

© 2016

Shunqiao Sun

ALL RIGHTS RESERVED

MIMO RADARS WITH SPARSE SENSING

By

SHUNQIAO SUN

A dissertation submitted to the

Graduate School—New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Electrical and Computer Engineering

Written under the direction of

Athina P. Petropulu

And approved by

New Brunswick, New Jersey

January, 2016

ABSTRACT OF THE DISSERTATION

MIMO Radars with Sparse Sensing

By SHUNQIAO SUN

Dissertation Director:

Athina P. Petropulu

Multi-input and multi-output (MIMO) radars achieve high resolution of arrival direction by transmitting orthogonal waveforms, performing matched filtering at the receiver end and then jointly processing the measurements of all receive antennas. This dissertation studies the use of compressive sensing (CS) and matrix completion (MC) techniques as means of reducing the amount of data that need to be collected by a MIMO radar system, without sacrificing the system's good resolution properties. MIMO radars with sparse sensing are useful in networked radar scenarios, in which the joint processing of the measurements is done at a fusion center, which might be connected to the receive antennas via a wireless link. In such scenarios, reduced amount of data translates into bandwidth and power saving in the receiver-fusion center link.

First, we consider previously defined CS-based MIMO radar schemes, and propose optimal transmit antenna power allocation and transmit waveform design schemes that improve target localization performance. The optimization criterion is to minimize the coherence between the columns of the sensing matrix. In addition, we propose a clutter suppression scheme based on the Capon beamforming in the CS-based MIMO radars.

Second, we propose a novel MIMO radar approach based on matrix completion, termed as MIMO-MC, in which each receive node either performs matched filtering

with a small number of randomly selected dictionary waveforms, or obtains sub-Nyquist samples of the received target echoes at randomly sampled instants, and forwards the results to a fusion center. Based on the received samples, and with knowledge of the sampling scheme, the fusion center partially fills a matrix, referred to as the data matrix and subsequently applies matrix completion techniques to estimate the full matrix. The completed data matrix is used for target estimation with standard array signal processing methods. We show that MIMO-MC radars share the advantages of the CS-based radars, i.e., high resolution with reduced amounts of data, but unlike CS-based radars do not require grid discretization and thus are not sensitive to basis mismatch. For MIMO radars with uniform linear arrays, we investigate the relationship between the coherence of the data matrix and the transmit waveforms, and formulate an optimal waveform design problem. This is an optimization problem on the complex Stiefel manifold, which is then solved via the modified steepest descent and the modified Newton algorithms with nonmonotone line search methods. We also propose transmit and receive beamforming schemes to significantly reduce the sampling rate at the receiver end in MIMO-MC radars.

Acknowledgements

The graduate study at Rutgers is a fruitful and rewarding experience to me. The days and nights spent here have constituted one of the most memorable periods of my life. First of all, I would like to express my deepest gratitude to my advisor, Prof. Athina P. Petropulu, for her generous support and guidance throughout the past years. Prof. Petropulu not only continuously encourages me to push the boundaries of innovation but also always gives me a helping hand whenever I am looking for advices in research or life. Her great passion in research, unexhausted energy, constant encouragement, and dedication to students have set up an inspirational model for me in both professional career and personal life.

I am also very grateful to Prof. Waheed U. Bajwa, who also served on my Ph.D. dissertation committee and oral qualifying exam committee, for his insightful guidance and suggestions during our collaborative research. I would also express my gratitude to other members of my Ph.D. dissertation committee, Prof. Wade Trappe and Prof. Michael N. Katehakis, for their precious time in reviewing my dissertation and their valuable suggestions. I would also like to thank Prof. Sophocles J. Orfanidis served on my dissertation proposal committee and oral qualifying exam committee. Prof. Orfanidis gave me many valuable comments on the physical model of my thesis. I am also very grateful to the other members served on my oral qualifying exam committee, Prof. Lawrence Rabiner, Prof. Peddapullaiah Sannuti, Prof. David Daut and Prof. Predrag Spasojevic.

I am also deeply indebted to current and former members of the Communications and Signal Processing Laboratory (CSPL), including Yao Yu, Jiangyuan Li, Yupeng Liu, Shuangyu Luo, Bo Li, Dionysios Kalogerias, Xiaqing Yang, Chrysanthi Koumpouzi, Wenjun Yan, Ning Xie and Jing Xu for their all kinds of help and friendship. In addition,

I would like to thank my co-author, Dr. Rabinder N. Madan, without whom this thesis would be incomplete. Furthermore, I am very grateful to the graduate director, Prof. Zoran Gajic, as well as current and former ECE staff, John Scafidi, Noraida Martinez, Ora Titus and John P. McCarthy, who have given me numerous help.

Specially, I want to thank my parents for their unconditional love, support and encouragement. Every step of my life could not be made without their constant support and dedication. I am proud to be their son.

Last but not least, I am very thankful for all the love and blessings that God has put into my life.

Dedication

To my parents

Table of Contents

Abstract	ii
Acknowledgements	iv
Dedication	vi
List of Tables	xii
List of Figures	xiii
1. Introduction	1
1.1. Radar Basics	1
1.2. MIMO Radars	6
1.2.1. Widely Separated MIMO Radars	7
1.2.2. Colocated MIMO Radars	7
1.2.3. MIMO Radar Ambiguity Function	10
1.3. Compressive Sensing (CS) Based MIMO Radars	11
1.3.1. Compressive Sensing	11
1.3.2. CS-Based Colocated MIMO Radars	12
1.3.3. Basis Incoherence in CS-Based MIMO Radars	14
1.3.4. Basis Mismatch in CS-Based MIMO Radars	15
1.4. Matrix Completion Theory	17
1.5. Contributions of the Dissertation	20
1.5.1. Performance Optimization and Clutter Suppression in CS-Based MIMO Radars	20
1.5.2. MIMO-MC Radar: A MIMO Radar Approach Based on Matrix Completion	21

1.5.3. Coherence Analysis and Optimal Waveform Design in MIMO Radars with Matrix Completion	22
1.5.4. Transmit and Receive Beamforming in MIMO Radars with Matrix Completion	23
1.6. Outline of the Dissertation	24
1.7. Notation	25
2. Power Allocation and Waveform Design in Compressive Sensing Based MIMO Radars	27
2.1. Introduction	27
2.2. Compressive Sensing (CS) Based MIMO Radars	29
2.2.1. CS-Based Colocated MIMO Radars	29
2.2.2. CS-Based Widely Separated MIMO Radars	29
2.3. Power Allocation and Waveform Design for Colocated CS-MIMO Radars	32
2.3.1. Power Allocation	32
2.3.2. Waveform Design	35
2.4. Power Allocation for Widely Separated CS-MIMO Radars	35
2.5. Simulations	37
2.5.1. Simulation Results for Colocated CS-MIMO Radars	37
Power Allocation	37
Waveform Design	39
2.5.2. Power Allocation for Widely Separated CS-MIMO Radars	40
2.6. Summary	43
3. Clutter Suppression in Colocated Compressive Sensing Based MIMO Radars	45
3.1. Introduction	45
3.2. Clutter Rejection in CS-Based MIMO Radars	46
3.3. Power Allocation in CS-based Clutter Suppression	49
3.4. Simulations	51

3.5. Summary	55
4. A MIMO Radar Approach Based on Matrix Completion	56
4.1. Introduction	57
4.2. Colocated MIMO Radar System Model	58
4.2.1. MIMO Radars without Matched Filtering	61
4.3. The proposed MIMO-MC radar approach	61
4.3.1. MIMO-MC with Sampling Scheme I	62
4.3.2. MIMO-MC with Sampling Scheme II	66
4.3.3. Discussion of MC in Sampling Schemes I and II	69
4.3.4. Target Parameters Estimation with Subspace Methods	70
4.4. Numerical Results	72
4.4.1. Matrix Recovery Error under Noisy Observations	72
4.4.2. DOA Resolution with Matrix Completion	75
4.4.3. Comparisons of Sampling Schemes I and II	76
4.5. Summary	78
5. Waveform Design for MIMO Radars with Matrix Completion	80
5.1. Introduction	80
5.2. Background of MIMO-MC Radar	83
5.3. Data Matrix Coherence Analysis and Optimal Waveform Conditions	85
5.3.1. The Coherence of Data Matrix in MIMO-MC Radar	85
5.3.2. Remarks	90
5.3.3. Comparative Study of Two Orthogonal Waveforms	91
5.4. Waveform Design under Spatial Power Spectra Constraints	93
5.4.1. Problem Formulation	93
5.4.2. Derivative and Hessian of Cost Function $f(\mathbf{S})$	94
5.4.3. Modified Steepest Descent on the Complex Stiefel Manifold	95
5.4.4. Modified Newton Algorithm on the Complex Stiefel Manifold	97

5.5. Performance Analysis of the Proposed Waveform When Targets Fall In Different Range Bins	99
5.6. Numerical Results	103
5.6.1. Performance Comparison of Waveform Design Methods	104
5.6.2. Spatial Power Spectra of Optimized Waveform Snapshots	106
5.6.3. Coherence Properties Under Optimized Waveforms	106
Targets in the Same Range Bin	106
Targets in Different Range Bins	109
5.6.4. Matrix Recovery Error Performance	110
Targets in the Same Range Bin	110
Targets in Different Range Bins	111
5.6.5. Target Estimation Performance via MC	112
5.7. Summary	115
6. On Transmit Beamforming in MIMO Radars with Matrix Completion	117
6.1. Introduction	117
6.2. On the Rank-1 beamforming	119
6.3. On the Multi-Rank Beamforming	121
6.4. DOA Estimation based on Matrix Completion	123
6.5. Simulations	124
6.6. Summary	126
7. On Receive Beamforming via Matrix Completion	127
7.1. Introduction	127
7.2. Fundamentals of Classical Beamforming	129
7.3. Covariance Matrix Estimation with Matrix Completion	130
7.3.1. Training Data Matrix Estimation via Matrix Completion	131
7.3.2. Training Data Matrix Estimation Error Analysis	132
7.4. Robust Low Rank Beamforming	133
7.5. Numerical Results	134

7.6. Summary	137
8. Conclusions and Future Research Directions	139
8.1. Conclusions	139
8.2. Future Research Directions	141
8.2.1. Waveform Design for Range Compression in MIMO-MC Radars	141
8.2.2. Waveform Design for Transmit Beamforming in MIMO-MC Radars	141
8.2.3. Clutter Suppression in MIMO-MC Radars	142
Appendix A. Proof of $\mu_0 = \mu_1 = 1$ for rank-1 matrix \mathbf{Z}_q^{MF}	143
Appendix B. Proof of Lemma 1	145
Appendix C. Derivative and Hessian of $f(\mathbf{S})$	146
Appendix D. Proof of Lemma 2	151
Appendix E. Proof of Lemma 3	152
Appendix F. Proof of Lemma 4	153
Appendix G. Proof of Lemma 5	154
Appendix H. Useful Equations	155
References	156

List of Tables

1.1. Notations	26
4.1. List of parameters used in the signal model	60

List of Figures

1.1. Radar with M_r receive antennas.	4
1.2. The matched filter output for a rectangular pulse with duration T . . .	5
1.3. The matched filter output for two closely spaced targets.	5
1.4. The colocated MIMO radar system is equipped with M_t transmit antennas and M_r receive antennas. Both transmit and receive arrays follow the ULA model.	8
1.5. The basis mismatch due to antenna position errors: (a) The true and estimated positions of $M_t = 20$ TX and $N_r = 6$ RX antennas; (b) The performance of CS when there is basis mismatch due to antenna position errors within $\frac{\lambda}{5}$	16
1.6. The CS recovery performance in the off-grid target scenario using orthogonal matching pursuit: (a) grid size as 0.1° ; (b) grid size as 0.01° . .	17
2.1. SCSM versus M_t under optimal (PPA), uniform (PPA) and random (RPA) power allocation in CS-based colocated MIMO radars. Case (I) $N_r = 6$; Case (II) $N_r = 12$; in both cases $L = 32$	38
2.2. The ROCs of angle estimates under power allocation in CS-based colocated MIMO radars with $L = 32$	39
2.3. The TX distribution in CS colocated MIMO radars with $M_t = 30$, $N_r = 12$ and $L = 32$	40
2.4. The ROCs of angle estimates under waveform design in CS-based colocated MIMO radars with $M_t = N_r = 10$, $L = 16$ and $M = 11$	41
2.5. SEGI versus M_t under optimal and uniform power allocation in CS-based widely separated MIMO radars with $L = 20$	42

2.6. ROC of target location estimates under power allocation in CS-based widely separated MIMO radar with four combinations of M_t and N_r . . .	42
2.7. The TX distribution in CS colocated MIMO radar with $M_t = 30$, $N_r = 2$ and $L = 32$	43
3.1. The diagram for the proposed CS-Capon method	49
3.2. Contour of the beamformed receive data matrix for $N_r = 20$ and $N_r = 200$, respectively.	51
3.3. SCNR comparisons with/without beamforming with $N_t = 20$	52
3.4. SCSM versus M_t under optimal (PPA), uniform (UPA) and random (RPA) power allocation in CS-Capon method with $N_r = 20$	53
3.5. ROC performances under optimal (PPA), uniform (UPA) power allocation in CS-Capon as well as Capon method with $N_t = 50$ and $N_r = 20$	54
4.1. Two sampling schemes in the colocated MIMO radars system: (a) Sampling scheme I; (b) Sampling scheme II.	62
4.2. Structure of the random matched filter bank (RMFB).	63
4.3. Scheme I, $K = 2$ targets: (a) the probability of $\Pr(\max(\mu(U), \mu(V)) > \mu_0)$ of \mathbf{Z}_q^{MF} for $\Delta\theta = 5^\circ$; (b) the average $\max(\mu(U), \mu(V))$ of \mathbf{Z}_q^{MF} as function of number of transmit and receive antennas, and for $\Delta\theta = 5^\circ$; (c) the average $\max(\mu(U), \mu(V))$ of \mathbf{Z}_q^{MF} as function of DOA separation.	64
4.4. Scheme I, $K = 2$ targets: The $\max(\mu(U), \mu(V))$ in terms of M_r for $\Delta\theta = 0^\circ$, $M_t = M_r$	65
4.5. Scheme II, $K = 2$ targets, and G-Orth waveforms: (a) The probability of $\Pr(\max(\mu(U), \mu(V)) > \mu_0)$ of \mathbf{Z}_q for $\Delta\theta = 5^\circ$ and $N = 256$; (b) The average $\max(\mu(U), \mu(V))$ of \mathbf{Z}_q as function of N , for $\Delta\theta = 5^\circ$ and different values of M_t, M_r ; (c) The average $\max(\mu(U), \mu(V))$ of \mathbf{Z}_q as function of $\Delta\theta$, for $N = 128, 256$, and different combinations of M_r, M_t	67
4.6. The maximal power spectrum of the orthogonal waveforms over $N = 32$ snapshots for $M_t = 10$	68

4.7.	The comparison of matrix completion in terms of relative recovery errors with $M_r = 128, M_t = 10, N = 32, \text{SNR} = 25\text{dB}$. There are $K = 2$ targets located at (a) 20° and 40° ; (b) 0° and 80°	69
4.8.	Scheme I, $K = 2$ targets: the relative recovery error for \mathbf{Z}_q^{MF} under different values of DOA separation. $M_r = M_t = 40$	73
4.9.	Scheme II, $K = 2$ targets: the relative recovery errors for \mathbf{Z}_q under Hadamard and Gaussian Orthogonal waveforms, and different values of $\Delta\theta$. $M_r = M_t = 40, N = 256$	74
4.10.	Scheme I: DOA resolution. The parameter are set as $M_r = M_t = 20, p_1 = 0.5$ and $\text{SNR} = 10, 25\text{dB}$, respectively.	76
4.11.	Scheme II, $K = 2, M_r = M_t = 20, N = 256, p_2 = 0.5, \text{SNR} = 10, 25\text{dB}$. DOA resolution with (a) G-Orth waveforms; (b) with Hadamard waveforms.	77
4.12.	Comparisons of the relative recovery errors in terms of number of N (M_t) for $M_r = 40, 80$. The matrix occupancy is set to $p_1 = p_2 = 0.5$. Two targets are generated at random in $[-90^\circ, 90^\circ]$ with DOA separation $\Delta\theta = 5^\circ, 30^\circ$	78
5.1.	The power spectra $ S_i(\alpha_k^t) ^2$ for $\alpha_k^t \in [-\frac{1}{2}, \frac{1}{2}]$ with $M_t = 40$ and $N = 64$. (a) Hadamard waveform; (b) Magnitudes of $ S_i(\alpha_1^t) ^2 + S_i(\alpha_2^t) ^2$ of Hadamard waveform; (c) G-Orth waveform; (d) Magnitudes of $ S_i(\alpha_1^t) ^2 + S_i(\alpha_2^t) ^2$ of G-Orth waveform.	92
5.2.	Comparison of MC error w.r.t. Hadamard and G-Orth waveforms as function of the matrix occupancy ratio.	92
5.3.	Illustration of transmit waveform spectra in the order of the transmit waveform arrival at the receiver end.	102

5.4.	Objective function of (5.29) vs. iterations: (a) DOA space $[-10^\circ : 1^\circ : 10^\circ]$ under modified steepest descent algorithm; (b) DOA space $[-90^\circ : \Delta\theta : 90^\circ]$, $\Delta\theta = 1^\circ, 5^\circ$ under the modified steepest descent algorithm; (c) comparison of the modified steepest descent algorithm and the modified Newton algorithm for $M_t = 20, N = 32$ and DOA space $[-5^\circ : 1^\circ : 5^\circ]$	105
5.5.	Power spectra $ S_i(\alpha_k^t) ^2$ for $\alpha_k^t \in [-0.0868, 0.0868]$ with $M_t = 40$ and $N = 64$ and DOA space $[-10^\circ : 1^\circ : 10^\circ]$. (a) Hadamard waveforms; (b) Optimized waveforms using Hadamard waveforms as initialization; (c) G-Orth waveforms; (d) Optimized waveforms using G-Orth waveforms as initial.	107
5.6.	Coherence $\mu(V)$ and its bound defined in (5.26) for targets in the same range bin. The $K = 4$ targets are located at $[-10^\circ, -5^\circ, 0^\circ, 1^\circ]$	108
5.7.	Coherence $\mu(V)$ for $K = 4$ targets located at $[-10^\circ, -5^\circ, 0^\circ, 1^\circ]$. (a) $\mu(V)$ versus N_1 with $M_t = 40, N = 64$ for targets in different range bins; (b) $\mu(V)$ and its bound defined in (5.63) with $N_1 = N/2$ for targets in different range bins.	109
5.8.	MC error vs. p for targets located at $[-1^\circ, 0^\circ, 0.1^\circ, 0.5^\circ]$. (a) all targets fall in the same range bin; (b) the ranges are $[2020, 3820, 3820, 6220]$ m.	111
5.9.	Angle-Range images using MIMO-MC radars for $K = 4$ targets located at $[-1^\circ, 0^\circ, 0.1^\circ, 0.5^\circ]$ with ranges as $[2020, 3820, 3820, 6220]$ m: (a) Hadamard; (b) G-Orth; (c) Optimized waveforms. The other parameters are $M_t = 20, N = 64, M_r = 20, Q = 40, p = 0.35$	113
5.10.	Probability of resolution comparison for optimized and G-Orth waveforms for $p = 0.35$	114
5.11.	Pseudo-spectra of speed estimation using MUSIC with $Q = 40$ pulses. The left and right side of the figure correspond to $p = 0.3$ and $p = 0.5$, respectively.	114
6.1.	Transmit beampattern under rank-2 beamformer with $M_t = 30$ for directions $[-11^\circ, 2^\circ]$;	124

6.2. Simulations: (a) RMSE versus SNR with $M_r = 60, M_t = 30$; (b) Probability of target resolution with $M_r = 60, M_t = 20$ and SNR = 25dB. The proposed approach is based on subsampling by 50%.	125
7.1. The relative recovery error. The x-axis is the number of samples per degree of freedom (df).	135
7.2. The relative errors of $\hat{\mathbf{S}}$ and $\hat{\mathbf{R}}_{\text{mc}}$ versus INR.	135
7.3. The output SINR versus the number of snapshots.	136
7.4. The output SINR versus SNR.	137
7.5. Beampatterns comparison.	138

Chapter 1

Introduction

In this chapter, we provide the necessary background on multi-input multi-output (MIMO) radars, sparse sensing in MIMO radars based on compressed sensing and matrix completion theory.

1.1 Radar Basics

A radar is a configuration of transmit and receive antennas, used for detecting targets and estimating their parameters, namely, range, direction of arrival (DOA) and speed [1]. It operates by transmitting a particular type of waveform and then processing the target returns as received by its receive antennas.

A pulse radar sends out signals in short bursts or pulses [2]. At the receivers, matched filters are applied to yield maximum signal-to-noise ratio (SNR) output, assuming that the received signal can be modeled as target returns in additive white Gaussian noise (AWGN). Let us consider a radar system in which both transmit and receive antennas are stationary and colocated. At the transmitter, the transmitted waveform is typically modeled as [2]

$$\bar{x}(t) = a(t) \sin(2\pi ft + \varphi(t)), \quad (1.1)$$

where f is the carrier frequency in Hertz. The term $a(t)$ represents amplitude modulation of the carrier. In a pulsed radar, this is typically a rectangular function. The term $\varphi(t)$ models any phase or frequency modulation of the carrier.

The real-valued waveform of (1.1) is more conveniently modeled by its complex

equivalent

$$\bar{x}(t) = s(t) e^{-j2\pi ft}, \quad (1.2)$$

where

$$s(t) = a(t) e^{-j\varphi(t)} \quad (1.3)$$

is the complex baseband signal. The baseband signal $s(t)$ can be obtained by filtering $\bar{x}(t)$ with a baseband filter, which is referred to as demodulation process [3].

- Range estimation: Consider the simple case of a stationary target. The noise-free target return at the receiver, $r(t)$, can be written as [4]

$$r(t) = \beta s(t - \tau_0) e^{-j2\pi f(t - \tau_0)}, \quad (1.4)$$

where β is the target scattering coefficient. The round trip delay is τ_0 , based on which, the distance between the radar and the target, i.e., range, is

$$d = \frac{\tau_0 c}{2}, \quad (1.5)$$

where c is the speed of light.

- Doppler estimation: A moving target introduces Doppler shift in the carrier frequency. Assume the target is moving with velocity v towards radar and $\frac{v}{c} \ll 1$. The received signal can be written as [4]

$$r(t) = \beta s\left(t - \tau_0 + \frac{2vt}{c}\right) e^{-j2\pi f(t - \tau_0 + \frac{2vt}{c})}. \quad (1.6)$$

Assume the transit signal does not change appreciably in a time equal to $\frac{1}{B}$, i.e.,

$$\frac{2vT}{c} \ll \frac{1}{B}, \quad (1.7)$$

or equivalently,

$$BT \ll \frac{c}{2v}. \quad (1.8)$$

Then, we can ignore the time-scale change of the transmitted signal. And the complex envelope of the receive signal can be simplified as [4]

$$r(t) \approx \beta s(t - \tau_0) e^{-j2\pi(f t - f \tau_0 + \nu t)}, \quad (1.9)$$

where $\nu = \frac{2vf}{c}$ is the Doppler shift. To extract the Doppler information, pulse radar needs to transmit multiple pulses.

- DOA estimation: Multiple receive antennas are required to extract the DOA information. Let us consider a uniform linear array (ULA) with the inter-spacing between receive antenna is d_r , shown in Fig. 1.1. Assume a stationary target is located in the far field and its distance to the nearest receive antenna is d and the round trip delay is $\tau_0 = \frac{2d}{c}$. The noise free receive signal at the m -th receive antenna can be written as

$$r_i(t) = \beta s \left(t - \tau_0 - \frac{(m-1)d_r \sin \theta}{c} \right) e^{-j2\pi f \left(t - \tau_0 - \frac{(m-1)d_r \sin \theta}{c} \right)}. \quad (1.10)$$

Let T and B denote the duration and bandwidth of the transmit signal, respectively. Assume that the transmitted signal is narrowband, i.e.,

$$\frac{1}{T} \ll f. \quad (1.11)$$

For example, suppose $d_r = \frac{\lambda}{2}$, where λ is the wavelength. Then, it holds that

$$\frac{(m-1)d_r \sin \theta}{c} = \frac{(m-1)\lambda \sin \theta}{2c} = \frac{(m-1) \sin \theta}{2f} \ll T. \quad (1.12)$$

Consequently, we can ignore the delay term $\frac{(m-1)d_r \sin \theta}{c}$ and the receive signal at the m -th receiver can be written as

$$r_m(t) \approx \beta s(t - \tau_0) e^{-j2\pi f \left(t - \tau_0 - \frac{(m-1)d_r \sin \theta}{c} \right)}. \quad (1.13)$$

Let $\alpha = \frac{fd_r \sin \theta}{c} = \frac{d_r \sin \theta}{\lambda}$, where λ is the wavelength of the signal waveform; α is referred to as the spatial frequency [3]. All the antennas in the receive array perform sampling simultaneously following the Nyquist rule. The collection of samples from all antennas in the receive array at a single time instant is referred to as snapshot. At a given time instant, the snapshot can be viewed as a complex sinusoid of frequency α (spatial frequency). If we apply the fast Fourier transform (FFT) on a single snapshot, the peak locations in the FFT index will correspond to the spatial frequency α , based on which the DOA θ can be obtained as long as $\theta \in [-90^\circ, 90^\circ]$. More accurate DOA information can be extracted using subspace

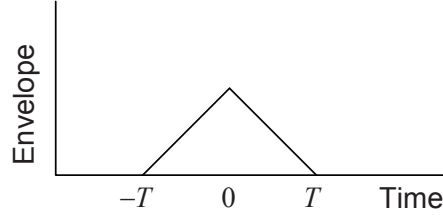


Figure 1.2: The matched filter output for a rectangular pulse with duration T .

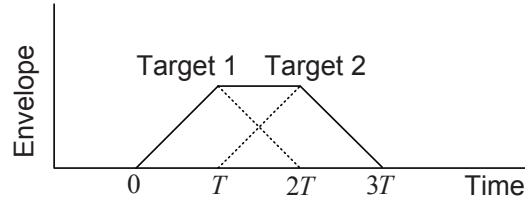


Figure 1.3: The matched filter output for two closely spaced targets.

where

$$A(\tau, \nu) = \int_{-\infty}^{\infty} s(t) s^*(t - \tau) e^{-j2\pi\nu t} dt \quad (1.17)$$

is referred to as the *ambiguity function* of the radar waveform $s(t)$ [6].

The ambiguity function is a major tool for radar signal analysis. The ambiguity function evaluated at $(\tau, \nu) = (0, 0)$ is equal to the matched filter output that is perfectly matched to the signal reflected from the target. The ambiguity function along the zero-Doppler axis (corresponding Doppler shift $\nu = 0$) is the autocorrelation function of the waveform

$$A(\tau, 0) = \int_{-\infty}^{\infty} s(t) s^*(t - \tau) dt = R_s(\tau). \quad (1.18)$$

Consider a single rectangular pulse signal with duration T . There would be a single peak along the zero-Doppler axis as well as zero-delay axis.

Figure 1.2 shows the output of a perfect matched filter with the input as a rectangular pulse, which is a triangular envelope of width $2T$ with a central peak. If the second target is located from the first one with a time delay of T seconds. At a radar receiver, the matched filter output of these two closely spaced targets is plotted in Fig. 1.3. It can be found that the two peaks are resolvable if the the range between two

targets, i.e., the range resolution, satisfies

$$\Delta R \geq \frac{cT}{2}. \quad (1.19)$$

Therefore, the range resolution is proportional to the width T . Shortening the width of pulse would increase the range resolution. The ideal ambiguity function along the zero-Doppler axis, i.e., the autocorrelation function, would be a delta function. On the other hand, for the radar pulse with imperfect autocorrelation function, a weak target located at the neighborhood of a strong target could be masked by the sidelobe of the ambiguity function centered around the strong target. Therefore, radar pulses that have good correlation properties not only help to improve the range accuracy but also help to distinguish closely located multiple targets.

As discussed above, the range resolution can be improved by using short pulses [7]. However, by transmitting short pulses, the average transmitted power would be reduced since the maximal transmit power of radar is limited and fixed. It is desirable to increase the pulse width, i.e., to increase the average transmitted power, or equivalently increase the average SNR, while maintaining good range resolution. This goal is achieved by employing the pulse compression techniques. Pulse compression allows us to increase the average transmitted power by transmitting a relative longer pulse, while obtain range resolution corresponding to a short pulse [7]. The pulse compression can be achieved either by linear frequency modulation (LFM) or phase coding [7], such as Barker sequences [8].

In the additive white Gaussian noise, the output of the matched filter maximizes the SNR. To improve the orthogonality of the signals at the receiver end, it is possible to use mismatched filters [9]. Mismatched filters have been introduced in radars to reduce the sidelobes at the cost of a reduce of the SNR at the filter output [9].

1.2 MIMO Radars

Inspired by the idea of MIMO communications, MIMO radar systems have received significant attention during the past decade [10] [11]. In phased-array radar, multiple

transmit antennas emit a scaled version of a single waveform [12], i.e., coherent waveforms are transmitted to illuminate certain directions of interests with maximum target illuminating power [13]. However, scanning a wide sector requires significant amount of time, and during each scan cycle only a few pulses reach the targets. Unlike phased-array radars, MIMO radars transmit independent waveforms through their transmit antennas. Therefore, the beampattern of MIMO radars is omnidirectional, which indicates that the whole search space would be illuminated simultaneously during the search mode. At the receivers, these waveforms can be separated with the help of a matched filter bank constructed based on the transmit waveforms. MIMO radars are configured in two scenarios, i.e., widely separated MIMO radars [10] [14] [15] [16] [17] and colocated MIMO radars [18] [19].

1.2.1 Widely Separated MIMO Radars

In widely separated MIMO radar systems, the transmit antennas are deployed far apart from each other as compared with their distance to the target [10] [14] [15] [16] [17]. The transmit antennas emit independent waveforms from decorrelated transmitters. These waveforms are transmitted through independent paths to the target. As a result, each target return contains independent information about the target. This configuration enables the radar to view the target from different directions simultaneously. Diversity gain could be achieved by combining these independent target returns and MIMO radar systems could reduce the radar cross section (RCS) scintillation of the target.

Non-coherent and coherent methods can be applied to process the radar observations. Only the signal envelope information is utilized in non-coherent methods and therefore only time synchronization between transmit and receive antennas is required. On the other hand, coherent methods utilize both envelop and phase information, which requires both time and phase synchronization.

1.2.2 Colocated MIMO Radars

In the colocated MIMO radar systems, the antennas are deployed close to each other and thus the RCS corresponding to different paths could be viewed as the same [18] [19].

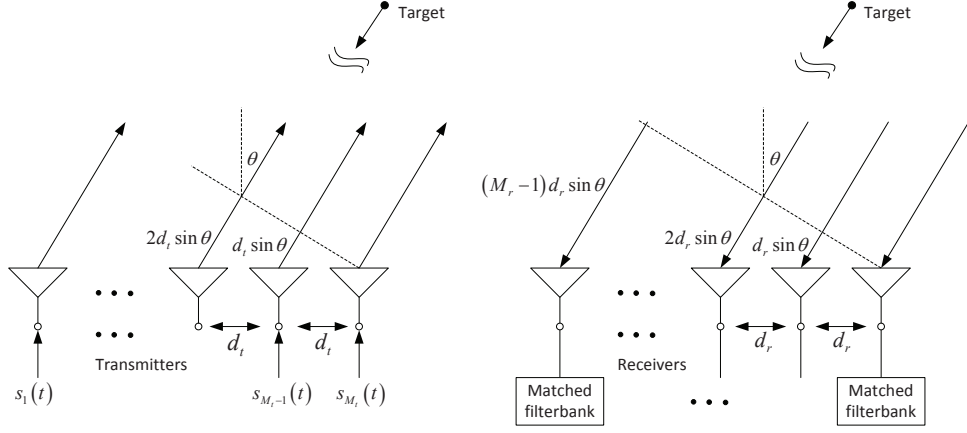


Figure 1.4: The colocated MIMO radar system is equipped with M_t transmit antennas and M_r receive antennas. Both transmit and receive arrays follow the ULA model.

Suppose the number of transmit and receive antennas are M_t and M_r . By transmitting independent waveforms at the transmit end and using matched filter bank to separate them at each receiver, the MIMO radar system is equivalent to a long virtual array with $M_t M_r$ elements [19]. This suggests that the colocated MIMO radars can provide a much higher degree of freedom (DOF) with only a small number of transmit and receive antennas. This enables the colocated MIMO radars to achieve superior spatial resolution as compared with traditional radar systems.

Let us consider a stationary point target located in the far field of the antennas; the target's direction of departure (DOD) and direction of arrival angles are the same, denoted by θ . The transmit and receive antennas are configured in a uniform linear array (ULA), as shown in Fig. 1.4. Let d_t and d_r denote the inter-spacing between the transmit and receive array, respectively. Let $s_1(n)$ be the discrete time baseband signal emitted from the m -th transmit antenna. Let $r_l(n)$ denote the received signal at the l -th receive antenna and let

$$\mathbf{r}(n) = [r_1(n), \dots, r_{M_r}(n)]^T, n = 1, \dots, N, \quad (1.20)$$

where N denotes the number of samples during the transmit pulse. Let

$$\mathbf{a}(\theta) = [e^{-j2\pi f\tau_1(\theta)}, e^{-j2\pi f\tau_2(\theta)}, \dots, e^{-j2\pi f\tau_{M_t}(\theta)}]^T, \quad (1.21)$$

$$\mathbf{b}(\theta) = [e^{-j2\pi f\tilde{\tau}_1(\theta)}, e^{-j2\pi f\tilde{\tau}_2(\theta)}, \dots, e^{-j2\pi f\tilde{\tau}_{M_r}(\theta)}]^T, \quad (1.22)$$

where $\tau_m(\theta)$ is the propagation time of the transmitted signal between the m -th transmit antenna and the target located at θ , and $\tilde{\tau}_l(\theta)$ is the time for the corresponding reflected signal to reach the l -th receive antenna. Under the point target assumption, the noise free received data vector can be written as [19]

$$\mathbf{r}(n) = \sum_{k=1}^K \beta_k \mathbf{b}(\theta_k) \mathbf{a}^T(\theta_k) \mathbf{s}(n), \quad (1.23)$$

where $\mathbf{s}(n) = [s_1(n), \dots, s_{M_r}(n)]^T$. If the transmitted waveforms from the M_t antennas are linearly independent of each other, i.e.,

$$\text{rank}\{\mathbf{s}(1), \dots, \mathbf{s}(N)\} = M_t, \quad (1.24)$$

the identifiability equation

$$\sum_{k=1}^K \tilde{\beta}_k \mathbf{b}(\tilde{\theta}_k) \mathbf{a}^T(\tilde{\theta}_k) \mathbf{s}(n) = \sum_{k=1}^K \beta_k \mathbf{b}(\theta_k) \mathbf{a}^T(\theta_k) \mathbf{s}(n) \quad (1.25)$$

would be equivalent to [19]

$$\sum_{k=1}^K \tilde{\beta}_k \mathbf{b}(\tilde{\theta}_k) \mathbf{a}^T(\tilde{\theta}_k) = \sum_{k=1}^K \beta_k \mathbf{b}(\theta_k) \mathbf{a}^T(\theta_k), \quad (1.26)$$

or $\tilde{\mathbf{A}}\tilde{\boldsymbol{\beta}} = \mathbf{A}\boldsymbol{\beta}$, where

$$\tilde{\mathbf{A}} = [\mathbf{a}(\tilde{\theta}_1) \otimes \mathbf{b}(\tilde{\theta}_1), \dots, \mathbf{a}(\tilde{\theta}_K) \otimes \mathbf{b}(\tilde{\theta}_K)] \quad (1.27)$$

$$\mathbf{A} = [\mathbf{a}(\theta_1) \otimes \mathbf{b}(\theta_1), \dots, \mathbf{a}(\theta_K) \otimes \mathbf{b}(\theta_K)] \quad (1.28)$$

and $\tilde{\boldsymbol{\beta}} = [\tilde{\beta}_1, \dots, \tilde{\beta}_K]^T$, $\boldsymbol{\beta} = [\beta_1, \dots, \beta_K]^T$. When $\frac{d_t}{d_r} = M_r$, the vector $\mathbf{a}(\theta) \otimes \mathbf{b}(\theta)$ would contain $M_t M_r$ distinct elements, i.e.,

$$\mathbf{a}(\theta) \otimes \mathbf{b}(\theta) = [1, e^{-j\omega}, \dots, e^{-j(M_t M_r - 1)\omega}]^T, \quad (1.29)$$

where $\omega = 2\pi f \tau(\theta)$ and $\tau(\theta)$ is the inter-element delay difference in the receive ULA. In other words, a virtual array of $M_t M_r$ elements is achieved with only $M_t + M_r$ elements [19]. As a result, colocated MIMO radars is capable to achieve superior angular resolution as compared to phased-array radars with the same number of antennas.

For the matched filter to work at the receive antennas, the transmit waveforms must be orthogonal. However, due to the time delay and Doppler shift, the waveforms orthogonality might be lost at the receiver end. The mismatched filter has been proposed to address this issue [20].

1.2.3 MIMO Radar Ambiguity Function

The MIMO radar ambiguity function under the arbitrary array configurations are given in [21]. Here, we state the ambiguity function of MIMO radars under ULA configurations. We consider a MIMO radar system with ULA transmit/receive arrays of inter-element spacing d_t and d_r , respectively. Let $\gamma \triangleq d_t/d_r$. For a single target at direction θ , we define the normalized spatial frequency as

$$\alpha = \frac{d_r}{\lambda} \sin \theta. \quad (1.30)$$

The demodulated receive signal in the l -th receive antenna is proportional to

$$r_l^{\tau, \nu, f}(t) \approx \sum_{m=0}^{M_t-1} s_m(t - \tau) e^{j2\pi\nu t} e^{j2\pi f(\gamma m + l)}. \quad (1.31)$$

At the receiver, the target information is captured with a matched filter with the parameters (τ', ν', α') . The matched filter output is [22]

$$\begin{aligned} & \sum_{l=0}^{M_r-1} \int_{-\infty}^{\infty} r_l^{\tau, \nu, \alpha}(t) \cdot \left(r_l^{\tau', \nu', \alpha'}\right)^*(t) dt \\ &= \left(\sum_{l=0}^{M_r-1} e^{j2\pi(\alpha - \alpha')l} \right) \left(\sum_{m=0}^{M_t-1} \sum_{m'=0}^{M_t-1} \int_{-\infty}^{\infty} s_m(t - \tau) s_{m'}^*(t - \tau') e^{j2\pi(\nu - \nu')t} dt \cdot e^{j2\pi(\alpha m - \alpha' m')\gamma} \right) \end{aligned} \quad (1.32)$$

The MIMO radar ambiguity function under ULA configurations is defined as [22]

$$A(\tau, \nu, \alpha, \alpha') \triangleq \sum_{m=0}^{M_t-1} \sum_{m'=0}^{M_t-1} A_{m, m'}(\tau, \nu) e^{j2\pi(\alpha m - \alpha' m')\gamma}, \quad (1.33)$$

where

$$A_{m, m'}(\tau, \nu) \triangleq \int_{-\infty}^{\infty} s_m(t) s_{m'}^*(t + \tau) e^{j2\pi\nu t} dt \quad (1.34)$$

is called the cross ambiguity function.

In the pulse MIMO radars, the transmitted signal is

$$s_m(t) = \sum_{q=0}^{Q-1} \phi_m(t - qT_{PRI}). \quad (1.35)$$

The duration of pulse $\phi_m(t)$, i.e., T_ϕ , satisfies $T_\phi \ll T_{PRI}$. Assume that $T_\phi \nu \approx 0$. The ambiguity function for the pulse MIMO radars is [22]

$$A(\tau, \nu, \alpha, \alpha') = \sum_{m=0}^{M_t-1} \sum_{m'=0}^{M_t-1} R_{m,m'}^\phi(\tau) e^{j2\pi(\alpha m - \alpha' m')\gamma} \cdot \sum_{q=0}^{Q-1} e^{j2\pi\nu q T_{PRI}}, \quad (1.36)$$

for $|\tau| < T_{PRI} - T_\phi$. Here, $R_{m,m'}^\phi(\tau)$ is the cross correlation between $\phi_m(t)$ and $\phi_{m'}(t)$, defined as

$$R_{m,m'}^\phi(\tau) \triangleq \int_0^{T_\phi} \phi_m(t) \phi_{m'}^*(t + \tau) dt. \quad (1.37)$$

The pulse MIMO radar ambiguity function shown in (1.36) indicates that the Doppler processing is separable from the correlation function under the assumption of slow moving targets and short pulses [22]. Consequently, the choice of the waveforms $\{\phi_m(t)\}$ does not affect the Doppler resolution.

1.3 Compressive Sensing (CS) Based MIMO Radars

Compressive sensing (CS) [23] [24] [25], is a relatively recent development for finding sparse solutions to under-determined linear systems. CS theory states that a K -sparse signal \mathbf{x} of length N can be recovered exactly with high probability from $\mathcal{O}(K \log N)$ linearly compressed measurements. CS finds applications in MIMO radar systems [26] [27] [28] [29] [30]. In CS based MIMO radars, the target parameters are estimated by exploiting the sparsity of targets in the angle, Doppler and range space, referred to as the *target space*. It is assumed that the targets are sparsely located in the target space. Each receive antenna performs a sub-Nyquist sampling and forwards the samples to the fusion center, where the target is estimated via sparse signal recovery techniques, such as the Dantzig selector [29], orthogonal matched pursuing (OMP) [31].

1.3.1 Compressive Sensing

A K -sparse signal \mathbf{x} of length N can be represented as

$$\mathbf{x} = \Psi \mathbf{s}, \quad (1.38)$$

where Ψ is the $N \times N$ basis matrix spanning the space under which \mathbf{x} is sparse; \mathbf{s} is the coefficient vector whose K elements have large magnitude and the magnitudes of the remaining elements are negligible or zeros. If $K \ll N$, the signal \mathbf{x} is sparse and compressible.

Suppose we collect linear measurements of \mathbf{x} as

$$\mathbf{y} = \Phi \mathbf{x} = \Theta \mathbf{s}, \quad (1.39)$$

where Φ is an $M \times N$ matrix, referred as the measurement matrix and $\Theta = \Phi \Psi$, referred as the sensing matrix in the CS society, respectively.

To recover the signal \mathbf{x} with samples \mathbf{y} of length M and $M \ll N$, it requires to solve an under-determined linear system. To find the sparsest solution of \mathbf{x} , an optimization problem is formulated as

$$\begin{aligned} \min_{\mathbf{s}} \quad & \|\mathbf{s}\|_0 \\ \text{s.t.} \quad & \mathbf{y} = \Theta \mathbf{s}, \end{aligned} \quad (1.40)$$

where $\|\cdot\|_0$ denotes the ℓ_0 norm which is number of nonzero elements of the argument. The ℓ_0 norm optimization problem is not convex. A convex relaxation is to solve the ℓ_1 norm optimization problem, defined as

$$\begin{aligned} \min_{\mathbf{s}} \quad & \|\mathbf{s}\|_1 \\ \text{s.t.} \quad & \mathbf{y} = \Theta \mathbf{s}, \end{aligned} \quad (1.41)$$

where $\|\cdot\|_1$ denotes the ℓ_1 norm which is the sum of absolute value of all the elements in the argument. Consequently, the optimization problem (1.41) a linear programming (LP) optimization problem. Under noisy measurement cases, the recovery problem can be formulated as a second-order cone programming (SOCP) problem. When the size of the optimization is relatively large, greedy algorithms, such as orthogonal matched pursuing [31], are also used to find the support of the sparse vector.

1.3.2 CS-Based Colocated MIMO Radars

Let us consider a MIMO radar system consisting of M_t TX antennas and N_r RX antennas. In the far field of the antennas there are K targets that need to be estimated.

For simplicity, we will assume that the targets are not moving, thus the only parameters that needs to be estimated are the target azimuth angles θ_k , $k = 1, \dots, K$. The results can be easily extended to the case of moving targets.

