) and then substituting the ML estimate of the interference covariance matrix in place

⁴With a slight abuse of notation, the same symbol is used to denote the detection threshold and its possible modifications introduced later, see, e.g., (2.18).

of the exact one (step 2)) is considered

$$\frac{\max_{\substack{a, \\ |\Delta u| \leq \alpha, |\Delta v| \leq \beta}} f_{H_1}(\mathbf{r} | \mathbf{M} = K^{-1} \mathbf{S}, a, \Delta \boldsymbol{\theta})}{f_{H_0}(\mathbf{r} | \mathbf{M} = K^{-1} \mathbf{S})} \underset{H_0}{\overset{H_1}{\geq}} T_1, \quad (2.16)$$

where

$$\mathbf{S} = \sum_{k=1}^K \mathbf{r}_k \mathbf{r}_k^\dagger \quad (2.17)$$

is proportional, via K , to the conventional secondary data sample covariance matrix. The motivation for considering both (2.15) and (2.16) stems from the observation that none of them can be a-priori claimed to be better than the other.

Following the same line of reasoning as in [60] and [93], it is not difficult to show that (2.15) and (2.16) can be cast respectively as

$$\max_{\substack{|\Delta u| \leq \alpha \\ |\Delta v| \leq \beta}} \frac{1}{1 + \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{r}} \frac{|\mathbf{r}^\dagger \mathbf{S}^{-1} (\mathbf{p} + \mathbf{H} \Delta \boldsymbol{\theta})|^2}{(\mathbf{p} + \mathbf{H} \Delta \boldsymbol{\theta})^\dagger \mathbf{S}^{-1} (\mathbf{p} + \mathbf{H} \Delta \boldsymbol{\theta})} \underset{H_0}{\overset{H_1}{\geq}} T, \quad (2.18)$$

and

$$\max_{\substack{|\Delta u| \leq \alpha \\ |\Delta v| \leq \beta}} \frac{|\mathbf{r}^\dagger \mathbf{S}^{-1} (\mathbf{p} + \mathbf{H} \Delta \boldsymbol{\theta})|^2}{(\mathbf{p} + \mathbf{H} \Delta \boldsymbol{\theta})^\dagger \mathbf{S}^{-1} (\mathbf{p} + \mathbf{H} \Delta \boldsymbol{\theta})} \underset{H_0}{\overset{H_1}{\geq}} T_1. \quad (2.19)$$

Remarkably, once the presence of a target is declared, i.e., the decision statistic on left hand side of (2.18) exceeds the detection threshold T , its angular estimate is obtained as a by-product from the decision statistic computation (2.18) and (2.19). Note that the evaluation of the decision rule in (2.18) as well as in (2.19) involves a non-convex fractional quadratic optimization problem. To handle it, different solution techniques are now devised, which represent the main technical contribution of this work from an optimization theory point of view. The 1-D case is studied in Subsection 2.2.1 whereas the 2-D case is analyzed in Subsection 2.2.2. Before proceeding further, it is worth observing that the decision statistic in (2.18)

is upper bounded by

$$\begin{aligned}\hat{t}_{upper}(\mathbf{r}, \mathbf{S}) &= \max_{\Delta\boldsymbol{\theta} \in \mathbb{C}^2} \frac{1}{1 + \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{r}} \frac{|\mathbf{r}^\dagger \mathbf{S}^{-1} (\mathbf{p} + \mathbf{H} \Delta\boldsymbol{\theta})|^2}{|\mathbf{S}^{-1/2} (\mathbf{p} + \mathbf{H} \Delta\boldsymbol{\theta})|^2} = \\ &= \frac{\mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{H}_{SD} \left(\mathbf{H}_{SD}^\dagger \mathbf{S}^{-1} \mathbf{H}_{SD} \right)^{-1} \mathbf{H}_{SD}^\dagger \mathbf{S}^{-1} \mathbf{r}}{1 + \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{r}},\end{aligned}$$

with $\mathbf{H}_{SD} = [\mathbf{p}, \mathbf{H}]$, namely,

$$\max_{\substack{|\Delta u| \leq \alpha \\ |\Delta v| \leq \beta}} \frac{1}{1 + \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{r}} \frac{|\mathbf{r}^\dagger \mathbf{S}^{-1} (\mathbf{p} + \mathbf{H} \Delta\boldsymbol{\theta})|^2}{|\mathbf{S}^{-1/2} (\mathbf{p} + \mathbf{H} \Delta\boldsymbol{\theta})|^2} < \hat{t}_{upper}(\mathbf{r}, \mathbf{S}). \quad (2.20)$$

Being the probability density function of $\hat{t}_{upper}(\mathbf{r}, \mathbf{S})$, under the H_0 hypothesis, functionally independent of \mathbf{M} , it follows that the decision rule in (2.18) ensures the bounded CFAR property. Indeed, for any given upper bound to the desired false alarm probability, a universal threshold, namely, just depending on the system parameters (i.e., pointing direction, number of antennas, and sample support size), can be set in (2.18) to fulfill the upper bound constraint. Leveraging (2.20), this property holds true even if a sub-optimal maximization is performed in (2.18) and thus an approximated implementation of the decision statistic in (2.18) is considered. Finally, following the same line of reasoning as before, it is not difficult to show that the two-step detector (2.19) and its possible sub-optimal implementations, i.e., relying on sub-optimal solution techniques to handle the involved maximization problem, still exhibit the bounded CFAR property.

2.2.1 Constrained GLRT Detector for 1-D Scenario

Handling the optimization problem involved in (2.18) and (2.19) for the 1-D case is tantamount to solving

$$\max_{|\Delta u| \leq \alpha} \frac{(\mathbf{p} + \dot{\mathbf{p}}_u \Delta u)^\dagger \mathbf{S}^{-1} \mathbf{r} \mathbf{r}^\dagger \mathbf{S}^{-1} (\mathbf{p} + \dot{\mathbf{p}}_u \Delta u)}{(\mathbf{p} + \dot{\mathbf{p}}_u \Delta u)^\dagger \mathbf{S}^{-1} (\mathbf{p} + \dot{\mathbf{p}}_u \Delta u)}. \quad (2.21)$$

In order to proceed, let us define the “whitened” quantities

$$\bar{\mathbf{p}} = \mathbf{S}^{-1/2} \mathbf{p}, \quad \dot{\bar{\mathbf{p}}}_u = \mathbf{S}^{-1/2} \dot{\mathbf{p}}_u, \quad \bar{\mathbf{r}} = \mathbf{S}^{-1/2} \mathbf{r}, \quad (2.22)$$

and recast Problem (2.21) as

$$\max_{|\Delta u| \leq \alpha} \frac{(\bar{\mathbf{p}} + \dot{\mathbf{p}}_u \Delta u)^\dagger \bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger (\bar{\mathbf{p}} + \dot{\mathbf{p}}_u \Delta u)}{(\bar{\mathbf{p}} + \dot{\mathbf{p}}_u \Delta u)^\dagger (\bar{\mathbf{p}} + \dot{\mathbf{p}}_u \Delta u)}. \quad (2.23)$$

The following proposition establishes a procedure to obtain a closed-form optimal solution to (2.23).

Proposition 1. *An optimal solution Δu^* to (2.23) is*

$$\Delta u^* = \arg \max_{\Delta u \in \{\Delta u_1, \Delta u_2\} \cup \mathcal{B}} \frac{(\bar{\mathbf{p}} + \dot{\mathbf{p}}_u \Delta u)^\dagger \bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger (\bar{\mathbf{p}} + \dot{\mathbf{p}}_u \Delta u)}{(\bar{\mathbf{p}} + \dot{\mathbf{p}}_u \Delta u)^\dagger (\bar{\mathbf{p}} + \dot{\mathbf{p}}_u \Delta u)}, \quad (2.24)$$

where $\Delta u_i = (-1)^i \alpha$, $i = 1, 2$, and \mathcal{B} is the finite set (whose cardinality is at most 2) containing the real roots (with absolute value less than α) of the quadratic equation

$$a' \Delta u^2 + b' \Delta u + c' = 0, \quad (2.25)$$

with

$$a' = \left| \bar{\mathbf{r}}^\dagger \dot{\mathbf{p}}_u \right|^2 \operatorname{Re}\{\bar{\mathbf{p}}^\dagger \dot{\mathbf{p}}_u\} - \|\dot{\mathbf{p}}_u\|^2 \operatorname{Re}\{\bar{\mathbf{p}}^\dagger \bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger \dot{\mathbf{p}}_u\}, \quad (2.26)$$

$$b' = \|\bar{\mathbf{p}}\|^2 \left| \bar{\mathbf{r}}^\dagger \dot{\mathbf{p}}_u \right|^2 - \|\dot{\mathbf{p}}_u\|^2 \left| \bar{\mathbf{r}}^\dagger \bar{\mathbf{p}} \right|^2, \quad (2.27)$$

$$c' = \|\bar{\mathbf{p}}\|^2 \operatorname{Re}\{\bar{\mathbf{p}}^\dagger \bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger \dot{\mathbf{p}}_u\} - \left| \bar{\mathbf{r}}^\dagger \bar{\mathbf{p}} \right|^2 \operatorname{Re}\{\bar{\mathbf{p}}^\dagger \dot{\mathbf{p}}_u\}. \quad (2.28)$$

Proof. See Appendix A.2. □

Exploiting the above results, it follows that the decision rule (2.18) for the 1-D case, referred to in the following as **GLRT-LAM**, can be expressed in closed form as

$$t_{\text{GLRT-LAM}} = \frac{1}{1 + \|\bar{\mathbf{r}}\|^2} \frac{\left| \bar{\mathbf{r}}^\dagger (\bar{\mathbf{p}} + \dot{\mathbf{p}}_u \widehat{\Delta u}^*) \right|^2}{\left\| \bar{\mathbf{p}} + \dot{\mathbf{p}}_u \widehat{\Delta u}^* \right\|^2} \underset{H_0}{\overset{H_1}{\geq}} T, \quad (2.29)$$

where $\widehat{\Delta u}^*$ is given by Proposition 1 (therein denoted by Δu^*) and represents the output estimate of the target **DOA** displacement, provided

that a detection is declared. The computational complexity required to implement the **GLRT-LAM** is $\mathcal{O}(KN^2)$, namely it is dominated by SCM evaluation. Finally, the **AMF** counterpart to (2.29), denoted as **GLRT-LAM-AMF**, is given by

$$t_{\text{GLRT-LAM-AMF}} = \frac{\left| \bar{\mathbf{r}}^\dagger (\bar{\mathbf{p}} + \dot{\hat{\mathbf{p}}}_u \hat{\Delta} u^*) \right|^2}{\left\| \bar{\mathbf{p}} + \dot{\hat{\mathbf{p}}}_u \hat{\Delta} u^* \right\|^2} \underset{H_0}{\overset{H_1}{\geq}} T_1, \quad (2.30)$$

2.2.2 Constrained GLRT Detector for 2-D Scenario

For the **2-D** array sensing scenario, the optimization problem to solve boils down to

$$\max_{\substack{|\Delta u| \leq \alpha \\ |\Delta v| \leq \beta}} \frac{(\mathbf{p} + \mathbf{H}\Delta\boldsymbol{\theta})^\dagger \mathbf{S}^{-1} \mathbf{r} \mathbf{r}^\dagger \mathbf{S}^{-1} (\mathbf{p} + \mathbf{H}\Delta\boldsymbol{\theta})}{(\mathbf{p} + \mathbf{H}\Delta\boldsymbol{\theta})^\dagger \mathbf{S}^{-1} (\mathbf{p} + \mathbf{H}\Delta\boldsymbol{\theta})}, \quad (2.31)$$

which can be equivalently expressed as

$$\max_{\substack{|\Delta u| \leq \alpha \\ |\Delta v| \leq \beta}} \frac{(\bar{\mathbf{p}} + \bar{\mathbf{H}}\Delta\boldsymbol{\theta})^\dagger \bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger (\bar{\mathbf{p}} + \bar{\mathbf{H}}\Delta\boldsymbol{\theta})}{(\bar{\mathbf{p}} + \bar{\mathbf{H}}\Delta\boldsymbol{\theta})^\dagger (\bar{\mathbf{p}} + \bar{\mathbf{H}}\Delta\boldsymbol{\theta})}, \quad (2.32)$$

where $\bar{\mathbf{p}}$ and $\bar{\mathbf{r}}$ are defined in (2.22), whereas $\bar{\mathbf{H}}$ is given by

$$\bar{\mathbf{H}} = \mathbf{S}^{-1/2} \mathbf{H}. \quad (2.33)$$

In the following, two optimization procedures are considered to handle Problem (2.32), which allow to localize the target within the antenna beam if its presence is declared. The former reaches the global optimum via the Dinkelbach algorithm [99]. The latter relies on a **CD** method [19] and converges to a stationary point without any theoretical guarantee to end up in a global maximizer of Problem (2.32). The second possibly sub-optimal approach exhibits in general a faster convergence than the Dinkelbach-based procedure, which can be a valuable feature from a practical point of view.

Dinkelbach-based DOA displacements estimate

To obtain the global optimal solution to Problem (2.32), some results from the fractional programming theory [99], [33], [16] are exploited, which are summarized here in the form of a lemma.

Lemma 2.2.1. [99] *Consider the fractional programming problem*

$$\max_{\mathbf{x} \in \mathcal{S}} Q(\mathbf{x}) = N(\mathbf{x})/D(\mathbf{x}), \quad (2.34)$$

where $\mathcal{S} \subseteq \mathbb{R}^n$ is a nonempty and compact set and $N(\mathbf{x}), D(\mathbf{x}) : \mathcal{S} \rightarrow \mathbb{R}$ are continuous functions, with $D(\mathbf{x})$ strictly positive. Then \mathbf{x}^* is an optimal solution to (2.34) if and only if it maximizes

$$N(\mathbf{x}) - Q(\mathbf{x}^*)D(\mathbf{x}). \quad (2.35)$$

Furthermore, the function

$$F(q) = \max_{\mathbf{x} \in \mathcal{S}} N(\mathbf{x}) - qD(\mathbf{x}), \quad q \in \mathbb{R}. \quad (2.36)$$

is continuous, convex, and strictly decreasing on \mathbb{R} with $F(q) > 0$ if $q < q^* = Q(\mathbf{x}^*)$ and $F(q) < 0$ if $q > q^*$. ■

Based on Lemma 2.2.1, an optimal solution to (2.34) can be found determining the unique root of (2.36), possibly via the bisection method, and computing the corresponding maximizer. This procedure, proposed by Dinkelbach [99], is summarized in **Algorithm 1**. Evidently, Problem (2.32) fulfils the conditions of Lemma 2.2.1 with

$$\mathbf{x} = \Delta\boldsymbol{\theta} \in \mathbb{R}^2, \quad (2.37)$$

$$N(\mathbf{x}) = (\bar{\mathbf{p}} + \bar{\mathbf{H}}\mathbf{x})^\dagger \bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger (\bar{\mathbf{p}} + \bar{\mathbf{H}}\mathbf{x}), \quad (2.38)$$

$$D(\mathbf{x}) = (\bar{\mathbf{p}} + \bar{\mathbf{H}}\mathbf{x})^\dagger (\bar{\mathbf{p}} + \bar{\mathbf{H}}\mathbf{x}), \quad (2.39)$$

and

$$\mathcal{S} = \{\mathbf{x} \in \mathbb{R}^2 : |\mathbf{x}(1)| \leq \alpha, |\mathbf{x}(2)| \leq \beta\}. \quad (2.40)$$

In fact, the feasible set (2.40) is nonempty and compact, (2.38) and (2.39) are continuous functions with (2.39) strictly greater than zero over \mathcal{S} . As

a result, **Algorithm 1** can be applied to solve Problem (2.32), where step 4 becomes

$$\begin{aligned} \Delta\theta_n^* = \arg \max_{\Delta\theta \in \mathcal{S}} & (\bar{\mathbf{p}} + \bar{\mathbf{H}}\Delta\theta)^\dagger \bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger (\bar{\mathbf{p}} + \bar{\mathbf{H}}\Delta\theta) \\ & - q_n (\bar{\mathbf{p}} + \bar{\mathbf{H}}\Delta\theta)^\dagger (\bar{\mathbf{p}} + \bar{\mathbf{H}}\Delta\theta). \end{aligned} \quad (2.41)$$

The procedure devised to determine an optimal point $\Delta\theta_n^*$ is summarized in **Algorithm 2** (analytical details are reported in Appendix A.3). Specifically, in step 2, the candidate optimal solutions that lie within the interior of the feasible set are determined. Besides, in steps 3 and 4 the candidate optimal solutions belonging to the boundary of the feasible set, i.e., the four edges of the box, are computed. Finally, Step 5 derives the global optimal solution, selecting the best among all the obtained candidates.

Note that $F(0) > 0$ and $F(\|\bar{\mathbf{r}}\|^2) < 0$ with probability one⁵ provided that $NM > 3$. As a consequence, the bisection method involved in **Algorithm 1** can be initialized with $q_{lb} = 0$ and $q_{ub} = \|\bar{\mathbf{r}}\|^2$ to solve Problem (2.32).

Now, denoting by $\hat{\Delta}\theta_{DO}^*$ the DOA displacements estimated via **Algorithm 1** tailored to the problem at hand, the decision rule (2.18) becomes

$$t_{\text{GLRT-LAM-DO}} = \frac{1}{1 + \|\bar{\mathbf{r}}\|^2} \frac{\left| \bar{\mathbf{r}}^\dagger (\bar{\mathbf{p}} + \bar{\mathbf{H}}\hat{\Delta}\theta_{DO}^*) \right|^2}{\left\| \bar{\mathbf{p}} + \bar{\mathbf{H}}\hat{\Delta}\theta_{DO}^* \right\|^2} \underset{H_0}{\overset{H_1}{\gtrsim}} T, \quad (2.42)$$

which will be denoted hereafter as **GLRT-LAM-DO**. Finally, the **AMF** version of (2.42), referred to as **GLRT-LAM-DO-AMF**, is given by

$$t_{\text{GLRT-LAM-DO-AMF}} = \frac{\left| \bar{\mathbf{r}}^\dagger (\bar{\mathbf{p}} + \bar{\mathbf{H}}\hat{\Delta}\theta_{DO}^*) \right|^2}{\left\| \bar{\mathbf{p}} + \bar{\mathbf{H}}\hat{\Delta}\theta_{DO}^* \right\|^2} \underset{H_0}{\overset{H_1}{\gtrsim}} T_1. \quad (2.43)$$

⁵Both $F(0) = \arg \max_{\Delta\theta \in \mathcal{S}} \|\bar{\mathbf{r}}^\dagger (\bar{\mathbf{p}} + \bar{\mathbf{H}}\Delta\theta)\|^2 = 0$ and $F(\|\bar{\mathbf{r}}\|^2) = \arg \max_{\Delta\theta \in \mathcal{S}} \|\bar{\mathbf{r}}^\dagger (\bar{\mathbf{p}} + \bar{\mathbf{H}}\Delta\theta)\|^2 - \|\bar{\mathbf{r}}\|^2 \|\bar{\mathbf{p}} + \bar{\mathbf{H}}\Delta\theta\|^2 = 0$ force $\bar{\mathbf{r}}$ to lie in a specific subspace whose dimension is less than NM . This is an event that occurs with zero probability.

Algorithm 1 Dinkelbach's Optimization (DO) Algorithm

Input: $\mathcal{S} \subseteq \mathbb{R}^n$, $N(\mathbf{x})$, $D(\mathbf{x})$, q_{ub} , q_{lb} and ε_{DO} .**Output:** A solution \mathbf{x}^* to (2.34).

- 1: set $n = 0$.
 - 2: **do**
 - 3: $q_n = (q_{lb} + q_{ub})/2$;
 - 4: find $\mathbf{x}_n^* = \arg \max_{\mathbf{x} \in \mathcal{S}} \{N(\mathbf{x}) - q_n D(\mathbf{x})\}$;
 - 5: let $F(q_n) = \{N(\mathbf{x}_n^*) - q_n D(\mathbf{x}_n^*)\}$;
 - 6: **if** $F(q_n) \geq 0$ set $q_{lb} = q_n$, otherwise $q_{ub} = q_n$;
 - 7: $n = n + 1$;
 - 8: **until** $F(q_n) = 0$ or $(q_{ub} - q_{lb})/2 < \varepsilon_{DO}$
 - 9: output $\mathbf{x}^* = \mathbf{x}_n^*$.
-

GLRT-LAM-DO and GLRT-LAM-DO-AMF involve $\mathcal{O}(K(NM)^2 + N_{it,b})$ operations, where $N_{it,b}$ is the number of iterations required by the bisection algorithm to converge. In fact, the complexity of the Sample Covariance Matrix (SCM) inverse computation is $\mathcal{O}(K(NM)^2)$ and $\mathcal{O}(1)$ operations are necessary at each execution of **Algorithm 2**.

Coordinate Descent DOA displacements estimate

Exploiting the CD framework, in this subsection another method is proposed to handle Problem (2.32). The idea is to alternate between the maximizations over each entry of $\Delta\theta = [\Delta u, \Delta v]^T$, namely optimizing one variable at a time while keeping the other fixed. Note that in the presence of two blocks/variables the alternating update rule involved in the CD approach is equivalent to the Maximum Block Improvement (MBI) policy [29]. As a result, any limit point of the sequence of solutions produced by the CD procedure is a stationary point for Problem (2.32). In **Algo-**

Algorithm 2 Solution to Problem (2.41)

Input: $\bar{r}, \bar{p}, \dot{\bar{p}}_u, \dot{\bar{p}}_v, q_n, \alpha, \beta$.

Output: A solution $\Delta\theta_n^*$ to (2.41).

- 1: let $\mathcal{V} = \emptyset$;
- 2: compute the unconstrained stationary point $\Delta\theta_1$ of (2.41) (see equation (A.3)) and set

$$\mathcal{V} = \mathcal{V} \cup \{\Delta\theta_1\} \cap \mathcal{S};$$

- 3: restrict the objective of (2.41) to the right (left) edge of \mathcal{S} , i.e., $\Delta u = \alpha$ ($\Delta u = -\alpha$), and compute the corresponding stationary point $\Delta\theta_2$ ($\Delta\theta_3$) (see equation (A.11)); hence, set

$$\mathcal{V} = \mathcal{V} \cup \{\Delta\theta_2, \Delta\theta_3\};$$

- 4: restrict the objective of (2.41) to the upper (lower) edge of \mathcal{S} , i.e., $\Delta v = \beta$ ($\Delta v = -\beta$), and compute the corresponding stationary point $\Delta\theta_4$ ($\Delta\theta_5$) (see equation (A.12)); hence, set

$$\mathcal{V} = \mathcal{V} \cup \{\Delta\theta_4, \Delta\theta_5\};$$

- 5:

$$\Delta\theta_n^* = \arg \max_{\Delta\theta \in \mathcal{V}} N(\Delta\theta) - q_n D(\Delta\theta),$$

where $N(\Delta\theta)$ and $D(\Delta\theta)$ are given in (2.38) and (2.39), respectively, and \mathcal{V} is a set with cardinality at most 5;

- 6: output $\Delta\theta_n^*$.
-

Algorithm 3 CD Optimization (CDO) Algorithm**Input:** $\bar{\mathbf{r}}, \bar{\mathbf{p}}, \dot{\bar{\mathbf{p}}}_u, \dot{\bar{\mathbf{p}}}_v, q_{n_s}, \alpha, \beta, \varepsilon_{CDO}$.**Output:** A solution $\hat{\Delta\boldsymbol{\theta}}_{CDO}^*$ to Problem (2.32).

- 1: set $n = 0$, $\Delta\boldsymbol{\theta}^{(n)} = [\Delta u^{(n)}, \Delta v^{(n)}]^T = \mathbf{0}$,
 $\bar{\mathbf{p}}_{\Delta v}^{(n)} = \bar{\mathbf{p}} + \dot{\bar{\mathbf{p}}}_v \Delta v^{(n)}$,
 $\text{obj}^{(n)} = \frac{(\bar{\mathbf{p}} + \bar{\mathbf{H}} \Delta\boldsymbol{\theta}^{(n)})^\dagger \bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger (\bar{\mathbf{p}} + \bar{\mathbf{H}} \Delta\boldsymbol{\theta}^{(n)})}{(\bar{\mathbf{p}} + \bar{\mathbf{H}} \Delta\boldsymbol{\theta}^{(n)})^\dagger (\bar{\mathbf{p}} + \bar{\mathbf{H}} \Delta\boldsymbol{\theta}^{(n)})}$;

2: **repeat**3: $n = n + 1$;4: Δu optimization, i.e.,

$$\Delta u^* = \arg \max_{|\Delta u| \leq \alpha} \frac{(\bar{\mathbf{p}}_{\Delta v}^{(n-1)} + \dot{\bar{\mathbf{p}}}_u \Delta u)^\dagger \bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger (\bar{\mathbf{p}}_{\Delta v}^{(n-1)} + \dot{\bar{\mathbf{p}}}_u \Delta u)}{(\bar{\mathbf{p}}_{\Delta v}^{(n-1)} + \dot{\bar{\mathbf{p}}}_u \Delta u)^\dagger (\bar{\mathbf{p}}_{\Delta v}^{(n-1)} + \dot{\bar{\mathbf{p}}}_u \Delta u)},$$

and set $\bar{\mathbf{p}}_{\Delta u}^{(n)} = \bar{\mathbf{p}} + \dot{\bar{\mathbf{p}}}_u \Delta u^*$;5: Δv optimization, i.e.,

$$\Delta v^* = \arg \max_{|\Delta v| \leq \beta} \frac{(\bar{\mathbf{p}}_{\Delta u}^{(n)} + \dot{\bar{\mathbf{p}}}_v \Delta v)^\dagger \bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger (\bar{\mathbf{p}}_{\Delta u}^{(n)} + \dot{\bar{\mathbf{p}}}_v \Delta v)}{(\bar{\mathbf{p}}_{\Delta u}^{(n)} + \dot{\bar{\mathbf{p}}}_v \Delta v)^\dagger (\bar{\mathbf{p}}_{\Delta u}^{(n)} + \dot{\bar{\mathbf{p}}}_v \Delta v)},$$

and set $\bar{\mathbf{p}}_{\Delta v}^{(n)} = \bar{\mathbf{p}} + \dot{\bar{\mathbf{p}}}_v \Delta v^*$;6: $\Delta\boldsymbol{\theta}^{(n)} = [\Delta u^*, \Delta v^*]^T$ and

$$\text{obj}^{(n)} = \frac{(\bar{\mathbf{p}} + \bar{\mathbf{H}} \Delta\boldsymbol{\theta}^{(n)})^\dagger \bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger (\bar{\mathbf{p}} + \bar{\mathbf{H}} \Delta\boldsymbol{\theta}^{(n)})}{(\bar{\mathbf{p}} + \bar{\mathbf{H}} \Delta\boldsymbol{\theta}^{(n)})^\dagger (\bar{\mathbf{p}} + \bar{\mathbf{H}} \Delta\boldsymbol{\theta}^{(n)})}$$
;

7: **until** $|\text{obj}^{(n)} - \text{obj}^{(n-1)}| < \varepsilon_{CDO}$ 8: output $\hat{\Delta\boldsymbol{\theta}}_{CDO}^* = \Delta\boldsymbol{\theta}^{(n)}$.

Algorithm 3, the CD-based solution technique specific for Problem (2.32) is reported. Note that the optimizations required at steps 4 and 5 can be performed resorting to Proposition 1. Otherwise stated, closed-form optimal

solutions are available. Leveraging the output $\widehat{\Delta}\boldsymbol{\theta}_{CDO}^*$ of **Algorithm 3**, the following approximated versions of (2.42) and (2.43), referred to as **GLRT-LAM-CDO** and **GLRT-LAM-CDO-AMF**, are obtained

$$t_{\text{GLRT-LAM-CDO}} = \frac{1}{1 + \|\bar{\mathbf{r}}\|^2} \frac{\left| \bar{\mathbf{r}}^\dagger (\bar{\mathbf{p}} + \bar{\mathbf{H}} \widehat{\Delta}\boldsymbol{\theta}_{CDO}^*) \right|^2}{\left\| \bar{\mathbf{p}} + \bar{\mathbf{H}} \widehat{\Delta}\boldsymbol{\theta}_{CDO}^* \right\|^2} \underset{H_0}{\overset{H_1}{\gtrless}} T, \quad (2.44)$$

and

$$t_{\text{GLRT-LAM-CDO-AMF}} = \frac{\left| \bar{\mathbf{r}}^\dagger (\bar{\mathbf{p}} + \bar{\mathbf{H}} \widehat{\Delta}\boldsymbol{\theta}_{CDO}^*) \right|^2}{\left\| \bar{\mathbf{p}} + \bar{\mathbf{H}} \widehat{\Delta}\boldsymbol{\theta}_{CDO}^* \right\|^2} \underset{H_0}{\overset{H_1}{\gtrless}} T_1. \quad (2.45)$$

The implementation of **GLRT-LAM-CDO** and **GLRT-LAM-CDO-AMF** require $\mathcal{O}(K(NM)^2 + N_{it,CD})$ operations, where $N_{it,CD}$ denotes the number of iterations of the **CD** method up to convergence. Precisely $\mathcal{O}(K(NM)^2)$ are due to **SCM** and $\mathcal{O}(1)$ operations are necessary for each iteration of the **CD** method.

2.3 Performance Analysis

This section is aimed at assessing the performance of the proposed strategies for joint target detection and angle estimation in comparison with some counterparts available in the open literature specifically designed either for detection or **DOA** evaluation.

In the reported case studies the disturbance covariance matrix is modeled as $\mathbf{M} = \mathbf{M}_J + \sigma_a^2 \mathbf{I}$ where σ_a^2 is the white noise power level (assumed without loss of generality equal to 0 dB) and \mathbf{M}_J refers to the jamming signals covariance contribution. Specifically, denoting by J_{NB} and J_{WB} the number of narrow-band and wide-band jammers, $\mathbf{M}_J = \mathbf{M}_1 + \mathbf{M}_2$, where

$$\mathbf{M}_1 = \sum_{j=1}^{J_{NB}} \sigma_j^2 \mathbf{p}_J(u_j, v_j) \mathbf{p}_J(u_j, v_j)^\dagger, \quad (2.46)$$

with $\mathbf{p}_J(u_j, v_j)$ the steering vector and σ_j^2 the power of the j -th jammer,

while

$$\mathbf{M}_2 = \sum_{h=1}^{J_{WB}} \bar{\sigma}_h^2 \frac{1}{B_h} \int_{-\frac{B_h}{2}}^{\frac{B_h}{2}} \mathbf{p}_J(u_h, v_h) \mathbf{p}_J(u_h, v_h)^\dagger df \quad (2.47)$$

with $\mathbf{p}_J(u_h, v_h)$, $\bar{\sigma}_h^2$, (u_h, v_h) , and B_h , the steering vector, the power, the DOA parameters, and the actual bandwidth associated with the h -th interferer, respectively.

In the following, three different interfering environments are analyzed:

- Scenario 1: two narrow-band jammers located at $u_1 = v_1 = 0.1$ and $u_2 = v_2 = 0.3$, with Jammer to Noise Ratio (JNR) given by $\text{JNR}_1 = 30$ dB and $\text{JNR}_2 = 40$ dB, respectively ($\sigma_j^2 = \text{JNR}_j \sigma_a^2$, $j = 1, 2$).
- Scenario 2: one narrow-band jammer at $u_1 = v_1 = 0.1$, with $\text{JNR}_1 = 30$ dB ($\sigma_1^2 = \text{JNR}_1 \sigma_a^2$), and one wide-band jammer ($B_f = 0.3$) at $u_2 = v_2 = 0.3$ with $\text{JNR}_2 = 40$ dB ($\bar{\sigma}_2^2 = \text{JNR}_2 \sigma_a^2$).
- Scenario 3: two narrow-band jammers located at $u_1 = v_1 = 0.2$ and $u_2 = v_2 = 0.3$, with $\text{JNR}_1 = 30$ dB and $\text{JNR}_2 = 40$ dB, respectively ($\sigma_j^2 = \text{JNR}_j \sigma_a^2$, $j = 1, 2$).

As already claimed, both detection and angle estimation capabilities of the proposed processors are analyzed. As to the former, the metric used to assess the performance is the Probability of Detection (P_D) estimated via standard Monte Carlo counting techniques over 10^4 independent trials. The threshold is set in order to guarantee a P_{fa} of 10^{-4} and it is evaluated using $100/P_{fa}$ independent trials. The decision statistics

$$t_{\text{GLRT}} = \frac{|\mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{p}|^2}{(1 + \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{r}) \mathbf{p}^\dagger \mathbf{S}^{-1} \mathbf{p}},$$

$$t_{\text{AMF}} = \frac{|\mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{p}|^2}{\mathbf{p}^\dagger \mathbf{S}^{-1} \mathbf{p}},$$

$$t_{\text{SD}} = \frac{\mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{H}_{SD} \left(\mathbf{H}_{SD}^\dagger \mathbf{S}^{-1} \mathbf{H}_{SD} \right)^{-1} \mathbf{H}_{SD}^\dagger \mathbf{S}^{-1} \mathbf{r}}{1 + \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{r}},$$

$$t_{\text{SD-AMF}} = \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{H}_{SD} \left(\mathbf{H}_{SD}^\dagger \mathbf{S}^{-1} \mathbf{H}_{SD} \right)^{-1} \mathbf{H}_{SD}^\dagger \mathbf{S}^{-1} \mathbf{r},$$

referred to as **GLRT** [60], **AMF** [93], Subspace Detector (**SD**) [63], and **SD-AMF** [39], [87], respectively, are considered for comparison purposes. **GLRT** and **AMF** detectors consider the nominal steering vector \mathbf{p} as useful signal directions (they usually operate in mismatched conditions and are thus also referred to as mismatched detectors), while **SD** and **SD-AMF** assume as useful signal directions those given by the columns of \mathbf{H}_{SD} . Finally, to assess the limits of the proposed algorithms, the **GLRT** and the **AMF** receivers with a perfect knowledge of the target **DOA** parameters (indicated as **GLRT-bench** and **AMF-bench**, respectively) are included as benchmarks.

In regard to the estimation performance, the Mean Square Error (**MSE**) is considered as the figure of merit. Again, Monte Carlo counting techniques are used to compute the **MSE** as

$$\widehat{\text{MSE}} = \frac{1}{M_C} \sum_{i=1}^{M_C} \left\| \Delta\boldsymbol{\theta}_0 - \widehat{\Delta\boldsymbol{\theta}}_i \right\|^2, \quad (2.48)$$

where $M_C = 10^4$ indicates the number of Monte Carlo independent trials, $\Delta\boldsymbol{\theta}_0 \in \mathbb{R}^2$ is the actual **DOA** displacements vector and $\widehat{\Delta\boldsymbol{\theta}}_i$ refers to the estimate provided at the i -th trial by a given technique. As performance benchmark, the **CRLB** for **DOA** displacements (see Appendix A.4) is reported too.

Two different simulation setups are considered in the following subsections to shed light on the performance limit of the proposed radar processors: a) linearized array manifold signal model; b) actual array manifold. Finally, the Signal to Interference plus Noise Ratio (**SINR**) is defined as

$$\text{SINR} = |a|^2 \mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{p}. \quad (2.49)$$

2.3.1 Linearized Array Manifold Signal Model

The performance of a radar system equipped with either a **1-D** or a **2-D** array pointing at the boresight direction is studied. The former employs a Uniform Linear Array (**ULA**) with $N = 16$ and $d_x = \lambda_0/2$. The latter exploits a **URA** with $N = M = 5$ and $d_x = d_y = \lambda_0/2$; in both cases the reference system is centered at the bottom-left corner. Within this

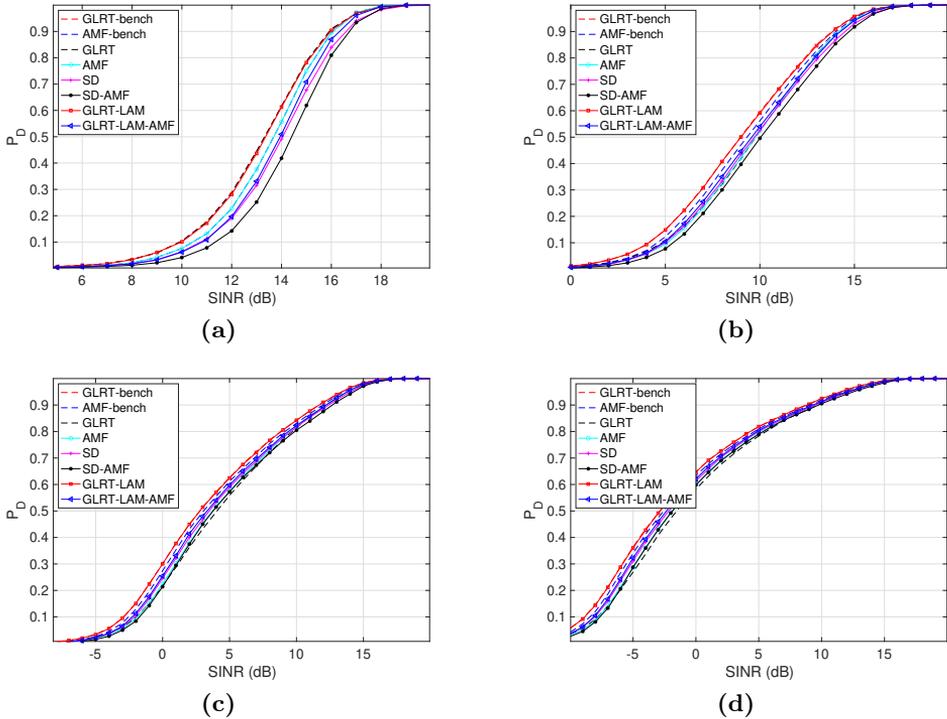


Figure 2.3. Detection performance for a ULA with 16 antennas in Scenario 1 for different mismatches, assuming $K = 32$ and $\alpha = 0.5$. The actual DOA displacement is drawn from a uniform distribution over $[-\Delta, \Delta]$: (a) $\Delta = 0$, (b) $\Delta = 0.1$, (c) $\Delta = 0.25$, and (d) $\Delta = 0.5$.

subsection, the data from the cell under test are modeled as

$$\mathbf{r} = a(\mathbf{p} + \dot{\mathbf{p}}_u \Delta u + \dot{\mathbf{p}}_v \Delta v) + \mathbf{n}, \quad (2.50)$$

namely according to the linearized array manifold⁶, in order to assess the capabilities of the devised signal processing techniques under nominal design conditions.

Fig. 2.3 shows the P_D curves for the 1-D case, assuming the interference environment of Scenario 1, with $K = 2N = 32$ secondary data. Therein,

⁶In the 1-D case the useful signal contribution becomes $\mathbf{r} = a(\mathbf{p} + \dot{\mathbf{p}}_u \Delta u) + \mathbf{n}$.

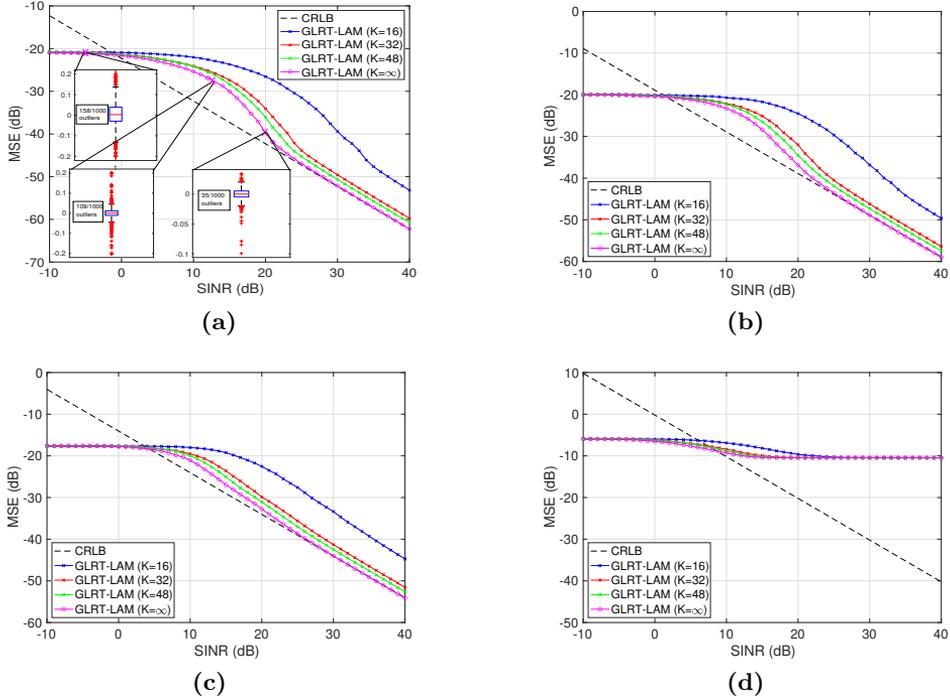


Figure 2.4. Estimation performance for a ULA with 16 antennas in Scenario 1 assuming $\alpha = 0.2$ and considering different sample support sizes, i.e., $K = [16, 32, 48, \infty]$, and several target locations: (a) $\Delta u = 0$, (b) $\Delta u = 0.05$, (c) $\Delta u = 0.1$, and (d) $\Delta u = 0.5$.

the design parameter α is set at 0.5. The actual DOA displacement is drawn from a uniform distribution over $[-\Delta, \Delta]$ and each subfigure refers to a specific value of Δ . Specifically, Fig. 2.3(a) considers $\Delta = 0$, i.e., the target is exactly matched to the array pointing direction, whereas Figs. 2.3(b), 2.3(c), and 2.3(d), consider $\Delta = 0.1, 0.25$, and 0.5 , respectively.

Inspection of the figures highlights that the GLRT-LAM detector exhibits performance very close to the GLRT-bench and outperforms all the other counterparts (including the AMF-bench) regardless of the operating conditions. Besides, GLRT-LAM-AMF experiences a performance degradation (about 0.5 dB at $P_D = 0.9$, in the worst case) with respect to the corresponding benchmark. However, it achieves higher P_D levels than the

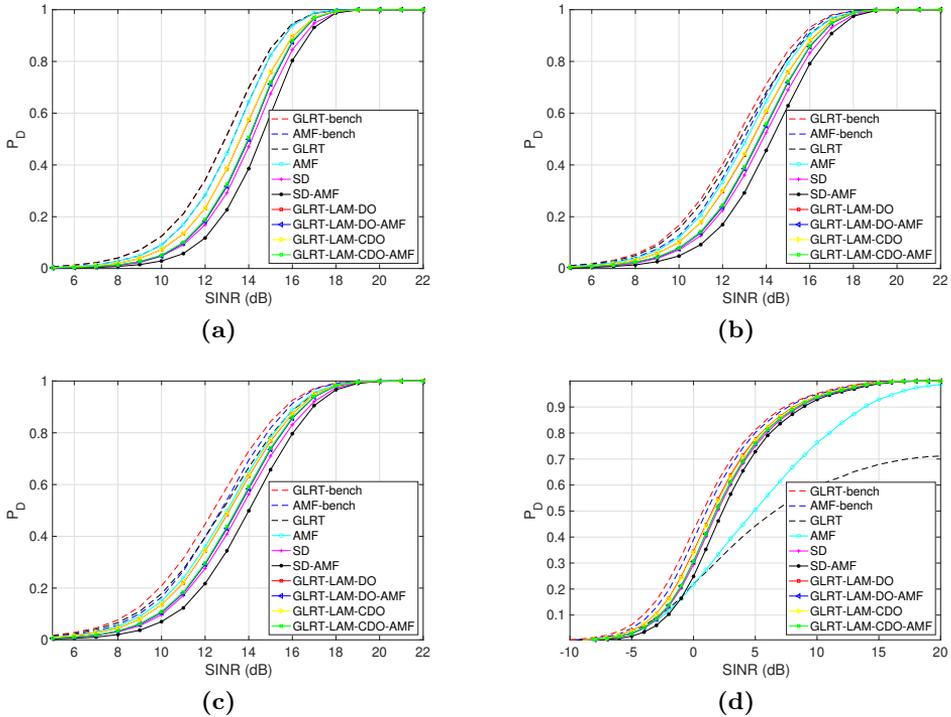


Figure 2.5. Detection performance for a URA with 5×5 antennas, $K = 50$ secondary data, and different mismatch conditions, in Scenario 2. The constraint parameters α and β are set at 0.5. The target's location $(\Delta u, \Delta v)$ is modeled as a pair of statistically independent uniform random variables in the region $[-\Delta_1, \Delta_1] \times [-\Delta_2, \Delta_2]$: (a) $(\Delta_1, \Delta_2) = (0, 0)$, (b) $(\Delta_1, \Delta_2) = (0, 0.05)$, (c) $(\Delta_1, \Delta_2) = (0.05, 0.05)$, and (d) $(\Delta_1, \Delta_2) = (0.5, 0.5)$.

mismatched detectors (of course apart from the case of $\Delta = 0$) and the subspace receivers in all the configurations, revealing the effectiveness of the method to estimate the actual steering vector.

To assess the estimation capabilities of the proposed constrained MLE $\hat{\Delta u}^*$ (see Proposition 1), the MSE versus SINR is displayed in Fig. 2.4 for different sample support sizes⁷, i.e., $K = [16, 32, 48, \infty]$, and target locations, i.e., $\Delta u \in \{0, 0.05, 0.1, 0.5\}$. In this case, $\alpha = 0.2$ and the

⁷ $K = \infty$ is tantamount to considering the exact covariance matrix.

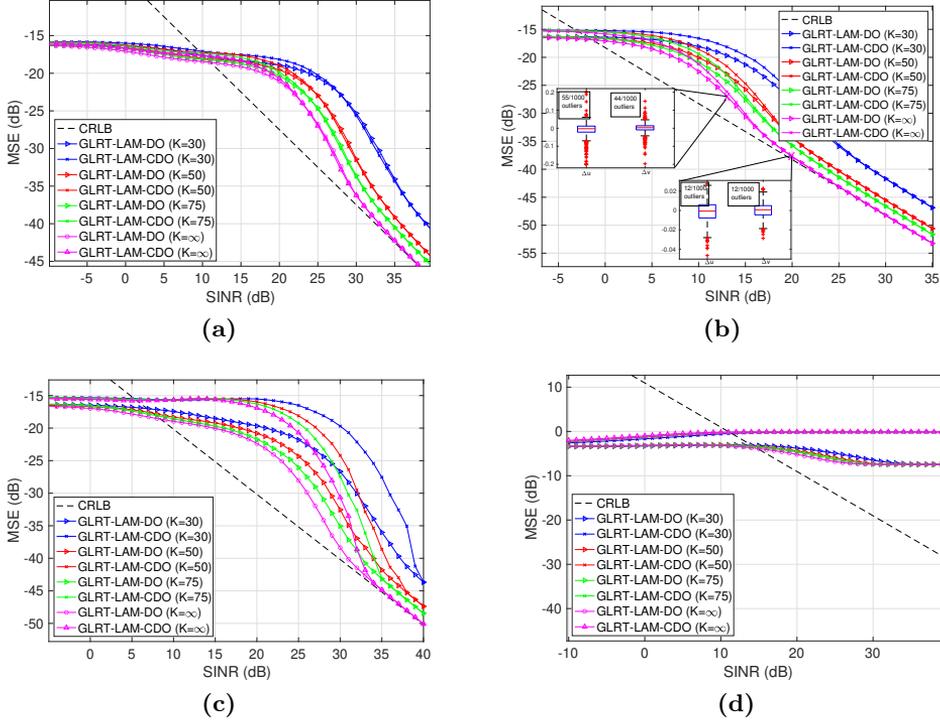


Figure 2.6. Estimation performance for a URA with 5×5 antennas in Scenario 2 for different sample support sizes, i.e., $K = [30, 50, 70, \infty]$ and several target locations, i.e., (a) $(\Delta u, \Delta v) = (0, 0)$, (b) $(\Delta u, \Delta v) = (0, 0.05)$, (c) $(\Delta u, \Delta v) = (0.05, 0.05)$, and (d) $(\Delta u, \Delta v) = (0.5, 0.5)$. Therein $\alpha = \beta = 0.2$. In Fig. (b) are also reported two box-and-whisker plots at 13 dB and 20 dB, respectively.

interfering setup of Scenario 1 is analyzed. Figs. 2.4(a), 2.4(b), 2.4(c), and 2.4(d) refer to $\Delta u = 0$, $\Delta u = 0.05$, $\Delta u = 0.1$, and $\Delta u = 0.5$, respectively. As expected, the MSE curves decrease with the SINR and the higher K the lower the estimation error (in the mean square sense), being better and better the accuracy of interference covariance matrix estimate.

The results clearly show the effectiveness of the proposed estimator. Indeed, in the high SINR regime, the performance becomes closer to the CRLB benchmark as K increases; of course, this happens when the actual

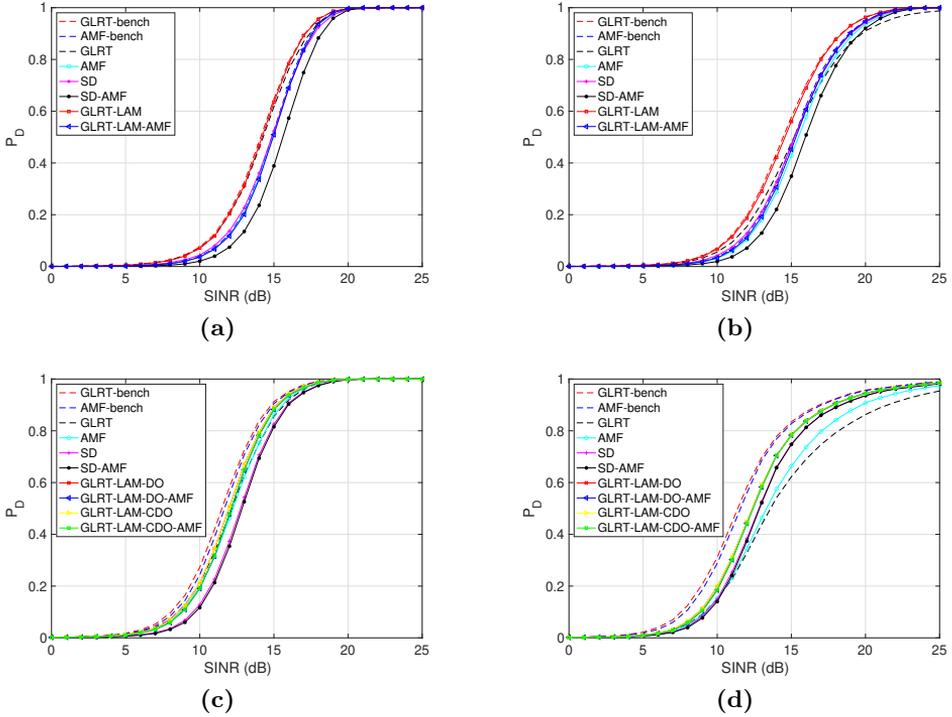


Figure 2.7. Detection performance for the actual array manifold assuming the interfering environment of Scenario 3. Figs. (a) and (b) report detection performance for a symmetric ULA with 9 antennas with $K = 18$, target location uniformly distributed over $[-\Delta, \Delta]$, and $\alpha = \Delta$: (a) $\Delta = 0.0523$ and (b) $\Delta = 0.891/N$. Figs. (c) and (d) correspond to a symmetric URA with 5×5 antennas, with $K = 75$, target location offset Δu and Δv modeled as a pair of statistically independent uniformly distributed random variables over $[-\Delta_1, \Delta_1]$ and $[-\Delta_2, \Delta_2]$, respectively, and $\alpha = \Delta_1$, $\beta = \Delta_2$: (c) $\Delta_1 = \Delta_2 = 0.1$ and (d) $\Delta_1 = \Delta_2 = 0.891/5 = 0.1782$.

target displacement belongs to the assumed uncertainty region. Otherwise, see Fig. 2.4(d), the MSE curves reach an error-floor of $(0.5 - 0.2)^2 = -10.4576$ dB. In this last situation, the devised technique reaches the feasible value closest to the actual target displacement, further corroborating the estimation capabilities of the devised strategy. At low SINR, smaller values than the CRLB benchmark are observed indicating that the pro-

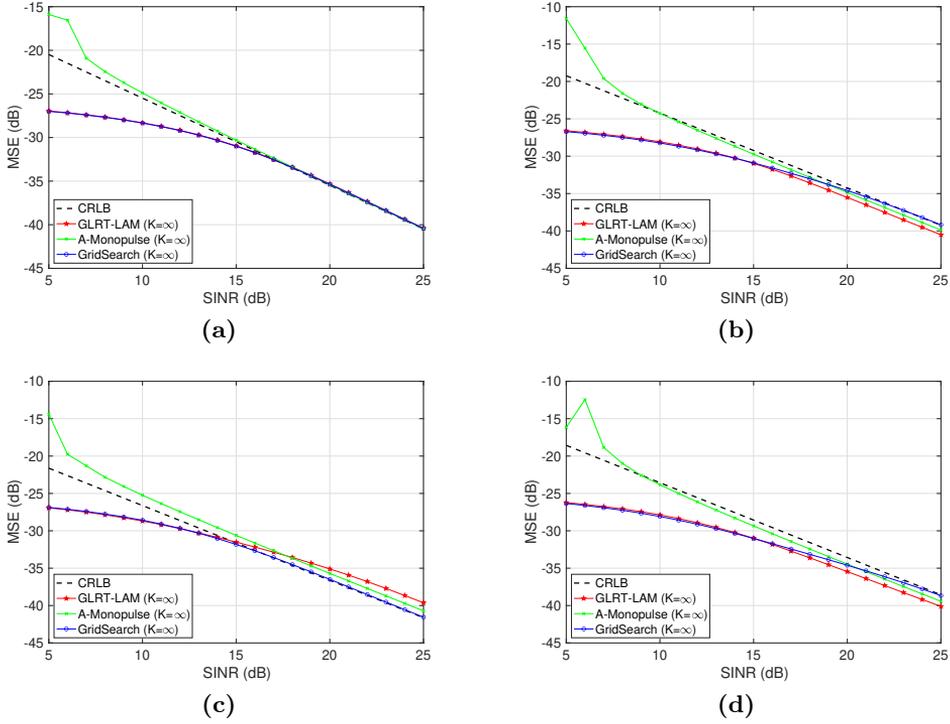


Figure 2.8. Estimation performance for a symmetric ULA with 9 antennas in Scenario 3 assuming $K = \infty$ and $\alpha = 0.0523$. The target is located at Δu with (a) $\Delta u = 0$, (b) $\Delta u = 0.01745$, (c) $\Delta u = -0.01745$, and (d) $\Delta u = 0.02618$.

posed estimator exhibits a bias under this **SINR** regime as well as an upper bound to the mean square values induced by the enforced constraint. Indeed, the **MSE** of the proposed estimator complies with

$$E[(\hat{\theta} - \theta)^2] \leq \max[(\theta - \alpha)^2, (\theta + \alpha)^2], \quad (2.51)$$

namely, the **MSE** exhibits a **SINR**-independent upper bound. On the contrary, equation (A.23) shows that the **CRLB** is unbounded above as the **SINR** goes to zero. As a result, the performance of our estimator is always better than the **CRLB** at low **SINR**. Similar considerations hold true for the **MSE** behaviour at low **SINR**, in all the subsequent case studies. Fi-

nally, it is also worth noting that the box-and-whisker plots [103] of the two estimation errors reported in Fig. 2.4(a), reveal the presence of a large number of outliers (points above and below the black whiskers) at low SINR. This behaviour provides a further explanation of the increasing departure of the MSE curves from the CRLB when the SINR ranges in the so called “below threshold region”.

Fig. 2.5 displays P_D curves for the 2-D case, assuming $K = 2MN = 50$, $\alpha = \beta = 0.5$, and Scenario 2. Therein, the displacements vector $(\Delta u, \Delta v)$ is drawn from a uniform distribution over $[-\Delta_1, \Delta_1] \times [-\Delta_2, \Delta_2]$. Specifically, Figs. 2.5(a), 2.5(b), 2.5(c), and 2.5(d) refer to $(\Delta_1, \Delta_2) = (0, 0)$, $(\Delta_1, \Delta_2) = (0, 0.05)$, $(\Delta_1, \Delta_2) = (0.05, 0.05)$, and $(\Delta_1, \Delta_2) = (0.5, 0.5)$, respectively. The results show the power of the GLRT-LAM-DO and GLRT-LAM-CDO detector. Indeed, the two algorithms attain the same P_D levels with performance very close to the clairvoyant GLRT-bench, with a loss smaller than 1 dB. Furthermore, apart from the case of $(\Delta_1, \Delta_2) = (0, 0)$, GLRT-LAM-DO and GLRT-LAM-CDO outperform the mismatched detector with performance gains higher and higher as the actual DOA offset region enlarges. Finally, they achieve P_D levels higher than the SD in all the configurations, revealing the capabilities of the new methods to benefit from the underlying structure of the weights. Similar results are achieved by the AMF version of the proposed detectors with respect to their relative counterparts.

Fig. 2.6 reports the MSE versus SINR for different sample support sizes, i.e., $K = [30; 50; 75; \infty]$, and $\alpha = \beta = 0.2$, with reference to Scenario 2. The target is located at $(\Delta u, \Delta v)$ and Figs. 2.6(a), 2.6(b), 2.6(c), and 2.6(d) refer to $(\Delta u, \Delta v) = (0, 0)$, $(\Delta u, \Delta v) = (0, 0.05)$, $(\Delta u, \Delta v) = (0.05, 0.05)$, and $(\Delta u, \Delta v) = (0.5, 0.5)$, respectively. As already seen in the 1-D case, the MSE curves decrease with the SINR and the higher K the lower the estimation error (in the mean square sense). In the high SINR regime, the performance approaches the CRLB benchmark as K increases, provided that the actual target displacement lies within the assumed uncertainty region. However, it is also worth pointing out that the MSE lower bound in the scenario of Fig. 2.6(d) is equal to $(0.5 - 0.2)^2 + (0.5 - 0.2)^2 = -7.4473$ dB and coincides with the error-floor level achieved by the devised techniques. Otherwise stated, in this case at high SINR the feasible point closer to the actual target displacement vector is returned as es-

timate. Besides, Fig. 2.6(c) outlines a departure of the MSE curves of GLRT-LAM-CDO and GLRT-LAM-DO. Not surprisingly, they pinpoint the limits of the sub-optimal optimization approach used in GLRT-LAM-CDO with respect to the optimal one used in GLRT-LAM-DO. Finally, the MSE behaviour at low SINR, where smaller values than the CRLB benchmark are achieved, indicates the presence of a bias. As in the 1-D case, the box-and-whisker plots for two operating SINR points, obtained via GLRT-LAM-DO with $K = \infty$ and reported in Fig. 2.6(b), confirm the role played by the outliers in the departure of the MSE curves from the CRLB.