Let us assume that the transmit and receive antennas are closely spaced, with the i -th transmit/receive antenna placed at location $(r_i^t, \alpha_i^t)/(r_i^r, \alpha_i^r)$ (in polar coordinates). Let L denote the number of T_s -spaced samples of the transmitted waveforms. The effect of the compressive receiver in Fig. 1 of [29] is equivalent to pre-multiplying by matrix Φ a T_s -sampled version of the received signal. The size of Φ is $M \times L$.

Under the narrowband transmitted signal assumption, the received baseband signal at the l -th receive antenna can be approximated by

$$\mathbf{r}_l \approx \sum_{k=1}^K \beta_k e^{j \frac{2\pi f}{c} \eta_l^r(\theta_k)} \Phi \mathbf{X} \mathbf{v}_t(\theta_k) + \Phi \mathbf{n}_l, \quad (1.42)$$

where \mathbf{X} is an $L \times M_t$ matrix that contains the transmit waveforms as its columns; β_k is the reflection coefficient of the k -th target;

$$\mathbf{v}_t(\theta_k) = [e^{j \frac{2\pi f}{c} \eta_1^t(\theta_k)}, \dots, e^{j \frac{2\pi f}{c} \eta_{M_t}^t(\theta_k)}]^T \quad (1.43)$$

is the transmit steering vector associated with angle θ_k ; and $\eta_i^{t/r}(\theta_k) = r_i^{t/r} \cos(\theta_k - \alpha_i^{t/r})$; \mathbf{n}_l is the interference at the l -th receiver, arising due to the interference and thermal noise.

Let us discretize the angle space into N discrete angles $[a_1, \dots, a_N]$. The discretization step is small enough so that each target falls on some angle grid point. Then, (1.42) can be rewritten as

$$\mathbf{r}_l = \underbrace{\Phi \Psi_l}_{\Theta_l} \mathbf{s} + \Phi \mathbf{n}_l, \quad (1.44)$$

where $\Psi_l = [e^{j \frac{2\pi f}{c} \eta_l^r(a_1)} \mathbf{X} \mathbf{v}_t(a_1), \dots, e^{j \frac{2\pi f}{c} \eta_l^r(a_N)} \mathbf{X} \mathbf{v}_t(a_N)]$; $\mathbf{s} = [s_1, \dots, s_N]^T$, with s_n being zero if there is no target at angle a_n , otherwise being equal to the reflection coefficient of the target at that angle.

By stacking the received data from each antenna into a long vector, we form \mathbf{y} , for

which it holds

$$\begin{aligned} \mathbf{y} &= [\mathbf{r}_1^T, \dots, \mathbf{r}_{N_r}^T]^T \\ &= \underbrace{\left[(\Phi \Psi_1)^T, \dots, (\Phi \Psi_{N_r})^T \right]^T}_{\Theta} \mathbf{s} + \underbrace{\left[(\Phi \mathbf{n}_1)^T, \dots, (\Phi \mathbf{n}_{N_r})^T \right]^T}_{\mathbf{z}}. \end{aligned} \quad (1.45)$$

According to the CS formulation, Θ is the sensing matrix and Ψ_l is the basis matrix for the l -th antenna. If the number of targets is small as compared to N , then \mathbf{s} is a sparse vector, with the locations of its non-zero elements providing information on the target angles. A variety of CS methods can be applied to the recovery of \mathbf{s} , e.g., basis pursuit [32], matching pursuit [31] and Lasso methods [33].

1.3.3 Basis Incoherence in CS-Based MIMO Radars

The widely used analysis tool for sparse signal processing is the uniform uncertainty principle (UUP) defined in [34] [35] [36]. UUP states that the $MN_r \times N$ sensing matrix Θ satisfies the “restricted isometry properties (RIP)”. Let $\Theta_T, T \subset \{1, \dots, MN_r\}$ be the matrix constructed from the columns of Θ according to the indices T . The S -restricted isometry constant δ_S of Θ defined in [36] is the smallest number such that

$$(1 - \delta_S) \|\mathbf{c}\|_2^2 \leq \|\Theta_T \mathbf{c}\|_2^2 \leq (1 + \delta_S) \|\mathbf{c}\|_2^2 \quad (1.46)$$

for all subsets T with $|T| \leq S$ and coefficient sequence $(\mathbf{c}_j)_{j \in T}$. The restricted isometry property requires that every set of columns with cardinality less than S is approximately orthonormal. It is shown in [36] that if S satisfies $\delta_S + \delta_{2S} + \delta_{3S} < 1$, then solving (1.41) could recover any sparse signal \mathbf{s} with supporting size $|T_0| \leq S$.

Verifying that a given matrix meets the RIP condition is not computationally feasible [31]. The other commonly used framework is the mutual incoherence properties (MIP) [37]. The mutual incoherence is defined as

$$\mu = \max_{i \neq j} |\langle \Theta_i, \Theta_j \rangle|. \quad (1.47)$$

To satisfy the MIP, the parameter μ needs to be small [31]. Suppose the vector \mathbf{s} is K -sparse, i.e., $|\text{supp}(\mathbf{s})| \leq K$. Then, it was shown in [38] that $\mu < \frac{1}{2K-1}$ is a sufficient condition to recover K -sparse signal correctly using orthogonal matching pursuit (OMP)

in the noiseless case. In the presence of noise, it was shown in [31] that under the MIP $\mu < \frac{1}{2K-1}$ and a minimum magnitude condition on the nonzero element in the vector \mathbf{s} , the sparse signal can be recovered exactly using the OMP in bounded noise case, and with a high probability in Gaussian noise case.

One can see that, to make μ small, the columns of the sensing matrix Θ need to be as orthogonal as possible. Based on this conclusion, in Chapter 2 we propose power allocation and waveform design algorithms to improve the performance of CS-based MIMO radars by minimizing the difference between the Gram of sensing matrix and an identity matrix.

1.3.4 Basis Mismatch in CS-Based MIMO Radars

When applying CS in MIMO radars to do target estimation, the target space (DOA, range and velocity) needs to be discretized into a fine grid, based on which the basis matrix is constructed. Other information that is used to construct the basis matrix includes transmit waveforms, radar antenna positions (see equation (1.44)). However, the performance of CS-based MIMO radars degrades when there are errors in the constructed basis matrix, a case also known as basis mismatch [39] [40].

The first possible scenario resulting in basis mismatch can be due to the small errors in the antenna positions. For example, consider the scenario of a radar formed by antennas placed on the backpacks of soldiers. The measurement data from each receive antenna is forwarded to a fusion center via a wireless link. The positioning of each soldier may contain small errors, especially when the soldiers are moving. Figure 1.5 (a) shows the random positions of $M_t = 20$ TX and $N_r = 6$ RX antennas in 2D Cartesian coordinates. The carrier frequency is set to $f = 0.1\text{GHz}$, and thus the wavelength equals $\lambda = 3\text{m}$. In each dimension, the position error of each antenna is set to a random number in $\lambda/5 = 0.6\text{m}$ with a random sign. There are $K = 3$ targets located in a far field of antennas with DOAs as $[1.1^\circ, 3.4^\circ, 4.5^\circ]$. The target reflection coefficients are all set to 1. Hadamard waveforms with $M_t = 20$ and $L = 32$ are transmitted. To apply the CS theory to estimate the target DOAs, we discretize the angle space $[0^\circ, 5^\circ]$ on a grid with spacing of 0.1° . At each receive antenna, a

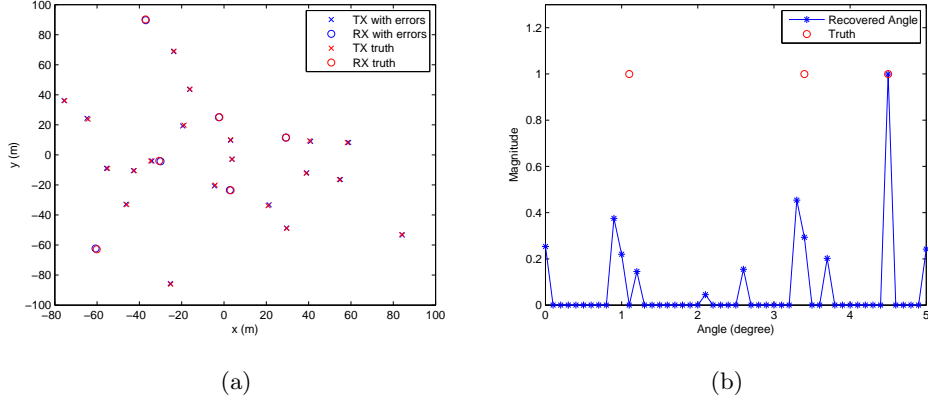


Figure 1.5: The basis mismatch due to antenna position errors: (a) The true and estimated positions of $M_t = 20$ TX and $N_r = 6$ RX antennas; (b) The performance of CS when there is basis mismatch due to antenna position errors within $\frac{\lambda}{5}$.

measurement matrix $\Phi \in \mathbb{C}^{M \times L}$ with Gaussian elements is used to pre-multiply the samples. Here, $M = \text{round}(p * L)$ and $p = 0.7$ is the portion of selected samples. The received data is corrupted with thermal noise and the signal-to-noise ratio (SNR) is set to 0dB. Figure 1.5 (b) shows the CS recovery performance using the Dantzig selector. One can see that the true target directions cannot be found via CS, i.e., a small basis mismatch error could yield big recovery errors, as shown in [39]. The antenna position errors would results in phase error in radar systems. Our simulations show that when the antenna position errors are larger than $\lambda/10$, there would be large target estimation error with CS.

Another basis mismatch scenario is when the targets fall between grid points. Continuing on setting in Fig. 1.5, let us assume that the antenna positions are perfectly known. There are $K = 3$ targets located in a far field of antennas with DOAs as $[1.15^\circ, 3.45^\circ, 4.55^\circ]$. Figure 1.6 (a) shows the CS performance using the OMP algorithm with grid size 0.1° . It can be found that the recovered target angles are the grid points which are the closest to the true angles. We next increase the grid density by setting the grid size to 0.01° . Consequently, the targets are on the grid under the new discretization. However, Figure 1.6 (b) shows that the recovered target angles do not coincide with the truth as one would expect. This is because the grid density cannot be

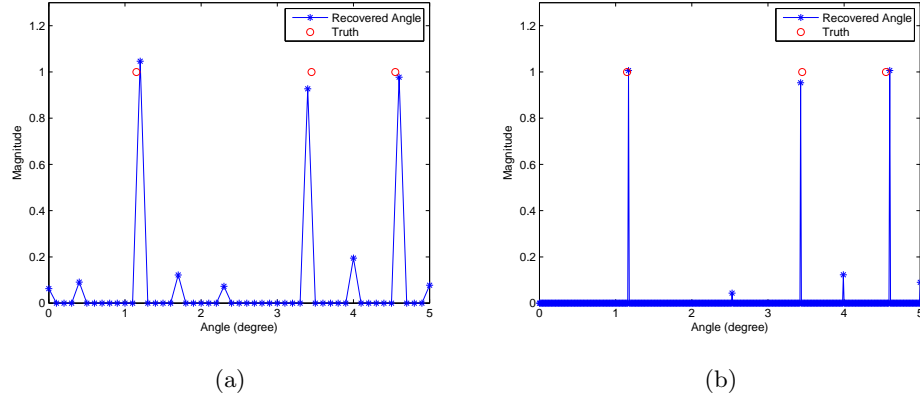


Figure 1.6: The CS recovery performance in the off-grid target scenario using orthogonal matching pursuit: (a) grid size as 0.1°; (b) grid size as 0.01°.

increased without decreasing the incoherence of the basis matrix of CS, which results in worse CS performance.

In this thesis, we propose a novel MIMO radar approach using matrix completion (MC), termed as MIMO-MC, which exploits the low-rank structure of the data matrix arising in MIMO radars. The new approach shares the same advantage of CS-based MIMO radars, i.e., significant reduction of samples required for high resolution target estimation, but does not require discretization of the target space. In the following, we give a brief introduction of matrix completion theory.

1.4 Matrix Completion Theory

Matrix completion is of interest in cases in which we are constrained to observe only a subset of the entries of an $n_1 \times n_2$ matrix, because for example, the cost of collecting all entries of a high dimensional matrix is high. If a matrix is low rank and satisfies certain conditions [41], it can be recovered *exactly* based on observations of a small number of its randomly selected entries. In this section we provide a brief overview of the problem of recovering a rank r matrix $\mathbf{M} \in \mathbb{C}^{n_1 \times n_2}$ based on partial knowledge of its entries using the method of [41] [42] [43].

Let us define the observation operation $\mathbf{Y} = \mathcal{P}_\Omega(\mathbf{M})$ as

$$[\mathbf{Y}]_{ij} = \begin{cases} [\mathbf{M}]_{ij}, & (i, j) \in \Omega \\ 0, & \text{otherwise} \end{cases} \quad (1.48)$$

where Ω is the set of indices of observed entries with cardinality m . According to [42], when \mathbf{M} is low-rank and meets certain conditions (see **(A0)** and **(A1)**, later in this section), \mathbf{M} can be estimated by solving a nuclear norm optimization problem

$$\begin{aligned} \min \quad & \|\mathbf{X}\|_* \\ \text{s.t.} \quad & \mathcal{P}_\Omega(\mathbf{X}) = \mathcal{P}_\Omega(\mathbf{M}) \end{aligned} \quad (1.49)$$

where $\|\cdot\|_*$ denotes the nuclear norm, i.e., the sum of singular values of \mathbf{X} .

In practice, the observations are typically corrupted by noise, i.e., $[\mathbf{Y}]_{ij} = [\mathbf{M}]_{ij} + [\mathbf{E}]_{ij}$, $(i, j) \in \Omega$, where, $[\mathbf{E}]_{ij}$ represents noise. In that case, it holds that $\mathcal{P}_\Omega(\mathbf{Y}) = \mathcal{P}_\Omega(\mathbf{M}) + \mathcal{P}_\Omega(\mathbf{E})$, and the completion of \mathbf{M} is done by solving the following optimization problem [43]

$$\begin{aligned} \min \quad & \|\mathbf{X}\|_* \\ \text{s.t.} \quad & \|\mathcal{P}_\Omega(\mathbf{X} - \mathbf{Y})\|_F \leq \delta. \end{aligned} \quad (1.50)$$

Assuming that the noise is zero-mean, white, $\delta > 0$ is a parameter related to the noise variance, σ^2 , as $\delta^2 = (m + \sqrt{8m})\sigma^2$ [43].

The conditions for successful matrix completion involve the notion of coherence, which is defined next [41].

Definition 1. Let U be a subspace of \mathbb{C}^{n_1} of dimension r that is spanned by the set of orthogonal vectors $\{\mathbf{u}_i \in \mathbb{C}^{n_1}\}_{i=1,\dots,r}$, P_U be the orthogonal projection onto U , i.e., $P_U = \sum_{1 \leq i \leq r} \mathbf{u}_i \mathbf{u}_i^H$, and \mathbf{e}_i be the standard basis vector whose i th element is 1. The coherence of U is defined as

$$\mu(U) = \frac{n_1}{r} \max_{1 \leq i \leq n_1} \|P_U \mathbf{e}_i\|^2 \in \left[1, \frac{n_1}{r}\right]. \quad (1.51)$$

Let the compact singular value decomposition (SVD) of \mathbf{M} be $\mathbf{M} = \sum_{k=1}^r \rho_k \mathbf{u}_k \mathbf{v}_k^H$, where $\rho_k, k = 1, \dots, r$ are the singular values, and \mathbf{u}_k and \mathbf{v}_k the corresponding left and

right singular vectors, respectively. Let U, V be the subspaces spanned by \mathbf{u}_k and \mathbf{v}_k , respectively. Matrix \mathbf{M} has coherence with parameters μ_0 and μ_1 if

(A0) $\max(\mu(U), \mu(V)) \leq \mu_0$ for some positive μ_0 .

(A1) The maximum element of the $n_1 \times n_2$ matrix $\sum_{1 \leq i \leq r} \mathbf{u}_i \mathbf{v}_i^H$ is bounded by $\mu_1 \sqrt{r/(n_1 n_2)}$ in absolute value, for some positive μ_1 .

In fact, it was shown in [41] that if (A0) holds, then (A1) also holds with $\mu_1 \leq \mu_0 \sqrt{r}$.

Now, suppose that matrix $\mathbf{M} \in \mathbb{C}^{n_1 \times n_2}$ satisfies (A0) and (A1). The following lemma gives a probabilistic bound for the number of entries, m , needed to estimate \mathbf{M} .

Theorem 1. [41] *Suppose that we observe m entries of the rank- r matrix $M \in \mathbb{C}^{n_1 \times n_2}$, with matrix coordinates sampled uniformly at random. Let $n = \max\{n_1, n_2\}$. There exist constants C and c such that if*

$$m \geq C \max\left\{\mu_1^2, \mu_0^{1/2} \mu_1, \mu_0 n^{1/4}\right\} n r \beta \log n$$

for some $\beta > 2$, the minimizer to the program of (1.49) is unique and equal to \mathbf{M} with probability at least $1 - cn^{-\beta}$.

For $r \leq \mu_0^{-1} n^{1/5}$ the bound can be improved to

$$m \geq C \mu_0 n^{6/5} r \beta \log n,$$

without affecting the probability of success.

Theorem 1 implies that the lower the coherence parameter μ_0 , the fewer entries of \mathbf{M} are required to estimate \mathbf{M} . The smallest possible value for μ_0 is 1.

Further, [43] establishes that, when observations are corrupted with white zero-mean Gaussian noise with variance σ^2 , when solving (1.50), the recovery error is bounded as

$$\|\mathbf{M} - \hat{\mathbf{M}}\|_F \leq 4 \sqrt{\frac{1}{p} (2 + p) \min(n_1, n_2) \delta} + 2\delta, \quad (1.52)$$

where $p = \frac{m}{n_1 n_2}$ is the fraction of observed entries, and $\delta^2 = (m + \sqrt{8m}) \sigma^2$.

There are several MC techniques in the literature [41–46]. For example, in [41–43], recovery can be performed by solving a nuclear norm optimization problem, which basically finds the matrix with the smallest nuclear norm out of all possible matrices

that fit the observed entries. Other matrix completion techniques are based on non-convex optimization using matrix manifolds, such as Grassmann manifold [44, 45], and Riemann manifolds [46].

1.5 Contributions of the Dissertation

1.5.1 Performance Optimization and Clutter Suppression in CS-Based MIMO Radars

By exploring sparsity in the target space, CS-based MIMO radar systems achieve either the same localization performance as traditional methods but with significantly fewer measurements, or significantly improved performance with the same number of measurements. The recovery requires that the product of the measurement matrix and the sparsifying basis matrix, referred to as the sensing matrix, satisfies the uniform uncertainty principle (UUP); in other words, the sensing matrix exhibits low correlation between its columns. To improve the CS-based MIMO radars performance, the columns of the sensing matrix should be as orthogonal as possible. We propose algorithms on power allocation among transmit antennas to improve the orthogonality of the columns of the sensing matrix arising in both colocated and widely separated MIMO radars. Furthermore, we propose waveform design algorithm to improve the performance CS-based colocated MIMO radars by minimizing the coherence between the target returns from different search cells, or equivalently, the coherence of the columns of the sensing matrix.

In the presence of clutter, the sparsity assumption does not hold. Under the assumption that the clutter covariance is known, we propose an approach to suppress clutter in the context of CS-based MIMO by applying the Capon beamforming at the fusion center on compressively obtained data, which are forwarded by the receive antennas. Subsequently, the target is estimated using CS theory, by exploiting the sparsity of the beamformed signals. Furthermore, we propose algorithm of power allocation among transmit antennas to further improve the CS performance in the clutter scenario.

This work has been published in

- Y. Yu, S. Sun, R. N. Madan, and A. P. Petropulu, “Power allocation and waveform design for the compressive sensing based MIMO radars,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 2, pp. 898-909, 2014.
- Y. Yu, S. Sun and A. P. Petropulu, “A Capon beamforming method for clutter suppression in colocated compressive sensing based MIMO radars,” *SPIE Defense, Security, and Sensing*, Baltimore, MD, April 29-May 3, 2013.

1.5.2 MIMO-MC Radar: A MIMO Radar Approach Based on Matrix Completion

We show that in a typical MIMO radar scenario with large scale transmit and receive arrays, the data matrix constructed at the fusion center based on the measurements of all antennas, is low-rank, and thus can be recovered based on knowledge of a small subset of its entries via matrix completion techniques. The recovered data matrix, in conjunction with standard array processing schemes leads to target detection and parameter estimation. Leveraging the low-rank property of that matrix, we propose a new MIMO radar approach, termed, MIMO-MC radar, in which each receive node either performs matched filtering with a small number of randomly selected dictionary waveforms, or obtains sub-Nyquist samples of the target returns at random sampling instants, and forwards the results to a fusion center. Based on the received samples, and with knowledge of the sampling scheme, the fusion center partially fills the data matrix and subsequently applies MC techniques to estimate the full matrix. MIMO-MC radars share the advantages of the CS-based MIMO radars, i.e., high resolution with reduced amounts of data, but do not require grid discretization. Thus, the proposed approach does not suffer from the target off-grid issues of CS-based MIMO radars.

This work has been published in

- S. Sun, W. U. Bajwa, and A. P. Petropulu, “MIMO-MC radar: A MIMO radar approach based on matrix completion,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, no. 3, pp. 1839-1852, 2015.
- S. Sun, A. P. Petropulu, and W. U. Bajwa, “Target estimation in colocated MIMO

radar via matrix completion,” in Proc. of *IEEE 38th International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Vancouver, Canada, May 2013.

- S. Sun, A. P. Petropulu, and W. U. Bajwa, “High-resolution networked MIMO radar based on sub-Nyquist observations,” in *Signal Processing with Adaptive Sparse Structured Representations Workshop (SPARS)*, EPFL, Lausanne, Switzerland, July 8-11, 2013.
- D. S. Kalogerias, S. Sun, and A. P. Petropulu, “Sparse sensing in colocated MIMO radar: A matrix completion approach,” in Proc. of *IEEE 13th International Symposium on Signal Processing and Information Technology (ISSPIT)*, Athens, Greece, Dec. 12-15, 2013.

1.5.3 Coherence Analysis and Optimal Waveform Design in MIMO Radars with Matrix Completion

We first study the applicability of matrix completion (MC) theory on the data matrix arising in colocated MIMO radars using uniform linear arrays. We show that the data matrix coherence, and consequently the performance of MC, is directly related to the transmit waveforms. Among orthogonal waveforms, the optimum choices are those for which, any snapshot across the transmit array has a flat spectrum.

We then propose waveform design methods for MIMO-MC radars. The problem of waveform design is formulated as an optimization problem on the complex Stiefel manifold, and is solved via the modified steepest descent method, or the modified Newton algorithm with nonmonotone line search.

Although the optimal waveforms are designed for the case of targets falling in the same range bin, we conduct the sensitivity analysis of the proposed optimal waveform to assess the performance degradation when the targets fall in different range bins. Specially, we show that when transmit waveforms have ideal correlation properties, the coherence upper bound of the data matrix in MIMO-MC radar is minimal.

This work has been published in

- S. Sun and A. P. Petropulu, “Waveform design for MIMO radars with matrix completion,” *IEEE Journal of Selected Topics in Signal Processing*, to appear in the December Issue, 2015.
- S. Sun and A. P. Petropulu, “On waveform conditions and range compression in MIMO radars using matrix completion,” in Proc. of *49th Annual Asilomar Conference on Signals, Systems, and Computers (Asilomar)*, Pacific Grove, CA, Nov. 8-11, 2015.
- S. Sun and A. P. Petropulu, “On waveform design for MIMO radar with matrix completion,” in Proc. of *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Information Processing for Big Data Symposium, Atlanta, GA, Dec. 3-5, 2014.
- S. Sun and A. P. Petropulu, “On the applicability of matrix completion on MIMO radars,” in Proc. of *48th Annual Asilomar Conference on Signals, Systems, and Computers (Asilomar)*, Pacific Grove, CA, Nov. 2-5, 2014.

1.5.4 Transmit and Receive Beamforming in MIMO Radars with Matrix Completion

In the tracking mode of MIMO radars, the transmit antennas transmit correlated waveforms to illuminate certain directions. We propose a matrix completion based colocated MIMO radar (MIMO-MC) approach that employs transmit beamforming. Each receive antenna performs sub-Nyquist sampling of the target returns at uniformly random times. Based on the forwarded samples, the fusion center partially fills a matrix, recovers the Nyquist rate samples via matrix completion, and subsequently proceeds with target estimation via standard techniques. The performance of matrix completion depends on the matrix coherence. We derive the relations between transmit waveforms and matrix coherence. Specifically, it is shown that, for a rank-1 beamformer, the coherence is optimal, i.e., 1, if and only if the waveforms are unimodular. For a multi-rank beamformer, the coherence of the row space of the data matrix is optimal if the waveform power is constant across each snapshot.

We propose a receive beamforming scheme using matrix completion in a radar array with a relative large scale sensors. Beamforming methods rely on training data to estimate the covariance matrix of the interference pulse noise, which requires a large numbers of training snapshots to maintain good performance. In a distributed array, in which the array nodes are connected to a fusion center via a wireless link, the estimation of the covariance matrix would require the communication of large amounts of data, and thus would consume significant power. We propose a matrix completion based approach that enables good beamforming performance while requiring substantially fewer data to be transmitted to the fusion center. The main idea is based on the fact that when the number of signal and interference sources is much smaller than the number of array sensors, the training data matrix is low rank. Thus, the training data matrix can be recovered via matrix completion based on sub-Nyquist samples of the array sensors. Following the recovery of the training data matrix, and to cope with the errors introduced during the matrix completion process, we propose a robust optimization approach, which obtains the beamforming weight vector by optimizing the worst-case performance.

This work has been published in

- S. Sun and A. P. Petropulu, “On transmit beamforming in MIMO radar with matrix completion,” in *Proc. of IEEE 40th International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Brisbane, Australia, April 2015.
- S. Sun and A. P. Petropulu, “Robust beamforming via matrix completion,” in *Proc. of 47th Annual Conference on Information Sciences and Systems (CISS)*, Baltimore, MD, March 20-22, 2013.

1.6 Outline of the Dissertation

The dissertation is organized as follows.

In Chapter 2, based on the UUP condition introduced in compressive sensing theory, we propose power allocation and waveform design algorithms to minimize the coherence of the sensing matrix arising in both colocated and widely separated MIMO radars to

improve the target estimation performance of CS-based MIMO radars.

In Chapter 3, we propose a Capon beamforming approach to reject the clutter in the compressive sensing based MIMO radars.

In Chapter 4, we propose a new radar approach based on matrix completion, termed as MIMO-MC radar.

In Chapter 5, we conduct the coherence analysis of the data matrix arising in MIMO radars and the condition of optimal waveform in term of matrix coherence minimization is derived for MIMO radars using matrix completion.

In Chapter 6, we derive the optimal waveform conditions of the transmit beamforming in MIMO-MC radar, in terms of matrix coherence and subsequently matrix completion performance, for both rank-1 and multi-rank beamformers.

In Chapter 7, to overcome the large amount data collection requirement for estimation of the covariance matrix arising in the standard array beamforming processing, we propose a robust receive beamforming scheme for large scale array with matrix completion.

Finally, Chapter 8 contains conclusions and possible future research directions.

1.7 Notation

We use lower-case and upper-case letters in bold denote vectors and matrices, respectively. See Table 1.1 for other notations used in the dissertation.

Table 1.1: Notations

$\Re \{\cdot\}$:	the real part of $\{\cdot\}$
$\Im \{\cdot\}$:	the imaginary part of $\{\cdot\}$
$\mathbf{1}_L$:	the vector of length L with each element as 1
$\ \mathbf{a}\ _2$:	the Euclidean norm of a vector \mathbf{a}
\mathbf{A}^* :	the complex conjugate of a matrix \mathbf{A}
\mathbf{A}^T :	the transpose of a matrix \mathbf{A}
\mathbf{A}^H :	the conjugate transpose of a matrix \mathbf{A}
$\text{tr}(\mathbf{A})$:	the trace of a matrix \mathbf{A}
$\lambda_{\min}(\mathbf{A})$:	the minimal singular value of a matrix \mathbf{A}
$\ \mathbf{A}\ _F$:	the Frobenius norm of a matrix \mathbf{A}
$\ \mathbf{A}\ _*$:	the nuclear norm of a matrix \mathbf{A} , i.e., the sum of singular values
$\text{vec}(\mathbf{A})$:	the vectorization of a matrix \mathbf{A}
$\mathbf{A} \otimes \mathbf{B}$:	the Kronecker product of two matrices \mathbf{A} and \mathbf{B}
$\mathbf{A} \odot \mathbf{B}$:	the Hadamard product of two matrices \mathbf{A} and \mathbf{B}
\mathbf{I}_M :	the identity matrix of dimension $M \times M$

Chapter 2

Power Allocation and Waveform Design in Compressive Sensing Based MIMO Radars

Compressive sensing (CS) based multi-input multi-output (MIMO) radars systems, by exploring sparsity in the target space achieve either the same localization performance as traditional methods but with significantly fewer measurements, or significantly improved performance with the same number of measurements. This chapter investigates the performance gain of CS-MIMO radars, stemming from optimal power allocation among the transmit antennas, or optimal waveform design. In both cases, the optimization criterion is the minimization of the coherence between the target returns from different search cells, or equivalently, the coherence of the columns of the sensing matrix.

2.1 Introduction

Due to their potential to improve target detection, multiple-input and multiple-output (MIMO) radar systems have received considerable attention in recent years. Unlike traditional phased-array radar, a MIMO radar transmits multiple independent waveforms from its antennas. Depending on the transmit (TX) and receive (RX) antenna configuration, MIMO radar systems are classified as widely separated [17] and colocated [18]. The former view the target from multiple uncorrelated directions and thus achieve improved target detection performance benefiting from spatial diversity. The latter exploit waveform diversity to form a long virtual array, much longer than traditional radar systems with the same number of TX and RX antennas, and as a result enjoy superior spatial resolution.

Compressive sensing (CS) is a relatively recent development for finding sparse solutions to underdetermined linear systems [23] [24] [25]. CS theory states that a K -sparse signal \mathbf{x} of length N can be recovered exactly with high probability from $\mathcal{O}(K \log N)$ linearly compressed measurements. The recovery requires that the product of the measurement matrix and the sparsifying basis matrix, referred to as the sensing matrix, satisfies the uniform uncertainty principle (UUP) [25] [34] [35]; in other words, the sensing matrix exhibits low correlation between its columns.

CS in the context of MIMO radars has been studied in [47] [26] [29] [48] [49] [50]; [47] [26] [29] considered the application of CS to colocated MIMO radars with point targets, while [48] [49] [50] considered the application of CS to widely separated MIMO radars with extended targets. Both cases of CS-MIMO radars exploit the sparsity of targets in the target space and enable target estimation based on a small number of samples obtained at the RX antennas.

There are several techniques to further improve the detection performance of CS-MIMO radars. For example, significant gain in range resolution can be achieved by using a step-frequency approach during transmission [51] [52] [53]. We can also improve the detection performance by using a measurement matrix that minimizes the coherence of the sensing matrix and/or the signal-to-inference ratio (SIR) [54]. In this paper, we investigate power allocation and/or waveform optimization as means of improving the detection performance. Previous works [55] and [56] have discussed power allocation for traditional MIMO radars. In [55], the authors proposed to minimize the total transmitted power such that a predefined Cramer-Rao bound (CRB) is met, or to minimize the CRB by optimizing power allocation among the transmit radars for a given total power budget. A power allocation scheme was proposed in [56] for correlated MIMO radar systems in the presence of Rician scattering. Less power was allocated to the antennas that are correlated, or that suffer low line-of-sight reflectivity, so that the total available power was spread across uncorrelated branches and strong reflectors. In [49], an energy allocation scheme for CS-based widely separated MIMO radars was proposed that determines the transmit energies for the next set of transmit pulses based on the estimates of targets obtained from the previously received signals. The

goal of [49] is to maximize the minimum target returns so that the probability of missing weak targets is reduced. In this paper, since UUP indicates that the sensing matrix should be as orthogonal as possible in order to guarantee reliable performance, we allocate the power among the transmit antennas so that we minimize the difference between the Gram of the sensing matrix, i.e., $\Theta^H \Theta$ and an identity matrix. This improves the condition of the sensing matrix and enables performance enhancement. Based on the same optimization criterion, we also propose waveform design, aiming at improving the orthogonality of the sensing matrix.

2.2 Compressive Sensing (CS) Based MIMO Radars

Let us consider a MIMO radar system consisting of M_t TX antennas and N_r RX antennas. In the far field of the antennas there are K targets that need to be estimated. For simplicity, we will assume that the targets are not moving, thus the only parameters that needs to be estimated are the target azimuth angles θ_k , $k = 1, \dots, K$. The results can be easily extended to the case of moving targets. In the following, we will consider separate the case of colocated and widely separated antennas.

2.2.1 CS-Based Colocated MIMO Radars

Let us assume that the transmit and receive antennas are closely spaced, with the i -th trasnmit/receive antenna placed at location $(r_i^t, \alpha_i^t)/(r_i^r, \alpha_i^r)$ (in polar coordinates). Let L denote the number of T_s -spaced samples of the transmitted waveforms. The effect of the compressive receiver in Fig. 1 of [29] is equivalent to pre-multiplying by matrix Φ a T_s -sampled version of the received signal. The size of Φ is $M \times L$. The details of CS-based colocated MIMO radars formulation can be found in Section 1.3.2 of Chapter 1.

2.2.2 CS-Based Widely Separated MIMO Radars

Let us assume a widely separated antenna scenario with K targets. Typically, in this scenario each target is viewed as a collection of multiple independent and isotropic

scatterers. However, if the TX waveforms are sufficiently narrowbanded, the target scatterers are unresolvable, and thus each extended target can be modeled as a point target located at the target gravity center, denoted here by (x_k, y_k) (in cartesian coordinates). Let (x_i^t, y_i^t) and (x_l^r, y_l^r) denote the locations of the i -th TX and the l -th RX antenna, respectively.

The i -th antenna transmits the signal $x_i(t)$. Assuming that the antennas transmits on different channels, the target returns due to each transmit antenna can be separated. The baseband signal at the l -th receive antenna, arising due to the transmission of the i -th antenna equals [17]

$$z_{il}(t) = \sum_{k=1}^K h_k^{il} x_i(t - \tau_{ik} - \tau_{kl}) + n_{il}(t), \quad (2.1)$$

where $\tau_{ik} = d_{ik}^t/c$ is the propagation delay between the i -th transmit antenna and the gravity center of the k -th target, with $d_{ik}^t = \sqrt{(x_i^t - x_k)^2 + (y_i^t - y_k)^2}$; τ_{kl} is the propagation delay between the gravity center of the k -th target and the l -th receive antenna; $h_k^{il} = \sum_{q=1}^Q h_{qk}^{il}$ represents the channel gain associated with the k -th target and the TX-RX antenna pair (i, l) . $n_{il}(t)$ denotes interference and noise.

Let us discretize the target state space into N grid points, i.e., $[(x_n, y_n)], n = 1, \dots, N$ and let s_n^{il} denote the coefficient associated with the n -th grid point for the TX-RX antenna pair (i, l) . By choosing the measurement matrix Φ as an identity matrix, the received signal $z_{il}(t)$ can be rewritten as a linear combination of target returns reflected from all grid points, i.e.,

$$\begin{aligned} r_{il}(t) &= \sum_{n=1}^N s_n^{il} x_i(t - \tau_{in} - \tau_{nl}) + n_{il}(t) \\ &= \mathbf{p}_{il}^T(t) \mathbf{s}^{il} + n_{il}(t), \end{aligned} \quad (2.2)$$

where $\mathbf{p}_{il}(t) = [x_i(t - \tau_{i1} - \tau_{1l}), \dots, x_i(t - \tau_{iN} - \tau_{Nl})]^T$ and $\mathbf{s}^{il} = [s_1^{il}, \dots, s_N^{il}]^T$. If the k -th target is located at (x_n, y_n) , the coefficient s_n^{il} equals h_k^{il} ; otherwise, it equals zero.

On denoting by L the number of T_s -spaced samples, we stack L samples corresponding to the TX-RX antenna pair (i, l) into a vector \mathbf{r}_{il} as

$$\begin{aligned}\mathbf{r}_{il} &= [z_{il}(0T_s), \dots, z_{il}((L-1)T_s)]^T \\ &= \mathbf{\Psi}_{il}\mathbf{s}^{il} + \mathbf{n}_{il},\end{aligned}\tag{2.3}$$

where $\mathbf{\Psi}_{il} = [\mathbf{p}_{il}(0T_s), \dots, \mathbf{p}_{il}((L-1)T_s)]^T$ and $\mathbf{n}_{il} = [n_{il}(0T_s), \dots, n_{il}((L-1)T_s)]^T$.

We can stack the received samples from all the pairs of TX and RX antenna into a vector \mathbf{y} of length $M_t N_r L$, i.e.,

$$\begin{aligned}\mathbf{y} &= [\mathbf{r}_{11}^T, \dots, \mathbf{r}_{1N_r}^T, \mathbf{r}_{M_t 1}^T, \dots, \mathbf{r}_{M_t N_r}^T]^T \\ &= \text{diag}\{[\mathbf{\Psi}_{11}, \mathbf{\Psi}_{12}, \dots, \mathbf{\Psi}_{M_t N_r}]\} \mathbf{s} + \mathbf{n},\end{aligned}\tag{2.4}$$

where $\mathbf{s} = [(\mathbf{s}^{11})^T, (\mathbf{s}^{12})^T, \dots, (\mathbf{s}^{M_t N_r})^T]^T$.

Note that for all pairs (i, l) , the vector \mathbf{s}^{il} contains zeros everywhere except at locations corresponding to the grid points occupied by targets. Thus, \mathbf{s} is a sparse vector. If there is a target at the n -th grid point, all n -th entries of \mathbf{s}^{il} corresponding to all TX-RX antenna pairs, i.e., s_n^{il} , $i = 1, \dots, M_t$, $l = 1, \dots, N_r$, are non-zero. Therefore, by appropriately rearranging the columns of the basis matrix, the non-zero elements of \mathbf{s} corresponding to different pairs and the same target can be clustered together, which makes \mathbf{s} to appear as group sparse. On letting \mathbf{u}_{il}^n denote the n -th column of $\mathbf{\Psi}_{il}$, the columns of the basis matrix $\mathbf{\Psi}_g$ that induce group sparsity can be arranged as follows [50], [49]:

$$\mathbf{\Psi}_g = [\tilde{\mathbf{\Psi}}_1, \dots, \tilde{\mathbf{\Psi}}_N]\tag{2.5}$$

where $\tilde{\mathbf{\Psi}}_n = \text{diag}\{[\mathbf{u}_{11}^n, \dots, \mathbf{u}_{1N_r}^n, \mathbf{u}_{21}^n, \dots, \mathbf{u}_{M_t N_r}^n]\}$.

The sparse vector \mathbf{s} associated with $\mathbf{\Psi}_g$ contains K groups of non-zeros entries and each of group is of length $M_t N_r$. The group sparsity of \mathbf{s} can be exploited using a group Lasso approach [57] [58], i.e.,

$$\arg \min_{\mathbf{s}} \underbrace{\frac{1}{2} \|\mathbf{\Phi}_g^H (\mathbf{y} - \mathbf{\Psi}_g \mathbf{s})\|_2^2}_{f_1(\mathbf{s})} + \lambda \underbrace{\sum_{n=1}^N \|\mathbf{s}_n\|_2}_{f_2(\mathbf{s})}\tag{2.6}$$

where $\mathbf{s}_n = [s_n^{11}, \dots, s_n^{M_t N_r}]$. $f_2(\mathbf{s})$ can be recast the ℓ_1 norm of vector $\lambda[\|\mathbf{s}_1\|_2, \dots, \|\mathbf{s}_N\|_2]^T$. Minimization of $f_2(\mathbf{s})$ produces a group-sparse solution [57] [58]. Due to the nature of ℓ_2 norm, all entries of the n -th group \mathbf{s}_n will be zero if $\|\mathbf{s}_n\|_2$ is zero, and will be non-zero otherwise. Since $f_2(s)$ is non-smooth, it is not trivial to directly solve (2.6). Instead of minimizing $f_1(s)$ and $f_2(s)$ simultaneously, the proximal gradient algorithm [59] [60] proceeds by dealing with $f_1(s)$ and $f_s(s)$ individually in an iterative way. Let $\hat{\mathbf{s}}$ denote the solution to (2.6). We can formulate the target indicator vector, \mathbf{d} , so that its n -th entry equals $\|\hat{\mathbf{s}}_n\|_2^2$. The peaks of \mathbf{d} will provide the target information.

2.3 Power Allocation and Waveform Design for Colocated CS-MIMO Radars

2.3.1 Power Allocation

Suppose that the total power allocated to the transmit antennas is fixed, equal to P_t . We next determine how P_t should be distributed among the transmit antennas so that it helps the CS recovery.

UUP [25] [34] [35] indicates that for the recovery of the sparse vector with high probability, the sensing matrix should be orthogonal. This is impossible for a fat measurement matrix, however, we can force the sensing matrix Θ to be as orthogonal as possible, i.e., by minimizing the difference between $\Theta^H \Theta$ and an identity matrix of size N . Let $\mathbf{p} = [p_1, \dots, p_{M_t}]^T$ denote the transmit power allocated to M_t TX antennas, with $\sum_{i=1}^{M_t} p_i = P_t$. Then, the k -th column of the sensing matrix equals

$$\mathbf{u}_k = \left[e^{j2\pi \frac{\eta_1^r(a_k)f}{c}}, \dots, e^{j2\pi \frac{\eta_{N_r}^r(a_k)f}{c}} \right]^T \otimes (\Phi \mathbf{X} \mathbf{V}(a_k) \tilde{\mathbf{p}}) \quad (2.7)$$

where $\mathbf{V}(a_k) = \text{diag}\{\mathbf{v}_t(a_k)\}$ and $\tilde{\mathbf{p}} = \sqrt{\mathbf{p}}$.

In particular, the power allocation problem can be formulated as the following optimization problem:

$$\begin{aligned} \min_{\tilde{\mathbf{p}}} \quad & \sum_{k \neq k'} |\mathbf{u}_{k'}^H \mathbf{u}_k|^2 + \sum_k |\mathbf{u}_k^H \mathbf{u}_k - P_t N_r|^2 \\ \text{s.t.} \quad & \tilde{\mathbf{p}}^H \tilde{\mathbf{p}} = P_t, \quad \tilde{\mathbf{p}} \geq \mathbf{0}_{M_t \times 1}, \quad \tilde{\mathbf{p}} \leq \sqrt{P_m} \mathbf{1}_{M_t \times 1} \end{aligned} \quad (2.8)$$

where the objective function is the squared difference between $\Theta^H \Theta$ and $P_t N_r \mathbf{I}_N$, with the first term, denoted by SCSM, being the sum of the square magnitude of the cross-correlation of column pairs $\mathbf{u}_k, \mathbf{u}_{k'}$ of the sensing matrix. P_m is the maximum transmit power of each TX antenna. It holds that

$$\begin{aligned} |\mathbf{u}_{k'}^H \mathbf{u}_k|^2 &= u_{kk'} |\tilde{\mathbf{p}}^H \mathbf{B}_{kk'} \tilde{\mathbf{p}}|^2 \\ &= u_{kk'} [(\tilde{\mathbf{p}}^H \mathbf{B}_{kk'} \tilde{\mathbf{p}})^2 + (\tilde{\mathbf{p}}^H \mathbf{B}_{kk'} \tilde{\mathbf{p}})^2] \end{aligned} \quad (2.9)$$

where $u_{kk'} = \left| \sum_{l=1}^{N_r} e^{j2\pi \frac{(\eta_l^r(a_k) - \eta_l^r(a_{k'}))f}{c}} \right|^2$, $\mathbf{B}_{kk'} = \mathbf{V}^H(a_k) \mathbf{X}^H \Phi^H \Phi \mathbf{X} \mathbf{V}(a_k)$, $\mathbf{B}_{kk'} = \frac{\mathbf{B}_{kk'} + \mathbf{B}_{kk'}^H}{2}$ and $\mathbf{B}_{kk'} = \frac{\mathbf{B}_{kk'} - \mathbf{B}_{kk'}^H}{2j}$. It is easy to see that $\mathbf{B}_{kk'}$ is not a positive semidefinite (PSD) matrix unless $k = k'$ and thus the objective function is nonconvex. However, we can make the objective function convex via the following trick:

$$\begin{aligned} (\tilde{\mathbf{p}}^T \mathbf{B}_{kk'} \tilde{\mathbf{p}})^2 &= (\tilde{\mathbf{p}}^T (\underbrace{\mathbf{B}_{kk'}}_{\mathbf{C}_{kk'}} + \frac{b}{P_t} \mathbf{I}) \tilde{\mathbf{p}} - b)^2 \\ &= (\tilde{\mathbf{p}}^T \mathbf{C}_{kk'} \tilde{\mathbf{p}})^2 + \underbrace{\tilde{\mathbf{p}}^T (-2b \mathbf{C}_{kk'} + \frac{d}{P_t} \mathbf{I}) \tilde{\mathbf{p}}}_{\mathbf{D}_{kk'}} + C_r \end{aligned} \quad (2.10)$$

where b and d are nonnegative real scalars that let $\mathbf{C}_{kk'}$ and $\mathbf{D}_{kk'}$ be PSD matrices, i.e., $\frac{b}{P_t} + \lambda_{\min}(\mathbf{B}_{kk'}) \geq 0$ and $\frac{d}{P_t} + \lambda_{\min}(-2b \mathbf{C}_{kk'}) \geq 0$. C is a constant that will not affect the objective function.

(2.10) is convex since $\mathbf{C}_{kk'}$ and $\mathbf{D}_{kk'}$ are PSD matrices. By performing the same trick on $(\tilde{\mathbf{p}}^H \mathbf{B}_{kk'} \tilde{\mathbf{p}})^2$, we can obtain

$$(\tilde{\mathbf{p}}^T \mathbf{B}_{kk'} \tilde{\mathbf{p}})^2 = (\tilde{\mathbf{p}}^T \mathbf{C}_{kk'} \tilde{\mathbf{p}})^2 + \tilde{\mathbf{p}}^T (\mathbf{D}_{kk'}) \tilde{\mathbf{p}} + C_i \quad (2.11)$$

In the same way, the second term in the objective function (2.8) can be rewritten as

$$\begin{aligned} |\mathbf{u}_k^H \mathbf{u}_k - P_t N_r|^2 &= N_r^2 (\tilde{\mathbf{p}}^T (\mathbf{B}_{kk} - \mathbf{I}) \tilde{\mathbf{p}})^2 \\ &= (\tilde{\mathbf{p}}^T \mathbf{C}_{kk} \tilde{\mathbf{p}})^2 + \tilde{\mathbf{p}}^T (\mathbf{D}_{kk}) \tilde{\mathbf{p}} + C \end{aligned} \quad (2.12)$$

Then the objective function of (2.8) can be transformed into a convex function as follows

$$\begin{aligned} \min_{\tilde{\mathbf{p}}} \quad & \sum_{k \neq k'} (\tilde{\mathbf{p}}^T \mathbf{C}_{kk'} \tilde{\mathbf{p}})^2 + (\tilde{\mathbf{p}}^T \mathbf{C}_{kk'} \tilde{\mathbf{p}})^2 + \tilde{\mathbf{p}}^T (\mathbf{D}_{kk'} + \mathbf{D}_{kk'}) \tilde{\mathbf{p}} + \sum_k (\tilde{\mathbf{p}}^T \mathbf{C}_{kk} \tilde{\mathbf{p}})^2 + \tilde{\mathbf{p}}^T (\mathbf{D}_{kk}) \tilde{\mathbf{p}} \\ \text{s.t.} \quad & \tilde{\mathbf{p}}^H \tilde{\mathbf{p}} = P_t, \quad \tilde{\mathbf{p}} \geq \mathbf{0}_{M_t \times 1}, \quad \tilde{\mathbf{p}} \leq \sqrt{P_m} \mathbf{1}_{M_t \times 1} \end{aligned} \quad (2.13)$$

Looking at the constraints of (2.13), we can see that the first constraint is nonconvex. We can approximate that constraint by its local affine approximation, i.e.,

$$(\tilde{\mathbf{p}})^T \tilde{\mathbf{p}} \approx (\tilde{\mathbf{p}}^j)^T \tilde{\mathbf{p}}^j + 2(\tilde{\mathbf{p}}^j)^T (\tilde{\mathbf{p}} - \tilde{\mathbf{p}}^j) \quad (2.14)$$

where $(\tilde{\mathbf{p}}^j)$ is the estimate of $\tilde{\mathbf{p}}$ at the j -th iteration. The solution to (2.13) can be obtained in an iterative fashion as follows.

1. Set $\tilde{\mathbf{p}}^{(0)} = [1, 1, \dots, 1]^T$ at the initial iteration;
2. At the j -th iteration, $\tilde{\mathbf{p}}^{(j)}$ is obtained by solving

$$\begin{aligned} \min_{\tilde{\mathbf{p}}} \quad & \sum_{k \neq k'} |\mathbf{u}_{k'}^H \mathbf{u}_k|^2 + \sum_k |\mathbf{u}_k^H \mathbf{u}_k - P_t N_r|^2 \\ \text{s.t.} \quad & (\tilde{\mathbf{p}}^{(j-1)})^H \tilde{\mathbf{p}}^{(j-1)} + 2(\tilde{\mathbf{p}}^{(j-1)})^H (\tilde{\mathbf{p}} - \tilde{\mathbf{p}}^{(j-1)}) = P_t, \\ & \tilde{\mathbf{p}} \geq \mathbf{0}_{M_t \times 1}, \quad \tilde{\mathbf{p}} \leq \sqrt{P_m} \mathbf{1}_{M_t \times 1} \end{aligned} \quad (2.15)$$

3. If the stop criterion is not satisfied, go to 2); otherwise, output $\tilde{\mathbf{p}} = \tilde{\mathbf{p}}^{(j)}$.

In applications, the implementation of the measurement matrix may increase the complexity of the analog circuit. Therefore, we can skip the step of linear compression and directly collect a small number of samples at the RX antennas, i.e., L is small.

If the TX antennas transmit orthogonal waveforms, i.e., $\mathbf{X}^H \mathbf{X} = \mathbf{I}$, then \mathbf{B}_{kk} is a diagonal matrix. Then (2.8) can be reduced to a simple convex problem as follows:

$$\begin{aligned} \min_{\mathbf{p}} \quad & \mathbf{p}^T \left(\sum_{k \neq k'} u_{kk'} \mathbf{b}_{kk'}^* \mathbf{b}_{kk'}^T \right) \mathbf{p} \\ \text{s.t.} \quad & \mathbf{1}_{M_t \times 1}^T \mathbf{p} = P_t, \quad \mathbf{p} \geq \mathbf{0}, \quad \mathbf{p} \leq P_m \mathbf{1}_{M_t \times 1} \end{aligned} \quad (2.16)$$

where $\mathbf{b}_{kk'}$ is the diagonal vector of $\mathbf{B}_{kk'}$.

We should note here that the power allocation vector obtained from (2.8) or (2.16) will not affect the SIR since the total transmitted power is fixed. Therefore, the performance gain stems from the improved sensing matrix only. If some TX antenna contributes to the sensing matrix negatively, the power allocated to those TX antennas will be very small. This indicates that the proposed scheme will reduce the number of active TX antennas as compared to the uniform power allocation.

2.3.2 Waveform Design

Along the lines of optimal power allocation, we can also optimize the transmit waveforms to further improve the performance of the CS-based MIMO radars. Let us again consider as design criterion the minimization of the difference between the Gram of the sensing matrix and an identity matrix for a fixed total transmit power. The k -th column of the sensing matrix becomes

$$\mathbf{u}_k = \left[e^{j2\pi \frac{\eta_1^r(a_k)f}{c}}, \dots, e^{j2\pi \frac{\eta_{N_r}^r(a_k)f}{c}} \right]^T \otimes (\Phi \tilde{\mathbf{V}}(a_k) \mathbf{x}) \quad (2.17)$$

where $\tilde{\mathbf{V}}(a_k) = \mathbf{I}_L \otimes \mathbf{v}_t^T(a_k)$ and $\mathbf{x} = \text{vec}(\mathbf{X}^T)$. Since the total power is set to P_t , it holds that $\mathbf{x}^H \mathbf{x} = P_t$.

We can formulate the following optimization problem:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \sum_{k \neq k'} |\mathbf{u}_{k'}^H \mathbf{u}_k|^2 + \sum_k |\mathbf{u}_k^H \mathbf{u}_k - P_t N_r|^2 \\ \text{s.t.} \quad & \mathbf{x}^H \mathbf{x} = P_t \end{aligned} \quad (2.18)$$

It is worth noting that

- (2.18) is not convex and thus we can only find a local minimum that depends on the initial waveforms.
- The method for waveform design is quite similar to that for power allocation except that the formulation is a little different. However, the number of variables for waveform design is much larger than that for power allocation, i.e., the former is $M_t L$ while the latter is M_t .

2.4 Power Allocation for Widely Separated CS-MIMO Radars

Suppose that the total power is set to P_t . Let \mathbf{p} denote the transmit power allocated to M_t TX antennas, and thus $\sum_{i=1}^{M_t} p_i = P_t$. Again, we force Φ_g to be as orthogonal as possible by minimizing the difference between the Gram matrix of the sensing matrix, i.e., $\Phi_g^H \Phi_g$ and an identity matrix of size $M_t N_r N$. In particular, we formulate the

following optimization problem:

$$\begin{aligned} \min_{\mathbf{p}} \quad & \|\Phi_g^H \Phi_g - \mathbf{I}\|_F^2 \\ \text{s.t.} \quad & \mathbf{1}_{M_t \times 1}^T \mathbf{p} = P_t, \quad \mathbf{p} \geq \mathbf{0}_{M_t \times 1}, \quad \mathbf{p} \leq P_m \mathbf{1}_{M_t \times 1} \end{aligned} \quad (2.19)$$

where P_m is the maximum transmitted power accepted by a TX antenna.

The Gram matrix $\Phi_g^H \Phi_g$ can be rewritten as

$$\begin{aligned} \Phi_g^H \Phi_g &= \begin{bmatrix} \tilde{\Psi}_1, \dots, \tilde{\Psi}_N \end{bmatrix}^H \begin{bmatrix} \tilde{\Psi}_1, \dots, \tilde{\Psi}_N \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{D}_p \mathbf{D}_{11} & \cdots & \mathbf{D}_p \mathbf{D}_{1N} \\ \vdots & \cdots & \vdots \\ \mathbf{D}_p \mathbf{D}_{1N} & \cdots & \mathbf{D}_p \mathbf{D}_{NN} \end{bmatrix} \end{aligned} \quad (2.20)$$

where $\mathbf{D}_{kk'} = \tilde{\Psi}_k^H \tilde{\Psi}_{k'}$, $k, k' = 1, \dots, N$ are diagonal matrices due to the special sparse structure of $\tilde{\Psi}_N$ as shown in (2.5) and $\mathbf{D}_p = \text{diag}\{\mathbf{p} \otimes \mathbf{1}_{N_r \times 1}\}$. Since the waveform of each TX antenna is of unit power, $\mathbf{D}_{kk} = \mathbf{I}$, $k = 1, \dots, N$.

Let $\mathbf{d}_i^{kk'}$ denote the i -th group of the diagonal elements of $\mathbf{D}_{kk'}$, whose length is N_r . Then the optimization problem in (2.20) can be further written as

$$\begin{aligned} \min_{\mathbf{p}} \quad & \mathbf{p}^T \mathbf{D} \mathbf{p} + N_r \mathbf{p}^T \mathbf{p} \\ \text{s.t.} \quad & \mathbf{1}_{M_t \times 1}^T \mathbf{p} = P_t, \quad \mathbf{p} \geq \mathbf{0}_{M_t \times 1}, \quad \mathbf{p} \leq P_m \mathbf{1}_{M_t \times 1} \end{aligned} \quad (2.21)$$

where \mathbf{D} is a diagonal matrix whose i -th diagonal element equals $\sum_{k \neq k'} |\mathbf{d}_i^{kk'}|_2^2$. The first term of the objective function sums up the squared correlation of cross column pairs in the sensing matrix. The second term represents the squared error between the column norm of the sensing matrix and 1, in which the constant terms have been removed.

One may wonder whether minimizing (2.21) requires that the column norms of the sensing matrix approach 1, or equivalently, \mathbf{p} approaches a unit vector. For large N , the second term in the objective function is small as compared to the first term, so that the solution $\mathbf{p} = \mathbf{1}$ can be avoided. From the perspective of the CS recovery methods, the column norm of the sensing matrix for the group Lasso method is not as important as in the Lasso method, since the group Lasso approach forces the coefficients

within a group to be nonzero or zero simultaneously. This indicates that the sum of the squared norm of the columns within a group (SSNCG), rather than of a single column of the sensing matrix weighs in the group Lasso approach. Due to the special structure of the sensing matrix, as seen in (2.5), the SSNCG of different groups is the same, independent of the power vector \mathbf{p} , and lies between $\frac{P_t^2}{M_t}$ and $CP_m^2 + (P_t - CP_m)^2$, where $C = \lfloor P_t/P_m \rfloor$. Therefore, the column norm of the sensing matrix does not have to approach 1 when the group Lasso method is used. It is also reasonable to remove $\mathbf{p}^T \mathbf{p}$ from the objective function.

The power allocation vector obtained from (2.21) will not affect the SIR since the total transmit power is fixed. Therefore, the performance gain stems from the improved sensing matrix only. If some TX antenna contributed to the sensing matrix negatively, the power allocated to those TX antennas would be reduced.

2.5 Simulations

2.5.1 Simulation Results for Colocated CS-MIMO Radars

Power Allocation

We consider a MIMO radar system with TX and RX antennas, uniformly located on a disk of radius 10m. The carrier frequency is $f = 5\text{GHz}$. Each TX antenna uses an orthogonal waveform sequence of length $L = 32$ and unit power. The received signal is corrupted by zero-mean Gaussian noise of unit power. The SNR is set to 0dB. Three targets are present on the angle grid $[0^\circ, 0.1^\circ, \dots, 4.9^\circ, 5^\circ]$. The total transmitted power is set to M_t and the maximum transmit power for each antenna is 9W. In the simulations, the identity matrix is used as the measurement matrix Φ . The power allocation vector is obtained by solving the problem of (2.16).

Figure 2.1 demonstrates the average coherence of a column pair of the sensing matrix, corresponding to the proposed power allocation (PPA) scheme, computed over 100 independent runs and for $M_t = 5, 10, 15, 20, 25, 30$. For comparison purposes, the corresponding results for uniform power allocation (UPA) and random power allocation (RPA) are also shown in the same figure. It can be seen from Fig. 2.1 that the PPA

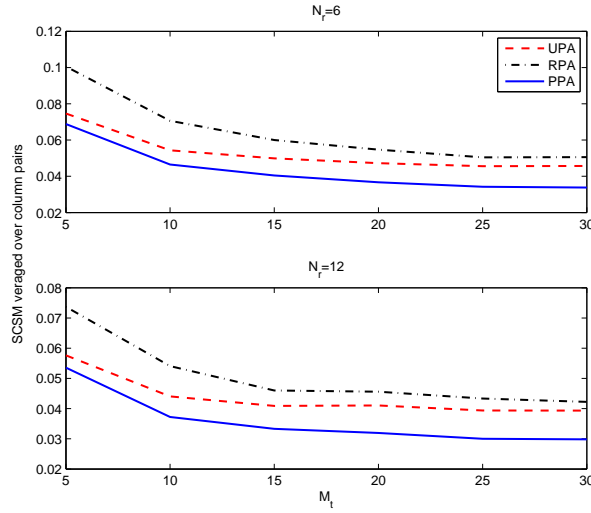


Figure 2.1: SCSM versus M_t under optimal (PPA), uniform (UPA) and random (RPA) power allocation in CS-based colocated MIMO radars. Case (I) $N_r = 6$; Case (II) $N_r = 12$; in both cases $L = 32$.

scheme can reduce the SCSM as compared to the UPA and RPA schemes. Figure 2.1 also shows that the SCSM can be reduced by increasing the number of RX antennas. The performance gain with the increase of the number of TX/RX antennas is more prominent at low SNR. This is because using more TX/RX antennas effectively increases the array aperture. In addition, an increase in the number of TX antennas can also improve the SNR of the received signal at the RX antennas, as in our simulations the transmit power for each TX antenna is fixed.

Figure 2.2 shows the receiver operating characteristic (ROC) curves of the angle estimates, obtained based on 500 independent runs. In each run, three targets are randomly generated on the angle grid. Here, the probability of detection (PD) is the percentage of cases in which all the targets are detected. The probability of false alarm (PFA) is defined as the percentage of cases in which false targets are detected. The cases of $M_t = 12$ and $N_r = 6, 12$ are shown in Fig. 2.2. One can see that the PPA scheme can improve the ROC performance as compared to the UPA scheme. Again, an increase in the number of RX antennas can improve the detection performance. Increasing the number of RX antennas cannot boost SNR and thus the performance

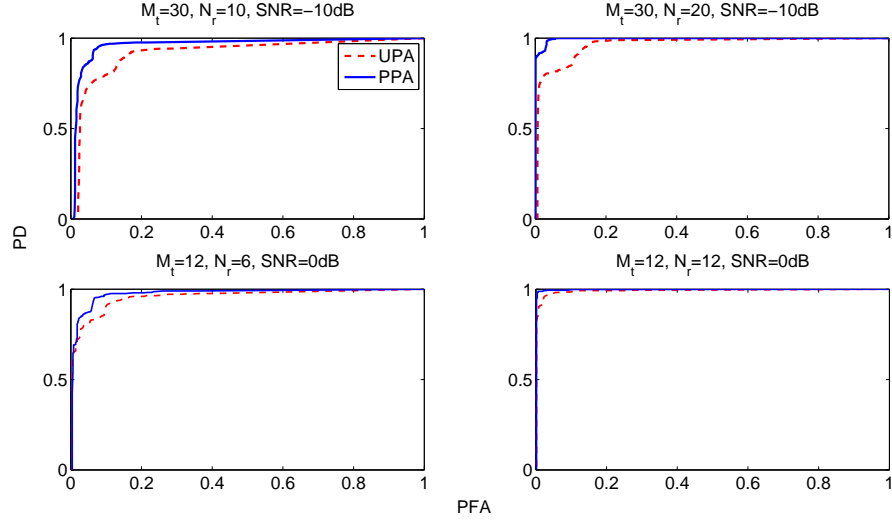


Figure 2.2: The ROCs of angle estimates under power allocation in CS-based colocated MIMO radars with $L = 32$.

gain comes from the improvement of the sensing matrix as shown in Fig. 2.1.

In addition to performance improvement, the PPA scheme also reduces the number of active TX antennas. For example, in the case of $M_t = 30$ and $N_r = 12$, 11 TX antennas are allocated power less than $0.0001W$ on an average of 500 runs. This indicates that the PPA scheme only requires 19 TX antennas to be active while all 30 TX antennas are needed for the UPA scheme. Figure 2.3 (a) shows the distribution of TX antennas which were assigned power less than $0.0001W$ in one run; the non-needed antennas are marked on the figure. Figure 2.3 (b) demonstrates the power allocation results in this case.

Waveform Design

We consider a MIMO radar system with 10 TX and RX antennas uniformly located on a disk of radius 10m. The carrier frequency is $f = 5\text{GHz}$. The received signal is corrupted by zero-mean Gaussian noise of unit power. Three targets are present on the angle grid $[0^\circ, 0.2^\circ, \dots, 2^\circ]$. The total transmitted power is set to M_t . The initial waveform $\mathbf{x}^{(0)}$ uses an orthogonal Hadamard sequence of length $L = 16$ and unit power. The measurement matrix Φ is chosen as a random matrix with $M = \text{round}(0.7L)$,

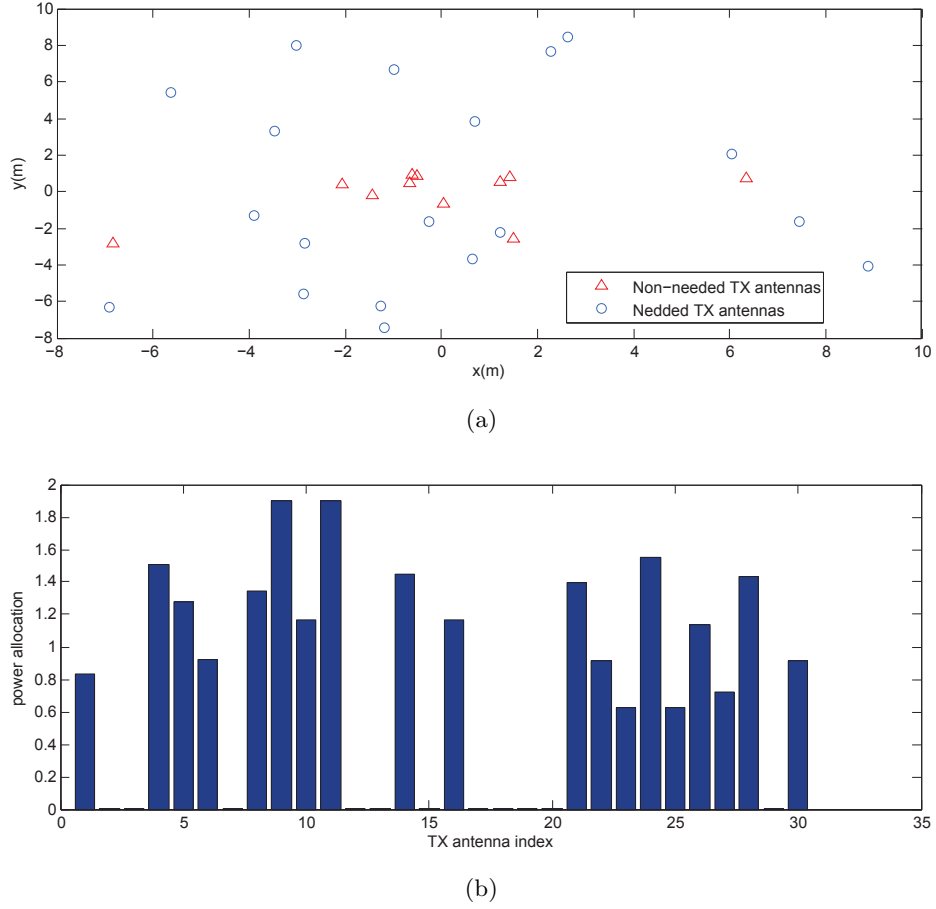


Figure 2.3: The TX distribution in CS colocated MIMO radars with $M_t = 30$, $N_r = 12$ and $L = 32$.

which means each RX antenna forwards $M = 11$ samples. The ROC curves of the angle estimates produced by the proposed waveform-design method is shown in Fig. 2.4 which demonstrates the performance improvement due to the designed waveform.

2.5.2 Power Allocation for Widely Separated CS-MIMO Radars

In this section, we demonstrate the performance of the PPA and the UPA schemes based on the group Lasso method in the context of widely separated MIMO radars. We consider a MIMO radar system with TX and RX radars that are uniformly located on a circle of radius 6000m and 3000m, respectively. The carrier frequency is $f = 5\text{GHz}$. Each transmit radar uses orthogonal Hadamard waveform sequences of

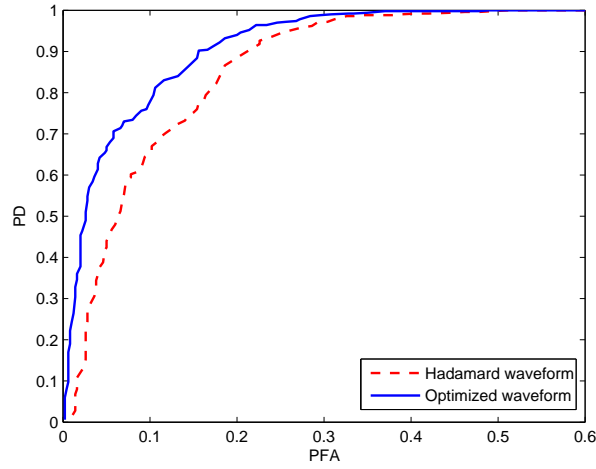


Figure 2.4: The ROCs of angle estimates under waveform design in CS-based colocated MIMO radars with $M_t = N_r = 10$, $L = 16$ and $M = 11$.

length $L = 20$, and unit power. In the simulations, the measurement matrix Φ is chosen as an identity matrix. Three targets are assumed to be present in the search space $[1000, 1050, \dots, 1200]\text{m} \times [1000, 1050, \dots, 1200]\text{m}$, and in each run, the targets are randomly located on grid points. The target reflectivity is a Gaussian random variable with unit variance. The total power varies with the number of TX antennas and is set to M_t . The power threshold for each TX antenna is 2.

Figure 2.5 compares the PPA and the UPA schemes in terms of the squared difference between the Gramian of the sensing matrix and the identity matrix (SEGI), averaged over the NM_tN_r elements of the Gramian matrix of the sensing matrix. It can be seen from Fig. 2.5 that the PPA scheme reduces SEGI as compared to UPA, and also that the SEGI decreases with the number of TX and RX antennas.

Figure 2.6 shows the ROC curves of the angle estimates, obtained based on 1000 independent runs. In each run, three targets are randomly generated on the angle grid of interest. Four cases are shown, i.e., $(M_t = 10, N_r = 1, \text{SNR} = 10\text{dB})$, $(M_t = 5, N_r = 2, \text{SNR} = 10\text{dB})$, $(M_t = 10, N_r = 1, \text{SNR} = 20\text{dB})$ and $(M_t = 5, N_r = 2, \text{SNR} = 20\text{dB})$. One can see that the PPA scheme can improve the performance in terms of the ROCs as compared to the UPA scheme. With the same number of TX and RX antennas,

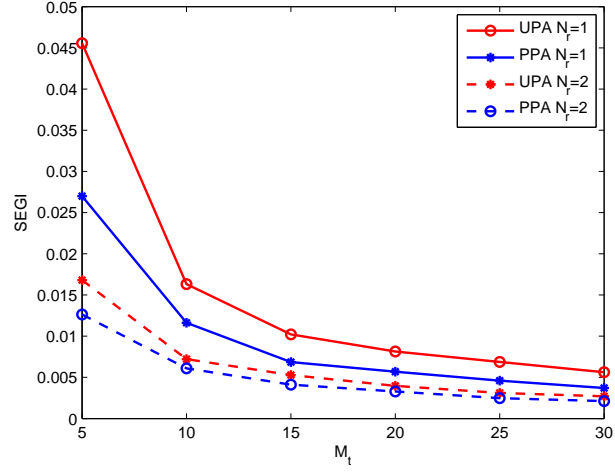


Figure 2.5: SEGI versus M_t under optimal and uniform power allocation in CS-based widely separated MIMO radars with $L = 20$.

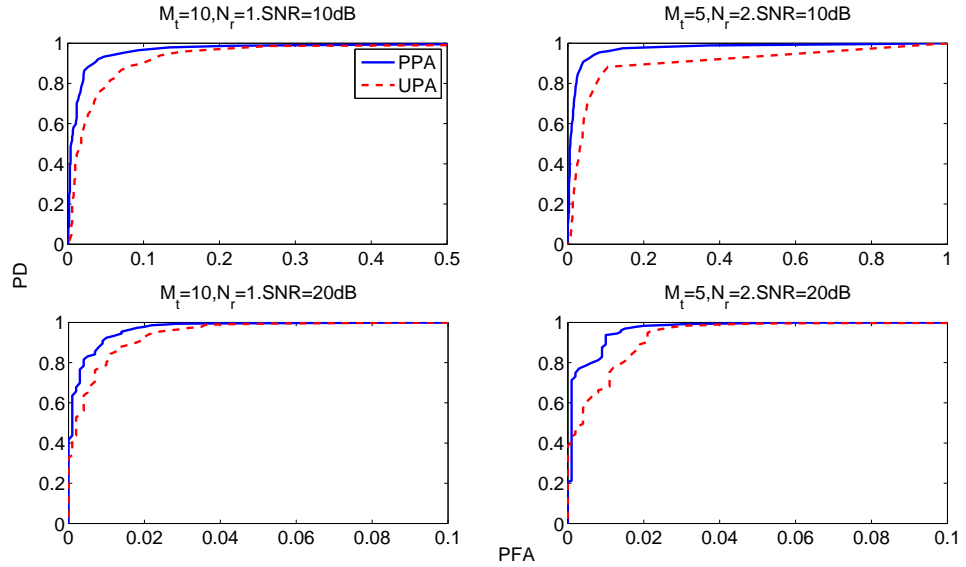


Figure 2.6: ROC of target location estimates under power allocation in CS-based widely separated MIMO radar with four combinations of M_t and N_r .

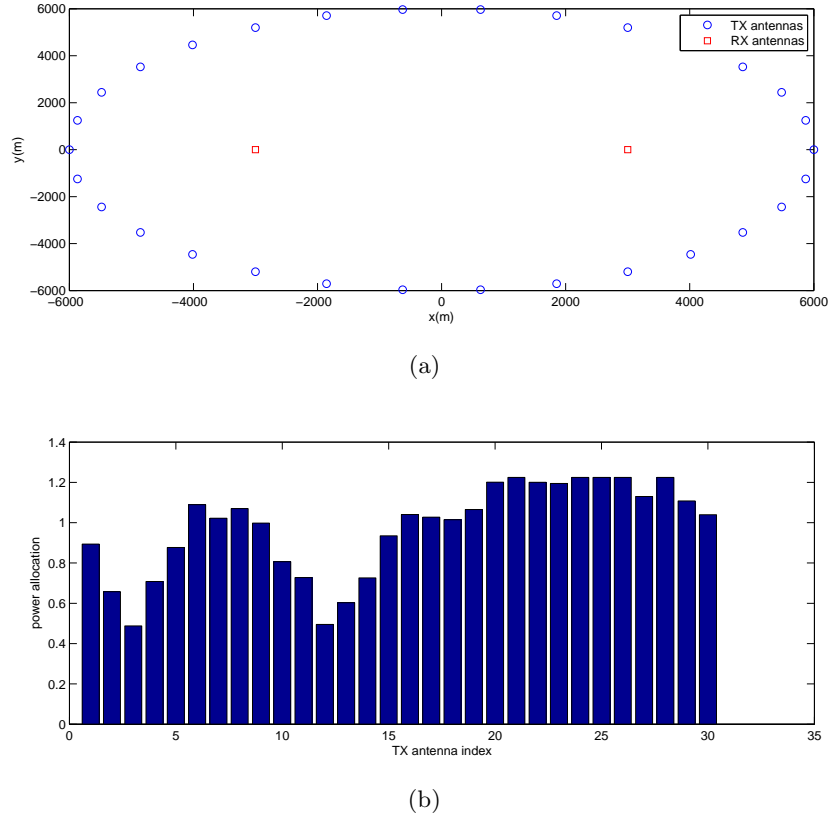


Figure 2.7: The TX distribution in CS colocated MIMO radar with $M_t = 30$, $N_r = 2$ and $L = 32$.

the performance improvement of the PPA scheme diminishes at higher SNR scenarios. Figure 2.7 shows the distribution of TX and RX antennas in one run where $M_t = 10$, $N_r = 1$, and power allocation results based on the PPA scheme in this case. Unlike in the colocated MIMO radars, all TX antennas are needed for widely separated MIMO radars. This is because the antennas in this case see different aspects of the target and so the received signals due to each TX-RX antenna pairs are separated and stacked together at the fusion center.

2.6 Summary

We have proposed power allocation schemes for colocated and widely separated CS-based MIMO radar systems. The proposed schemes aims at rendering the sensing

matrix as orthogonal as possible. It has been shown that the estimation of DOA can be improved by allocating power adaptively to TX antennas given a total transmitted power. Furthermore, the proposed scheme can reduce the number of active TX antennas as compared to the uniform power allocation scheme. This is because the TX antennas that cause the sensing matrix to be more correlated are eliminated by the proposed power allocation scheme. Along the lines of optimal power allocation, we also developed a waveform design method that further improves the performance of CS-MIMO radars in the colocated case. The method proceeds by minimizing the difference between the Gram matrix of the sensing matrix and an identity matrix with respect to the vector of the transmitted waveforms, while the total transmitted power is fixed. Simulations showed substantial detection performance improvement when using the optimally designed waveform as compared to the already good performance corresponding to the Hadamard waveform. The improvement is clearly observed at the very low probability of false alarm range, which is the only region in which the Hadamard waveform did not yield good performance.

Chapter 3

Clutter Suppression in Colocated Compressive Sensing Based MIMO Radars

In this chapter, we proposed a Capon beamforming based clutter suppression scheme in the context of compressive sensing based colocated MIMO radars. Capon beamforming is applied at the fusion center on compressively obtained data, which are forwarded by the receive antennas. Subsequently, the target is estimated using CS theory, by exploiting the sparsity of the beamformed signals. The power allocation scheme is proposed to further improve the target estimation performance by minimizing the coherence between the columns of the sensing matrix.

3.1 Introduction

CS-based colocated MIMO radars [47] [26] [29] exploit the sparsity of targets in the target space and enable target estimation based on a small number of samples obtained at the RX antennas. The assumed sparsity, however, diminishes in the presence of clutter, and as result the performance of CS-based MIMO radars in the presence of clutter degrades. In this chapter we proposed a scheme to suppress clutter in the context of colocated MIMO radars. When clutter and signal of interest have different Doppler shifts, clutter suppression can be relatively easily done, especially when the source of clutter is static while the target is moving [61] [62] [63]. In this chapter, we will consider the more challenging case, in which Doppler differences cannot be exploited to separate clutter from signal of interest, e.g., when both the target and the source of clutter are static. We will assume that the clutter covariance is known. Typically, clutter has high space correlation to its neighboring range cells [64] [65].

Therefore, the received signals at neighboring range cells can be used to estimate the clutter covariance matrix of the current range cell. The proposed schemes consists of first applying at the fusion center Capon beamforming on the received compressively obtained data, which are forwarded by the receive antennas, and second, recovering the target using CS theory, by exploiting the sparsity of the beamformed signals. Further, based on the uniform uncertainty principle (UUP) [25] [34] [35], we propose a power allocation scheme to improve the CS performance after beamforming.

The proposed scheme has been shown to significantly improve detection performance in the presence of strong clutter.

3.2 Clutter Rejection in CS-Based MIMO Radars

Let us consider a MIMO radar system consisting of M_t TX antennas and N_r RX antennas that are closely spaced in an arbitrary configuration. Let L denote the number of T_s -spaced samples of the transmitted waveforms. The size of measurement matrix is Φ is $M \times L$. The details of CS-based colocated MIMO radars have been introduced in Section (1.3.2) of Chapter 2. Let us form the $(M \times N_r)$ matrix $\mathbf{Y} = [\mathbf{r}_1, \dots, \mathbf{r}_{N_r}]$, where \mathbf{r}_l defined in (1.44) is a vector containing the compressed samples forwarded to the fusion center by the l -th receive antenna.

It holds that

$$\mathbf{Y}^T = \sum_{k=1}^K \beta_k \mathbf{v}_r(\theta_k) \mathbf{v}_t^T(\theta_k) \mathbf{X}^T \Phi^T + \mathbf{Z} \quad (3.1)$$

where

- \mathbf{Y} : the received signal at N_r RX antennas, of which the i -th column contains the received signal from the i -th RX antenna
- $\mathbf{v}_t(\theta_k)$: the transmit steering vector at the direction of θ_k , defined in (1.43)
- $\mathbf{v}_r(\theta_k)$: the receive steering vector at the direction of θ_k , defined similarly to $\mathbf{v}_t(\theta_k)$
- \mathbf{Z} : the clutter matrix, whose covariance \mathbf{R}_z is assumed to be known

The SCNR of compressed receiving data without beamforming is

$$SCNR_{CS} = \frac{\left\| \text{vec} \left(\sum_{k=1}^K \beta_k \mathbf{v}_r(\theta_k) \mathbf{v}_t^T(\theta_k) \mathbf{X}^T \Phi^T \right) \right\|^2}{\left\| \text{vec}(\mathbf{Z} \Phi^T) \right\|^2}. \quad (3.2)$$

The Capon beamformer, \mathbf{w} , allows the signal from a particular direction θ to pass undistorted, while it minimizes the power coming from all other directions. The estimation of \mathbf{w} is formulated as

$$\min_{\mathbf{w}} \mathbf{w}^H \mathbf{R}_z \mathbf{w} \text{ s.t. } \mathbf{w}^H \mathbf{v}_r(\theta) = 1. \quad (3.3)$$

The solution of (3.3) is

$$\mathbf{w}(\theta) = \frac{\mathbf{R}^{-1} \mathbf{v}_r(\theta)}{\mathbf{v}_r^H(\theta) \mathbf{R}^{-1} \mathbf{v}_r(\theta)}. \quad (3.4)$$

Let \mathbf{w}_n denote the Capon beamformer that focuses on the discrete angle a_n of the N -point angle grid. Let us apply \mathbf{w}_n on the received data matrix \mathbf{Y} to produce $\tilde{\mathbf{y}}_n$, i.e.,

$$\begin{aligned} \tilde{\mathbf{y}}_n &= (\mathbf{w}_n^H \mathbf{Y}^T)^T = \mathbf{Y} \mathbf{w}_n^* \\ &= \tilde{\mathbf{\Theta}}_n \mathbf{s} + \mathbf{Z}^T \mathbf{w}_n^* \quad n = 1, \dots, N \end{aligned} \quad (3.5)$$

where

$$\tilde{\mathbf{\Theta}}_n = \Phi \mathbf{X} [\mathbf{v}_t(a_1) \mathbf{v}_r^T(a_1) \mathbf{w}_n^*, \dots, \mathbf{v}_t(a_N) \mathbf{v}_r^T(a_N) \mathbf{w}_n^*]$$

and \mathbf{s} is a sparse vector, the non-zero elements of which indicate the target locations.

The vector $\tilde{\mathbf{y}}_n$ ($M \times 1$) is the beamformer output corresponding to angle a_n , computed based on M snapshots of the data received at the fusion center. If there is a target at a_n then all elements of $\tilde{\mathbf{y}}_n$ will have a large magnitude, otherwise, they will

have low magnitudes. The SCNR of the signal \mathbf{y}_n after beamforming is

$$\begin{aligned}
SCNR_n &= \frac{\mathbf{w}_n^H \mathbf{R}_Y \mathbf{w}_n}{\mathbf{w}_n^H \mathbf{Z} \Phi^T (\Phi^T)^H \mathbf{Z}^H \mathbf{w}_n} \\
&= \mathbf{w}_n^H \mathbf{R}_Y \mathbf{w}_n \left(\mathbf{v}_r^H(\theta_n) \tilde{\mathbf{R}}^{-1} \mathbf{v}_r(\theta_n) \right) \\
&= \frac{\mathbf{v}_r^H(\theta_n) \tilde{\mathbf{R}}^{-1}}{\mathbf{v}_r^H(\theta_n) \tilde{\mathbf{R}}^{-1} \mathbf{v}_r(\theta_n)} \mathbf{R}_Y \mathbf{w}_n \left(\mathbf{v}_r^H(\theta_n) \tilde{\mathbf{R}}^{-1} \mathbf{v}_r(\theta_n) \right) \\
&= \mathbf{v}_r^H(\theta_n) \tilde{\mathbf{R}}^{-1} \mathbf{R}_Y \mathbf{w}_n \\
&= \frac{\mathbf{v}_r^H(\theta_n) \tilde{\mathbf{R}}^{-1} \mathbf{R}_Y \tilde{\mathbf{R}}^{-1} \mathbf{v}_r(\theta_n)}{\mathbf{v}_r^H(\theta_n) \tilde{\mathbf{R}}^{-1} \mathbf{v}_r(\theta_n)}, \tag{3.6}
\end{aligned}$$

where $\mathbf{R}_Y = \left(\sum_{k=1}^K \beta_k \mathbf{v}_r(\theta_k) \mathbf{v}_t^T(\theta_k) \mathbf{X}^T \Phi^T \right) \left(\sum_{k=1}^K \beta_k^* (\Phi^T)^H (\mathbf{X}^T)^H \mathbf{v}_t^*(\theta_k) \mathbf{v}_r^H(\theta_k) \right)$.

Stacking $\tilde{\mathbf{y}}_n, n = 1, \dots, N$ into a vector, i.e., $\tilde{\mathbf{y}}$, which has a block-sparse like appearance (see Fig. 3.2), and which can be expressed as:

$$\begin{aligned}
\tilde{\mathbf{y}} &= [\tilde{\mathbf{y}}_1^T, \dots, \tilde{\mathbf{y}}_N^T]^T \\
&= \begin{bmatrix} \sum_{k=1}^N s_k \Phi \mathbf{X} \mathbf{v}_t(\theta_k) \mathbf{v}_r^T(\theta_k) \mathbf{w}_1^* \\ \vdots \\ \sum_{k=1}^N s_k \Phi \mathbf{X} \mathbf{v}_t(\theta_k) \mathbf{v}_r^T(\theta_k) \mathbf{w}_N^* \end{bmatrix} + \begin{bmatrix} \mathbf{Z}^T \mathbf{w}_1^* \\ \vdots \\ \mathbf{Z}^T \mathbf{w}_N^* \end{bmatrix} \\
&= \underbrace{\begin{bmatrix} \tilde{\Theta}_1^T, \dots, \tilde{\Theta}_N^T \end{bmatrix}^T}_{\tilde{\Theta}} \mathbf{s} + \underbrace{\begin{bmatrix} \mathbf{w}_1^H \mathbf{Z}, \dots, \mathbf{w}_N^H \mathbf{Z} \end{bmatrix}^T}_{\tilde{\mathbf{Z}}} \\
&= \tilde{\Theta} \mathbf{s} + \tilde{\mathbf{Z}}. \tag{3.7}
\end{aligned}$$

Here, $\tilde{\mathbf{y}}$ provides a high resolution target picture. However, by exploiting the structure of (3.7), we can further improve resolution. This can be achieved by compressing $\tilde{\mathbf{y}}$ through a matrix $\tilde{\Phi}$ and applying CS theory to recover the vector \mathbf{s} . We should note that the compression through $\tilde{\Phi}$ is not necessary but it reduces the dimensionality of the problem and thus the complexity of the recovery. If the measurement matrix $\tilde{\Phi}$ is

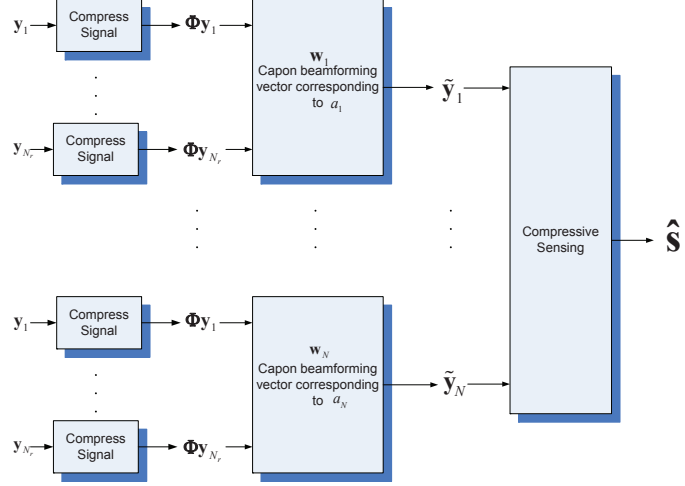


Figure 3.1: The diagram for the proposed CS-Capon method

an identity matrix, then the SCNR of CS-Capon method can be written as

$$\begin{aligned}
 & SCNR_{CS-Capon} \\
 &= \frac{\sum_{n=1}^N \mathbf{w}_n^T \left(\sum_{k=1}^K \beta_k^* \mathbf{v}_r^*(\theta_k) \mathbf{v}_t^H(\theta_k) \mathbf{X}^H \Phi^H \right) \left(\sum_{k=1}^K \beta_k \Phi \mathbf{X} \mathbf{v}_t(\theta_k) \mathbf{v}_r^T(\theta_k) \right) \mathbf{w}_n^*}{\sum_{n=1}^N \mathbf{w}_n^T \mathbf{Z}^* \Phi^H \Phi \mathbf{Z}^T \mathbf{w}_n^*} \\
 &= \frac{\sum_{n=1}^N \mathbf{w}_n^H \left(\sum_{k=1}^K \beta_k \mathbf{v}_r(\theta_k) \mathbf{v}_t^T(\theta_k) \mathbf{X}^T \Phi^T \right) \left(\sum_{k=1}^K \beta_k^* (\Phi^T)^H (\mathbf{X}^T)^H \mathbf{v}_t^*(\theta_k) \mathbf{v}_r^H(\theta_k) \right) \mathbf{w}_n}{\sum_{n=1}^N \mathbf{w}_n^H \mathbf{Z} \Phi^T (\Phi^T)^H \mathbf{Z}^H \mathbf{w}_n} \\
 &= \frac{\sum_{n=1}^N \mathbf{w}_n^H \mathbf{R}_Y \mathbf{w}_n}{\sum_{n=1}^N \mathbf{w}_n^H \mathbf{Z} \Phi^T (\Phi^T)^H \mathbf{Z}^H \mathbf{w}_n}. \tag{3.8}
 \end{aligned}$$

It is easy to verify that

$$\min \{SCNR_1, \dots, SCNR_N\} \leq SCNR_{CS-Capon} \leq \max \{SCNR_1, \dots, SCNR_N\}$$

3.3 Power Allocation in CS-based Clutter Suppression

The CS-based recovery requires that the product of the measurement matrix and the sparsifying basis matrix, referred to as the sensing matrix, satisfies the uniform uncertainty principle (UUP) [25] [34] [35]; in other words, the sensing matrix exhibits low

correlation between its columns. Let \mathbf{p} denote the transmit power vector for M_t transmit antennas and suppose the total power is P_t . Let \mathbf{u}_k denote the k -th column of the sensing matrix $\tilde{\Theta}$ defined in (3.7) and it can be written as

$$\mathbf{u}_k = [c_{k1}, \dots, c_{kN}] \otimes \Phi \mathbf{X} \mathbf{V}_t(a_k) \tilde{\mathbf{p}}, \quad (3.9)$$

where $\mathbf{V}_t(a_k) = \text{diag}(\mathbf{v}_t(a_k))$, $\tilde{\mathbf{p}} = \sqrt{P_t} \mathbf{p}$, $c_{kn} = \mathbf{v}_r^T(a_k) \mathbf{w}_n^*$ for $n = 1, \dots, N$.