Table 2.1. Computational Complexity of the considered processors.

Methods	Computational Costs
GLRT	$\mathcal{O}(KL^2)$
AMF	$\mathcal{O}(KL^2)$
SD	$\mathcal{O}(KL^2)$
SD-AMF	$\mathcal{O}(KL^2)$
GLRT-LAM	$\mathcal{O}(KN^2)$
GLRT-LAM-AMF	$\mathcal{O}(KN^2)$
GLRT-LAM-DO	$\mathcal{O}(K(NM)^2 + N_{it,b})$
GLRT-LAM-DO-AMF	$\mathcal{O}(K(NM)^2 + N_{it,b})$
GLRT-LAM-CDO	$\mathcal{O}(K(NM)^2 + N_{it,CD})$
GLRT-LAM-CDO-AMF	$\mathcal{O}(K(NM)^2 + N_{it,CD})$

A computational complexity comparison between GLRT-LAM-DO and GLRT-LAM-CDO is addressed in terms of the average number of iterations involved in the computation of the decision statistics. Assuming as exit condition $\varepsilon_{DO} = 10^{-5}$ and $\varepsilon_{CDO} = 10^{-5}$, the average number of iterations over 100 trials, related to the simulation setup of Fig. 2.6(b), is given by 20 and 10 for GLRT-LAM-DO and GLRT-LAM-CDO, respectively. As a consequence, GLRT-LAM-CDO exhibits a reduced computational burden as compared with GLRT-LAM-DO, at the price of a slight performance degradation. Finally, the computational complexity of all the analyzed architectures is provided in Table 2.1, where $L = N$ for the 1-D case whereas $L = NM$, for the 2-D case of the GLRT, AMF, SD, and SD-AMF

receivers. Table 2.1 emphasizes that the growth curve of the number of operations with respect to the problem size is ruled by the law KL^2 .

2.3.2 Actual Array Manifold

The considered radar system employs either a **ULA** of $N = 9$ elements or a **URA** with 25 elements arranged in a 5×5 square matrix. For both **1-D** and **2-D** setup, the array points the beam in the boresight direction and the spacing among the elements of the array is $\lambda_0/2$. Moreover, the reference system center coincides with the center of the array. The useful signal contribution is generated according to the actual array manifold, i.e., the steering vector is

$$\mathbf{p}(\bar{u} + \Delta u, \bar{v} + \Delta v), \quad (2.52)$$

with $\bar{u} = \bar{v} = 0$.

For a comparative analysis in terms of angular estimation, the following competitors are considered in the sequel:

- Adaptive Monopulse Procedure (**AMP**), referred to in the following as A-Monopulse and reported in Appendix A.1 as **Algorithm 13**.
- **MLE** implementation via grid search (only in the **1-D** case), defined as follow

$$\hat{\Delta}_{uML} = \arg \max_{\Delta u \in \mathcal{G}} \frac{|\mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{p}(\bar{u} + \Delta u)|^2}{\mathbf{p}^\dagger(\bar{u} + \Delta u) \mathbf{S}^{-1} \mathbf{p}(\bar{u} + \Delta u)}, \quad (2.53)$$

where

$$\mathcal{G} = \{-\alpha + i \alpha/200, \quad i = 0, \dots, 400\}. \quad (2.54)$$

For comparison purposes, the **CRLB** [105, p. 927, eq. 8.34] is also considered.

Fig. 2.7 presents detection performance both for the **1-D** and **2-D** scenarios. The former is considered in Figs. 2.7(a) and 2.7(b) under the interfering setup described in Scenario 3 and assuming $K = 18$. The actual **DOA** displacement is drawn from a uniform distribution over $[-\Delta, \Delta]$, where in Fig. 2.7(a) $\alpha = \Delta = \sin(3\pi/180) = 0.0523$ while in Fig. 2.7(b) $\alpha = \Delta = 0.891/9 = 0.099$. The **2-D** case is displayed in Figs. 2.7(c) and

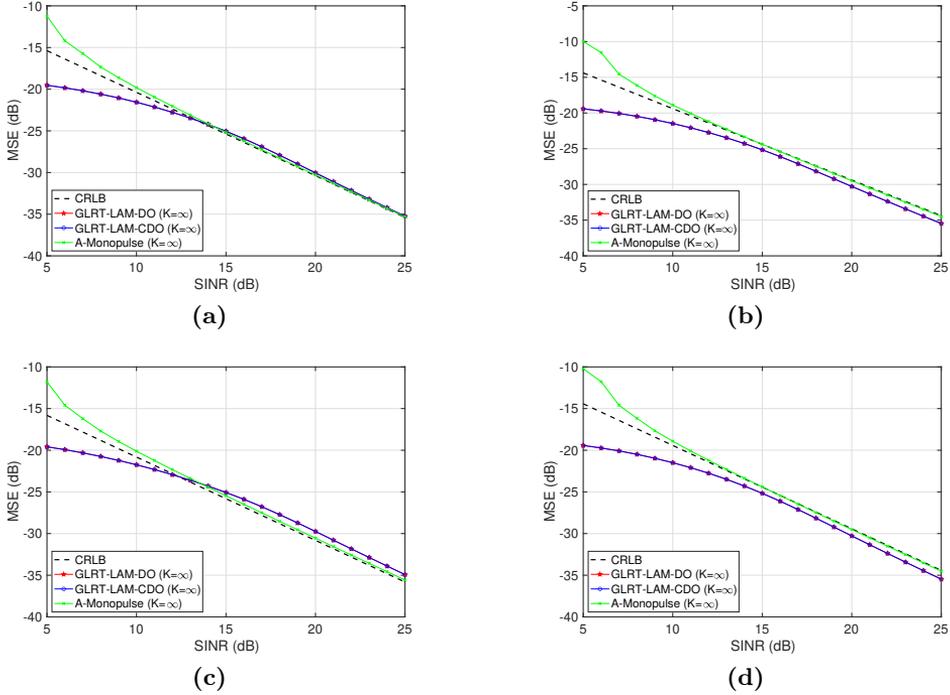


Figure 2.9. Estimation performance for a symmetric URA with 5×5 antennas in Scenario 3 assuming $K = \infty$, $\alpha = 0.1$, $\beta = 0.1$, and target located at $(\Delta u, \Delta v)$, with (a) $(\Delta u, \Delta v) = (0, 0)$, (b) $(\Delta u, \Delta v) = (0.0349, 0)$, (c) $(\Delta u, \Delta v) = (-0.01745, 0)$, and (d) $(\Delta u, \Delta v) = (0.0349, -0.001218)$.

2.7(d) assuming $K = 75$ and the disturbance environment corresponding to Scenario 3. The target location offsets Δu and Δv are modeled as statistically independent random variable with $\Delta u \sim U(-\Delta_1, \Delta_1)$ and $\Delta v \sim U(-\Delta_2, \Delta_2)$. Fig. 2.7(c) refers to $\alpha = \beta = \Delta_1 = \Delta_2 = 0.1$ and Fig. 2.7(d) considers $\alpha = \beta = \Delta_1 = \Delta_2 = 0.891/5 = 0.1782$, respectively. Otherwise stated, in all the figures the design parameters α and β are matched to actual DOA uncertainty.

Inspection of the results shows that the proposed one-step GLRT detectors (GLRT-LAM for the 1-D case, GLRT-LAM-DO and GLRT-LAM-CDO for the 2-D case) ensure a performance level very close to the clairvoyant GLRT and outperform the counterparts for the considered simula-

tion scenarios. Specifically, the P_D curves of the GLRT-LAM and GLRT-LAM-AMF almost overlap with the corresponding benchmark limits. For the 2-D case, GLRT-LAM-DO and GLRT-LAM-CDO experience a slight performance degradation lower than 1 dB at $P_D=0.9$ with respect to the optimal receiver.

In Fig. 2.8 the DOA estimation capabilities of GLRT-LAM, A-Monopulse, and MLE via grid search, are analyzed for the 1-D case, assuming $K = \infty$, $\alpha = 0.0523$, and the interference environment of Scenario 3. Figs. 2.8(a), 2.8(b), 2.8(c), and 2.8(d) refer to $\Delta u = 0$, $\Delta u = 0.01745$, $\Delta u = -0.01745$, and $\Delta u = 0.02618$, respectively. The exploration of the curves reveals that GLRT-LAM outperforms the A-Monopulse for a wide range of SINR values and provides estimation performance almost overlapped to that of the MLE via grid search method up to a SINR of 20 dB, corroborating the strength of the new devised method. Furthermore, the higher the SINR the lower the MSE of all the estimators regardless of the setup (but for the A-Monopulse in Fig. 2.8(d) at low SINR) and performance levels comparable with the CRLB benchmark are achieved at the high SINR regime. In this respect, note that possible deviations from the CRLB of the MSE curves (at high SINR) may arise due to the bias of the estimators. Indeed, in the presence of a biased estimator, the general bound to consider is given in [104, p. 147], which can be also lower than the conventional CRLB for unbiased estimators.

Moreover, the GLRT-LAM and the MLE via grid search achieve MSE values smaller than the CRLB at low SINR, reflecting the presence of a bias in the estimators as well as the limit to the error imposed by the constraint.

The estimation performance for the 2-D case of GLRT-LAM-DO, GLRT-LAM-CDO, and A-Monopulse is analyzed in Fig. 2.9, considering $K = \infty$, $\alpha = 0.1$, $\beta = 0.1$, where each subfigure refers to a specific target location in the $u - v$ plane. It can be observed that in all the reported case studies GLRT-LAM-DO and GLRT-LAM-CDO achieve the same MSE values and exhibit better estimation capabilities than the A-Monopulse (apart from Fig. 2.9(b) where a slight loss appears), showing the benefits of the new signal processing strategies. Remarks similar to those made for Fig. 2.8 hold true with reference to the comparison of the estimators with the CRLB benchmark performance.

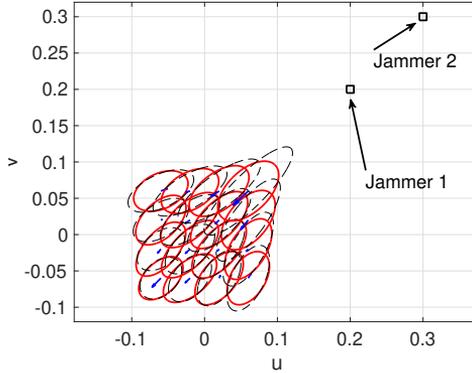


Figure 2.10. Bias and variance for a URA with 5×5 elements in Scenario 3 assuming $K = \infty$ and $\alpha = \beta = 0.1$. Dashed black ellipses refer to CRLB limit, variances are shown with solid red ellipses whereas the bias vectors are represented by solid blue arrows.

Finally, the average number of iterations (over 100 trials) required by [GLRT-LAM-DO](#) and [GLRT-LAM-CDO](#) in order to converge is 20 and 10, respectively, assuming $\varepsilon_{DO} = \varepsilon_{CDO} = 10^{-5}$ and the simulation setup of [2.9\(b\)](#). These results confirm that [Algorithm 3](#) is usually less demanding than [Algorithm 1](#), whilst ensuring satisfactory performance.

To further shed light on [Algorithm 1](#) estimation performance, [Fig. 2.10](#) displays the bias and variance ellipses granted by the proposed angle estimator for a grid of 16 displacements between the actual pointing direction and the array steering. The ellipses corresponding to the CRLB are also reported for comparison purposes. The simulation assumes the interference environment of Scenario 3, with $K = \infty$. The results highlight that for (u, v) -displacement in the left bottom corner of the grid (i.e., angle directions far from the jamming DOAs) the variance and CRLB ellipses match quite well. However, when displacements belongs to the upper right corner of the grid, some departures of the variance ellipse from the CRLB are experienced. Besides, as the pointing vector moves towards the jamming DOA, also the bias of the estimator increases.

2.3.3 Performance Analysis Employing Two-stage Processors

It is also worth pointing out that the proposed joint detection and estimation strategies can be interpreted as two-stage architectures, where first the target direction of arrival is estimated and then a bespoke adaptive detector is applied. Therefore, in the following, the comparison with other two-stage processors (in the sense of performing first the angle estimation with a given technique and then exploiting the resulting estimate within a decision statistic) is reported and discussed.

Specifically, two more terms of comparison, namely MON-KELLY and MON-AMF detectors are now considered. These receivers capitalize the monopulse angle estimate to predict the useful target steering vector of Kelly's and AMF's detectors, respectively. Specifically, denoting by $\hat{\boldsymbol{\theta}}_{NME}^* = [\hat{\Delta}u_{NME}, \hat{\Delta}v_{NME}]^T$ the monopulse estimate of the DOA displacements (obtained according to **Algorithm 13**), the decision statistics are formally defined as

$$t_{\text{MON-KELLY}} = \frac{\left| \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{p}(\bar{u} + \hat{\Delta}u_{NME}, \bar{v} + \hat{\Delta}v_{NME}) \right|^2}{(1 + \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{r}) \left| \mathbf{S}^{-1/2} \mathbf{p}(\bar{u} + \hat{\Delta}u_{NME}, \bar{v} + \hat{\Delta}v_{NME}) \right|^2} \quad (2.55)$$

and

$$t_{\text{MON-AMF}} = \frac{\left| \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{p}(\bar{u} + \hat{\Delta}u_{NME}, \bar{v} + \hat{\Delta}v_{NME}) \right|^2}{\left| \mathbf{S}^{-1/2} \mathbf{p}(\bar{u} + \hat{\Delta}u_{NME}, \bar{v} + \hat{\Delta}v_{NME}) \right|^2}. \quad (2.56)$$

Fig. 2.11 shows P_D versus SINR for the actual array manifold in Scenario 3 and with the threshold set in order to guarantee $P_{fa} = 10^{-4}$ (evaluated using $100/P_{fa}$ independent trials). Figs. 2.11(a) and 2.11(b) report the detection performance for a symmetric ULA with 9 antennas, $K = 18$, target location uniformly distributed over $[-\Delta, \Delta]$, and $\alpha = \Delta$: 2.11(a) $\Delta = 0.0523$ and 2.11(b) $\Delta = 0.891/N$. Figs. 2.11(c) and 2.11(d) show the detection performance for a symmetric URA with 5×5 antennas, assuming $K = 75$ and target location offset vector $(\Delta u, \Delta v)$ modeled as a uniform random variable over $[-\Delta_1, \Delta_1] \times [-\Delta_2, \Delta_2]$, with $\alpha = \Delta_1, \beta = \Delta_2$: 2.11(c)

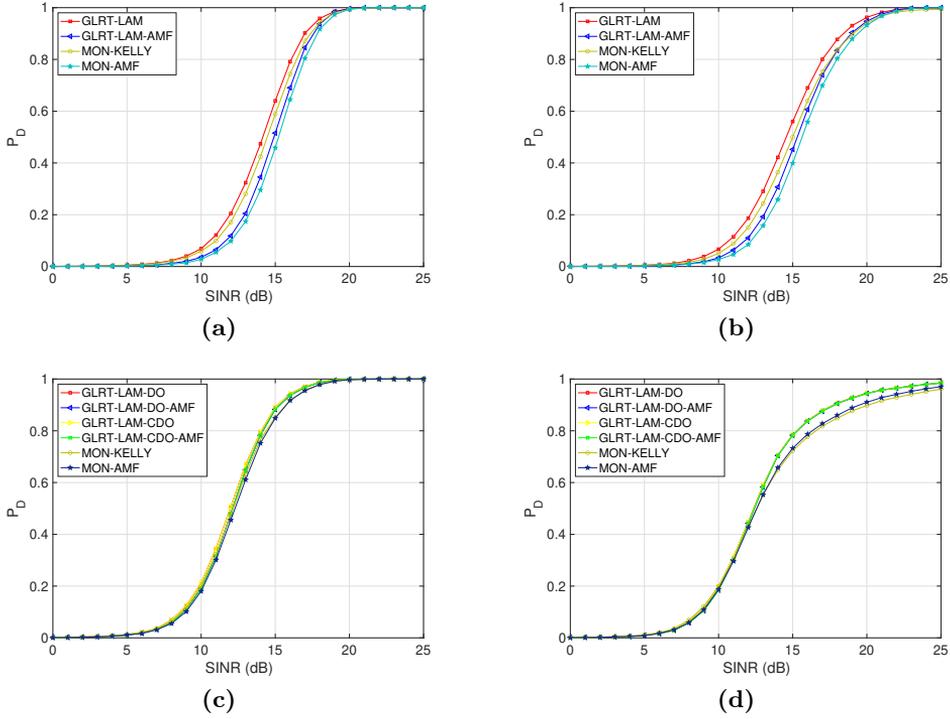


Figure 2.11. Detection performance for the actual array manifold assuming the interfering environment of Scenario 3. Figs. (a) and (b) consider a symmetric ULA with 9 antennas, $K = 18$ and target location uniformly distributed over $[-\Delta, \Delta]$ with $\alpha = \Delta$: (a) $\Delta = 0.0523$ and (b) $\Delta = 0.891/N$. Figs. (c) and (d) correspond to a symmetric URA with 5×5 antennas, assuming $K = 75$ and target location offset Δu and Δv modeled as a pair of statistically independent uniformly distributed random variables over $[-\Delta_1, \Delta_1]$ and $[-\Delta_2, \Delta_2]$, respectively. Besides, $\alpha = \Delta_1$, $\beta = \Delta_2$: (c) $\Delta_1 = \Delta_2 = 0.1$ and (d) $\Delta_1 = \Delta_2 = 0.891/5 = 0.1782$.

$\Delta_1 = \Delta_2 = 0.1$ and 2.11(d) $\Delta_1 = \Delta_2 = 0.891/5 = 0.1782$.

The results clearly confirm the effectiveness of the new devised joint detection and angle estimation procedures relying on **Algorithms 1 and 3**, with some detection performance improvement with respect to the considered heuristic two-stage counterparts.

2.4 Conclusion

This chapter has considered simultaneous detection and target angle localization for a multichannel phased array radar. Signal processing architectures have been proposed which, after target detection, are able to provide directly estimates of the target angular offsets from the array pointing direction. Two estimation procedures respectively based on the Dinkelbach's algorithm and a CD method have been devised. The former provides the optimal ML estimates of the unknown displacements. The latter generally exhibits a faster convergence than the former but the optimality of the solution (in the ML sense) cannot be claimed. A comparative analysis has been conducted with other techniques available in the open literature which are either tailored for detection or target angle localization. Benchmark limits have been also considered. The results highlight the interplay of the different design parameters and show that the new algorithms provide adequate performance, thus representing viable solutions for practical implementations.

Chapter 3

Structured Covariance Matrix Estimation with Missing-Data via Expectation-Maximization

Structured covariance matrix estimation in the presence of missing-(complex) data is addressed in this chapter¹ with emphasis on radar signal processing applications. In that regard, the array model for the missing-data scenario is specified and the problem of computing the maximum likelihood estimate of a structured covariance matrix is formulated. A general procedure to optimize the observed-data likelihood function is developed resorting to the expectation-maximization algorithm. The corresponding convergence properties are thoroughly established and the rate of convergence is analyzed. The estimation technique is contextualized for two practically relevant radar problems: beamforming and detection of the number of sources. In the former case an adaptive beamformer leveraging the [EM](#)-based estimator is presented; in the latter, detection techniques generalizing the classic Akaike information criterion, minimum description length, and Hannan–Quinn information criterion, are introduced. Numerical results are finally presented to corroborate the theoretical study.

¹©2021 IEEE. Reprinted, with permission, from A. Aubry, A. De Maio, S. Marano and M. Rosamilia, “Structured Covariance Matrix Estimation With Missing-(Complex) Data for Radar Applications via Expectation-Maximization,” *IEEE Transactions on Signal Processing*, vol. 69, pp. 5920-5934, 2021.

Summarizing, the main contributions of this chapter are:

- a) the development of an EM-based technique for the estimation of a structured covariance matrix in the presence of missing-(complex) data, considering uncertainty sets of practical interest for radar signal processing applications;
- b) the study of the convergence properties for the resulting iterative procedure according to B-stationarity, as well as the computation of the convergence rate;
- c) the application of the methodology to the context of two fundamental radar problems, i.e., beamforming and detection of the number of sources;
- d) the presentation of numerical results aimed at corroborating the theoretical achievements.

The rest of the chapter is organized as follows. Section 3.1 introduces the system model and defines some covariance matrix uncertainty sets of practical relevance. Section 3.2 formulates the structured covariance matrix estimation problem in the presence of missing-data and presents tailored iterative solution methods leveraging possible a-priori structural information. Besides, it addresses convergence issues about the proposed techniques. In Section 3.3, the performance of the estimators is analyzed in the context of adaptive beamforming and detection of number of sources. Finally, Section 3.4 draws some conclusions and highlights possible future research avenues.

3.1 Problem Formulation

Let us consider a radar system that collects spatial data via a narrow-band array composed of N antennas and operating in the presence of noise and interference, with unknown spectral characteristics. Let us suppose that a set of spatial snapshots \mathbf{r}_i , $i = 1, \dots, K$, modeled as Independent and Identically Distributed (IID) zero-mean circularly symmetric Gaussian random vectors (homogeneous environment) with unknown but structured

covariance matrix, is available. Specifically

$$\mathbf{r}_i \sim CN(0, \mathbf{M}), \mathbf{M} \in \mathcal{C} \subseteq \mathbb{H}_{++}^N, i = 1, \dots, K, \quad (3.1)$$

where \mathcal{C} denotes the subset of covariance matrices that can generate the observables. Enforcing \mathbf{M} to belong to \mathcal{C} is tantamount to exploiting some problem structure stemming from some a-priori knowledge about the operating environment (i.e., the number and the DOAs of active jammers, gathered via an Electronic Support Measurement (ESM) system [1]). Moreover, the array geometry (i.e., uniform linear array or other regular structures) [105] and the characteristics of the acquisition system (i.e., knowledge of the thermal noise level, or the available number of bits that results in an upper bound to the covariance condition number) induce a specific data covariance structure, to be wisely capitalized in order to boost radar sensing capabilities.

Some practical examples of covariance matrix uncertainty sets are now illustrated.

1. Structured covariance matrix with a lower bound on the white disturbance power level [101, 7]

$$\mathcal{C} = \left\{ \begin{array}{l} \mathbf{M} = \sigma_n^2 \mathbf{I} + \mathbf{r}_e \\ \mathbf{r}_e \succeq \mathbf{0} \\ \sigma_n^2 \geq \sigma^2 \end{array} \right., \quad (3.2)$$

where \mathbf{r}_e accounts for colored interference and clutter, $\sigma_n^2 > 0$ is the power of the white disturbance term, and $\sigma^2 > 0$ is a known lower bound on the white disturbance power.

2. Structured covariance matrix with a condition number constraint [10]

$$\mathcal{C} = \left\{ \begin{array}{l} \mathbf{M} = \sigma_n^2 \mathbf{I} + \mathbf{r}_e \\ \mathbf{r}_e \succeq \mathbf{0} \\ \sigma_n^2 \geq \sigma^2 \\ \frac{\lambda_{\max}(\mathbf{M})}{\lambda_{\min}(\mathbf{M})} \leq K_{max} \end{array} \right., \quad (3.3)$$

where \mathbf{r}_e , σ_n^2 , and σ^2 are defined as in (3.2), whereas K_{max} is an upper bound to the covariance condition number.

3. Structured covariance matrix with a rank constraint and a lower bound on the white disturbance power level [58]

$$\mathcal{C} = \begin{cases} \mathbf{M} = \sigma_n^2 \mathbf{I} + \mathbf{r}_e \\ \mathbf{r}_e \succeq \mathbf{0} \\ \text{Rank}(\mathbf{r}_e) \leq r \\ \sigma_n^2 \geq \sigma^2 \end{cases}, \quad (3.4)$$

where \mathbf{r}_e , σ_n^2 , and σ^2 are defined as in (3.2), whereas r is the maximum rank of \mathbf{r}_e .

4. Structured covariance matrix with a rank constraint [105, 22]

$$\mathcal{C} = \begin{cases} \mathbf{M} = \sigma_n^2 \mathbf{I} + \mathbf{V} \mathbf{S}_f \mathbf{V}^\dagger \\ \mathbf{V} \mathbf{S}_f \mathbf{V}^\dagger \succeq \mathbf{0} \\ \text{Rank}(\mathbf{V} \mathbf{S}_f \mathbf{V}^\dagger) \leq d \\ \sigma_n^2 > 0 \end{cases}, \quad (3.5)$$

where d is an upper bound to $\text{Rank}(\mathbf{V} \mathbf{S}_f \mathbf{V}^\dagger)$, σ_n^2 is defined as in (3.2), \mathbf{V} is an $N \times d$ array manifold matrix (which can be modeled either as a known or as an unknown parameter), \mathbf{S}_f denotes the $d \times d$ diagonal sources covariance matrix, whereas $d \leq N$ is the number of sources.

5. Structured covariance matrix with a centro-Hermitian symmetry [26]

$$\mathcal{C} = \begin{cases} \mathbf{M} = \mathbf{J} \mathbf{M}^* \mathbf{J} \\ \mathbf{M} \succeq \mathbf{0} \end{cases}, \quad (3.6)$$

with \mathbf{J} the $N \times N$ permutation matrix given by

$$\mathbf{J} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & 1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \quad (3.7)$$

6. Structured covariance matrix with a Toeplitz structure [122, 47, 43,
-

44]

$$\mathcal{C} = \begin{cases} \mathbf{M} \in \mathbb{T}^N \\ \mathbf{M} \succeq \mathbf{0} \end{cases}, \quad (3.8)$$

7. Structured covariance matrix with a Kronecker product structure [53, 110]

$$\mathcal{C} = \begin{cases} \mathbf{M} = \mathbf{r}_A \otimes \mathbf{r}_B \\ \mathbf{r}_A \succeq \mathbf{0} \\ \mathbf{r}_B \succeq \mathbf{0} \\ \mathbf{M} \succeq \mathbf{0} \end{cases}, \quad (3.9)$$

8. Structured real covariance matrix with a spectral symmetry [61, 27]

$$\mathcal{C} = \begin{cases} \mathbf{M} = \mathbf{M}^T \\ \mathbf{M} \succeq \mathbf{0} \end{cases}. \quad (3.10)$$

Besides, any combination of the above uncertainty sets (corresponding to their intersection) constitutes additional interesting examples.

The estimation problem (object of this chapter) demands a data model endowed with the capability to handle missing-data arising from the lack of some entries within specific spatial snapshots. To this end, each observed snapshot is modeled as

$$\mathbf{y}_i = \mathbf{A}_i \mathbf{r}_i, \quad i = 1, \dots, K, \quad (3.11)$$

where \mathbf{A}_i is the $p_i \times N$ selection matrix, constructed by extracting from \mathbf{I} the $p_i \leq N$ rows corresponding to the available observations at the i -th snapshot. In the following, the vectors \mathbf{r}_i and \mathbf{y}_i will be referred to as *complete* and *observed* data, respectively.

3.2 Covariance Matrix Estimation Procedure

This section is devoted to the derivation of a covariance matrix estimation procedure in the presence of missing-data accounting for model structures via suitable constraints. The problem is of primary importance for many applications in the field of radar signal processing [3, 109, 72, 71, 121, 106, 63, 38] and, in most cases, a ML estimator is usually demanded at

least due to its favorable asymptotic properties. For the missing-data case, the constrained **ML** estimate of the covariance matrix, given the observed-data, can be formulated as

$$\hat{\mathbf{M}}(\boldsymbol{\theta}) = \arg \max_{\mathbf{M}(\boldsymbol{\theta}) \in \mathcal{C}} \mathcal{L}_y(\mathbf{M}(\boldsymbol{\theta}) | \mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K), \quad (3.12)$$

with

$$\begin{aligned} \mathcal{L}_y(\mathbf{M}(\boldsymbol{\theta}) | \mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K) = & - \sum_{i=1}^K p_i \ln(\pi) - \sum_{i=1}^K \left[\ln(\det(\mathbf{A}_i \mathbf{M}(\boldsymbol{\theta}) \mathbf{A}_i^\dagger)) \right. \\ & \left. + \text{tr}\{(\mathbf{A}_i \mathbf{M}(\boldsymbol{\theta}) \mathbf{A}_i^\dagger)^{-1} \mathbf{y}_i \mathbf{y}_i^\dagger\} \right] \end{aligned} \quad (3.13)$$

the observed-data log-likelihood, $\mathbf{Y} = \{\mathbf{y}_1, \dots, \mathbf{y}_K\}$ the set of observed-data, and $\boldsymbol{\theta} \in \mathbb{R}^V$ the vector of the unknown parameters defining the underlying structure of \mathbf{M} . This is tantamount to solving

$$\hat{\boldsymbol{\theta}}_{ML} = \arg \max_{\boldsymbol{\theta}: \mathbf{M}(\boldsymbol{\theta}) \in \mathcal{C}} \mathcal{L}_y(\boldsymbol{\theta} | \mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K). \quad (3.14)$$

Computing $\hat{\boldsymbol{\theta}}_{ML}$ (or equivalently $\hat{\mathbf{M}}(\boldsymbol{\theta})$) is, in general, a difficult problem for which an analytic closed-form solution could not be available [24]. Besides, an optimization procedure based on a Multi-Dimensional Grid Search (**MDGS**) in the unknown parameter space could be computationally prohibitive. This motivates the interest toward iterative approximated procedures characterized by a more affordable computational cost than **ML** evaluation via **MDGS**.

3.2.1 EM Algorithm

EM is a widely adopted iterative technique to obtain approximate **ML** estimates of parameters from incomplete-data² [42, 105, 56]. The algorithm is composed of two steps. In the former, referred to as expectation (E) step, the conditional expectation of the complete-data likelihood, given

²In situations where direct access to the *complete* set of observations is not available, part of the data are missing or, more in general, data undergo a many-to-one mapping before becoming available to the observer.

the observed-data and the current estimate of the parameters, is evaluated (E-step score function). In the latter, referred to as the maximization (M) step, the E-step score function (corresponding to current estimate of the parameters) is maximized with respect to the unknowns. The EM starts with an initial guess of the parameters, i.e., $\boldsymbol{\theta}^{(0)}$, and iterates between E and M steps. The procedure can also be interpreted as a minorization-maximization optimization technique where the surrogate function stems from the Jensen inequality [54]. With reference to the estimation problem in (3.14), at the h -th iteration, the E-step consists in the evaluation of the score function

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(h-1)}) = \mathbb{E}[\mathcal{L}_r(\boldsymbol{\theta}) | \mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K, \hat{\mathbf{M}}(\boldsymbol{\theta}^{(h-1)})], \quad (3.15)$$

where

$$\mathcal{L}_r(\boldsymbol{\theta}) = -K[N \ln(\pi) + \ln(\det(\mathbf{M}(\boldsymbol{\theta}))) + \text{tr}\{\mathbf{M}(\boldsymbol{\theta})^{-1} \mathbf{S}\}] \quad (3.16)$$

is the complete-data log-likelihood,

$$\mathbf{S} = \frac{1}{K} \sum_{i=1}^K \mathbf{r}_i \mathbf{r}_i^\dagger \quad (3.17)$$

is the SCM of the complete-data, and $\hat{\mathbf{M}}(\boldsymbol{\theta}^{(h-1)})$ is the estimate of the covariance matrix at the $(h-1)$ -iteration. Computing the conditional expectation involved in (3.15) yields

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(h-1)}) = -K[N \ln(\pi) + \ln(\det(\mathbf{M}(\boldsymbol{\theta}))) + \text{tr}\{\mathbf{M}(\boldsymbol{\theta})^{-1} \boldsymbol{\Sigma}^{(h-1)}\}], \quad (3.18)$$

where

$$\boldsymbol{\Sigma}^{(h-1)} = \frac{1}{K} \sum_{i=1}^K \mathbf{C}_i^{(h-1)} \quad (3.19)$$

is the sample mean of the conditional correlation matrices

$$\mathbf{C}_i^{(h-1)} = \mathbb{E}[\mathbf{r}_i \mathbf{r}_i^\dagger | \mathbf{y}_i, \mathbf{A}_i, \hat{\mathbf{M}}(\boldsymbol{\theta}^{(h-1)})], \quad i = 1, \dots, K. \quad (3.20)$$

A closed-form expression to

$$\mathbb{E}[\mathbf{r}_i \mathbf{r}_i^\dagger | \mathbf{y}_i, \mathbf{A}_i, \mathbf{M}] = \mathbf{C}_i, \quad (3.21)$$

is given by (the interested reader may refer to Appendix B.1 for the detailed derivation)

$$\mathbf{C}_i = (\mathbf{A}_i^\dagger \mathbf{y}_i + \bar{\mathbf{A}}_i^\dagger \boldsymbol{\mu}_i) (\mathbf{A}_i^\dagger \mathbf{y}_i + \bar{\mathbf{A}}_i^\dagger \boldsymbol{\mu}_i)^\dagger + \bar{\mathbf{A}}_i^\dagger \mathbf{G}_i \bar{\mathbf{A}}_i \quad (3.22)$$

with $\bar{\mathbf{A}}_i$ the $N - p_i \times N$ selection matrix complementary to \mathbf{A}_i (obtained removing from \mathbf{I} the p_i rows corresponding to \mathbf{A}_i),

$$\boldsymbol{\mu}_i = \bar{\mathbf{A}}_i \mathbf{M} \mathbf{A}_i^\dagger (\mathbf{A}_i \mathbf{M} \mathbf{A}_i^\dagger)^{-1} \mathbf{y}_i \quad (3.23)$$

and

$$\mathbf{G}_i = \bar{\mathbf{A}}_i \mathbf{M} \bar{\mathbf{A}}_i^\dagger - \bar{\mathbf{A}}_i \mathbf{M} \mathbf{A}_i^\dagger (\mathbf{A}_i \mathbf{M} \mathbf{A}_i^\dagger)^{-1} \mathbf{A}_i \mathbf{M} \bar{\mathbf{A}}_i^\dagger. \quad (3.24)$$

After an E-step, an M-step is performed, corresponding to the maximization of the score function (3.18), namely the estimate of the parameters is updated according to

$$\boldsymbol{\theta}^{(h)} = \underset{\boldsymbol{\theta}: \mathbf{M}(\boldsymbol{\theta}) \in \mathcal{C}}{\operatorname{argmax}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(h-1)}). \quad (3.25)$$

The following proposition outlines the main features of the sequence of estimates.

Proposition 2. *Provided that $\mathbf{M}(\boldsymbol{\theta}^{(0)}) \succ \mathbf{0}$, $K \geq N$, and $\mathcal{C} = \mathcal{B} \cap \mathbb{H}_{++}^N$, with \mathcal{B} a closed set of \mathbb{H}^N , then*

- $\mathbf{M}(\boldsymbol{\theta}^{(h)}) \succ \mathbf{0}$, for all $h \geq 0$ and $\mathcal{L}_y(\mathbf{M}(\boldsymbol{\theta}^{(h)}) | \mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K)$ is a monotonically increasing sequence;
- if $\mathcal{B} \in \mathbb{H}_{++}^N$ is a closed set of positive definite matrices, then $\mathbf{M}(\boldsymbol{\theta}^{(h)})$, $h \geq 0$, is a bounded sequence and $\mathcal{L}_y(\mathbf{M}(\boldsymbol{\theta}^{(h)}) | \mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K)$, $h \geq 0$, converges to a finite value. Besides, supposing $\mathbf{M}(\boldsymbol{\theta})$ a norm coercive differentiable mapping, any limit point $\boldsymbol{\theta}^*$ to $\boldsymbol{\theta}^{(h)}$ is a B-stationary point³ [55, 85, 124, 115] to Problem (3.14).

³Substantially, a B-stationary point is any element of the feasible set with the prop-

Algorithm 4 EM Covariance Matrix Estimation Procedure

Require: $N, K, \mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K, \mathcal{C}, \boldsymbol{\theta}^{(0)}, \varepsilon_1, \varepsilon_2$.**Ensure:** EM solution $\hat{\boldsymbol{\theta}}$ to Problem (3.14).**Initialization** $h = 0;$ $P^{(0)} = \mathcal{L}_y(\mathbf{M}(\boldsymbol{\theta}^{(0)})|\mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K);$ **repeat**1: $h = h + 1;$ 2: E-Step: Compute $\boldsymbol{\Sigma}^{(h-1)}$ given by (3.19);3: M-Step: Find $\boldsymbol{\theta}^{(h)}$ using (3.25);4: Compute $P^{(h)}$ using (3.26);**until** $|P^{(h)} - P^{(h-1)}| > \varepsilon_1$ or $\|\boldsymbol{\theta}^{(h)} - \boldsymbol{\theta}^{(h-1)}\| > \varepsilon_2;$ Output $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}^{(h)}$.

Proof. See Appendix B.2. □

A summary of the EM procedure is reported in **Algorithm 4**, where, leveraging the results of Proposition 2, the exit condition of the procedure is set as $|P^{(h)} - P^{(h-1)}| \leq \varepsilon_1$ or $\|\boldsymbol{\theta}^{(h)} - \boldsymbol{\theta}^{(h-1)}\| \leq \varepsilon_2$, where $\varepsilon_1, \varepsilon_2 > 0$ and

$$P^{(h)} = \mathcal{L}_y(\mathbf{M}(\boldsymbol{\theta}^{(h)})|\mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K). \quad (3.26)$$

Remark 1. Before proceeding further, a useful digression on the convergence rate of **Algorithm 4** is now in order. As shown in [42], assuming that $\boldsymbol{\theta}^{(h)}$ converges to the ML estimate $\hat{\boldsymbol{\theta}}_{ML}$, then the rate of convergence is ruled by the spectral radius $\rho(\mathbf{r}^{EM})$ of the rate matrix

$$\mathbf{r}^{EM} = \mathbf{I} - \mathbf{F}_{obs}^{\frac{1}{2}} \mathbf{F}_{EM}^{-1} \mathbf{F}_{obs}^{\frac{1}{2}}, \quad (3.27)$$

where

$$\mathbf{F}_{obs} = -\nabla_{\boldsymbol{\theta}} \nabla_{\boldsymbol{\theta}}^T \mathcal{L}_y(\boldsymbol{\theta}) \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{ML}} \quad (3.28)$$

erty that along any limiting feasible direction the objective function is locally not increasing; please, see [55, 85, 124, 115] and Appendix B.2 for a rigorous definition.

is the observed information matrix and

$$\mathbf{F}_{EM} = \mathbb{E} \left[-\nabla_{\boldsymbol{\theta}} \nabla_{\boldsymbol{\theta}}^T \mathcal{L}_r(\boldsymbol{\theta}) | \mathbf{Y}, \boldsymbol{\theta} \right] \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{ML}} \quad (3.29)$$

is the expected complete information matrix. The interested reader may refer to Appendix B.3 for the computation of (3.28) and (3.29) with reference to (3.13) and (3.16).

Just to provide a study example, let us consider $N = 10$ and covariance matrix belonging to (3.5). Denoting by $\mathbf{v}(\theta_0)$ the steering vector in the direction of interest θ_0 , the true parameters are $\mathbf{V} = [\mathbf{v}(\theta_1), \mathbf{v}(\theta_2)]$ with $\theta_1 = -7^\circ$ and $\theta_2 = 15^\circ$, $\mathbf{S}_f = \mathbf{diag}(\lambda_1, \lambda_2)$ with $\lambda_1 = \lambda_2 = 10$, and $\sigma_n^2 = 1$. The presence of missing-data is emulated assuming that almost 37% of the snapshots undergoes information loss mechanisms. As to the missing-data pattern, the selection matrices obtained skipping the zero rows of

1. $\mathbf{diag}(\mathbf{1}_N - \mathbf{e}_1 - \mathbf{e}_3)$,
2. $\mathbf{diag}(\mathbf{1}_N - \mathbf{e}_2 - \mathbf{e}_5)$,
3. $\mathbf{diag}(\mathbf{1}_N - \mathbf{e}_4 - \mathbf{e}_7)$,
4. $\mathbf{diag}(\mathbf{1}_N - \mathbf{e}_6 - \mathbf{e}_8)$,
5. $\mathbf{diag}(\mathbf{1}_N - \mathbf{e}_9 - \mathbf{e}_{10})$,

are cyclically used (according to the reported order) to choose the observations at the different snapshots experiencing missing-data.

Figs. 3.1 (a) and (b) display the average convergence rate and the average number of iterations, respectively, required by **Algorithm 4** (with $\varepsilon_1 = \varepsilon_2 = 10^{-7}$ and initialized using the observed-data sample covariance matrix \mathbf{S}_y) to achieve convergence, versus the number of snapshots. Specifically, $\mathbf{S}_y = 1/K \sum_{i=1}^K \tilde{\mathbf{y}}_i \tilde{\mathbf{y}}_i^\dagger$, with $\tilde{\mathbf{y}}_i$ obtained by filling the missing values of \mathbf{y}_i with zero-elements, $i = 1, \dots, K$. The results rely on standard Monte Carlo counting techniques over 100 independent trials.

Inspection of the figures outlines that a lower value of $\rho(\mathbf{r}^{EM})$ is associated with a faster convergence of **Algorithm 4**.

In Fig. 3.1 (c), for a given trial, the distance between the ML estimate and the EM solution at the h -th M-step, i.e., $\|\boldsymbol{\theta}^{(h)} - \hat{\boldsymbol{\theta}}_{ML}\|$, is plotted versus the number of iterations, assuming $K = 40, 60, 80, 100$. This analysis

confirms that increasing the number of available snapshots the resulting estimate of **Algorithm 4** is closer and closer to $\hat{\boldsymbol{\theta}}_{ML}$. Besides, a larger sample support size K is connected with a smaller number of iterations required for **Algorithm 4** to converge, confirming the trend of Fig. 3.1 (b).

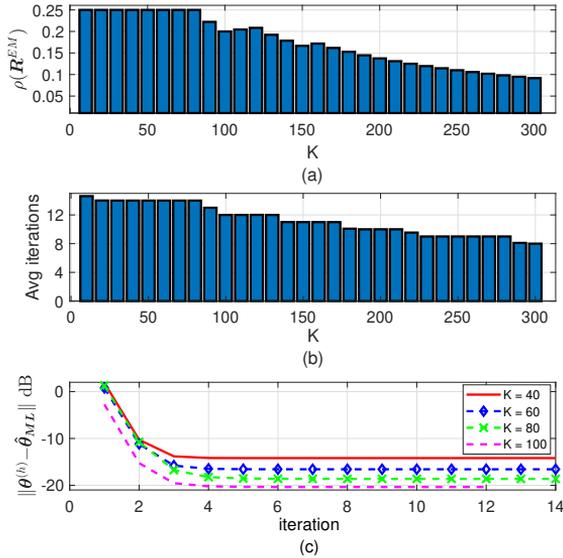


Figure 3.1. Convergence rate analysis for the case study discussed in the main text, with $N = 10$. Fig. (a) displays the average rate of convergence versus the number of snapshots, while Fig. (b) displays the average number of iterations versus the number of snapshots. The norm difference $\|\boldsymbol{\theta}^{(h)} - \hat{\boldsymbol{\theta}}_{ML}\|$ in dB versus the number of iterations for **Algorithm 4** is reported in Fig. (c), assuming $K = 40, 60, 80, 100$.

Remark 2. It is worth pointing out that the main advantage connected with the use of an EM algorithm occurs when the optimization involved in (3.25) is more tractable than the direct maximization of the observed-data likelihood (3.14). It is clear that the crucial point to devise an EM-based constrained covariance estimation procedure is the capability to obtain an optimal solution to (3.25) with an affordable computational effort. Besides, it is important to remark that different system constraints generally induce distinct feasible sets that generally result in different solutions $\boldsymbol{\theta}^{(h)}$. In particular, for the special case of unconstrained estima-

tion [105],

$$\hat{\mathbf{M}}(\boldsymbol{\theta}^{(h)}) = \boldsymbol{\Sigma}^{(h-1)} \quad (3.30)$$

is the maximizer of $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(h-1)})$ and therefore it constitutes the updated estimate $\boldsymbol{\theta}^{(h)}$.

In the following subsections, two well-known radar applications are analyzed in the missing-data scenario: adaptive beamforming and detection of the sources number. In particular, each application is presented and the underlying structured covariance matrix estimation problem is discussed. Then, EM-based solution methods, leveraging problem structure at different extents, are devised.

3.2.2 Adaptive Beamforming

Let us consider the Minimum Variance Distortionless Response (MVDR) (also known as the Capon) beamformer [105]

$$\mathbf{w} = \frac{\mathbf{M}^{-1}\mathbf{v}(\theta_0)}{\mathbf{v}(\theta_0)^\dagger \mathbf{M}^{-1}\mathbf{v}(\theta_0)}. \quad (3.31)$$

In a practical scenario the covariance matrix must be estimated from the incoming data leading to an adaptive weight vector. It is crystal clear that obtaining an accurate estimate of the unknown interference covariance matrix is a crucial task affecting the performance of the resulting adaptive beamformer. In a typical case where a set of $K \geq N$ secondary data $\{\mathbf{r}_i\}$, $i = 1, \dots, K$, is available, the unstructured ML estimate of \mathbf{M} is given by the SCM \mathbf{S} (often with a diagonal loading), defined as in (3.17) [105]. Therefore, \mathbf{S} (or possibly a diagonally loaded version) is employed in place of \mathbf{M} in (3.31), obtaining the MVDR adaptive beamformer.

Let us now focus on a missing-data context where the problem of computing the ML estimate of the covariance matrix from the observed-data is described in (3.12) and a viable estimation procedure is reported in **Algorithm 4**. As a consequence, following the same guideline as in the definition of the MVDR adaptive beamformer⁴, it is possible to gain adap-

⁴The analysis developed in the following can be also naturally extended to other kinds of beamformers.

tivity under the missing-data scenario using

$$\mathbf{w}_{EM} = \frac{\hat{\mathbf{M}}_{EM}^{-1} \mathbf{v}(\theta_0)}{\mathbf{v}(\theta_0)^\dagger \hat{\mathbf{M}}_{EM}^{-1} \mathbf{v}(\theta_0)}, \quad (3.32)$$

where $\hat{\mathbf{M}}_{EM}$ denotes the estimate of the covariance matrix obtained via the EM procedure described in **Algorithm 4**.

As highlighted in the previous subsection, tailored solutions to the M-step of **Algorithm 4** can be devised under the assumption of \mathbf{M} belonging to a specific covariance matrix uncertainty set. In this respect, some case studies are discussed in the following.

Unconstrained Estimation

The special case of unconstrained estimation has been described in the previous subsection and a solution to the resulting M-step is given by (3.30).

Constraint on the lower bound of the white noise power level

The Fast Maximum Likelihood (FML) procedure [101, 7] provides the M-step solution when \mathbf{M} belongs to the uncertainty set (3.2), i.e., a lower bound on the thermal noise power level is a-priori available. Specifically, denoting by $\mathbf{U}\mathbf{\Lambda}_\Sigma\mathbf{U}^\dagger$ the EigenValue Decomposition (EVD) of $\Sigma^{(h-1)}$ and by $\tilde{\lambda}_v$, $v = 1, \dots, N$ its eigenvalues, at the M-step update under the uncertainty set (3.2), is given by

$$\hat{\mathbf{M}}(\boldsymbol{\theta}^{(h)}) = \mathbf{U}\mathbf{\Lambda}_{FML}\mathbf{U}^\dagger, \quad (3.33)$$

with

$$\mathbf{\Lambda}_{FML} = \mathbf{diag}(\lambda_{1,FML}, \dots, \lambda_{N,FML}) \quad (3.34)$$

and $\lambda_{v,FML} = \max(\tilde{\lambda}_v, \sigma^2)$, $v = 1, \dots, N$.

Notably, this technique ensures that all the eigenvalues of $\hat{\mathbf{M}}(\boldsymbol{\theta}^{(h)})$ are greater than or equal to the lower bound on the power noise level. This is tantamount to projecting (according to the Frobenius norm) $\Sigma^{(h-1)}$ onto the set of the positive definite matrices greater than or equal to $\sigma^2 \mathbf{I}$ [7].

Centro-Hermitianity constraint

In many scenarios of practical interests (standard rectangular, hexagonal, uniform circular or cylindrical array), the covariance matrix exhibits a centro-Hermitian structure [105], which is equivalent to consider \mathbf{M} belonging to (3.6). Capitalizing on the problem structure, the M-step solution can be obtained using the Forward-Backward (FB) averaged sample covariance matrix procedure [105], resulting into

$$\hat{\mathbf{M}}(\boldsymbol{\theta}^{(h)}) = \boldsymbol{\Sigma}_{FB}, \quad (3.35)$$

where

$$\boldsymbol{\Sigma}_{FB} = \frac{1}{2}(\boldsymbol{\Sigma}^{(h-1)} + \mathbf{J}\boldsymbol{\Sigma}^{(h-1)*}\mathbf{J}). \quad (3.36)$$

Lower bound constraint on the white noise power level plus Centro-Hermitianity

This is tantamount to considering the uncertainty set characterizing the centro-Hermitian covariance matrices with a lower bound on the white disturbance power level

$$\mathcal{C} = \begin{cases} \mathbf{M} = \sigma_n^2 \mathbf{I} + \mathbf{r}_e \\ \mathbf{M} = \mathbf{J}\mathbf{M}^*\mathbf{J} \\ \mathbf{r}_e \succeq \mathbf{0} \\ \sigma_n^2 \geq \sigma^2 \end{cases}, \quad (3.37)$$

where \mathbf{r}_e , σ_n^2 and σ^2 are defined as in (3.2), whereas \mathbf{J} is given by (3.7). In this situation, denoting by $\mathbf{U}_{FB} \boldsymbol{\Lambda}_{FB} \mathbf{U}_{FB}^\dagger$ the EVD of $\boldsymbol{\Sigma}_{FB}$ defined in (3.36), with $\boldsymbol{\Lambda}_{FB} = \text{diag}(\tilde{\lambda}_{1,FB}, \dots, \tilde{\lambda}_{N,FB})$, it follows that the M-step update is now given by

$$\hat{\mathbf{M}}(\boldsymbol{\theta}^{(h)}) = \mathbf{U}_{FB} \boldsymbol{\Lambda}_{FML-FB} \mathbf{U}_{FB}^\dagger, \quad (3.38)$$

where

$$\boldsymbol{\Lambda}_{FML-FB} = \text{diag}(\lambda_{1,FML-FB}, \dots, \lambda_{N,FML-FB}), \quad (3.39)$$

with $\lambda_{v,FML-FB} = \max(\tilde{\lambda}_{v,FB}, \sigma^2)$, $v = 1, \dots, N$.

The overall procedure used to find the proposed adaptive Capon beam-

Algorithm 5 Adaptive beamforming in the context of missing-data

Input: $N, K, \mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K, \mathcal{C}, \boldsymbol{\theta}^{(0)}, \varepsilon_1, \varepsilon_2$.

Output: EM-based adaptive beamformer \mathbf{w}_{EM} .

- 1: find $\hat{\mathbf{M}}_{EM}$ via **Algorithm 4** using the appropriate bespoke solution (3.30), (3.33), (3.35), (3.38) to the M-step;
 - 2: compute \mathbf{w}_{EM} using (3.32);
 - 3: output \mathbf{w}_{EM} .
-

former in the context of missing-data is summarized in **Algorithm 5**.

3.2.3 Detection of Number of Sources

Let us consider d uncorrelated narrow-band sources impinging the array from distinct directions $\{\theta_s\}, s = 1, \dots, d < N$. After amplification, down-conversion, and digital sampling, the i -th received complete spatial snapshot \mathbf{r}_i is given by

$$\mathbf{r}_i = \mathbf{V} \mathbf{s}_i + \mathbf{n}_i, \quad i = 1, \dots, K, \quad (3.40)$$

where $\mathbf{V} = [\mathbf{v}(\theta_1), \mathbf{v}(\theta_2), \dots, \mathbf{v}(\theta_d)] \in \mathbb{C}^{N \times d}$ is the array manifold matrix (assumed full-rank), $\mathbf{s}_i, i = 1, \dots, K$, are IID zero-mean Gaussian random vectors of sources amplitudes (independent of each other) with powers $\sigma_s^2, s = 1, \dots, d$, respectively, and \mathbf{n}_i are IID zero-mean circularly symmetric Gaussian random vectors with power σ_n^2 , assumed statistically independent from the sources.

For the case at hand, the covariance matrix of the received signal can be assumed belonging to (3.5). Resorting to the EVD, the complete-data covariance matrix takes on the convenient form

$$\mathbf{M} = \sum_{v=1}^d \lambda_v \boldsymbol{\Phi}_v \boldsymbol{\Phi}_v^\dagger + \sum_{v=d+1}^N \lambda_v \boldsymbol{\Phi}_v \boldsymbol{\Phi}_v^\dagger, \quad v = 1, \dots, N, \quad (3.41)$$

where λ_v and $\boldsymbol{\Phi}_v, v = 1, \dots, N$, denote the eigenvalues and the corresponding eigenvectors of \mathbf{M} , respectively, with $\lambda_1 \geq \lambda_2 \geq \dots \lambda_d \geq \lambda_{d+1} =$

$\lambda_{d+2} = \dots \lambda_N = \sigma_n^2$. As a consequence, denoting by $\Phi_{v,R}^T$ and $\Phi_{v,I}^T$ the vectors of the real and imaginary components of Φ_v , for a given d , the vector of the unknown parameters (underlying the covariance structure) is in one-to-one mapping with [105, p. 831]

$$\theta_d = [\lambda_1, \dots, \lambda_d, \sigma_n^2, \Phi_{1,R}^T, \Phi_{1,I}^T, \dots, \Phi_{d,R}^T, \Phi_{d,I}^T]^T, \quad (3.42)$$

which explicitly reveals the role of d in controlling the degrees of freedom of the covariance matrix, where $U_d = [\Phi_{1,R} + j\Phi_{1,I}, \dots, \Phi_{d,R} + j\Phi_{d,I}]$ is such that $U_d^\dagger U_d = I$.

The approach pursued in this subsection follows from [105, 3, 109, 50], where a source number detection algorithm, based on a data-adaptive test statistic plus a penalty function related to the degrees of freedom, is devised. Specifically, denoting by

$$\hat{\theta}_{d,ML} = [\hat{\lambda}_1, \dots, \hat{\lambda}_d, \hat{\sigma}_n^2, \hat{\Phi}_{1,R}^T, \hat{\Phi}_{1,I}^T, \dots, \hat{\Phi}_{d,R}^T, \hat{\Phi}_{d,I}^T]^T \quad (3.43)$$

the ML estimate of θ_d , the problem of detecting the number of sources can be formulated as

$$\hat{d} = \underset{d=0, \dots, K_1}{\operatorname{argmax}} \mathcal{L}_r(\hat{\theta}_{d,ML}) - T(d), \quad (3.44)$$

where $K_1 \leq N - 1$ is an upper bound to the number of sources, $\mathcal{L}_r(\hat{\theta}_{d,ML})$ is the statistic given by the complete-data log-likelihood (3.16) evaluated at $\hat{\theta}_{d,ML}$, and $T(d)$ is a penalty term accounting for the number of free parameters in the assumed model. In particular, taking the negative value of $\mathcal{L}_r(\hat{\theta}_{d,ML})$ and dropping the terms functionally independent from d , the following decision statistic is obtained [4]

$$L(d, \hat{\lambda}_1, \dots, \hat{\lambda}_N) = K(N - d) \ln \left\{ \frac{\frac{1}{N-d} \sum_{v=d+1}^N \hat{\lambda}_v}{\left(\prod_{v=d+1}^N \hat{\lambda}_v \right)^{\frac{1}{N-d}}} \right\}. \quad (3.45)$$

Exploiting the above result, problem (3.44) is equivalently recast as⁵

$$\hat{d} = \underset{d=0,\dots,K_1}{\operatorname{argmin}} L(d, \hat{\lambda}_1, \dots, \hat{\lambda}_N) + p(d), \quad (3.46)$$

where $p(d)$ is a specific penalty function. In the following, three detection tests, Akaike Information Criterion (AIC) [3], Minimum Description Length (MDL) [109], and Hannan–Quinn Information Criterion (HQC) [50], are considered. Each test is characterized by a different penalty function $p(d)$; in particular

$$p(d) = \begin{cases} d(2N - d), & \text{AIC} \\ 1/2 [d(2N - d) + 1] \ln K, & \text{MDL} \\ [d(2N - d) + 1] \ln(\ln(K)) & \text{HQC} \end{cases}. \quad (3.47)$$

Let us now frame the decision statistic in missing-data context. Accordingly, the criterion (3.44) can be modified as:

$$\hat{d} = \underset{d=0,\dots,K_1}{\operatorname{argmax}} \mathcal{L}_y(\hat{\boldsymbol{\theta}}_{d,ML} | \mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K) - T(d). \quad (3.48)$$

This requires, for a given d , the computation of the ML estimate $\hat{\boldsymbol{\theta}}_{d,ML}$ from the observed-data. To this end, a viable technique is represented by **Algorithm 4** applied to a covariance uncertainty set including the fixed rank constraint in (3.5). Two relevant case studies are thus developed in the following, providing tailored solutions to the M-step.