The power allocation for CS-Capon can be formulated as the following optimization problem

$$\begin{aligned} \min_{\tilde{\mathbf{p}}} \quad & \left\| \tilde{\Theta}^H \tilde{\Theta} - \mathbf{I} \right\|_F^2 \\ \text{s.t.} \quad & \tilde{\mathbf{p}}^H \tilde{\mathbf{p}} = P_t, \tilde{\mathbf{p}} \geq \mathbf{0}_{M_t \times 1}, \tilde{\mathbf{p}} \leq \sqrt{P_m} \mathbf{1}_{M_t \times 1} \end{aligned} \quad (3.10)$$

The optimization problem (3.10) can be further formulated as

$$\begin{aligned} \min_{\tilde{\mathbf{p}}} \quad & \sum_{k \neq k'} |\mathbf{u}_k^H \mathbf{u}_{k'}|^2 + \sum_k |\mathbf{u}_k^H \mathbf{u}_k - 1|^2 \\ \text{s.t.} \quad & \tilde{\mathbf{p}}^H \tilde{\mathbf{p}} = P_t, \tilde{\mathbf{p}} \geq \mathbf{0}_{M_t \times 1}, \tilde{\mathbf{p}} \leq \sqrt{P_m} \mathbf{1}_{M_t \times 1} \end{aligned} \quad (3.11)$$

The correlation of the k -th and k' -th column in the sensing matrix can be written and simplified as

$$|\mathbf{u}_k^H \mathbf{u}_{k'}|^2 = u_{kk'} |\tilde{\mathbf{p}}^H \mathbf{B}_{kk'} \tilde{\mathbf{p}}|^2 \quad (3.12)$$

where $u_{kk'} = \left| \sum_{n=1}^N c_{kn}^* c_{k'n} \right|^2$, $\mathbf{B}_{kk'} = \mathbf{V}^H(a_{k'}) \mathbf{X}^H \Phi^H \Phi \mathbf{X} \mathbf{V}(a_k)$. It can be verified that matrix $\mathbf{B}_{kk'}$ is not positive semidefinite. To make sure the objective function in the optimization problem is convex, we use the same trick as that in [30] to rewritten the objective function. The first constraint in (3.11) is not convex. We use it local affine function to approximate it. The optimization problem (3.11) is solved recursively.

In practice, we may skip the compression step and collect a few samples at each RX antenna. If the transmit waveform is orthogonal, i.e., $\mathbf{X}^H \mathbf{X} = \mathbf{I}$, then $\mathbf{B}_{kk'}$ is a diagonal matrix, then the problem (3.11) has a simplified convex version

$$\begin{aligned} \min_{\mathbf{p}} \quad & \mathbf{p}^T \left(\sum_{k \neq k'} u_{kk'} \mathbf{b}_{kk'}^* \mathbf{b}_{kk'}^T \right) \mathbf{p} \\ \text{s.t.} \quad & \mathbf{1}_{M_t \times 1}^T \mathbf{p} = P_t, \mathbf{p} \geq \mathbf{0}, \mathbf{p} \leq P_m \mathbf{1}_{M_t \times 1} \end{aligned} \quad (3.13)$$

where $\mathbf{b}_{kk'}$ is the diagonal vector of $\mathbf{B}_{kk'}$.

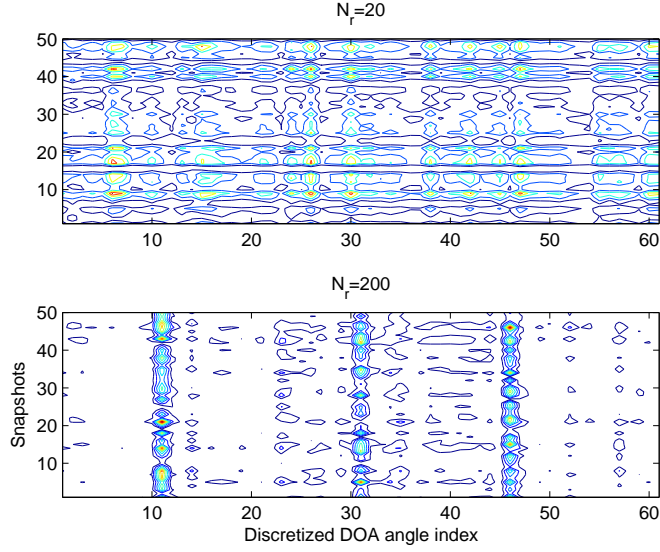


Figure 3.2: Contour of the beamformed receive data matrix for $N_r = 20$ and $N_r = 200$, respectively.

3.4 Simulations

We consider a MIMO radar system with TX and RX antennas uniformly located on a disk of radius 10m. The carrier frequency is $f = 5\text{GHz}$. Each TX antenna uses Hadamard orthogonal waveform sequence of length $L = 128$ and unit power. The received signal is corrupted by zero-mean Gaussian noise of unit power. Three targets are present on the angle grid $[-30^\circ, -29^\circ, \dots, 30^\circ]$. The signal-to-noise ratio (SNR) is set as 0dB. The clutter signal is constructed as the sum of returns reflected by 3000 reflectors, located in the DOA angle space $[-30^\circ, -30^\circ + 60/1500^\circ, \dots, 30^\circ]$. The reflection coefficient of each reflector is 0.15. Random matrix is used as the measurement matrix $\Phi \in \mathcal{R}^{M \times L}$ with $M = 50$.

Figure 3.2 shows the contour of the beamformed receive data matrix $\mathbf{Y} = [\tilde{\mathbf{y}}_1, \dots, \tilde{\mathbf{y}}_N]$ for $N_r = 20$ and $N_r = 200$, respectively. In both cases, the number of TX antennas is set as $M_t = 20$. Assume there are $K = 3$ targets reflected from directions $[-20^\circ, 0^\circ, 15^\circ]$ and the target reflect coefficients are set as $[1, 1, 1]$. It can be found from Fig. 3.2 that, as N_r is large enough, there are $K = 3$ columns with large amplitude in matrix \mathbf{Y} whose column indexes are corresponding to the indexes of each target's DOA in the

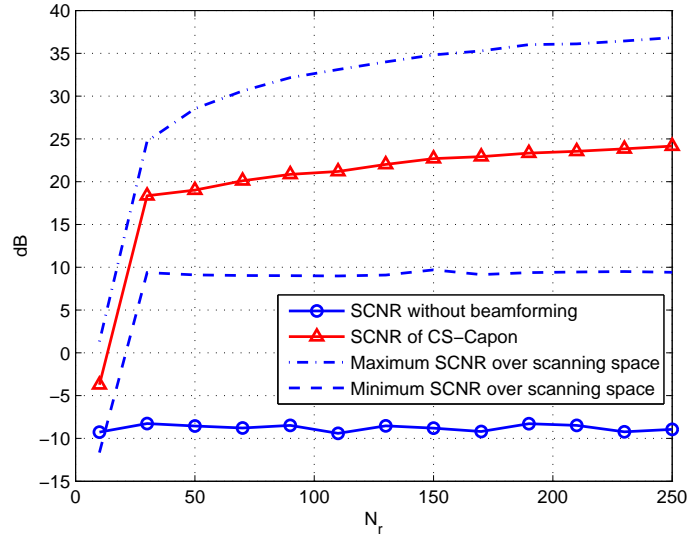


Figure 3.3: SCNR comparisons with/without beamforming with $N_t = 20$.

discretized angle space. This indicates that the stacked receive data vector $\tilde{\mathbf{y}}$ defined in (3.7) is relatively block-sparse after applying the Capon-beamformer.

Figure 3.3 shows the averaged signal-to-clutter-plus-noise ratio (SCNR) versus the number of RX antennas. The SCNR without beamforming is obtained according to (3.2). The SCNR of CS-Capon is defined according to (3.8). Here, we also plot the maximum and minimum of SCNR for all beamformed receive vector $\tilde{\mathbf{y}}_n$, $n = 1, \dots, N$. In simulations, $K = 3$ targets are generated randomly on the DOA grid $[-30^\circ, -29^\circ, \dots, 30^\circ]$ and the target reflect coefficients are set as $[1, 1, 1]$. The clutter signal is constructed as the sum of returns reflected by 3000 reflectors, located in the DOA angle space $[-30^\circ, -30^\circ + 60/1500^\circ, \dots, 30^\circ]$. The reflection coefficient of each reflector is 0.15. The number of TX antennas is set as $N_t = 20$. For each N_r , 100 runs have been carried out and average SCNR is calculated. It can be found from Fig. 3.3 that, without beamforming, the SCNR of the received signal defined in (3.1) is around -10 dB and does not change much with the increase of N_r . With beamforming, the SCNR of the CS-Capon estimator defined in (3.8) is around 20 dB when $N_r = 50$ and would go up with the increase of N_r . This indicates that the proposed CS-Capon

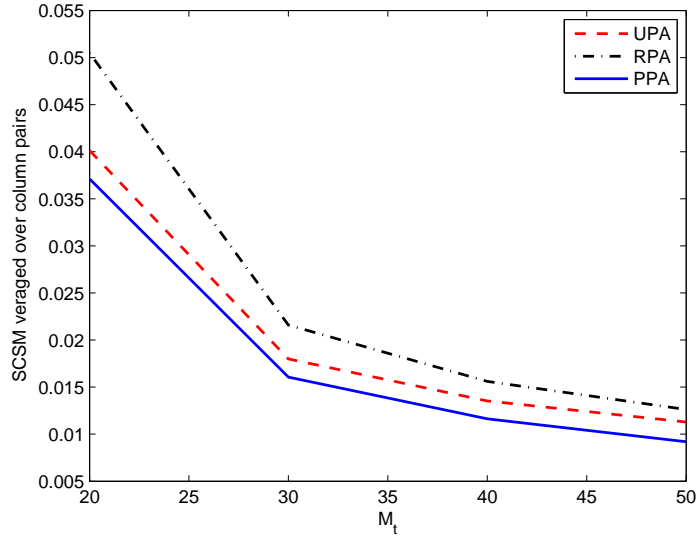


Figure 3.4: SCSM versus M_t under optimal (PPA), uniform (UPA) and random (RPA) power allocation in CS-Capon method with $N_r = 20$.

method could achieve a significant gain in SCNR performance, i.e., a strong clutter suppression capability. We can also find from Fig. 3.3 that, the SCNR of the CS-Capon estimator falls in the middle between the maximum and minimum SCNR of all beamformed vector $\tilde{\mathbf{y}}_n$, $n = 1, \dots, N$. This is because SCNR of the CS-Capon estimator defined in (3.8) is a divisor of the averaged beamformed vector in the numerator and denominator, respectively.

Figure 3.4 shows the distribution of the sum of the square magnitude (SCSM) of the cross-correlation of column pairs in the sensing matrix of the CS-Capon estimator versus M_t for optimal (PPA), uniform (UPA) and random (RPA) power allocation schemes, respectively. For simplicity, in the simulations, the identity matrix is used as the measurement matrix and the optimal power allocation is done by solving the convex optimization problem (3.13). The number of RX antennas is set as $N_r = 20$. The total power budget for all TX antennas equals M_t . The DOA angle search space is discretized into $[-8^\circ, -8^\circ + 0.2^\circ, \dots, 8^\circ]$. The clutter signal is constructed as the sum of returns reflected by 300 reflectors, located in the DOA angle space $[-8^\circ, -8^\circ + 16/300^\circ, \dots, 8^\circ]$. The reflection coefficient of each reflector is 0.15. It can be found from Fig. 3.4 that

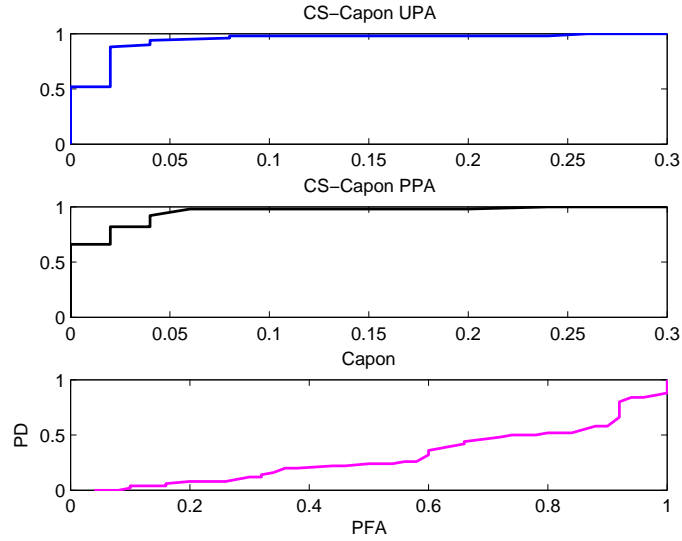


Figure 3.5: ROC performances under optimal (PPA), uniform (UPA) power allocation in CS-Capon as well as Capon method with $N_t = 50$ and $N_r = 20$.

the coherence under PPA is smaller than the UPA and RPA. With the increase of M_t , the coherence under all power allocation schemes would decrease.

Figure 3.5 plots the ROC curves of the angle estimates that are obtained based on 50 random and independent runs. In each run, $K = 3$ targets are randomly generated on the angle grid of interest $[-8^\circ, -8^\circ + 0.2^\circ, \dots, 8^\circ]$ with random reflection coefficients drawing from $[0.4, 1]$, respectively. In the simulations, $N_t = 50$, $N_r = 20$, and SNR is set to 0dB. Identity matrix is used as the measurement matrix. One can see that the proposed CS-Capon method outperforms the CS method and Capon method. In addition, the optimal power allocation could further improve the detection performance in the CS-Capon estimator due to its capability to reduce the coherence of sensing matrix.

Note: If the reflection coefficient of each target is identical, Capon performs as well as the proposed CS-Capon method. While the reflection coefficients of three targets are different, the ripples generated by Capon may mask the targets with smaller reflectivity.

3.5 Summary

We have considered the scenario that strong clutter is present in the context of colocated MIMO radars and thus sparsity condition might not hold for compressive sensing based MIMO radars. By applying the Capon beamforming to the compressed received data, we showed that the beamformed compressed data is block sparse and its signal-to-clutter-noise ratio (SCNR) is improved greatly. Compressive sensing is subsequently applied to the beamformed compressed data for target estimation. Based on the uniform uncertainty principle, we further proposed a power allocation scheme to minimize the coherence between the columns of the sensing matrix in CS, and thus to improve the CS performance. Simulation results showed accurate target estimation could be achieved using CS when there is strong clutter in the colocated MIMO radar scenarios.

Chapter 4

A MIMO Radar Approach Based on Matrix Completion

In a typical MIMO radar scenario, the receive nodes transmit to a fusion center either samples of the target returns, or the results of matched filtering with the transmit waveforms. Based on the data it receives from multiple antennas, the fusion center formulates a matrix, referred to as the data matrix, which, in conjunction with standard array processing schemes leads to target detection and parameter estimation. In this chapter, it is shown that under certain conditions, the data matrix is low-rank, and thus can be recovered based on knowledge of a small subset of its entries via matrix completion (MC) techniques. Leveraging the low-rank property of that matrix, we propose a new MIMO radar approach, termed, MIMO-MC radar, in which each receive node either performs matched filtering with a small number of randomly selected dictionary waveforms or obtains sub-Nyquist samples of the target returns at random sampling instants, and forwards the results to a fusion center. Based on the received samples, and with knowledge of the sampling scheme, the fusion center partially fills the data matrix and subsequently applies MC techniques to estimate the full matrix. MIMO-MC radars share the advantages of the MIMO radars with compressive sensing, (MIMO-CS), i.e., high resolution with reduced amounts of data, but unlike MIMO-CS radars do not require grid discretization. The MIMO-MC radar concept is illustrated through a linear uniform array configuration, and its target estimation performance is demonstrated via simulations.

4.1 Introduction

MIMO radar systems have received considerable attention in recent years due to their superior resolution [17] [18] [66]. The MIMO radars using compressed sensing (MIMO-CS) maintain the MIMO radars advantages, while significantly reducing the required measurements per receive antenna [28] [29]. In MIMO-CS radars, the target parameters are estimated by exploiting the sparsity of targets in the angle, Doppler and range space, referred to as the *target space*; the target space is discretized into a fine grid, based on which a compressive sensing matrix is constructed, and the target is estimated via sparse signal recovery techniques, such as the Dantzig selector [29]. However, the performance of CS-based MIMO radars degrades when targets fall between grid points, a case also known as basis mismatch [39] [40].

In this chapter, a novel approach to lower-complexity, higher-resolution radar is proposed, termed MIMO-MC radars, which stands for MIMO radars using matrix completion (MC). MIMO-MC radars achieve the advantages of MIMO-CS radars without requiring grid discretization. Matrix completion is of interest in cases in which we are constrained to observe only a subset of the entries of an $n_1 \times n_2$ matrix, because the cost of collecting all entries of a high dimensional matrix is high. If a matrix is low rank and satisfies certain conditions [41], it can be recovered *exactly* based on observations of a small number of its randomly selected entries. There are several MC techniques in the literature [41–46]. For example, in [41–43], recovery can be performed by solving a nuclear norm optimization problem, which basically finds the matrix with the smallest nuclear norm out of all possible matrices that fit the observed entries. Other matrix completion techniques are based on non-convex optimization using matrix manifolds, such as Grassmann manifold [44, 45], and Riemann manifolds [46].

In a typical MIMO radar scenario [66], transmit nodes transmit orthogonal waveforms, while each receive node performs matched filtering with the known set of transmit waveforms, and forwards the results to the fusion center. Based on the data it receives from multiple antennas, the fusion center formulates a matrix, which, in conjunction with standard array processing schemes, such as MUSIC [5], leads to target detection

and estimation. In this paper, we show that under certain conditions, the data matrix at the fusion center is low-rank, which means that it can be recovered based on knowledge of a small subset of its entries via matrix completion (MC) techniques. Leveraging the low-rank property of that matrix, we propose MIMO-MC radar, in which, each receive antenna either performs matched filtering with a small number of dictionary waveforms or obtains sub-Nyquist samples of the received signal and forwards the results to a fusion center. Based on the samples forwarded by all receive nodes, and with knowledge of the sampling scheme, the fusion center applies MC to estimate the full matrix. Although the proposed ideas apply to arbitrary transmit and receive array configurations, in which the antennas are not physically connected, in this paper we illustrate the idea through a linear uniform array configuration. The properties and performance of the proposed scheme are demonstrated via simulations. Compared to MIMO-CS radars, MIMO-MC radars have the same advantage in terms of reduction of samples needed for accurate estimation, while they avoid the basis mismatch issue, which is inherent in MIMO-CS radar systems.

4.2 Colocated MIMO Radar System Model

Let us consider a MIMO pulse radar system that employs colocated transmit and receive antennas, as shown in Fig. 1.4. We use M_t and M_r to denote the numbers of transmit and receive antennas, respectively. Although our results can be extended to an arbitrary antenna configuration, we illustrate the ideas for uniform linear arrays (ULAs). The inter-element spacing in the transmit and receive arrays is denoted by d_t and d_r , respectively. The pulse duration is T_p , and the pulse repetition interval is T_{PRI} . The waveform of the i th transmit antenna is $s_i(\tau) = \sqrt{\frac{E}{M_t}}\phi_i(\tau)$, where E is the total energy for all the transmit antennas, and $\phi_i(\tau)$, $i = 1, \dots, M_t$ are orthonormal. The waveforms are transmitted over a carrier with wavelength λ . Let us consider a scenario with K point targets in the far field at angles θ_k , $k = 1, \dots, K$, each moving with speed v_k .

The following assumptions are made:

- The transmit waveforms are narrowband, i.e., $\frac{1}{T_p} \ll \frac{c}{\lambda}$, where c is the speed of light.
- The target reflection coefficients $\{\beta_k\}, k = 1, \dots, K$ are complex and remain constant during a number of pulses, Q . Also, all parameters related to the array configuration remain constant during the Q pulses.
- The delay spread in the receive signals is smaller than the temporal support of pulse T_p .
- The Doppler spread of the receive signals is much smaller than the bandwidth of the pulse, i.e., $\frac{2\vartheta}{\lambda} \ll \frac{1}{T_p}$.

Under the narrowband transmit waveform assumption, the delay spread in the baseband signals can be ignored. For slowly moving targets, the Doppler shift within a pulse can be ignored, while the Doppler changes from pulse to pulse. Thus, if we express time as $t = qT_{PRI} + \tau$, where q is the pulse index (or slow time) and $\tau \in [0, T_p]$ is the time within a pulse (or fast time), the Doppler shift will depend on q only, and the received signal at the l -th receive antenna can be approximated as [66]

$$x_l \left(qT_{PRI} + \tau + \frac{2d}{c} \right) \approx \sum_{k=1}^K \beta_k e^{j \frac{2\pi}{\lambda} (2\vartheta_k (q-1)T_{PRI} + (l-1)d_r \sin(\theta_k))} \mathbf{a}^T(\theta_k) \mathbf{s}(\tau) + w_l \left(qT_{PRI} + \tau + \frac{2d}{c} \right), \quad (4.1)$$

where d is the distance of the range bin of interest; w_l contains both interference and noise;

$$\mathbf{a}(\theta_k) = \left[1, e^{j \frac{2\pi}{\lambda} d_t \sin(\theta_k)}, \dots, e^{j \frac{2\pi}{\lambda} (M_t-1) d_t \sin(\theta_k)} \right]^T, \quad (4.2)$$

and $\mathbf{s}(\tau) = [s_1(\tau), \dots, s_{M_t}(\tau)]^T$. For convenience, the signal parameters are summarized in Table 4.1.

At the l -th receive node, for $(l = 1, \dots, M_r)$, a matched filter bank [66] is used to extract the returns due to each transmit antenna [66] (see Fig. 4.1 (a)). Consider a filter bank composed of M_t filters, corresponding to the M_t orthogonal transmit waveforms. The receive node performs M_t correlation operations and the maximum of each matched

Table 4.1: List of parameters used in the signal model

d_t	spacing between the transmit antennas
d_r	spacing between the receive antennas
M_t	number of transmit antennas
M_r	number of receive antennas
Q	number of pulses in a coherent processing interval
T_{PRI}	radar pulse repetition interval
q	index of radar pulse (slow time)
τ	time in one pulse (fast time)
v	speed of target
ϕ_m	baseband waveform
d	distance of range bin of interest
c	speed of light
θ	direction of arrival of the target
β	target reflect coefficient
λ	wavelength of carrier signal
w_l	interference and white noise in the l th antenna
T_p	duration of one pulse
T_s	Nyquist sampling period

filter is forwarded to the fusion center. At the fusion center, the received signal due to the i -th matched filter of the l -th receive node, during the q -th pulse, can be expressed as

$$x_q(l, i) = \sum_{k=1}^K \beta_k e^{j \frac{2\pi}{\lambda} (2v_k(q-1)T_{PRI} + (l-1)d_r \sin(\theta_k) + (i-1)d_t \sin(\theta_k))} + w_q(l, i) \quad (4.3)$$

for $l = 1, \dots, M_r$, $i = 1, \dots, M_t$, and $q = 1, \dots, Q$, where $w_q(l, i)$ is the corresponding interference plus white noise.

Based on the data from all receive antennas, the fusion center can construct a matrix \mathbf{X}_q^{MF} , of size $M_r \times M_t$, whose (l, i) element equals $x_q(l, i)$. That matrix can be expressed

as

$$\mathbf{X}_q^{MF} = \underbrace{\mathbf{B}\mathbf{\Sigma}\mathbf{D}_q\mathbf{A}^T}_{\mathbf{Z}_q^{MF}} + \mathbf{W}_q^{MF}, \quad (4.4)$$

where \mathbf{W}_q^{MF} is the filtered noise; $\mathbf{\Sigma} = \text{diag}([\beta_1, \dots, \beta_K])$; $\mathbf{D}_q = \text{diag}(\mathbf{d}_q)$, with $\mathbf{d}_q = [e^{j\frac{2\pi}{\lambda}2\vartheta_1(q-1)T_{PRI}}, \dots, e^{j\frac{2\pi}{\lambda}2\vartheta_K(q-1)T_{PRI}}]^T$; \mathbf{A} is the $M_t \times K$ transmit steering matrix, defined as $\mathbf{A} = [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_K)]$; \mathbf{B} is the $M_r \times K$ dimensional receive steering matrix, defined in a similar fashion based on the receive steering vectors

$$\mathbf{b}(\theta_k) = \left[1, e^{j\frac{2\pi}{\lambda}d_r \sin(\theta_k)}, \dots, e^{j\frac{2\pi}{\lambda}(M_r-1)d_r \sin(\theta_k)}\right]^T. \quad (4.5)$$

4.2.1 MIMO Radars without Matched Filtering

In this scenario, each receive antenna performs Nyquist sampling to obtains N samples from the target returns and forwards them to the fusion center (see Fig. 4.1 (b)). If the data forwarded by the l -th antenna ($l = 1, \dots, M_r$) during pulse q are inserted in the l -th row of an $M_r \times N$ matrix, \mathbf{X}_q , then, an equation similar to (4.4) holds, except that now the transmit waveforms also appear in the expression, i.e., [67]

$$\mathbf{X}_q = \underbrace{\mathbf{B}\mathbf{\Sigma}\mathbf{D}_q\mathbf{A}^T}_{\mathbf{Z}_q} \mathbf{S} + \mathbf{W}_q, \quad (4.6)$$

where $\mathbf{S} = [\mathbf{s}(0T_s), \dots, \mathbf{s}((N-1)T_s)] \in \mathbb{C}^{M_t \times N}$.

4.3 The proposed MIMO-MC radar approach

Looking at (4.4), if $M_t > K$ and $M_r > K$, both matrices $\mathbf{\Sigma}$ and \mathbf{D}_q are rank- K . Thus, the rank of the noise free matrix $\mathbf{Z}_q^{MF} \in \mathbb{C}^{M_r \times M_t}$ is K , which implies that matrix \mathbf{Z}_q^{MF} is low-rank if both M_t and M_r are much larger than K .

Similarly, looking at (4.6), both matrices $\mathbf{\Sigma}$ and \mathbf{D}_q are rank- K . The rank of matrix \mathbf{S} is $\min\{M_t, N\}$. Let us assume that $N > M_t$. For $M_t > K$, the rank of the noise free data matrix $\mathbf{Z}_q \in \mathbb{C}^{M_r \times N}$ is K . In other words, for $M_r \gg K$ the data matrix \mathbf{Z}_q is low-rank.

Therefore, in both sampling schemes, assuming that the conditions **(A0)** and **(A1)** are satisfied, the fusion center matrix can be recovered from a small number of its

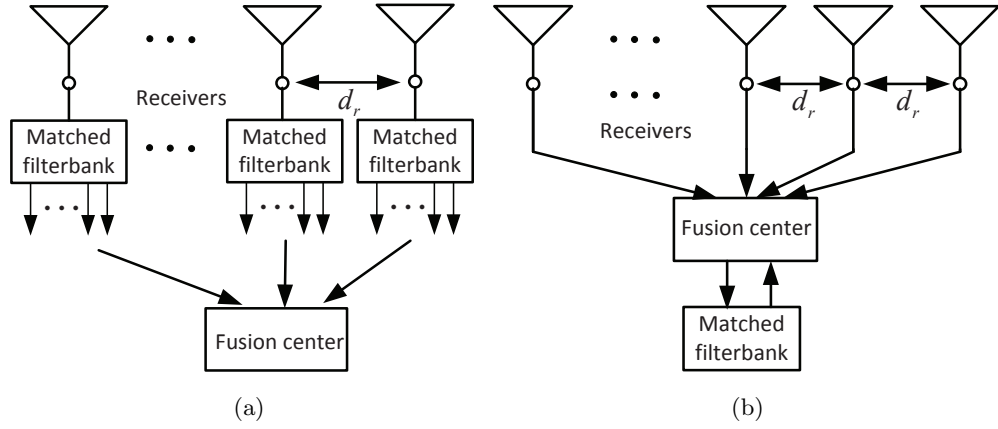


Figure 4.1: Two sampling schemes in the colocated MIMO radars system: (a) Sampling scheme I; (b) Sampling scheme II.

entries. The estimated matrices corresponding to several pulses can be used to estimate the target parameters via MUSIC [5], for example.

In the following, we leverage the low-rank property of the data matrices at the fusion center to propose a new MIMO radar approach. Since both \mathbf{Z}_q and \mathbf{Z}_q^{MF} are formulated based on different sampling schemes at the receive nodes, we will study two cases, namely, sampling scheme I, which gives rise to \mathbf{Z}_q^{MF} , and sampling scheme II, which gives rise to \mathbf{Z}_q .

4.3.1 MIMO-MC with Sampling Scheme I

Suppose that the l th receive node uses a random matched filter bank (RMFB), as shown in Fig. 4.2, in which, a random switch unit is used to turn on and off each matched filter. Suppose that L_1 matched filters are selected at random out of the M_t available filters, according to the output of a random number generator, returning L_1 integers in $[0, M_t - 1]$ based on the seed s_l . Let \mathcal{J}^l denote the set of indices of the selected filters. The same random generator algorithm is also available to the fusion center. The l th receive antenna forwards the L_1 samples along with the seed s_l to the fusion center. Based on the seed s_l , the fusion center generates the indices \mathcal{J}^l . Then, it places the j -th sample of the l -th antenna in the $M_r \times M_t$ matrix \mathbf{Z}_q^{MF} at location $(l, \mathcal{J}^l(j))$. In total, $L_1 M_t$ entries of the matrix are filled. The fusion center declares the rest of

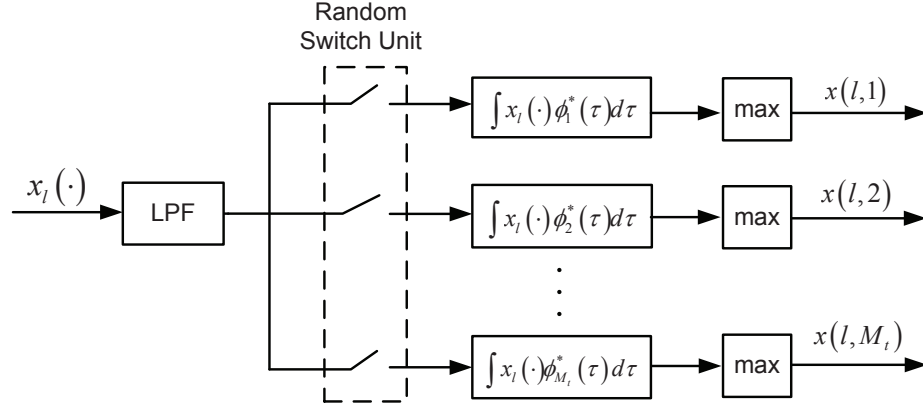


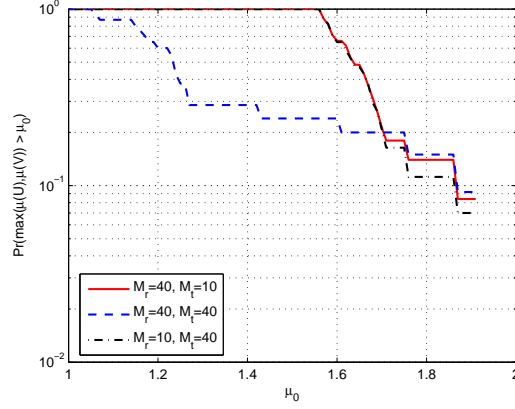
Figure 4.2: Structure of the random matched filter bank (RMFB).

the entries as “missing” and assuming that \mathbf{Z}_q^{MF} meets **(A0)** and **(A1)**, applies MC techniques to estimate the full data matrix.

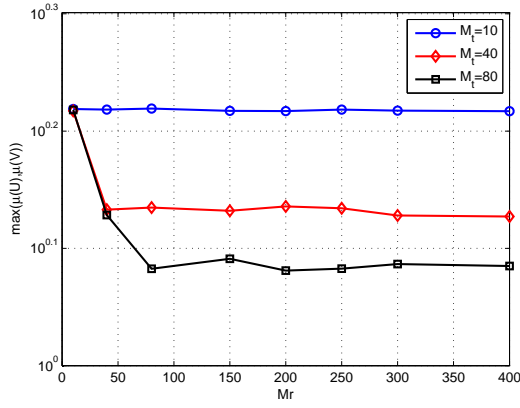
Since the samples forwarded by the receive nodes are obtained in a random sampling fashion, the filled entries of \mathbf{Z}_q^{MF} will correspond to a uniformly random sampling of \mathbf{Z}_q^{MF} . In order to show that \mathbf{Z}_q^{MF} indeed satisfies **(A0)**, and as a result **(A1)**, we need to show that the maximum coherence of the spaces spanned by the left and right singular vectors of \mathbf{Z}_q^{MF} is bounded by a number, μ_0 . The smaller that number, the fewer samples of \mathbf{Z}_q^{MF} will be required for estimating the matrix. The theoretical analysis is pursued separately in [68]. Here, we confirm the applicability of MC techniques via simulations.

We consider a scenario with $K = 2$ point targets. The DOA of the first target, θ_1 , is taken to be uniformly distributed in $[-90^\circ, 90^\circ]$, while the DOA of the second target is taken to be $\theta_2 = \theta_1 + \Delta\theta$. The target speeds are taken to be uniformly distributed in $[0, 500]$ m/s, and the target reflectivities, β_k are taken to be zero-mean Gaussian. Both the transmit and receive arrays follow the ULA model with $d_t = d_r = \frac{\lambda}{2}$. The carrier frequency is taken as $f = 1 \times 10^9$ Hz.

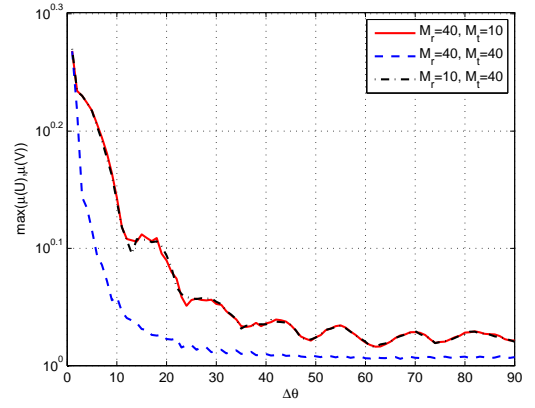
The left and right singular vectors of \mathbf{Z}_q^{MF} were computed for 500 independent realizations of θ_1 and target speeds. The probability that $\max(\mu(U), \mu(V)) > \mu_0$ among all the runs is shown in Fig. 4.3 (a) for $\Delta\theta = 5^\circ$ and different values of M_r, M_t . One can see from the figure that in all cases, the probability that the coherence is



(a)



(b)



(c)

Figure 4.3: Scheme I, $K = 2$ targets: (a) the probability of $\Pr(\max(\mu(U), \mu(V)) > \mu_0)$ of \mathbf{Z}_q^{MF} for $\Delta\theta = 5^\circ$; (b) the average $\max(\mu(U), \mu(V))$ of \mathbf{Z}_q^{MF} as function of number of transmit and receive antennas, and for $\Delta\theta = 5^\circ$; (c) the average $\max(\mu(U), \mu(V))$ of \mathbf{Z}_q^{MF} as function of DOA separation.

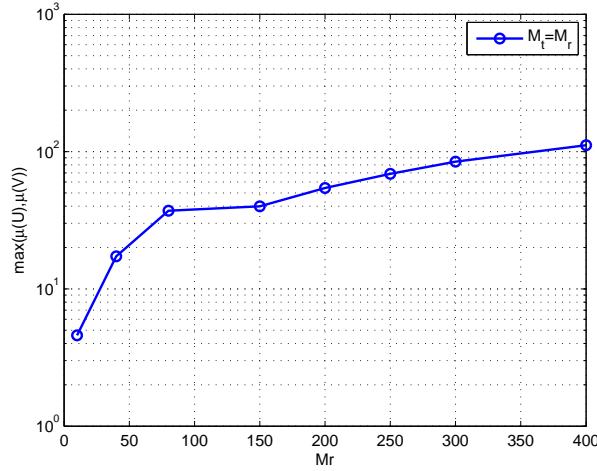


Figure 4.4: Scheme I, $K = 2$ targets: The $\max(\mu(U), \mu(V))$ in terms of M_r for $\Delta\theta = 0^\circ$, $M_t = M_r$.

bounded by a number less than 2 is very high, while the bound gets tighter as the number of receive or transmit antennas increases. On the average, over all independent realizations, the $\max(\mu(U), \mu(V))$ corresponding to different number of receive and transmit antennas and fixed $\Delta\theta$, appears to decrease as the number of transmit and receive antennas increases (see Fig. 4.3 (b)). Also, the maximum appears to decrease as $\Delta\theta$ increases, reaching 1 for large $\Delta\theta$ (see Fig. 4.3 (c)). The rate at which the maximum reaches 1 increases as the number of antennas increases.

It is interesting to see what happens at the limit $\Delta\theta = 0$, i.e., when the two targets are on a line in the angle plane. Computing the coherence based on the assumption of rank 2, i.e., using two eigenvectors, the coherence shown in Fig. 4.4 appears unbounded as M_r changes. However, in this case, the true rank of \mathbf{Z}_q^{MF} is 1, and \mathbf{Z}_q^{MF} has the best possible coherence. Indeed, as it is shown in the Appendix A, for a rank-1 \mathbf{Z}_q^{MF} , it holds that $\mu_0 = \mu_1 = 1$. Consequently, according to Theorem 1, the required number of entries to estimate \mathbf{Z}_q^{MF} is minimal. This explains why in Fig. 4.8 (discussed further in Section 4.4) the relative recovery error of \mathbf{Z}_q^{MF} goes to the reciprocal of SNR faster when the two targets have the same DOA. Of course, in this case, the two targets with the same DOA appear as one, and cannot be separated in the angle space unless other parameters, e.g., speed or range are used. For multiple targets, i.e., for $K \geq 3$, if there

are n ($n < K$) targets with the same DOA, the rank of \mathbf{Z}_q^{MF} is $K - n$, which yields a low coherence condition since these $K - n$ DOAs are separated.

4.3.2 MIMO-MC with Sampling Scheme II

Suppose that the Nyquist rate samples of signals at the receive nodes correspond to sampling times $t_i = iT_s$, $i = 0, \dots, N - 1$ with $N = T_p/T_s$. Instead of the receive nodes sampling at the Nyquist rate, let the l -th receive antenna sample at times $\tau_j^l = jT_s$, $j \in \mathcal{J}^l$, where \mathcal{J}^l is the output of a random number generator, containing L_2 integers in the interval $[0, N - 1]$ according to a unique seed s_l . The l -th receive antenna forwards the L_2 samples along with the seed s_l to the fusion center. Under the assumption that the fusion center and the receive nodes use the same random number generator algorithm, the fusion center places the j -th sample of the l -th antenna in the $M_r \times N$ matrix \mathbf{Z}_q at location $(l, \mathcal{J}^l(j))$, and declares the rest of the samples as “missing”. Therefore, under conditions (A0) and (A1), \mathbf{Z}_q can be estimated based on $m = L_2 M_r$ elements, for m sufficiently large.

The left singular vectors of \mathbf{Z}_q are the eigenvectors of $\mathbf{Z}_q \mathbf{Z}_q^H = \mathbf{H} \mathbf{S} \mathbf{S}^H \mathbf{H}^H$, where $\mathbf{H} = \mathbf{B} \mathbf{\Sigma} \mathbf{D}_q \mathbf{A}^T$. The right singular vectors of \mathbf{Z}_q are the eigenvectors of $\mathbf{S}^H \mathbf{H}^H \mathbf{H} \mathbf{S}$. Since the transmit waveforms are orthogonal, it holds that $\mathbf{S} \mathbf{S}^H = \mathbf{I}$ [20]. Thus, the left singular vectors are only determined by matrix \mathbf{H} , while the right singular vectors are affected by both transmit waveforms and matrix \mathbf{H} .

Again, to check whether \mathbf{Z}_q satisfies the conditions for MC, we resort to simulations. In particular, we show that the maximum coherence of \mathbf{Z}_q is bounded by a small positive number μ_0 . Assume there are $K = 2$ targets. The DOA of the first target, θ_1 , is uniformly distributed in $[-90^\circ, 90^\circ]$ and the DOA of the second target is set as $\theta_1 + \Delta\theta$. The corresponding speeds are uniformly distributed in $[150, 450]$ m/s. The target reflectivities, β_k , are zero-mean, Gaussian distributed. The transmit waveforms are taken to be complex Gaussian orthogonal (G-Orth). The carrier frequency is $f = 10^9$ Hz, resulting in $\lambda = c/f = 0.3$ m. The inter-spacing between transmit and receive antennas is set as $d_t = d_r = \lambda/2$, respectively.

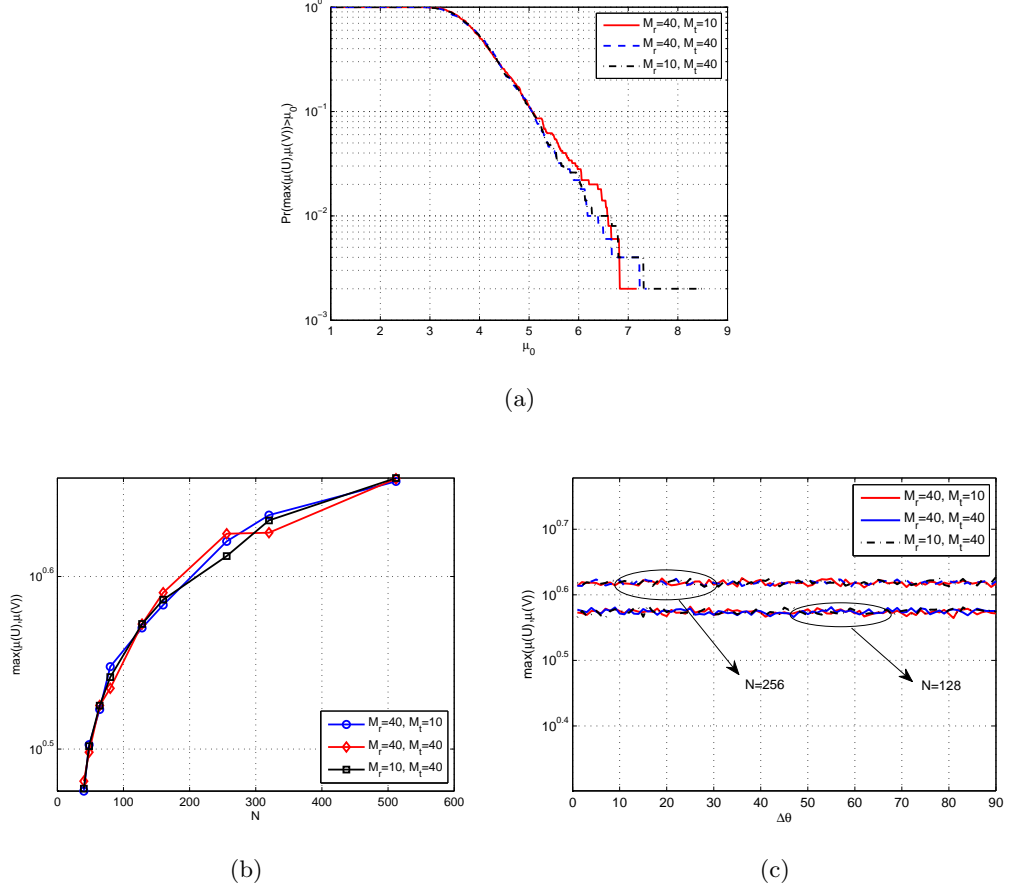


Figure 4.5: Scheme II, $K = 2$ targets, and G-Orth waveforms: (a) The probability of $\Pr(\max(\mu(U), \mu(V)) > \mu_0)$ of \mathbf{Z}_q for $\Delta\theta = 5^\circ$ and $N = 256$; (b) The average $\max(\mu(U), \mu(V))$ of \mathbf{Z}_q as function of N , for $\Delta\theta = 5^\circ$ and different values of M_t, M_r ; (c) The average $\max(\mu(U), \mu(V))$ of \mathbf{Z}_q as function of $\Delta\theta$, for $N = 128, 256$, and different combinations of M_r, M_t .

The left and right singular vectors of \mathbf{Z}_q are computed for 500 independent realizations of θ_1 and target speeds. Among all the runs, the probability that the $\max(\mu(U), \mu(V)) > \mu_0$ is shown in Fig. 4.5, for different values of M_t, M_r , $\Delta\theta = 5^\circ$, and $N = 256$. One can see from the figure that in all cases, the probability that the coherence is bounded by a number less than 7 is very high, while the bound gets tighter as the number of receive or transmit antennas increases. On average, over all independent realizations, the $\max(\mu(U), \mu(V))$ corresponding to different values of M_t, M_r and a fixed $\Delta\theta$ appears to increase with N , (see Fig. 4.5 (b)), while the increase is not

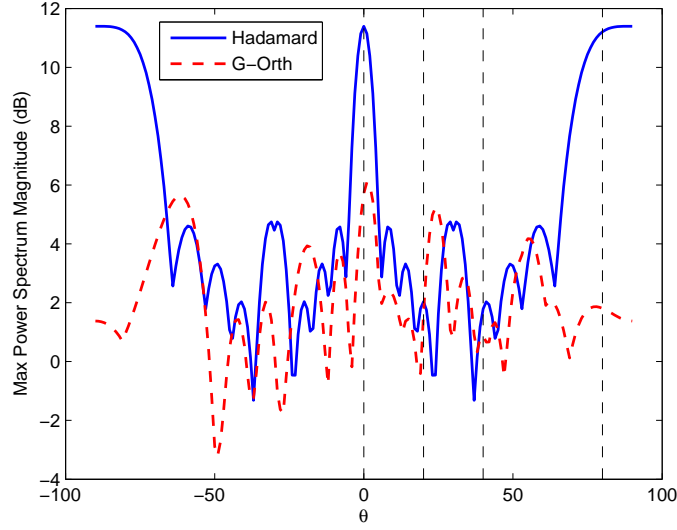


Figure 4.6: The maximal power spectrum of the orthogonal waveforms over $N = 32$ snapshots for $M_t = 10$.

affected by the number of transmit and receive antennas. The average maximum does not appear to change as $\Delta\theta$ increases, and this holds for various values of M_t, N (see Fig. 4.5 (c)).

Based on our simulations, the MC reconstruction depends on the waveform. In particular, the coherence bound is related to the power spectrum of each column of the waveform matrix (each column can be viewed as a waveform snapshot across the transmit antennas). Let $\tilde{S}_i(\omega)$ denote the power spectrum of the i -th column of $\mathbf{S} \in \mathbb{C}^{M_t \times N}$. If $\tilde{S}_i(\omega)$ is similar for different i 's, the MC recovery performance improves with increasing M_t (or equivalently, the coherence bound decreases) and does not depend on N ; otherwise, the performance worsens with increasing N (i.e., the coherence bound increases). When the $\tilde{S}_i(\omega)$ has peaks at certain ω 's that occur close to targets, the performance worsens. In Fig. 4.6, we show the maximum power spectra values corresponding to Hadamard and G-Orth waveforms for $M_t = 10$ and $N = 32$. It can be seen in Fig. 4.6 that the maximum power spectrum values corresponding to the Hadamard waveform have strong peaks at certain ω 's, while those for the G-Orth waveforms fluctuate around a low value. Suppose that there are two targets at angles $\theta_1 = 20^\circ$ and

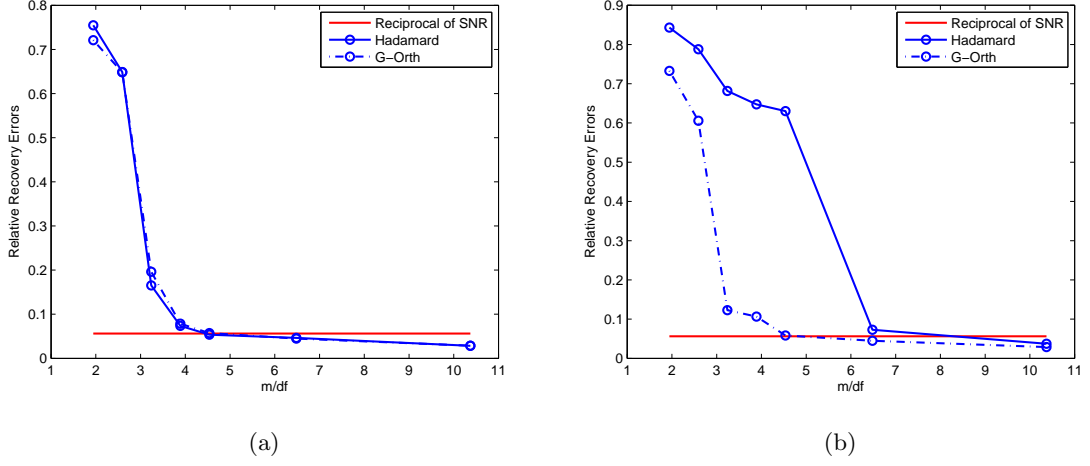


Figure 4.7: The comparison of matrix completion in terms of relative recovery errors with $M_r = 128, M_t = 10, N = 32, \text{SNR} = 25\text{dB}$. There are $K = 2$ targets located at (a) 20° and 40° ; (b) 0° and 80° .

$\theta_2 = 40^\circ$, corresponding to $\omega_1 = \frac{1}{2} \sin\left(\frac{\pi}{9}\right)$ and $\omega_2 = \frac{1}{2} \sin\left(\frac{2\pi}{9}\right)$, respectively. From Fig. 4.6 one can see that the targets fall under low power spectral values for both waveform cases. The corresponding MC recovery error, computed based on 50 independent runs is shown in Fig. 4.7 (a). One can see that the error is the same for both waveforms. As another case, suppose that the two targets are at angles $0^\circ, 80^\circ$, corresponding to $\omega_1 = 0, \omega_2 = \frac{1}{2} \sin\left(\frac{4\pi}{9}\right)$, respectively. Based on Fig. 4.6, one can see that ω_1 and ω_2 fall under high spectral peaks in the case of Hadamard waveforms. The corresponding MC recovery error is shown in Fig. 4.7(b), where one can see that Hadamard waveforms yield higher error.

4.3.3 Discussion of MC in Sampling Schemes I and II

To apply the matrix completion techniques in colocated MIMO radars, the data matrices $\mathbf{Z}_q \in \mathbb{C}^{M_r \times N}$ and $\mathbf{Z}_q^{MF} \in \mathbb{C}^{M_r \times M_t}$ need to be low-rank, and satisfy the coherence conditions with small $\mu_i, i = 0, 1$.

We have already shown that the rank of the above two matrices equals the number of targets. In sampling scheme I, to ensure that matrix \mathbf{Z}_q^{MF} is low-rank, both M_t and M_r need to be much larger than K , in other words, a large transmit as well as a

large receive array are required. This, along with the fact that each receiver needs a filter bank, make scheme I more expensive in terms of hardware. However, the matched filtering operation improves the SNR in the received signals. Although in this paper we use the ULA model to illustrate the idea of MIMO-MC radar, the idea can be extended to arbitrary antenna configurations. One possible scenario with a large number of antennas is a networked radar system [69] [70], in which the antennas are placed on the nodes of a network. In such scenarios, a large number of colocated or widely separated sensors could be deployed to collaboratively perform target detection.

In sampling scheme II, assuming that more samples (N) are obtained than existing targets (K), \mathbf{Z}_q will be low-rank as long as there are more receive antennas than targets, i.e., $M_r \gg K$. For this scheme, there is no condition on the number of transmit antennas M_t if G-Orth waveform is applied.

Based on Figs. 4.3 and 4.5, it appears that the average coherence bound, μ_0 , corresponding to \mathbf{Z}_q is larger than that of \mathbf{Z}_q^{MF} . This indicates that the coherence under scheme II is larger than that under scheme I, which means that for scheme II, more observations at the fusion center are required to recover the data matrix with missing entries.

4.3.4 Target Parameters Estimation with Subspace Methods

In this section we describe the MUSIC-based method that will be applied to the estimated data matrices at the fusion center to yield target information.

Let $\hat{\mathbf{Z}}_q$ denote the estimated data matrix for sampling scheme II, during pulse q . Let us perform matched filtering on $\hat{\mathbf{Z}}_q$ to obtain

$$\mathbf{Y}_q = \frac{1}{L} \hat{\mathbf{Z}}_q \tilde{\mathbf{S}}^H = \mathbf{B} \Sigma \mathbf{D}_q \mathbf{A}^T + \tilde{\mathbf{W}}_q, \quad (4.7)$$

where $\tilde{\mathbf{W}}_q$ is noise whose distribution is a function of the additive noise and the nuclear norm minimization problem in (1.50). For sampling scheme I, a similar equation holds for the recovered matrix without further matched filtering.

Then, let us stack the matrices into vector $\mathbf{y}_q = \text{vec}(\mathbf{Y}_q)$, for sampling scheme II, or $\mathbf{y}_q = \text{vec}(\hat{\mathbf{Z}}_q^{MF})$, for sampling scheme I. Based on Q pulses, the following matrix

can be formed: $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_Q] \in \mathbb{C}^{M_t M_r \times Q}$, for which it holds that

$$\mathbf{Y} = \mathbf{V}(\theta) \tilde{\mathbf{X}} + \mathbf{W}, \quad (4.8)$$

where $\tilde{\mathbf{X}} = [\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_Q]$ is a $K \times Q$ matrix containing target reflect coefficient and Doppler shift information; $\tilde{\mathbf{x}}_q = [\tilde{x}_{1,q}, \dots, \tilde{x}_{K,q}]^T$ and $\tilde{x}_{k,q} = \beta_k e^{j \frac{2\pi}{\lambda} 2\vartheta_k (q-1) T_{PRI}}$; $\mathbf{V}(\theta) = [\mathbf{v}(\theta_1), \dots, \mathbf{v}(\theta_K)]$ is a $M_t M_r \times K$ matrix with columns

$$\mathbf{v}(\theta) = \mathbf{a}(\theta) \otimes \mathbf{b}(\theta) \quad (4.9)$$

and $\mathbf{W} = [\text{vec}(\tilde{\mathbf{W}}_1), \dots, \text{vec}(\tilde{\mathbf{W}}_Q)]$.

The sample covariance matrix can be obtained as

$$\hat{\mathbf{R}} = \frac{1}{Q} \sum_{n=1}^Q \mathbf{y}_n \mathbf{y}_n^H = \frac{1}{Q} \mathbf{Y} \mathbf{Y}^H. \quad (4.10)$$

According to [5], the pseudo-spectrum of MUSIC estimator can be written as

$$P(\theta) = \frac{1}{\mathbf{v}^H(\theta) \mathbf{E}_n \mathbf{E}_n^H \mathbf{v}(\theta)} \quad (4.11)$$

where \mathbf{E}_n is a matrix containing the eigenvectors of the noise subspace of $\hat{\mathbf{R}}$. The DOAs of target can be obtained by finding the peak locations of the pseudo-spectrum (4.11).

For joint DOA and speed estimation, we reshape \mathbf{Y} into $\tilde{\mathbf{Y}} \in \mathbb{C}^{Q M_t \times M_r}$ and get

$$\tilde{\mathbf{Y}} = \mathbf{F} \Sigma [\mathbf{b}(\theta_1), \dots, \mathbf{b}(\theta_K)] + \mathbf{W}, \quad (4.12)$$

where

$$\begin{aligned} \mathbf{F} &= [\mathbf{d}(\vartheta_1) \otimes \mathbf{a}(\theta_1), \dots, \mathbf{d}(\vartheta_K) \otimes \mathbf{a}(\theta_K)], \\ \mathbf{d}(\vartheta) &= \left[1, e^{j \frac{2\pi}{\lambda} 2\vartheta T_{PRI}}, \dots, e^{j \frac{2\pi}{\lambda} 2\vartheta (Q-1) T_{PRI}} \right]^T. \end{aligned}$$

The sampled covariance matrix of the receive data signal can then be obtained as $\hat{\mathbf{R}}_{\tilde{\mathbf{Y}}} = \frac{1}{M_r} \tilde{\mathbf{Y}} \tilde{\mathbf{Y}}^H$, based on which DOA and speed joint estimation can be implemented using 2D-MUSIC. The pseudo-spectrum of 2D-MUSIC estimator is

$$P(\theta, \vartheta) = \frac{1}{[\mathbf{d}(\vartheta) \otimes \mathbf{a}(\theta)]^H \mathbf{E}_n \mathbf{E}_n^H [\mathbf{d}(\vartheta) \otimes \mathbf{a}(\theta)]} \quad (4.13)$$

where $\mathbf{E}_n \in \mathbb{C}^{Q M_t \times (Q M_t - K)}$ is the matrix constructed by the eigenvectors corresponding to the noise-subspace of $\hat{\mathbf{R}}_{\tilde{\mathbf{Y}}}$.

4.4 Numerical Results

In this section we demonstrate the performance of the proposed approaches in terms of matrix recovery error and DOA resolution.

We use ULAs for both transmitters and receivers. The inter-node distance for the transmit array is set to $M_r\lambda/2$, while for the receive antennas is set as $\lambda/2$. Therefore, the degrees of freedom of the MIMO radars is M_rM_t [19], i.e., a high resolution could be achieved with a small number of transmit and receive antennas. The carrier frequency is set to $f = 1 \times 10^9 \text{Hz}$, which is a typical radar frequency. The noise introduced in both sampling schemes is white Gaussian with zero mean and variance σ^2 . The data matrix recovery is done using the singular value thresholding (SVT) algorithm [71]. Nuclear norm optimization is a convex optimization problem. There are several algorithms available to solve this problem, such as TFOCS [72]. Here, we chose the SVT algorithm because it is a simple first order method and is suitable for a large size problem with a low-rank solution. During every iteration of SVT, the storage space is minimal and computation cost is low.

We should note that in the SVT algorithm, the matrix rank, or equivalently, the number of targets, is not required to be known a priori. The only requirement is that the number of targets is much smaller than the number of TX/RX antennas, so that the receive data matrix is low-rank. To make sure the iteration sequences of SVT algorithm converge to the solution of the nuclear norm optimization problem, the thresholding parameter τ should be large enough. In the simulation, τ is chosen empirically and set to $\tau = 5\zeta$, where ζ is the dimension of the low-rank matrix that needs to be recovered.

4.4.1 Matrix Recovery Error under Noisy Observations

We consider a scenario with two targets. The first target DOA, θ_1 is generated at random in $[-90^\circ, 90^\circ]$, and the second target DOA, is taken as $\theta_2 = \theta_1 + \Delta\theta$. The target reflection coefficients are set as complex random, and the corresponding speeds are taken at random in $[0, 500] \text{m/s}$. The SNR at each receive antenna is set to 25dB.

In the following, we compute the matrix recovery error as function of the number

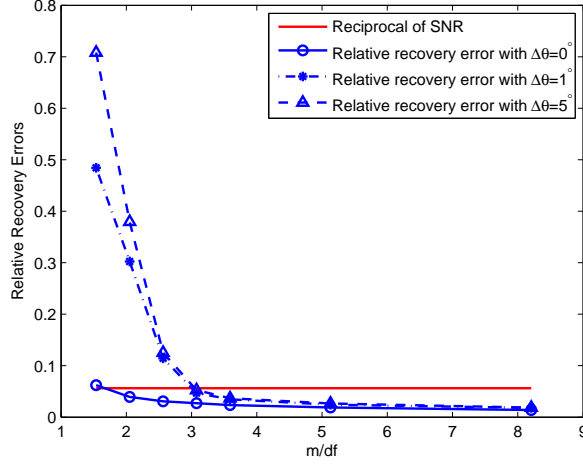


Figure 4.8: Scheme I, $K = 2$ targets: the relative recovery error for \mathbf{Z}_q^{MF} under different values of DOA separation. $M_r = M_t = 40$.

of samples, m , per degrees of freedom, df, i.e., m/df , a quantity also used in [43]. A matrix of size $n_1 \times n_2$ with rank r , has $r(n_1 + n_2 - r)$ degrees of freedom [41]. Let $\phi_{\hat{\mathbf{Z}}}$ denote the relative matrix recovery error, defined as:

$$\phi_{\hat{\mathbf{Z}}} = \left\| \hat{\mathbf{Z}} - \mathbf{Z} \right\|_F / \left\| \mathbf{Z} \right\|_F, \quad (4.14)$$

where we use \mathbf{Z} to denote the data matrix in both sampling schemes, and $\hat{\mathbf{Z}}$ to denote the estimated data matrix.

Figure 4.8 shows $\phi_{\hat{\mathbf{Z}}}$ under sampling scheme I, versus the number of samples per degree of freedom for the same scenario as above. The number of transmit/receiver antennas is set as $M_t = M_r = 40$. It can be seen from Fig. 4.8 that when m/df increases from 2 to 4, or correspondingly, the matrix occupancy ratio increases from $p_1 \approx 0.2$ to ≈ 0.4 , the relative error $\phi_{\hat{\mathbf{Z}}}$ drops sharply to the reciprocal of the matched filter SNR level, i.e., a “phase transition” [44] occurs. It can be seen in Fig. 4.8 that, when the two targets have the same DOA, the relative recovery error is the smallest. This is because in that case the data matrix has the optimum coherence parameter, i.e., $\mu_0 = 1$. As the DOA separation between the two target increases, the relative recovery error of the data matrix in the transition phase increases. In the subsequent DOA resolution simulations, we set the matrix occupancy ratio as $p_1 = \frac{L_1 M_r}{M_t M_r} = 0.5$,

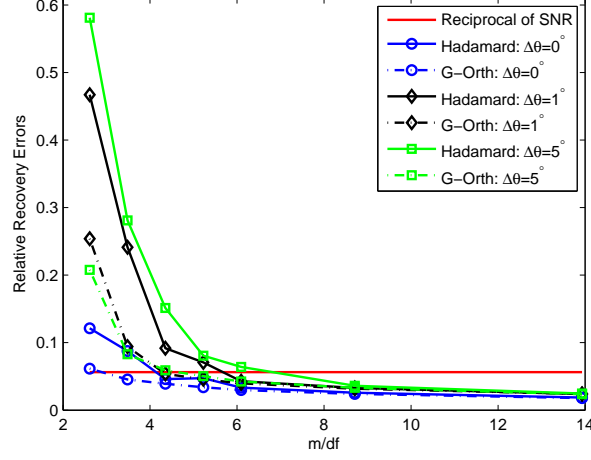


Figure 4.9: Scheme II, $K = 2$ targets: the relative recovery errors for \mathbf{Z}_q under Hadamard and Gaussian Orthogonal waveforms, and different values of $\Delta\theta$. $M_r = M_t = 40$, $N = 256$.

which corresponds to $m/\text{df} \approx 5$, to ensure that the relative recovery error has dropped to the reciprocal of SNR level.

Figure 4.9 shows the relative recovery errors, $\phi_{\hat{\mathbf{Z}}}$, for data matrix \mathbf{Z}_q (sampling scheme II), corresponding to Hadamard or Gaussian orthogonal (G-Orth) transmit waveforms, and the number of Nyquist samples is taken to be $N = 256$. Different values of DOA separation for the two targets are considered, i.e., $\Delta\theta = 0^\circ, 1^\circ, 5^\circ$, respectively.

The results are averaged over 100 independent angle and speed realizations; in each realization the L_2 samples are obtained at random among the N Nyquist samples at each receive antenna. The results of Fig. 4.9 indicate that, for the same $\Delta\theta$, as m/df increases, the relative recovery error, $\phi_{\hat{\mathbf{Z}}}$, under Gaussian orthogonal waveforms (dash lines) reduces to the reciprocal of the SNR faster than under Hadamard waveforms (solid lines). A plausible reason for this is that under G-Orth waveforms, the average coherence parameter of \mathbf{Z}_q is smaller as compared with that under Hadamard waveforms. Under Gaussian orthogonal waveforms, the error $\phi_{\hat{\mathbf{Z}}}$ decreases as $\Delta\theta$ increases. On the other hand, for Hadamard waveforms the relative recovery error appears to increase with an increasing $\Delta\theta$, a behavior that diminishes in the region to the right of

the point of “phase transition”. However, the behavior of the error at the left of the “phase transition” point is not of interest as the matrix completion errors are pretty high and DOA estimation is simply not possible. At the right of the “phase transition” point, the observation noise dominates in the DOA estimation performance.

In both waveforms, the minimum error is achieved when $\Delta\theta = 0^\circ$, i.e., when the two targets have the same DOA, in which case the rank of data matrix \mathbf{Z}_q is rank-1. The above observations suggest that the waveforms do affect performance, and optimal waveform design would be an interesting problem. The waveform selection problem could be formulated as an optimization problem under the orthogonal and narrow-band constraints. We plan to pursue this in our future work.

It can be seen from Fig. 4.8 and Fig. 4.9 that in the noisy cases, as the matrix occupancy ratio increases, the relative recovery errors of the matrices decreases to the reciprocal of SNR.

4.4.2 DOA Resolution with Matrix Completion

In this section we study the probability that two DOAs will be resolved based on the proposed techniques. Two targets are generated at 10° and $10^\circ + \Delta\theta$, where $\Delta\theta = [0.05^\circ, 0.08^\circ, 0.1^\circ, 0.12^\circ, 0.15^\circ, 0.18^\circ, 0.2^\circ, 0.22^\circ, 0.25^\circ, 0.3^\circ]$. The corresponding target speeds are set to 150 and 400 m/s. We set $M_t = M_r = 20$ and $Q = 5$. The DOA information is obtained by finding the peak locations of the pseudo-spectrum (4.11). If the DOA estimates $\hat{\theta}_i$, $i = 1, 2$ satisfy $|\theta_i - \hat{\theta}_i| \leq \varepsilon\Delta\theta$, $\varepsilon = 0.1$, we declare the estimation a success. The probability of DOA resolution is then defined as the fraction of successful events in 200 iterations. For comparison, we also plot the probability curves with full data matrix observations.

First, for scheme I, $L_1 = 10$ matched filters are independently selected at random at each receive antenna, resulting matrix occupancy ratio of $p_1 = 0.5$. The corresponding probability of DOA resolution is shown in Fig. 4.10. As expected, the probability of DOA resolution increase as the SNR increases. The performance of DOA resolution based on the full set of observations has similar behavior. When $\text{SNR} = 25\text{dB}$, the performance of MC-based DOA estimation is close to that with the full data matrix.

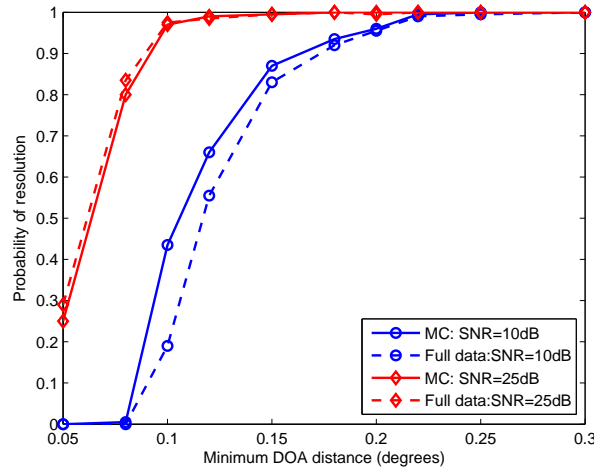


Figure 4.10: Scheme I: DOA resolution. The parameter are set as $M_r = M_t = 20$, $p_1 = 0.5$ and SNR = 10, 25dB, respectively.

Interestingly, for SNR = 10dB, the MC-based result is better than that corresponding to a full data matrix. Most likely, the MC acts like a low-rank approximation of \mathbf{Z}_q^{MF} , and thus eliminates some of the noise.

The probabilities of DOA resolution of DOA estimates under scheme II, with G-Orth and Hadamard waveforms are plotted in Fig. 4.11 (a) and (b), respectively. The parameters are set to $N = 256$ and $p_2 = 0.5$, i.e., each receive antenna uniformly selects $L_2 = 128$ samples at random to forward. Similarly, the simulation results show that under scheme II, the performance at SNR = 10dB is slightly better than that with full data access. In addition, it can be seen that the performance with G-Orth waveforms is better than with Hadamard waveforms. This is because the average coherence of \mathbf{Z}_q under Hadamard waveforms is higher than that with G-Orth waveforms. As shown in Fig. 4.11, increasing the SNR from 10dB to 25dB can greatly improve the DOA estimation performance, as it benefits both the matrix completion and the performance of subspace based DOA estimation method, i.e., MUSIC (see chapt. 9 in [5]).

4.4.3 Comparisons of Sampling Schemes I and II

Comparing the two sampling methods based on the above figures (see Figs. 4.10, and 4.11 (a),(b)) we see that although the performance is the same, sampling scheme I uses

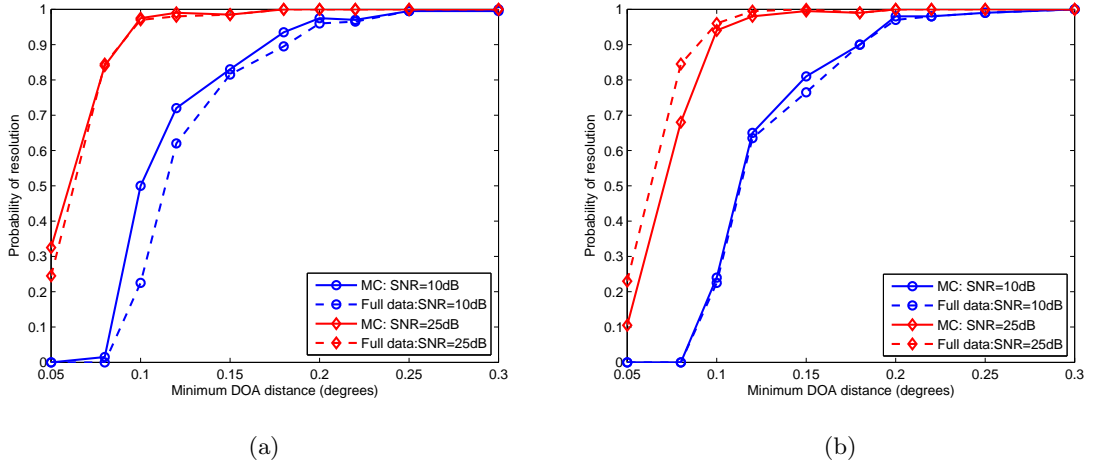


Figure 4.11: Scheme II, $K = 2$, $M_r = M_t = 20$, $N = 256$, $p_2 = 0.5$, SNR = 10, 25dB. DOA resolution with (a) G-Orth waveforms; (b) with Hadamard waveforms.

fewer samples, i.e., 10×20 samples, as compared to sampling scheme II, which uses 128×20 samples. To further elaborate on this observation, we compare the performance of the two sampling schemes when they both forward to the fusion center the same number of samples. The parameters are set to SNR = 25dB, $p_1 = p_2 = 0.5$ and $M_t = N$. Therefore, in both schemes, the number of samples forwarded by each receive antenna was the same. The number of transmit antenna was set as $M_r = 40$ and 80, respectively. Gaussian orthogonal transmit waveforms are used. Two targets are generated at random in $[-90^\circ, 90^\circ]$ at two different DOA separations, i.e., $\Delta\theta = 5^\circ, 30^\circ$. The results are averaged over 100 independent realizations; in each realization, the targets are independently generated at random and the sub-sampling at each receive antenna is also independent between realizations. The relative recovery error comparison is plotted in Fig.4.12.

It can be seen in Fig. 4.12 that as N (or equivalently M_t) increases, the relative recovery error corresponding to \mathbf{Z}_q and \mathbf{Z}_q^{MF} decreases proportionally to the reciprocal of the observed SNR. The relative recovery error under scheme I drops faster than under scheme II for both $M_r = 40$ and $M_r = 80$ cases. This indicates that scheme I has a better performance than scheme II for the same number of samples.

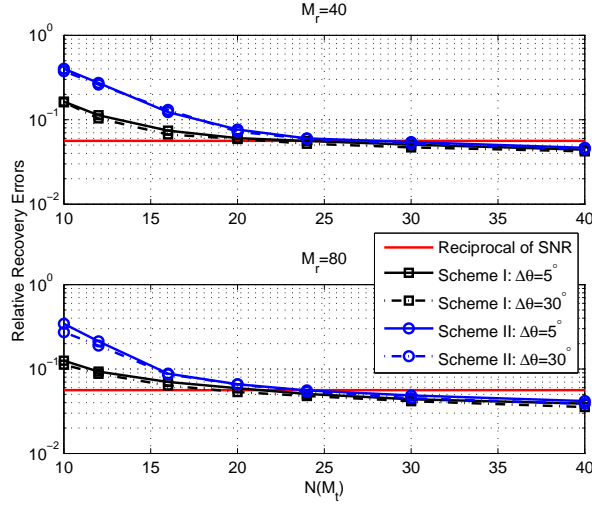


Figure 4.12: Comparisons of the relative recovery errors in terms of number of N (M_t) for $M_r = 40, 80$. The matrix occupancy is set to $p_1 = p_2 = 0.5$. Two targets are generated at random in $[-90^\circ, 90^\circ]$ with DOA separation $\Delta\theta = 5^\circ, 30^\circ$.

4.5 Summary

We have proposed MIMO-MC radars, which is a novel MIMO radar approach for high resolution target parameter estimation that involves small amounts of data. Each receive antenna either performs matched filtering with a small number of dictionary waveforms (scheme I) or obtains sub-Nyquist samples of the received signal (scheme II) and forwards the results to a fusion center. Based on the samples forwarded by all receive nodes, and with knowledge of the sampling scheme, the fusion center applies MC techniques to estimate the full matrix, which is then used in the context of existing array processing techniques, such as MUSIC, to obtain target information. Although ULAs have been considered, the proposed ideas can be generalized to arbitrary configurations. MIMO-MC radars are best suited for sensor networks with large numbers of nodes. Unlike MIMO-CS radars, there is no need for target space discretization, which avoids basis mismatch issues. It has been confirmed with simulations that the coherence of the data matrix at the fusion center meets the conditions for MC techniques to be applicable. The coherence of the matrix is always bounded by a small number. For scheme I, that number approaches 1 as the number of transmit and receive antennas

increases and as the targets separation increases. For scheme II, the coherence does not depend as much on the number of transmit and receive antennas, or the target separation, but it does depend on N , the number of Nyquist samples within one pulse, which is related to the bandwidth of the signal; the coherence increases as N increases. Comparing the two sampling schemes, scheme I has better performance than scheme II for the same number of forwarded samples.

Chapter 5

Waveform Design for MIMO Radars with Matrix Completion

It was shown in Chapter 4 that MIMO radars with sparse sensing and matrix completion can significantly reduce the volume of data required for accurate target detection and estimation. Based on sparsely sampled target returns, forwarded by the receive antennas to a fusion center, a matrix, referred to as the data matrix, can be partially filled, and subsequently completed via MC techniques. The completed data matrix can then be used in standard array processing methods to estimate the target parameters. This chapter studies the applicability of MC theory on the data matrix arising in colocated MIMO radars using uniform linear arrays. It is shown that the data matrix coherence, and consequently the performance of MC, is directly related to the transmit waveforms. Among orthogonal waveforms, the optimum choices are those for which, any snapshot across the transmit array has a flat spectrum. The problem of waveform design is formulated as an optimization problem on the complex Stiefel manifold, and is solved via the modified steepest descent method, or the modified Newton algorithm with nonmonotone line search. Although the optimal waveforms are designed for the case of targets falling in the same range bin, sensitivity analysis is conducted to assess the performance degradation when the targets fall in different range bins.

5.1 Introduction

Unlike traditional phased-array radars which transmit fully correlated signals through their transmit antennas, MIMO radars [18] transmit mutually orthogonal signals. The orthogonality allows the receive antennas to separate the transmitted signals via matched

filtering. The target parameters are obtained by processing the phase shifts of the received signals. MIMO radars offer a high degree of freedom [19] and consequently, enable improved resolution. Since the transmitted signals are orthogonal, the transmit beam is not focused on a particular direction [73], thus resulting in decrease of illumination power. Pulse compression techniques are typically used to improve the range resolution as well as strengthen the receive signal power. The transmit pulse can be coded or modulated, e.g., phase-coded pulse or linear frequency modulated (LFM) pulse [74].

The idea of using low-rank matrix completion (MC) techniques in MIMO radars (termed as MIMO-MC radars) was first proposed in [75], [76] as means of reducing the volume of data required for accurate target detection and estimation. In particular, [76] considers a colocated pulse MIMO radar scenario with uniform linear arrays (ULAs) at the transmitter and the receiver. Each receive antenna samples the target returns and forwards the obtained samples to a fusion center. Based on the data received, the fusion center formulates a matrix, referred to as “data matrix”, which can then be used in standard array processing methods for target detection and estimation. If the data samples are obtained in a Nyquist fashion, and the number of targets is small relative to the number of transmit and receive antennas, the data matrix is low-rank [75]. Thus, it can be recovered from a small subset of its uniformly spaced elements via matrix completion (MC) techniques. By exploiting the latter fact, the MIMO-MC radar receive antennas obtain a small number of samples at uniformly random sampling times. Based on knowledge of the sampling instances, the fusion center populates the data matrix in a uniformly sparse fashion, and subsequently recovers the full matrix via MC techniques. This is referred to as Scheme II in [76]. Alternatively, the receive antennas can perform matched filtering with a randomly selected set of transmit waveforms, and forward the results to the fusion center, partially populating the data matrix. The data matrix is subsequently completed via MC. This is referred to as Scheme I in [76]. Both Schemes I and II reduce the number of samples that need to be forwarded to the fusion center. If the transmission occurs in a wireless fashion, this translates to savings in power and bandwidth. As compared to MIMO radars based on sparse signal

recovery [26], [27], [29], it has been shown in [76] that MIMO-MC radars achieve similar performance but without requiring a target space grid. This is because MC techniques do not require building a basis matrix based on the discretized target space. However, this advantage is achieved at a cost of higher computation complexity arising from the need of both matrix recovery and subspace estimation methods.

Details on the general topic of matrix completion and the conditions for matrix recovery can be found in [41], [42], [43]. The conditions for the applicability of MC on Scheme I of [76] ULA can be found in [68]. In that case, and under ideal conditions, the transmit waveforms do not affect the MC performance. On the other hand, for Scheme II of [76], it was shown in [76], [77] that the transmit waveforms affect the matrix completion performance, as they directly affect the data matrix coherence [42]; a larger coherence implies that more samples need to be collected for reliable target estimation. That observation motivates the work in this paper, where we explore the relationship between matrix coherence and transmit waveforms for Scheme II of [76], and design waveforms that result in the lowest possible coherence. Waveform design in the context of MIMO radars has been extensively studied. For example, in [78], [79], the waveforms are designed to maximize the mutual information between target impulse response (which are assumed known) and reflected signals. In [80], waveforms with good correlation properties are designed under the unimodular constraint. In [81], clutter mitigation is considered in waveform design for ground moving-target indication (GMTI). Orthogonal phase-coded waveforms are designed in [82] via simulated annealing. Frequency hopping waveforms are designed in [22] to reduce the sidelobe of the radar ambiguity function [21]. For MIMO radars with sparse signal recovery, e.g., in [83], [30], the goal of waveform design is to reduce the coherence of sensing matrix.

In this chapter, we aim to design waveforms so that the coherence of the receive data matrix attains its lowest possible value. Since the waveforms are constrained to be orthogonal, we formulate the design problem as an optimization problem on the complex Stiefel manifold [84], and for its solution employ the modified steepest descent algorithm [85] and the modified Newton algorithm. The local gradient and Hessian of the cost functions are derived in closed forms. We also provide sensitivity analysis

of the optimized waveforms in the scenario of targets falling in different range bins. In particular, we show that for relatively small delays, i.e., of the order of the radar pulse duration, the matrix coherence increases only slightly as the maximal range delay increases. Numerical results show that good angle and speed estimation performance could be achieved under the optimized waveforms.

5.2 Background of MIMO-MC Radar

We consider the problem formulation proposed in [76] for scheme II. The scenario involves narrowband orthogonal transmit waveforms, transmitted in pulses with pulse repetition interval T_{PRI} and carrier wavelength λ , K far-field targets at angles θ_k , and ULAs for transmission and reception, equipped with M_t transmit and M_r receive antennas, respectively, and inter-element spacing d_t and d_r , respectively (see Subsection B of Section II in [76]).

During each pulse, the m -th, $m \in \mathbb{N}_{M_t}^+$ antenna transmits a coded waveform containing N symbols $\{s_m(n)\}, n = 1, \dots, N$ of duration T_b each, which can be written in the baseband as

$$\phi_m(t) = \sum_{n=1}^N s_m(n) \Lambda \left[\frac{t - (n-1)T_b}{T_b} \right], t \in [0, T_\phi], \quad (5.1)$$

where $s_m(n) = a_m(n) e^{j\varphi_m(n)}$, with $\{\varphi_m(n)\}$ uniformly distributed in $[-\pi, \pi]$, and $\{a_m(n)\}$ taking arbitrary positive values; $\Lambda(t)$ is a rectangular pulse and $T_\phi = NT_b$ is the duration of the entire pulse. In this chapter, we relax the constant amplitude requirement to exploit more degrees of freedom for waveform design. A similar relaxation was also exploited in [83] [30]. We admit that this relaxation would put some difficulties in the radar transmitters. We will assume that the waveforms are sufficiently narrowband, i.e., $\frac{1}{T_\phi} \ll \frac{c}{\lambda}$ and targets are slow moving, i.e., $\frac{2\vartheta}{\lambda} \ll \frac{1}{T_\phi}$, where c is the speed of light and ϑ is target speed.

In the following, we do the analysis for a scenario in which all targets are in the same range bin [18] [19]; in Section 5.5, we will study sensitivity issues arising from targets appearing in different range bins. Suppose that the receive antennas sample the target echoes with sampling interval T_b and forward their samples to a fusion center.

Let \mathbf{X} be the matrix formulated at the fusion center based on receive antenna samples data with each antenna contributing a row to \mathbf{X} . It holds that

$$\mathbf{X} = \mathbf{W} + \mathbf{J}, \quad (5.2)$$

where \mathbf{J} is an interference/noise matrix and

$$\mathbf{W} = \mathbf{B} \mathbf{D} \mathbf{A}^T \mathbf{S}^T, \quad (5.3)$$

where $\mathbf{A} \in \mathbb{C}^{M_t \times K}$ is the transmit steering matrix (respectively defined is $\mathbf{B} \in \mathbb{C}^{M_r \times K}$) with $[\mathbf{A}]_{mk} = e^{-j2\pi(m-1)\alpha_k^t}$ and $\alpha_k^t \triangleq \frac{d_t \sin(\theta_k)}{\lambda}$, $(m, k) \in \mathbb{N}_{M_t}^+ \times \mathbb{N}_K^+$, where θ_k is the angle of the k -th target, or equivalently, α_k^t is the spatial frequency corresponding to the k -th target.

$$\mathbf{D} \triangleq \text{diag} \left(\begin{bmatrix} \beta_1 \zeta_{q1} & \beta_2 \zeta_{q2} & \cdots & \beta_K \zeta_{qK} \end{bmatrix} \right), \quad (5.4)$$

where $\zeta_{qk} = e^{j2\pi\nu_k(q-1)T_{PRI}}$, with q denoting the pulse index, $\nu_k = \frac{2\vartheta_k}{\lambda}$ denoting the Doppler shift of the k -th target, and $\{\beta_k\}_{k \in \mathbb{N}_K^+}, \{\vartheta_k\}_{k \in \mathbb{N}_K^+}$ denoting contain target reflection coefficients and speeds, respectively. $\mathbf{S} = [\mathbf{s}(1), \dots, \mathbf{s}(N)]^T \in \mathbb{C}^{N \times M_t}$ with $\mathbf{s}(i) = [s_1(i), \dots, s_{M_t}(i)]^T$, is the sampled waveform matrix, with its vertical dimension corresponding to sampling along time and its horizontal dimension corresponding to sampling across the array (sampling in space). The i -th row of \mathbf{S} can be thought of as the snapshot of the waveforms across the transmit antennas at sampling time i . Due to the assumed orthogonality of the waveforms, it holds that $\mathbf{S}^H \mathbf{S} = \mathbf{I}_{M_t}$ when $N \geq M_t$ [67].

When both M_r, M_t as well as N are larger than K , the noise free data matrix \mathbf{W} is rank- K and can be recovered from a small number of its entries via matrix completion. This fact motivated the approach of [75] [76], which calls for subsampling the target echoes at the receive antennas in a uniformly pseudo-random fashion, partially filling the data matrix at the fusion center, and then completing the data matrix via MC techniques. This approach reduces the number of samples that need to be forwarded to the fusion center. It turns out that the applicability of MC depends on the transmit waveforms. In the next section, we derive the necessary and sufficient conditions the

transmit waveforms must satisfy so that the coherence of \mathbf{W} is asymptotical optimal, i.e., it approaches 1 as M_t increases.

5.3 Data Matrix Coherence Analysis and Optimal Waveform Conditions

In this section, we analyze the coherence of \mathbf{W} , for the MIMO radar system defined in the previous section. In particular, (1) we provide sufficient and necessary conditions for the optimal transmit waveforms under which the coherence of \mathbf{W} attains its lowest possible value; (2) under those conditions, we show asymptotic optimality of the coherence of \mathbf{W} w.r.t. the number of transmit/receive antennas; (3) we show that the coherence of \mathbf{W} does not depend on the Doppler shift.

5.3.1 The Coherence of Data Matrix in MIMO-MC Radar

Let $S_i(\alpha_k^t)$ denote the discrete-time Fourier transform (DTFT) of the i -th snapshot of the transmit waveforms evaluated at spatial frequency α_k^t , i.e.,

$$S_i(\alpha_k^t) = \sum_{m=1}^{M_t} s_m(i) e^{-j2\pi(m-1)\alpha_k^t}, \quad (5.5)$$

where $\{\{s_m(i)\}_{m \in \mathbb{N}_{M_t}^+}\}$ are the elements in the i -th row of \mathbf{S} .