Fixed rank constraint

Let us exploit the knowledge that \mathbf{M} belongs to (3.5). Specifically, for a given d , the M-step at the h -th iteration is cast as

$$\boldsymbol{\theta}_d^{(h)} = \underset{\boldsymbol{\theta}_d}{\operatorname{argmax}} Q(\boldsymbol{\theta}_d, \boldsymbol{\theta}_d^{(h-1)}), \quad (3.49)$$

⁵It is also worth pointing out that (3.46) can be generalized to the case of covariance matrices with additional structured constraints.

where $\boldsymbol{\theta}_d$ is defined as in (3.42). The maximizer of Problem (3.49) is given by [105]

$$\boldsymbol{\theta}_d^{(h)} = [\tilde{\lambda}_1, \dots, \tilde{\lambda}_d, \tilde{\sigma}_n^2, \tilde{\boldsymbol{\Phi}}_{1,R}, \tilde{\boldsymbol{\Phi}}_{1,I}, \dots, \tilde{\boldsymbol{\Phi}}_{d,R}, \tilde{\boldsymbol{\Phi}}_{d,I}]^T, \quad (3.50)$$

where $\tilde{\lambda}_v$ and $\tilde{\boldsymbol{\Phi}}_v$, $v = 1, \dots, d$ are the d greatest eigenvalues and the corresponding eigenvectors of $\boldsymbol{\Sigma}^{(h-1)}$, with $\tilde{\boldsymbol{\Phi}}_{v,R}$, and $\tilde{\boldsymbol{\Phi}}_{v,I}$ the real and imaginary components of $\tilde{\boldsymbol{\Phi}}_v$, whereas

$$\tilde{\sigma}_n^2 = \frac{1}{N-d} \sum_{v=N-d}^N \tilde{\lambda}_v \quad (3.51)$$

is the arithmetic mean of the $N-d$ smallest eigenvalues of $\boldsymbol{\Sigma}^{(h-1)}$.

Exploiting the above results, the h -th estimate of the covariance matrix is given by

$$\hat{\boldsymbol{M}}(\boldsymbol{\theta}_d^{(h)}) = \boldsymbol{U} \boldsymbol{\Lambda}_S \boldsymbol{U}^\dagger + \tilde{\sigma}_n^2 \boldsymbol{I}, \quad (3.52)$$

where

$$\boldsymbol{\Lambda}_S = \mathbf{diag}(\tilde{\lambda}_1 - \tilde{\sigma}_n^2, \dots, \tilde{\lambda}_d - \tilde{\sigma}_n^2), \quad (3.53)$$

and

$$\boldsymbol{U} = [\tilde{\boldsymbol{\Phi}}_1, \dots, \tilde{\boldsymbol{\Phi}}_d]. \quad (3.54)$$

Hence, taking the negative value and dropping the constant terms of the observed-data log-likelihood, the order estimate is given by

$$\hat{d}_{EM} = \underset{d=0, \dots, K_1}{\operatorname{argmin}} L_y(\hat{\boldsymbol{\theta}}_d) + p(d), \quad (3.55)$$

where

$$L_y(\hat{\boldsymbol{\theta}}_d) = \sum_{i=1}^K \ln(\det(\boldsymbol{A}_i \hat{\boldsymbol{M}}(\hat{\boldsymbol{\theta}}_d) \boldsymbol{A}_i^\dagger)) + \operatorname{tr}\{(\boldsymbol{A}_i \hat{\boldsymbol{M}}(\hat{\boldsymbol{\theta}}_d) \boldsymbol{A}_i^\dagger)^{-1} \boldsymbol{y}_i \boldsymbol{y}_i^\dagger\} \quad (3.56)$$

with $\hat{\boldsymbol{\theta}}_d$ the final estimate of **Algorithm 4** and $p(d)$ a specific penalty function (3.47) related to the **AIC**, **MDL** or **HQC** tests.

The overall procedure to find the sources number in the context of missing-data is summarized in **Algorithm 6**.

Algorithm 6 Detection of number of sources in the context of missing-data and fixed rank constraint

Input: $N, K, \mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K, \boldsymbol{\theta}^{(0)}, K_1, p(d), \varepsilon_1, \varepsilon_2$.

Output: A solution \hat{d}_{EM} to Problem (3.48).

1: for $\tilde{d} = 0, \dots, K_1$ do

a) compute the estimate $\hat{\boldsymbol{\theta}}_{\tilde{d}}$ via **Algorithm 4** using (3.50) as solution to the M-step with $d = \tilde{d}$;

b) compute $L_y(\hat{\boldsymbol{\theta}}_{\tilde{d}})$ in (3.56) using the estimate $\hat{\boldsymbol{\theta}}_{\tilde{d}}$.

end for

2: evaluate

$$\hat{d}_{EM} = \underset{\tilde{d}=0, \dots, K_1}{\operatorname{argmin}} L_y(\hat{\boldsymbol{\theta}}_{\tilde{d}}) + p(\tilde{d});$$

3: output \hat{d}_{EM} .

Rank constraint and centro-Hermitianity

Let us assume that the covariance matrix belongs to both the uncertainty sets (3.5) and (3.6), i.e.

$$\mathcal{C} = \begin{cases} \mathbf{M} = \sigma_n^2 \mathbf{I} + \mathbf{V} \mathbf{S}_f \mathbf{V}^\dagger \\ \mathbf{M} = \mathbf{J} \mathbf{M}^* \mathbf{J} \\ \mathbf{V} \mathbf{S}_f \mathbf{V}^\dagger \succeq \mathbf{0} \\ \operatorname{Rank}(\mathbf{V} \mathbf{S}_f \mathbf{V}^\dagger) \leq d \\ \sigma_n^2 > 0 \end{cases}, \quad (3.57)$$

where \mathbf{V} , \mathbf{S}_f , d , and σ_n^2 are defined as in (3.5), whereas \mathbf{J} is given by (3.7).

Therefore, for a given d , the maximizer of $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(h-1)})$ is given by [57]

$$\boldsymbol{\theta}_{d,FB}^{(h)} = [\tilde{\lambda}_{1,FB}, \dots, \tilde{\lambda}_{d,FB}, \tilde{\sigma}_{n,FB}^2, \tilde{\boldsymbol{\Phi}}_{1,FB,R}, \tilde{\boldsymbol{\Phi}}_{1,FB,I}, \dots, \tilde{\boldsymbol{\Phi}}_{d,FB,R}, \tilde{\boldsymbol{\Phi}}_{d,FB,I}]^\top, \quad (3.58)$$

where $\tilde{\lambda}_{v,FB}$ and $\tilde{\boldsymbol{\Phi}}_{v,FB}$, $v = 1, \dots, d$ are the d greatest eigenvalues and the corresponding eigenvectors of $\boldsymbol{\Sigma}_{FB}$, defined as in (3.36), with $\tilde{\boldsymbol{\Phi}}_{v,FB,R}$

and $\tilde{\Phi}_{v,FB,I}$ the real and imaginary components of $\tilde{\Phi}_{v,FB}$, respectively, and

$$\tilde{\sigma}_{n,FB}^2 = \frac{1}{N-d} \sum_{v=N-d}^N \tilde{\lambda}_{v,FB} \quad (3.59)$$

is the arithmetic mean of the $N-d$ smallest eigenvalues of Σ_{FB} . As a consequence,

$$\hat{M}(\hat{\theta}_{d,FB}) = \sum_{v=1}^d (\tilde{\lambda}_{v,FB} - \tilde{\sigma}_{n,FB}^2) \tilde{\Phi}_{v,FB} \tilde{\Phi}_{v,FB}^\dagger + \tilde{\sigma}_{n,FB}^2 \mathbf{I} \quad (3.60)$$

with $\hat{\theta}_{d,FB}$ the resulting estimate of **Algorithm 4**.

Along the same line as the previous case, the statistic is computed for each possible d , to get the order estimate

$$\hat{d}_{EM-FB} = \underset{d=0,\dots,K_1}{\operatorname{argmin}} L_y(\hat{\theta}_{d,FB}) + \frac{1}{2}p(d), \quad (3.61)$$

where $L_y(\hat{\theta}_{d,FB})$ is given by (3.56) evaluated in correspondence of the estimate (3.60) and $p(d)$ is one of the penalty functions in (3.47) [117].

3.3 Performance Analysis

In this section, the performance of the proposed estimation strategy, framed in the context of adaptive beamforming and detection of number of sources, is analyzed. For both applications it is considered a radar system equipped with a **ULA** pointing in the bore-sight direction ($\theta_0 = 0$). The array is composed of $N = 20$ antennas with inter-element spacing $d_x = \lambda/2$, where λ denotes the radar operating wavelength. Moreover, two different values for the probability p_m of missing an observation are considered, i.e., $p_m = 0.1$ or $p_m = 0.3$. For a given p_m , the selection matrix A_i of the i -th snapshot is constructed from the diagonal matrix D_i whose diagonal entries are statistically **IID** Bernoulli random variables with parameter $1 - p_m$, skipping rows containing all zeros. Besides, the computation of the observed-data **SCM** $\mathbf{S}_y = 1/K \sum_{i=1}^K \tilde{\mathbf{y}}_i \tilde{\mathbf{y}}_i^\dagger$ is performed employing $\tilde{\mathbf{y}}_i = D_i \mathbf{r}_i$, $i = 1, \dots, K$.

3.3.1 Adaptive Beamforming

The performance of the adaptive beamformer is analyzed in terms of beampattern shape and normalized average Signal to Interference Power Ratio (S/I) versus the number of snapshots. Standard Monte Carlo counting techniques over 100 independent trials to compute the former performance metric and 500 independent trials for the latter are used.

In the reported case studies the disturbance covariance matrix is modeled as $\mathbf{M} = \mathbf{M}_J + \sigma_a^2 \mathbf{I}$, where σ_a^2 is the white noise power level (assumed without loss of generality equal to 0 dB) and \mathbf{M}_J is the jamming covariance contribution. Specifically, denoting by J_{NB} and J_{WB} the number of narrow-band and wide-band jammers (assumed separated in space), $\mathbf{M}_J = \mathbf{M}_1 + \mathbf{M}_2$, where [45]

$$\mathbf{M}_1 = \sum_{l=1}^{J_{NB}} \sigma_l^2 \mathbf{v}(\theta_l) \mathbf{v}(\theta_l)^\dagger, \quad (3.62)$$

with

$$\mathbf{v}(\theta_l) = [1, e^{j\frac{2\pi}{\lambda} d_x \sin(\theta_l)}, \dots, e^{j(N-1)\frac{2\pi}{\lambda} d_x \sin(\theta_l)}]^\text{T} \in \mathbb{C}^N \quad (3.63)$$

the steering vector in the direction θ_l of the l -th jammer and σ_l^2 the power of the l -th jammer, while

$$\mathbf{M}_2(n, m) = \sum_{r=1}^{J_{WB}} \bar{\sigma}_r^2 \text{sinc}[0.5B_{f_r} (n - m)\zeta_r] e^{j(n-m)\zeta_r}, \quad (3.64)$$

with $(n, m) \in \{0, \dots, N - 1\}^2$ and $\zeta_r = \pi \sin \theta_r$; moreover in (3.64), $\bar{\sigma}_r^2$, θ_r , and B_{f_r} , represent the power, the DOA, and the fractional bandwidth B_r/f_0 (with B_r the actual bandwidth and f_0 the carrier frequency) associated with the r -th interferer. The sinc function appearing in (3.64) is defined as $\text{sinc}(x) = \sin(x)/(x)$.

In the following, two different interfering environments are analyzed:

- Scenario 1: five narrow-band jammers located at $\theta_l = 10 + 10l$ degrees, $l = 1, \dots, 5$ with JNR given by $JNR_l = 30$ dB ($\sigma_l^2 = JNR_l \sigma_a^2$, $l = 1, \dots, 5$).

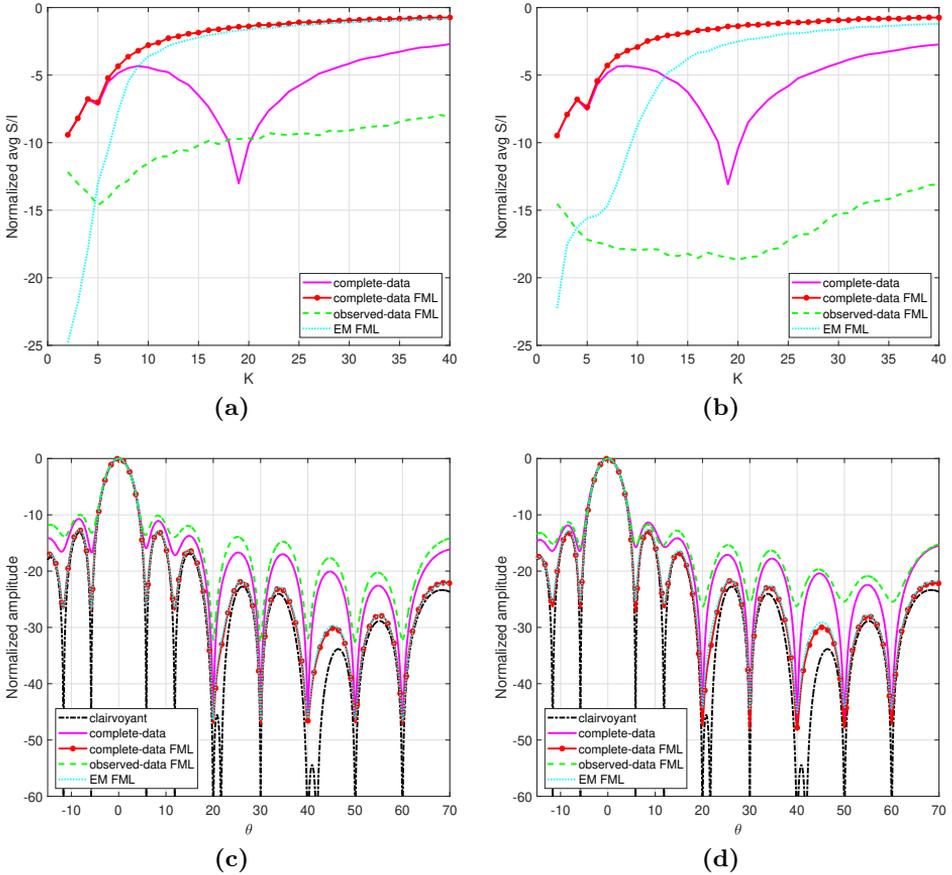


Figure 3.2. Adaptive beamformer performance for a ULA with 20 antennas in Scenario 1. Figs. (a) and (c) consider $p_m = 0.1$ while Figs. (b) and (d) consider $p_m = 0.3$. Figs. (a) and (b) display the normalized average S/I versus number of snapshots, while Figs. (c) and (d) display the resulting beampattern with $K = 60$ (therein, the red-Xs along the θ -axis denote the sources directions).

- Scenario 2: five wide-band jammers ($B_f = 0.03$) located at $\theta_r = 10 + 10r$ degrees, $r = 1, \dots, 5$ with $JNR_r = 30$ dB ($\sigma_j^2 = JNR_r \sigma_a^2$, $r = 1, \dots, 5$).

The performance of the adaptive beamformer, assuming either $p_m =$

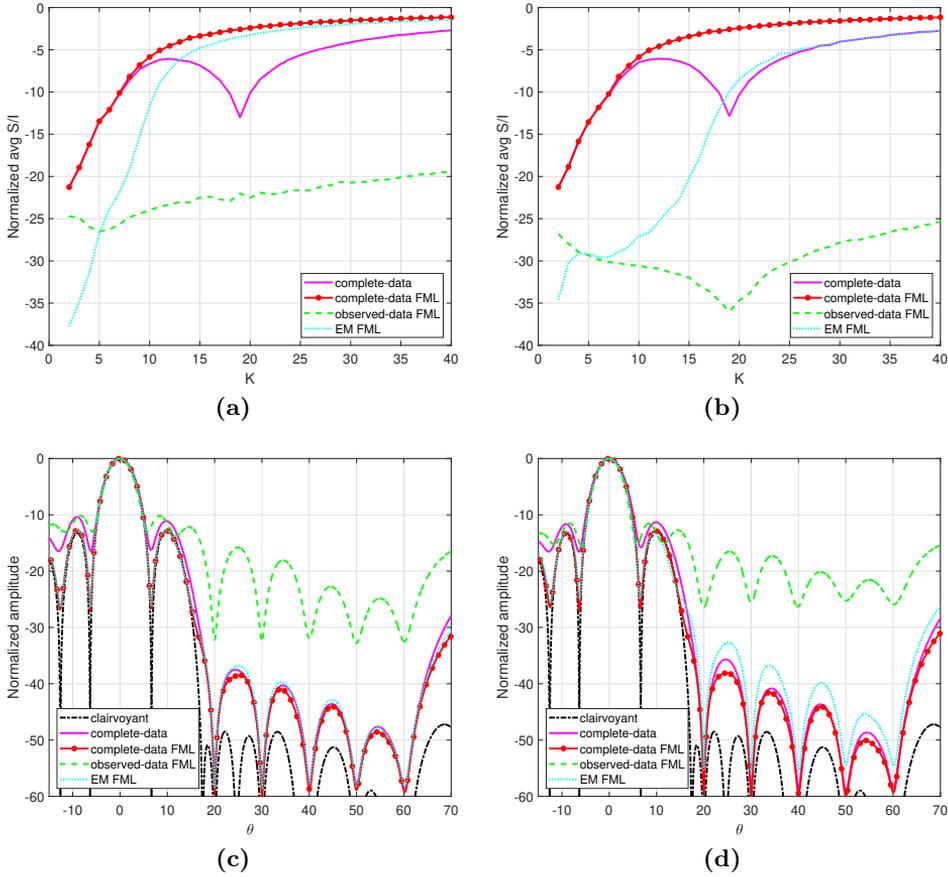


Figure 3.3. Adaptive beamformer performance for a ULA with 20 antennas in Scenario 2. Figs. (a) and (c) consider $p_m = 0.1$ while Figs. (b) and (d) consider $p_m = 0.3$. Figs. (a) and (b) display the normalized average S/I versus number of snapshots, while Figs. (c) and (d) display the resulting beampattern with $K = 60$ (therein, the red-Xs along the θ -axis denote the sources directions).

0.1 or $p_m = 0.3$, is analyzed in terms of normalized average S/I in Figs. 3.2(a), 3.2(b), 3.3(a), and 3.3(b). The resulting beampatterns (assuming $K = 60$), are displayed in Figs. 3.2(c), 3.2(d), 3.3(c), and 3.3(d). In particular, Figs. 3.2 and 3.3 refer to the interference environments of Scenario 1 and

2, respectively.

The proposed strategy employs the **EM** procedure assuming the uncertainty set (3.2) with the **FML** computed from \mathbf{S}_y , used to initialize the **EM** procedure. The beampattern and the normalized average **S/I** obtained using the **SCM** of the complete-data (as well as its variant based on **FML**) and the **FML** of \mathbf{S}_y , are considered for comparison. As performance benchmark, the clairvoyant beampattern, based on a perfect knowledge of the covariance matrix, is reported too. A close inspection of the results under the interference environment of Scenario 1 shows that for $p_m = 0.1$ and $K \geq N$ the performance of the proposed procedure comes closer and closer to the complete-data **FML** whereas it exhibits for $p_m = 0.3$ a slight degradation in terms of normalized average **S/I** in the order of 1 dB for $K > N$, with respect to the complete-data benchmark. The effectiveness of the proposed algorithm is also confirmed by the more challenging interference environment of Scenario 2, where the performance is very close to the complete-data **FML** for $p_m = 0.1$ and experiences a maximum degradation, in terms of normalized average **S/I**, lower than 6 dB, for $p_m = 0.3$ and $K \geq N$. Nevertheless, for all the configurations, the **S/I** of the **EM**-based beampattern approaches the complete-data performance as K increases and this represents an indirect proof that the quality of the proved covariance estimation procedure improves when more and more snapshots are available for the estimation process.

As to the beampattern analysis, the inspection of the figures reveals that the **EM FML** is able to correctly nullify the jammers while preserving low side-lobes levels.

Finally, Fig. 3.4 compares the performance of **EM FML** and **EM FML-FB**, highlighting the capability of **FML-FB** to benefit from the underlying structure of the covariance matrix.

3.3.2 Detection of Number of Sources

In the following, equal-power signals impinging on the array from different directions θ_v are considered. The values of the parameters involved in the three analyzed scenarios, each related to a different number of sources, are listed in Table 3.1.

Specifically, $\text{SSBW} = 0.891/N$ denotes the 3dB Single-Side Beam-Width (**SSBW**) of the considered **ULA** [112], whereas $u_v = \sin(\theta_v)$ is the

Table 3.1. Simulation Parameters

d	$u_1 = \sin(\theta_1)$	$u_2 = \sin(\theta_2)$	$u_3 = \sin(\theta_3)$	$u_4 = \sin(\theta_4)$
2	-1/2 SSBW	1/2 SSBW		
3	-1/2 SSBW	1/2 SSBW	3/2 SSBW	
4	-1/2 SSBW	1/2 SSBW	3/2 SSBW	-3/2 SSBW

target angular location of the v -th source in the space of directional cosine [105]. Therefore, the covariance matrix is modeled as $\mathbf{M} = \mathbf{M}_S + \sigma_n^2 \mathbf{I}$, where σ_n^2 is the white noise power level (assumed without loss of generality equal to 0 dB) and \mathbf{M}_S refers to the useful covariance contribution, given by

$$\mathbf{M}_S = \sigma_s^2 \sum_{v=1}^d \mathbf{v}(\theta_v) \mathbf{v}(\theta_v)^\dagger, \quad (3.65)$$

with σ_s^2 the power of each signal of interest and $\mathbf{v}(\theta_v)$ is defined as in (3.63).

The metric used to assess the detection performance is the P_D , namely the probability that $\hat{d} = d$ [105], which is estimated via standard Monte Carlo counting techniques over 500 independent trials⁶. Moreover, the Array Signal-to-Noise Ratio (ASNR) is defined as

$$\text{ASNR} = N \frac{\sigma_s^2}{\sigma_n^2}. \quad (3.66)$$

Finally, the detection algorithm assumes $K = 100$ and a maximum number of sources equal to $N/2 = 10$.

The detection performance is reported in Fig. 3.5 assuming $p_m \in \{0.1, 0.3\}$ and $K = 100$. In particular, denoting by d the actual number of sources, Figs. 3.5 (a), (b), and (c) assume $d = 2$, Figs. 3.5 (d), (e), and (f) $d = 3$, whereas Figs. 3.5 (g), (h), and (i) $d = 4$. Moreover, Figs. 3.5 (a), (d), and (g) refer to AIC, Figs. 3.5 (b), (e), and (h) consider MDL whereas Figs. 3.5 (c), (f), and (i) display HQC. The proposed strategy employs the EM procedure assuming the uncertainty set (3.5) and consid-

⁶Notice that a rank-deficient \mathbf{S}_y , due to a possible selection matrix configuration, is a non-zero probability event. Such realizations are excluded from the Monte Carlo trials.

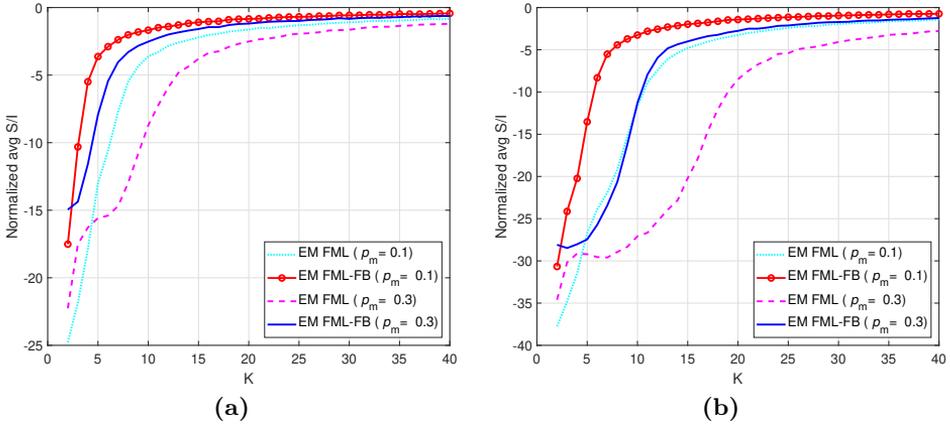


Figure 3.4. Normalized average S/I versus number of snapshots for a ULA with 20 antennas. Fig. (a) considers Scenario 1 while Fig. (b) Scenario 2.

ering as a starting guess, at each trial, the **SCM** of $2N$ **IID** white zero-mean circularly symmetric Gaussian random vectors of size N .

The results highlight that for $p_m = 0.1$ the **EM** approach leads to a performance very close to the complete-data case (with a loss smaller than 1 dB), and outperforms the basic approach of replacing the missing observations in the complete-data with zeros (dashed brown curves), in most of the analyzed case. In fact, a close inspection of the curves shows that only when $d = 4$, low **ASNR**, and with reference to the **AIC** (Fig. 3.5 (g)), the basic approach performs better than **EM**-based technique. This results is not surprising due to the overestimation behavior of the **AIC** [105]. Besides, the basic strategy may not provide a monotonic behaviour with respect to the **ASNR**, reflecting the reasonable larger and larger discrepancy between the actual covariance matrix and that heuristically computed.

As expected, the **EM**-based order selection procedure experiences a performance degradation at $p_m = 0.3$, as compared with the complete-data counterpart. Remarkably, the gap between the **EM** and the complete-data curves, for $p_m = 0.3$, is less than 3 dB in the worst case, whereas at the high **ASNR** regime it is almost absent. As in the case $p_m = 0.1$, **EM**-based strategy outperforms the basic counterpart, with the only exception of **AIC** with 4 sources, reported in Fig. 3.5 (g).

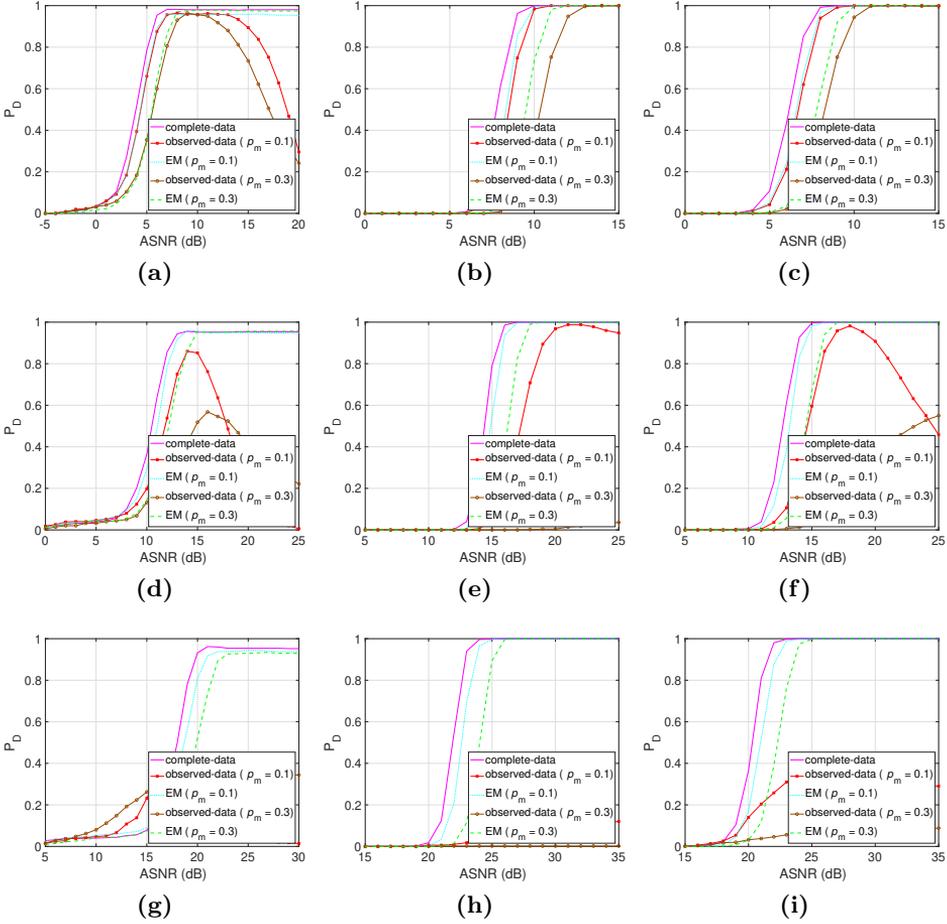


Figure 3.5. Detection performance for a ULA with 20 antennas assuming $K = 100$ and $p_m \in \{0.1, 0.3\}$. Figs. (a), (b), and (c) assume $d = 2$, Figs. (d), (e), and (f) assume $d = 3$, whereas Figs. (g), (h), and (i) assume $d = 4$ equal-power signals impinging the array, respectively, with signal separation corresponding to $0.891/N$. Moreover, Figs. (a), (d), and (g) consider AIC, Figs. (b), (e), and (h) consider MDL, whereas Figs. (c), (f), and (i) consider HQC.

Finally, the detection performance using **EM** and **EM-FB** (both initialized, at each trial, with the **SCM** of $2N$ IID white zero-mean circularly

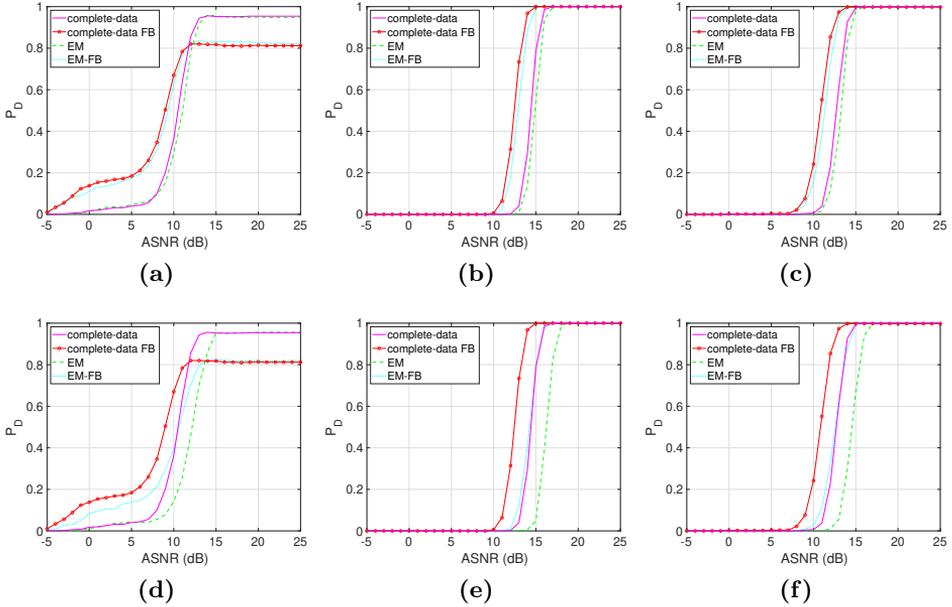


Figure 3.6. Comparison of the P_D using EM and EM-FB estimation strategies for a ULA with 20 antennas assuming 3 equal-power signals impinging the array with signal separation corresponding to $0.891/N$. Figs. (a), (b), (c) consider $p_m = 0.1$, whereas $p_m = 0.3$ is assumed in Figs. (d), (e), (f). Besides, Figs. (a) and (d), Figs. (b) and (e), and Figs. (c) and (f) consider AIC, MDL, and HQC, respectively.

symmetric Gaussian random vectors of size N) is compared in Fig. 3.6. Inspection of the curves pinpoints that capitalizing on the centro-Hermitian structure, **EM-FB** achieves higher P_D levels than the unstructured **EM** in all the considered scenarios, except for the **AIC** at high **ASNR** regime where an expected saturation is experienced [105].

3.4 Conclusion

This chapter has considered the problem of structured covariance matrix estimation in the presence of missing-(complex) data with special attention to a radar signal processing background. After providing a sub-

stantial motivation on the study and specifying some constraint sets of particular interest for the covariance matrix, the missing-data model is described assuming Gaussian observations. Hence, the ML covariance estimation problem is formulated as the maximization of the observed-data log-likelihood. To circumvent the analytical difficulties which are usually connected with the direct optimization of the mentioned function, an iterative maximization procedure based on the EM algorithm is developed and its convergence properties are established. Besides, a closed-form expression is computed for the convergence rate. The theoretical results are capitalized for some specific structural covariance models with reference to two radar applications: adaptive beamforming and detection of the number of sources. General procedures are suggested to construct adaptive beamformers and to detect the number of active sources in a collection of snapshots when missing observations are present. At the analysis stage, extensive numerical results have been discussed to show the effectiveness of the bespoke strategies to handle missing-data scenarios.

Chapter 4

Adaptive Radar Detection in the Presence of Missing-Data

This chapter¹ addresses the problem of adaptive radar detection in a context with missing-data where the complete observations (i.e., downstream information loss mechanisms) are characterized by homogeneous Gaussian disturbance with an unknown but possibly structured covariance matrix. The detection problem, formulated as a composite hypothesis test, is tackled by resorting to sub-optimal design strategies, leveraging the [GLR](#) criterion demanding appropriate [MLEs](#) of the unknowns under both hypotheses. Capitalizing on some possible a-priori knowledge about the interference covariance matrix structure, the optimization problems involved in the [MLE](#) computation are handled by employing the [EM](#) algorithm or its [ECM](#) and [M-EM](#) variants. At the analysis stage, the performance of the devised architectures is assessed both via Monte Carlo simulations and on measured data for some covariance matrix structures of practical interest.

The rest of the chapter is organized as follows. The data model and target detection problem with missing-data are presented in Section [4.1](#). In Section [4.2](#), the detection problem is addressed resorting to sub-optimal design criteria, i.e., one-step and two-step [GLR](#), which demand the opti-

¹©2022 IEEE. Reprinted, with permission, from A. Aubry, V. Carotenuto, A. De Maio, M. Rosamilia and S. Marano, “Adaptive Radar Detection in the Presence of Missing-data,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 58, no. 4, pp. 3283-3296, Aug. 2022.

mization of appropriate observed-data likelihood functions under the two hypotheses. Hence, in Section 4.2, an EM-based framework is also devised to tackle the resulting optimization problems and derive practical detectors. The performance of the mentioned detectors is analyzed in Section 4.3, whereas conclusions are drawn in Section 4.4.

4.1 Problem Formulation

Let us consider a radar system collecting spatial data via a linear array composed of N antennas and operating in the presence of noise and interference, with unknown spectral characteristics.

Under the ideal conditions of complete access to the set of space-time observations, the problem of detecting a prospective target located at range R and elevation θ_0 with respect to the array boresight (under the narrow-band radar probing signal assumption), can be formulated as the following composite binary hypothesis testing problem

$$\begin{cases} H_0 : \begin{cases} \mathbf{r} = \mathbf{n} \\ \mathbf{r}_i = \mathbf{n}_i, \quad i = 1, \dots, K \end{cases} \\ H_1 : \begin{cases} \mathbf{r} = \alpha \mathbf{p} + \mathbf{n} \\ \mathbf{r}_i = \mathbf{n}_i, \quad i = 1, \dots, K \end{cases} \end{cases} \quad (4.1)$$

where \mathbf{r} is the primary data, α is an unknown complex parameter which accounts for the target reflectivity and the channel propagation effects, whereas \mathbf{p} denotes the spatial steering vector evaluated at θ_0 , which is assumed known at the design stage. Besides, a set of secondary data \mathbf{r}_i , $i = 1, \dots, K$, free of the useful signal and with the same covariance matrix as the primary data (homogeneous environment) [60, 93, 14, 101], is supposed available. The interference plus noise components \mathbf{n} and \mathbf{n}_i , $i = 1, \dots, K$, are modeled as IID zero-mean circularly symmetric Gaussian random vectors, with unknown (but possibly structured) covariance matrix given by

$$\mathbf{M}(\boldsymbol{\theta}) = \mathbb{E}[\mathbf{n}\mathbf{n}^\dagger] = \mathbb{E}[\mathbf{n}_i\mathbf{n}_i^\dagger], \quad i = 1, \dots, K \quad (4.2)$$

where $\boldsymbol{\theta}$ denotes the vector of the unknowns parameterizing the structure of \mathbf{M} .

Let us now frame the detection problem in a context with missing-data caused by random failures of some array elements [123, 68, 111, 116] or possible transmission-reception faults experienced by distributed radar systems [51] or wirelessly networked aperiodic digital phased array radars [46], etc. For the case at hand, the observed primary data is modeled as

$$\mathbf{z} = \mathbf{A}\mathbf{r} \quad (4.3)$$

where \mathbf{A} is an appropriate $p \times N$ selection matrix; specifically, denoting by $\kappa_1, \kappa_2, \dots, \kappa_{N-p} \in \{1, \dots, N\}$ the indices of the channels where a missing-data occurs in the snapshot from the CUT, \mathbf{A} is obtained from the $N \times N$ identity matrix, by deleting the rows indexed by $(\kappa_1, \kappa_2, \dots, \kappa_{N-p})$. Similarly, each secondary observed snapshot can be modeled as

$$\mathbf{z}_i = \mathbf{A}_i \mathbf{r}_i, \quad i = 1, \dots, K \quad (4.4)$$

with \mathbf{A}_i the $p_i \times N$ selection matrix of the i -th snapshot which is defined similarly to \mathbf{A} . In the following, the vectors $\mathbf{r}, \mathbf{r}_i, i = 1, \dots, K$, and $\mathbf{z}, \mathbf{z}_i, i = 1, \dots, K$, will be referred to as the *complete* and the *observed* data, respectively. Accordingly, the variables p and p_i indicate the number of the actual available channels, i.e., the number of observed elements, in the primary \mathbf{r} and the i -th secondary snapshot $\mathbf{r}_i, i = 1, \dots, K$, respectively².

Hence, leveraging the observed-data model in (4.3) and (4.4), the target detection problem in the presence of missing-data can be cast as

$$\begin{cases} H_0 : \begin{cases} \mathbf{z} = \mathbf{A}\mathbf{n} \\ \mathbf{z}_i = \mathbf{A}_i \mathbf{n}_i, \quad i = 1, \dots, K \end{cases} \\ H_1 : \begin{cases} \mathbf{z} = \alpha \mathbf{A}\mathbf{p} + \mathbf{A}\mathbf{n} \\ \mathbf{z}_i = \mathbf{A}_i \mathbf{n}_i, \quad i = 1, \dots, K \end{cases} \end{cases} \quad (4.5)$$

where the unknowns are $\boldsymbol{\theta}$ under H_0 and $\alpha, \boldsymbol{\theta}$ under H_1 .

²From a physical point of view, the selection matrices \mathbf{A} and \mathbf{A}_i , associated with \mathbf{r} and $\mathbf{r}_i, i = 1, \dots, K$, respectively, provide the components of the complete-data vectors into the observed-data space. As a result, the number of rows p (p_i) of the selection matrix \mathbf{A} (\mathbf{A}_i) denotes the dimension of the complex space where the observed-data in the considered (i -th secondary) snapshot lies.

4.2 Design of Decision Rules

Pursuing the classical approach based on the Neyman-Pearson criterion, the optimal decision statistic to the hypotheses testing problem (4.5), i.e., maximizing P_D for a given P_{fa} , could be devised. Unluckily, the resulting LRT, relies on the complete knowledge of the Probability Density Function (PDF)s under both hypotheses which requires a perfect knowledge of the parameters α and θ , reasonably not available in practical situations. As a result, a UMP test for the aforementioned problem does not exist. Hence, practically implementable receivers have to be designed resorting to sub-optimal criteria, such as GLR, which demands the ML estimation of the parameters under both hypotheses.

In this respect, it is worth pointing out that the existence of affordable low-complexity optimal solutions to the optimization problems involved in the estimation process under the two hypotheses is essential for the design of practically implementable detectors. Unfortunately, quite often, closed-form solutions are not available [24].

In light of the above considerations, the EM framework represents a viable means to determine approximated MLE of the parameters from the observed-data. Specifically, it alternates between an expectation (E)-step (in which the conditional expectation of the more analytically tractable complete-data likelihood is evaluated using the current estimate of the parameters) and a maximization (M)-step, in which the E-step score function is optimized in order to update the estimates. The above steps are then repeated until a convergence condition is achieved.

An iterative EM-based solution $\hat{\theta}_{EM,0}$ to the optimization problem under H_0 has been devised in [9]. More specifically, accounting for some possible a-priori knowledge on the covariance matrix structure, estimator devised in [9] involves only closed-form updates, at each iteration of the procedure, for a wide class of covariance structures. Detailed analysis on the convergence properties, as well as on the convergence rate, are also provided in [9].

The work in [9], however, does not address the optimization problem under H_1 , which is necessary to solve the detection problem (4.5). This challenging step is here addressed, yielding an EM-based framework for the joint estimation of α and \mathbf{M} . From an optimization theory point of

view this represents the main innovation of this chapter.

4.2.1 Parameters estimation under H_1

The EM procedure starts with an initial guess of the parameters, i.e., $\bar{\boldsymbol{\theta}}^{(0)} = [\alpha^{(0)}, \boldsymbol{\theta}^{(0)\top}]^\top$, and iterates between the E-step and the M-step, until convergence [42]. Specifically, at the h -th iteration, the E-step involves the evaluation of the score function

$$\begin{aligned} Q & \left(\alpha, \boldsymbol{\theta} | \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)} \right) \\ & = \mathbb{E}[\mathcal{L}_r(\alpha, \boldsymbol{\theta}, H_1) | \mathbf{z}, \mathbf{Z}, \mathbf{A}, \mathbf{A}_1, \dots, \mathbf{A}_K, \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)}, H_1] \end{aligned} \quad (4.6)$$

where

- $\mathbf{Z} = \{\mathbf{z}_1, \dots, \mathbf{z}_K\}$ is the set of observed secondary data;
- $\alpha^{(h-1)}$ and $\boldsymbol{\theta}^{(h-1)}$ are the estimates at the $(h-1)$ -th iteration;
- $\mathcal{L}_r(\alpha, \boldsymbol{\theta}, H_1)$ is the complete-data log-likelihood given by

$$\begin{aligned} \mathcal{L}_r(\alpha, \boldsymbol{\theta}, H_1) & = -(K+1) [N \ln(\pi) + \ln(\det(\mathbf{M}(\boldsymbol{\theta})))] \\ & \quad - \text{tr} \left\{ \mathbf{M}(\boldsymbol{\theta})^{-1} [(\mathbf{r} - \alpha \mathbf{p})(\mathbf{r} - \alpha \mathbf{p})^\dagger + \mathbf{S}] \right\} \end{aligned} \quad (4.7)$$

- $\mathbf{S} = \sum_{i=1}^K \mathbf{r}_i \mathbf{r}_i^\dagger$ is proportional, via K , to the conventional secondary data SCM.

Computing the conditional expectation involved in (4.6) yields (see Appendix C.1 for details on the statistical expectation evaluation)

$$\begin{aligned} Q & \left(\alpha, \boldsymbol{\theta} | \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)} \right) = -(K+1) [N \ln(\pi) + \ln(\det(\mathbf{M}(\boldsymbol{\theta})))] \\ & \quad - \text{tr} \left\{ \mathbf{M}(\boldsymbol{\theta})^{-1} [(\boldsymbol{\mu}^{(h-1)} - \alpha \mathbf{p})(\boldsymbol{\mu}^{(h-1)} - \alpha \mathbf{p})^\dagger + \boldsymbol{\Sigma}^{(h-1)}] \right\} \end{aligned} \quad (4.8)$$

where (the detailed expression is provided in (C.6) and (C.7))

$$\boldsymbol{\mu}^{(h-1)} = \mathbb{E}[\mathbf{r} | \mathbf{z}, \mathbf{A}, \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)}, H_1] \quad (4.9)$$

and (see (C.8)-(C.16) for the detailed derivation)

$$\begin{aligned} \boldsymbol{\Sigma}^{(h-1)} = & \left(\sum_{i=1}^K \mathbb{E}[\mathbf{r}_i \mathbf{r}_i^\dagger | \mathbf{z}_i, \mathbf{A}_i, \boldsymbol{\theta}^{(h-1)}] \right) - \boldsymbol{\mu}^{(h-1)} \boldsymbol{\mu}^{(h-1)\dagger} \\ & + \mathbb{E}[\mathbf{r} \mathbf{r}^\dagger | \mathbf{z}, \mathbf{A}, \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)}, H_1]. \end{aligned} \quad (4.10)$$

After the computation of the E-step, the M-step is performed, i.e., the score function (4.8) is maximized providing the following updated estimate of the unknowns

$$\left(\alpha^{(h)}, \boldsymbol{\theta}^{(h)} \right) = \underset{\alpha, \boldsymbol{\theta}: \mathbf{M}(\boldsymbol{\theta}) \in \mathcal{C}}{\operatorname{argmax}} Q \left(\alpha, \boldsymbol{\theta} | \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)} \right). \quad (4.11)$$

Still, as in the H_0 case analyzed in **Chapter 3**, different solution strategies to the optimization problem (4.11) are connected to diverse feasible sets \mathcal{C} . In this regard, some relevant cases of interest are analyzed in the following.

Unconstrained estimation

For this special and relevant case, the optimal solution to the M-step is available in closed-form, i.e., [60]

$$\alpha^{(h)} = \frac{\mathbf{p}^\dagger [\boldsymbol{\Sigma}^{(h-1)}]^{-1} \boldsymbol{\mu}^{(h-1)}}{\mathbf{p}^\dagger [\boldsymbol{\Sigma}^{(h-1)}]^{-1} \mathbf{p}} \quad (4.12)$$

and

$$\mathbf{M}(\boldsymbol{\theta}^{(h)}) = \frac{(\boldsymbol{\mu}^{(h-1)} - \alpha^{(h)} \mathbf{p})(\boldsymbol{\mu}^{(h-1)} - \alpha^{(h)} \mathbf{p})^\dagger + \boldsymbol{\Sigma}^{(h-1)}}{K + 1}. \quad (4.13)$$

Centro-Hermitianity constraint

As already described in 3.2.2, Centro-Hermitian is a particular matrix structure, commonly satisfied by covariance matrices encountered in many radar signal processing applications, e.g., radar systems utilizing standard rectangular, hexagonal, uniform circular, or cylindrical array [105]. Enforcing this structure is tantamount to considering \mathbf{M} belonging to the constraint set (3.6).

Following the same line of reasoning as in [25] and modeling \mathbf{p} as a persymmetric vector [74], i.e., $\mathbf{p} = \mathbf{J}\mathbf{p}^*$, the E-step (4.8) can be recast as

$$Q\left(\alpha, \boldsymbol{\theta} | \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)}\right) = - (K + 1) [N \ln(\pi) + \ln(\det(\mathbf{M}(\boldsymbol{\theta})))] \\ - \text{tr} \left\{ \mathbf{M}(\boldsymbol{\theta})^{-1} [\mathbf{P} + \boldsymbol{\Sigma}_{FB}^{(h-1)}] \right\} \quad (4.14)$$

where

$$\mathbf{P} = \frac{1}{2} \left[(\boldsymbol{\mu}^{(h-1)} - \alpha \mathbf{p})(\boldsymbol{\mu}^{(h-1)} - \alpha \mathbf{p})^\dagger \right. \\ \left. + \mathbf{J} \left((\boldsymbol{\mu}^{(h-1)} - \alpha \mathbf{p})(\boldsymbol{\mu}^{(h-1)} - \alpha \mathbf{p})^\dagger \right)^* \mathbf{J} \right] \quad (4.15)$$

and

$$\boldsymbol{\Sigma}_{FB}^{(h-1)} = \frac{1}{2} \left[\boldsymbol{\Sigma}^{(h-1)} + \mathbf{J} \left(\boldsymbol{\Sigma}^{(h-1)} \right)^* \mathbf{J} \right] \quad (4.16)$$

which plays the role of a FB averaged estimator [105]. Expression (4.14) allows the computation of the optimizers for the M-step in closed form as

$$\alpha^{(h)} = \frac{\mathbf{p}^\dagger [\boldsymbol{\Sigma}_{FB}^{(h-1)}]^{-1} \boldsymbol{\mu}^{(h-1)}}{\mathbf{p}^\dagger [\boldsymbol{\Sigma}_{FB}^{(h-1)}]^{-1} \mathbf{p}} \quad (4.17)$$

and

$$\hat{\mathbf{M}}(\boldsymbol{\theta}^{(h)}) = \frac{\mathbf{P}^{(h)} + \boldsymbol{\Sigma}_{FB}^{(h-1)}}{K + 1} \quad (4.18)$$

with

$$\mathbf{P}^{(h)} = \frac{1}{2} \left[(\boldsymbol{\mu}^{(h-1)} - \alpha^{(h)} \mathbf{p})(\boldsymbol{\mu}^{(h-1)} - \alpha^{(h)} \mathbf{p})^\dagger \right. \\ \left. + \mathbf{J} \left((\boldsymbol{\mu}^{(h-1)} - \alpha^{(h)} \mathbf{p})(\boldsymbol{\mu}^{(h-1)} - \alpha^{(h)} \mathbf{p})^\dagger \right)^* \mathbf{J} \right]. \quad (4.19)$$

General structured covariance matrix

For the case of an arbitrary constraint set \mathcal{C} , closed-form expressions for the joint estimation of α and $\boldsymbol{\theta}$, as involved in each M-step, could not be available. To this end, variations of the plain EM strategy are demanded,

for which the resulting update step is easier to handle. Two relevant methods are analyzed in the following, i.e., the ECM and M-EM [77, 73] which turn out to be very useful when the marginal optimization of the E-step score function over a single or sub-groups of unknowns can be conducted in closed form. In the former, which is a particular Generalized EM (GEM) algorithm [77, 97], the E-step is given by (4.8), while the M-step demands a sequence of Conditional Maximizations (CM), where in each of them a parameter is optimized while the others are held fixed. Formally, the CM over α is cast as

$$\alpha^{(h)} = \operatorname{argmax}_{\alpha} Q \left(\alpha, \boldsymbol{\theta}^{(h-1)} \mid \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)} \right) \quad (4.20)$$

whose solution is given by (4.12), whereas

$$\boldsymbol{\theta}^{(h)} = \operatorname{argmax}_{\boldsymbol{\theta}: \mathcal{M}(\boldsymbol{\theta}) \in \mathcal{C}} Q \left(\alpha^{(h)}, \boldsymbol{\theta} \mid \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)} \right) \quad (4.21)$$

is the CM related to the parameter vector $\boldsymbol{\theta}$.

As to the M-EM procedure, each iteration consists of several “cycles”, each focused on the optimization over a single parameter. Specifically, in the m -th cycle of the i -th iteration, the m -th variable is updated by maximizing the E-step score function computed with respect to the currently available parameters estimate. For the problem at hand, two cycles are considered. In the first, the E-step is given by evaluating the score function (4.6) at the point $(\alpha, \boldsymbol{\theta}^{(h-1)})$, i.e.,

$$\begin{aligned} Q \left(\alpha, \boldsymbol{\theta}^{(h-1)} \mid \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)} \right) = \\ \mathbb{E}[\mathcal{L}_r(\alpha, \boldsymbol{\theta}^{(h-1)}, H_1) \mid \mathbf{z}, \mathbf{Z}, \mathbf{A}, \mathbf{A}_1, \dots, \mathbf{A}_K, \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)}, H_1] \end{aligned} \quad (4.22)$$

whereas the M-step yields

$$\alpha^{(h)} = \operatorname{argmax}_{\alpha} Q \left(\alpha, \boldsymbol{\theta}^{(h-1)} \mid \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)} \right) \quad (4.23)$$

with the optimal solution provided by (4.12).

As to the second, the E-step is cast as the score function (4.6) evaluated

at the point $(\alpha^{(h)}, \boldsymbol{\theta})$ given the knowledge of $\alpha^{(h)}$ and $\boldsymbol{\theta}^{(h-1)}$, i.e.,

$$\begin{aligned} Q\left(\alpha^{(h)}, \boldsymbol{\theta} \mid \alpha^{(h)}, \boldsymbol{\theta}^{(h-1)}\right) = \\ \mathbb{E}[\mathcal{L}_r(\alpha^{(h)}, \boldsymbol{\theta}, H_1) \mid \mathbf{z}, \mathbf{Z}, \mathbf{A}, \mathbf{A}_1, \dots, \mathbf{A}_K, \alpha^{(h)}, \boldsymbol{\theta}^{(h-1)}, H_1] \end{aligned} \quad (4.24)$$

while the M-step is given by

$$\boldsymbol{\theta}^{(h)} = \underset{\boldsymbol{\theta}: \mathbf{M}(\boldsymbol{\theta}) \in \mathcal{C}}{\operatorname{argmax}} Q\left(\alpha^{(h)}, \boldsymbol{\theta} \mid \alpha^{(h)}, \boldsymbol{\theta}^{(h-1)}\right). \quad (4.25)$$

A case study is analyzed in the following.

Constraint on the lower bound of the white noise power level

Let us consider \mathbf{M} belonging to the uncertainty set (3.2). Denoting by $\mathbf{U}\boldsymbol{\Lambda}_\Sigma\mathbf{U}^\dagger$ the EVD of

$$\boldsymbol{\Sigma}_1^{(h-1)} = \frac{(\boldsymbol{\mu}^{(h-1)} - \alpha^{(h)}\mathbf{p})(\boldsymbol{\mu}^{(h-1)} - \alpha^{(h)}\mathbf{p})^\dagger + \boldsymbol{\Sigma}^{(h-1)}}{K+1}$$

and by $\tilde{\lambda}_v$, $v = 1, \dots, N$, its eigenvalues, the FML procedure [101, 7] (see also 3.2.2) provides the solution to the optimization problems (4.21) and (4.25), i.e.,

$$\hat{\mathbf{M}}(\boldsymbol{\theta}^{(h)}) = \mathbf{U}\boldsymbol{\Lambda}_{FML}\mathbf{U}^\dagger \quad (4.26)$$

with

$$\boldsymbol{\Lambda}_{FML} = \operatorname{diag}(\lambda_{1,FML}, \dots, \lambda_{N,FML}) \quad (4.27)$$

and $\lambda_{v,FML} = \max(\tilde{\lambda}_v, \sigma^2)$, $v = 1, \dots, N$.

Remark: Unitary invariant constraints. In many practical cases, the covariance matrix belongs to the feasible set of covariance matrices defined via unitary invariant continuous functions of the matrix entries [11]. Interestingly, many of these uncertainty sets can be described in terms of convex functions of the covariance matrix eigenvalues, paving the way for the development of tailored solutions to the ML estimation problems (4.21) and (4.25). However, even for some non-convex uncertainty sets, efficient algorithms can still be derived [11]. Uncertainty sets defined via unitary invariant functions encompass those resulting from an upper bound on the covariance condition number or a constraint on the maximum number of

uncorrelated interfering sources, just to mention a few [11].

4.2.2 Decision rules

This subsection provides practical detectors stemming from³ the GLR [60] and AMF [93] design criteria. Specifically, the following detectors are considered:

1. GLRT detector

$$\tau_{\text{GLRT-EM}} = \frac{f_{\mathbf{z}}(\mathbf{z}, \mathbf{Z}; \bar{\boldsymbol{\theta}}_1, \mathbf{A}, \mathbf{A}_1, \dots, \mathbf{A}_K, H_1)}{f_{\mathbf{z}}(\mathbf{z}, \mathbf{Z}; \bar{\boldsymbol{\theta}}_0, \mathbf{A}, \mathbf{A}_1, \dots, \mathbf{A}_K, H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} T \quad (4.28)$$

where $f_{\mathbf{z}}(\cdot)$ represents the likelihood function of the observations (under the appropriate hypothesis), $\bar{\boldsymbol{\theta}}_0 = [0, \hat{\boldsymbol{\theta}}_{EM,0}]$, $\bar{\boldsymbol{\theta}}_1 = [\hat{\alpha}_{EM}, \hat{\boldsymbol{\theta}}_{EM,1}]$ with $\hat{\boldsymbol{\theta}}_{EM,0}$ and $\hat{\boldsymbol{\theta}}_{EM,1}$ the estimates of $\boldsymbol{\theta}$ under H_0 and H_1 , respectively, provided by the bespoke EM-based procedures, and T is an appropriate detection threshold⁴ set to ensure a desired P_{fa} . Equation (4.28) is statistically equivalent to

$$\begin{aligned} \tau_{\text{GLRT-EM}} &= \mathcal{L}_{\mathbf{z}}(\mathbf{z}, \mathbf{Z}; \bar{\boldsymbol{\theta}}_1, \mathbf{A}, \mathbf{A}_1, \dots, \mathbf{A}_K, H_1) \\ &\quad - \mathcal{L}_{\mathbf{z}}(\mathbf{z}, \mathbf{Z}; \bar{\boldsymbol{\theta}}_0, \mathbf{A}, \mathbf{A}_1, \dots, \mathbf{A}_K, H_0) \underset{H_0}{\overset{H_1}{\gtrless}} T \end{aligned} \quad (4.29)$$

where, for $h = 0, 1$,

$$\begin{aligned} &\mathcal{L}_{\mathbf{z}}(\mathbf{z}, \mathbf{Z}; \bar{\boldsymbol{\theta}}_h, \mathbf{A}, \mathbf{A}_1, \dots, \mathbf{A}_K, \mathcal{H}_h) \\ &= \log f_{\mathbf{z}}(\mathbf{z}, \mathbf{Z}; \bar{\boldsymbol{\theta}}_h, \mathbf{A}, \mathbf{A}_1, \dots, \mathbf{A}_K, \mathcal{H}_h) \\ &= - \left(p + \sum_{i=1}^K p_i \right) \ln(\pi) - \ln(\det(\mathbf{A}\mathbf{M}(\boldsymbol{\theta})\mathbf{A}^T)) \\ &\quad - \text{tr}\{(\mathbf{A}\mathbf{M}(\boldsymbol{\theta})\mathbf{A}^T)^{-1}\mathbf{C}_h\} \\ &\quad - \sum_{i=1}^K \left[\ln(\det(\mathbf{A}_i\mathbf{M}(\boldsymbol{\theta})\mathbf{A}_i^T)) + \text{tr}\{(\mathbf{A}_i\mathbf{M}(\boldsymbol{\theta})\mathbf{A}_i^T)^{-1}\mathbf{z}_i\mathbf{z}_i^\dagger\} \right] \end{aligned} \quad (4.30)$$

³Notice that other sub-optimal criteria, such as Rao [36] and Wald [59] tests, can be pursued as well.

⁴With a slight abuse of notation, the same symbol is used to denote the detection threshold and its possible modifications introduced later, see, e.g., (4.29) and (4.31).

with $\mathbf{C}_0 = \mathbf{z}\mathbf{z}^\dagger$ and $\mathbf{C}_1 = (\mathbf{z} - \hat{\alpha}_{EM}\mathbf{A}\mathbf{p})(\mathbf{z} - \hat{\alpha}_{EM}\mathbf{A}\mathbf{p})^\dagger$.