Before we proceed we provide a lemma that will be useful in the subsequent theorems.

Lemma 1. *For the MIMO radar system described in Section 5.2 with $d_t = \lambda/2$, and K targets randomly located at angles $\{\theta_k \in [-\frac{\pi}{2}, \frac{\pi}{2}]\}_{k \in \mathbb{N}_K^+}$, or equivalently, at spatial frequencies $\{\alpha_k^t \in [-\frac{1}{2}, \frac{1}{2}]\}_{k \in \mathbb{N}_K^+}$, it holds that*

$$\sum_{i=1}^N \sum_{k=1}^K |S_i(\alpha_k^t)|^2 = KM_t. \quad (5.6)$$

Proof of Lemma 1. See the Appendix B for the proof. \square

The waveforms conditions under which the coherence of \mathbf{W} attains its lowest value are summarized in the following theorem.

Theorem 2. (*Optimal Waveform Conditions*): Consider the MIMO radar systems as defined in Section 5.2, with $d_t = \frac{\lambda}{2}$.

The necessary condition under which the coherence of \mathbf{W} attains its lowest possible value is

$$\sum_{k=1}^K |S_i(\alpha_k^t)|^2 = \frac{KM_t}{N}, \quad \forall i \in \mathbb{N}_N^+. \quad (5.7)$$

A sufficient condition for the coherence of \mathbf{W} to attain its lowest possible value, independent of the target angles is

$$|S_i(\alpha_l^t)|^2 = \frac{M_t}{N}, \quad \forall i \in \mathbb{N}_N^+ \text{ and } \forall \alpha_l^t \in \left[-\frac{1}{2}, \frac{1}{2}\right]. \quad (5.8)$$

Proof of Theorem 2. To make the proof more tractable, we break it into two parts. In the first part, we characterize the SVD of the matrix \mathbf{W} in order to identify the actions that are needed in order to bound its coherence. In the second part, we derive the optimal conditions of the coded orthogonal waveforms.

1) *Characterization of the SVD of \mathbf{W} :* The compact SVD of \mathbf{W} can be expressed as

$$\mathbf{W} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H, \quad (5.9)$$

where $\mathbf{U} \in \mathbb{C}^{M_r \times K}$, $\mathbf{V} \in \mathbb{C}^{N \times K}$ such that $\mathbf{U}^H \mathbf{U} = \mathbf{I}_K$, $\mathbf{V}^H \mathbf{V} = \mathbf{I}_K$, and $\mathbf{\Lambda} \in \mathbb{R}^{K \times K}$ is a diagonal matrix containing the singular values of \mathbf{W} .

Consider the QR decomposition of \mathbf{B} , i.e., $\mathbf{B} = \mathbf{Q}_r \mathbf{R}_r$, with $\mathbf{Q}_r \in \mathbb{C}^{M_r \times K}$, such that $\mathbf{Q}_r^H \mathbf{Q}_r = \mathbf{I}_K$ and $\mathbf{R}_r \in \mathbb{C}^{K \times K}$ an upper triangular matrix. Similarly, consider the QR decomposition of $\mathbf{S}\mathbf{A}$, i.e., $\mathbf{S}\mathbf{A} = \mathbf{Q}_s \mathbf{R}_s$, with $\mathbf{Q}_s \in \mathbb{C}^{N \times K}$, such that $\mathbf{Q}_s^H \mathbf{Q}_s = \mathbf{I}_K$ and $\mathbf{R}_s \in \mathbb{C}^{K \times K}$ an upper triangular matrix. The matrix $\mathbf{R}_r \mathbf{D} \mathbf{R}_s^T \in \mathbb{C}^{K \times K}$ is rank- K and its SVD can be expressed as $\mathbf{R}_r \mathbf{D} \mathbf{R}_s^T = \mathbf{Q}_1 \mathbf{\Delta} \mathbf{Q}_2^H$. Here, $\mathbf{Q}_1 \in \mathbb{C}^{K \times K}$ is such that $\mathbf{Q}_1 \mathbf{Q}_1^H = \mathbf{Q}_1^H \mathbf{Q}_1 = \mathbf{I}_K$ (the same holds for \mathbf{Q}_2) and $\mathbf{\Delta} \in \mathbb{R}^{K \times K}$ is non-zero diagonal, containing the singular values of $\mathbf{R}_r \mathbf{D} \mathbf{R}_s^T$. Therefore, it holds that

$$\mathbf{W} = \mathbf{Q}_r \mathbf{Q}_1 \mathbf{\Delta} \mathbf{Q}_2^H \mathbf{Q}_s^T = \mathbf{Q}_r \mathbf{Q}_1 \mathbf{\Delta} (\mathbf{Q}_s^* \mathbf{Q}_2)^H, \quad (5.10)$$

which is a valid SVD of \mathbf{W} since $(\mathbf{Q}_r \mathbf{Q}_1)^H \mathbf{Q}_r \mathbf{Q}_1 = \mathbf{I}_K$ and $(\mathbf{Q}_s^* \mathbf{Q}_2)^H \mathbf{Q}_s^* \mathbf{Q}_2 = \mathbf{I}_K$. Via the uniqueness of the singular values of a matrix, it holds that $\mathbf{\Lambda} = \mathbf{\Delta}$, thus $\mathbf{U} = \mathbf{Q}_r \mathbf{Q}_1$ and $\mathbf{V} = \mathbf{Q}_s^* \mathbf{Q}_2$.

Let \mathbf{Q}_r^i denote the i -th row of \mathbf{Q}_r . The coherence of the row space of \mathbf{W} is

$$\mu(U) = \frac{M_r}{K} \sup_{i \in \mathbb{N}_{M_r}^+} \|\mathbf{Q}_r^i \mathbf{Q}_1\|_2^2 = \frac{M_r}{K} \sup_{i \in \mathbb{N}_{M_r}^+} \|\mathbf{Q}_r^i\|_2^2 \quad (5.11)$$

It can be seen from (5.11) that $\mu(U)$ is determined by \mathbf{Q}_r , which is only related to the receive steering matrix \mathbf{B} and is independent of the transmit waveform \mathbf{S} . In the MIMO radar systems under ULA configuration with $d_r = \frac{\lambda}{2}$, under the assumption that the target angles set $\{\theta_k\}_{k \in \mathbb{N}_N^+}$ are distinct with minimal spatial frequency separation ξ , it was shown in [68] that

$$\mu(U) \leq \frac{\sqrt{M_r}}{\sqrt{M_r} - (K-1) \sqrt{\beta_{M_r}(\xi_r)}}, \quad (5.12)$$

where $\beta_{M_r}(\xi_r)$ is the Fejér kernel (see (5.23)).

Let $\mathbf{Q}_s^{*(i)}$ and $\mathbf{S}^{*(i)}$, $i \in \mathbb{N}_N^+$ denote the i -th row of \mathbf{Q}_s^* and \mathbf{S}^* , respectively. For the coherence of the row space of \mathbf{W} we have

$$\begin{aligned} \mu(V) &= \frac{N}{K} \sup_{i \in \mathbb{N}_N^+} \|\mathbf{Q}_s^{*(i)} \mathbf{Q}_2\|_2^2 = \frac{N}{K} \sup_{i \in \mathbb{N}_N^+} \|\mathbf{Q}_s^{*(i)}\|_2^2 \\ &= \frac{N}{K} \sup_{i \in \mathbb{N}_N^+} \|\mathbf{S}^{*(i)} \mathbf{A}^* (\mathbf{R}_s^*)^{-1}\|_2^2 \\ &\leq \frac{N}{K} \sup_{i \in \mathbb{N}_N^+} \frac{\|\mathbf{S}^{*(i)} \mathbf{A}^*\|_2^2}{\sigma_{\min}^2(\mathbf{R}_s^*)}, \end{aligned} \quad (5.13)$$

where

$$\begin{aligned} \sigma_{\min}^2(\mathbf{R}_s^*) &= \lambda_{\min} \left((\mathbf{R}_s^*)^H \mathbf{R}_s^* \right) = \lambda_{\min} \left(\mathbf{R}_s^H \mathbf{R}_s \right) \\ &= \lambda_{\min} \left(\mathbf{R}_s^H \mathbf{Q}_s^H \mathbf{Q}_s \mathbf{R}_s \right) = \lambda_{\min} \left((\mathbf{S}\mathbf{A})^H \mathbf{S}\mathbf{A} \right) \\ &= \lambda_{\min} \left(\mathbf{A}^H \mathbf{A} \right). \end{aligned} \quad (5.14)$$

Here, we use the symbol $\lambda_{\min}(\cdot)$ to denote the minimal eigenvalue of a matrix. In addition, we apply the fact that the eigenvalues of a Hermitian matrix are real, and the eigenvalues of \mathbf{X}^* are the complex conjugate of the eigenvalues of \mathbf{X} . Thus, if \mathbf{X} is Hermitian, its eigenvalues are equal to eigenvalues of \mathbf{X}^* .

In the MIMO radar systems under ULA configuration with $d_t = \frac{\lambda}{2}$, under the assumption that the target angles are distinct with minimal spatial frequency separation

ξ_t , it was shown in [68] that

$$\lambda_{\min}(\mathbf{A}^H \mathbf{A}) \geq M_t - (K - 1) \sqrt{M_t \beta_{M_t}(\xi_t)} \quad (5.15)$$

where $\beta_{M_t}(\xi_t)$ is the kernel defined in (5.23). Therefore, regarding the coherence of the column space of \mathbf{W} , we have

$$\mu(V) \leq \frac{N}{K} \sup_{i \in \mathbb{N}_N^+} \frac{\|\mathbf{S}^{*(i)} \mathbf{A}^*\|_2^2}{M_t - (K - 1) \sqrt{M_t \beta_{M_t}(\xi_t)}}. \quad (5.16)$$

Next, we focus on finding the minimum of the supremum of $\|\mathbf{S}^{*(i)} \mathbf{A}^*\|_2^2$ over $i \in \mathbb{N}_N^+$, which results in the optimal waveform conditions.

2) *Optimal Waveform Conditions:* The transmit steering matrix \mathbf{A} has the Vandermonde form under ULA configuration. Consequently, it holds that

$$\begin{aligned} & \|\mathbf{S}^{*(i)} \mathbf{A}^*\|_2^2 \\ &= \mathbf{S}^{*(i)} \mathbf{A}^* (\mathbf{S}^{*(i)} \mathbf{A}^*)^H \\ &= \sum_{k=1}^K \sum_{m=1}^{M_t} s_m(i) e^{-j2\pi(m-1)\alpha_k^t} \sum_{m'=1}^{M_t} s_{m'}^*(i) e^{j2\pi(m'-1)\alpha_k^t} \\ &= \sum_{k=1}^K \left| \sum_{m=1}^{M_t} s_m(i) e^{-j2\pi(m-1)\alpha_k^t} \right|^2 = \sum_{k=1}^K |S_i(\alpha_k^t)|^2, \end{aligned} \quad (5.17)$$

Therefore, the coherence bound of the row space of \mathbf{W} is

$$\mu(V) \leq \frac{N}{K} \frac{\sup_{i \in \mathbb{N}_N^+} \sum_{k=1}^K |S_i(\alpha_k^t)|^2}{M_t - (K - 1) \sqrt{M_t \beta_{M_t}(\xi_t)}}. \quad (5.18)$$

The lowest possible coherence bound of $\mu(V)$ can be achieved by finding waveforms that minimize $\sup_{i \in \mathbb{N}_N^+} \sum_{k=1}^K |S_i(\alpha_k^t)|^2$. This can be formulated as a min-max optimization problem subject to the constraint given in Lemma 1, i.e.,

$$\begin{aligned} & \min_{\mathbf{S}} \left(\max_{i \in \mathbb{N}_N^+} \sum_{k=1}^K |S_i(\alpha_k^t)|^2 \right) \\ & \text{s.t.} \quad \sum_{i=1}^N \sum_{k=1}^K |S_i(\alpha_k^t)|^2 = K M_t \end{aligned} \quad (5.19)$$

Since $\sum_{k=1}^K |S_i(\alpha_k^t)|^2 \geq 0$, for $i \in \mathbb{N}_N^+$, the optimal solution of the min-max optimization problem is

$$\sum_{k=1}^K |S_i(\alpha_k^t)|^2 = \frac{KM_t}{N}, \quad \forall i \in \mathbb{N}_N^+. \quad (5.20)$$

The solution, as shown in (5.20) depends on the specific target spatial angles $\{\alpha_k^t\}_{k \in \mathbb{N}_K^+}$. Since these angles are not known, we need to consider every possible angle θ_l in the angle space $[-\frac{\pi}{2}, \frac{\pi}{2}]$, or every $\alpha_l^t \in [-\frac{1}{2}, \frac{1}{2}]$. Thus, the optimal waveforms should sufficiently satisfy:

$$|S_i(\alpha_l^t)|^2 = \frac{M_t}{N}, \quad \forall i \in \mathbb{N}_N^+ \text{ and } \forall \alpha_l^t \in \left[-\frac{1}{2}, \frac{1}{2}\right]. \quad (5.21)$$

The condition of (5.21) indicates that the power spectrum of each snapshot should be flat in the spatial frequency range $\alpha_l^t \in [-\frac{1}{2}, \frac{1}{2}]$, and thus each waveform snapshot must be white noise type sequence with variance M_t/N . This completes the proof of Theorem 2. \square

Under the optimal waveforms conditions stated in the Theorem 2, the coherence of matrix \mathbf{W} is asymptotical optimal, as stated in the following theorem.

Theorem 3. (*Coherence of \mathbf{W}*): Consider the MIMO radar system as presented in Section 5.2 and K distinct targets. Let the minimum spatial frequency separation of the targets be

$$x = \min_{(i,j) \in \mathbb{N}_K^+ \times \mathbb{N}_K^+, i \neq j} \frac{d_h}{\lambda} (\sin \theta_i - \sin \theta_j), \quad h \in \{t, r\}. \quad (5.22)$$

and assume that $|x| \geq \xi_h \neq 0, h \in \{t, r\}$. Let us also define

$$\beta_{M_h}(x) = \frac{1}{M_h} \frac{\sin^2(\pi M_h x)}{\sin^2(\pi x)}, \quad (5.23)$$

For $d_t = d_r = \frac{\lambda}{2}$ and under the optimal waveform conditions stated in Theorem 2, as long as

$$K \leq \min_{h \in \{t, r\}} \left\{ \sqrt{\frac{M_h}{\beta_{M_h}(\xi_h)}} \right\}, \quad (5.24)$$

the matrix \mathbf{W} obeys the conditions (A0) and (A1) with

$$\mu_0 \triangleq \max_{h \in \{t, r\}} \left\{ \frac{\sqrt{M_h}}{\sqrt{M_h} - (K-1) \sqrt{\beta_{M_h}(\xi_h)}} \right\}, \quad (5.25)$$

and $\mu_1 \triangleq \sqrt{K} \mu_0$ with probability 1.

Proof of Theorem 3. Following Theorem 2, for waveforms that satisfy the necessary condition (5.20), it holds that

$$\begin{aligned} \mu(V) &\leq \inf_{\mathbf{s}} \left(\frac{N}{K} \sup_{i \in \mathbb{N}_N^+} \frac{\sum_{k=1}^K |S_i(\alpha_k^t)|^2}{M_t - (K-1) \sqrt{M_t \beta_{M_t}(\xi_t)}} \right) \\ &= \frac{\sqrt{M_t}}{\sqrt{M_t} - (K-1) \sqrt{\beta_{M_t}(\xi_t)}}. \end{aligned} \quad (5.26)$$

Consequently, under the optimal waveforms conditions, and via inequality (5.12), we get (5.25). It was shown in [41] that in the general case, $\mu_1 = \mu_0 \sqrt{K}$ always holds true. Consequently, the conditions (A0) and (A1) hold. \square

5.3.2 Remarks

1) Asymptotic Optimal Coherence of \mathbf{W}

It should be noted that kernel $\beta_{M_h}(x), h \in \{t, r\}$ is a periodic function of x .

For $d_t = d_r = \frac{\lambda}{2}$, the spatial frequency separation corresponding to both transmit and receive arrays, satisfy $|x| \in (0, \frac{1}{2}]$. If $\xi \triangleq \max\{\xi_r, \xi_t\} \neq 0$, we can find a small constant ξ and $0 < \xi < \frac{1}{\min\{M_t, M_r\}}$ such that the Dirichlet kernel $\frac{\sin(\pi M_h \xi)}{\sin(\pi \xi)} = \mathcal{O}(1)$ and the kernel $\beta_{M_h}(\xi)$ satisfies $\sqrt{\beta_{M_h}(\xi)} = \frac{\sin(\pi M_h \xi)}{\sqrt{M_h} \sin(\pi \xi)} = \mathcal{O}\left(\frac{1}{\sqrt{M_h}}\right)$. Consequently, the values of $\beta_{M_h}(\xi)$ decrease as $M_h, h \in \{t, r\}$ increase. Then for any fixed K , if $\sqrt{M_h} \geq K \sqrt{\beta_{M_h}(\xi)}, h \in \{t, r\}$, or equivalently

$$M_h \geq K \frac{\sin(\pi M_h \xi)}{\sin(\pi \xi)} = \mathcal{O}(K), \quad (5.27)$$

both (5.12) and (5.26) hold. Consequently, under the optimal waveform conditions, it holds that

$$\lim_{M_t \rightarrow \infty} \mu(V) \leq \lim_{M_t \rightarrow \infty} \frac{\sqrt{M_t}}{\sqrt{M_t} - (K-1) \sqrt{\beta_{M_t}(\xi)}} = 1.$$

Since $\mu(V) \geq 1$, via the coherence definition, it must hold that under the optimal waveform conditions $\mu(V) = 1$ in the limit w.r.t. M_t . Similarly, it must hold that $\mu(U) = 1$ in the limit w.r.t. M_r . As a result, the coherence of \mathbf{W} is asymptotically optimal.

It should be noted that the spatial frequency separation requirement is not restrictive. For example, in a ULA with M antennas, the spatial frequency separation of targets should be larger than the resolution of the array, i.e., $\frac{1}{M}$. As it can be seen in the proof of Theorem 3, Theorem 3 holds even when the spatial frequency separation of the targets is less than the resolution of the array.

2) Coherence and Doppler Shift

It can be easily seen from Theorem 3 that the coherence of \mathbf{W} does not depend on the Doppler shift $\{\nu_k\}_{k \in \mathbb{N}_K^+}$ under the assumption that targets are slow moving.

5.3.3 Comparative Study of Two Orthogonal Waveforms

This section provides a comparative study of Hadamard and randomly generated Gaussian Orthogonal (G-Orth) waveforms.

Let us consider a ULA transmit array with carrier frequency $f_c = 1 \times 10^9 \text{ Hz}$, $d_t = \frac{\lambda}{2}$, $M_t = 40$ and $N = 64$. The corresponding waveform matrix $\mathbf{S} \in \mathbb{C}^{N \times M_t}$ contains Hadamard or G-Orth waveforms.

Two targets are considered at $\theta_1 = 0^\circ, \theta_2 = 20^\circ$, or equivalently, $\alpha_1^t = 0, \alpha_2^t = \frac{1}{2} \sin(\frac{\pi}{9})$. The target angle search space is $[-90^\circ, 90^\circ]$ and the corresponding spatial frequency range is $\alpha_k^t \in [-\frac{1}{2}, \frac{1}{2}]$. The power spectra of the waveform snapshots $i \in \mathbb{N}_N^+$ corresponding to Hadamard and G-Orth waveforms are shown in Fig. 5.1.

It can be seen in Fig. 5.1 (a) that the power spectra of the snapshots corresponding to the Hadamard case have high values at some specific snapshots and angles, while those of G-Orth case are spread out across all snapshots and angles (see Fig. 5.1 (c)). In addition, the maximum power spectral value for the Hadamard case is much higher than that of the G-Orth case. Thus, based on (5.18), the coherence bound $\mu(V)$ corresponding to the Hadamard waveforms is larger than that for the G-Orth waveforms. The lower coherence bound result is better MC recovery performance for the

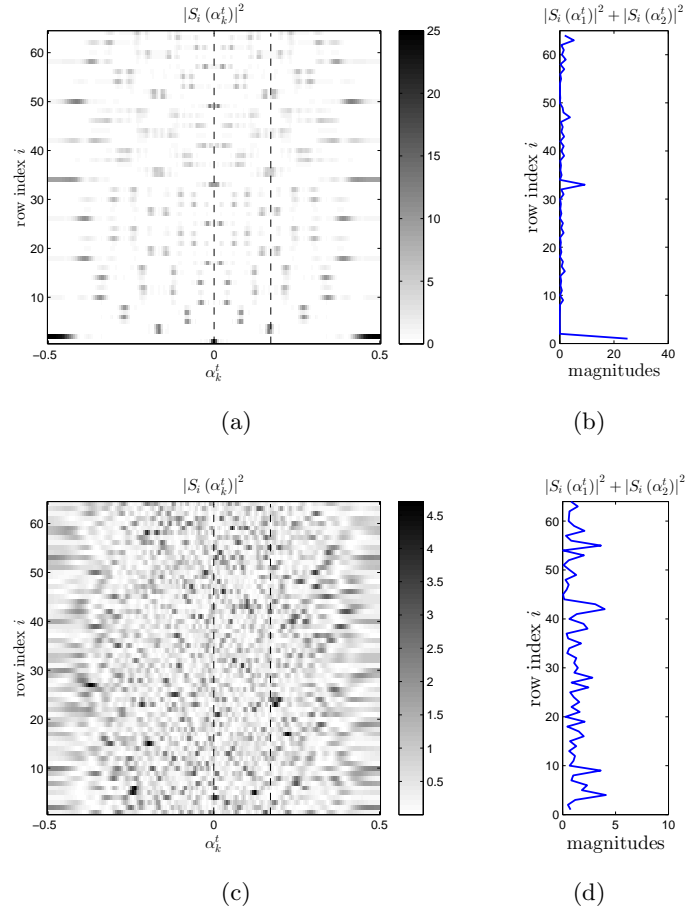


Figure 5.1: The power spectra $|S_i(\alpha_k^t)|^2$ for $\alpha_k^t \in [-\frac{1}{2}, \frac{1}{2}]$ with $M_t = 40$ and $N = 64$. (a) Hadamard waveform; (b) Magnitudes of $|S_i(\alpha_1^t)|^2 + |S_i(\alpha_2^t)|^2$ of Hadamard waveform; (c) G-Orth waveform; (d) Magnitudes of $|S_i(\alpha_1^t)|^2 + |S_i(\alpha_2^t)|^2$ of G-Orth waveform.

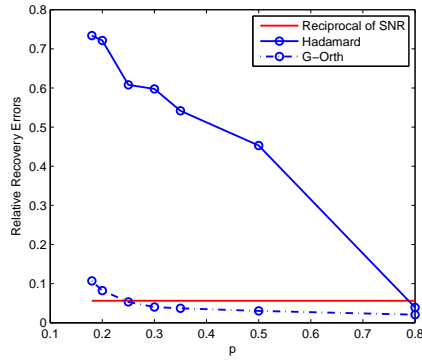


Figure 5.2: Comparison of MC error w.r.t. Hadamard and G-Orth waveforms as function of the matrix occupancy ratio.

G-orth waveforms; this can be seen in Fig. 5.2 where the horizontal axis is the portion of entries of \mathbf{W} that were available. In the simulations, the data matrix corresponding the two targets are recovered via the SVT algorithm [71] for $M_r = 128, M_t = 40, N = 64$ and $\text{SNR} = 25\text{dB}$. The simulation results are averaged over 50 independent runs. In each run, the G-Orth waveforms are randomly generated.

5.4 Waveform Design under Spatial Power Spectra Constraints

Theorem 2 states that, among the class of orthogonal waveforms, and for MIMO radars using ULAs, the optimal waveform matrix should have rows that are white-noise type functions, i.e., the waveform snapshots across the transmit antennas should be white. In this section, we propose a scheme to optimally design the transmit waveform matrix for MIMO-MC radars. In particular, the design problem is formulated as optimization on matrix manifolds [86]. Due to the orthogonality constraint on the transmit waveforms, i.e., the columns of matrix \mathbf{S} , the matrix manifold is the complex Stiefel manifold, which is non-convex. The solution can be obtained via the modified steepest descent algorithm [85], or the modified Newton algorithm with a nonmonotone line search method [87]. The derivative and Hessian of the objective function w.r.t. the waveform matrix are obtained in a closed form.

5.4.1 Problem Formulation

Let us discretize the angle space $[-\frac{\pi}{2}, \frac{\pi}{2}]$ into L phases $\{\theta_l\}_{l \in \mathbb{N}_L^+}$, corresponding to the spatial frequencies $\{\alpha_l^t\}_{l \in \mathbb{N}_L^+}$. Let $c_{il} = \mathbf{S}^{*(i)} \mathbf{A}^*(\theta_l)$ for $i \in \mathbb{N}_N^+$. According to the optimal condition (5.21), it holds that

$$|c_{il}|^2 = |S_i(\alpha_l^t)|^2 = \frac{M_t}{N}, \quad i \in \mathbb{N}_N^+, \quad l \in \mathbb{N}_L^+. \quad (5.28)$$

Define $\mathbf{A}^* = [\mathbf{A}^*(\theta_1), \dots, \mathbf{A}^*(\theta_L)]$ and $\mathbf{F} = \mathbf{S}^* \mathbf{A}^*$. It holds that $[\mathbf{F} \odot \mathbf{F}^*]_{il} = |c_{il}|^2$. Based on (5.21), let us define the objective function

$$f(\mathbf{S}) = \left\| \mathbf{F} \odot \mathbf{F}^* - \frac{M_t}{N} \mathbf{1}_N \mathbf{1}_L^T \right\|_F^2. \quad (5.29)$$

The waveform design problem is formulated as

$$\begin{aligned} \min \quad & f(\mathbf{S}) \\ \text{s.t.} \quad & \mathbf{S}^H \mathbf{S} = \mathbf{I}_{M_t}. \end{aligned} \quad (5.30)$$

Due to the orthogonal constraint, \mathbf{S} belongs to the complex Stiefel manifold $\mathcal{S}(N, M_t)$, defined as

$$\mathcal{S}(N, M_t) = \{\mathbf{S} \in \mathbb{C}^{N \times M_t} : \mathbf{S}^H \mathbf{S} = \mathbf{I}_{M_t}\}. \quad (5.31)$$

The nonconvexity of the orthogonal constraint on the complex Stiefel manifold makes the waveform design problem challenging. In the following we adopt the modified steepest descent algorithm [85], or the modified Newton algorithm on the Stiefel manifold to solve the problem of (5.30).

5.4.2 Derivative and Hessian of Cost Function $f(\mathbf{S})$

In this subsection, we will address the derivative and Hessian of the cost function $f(\mathbf{S})$ defined in (5.29) w.r.t. the variables \mathbf{S} . First, based on the second order Taylor series approximation (see [85]), the cost function $f : \mathbb{C}^{N \times M_t} \rightarrow \mathbb{R}$ can be written as

$$\begin{aligned} f(\mathbf{S} + \delta \mathbf{Z}) = & f(\mathbf{S}) + \delta \Re \{ \text{tr}(\mathbf{Z}^H \mathbf{D}_{\mathbf{S}}) \} \\ & + \frac{\delta^2}{2} \text{vec}(\mathbf{Z})^H \mathbf{H}_{\mathbf{S}} \text{vec}(\mathbf{Z}) \\ & + \frac{\delta^2}{2} \Re \left\{ \text{vec}(\mathbf{Z})^T \mathbf{C}_{\mathbf{S}} \text{vec}(\mathbf{Z}) \right\} + \mathcal{O}(\delta^3), \end{aligned} \quad (5.32)$$

where $\mathbf{D}_{\mathbf{S}} \in \mathbb{C}^{N \times M_t}$ is the derivative of f evaluated at \mathbf{S} , and the matrix pair $\mathbf{H}_{\mathbf{S}}, \mathbf{C}_{\mathbf{S}} \in \mathbb{C}^{NM_t \times NM_t}$ are the Hessian of f evaluated at \mathbf{S} . To ensure uniqueness, we require $\mathbf{H}_{\mathbf{S}} = \mathbf{H}_{\mathbf{S}}^H, \mathbf{C}_{\mathbf{S}} = \mathbf{C}_{\mathbf{S}}^T$.

The complex-valued derivative $\mathbf{D}_{\mathbf{S}}$ is used in the modified steepest descent method. To calculate the Newton direction with the standard Newton method [88], we will use the second order Taylor series expansion of the function $f : \mathbb{R}^{2NM_t} \rightarrow \mathbb{R}$ in the well-known vector form with real-valued elements as

$$f(\mathbf{s} + \delta \mathbf{z}) = f(\mathbf{s}) + \delta \mathbf{z}^T \mathbf{d} + \frac{\delta^2}{2} \mathbf{z}^T \mathbf{H} \mathbf{z} + \mathcal{O}(\delta^3), \quad (5.33)$$

where the vector $\mathbf{s} \in \mathbb{R}^{2NM_t}$ is defined as

$$\mathbf{s} \triangleq \begin{pmatrix} \mathbf{s}^{\text{re}} \\ \mathbf{s}^{\text{im}} \end{pmatrix} \triangleq \begin{pmatrix} \Re \{\text{vec}(\mathbf{S})\} \\ \Im \{\text{vec}(\mathbf{S})\} \end{pmatrix}. \quad (5.34)$$

In the above, $\mathbf{d} \in \mathbb{R}^{2NM_t}$ is the derivative of $f(\mathbf{s})$ evaluated at \mathbf{s} , similarly defined in terms of its real and imaginary parts as \mathbf{s} in equation (5.34), and $\mathbf{H} \in \mathbb{R}^{2NM_t \times 2NM_t}$ is the Hessian of $f(\mathbf{s})$ evaluated at \mathbf{s} (for definitions, see Section 5.4.4). The derivatives and Hessians developed in the above two Taylor series expansion forms of (5.32) and (5.33) can be transformed into each other [89].

In the following, we list the derivative and Hessian of the cost function $f(\mathbf{S})$. The derivation details are given in Appendix C.

$$\mathbf{D}_{\mathbf{S}} = 2 \{ [(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{S} \mathbf{A}) - \mathbf{N}] \odot \mathbf{Y}^T \} \mathbf{A}^H, \quad (5.35)$$

$$\mathbf{H}_{\mathbf{S}} = 4 \mathbf{P}_{M_t \times N} (\mathbf{I}_N \otimes \mathbf{A}^*) \text{diag} \left(\text{vec} \left(2 \tilde{\mathbf{H}}^T - \mathbf{N}^T \right) \right) (\mathbf{I}_N \otimes \mathbf{A}^T) \mathbf{P}_{N \times M_t}, \quad (5.36)$$

$$\mathbf{C}_{\mathbf{S}} = 4 \mathbf{P}_{M_t \times N} (\mathbf{I}_N \otimes \mathbf{A}) \text{diag} (\text{vec} (\mathbf{Y}^* \odot \mathbf{Y}^*)) (\mathbf{I}_N \otimes \mathbf{A}^T) \mathbf{P}_{N \times M_t}, \quad (5.37)$$

where $\mathbf{N} = \frac{M_t}{N} \mathbf{1}_N \mathbf{1}_L^T$, $\tilde{\mathbf{H}} = (\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{S} \mathbf{A}) \in \mathbb{R}^{N \times L}$ and $\mathbf{Y} = \mathbf{A}^T \mathbf{S}^T \in \mathbb{C}^{L \times N}$. In the above, $\mathbf{P}_{N \times M_t}$ is a commutation matrix, such that

$$\text{vec} (\mathbf{Z}^T) = \mathbf{P}_{N \times M_t} \text{vec} (\mathbf{Z}), \quad (5.38)$$

which can be expressed as [90]

$$\mathbf{P}_{N \times M_t} = \sum_{m=1}^N \sum_{n=1}^{M_t} (\mathbf{E}_{mn} \otimes \mathbf{E}_{mn}^T), \quad (5.39)$$

where \mathbf{E}_{mn} is a matrix of dimension $N \times M_t$ with 1 at its mn -th position and zeros elsewhere. It holds that $\mathbf{P}_{N \times M_t} = \mathbf{P}_{M_t \times N}^T$. It is easy to verify that $\mathbf{H}_{\mathbf{S}} = \mathbf{H}_{\mathbf{S}}^H$, $\mathbf{C}_{\mathbf{S}} = \mathbf{C}_{\mathbf{S}}^T$.

5.4.3 Modified Steepest Descent on the Complex Stiefel Manifold

Here, we apply the modified steepest descent method of [85] to solve the optimization problem of (5.30). Let $\mathcal{T}_{\mathbf{S}}(N, M_t)$ denote the tangent space, i.e., the plane that is tangent to the complex Stiefel manifold at point $\mathbf{S} \in \mathcal{S}(N, M_t)$ [84]. The inner product

in the tangent space is defined using the canonical metric [84] in the complex-value case, i.e.,

$$\langle \mathbf{Z}_1, \mathbf{Z}_2 \rangle = \Re \left\{ \text{tr} \left[\mathbf{Z}_2^H \left(\mathbf{I} - \frac{1}{2} \mathbf{S} \mathbf{S}^H \right) \mathbf{Z}_1 \right] \right\}, \quad (5.40)$$

for $\mathbf{Z}_1, \mathbf{Z}_2 \in \mathcal{T}_{\mathbf{S}}(N, M_t)$.

Let $\mathbf{Z}^k \in \mathcal{T}_{\mathbf{S}}(N, M_t)$ be the steepest descent at point $\mathbf{S}^k \in \mathcal{S}(N, M_t)$ in the k -th iteration. The steepest descent algorithm starts from \mathbf{S}^k and moves along \mathbf{Z}^k with a step size δ , i.e.,

$$\mathbf{S}^{k+1} = \mathbf{S}^k + \delta \mathbf{Z}^k. \quad (5.41)$$

To preserve the orthogonality during the update steps, the new point \mathbf{S}^{k+1} is projected back to the complex Stiefel manifold, i.e., $\mathbf{S}^{k+1} = \Pi(\mathbf{S}^k + \delta \mathbf{Z}^k)$, where Π is the projection operator. For a matrix $\mathbf{S} \in \mathbb{C}^{N \times M_t}$ with $N \geq M_t$ and with SVD $\mathbf{S} = \tilde{\mathbf{U}} \mathbf{\Sigma} \tilde{\mathbf{V}}^H$, the point in the Stiefel manifold that is nearest to \mathbf{S} in the Frobenius norm sense is given by $\Pi(\mathbf{S}) = \tilde{\mathbf{U}} \mathbf{I}_{N, M_t} \tilde{\mathbf{V}}^H$ [85].

The modified steepest descent is defined as follows [85]. Let $g(\mathbf{Z}^k) = f(\Pi(\mathbf{S}^k + \mathbf{Z}^k))$ be the local cost function for $\mathbf{S}^k \in \mathcal{S}(N, M_t)$. The gradient of $g(\mathbf{Z}^k)$ at $\mathbf{Z}^k = \mathbf{0}$ under the canonical inner product (5.40) is

$$\tilde{\nabla}_{\mathbf{S}} f(\mathbf{S}^k) = \nabla_{\mathbf{S}} f(\mathbf{S}^k) - \mathbf{S}^k \left(\nabla_{\mathbf{S}} f(\mathbf{S}^k) \right)^H \mathbf{S}^k, \quad (5.42)$$

where $\nabla_{\mathbf{S}} f(\mathbf{S}) = \mathbf{D}_{\mathbf{S}}$ denotes the derivative of $f(\mathbf{S})$ (see (5.35)). Then, the modified steepest descent is $\mathbf{Z}^k = -\tilde{\nabla}_{\mathbf{S}} f(\mathbf{S}^k)$.

The step size δ is chosen using a nonmonotone line search method based on [87], i.e., so that

$$f(\Pi(\mathbf{S}^k + \delta \mathbf{Z}^k)) \leq C_k + \beta \delta \langle \tilde{\nabla}_{\mathbf{S}} f(\mathbf{S}^k), \mathbf{Z}^k \rangle, \quad (5.43)$$

$$\langle \tilde{\nabla}_{\mathbf{S}} f(\Pi(\mathbf{S}^k + \delta \mathbf{Z}^k)), \mathbf{Z}^k \rangle \geq \sigma \langle \tilde{\nabla}_{\mathbf{S}} f(\mathbf{S}^k), \mathbf{Z}^k \rangle. \quad (5.44)$$

Here, C_k is taken to be a convex combination of function values $f(\mathbf{S}^0), f(\mathbf{S}^1), \dots, f(\mathbf{S}^k)$, i.e.,

$$C_{k+1} = \left[\eta Q_k C_k + f(\mathbf{S}^{k+1}) \right] / Q_{k+1}, \quad (5.45)$$

where $Q_{k+1} = \eta Q_k + 1$, $C_0 = f(\mathbf{S}^0)$ and $Q_0 = 1$. In the above, the parameter η controls the degree of nonmonotonicity. When $\eta = 0$, the line search is the usual monotone Wolfe or Armijo line search [91]. When $\eta = 1$, then

$$C_k = \frac{1}{k+1} \sum_{i=0}^k f(\mathbf{S}^{i+1}). \quad (5.46)$$

The modified steepest descent algorithm is summarized in Algorithm 1.

Algorithm 1 Modified steepest descent algorithm

- 1: **Initialize:** Choose $\mathbf{S}^0 \in \mathcal{S}(N, M_t)$ and parameters $\alpha, \eta, \epsilon \in (0, 1)$, $0 < \beta < \sigma < 1$.
Set $\delta = 1, C_0 = f(\mathbf{S}^0), Q_0 = 1, k = 0$.
 - 2: **Descent direction update:** Compute the descent direction as $\mathbf{Z}^k = -\tilde{\nabla}_{\mathbf{S}} f(\mathbf{S}^k)$ via equation (5.42).
 - 3: **Convergence test:** If $\langle \mathbf{Z}^k, \mathbf{Z}^k \rangle \leq \epsilon$, then stop.
 - 4: **Line search update:** Compute $\mathbf{S}^{k+1} = \Pi(\mathbf{S}^k + \delta \mathbf{Z}^k)$ and $\tilde{\nabla}_{\mathbf{S}} f(\mathbf{S}^{k+1})$. If $f(\mathbf{S}^{k+1}) \geq \beta \delta \langle \tilde{\nabla}_{\mathbf{S}} f(\mathbf{S}^k), \mathbf{Z}^k \rangle + C_k$ and $\langle \tilde{\nabla}_{\mathbf{S}} f(\mathbf{S}^{k+1}), \mathbf{Z}^k \rangle \leq \sigma \langle \tilde{\nabla}_{\mathbf{S}} f(\mathbf{S}^k), \mathbf{Z}^k \rangle$, then set $\delta = \alpha \delta$ and repeat Step 4.
 - 5: **Cost update:** $Q_{k+1} = \eta Q_k + 1$, $C_{k+1} = [\eta Q_k C_k + f(\mathbf{S}^{k+1})] / Q_{k+1}$.
 - 6: Perform update $\mathbf{S}^{k+1} = \Pi(\mathbf{S}^k + \delta \mathbf{Z}^k)$, $k = k + 1$. Go to Step 2.
-

5.4.4 Modified Newton Algorithm on the Complex Stiefel Manifold

With expressions for the derivative and Hessian of the cost function given in (5.35), (5.36) and (5.37), we can now formulate the Newton method [88] to solve the waveform design problem of (5.30).

First, the Newton search direction is calculated as follows. Let $\mathbf{Z}^k \in \mathbb{C}^{N \times M_t}$ denote the Newton search direction in the k -th iteration. In a similar way as in [92], we arrange the complex-valued elements of \mathbf{Z}^k into a real-valued vector \mathbf{z}^k of length $2NM_t$, defined as

$$\mathbf{z}^k \triangleq \begin{pmatrix} \mathbf{z}_{\text{re}}^k \\ \mathbf{z}_{\text{im}}^k \end{pmatrix} \triangleq \begin{pmatrix} \Re \{ \text{vec}(\mathbf{Z}^k) \} \\ \Im \{ \text{vec}(\mathbf{Z}^k) \} \end{pmatrix}. \quad (5.47)$$

Let us also define the real-valued vector

$$\mathbf{d}^k \triangleq \begin{pmatrix} \Re \{ \text{vec}(\mathbf{D}_{\mathbf{S}^k}) \} \\ \Im \{ \text{vec}(\mathbf{D}_{\mathbf{S}^k}) \} \end{pmatrix}, \quad (5.48)$$

and the real-valued matrix

$$\mathbf{H}^k \triangleq \begin{bmatrix} \Re \{ \mathbf{H}_{\mathbf{S}^k} + \mathbf{C}_{\mathbf{S}^k} \} & -\Im \{ \mathbf{H}_{\mathbf{S}^k} + \mathbf{C}_{\mathbf{S}^k} \} \\ \Im \{ \mathbf{H}_{\mathbf{S}^k} - \mathbf{C}_{\mathbf{S}^k} \} & \Re \{ \mathbf{H}_{\mathbf{S}^k} - \mathbf{C}_{\mathbf{S}^k} \} \end{bmatrix}. \quad (5.49)$$

Following the standard Newton method, the vector \mathbf{z}^k is computed as

$$\mathbf{z}^k = -[\mathbf{H}^k + \sigma_k \mathbf{I}]^{-1} \mathbf{d}^k, \quad (5.50)$$

where $\sigma_k \geq 0$ is chosen to make the matrix $\mathbf{H}^k + \sigma_k \mathbf{I}$ positive definite. Consequently, the complex-valued Newton search direction can be found as $\mathbf{Z}^k = \text{mat}_{N \times M_t}(\mathbf{z}_{\text{re}}^k + j\mathbf{z}_{\text{im}}^k)$, corresponding to the inverse vector operation defined in (5.47).

In the k -th iteration, the standard Newton method performs the update

$$\mathbf{S}^{k+1} = \mathbf{S}^k + \delta \mathbf{Z}^k, \quad (5.51)$$

where δ is the step size. Since the waveforms are on the complex Stiefel manifold, to preserve the orthogonality in the modified Newton method, the new point \mathbf{S}^{k+1} is projected back to the complex Stiefel manifold, i.e.,

$$\mathbf{S}^{k+1} = \Pi(\mathbf{S}^k + \delta \mathbf{Z}^k). \quad (5.52)$$

It should be pointed out that the Hessian matrix defined in (5.49) is not always positive definite. In each step we choose σ_k such that $\mathbf{H}_k + \sigma_k \mathbf{I}$ is positive definite. If matrix \mathbf{H}_k has nonpositive eigenvalues, σ_k should be larger than $-\lambda_{\min}(\mathbf{H}_k)$, where λ_{\min} is the minimal eigenvalue of \mathbf{H}_k . In the local area of the minimum, the modified Newton update will approach the pure Newton step. However, if σ_k is chosen very large, the modified Newton search direction will be close to the negative steepest descent. Last, the step size δ could be obtained using a similar nonmonotone line search method [87]. Throughout the modified Newton method, the inner product of two matrices $\mathbf{Z}_1, \mathbf{Z}_2 \in \mathbb{C}^{N \times M_t}$ is defined as $\langle \mathbf{Z}_1, \mathbf{Z}_2 \rangle = \text{tr}(\mathbf{Z}_1^H \mathbf{Z}_2)$. The modified Newton algorithm is summarized in Algorithm 2.

Algorithm 2 Modified Newton algorithm

- 1: **Initialize:** Choose $\mathbf{S}^0 \in \mathcal{S}(N, M_t)$ and parameters $\alpha, \eta, \epsilon \in (0, 1)$, $0 < \beta < \sigma < 1$.
Set $\delta = 1, C_0 = f(\mathbf{S}^0), Q_0 = 1, k = 0$.
 - 2: Compute the derivative $\mathbf{D}_{\mathbf{S}^k}$ with equation (5.35) as well as Hessian $\mathbf{H}_{\mathbf{S}^k}, \mathbf{C}_{\mathbf{S}^k}$ with equations (5.36) and (5.37).
 - 3: **Convergence test:** If $\langle \mathbf{D}_{\mathbf{S}^k}, \mathbf{D}_{\mathbf{S}^k} \rangle \leq \epsilon$, then stop.
 - 4: **Newton search direction computation:** Compute the real-valued vector \mathbf{d}^k with (5.48), as well as the real-valued matrix \mathbf{H}^k with (5.49). Compute the vector \mathbf{z}^k with (5.50). The Newton search direction is arranged as $\mathbf{Z}^k = \text{mat}_{N \times M_t}(\mathbf{z}_{\text{re}}^k + j\mathbf{z}_{\text{im}}^k)$.
 - 5: **Line search update:** Compute $\mathbf{S}^{k+1} = \Pi(\mathbf{S}^k + \delta \mathbf{Z}^k)$ and $\mathbf{D}_{\mathbf{S}^{k+1}}$. If $f(\mathbf{S}^{k+1}) \geq \beta \delta \Re\{\langle \mathbf{D}_{\mathbf{S}^{k+1}}, \mathbf{Z}^k \rangle\} + C_k$ and $\Re\{\langle \mathbf{D}_{\mathbf{S}^{k+1}}, \mathbf{Z}^k \rangle\} \leq \sigma \Re\{\langle \mathbf{D}_{\mathbf{S}^k}, \mathbf{Z}^k \rangle\}$, then set $\delta = \alpha \delta$ and repeat Step 5.
 - 6: **Cost update:** $Q_{k+1} = \eta Q_k + 1$, $C_{k+1} = [\eta Q_k C_k + f(\mathbf{S}^{k+1})]/Q_{k+1}$.
 - 7: Perform update $\mathbf{S}^{k+1} = \Pi(\mathbf{S}^k + \delta \mathbf{Z}^k)$, $k = k + 1$. Go to Step 2.
-

5.5 Performance Analysis of the Proposed Waveform When Targets Fall In Different Range Bins

In the above derivation of the optimum waveform we used the model of [18], which is valid for targets falling in the same range bin. When the targets fall in different range bins, the model has to be modified to account for different delays in the transmit waveforms, corresponding to the different targets. In that case, the orthogonality of the delayed waveforms cannot be guaranteed.

In this section, we consider the scenario of targets falling in different range bins, and determine the effect on performance when using the waveforms designed in Section 5.4, i.e., waveforms optimized under the simplified scenario of targets falling in the same range bin. In the following we will assume that the orthogonal waveforms satisfy (5.21).

To model this case we will adopt the model of [51], modified for a ULA array configuration and assuming that the Doppler shifts within a pulse can be ignored because the targets are moving slowly. Let $d_{\max}(0)$ and $d_{\min}(0)$ be the maximum and minimum

ranges of the K far field targets at the initial sampling time. The maximum range delay normalized by T_s is

$$N_1 = \left\lfloor \frac{2(d_{\max}(0) - d_{\min}(0))}{cT_s} \right\rfloor. \quad (5.53)$$

Let us set the length of the sampling window at the receiver end to $\tilde{N} = N + N_1$, where $N_1 < N$ and $\tilde{N} < T_{PRI}$; we will assume that the target reflections fall within this sampling window. During the q -th pulse, the k -th target echo received at the l -th receive antenna and the corresponding demodulated baseband signal are given in equations (3) and (4) of [51], respectively. For our scenario, the noise-free sampled received data at the l -th receive antenna during the q -th pulse can be approximated as

$$\mathbf{r}_l^T \approx \sum_{k=1}^K \beta_k \zeta_{qk} e^{\frac{-j2\pi 2d_k(0)}{\lambda}} e^{-j2\pi(l-1)\alpha_k^r} \mathbf{a}^T(\theta_k) \mathbf{S}^T \mathbf{C}_{\tau_k}^T, \quad (5.54)$$

where $\alpha_k^r \triangleq \frac{d_r \sin(\theta_k)}{\lambda}$, $\tau_k = \left\lfloor \frac{2(d_k(0) - d_{\min}(0))}{cT_s} \right\rfloor$ with $d_k(0)$ denoting the range of the k -th target at the initial sampling time; ζ_{qk} is the Doppler shift defined in (5.4); $\mathbf{C}_{\tau_k} = \begin{bmatrix} \mathbf{0}_{N \times \tau_k} & \mathbf{I}_N & \mathbf{0}_{N \times (N_1 - \tau_k)} \end{bmatrix}^T \in \mathbb{C}^{\tilde{N} \times N}$; $\mathbf{a}(\theta_k) = [1, e^{-j2\pi\alpha_k^t}, \dots, e^{-j2\pi(M_t-1)\alpha_k^t}]^T$ is the transmit steering vector. By collecting samples from all receive antennas at the fusion center, the noise-free data matrix can be constructed as

$$\mathbf{W} = [\mathbf{r}_1, \dots, \mathbf{r}_{M_r}]^T = \mathbf{B} \mathbf{\Sigma} \mathbf{\Gamma}^T, \quad (5.55)$$

where

$$\begin{aligned} \mathbf{\Gamma} &= [\mathbf{C}_{\tau_1} \mathbf{S} \mathbf{a}(\theta_1), \dots, \mathbf{C}_{\tau_K} \mathbf{S} \mathbf{a}(\theta_K)] \in \mathbb{C}^{\tilde{N} \times K}, \\ \mathbf{\Sigma} &= \text{diag} \left\{ \left[\beta_1 \zeta_{q1} e^{\frac{-j2\pi 2d_1(0)}{\lambda}}, \dots, \beta_K \zeta_{qK} e^{\frac{-j2\pi 2d_K(0)}{\lambda}} \right] \right\}. \end{aligned}$$

We should note that although \mathbf{W} and $\mathbf{\Sigma}$ do depend on the pulse index, this dependence will not be shown for notational convenience.

To analyze the coherence $\mu(V)$ of \mathbf{W} , we do a QR decomposition of $\mathbf{\Gamma}$, along the lines of Section 5.3, i.e., $\mathbf{\Gamma} = \mathbf{Q}_s \mathbf{R}_s$ with $\mathbf{Q}_s \in \mathbb{C}^{\tilde{N} \times K}$ such that $\mathbf{Q}_s^H \mathbf{Q}_s = \mathbf{I}_K$. Let $\mathbf{Q}_s^{*(i)}$ denote the i -th row of \mathbf{Q}_s^* . It holds that

$$\mu(V) = \frac{\tilde{N}}{K} \sup_{i \in \mathbb{N}_N^+} \left\| \mathbf{Q}_s^{*(i)} \right\|^2 \leq \frac{\tilde{N}}{K} \sup_{i \in \mathbb{N}_N^+} \frac{\left\| \mathbf{\Gamma}^{*(i)} \right\|^2}{\sigma_{\min}^2(\mathbf{R}_s^*)}, \quad (5.56)$$

where $\sigma_{\min}^2(\mathbf{R}_s^*) = \lambda_{\min}(\mathbf{R}_s^H \mathbf{Q}_s^H \mathbf{Q}_s \mathbf{R}_s) = \lambda_{\min}(\Phi)$ and $\Phi = \Gamma^H \Gamma$ with $[\Phi]_{k_1 k_2} = \mathbf{a}^H(\theta_{k_1}) \mathbf{S}^H \mathbf{C}_{\tau_{k_1}}^T \mathbf{C}_{\tau_{k_2}} \mathbf{S} \mathbf{a}(\theta_{k_2})$. It holds that $\mathbf{S}^H \mathbf{C}_{\tau_{k_1}}^T \mathbf{C}_{\tau_{k_2}} \mathbf{S} = \mathbf{S}^H \mathbf{J}_{\tau_{k_2} - \tau_{k_1}} \mathbf{S} = \mathbf{R}_{\tau_{k_2} - \tau_{k_1}}^T$, where \mathbf{R}_n is the waveform covariance matrix [80] with $[\mathbf{R}_n]_{m_1 m_2} = \sum_{k=n+1}^N s_{m_1}(k) s_{m_2}^*(k-n)$ and \mathbf{J}_n is a shifting matrix defined as [80]

$$\mathbf{J}_n = \begin{bmatrix} \overbrace{0 \cdots 0}^n 1 & & 0 \\ & \ddots & \\ & & 1 \\ 0 & & & \end{bmatrix}_{N \times N} = \mathbf{J}_{-n}^T, \quad (5.57)$$

for $n = 0, \dots, N_1$. Since the waveforms are orthogonal it holds that $\mathbf{R}_0 = \mathbf{I}_{M_t}$, and Φ becomes

$$\Phi = \begin{bmatrix} M_t & \cdots & \mathbf{a}^H(\theta_1) \mathbf{R}_{\tau_K - \tau_1}^* \mathbf{a}(\theta_K) \\ \vdots & \ddots & \vdots \\ \mathbf{a}^H(\theta_K) \mathbf{R}_{\tau_K - \tau_1}^T \mathbf{a}(\theta_1) & \cdots & M_t \end{bmatrix}. \quad (5.58)$$

In order to compute bounds for $\lambda_{\min}(\Phi)$, we need to compute $\text{tr}(\Phi)$ and $\text{tr}(\Phi^2)$ (see [93]). It can be seen than $\text{tr}(\Phi) = K M_t$ and $\text{tr}(\Phi^2) = K M_t^2 + 2 \sum_{k_2=k_1+1}^K \sum_{k_1=1}^K |[\Phi]_{k_1 k_2}|^2$ with

$$\begin{aligned} [\Phi]_{k_1 k_2} &= \mathbf{a}^H(\theta_{k_1}) \mathbf{R}_{\tau_{k_2} - \tau_{k_1}}^* \mathbf{a}(\theta_{k_2}) \\ &= \sum_{m_2=1}^{M_t} \sum_{m_1=1}^{M_t} e^{j2\pi[(m_1-1)\alpha_{k_1}^t - (m_2-1)\alpha_{k_2}^t]} r_{m_1 m_2}^*(\tau_{k_2} - \tau_{k_1}). \end{aligned} \quad (5.59)$$

The lower bound of $\lambda_{\min}(\Phi)$, can be found as (see Theorem 2.1 in [93])

$$\begin{aligned} \lambda_{\min}(\Phi) &\geq \frac{\text{tr}(\Phi)}{K} - \sqrt{K-1} \sqrt{\frac{\text{tr}(\Phi^2)}{K} - \left(\frac{\text{tr}(\Phi)}{K}\right)^2} \\ &= M_t - \sqrt{1 - \frac{1}{K}} \sqrt{2 \sum_{k_2=k_1+1}^K \sum_{k_1=1}^K |[\Phi]_{k_1 k_2}|^2}. \end{aligned} \quad (5.60)$$

Next, we find the minimum supremum of $\|\mathbf{\Gamma}^{*(i)}\|^2$ over $i \in \mathbb{N}_N^+$. Without loss of generality, we assume that the ranges of targets are ordered as $d_1(0) \leq \dots \leq d_K(0)$. Thus, $\tau_1 = 0, \tau_K = N_1$. To better understand the term $\|\mathbf{\Gamma}^{*(i)}\|^2$, in Fig. 5.3 we give an

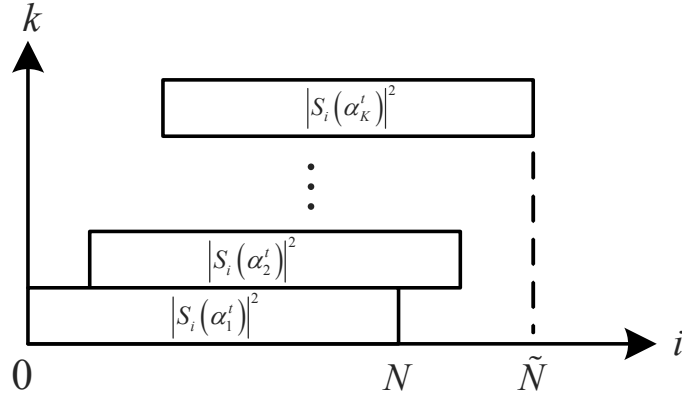


Figure 5.3: Illustration of transmit waveform spectra in the order of the transmit waveform arrival at the receiver end.

illustration of the power spectra of the transmitted waveforms corresponding to spatial frequency α_k^t , in the order in which the waveforms arrive at the receiver end. It holds that

$$\left\| \mathbf{\Gamma}^{*(i)} \right\|^2 = \begin{cases} |S_i(\alpha_1^t)|^2, 1 \leq i \leq \tau_2 \\ \sum_{k=1}^2 |S_{i-\tau_k}(\alpha_k^t)|^2, \tau_2 < i \leq \tau_3 \\ \vdots \\ \sum_{k=1}^K |S_{i-\tau_k}(\alpha_k^t)|^2, N_1 < i \leq N \\ \sum_{k=2}^K |S_{i-\tau_k}(\alpha_k^t)|^2, N < i \leq N + \tau_2 \\ \vdots \\ |S_{i-N_1}(\alpha_K^t)|^2, N + \tau_{K-1} < i \leq N + N_1 \end{cases} \quad (5.61)$$

It is easy to verify that

$$\sum_{i=1}^{\tilde{N}} \left\| \mathbf{\Gamma}^{*(i)} \right\|^2 = \sum_{i=1}^N \sum_{k=1}^K |S_i(\alpha_k^t)|^2 = KM_t. \quad (5.62)$$

It can be seen that the maximum value of $\left\| \mathbf{\Gamma}^{*(i)} \right\|^2$ over $i \in \mathbb{N}_{\tilde{N}}^+$ is determined by the spatial power spectra $S_i(\alpha_k^t)$, $i \in \mathbb{N}_N^+$ as well as the delay corresponding to each target. Based on the conditions of (5.21), it holds that $\sup_{i \in \mathbb{N}_{\tilde{N}}^+} \left\| \mathbf{\Gamma}^{*(i)} \right\|^2 = \sum_{k=1}^K |S_{i-\tau_k}(\alpha_k^t)|^2 = \frac{KM_t}{N}$.

Combining the above we get

$$\mu(V) \leq \frac{\left(1 + \frac{N_1}{N}\right) M_t}{M_t - \sqrt{1 - \frac{1}{K}} \sqrt{2 \sum_{k_2=k_1+1}^K \sum_{k_1=1}^K |[\Phi]_{k_1 k_2}|^2}}. \quad (5.63)$$

One can see that as M_t increases, the above bound tends to $1 + \frac{N_1}{N} < 2$.

Corollary 1. *If the orthogonal waveforms not only satisfied the sufficient condition stated in (5.21) but were also designed to have zero auto and cross-correlations for maximal normalized delay range N_1 , i.e., $\mathbf{R}_n = \mathbf{0}_{M_t}$, $n = 1, \dots, N_1$, then it would hold that $\Phi = M_t \mathbf{I}_{M_t}$ and thus $\lambda_{\min}(\Phi) = M_t$. As a result,*

$$\mu(V) \leq \left(1 + \frac{N_1}{N}\right) < 2. \quad (5.64)$$

Proof. The conclusion is straightforward from the above analysis and the proof is omitted here. \square

The above analysis shows that, if the proposed waveforms were to be used in a scenario in which the targets fall in different range bins, the coherence $\mu(V)$ would increase slightly as the maximum normalized range delay increases. Fortunately, the coherence $\mu(V)$ is bounded by 2 for most of the possible values of $N_1 < N$ (see Fig. 5.7 (b) for an example). Our analysis also indicates that if the proposed waveforms were designed to additionally have good correlation properties, they would yield the lowest upper bound of coherence $\mu(V)$. One could design waveforms with good correlation properties for the scenario in which targets fall into different range bins. The condition of (5.21) can be combined with waveform correlation requirements in waveform design in MIMO radars [80] [94]. This kind of design will be subject of future study.

5.6 Numerical Results

In this section, we provide numerical results to demonstrate the performance of the proposed waveform matrix design schemes. Both transmit and receive antennas are configured as ULAs with $d_t = \lambda/2$, $d_r = M_t \lambda/2$ and carrier frequency $f_c = 5 \times 10^9 \text{ Hz}$. The pulse repetition interval is $T_{PRI} = 1/4000 \text{ s}$ and pulse duration is $T_\phi = N \times 8 \times$

10^{-7} s. The target reflection coefficients $\{\beta_k\}_{k \in \mathbb{N}_K^+}$ remain constant during the coherent processing interval [74].

5.6.1 Performance Comparison of Waveform Design Methods

We first design the waveform matrix by applying the modified steepest descent method. We take $N = 64, M_t = 40$ for DOA space $[-10^\circ : 1^\circ : 10^\circ]$; correspondingly, $\alpha_k^t \in [-0.0868, 0.0868]$; the number of discretized angles is $L = 21$ and the steering matrix \mathbf{A} has a dimension 40×21 . In the nonmonotone line search, we chose $\beta = 0.01$ and $\sigma = 0.99$. These values are selected by trial and error to satisfy the Wolfe conditions (5.43) and (5.44). In addition, we set $\alpha = 0.5$ to adjust the step size, and $\epsilon = 10^{-5}$ as the stopping check value. The initial step size is set to $\delta = 0.1$. The iteration is initialized with a column-wise Hadamard matrix $\mathbf{S}^0 \in \mathcal{S}(64, 40)$, i.e., a matrix that has Hadamard sequences in its columns. The convergence of the proposed modified steepest descent algorithm for $\eta = 1, 0.5, 0$ is shown in Fig. 5.4 (a). As it can be seen from Fig. 5.4 (a), the objective value f under $\eta = 0.5$ decreases the fastest, while under $\eta = 1$ decreases very slowly for number of iterations less than 2000. The simulation results indicate that the performance of the line search method could be improved if historical objective values in each iteration are partially utilized, as indicated in (5.45). The simulation results also show that the value of the objective function, f , approaches its global minimal, i.e., 0. The corresponding optimal solution, \mathbf{S} , is not unique, and depends on the initial point and the step size. Based on extensive simulations, not shown here due to space constraints, all solutions result in very similar MC recovery performance (see simulations in [95]).

Since the complex Stiefel manifold is not a convex set, there is no guarantee that the algorithms will converge to the global minimum. In the problem of (5.30), the number of equations is $\frac{M_t(M_t+1)}{2} + NL$, which should be less than the total available combinations, $2M_tN$. Consequently, to make the objective function zero, it must hold that $L < 2M_t - \frac{M_t^2 + M_t}{2N}$. Our simulations show that for the entire DOA space $[-90^\circ : 1^\circ : 90^\circ]$, corresponding to $L = 181$, when the dimension of \mathbf{S} is relatively small, e.g., $N = 64, M_t = 40$ and therefore $L > 2M_t - \frac{M_t^2 + M_t}{2N} \approx 67$, the objective value gets stuck to

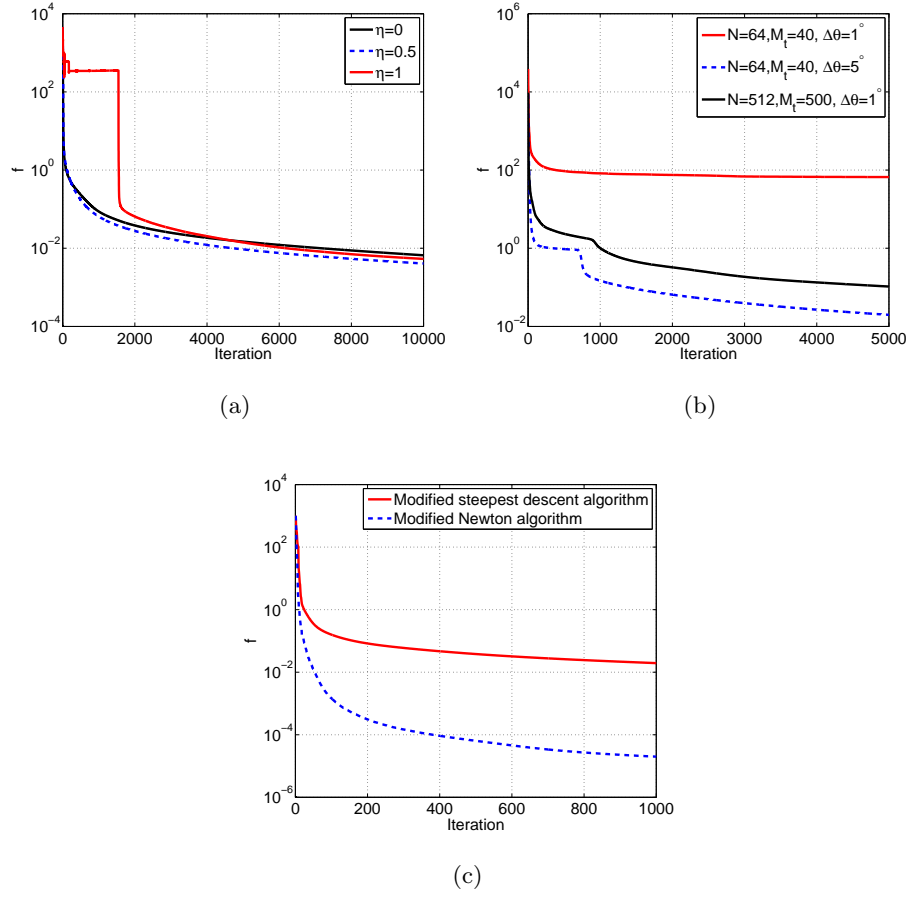


Figure 5.4: Objective function of (5.29) vs. iterations: (a) DOA space $[-10^\circ : 1^\circ : 10^\circ]$ under modified steepest descent algorithm; (b) DOA space $[-90^\circ : \Delta\theta : 90^\circ]$, $\Delta\theta = 1^\circ, 5^\circ$ under the modified steepest descent algorithm; (c) comparison of the modified steepest descent algorithm and the modified Newton algorithm for $M_t = 20, N = 32$ and DOA space $[-5^\circ : 1^\circ : 5^\circ]$.

local minima; however, if the spacing increases, for example to 5° , corresponding to $L = 37$, the iteration converges to the global minimum. If the dimension is relatively large, e.g., $N = 512, M_t = 500$, even for small spacing, i.e., 1° , the objective value converges to its global minimum (see Fig. 5.4 (b) for $\eta = 0$).

Next we provide a numerical example to compare the performance of the modified steepest descent algorithm and the modified Newton algorithm. Since the complexity of the Newton method increases with the size of the matrix, we do the comparison for $N = 32, M_t = 20$ and DOA space $[-5^\circ : 1^\circ : 5^\circ]$. A column-wise Hadamard waveform matrix $\mathbf{S}^0 \in \mathbb{C}^{32 \times 20}$ is used as initial search point for both algorithms. The performance comparison is illustrated in Fig. 5.4 (c), where it can be seen that the value of the objective function, $f(\mathbf{S})$, under the modified Newton algorithm decreases much faster than that under the modified steepest descent algorithm.

5.6.2 Spatial Power Spectra of Optimized Waveform Snapshots

The power spectra of the optimized waveform snapshots, i.e., the rows of the optimized waveform matrix \mathbf{S} , are plotted in Fig. 5.5 (b) and (d) for the same parameters defined in Section 5.6.1, i.e., $M_t = 40, N = 64$ and DOA space $[-10^\circ : 1^\circ : 10^\circ]$, corresponding to $\alpha_k^t \in [-0.0868, 0.0868]$. In the simulation, a waveform matrix that is either column-wise Gaussian Orthogonal (G-Orth) or Hadamard, is used as the initial search point, respectively. The power spectra of the rows of the initial waveform matrix fluctuate over different DOAs (see Fig. 5.5 (a) and (c), respectively). As can be seen in Fig. 5.5 (b) and (d), the optimized waveform matrix yields almost flat row power spectra, with value $\frac{M_t}{N} = 0.625$, which, according to Theorem 2, corresponds to the lower possible coherence value.

5.6.3 Coherence Properties Under Optimized Waveforms

Targets in the Same Range Bin

We first look at the coherence properties under the optimized waveforms for the scenario that all targets fall in the same range bin. In Fig. 5.6, we plot the coherence $\mu(V)$

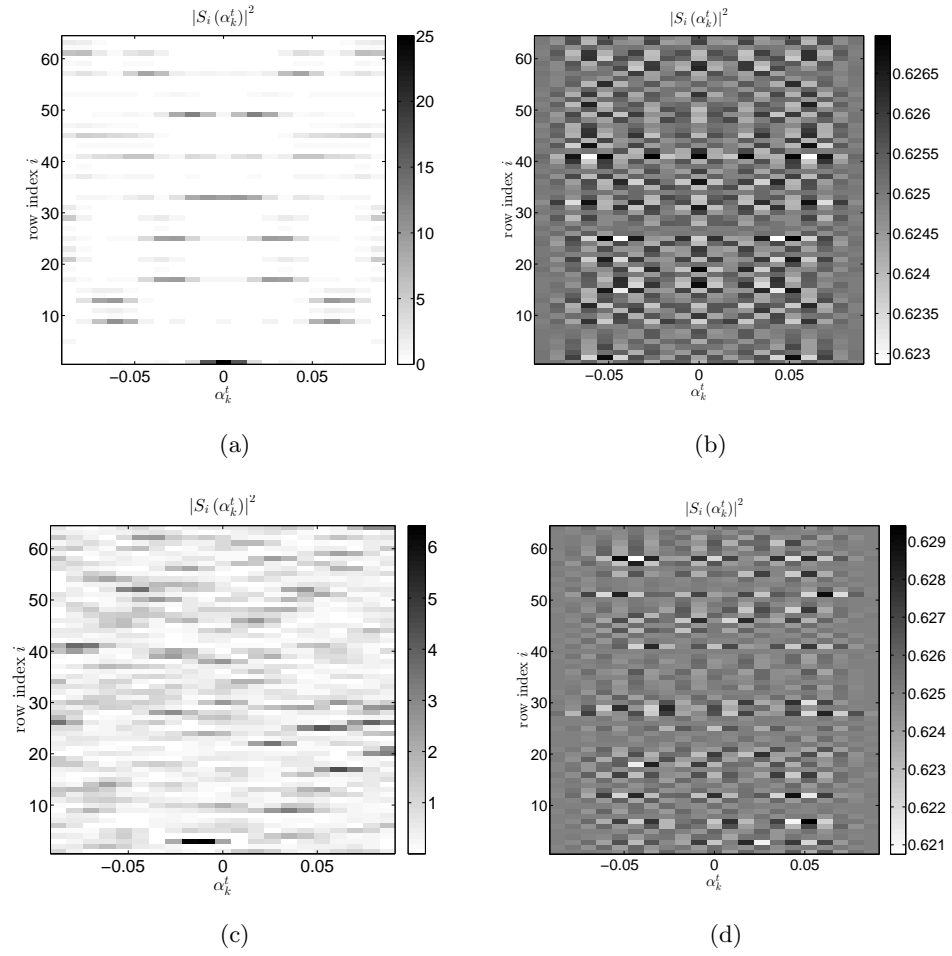


Figure 5.5: Power spectra $|S_i(\alpha_k^t)|^2$ for $\alpha_k^t \in [-0.0868, 0.0868]$ with $M_t = 40$ and $N = 64$ and DOA space $[-10^\circ : 1^\circ : 10^\circ]$. (a) Hadamard waveforms; (b) Optimized waveforms using Hadamard waveforms as initialization; (c) G-Orth waveforms; (d) Optimized waveforms using G-Orth waveforms as initial.

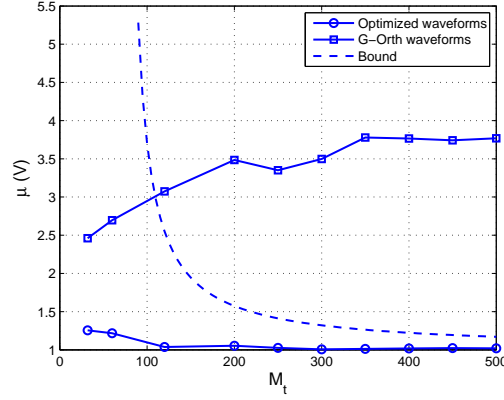


Figure 5.6: Coherence $\mu(V)$ and its bound defined in (5.26) for targets in the same range bin. The $K = 4$ targets are located at $[-10^\circ, -5^\circ, 0^\circ, 1^\circ]$.

of matrix \mathbf{W} and its bound, defined in (5.26), versus the number of transmit antennas, for $K = 4$ targets located at angles $[-10^\circ, -5^\circ, 0^\circ, 1^\circ]$. The optimized waveforms for different values of M_t are obtained by solving the problem of (5.30) via Algorithm 1 focusing on DOA space $[-10^\circ : 1^\circ : 10^\circ]$, i.e., $L = 21$. For comparison, the coherence $\mu(V)$ under the G-Orth waveform matrix is also plotted, where the results are averaged over 100 independent implementations, and in each implementation the waveforms are generated randomly. It can be found that the averaged coherence under G-Orth waveforms is higher than the coherence under optimized waveforms over the entire M_t range. On the other hand, under the optimized waveforms, our simulations show that for different number of targets, the coherence is always bounded by the bound of (5.26) and approaches its smallest value (not necessarily in a monotone way) when M_t increases. The simulation results in Fig. 5.6 confirm the conclusions in Theorem 3, i.e., when the waveforms satisfy the optimal waveform conditions stated in Theorem 2, the matrix coherence $\mu(V)$ is asymptotically optimal w.r.t. M_t . We should note, however, that the rather big coherence difference between the optimized and the G-Orth waveforms, does not translate in to substantial difference in terms of matrix recovery error. Indeed, the G-Orth waveforms perform very closely with the optimized ones when M_t becomes larger.

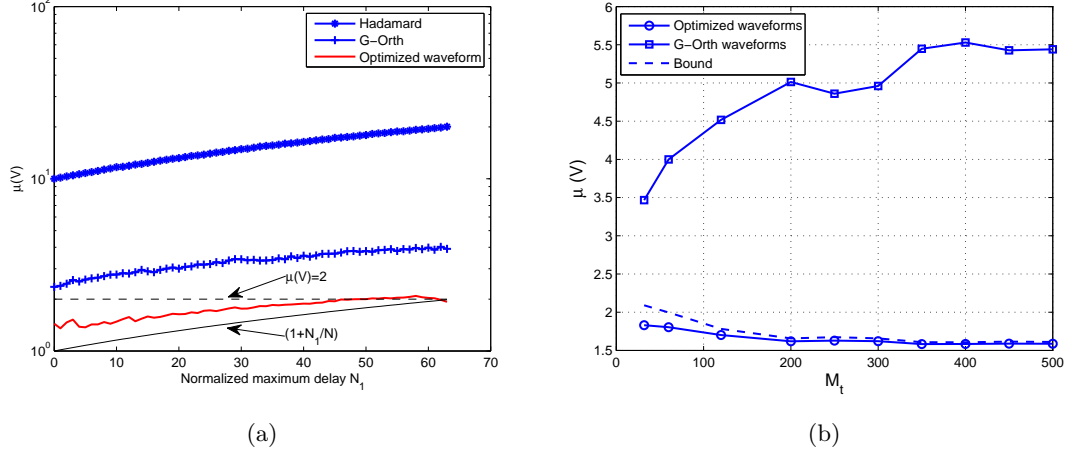


Figure 5.7: Coherence $\mu(V)$ for $K = 4$ targets located at $[-10^\circ, -5^\circ, 0^\circ, 1^\circ]$. (a) $\mu(V)$ versus N_1 with $M_t = 40, N = 64$ for targets in different range bins; (b) $\mu(V)$ and it bound defined in (5.63) with $N_1 = N/2$ for targets in different range bins.

Targets in Different Range Bins

Next, we conduct simulations considering the scenario of targets falling into different range bins, and test the coherence $\mu(V)$ for different maximum normalized delay N_1 and number of transmit antennas M_t .

In Fig. 5.7 (a), we plot the coherence $\mu(V)$ versus N_1 for $K = 4$ targets located at $[-10^\circ, -5^\circ, 0^\circ, 1^\circ]$ for $M_t = 40, N = 64$. The maximum range $d_{\max}(0)$ is set so that N_1 takes values from 0 to $N - 1$. The simulation results are averaged over 100 independent implementations, and in each implementation the ranges of the middle two targets are chosen randomly in $[d_{\min}(0), d_{\max}(0)]$. It can be seen from Fig. 5.7 (a) that $\mu(V)$ increases as N_1 becomes larger. In addition, $\mu(V)$ under the optimized waveforms is the lowest as compared to Hadamard and G-Orth waveforms, which is slightly higher than $1 + N_1/N$ and is less than 2 for most N_1 values. In Fig. 5.7 (b), we plot $\mu(V)$ versus M_t for $K = 4$ targets located at $[-10^\circ, -5^\circ, 0^\circ, 1^\circ]$, where $N_1 = N/2$. The results are averaged over 100 independent implementations, and in each implementation the two middle target ranges are chosen randomly in $[d_{\min}(0), d_{\max}(0)]$ and G-Orth waveforms are randomly generated. It can be seen from Fig. 5.7 (b) that the averaged $\mu(V)$ under the optimized waveforms decreases slightly as M_t increases and is bounded by (5.63).

Our simulation results show that when targets are not all in the same range bin, $\mu(V)$, computed under the optimized waveforms, is close to $1 + N_1/N$ for a wide range of M_t, N_1 values.

5.6.4 Matrix Recovery Error Performance

Here, we look at the MC performance as function of the portion of observed entries, p , corresponding to the optimized waveform matrix for $K = 4$ targets located $[-1^\circ, 0^\circ, 0.1^\circ, 0.5^\circ]$, respectively. For each configuration, both $M_r = 20$ and $M_r = 40$ antennas are applied, respectively. The targets are moving slowly with speeds $[1, 5, 10, 15]$ m/s. The signal-to-noise ratio (SNR), defined as the power of all receive signals at the receiver end over the power of noise, is set to 25dB. The simulation results are averaged over 50 independent runs, where in each run, the noise is randomly generated. The optimized waveform matrices with $M_t = 20, N = 64$ are obtained via Algorithm 1, focusing on DOA space $[-5^\circ : 0.1^\circ : 5^\circ]$. In the simulations, the data matrix is recovered via the SVT algorithm of [71].

Targets in the Same Range Bin

We take all targets to fall in the same range bin. Fig. 5.8 (a) shows the recovery error, suggesting that the optimized waveform matrix results in significantly better performance as compared to the column-wise Hadamard matrix, especially for small values of p . One can see that in order to achieve an error around 5%, MC with the optimized waveforms requires about 50% of the data matrix entries for $M_r = 20$, and 30% for $M_r = 40$. On the other hand, MC with a column-wise Hadamard matrix requires more than 60% of the data matrix entries for both $M_r = 20$ and $M_r = 40$. In the same figure, we also compare the optimized waveforms against column-wise G-Orth waveforms. One can see that the former result in lower MC recover error for smaller p 's, while their advantage diminishes for higher p 's. Based on our experience with simulations, the range of p over which the optimized waveforms have an advantage over the G-Orth waveforms shrinks as M_t increases. This observation suggests that for large number of transmit antennas, the G-Orth waveforms behave like optimal in the sense

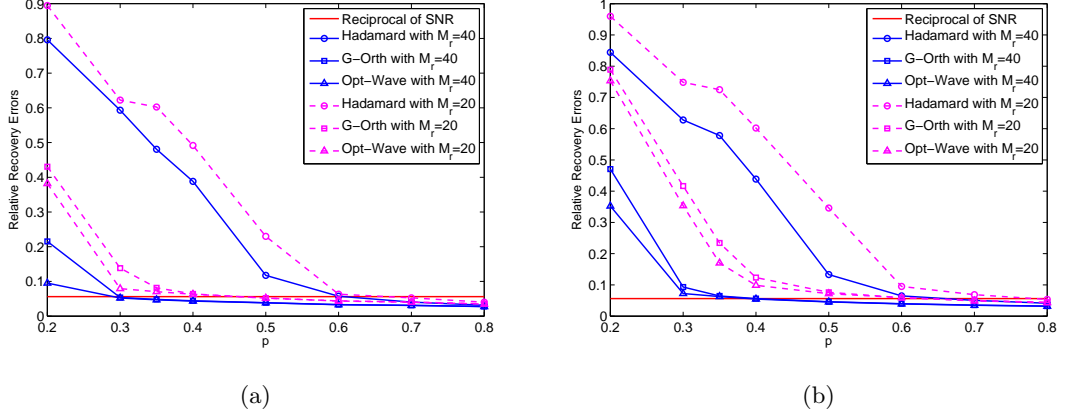


Figure 5.8: MC error vs. p for targets located at $[-1^\circ, 0^\circ, 0.1^\circ, 0.5^\circ]$. (a) all targets fall in the same range bin; (b) the ranges are $[2020, 3820, 3820, 6220]$ m.

that they achieve a comparable MC performance as the optimized waveforms. Due to the substantially lower computational cost involved, in such cases G-Orth waveforms would be preferable in a practical scenario.

Although the waveform design requires angle space discretization, the sensitivity due to targets falling off grids is rather low. Also, although the optimized waveforms are not unique, all solutions result in almost identical MC performance (see simulations in [95]).

Targets in Different Range Bins

Let us continue on the configuration in Fig. 5.8 (a) but set the $K = 4$ target ranges to $[2020, 3820, 3820, 220]$ m, corresponding to maximal normalized delay $N_1 = 35$. As shown in Fig. 5.8 (b), the recovery error under the optimized waveform matrix has a similar performance trend as the scenario considered in Fig. 5.8 (a), i.e., the optimized waveform matrix results in significant performance improvement as compared to the Hadamard matrix. As compared to Fig. 5.8 (a), a slightly larger portion of samples is required for MC under optimized waveforms to achieve a recovery error less than the inverse of SNR, i.e., the ratio of noise power over the targets' power. This is because the coherence in the scenario of targets falling in different range bins is slightly higher, as indicated in (5.63).

In both Fig. 5.8 (a) and (b), one can see that smaller p 's are required to achieve an error around 5% for all waveforms by increasing M_r .

5.6.5 Target Estimation Performance via MC

Continuing on the scenario in Fig. 5.8 (b), we look at the target estimation performance after the data matrix is recovered via MC. Since the targets fall in different range bins, the range compression (pulse compression) is first applied to the recovered data matrix. Then, the DOA and speed estimation follow using the subspace methods, such as MUSIC [5]. The details of DOA estimation using MUSIC method are addressed in equation (17) of [76].

Fig. 5.9 shows the Angle-Range image of the target scene of Fig. 5.8 (b), using the MUSIC method following MC recovery. Total $Q = 40$ pulses are transmitted and $p = 0.35$ portion samples are collected from $M_r = 20$ antennas. One can see that the performance of optimized and G-Orth waveforms is better than that of Hadamard waveform. False alarms at angles -1° and 0.5° in the wrong range bins are triggered under the Hadamard waveform; this is due to the waveform poor correlation properties. Furthermore, the two middle targets at range 3820m, are unresolvable under the Hadamard waveform. In this case, the performance under the optimized waveform is slightly better than that under G-Orth waveform in terms of magnitude, which would result in better DOA resolution. However, as p increases, the two waveforms perform comparably (see Fig. 5.8 (b)).

Next, we access the capability of the MC based method to resolve two closely located targets in the same range bin. We change the scenario in Fig. 5.9 by setting $\theta_3 = \theta_2 + \Delta\theta$ and keep the rest parameters unchanged, where θ_k denotes the angle of the k -th target. Two targets are considered to be resolved if $|\hat{\theta}_k - \theta_k| \leq \Delta\theta/2, k = 2, 3$, where $\hat{\theta}_k$ denotes the estimation of the k -th target [5]. Figure 5.10 shows the probability of resolution comparison between the optimized and G-Orth waveforms for $p = 0.35$. The results are obtained based on 50 independent runs, and in each run, the noise is randomly generated. The probability of resolution is calculated by counting the number of successful resolvable events over the total number of runs. It can be seen that the

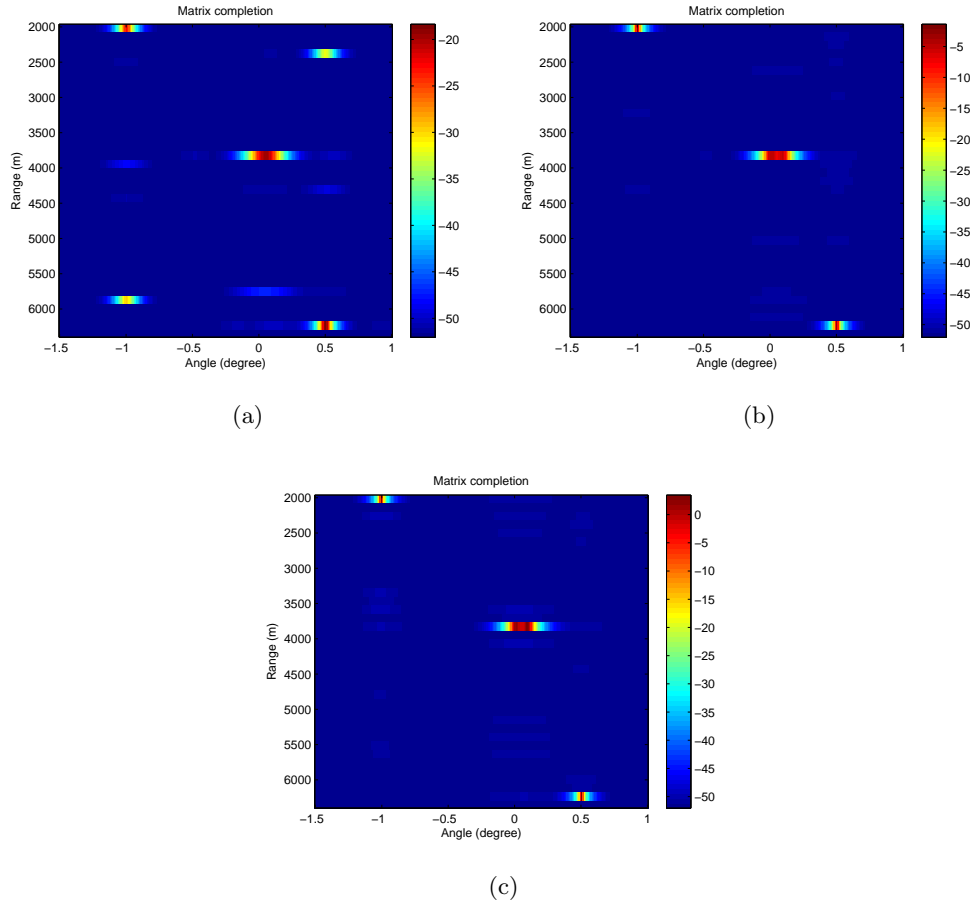


Figure 5.9: Angle-Range images using MIMO-MC radars for $K = 4$ targets located at $[-1^\circ, 0^\circ, 0.1^\circ, 0.5^\circ]$ with ranges as $[2020, 3820, 3820, 6220]$ m: (a) Hadamard; (b) G-Orth; (c) Optimized waveforms. The other parameters are $M_t = 20, N = 64, M_r = 20, Q = 40, p = 0.35$.

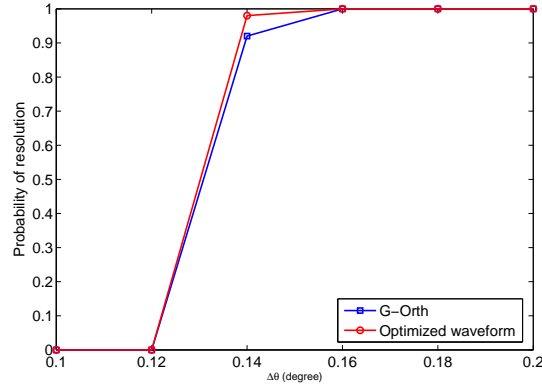


Figure 5.10: Probability of resolution comparison for optimized and G-Orth waveforms for $p = 0.35$.

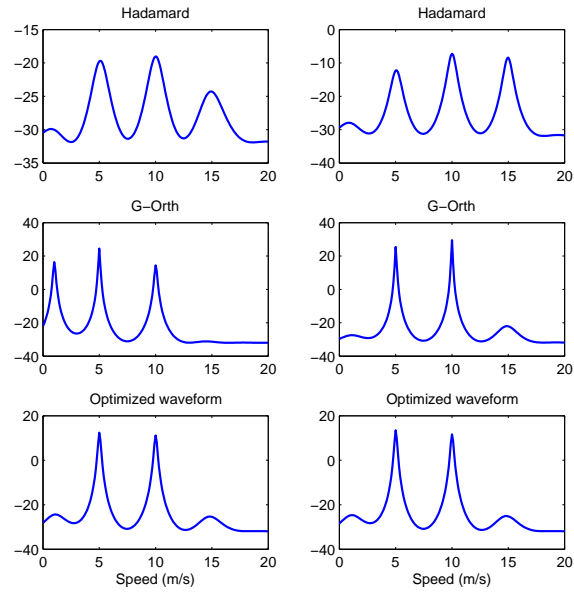


Figure 5.11: Pseudo-spectra of speed estimation using MUSIC with $Q = 40$ pulses. The left and right side of the figure correspond to $p = 0.3$ and $p = 0.5$, respectively.

optimized waveform results in a better probability of resolution as compared to the G-Orth waveform. High resolution could be achieved using MC under both optimized and G-Orth waveforms.