2. **AMF** counterpart to (4.28), also referred to as **2SGLRT** (it computes the **GLRT** of the observed primary data over the parameter α and then substitutes in the resulting **GLRT** the estimate of the covariance parameters obtained from the secondary data)

$$\tau_{\text{AMF-EM}} = \frac{\left| \mathbf{z}^\dagger \left(\mathbf{A}\mathbf{M}(\hat{\boldsymbol{\theta}}_{EM,2})\mathbf{A}^\text{T} \right)^{-1} \mathbf{A}\mathbf{p} \right|^2}{\mathbf{p}^\dagger \mathbf{A}^\text{T} \left(\mathbf{A}\mathbf{M}(\hat{\boldsymbol{\theta}}_{EM,2})\mathbf{A}^\text{T} \right)^{-1} \mathbf{A}\mathbf{p}} \underset{H_0}{\overset{H_1}{\gtrless}} T \quad (4.31)$$

where $\hat{\boldsymbol{\theta}}_{EM,2}$ indicates the estimate of $\boldsymbol{\theta}$ obtained via the **EM** algorithm fed by secondary data only.

4.3 Performance Analysis

In this section, the effectiveness of the detectors devised to counteract the presence of missing-data is assessed. Specifically, the observables (subjected to downstream information loss mechanisms) are considered gathered by a radar system employing a **ULA** pointing at the boresight direction ($\theta_0 = 0$). The array comprises $N = 16$ antennas, unless otherwise stated, separated by $d_x = \lambda_0/2$, with λ_0 the radar operating wavelength. For each complete-data snapshot, the information provided by the output of $L = 3$ randomly selected channels is assumed missed. Therefore, each selection matrix (including \mathbf{A} of the primary data and \mathbf{A}_i of the i -th secondary data snapshot, $i = 1, \dots, K$) is constructed removing, independently from the other snapshots, L rows from the identity matrix, with the subset of L indexes randomly picked up from $\{1, \dots, N\}$ without replacement.

The performance of the detectors is analyzed in terms of P_D estimated via standard Monte Carlo counting techniques over 10^4 independent trials. Besides, the detection thresholds of the receivers are set to guarantee $P_{fa} = 10^{-4}$ and are evaluated using $100/P_{fa}$ independent Monte Carlo trials.

In the reported case studies, the disturbance covariance matrix is modeled as $\mathbf{M} = \mathbf{M}_J + \sigma_a^2 \mathbf{I}$, where σ_a^2 is the noise power level (assumed without

Table 4.1. Simulation Parameters.

Parameter	L	d_x	θ_0	σ_a^2	JNR_1	JNR_2	θ_1	θ_2
Value	3	$\lambda_0/2$	0°	1	30 dB	40 dB	-10°	15°

loss of generality equal to 0 dB) and

$$\mathbf{M}_J = \sum_{l=1}^{J_{NB}} \sigma_l^2 \mathbf{p}(\theta_l) \mathbf{p}(\theta_l)^\dagger \quad (4.32)$$

is the covariance contribution of J_{NB} uncorrelated narrow-band jammers, with

$$\mathbf{p}(\theta_l) = [1, e^{j\frac{2\pi}{\lambda_0} d_x \sin(\theta_l)}, \dots, e^{j(N-1)\frac{2\pi}{\lambda_0} d_x \sin(\theta_l)}]^\text{T} \in \mathbb{C}^N \quad (4.33)$$

the steering vector in the direction θ_l of the l -th jammer and σ_l^2 the power, or JNR σ_l^2/σ_a^2 , of the l -th jammer. Moreover, the complete-data SINR is defined as [40]

$$\text{SINR} = |\alpha|^2 \mathbf{p}^\dagger \mathbf{M}^{-1} \mathbf{p}. \quad (4.34)$$

Finally, denoting by $\tilde{\mathbf{z}}$ and $\tilde{\mathbf{z}}_i$, $i = 1, \dots, K$, the observed primary and secondary data snapshots with missing values replaced by zero-elements, the devised EM-based estimation procedures are initialized with

$$\alpha^{(0)} = \frac{\mathbf{p}^\dagger \mathbf{S}_1^{-1} \tilde{\mathbf{z}}}{\mathbf{p}^\dagger \mathbf{S}_1^{-1} \mathbf{p}} \quad (4.35)$$

and

$$\hat{\mathbf{M}}(\boldsymbol{\theta}^{(0)}) = \mathbf{S}_1 \quad (4.36)$$

where⁵ $\mathbf{S}_1 = 1/K \sum_{i=1}^K \tilde{\mathbf{z}}_i \tilde{\mathbf{z}}_i^\dagger$. The values of the system parameters involved in the analyzed case studies are summarized in Table 4.1.

⁵It is worth noting that some selection matrix configuration can lead, with non-zero probability, to a rank-deficient \mathbf{S}_1 . The considered Monte Carlo trials do not include such realizations.

4.3.1 Unconstrained estimation

Fig. 4.1 reports the P_D curves of the devised EM-based detection strategies⁶ versus SINR for the case of unconstrained interference covariance, with $K = 48$ and $K = 64$ in Figs. 4.1(a) and 4.1(b), respectively. As benchmarks, the GLRT and AMF detectors, with direct access to the complete-data set, are included for comparison. Besides, two additional heuristic counterparts, comprising the Kelly's GLRT and AMF detectors computed on a data set where the missing values are replaced by Linear Interpolation (LI) (in the following referred to as GLRT-LI and AMF-LI, respectively) are reported too (see Appendix C.2 for details).

The curves highlight that the devised procedures attain almost the same P_D levels, with performance comparable to the benchmarks, i.e., a loss of 4 dB for $K = 48$ and smaller than 2 dB for $K = 64$ at $P_D = 0.9$. As expected, if the sample support size increases, higher P_D levels can be achieved, with performance closer and closer to the benchmarks. The results reveal that for the considered unstructured case, the devised two-step as strategy is an effective and less computational demanding detector compared with the GLRT architecture. Indeed, it only requires the estimation of the covariance matrix under H_0 . As to the LI-based methods, they achieve P_D levels far below the benchmarks, with a significant loss in all the analyzed scenarios. This confirms the requirement to develop appropriate detection procedures capable of dealing with missing-data.

4.3.2 Centro-Hermitianity constraint

Fig. 4.2 presents the detection performance of the proposed receivers assuming the uncertainty set (3.6) for a symmetric ULA composed of $N = 15$ antennas, i.e.,

$$\mathbf{p}(\theta) = [e^{j\frac{2\pi}{\lambda_0}x_0 \sin(\theta)}, e^{j\frac{2\pi}{\lambda_0}x_1 \sin(\theta)}, \dots, e^{j\frac{2\pi}{\lambda_0}x_{N-1} \sin(\theta)}]^\text{T} \in \mathbb{C}^N \quad (4.37)$$

where $x_i = d_x \left(i - \left(\frac{N-1}{2}\right)\right)$, $i = 0, 1, \dots, N - 1$.

Specifically, two different sample support sizes are considered, i.e., $K = 30$ in Fig. 4.2(a) and $K = 45$ in Fig. 4.2(b). For comparison purposes,

⁶For covariance regularization purposes, a diagonal loading of 10^{-2} is applied to $\mathbf{M}(\hat{\boldsymbol{\theta}}_{EM,0})$, $\mathbf{M}(\hat{\boldsymbol{\theta}}_{EM,1})$, and $\mathbf{M}(\hat{\boldsymbol{\theta}}_{EM,2})$.

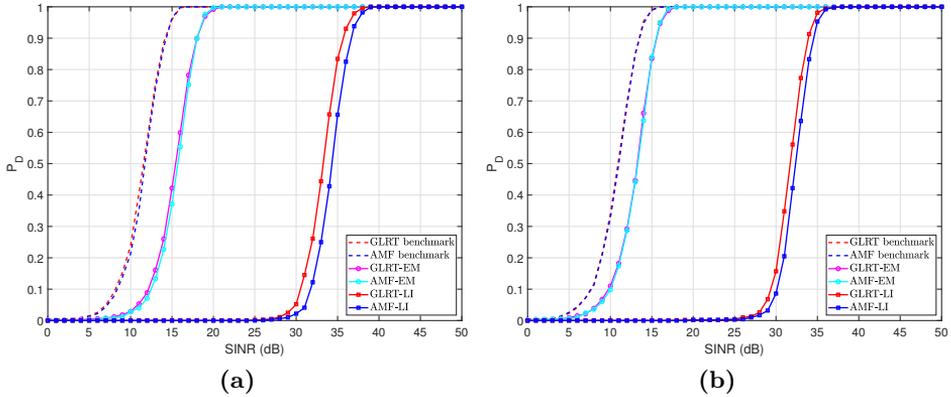


Figure 4.1. Detection performance for an ULA with $N = 16$ antennas and unconstrained estimation. Different sample support sizes are considered, i.e., (a) $K = 48$ and (b) $K = 64$.

tailored **GLRT** and **AMF** detectors, leveraging the Centro-Hermiticity (**CH**) structure for the estimation of the covariance matrix, are reported, too. Specifically, those computed on the complete-data set serve as benchmarks, whereas those evaluated on the observed-data set, with missing-data replaced by appropriately interpolated values, are considered as counterparts.

Inspection of the results shows that the proposed detectors ensure performance levels close to the benchmark with a gap between the curves less than 2 dB at $P_D = 0.9$ and $K = 30$. This is an indirect proof that capitalizing on the centro-Hermitian structure, accurate estimation of the unknowns could be obtained under both hypotheses, resulting in improved detection performance even with a reduced number of secondary data. Besides, both the devised one-step and two-step strategies achieve similar performance levels, with P_D values closer and closer to the benchmark as K increases, further corroborating the effectiveness of the bespoke detectors. Summarizing, the proposed detectors outperform all the considered (practically implantable) counterparts in the analyzed scenarios, confirming the capabilities of the devised adaptive architectures to operate in contexts with missing-data and structured covariance matrix.

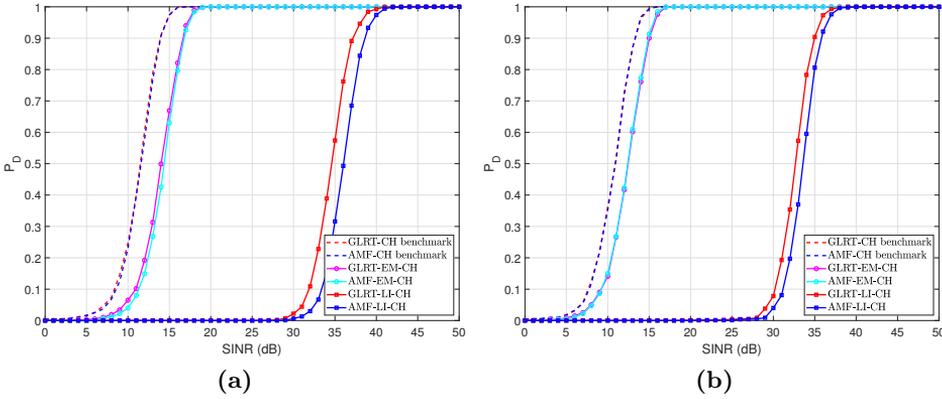


Figure 4.2. Detection performance for a symmetric ULA with $N = 15$ antennas assuming the covariance matrix with a centro-Hermitian structure, see (3.6). Different sample support sizes are considered, i.e., (a) $K = 30$ and (b) $K = 45$.

4.3.3 Lower bound of the white noise power level constraint

The performance of the devised detectors assuming a ULA with $N = 16$ antennas and the uncertainty set (3.2) is depicted in Fig. 4.3. Specifically, Figs. 4.3(a) and 4.3(b) consider $K = 24$ and $K = 48$, respectively. The performance of two GLRT detectors, using respectively the ECM and M-EM for the estimation of the parameters under H_1 hypothesis, are analyzed. In addition, a tailored (covariance structure aware) two-step receiver, using the EM-based structured procedure devised in [9] for the covariance matrix estimation under the H_0 hypothesis and referred to in the following as AMF-EM-FML, is also reported. Besides, the clairvoyant receiver, based on a perfect knowledge of the covariance matrix, is considered as benchmark.

The curves show that the devised architectures provide detection probabilities quite close to the optimum, highlighting the capabilities of the proposed detectors to leverage a-priori knowledge about the covariance matrix structure to keep the loss due to missing-data. The results are in line with the centro-Hermitian case and confirm the intuition that a-priori knowledge exploitation represents a viable means to perform an improved adap-

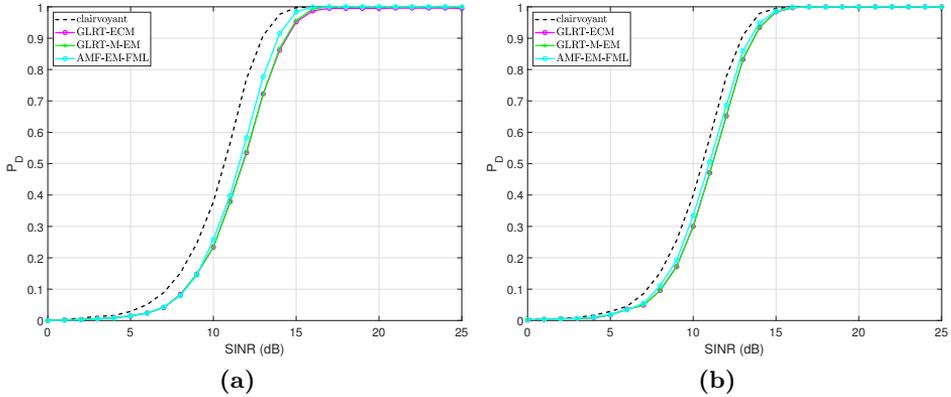


Figure 4.3. Detection performance for a ULA with $N = 16$ antennas assuming the uncertainty set in (3.2). Different sample support sizes are considered, i.e., (a) $K = 24$ and (b) $K = 48$.

tation process, especially in the presence of missing observations. More specifically, for $K = 24$, the loss with respect to the clairvoyant is less than 1 dB for the AMF-EM-FML and less than 2 dB for the two GLRT-based receivers. Besides, as K increases, the loss reduces progressively more and more, as depicted in Fig. 4.3(b).

4.3.4 Analysis on measured data

In this subsection, the performance of the devised detectors is analyzed on the measured data set collected in [94]. Specifically, the test-bed used for the acquisition process consists of

- a low-cost Software Defined Radio (SDR) coherent receiver made up of four RTL-SDR dongles (based on the RTL2832U chipset manufactured by Realtek [102]) that share the same clock source;
- a standard personal computer, used to calibrate the devices and run algorithms;
- a ULA comprising four dipole antennas with an inter-element space of $\lambda_0/2$.

Table 4.2. Targets Range and Angle Locations.

Range cell	30	40	78	85	120	145	180	245	280
Angle	7°	9°	9°	8°	9°	9°	8°	7°	9°

The data recording process has been conducted in an anechoic chamber using two **SDR** transmitters (each feeding a horn antenna with an azimuth beamwidth of 90° , at the considered operating frequency 1 GHz) to mimic the presence of two jammers occupying different spectral intervals and located at $\theta_1 = 0.5^\circ$ and $\theta_2 = 14.5^\circ$, respectively [6].

In [94], the data set have been used to validate the effectiveness of the covariance matrix estimator proposed in [9] on measured data. Here, the measured data set is used to validate the robustness to missing-data endowed by the proposed detection strategies. To this end, 9 prospective point-like targets causing echo returns with a **SINR** of 20 dB are synthetically simulated and injected in the dataset. The angle positions and ranges of the considered targets are reported in Table 4.2. Furthermore, for the analyzed case studies, the missing-data context is emulated considering a missing element at the output of a channel (chosen at random) for each observed snapshot.

Figs. 4.4 and 4.5 display the decision statistics of one-step and two-step strategies versus time/range resolution cell index, for a window of 300 bins. Besides, K snapshots, selected from a distinct but homogeneous temporal window, are used as a secondary data set⁷. Fig. 4.4 report the behavior of the considered detectors in the unstructured case, assuming $K = 40$. The output of the **GLRT-EM** detector is depicted in Fig. 4.4(a), whereas the output of the two-step counterpart strategy is shown in Fig. 4.4(b). Inspection of the results clearly reveals the presence of the targets, with peak levels higher than the interference-only floor level. Precisely, the **GLRT-EM** exhibits peaks greater than the floor of 2 dB in the worst case, whereas 5 dB peak gains are reached by the **AMF-EM**. The results corroborate the effectiveness of the proposed detectors highlighting their robustness on

⁷Notice that the choice of using secondary data from a tailored temporal window is to ensure that the homogeneous environment assumption is met.

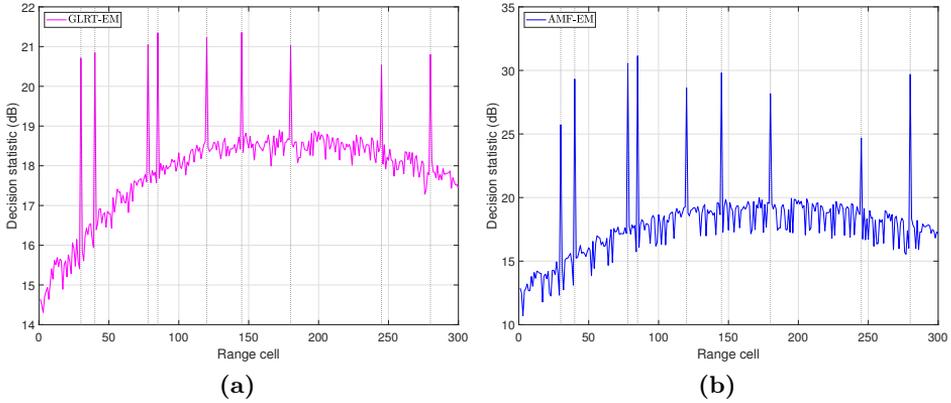


Figure 4.4. Detection performance for a ULA with $N = 4$ antennas and unconstrained estimation assuming $K = 40$ secondary data. Target locations are indicated by black dotted lines.

measured data.

In Fig. 4.5, the analysis is conducted assuming a sample support size of $K = 24$ and the covariance matrix belonging to the uncertainty set in (3.2). Fig. 4.5(a) shows the one-step decision strategies based on the ECM and M-EM methods (whose performance curves are substantially overlapped), while the output of the AMF-EM-FML detector is illustrated in Fig. 4.5(b).

As in the unstructured case, it is evident the presence of visible peaks located in correspondence of the targets range. In particular, peak levels greater than interference-only floor of at least 1 dB are reached by the two GLRT detectors (apart from the target at 180 where the peak strength is about 0.5 dB), whereas the AMF-EM-FML detector provides peak gains larger than 5 dB. The results highlight that also on measured data reliable detection performance can be obtained with a reduced number of secondary data if bespoke a-priori structural knowledge is exploited, corroborating the capabilities of the devised detection strategies also on a measured data scenario.

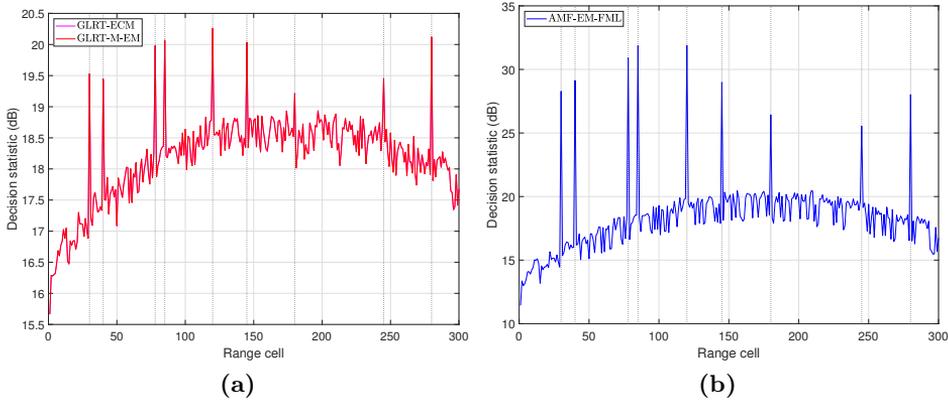


Figure 4.5. Detection performance for a ULA with $N = 4$ antennas assuming the uncertainty set in (3.2) and $K = 24$ secondary data. Target locations are indicated by black dotted lines.

4.4 Conclusion

In this chapter, adaptive detection architectures accounting for the presence of missing-data, have been proposed. Specifically, the problem of detecting a potential target echo return buried in Gaussian interference with a possibly structured covariance is formulated as a composite hypothesis test. The problem is handled via the GLR criterion, leading to the design of one-step and two-step GLRT detectors, which require the maximization of proper likelihood functions. Leveraging specific covariance structures, tailored estimation procedures relying upon the EM framework are developed. Specifically, for some covariance structures of practical interest, the optimization procedures involve only closed-form solutions at each iteration. Conversely, the case of a quite arbitrary constraint set is addressed resorting to ECM and M-EM frameworks, yielding more tractable optimization problems than classic EM strategy. The performance of the devised detection strategies has been assessed via Monte Carlo simulations for some a-priori structural covariance models. The results have highlighted the potentialities of the proposed detectors showing a performance level comparable to the benchmarks, which assume access to the entire set of observables. Besides, the effectiveness of the detectors

has been validated on measured data, collected in a controlled environment using an inexpensive four-channel receiver.

Chapter 5

Single-Snapshot Angle and Incremental Range Estimation for FDA-MIMO Radar

This chapter¹ investigates the problem of angle and incremental range (i.e., the target range offset with respect to the center of the cell under test) estimation with an FDA-MIMO radar exploiting as observable a single data snapshot. Starting from the observation that the ML estimation entails a two-dimensional grid search over the parameters of interest, three approximated ML techniques are designed resorting to the coordinate descent algorithm and the adaptive monopulse criterion (employing either real or complex slope/bias corrections). At the analysis stage, the estimation performance of the proposed methods, including the tapered and double-step monopulse versions, is assessed also in comparison with the CRLB. Numerical results corroborate the effectiveness of the considered estimation strategies in some diverse simulated scenarios.

The rest of the chapter is organized as follows. Section 5.1 presents the signal model for FDA-MIMO radar. In Section 5.2, the single snapshot angle and incremental range estimation problem is formulated. Besides,

¹©2021 IEEE. Reprinted, with permission, from L. Lan, M. Rosamilia, A. Aubry, A. De Maio and G. Liao, “Single-Snapshot Angle and Incremental Range Estimation for FDA-MIMO Radar,” IEEE Transactions on Aerospace and Electronic Systems, vol. 57, no. 6, pp. 3705-3718, Dec. 2021.

the ML estimator and the three aforementioned approximated methods are introduced. The CRLB for angle and incremental range is computed in Section 5.3, whereas performance analysis is addressed in Section 5.4. Finally, conclusions and possible future research developments are discussed in Section 5.5.

5.1 Signal Model for FDA-MIMO Radar

5.1.1 Transmitted Signal Model

Let us consider a colocated FDA-MIMO radar consisting of M transmit and N receive modules placed according to a ULA configuration in both transmission and reception (see Fig. 5.1) [66]. A frequency increment Δf is introduced element-by-element in the transmit array with the first array-element being the reference. Thus, the carrier frequency at the m -th transmitting element is

$$f_m = f_0 + (m - 1)\Delta f, \quad m = 1, 2, \dots, M, \quad (5.1)$$

where f_0 indicates the reference carrier. Each element transmits a specific base-band phase-modulated pulse, which is composed of P subpulses, and the resulting complex envelope of the radio frequency signal radiated by the m -th element can be expressed as

$$s_m(t) = \sqrt{\frac{E}{P}} x_m(t) e^{j2\pi f_m t}, \quad 0 \leq t \leq T_p, \quad (5.2)$$

where E is the transmitted energy, T_p is the radar pulse duration, and

$$x_m(t) = \frac{1}{\sqrt{\tau_b}} \sum_{p=1}^P \varphi_m(p) u\left[\frac{t - (p-1)\tau_b}{\tau_b}\right], \quad m = 1, 2, \dots, M, \quad (5.3)$$

$\tau_b = \frac{T_p}{P}$, $u(t)$ is the asymmetric rect function, i.e., $u(t) = 1$ as long as $0 \leq t \leq 1$ and zero elsewhere, and $\varphi_m(p) = e^{j\phi_m(p)}$, with $\phi_m(p) \in [0, 2\pi]$.

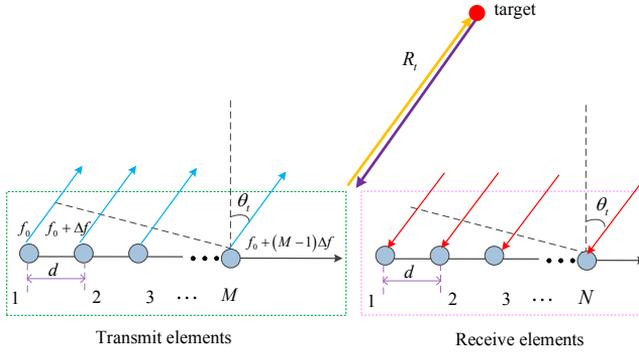


Figure 5.1. Signal transmission and reception in FDA-MIMO radar.

5.1.2 Received Signal Model

For a point-like target with a constant Radar Cross-Section (RCS) over the FDA-MIMO radar bandwidth, located in far-field at the angle θ_t and range R_t (see Fig. 5.1) [64, 65], the complex envelope of the signal received by the n -th radiating element ($n = 1, 2, \dots, N$) due to the signal transmitted by the m -th antenna ($m = 1, 2, \dots, M$) can be expressed as [66]

$$y_{m,n}(t) = \beta x_m(t - \tau_{m,n}) e^{j2\pi f_m(t - \tau_{m,n})} \approx \beta x_m(t - \tau_0) e^{j2\pi f_m(t - \tau_{m,n})}, \quad (5.4)$$

where $\tau_{m,n} = \frac{2R_t - d(n-1)\sin(\theta_t) - d(m-1)\sin(\theta_t)}{c}$ is the round-trip propagation time, β is the complex echo amplitude (accounting for the transmit amplitude, phase, target reflectivity, and channels propagation effects), d is the array's inter-element spacing, and c is the speed of light. The approximation relies on the narrowband assumption, i.e., $x_m(t - \tau_{m,n}) \approx x_m(t - \tau_0)$, with $\tau_0 = \frac{2R_t}{c}$ the customary envelope time delay.

After the pre-processing of Fig. 5.2, it can be shown that, under some mild technical conditions (see Appendix A of [66]), the received useful samples from the CUT can be stacked to form a $MN \times 1$ -dimensional vector

$$\mathbf{y}_S = \beta_1 \mathbf{b}(\theta_t) \otimes [\mathbf{c}(\theta_t) \odot \mathbf{a}(\Delta\tau)] = \beta_1 \mathbf{s}(\theta_t, \Delta\tau), \quad (5.5)$$

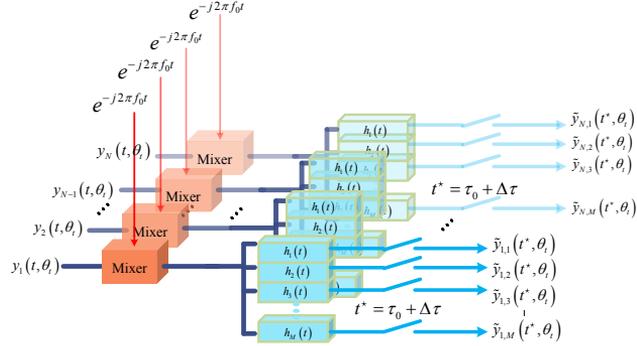


Figure 5.2. Signal processing at the receiver with multiple match-filtered waveforms.

where

- $\mathbf{s}(\theta_t, \Delta\tau) = \mathbf{b}(\theta_t) \otimes [\mathbf{c}(\theta_t) \odot \mathbf{a}(\Delta\tau)] \in \mathbb{C}^{MN}$ with $\Delta\tau$ the incremental delay w.r.t. the sampling time associated with the target range cell [66];
- $\mathbf{b}(\theta_t) = \left[1, e^{j2\pi \frac{d}{\lambda_0} \sin(\theta_t)}, \dots, e^{j2\pi \frac{d}{\lambda_0} (N-1) \sin(\theta_t)} \right]^T \in \mathbb{C}^N$ denotes the angle-dependent receive steering vector;
- $\mathbf{c}(\theta_t) = \mathbf{R}^T \mathbf{d}(\theta_t) \in \mathbb{C}^M$;
- $\mathbf{d}(\theta_t) = \left[1, e^{j2\pi \frac{d}{\lambda_0} \sin(\theta_t)}, \dots, e^{j2\pi \frac{d}{\lambda_0} (M-1) \sin(\theta_t)} \right]^T \in \mathbb{C}^M$ is the angle dependent transmit steering vector;
- $\mathbf{R} \in \mathbb{C}^{M \times M}$ denotes the transmit waveforms correlation matrix, i.e., $R_{m,l} = \int_0^{T_p} x_m(s) x_l^*(s) ds$, $(m, l) \in \{1, \dots, M\}^2$;
- $\mathbf{a}(\Delta\tau) = \left[1, e^{j2\pi \Delta f \Delta\tau}, \dots, e^{j2\pi \Delta f (M-1) \Delta\tau} \right]^T \in \mathbb{C}^M$ indicates the range-dependent steering vector.

5.2 Joint Angle-Range Estimation for FDA-MIMO Radar

This section investigates parameters estimation in **FDA-MIMO** radar. Let us assume the availability of a single data snapshot $\mathbf{z} \in \mathbb{C}^{MN}$ containing the superposition of the useful target signal and interference plus noise contribution [35, 79], namely the vector of observables can be cast as

$$\mathbf{z} = \beta_1 \mathbf{s}(\theta_t, \Delta\tau) + \mathbf{n}, \quad (5.6)$$

where $\mathbf{n} \in \mathbb{C}^{MN}$ is modeled as zero-mean complex circularly symmetric Gaussian random vector, i.e., $\mathbf{n} \sim CN(0, \mathbf{Q})$ with $\mathbf{Q} \in \mathbb{H}^{MN}$ the positive definite covariance matrix of the interference plus noise term. Therein, θ_t and $\Delta\tau$ describe the unknown angle and incremental range to be estimated, respectively. Now, letting $u = \sin(\theta_t)$ (complying with $|u| \leq 1$) and $\delta = 2\Delta f \Delta\tau$ (satisfying $|\delta| \leq \frac{\Delta f}{B}$), $\mathbf{s}(\theta_t, \Delta\tau)$ can be further expressed as

$$\mathbf{s}(\theta_t, \Delta\tau) = \mathbf{s}(u, \delta) = \mathbf{b}(u) \otimes [\mathbf{c}(u) \odot \mathbf{a}(\delta)], \quad (5.7)$$

where

- $\mathbf{b}(u) = [1, e^{j2\pi \frac{d}{\lambda_0} u}, \dots, e^{j2\pi \frac{d}{\lambda_0} (N-1)u}]^T \in \mathbb{C}^N$;
- $\mathbf{c}(u) = \mathbf{R}^T \mathbf{d}(u) \in \mathbb{C}^M$;
- $\mathbf{d}(u) = [1, e^{j2\pi \frac{d}{\lambda_0} u}, \dots, e^{j2\pi \frac{d}{\lambda_0} (M-1)u}]^T \in \mathbb{C}^M$;
- $\mathbf{a}(\delta) = [1, e^{j\pi\delta} \dots, e^{j\pi(M-1)\delta}]^T \in \mathbb{C}^M$.

In the following subsections, the **ML** estimator of u and δ is introduced and three low complexity methods are proposed to approximate the computation.

5.2.1 ML Estimation of u and δ

In this subsection, the **ML** estimation problem is formalized as the constrained maximization (w.r.t. the unknown parameters, i.e., β_1 , u and

δ [35]) of the likelihood function,

$$f(u, \delta, \beta_1; \mathbf{z}) = \frac{1}{\pi^{MN} \det(\mathbf{Q})} e^{-[(\mathbf{z} - \beta_1 \mathbf{s}(u, \delta))^\dagger \mathbf{Q}^{-1} (\mathbf{z} - \beta_1 \mathbf{s}(u, \delta))]} . \quad (5.8)$$

This is equivalent to minimizing the following quadratic form

$$\min_{\substack{\beta_1 \in \mathbb{C}, u \in [-1, 1], \\ \delta \in [-\frac{\Delta f}{B}, \frac{\Delta f}{B}]}} (\mathbf{z} - \beta_1 \mathbf{s}(u, \delta))^\dagger \mathbf{Q}^{-1} (\mathbf{z} - \beta_1 \mathbf{s}(u, \delta)). \quad (5.9)$$

Now, concentrating (5.9) over β_1 yields

$$\hat{\beta}_1 = \frac{\mathbf{s}^\dagger(u, \delta) \mathbf{Q}^{-1} \mathbf{z}}{\mathbf{s}^\dagger(u, \delta) \mathbf{Q}^{-1} \mathbf{s}(u, \delta)}. \quad (5.10)$$

Hence, substituting (5.10) into the objective function of (5.9) as well as dropping constant and irrelevant terms leads to

$$\max_{u \in [-1, 1], \delta \in [-\frac{\Delta f}{B}, \frac{\Delta f}{B}]} \frac{|\mathbf{s}^\dagger(u, \delta) \mathbf{Q}^{-1} \mathbf{z}|^2}{\mathbf{s}^\dagger(u, \delta) \mathbf{Q}^{-1} \mathbf{s}(u, \delta)}. \quad (5.11)$$

Finally, the ML estimates of u and δ can be obtained as maximizers of

$$P(u, \delta) = |\mathbf{w}_0^\dagger(u, \delta) \mathbf{z}|^2, \quad (5.12)$$

where $\mathbf{w}_0(u, \delta) = [\mathbf{s}^\dagger(u, \delta) \mathbf{Q}^{-1} \mathbf{s}(u, \delta)]^{-\frac{1}{2}} \mathbf{Q}^{-1} \mathbf{s}(u, \delta) \in \mathbb{C}^{MN}$.

5.2.2 Approximated Methods for Range and Angle Estimation

The ML rule can be practically implemented via a 2-D grid search. To reduce the computational cost required by the foregoing procedure, it is valuable to design approximated solution methods. To this end, this section is focused on designing: 1) a CD algorithm; 2) AMPs.

Coordinate Descent Algorithm

An approximation of the optimal 2-D search involved in (5.12) is developed via the CD method. This leads to a sequence of 1-D searches

obtained alternating between the optimization over each variable keeping the other fixed. The problem of finding the maximizer of each 1-D search is tackled using the grid search method where the feasible interval of interest is discretized in a finite set of points. Specifically, the 1-D searches w.r.t. u and δ are respectively conducted over the discretized intervals \mathcal{I}_u and \mathcal{I}_δ , defined as

$$\mathcal{I}_u = \left\{-1 + \frac{2i}{N_u}, i = 0, \dots, N_u\right\} \quad (5.13a)$$

and

$$\mathcal{I}_\delta = \left\{-\frac{\Delta f}{B} + \frac{2i}{N_\delta} \frac{\Delta f}{B}, i = 0, \dots, N_\delta\right\}, \quad (5.13b)$$

where $(N_u + 1)$ and $(N_\delta + 1)$ are the number of discrete points considered for the optimization over u and δ , respectively. It is also worth pointing out that it is not possible to establish *a-priori* which order of optimization leads to the best estimate, i.e., first optimize u and then optimize δ , or vice versa. To overcome this problem, the CD-based algorithm considered herein is applied twice, one for each possible initial search direction. Therefore, among the two obtained solutions, the one that maximizes (5.12) is chosen as estimate. The exit condition (for each updating policy) is set as $D^n \leq \varepsilon$ with $\varepsilon > 0$ and $D^n = |P^n - P^{n-1}|$, where

$$P^n = |(\mathbf{w}^n)^\dagger \mathbf{z}|^2 \quad (5.14)$$

indicates the objective function at the n -th iteration with

$$\mathbf{w}^n = \left[\mathbf{s}^\dagger \left(\hat{u}^n, \hat{\delta}^n \right) \mathbf{Q}^{-1} \mathbf{s} \left(\hat{u}^n, \hat{\delta}^n \right) \right]^{-\frac{1}{2}} \mathbf{Q}^{-1} \mathbf{s} \left(\hat{u}^n, \hat{\delta}^n \right). \quad (5.15)$$

Letting u_0 and δ_0 the nominal angle and range, the initial estimates are chosen as $\hat{u}^0 = u_0$ and $\hat{\delta}^0 = \delta_0$. The resulting method is summarized in **Algorithm 7**.

Note that in the presence of two blocks/variables, regardless of the initial search direction, the CD approach (starting from the second iteration) coincides with the MBI policy [29]. Therefore, invoking the convergence properties of MBI [29, 88, 12], any limit point resulting from **Algorithm 7** is a stationary point to Problem (5.12), although convergence to the optimal value cannot be claimed [19].

Algorithm 7 FDA-CD

Require: $u_0, \delta_0, \mathbf{z}, \mathbf{Q}, \mathbf{s}, \varepsilon$ **Ensure:** A solution $\hat{u}, \hat{\delta}$ to (5.11).**Initialization:** $n = 0, \hat{u}^0 = u_0, \hat{\delta}^0 = \delta_0, P^0 = P(\hat{u}^0, \hat{\delta}^0)$ **repeat** (optimization for initial search direction given by u)1. Find $\hat{u}^{n+1} = \arg \max_{u \in \mathcal{I}_u} P(u, \hat{\delta}^n)$;2. Find $\hat{\delta}^{n+1} = \arg \max_{\delta \in \mathcal{I}_\delta} P(\hat{u}^{n+1}, \delta)$ and set P^{n+1} as the corresponding maximum value;3. $n = n + 1$;**until** $|P^n - P^{n-1}| > \varepsilon$; $P_x = P^n$; $\hat{u}_x = \hat{u}^n$; $\hat{\delta}_x = \hat{\delta}^n$;**Initialization:** $n = 0$ **repeat** (optimization for initial search direction given by δ)1. Find $\hat{\delta}^{n+1} = \arg \max_{\delta \in \mathcal{I}_\delta} P(\hat{u}^n, \delta)$;2. Find $\hat{u}^{n+1} = \arg \max_{u \in \mathcal{I}_u} P(u, \hat{\delta}^{n+1})$ and set P^{n+1} as the corresponding maximum value;3. $n = n + 1$;**until** $|P^n - P^{n-1}| > \varepsilon$; $P_y = P^n$; $\hat{u}_y = \hat{u}^n$; $\hat{\delta}_y = \hat{\delta}^n$;**if** $P_x > P_y$ **then** Output $\hat{u} = \hat{u}_x$ and $\hat{\delta} = \hat{\delta}_x$.**else** Output $\hat{u} = \hat{u}_y$ and $\hat{\delta} = \hat{\delta}_y$.**end if**

Leveraging the output of **Algorithm 7** the estimates of u and δ obtained via the CD method are given by

$$u_{\text{CD}} = \hat{u} \quad (5.16)$$

and

$$\delta_{\text{CD}} = \hat{\delta}. \quad (5.17)$$

Adaptive Monopulse Procedure

The CD method is still time-consuming because of the 1-D searches involved during each iteration. To circumvent this drawback, in this subsection, the generalized monopulse approach is exploited [81] to approximate the optimal search in (5.12). To this end, let

$$\mathbf{h}_{\text{AMP}} = \mathbf{h}_0 + \mathbf{C}(\mathbf{r} - \boldsymbol{\mu}), \quad (5.18)$$

where $\mathbf{h}_{\text{AMP}} = [u_{\text{AMP}}, \delta_{\text{AMP}}]^T \in \mathbb{R}^2$ refers to the unknown parameters, $\mathbf{h}_0 = [u_0, \delta_0]^T \in \mathbb{R}^2$, $\mathbf{C} = \begin{bmatrix} C_{xu} & C_{x\delta} \\ C_{yu} & C_{y\delta} \end{bmatrix} \in \mathbb{R}^{2 \times 2}$ denotes a slope correction matrix, $\boldsymbol{\mu} = [\mu_x, \mu_y]^T \in \mathbb{R}^2$ represents a bias correction vector, and $\mathbf{r} = [r_x, r_y]^T \in \mathbb{R}^2$ refers to compressed measures, with r_x and r_y the monopulse ratios defined as

$$r_x = \text{Re} \left\{ \frac{\mathbf{d}_x^\dagger \mathbf{z}}{\mathbf{w}^\dagger \mathbf{z}} \right\}, \quad (5.19a)$$

$$r_y = \text{Re} \left\{ \frac{\mathbf{d}_y^\dagger \mathbf{z}}{\mathbf{w}^\dagger \mathbf{z}} \right\}. \quad (5.19b)$$

In (5.19a) and (5.19b) $\mathbf{w} = \mathbf{Q}^{-1} \mathbf{s}_0 \in \mathbb{C}^{MN}$ indicates the sum weight vector with $\mathbf{s}_0 = \mathbf{s}(u_0, \delta_0)$, $\mathbf{d}_x = \mathbf{Q}^{-1} \mathbf{s}_u \in \mathbb{C}^{MN}$ and $\mathbf{d}_y = \mathbf{Q}^{-1} \mathbf{s}_\delta \in \mathbb{C}^{MN}$ the difference beam weights, w.r.t. u and δ , respectively, where $\mathbf{s}_h = \frac{\partial \mathbf{s}}{\partial h} \Big|_{(u_0, \delta_0)}$, $h \in \{u, \delta\}$. Detailed expressions for \mathbf{s}_u and \mathbf{s}_δ are available in Appendix D.1.

The matrix \mathbf{C} and the vector $\boldsymbol{\mu}$ are determined from the vector-valued function $\mathbf{M}(u, \delta) \in \mathbb{R}^2$,

$$\mathbf{M}(u, \delta) = \mathbf{C}(\mathbb{E}[\mathbf{r}] - \boldsymbol{\mu}), \quad (5.20)$$

forcing the conditions

$$\mathbf{M}(u_0, \delta_0) = \mathbf{0}, \quad (5.21)$$

and

$$\mathbf{C} \mathbb{E} \left[\left(\frac{\partial \mathbf{r}}{\partial u}, \frac{\partial \mathbf{r}}{\partial \delta} \right) \right] \Big|_{(u_0, \delta_0)} = \mathbf{I}. \quad (5.22)$$

Under the Gaussianity assumption for the received data vector, the expectation of the monopulse ratios involved in (5.20), i.e., $\mathbb{E}[r_\alpha]$, $\alpha \in \{x, y\}$, can be expressed as²

$$\begin{aligned} \mathbb{E}[r_\alpha] &= \mathbb{E} \left[\operatorname{Re} \left\{ \frac{\mathbf{d}_\alpha^\dagger \mathbf{z}}{\mathbf{w}^\dagger \mathbf{z}} \right\} \right] \\ &= \mathbb{E} \left[\operatorname{Re} \left\{ \frac{\mathbf{d}_\alpha^\dagger \mathbf{z} \mathbf{z}^\dagger \mathbf{w}}{\mathbf{w}^\dagger \mathbf{z} \mathbf{z}^\dagger \mathbf{w}} \right\} \right] = \operatorname{Re} \left\{ \frac{\mathbf{d}_\alpha^\dagger \mathbb{E}[\mathbf{z} \mathbf{z}^\dagger] \mathbf{w}}{\mathbf{w}^\dagger \mathbb{E}[\mathbf{z} \mathbf{z}^\dagger] \mathbf{w}} \right\} \\ &= \frac{|\beta_1|^2 \operatorname{Re} \left\{ \mathbf{d}_\alpha^\dagger \mathbf{s} \mathbf{s}^\dagger \mathbf{w} \right\} + \operatorname{Re} \left\{ \mathbf{d}_\alpha^\dagger \mathbf{Q} \mathbf{w} \right\}}{|\beta_1|^2 \mathbf{w}^\dagger \mathbf{s} \mathbf{s}^\dagger \mathbf{w} + \mathbf{w}^\dagger \mathbf{Q} \mathbf{w}}, \quad \alpha \in \{x, y\}, \end{aligned} \quad (5.23)$$

where $\mathbf{s} = \mathbf{s}(u, \delta)$. For a sufficiently high value of $|\beta_1|^2$, the two terms $\mathbf{d}_\alpha^\dagger \mathbf{Q} \mathbf{w}$ and $\mathbf{w}^\dagger \mathbf{Q} \mathbf{w}$, at the numerator and the denominator of (5.23), respectively, can be disregarded. As a result, the bias $\mathbb{E}[\mathbf{r}] = [\mathbb{E}[r_x], \mathbb{E}[r_y]]^\top$ correction values, fulfilling (5.21), can be approximated as

$$\mu_\alpha \approx \operatorname{Re} \left\{ \frac{\mathbf{d}_\alpha^\dagger \mathbf{s}_0}{\mathbf{w}^\dagger \mathbf{s}_0} \right\}, \quad \alpha \in \{x, y\}. \quad (5.24)$$

Besides, the slope correction matrix is computed according to (5.22), i.e.,

$$\mathbf{C} = \begin{bmatrix} C_{xu} & C_{x\delta} \\ C_{yu} & C_{y\delta} \end{bmatrix} = \begin{bmatrix} \mathbb{E} \left[\left. \frac{\partial r_x}{\partial u} \right|_{(u_0, \delta_0)} \right] & \mathbb{E} \left[\left. \frac{\partial r_x}{\partial \delta} \right|_{(u_0, \delta_0)} \right] \\ \mathbb{E} \left[\left. \frac{\partial r_y}{\partial u} \right|_{(u_0, \delta_0)} \right] & \mathbb{E} \left[\left. \frac{\partial r_y}{\partial \delta} \right|_{(u_0, \delta_0)} \right] \end{bmatrix}^{-1}, \quad (5.25)$$

where $\mathbb{E} \left[\left. \frac{\partial r_\alpha}{\partial h} \right|_{(u_0, \delta_0)} \right]$, with $\alpha \in \{x, y\}$, $h \in \{u, \delta\}$, is approximated as (see [80, eq. 17])

$$\mathbb{E} \left[\left. \frac{\partial r_\alpha}{\partial h} \right|_{(u_0, \delta_0)} \right] \approx \frac{\operatorname{Re} \left\{ \mathbf{d}_\alpha^\dagger \mathbf{s}_h \mathbf{s}_0^\dagger \mathbf{w} + \mathbf{d}_\alpha^\dagger \mathbf{s}_0 \mathbf{s}_h^\dagger \mathbf{w} \right\}}{|\mathbf{w}^\dagger \mathbf{s}_0|^2} - \mu_\alpha 2 \operatorname{Re} \left\{ \frac{\mathbf{w}^\dagger \mathbf{s}_h}{\mathbf{w}^\dagger \mathbf{s}_0} \right\}, \quad (5.26)$$

²The “mean” is computed performing the expectation, w.r.t. the denominator, of conditional mean of the ratio given the denominator.

Algorithm 8 FDA-AMP**Require:** $u_0, \delta_0, \mathbf{Q}, \mathbf{s}_0, \mathbf{s}_u, \mathbf{s}_\delta, \mathbf{z}$ **Ensure:** A solution $\hat{u}_{\text{AMP}}, \hat{\delta}_{\text{AMP}}$ to (5.11).

1. Compute $\mathbf{w} = \mathbf{Q}^{-1}\mathbf{s}_0$, $\mathbf{d}_x = \mathbf{Q}^{-1}\mathbf{s}_u$, and $\mathbf{d}_y = \mathbf{Q}^{-1}\mathbf{s}_\delta$;
 2. Evaluate $r_\alpha = \text{Re} \left\{ \frac{\mathbf{d}_\alpha^\dagger \mathbf{z}}{\mathbf{w}^\dagger \mathbf{z}} \right\}$, $\alpha = x, y$ to obtain $\mathbf{r} = [r_x, r_y]^\text{T}$;
 3. Compute $\boldsymbol{\mu} = [\mu_x, \mu_y]^\text{T}$ via (5.24) and \mathbf{C} using (5.25);
 4. Determine \mathbf{h}_{AMP} using (5.18);
 5. Project the candidate solution \mathbf{h}_{AMP} onto $[-1, 1] \times [-\frac{\Delta f}{B}, \frac{\Delta f}{B}]$ to get an estimate $\hat{\mathbf{h}}_{\text{AMP}}$ complying with the problem constraints.
- Output $[\hat{u}_{\text{AMP}}, \hat{\delta}_{\text{AMP}}]^\text{T} = \hat{\mathbf{h}}_{\text{AMP}}$.

Summarizing, the AMP procedure is synthetically reported in **Algorithm 8**. It is also worth pointing out that, as shown in [79], by further executing the procedure, namely, re-applying the monopulse algorithm employing the estimates \hat{u}_{AMP} and $\hat{\delta}_{\text{AMP}}$ in place of the nominal u_0 and δ_0 , the potential bias could be reduced, leading to some possible performance improvements. More in details, the second iteration of **Algorithm 8** is performed using as input $\hat{u}_{\text{AMP}}, \hat{\delta}_{\text{AMP}}, \mathbf{Q}, \mathbf{s}_{\text{AMP}}, \mathbf{s}_{\text{AMP}-u}, \mathbf{s}_{\text{AMP}-\delta}, \mathbf{z}$, where

- $\mathbf{s}_{\text{AMP}} = \mathbf{s}(\hat{u}_{\text{AMP}}, \hat{\delta}_{\text{AMP}})$ refers to the receive steering vector associated with the estimated direction and incremental range.
- $\mathbf{s}_{\text{AMP}-u}$, and $\mathbf{s}_{\text{AMP}-\delta}$ denote the partial derivatives of $s(u, \delta)$ w.r.t. u and δ , respectively computed at $(\hat{u}_{\text{AMP}}, \hat{\delta}_{\text{AMP}})$.

The overall procedure will be referred to as the Double-Step Corrected (DSC)-AMP and could be also potentially iterated multiple times. The resulting estimates are denoted by $\hat{u}_{\text{DSC-AMP}}$ and $\hat{\delta}_{\text{DSC-AMP}}$, respectively.

AGMP-CC

As the generalized monopulse approach procedure in (5.18) assumes real slope and bias correction, a more general procedure calls for a complex slope and bias correction [84]. This is the rationale followed in this section which is focused on the Adaptive Generalized Monopulse Procedure with

Complex Correction (AGMP-CC), given by

$$\mathbf{h}_{\text{AGMP-CC}} = \mathbf{h}_0 + \text{Re} \left\{ \hat{\mathbf{C}}^\dagger (\hat{\mathbf{r}} - \hat{\boldsymbol{\mu}}) \right\}, \quad (5.27)$$

where $\mathbf{h}_{\text{AGMP-CC}} = [u_{\text{AGMP-CC}} \delta_{\text{AGMP-CC}}]^\text{T} \in \mathbb{R}^2$, $\mathbf{h}_0 = [u_0, \delta_0]^\text{T} \in \mathbb{R}^2$ denotes the nominal values of the unknowns, $\hat{\mathbf{C}} \in \mathbb{C}^{2 \times 2}$ denotes a complex slope correction matrix, $\hat{\boldsymbol{\mu}} \in \mathbb{C}^2$ indicates a bias correction vector, i.e., $\hat{\boldsymbol{\mu}} = [\tilde{\mu}_x, \tilde{\mu}_y]^\text{T}$ with $\tilde{\mu}_\alpha = \frac{\mathbf{d}_\alpha^\dagger \mathbf{s}_0}{\mathbf{w}^\dagger \mathbf{s}_0}$, $\alpha \in \{x, y\}$ (obtained following the same line of reasoning as in [84]), and $\hat{\mathbf{r}} \in \mathbb{C}^2$ represents the complex monopulse ratio vector, i.e., $\hat{\mathbf{r}} = [\tilde{r}_x, \tilde{r}_y]^\text{T}$ with

$$\tilde{r}_\alpha = \frac{\mathbf{d}_\alpha^\dagger \mathbf{z}}{\mathbf{w}^\dagger \mathbf{z}}, \quad \alpha \in \{x, y\}. \quad (5.28)$$

Now, substituting $\mathbf{z} = \beta_1 \mathbf{s} + \mathbf{n}$ into (5.28), and after some algebra and approximations as in [84], the following equation involving the actual unknowns, i.e., $\mathbf{h} = [u, \delta]^\text{T}$, is obtained

$$(\tilde{\mathbf{r}} - \tilde{\boldsymbol{\mu}}) = \mathbf{B}_1 (\mathbf{h} - \mathbf{h}_0) + \mathbf{p}_1, \quad (5.29)$$

where

- $\tilde{\mathbf{r}} \in \mathbb{C}^4$ represents the monopulse ratio vector whose entries are the complex monopulse ratios and their complex conjugates, i.e., $\tilde{\mathbf{r}} = [\hat{\mathbf{r}}^\text{T}, \hat{\mathbf{r}}^{*\text{T}}]^\text{T}$;
- $\tilde{\boldsymbol{\mu}} = [\hat{\boldsymbol{\mu}}^\text{T}, \hat{\boldsymbol{\mu}}^{*\text{T}}]^\text{T} \in \mathbb{C}^4$;
- $\mathbf{B}_1 = [\hat{\mathbf{B}}^\text{T}, \hat{\mathbf{B}}^{*\text{T}}]^\text{T} \in \mathbb{C}^{4 \times 2}$ where $\hat{\mathbf{B}} = \begin{bmatrix} B_{xu} & B_{x\delta} \\ B_{yu} & B_{y\delta} \end{bmatrix}$ with $B_{\alpha h} = \frac{\mathbf{d}_\alpha^\dagger \mathbf{s}_h \mathbf{w}^\dagger \mathbf{s}_0 - \mathbf{d}_\alpha^\dagger \mathbf{s}_0 \mathbf{w}^\dagger \mathbf{s}_h}{(\mathbf{w}^\dagger \mathbf{s}_0)^2}$ ($\alpha \in \{x, y\}, h \in \{u, \delta\}$);
- $\mathbf{p}_1 = [\hat{\mathbf{p}}^\text{T}, \hat{\mathbf{p}}^{*\text{T}}]^\text{T} \in \mathbb{C}^4$ where $\hat{\mathbf{p}} = [p_x, p_y]^\text{T}$ with $p_\alpha = \frac{\mathbf{d}_\alpha^\dagger \tilde{\mathbf{p}}}{\beta_1 \mathbf{w}^\dagger \mathbf{s}}$, $\alpha \in \{x, y\}$;
- $\tilde{\mathbf{p}} = \mathbf{n} - \left(\frac{\mathbf{w}^\dagger \mathbf{n}}{\mathbf{w}^\dagger \mathbf{s}} \right) \mathbf{s} \in \mathbb{C}^{MN}$.

Equation (5.29) poses an important constraint on the matrix $\hat{\mathbf{C}}$, so as to obtain in absence of noise and under the validity of (5.29) the exact values of the parameters via (5.27), i.e.,

$$\mathbf{C}_1^\dagger \mathbf{B}_1 = 2\mathbf{I}. \quad (5.30)$$

where $\mathbf{C}_1 = [\hat{\mathbf{C}}; \hat{\mathbf{C}}^*] \in \mathbb{C}^{4 \times 2}$.

After some algebra, (5.30) can be recast as

$$\text{Re} \left\{ \hat{\mathbf{C}}^\dagger \hat{\mathbf{B}} \right\} = \mathbf{I}. \quad (5.31)$$

Hence, by substituting (5.31) and (5.29) into (5.27), it yields

$$\mathbf{h}_{\text{AGMP-CC}} - \mathbf{h}_0 = \text{Re} \left\{ \hat{\mathbf{C}}^\dagger (\hat{\mathbf{r}} - \hat{\boldsymbol{\mu}}) \right\} = (\mathbf{h} - \mathbf{h}_0) + \text{Re} \left\{ \hat{\mathbf{C}}^\dagger \hat{\mathbf{p}} \right\}. \quad (5.32)$$

It can be seen that the estimates $\mathbf{h}_{\text{AGMP-CC}}$ will be closest to the true value, i.e., \mathbf{h} , when the noise term effect, i.e., $\text{Re} \left\{ \hat{\mathbf{C}}^\dagger \hat{\mathbf{p}} \right\}$, is minimised. Following [84], the mean square value of the noise error terms on the u and δ components, is minimized according to

$$\begin{cases} \min_{\hat{\mathbf{C}}_u} \hat{\mathbf{C}}_u^\dagger \mathbf{z}_1 \hat{\mathbf{C}}_u & \& \min_{\hat{\mathbf{C}}_\delta} \hat{\mathbf{C}}_\delta^\dagger \mathbf{z}_1 \hat{\mathbf{C}}_\delta \\ \text{s.t.} & \text{Re} \left\{ \hat{\mathbf{C}}^\dagger \hat{\mathbf{B}} \right\} = \mathbf{I}, \end{cases} \quad (5.33)$$

where $\hat{\mathbf{C}}_u \in \mathbb{C}^2$, $\hat{\mathbf{C}}_\delta \in \mathbb{C}^2$ denote the first and second columns of $\hat{\mathbf{C}}$ and the matrix $\mathbf{z}_1 \in \mathbb{C}^{2 \times 2}$ is constructed as [84]

$$\mathbf{z}_1 = \boldsymbol{\Omega}^\dagger \boldsymbol{\Lambda} \boldsymbol{\Omega}, \quad (5.34)$$

where $\boldsymbol{\Omega} = [\mathbf{d}_x, \mathbf{d}_y] \in \mathbb{C}^{MN \times 2}$ and $\boldsymbol{\Lambda} = \mathbb{E} \left[\tilde{\mathbf{p}}_0 \tilde{\mathbf{p}}_0^\dagger \right] \in \mathbb{C}^{MN \times MN}$ with $\tilde{\mathbf{p}}_0 = \hat{\mathbf{p}}|_{\mathbf{s}=\mathbf{s}_0} = \mathbf{n} - \left(\frac{\mathbf{w}^\dagger \mathbf{n}}{\mathbf{w}^\dagger \mathbf{s}_0} \right) \mathbf{s}_0$.

Standard optimization theory argumentation leads to the following expression for the optimal solution

$$\hat{\mathbf{C}}_{\text{opt}} = \mathbf{z}_1^{-1} \hat{\mathbf{B}} \left(\text{Re} \left\{ \hat{\mathbf{B}}^\dagger \mathbf{z}_1^{-1} \hat{\mathbf{B}} \right\} \right)^{-1}. \quad (5.35)$$

Algorithm 9 FDA-AGMP-CC**Require:** $u_0, \delta_0, \mathbf{Q}, \mathbf{s}_0, \mathbf{s}_u, \mathbf{s}_\delta, \mathbf{z}$ **Ensure:** A solution $\hat{u}_{\text{AGMP-CC}}, \hat{\delta}_{\text{AGMP-CC}}$ to (5.11) .

1. Compute $\mathbf{w} = \mathbf{Q}^{-1}\mathbf{s}_0$, $\mathbf{d}_x = \mathbf{Q}^{-1}\mathbf{s}_u$, and $\mathbf{d}_y = \mathbf{Q}^{-1}\mathbf{s}_\delta$;
2. Evaluate $\hat{\mathbf{r}} = [\tilde{r}_x, \tilde{r}_y]^\top$ with \tilde{r}_α given by (5.28);
3. Compute

- $\hat{\boldsymbol{\mu}} = [\tilde{\mu}_x, \tilde{\mu}_y]^\top$ with $\tilde{\mu}_\alpha = \frac{\mathbf{d}_\alpha^\dagger \mathbf{s}_0}{\mathbf{w}^\dagger \mathbf{s}_0}$, $\alpha \in \{x, y\}$.

- $\hat{\mathbf{B}} = \begin{bmatrix} B_{xu} & B_{x\delta} \\ B_{yu} & B_{y\delta} \end{bmatrix}$, with
 $B_{\alpha h} = \frac{\mathbf{d}_\alpha^\dagger \mathbf{s}_h \mathbf{w}^\dagger \mathbf{s}_0 - \mathbf{d}_\alpha^\dagger \mathbf{s}_0 \mathbf{w}^\dagger \mathbf{s}_h}{(\mathbf{w}^\dagger \mathbf{s}_0)^2}$, $\alpha \in \{x, y\}, h \in \{u, \delta\}$.

- \mathbf{z}_1 according to (5.34).

- $\hat{\mathbf{C}}_{\text{opt}}$ using (5.35).

4. Evaluate $\mathbf{h}_{\text{AGMP-CC}}$ using (5.36);

5. Project $\mathbf{h}_{\text{AGMP-CC}}$ onto $[-1, 1] \times [-\frac{\Delta f}{B}, \frac{\Delta f}{B}]$ to get an estimate $\hat{\mathbf{h}}_{\text{AGMP-CC}}$ complying with the problem constraints.

Output $[\hat{u}_{\text{AGMP-CC}}, \hat{\delta}_{\text{AGMP-CC}}]^\top = \hat{\mathbf{h}}_{\text{AGMP-CC}}$.

Finally, the estimates of u and δ can be obtained according to (5.27), i.e.,

$$\mathbf{h}_{\text{AGMP-CC}} = \text{Re} \left\{ \left(\text{Re} \left\{ \hat{\mathbf{B}}^\dagger \mathbf{z}_1^{-1} \hat{\mathbf{B}} \right\} \right)^{-1} \hat{\mathbf{B}}^\dagger \mathbf{z}_1^{-1} (\hat{\mathbf{r}} - \hat{\boldsymbol{\mu}}) \right\} + \mathbf{h}_0. \quad (5.36)$$

The complete procedure for the AGMP-CC is provided in **Algorithm 9**.

Along the same line of reasoning followed to introduce the DSC-AMP estimator, a refined version of the AGMP-CC procedure can be conceived via a second execution of **Algorithm 9**, with inputs induced by $\hat{u}_{\text{AGMP-CC}}$ and $\hat{\delta}_{\text{AGMP-CC}}$. This procedure will be denoted as the DSC-AGMP-CC and $\hat{u}_{\text{DSC-AGMP-CC}}$ and $\hat{\delta}_{\text{DSC-AGMP-CC}}$ indicate the resulting estimates.

5.2.3 Discussion on the Computational Complexity

In this subsection, the assessment of the computational burden involved by the proposed estimators is provided³. To this end, the following equivalent expression of the objective function (5.12), that can be derived according to the results of Appendix B of [66], is used for ML and CD methods

$$P(u, \delta) = \frac{|\mathbf{a}(\delta)^\dagger \mathbf{v}(u)|}{\mathbf{a}(\delta)^\dagger \hat{\mathbf{T}}(u) \mathbf{a}(\delta)}, \quad (5.37)$$

where

$$\mathbf{v}(u) = \mathbf{c}(u) \odot \tilde{\mathbf{z}}(u) \in \mathbb{C}^M, \quad (5.38)$$

with

- $\tilde{\mathbf{z}}(u) = \sum_{l=1}^N [\mathbf{b}(u)]_l^* \bar{\mathbf{z}}_l \in \mathbb{C}^M$.
- $\bar{\mathbf{z}} = \mathbf{Q}^{-1} \mathbf{z} = [\bar{\mathbf{z}}_1^T, \dots, \bar{\mathbf{z}}_N^T]^T \in \mathbb{C}^{MN}$, $\bar{\mathbf{z}}_l \in \mathbb{C}^M$

and

$$\hat{\mathbf{T}}(u) = \mathbf{C}^\dagger(u) \mathbf{T}(u) \mathbf{C}(u) \in \mathbb{H}^M, \quad (5.39)$$

with

- $\mathbf{C}(u) = \mathbf{diag}(\mathbf{c}(u)) \in \mathbb{C}^{M \times M}$.
- $\mathbf{T}(u) = \sum_{k=1}^N \sum_{l=1}^N \left([\mathbf{b}(u)]_l^* \tilde{\mathbf{Q}}_{l,k} [\mathbf{b}(u)]_k \right)$.
- $\mathbf{Q}^{-1} = \begin{bmatrix} \tilde{\mathbf{Q}}_{1,1} & \dots & \tilde{\mathbf{Q}}_{1,N} \\ \vdots & \ddots & \vdots \\ \tilde{\mathbf{Q}}_{N,1} & \dots & \tilde{\mathbf{Q}}_{N,N} \end{bmatrix}$, where $\tilde{\mathbf{Q}}_{l,k} \in \mathbb{C}^{M \times M}$, $(l, k) \in \{1, \dots, N\}^2$.

Therefore, given $\bar{\mathbf{z}}$ (which requires $\mathcal{O}((MN)^2)$ operations), for any given u (5.38) can be evaluated with a computational complexity $\mathcal{O}(MN)$. Indeed, $\mathcal{O}(M)$ operations are needed in the Hadamard product, while the

³Without loss of generality, it is assumed that \mathbf{Q}^{-1} is pre-computed off-line. If adaptive implementations of the proposed methods are considered, i.e., \mathbf{Q} is estimated resorting to secondary data, the term $(MN)^2 K$ has to be added in the computational complexity expressions, where $K \geq MN$ is the size of the secondary data set.

evaluation of $\hat{\mathbf{z}}(u)$ involves $\mathcal{O}(MN)$ operations. Besides, the computational complexity connected with (5.39) is $\mathcal{O}((MN)^2)$, where the main task is the evaluation of $\mathbf{T}(u)$, demanding $\mathcal{O}((MN)^2)$ operations.

ML estimator. Being the feasible set \mathcal{I} given by the Cartesian product of \mathcal{I}_u and \mathcal{I}_δ , the 2-D search is tantamount to performing N_u 1-D search, each at a given value of u , with δ the optimization variable. Exploiting (5.37), for $\hat{\mathbf{T}}(u)$ and $\mathbf{v}(u)$ given, each 1-D search requires $\mathcal{O}(N_\delta M^2)$ operations. Hence, the overall computational complexity is $\mathcal{O}(N_u((NM)^2 + N_\delta M^2))$, where $(NM)^2$ amounts for the computation of $\hat{\mathbf{T}}(u)$ and $\mathbf{v}(u)$.

FDA-CD. At each iteration of the CD method, the 1-D searches w.r.t. u and δ require $\mathcal{O}(N_u M^2)$ and $\mathcal{O}(N_\delta M^2)$ operations, respectively, provided that $\hat{\mathbf{T}}(u)$ and $\mathbf{v}(u)$ are pre-computed for each $u \in \mathcal{I}_u$. This last task involves $\mathcal{O}(N_u(MN)^2)$ operations. To proceed further, let us denote by $N_{it,CD} = N_{it,CD,x} + N_{it,CD,y}$ the total number of iterations of the CD method, where $N_{it,CD,x}$ and $N_{it,CD,y}$ refer to the number of iterations required by the first and second cycle, respectively (the former with initial search direction given by u , the latter by δ). Hence, the implementation of FDA-CD requires $\mathcal{O}(N_u(MN)^2 + N_{it,CD}(N_u + N_\delta)M^2)$ operations, where the first term is due the computation of $\hat{\mathbf{T}}(u)$ and $\mathbf{v}(u)$ for any $u \in \mathcal{I}_u$.

FDA-AMP and FDA-AGMP-CC. The evaluation of the terms \mathbf{w} , \mathbf{d}_x , and \mathbf{d}_y involved in step 1 of both FDA-AMP and FDA-AGMP-CC is the most demanding task and requires $\mathcal{O}((NM)^2)$ operations. Furthermore, as to FDA-AMP, $\mathcal{O}(NM)$ operations are connected with the computation of \mathbf{r} , $\boldsymbol{\mu}$, and \mathbf{C} , whereas $\mathcal{O}(1)$ operations are necessary to perform steps 4 and 5. On the other hand, with reference to FDA-AGMP-CC, $\mathcal{O}(NM)$ amounts for the computation of $\hat{\mathbf{r}}$, $\hat{\boldsymbol{\mu}}$, and $\hat{\mathbf{B}}$, while $\mathcal{O}((NM)^2)$ operations are necessary to determine \mathbf{z}_1 . Besides, the evaluation of $\hat{\mathbf{C}}_{opt}$ as well as the execution of steps 4 and 5 need $\mathcal{O}(1)$ operations. This implies that the computational complexity for both FDA-AMP and FDA-AGMP-CC is $\mathcal{O}((NM)^2)$.

Summarizing, the computational complexity connected with the implementation (either non-adaptive or adaptive) of the devised estimators is reported in Table 5.1. Before concluding this subsection, it is worth observing that the ML procedure is the most demanding. In fact, it exhibits a computational burden always higher than the monopulse-based

Table 5.1. Computational complexity of the considered estimators.

Methods	Computational Costs	
	Non-adaptive implementation	Adaptive implementation
ML	$\mathcal{O}(N_u((NM)^2 + N_\delta M^2))$	$\mathcal{O}(N_u((NM)^2 + N_\delta M^2) + (NM)^2 K)$
FDA-CD	$\mathcal{O}(N_u(MN)^2 + N_{it,CD}(N_u + N_\delta)M^2)$	$\mathcal{O}(N_u(MN)^2 + N_{it,CD}(N_u + N_\delta)M^2 + (NM)^2 K)$
FDA-AMP	$\mathcal{O}((NM)^2)$	$\mathcal{O}((NM)^2 K)$
FDA-AGMP-CC	$\mathcal{O}((NM)^2)$	$\mathcal{O}((NM)^2 K)$

procedures, and requires more operations than the CD method, as long as $N_u N_\delta > N_{it,CD}(N_u + N_\delta)$, which is a condition always met in our numerical experiments.

5.3 CRLB for FDA-MIMO radar

To shed light on the statistical efficiency of proposed estimators, the CRLB for angle and incremental range is derived. Let us first define three auxiliary vectors, i.e., $\boldsymbol{\xi} = \mathbf{Q}^{-\frac{1}{2}} \mathbf{s} \in \mathbb{C}^{MN}$, $\boldsymbol{\xi}_u = \mathbf{Q}^{-\frac{1}{2}} \mathbf{s}_u \in \mathbb{C}^{MN}$, and $\boldsymbol{\xi}_\delta = \mathbf{Q}^{-\frac{1}{2}} \mathbf{s}_\delta \in \mathbb{C}^{MN}$. The CRLB for angle and incremental range is derived assuming unknown β_1 .

In this respect, let us introduce the vector $\boldsymbol{\gamma} \in \mathbb{R}^4$ containing the real-valued unknown parameters, $u, \delta, \bar{\beta}_1 = \text{Re}\{\beta_1\}, \tilde{\beta}_1 = \text{Im}\{\beta_1\}$, i.e., $\boldsymbol{\gamma} = [u, \delta, \bar{\beta}_1, \tilde{\beta}_1]^T$. Hence, the CRLB for the unknown parameters is given by the diagonal elements of $\mathbf{D}_\boldsymbol{\gamma} = \mathbf{F}^{-1} \in \mathbb{R}^{4 \times 4}$, where the Fisher Information Matrix (FIM) $\mathbf{F} \in \mathbb{R}^{4 \times 4}$ can be computed via the Slepian-Bangs formula [105, p. 927, eq. 8.34], which yields

$$\begin{aligned}
\mathbf{F} &= 2 \text{Re} \left\{ \left(\frac{\partial \beta_1 \mathbf{s}}{\partial \boldsymbol{\gamma}^T} \right)^\dagger \mathbf{Q}^{-1} \left(\frac{\partial \beta_1 \mathbf{s}}{\partial \boldsymbol{\gamma}^T} \right) \right\} \\
&= 2 \text{Re} \left\{ [\beta_1 \mathbf{s}_u, \beta_1 \mathbf{s}_\delta, \mathbf{s}, j\mathbf{s}]^\dagger \left(\mathbf{Q}^{-\frac{1}{2}} \right)^\dagger \left(\mathbf{Q}^{-\frac{1}{2}} \right) [\beta_1 \mathbf{s}_u, \beta_1 \mathbf{s}_\delta, \mathbf{s}, j\mathbf{s}] \right\} \quad (5.40) \\
&= 2 \text{Re} \left\{ [\beta_1 \boldsymbol{\xi}_u, \beta_1 \boldsymbol{\xi}_\delta, \boldsymbol{\xi}, j\boldsymbol{\xi}]^\dagger [\beta_1 \boldsymbol{\xi}_u, \beta_1 \boldsymbol{\xi}_\delta, \boldsymbol{\xi}, j\boldsymbol{\xi}] \right\}.
\end{aligned}$$

Hence, the FIM can be expressed in block form as

$$\mathbf{F} = 2 \begin{bmatrix} \mathbf{F}_{11} & \mathbf{F}_{12} \\ \mathbf{F}_{21} & \mathbf{F}_{22} \end{bmatrix}, \quad (5.41)$$

where $\mathbf{F}_{11} \in \mathbb{R}^{2 \times 2}$, $\mathbf{F}_{12} \in \mathbb{R}^{2 \times 2}$, $\mathbf{F}_{21} \in \mathbb{R}^{2 \times 2}$, and $\mathbf{F}_{22} \in \mathbb{R}^{2 \times 2}$ are respectively given by

$$\mathbf{F}_{11} = \begin{bmatrix} |\beta_1|^2 \|\boldsymbol{\xi}_u\|^2 & |\beta_1|^2 \operatorname{Re} \left\{ \boldsymbol{\xi}_u^\dagger \boldsymbol{\xi}_\delta \right\} \\ |\beta_1|^2 \operatorname{Re} \left\{ \boldsymbol{\xi}_u^\dagger \boldsymbol{\xi}_\delta \right\} & |\beta_1|^2 \|\boldsymbol{\xi}_\delta\|^2 \end{bmatrix}, \quad (5.42a)$$

$$\mathbf{F}_{12} = \begin{bmatrix} \operatorname{Re} \left\{ \boldsymbol{\xi}_u^\dagger \boldsymbol{\xi} \beta_1^* \right\} & -\operatorname{Im} \left\{ \boldsymbol{\xi}_u^\dagger \boldsymbol{\xi} \beta_1^* \right\} \\ \operatorname{Re} \left\{ \boldsymbol{\xi}_\delta^\dagger \boldsymbol{\xi} \beta_1^* \right\} & -\operatorname{Im} \left\{ \boldsymbol{\xi}_\delta^\dagger \boldsymbol{\xi} \beta_1^* \right\} \end{bmatrix}, \quad (5.42b)$$

$$\mathbf{F}_{21} = \mathbf{F}_{12}^T \quad (5.42c)$$

$$\mathbf{F}_{22} = \|\boldsymbol{\xi}\|^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (5.42d)$$

Then, \mathbf{D}_γ can be calculated as the inverse of \mathbf{F} , i.e.,

$$\mathbf{D}_\gamma = \mathbf{F}^{-1} = \frac{1}{2} \begin{bmatrix} \mathbf{G}_1^{-1} & \mathbf{G}_2 \\ \mathbf{G}_3 & \mathbf{G}_4^{-1} \end{bmatrix}, \quad (5.43)$$

where

$$\begin{aligned} \mathbf{G}_1 &= \mathbf{F}_{11} - \mathbf{F}_{12} \mathbf{F}_{22}^{-1} \mathbf{F}_{21} \\ &= |\beta_1|^2 \begin{bmatrix} \|\boldsymbol{\xi}_u\|^2 - \frac{|\boldsymbol{\xi}_u^\dagger \boldsymbol{\xi}|^2}{\|\boldsymbol{\xi}\|^2} & \operatorname{Re} \left\{ \boldsymbol{\xi}_u^\dagger \boldsymbol{\xi}_\delta \right\} - \frac{\operatorname{Re} \left\{ \boldsymbol{\xi}_u^\dagger \boldsymbol{\xi} \boldsymbol{\xi}^\dagger \boldsymbol{\xi}_\delta \right\}}{\|\boldsymbol{\xi}\|^2} \\ \operatorname{Re} \left\{ \boldsymbol{\xi}_u^\dagger \boldsymbol{\xi}_\delta \right\} - \frac{\operatorname{Re} \left\{ \boldsymbol{\xi}_u^\dagger \boldsymbol{\xi} \boldsymbol{\xi}^\dagger \boldsymbol{\xi}_\delta \right\}}{\|\boldsymbol{\xi}\|^2} & \|\boldsymbol{\xi}_\delta\|^2 - \frac{|\boldsymbol{\xi}_\delta^\dagger \boldsymbol{\xi}|^2}{\|\boldsymbol{\xi}\|^2} \end{bmatrix} \\ &= \mathbf{F}_{11} - \mathbf{F}_{12} \mathbf{F}_{22}^{-1} \mathbf{F}_{21} \in \mathbb{R}^{2 \times 2}. \end{aligned} \quad (5.44)$$

Hence, the CRLB for u and δ (analytical details is reported in Appendix D.2) are given by

$$D_u = \frac{|\beta_1|^2}{2 \det(\mathbf{G}_1)} \left(\|\boldsymbol{\xi}_\delta\|^2 - \frac{|\boldsymbol{\xi}_\delta^\dagger \boldsymbol{\xi}|^2}{\|\boldsymbol{\xi}\|^2} \right), \quad (5.45)$$

Table 5.2. Simulation Parameters of FDA-MIMO Radar

Parameter	Symbol	Value
transmit elements	M	4
receive elements	N	10
bandwidth	B	1 MHz
frequency increment	Δf	0.5 MHz
nominal target angle	u_0	0
nominal target incremental range	δ_0	0
possible angle of the target	u	$-\frac{0.891}{2(N+M)}, 0, \frac{0.891}{2(N+M)}$
possible incremental range of the target	δ	$-\frac{\Delta f}{2B}, -\frac{\Delta f}{4B}, 0, \frac{\Delta f}{4B}, \frac{\Delta f}{2B}$
angle of the coherent repeater 1	u_1	$5 \frac{0.891}{(N+M)}$
incremental range of the coherent repeater 1	δ_1	$\frac{\Delta f}{3B}$
angle of the coherent repeater 2	u_2	$-6 \frac{0.891}{(N+M)}$
incremental range of the coherent repeater 2	δ_2	$-\frac{\Delta f}{8B}$

and

$$D_\delta = \frac{|\beta_1|^2}{2 \det(\mathbf{G}_1)} \left(\|\boldsymbol{\xi}_u\|^2 - \frac{|\boldsymbol{\xi}_u^\dagger \boldsymbol{\xi}|^2}{\|\boldsymbol{\xi}\|^2} \right). \quad (5.46)$$

5.4 Performance Analysis

In this section, numerical examples are provided to assess the performance of the proposed methods to estimate the target incremental range and angle of arrival with reference to an FDA-MIMO radar sensing system. To this end, a transmit ULA with $M = 4$ elements and a receive ULA with $N = 10$ elements, both pointing toward the boresight direction (i.e., $u_0 = 0$), are considered. Moreover, orthogonal baseband signals are radiated, i.e., $\mathbf{R} = \mathbf{I}$, whereas the spacing among the antennas is set to $d = \lambda_0/2$. Resorting to Monte Carlo technique, the performance of the proposed methods is evaluated for both range and angle estimations. As figure of merit, the Root Mean Square Error (RMSE) is considered, which is computed as

$$\widehat{\text{RMSE}}_u = \sqrt{\frac{1}{M_C} \sum_{i=1}^{M_C} \|u - \hat{u}_i\|^2}, \quad (5.47)$$

and

$$\widehat{\text{RMSE}}_{\delta} = \sqrt{\frac{1}{M_C} \sum_{i=1}^{M_C} \|\delta - \hat{\delta}_i\|^2}, \quad (5.48)$$

where $M_C = 500$ indicates the number of Monte Carlo independent trials, u and δ denote the actual DOA and incremental range of the target, whereas \hat{u}_i and $\hat{\delta}_i$ are the estimates provided at the i -th trial by a given technique.

The performance of the CD, AMP, and AGMP-CC algorithms are evaluated for several values of SINR, defined according to [40], as

$$\text{SINR} = |\beta_1|^2 \mathbf{s}^\dagger(u_0, \delta_0) \mathbf{Q}^{-1} \mathbf{s}(u_0, \delta_0). \quad (5.49)$$

Besides, $B = 1$ MHz and $\Delta f = 0.5$ MHz are considered, with the nominal parameters values set as $u_0 = 0$ and $\delta_0 = 0$, respectively. Tapered version of AMP and AGMP-CC, referred to as AMP_t and AGMP-CC_t , respectively, are considered too. In particular, Taylor and Bayliss tapers [81], both with SideLobe Level (SLL) = 30 dB and $\bar{n} = 4$, are used for sum and difference beamforming, respectively, where \bar{n} indicates the number of nearly constant-level sidelobes adjacent to the mainlobe. The DSC versions of AMP, AMP_t , and AGMP-CC_t algorithms, respectively denoted as DSC-AMP, DSC- AMP_t , and DSC- AGMP-CC_t , are also included in the reported analysis. Besides, since AGMP-CC is equivalent to AMP when tapering is not applied [84], only the AMP and DSC-AMP curves are displayed in the figures, without loss of generality. Finally, the CRLB for both angle and incremental range estimation is used as performance benchmarks.

In the following subsections, two different interference scenarios are examined. In the former, the useful signal is buried in white Gaussian noise; in the latter, white Gaussian noise plus two coherent repeaters, is considered. The values of the parameters involved in the analyzed case studies are listed in Table 5.2.

5.4.1 White Noise Interference Scenario

Within this subsection, the overall disturbance is assumed composed of white Gaussian interference only. Therefore, its covariance matrix is modeled as $\mathbf{Q} = \sigma_n^2 \mathbf{I}_{MN}$, where σ_n^2 is the power level, assumed without

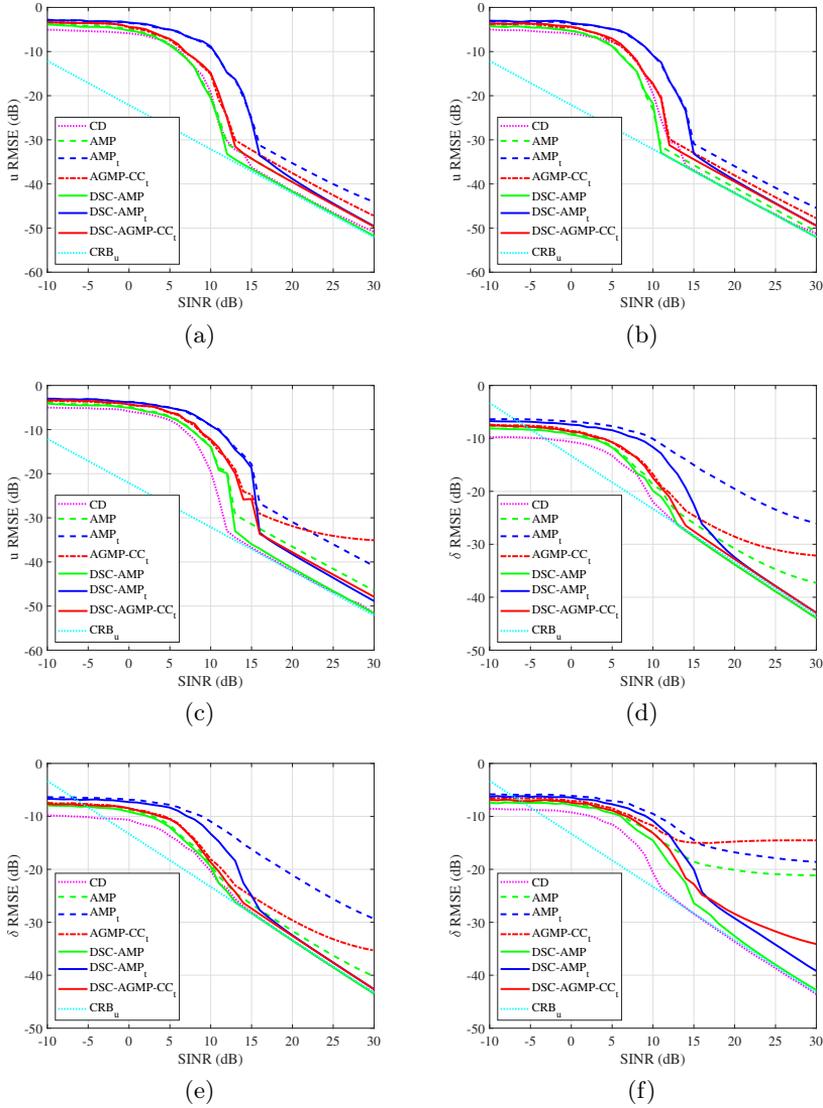


Figure 5.3. Comparison of RMSE (dB) assuming white noise for some u and δ : (a) and (d) $u = -0.891/(2(N+M))$, $\delta = -\Delta f/(4B)$, (b) and (e) $u = 0$, $\delta = \Delta f/(4B)$, (c) and (f) $u = 0.891/(2(N+M))$, $\delta = \Delta f/(2B)$. The RMSE analysis w.r.t. u is reported in (a), (b), and (c), whereas that w.r.t. δ is reported in (d), (e), and (f).

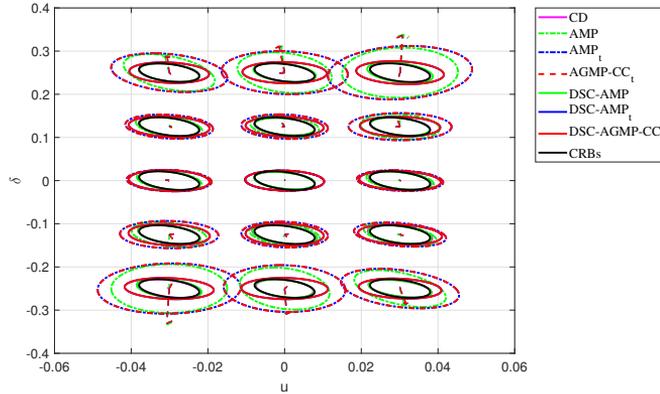


Figure 5.4. Bias and variance analysis assuming white noise for 15 different pairs (u, δ) and $\text{SINR} = 20\text{dB}$.

loss of generality equal to 0 dB.

Fig. 5.3 illustrates the **RMSE** versus **SINR** for three case studies assuming different values of the true angles and incremental ranges of the target. In particular, Figs. 5.3 (a) and (d) consider $u = -0.891/(2(N + M))$, $\delta = -\Delta f/(4B)$, Figs. 5.3 (b) and (e) assume $u = 0$, $\delta = \Delta f/(4B)$, while Figs. 5.3 (c) and (f) suppose $u = 0.891/(2(N + M))$, $\delta = \Delta f/(2B)$. The **RMSE** analysis w.r.t. u is reported in Figs. 5.3 (a), (b), and (c), whereas that w.r.t. δ in Figs. 5.3 (d), (e), and (f).

Inspection of the curves shows that the higher the **SINR** the lower the **RMSE** of all the estimators. Besides, some of them achieve performance levels comparable with the **CRLB** benchmark when the **SINR** is sufficiently high, for all the considered scenarios. Specifically, both the angle and incremental range estimates provided by the **CD** method are very close to their true values. Similar results hold for the **DSC-AMP**, with **RMSE** curves almost overlapped, especially for the high **SINR** regime, with those pertaining to **CD** technique. The plots also highlight that the **DSC** versions of the monopulse procedures outperform their corresponding single iteration (i.e., without **DSC**) counterparts for all the range of **SINR** values and all the considered experimental setups. Besides, the tapered procedures exhibit a performance degradation w.r.t. the unweighted counterparts. In particular, even with the further iteration, i.e., considering **DSC-AMP_t** and **DSC-AGMP-CC_t** at high **SINR** values, the **CRLB** is not attained, in

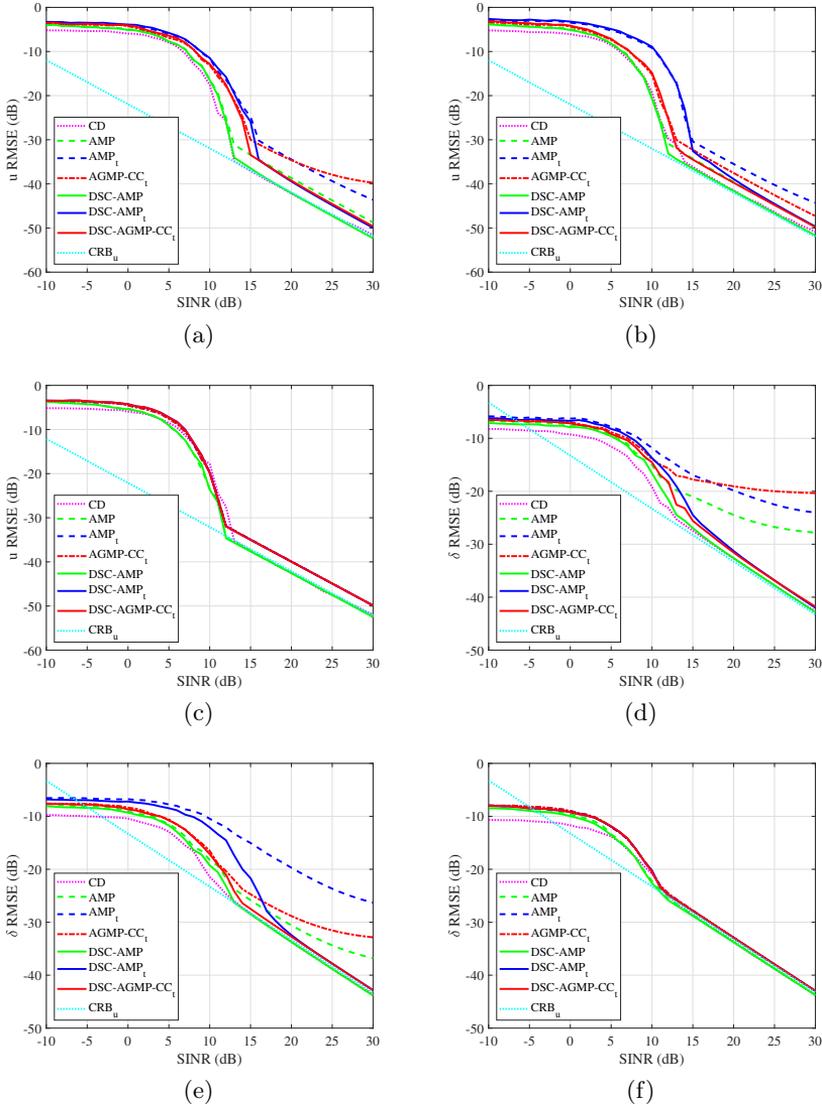


Figure 5.5. Comparison of RMSE (dB) for some u and δ : (a) and (d) $u = 0.891/(2(N+M))$, $\delta = -\Delta f/(2B)$, (b) and (e) $u = -0.891/(2(N+M))$, $\delta = -\Delta f/(4B)$, (c) and (f) $u = 0$, $\delta = 0$, assuming two coherent repeaters, with $\text{SINR} = 30\text{dB}$, located at $u_1 = 5(0.891/(N+M))$ and $u_2 = -6(0.891/(N+M))$ with incremental range of $\delta_1 = \Delta f/(3B)$ and $\delta_2 = -\Delta f/(8B)$, respectively. The RMSE analysis w.r.t. u is reported in (a), (b), and (c), whereas that w.r.t. δ is reported in (d), (e), and (f).

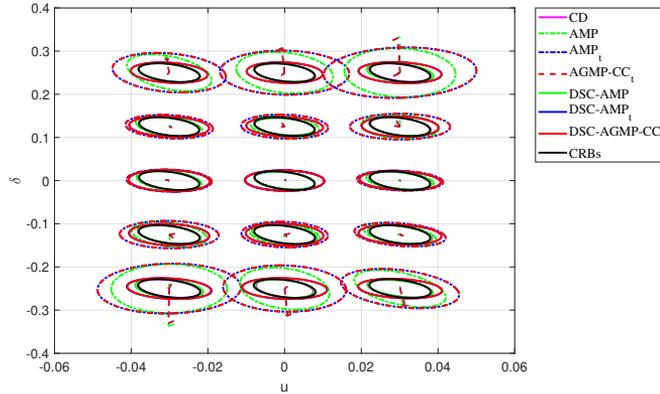


Figure 5.6. Bias and variance analysis for 15 different pairs (u, δ) assuming $\text{SINR} = 20\text{dB}$ and two coherent repeaters, with $\text{SINR} = 30\text{dB}$, located at $u_1 = 5(0.891/(N+M))$ and $u_2 = -6(0.891/(N+M))$ with incremental range of $\delta_1 = \Delta f/(3B)$ and $\delta_2 = -\Delta f/(8B)$, respectively.

both u and δ domains. Furthermore, at low SINR , smaller RMSE values than the CRLB benchmark are observed indicating that all the proposed estimators exhibit a bias under this SINR regime due to an upper bound to the mean square error induced by the enforced constraint.

To further shed light on performance of the different procedures, Fig. 5.4 displays the bias and variance ellipses of the proposed estimators for a grid of 15 points and $\text{SINR} = 20\text{dB}$. The ellipses corresponding to the CRLB are also reported for comparison. The simulation assumes the same interference environment as in Fig. 5.3. The results reveal that the AMP (or equivalently the AGMP-CC) method, as well as AMP_t and AGMP-CC_t , exhibit a bias in both the u and δ domains, with a much more marked effect on the δ component. As expected, the bias is almost corrected by the second iteration of the double-step implementation, i.e., DSC-AMP and DSC-AGMP-CC , thus leading to a performance very close to the CRLB . On the other hand, despite the second iteration, a small but noticeable bias persists in both DSC-AMP_t and DSC-AGMP-CC_t . Therefore, the bias and variance analysis confirms that all the tapered monopulse algorithms experience a bias in both u and δ domains, which is the main reason for the deviations of these estimators from the CRLB (at high SINR). Finally, under this SINR regime, no bias is exhibited by the CD method with

variance ellipses almost overlapped with CRLBs.

5.4.2 Coherent Repeaters In The Interference Scenario

The interference scenario considered in this subsection accounts for the presence of two coherent signals at the same range ring as the target (this can be the situation of stand-off or escort jamming configuration), with **SINR** equal to 30 dB, impinging on the array from $u_1 = 5$ ($0.891/(N+M)$) and $u_2 = -6$ ($0.891/(N+M)$), with $\delta_1 = \Delta f/(3B)$ and $\delta_2 = -\Delta f/(8B)$, respectively. As a consequence, for this specific environment, the interference covariance matrix is modeled as $\mathbf{Q} = \sigma_n^2 \mathbf{I}_{MN} + \sigma_c^2 \Sigma_c$ where σ_n^2 and σ_c^2 are the noise and interferers powers⁴, respectively, with $\sigma_c^2/\sigma_n^2 = 30$ dB. Besides,

$$\Sigma_c = \sum_{i=1}^2 \mathbf{s}(u_i, \delta_i) \mathbf{s}^\dagger(u_i, \delta_i), \quad (5.50)$$

where $\mathbf{s}(u_i, \delta_i)$, u_i , and δ_i are the steering vector, the angle and the incremental range of the i -th coherent repeater, respectively.

The **RMSE** versus **SINR** is displayed in Fig. 5.5, where in each subfigure different values of the true angle and incremental range of the target are considered. In particular, Figs. 5.5 (a) and (d) refer to $u = 0.891/(2(N+M))$, $\delta = -\Delta f/(2B)$, Figs. 5.5 (b) and (e) assume $u = -0.891/(2(N+M))$, $\delta = -\Delta f/(4B)$, while Figs. 5.5 (c) and (f) consider $u = 0$, $\delta = 0$. The **RMSE** analysis w.r.t. u is reported in Figs. 5.5 (a), (b), and (c), whereas that w.r.t. δ in Figs. 5.5 (d), (e), and (f). Inspection of the curves highlights that the considered estimators exhibit performance behaviors comparable to those obtained in the white noise only scenario. In other words, the methods correctly estimate the parameters of a target located in the main beam without experiencing significant performance degradation due to possible coherent interference. In particular, the **CD** and **DSC-AMP** procedures achieve similar **RMSE** levels with performance very close to **CRLB**, at high **SINR**. Furthermore, the bias and variance analysis reported in Fig. 5.6, for **SINR** = 20dB, does not show specific differences w.r.t. the noise-only case, corroborating the effectiveness of the **DSC** technique to reduce the bias and thus improve the performance.

⁴The white noise power level σ_n^2 can be again assumed, without loss of generality, equal to 0 dB.

5.5 Conclusion

The problem of target angle and incremental range estimation with an FDA-MIMO radar has been investigated using a single data snapshot. At the design stage, three estimators, with lower computational complexity than the 2-D grid search procedure required by the ML estimator, have been devised. Firstly, a CD algorithm has been proposed, which reduces the 2-D search to a sequence of 1-D problems alternating between the optimization over each variable while keeping the other fixed. Then, two approximated estimators, i.e., the AMP leveraging real slope and bias correction and the AGMP-CC capitalizing complex slope and bias correction, have been devised. The performance of the proposed estimators, also considering double-step and tapered versions, has been assessed in terms of RMSE versus SINR. Comparisons with benchmark limits, along with an extensive bias and variance analysis, have also been conducted. The results (for both white and colored interference) have pinpointed the effectiveness of the devised estimators to reliably estimate the angle and incremental range of the target in all the considered case studies. In particular, the CD method achieves a performance level very close to the theoretical CRLB when the SINR is sufficiently high.

Chapter 6

Adaptive Target Detection with Polarimetric FDA-MIMO Radar

The problem of adaptive radar detection with a polarimetric FDA-MIMO radar is addressed in this chapter¹. At the design stage, the target detection problem is formulated as a composite hypothesis test, with the unknowns given by the target angle, incremental range (target displacement with respect to the center of the occupied range cell), and scattering matrix, as well as the interference covariance matrix. The formulated detection problem is handled by resorting to sub-optimal design strategies based on the GLR criterion. The resulting detectors demand, under the H_1 hypothesis, the solution of a box-constrained optimization problem for which several iterative techniques, i.e., the LAM, the GPM, and the CD algorithms, are exploited. At the analysis stage, the performance of the proposed architectures, which ensure the bounded CFAR property, is evaluated via Monte Carlo simulations and compared with the benchmarks in both white and colored disturbance.

The chapter is organized as follows. The signal model for the polarimetric FDA-MIMO radar is presented in Section 6.1. In Section 6.2, the

¹©2022 IEEE. Reprinted, with permission, from L. Lan, M. Rosamilia, A. Aubry, A. De Maio, G. Liao and J. Xu, “Adaptive Target Detection with Polarimetric FDA-MIMO Radar,” IEEE Transactions on Aerospace and Electronic Systems, 2022.

detection problem for a polarimetric FDA-MIMO radar is formulated. Section 6.3 investigates optimization strategies for the computation of GLRT-based detectors and studies the bounded CFARness as well as the computational complexity of the derived algorithms. The performance analysis is addressed in Section 6.4, whereas conclusions are drawn in Section 6.5.

6.1 Signal Model For Polarimetric FDA-MIMO Radar

Let us consider a colocated FDA-MIMO radar consisting of M transmit and N receive pairs of crossed dipoles, which are placed in a uniform linear array configuration for both transmission and reception on the xy -plane with array's inter-element spacing equal to d (see Fig. 6.1). An elliptical electromagnetic polarization for both the transmitted and received signals [13] is considered. It is also assumed that a point-like target is located in the far-field at the angle θ_t and range R_t in the yz -plane.

For the FDA configuration, a frequency increment Δf is introduced at each crossed dipole composing the transmit array, with the first one being the reference. As a consequence, the carrier frequency at the m -th ($m = 1, \dots, M$) transmitting crossed dipole is [66]

$$f_m = f_0 + (m - 1)\Delta f, \quad m = 1, 2, \dots, M, \quad (6.1)$$

where f_0 indicates the reference carrier. Besides, each element transmits a specific base-band phase-modulated pulse composed of P subpulses. The resulting complex envelope of the radio frequency signal radiated by the m -th crossed dipole can be expressed as

$$s_m(t) = \sqrt{\frac{E_t}{P}} x_m(t) e^{j2\pi f_m t}, \quad 0 \leq t \leq T_p, \quad (6.2)$$

with E_t the transmitted energy, T_p the radar pulse duration, and

$$x_m(t) = \frac{1}{\sqrt{\tau_b}} \sum_{p=1}^P \varphi_m(p) u \left[\frac{t - (p-1)\tau_b}{\tau_b} \right], \quad m = 1, \dots, M, \quad (6.3)$$

where $\tau_b = \frac{T_p}{P}$, $u(t)$ is the asymmetric rect function, i.e., $u(t) = 1$ as long as $0 \leq t \leq 1$ and zero elsewhere, and $\varphi_m(p) = e^{j\phi_m(p)}$, with $\phi_m(p) \in [0, 2\pi]$ the p -th entry of the phase code on the m -th transmit array element.

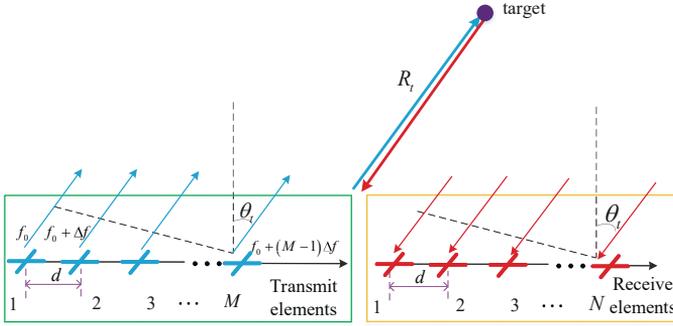


Figure 6.1. Illustration of the polarimetric FDA-MIMO radar system.

6.1.1 Received Signal for Polarimetric FDA-MIMO radar

Let us consider the far-field target having a constant polarimetric scattering matrix $\Sigma_t = \begin{bmatrix} \Sigma_{HH} & \Sigma_{VH} \\ \Sigma_{HV} & \Sigma_{VV} \end{bmatrix} \in \mathbb{C}^{2 \times 2}$ over the FDA-MIMO radar bandwidth, with Σ_{rt} , $r, t \in \{H, V\}$ the complex target scattering amplitude assuming a polarization ‘t’ on transmit and a polarization ‘r’ on receive².

As shown in Fig. 6.2, the H and V components received on the n -th polarimetric-spatial channel, i.e., $\mathbf{y}_n(t, \theta_t) = [y_n^H(t, \theta_t), y_n^V(t, \theta_t)]^T \in \mathbb{C}^2$, $n = 1, \dots, N$, are firstly multiplied by $e^{-j2\pi f_0 t}$, and then processed through a bank of M matched filters $h_l(t) = x_l^*(-t) e^{j2\pi \Delta f (l-1)t}$ ($l = 1, \dots, M$). Moreover, after sampling at the range gate of interest and under some mild technical conditions (see Appendix A of [66] for the single polarization case), the received polarimetric signal from the target can be

²For co-polarized channels $r = t$, and for cross-polarized channels $r \neq t$. For a reciprocal medium (e.g., no Faraday rotation of polarization), $\Sigma_{VH} = \Sigma_{HV}$ [91].

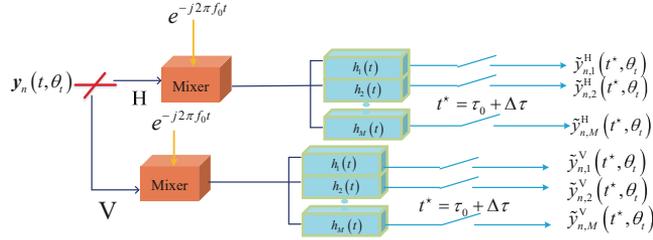


Figure 6.2. Signal processing chain at the n -th polarimetric receiver with multiple match-filters.

expressed as a $2MN \times 1$ -dimensional vector

$$\begin{aligned}
 \mathbf{y}_S(\theta_t, \Delta\tau) &= [\tilde{\mathbf{y}}_{1,1}^T(t^*, \theta_t), \dots, \tilde{\mathbf{y}}_{1,M}^T(t^*, \theta_t), \dots, \\
 &\quad \tilde{\mathbf{y}}_{N,1}^T(t^*, \theta_t), \dots, \tilde{\mathbf{y}}_{N,M}^T(t^*, \theta_t)]^T \\
 &= \boldsymbol{\alpha} \otimes (\mathbf{b}(\theta_t) \otimes [\mathbf{c}(\theta_t) \odot \mathbf{a}(\Delta\tau)]) \\
 &= \boldsymbol{\alpha} \otimes \mathbf{s}(\theta_t, \Delta\tau),
 \end{aligned} \tag{6.4}$$

where³

- $t^* = \tau_0 + \Delta\tau$ denotes the time instant when the data from the CUT are collected, with $\Delta\tau$ the unknown incremental delay w.r.t. the sampling time associated with the target range cell [66] and $\tau_0 = \frac{2R_t}{c}$ the envelope time delay, respectively;
- $\tilde{\mathbf{y}}_{n,l}(t^*, \theta_t) = [\tilde{y}_{n,l}^H(t^*, \theta_t), \tilde{y}_{n,l}^V(t^*, \theta_t)]^T \in \mathbb{C}^2, l = 1, \dots, M, n = 1, \dots, N$, is obtained by filtering the output of the n -th polarimetric channels (after multiplication by $e^{-j2\pi f_0 t}$) with the l -th matched filter and sampling the resulting signal in the CUT;
- $\boldsymbol{\alpha} = \boldsymbol{\Sigma}_t \mathbf{e} \sqrt{(E_t/P)} \in \mathbb{C}^2$;
- $\mathbf{e} = [E_H, E_V]^T \in \mathbb{C}^2$, with E_H and E_V the horizontal and vertical components of the electric field impinging on the target, respectively;

³Note that the vector $\boldsymbol{\alpha}$ can be also expressed in terms of the polarization auxiliary angle and phase difference [70, 78].

- $\mathbf{s}(\theta_t, \Delta\tau) = \mathbf{b}(\theta_t) \otimes [\mathbf{c}(\theta_t) \odot \mathbf{a}(\Delta\tau)] \in \mathbb{C}^{MN}$ the transmit-receive steering vector;
- $\mathbf{b}(\theta_t) = \left[1, e^{j2\pi \frac{d}{\lambda_0} \sin(\theta_t)}, \dots, e^{j2\pi \frac{d}{\lambda_0} (N-1) \sin(\theta_t)}\right]^T \in \mathbb{C}^N$ the angle-dependent receive steering vector, with $\lambda_0 = \frac{c}{f_0}$ the reference carrier wavelength;
- $\mathbf{c}(\theta_t) = \mathbf{R}^T \mathbf{d}(\theta_t) \in \mathbb{C}^M$ the angle-dependent transmit steering vector;
- $\mathbf{R} \in \mathbb{C}^{M \times M}$ is the transmit waveforms correlation matrix, i.e., $R_{m,l} = \int_0^{T_p} x_m(s) x_l^*(s) ds$, $(m, l) \in \{1, \dots, M\}^2$;
- $\mathbf{d}(\theta_t) = \left[1, e^{j2\pi \frac{d}{\lambda_0} \sin(\theta_t)}, \dots, e^{j2\pi \frac{d}{\lambda_0} (M-1) \sin(\theta_t)}\right]^T \in \mathbb{C}^M$ the angle-dependent transmit steering vector;
- $\mathbf{a}(\Delta\tau) = \left[1, e^{j2\pi \Delta f \Delta\tau}, \dots, e^{j2\pi \Delta f (M-1) \Delta\tau}\right]^T \in \mathbb{C}^M$ the range-dependent steering vector.

Furthermore, letting $u = \sin(\theta_t)$ and $\delta = 2\Delta f \Delta\tau$ (satisfying $|\delta| \leq \frac{\Delta f}{B}$), $\mathbf{s}(\theta_t, \Delta\tau)$ can be further expressed as [66, 67]

$$\mathbf{s}(\theta_t, \Delta\tau) = \mathbf{s}(u, \delta) = \mathbf{b}(u) \otimes [\mathbf{c}(u) \odot \mathbf{a}(\delta)], \quad (6.5)$$

where

- $\mathbf{b}(u) = \left[1, e^{j2\pi \frac{d}{\lambda_0} u}, \dots, e^{j2\pi \frac{d}{\lambda_0} (N-1) u}\right]^T \in \mathbb{C}^N$;
- $\mathbf{c}(u) = \mathbf{R}^T \mathbf{d}(u) \in \mathbb{C}^M$;
- $\mathbf{d}(u) = \left[1, e^{j2\pi \frac{d}{\lambda_0} u}, \dots, e^{j2\pi \frac{d}{\lambda_0} (M-1) u}\right]^T \in \mathbb{C}^M$;
- $\mathbf{a}(\delta) = \left[1, e^{j\pi \delta}, \dots, e^{j\pi (M-1) \delta}\right]^T \in \mathbb{C}^M$.

Hence, defining the steering matrix $\mathbf{H}(u, \delta) = \begin{bmatrix} \mathbf{s}(u, \delta) & \mathbf{0} \\ \mathbf{0} & \mathbf{s}(u, \delta) \end{bmatrix} \in \mathbb{C}^{2MN \times 2}$, the useful polarimetric target echo (6.4) can be recast as

$$\mathbf{y}_S(u, \delta) = \boldsymbol{\alpha} \otimes \mathbf{s}(u, \delta) = \mathbf{H}(u, \delta) \boldsymbol{\alpha}. \quad (6.6)$$

6.2 Formulation of the Detection Problem for Polarimetric FDA-MIMO Radar

This section investigates the problem of detecting a prospective target using a radar system equipped with a polarimetric FDA-MIMO. Let us define the received observation vector from the CUT (under the target presence) as $\mathbf{r} \in \mathbb{C}^{2MN}$, which contains the superposition of the polarimetric echo signal (6.6) of a prospective target and the interference plus noise contribution [35, 79]. Therefore, it can be cast as

$$\mathbf{r} = \mathbf{H}(u, \delta)\boldsymbol{\alpha} + \mathbf{m}, \quad (6.7)$$

where u , δ , and $\boldsymbol{\alpha}$ are regarded as unknown parameters and $\mathbf{m} \in \mathbb{C}^{2MN}$ is modeled as a zero-mean complex circularly symmetric Gaussian random vector, i.e., $\mathbf{m} \sim CN(0, \mathbf{M})$, with $\mathbf{M} \in \mathbb{H}^{2MN}$ the positive definite polarimetric covariance matrix of the interference plus noise term. Let us also assume that a set of $K \geq 2MN$ secondary data $\mathbf{r}_k \in \mathbb{C}^{2MN}$, $k = 1, 2, \dots, K$, free of useful target returns and with the same spectral characteristics as the interference from the CUT (homogeneous environment)⁴, i.e., $\mathbb{E}[\mathbf{r}_k \mathbf{r}_k^\dagger] = \mathbf{M}$, $k = 1, 2, \dots, K$, is available. As a consequence, the target detection problem can be formulated as a binary hypothesis test where the null hypothesis H_0 indicates the situation of target absence and H_1 represents the alternative, i.e.,

$$\begin{aligned} H_0 &: \begin{cases} \mathbf{r} = \mathbf{m} \\ \mathbf{r}_k = \mathbf{m}_k, \quad k = 1, 2, \dots, K \end{cases} \\ H_1 &: \begin{cases} \mathbf{r} = \mathbf{H}(u, \delta)\boldsymbol{\alpha} + \mathbf{m} \\ \mathbf{r}_k = \mathbf{m}_k, \quad k = 1, 2, \dots, K \end{cases} \end{aligned} \quad (6.8)$$

⁴It worth mentioning that often a data selection stage of the \mathbf{r}_k is also foreseen. This preprocessing is aimed at censoring from the training set data vectors containing possible outliers (sources of deviations from homogeneous assumption). The interested reader can refer to [49, 20, 30].

Notably, the joint PDFs of the observations under H_0 and H_1 are respectively given by

$$f(\mathbf{r}, \mathbf{r}_1, \dots, \mathbf{r}_K | \mathbf{M}; H_0) = \left\{ \frac{1}{\pi^{2MN} \det(\mathbf{M})} e^{-\text{tr}\{\mathbf{M}^{-1} \mathbf{R}_0\}} \right\}^{K+1} \quad (6.9a)$$

and

$$f(\mathbf{r}, \mathbf{r}_1, \dots, \mathbf{r}_K | \boldsymbol{\alpha}, u, \delta, \mathbf{M}; H_1) = \left\{ \frac{1}{\pi^{2MN} \det(\mathbf{M})} e^{-\text{tr}(\mathbf{M}^{-1} \mathbf{R}_1)} \right\}^{K+1}, \quad (6.9b)$$

with \mathbf{R}_0 and \mathbf{R}_1 defined as

$$\mathbf{R}_0 = \frac{\mathbf{r} \mathbf{r}^\dagger + \sum_{k=1}^K \mathbf{r}_k \mathbf{r}_k^\dagger}{K+1} \quad (6.10a)$$

and

$$\mathbf{R}_1 = \frac{(\mathbf{r} - \mathbf{H}(u, \delta) \boldsymbol{\alpha})(\mathbf{r} - \mathbf{H}(u, \delta) \boldsymbol{\alpha})^\dagger + \sum_{k=1}^K \mathbf{r}_k \mathbf{r}_k^\dagger}{K+1}. \quad (6.10b)$$

Invoking the Neyman-Pearson framework [39], the optimal solution to the hypothesis testing problem (6.8), i.e., the LRT, cannot be implemented due to the unknowns $\boldsymbol{\alpha}$, u , δ , and \mathbf{M} . In this respect, some sub-optimal and practically implementable architectures, based on the subspace GLRT [60] and 2SGLRT [93] frameworks, are proposed in subsection 6.2.1 and 6.2.2, respectively.

6.2.1 Detection with GLRT

In this subsection, the design of a detector based on subspace GLRT framework [87] is investigated. In particular, the GLRT over the unknowns computes the decision statistic

$$\Lambda_{\text{GLRT}} = \frac{\max_{\boldsymbol{\alpha}, u \in \mathcal{A}, \delta \in \mathcal{B}, \mathbf{M}} f(\mathbf{r}, \mathbf{r}_1, \dots, \mathbf{r}_K | \boldsymbol{\alpha}, u, \delta, \mathbf{M}; H_1)}{\max_{\mathbf{M}} f(\mathbf{r}, \mathbf{r}_1, \dots, \mathbf{r}_K | \mathbf{M}; H_0)}, \quad (6.11)$$

where, according to [67], \mathcal{A} denotes the uncertainty set associated with u , i.e., $[-1, 1]$, and \mathcal{B} that on δ , i.e., $\left[-\frac{\Delta f}{B}, \frac{\Delta f}{B}\right]$.

Maximizing both the numerator and the denominator of (6.11) over \mathbf{M} yields (according to standard argumentation) [60]

$$\tau_{\text{GLRT}} = \frac{1 + \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{r}}{1 + \min_{\boldsymbol{\alpha}, u \in \mathcal{A}, \delta \in \mathcal{B}} (\mathbf{r} - \mathbf{H}(u, \delta) \boldsymbol{\alpha})^\dagger \mathbf{S}^{-1} (\mathbf{r} - \mathbf{H}(u, \delta) \boldsymbol{\alpha})} \underset{H_0}{\overset{H_1}{\geq}} \xi, \quad (6.12)$$

where $\mathbf{S} = \sum_{k=1}^K \mathbf{r}_k \mathbf{r}_k^\dagger \in \mathbb{C}^{2MN \times 2MN}$ and ξ denotes the detection threshold, set to ensure the desired P_{fa} . Letting for notation simplicity $\mathbf{H}(u, \delta) = \mathbf{H}$, replacing $\boldsymbol{\alpha}$ with its ML estimate yields

$$\frac{1 + \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{r}}{1 + \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{r} - \max_{u \in \mathcal{A}, \delta \in \mathcal{B}} g(u, \delta)} \underset{H_0}{\overset{H_1}{\geq}} \xi, \quad (6.13)$$

where⁵

$$g(u, \delta) = \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{H} \left[\mathbf{H}^\dagger \mathbf{S}^{-1} \mathbf{H} \right]^{-1} \mathbf{H}^\dagger \mathbf{S}^{-1} \mathbf{r}. \quad (6.14)$$

Evidently, (6.13) is statistically equivalent to

$$\tau_{\text{GLRT}} = \frac{1}{1 + \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{r}} \max_{u \in \mathcal{A}, \delta \in \mathcal{B}} g(u, \delta) \underset{H_0}{\overset{H_1}{\geq}} \xi. \quad (6.15)$$

6.2.2 Detection with 2SGLRT

In this subsection, using the conventional approach of assuming the interference covariance matrix known at the design stage, an AMF-like detector (also known as the 2SGLRT) is derived. Precisely, the GLRT decision statistic for known covariance matrix is given by [93]

$$\Lambda = \frac{\max_{\boldsymbol{\alpha}, u \in \mathcal{A}, \delta \in \mathcal{B}} \bar{f}(\mathbf{r} | \boldsymbol{\alpha}, u, \delta, \mathbf{M}; H_1)}{\bar{f}(\mathbf{r} | \mathbf{M}; H_0)}, \quad (6.16)$$

⁵With some abuse of notation, the same ξ is used to denote the different detection thresholds for both the GLRT and the 2SGLRT decision statistics (6.12), (6.13), (6.15), (6.19), (6.20), (6.22), and (6.23).

where

$$\bar{f}(\mathbf{r}|\mathbf{M}; H_0) = \frac{1}{\pi^{2MN} \det(\mathbf{M})} e^{-\mathbf{r}^\dagger \mathbf{M}^{-1} \mathbf{r}} \quad (6.17a)$$

and

$$\bar{f}(\mathbf{r}|\boldsymbol{\alpha}, u, \delta, \mathbf{M}; H_1) = \frac{1}{\pi^{2MN} \det(\mathbf{M})} e^{-(\mathbf{r} - \mathbf{H}(u, \delta)\boldsymbol{\alpha})^\dagger \mathbf{M}^{-1} (\mathbf{r} - \mathbf{H}(u, \delta)\boldsymbol{\alpha})}. \quad (6.17b)$$

Hence, substituting the ML estimate of the unknown $\boldsymbol{\alpha}$, i.e.,

$$\hat{\boldsymbol{\alpha}} = \left[\mathbf{H}^\dagger \mathbf{M}^{-1} \mathbf{H} \right]^{-1} \mathbf{H}^\dagger \mathbf{M}^{-1} \mathbf{r}. \quad (6.18)$$

into (6.16), the following decision rule is obtained

$$\tau_{2\text{SGLRT}} = \max_{u \in \mathcal{A}, \delta \in \mathcal{B}} \mathbf{r}^\dagger \mathbf{M}^{-1} \mathbf{H} \left[\mathbf{H}^\dagger \mathbf{M}^{-1} \mathbf{H} \right]^{-1} \mathbf{H}^\dagger \mathbf{M}^{-1} \mathbf{r} \underset{H_0}{\overset{H_1}{\geq}} \xi. \quad (6.19)$$

Finally, using the SCM $\hat{\mathbf{M}} = \frac{1}{K} \mathbf{S}$ in lieu of \mathbf{M} yields

$$\tau_{2\text{SGLRT}} = \max_{u \in \mathcal{A}, \delta \in \mathcal{B}} g(u, \delta) \underset{H_0}{\overset{H_1}{\geq}} \xi. \quad (6.20)$$

6.3 Optimization Problem (6.20)

This section is devoted to the development of some optimization strategies to handle

$$\max_{u \in \mathcal{A}, \delta \in \mathcal{B}} g(u, \delta), \quad (6.21)$$

namely, to compute the ML estimate of the target location parameters under the H_1 hypothesis. To this end, let us preliminary observe that (6.21) falls in the class of box-constrained optimization problems. Moreover, an approximated solution based on a 2-D grid search is characterized by a high and often impractical computational cost if the grid size is sufficiently dense to get a close to optimum objective value. This motivates the design of some optimization methods pursuing reduced-complexity sub-optimal solutions. In this regard, three strategies are developed:

1. the LAM procedure [8];
2. the GPM [18];
3. the CD Method [19].

Denoting by $(\hat{u}, \hat{\delta})$ the estimate of the unknowns using one of the aforementioned LAM, GPM, and CD methods, it follows that the actual decision rules, based on the GLRT and 2SGLRT criteria, are given by

$$\tau_{\text{GLRT}} = \frac{1}{1 + \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{r}} g(\hat{u}, \hat{\delta}) \underset{H_0}{\overset{H_1}{\geq}} \xi \quad (6.22)$$

and

$$\tau_{\text{2SGLRT}} = g(\hat{u}, \hat{\delta}) \underset{H_0}{\overset{H_1}{\geq}} \xi, \quad (6.23)$$

respectively.

6.3.1 LAM Method

Resorting to the LAM approach proposed in Chapter 2, a sub-optimal and approximated solution method to the optimization problem in (6.21) is derived. To this end, let us observe that (6.21) is equivalent to

$$\min_{\boldsymbol{\alpha}, u \in \mathcal{A}, \delta \in \mathcal{B}} (\mathbf{r} - \mathbf{H}(u, \delta) \boldsymbol{\alpha})^\dagger \mathbf{S}^{-1} (\mathbf{r} - \mathbf{H}(u, \delta) \boldsymbol{\alpha}), \quad (6.24)$$

namely, given a maximizer u_1^*, δ_1^* of (6.21), an optimal solution to (6.24) is given by

$$\boldsymbol{\alpha}_2^* = \left[\mathbf{H}(u_1^*, \delta_1^*)^\dagger \mathbf{S}^{-1} \mathbf{H}(u_1^*, \delta_1^*) \right]^{-1} \mathbf{H}(u_1^*, \delta_1^*)^\dagger \mathbf{S}^{-1} \mathbf{r}, \quad (6.25a)$$

$$u_2^* = u_1^*, \quad (6.25b)$$

$$\delta_2^* = \delta_1^*. \quad (6.25c)$$

In a similar manner, given the minimizer $\boldsymbol{\alpha}_2^*, u_2^*, \delta_2^*$ of (6.24), an optimal solution to (6.21) is $u_1^* = u_2^*$ and $\delta_1^* = \delta_2^*$.

Now, given the nominal angle and incremental range \bar{u} and $\bar{\delta}$, the actual mismatch w.r.t. the true target parameters values can be accounted by performing a tailored linearization of the array steering matrix $\mathbf{H}(u, \delta)$ around \bar{u} and $\bar{\delta}$. To this end, denoting by $(\Delta u, \Delta \delta)$ the angle and range offsets, i.e., $\Delta u = u - \bar{u}$, $\Delta \delta = \delta - \bar{\delta}$, the array steering matrix can be approximated at the first order as

$$\begin{aligned} \mathbf{H}_a(\Delta \boldsymbol{\theta}) &= \mathbf{H}_a(\Delta u, \Delta \delta) \\ &\simeq \mathbf{H}(\bar{u}, \bar{\delta}) + \left. \frac{\partial \mathbf{H}(u, \delta)}{\partial u} \right|_{(\bar{u}, \bar{\delta})} \Delta u + \left. \frac{\partial \mathbf{H}(u, \delta)}{\partial \delta} \right|_{(\bar{u}, \bar{\delta})} \Delta \delta \\ &= \bar{\mathbf{H}} + \bar{\mathbf{H}}_u \Delta u + \bar{\mathbf{H}}_\delta \Delta \delta, \end{aligned} \quad (6.26)$$

with $\Delta \boldsymbol{\theta} = [\Delta u, \Delta \delta]^T \in \mathbb{R}^2$. Detailed expressions for $\mathbf{H}_u = \frac{\partial \mathbf{H}(u, \delta)}{\partial u}$ and $\mathbf{H}_\delta = \frac{\partial \mathbf{H}(u, \delta)}{\partial \delta}$ are reported in Appendix E.1. Hence, resorting to the linearization procedure, an optimized solution to (6.24) can be obtained solving

$$\min_{\boldsymbol{\alpha}, \Delta \boldsymbol{\theta} \in \mathcal{C}} (\mathbf{r} - \mathbf{H}_a(\Delta \boldsymbol{\theta}) \boldsymbol{\alpha})^\dagger \mathbf{S}^{-1} (\mathbf{r} - \mathbf{H}_a(\Delta \boldsymbol{\theta}) \boldsymbol{\alpha}), \quad (6.27)$$

where $\mathcal{C} = [-\kappa, \kappa] \times [-\rho, \rho]$, with $\kappa = 1$ and $\rho = \frac{\Delta f}{B}$, is the nonempty and compact feasible set⁶. In this respect, a CD-based procedure is proposed to solve the optimization problem (6.27) by alternatively optimizing $\boldsymbol{\alpha}$ and $\Delta \boldsymbol{\theta}$ up to convergence. In the following, the analytical solutions to the resulting optimization problems, at the h -th iteration, are derived.

a) Optimization w.r.t. $\boldsymbol{\alpha}$.

At the h -th iteration, the optimization problem w.r.t. $\boldsymbol{\alpha}$ is given by

$$\min_{\boldsymbol{\alpha}} (\mathbf{r} - \mathbf{H}_a(\Delta \boldsymbol{\theta}^{(h-1)}) \boldsymbol{\alpha})^\dagger \mathbf{S}^{-1} (\mathbf{r} - \mathbf{H}_a(\Delta \boldsymbol{\theta}^{(h-1)}) \boldsymbol{\alpha}), \quad (6.28)$$

⁶Note that the accuracy of the approximation can be improved considering a partition of the original feasible set \mathcal{C} into D subsets, $\mathcal{C}_i, i = 1, \dots, D$, and performing the linearization of the steering matrix around different nominal points $(\bar{u}_i, \bar{\delta}_i), i = 1, \dots, D$, each associated to a specific uncertainty set $\mathcal{C}_i, i = 1, \dots, D$.

whose optimal solution is

$$\boldsymbol{\alpha}^{(h)} = \left(\mathbf{H}_a^\dagger \left(\Delta\boldsymbol{\theta}^{(h-1)} \right) \mathbf{S}^{-1} \mathbf{H}_a \left(\Delta\boldsymbol{\theta}^{(h-1)} \right) \right)^{-1} \mathbf{H}_a^\dagger \left(\Delta\boldsymbol{\theta}^{(h-1)} \right) \mathbf{S}^{-1} \mathbf{r}. \quad (6.29)$$

b) Optimization w.r.t. $\Delta\boldsymbol{\theta}$.

At the h -th iteration, the optimization w.r.t. $\Delta\boldsymbol{\theta}$ yields

$$\begin{aligned} & \min_{\substack{|\Delta u| \leq \kappa \\ |\Delta \delta| \leq \rho}} (\mathbf{r} - (\bar{\mathbf{H}} + \bar{\mathbf{H}}_u \Delta u + \bar{\mathbf{H}}_\delta \Delta \delta) \boldsymbol{\alpha}^{(h)})^\dagger \mathbf{S}^{-1} \\ & \quad (\mathbf{r} - (\bar{\mathbf{H}} + \bar{\mathbf{H}}_u \Delta u + \bar{\mathbf{H}}_\delta \Delta \delta) \boldsymbol{\alpha}^{(h)}) \\ & = \min_{\Delta\boldsymbol{\theta} \in \mathcal{C}} (\tilde{\mathbf{r}} - \tilde{\mathbf{H}} \Delta\boldsymbol{\theta})^\dagger \mathbf{S}^{-1} (\tilde{\mathbf{r}} - \tilde{\mathbf{H}} \Delta\boldsymbol{\theta}) \\ & = \min_{\Delta\boldsymbol{\theta} \in \mathcal{C}} (\hat{\mathbf{r}} - \hat{\mathbf{H}} \Delta\boldsymbol{\theta})^\dagger (\hat{\mathbf{r}} - \hat{\mathbf{H}} \Delta\boldsymbol{\theta}), \end{aligned} \quad (6.30)$$

where

- $\tilde{\mathbf{r}} = \mathbf{r} - \bar{\mathbf{H}} \boldsymbol{\alpha}^{(h)}$;
- $\tilde{\mathbf{H}} = [\bar{\mathbf{H}}_u \boldsymbol{\alpha}^{(h)}, \bar{\mathbf{H}}_\delta \boldsymbol{\alpha}^{(h)}]$;
- $\hat{\mathbf{r}} = \mathbf{S}^{-1/2} \tilde{\mathbf{r}}$;
- $\hat{\mathbf{H}} = [\hat{\mathbf{h}}_u, \hat{\mathbf{h}}_\delta] = \mathbf{S}^{-1/2} \tilde{\mathbf{H}}$.

The following proposition provides an optimal point to (6.30).

Proposition 3. An optimal solution $\Delta\boldsymbol{\theta}^{(h)}$ to (6.30) is

$$\Delta\boldsymbol{\theta}_1 = \left[\Re\{\hat{\mathbf{H}}^\dagger \hat{\mathbf{H}}\} \right]^{-1} \Re\{\hat{\mathbf{H}}^\dagger \hat{\mathbf{r}}\} \quad (6.31)$$

if it is feasible. Otherwise, it can be computed as

$$\Delta\boldsymbol{\theta}^{(h)} = \arg \min_{\Delta\boldsymbol{\theta} \in \{\Delta\boldsymbol{\theta}_i\}_{i=2}^5} (\hat{\mathbf{r}} - \hat{\mathbf{H}} \Delta\boldsymbol{\theta})^\dagger (\hat{\mathbf{r}} - \hat{\mathbf{H}} \Delta\boldsymbol{\theta}), \quad (6.32)$$

where $\Delta\boldsymbol{\theta}_i$, $i = 2, \dots, 5$ are candidate solutions given by

$$\begin{aligned} \Delta\boldsymbol{\theta}_2 &= [\kappa, \Delta\delta_+^*]^\text{T}, & \Delta\boldsymbol{\theta}_3 &= [-\kappa, \Delta\delta_-^*]^\text{T}, \\ \Delta\boldsymbol{\theta}_4 &= [\Delta u_+^*, \rho]^\text{T}, & \Delta\boldsymbol{\theta}_5 &= [\Delta u_-^*, -\rho]^\text{T}, \end{aligned} \quad (6.33)$$

with

$$\Delta u_{\pm}^* = \min \left(\kappa, \max \left(\frac{\Re\{(\hat{\mathbf{r}}^\dagger - \hat{\mathbf{h}}_\delta^\dagger(\pm\rho)) \hat{\mathbf{h}}_u\}}{\|\hat{\mathbf{h}}_u\|^2}, -\kappa \right) \right) \quad (6.34)$$

and

$$\Delta \delta_{\pm}^* = \min \left(\rho, \max \left(\frac{\Re\{(\hat{\mathbf{r}}^\dagger - \hat{\mathbf{h}}_u^\dagger(\pm\kappa)) \hat{\mathbf{h}}_\delta\}}{\|\hat{\mathbf{h}}_\delta\|^2}, -\rho \right) \right). \quad (6.35)$$

Proof. See Appendix E.2. \square

A summary of the devised procedure to determine a solution to (6.27) is reported in **Algorithm 10**, where the exit condition is set as $|P^{(h)} - P^{(h-1)}| < \varepsilon_1$, with $\varepsilon_1 > 0$ and

$$P^{(h)} = (\mathbf{r} - \mathbf{H}_a(\Delta\boldsymbol{\theta}^{(h)})\boldsymbol{\alpha}^{(h)})^\dagger \mathbf{S}^{-1}(\mathbf{r} - \mathbf{H}_a(\Delta\boldsymbol{\theta}^{(h)})\boldsymbol{\alpha}^{(h)}). \quad (6.36)$$

Now, given the output of **Algorithm 10**, an estimate of the target angle and incremental range can be obtained as $\tilde{u}_{\text{LAM}} = \bar{u} + \Delta u^*$ and $\tilde{\delta}_{\text{LAM}} = \bar{\delta} + \Delta \delta^*$, respectively, with Δu^* and $\Delta \delta^*$ the resulting mismatch estimates w.r.t. the nominal values \bar{u} and $\bar{\delta}$, respectively. However, $[\tilde{u}_{\text{LAM}}, \tilde{\delta}_{\text{LAM}}]^\text{T}$ could not be a feasible solution to (6.21). To this end, the final estimate of the target location parameters undergoes a projection in the feasible set by $\hat{\mathbf{h}}_{\text{LAM}} = [\hat{u}_{\text{LAM}}, \hat{\delta}_{\text{LAM}}]^\text{T} = \mathbf{P}_{\mathcal{C}} \left([\tilde{u}_{\text{LAM}}, \tilde{\delta}_{\text{LAM}}]^\text{T} \right)$, where $\mathbf{P}_{\mathcal{C}}$ is the projection operator onto the constraint set \mathcal{C} , i.e.,

$$\mathbf{P}_{\mathcal{C}} \left([\tilde{u}, \tilde{\delta}]^\text{T} \right) = \begin{bmatrix} \min \{ \max \{ \tilde{u}, -\kappa \}, \kappa \} \\ \min \{ \max \{ \tilde{\delta}, -\rho \}, \rho \} \end{bmatrix}, \tilde{u}, \tilde{\delta} \in \mathbb{R}. \quad (6.37)$$

6.3.2 GPM

A procedure to tackle the maximization problem (6.21) is devised resorting to an ascent direction method. Specifically, an estimate of \mathbf{h} could be computed iteratively via the projected gradient technique [18]. Accordingly, at the k -th iteration, an updated estimate is obtained moving from the previously estimated point \mathbf{h}_{k-1} along the current ascent direction (given by the gradient evaluated at \mathbf{h}_{k-1}) and then projecting the

Algorithm 10 FDA-LAM

Input: $\mathbf{r}, \mathbf{S}, \bar{u}, \bar{\delta}, \kappa, \rho, \varepsilon_1$.**Output:** $\hat{\mathbf{h}}_{\text{LAM}}$.

1. Set $h = 0$, $\Delta\boldsymbol{\theta}^{(h)} = [\Delta u^{(h)}, \Delta\delta^{(h)}]^T = \mathbf{0}$,
 $P^{(h)} = \infty$.
2. **repeat**
3. $h = h + 1$;
4. Compute $\boldsymbol{\alpha}^{(h)}$ via (6.29);
5. Compute $\Delta\boldsymbol{\theta}^{(h)}$ via Proposition 3;
6. Evaluate

$$P^{(h)} = (\mathbf{r} - \mathbf{H}_a(\Delta\boldsymbol{\theta}^{(h)})\boldsymbol{\alpha}^{(h)})^\dagger \mathbf{S}^{-1} (\mathbf{r} - \mathbf{H}_a(\Delta\boldsymbol{\theta}^{(h)})\boldsymbol{\alpha}^{(h)})$$

7. **until** $|P^{(h)} - P^{(h-1)}| < \varepsilon_1$.
 8. $\Delta\boldsymbol{\theta}^* = [\Delta u^*, \Delta\delta^*]^T = \Delta\boldsymbol{\theta}^{(h)}$
 9. Output $\hat{\mathbf{h}}_{\text{LAM}} = \mathbf{P}_{\mathcal{C}}([\bar{u} + \Delta u^*, \bar{\delta} + \Delta\delta^*]^T)$.
-

resulting point onto the considered convex set \mathcal{C} . Precisely,

$$\mathbf{h}_k = \mathbf{P}_{\mathcal{C}}(\mathbf{h}_{k-1} + \eta_k \mathbf{g}_{\Delta}(\mathbf{h}_{k-1})), \quad (6.38)$$

where

- $\mathbf{h}_0 \in \mathbb{R}^2$ is the initial vector;
 - $\mathbf{h}_{k-1} = [u_{k-1}, \delta_{k-1}]^T \in \mathbb{R}^2$ with u_{k-1} and δ_{k-1} the estimates at the $(k-1)$ -th iteration;
 - η_k the step size;
 - $\mathbf{g}_{\Delta}(\mathbf{h}_{k-1}) = \left[\frac{\partial g(u, \delta)}{\partial u}, \frac{\partial g(u, \delta)}{\partial \delta} \right]^T \Big|_{(u_{k-1}, \delta_{k-1})}$ the gradient of $g(u, \delta)$ evaluated at u_{k-1} and δ_{k-1} (see Appendix E.3 for its computation);
-

- $\mathbf{P}_{\mathcal{C}}$ is defined as in (6.37).

Note that η_k can be either a constant stepsize $\bar{\eta} \in (0, \frac{2}{L})$, with L the smallest Lipschitz constant for $g(\mathbf{h}) = g(u, \delta)$, or chosen adaptively by means of the backtracking procedure [18], i.e.,

$$\begin{cases} \eta_k^{\text{cons}} = \bar{\eta}, & \text{constant stepsize} \\ \eta_k^{\text{back}} = s\beta^{i_k}, & \text{backtracking} \end{cases}, \quad (6.39)$$

where i_k the smallest nonnegative integer satisfying

$$g(\tilde{\mathbf{h}}^{(i_k)}) - g(\mathbf{h}_{k-1}) \geq \zeta s\beta^{i_k} \left\| \mathbf{G}_{\frac{1}{s\beta^{i_k}}}(\mathbf{h}_{k-1}) \right\|^2, \quad (6.40)$$

with $s > 0$, $\zeta \in (0, 1)$, and $\beta \in (0, 1)$ algorithm tuning parameters,

$$\tilde{\mathbf{h}}^{(i_k)} = \mathbf{P}_{\mathcal{C}}(\mathbf{h}_{k-1} + s\beta^{i_k} \mathbf{g}_{\Delta}(\mathbf{h}_{k-1})) \quad (6.41)$$

and

$$\mathbf{G}_{\varkappa}(\mathbf{h}) = \varkappa \left[\mathbf{h} - \mathbf{P}_{\mathcal{C}} \left(\mathbf{h} + \frac{1}{\varkappa} \mathbf{g}_{\Delta}(\mathbf{h}) \right) \right] \quad (6.42)$$

the gradient mapping with $\varkappa > 0$.

Algorithm 11 summarizes the gradient projection procedure, where the exit condition is set as $\|\mathbf{h}_k - \mathbf{h}_{k-1}\| < \varepsilon_2$, with $\varepsilon_2 > 0$. Besides, the resulting estimate is denoted by $\hat{\mathbf{h}}_{\text{GPM}} = [\hat{u}_{\text{GPM}}, \hat{\delta}_{\text{GPM}}]^{\text{T}}$.

Remark 1. Before concluding this subsection, the convergence of the GPM is examined via the following lemma [18, Th. 9.14], reformulated for the case of a maximization problem.

Lemma 6.3.1. *Consider the optimization problem*

$$\mathcal{P} \begin{cases} \max & g(\mathbf{h}) \\ \text{s.t.} & \mathbf{h} \in \mathcal{C} \end{cases}, \quad (6.43)$$

where $g(\mathbf{h}) \in C_L^{1,1}$ is bounded above and \mathcal{C} is a closed and convex set. Let $\{\mathbf{h}_k\}_{k \geq 0}$ be the sequence generated by the GPM w.r.t. problem \mathcal{P} using either constant or adaptive (backtracking) stepsize. Then

Algorithm 11 FDA-GPM

Input: $\mathbf{r}, \mathbf{S}, \bar{u}, \bar{\delta}, s, \beta, \zeta, \varepsilon_2$.**Output:** $\hat{\mathbf{h}}_{\text{GPM}}$.

1. Set $k = 0, \mathbf{h}_k = \mathbf{h}_0 = [\bar{u}, \bar{\delta}]^T$.
 2. **repeat**
 3. $k = k + 1$;
 4. Update η_k using either a constant or adaptive (via backtracking strategy) stepsize (6.39);
 5. Compute \mathbf{h}_k via (6.38);
 6. **until** $\|\mathbf{h}_k - \mathbf{h}_{k-1}\| < \varepsilon_2$.
 7. Output $\hat{\mathbf{h}}_{\text{GPM}} = [\hat{u}_{\text{GPM}}, \hat{\delta}_{\text{GPM}}]^T = \mathbf{h}_k$.
-

- the sequence $\{g(\mathbf{h}_k)\}_{k \geq 0}$ is nondecreasing and $g(\mathbf{h}_k) > g(\mathbf{h}_{k-1})$ unless \mathbf{h}_{k-1} is a stationary point of \mathcal{P} ;
- any limit point \mathbf{h}^* of $\{\mathbf{h}_k\}_{k \geq 0}$ is a stationary point of \mathcal{P} .

□

Exploiting the above result, the convergence of **Algorithm 11** to a stationary point of problem (6.21) is guaranteed, provided that $g(u, \delta) \in C_L^{1,1}$ (see Appendix E.4 for the detailed proof).

6.3.3 CD Method

The **CD** method [113, 19] is an iterative procedure that exploits a series of **1-D** searches, with one variable optimized at a time while keeping the other constant. Precisely, the searches w.r.t. u and δ are respectively conducted over \mathcal{I}_u and \mathcal{I}_δ (corresponding to the discretized versions of \mathcal{A}

and \mathcal{B} with $(N_u + 1)$ and $(N_\delta + 1)$ points, respectively), defined as [67]

$$\mathcal{J}_u = \left\{ -1 + \frac{2i}{N_u}, i = 0, \dots, N_u \right\} \quad (6.44a)$$

and

$$\mathcal{J}_\delta = \left\{ -\frac{\Delta f}{B} + \frac{2i}{N_\delta} \frac{\Delta f}{B}, i = 0, \dots, N_\delta \right\}. \quad (6.44b)$$

Remarkably, since the order of the optimization could lead to different solutions, both the instances of optimizing first u and then over δ , and first δ and then over u , are implemented, with the optimal estimate chosen between the resulting two solutions.

Denoting by $g^k = g(\hat{u}^k, \hat{\delta}^k)$ and setting the nominal values as initial estimates, i.e., $\hat{u}^0 = \bar{u}$ and $\hat{\delta}^0 = \bar{\delta}$, u and δ can be updated according to **Algorithm 12**, where the exit condition is set as $|g^k - g^{k-1}| < \varepsilon_3$, with $\varepsilon_3 > 0$. Hence, the final solution using the **CD** method is obtained as $\hat{\mathbf{h}}_{\text{CD}} = [\hat{u}_{\text{CD}}, \hat{\delta}_{\text{CD}}]^T$.

Note that for any initial search direction, i.e., along the u or δ domain, the **CD** approach (starting from the second iteration) coincides with the **MBI** policy [29]. Therefore, invoking the convergence properties of **MBI** [29, 88, 12], any limit point resulting from **Algorithm 12** (assuming an exact optimization at each step) is a stationary point to Problem (6.21), although its convergence to the optimal value cannot be claimed [19].

6.3.4 Bounded CFARness Analysis of the Derived Detectors and Computational Complexities

In this subsection, the bounded **CFAR** property of the proposed detection architectures is studied. First of all, from (6.12) it is straightforward to see that the **GLRT** decision statistic is upper bounded by a **CFAR** detector, i.e.,

$$\tau_{\text{GLRT}} \leq 1 + \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{r} = \tau_{\text{CFAR}}. \quad (6.45)$$

Similarly, with reference to the **2SGLRT** detector (6.20), the following inequality applies, i.e.,

$$1 + \tau_{2\text{SGLRT}} \leq \tau_{\text{CFAR}}. \quad (6.46)$$

Algorithm 12 FDA-CD

Input: \mathbf{r} , \mathbf{S} , \bar{u} , $\bar{\delta}$, \mathcal{J}_u , \mathcal{J}_δ , ε_3 .**Output:** $\hat{\mathbf{h}}_{\text{CD}}$.

1. Set $k = 0$, $\hat{u}^k = \bar{u}$, $\hat{\delta}^k = \bar{\delta}$, $g^k = g(\hat{u}^k, \hat{\delta}^k)$
 2. **repeat** (optimization for initial search direction given by u)
 3. Find $\hat{u}^{k+1} = \arg \max_{u \in \mathcal{J}_u} g(u, \hat{\delta}^k)$;
 4. Find $\hat{\delta}^{k+1} = \arg \max_{\delta \in \mathcal{J}_\delta} g(\hat{u}^{k+1}, \delta)$ and set g^{k+1} as the corresponding maximum value;
 5. $k = k + 1$;
 6. **until** $|g^k - g^{k-1}| < \varepsilon_3$;
 7. $g_x = g^k$; $\hat{u}_x = \hat{u}^k$; $\hat{\delta}_x = \hat{\delta}^k$;
 8. Set $k = 0$;
 9. **repeat** (optimization for initial search direction given by δ)
 10. Find $\hat{\delta}^{k+1} = \arg \max_{\delta \in \mathcal{J}_\delta} g(\hat{u}^k, \delta)$;
 11. Find $\hat{u}^{k+1} = \arg \max_{u \in \mathcal{J}_u} g(u, \hat{\delta}^{k+1})$ and set g^{k+1} as the corresponding maximum value;
 12. $k = k + 1$;
 13. **until** $|g^k - g^{k-1}| < \varepsilon_3$;
 14. $g_y = g^k$; $\hat{u}_y = \hat{u}^k$; $\hat{\delta}_y = \hat{\delta}^k$;
 15. **if** $g_x > g_y$
 16. Output $\hat{\mathbf{h}}_{\text{CD}} = [\hat{u}_{\text{CD}}, \hat{\delta}_{\text{CD}}]^T = [\hat{u}_x, \hat{\delta}_x]^T$.
 17. **else**
 18. Output $\hat{\mathbf{h}}_{\text{CD}} = [\hat{u}_{\text{CD}}, \hat{\delta}_{\text{CD}}]^T = [\hat{u}_y, \hat{\delta}_y]^T$.
 19. **end**
-

As a consequence, the devised decision architectures ensure the bounded CFAR property, i.e.,

$$P_{fa, \text{GLRT}} = \Pr(\tau_{\text{GLRT}} > \xi | H_0) \leq \Pr(\tau_{\text{CFAR}} > \xi | H_0), \quad (6.47)$$

and

$$P_{fa, 2\text{SGLRT}} = \Pr(\tau_{2\text{SGLRT}+1} > \xi | H_0) \leq \Pr(\tau_{\text{CFAR}} > \xi | H_0). \quad (6.48)$$

Hence, for each considered decision strategies, it is always possible to set a universal threshold ξ (based on just the system parameters) for which $\Pr(\tau_{\text{CFAR}} > \xi | H_0)$ achieves a desired upper bound \bar{P}_{fa} to the actual P_{fa} . Evidently, this property holds true also for an approximated implementation of the decision statistics in (6.12) and (6.20), when possible sub-optimal solution methods are employed to estimate the unknown target parameters α , u , and δ .

Finally, as long as the number of iterations involved in the three optimization procedures keeps quite limited, the computational complexity is dominated by the evaluation of the SCM \mathbf{S} , which demands $\mathcal{O}(M^2 N^2 K)$ operations, regardless of the adopted optimization strategy.

6.4 Simulation Results

In this section, numerical examples are provided to evaluate the performance of the proposed target detection methods for the polarimetric FDA-MIMO radar. In the following, orthogonal baseband signals are assumed, i.e., $\mathbf{R} = \mathbf{I}$.

The P_D is used as figure of merit to assess the detection performance, which is estimated resorting to 1000 independent Monte Carlo runs assuming a $P_{fa} = 10^{-4}$. Besides, the detection thresholds are computed using $100/P_{fa}$ independent trials. The SINR is defined as

$$\text{SINR} = \mathbb{E}[\alpha^\dagger \mathbf{H}^\dagger \mathbf{M}^{-1} \mathbf{H} \alpha] = \gamma^2 \text{tr}\{\boldsymbol{\Sigma}_t \mathbf{H}^\dagger \mathbf{M}^{-1} \mathbf{H}\}, \quad (6.49)$$

where $\mathbb{E}[\alpha \alpha^\dagger] = \gamma^2 \boldsymbol{\Sigma}_t$, with $\boldsymbol{\Sigma}_t = \begin{bmatrix} 1 & \epsilon_t \sqrt{\delta_t} \\ \epsilon_t^* \sqrt{\delta_t} & \delta_t \end{bmatrix} \in \mathbb{C}^{2 \times 2}$ and γ rules the target strength, respectively. Besides, in the reported results it is as-

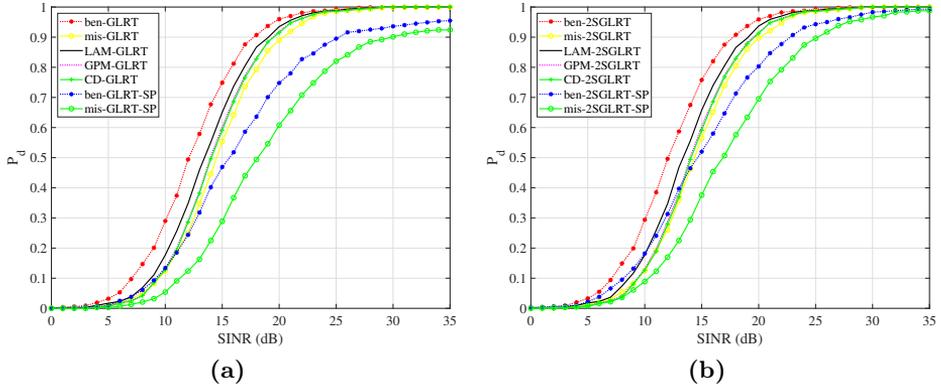


Figure 6.3. Detection performance for white noise, assuming $K = 320$ and the parameters of Table 6.1. Fig. (a) considers the GLRT-based detectors, whereas Fig. (b) reports the 2SGLRT-based counterparts.

Table 6.1. Simulation Parameters of the FDA-MIMO Radar

Parameter	Value	Parameter	Value
transmit elements M	4	receive elements N	10
bandwidth B	1 MHz	frequency increment Δf	0.5 MHz
angle of the target u	$\frac{0.891}{2(N+M)}$	incremental range of the target δ	$0.6 \frac{\Delta f}{2B}$
nominal angle of the target \bar{u}	0	nominal incremental range of the target $\bar{\delta}$	0
target polarimetric parameter ϵ_t	0.28	target polarimetric parameter δ_t	1

sumed that $\boldsymbol{\alpha} \sim CN(0, \gamma^2 \boldsymbol{\Sigma}_t)$. The parameters involved in the simulations are listed in Table 6.1, including both the nominal and the actual target angle and incremental range.

Two benchmark detectors (devised assuming a perfect knowledge of \mathbf{H} and resorting to the **GLRT** and **2SGLRT** decision statistics, respectively) are considered for comparison purposes, i.e.,

$$\Lambda_{\text{ben-GLRT}}(u, \delta) = \frac{1}{1 + \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{r}} \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{H} \left[\mathbf{H}^\dagger \mathbf{S}^{-1} \mathbf{H} \right]^{-1} \mathbf{H}^\dagger \mathbf{S}^{-1} \mathbf{r} \quad (6.50)$$

and

$$\Lambda_{\text{ben-2SGLRT}}(u, \delta) = \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{H} \left[\mathbf{H}^\dagger \mathbf{S}^{-1} \mathbf{H} \right]^{-1} \mathbf{H}^\dagger \mathbf{S}^{-1} \mathbf{r}. \quad (6.51)$$

Besides, two additional decision strategies, respectively based on the **GLRT** and **2SGLRT** criteria and using the nominal values of u and δ (thus referred to as mismatched detectors), i.e.,

$$\Lambda_{\text{mis-GLRT}}(\bar{u}, \bar{\delta}) = \frac{1}{1 + \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{r}} \mathbf{r}^\dagger \mathbf{S}^{-1} \bar{\mathbf{H}} \left[\bar{\mathbf{H}}^\dagger \mathbf{S}^{-1} \bar{\mathbf{H}} \right]^{-1} \bar{\mathbf{H}}^\dagger \mathbf{S}^{-1} \mathbf{r} \quad (6.52)$$

and

$$\Lambda_{\text{mis-2SGLRT}}(\bar{u}, \bar{\delta}) = \mathbf{r}^\dagger \mathbf{S}^{-1} \bar{\mathbf{H}} \left[\bar{\mathbf{H}}^\dagger \mathbf{S}^{-1} \bar{\mathbf{H}} \right]^{-1} \bar{\mathbf{H}}^\dagger \mathbf{S}^{-1} \mathbf{r}. \quad (6.53)$$

are also included. Moreover, the Single Polarization (**SP**) counterparts of (6.50), (6.51), (6.52), and (6.53), exploiting the data received by the HH channels, are also reported. Their decision statistics are respectively given by

$$\Lambda_{\text{ben-GLRT-SP}} = \frac{|\mathbf{s}^\dagger(u, \delta) \mathbf{S}^{-1} \mathbf{r}|^2}{(1 + \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{r}) \mathbf{s}^\dagger(u, \delta) \mathbf{S}^{-1} \mathbf{s}(u, \delta)}, \quad (6.54)$$

$$\Lambda_{\text{ben-2SGLRT-SP}} = \frac{|\mathbf{s}^\dagger(u, \delta) \mathbf{S}^{-1} \mathbf{r}|^2}{\mathbf{s}^\dagger(u, \delta) \mathbf{S}^{-1} \mathbf{s}(u, \delta)}, \quad (6.55)$$

$$\Lambda_{\text{mis-GLRT-SP}} = \frac{|\mathbf{s}^\dagger(\bar{u}, \bar{\delta}) \mathbf{S}^{-1} \mathbf{r}|^2}{(1 + \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{r}) \mathbf{s}^\dagger(\bar{u}, \bar{\delta}) \mathbf{S}^{-1} \mathbf{s}(\bar{u}, \bar{\delta})}, \quad (6.56)$$

and

$$\Lambda_{\text{mis-2SGLRT-SP}} = \frac{|\mathbf{s}^\dagger(\bar{u}, \bar{\delta}) \mathbf{S}^{-1} \mathbf{r}|^2}{\mathbf{s}^\dagger(\bar{u}, \bar{\delta}) \mathbf{S}^{-1} \mathbf{s}(\bar{u}, \bar{\delta})}. \quad (6.57)$$

In the considered simulations, the parameters ε_1 , ε_2 , and ε_3 , involved in **Algorithms 10, 11, and 12**, respectively, are set as $\varepsilon_1 = \varepsilon_2 = \varepsilon_3 = 10^{-4}$. For **Algorithm 11**, $s = 1, \beta = 0.5, \zeta = 0.5$, whereas for **Algorithm 12**, $N_u = N_\delta = 250$. Additionally, 9 initial points $(\bar{u}, \bar{\delta})$, picked up within the constraint set \mathcal{C} , i.e., $\{-1/2, 0, 1/2\} \times \{-\frac{\Delta f}{2B}, 0, \frac{\Delta f}{2B}\}$ are used for the execution of the algorithms⁷. Then, the best achieved estimate in likelihood sense is selected.

⁷It is worth noting that several feasible initial vectors can be employed for the execution of the algorithms to minimize the risk of being trapped in a local maximum.

In the reported simulations, three different interference scenarios are examined. In the first one, the useful signal is buried in white Gaussian noise; in the last two, clutter disturbance is considered.

White noise interference case

Figs. 6.3(a) and 6.3(b) show the P_D curves of the proposed detectors (assuming $K = 320$ secondary data) in a scenario dominated by white Gaussian noise, i.e., with the covariance matrix modeled as

$$\mathbf{M} = \sigma_n^2 \mathbf{I} \in \mathbb{H}^{2MN}, \quad (6.58)$$

where, without loss of generality, the noise power level σ_n^2 is assumed to be 0 dB.

Inspection of the figures highlights that the LAM detectors achieve the best performance, with a loss w.r.t. the benchmarks of almost 1 dB at $P_D = 0.9$. Furthermore, the GPM and CD methods exhibit nearly an identical behavior, with a gap smaller than 1 dB when compared with the LAM curve. The results confirm both the capabilities of the considered linearization technique to approximate the actual steering vector correctly and of the iterative ascent algorithms to provide close-to-optimum solutions to the optimization problem (6.21). A clear performance advantage over the mismatched detectors (which rely on the nominal parameters to derive the decision statistics) is highlighted, corroborating the effectiveness of the proposed strategies. The plots also reveal that, at $P_D = 0.9$, the benchmark SP detectors experience a performance degradation w.r.t. the full polarized counterparts of 7 dB for the GLRT-based detector and of 5 dB for the 2SGLRT one, pinpointing the advantage of the devised architectures to leverage the polarimetric diversity. Finally, a direct comparison between Figs. 6.3(a) and 6.3(b) does not reveal significant performance differences between the GLRT and 2SGLRT methodologies.

Clutter interference case

In this situation, the useful target echo return is assumed buried in clutter plus noise, with covariance matrix $\mathbf{M} \in \mathbb{H}^{2MN}$ modeled as

$$\mathbf{M} = \sigma_c^2 \boldsymbol{\Sigma}_c \otimes \mathbf{M}_c + \sigma_n^2 \mathbf{I}, \quad (6.59)$$

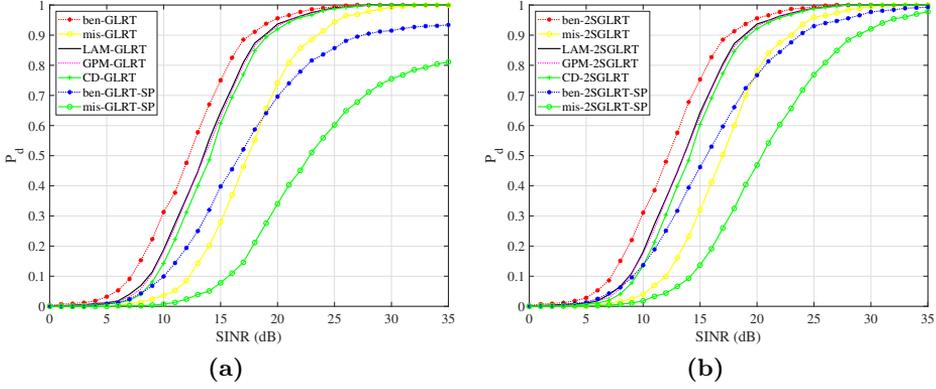


Figure 6.4. Detection performance for a mixed clutter environment (see Table 6.2) with $\rho_c = 0.4$, assuming $K = 320$ and the parameters of Table 6.1. Fig. (a) considers the GLRT-based detectors, whereas Fig. (b) reports the 2SGLRT counterparts.

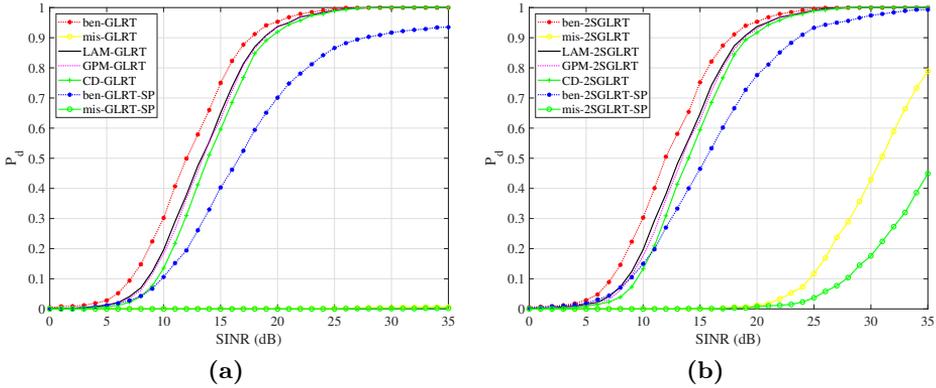


Figure 6.5. Detection performance for a mixed clutter environment (see Table 6.2) with $\rho_c = 0.9$, assuming $K = 320$ and the parameters of Table 6.1. Fig. (a) considers the GLRT-based detectors, whereas Fig. (b) reports the 2SGLRT counterparts.

where σ_c^2 is the clutter power level, $\mathbf{M}_c \in \mathbb{C}^{MN \times MN}$ is an exponentially-shaped matrix with 1-lag correlation coefficient ρ_c accounting for the covariance between the returns from the same polarimetric channel, and

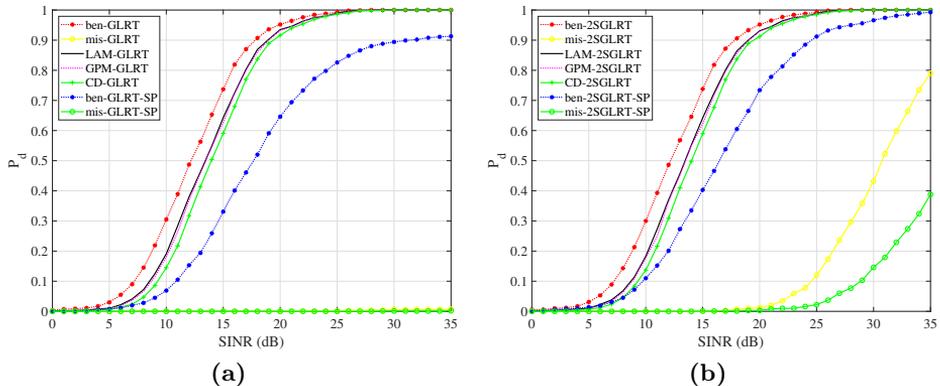


Figure 6.6. Detection performance for a clutter trees environment (see Table 6.2) with $\rho_c = 0.9$, assuming $K = 320$ and the parameters of Table 6.1. Fig. (a) considers the GLRT-based detectors, whereas Fig. (b) reports the 2SGLRT counterparts.

$\Sigma_c = \begin{bmatrix} 1 & \epsilon_c \sqrt{\delta_c} \\ \epsilon_c^* \sqrt{\delta_c} & \delta_c \end{bmatrix} \in \mathbb{C}^{2 \times 2}$ indicates the clutter normalized polarimetric scattering matrix. Besides, without loss of generality $\sigma_n^2 = 0$ dB is assumed and a Clutter to Noise Ratio (CNR), i.e., $\text{CNR} = \sigma_c^2 / \sigma_n^2$, of 30 dB is considered. Typical values of ϵ_c and δ_c for different clutter models are summarized in Table 6.2 (see also [83, 82, 37]).

The detection performance for two different clutter environments and a sample support size of⁸ $K = 320$ are reported in Figs. 6.4, 6.5, and 6.6. Specifically, assuming $\rho_c = 0.4$ and mixed clutter (see Table 6.2), the P_D versus SINR curves of the GLRT-based detectors are displayed in Fig. 6.4(a), whereas the 2SGLRT counterparts are given in Fig. 6.4(b). The results show that the proposed adaptive techniques achieve almost the same detection performance in the high SINR regime, with a degradation, w.r.t. the benchmarks, of 1 dB at $P_D = 0.9$. Precisely, the LAM-based and GPM-based detectors almost achieve the same (and closest to the opti-

⁸Notice that, although the use of $K = 320$ secondary data is rather unlikely in practice, it allows for a good estimation of the covariance matrix. In this case, the target detection performance of the considered algorithms can be accurately evaluated without being affected by possible errors in the estimation of noise/interference, which is the main goal of the presented investigation.

mum) performance, with a very slight advantage over the CD counterparts. This behavior represents a first confirmation of the effectiveness of the considered techniques to handle a challenging clutter-dominated scenario. On the contrary, the mismatched detectors exhibit a clear performance degradation w.r.t. the benchmarks, with a gap between the curves in the order of 6 dB. A further inspection of the figures also highlights that, for this challenging scenario, the SP detectors are not able to provide an adequate detection performance, further corroborating the need to devise appropriate detection strategies leveraging the polarimetric diversity techniques.

In Fig. 6.5, the same scenario as Fig. 6.4 but for $\rho_c = 0.9$, is considered. Inspection of the figure reveals that the proposed detectors achieve almost the same performance. Nonetheless, the mismatched detectors exhibit a noticeable performance degradation in a clutter environment with $\rho = 0.9$ w.r.t. the $\rho = 0.4$ case, especially with reference to the GLRT configuration. In this regard, it is worth analyzing the similarity θ_{sim} between the true and the nominal array steering matrices, measured as

$$\theta_{\text{sim}} = \frac{|\text{tr}\{\check{\mathbf{H}}_t^\dagger \check{\mathbf{H}}\}|}{\sqrt{\text{tr}\{\check{\mathbf{H}}_t^\dagger \check{\mathbf{H}}_t\}} \sqrt{\text{tr}\{\check{\mathbf{H}}^\dagger \check{\mathbf{H}}\}}}, \quad (6.60)$$

with $\check{\mathbf{H}}_t = \mathbf{M}^{-1/2} \mathbf{H}(u, \delta)$ and $\check{\mathbf{H}} = \mathbf{M}^{-1/2} \mathbf{H}(\bar{u}, \bar{\delta})$. Specifically, for the case of $\rho = 0.4$, the similarity is $\theta_{\text{sim}} = 0.5715$, whereas for the $\rho = 0.9$ scenario, $\theta_{\text{sim}} = 0.0930$, which corroborates the larger performance loss experienced by the mismatched detectors in the latter scenario w.r.t. the former one.

In Fig. 6.6 an additional clutter model, i.e., clutter from trees (whose parameters are reported in Table 6.2) with $\rho_c = 0.9$ is considered for the interference simulation. Analysis of the curves highlights the absence of noticeable differences when compared with the mixed clutter case of Fig. 6.5.

In Fig. 6.7, the performance of the detectors is compared for different sample support sizes in a scenario of mixed clutter interference with $\rho_c = 0.9$. In this regard, Figs. 6.7(a), 6.7(c), and 6.7(e) report the performance of the GLRT-based detectors, whereas Figs. 6.7(b), 6.7(d), and 6.7(f) those related to the 2SGLRT-based strategies. Specifically, assuming $K = 320$, $K = 240$, and $K = 120$ secondary data to estimate the covari-

Table 6.2. Typical parameters for polarimetric clutter characterization [37]

Clutter Models	δ_c	ϵ_c	ρ_c
Trees	0.89	0.64	0.9
Mixed	1.08	0.57	0.9

ance matrix, the case of LAM-based detectors is displayed in Figs. 6.7(a) and 6.7(b), that of the GPM-based strategies in Figs. 6.7(c) and 6.7(d), and the situation related to the CD-based detectors is analyzed in Figs. 6.7(e) and 6.7(f). As expected, capitalizing on a larger number of secondary data leads to a clear performance gain, owing to a better estimate of the covariance matrix. Remarkably, even with a reduced sample support size, the performance loss between the proposed receivers and the corresponding benchmarks (either with $K = 320$, $K = 240$, or $K = 120$) remains almost acceptable, with a gap smaller than 2 dB for $P_D = 0.9$, for all the considered architectures. Besides, regardless of the number of secondary data, all the proposed detectors achieve a better detection performance than the SP benchmarks, stressing again the benefit of the polarimetric domain.

To summarize, the detectors based on either the LAM or GPM procedures provide the best solutions (in terms of P_D for a given P_{fa}) to the target detection problem for all the analyzed case studies. Additionally, the proposed methods exhibit substantially the same computational complexity provided the number of iterations is small enough.

6.5 Conclusion

This chapter has considered the design of adaptive detectors for a polarimetric FDA-MIMO radar in a scenario characterized by the presence of interference with unknown spectral properties. The detection problem, formulated in terms of a binary hypothesis test, has been handled by resorting either to the GLRT or to the 2SGLRT criterion, demanding the ML estimation of the interference covariance matrix and, under the H_1 hypothesis, also of the unknown target parameters. To get computationally efficient even though sub-optimal solutions to the resulting optimization problem, three iterative strategies have been proposed. First, a procedure has been developed leveraging a LAM technique, which capitalizes on the target lo-

cation offsets w.r.t. the nominal array steering. Then, a GPM technique has been considered, that iteratively updates the parameters according to the gradient of the objective function. Finally, a CD approach has been investigated, which sequentially optimizes one variable while keeping the other fixed. Noticeably, the proposed detectors ensure the so-called bounded CFAR property. At the analysis stage, the performance of the receivers has been numerically assessed also in comparisons with benchmarks. The results, for both white and clutter interference (either mixed or trees), have pinpointed the effectiveness of the devised architectures to provide a detection performance close to the clairvoyant structure. Furthermore, a clear performance advantage over both mismatched detectors and the single polarization counterparts has been underlined, corroborating the strength of the proposed detectors.

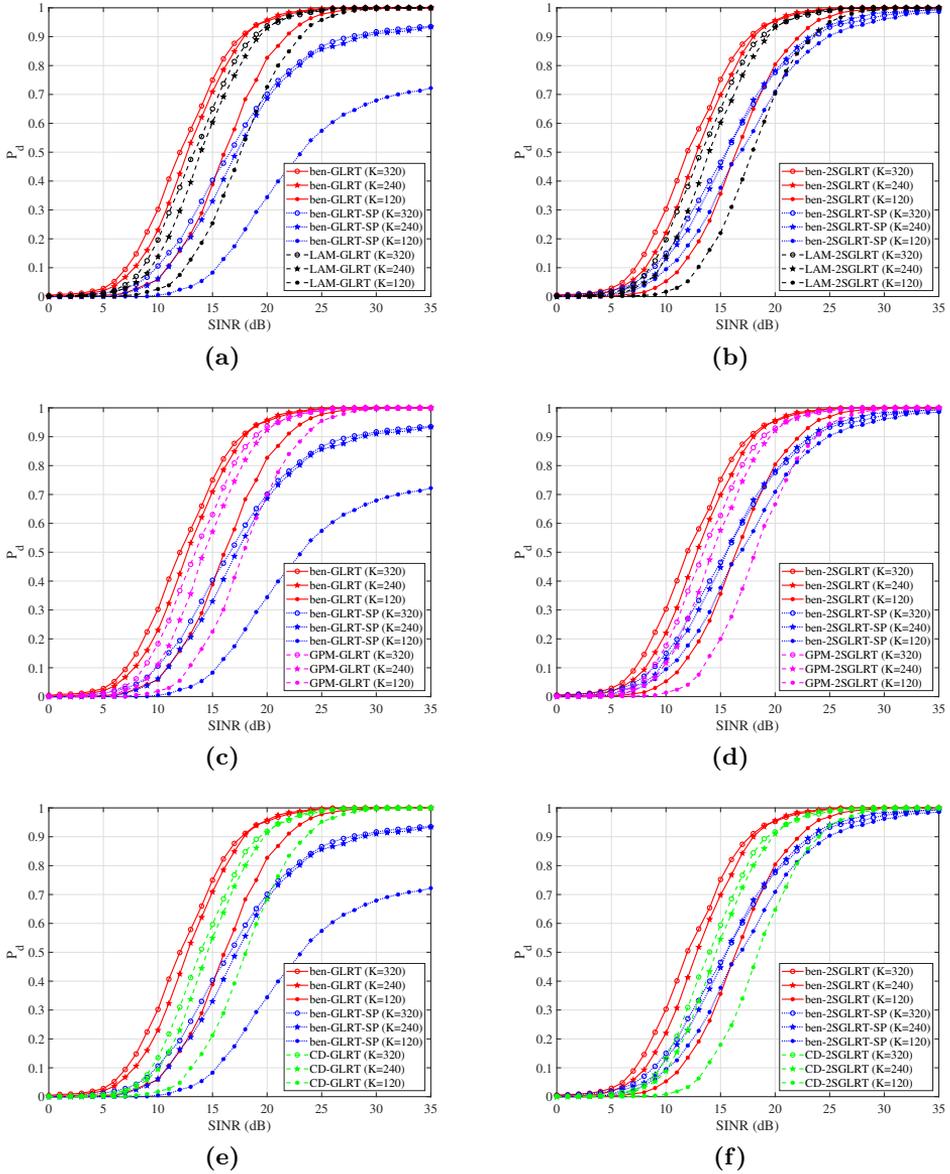


Figure 6.7. Detection performance of the LAM, GPM, and CD methods for different sample support sizes, i.e., $K = 320$, $K = 240$, $K = 120$, assuming a mixed clutter environment (see Table 6.2) with $\rho_c = 0.9$ and the parameters of Table 6.1. Figs. (a), (c), and (e) report the performance of the GLRT-based decision statistics; Figs. (b), (d), and (f) that of the 2SGLRT rules.

Conclusions

This dissertation considered the design of novel signal processing architectures for either multichannel phased array or **FDA-MIMO** radars operating in an adverse environment characterized by the presence of different kinds of interference.

In particular, in the context of a radar system equipped with a multichannel phased array (**ULA** or **URA**), **Chapter 2** presented the problem of simultaneous target detection and angle estimation. The approach pursued in this work relied on the idea of approximating the steering vector of the received echo signal via the **LAM** at the nominal array looking direction, with a resulting functional dependency of the array manifold on the directional cosine offset. This paved the way for single-pulse spatial processing aimed at providing simultaneous target detection and angle estimation. Considering fast architectures for real time operations, the resulting processing can be implemented for every search beam of a multifunction phased array radar. Otherwise, it can be employed in a target confirmation phase where, after a first detection is triggered by a standard detector, one needs to confirm the target presence (thus lowering the P_{fa}) and provide angle estimates. At the analysis stage, the new **GLRT**-based signal processing architectures were compared with classic detectors such as Kelly's **GLRT**, **AMF**, and **SD**, in terms of detection performance. Besides, the **MSEs** of the angular estimates were compared with the **CRLB**, **ML**, and adaptive monopulse. The results highlighted that the bespoke new methodology is a very effective candidate to solve the problem of joint

target detection and angular estimation, providing close-to-optimum detection performance and high quality angular estimates in many scenarios of practical relevance for modern phased array radars.

In **Chapter 3**, an **EM**-based technique was devised for the estimation of a structured covariance matrix in the presence of missing data, considering uncertainty sets of practical interest for radar signal processing applications. The convergence properties for the resulting iterative procedure was studied along with the computation of the convergence rate. Then, the derived methodology was applied to the context of two fundamental radar problems, i.e., beamforming and detection of the number of sources. Finally, numerical results were presented aimed at corroborating the theoretical achievements.

Chapter 4 investigated the target detection problem in a missing-data context comprising Gaussian interference with an unknown and possibly structured covariance matrix. Therefore, under the assumption of a given structure (known a-priori) of the unknown interference covariance matrix, a **GLR**-based detector was devised. The implementation of this approach included the **MLE** of the unknown parameters under both detection hypotheses, which were addressed by **EM**-based procedures. At the analysis stage, the performance of the devised detectors was assessed in terms of P_D versus **SINR** on both simulated and measured data. This latter evaluation was critical for validating the robustness of the proposed detection strategies on real data, including potential mismatches (due to hardware imperfections) that were not taken into account at the design stage. For comparison purposes, two additional detectors were considered. The former was a benchmark that assumed direct access to the complete-data set, whereas the latter replaced the missing values via linear interpolation.

Chapter 5 considered the problem of joint target angle and incremental range estimation using an **FDA-MIMO** radar in Gaussian interference with known spectral properties. Specifically, an **ML** estimator for the unknown target parameters (i.e., angle, incremental range, and echo-amplitude) was derived leveraging three approximated methods: the **CD**, the adaptive monopulse, and the generalized monopulse. At the analysis stage, the **CRLB** for the angle and incremental range estimation via an **FDA-MIMO** radar were derived. The **RMSEs** of the estimates versus input **SINR** were provided. Moreover, an extensive bias and variance analysis

was developed to show the effectiveness of the considered approximated estimation methods (also including tapered and double-step versions of the monopulse procedures).

Chapter 6 addressed the target detection problem with a polarimetric FDA-MIMO radar, leveraging the methodology studied in **Chapter 5**. In particular, two practical detection architectures were derived resorting to the GLRT and the 2SGLRT criteria. However, their implementation demanded the solution of an optimization problem involving the concentrated log-likelihood functions under the target presence hypothesis. Considering feasible detectors, three strategies were devised, namely, LAM method, GPM, and CD procedure. At the analysis stage, the detection performance of the proposed architectures was evaluated in white noise and clutter (mixed or trees) environment and compared with both benchmark and mismatched receivers. These cases rely on the true and nominal target location parameters, respectively. The results highlighted that the proposed radar architectures can achieve a better P_D than their counterparts (also including the single polarization receiver), thus representing a viable solution for practical implementations.

Future research efforts include the following points:

- The extension of the LAM approach, derived in **Chapter 2**, to account for some deviations from the proposed homogeneous disturbance model. For example, non-Gaussian interference, the occurrence of clutter discretely and/or multiple targets (including false targets) within a specific range cell, as well as the design of alternative decision criteria such as the Wald test [59, 107, 69] possibly accounting for rejection of signals outside a specific region in the $u-v$ space [86, 15].
 - The application of the methodology developed in **Chapter 3** for a time processing background where some pulses of the received train may experience unwanted sporadic radio frequency interference. This means that some slow time samples from some given range cells are missed and the lack of this data has to be properly accounted for at the signal processing design stage.
 - The derivation of bespoke detectors, based on the framework studied in **Chapter 4**, accounting for distributed (range-spread) targets, as
-

well as the generalization of the architectures devised in **Chapter 4** to partially homogeneous and heterogeneous environments (also including other covariance structures of practical relevance). Moreover, it would be of great interest to consider the case of a multistatic radar where different sensors observe the same scene whereby the missing-data are diverse from sensor to sensor.

- The design of estimators for an **FDA-MIMO** radar, tailored for specific jammer and/or clutter scenarios, as well as the extension of the approach developed in **Chapter 5** to the case of multiple targets via compressed sensing techniques. Furthermore, the analysis of the architectures derived in **Chapter 5** on real **FDA-MIMO** radar data represents definitely another research topic of primary concern.
- Finally, the study of additional decision strategies to tackle the detection problem analyzed in **Chapter 6**, according to other sub-optimal criteria (e.g., Rao [36] and Wald [59] tests), as well as the extension of the framework proposed in **Chapter 6** to different scenarios including compound Gaussian disturbance [31], structured interference covariance matrix, and the case of extended targets occupying multiple range cells.

Appendix **A**

Appendix to Chapter 2

A.1 Adaptive Monopulse Estimation Algorithm

In **Algorithm 13** the steps involved in the derivation of the Adaptive Monopulse Estimator proposed in [79] are summarized.

Algorithm 13 Adaptive Monopulse Estimation Algorithm [79, 112]

Input: \mathbf{r} , \mathbf{S} , \mathbf{p} , $\dot{\mathbf{p}}_u$, $\dot{\mathbf{p}}_v$.

Output: DOA estimate $\hat{\boldsymbol{\theta}}_{NME}^*$.

- 1: let $r_x = \text{Re} \left\{ \frac{\dot{\mathbf{p}}_u^\dagger \mathbf{S}^{-1} \mathbf{r}}{\mathbf{p}^\dagger \mathbf{S}^{-1} \mathbf{r}} \right\}$ and $r_y = \text{Re} \left\{ \frac{\dot{\mathbf{p}}_v^\dagger \mathbf{S}^{-1} \mathbf{r}}{\mathbf{p}^\dagger \mathbf{S}^{-1} \mathbf{r}} \right\}$;
 - 2: compute the components of the bias correction

$$\mu_x = \frac{\text{Re} \{ \dot{\mathbf{p}}_u^\dagger \mathbf{S}^{-1} \mathbf{p} \}}{\mathbf{p}^\dagger \mathbf{S}^{-1} \mathbf{p}}, \text{ and } \mu_y = \frac{\text{Re} \{ \dot{\mathbf{p}}_v^\dagger \mathbf{S}^{-1} \mathbf{p} \}}{\mathbf{p}^\dagger \mathbf{S}^{-1} \mathbf{p}};$$
 - 3: compute $F_u = 2r_x - 2\mu_x$, $F_v = 2r_y - 2\mu_y$;

$$F_{uu} = 2\mu_x^2 - 2 \frac{\dot{\mathbf{p}}_u^\dagger \mathbf{S}^{-1} \dot{\mathbf{p}}_u}{\mathbf{p}^\dagger \mathbf{S}^{-1} \mathbf{p}}, \quad F_{uv} = 2\mu_x \mu_y - 2 \frac{\text{Re} \{ \dot{\mathbf{p}}_u^\dagger \mathbf{S}^{-1} \dot{\mathbf{p}}_v \}}{\mathbf{p}^\dagger \mathbf{S}^{-1} \mathbf{p}},$$

$$F_{vv} = 2\mu_y^2 - 2 \frac{\dot{\mathbf{p}}_v^\dagger \mathbf{S}^{-1} \dot{\mathbf{p}}_v}{\mathbf{p}^\dagger \mathbf{S}^{-1} \mathbf{p}};$$
 - 4: output $\hat{\boldsymbol{\theta}}_{NME}^* = - \begin{pmatrix} F_{uu} & F_{uv} \\ F_{uv}^* & F_{vv} \end{pmatrix}^{-1} \begin{pmatrix} F_u \\ F_v \end{pmatrix}$.
-

A.2 Proof of Proposition 1

Proof. Since (2.23) is a continuous function and $[-\alpha, \alpha]$ a compact set, Weierstrass Theorem [19] ensures the existence of a maximizer. Candidate optimal solutions are the boundary points, i.e., $-\alpha, \alpha$, as well as the feasible points satisfying the first order optimality condition, i.e., nulling the derivative of the objective function. Computing the derivative of (2.23) and setting it to zero, the second-order polynomial in (2.25) is obtained. Hence, in the worst case, the optimal solution to Problem (2.23) has to be searched among four points, i.e., the roots of (2.25) and the boundary points $-\alpha$ and α . \square

A.3 Solution to Problem (2.41)

Being the objective function in (2.41) continuous and the feasible set non-empty and compact, the existence of a global maximizer is guaranteed by Weierstrass Theorem [19]. Furthermore, the objective function is continuously differentiable. As a result, any optimal solution $\Delta\theta^*$ is either a feasible stationary point, i.e., $\Delta\theta^* \in \mathcal{S}$ and satisfies the first order optimality condition, or it belongs to the boundary of the feasible set.

The stationary points of the objective function in (2.41) are obtained by nulling the objective gradient, namely by solving the equation

$$\nabla \left[(\bar{\mathbf{p}} + \bar{\mathbf{H}} \Delta\theta)^\dagger (\bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger - q_n \mathbf{I}) (\bar{\mathbf{p}} + \bar{\mathbf{H}} \Delta\theta) \right] = 0. \quad (\text{A.1})$$

After some straightforward calculations, equation (A.1) can be expressed as

$$\text{Re} \left\{ \bar{\mathbf{H}}^\dagger (\bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger - q_n \mathbf{I}) \bar{\mathbf{p}} \right\} + \text{Re} \left\{ \bar{\mathbf{H}}^\dagger (\bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger - q_n \mathbf{I}) \bar{\mathbf{H}} \right\} \Delta\theta = 0 \quad (\text{A.2})$$

implying that¹ the unique stationary point is

$$\Delta\theta_1 = -\text{Re}^{-1} \left\{ \bar{\mathbf{H}}^\dagger (\bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger - q_n \mathbf{I}) \bar{\mathbf{H}} \right\} \text{Re} \left\{ \bar{\mathbf{H}}^\dagger (\bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger - q_n \mathbf{I}) \bar{\mathbf{p}} \right\}, \quad (\text{A.3})$$

which is a candidate optimal solution if it is also feasible for (2.41), i.e. $\Delta\theta_1 \in \mathcal{S}$.

Let us now focus on the study of the objective function restricted to the boundary of \mathcal{S} , determining, for each edge, the maximizer.

Case A. *Left and right edges* ($\Delta u = \pm\alpha$).

Let us define $\bar{\mathbf{p}}_{\pm\alpha} = \bar{\mathbf{p}} \pm \alpha \dot{\bar{\mathbf{p}}}_u$. The restriction of (2.41) to $\Delta u = \pm\alpha$, yields

$$\max_{|\Delta v| \leq \beta} (\bar{\mathbf{p}}_{\pm\alpha} + \dot{\bar{\mathbf{p}}}_v \Delta v)^\dagger (\bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger - q_n \mathbf{I}) (\bar{\mathbf{p}}_{\pm\alpha} + \dot{\bar{\mathbf{p}}}_v \Delta v). \quad (\text{A.4})$$

The expression appearing in (A.4) can be written as the following quadratic function

$$\rho' \Delta v^2 + 2\tau'_\pm \Delta v + \nu'_\pm, \quad (\text{A.5})$$

where

$$\rho' = \dot{\bar{\mathbf{p}}}_v^\dagger (\bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger - q_n \mathbf{I}) \dot{\bar{\mathbf{p}}}_v, \quad (\text{A.6})$$

$$\tau'_\pm = \text{Re} \{ \dot{\bar{\mathbf{p}}}_v^\dagger (\bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger - q_n \mathbf{I}) \bar{\mathbf{p}}_{\pm\alpha} \}, \quad (\text{A.7})$$

$$\nu'_\pm = \bar{\mathbf{p}}_{\pm\alpha}^\dagger (\bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger - q_n \mathbf{I}) \bar{\mathbf{p}}_{\pm\alpha}. \quad (\text{A.8})$$

Now, if $\rho' < 0$, (A.4) is a convex optimization problem and the optimal solution is

$$\Delta v_\pm^* = \min(\beta, \max(-\tau'_\pm / \rho', -\beta)), \quad (\text{A.9})$$

¹To avoid unnecessary complications, it is assumed that the matrix $\text{Re} \{ \bar{\mathbf{H}}^\dagger (\bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger - q_n \mathbf{I}) \bar{\mathbf{H}} \}$ is full-rank. The rank-deficient event never occurred in our Monte Carlo trials.

However, it is straightforward to account for this situation too. Specifically, if (A.2) admits a solution, the unique candidate optimal solution is a maximizer of the one-dimensional optimization problem resulting from the restriction of (2.41) to the line $\Delta\theta = \gamma \text{Re} \{ \bar{\mathbf{H}}^\dagger (\bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger - q_n \mathbf{I}) \bar{\mathbf{p}} \}$ with $\gamma \in \mathbb{R}$ the optimization variable. An optimized $\Delta\theta_1 = \gamma^* \text{Re} \{ \bar{\mathbf{H}}^\dagger (\bar{\mathbf{r}} \bar{\mathbf{r}}^\dagger - q_n \mathbf{I}) \bar{\mathbf{p}} \}$ can be obtained following the approach developed to determine boundary maximizers (see **Case A** and **Case B**).

otherwise²

$$\Delta v_{\pm}^* = \beta \tau'_{\pm} / |\tau'_{\pm}|. \quad (\text{A.10})$$

Hence, the two candidate optimal solutions associated with the right and left edges are given by

$$\Delta \theta_2 = [\alpha, \Delta v_{+}^*] \quad \text{and} \quad \Delta \theta_3 = [-\alpha, \Delta v_{-}^*]. \quad (\text{A.11})$$

Case B. *Upper and lower edges* ($\Delta v = \pm\beta$).

Denoting by $\bar{\mathbf{p}}_{\pm\beta} = \bar{\mathbf{p}} \pm \beta \dot{\bar{\mathbf{p}}}_v$ and following the same line of reasoning as for **Case A**, the candidate optimal solutions take the form

$$\Delta \theta_4 = [\Delta u_{+}^*, \beta] \quad \text{and} \quad \Delta \theta_5 = [\Delta u_{-}^*, -\beta], \quad (\text{A.12})$$

where,

$$\Delta u_{\pm}^* = \min(\alpha, \max(-\tau''_{\pm}/\rho'', -\alpha)), \quad (\text{A.13})$$

with

$$\tau''_{\pm} = \text{Re}\{\dot{\bar{\mathbf{p}}}_u^{\dagger} (\bar{\mathbf{r}} \bar{\mathbf{r}}^{\dagger} - q_n \mathbf{I}) \bar{\mathbf{p}}_{\pm\beta}\}, \quad (\text{A.14})$$

otherwise³

$$\Delta u_{\pm}^* = \alpha \tau''_{\pm} / |\tau''_{\pm}|. \quad (\text{A.15})$$

Based on the above results, a closed form optimal solution to (2.41) is

$$\begin{aligned} \Delta \theta^* = \arg \max_{\Delta \theta \in \{\Delta \theta_i\}_{i=1}^5 \cap \mathcal{S}} & (\bar{\mathbf{p}} + \bar{\mathbf{H}} \Delta \theta)^{\dagger} \bar{\mathbf{r}} \bar{\mathbf{r}}^{\dagger} (\bar{\mathbf{p}} + \bar{\mathbf{H}} \Delta \theta) \\ & - q_n (\bar{\mathbf{p}} + \bar{\mathbf{H}} \Delta \theta)^{\dagger} (\bar{\mathbf{p}} + \bar{\mathbf{H}} \Delta \theta). \end{aligned} \quad (\text{A.16})$$

A.4 CRLB Derivation

In the following, the **CRLB** evaluation for the estimation of the parameters Δu and Δv is conducted, assuming either the linearized model $\mathbf{p}_a = \mathbf{p} + \dot{\mathbf{p}}_u \Delta u + \dot{\mathbf{p}}_v \Delta v$ or the actual manifold $\mathbf{p}(\bar{u} + \Delta u, \bar{v} + \Delta v)$. Hereafter, $\dot{\mathbf{p}}_u$ and $\dot{\mathbf{p}}_v$ denote the first order derivative with respect to Δu and

²If $\tau'_{\pm} = 0$, $\Delta v_{\pm}^* = \beta$ can be considered without loss of generality.

³If $\tau''_{\pm} = 0$, $\Delta u_{\pm}^* = \alpha$ can be considered without loss of generality.

Δv , respectively, of the considered steering vector model evaluated at the actual DOA offsets. The vector of the unknown parameters Θ is partitioned as follows:

$$\Theta = [\Theta_u, \Theta_w]^T, \quad (\text{A.17})$$

where

$$\Theta_u = [a_r, a_i]^T \quad (\text{A.18})$$

contains the unwanted (nuisance) parameters (signal complex amplitude) and

$$\Theta_w = \Delta\theta = \begin{cases} \Delta u, & (\text{1-D}) \\ [\Delta u, \Delta v]^T, & (\text{2-D}) \end{cases} \quad (\text{A.19})$$

contains those of interest.

Resorting to the Slepian–Bangs formula [105, p. 927, eq. 8.34], the FIM can be written as a block partitioned matrix

$$\mathbf{J}(\Theta) = \begin{bmatrix} \mathbf{J}_{\Theta_u\Theta_u}(\Theta) & \mathbf{J}_{\Theta_u\Theta_w}(\Theta) \\ \mathbf{J}_{\Theta_w\Theta_u}(\Theta) & \mathbf{J}_{\Theta_w\Theta_w}(\Theta) \end{bmatrix}, \quad (\text{A.20})$$

where

$$\mathbf{J}_{\Theta_u\Theta_u}(\Theta) = 2\mathbf{p}_a^\dagger(\Delta\theta)\mathbf{M}^{-1}\mathbf{p}_a(\Delta\theta)\mathbf{I}_2, \quad (\text{A.21})$$

$$\mathbf{J}_{\Theta_u\Theta_w}(\Theta) = \begin{cases} 2\operatorname{Re}\left\{ (a\mathbf{p}_a^\dagger(\Delta\theta)\mathbf{M}^{-1}\dot{\mathbf{p}}_u) \begin{bmatrix} 1 \\ -j \end{bmatrix} \right\} & (\text{1-D}) \\ 2\operatorname{Re}\left\{ \begin{bmatrix} 1 \\ -j \end{bmatrix} \begin{bmatrix} a\mathbf{p}_a^\dagger(\Delta\theta)\mathbf{M}^{-1}\dot{\mathbf{p}}_u \\ a\mathbf{p}_a^\dagger(\Delta\theta)\mathbf{M}^{-1}\dot{\mathbf{p}}_v \end{bmatrix}^T \right\} & (\text{2-D}) \end{cases}, \quad (\text{A.22})$$

and

$$\mathbf{J}_{\Theta_w\Theta_w}(\Theta) = \begin{cases} 2|a|^2\dot{\mathbf{p}}_u^\dagger\mathbf{M}^{-1}\dot{\mathbf{p}}_u & (\text{1-D}) \\ 2|a|^2 \begin{bmatrix} \dot{\mathbf{p}}_u^\dagger\mathbf{M}^{-1}\dot{\mathbf{p}}_u & \operatorname{Re}\{\dot{\mathbf{p}}_u^\dagger\mathbf{M}^{-1}\dot{\mathbf{p}}_v\} \\ \operatorname{Re}\{\dot{\mathbf{p}}_u^\dagger\mathbf{M}^{-1}\dot{\mathbf{p}}_v\} & \dot{\mathbf{p}}_v^\dagger\mathbf{M}^{-1}\dot{\mathbf{p}}_v \end{bmatrix} & (\text{2-D}) \end{cases}. \quad (\text{A.23})$$

Finally, the CRLB for the parameters of interest is

$$\mathbf{C}_{CR}(\Theta_w) = [\mathbf{J}_{\Theta_w\Theta_w}(\Theta) - \mathbf{J}_{\Theta_w\Theta_u}(\Theta)\mathbf{J}_{\Theta_u\Theta_u}^{-1}(\Theta)\mathbf{J}_{\Theta_u\Theta_w}(\Theta)]^{-1}. \quad (\text{A.24})$$

Appendix B

Appendix to Chapter 3

B.1 Proof of (3.22)

Following the same line of reasoning as in [56], for any $i \in \{1, \dots, K\}$

$$\mathbf{C}_i = \mathbb{E}[\mathbf{r}_i \mathbf{r}_i^\dagger | \mathbf{A}_i \mathbf{r}_i, M] = \mathbf{B}_i^\top \mathbb{E}[\mathbf{B}_i \mathbf{r}_i \mathbf{r}_i^\dagger \mathbf{B}_i^\top | \mathbf{A}_i \mathbf{r}_i, M] \mathbf{B}_i, \quad (\text{B.1})$$

where

$$\mathbf{B}_i = [\mathbf{A}_i^\top \bar{\mathbf{A}}_i^\top]^\top \quad (\text{B.2})$$

which satisfies

$$\mathbf{B}_i^\top \mathbf{B}_i = \mathbf{I}. \quad (\text{B.3})$$

To provide a closed form expression to \mathbf{C}_i , let us observe that

$$\begin{aligned} & \mathbb{E}[\mathbf{B}_i \mathbf{r}_i \mathbf{r}_i^\dagger \mathbf{B}_i^\top | \mathbf{A}_i \mathbf{r}_i, M] \\ &= \mathbb{E} \left[\begin{bmatrix} \mathbf{A}_i \mathbf{r}_i \\ \bar{\mathbf{A}}_i \mathbf{r}_i \end{bmatrix} \begin{bmatrix} \mathbf{r}_i^\dagger \mathbf{A}_i^\dagger & \mathbf{r}_i^\dagger \bar{\mathbf{A}}_i^\dagger \end{bmatrix} | \mathbf{A}_i \mathbf{r}_i, M \right] \\ &= \mathbb{E} \left[\begin{bmatrix} \mathbf{A}_i \mathbf{r}_i \mathbf{r}_i^\dagger \mathbf{A}_i^\dagger & \mathbf{A}_i \mathbf{r}_i \mathbf{r}_i^\dagger \bar{\mathbf{A}}_i^\dagger \\ \bar{\mathbf{A}}_i \mathbf{r}_i \mathbf{r}_i^\dagger \mathbf{A}_i^\dagger & \bar{\mathbf{A}}_i \mathbf{r}_i \mathbf{r}_i^\dagger \bar{\mathbf{A}}_i^\dagger \end{bmatrix} | \mathbf{A}_i \mathbf{r}_i, M \right] \\ &= \begin{bmatrix} \mathbf{y}_i \mathbf{y}_i^\dagger & \mathbf{y}_i \mathbb{E} \left[\mathbf{r}_i^\dagger \bar{\mathbf{A}}_i^\dagger | \mathbf{A}_i \mathbf{r}_i, M \right] \\ \mathbb{E} \left[\bar{\mathbf{A}}_i \mathbf{r}_i | \mathbf{A}_i \mathbf{r}_i, M \right] \mathbf{y}_i^\dagger & \mathbb{E} \left[\bar{\mathbf{A}}_i \mathbf{r}_i \mathbf{r}_i^\dagger \bar{\mathbf{A}}_i^\dagger | \mathbf{A}_i \mathbf{r}_i, M \right] \end{bmatrix} \end{aligned} \quad (\text{B.4})$$

$$= \begin{bmatrix} \mathbf{y}_i \mathbf{y}_i^\dagger & \mathbf{y}_i \boldsymbol{\mu}_i^\dagger \\ \boldsymbol{\mu}_i \mathbf{y}_i^\dagger & \mathbb{E} \left[\bar{\mathbf{A}}_i \mathbf{r}_i \mathbf{r}_i^\dagger \bar{\mathbf{A}}_i^\dagger | \mathbf{A}_i \mathbf{r}_i, M \right] \end{bmatrix},$$

where

$$\boldsymbol{\mu}_i = \mathbb{E} \left[\bar{\mathbf{A}}_i \mathbf{r}_i | \mathbf{A}_i \mathbf{r}_i, M \right] = \bar{\mathbf{A}}_i M \mathbf{A}_i^\dagger (\mathbf{A}_i M \mathbf{A}_i^\dagger)^{-1} \mathbf{y}_i \quad (\text{B.5})$$

and

$$\begin{aligned} & \mathbb{E} \left[\bar{\mathbf{A}}_i \mathbf{r}_i \mathbf{r}_i^\dagger \bar{\mathbf{A}}_i^\dagger | \mathbf{A}_i \mathbf{r}_i, M \right] \\ &= \mathbf{G}_i + \bar{\mathbf{A}}_i M \mathbf{A}_i^\dagger (\mathbf{A}_i M \mathbf{A}_i^\dagger)^{-1} \mathbf{y}_i \mathbf{y}_i^\dagger (\mathbf{A}_i M \mathbf{A}_i^\dagger)^{-1} \mathbf{A}_i M \bar{\mathbf{A}}_i^\dagger, \end{aligned} \quad (\text{B.6})$$

with

$$\mathbf{G}_i = \bar{\mathbf{A}}_i M \bar{\mathbf{A}}_i^\dagger - \bar{\mathbf{A}}_i M \mathbf{A}_i^\dagger (\mathbf{A}_i M \mathbf{A}_i^\dagger)^{-1} \mathbf{A}_i M \bar{\mathbf{A}}_i^\dagger. \quad (\text{B.7})$$

Exploiting the above results, it follows that

$$\begin{aligned} \mathbf{C}_i &= \mathbf{B}_i^\top \mathbb{E} [\mathbf{B}_i \mathbf{r}_i \mathbf{r}_i^\dagger \mathbf{B}_i^\top | \mathbf{A}_i \mathbf{r}_i, M] \mathbf{B}_i \\ &= \mathbf{A}_i^\dagger \mathbf{y}_i \mathbf{y}_i^\dagger \mathbf{A}_i + \bar{\mathbf{A}}_i^\dagger \bar{\mathbf{A}}_i M \mathbf{A}_i^\dagger (\mathbf{A}_i M \mathbf{A}_i^\dagger)^{-1} \mathbf{y}_i \mathbf{y}_i^\dagger \mathbf{A}_i \\ &\quad + \mathbf{A}_i^\dagger \mathbf{y}_i \mathbf{y}_i^\dagger (\mathbf{A}_i M \mathbf{A}_i^\dagger)^{-1} \mathbf{A}_i M \bar{\mathbf{A}}_i^\dagger \bar{\mathbf{A}}_i \\ &\quad + \bar{\mathbf{A}}_i^\dagger \bar{\mathbf{A}}_i M \bar{\mathbf{A}}_i^\dagger \bar{\mathbf{A}}_i - \bar{\mathbf{A}}_i^\dagger \bar{\mathbf{A}}_i M \mathbf{A}_i^\dagger (\mathbf{A}_i M \mathbf{A}_i^\dagger)^{-1} \mathbf{A}_i M \bar{\mathbf{A}}_i^\dagger \bar{\mathbf{A}}_i \\ &\quad + \bar{\mathbf{A}}_i^\dagger \bar{\mathbf{A}}_i M \mathbf{A}_i^\dagger (\mathbf{A}_i M \mathbf{A}_i^\dagger)^{-1} \mathbf{y}_i \mathbf{y}_i^\dagger (\mathbf{A}_i M \mathbf{A}_i^\dagger)^{-1} \mathbf{A}_i M \bar{\mathbf{A}}_i^\dagger \bar{\mathbf{A}}_i. \end{aligned}$$

Finally, after some algebraic manipulations,

$$\mathbf{C}_i = (\mathbf{A}_i^\dagger \mathbf{y}_i + \bar{\mathbf{A}}_i^\dagger \boldsymbol{\mu}_i) (\mathbf{A}_i^\dagger \mathbf{y}_i + \bar{\mathbf{A}}_i^\dagger \boldsymbol{\mu}_i)^\dagger + \bar{\mathbf{A}}_i^\dagger \mathbf{G}_i \bar{\mathbf{A}}_i \quad (\text{B.8})$$

■

B.2 Proof of Proposition 2

This Appendix is organized in two parts: in Subsection B.2.1 the proof of the first item of Proposition 2 is provided, whereas Subsection B.2.2 deals with second claim.

B.2.1 Proof of the first item

As first step toward the proof, let us observe that a maximizer to Problem (3.25) exists, provided that $\Sigma^{(h-1)} \succ \mathbf{0}$. Indeed, according to [24], if $\Sigma^{(h-1)} \succ \mathbf{0}$ there exist two positive constants a and b such that Problem (3.25) is equivalent to

$$\begin{cases} \max_{\mathbf{M}} & -\ln(\det(\mathbf{M})) - \text{tr}\{\mathbf{M}^{-1}\Sigma^{(h-1)}\} \\ \text{s.t.} & \mathbf{M} \in \mathcal{C} \\ & \lambda_{\min}(\mathbf{M}) \geq a \\ & \text{tr}\{\mathbf{M}\} \leq b \end{cases} \quad (\text{B.9})$$

and any optimal solution $\bar{\boldsymbol{\theta}}$ to (3.25) must comply with $\lambda_{\min}(\mathbf{M}(\bar{\boldsymbol{\theta}})) \geq a$ and $\text{tr}\{\mathbf{M}(\bar{\boldsymbol{\theta}})\} \leq b$. As a result, being $\mathcal{C} \cap \{\mathbf{M} \succeq \mathbf{0} : \lambda_{\min}(\mathbf{M}) \geq a\} \cap \{\mathbf{M} \succeq \mathbf{0} : \text{tr}\{\mathbf{M}\} \leq b\}$ a compact set of positive definite matrices, Problem (B.9) admits a global optimal solution $\mathbf{M}^* \succ \mathbf{0}$, due to Weierstrass Theorem, and any optimal solution is positive definite. Finally, since $\mathbf{M}^* \in \mathcal{C}$, there exists $\boldsymbol{\theta}^*$ such that $\mathbf{M}^* = \mathbf{M}(\boldsymbol{\theta}^*)$, i.e., $\boldsymbol{\theta}^{(h)} = \boldsymbol{\theta}^*$ solves Problem (3.25) and the M-step at the h -th iteration is well-defined, being $\hat{\mathbf{M}}(\boldsymbol{\theta}^{(h)}) \succ \mathbf{0}$.

Let us now show that

$$\Sigma^{(h-1)} = \frac{1}{K} \mathbb{E}[DD^\dagger | \mathbf{Y}, \{\mathbf{A}_i\}_{i=1}^N, \hat{\mathbf{M}}(\boldsymbol{\theta}^{(h-1)})] \succ \mathbf{0},$$

almost surely if $\hat{\mathbf{M}}(\boldsymbol{\theta}^{(h-1)}) \succ \mathbf{0}$ and $K \geq N$, where $\mathbf{D} = [\mathbf{r}_1, \dots, \mathbf{r}_K] \in \mathbb{C}^{N,K}$. To this end, note that the random matrix \mathbf{D} conditioned on \mathbf{Y} is full rank with probability one, for any (but for a zero-measure set) realization of \mathbf{Y} . As a consequence, for any $\mathbf{v} \in \mathbb{C}^N$, the random variable (conditioned on \mathbf{Y})

$$\mathbf{v}^\dagger \left(DD^\dagger \right)_{|\mathbf{Y}, \{\mathbf{A}_i\}_{i=1}^N, \hat{\mathbf{M}}(\boldsymbol{\theta}^{(h-1)})} \mathbf{v}$$

is greater than zero with probability one, which implies

$$\frac{1}{K} \mathbb{E}[DD^\dagger | \mathbf{Y}, \{\mathbf{A}_i\}_{i=1}^N, \hat{\mathbf{M}}(\boldsymbol{\theta}^{(h-1)})] \succ \mathbf{0}.$$

Thus, it follows that $\hat{\mathbf{M}}(\boldsymbol{\theta}^{(0)}) \succ \mathbf{0}$ ensures $\hat{\mathbf{M}}(\boldsymbol{\theta}^{(h)}) \succ \mathbf{0}$ for all $h \geq 1$,

namely the M-step is well defined for all $h \geq 1$, because of

$$\hat{\mathbf{M}}(\boldsymbol{\theta}^{(h-1)}) \succ \mathbf{0} \Rightarrow \boldsymbol{\Sigma}^{(h-1)} \succ \mathbf{0} \Rightarrow \hat{\mathbf{M}}(\boldsymbol{\theta}^{(h)}) \succ \mathbf{0}, \forall h \geq 1.$$

Finally, the monotonically increasing behavior of the observed-data likelihood function results from the well known properties of EM iterations [42]. The first item of the proposition is thus proved.

B.2.2 Proof of the second item

Before proceeding with the proof, let us introduce the definitions of Bouligand tangent cone and B-stationary point [55, 85].

Definition 1. Given a set $\mathcal{Z} \subseteq \mathbb{R}^M$, the Bouligand tangent cone $\mathcal{I}_{\mathcal{Z}}(\mathbf{z}_0)$ of \mathcal{Z} at $\mathbf{z}_0 \in \mathcal{Z}$, is defined as

$$\mathcal{I}_{\mathcal{Z}}(\mathbf{z}_0) = \left\{ \mathbf{d} \in \mathbb{R}^M : \begin{array}{l} \text{there exist two sequences } \mathbf{z}^{(k)} \in \mathcal{Z} \rightarrow \mathbf{z}_0 \\ \text{and } \tau^{(k)} \in \mathbb{R}^{++} \rightarrow 0 \text{ such that } \mathbf{d} = \lim_{k \rightarrow \infty} \frac{\mathbf{z}^{(k)} - \mathbf{z}_0}{\tau^{(k)}} \end{array} \right\}$$

Definition 2. Given an optimization problem

$$\mathcal{P} \begin{cases} \max_{\mathbf{z}} & f(\mathbf{z}) \\ \text{s.t.} & \mathbf{z} \in \mathcal{Z} \subseteq \mathbb{R}^M \end{cases} \quad (\text{B.10})$$

a feasible solution $\mathbf{z}^* \in \mathcal{Z}$ is a B-stationary point to Problem \mathcal{P} if

$$\lim_{\tau_1 \downarrow 0} \frac{f(\mathbf{z}^* + \tau_1 \mathbf{d}) - f(\mathbf{z}^*)}{\tau_1} \leq 0, \quad \forall \mathbf{d} \in \mathcal{I}_{\mathcal{Z}}(\mathbf{z}^*) \quad (\text{B.11})$$

In a nutshell, the Bouligand tangent cone generalizes the concept of feasible directions emanating from a point, allowing the extension of the stationarity notion (via the B-stationarity) to a broader class of optimization problems.

Let us now focus on the convergence of $\mathcal{L}_y(\boldsymbol{\theta}^{(h)} | \mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K)$ and

$\hat{\mathbf{M}}(\boldsymbol{\theta}^{(h)})$. In this respect, note that

$$\begin{aligned}
& \mathcal{L}_y(\boldsymbol{\theta}|\mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K) \\
&= - \sum_{i=1}^K \ln(\det(\mathbf{A}_i \mathbf{M}(\boldsymbol{\theta}) \mathbf{A}_i^\dagger)) - \text{tr}\{(\mathbf{A}_i \mathbf{M}(\boldsymbol{\theta}) \mathbf{A}_i^\dagger)^{-1} \mathbf{y}_i \mathbf{y}_i^\dagger\} \\
&\leq - \sum_{i=1}^K p_i \ln(\lambda_{\min}(\mathbf{A}_i \mathbf{M}(\boldsymbol{\theta}) \mathbf{A}_i^\dagger)) - \frac{\|\mathbf{y}_i\|^2}{\text{tr}\{\mathbf{A}_i \mathbf{M}(\boldsymbol{\theta}) \mathbf{A}_i^\dagger\}} \\
&\leq - \sum_{i=1}^K p_i \ln(\lambda_{\min}(\mathbf{M}(\boldsymbol{\theta}))) - \frac{\|\mathbf{y}_i\|^2}{\text{tr}\{\mathbf{M}(\boldsymbol{\theta})\}}
\end{aligned}$$

where the last inequality stems from the Eigenvalue Interlacing Theorem [52]. Hence, being \mathcal{C} a closed set of positive definite matrices, it follows that

$$\min_{\mathbf{M} \in \mathcal{C}} \lambda_{\min}(\mathbf{M}) = \delta > 0,$$

which entails

$$\mathcal{L}_y(\boldsymbol{\theta}|\mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K) \leq - \left(\sum_{i=1}^K p_i \right) \log(\delta) - \frac{\sum_{i=1}^K \|\mathbf{y}_i\|^2}{\text{tr}\{\mathbf{M}(\boldsymbol{\theta})\}} \quad (\text{B.12})$$

As a result, the log-likelihood diverges to $-\infty$ as $\text{tr}\{\mathbf{M}\} \rightarrow +\infty$ and, following the same line of reasoning as that leading to the equivalent formulation of the M-step in (B.9), an optimal solution $\hat{\mathbf{M}}(\hat{\boldsymbol{\theta}}_{ML}) \succ \mathbf{0}$ to Problem (3.14) exists. Now, since $\mathcal{L}_y(\mathbf{M}(\boldsymbol{\theta}^{(h)})|\mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K)$ defines an increasing sequence of values

1. $\mathcal{L}_y(\boldsymbol{\theta}^{(h)}|\mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K)$ converges to a finite value, as it is bounded above by $\mathcal{L}_y(\hat{\mathbf{M}}(\hat{\boldsymbol{\theta}}_{ML})|\mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K)$
2. $\mathbf{M}(\boldsymbol{\theta}^{(h)})$, $h \geq 1$ is a bounded sequence, due to (B.12) and $\|\mathbf{M}\|_F \leq \text{tr}(\mathbf{M})$.

Let now us assess the convergence properties of $\boldsymbol{\theta}^{(h)}$, $h \geq 1$. To proceed further, observe that $\boldsymbol{\theta}^{(h)}$, $h \geq 1$, is a bounded sequence, being $\mathbf{M}(\boldsymbol{\theta})$ a

norm coercive mapping¹. Hence, let $\boldsymbol{\theta}^*$ be a limit point of $\boldsymbol{\theta}^{(h)}$, $h \geq 1$, whose existence is ensured by the boundedness of $\boldsymbol{\theta}^{(h)}$, $h \geq 1$. Owing to the interpretation of the EM procedure as a minorization-maximization optimization technique [54], it can be shown that $\boldsymbol{\theta}^*$ is a global optimal solution to

$$\operatorname{argmax}_{\boldsymbol{\theta}: \mathbf{M}(\boldsymbol{\theta}) \in \mathcal{C}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^*) \quad (\text{B.13})$$

Indeed, denoting by $g(\boldsymbol{\theta}) = \mathcal{L}_y(\boldsymbol{\theta}|\mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K) - Q(\boldsymbol{\theta}, \boldsymbol{\theta})$, it follows that

$$\begin{aligned} \mathcal{L}_y(\boldsymbol{\theta}|\mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K) &= Q(\boldsymbol{\theta}, \boldsymbol{\theta}) + g(\boldsymbol{\theta}), \quad \forall \boldsymbol{\theta} \\ \mathcal{L}_y(\boldsymbol{\theta}|\mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K) &\geq Q(\boldsymbol{\theta}, \boldsymbol{\theta}_1) + g(\boldsymbol{\theta}_1), \quad \forall \boldsymbol{\theta}, \forall \boldsymbol{\theta}_1 \end{aligned}$$

and

$$\begin{aligned} &Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(i_{(h)})}) + g(\boldsymbol{\theta}^{(i_{(h)})}) \\ &\leq Q(\boldsymbol{\theta}^{(i_{(h)+1})}, \boldsymbol{\theta}^{(i_{(h)})}) + g(\boldsymbol{\theta}^{(i_{(h)})}) \\ &\leq \mathcal{L}_y(\boldsymbol{\theta}^{(i_{(h)+1})}|\mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K) \\ &\leq \mathcal{L}_y(\boldsymbol{\theta}^{(i_{(h+1)})}|\mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K) = Q(\boldsymbol{\theta}^{(i_{(h+1)})}, \boldsymbol{\theta}^{(i_{(h+1)})}) + g(\boldsymbol{\theta}^{(i_{(h+1)})}) \end{aligned} \quad (\text{B.14})$$

where $i_{(h)}$ indexes the subsequence extracted by $\boldsymbol{\theta}^{(h)}$ that converges to $\boldsymbol{\theta}^*$, i.e., $\boldsymbol{\theta}^{(i_{(h)})} \rightarrow \boldsymbol{\theta}^*$, as $h \rightarrow \infty$. Now, since $Q(\boldsymbol{\theta}, \boldsymbol{\theta}_1) + g(\boldsymbol{\theta}_1)$ is a continuous function with respect to $(\boldsymbol{\theta}, \boldsymbol{\theta}_1)$,

$$Q(\boldsymbol{\theta}, \boldsymbol{\theta}^*) + g(\boldsymbol{\theta}^*) \leq Q(\boldsymbol{\theta}^*, \boldsymbol{\theta}^*) + g(\boldsymbol{\theta}^*).$$

As a consequence, since $\boldsymbol{\theta}^*$ is a global optimal solution to Problem (B.13) and $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^*) + g(\boldsymbol{\theta}^*)$ is a differentiable function with respect to $\boldsymbol{\theta}$ in a neighborhood of $\boldsymbol{\theta}^*$ (being it the composition of differentiable functions and $\mathbf{M}(\boldsymbol{\theta}^*) \succ \mathbf{0}$), it follows that $\boldsymbol{\theta}^*$ is a B-stationary point to Problem (B.13) [55, 85], namely

$$\lim_{\tau \downarrow 0} \frac{Q(\boldsymbol{\theta}^* + \tau \mathbf{d}, \boldsymbol{\theta}^*) - Q(\boldsymbol{\theta}^*, \boldsymbol{\theta}^*)}{\tau} \leq 0, \quad \forall \mathbf{d} \in \mathcal{T}_{\mathcal{F}}(\boldsymbol{\theta}^*) \quad (\text{B.15})$$

¹A mapping $\mathbf{M}(\cdot) : \mathbf{x} \in \mathbb{R}^{M_2} \rightarrow \mathbf{M}(\mathbf{x}) \in \mathbb{C}^{M_1, M_1}$ is said norm coercive if $\|\mathbf{x}\| \rightarrow \infty$ implies $\|\mathbf{M}(\mathbf{x})\|_F \rightarrow \infty$.

where $\mathcal{T}_{\mathcal{F}}(\boldsymbol{\theta}^*)$ is the Bouligand tangent cone of $\mathcal{F} = \{\boldsymbol{\theta} : \boldsymbol{\theta} \in \mathcal{C}\}$ at $\boldsymbol{\theta}^*$.

Finally, being $\mathcal{L}_y(\boldsymbol{\theta}|\mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K)$ a differentiable function in a neighborhood of $\boldsymbol{\theta}^*$, leveraging Lemma 2 in [12] it holds true

$$\begin{aligned} \lim_{\tau \downarrow 0} \frac{\mathcal{L}_y(\boldsymbol{\theta}^* + \tau \mathbf{d} | \mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K) - \mathcal{L}_y(\boldsymbol{\theta}^* | \mathbf{Y}, \mathbf{A}_1, \dots, \mathbf{A}_K)}{\tau} = \\ \lim_{\tau \downarrow 0} \frac{Q(\boldsymbol{\theta}^* + \tau \mathbf{d}, \boldsymbol{\theta}^*) - Q(\boldsymbol{\theta}^*, \boldsymbol{\theta}^*)}{\tau}, \quad \forall \mathbf{d}, \end{aligned}$$

which implies that $\boldsymbol{\theta}^*$ is a B-stationary point to Problem (3.14).

B.3 Derivation of (3.28) and (3.29)

In the following, closed-form expression of \mathbf{F}_{EM} and \mathbf{F}_{obs} are derived with reference to (3.13) and (3.16). In this respect, note that the (l, m) -th element of $\mathbf{F}_{EM} = \mathbb{E} \left[-\nabla_{\boldsymbol{\theta}} \nabla_{\boldsymbol{\theta}}^T \mathcal{L}_r(\boldsymbol{\theta}) | \mathbf{Y}, \boldsymbol{\theta} \right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{ML}}$ is given by

$$\mathbf{F}_{EM}(l, m) = \mathbb{E} \left[-\frac{\partial^2 \mathcal{L}_r(\boldsymbol{\theta})}{\partial \theta_l \partial \theta_m} | \mathbf{Y}, \boldsymbol{\theta} \right]_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{ML}}, \quad (\text{B.16})$$

where $(l, m) \in \{1, \dots, V\}^2$ and $\mathcal{L}_r(\boldsymbol{\theta})$ is the complete-data log-likelihood given by (3.16). Furthermore,

$$-\frac{\partial^2 \mathcal{L}_r(\boldsymbol{\theta})}{\partial \theta_l \partial \theta_m} = K \frac{\partial^2 \ln(\det(\mathbf{M}(\boldsymbol{\theta})))}{\partial \theta_l \partial \theta_m} + K \frac{\partial^2 \text{tr} \{ \mathbf{M}(\boldsymbol{\theta})^{-1} \mathbf{S} \}}{\partial \theta_l \partial \theta_m}, \quad (\text{B.17})$$

with \mathbf{S} given in (3.17), whereas [105, A.393]

$$\begin{aligned} & \frac{\partial^2 \ln(\det(\mathbf{M}(\boldsymbol{\theta})))}{\partial \theta_l \partial \theta_m} \\ &= \text{tr} \left\{ -\mathbf{M}(\boldsymbol{\theta})^{-1} \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_l} \mathbf{M}(\boldsymbol{\theta})^{-1} \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_m} + \mathbf{M}(\boldsymbol{\theta})^{-1} \frac{\partial^2 \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_l \partial \theta_m} \right\} \end{aligned} \quad (\text{B.18})$$

and [105, A.391], [105, A.392]

$$\begin{aligned}
& \frac{\partial^2 \text{tr} \{ \mathbf{M}(\boldsymbol{\theta})^{-1} \mathbf{S} \}}{\partial \theta_l \partial \theta_m} \\
&= \frac{\partial}{\partial \theta_l} \left[-\text{tr} \left\{ \mathbf{M}(\boldsymbol{\theta})^{-1} \mathbf{S} \mathbf{M}(\boldsymbol{\theta})^{-1} \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_m} \right\} \right] \\
&= -\text{tr} \left\{ \mathbf{M}(\boldsymbol{\theta})^{-1} \mathbf{S} \mathbf{M}(\boldsymbol{\theta})^{-1} \frac{\partial^2 \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_l \partial \theta_m} + \right. \\
&\quad \left. + \frac{\partial \mathbf{M}(\boldsymbol{\theta})^{-1}}{\partial \theta_l} \mathbf{S} \mathbf{M}(\boldsymbol{\theta})^{-1} \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_m} + \right. \\
&\quad \left. + \mathbf{M}(\boldsymbol{\theta})^{-1} \mathbf{S} \frac{\partial \mathbf{M}(\boldsymbol{\theta})^{-1}}{\partial \theta_l} \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_m} \right\} \tag{B.19} \\
&= -\text{tr} \left\{ \mathbf{M}(\boldsymbol{\theta})^{-1} \mathbf{S} \mathbf{M}(\boldsymbol{\theta})^{-1} \frac{\partial^2 \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_l \partial \theta_m} + \right. \\
&\quad \left. - \mathbf{M}(\boldsymbol{\theta})^{-1} \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_l} \mathbf{M}(\boldsymbol{\theta})^{-1} \mathbf{S} \mathbf{M}(\boldsymbol{\theta})^{-1} \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_m} + \right. \\
&\quad \left. - \mathbf{M}(\boldsymbol{\theta})^{-1} \mathbf{S} \mathbf{M}(\boldsymbol{\theta})^{-1} \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_l} \mathbf{M}(\boldsymbol{\theta})^{-1} \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_m} \right\}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
& -\frac{\partial^2 \mathcal{L}_r(\boldsymbol{\theta})}{\partial \theta_l \partial \theta_m} \\
&= K \text{tr} \left\{ -\mathbf{M}(\boldsymbol{\theta})^{-1} \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_l} \mathbf{M}(\boldsymbol{\theta})^{-1} \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_m} + \mathbf{M}(\boldsymbol{\theta})^{-1} \frac{\partial^2 \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_l \partial \theta_m} \right\} \\
&\quad + K \text{tr} \left\{ \left[\mathbf{M}(\boldsymbol{\theta})^{-1} \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_l} \mathbf{M}(\boldsymbol{\theta})^{-1} \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_m} - \mathbf{M}(\boldsymbol{\theta})^{-1} \frac{\partial^2 \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_l \partial \theta_m} + \right. \right. \\
&\quad \left. \left. + \mathbf{M}(\boldsymbol{\theta})^{-1} \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_m} \mathbf{M}(\boldsymbol{\theta})^{-1} \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_l} \right] \mathbf{M}(\boldsymbol{\theta})^{-1} \mathbf{S} \right\}
\end{aligned}$$

Finally, substituting the above equation into (B.16) leads to

$$\mathbf{F}_{EM}(l, m) = K \text{tr} \left\{ -\mathbf{M}(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_l} \mathbf{M}(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_m} + \right.$$

$$\begin{aligned}
& + \mathbf{M}(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial^2 \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_l \partial \theta_m} \Big\} \\
& + K \operatorname{tr} \left\{ \left[\mathbf{M}(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_l} \mathbf{M}(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_m} + \right. \right. \\
& \quad \left. \left. - \mathbf{M}(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial^2 \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_l \partial \theta_m} \right. \right. \quad (B.20) \\
& \left. \left. + \mathbf{M}(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_m} \mathbf{M}(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial \mathbf{M}(\boldsymbol{\theta})}{\partial \theta_l} \right] \right. \\
& \quad \left. \mathbf{M}(\hat{\boldsymbol{\theta}}_{ML})^{-1} \boldsymbol{\Sigma}^* \right\}, \quad (l, m) \in \{1, \dots, V\}^2
\end{aligned}$$

with

$$\boldsymbol{\Sigma}^* = \frac{1}{K} \sum_{i=1}^K \mathbb{E}[\mathbf{r}_i \mathbf{r}_i^\dagger | \mathbf{y}_i, \mathbf{A}_i, \hat{\mathbf{M}}(\hat{\boldsymbol{\theta}}_{ML})]. \quad (B.21)$$

Let us now derive the expression of \mathbf{F}_{obs} . The (l, m) -th element of $\mathbf{F}_{obs} = -\nabla_{\boldsymbol{\theta}} \nabla_{\boldsymbol{\theta}}^T \mathcal{L}_y(\boldsymbol{\theta})|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{ML}}$ is given by

$$\begin{aligned}
\mathbf{F}_{obs}(l, m) &= - \frac{\partial^2 \mathcal{L}_r(\boldsymbol{\theta})}{\partial \theta_l \partial \theta_m} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{ML}} \\
&= \sum_{i=1}^K \frac{\partial^2 \ln(\det(\mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})))}{\partial \theta_l \partial \theta_m} + \frac{\partial^2 \operatorname{tr} \left\{ \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \mathbf{y}_i \mathbf{y}_i^\dagger \right\}}{\partial \theta_l \partial \theta_m}, \\
& \quad (l, m) \in \{1, \dots, V\}^2 \quad (B.22)
\end{aligned}$$

with [105, A.393]

$$\begin{aligned}
& \frac{\partial^2 \ln(\det(\mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})))}{\partial \theta_l \partial \theta_m} \\
&= \operatorname{tr} \left\{ -\mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial \mathbf{M}_i(\boldsymbol{\theta})}{\partial \theta_l} \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial \mathbf{M}_i(\boldsymbol{\theta})}{\partial \theta_m} + \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial^2 \mathbf{M}_i(\boldsymbol{\theta})}{\partial \theta_l \partial \theta_m} \right\} \quad (B.23)
\end{aligned}$$

and [105, A.391], [105, A.392]

$$\begin{aligned}
& \frac{\partial^2 \text{tr} \left\{ \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \mathbf{y}_i \mathbf{y}_i^\dagger \right\}}{\partial \theta_l \partial \theta_m} \\
&= \frac{\partial}{\partial \theta_l} \left[-\text{tr} \left\{ \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \mathbf{y}_i \mathbf{y}_i^\dagger \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial \mathbf{M}_i(\boldsymbol{\theta})}{\partial \theta_m} \right\} \right] \\
&= -\text{tr} \left\{ \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \mathbf{y}_i \mathbf{y}_i^\dagger \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial^2 \mathbf{M}_i(\boldsymbol{\theta})}{\partial \theta_l \partial \theta_m} \right. \\
&\quad - \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial \mathbf{M}_i(\boldsymbol{\theta})}{\partial \theta_l} \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \mathbf{y}_i \mathbf{y}_i^\dagger \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial \mathbf{M}_i(\boldsymbol{\theta})}{\partial \theta_m} \\
&\quad \left. - \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \mathbf{y}_i \mathbf{y}_i^\dagger \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial \mathbf{M}_i(\boldsymbol{\theta})}{\partial \theta_l} \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial \mathbf{M}_i(\boldsymbol{\theta})}{\partial \theta_m} \right\}. \tag{B.24}
\end{aligned}$$

Exploiting the above results,

$$\begin{aligned}
\mathbf{F}_{obs}(l, m) &= \sum_{i=1}^K \text{tr} \left\{ -\mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial \mathbf{M}_i(\boldsymbol{\theta})}{\partial \theta_l} \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial \mathbf{M}_i(\boldsymbol{\theta})}{\partial \theta_m} + \right. \\
&\quad \left. + \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial^2 \mathbf{M}_i(\boldsymbol{\theta})}{\partial \theta_l \partial \theta_m} \right\} \\
&\quad - \text{tr} \left\{ \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \mathbf{y}_i \mathbf{y}_i^\dagger \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial^2 \mathbf{M}_i(\boldsymbol{\theta})}{\partial \theta_l \partial \theta_m} \right. \\
&\quad - \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial \mathbf{M}_i(\boldsymbol{\theta})}{\partial \theta_l} \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \mathbf{y}_i \mathbf{y}_i^\dagger \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial \mathbf{M}_i(\boldsymbol{\theta})}{\partial \theta_m} \\
&\quad \left. - \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \mathbf{y}_i \mathbf{y}_i^\dagger \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial \mathbf{M}_i(\boldsymbol{\theta})}{\partial \theta_l} \mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML})^{-1} \frac{\partial \mathbf{M}_i(\boldsymbol{\theta})}{\partial \theta_m} \right\}, \tag{B.25}
\end{aligned}$$

with $(l, m) \in \{1, \dots, V\}^2$ and $\mathbf{M}_i(\hat{\boldsymbol{\theta}}_{ML}) = \mathbf{A}_i \mathbf{M}(\hat{\boldsymbol{\theta}}_{ML}) \mathbf{A}_i^\dagger$.

Appendix C

Appendix to Chapter 4

C.1 Closed-form expression of the score function (3.18)

Let us first rewrite equation (3.16) as

$$\begin{aligned} \mathcal{L}_{\mathbf{r}}(\alpha, \boldsymbol{\theta}, H_1) &= -(K+1) [N \ln(\pi) + \ln(\det(\mathbf{M}(\boldsymbol{\theta})))] \\ &\quad - \text{tr} \left\{ \mathbf{M}(\boldsymbol{\theta})^{-1} \left[\mathbf{r}\mathbf{r}^\dagger + |\alpha|^2 \mathbf{p}\mathbf{p}^\dagger - 2 \text{Re}\{\{\alpha \mathbf{p}\mathbf{r}^\dagger\} + \mathbf{S}\} \right] \right\}. \end{aligned} \quad (\text{C.1})$$

It follows that

$$\begin{aligned} Q\left(\alpha, \boldsymbol{\theta} | \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)}\right) &= -(K+1) [N \ln(\pi) + \ln(\det(\mathbf{M}(\boldsymbol{\theta})))] \\ &\quad - \text{tr} \left\{ \mathbf{M}(\boldsymbol{\theta})^{-1} \left[\mathbb{E}[\mathbf{r}\mathbf{r}^\dagger | \mathbf{z}, \mathbf{A}, \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)}, H_1] \right. \right. \\ &\quad \left. \left. + |\alpha|^2 \mathbf{p}\mathbf{p}^\dagger - 2 \text{Re}\{\{\alpha \mathbf{p}\boldsymbol{\mu}^{(h-1)\dagger}\} + \mathbf{X}^{(h-1)}\} \right] \right\} \\ &= -(K+1) [N \ln(\pi) + \ln(\det(\mathbf{M}(\boldsymbol{\theta})))] - \text{tr} \left\{ \mathbf{M}(\boldsymbol{\theta})^{-1} \right. \\ &\quad \left. \left[(\boldsymbol{\mu}^{(h-1)} - \alpha \mathbf{p})(\boldsymbol{\mu}^{(h-1)} - \alpha \mathbf{p})^\dagger + \boldsymbol{\Sigma}^{(h-1)} \right] \right\} \end{aligned} \quad (\text{C.2})$$

where

$$\mathbf{X}^{(h-1)} = \sum_{i=1}^K \mathbb{E}[\mathbf{r}_i \mathbf{r}_i^\dagger | \mathbf{z}_i, \mathbf{A}_i, \boldsymbol{\theta}^{(h-1)}] \quad (\text{C.3})$$

and

$$\boldsymbol{\Sigma}^{(h-1)} = \mathbf{X}^{(h-1)} - \boldsymbol{\mu}^{(h-1)} \boldsymbol{\mu}^{(h-1)\dagger} + \mathbb{E}[\mathbf{r}\mathbf{r}^\dagger | \mathbf{z}, \mathbf{A}, \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)}, H_1]. \quad (\text{C.4})$$

To proceed further, let us denote

$$\mathbf{B} = [\mathbf{A}^\text{T} \bar{\mathbf{A}}^\text{T}]^\text{T} \quad (\text{C.5})$$

where $\bar{\mathbf{A}}$ is the $N - p \times N$ selection matrix complementary to \mathbf{A} (obtained removing from \mathbf{I} the p rows not removed in the definition of \mathbf{A}) and $\mathbf{B}^\text{T} \mathbf{B} = \mathbf{I}$. Hence [56],

$$\begin{aligned} \boldsymbol{\mu}^{(h-1)} &= \mathbf{B}^\text{T} \mathbb{E}[\mathbf{B}\mathbf{r} | \mathbf{z}, \mathbf{A}, \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)}, H_1] \\ &= \mathbf{B}^\text{T} [\mathbf{z}^\text{T}, \boldsymbol{\zeta}^{(h-1)\text{T}}]^\text{T} \end{aligned} \quad (\text{C.6})$$

where

$$\begin{aligned} \boldsymbol{\zeta}^{(h-1)} &= \mathbb{E}[\bar{\mathbf{A}}\mathbf{r} | \mathbf{z}, \mathbf{A}, \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)}, H_1] \\ &= \bar{\mathbf{A}}\mathbf{M}(\boldsymbol{\theta}^{(h-1)})\mathbf{A}^\text{T}(\mathbf{A}\mathbf{M}(\boldsymbol{\theta}^{(h-1)})\mathbf{A}^\text{T})^{-1}(\mathbf{z} - \alpha^{(h-1)}\mathbf{A}\mathbf{p}) + \alpha^{(h-1)}\bar{\mathbf{A}}\mathbf{p}. \end{aligned} \quad (\text{C.7})$$

Besides,

$$\begin{aligned} &\mathbb{E}[\mathbf{r}\mathbf{r}^\dagger | \mathbf{z}, \mathbf{A}, \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)}, H_1] \\ &= \mathbf{B}^\text{T} \mathbb{E}[\mathbf{B}\mathbf{r}\mathbf{r}^\dagger \mathbf{B}^\text{T} | \mathbf{z}, \mathbf{A}, \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)}, H_1] \mathbf{B}. \end{aligned} \quad (\text{C.8})$$

As to the expectation term,

$$\begin{aligned} &\mathbb{E}[\mathbf{B}\mathbf{r}\mathbf{r}^\dagger \mathbf{B}^\text{T} | \mathbf{z}, \mathbf{A}, \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)}, H_1] \\ &= \begin{bmatrix} \mathbf{z}\mathbf{z}^\dagger & \mathbf{z}\boldsymbol{\zeta}^{(h-1)\dagger} \\ \boldsymbol{\zeta}^{(h-1)}\mathbf{z}^\dagger & \mathbb{E}[\bar{\mathbf{A}}\mathbf{r}\mathbf{r}^\dagger \bar{\mathbf{A}}^\text{T} | \mathbf{z}, \mathbf{A}, \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)}, H_1] \end{bmatrix} \end{aligned} \quad (\text{C.9})$$

where

$$\mathbb{E}[\bar{\mathbf{A}}\mathbf{r}\mathbf{r}^\dagger \bar{\mathbf{A}}^\text{T} | \mathbf{z}, \mathbf{A}, \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)}, H_1] = \mathbf{G} + \boldsymbol{\zeta}^{(h-1)} \boldsymbol{\zeta}^{(h-1)\dagger} \quad (\text{C.10})$$

with

$$\begin{aligned} \mathbf{G} &= \bar{\mathbf{A}}\mathbf{M}(\boldsymbol{\theta}^{(h-1)})\bar{\mathbf{A}}^T - \bar{\mathbf{A}}\mathbf{M}(\boldsymbol{\theta}^{(h-1)})\mathbf{A}^T \\ &\quad (\mathbf{A}\mathbf{M}(\boldsymbol{\theta}^{(h-1)})\mathbf{A}^T)^{-1}\mathbf{A}\mathbf{M}(\boldsymbol{\theta}^{(h-1)})\bar{\mathbf{A}}^T. \end{aligned} \quad (\text{C.11})$$

Therefore,

$$\begin{aligned} &\mathbf{B}^T\mathbb{E}[\mathbf{B}\mathbf{r}\mathbf{r}^\dagger\mathbf{B}^T|\mathbf{z}, \mathbf{A}, \alpha^{(h-1)}, \boldsymbol{\theta}^{(h-1)}, H_1]\mathbf{B} \\ &= \mathbf{A}^T\mathbf{z}\mathbf{z}^\dagger\mathbf{A} + \bar{\mathbf{A}}^T\boldsymbol{\zeta}^{(h-1)}\mathbf{z}^\dagger\mathbf{A} + \mathbf{A}^T\mathbf{z}\boldsymbol{\zeta}^{(h-1)\dagger}\bar{\mathbf{A}} \\ &\quad + \bar{\mathbf{A}}^T(\mathbf{G} + \boldsymbol{\zeta}^{(h-1)}\boldsymbol{\zeta}^{(h-1)\dagger})\bar{\mathbf{A}} \\ &= (\mathbf{A}^T\mathbf{z} + \bar{\mathbf{A}}^T\boldsymbol{\zeta}^{(h-1)})(\mathbf{A}^T\mathbf{z} + \bar{\mathbf{A}}^T\boldsymbol{\zeta}^{(h-1)})^\dagger + \bar{\mathbf{A}}^T\mathbf{G}\bar{\mathbf{A}}. \end{aligned} \quad (\text{C.12})$$

Finally, denoting by $\bar{\mathbf{A}}_i$ the $N - p_i \times N$ selection matrix defined similarly to $\bar{\mathbf{A}}$, equation (C.3) can be recast as

$$\mathbf{X}^{(h-1)} = \sum_{i=1}^K \mathbf{C}_i^{(h-1)} \quad (\text{C.13})$$

where

$$\mathbf{C}_i^{(h-1)} = (\mathbf{A}_i^T + \bar{\mathbf{A}}_i^T\boldsymbol{\Gamma}_i)\mathbf{z}_i\mathbf{z}_i^\dagger(\mathbf{A}_i^T + \bar{\mathbf{A}}_i^T\boldsymbol{\Gamma}_i)^\dagger + \bar{\mathbf{A}}_i^T\mathbf{G}_i\bar{\mathbf{A}}_i \quad (\text{C.14})$$

with

$$\boldsymbol{\Gamma}_i = \bar{\mathbf{A}}_i\mathbf{M}(\boldsymbol{\theta}^{(h-1)})\mathbf{A}_i^T(\mathbf{A}_i\mathbf{M}(\boldsymbol{\theta}^{(h-1)})\mathbf{A}_i^T)^{-1} \quad (\text{C.15})$$

and

$$\mathbf{G}_i = \bar{\mathbf{A}}_i\mathbf{M}(\boldsymbol{\theta}^{(h-1)})\bar{\mathbf{A}}_i^T - \boldsymbol{\Gamma}_i\mathbf{A}_i\mathbf{M}(\boldsymbol{\theta}^{(h-1)})\bar{\mathbf{A}}_i^T. \quad (\text{C.16})$$

C.2 Detailed expressions for LI-based detectors

In order to formally define the adopted LI procedure, let us observe that given two vectors (x_1, y_1) and (x_2, y_2) belonging to \mathbb{R}^2 , the equation

of the line connecting these two points is

$$y = y_1 + \frac{(y_2 - y_1)}{x_2 - x_1}(x - x_1). \quad (\text{C.17})$$

Hence, the value at a point x^* can be predicted via interpolation as

$$y^* = y_1 + \frac{(y_2 - y_1)}{x_2 - x_1}(x^* - x_1). \quad (\text{C.18})$$

For the case at hand, based on the location \hat{x} of the missing element within the snapshot, the two points (x_1, y_1) and (x_2, y_2) with closest spatial locations to \hat{x} are used to interpolate the value at point \hat{x} . For the two special cases when \hat{x} is located at the beginning (the end) of the array, the sample points are chosen as the first two (last two) elements of the observed-data snapshot, respectively. Moreover, due to complex-valued nature of the data, each missing element is estimated by linearly interpolating real and imaginary components separately. Then, denoting by $\check{\mathbf{z}}_i \in \mathbb{C}^N, i = 1, \dots, K$, the interpolated secondary data snapshot, the following detectors, based on the [GLR](#) and [AMF](#) criteria, respectively, can be implemented, i.e.,

$$\tau_{\text{GLRT-LI}} = \frac{1}{1 + \mathbf{z}^\dagger (\mathbf{A}\check{\mathbf{S}}\mathbf{A}^\text{T})^{-1} \mathbf{z}} \frac{\left| \mathbf{z}^\dagger (\mathbf{A}\check{\mathbf{S}}\mathbf{A}^\text{T})^{-1} \mathbf{A}\mathbf{p} \right|^2}{\mathbf{p}^\dagger \mathbf{A}^\text{T} (\mathbf{A}\check{\mathbf{S}}\mathbf{A}^\text{T})^{-1} \mathbf{A}\mathbf{p}} \quad (\text{C.19})$$

and

$$\tau_{\text{AMF-LI}} = \frac{\left| \mathbf{z}^\dagger (\mathbf{A}\check{\mathbf{S}}\mathbf{A}^\text{T})^{-1} \mathbf{A}\mathbf{p} \right|^2}{\mathbf{p}^\dagger \mathbf{A}^\text{T} (\mathbf{A}\check{\mathbf{S}}\mathbf{A}^\text{T})^{-1} \mathbf{A}\mathbf{p}} \quad (\text{C.20})$$

where $\check{\mathbf{S}} = \sum_{i=1}^K \check{\mathbf{z}}_i \check{\mathbf{z}}_i^\dagger$.

Appendix **D**

Appendix to Chapter 5

D.1 Expressions of the s_u and s_δ

According to (5.7), the first derivatives of \mathbf{s} w.r.t. u and δ , evaluated respectively at u_0 and δ_0 , can be calculated as

$$\begin{aligned} \mathbf{s}_u = \left. \frac{\partial \mathbf{s}}{\partial u} \right|_{(u_0, \delta_0)} &= \left. \frac{\partial \mathbf{b}(u)}{\partial u} \right|_{u_0} \otimes [\mathbf{c}(u_0) \odot \mathbf{a}(\delta_0)] \\ &+ \mathbf{b}(u_0) \otimes \left[\left. \frac{\partial \mathbf{c}(u)}{\partial u} \right|_{u_0} \odot \mathbf{a}(\delta_0) \right], \end{aligned} \quad (\text{D.1})$$

and

$$\mathbf{s}_\delta = \left. \frac{\partial \mathbf{s}}{\partial \delta} \right|_{(u_0, \delta_0)} = \mathbf{b}(u_0) \otimes \left[\mathbf{c}(u_0) \odot \left. \frac{\partial \mathbf{a}(\delta)}{\partial \delta} \right|_{\delta_0} \right], \quad (\text{D.2})$$

where

$$\left. \frac{\partial \mathbf{b}(u)}{\partial u} \right|_{u_0} = j2\pi \frac{d}{\lambda_0} \mathbf{E}_T \mathbf{b}(u_0), \quad (\text{D.3})$$

$$\left. \frac{\partial \mathbf{c}(u)}{\partial u} \right|_{u_0} = \left. \frac{\partial (\mathbf{R}^T \mathbf{d}(u))}{\partial u} \right|_{u_0} = \mathbf{R}^T \left. \frac{\partial \mathbf{d}(u)}{\partial u} \right|_{u_0} = j2\pi \frac{d}{\lambda_0} \mathbf{R}^T \mathbf{E}_R \mathbf{d}(u_0), \quad (\text{D.4})$$

$$\left. \frac{\partial \mathbf{a}(\delta)}{\partial \delta} \right|_{\delta_0} = j\pi \mathbf{E}_R \mathbf{a}(\delta_0), \quad (\text{D.5})$$

with $\mathbf{E}_R = \mathbf{diag}([0, 1, \dots, M-1]^T)$ and $\mathbf{E}_T = \mathbf{diag}([0, 1, \dots, N-1]^T)$.

D.2 Computations of D_u and D_δ

Let us consider

$$\mathbf{V} = \mathbf{F}_{12}\mathbf{F}_{22}^{-1}\mathbf{F}_{21} = \|\boldsymbol{\xi}\|^{-2} \begin{bmatrix} V_{11} & V_{12} \\ V_{12} & V_{22} \end{bmatrix} \in \mathbb{R}^{2 \times 2}, \quad (\text{D.6})$$

The entries of \mathbf{V} are given by

$$V_{11} = \text{Re}^2 \left\{ \boldsymbol{\xi}_u^\dagger \boldsymbol{\xi} \beta_1^* \right\} + \text{Im}^2 \left\{ \boldsymbol{\xi}_u^\dagger \boldsymbol{\xi} \beta_1^* \right\} = |\beta_1|^2 \left| \boldsymbol{\xi}_u^\dagger \boldsymbol{\xi} \right|^2, \quad (\text{D.7a})$$

$$\begin{aligned} V_{12} &= \text{Re} \left\{ \boldsymbol{\xi}_u^\dagger \boldsymbol{\xi} \beta_1^* \right\} \text{Re} \left\{ \boldsymbol{\xi}^\dagger \boldsymbol{\xi}_\delta \beta_1 \right\} + \text{Im} \left\{ \boldsymbol{\xi}_u^\dagger \boldsymbol{\xi} \beta_1^* \right\} \text{Im} \left\{ \boldsymbol{\xi}^\dagger \boldsymbol{\xi}_\delta \beta_1 \right\} \\ &= \text{Re} \left\{ \boldsymbol{\xi}_u^\dagger \boldsymbol{\xi} \beta_1^* \boldsymbol{\xi}^\dagger \boldsymbol{\xi}_\delta \beta_1 \right\} = |\beta_1|^2 \text{Re} \left\{ \boldsymbol{\xi}_u^\dagger \boldsymbol{\xi} \boldsymbol{\xi}^\dagger \boldsymbol{\xi}_\delta \right\}, \end{aligned} \quad (\text{D.7b})$$

$$V_{22} = \text{Re}^2 \left\{ \boldsymbol{\xi}_\delta^\dagger \boldsymbol{\xi} \beta_1^* \right\} + \text{Im}^2 \left\{ \boldsymbol{\xi}_\delta^\dagger \boldsymbol{\xi} \beta_1^* \right\} = |\beta_1|^2 \left| \boldsymbol{\xi}_\delta^\dagger \boldsymbol{\xi} \right|^2. \quad (\text{D.7c})$$

Hence, \mathbf{G}_1 is derived as

$$\mathbf{G}_1 = \mathbf{F}_{11} - \mathbf{V} = \begin{bmatrix} G_{11} & G_{12} \\ G_{12} & G_{22} \end{bmatrix}, \quad (\text{D.8})$$

with

$$G_{11} = |\beta_1|^2 \|\boldsymbol{\xi}_u\|^2 - \frac{|\beta_1|^2}{\|\boldsymbol{\xi}\|^2} \left| \boldsymbol{\xi}_u^\dagger \boldsymbol{\xi} \right|^2, \quad (\text{D.9})$$

$$G_{12} = |\beta_1|^2 \text{Re} \left\{ \boldsymbol{\xi}_u^\dagger \boldsymbol{\xi}_\delta \right\} - \frac{|\beta_1|^2}{\|\boldsymbol{\xi}\|^2} \text{Re} \left\{ \boldsymbol{\xi}_u^\dagger \boldsymbol{\xi} \boldsymbol{\xi}^\dagger \boldsymbol{\xi}_\delta \right\}, \quad (\text{D.10})$$

and

$$G_{22} = |\beta_1|^2 \|\boldsymbol{\xi}_\delta\|^2 - \frac{|\beta_1|^2}{\|\boldsymbol{\xi}\|^2} \left| \boldsymbol{\xi}_\delta^\dagger \boldsymbol{\xi} \right|^2. \quad (\text{D.11})$$

As a result, the CRLB for u and δ is the diagonal elements of $\frac{1}{2}\mathbf{G}_1^{-1}$, which have been provided in (5.45) and (5.46), respectively, where

$$\begin{aligned} \mathbf{G}_1^{-1} &= \frac{1}{\det(\mathbf{G}_1)} \begin{bmatrix} G_{22} & -G_{12} \\ -G_{21} & G_{11} \end{bmatrix} \\ &= \frac{|\beta_1|^2}{\det(\mathbf{G}_1)} \begin{bmatrix} \|\boldsymbol{\xi}_\delta\|^2 - \frac{|\boldsymbol{\xi}_\delta^\dagger \boldsymbol{\xi}|^2}{\|\boldsymbol{\xi}\|^2} & -\operatorname{Re}\{\boldsymbol{\xi}_u^\dagger \boldsymbol{\xi}_\delta\} + \frac{\operatorname{Re}\{\boldsymbol{\xi}_u^\dagger \boldsymbol{\xi} \boldsymbol{\xi}^\dagger \boldsymbol{\xi}_\delta\}}{\|\boldsymbol{\xi}\|^2} \\ -\operatorname{Re}\{\boldsymbol{\xi}_u^\dagger \boldsymbol{\xi}_\delta\} + \frac{\operatorname{Re}\{\boldsymbol{\xi}_u^\dagger \boldsymbol{\xi} \boldsymbol{\xi}^\dagger \boldsymbol{\xi}_\delta\}}{\|\boldsymbol{\xi}\|^2} & \|\boldsymbol{\xi}_u\|^2 - \frac{|\boldsymbol{\xi}_u^\dagger \boldsymbol{\xi}|^2}{\|\boldsymbol{\xi}\|^2} \end{bmatrix} \end{aligned} \quad (\text{D.12})$$

with $\det(\mathbf{G}_1) = G_{11}G_{22} - G_{12}G_{21}$.

Appendix **E**

Appendix to Chapter 6

E.1 Expressions for H_u and H_δ

Denoting by $\mathbf{s}_u \in \mathbb{C}^{MN}$ and $\mathbf{s}_\delta \in \mathbb{C}^{MN}$ the derivative functions of $\mathbf{s}(u, \delta)$ w.r.t. u and δ , respectively, they are given by [67]

$$\begin{aligned} \mathbf{s}_u = \frac{\partial \mathbf{s}(u, \delta)}{\partial u} &= \frac{\partial \mathbf{b}(u)}{\partial u} \otimes [\mathbf{c}(u) \odot \mathbf{a}(\delta)] \\ &+ \mathbf{b}(u) \otimes \left[\frac{\partial \mathbf{c}(u)}{\partial u} \odot \mathbf{a}(\delta) \right] \end{aligned} \quad (\text{E.1})$$

and

$$\mathbf{s}_\delta = \frac{\partial \mathbf{s}(u, \delta)}{\partial \delta} = \mathbf{b}(u) \otimes \left[\mathbf{c}(u) \odot \frac{\partial \mathbf{a}(\delta)}{\partial \delta} \right], \quad (\text{E.2})$$

where

- $\frac{\partial \mathbf{b}(u)}{\partial u} = j2\pi \frac{d}{\lambda_0} \mathbf{E}_T \mathbf{b}(u)$;
- $\frac{\partial \mathbf{c}(u)}{\partial u} = j2\pi \frac{d}{\lambda_0} \mathbf{R}^T \mathbf{E}_R \mathbf{d}(u)$;
- $\frac{\partial \mathbf{a}(\delta)}{\partial \delta} = j\pi \mathbf{E}_R \mathbf{a}(\delta)$.

with

- $\mathbf{E}_T = \text{diag}([0, 1, \dots, N-1]^T) \in \mathbb{R}^{N \times N}$;

- $\mathbf{E}_R = \text{diag}([0, 1, \dots, M-1]^T) \in \mathbb{R}^{M \times M}$.

Therefore,

$$\mathbf{H}_u = \begin{bmatrix} \mathbf{s}_u & \mathbf{0} \\ \mathbf{0} & \mathbf{s}_u \end{bmatrix} \quad (\text{E.3})$$

and

$$\mathbf{H}_\delta = \begin{bmatrix} \mathbf{s}_\delta & \mathbf{0} \\ \mathbf{0} & \mathbf{s}_\delta \end{bmatrix}, \quad (\text{E.4})$$

E.2 Proof of Proposition 3

Proof. The existence of the global minimizer of (6.30) is guaranteed by the Weierstrass theorem [19], being the objective function (6.30) continuous and the feasible set \mathcal{C} non-empty and compact. Besides, the stationary points of the objective function in (6.30) are obtained by nulling the objective gradient, namely by solving the equation

$$\nabla_{\Delta\theta} \left[(\hat{\mathbf{r}} - \hat{\mathbf{H}}\Delta\theta)^\dagger (\hat{\mathbf{r}} - \hat{\mathbf{H}}\Delta\theta) \right] = 0. \quad (\text{E.5})$$

After some algebra, equation (E.5) can be expressed as

$$\Re\{\hat{\mathbf{H}}^\dagger \hat{\mathbf{r}}\} - \Re\{\hat{\mathbf{H}}^\dagger \hat{\mathbf{H}}\} \Delta\theta = 0, \quad (\text{E.6})$$

meaning that the unique stationary point is obtained as

$$\Delta\theta_1 = \left[\Re\{\hat{\mathbf{H}}^\dagger \hat{\mathbf{H}}\} \right]^{-1} \Re\{\hat{\mathbf{H}}^\dagger \hat{\mathbf{r}}\}. \quad (\text{E.7})$$

Remarkably, if it is feasible, i.e., $\Delta\theta_1 \in \mathcal{C}$, it is the optimal solution to the optimization problem (6.30). Otherwise, an optimal solution can be found restricting the objective function to the boundaries of \mathcal{C} and determining, for each edge, the corresponding minimizer.

Case C. Left and right edges ($\Delta u = \pm\kappa$).

Denoting by $\hat{\mathbf{r}}_{\pm\kappa} = \hat{\mathbf{r}} - \hat{\mathbf{h}}_u(\pm\kappa)$, the optimization problem (6.30), restricted to $\Delta u = \pm\kappa$, boils down to

$$\min_{|\Delta\delta| \leq \rho} (\hat{\mathbf{r}}_{\pm\kappa} - \hat{\mathbf{H}}_\delta \Delta\delta)^\dagger (\hat{\mathbf{r}}_{\pm\kappa} - \hat{\mathbf{H}}_\delta \Delta\delta). \quad (\text{E.8})$$

The objective function of (E.8) can be written as

$$a' \Delta\delta^2 - 2b'_{\pm\kappa} \Delta\delta + c'_{\pm\kappa}, \quad (\text{E.9})$$

where

$$a' = \|\hat{\mathbf{H}}_\delta\|^2, \quad (\text{E.10a})$$

$$b'_{\pm\kappa} = \Re\{\hat{\mathbf{r}}_{\pm\kappa}^\dagger \hat{\mathbf{h}}_\delta\}, \quad (\text{E.10b})$$

$$c'_{\pm\kappa} = \|\hat{\mathbf{r}}_{\pm\kappa}\|^2. \quad (\text{E.10c})$$

Since $a' > 0$, (E.8) is a convex optimization problem, whose optimal solution is given by

$$\Delta\delta_\pm^* = \min(\rho, \max(b'_{\pm\kappa}/a', -\rho)). \quad (\text{E.11})$$

Therefore, the two candidate optimal solutions associated with the right and left edges are given by

$$\Delta\boldsymbol{\theta}_2 = [\kappa, \Delta\delta_+^*] \quad (\text{E.12})$$

and

$$\Delta\boldsymbol{\theta}_3 = [-\kappa, \Delta\delta_-^*]. \quad (\text{E.13})$$

Case D. Upper and lower edges ($\Delta\delta = \pm\rho$).

Similarly to the **Case A**, denoting by $\hat{\mathbf{r}}_{\pm\rho} = \hat{\mathbf{r}} - \hat{\mathbf{h}}_\delta(\pm\rho)$, the candidate optimal solutions associated with the superior and inferior edges are given by

$$\Delta\boldsymbol{\theta}_4 = [\Delta u_+^*, \rho] \quad (\text{E.14})$$

and

$$\Delta\boldsymbol{\theta}_5 = [\Delta u_-^*, -\rho], \quad (\text{E.15})$$

where

$$\Delta u_\pm^* = \min(\kappa, \max(b''_{\pm\rho}/a'', -\kappa)) \quad (\text{E.16})$$

with

$$a'' = \|\hat{\mathbf{h}}_u\|^2, \quad (\text{E.17a})$$

$$b''_{\pm\rho} = \Re\{\hat{\mathbf{r}}_{\pm\rho}^\dagger \hat{\mathbf{h}}_u\}. \quad (\text{E.17b})$$

As a consequence, an optimal solution to the optimization problem (6.30)

is given by

$$\Delta\boldsymbol{\theta}^{(h)} = \arg \min_{\Delta\boldsymbol{\theta} \in \{\Delta\boldsymbol{\theta}_i\}_{i=1}^5} (\hat{\mathbf{r}} - \hat{\mathbf{H}}\Delta\boldsymbol{\theta})^\dagger (\hat{\mathbf{r}} - \hat{\mathbf{H}}\Delta\boldsymbol{\theta}). \quad (\text{E.18})$$

□

E.3 Computation of $g_\Delta(u, \delta)$

Denoting by $\bar{\mathbf{r}} = \mathbf{S}^{-\frac{1}{2}}\mathbf{r} \in \mathbb{C}^{2MN}$ and $\check{\mathbf{H}} = \mathbf{S}^{-\frac{1}{2}}\mathbf{H} \in \mathbb{C}^{2MN \times 2}$, $g(u, \delta)$ can be rewritten as

$$g(u, \delta) = \bar{\mathbf{r}}^\dagger \check{\mathbf{H}} \left[\check{\mathbf{H}}^\dagger \check{\mathbf{H}} \right]^{-1} \check{\mathbf{H}}^\dagger \bar{\mathbf{r}} = \bar{\mathbf{r}}^\dagger \bar{\mathbf{P}}_{\check{\mathbf{H}}} \bar{\mathbf{r}} \quad (\text{E.19})$$

where $\bar{\mathbf{P}}_{\check{\mathbf{H}}} = \check{\mathbf{H}} \left[\check{\mathbf{H}}^\dagger \check{\mathbf{H}} \right]^{-1} \check{\mathbf{H}}^\dagger \in \mathbb{C}^{2MN \times 2MN}$.

Therefore, the first derivatives of $g(u, \delta)$ w.r.t. u and δ , are given by [Eq. A.394, [105]]

$$\frac{\partial g(u, \delta)}{\partial x} = \bar{\mathbf{r}}^\dagger \frac{\partial \bar{\mathbf{P}}_{\check{\mathbf{H}}}}{\partial x} \bar{\mathbf{r}}, \quad x \in \{u, \delta\}, \quad (\text{E.20})$$

where

- $\frac{\partial \bar{\mathbf{P}}_{\check{\mathbf{H}}}}{\partial x} = \bar{\mathbf{P}}_{\check{\mathbf{H}}}^\perp \check{\mathbf{H}}_x \check{\mathbf{H}}^\dagger + \left(\bar{\mathbf{P}}_{\check{\mathbf{H}}}^\perp \check{\mathbf{H}}_x \check{\mathbf{H}}^\dagger \right)^\dagger$ with $\check{\mathbf{H}}_x = \mathbf{S}^{-1/2} \mathbf{H}_x \in \mathbb{C}^{2MN \times 2}$, $x \in u, \delta$;
- $\bar{\mathbf{P}}_{\check{\mathbf{H}}}^\perp = \mathbf{I} - \bar{\mathbf{P}}_{\check{\mathbf{H}}} \in \mathbb{C}^{2MN \times 2MN}$.

As a consequence, the gradient of $g(u, \delta)$ is given by

$$\mathbf{g}_\Delta(u, \delta) = \left[\frac{\partial g(u, \delta)}{\partial u}, \frac{\partial g(u, \delta)}{\partial \delta} \right]^\text{T}. \quad (\text{E.21})$$

E.4 Proof of $g(u, \delta) \in C_L^{1,1}$

Before proceeding with the proof, let us first introduce the following lemma [18], which provides a sufficient condition for a function to be in $C_L^{1,1}$.

Lemma E.4.1. *Let $f(\mathbf{x})$ belongs to C^2 with $\mathbf{f}_{\Delta\Delta}(\mathbf{x})$ its Hessian matrix; if there exists $L > 0$ such that $\|\mathbf{f}_{\Delta\Delta}(\mathbf{x})\| \leq L$ holds true for any $\mathbf{x} \in \mathbb{R}^n$, then $f(\mathbf{x}) \in C_L^{1,1}$. \square*

Therefore, in order to exploit the above result, in the following, the function $g(u, \delta)$ is proved to fulfill the conditions required by *Lemma E.4.1*. To this end, let us report again, for easy of reference, the detailed expression of $\mathbf{s}(u, \delta)$, i.e.,

$$\mathbf{s}(u, \delta) = \mathbf{b}(u) \otimes [\mathbf{c}(u) \odot \mathbf{a}(\delta)] \in \mathbb{C}^{MN} \quad (\text{E.22})$$

with

$$\mathbf{b}(u) = \left[1, e^{j2\pi \frac{d}{\lambda_0} u}, \dots, e^{j2\pi \frac{d}{\lambda_0} (N-1)u} \right]^T \in \mathbb{C}^N, \quad (\text{E.23a})$$

$$\mathbf{c}(u) = \mathbf{R}^T \mathbf{d}(u) \in \mathbb{C}^M, \quad (\text{E.23b})$$

$$\mathbf{d}(u) = \left[1, e^{j2\pi \frac{d}{\lambda_0} u}, \dots, e^{j2\pi \frac{d}{\lambda_0} (M-1)u} \right]^T \in \mathbb{C}^M, \quad (\text{E.23c})$$

$$\mathbf{a}(\delta) = \left[1, e^{j\pi\delta}, \dots, e^{j\pi(M-1)\delta} \right]^T \in \mathbb{C}^M. \quad (\text{E.23d})$$

From equations (E.23), it is straightforward to see that $\mathbf{b}(u)$, $\mathbf{c}(u)$, $\mathbf{d}(u)$, and $\mathbf{a}(\delta)$ belong to C^2 and hence $\mathbf{s}(u, \delta)$ belongs to C^2 . Therefore, assuming $\mathbf{S} \succ 0$, which occurs almost surely being $K \geq 2MN$,

$$g(u, \delta) = \mathbf{r}^\dagger \mathbf{S}^{-1} \mathbf{H} \left[\mathbf{H}^\dagger \mathbf{S}^{-1} \mathbf{H} \right]^{-1} \mathbf{H}^\dagger \mathbf{S}^{-1} \mathbf{r} \quad (\text{E.24})$$

belongs to C^2 since

$$\mathbf{H} = \mathbf{H}(u, \delta) = \begin{bmatrix} \mathbf{s}(u, \delta) & \mathbf{0} \\ \mathbf{0} & \mathbf{s}(u, \delta) \end{bmatrix} \in \mathbb{C}^{2MN \times 2} \quad (\text{E.25})$$

belongs to C^2 and it is full rank.

Then, it can be verified that $\|\mathbf{g}_{\Delta\Delta}(u, \delta)\|$ is continuous and periodic, with

$$\mathbf{g}_{\Delta\Delta} = \begin{bmatrix} g_{uu}(u, \delta) & g_{u\delta}(u, \delta) \\ g_{\delta u}(u, \delta) & g_{\delta\delta}(u, \delta) \end{bmatrix},$$

since $\|\mathbf{A}\|$ is a continuous function w.r.t. $\mathbf{A} \in \mathbb{H}^N$ and the second derivatives of $g(u, \delta)$, i.e., $g_{xy}(u, \delta)$, $x, y \in \{u, \delta\}$ are continuous and periodic functions with period $T_u = \lambda_0/d$ and $T_\delta = 2$ along u and δ directions, respectively (see Appendix E.5). Therefore, invoking the Weierstrass theorem [19],

$$\|\mathbf{g}_{\Delta\Delta}(u, \delta)\| \leq \max_{u \in [0, T_u], \delta \in [0, T_\delta]} \|\mathbf{g}_{\Delta\Delta}(u, \delta)\| = L < \infty. \quad (\text{E.26})$$

E.5 Proof that the second derivatives of $g(u, \delta)$ are periodic functions

This Appendix is composed of two parts: first, $\mathbf{s}(u, \delta)$ and $g(u, \delta)$ are proved to be periodic functions. Then, the periodicity of the first and second order derivatives of $g(u, \delta)$ is analyzed.

From equations (E.23), it is easy to verify that $\mathbf{s}(u, \delta)$ is a periodic function with period T_u and T_δ along u and δ directions, respectively, being

- $\mathbf{b}(u)$ periodic with period T_u ;
- $\mathbf{d}(u)$ and $\mathbf{c}(u)$ periodic with period T_u ;
- $\mathbf{a}(\delta)$ periodic with period T_δ .

As a consequence, $g(u, \delta)$ shares the same periodicity as $\mathbf{s}(u, \delta)$.

Now, the periodicity of the first and second order derivatives can be established resorting to the following lemma.

Lemma E.5.1. Let $f(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}$ be a continuous, derivable, and periodic function with period T_x and T_y along the x and y directions, respectively. Then, $f_x(x, y) = \frac{\partial f(x, y)}{\partial x}$ and $f_y(x, y) = \frac{\partial f(x, y)}{\partial y}$ are periodic functions with the same periodicities as $f(x, y)$.

Proof. Denoting by $x_1 = x + N_1T_x$ and $y_1 = y + N_2T_y$, with $N_1, N_2 \in \mathbb{N}$,

$$\begin{aligned} f_x(x_1, y_1) &= f_x(x + N_1T_x, y + N_2T_y) \\ &= \lim_{h \rightarrow 0} \frac{f(x + N_1T_x + h, y + N_2T_y) - f(x + N_1T_x, y + N_2T_y)}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(x + h, y) - f(x, y)}{h} = f_x(x, y). \end{aligned}$$

Therefore, $f_x(x, y)$ is periodic as $f(x, y)$. Along the same line, it is straight-forward to prove that $f_y(x, y)$ shares the same periodicity as $f(x, y)$. \square

Exploiting the above result, $g_u(u, \delta)$ and $g_\delta(u, \delta)$ are periodic functions with the same period as $g(u, \delta)$. Then, resorting again to Lemma E.5.1, $g_{uu}(u, \delta)$ and $g_{u\delta}(u, \delta)$, as well as $g_{\delta u}(u, \delta)$ and $g_{\delta\delta}(u, \delta)$, share the same periodicity as $g_u(u, \delta)$ and $g_\delta(u, \delta)$, respectively. As a result, the first and second order derivatives of $g(u, \delta)$ are periodic functions as $g(u, \delta)$.

Bibliography

- [1] D. Adamy. *EW 101 : a first course in electronic warfare*. Artech House, Norwood, MA, 2001.
- [2] M. A. Ainslie. *Principles of sonar performance modelling*, volume 707. Springer, 2010.
- [3] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974.
- [4] T. W. Anderson. Asymptotic theory for principal component analysis. *Ann. Math. Statist.*, 34(1):122–148, 03 1963.
- [5] P. Antonik, M. C. Wicks, H. D. Griffiths, and C. J. Baker. Frequency diverse array radars. In *2006 IEEE Conference on Radar*, pages 3–pp. IEEE, 2006.
- [6] A. Aubry, V. Carotenuto, A. De Maio, M. A. Govoni, and A. Farina. Experimental analysis of block-sparsity-based spectrum sensing techniques for cognitive radar. *IEEE Transactions on Aerospace and Electronic Systems*, 57(1):355–370, 2021.
- [7] A. Aubry, A. De Maio, and V. Carotenuto. Optimality claims for the FML covariance estimator with respect to two matrix norms. *IEEE Transactions on Aerospace and Electronic Systems*, 49(3):2055–2057, 2013.
- [8] A. Aubry, A. De Maio, S. Marano, and M. Rosamilia. Single-pulse simultaneous target detection and angle estimation in a multichannel phased array radar. *IEEE Transactions on Signal Processing*, 68:6649–6664, 2020.
- [9] A. Aubry, A. De Maio, S. Marano, and M. Rosamilia. Structured covariance matrix estimation with missing-(complex) data for radar applications via expectation-maximization. *IEEE Transactions on Signal Processing*, 69:5920–5934, 2021.

-
- [10] A. Aubry, A. De Maio, L. Pallotta, and A. Farina. Maximum likelihood estimation of a structured covariance matrix with a condition number constraint. *IEEE Transactions on Signal Processing*, 60(6):3004–3021, 2012.
- [11] A. Aubry, A. De Maio, L. Pallotta, and A. Farina. Radar detection of distributed targets in homogeneous interference whose inverse covariance structure is defined via unitary invariant functions. *IEEE Transactions on Signal Processing*, 61(20):4949–4961, 2013.
- [12] A. Aubry, A. De Maio, A. Zappone, M. Razaviyayn, and Z. Luo. A new sequential optimization procedure and its applications to resource allocation for wireless systems. *IEEE Transactions on Signal Processing*, 66(24):6518–6533, 2018.
- [13] C. A. Balanis. *Antenna theory: Analysis and design*. New York: Harper & Row Publishers, Inc., 1982.
- [14] F. Bandiera, O. Besson, and G. Ricci. An ABORT-like detector with improved mismatched signals rejection capabilities. *IEEE Transactions on Signal Processing*, 56(1):14–25, 2008.
- [15] F. Bandiera, A. De Maio, and G. Ricci. Adaptive CFAR radar detection with conic rejection. *IEEE Transactions on Signal Processing*, 55(6):2533–2541, 2007.
- [16] A. I. Barros, J. B. Frenk, S. Schaible, and S. Zhang. A new algorithm for generalized fractional programs. *Math. Program.*, 72(2):147–175, February 1996.
- [17] A. Basit, W. Khan, S. Khan, and I. M. Qureshi. Development of frequency diverse array radar technology: a review. *IET Radar, Sonar & Navigation*, 12(2):165–175, 2018.
- [18] A. Beck. *Introduction to nonlinear optimization: Theory, algorithms, and applications with MATLAB*. SIAM, 2014.
- [19] D. P. Bertsekas. *Nonlinear programming*. Athena Scientific, Belmont, Massachusetts, 1999.
- [20] S. D. Blunt and K. Gerlach. Efficient robust AMF using the FRACTA algorithm. *IEEE Transactions on Aerospace and Electronic Systems*, 41(2):537–548, 2005.
- [21] S. Bose and A. O. Steinhardt. A maximal invariant framework for adaptive detection with structured and unstructured covariance matrices. *IEEE Transactions on Signal Processing*, 43(9):2164–2175, 1995.
-

-
- [22] Y. Bresler. Maximum likelihood estimation of a linearly structured covariance with application to antenna array processing. In *Fourth Annual ASSP Workshop on Spectrum Estimation and Modeling*, pages 172–175, 1988.
- [23] E. Brookner. MIMO versus conventional radar performance against jammers. In *2017 IEEE Radar Conference (RadarConf)*, pages 0703–0708, 2017.
- [24] J. P. Burg, D. G. Luenberger, and D. L. Wenger. Estimation of structured covariance matrices. *Proceedings of the IEEE*, 70(9):963–974, 1982.
- [25] L. Cai and H. Wang. A persymmetric multiband GLR algorithm. *IEEE Transactions on Aerospace and Electronic Systems*, 28(3):806–816, 1992.
- [26] A. Cantoni and P. Butler. Properties of the eigenvectors of persymmetric matrices with applications to communication theory. *IEEE Transactions on Communications*, 24(8):804–809, 1976.
- [27] V. Carotenuto, A. De Maio, D. Orlando, and P. Stoica. Model order selection rules for covariance structure classification in radar. *IEEE Transactions on Signal Processing*, 65(20):5305–5317, 2017.
- [28] B. Chen, X. Chen, Y. Huang, and J. Guan. Transmit beampattern synthesis for the FDA radar. *IEEE Antennas and Wireless Propagation Letters*, 17(1):98–101, 2017.
- [29] B. Chen, S. He, Z. Li, and S. Zhang. Maximum block improvement and polynomial optimization. *SIAM J. on Optimization*, 22(1):87–107, January 2012.
- [30] P. Chen, W. L. Melvin, and M. C. Wicks. Screening among multivariate normal data. *Journal of Multivariate Analysis*, 69(1):10–29, 1999.
- [31] E. Conte, A. De Maio, and C. Galdi. Statistical analysis of real clutter at different range resolutions. *IEEE Transactions on Aerospace and Electronic Systems*, 40(3):903–918, 2004.
- [32] E. Conte, M. Lops, and G. Ricci. Asymptotically optimum radar detection in compound-gaussian clutter. *IEEE Transactions on Aerospace and Electronic Systems*, 31(2):617–625, 1995.
- [33] J. P. Crouzeix and J. A. Ferland. Algorithms for generalized fractional programming. *Math. Program.*, 52(1–3):191–207, May 1991.
- [34] C. Cui, J. Xiong, W.-Q. Wang, and W. Wu. Localization performance analysis of fda radar receiver with two-stage estimator. *IEEE Transactions on Aerospace and Electronic Systems*, 54(6):2873–2887, 2018.
-

-
- [35] R. C. Davis, L. E. Brennan, and L. S. Reed. Angle estimation with adaptive arrays in external noise fields. *IEEE Transactions on Aerospace and Electronic Systems*, AES-12(2):179–186, 1976.
- [36] A. De Maio. Rao test for adaptive detection in Gaussian interference with unknown covariance matrix. *IEEE Transactions on Signal Processing*, 55(7):3577–3584, 2007.
- [37] A. De Maio and G. Alfano. Polarimetric adaptive detection in non-gaussian noise. *Signal Processing*, 83(2):297–306, 2003.
- [38] A. De Maio, S. De Nicola, A. Farina, and S. Iommelli. Adaptive detection of a signal with angle uncertainty. *IET Radar, Sonar Navigation*, 4(4):537–547, 2010.
- [39] A. De Maio and M. S. Greco. *Modern Radar Detection Theory*. Electromagnetics and Radar. Institution of Engineering and Technology, 2015.
- [40] A. De Maio and D. Orlando. Adaptive radar detection of a subspace signal embedded in subspace structured plus gaussian interference via invariance. *IEEE Transactions on Signal Processing*, 64(8):2156–2167, 2016.
- [41] A. De Maio and D. Orlando. A survey on two-stage decision schemes for point-like targets in gaussian interference. *IEEE Aerospace and Electronic Systems Magazine*, 31(4):20–29, 2016.
- [42] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39(1):1–38, 1977.
- [43] X. Du, A. Aubry, A. De Maio, and G. Cui. Toeplitz structured covariance matrix estimation for radar applications. *IEEE Signal Processing Letters*, 27:595–599, 2020.
- [44] X. Du, A. Aubry, A. De Maio, and G. Cui. Toeplitz structured covariance matrix estimation for radar applications. In *2020 IEEE 11th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pages 1–5, 2020.
- [45] A. Farina. *Antenna-based Signal Processing Techniques for Radar Systems*. Antennas and Propagation Library. Artech House, 1992.
- [46] J. S. Gomez Noris. Wireless networks for beamforming in distributed phased array radar. Master’s thesis, Naval Postgraduate School, Monterey, CA, 2007.
- [47] J. R. Guerci. *Space-time adaptive processing for radar*. Artech House, 2014.
-

-
- [48] R. Gui, W.-Q. Wang, Y. Pan, and J. Xu. Cognitive target tracking via angle-range-doppler estimation with transmit subaperturing fda radar. *IEEE Journal of Selected Topics in Signal Processing*, 12(1):76–89, 2018.
- [49] S. Han, A. De Maio, V. Carotenuto, L. Pallotta, and X. Huang. Censoring outliers in radar data: An approximate ml approach and its analysis. *IEEE Transactions on Aerospace and Electronic Systems*, 55(2):534–546, 2019.
- [50] E. J. Hannan and B. G. Quinn. The determination of the order of an autoregression. *Journal of the Royal Statistical Society. Series B (Methodological)*, 41(2):190–195, 1979.
- [51] R. C. Heimiller, J. E. Belyea, and P. G. Tomlinson. Distributed array radar. *IEEE Transactions on Aerospace and Electronic Systems*, AES-19(6):831–839, 1983.
- [52] R. A. Horn. *Matrix Analysis: Second Edition*. Cambridge University Press, oct 2012.
- [53] R. A. Horn and C. R Johnson. *Topics in matrix analysis*. Cambridge university press, 1994.
- [54] D. R. Hunter and K. Lange. A tutorial on MM algorithms. *The American Statistician*, 58(1):30–37, 2004.
- [55] J. S. Pang, M. Razaviyayn, and A. Alvarado. Computing B-stationary points of nonsmooth DC programs. *Math. Oper. Res.*, 42(1):95–118, January 2017.
- [56] M. Jamshidian and P. M. Bentler. ML estimation of mean and covariance structures with missing data using complete data routines. *Journal of Educational and Behavioral Statistics*, 24(1):21–41, 1999.
- [57] M. Jansson and P. Stoica. Analysis of forward-only and forward-backward sample covariances. In *Acoustics, Speech, and Signal Processing, IEEE International Conference on*, volume 5, pages 2825–2828, Los Alamitos, CA, USA, mar 1999. IEEE Computer Society.
- [58] B. Kang, V. Monga, and M. Rangaswamy. Rank-constrained maximum likelihood estimation of structured covariance matrices. *IEEE Transactions on Aerospace and Electronic Systems*, 50(1):501–515, 2014.
- [59] S. M. Kay. *Fundamentals of Statistical Signal Processing: Detection theory*. Prentice Hall Signal Processing Series. Prentice-Hall PTR, 1998.
- [60] E. J. Kelly. An Adaptive Detection Algorithm. *IEEE Transactions on Aerospace and Electronic Systems*, AES-22(2):115–127, 3 1986.
-

-
- [61] R. Klemm. *Principles of space-time adaptive processing*. Number 12. IET, 2002.
- [62] R. Klemm. *Novel Radar Techniques and Applications*. Radar, Sonar and Navigation. Institution of Engineering & Technology, 2018.
- [63] S. Kraut, L. L. Scharf, and L. T. McWhorter. Adaptive subspace detectors. *IEEE Transactions on Signal Processing*, 49(1):1–16, 2001.
- [64] L. Lan, A. Marino, A. Aubry, A. De Maio, G. Liao, and J. Xu. Design of adaptive detectors for FDA-MIMO radar. In *2020 IEEE 11th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pages 1–5, 2020.
- [65] L. Lan, A. Marino, A. Aubry, A. De Maio, G. Liao, and J. Xu. Design of GLR-based detectors for FDA-MIMO radar. In *2020 IEEE 7th International Workshop on Metrology for AeroSpace (MetroAeroSpace)*, pages 17–21, 2020.
- [66] L. Lan, A. Marino, A. Aubry, A. De Maio, G. Liao, J. Xu, and Y. Zhang. GLRT-based adaptive target detection in FDA-MIMO radar. *IEEE Transactions on Aerospace and Electronic Systems*, 57(1):597–613, 2021.
- [67] L. Lan, M. Rosamilia, A. Aubry, A. De Maio, and G. Liao. Single-snapshot angle and incremental range estimation for FDA-MIMO radar. *IEEE Transactions on Aerospace and Electronic Systems*, 57(6):3705–3718, 2021.
- [68] L. Lei, G. Zhang, and R. J. Doviak. Design and simulations for a cylindrical polarimetric phased array weather radar. In *2011 IEEE RadarCon (RADAR)*, pages 671–674, 2011.
- [69] A. J. Lemonte and S. L. P. Ferrari. The local power of the gradient test. *Annals of the Institute of Statistical Mathematics*, 64(2):373–381, October 2010.
- [70] J. Li and R.T. Compton. Angle and polarization estimation using ESPRIT with a polarization sensitive array. *IEEE Transactions on Antennas and Propagation*, 39(9):1376–1383, 1991.
- [71] J. Li and P. Stoica. *Robust Adaptive Beamforming*. Wiley Series in Telecommunications and Signal Processing. Wiley, 2005.
- [72] D. A. Linebarger, R. D. DeGroat, E. M. Dowling, P. Stoica, and G. L. Fudge. Incorporating a priori information into music-algorithms and analysis. *Signal Processing*, 46(1):85 – 104, 1995.
- [73] C. Liu and D. B. Rubin. The ECME algorithm: A simple extension of EM and ECM with faster monotone convergence. *Biometrika*, 81(4):633–648, 1994.
-

-
- [74] J. Liu, W. Liu, H. Liu, B. Chen, X. Xia, and F. Dai. Average SINR calculation of a persymmetric sample matrix inversion beamformer. *IEEE Transactions on Signal Processing*, 64(8):2135–2145, 2016.
- [75] W. Liu, J. Liu, C. Hao, Y. Gao, and Y.-L. Wang. Multichannel adaptive signal detection: Basic theory and literature review. *Science China Information Sciences*, 65(2):1–40, 2022.
- [76] Y. Ma, P. Wei, and H. Zhang. General focusing beamformer for FDA: Mathematical model and resolution analysis. *IEEE Transactions on Antennas and Propagation*, 67(5):3089–3100, 2019.
- [77] X. L. Meng and D. B. Rubin. Maximum likelihood estimation via the ECM algorithm: A general framework. *Biometrika*, 80(2):267–278, 06 1993.
- [78] A. Nehorai and E. Paldi. Vector-sensor array processing for electromagnetic source localization. *IEEE Transactions on Signal Processing*, 42(2):376–398, 1994.
- [79] U. Nickel. Monopulse estimation with adaptive arrays. *IEE Proceedings F - Radar and Signal Processing*, 140(5):303–308, Oct 1993.
- [80] U. Nickel. Monopulse estimation with subarray-adaptive arrays and arbitrary sum and difference beams. *IEE Proceedings - Radar, Sonar and Navigation*, 143:232–238(6), August 1996.
- [81] U. Nickel. Overview of generalized monopulse estimation. *IEEE Aerospace and Electronic Systems Magazine*, 21(6):27–56, 2006.
- [82] L. M. Novak, M. C. Burl, and W. W. Irving. Optimal polarimetric processing for enhanced target detection. *IEEE Transactions on Aerospace and Electronic Systems*, 29(1):234–244, 1993.
- [83] L. M. Novak, M. B. Sechtin, and M. J. Cardullo. Studies of target detection algorithms that use polarimetric radar data. *IEEE Transactions on Aerospace and Electronic Systems*, 25(2):150–165, 1989.
- [84] A. S. Paine. Minimum variance monopulse technique for an adaptive phased array radar. *IEE Proceedings - Radar, Sonar and Navigation*, 145(6):374–380, 1998.
- [85] J. S. Pang. Partially B-regular optimization and equilibrium problems. *Mathematics of Operations Research*, 32(3):687–699, 2007.
- [86] N. B. Pulsone and C. M. Rader. Adaptive beamformer orthogonal rejection test. *IEEE Transactions on Signal Processing*, 49(3):521–529, 2001.
-

-
- [87] R. S. Raghavan, N. Pulsone, and D. J. McLaughlin. Performance of the GLRT for adaptive vector subspace detection. *IEEE Transactions on Aerospace and Electronic Systems*, 32(4):1473–1487, 1996.
- [88] M. Razaviyayn, M. Hong, and Z.-Q. Luo. A unified convergence analysis of block successive minimization methods for nonsmooth optimization. *SIAM Journal on Optimization*, 23(2):1126–1153, 2013.
- [89] M. A. Richards, W. L. Melvin, J. Scheer, J. A. Scheer, and W. A. Holm. *Principles of Modern Radar: Radar Applications, Volume 3*. Electromagnetics and Radar. Institution of Engineering and Technology, 2013.
- [90] M. A. Richards, J. A. Scheer, and W. A. Holm. *Principles of Modern Radar: Basic principles*. Radar, Sonar & Navigation. Stevenage, UK: Institution of Engineering and Technology, 2010.
- [91] M. A. Richards, J. A. Scheer, W. L. Melvin, and J. Scheer. *Principles of Modern Radar: Advanced Techniques, Volume 2*. EBSCO ebook academic collection. Institution of Engineering and Technology, 2012.
- [92] C. D. Richmond. Performance of the adaptive sidelobe blanker detection algorithm in homogeneous environments. *IEEE Transactions on Signal Processing*, 48(5):1235–1247, 2000.
- [93] F. C. Robey, D. R. Fuhrmann, E. J. Kelly, and R. Nitzberg. A CFAR adaptive matched filter detector. *IEEE Transactions on Aerospace and Electronic Systems*, 28(1):208–216, 1992.
- [94] M. Rosamilia, A. Aubry, V. Carotenuto, and A. De Maio. Experimental analysis of structured covariance estimators with missing data. In *2021 IEEE 8th International Workshop on Metrology for AeroSpace (MetroAeroSpace)*, pages 271–276, 2021.
- [95] S. Sabatini and M. Tarantino. *Multifunction Array Radar: System Design and Analysis*. Artech House radar library. Artech House, 1994.
- [96] P. F. Sannmartino, C. J. Baker, and H. D. Griffiths. Frequency diverse mimo techniques for radar. *IEEE Transactions on Aerospace and Electronic Systems*, 49(1):201–222, 2013.
- [97] J. Sexton and A. R. Swensen. ECM algorithms that converge at the rate of EM. *Biometrika*, 87(3):651–662, 2000.
- [98] H. Shao, J. Dai, J. Xiong, H. Chen, and W.-Q. Wang. Dot-shaped range-angle eampattern synthesis for frequency diverse array. *IEEE antennas and wireless propagation letters*, 15:1703–1706, 2016.
-

-
- [99] S. Siegfried. Fractional Programming - 2. on Dinkelbach'S Algorithm. *Management Science*, 22(8):868–873, 4 1976.
- [100] M. I. Skolnik. *Radar Handbook, Third Edition*. Electronics electrical engineering. McGraw-Hill Education, 2008.
- [101] M. Steiner and K. Gerlach. Fast converging adaptive processor or a structured covariance matrix. *IEEE Transactions on Aerospace and Electronic Systems*, 36(4):1115–1126, 2000.
- [102] R. W. Stewart, K. W. Barlee, and D. S. Atkinson. *Software Defined Radio using MATLAB & Simulink and the RTL-SDR*. Strathclyde Academic Media, 2015.
- [103] J. W. Tukey. *Exploratory Data Analysis*. Exploratory Data Analysis. Addison Wesley Publishing Company, 1970.
- [104] H. L. Van Trees. *Detection, Estimation, and Linear Modulation Theory: Part I*. Detection, Estimation, and Modulation Theory. Wiley, New York, 2001.
- [105] H. L. Van Trees. *Optimum Array Processing: Part IV*. Detection, Estimation, and Modulation Theory. Hoboken, NJ, USA: Wiley, 2004.
- [106] S. A. Vorobyov, A. B. Gershman, and Z.-Q.Luo. Robust adaptive beamforming using worst-case performance optimization: a solution to the signal mismatch problem. *IEEE Transactions on Signal Processing*, 51(2):313–324, 2003.
- [107] A. Wald. Tests of statistical hypotheses concerning several parameters when the number of observations is large. *Transactions of the American Mathematical Society*, 54(3):426–426, March 1943.
- [108] W.-Q. Wang. Overview of frequency diverse array in radar and navigation applications. *IET Radar, Sonar & Navigation*, 10(6):1001–1012, 2016.
- [109] M. Wax and I. Ziskind. Detection of the number of coherent signals by the MDL principle. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(8):1190–1196, 1989.
- [110] K. Werner, M. Jansson, and P. Stoica. On estimation of covariance matrices with kronecker product structure. *IEEE Transactions on Signal Processing*, 56(2):478–491, 2008.
- [111] A. J. Wileman and S. Perinpanayagam. Failure mechanisms of radar and RF systems. *Procedia CIRP*, 11:56 – 61, 2013. 2nd International Through-life Engineering Services Conference.
-

-
- [112] W. D. Wirth and Institution of Electrical Engineers. *Radar Techniques Using Array Antennas*. Electromagnetics and Radar. Stevenage, UK: Institution of Engineering and Technology, 2001.
- [113] S. J. Wright. Coordinate descent algorithms. *Mathematical Programming*, 151(1):3–34, 2015.
- [114] C. F. J. Wu. On the convergence properties of the EM algorithm. *The Annals of Statistics*, 11(1):95–103, 1983.
- [115] L. Wu, P. Babu, and D. P. Palomar. Transmit waveform/receive filter design for MIMO radar with multiple waveform constraints. *IEEE Transactions on Signal Processing*, 66(6):1526–1540, 2018.
- [116] C. Xiong, G. Xiao, Y. Hou, and M. Hameed. A compressed sensing-based element failure diagnosis method for phased array antenna during beam steering. *IEEE Antennas and Wireless Propagation Letters*, 18(9):1756–1760, 2019.
- [117] G. Xu, R. H. Roy, and T. Kailath. Detection of number of sources via exploitation of centro-symmetry property. *IEEE Transactions on Signal Processing*, 42(1):102–112, 1994.
- [118] Y. Xu, X. Shi, W. Li, J. Xu, and L. Huang. Low-sidelobe range-angle beamforming with FDA using multiple parameter optimization. *IEEE Transactions on Aerospace and Electronic Systems*, 55(5):2214–2225, 2018.
- [119] T. Yang, A. De Maio, J. Zheng, T. Su, V. Carotenuto, and A. Aubry. An adaptive radar signal processor for uavs detection with super-resolution capabilities. *IEEE Sensors Journal*, 21(18):20778–20787, 2021.
- [120] Z. Yang, J. Li, P. Stoica, and L. Xie. Sparse methods for direction-of-arrival estimation. In *Academic Press Library in Signal Processing, Volume 7*, pages 509–581. Elsevier, 2018.
- [121] K. B. Yu and D. J. Murrow. Adaptive digital beamforming for angle estimation in jamming. *IEEE Transactions on Aerospace and Electronic Systems*, 37(2):508–523, 2001.
- [122] M. A. Zatman. ABF limitations when using either toeplitz covariance matrix estimators or the parametric vector ar technique. In *Conference Record of Thirty-Fifth Asilomar Conference on Signals, Systems and Computers (Cat. No. 01CH37256)*, volume 2, pages 1111–1115. IEEE, 2001.
- [123] Y. Zhang, J. P. Lie, B. P. Ng, and C. M. S. See. Robust minimum ℓ_1 -norm adaptive beamformer against intermittent sensor failure and steering vector error. *IEEE Transactions on Antennas and Propagation*, 58(5):1796–1801, 2010.
-

-
- [124] L. Zhao and D. P. Palomar. Maximin joint optimization of transmitting code and receiving filter in radar and communications. *IEEE Transactions on Signal Processing*, 65(4):850–863, 2017.
-

Author's Publications

Published journal papers

- [J1] A. Aubry, A. De Maio, S. Marano, and M. Rosamilia, "Single-Pulse Simultaneous Target Detection and Angle Estimation in a Multichannel Phased Array Radar," *IEEE Transactions on Signal Processing*, vol. 68, pp. 6649-6664, 2020.
- [J2] L. Lan, M. Rosamilia, A. Aubry, A. De Maio, and G. Liao, "Single-Snapshot Angle and Incremental Range Estimation for FDA-MIMO Radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 6, pp. 3705-3718, Dec. 2021.
- [J3] A. Aubry, A. De Maio, and M. Rosamilia, "Reconfigurable Intelligent Surfaces for N-LOS Radar Surveillance," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 10, pp. 10735-10749, Oct. 2021.
- [J4] A. Aubry, A. De Maio, S. Marano, and M. Rosamilia, "Structured Covariance Matrix Estimation With Missing-(Complex) Data for Radar Applications via Expectation-Maximization," *IEEE Transactions on Signal Processing*, vol. 69, pp. 5920-5934, 2021.
- [J5] A. Aubry, V. Carotenuto, A. De Maio, M. Rosamilia, and S. Marano, "Adaptive Radar Detection in the Presence of Missing-data," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 58, no. 4, pp. 3283-3296, Aug. 2022.
- [J6] L. Lan, M. Rosamilia, A. Aubry, A. De Maio, G. Liao, and J. Xu, "Adaptive Target Detection with Polarimetric FDA-MIMO Radar," *IEEE Transactions on Aerospace and Electronic Systems*, Early Access, 2022.

- [J7] M. Rosamilia, A. Balleri, A. De Maio, A. Aubry, and V. Carotenuto, "Radar Detection Performance Prediction using Measured UAVs RCS Data", *IEEE Transactions on Aerospace and Electronic Systems*, Early Access, 2022.
- [J8] A. Aubry, A. De Maio, L. Lan, and M. Rosamilia, "Adaptive Radar Detection and Bearing Estimation in the Presence of Unknown Mutual Coupling", submitted to *IEEE Transactions on Signal Processing*, 2022.

Published conference papers

- [C1] M. Rosamilia, A. Aubry, A. De Maio, and S. Marano, "Simultaneous Radar Detection and Constrained Target Angle Estimation via Dinkelbach Algorithm," *2020 IEEE Radar Conference (RadarConf20)*, 2020, pp. 1-6.
- [C2] M. Rosamilia, A. Aubry, V. Carotenuto, and A. De Maio, "Experimental Analysis of Structured Covariance Estimators with Missing data," *2021 IEEE 8th International Workshop on Metrology for AeroSpace (MetroAeroSpace)*, 2021, pp. 271-276.
- [C3] A. Aubry, A. De Maio, and M. Rosamilia, "RIS-Aided Radar Sensing in N-LOS Environment," *2021 IEEE 8th International Workshop on Metrology for AeroSpace (MetroAeroSpace)*, 2021, pp. 277-282.
- [C4] M. Rosamilia, L. Lan, A. Aubry, and A. De Maio, "Polarimetric FDA-MIMO Radar Detection," *2022 IEEE Radar Conference (RadarConf22)*, 2022, pp. 1-6.
- [C5] M. Rosamilia, A. Aubry, A. Balleri, V. Carotenuto and A. De Maio, "RCS Measurements of UAVs and Their Statistical Analysis," *2022 IEEE 9th International Workshop on Metrology for AeroSpace (MetroAeroSpace)*, 2022, pp. 179-184.
- [C6] M. Rosamilia, A. Aubry, A. Balleri, V. Carotenuto and A. De Maio, "Performance Prediction of the Coherent Radar Detector on Measured UAVs Data," *Radar 2022 International Conference on Radar Systems*, 2022.
-