Figure 5.11 shows the pseudo-spectra of speed estimation using MUSIC for the scenario of Fig. 5.8 (b) for $M_r = 40$ and $Q = 40$. Range compression corresponding to range 3820m is applied first, and MUSIC is subsequently used to extract the speed information of the middle two targets, i.e, $\vartheta_2 = 5\text{m/s}$ and $\vartheta_3 = 10\text{m/s}$, where ϑ_k denotes the speed of the k -th target. It can be seen that the Hadamard waveform with both $p = 0.3$ and $p = 0.5$ yields the worst performance, exhibiting a false peak corresponding to speed of 15m/s at range 6220m. Under the G-Orth waveform, a false peak is found at $p = 0.3$, corresponding to speed of 1m/s at range 2020m. The optimized waveform results in the best performance at $p = 0.3$. The performance under the optimized and G-Orth waveforms is comparable at $p = 0.5$, in which case the recovery error under both waveforms drops below the noise level (see Fig. Fig. 5.8 (b)).

5.7 Summary

In this chapter, we have presented an analysis of the coherence of the data matrix arising in MIMO-MC radar with ULA configurations and transmitting orthogonal waveforms. We have shown that, the data matrix attains its lowest possible coherence if the waveform snapshots across the transmit array have flat power spectra for all time instances. The waveform design problem has been approached as an optimization problem on the complex Stiefel manifold and has been solved via the modified steepest descent algorithm and the modified Newton algorithm. The numerical results have shown that as the number of antennas increases, the optimized waveforms result in optimal data matrix coherence, i.e., 1, and thus, only a small portion of samples are needed for the data matrix recovery. For a particular array, the optimal waveforms depend on the target space to be investigated; for different regions of the target space, the corresponding optimal waveforms can be constructed a priori. Since their construction involves high computational complexity, the optimal waveforms can be used as benchmark against

easily constructed waveforms. For example, our simulations revealed that as the number of transmit antennas increases, simply transmitting G-Orth waveforms results in comparable matrix recovery performance as transmitting optimized waveforms. Thus, given the cost of computing the optimized waveforms, certain applications and under certain conditions may treat G-Orth waveforms as optimal. Although the optimal waveforms are designed based on the assumption that the targets fall in the same range bin, our analysis and simulations showed that they cause only small amount of performance degradation for relatively small delays, i.e., of the order of the symbol interval when used in a scenario in which the targets appear in different range bins.

Chapter 6

On Transmit Beamforming in MIMO Radars with Matrix Completion

This chapter proposes a matrix completion based colocated MIMO radars (MIMO-MC) approach that employs transmit beamforming. The transmit antennas transmit correlated waveforms to illuminate certain directions. Each receive antenna performs sub-Nyquist sampling of the target returns at uniformly random times, and forwards the samples to a fusion center along with information on the sampling times. Based on the forwarded samples, the fusion center partially fills a matrix, recovers the Nyquist rate samples via matrix completion, and subsequently proceeds with target estimation via standard techniques. The performance of matrix completion depends on the matrix coherence. This chapter derives the relations between transmit waveforms and matrix coherence. Specifically, it is shown that, for a rank-1 beamformer, the coherence is optimal, i.e., 1, if and only if the waveforms are unimodular. For a multi-rank beamformer, the coherence of the row space of the data matrix is optimal if the waveform power is constant across each snapshot. Simulation results show that the proposed scheme achieves high resolution with a significantly reduced number of samples.

6.1 Introduction

A colocated MIMO radar approach based on matrix completion (MC) [41] [42] [43] (MIMO-MC radar) has been recently proposed in [76] [75] to achieve the high resolution of MIMO radars while requiring significantly fewer samples to be collected and forwarded to a fusion center. Based In MIMO-MC radars, each receive antenna obtains

samples at uniformly random times and forwards them to a fusion center, which partially fills a matrix, referred to as the data matrix. The matrix can be subsequently recovered via matrix completion techniques. As shown in [76] [77], [68] and [96], [97] the transmit/receive array configuration as well as transmit waveforms affect the matrix coherence and thus the MC performance.

Most of the work on MIMO radars assumes the transmission of uncorrelated waveforms from the transmit antennas. However, for MIMO radars operating in tracking mode, that considers correlated transmit waveforms [98] [99]. For example, in [98], the transmit waveforms correlation is designed so that a desired transmit beampattern is achieved. In [99], the authors proposed a phased-MIMO radar approach by dividing the transmit array into multiple sub-arrays, with each sub-array coherently transmitting a waveform which is orthogonal to waveforms transmitted by other sub-arrays. Thus, in each sub-array, beamforming is achieved. A multi-rank beamformer for MIMO radars has been recently proposed in [100], which, unlike [98] does not require solving a complicated optimization problem. The multi-rank beamformer is taken as the combination of rank-1 beamformers with the corresponding multiple waveforms chosen to be orthogonal.

In this chapter we consider the same MIMO radars transmit beamforming framework as in [100]. When the number of illuminated targets is much smaller than the size of receive array, the data matrix formulated by the fusion center based on Nyquist-rate samples at the receive antennas is low-rank. Therefore, Nyquist sampling is not required at each receive antenna. Instead, the antenna can uniformly at random select samples and forward them to the fusion center, thus partially filling the data matrix. By applying MC, the fusion center can recover the full matrix. Based on the recovered data matrix, various methods, e.g., MUSIC [5], can be employed for target estimation. The advantages of sending fewer samples to the fusion center include power and bandwidth savings. The focus of this chapter is to determine the suitability of MC in this scenario. For this purpose, we conduct matrix coherence analysis and derive the optimal waveform conditions for both rank-1 and multi-rank beamformers.

6.2 On the Rank-1 beamforming

Let $\mathbf{s} \in \mathbb{C}^{N \times 1}$ be the waveform sequence transmitted by each of the M_t antennas over one pulse. Let $\mathbf{w} \in \mathbb{C}^{M_t \times 1}$ denote the transmit beamformer. Then, according to [73], the rank-1 beamformer equals $\mathbf{w} = \mathbf{a}(\theta)/\|\mathbf{a}(\theta)\|$, where $\mathbf{a}(\theta)$ denotes the transmit steering vector corresponding to direction θ . High angle resolution can be achieved under the rank-1 beamformer [100] by doing joint transmit and receive beamforming, which shows great advantage of MIMO radars over phased-array radar for single target tracking.

Under the narrow-band assumption, the noise free receive data matrix collected at the fusion center containing the samples of target reflections equals $\mathbf{X} = \mathbf{b}(\theta) \beta \zeta \mathbf{a}(\theta)^T \tilde{\mathbf{S}}$, where $\mathbf{b}(\theta)$ is the receive steering vector w.r.t. direction θ . The transmit signal matrix equals $\tilde{\mathbf{S}} = \mathbf{w} \mathbf{s}^T$. In addition, β and ζ are the target reflection coefficient and Doppler shift, respectively.

The matrix \mathbf{X} is low rank, and as long as its left and right subspaces coherence is low, it can be recovered from a small number of its entries, selected uniformly at random.

In the following theorem we state the conditions so that the coherence of \mathbf{X} achieves its smallest possible value of 1.

Theorem 4. *Under a ULA configuration, when the MIMO radar antennas transmit the same waveform and a rank-1 beamformer is used, i.e., $\mathbf{w} = \mathbf{a}(\theta)/\|\mathbf{a}(\theta)\|$, the coherence of the matrix \mathbf{X} achieves its lowest value, i.e., $\mu(U) = \mu(V) \equiv 1$ if and only if the waveform sequence is unimodular.*

Proof. The data matrix $\mathbf{X} \in \mathbb{C}^{M_r \times N}$ as defined above is rank-1. Let its compact singular value decomposition (SVD) be $\mathbf{X} = \mathbf{u} \sigma \mathbf{v}^H$, where $\mathbf{u} \in \mathbb{C}^{M_r \times 1}$, $\mathbf{v} \in \mathbb{C}^{N \times 1}$ with $\mathbf{u}^H \mathbf{u} = 1, \mathbf{v}^H \mathbf{v} = 1$ and σ the corresponding singular value. Consider the QR decomposition of $\mathbf{b}(\theta)$ given by $\mathbf{b}(\theta) = \mathbf{q}_r r_r$, where

$$\mathbf{q}_r = \frac{1}{\sqrt{M_r}} \begin{bmatrix} 1 & e^{j \frac{2\pi}{\lambda} d_r \sin \theta} & \dots & e^{j \frac{2\pi}{\lambda} (M_r-1) d_r \sin \theta} \end{bmatrix}^T$$

such that $\mathbf{q}_r^H \mathbf{q}_r = 1$ and $r_r = \sqrt{M_r}$ where λ denotes the wavelength. Similarly, we consider the QR decomposition of $\tilde{\mathbf{S}}^T \mathbf{a}(\theta)$ given by $\tilde{\mathbf{S}}^T \mathbf{a}(\theta) = \frac{\mathbf{s} \mathbf{a}(\theta)^T \mathbf{a}(\theta)}{\|\mathbf{a}(\theta)\|} = \mathbf{q}_s r_s$, where $\mathbf{q}_s \in \mathbb{C}^{N \times 1}$ such that $\mathbf{q}_s^H \mathbf{q}_s = 1$, and r_s is a real number. Then, $\mathbf{X} = \mathbf{q}_r r_r \beta \zeta r_s \mathbf{q}_s^T$. The SVD of the complex number $r_r \beta \zeta r_s$ can be written as $r_r \beta \zeta r_s = q_1 \rho q_2^*$, where $|q_1| = |q_2| = 1$ and ρ is a real number. Therefore, $\mathbf{X} = \mathbf{q}_r q_1 \rho q_2^* \mathbf{q}_s^T = \mathbf{q}_r q_1 \rho (\mathbf{q}_s^* q_2)^H$, which is a valid SVD of \mathbf{X} since $(\mathbf{q}_r q_1)^H \mathbf{q}_r q_1 = 1$, $(\mathbf{q}_s^* q_2)^H \mathbf{q}_s^* q_2 = 1$. By the uniqueness of singular values of a matrix, it holds that $\sigma \equiv \rho$. Therefore, we can set $\mathbf{u} = \mathbf{q}_r q_1$, $\mathbf{v} = \mathbf{q}_s^* q_2$.

Let $\mathbf{q}_r^{(i)}$ denote the i -th element of \mathbf{q}_r . The coherence of the column space of \mathbf{X} is

$$\mu(U) = \frac{M_r}{1} \sup_{i \in \mathbb{N}_{M_r}^+} \left| \mathbf{q}_r^{(i)} q_1 \right|^2 = M_r \sup_{i \in \mathbb{N}_{M_r}^+} \left| \mathbf{q}_r^{(i)} \right|^2 \equiv 1. \quad (6.1)$$

Let $\mathbf{q}_s^{*(i)}$, s_i denote the i -th element of \mathbf{q}_s^* and \mathbf{s} , respectively. The coherence of the row space of \mathbf{X} is

$$\begin{aligned} \mu(V) &= \frac{N}{1} \sup_{i \in \mathbb{N}_N^+} \left| \mathbf{q}_s^{*(i)} q_2 \right|^2 = N \sup_{i \in \mathbb{N}_N^+} \left| \mathbf{q}_s^{(i)} \right|^2 \\ &= N \sup_{i \in \mathbb{N}_N^+} \left| \frac{s_i \mathbf{a}(\theta)^T \mathbf{a}(\theta)}{\|\mathbf{a}(\theta)\| r_s} \right|^2 \\ &= N \sup_{i \in \mathbb{N}_N^+} \frac{\mathbf{a}(\theta)^H \mathbf{a}(\theta)^* |s_i|^2 \mathbf{a}(\theta)^T \mathbf{a}(\theta)}{\|\mathbf{a}(\theta)\|^2 r_s^2}. \end{aligned} \quad (6.2)$$

Here, it holds that

$$\begin{aligned} r_s^2 &= r_s \mathbf{q}_s^H \mathbf{q}_s r_s = (\mathbf{q}_s r_s)^H \mathbf{q}_s r_s \\ &= \frac{\mathbf{a}(\theta)^H \mathbf{a}(\theta)^* \mathbf{s}^H \mathbf{s} \mathbf{a}(\theta)^T \mathbf{a}(\theta)}{\|\mathbf{a}(\theta)\|^2}. \end{aligned} \quad (6.3)$$

Consequently,

$$\mu(V) = N \sup_{i \in \mathbb{N}_N^+} \frac{|s_i|^2}{\mathbf{s}^H \mathbf{s}}. \quad (6.4)$$

Since $\sum_{i=1}^N |s_i|^2 = \mathbf{s}^H \mathbf{s}$ and $|s_i|^2 \geq 0$, the minimum possible value of $\mu(V)$ could achieve the minimum value, i.e., 1, if and only if $|s_i|^2 = \frac{1}{N} \mathbf{s}^H \mathbf{s}$ for any $i \in \mathbb{N}_N^+$. This condition suggests that the transmit power in each snapshot, i.e., $|s_i|^2$, should equal the total transmit power $\mathbf{s}^H \mathbf{s}$ divided by N . In other words, the transmit waveform should be unimodular. Consequently, it holds that $\mu(U) = \mu(V) \equiv 1$, which completes the proof. \square

It is interesting to note that the coherence is optimal independent of the the beamforming vector.

6.3 On the Multi-Rank Beamforming

According to [100], to track multiple targets at directions $\{\theta_k\}_{k \in \mathbb{N}_K^+}$, a rank- K beamformer $\mathbf{W} = [\mathbf{w}_1 \cdots \mathbf{w}_K] \in \mathbb{C}^{M_t \times K}$ can be used, where $\mathbf{w}_k = \frac{\mathbf{a}(\theta_k)}{\|\mathbf{a}(\theta_k)\|}$ is the beamformer focussing on direction θ_k . The sampled transmitted signal matrix equals $\tilde{\mathbf{S}} = \sqrt{\frac{M_t}{K}} \mathbf{W} \mathbf{S}^T$, where $\sqrt{\frac{M_t}{K}}$ is a factor to satisfy that the total transmit energy is M_t ; $\mathbf{S} \in \mathbb{C}^{N \times K}$ contains sampled orthogonal waveforms so that $\mathbf{S}^H \mathbf{S} = \mathbf{I}_K$. The transmit beampattern in direction ϕ is the sum of K rank-1 beampatterns, i.e.,

$$\begin{aligned} P_T(\phi) &= \frac{M_t}{K} \mathbf{a}(\phi)^H \mathbf{W} \mathbf{W}^H \mathbf{a}(\phi) \\ &= \frac{M_t}{K} \sum_{k=1}^K \mathbf{a}(\phi)^H \mathbf{w}_k \mathbf{w}_k^H \mathbf{a}(\phi). \end{aligned} \quad (6.5)$$

Under the narrow-band assumption, the noise free receive data matrix is [76]

$$\mathbf{X} = \mathbf{B} \mathbf{D} \mathbf{A}^T \tilde{\mathbf{S}}, \quad (6.6)$$

where $\mathbf{A} \in \mathbb{C}^{M_t \times K}$ is the transmit steering matrix defined as $\mathbf{A} = [\mathbf{a}(\theta_1) \cdots \mathbf{a}(\theta_K)]$; $\mathbf{B} \in \mathbb{C}^{M_r \times K}$ is the receive steering matrix, and $\mathbf{D} \in \mathbb{C}^{K \times K}$ is a diagonal matrix containing target reflection coefficients and Doppler shifts. It can be shown that \mathbf{X} is a low-rank matrix. Thus, depending on how low its coherence is, it can be recovered based on a small, uniformly sampled subset of its elements. On the coherence of \mathbf{X} , we have the following theorem.

Theorem 5. *Consider an ULA configuration and a MIMO radar applying the rank- K beamformer $\mathbf{W} = [\mathbf{w}_1 \cdots \mathbf{w}_K] \in \mathbb{C}^{M_t \times K}$ to K orthogonal waveforms $\mathbf{S} \in \mathbb{C}^{N \times K}$.*

The coherence of the row space of \mathbf{X} is optimal, i.e., $\mu(V) \equiv 1$, if and only if

$$\mathbf{S}^{(i)} \left(\mathbf{S}^{(i)} \right)^H = \frac{K}{N}, \quad \forall i \in \mathbb{N}_N^+, \quad (6.7)$$

where $\mathbf{S}^{(i)} \in \mathbb{C}^{1 \times K}$ denotes the i -th row of \mathbf{S} .

Proof. The compact SVD of \mathbf{X} can be expressed as $\mathbf{X} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H$, where $\mathbf{U} \in \mathbb{C}^{M_r \times K}$, $\mathbf{V} \in \mathbb{C}^{N \times K}$ such that $\mathbf{U}\mathbf{U}^H = \mathbf{I}_K$, $\mathbf{V}\mathbf{V}^H = \mathbf{I}_K$, and $\mathbf{\Lambda} \in \mathbb{R}^{K \times K}$ is a diagonal matrix containing the singular values of \mathbf{X} . Consider the QR decomposition of \mathbf{B} , i.e., $\mathbf{B} = \mathbf{Q}_r \mathbf{R}_r$, where $\mathbf{Q}_r \in \mathbb{C}^{M_r \times K}$ such that $\mathbf{Q}_r^H \mathbf{Q}_r = \mathbf{I}_K$, and $\mathbf{R}_r \in \mathbb{C}^{K \times K}$ is an upper triangular matrix. Similarly, we consider the QR decomposition of $\tilde{\mathbf{S}}^T \mathbf{A}$ given by $\tilde{\mathbf{S}}^T \mathbf{A} = \mathbf{Q}_s \mathbf{R}_s$, where $\mathbf{Q}_s \in \mathbb{C}^{N \times K}$ such that $\mathbf{Q}_s^H \mathbf{Q}_s = \mathbf{I}_K$ and $\mathbf{R}_s \in \mathbb{C}^{K \times K}$ is an upper triangular matrix. Then, $\mathbf{X} = \mathbf{Q}_r \mathbf{R}_r \mathbf{D} \mathbf{R}_s^T \mathbf{Q}_s^T$ and the matrix $\mathbf{R}_r \mathbf{D} \mathbf{R}_s^T \in \mathbb{C}^{K \times K}$ is rank- K whose SVD is given as $\mathbf{R}_r \mathbf{D} \mathbf{R}_s^T = \mathbf{Q}_1 \mathbf{\Delta} \mathbf{Q}_2^H$. Here, $\mathbf{Q}_1 \in \mathbb{C}^{K \times K}$ is such that $\mathbf{Q}_1 \mathbf{Q}_1^H = \mathbf{Q}_1^H \mathbf{Q}_1 = \mathbf{I}_K$ (the same holds for \mathbf{Q}_2) and $\mathbf{\Delta} \in \mathbb{R}^{K \times K}$ is non-zero diagonal, containing the singular values of $\mathbf{R}_r \mathbf{D} \mathbf{R}_s^T$. Therefore,

$$\mathbf{X} = \mathbf{Q}_r \mathbf{Q}_1 \mathbf{\Delta} \mathbf{Q}_2^H \mathbf{Q}_s^T = \mathbf{Q}_r \mathbf{Q}_1 \mathbf{\Delta} (\mathbf{Q}_s^* \mathbf{Q}_2)^H, \quad (6.8)$$

which is a valid SVD of \mathbf{X} since $(\mathbf{Q}_r \mathbf{Q}_1)^H \mathbf{Q}_r \mathbf{Q}_1 = \mathbf{I}_K$ and $(\mathbf{Q}_s^* \mathbf{Q}_2)^H \mathbf{Q}_s^* \mathbf{Q}_2 = \mathbf{I}_K$. By the uniqueness of the singular values of a matrix, it holds that $\mathbf{\Lambda} \equiv \mathbf{\Delta}$. Therefore, we can set $\mathbf{U} = \mathbf{Q}_r \mathbf{Q}_1$ and $\mathbf{V} = \mathbf{Q}_s^* \mathbf{Q}_2$.

Let $\mathbf{Q}_s^{*(i)}, \mathbf{S}^{(i)} \in \mathbb{C}^{1 \times K}$ denote the i -th row of \mathbf{Q}_s^* and \mathbf{S} , respectively. Regarding the coherence of the row space of \mathbf{X} , we have

$$\begin{aligned} \mu(V) &= \frac{N}{K} \sup_{i \in \mathbb{N}_N^+} \left\| \mathbf{Q}_s^{*(i)} \mathbf{Q}_2 \right\|^2 = \frac{N}{K} \sup_{i \in \mathbb{N}_N^+} \left\| \mathbf{Q}_s^{(i)} \right\|^2 \\ &= \frac{N}{K} \sup_{i \in \mathbb{N}_N^+} \left\| \sqrt{\frac{M_t}{K}} \mathbf{S}^{(i)} \mathbf{W}^T \mathbf{A} \mathbf{R}_s^{-1} \right\|^2 \\ &= \frac{N}{K} \sup_{i \in \mathbb{N}_N^+} \frac{M_t}{K} \mathbf{S}^{(i)} \mathbf{W}^T \mathbf{A} \mathbf{R}_s^{-1} (\mathbf{R}_s^{-1})^H \mathbf{A}^H \mathbf{W}^* (\mathbf{S}^{(i)})^H. \end{aligned}$$

Here, since the waveforms are orthogonal, i.e., $\mathbf{S}^H \mathbf{S} = \mathbf{I}_K$, it holds that

$$\begin{aligned} \mathbf{R}_s^{-1} (\mathbf{R}_s^{-1})^H &= (\mathbf{R}_s^H \mathbf{R}_s)^{-1} = (\mathbf{R}_s^H \mathbf{Q}_s^H \mathbf{Q}_s \mathbf{R}_s)^{-1} \\ &= \frac{K}{M_t} (\mathbf{A}^H \mathbf{W}^* \mathbf{S}^H \mathbf{S} \mathbf{W}^T \mathbf{A})^{-1} \\ &= \frac{K}{M_t} (\mathbf{A}^H \mathbf{W}^* \mathbf{W}^T \mathbf{A})^{-1} \\ &= \frac{K}{M_t} (\mathbf{W}^T \mathbf{A})^{-1} (\mathbf{A}^H \mathbf{W}^*)^{-1}. \end{aligned} \quad (6.9)$$

Consequently, $\mu(V) = \frac{N}{K} \sup_{i \in \mathbb{N}_N^+} \mathbf{S}^{(i)} (\mathbf{S}^{(i)})^H$. In addition, it holds that $\sum_{i=1}^N \mathbf{S}^{(i)} (\mathbf{S}^{(i)})^H = \mathbf{I}_K$ for orthogonal waveforms. Therefore, to find the lowest possible value of $\mu(V)$, we

solve the following optimization problem

$$\begin{aligned} \min_{i \in \mathbb{N}_N^+} & \left(\max_{i \in \mathbb{N}_N^+} \mathbf{S}^{(i)} (\mathbf{S}^{(i)})^H \right) \\ \text{s.t.} & \sum_{i=1}^N \mathbf{S}^{(i)} (\mathbf{S}^{(i)})^H = K. \end{aligned} \quad (6.10)$$

Since $\mathbf{S}^{(i)} (\mathbf{S}^{(i)})^H \geq 0$, the solutions of the above problem are $\mathbf{S}^{(i)} (\mathbf{S}^{(i)})^H = \frac{K}{N}, \forall i \in \mathbb{N}_N^+$. Consequently, the lowest possible value of row space of \mathbf{X} is achieved as $\mu(V) \equiv 1$, which completes the proof. \square

Theorem 5 indicates that the energy of the K orthogonal waveforms should be constant during each snapshot. It is interesting to note that the coherence of the row space of \mathbf{X} is independent of the multi-rank beamformer. Therefore, the analysis results hold for all kinds of multi-rank beamformers obtained via different methods, e.g., multi-rank beamformer for the approximation of a desired beampattern, proposed in [101]. It should be noted that the coherence of the column space of \mathbf{X} , i.e., $\mu(U)$ coincides with the results in MIMO-MC radar and interested readers can refer to [68] for detail discussions.

6.4 DOA Estimation based on Matrix Completion

At the fusion center, for each pulse, the data matrix \mathbf{X} is recovered via MC using a small portion of samples collected uniformly at random. Let $\tilde{\mathbf{X}}$ denote the recovered data matrix. Subsequently, matched filtering is applied on $\tilde{\mathbf{X}}$ to obtain

$$\mathbf{Y}_q = \tilde{\mathbf{X}} \mathbf{S}^* = \sqrt{\frac{M_t}{K}} \mathbf{B} \mathbf{D} \mathbf{A}^T \mathbf{W} + \mathbf{Z}_q, \quad (6.11)$$

where q is the pulse index and \mathbf{Z}_q represents noise. Stacking the matrix (6.11) into a $K M_r \times 1$ vector, we get

$$\mathbf{y}_q = \text{vec}(\mathbf{Y}_q) = \sqrt{\frac{M_t}{K}} \sum_{k=1}^K d_k (\mathbf{W}^T \mathbf{a}(\theta_k)) \otimes \mathbf{b}(\theta_k) + \mathbf{z}_q,$$

where d_k denotes the reflection coefficient and Doppler shift w.r.t. the k -th target; $\mathbf{z}_q = \text{vec}(\mathbf{Z}_q)$. With Q pulses data, the sample covariance matrix is obtained as $\mathbf{R} =$

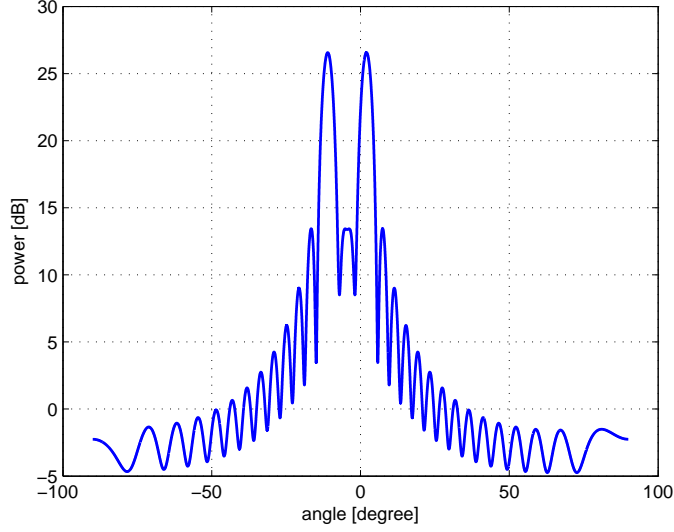


Figure 6.1: Transmit beampattern under rank-2 beamformer with $M_t = 30$ for directions $[-11^\circ, 2^\circ]$;

$\frac{1}{Q} \sum_{q=1}^Q \mathbf{y}_q \mathbf{y}_q^H$. The pseudo-spectrum of MUSIC estimator is [5]

$$P(\theta) = \frac{1}{\mathbf{c}^H(\theta) \mathbf{E}_n \mathbf{E}_n^H \mathbf{c}(\theta)}, \quad (6.12)$$

where $\mathbf{c}(\theta) = (\mathbf{W}^T \mathbf{a}(\theta)) \otimes \mathbf{b}(\theta)$ and $\mathbf{E}_n \in \mathbb{C}^{KM_r \times (KM_r - K)}$ is a matrix containing the eigenvectors of the noise subspace of \mathbf{R} . The angle of the targets can be obtained by the finding the peak locations of the pseudo-spectrum (6.12).

6.5 Simulations

In the following, we use simulation to test the performance of the proposed approach. Throughout the simulations, the transmit/receive arrays are configured as ULA with $d_t = \lambda/2$ and $d_r = M_t \lambda/2$. The first K waveforms of the Hadamard sequences are used for transmit beamforming. The number of pulses is set to $Q = 3$. The number of Nyquist samples in one pulse is $N = 128$. The data matrix \mathbf{X} is recovered via the SVT algorithm [71] using only $p = 50\%$ of its entries. The obtained results are averaged over 100 independent runs. First, a rank-2 beamformer with $M_t = 30$ is applied to illuminate $K = 2$ targets at angles $\theta_1 = -11^\circ$ and $\theta_2 = 2^\circ$. The transmit beampattern

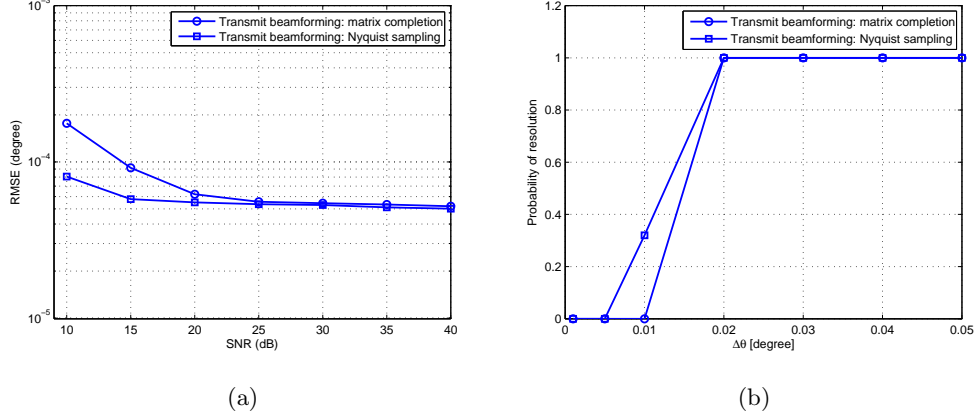


Figure 6.2: Simulations: (a) RMSE versus SNR with $M_r = 60, M_t = 30$; (b) Probability of target resolution with $M_r = 60, M_t = 20$ and SNR = 25dB. The proposed approach is based on subsampling by 50%.

of the beamformer is shown in Fig. 6.1. The root mean square error (RMSE) of the direction of arrival (DOA) estimation for these two targets is plotted in Fig. 6.2 (a) for $M_r = 60$. It can be found that the RMSE of DOA estimation using MC or Nyquist sampling decreases as the signal-to-noise ratio (SNR) becomes larger (the SNR is defined at the fusion center before the matched filtering operation). Interestingly, as $\text{SNR} \geq 25\text{dB}$, these two RMSE curves become almost identical. This is because the recovery error of \mathbf{X} introduced by MC is quite small when the SNR is high [43]. Next, we access the capability of the proposed scheme to resolve two closely located targets. We take the first target to be in direction $\theta_1 = 10^\circ$ and the second in $\theta_2 = \theta_1 + \Delta\theta$. The targets are considered to be resolved if $|\hat{\theta}_k - \theta_k| \leq \Delta\theta/2, k = 1, 2$, where $\hat{\theta}_k$ denotes the estimation of the k -th target [5]. The probability of resolution under $M_r = 60, M_t = 20$ and SNR = 25dB is plotted in Fig. 6.2 (b). It can be found that the proposed MC based scheme has the same resolution of $\Delta\theta = 0.02^\circ$ as the method with Nyquist sampling. Therefore, a comparable DOA estimation performance is achieved under the proposed scheme as the method that uses Nyquist sampling.

6.6 Summary

In this chapter, we have proposed a MIMO-MC radar approach with transmit beamforming. Each receive antenna performs sub-Nyquist sampling and the full data matrix is recovered at the fusion center via MC. Analysis results have shown that the matrix coherence is independent of the beamformer. The row space coherence of data matrix is optimal if and only if the transmit orthogonal waveforms have constant power during all snapshots. The simulation results show that the proposed scheme could achieve super resolution at a low sub-Nyquist sampling rate.

Chapter 7

On Receive Beamforming via Matrix Completion

Beamforming methods rely on training data to estimate the covariance matrix of the interference pulse noise. Their convergence slows down if the signal of interest is present in the training data, thus requiring a large numbers of training snapshots to maintain good performance. In a distributed array, in which the array nodes are connected to a fusion center via a wireless link, the estimation of the covariance matrix would require the communication of large amounts of data, and thus would consume significant power. We propose an approach that enables good beamforming performance while requiring substantially fewer data to be transmitted to the fusion center. The main idea is based on the fact that when the number of signal and interference sources is much smaller than the number of array sensors, the training data matrix is low rank. Thus, based on matrix completion theory, under certain conditions, the training data matrix can be recovered from a subset of its elements, i.e., based on sub-Nyquist samples of the array sensors. Following the recovery of the training data matrix, and to cope with the errors introduced during the matrix completion process, we propose a robust optimization approach, which obtains the beamforming weight vector by optimizing the worst-case performance. Numerical results show that combination of matrix completion and robust optimization is very successful in suppressing interference and achieving a near-optimal beamforming performance with only partial training data.

7.1 Introduction

Adaptive beamforming has been widely used in wireless communications, radar and sonar for signal estimation. The adaptive beamforming method relies on the covariance

matrix of the interference-plus noise, which needs to be estimated based on training data, prior to applying the method for signal estimation. However, in a practical setting, such as passive source localization applications, the training data always contain the signal of interest. In that case, the convergence rates of the adaptive beamforming algorithm are significantly reduced, and can be improved only by considering very long training data [102]. If those data were collected by distributed nodes and they need to be forwarded to a fusion center for the computation of the covariance matrix, a lot of communication power would be required.

When the number of signals and interference sources is much smaller than the number of sensors in the array, the training data matrix is low rank. This means that, under certain conditions, even if some entries of the training data matrix are missing, the full matrix can be recovered via matrix completion techniques [41] [43]. Based on the above observation, we propose a scheme that significantly reduces the number of training data needed for estimating the sample covariance matrix. The idea is that during the training phase, each sensor carries out a uniformly random sub-Nyquist sampling, and forwards the samples to a fusion center. The full training data matrix can then be recovered using matrix completion.

The matrix completion step introduces errors when noisy observation is considered. Assuming that the training data matrix satisfies the *restricted isometry property* [36] [103], the relative matrix recovery error is bounded by a number of the order of the observed inverse signal-and-interference-to-noise ratio. Based on that bound, we propose a robust adaptive beamforming method with robust optimization [104], which obtains the beamforming weight vector by optimizing the worst-case performance. Numerical results show that combination of matrix completion and adaptive beamforming is very successful in suppressing interference and achieving a near-optimal signal-to-interference-plus-noise ratio (SINR) output with only partial training data.

7.2 Fundamentals of Classical Beamforming

In array signal processing, it is often desired to estimate the signal $s(k)$ under the presence of interference and noise with the help of an array of M sensors. The array observations $\mathbf{x}(k) \in \mathbb{C}^{M \times 1}$ can be written as [105]

$$\mathbf{x}(k) = s(k) \mathbf{a} + \mathbf{i}(k) + \mathbf{n}(k), \quad (7.1)$$

where $\mathbf{i}(k)$, $\mathbf{n}(k)$ are the interference and noise, respectively. Here, k is the time index and \mathbf{a} is the signal steering vector. The output of a narrow-band beamformer is

$$\mathbf{y}(k) = \mathbf{w}^H \mathbf{x}(k), \quad (7.2)$$

where $\mathbf{w} \in \mathbb{C}^{M \times 1}$ is the complex vector containing the beamforming weights. The SINR is written as [105]

$$\text{SINR} = \frac{\sigma_s^2 |\mathbf{w}^H \mathbf{a}|^2}{\mathbf{w}^H \mathbf{R}_{\mathbf{i}+\mathbf{n}} \mathbf{w}}, \quad (7.3)$$

where

$$\mathbf{R}_{\mathbf{i}+\mathbf{n}} = E \left\{ [\mathbf{i}(k) + \mathbf{n}(k)] [\mathbf{i}(k) + \mathbf{n}(k)]^H \right\} \quad (7.4)$$

is the interference-plus-noise covariance matrix and σ_s^2 is the signal power. In the minimum variance distortionless response (MVDR) beamformer, the weight vector is obtained by minimizing the output interference-plus-noise power while keeping the signal from a desirable direction distortionless [105]. This leads to the solution

$$\mathbf{w}_{\text{opt}} = \frac{\mathbf{R}_{\mathbf{i}+\mathbf{n}}^{-1} \mathbf{a}}{\mathbf{a}^H \mathbf{R}_{\mathbf{i}+\mathbf{n}}^{-1} \mathbf{a}}. \quad (7.5)$$

and optimal SINR

$$\text{SINR}_{\text{opt}} = \sigma_s^2 \mathbf{a}^H \mathbf{R}_{\mathbf{i}+\mathbf{n}}^{-1} \mathbf{a}. \quad (7.6)$$

In practice, the exact interference-plus-noise covariance matrix $\mathbf{R}_{\mathbf{i}+\mathbf{n}}$ is not available. Instead, the sample covariance matrix is used, i.e.,

$$\hat{\mathbf{R}} = \frac{1}{L} \mathbf{X} \mathbf{X}^H \quad (7.7)$$

where $\mathbf{X} = [\mathbf{x}(1), \dots, \mathbf{x}(L)] \in \mathbb{C}^{M \times L}$ and L is the number of snapshots. Thus, the optimization problem of MVDR beamformer and the corresponding solution can be rewritten by replacing \mathbf{R} with $\hat{\mathbf{R}}$ which is known as Capon method [106], and the solution, referred to as the sample matrix inversion (SMI) beamformer, is

$$\mathbf{w}_{\text{SMI}} = \frac{\hat{\mathbf{R}}^{-1} \mathbf{a}}{\mathbf{a}^H \hat{\mathbf{R}}^{-1} \mathbf{a}}. \quad (7.8)$$

When there is no signal in the training samples, as L increases the SINR under the weight vector (7.8) converges very fast to the optimal value defined in (7.6). If $L \geq 2M$, the average performance losses are less than 3dB. However, when the signals are present in the training samples, the convergence rate to (7.6) is much slower. Usually, in the later case, $L \gg M$ is required [102].

7.3 Covariance Matrix Estimation with Matrix Completion

Let us assume the the number of targets and interference sources is K , and the number of sensors in the array is much larger than K , i.e., $M \gg K$. To achieve high output SINR, the Capon beamformer requires accurate estimation of the sample covariance matrix, which in turn requires a number of snapshots much larger than the number of sensors, i.e., $L \gg M$. Thus, for a large array, a large number of training data need to be collected.

The training data matrix in the L snapshots can be rewritten as

$$\mathbf{X} = \mathbf{S} + \mathbf{Z}, \quad (7.9)$$

where $\mathbf{S} = [s(1)\mathbf{a} + \mathbf{i}(1), \dots, s(L)\mathbf{a} + \mathbf{i}(L)]$ is rank K and contains both signal and interference components. Here, $\mathbf{Z} = [\mathbf{n}(1), \dots, \mathbf{n}(L)]$ is the noise matrix. Under the above assumption, matrix \mathbf{S} is low rank. When the noise level is low, matrix \mathbf{X} and the corresponding covariance matrix, $\hat{\mathbf{R}}$, are approximately low rank. Matrix completion techniques can thus be applied to recover the training data matrix \mathbf{S} based on partial observations by exploiting its low rank structure.

7.3.1 Training Data Matrix Estimation via Matrix Completion

In [107], the authors applied matrix completion techniques to estimate the sample covariance matrix $\hat{\mathbf{R}}$ in a distributed way. However, the estimation of sample covariance matrix directly based on partial observations of covariance matrix entries does not reduce the number of samples at each sensor and the amount of local message passing is also large. This can be seen as follows. In order to recover a matrix, at least one observation per row and one observation per column are required [41]. Therefore, to recover $\hat{\mathbf{R}}$ directly, each row and each column of $\hat{\mathbf{R}}$ need to have at least one non-zero element; corresponding to element (i, j) , nodes i and j in the array need to do Nyquist sampling to obtain L samples and then share those samples. Since $\hat{\mathbf{R}}$ needs to be uniformly populated, many nodes need to do Nyquist sampling and then share data.

Instead of estimating $\hat{\mathbf{R}}$ directly, in this paper, we recover the training data matrix \mathbf{S} with only partial samples from each sensor. The recovered $\hat{\mathbf{S}}$ is then used to construct the sample covariance matrix.

It is difficult to show analytically that the singular vectors of the training matrix meet the conditions **(A0)**, and **(A1)**. However, for the case of a linear uniform array and uncorrelated signal and interference, extensive simulation results indicate that the maximal element values in both left and right singular vectors of matrix \mathbf{S} is bounded by a small number with high probability as M and L are large, i.e., the conditions **(A0)**, and **(A1)** hold (see also [75]). Therefore, matrix \mathbf{S} satisfies the *incoherence properties*.

To recover the data matrix \mathbf{S} in (7.9) with matrix completion, during the L snapshots, each sensor carries out uniformly random sampling and then forwardd the sub-Nyquist samples to a fusion center. At the fusion center, the observation can be written in the vector form

$$\mathbf{b} = \mathcal{A}(\mathbf{S}) + \mathbf{z}, \quad (7.10)$$

where \mathcal{A} is a linear transformation mapping $M \times L$ matrices into \mathbb{R}^m , i.e., $\mathcal{A}(\mathbf{S}) = [\mathbf{S}]_{ij}, (i, j) \in \Omega$. The adjoint of \mathcal{A} is denoted as \mathcal{A}^* and $\mathcal{P}_\Omega(\mathbf{X}) = \mathcal{A}^*(\mathbf{b})$. Here, \mathbf{z} is a noise vector and its distribution has been described in (7.9).

Considering the corrupted observation case, the matrix recovery is done by solving

the following nuclear norm optimization problem with quadratic constraint

$$\min \|\mathbf{W}\|_* \quad \text{s.t.} \quad \|\mathbf{b} - \mathcal{A}(\mathbf{W})\| \leq \delta. \quad (7.11)$$

The objective in problem (7.11) can be replaced by the approximate function $\alpha\|\mathbf{W}\|_* + \frac{1}{2}\|\mathbf{W}\|_F$, where $\frac{1}{2}\|\mathbf{W}\|_F$ is a smooth part and the parameter α controls the trade off between the accuracy of approximation and the performance of the algorithm [72]. Then, the problem of (7.11) is casted as a conic programming problem [71]

$$\min \alpha\|\mathbf{W}\|_* + \frac{1}{2}\|\mathbf{W}\|_F \quad \text{s.t.} \quad \|\mathbf{b} - \mathcal{A}(\mathbf{W})\| \leq \delta. \quad (7.12)$$

The problem of (7.12) can be solved with the singular value thresholding (SVT) algorithm in an iterative fashion (see [71]); the iteration converges to that of the original matrix completion problem as $\alpha \rightarrow \infty$ [71].

The recovered training data matrix $\hat{\mathbf{S}}$ is the optimal solution \mathbf{W}_{opt} of problem (7.12). Then, the sample covariance matrix is obtained as

$$\hat{\mathbf{R}}_{\text{mc}} = \frac{1}{L} \hat{\mathbf{S}} \hat{\mathbf{S}}^H. \quad (7.13)$$

7.3.2 Training Data Matrix Estimation Error Analysis

Define the signal-plus-interference-to-noise ratio over the observed data matrix as

$$\eta = \|P_{\Omega}(\mathbf{S})\|_F / \|P_{\Omega}(\mathbf{Z})\|_F.$$

Define the relative recovery error of the sample data matrix $\hat{\mathbf{S}}$ as $\phi_{\hat{\mathbf{S}}} = \|\hat{\mathbf{S}} - \mathbf{S}\|_F / \|\mathbf{S}\|_F$.

Then, we have the following lemmas.

Lemma 2. *The relative recovery error bound of the training data matrix is on the order of bound of $\frac{1}{\eta}$, i.e., $\text{bound}(\phi_{\hat{\mathbf{S}}}) = (1 \pm \varepsilon) \text{bound}\left(\frac{1}{\eta}\right)$, where ε is a small number.*

The proof is given in Appendix D.

Lemma 3. *The error bound of the recovered data matrix $\hat{\mathbf{S}}$ to the original noisy training data matrix \mathbf{X} is on the order of noise level, i.e., $\|\hat{\mathbf{S}} - \mathbf{X}\|_F \leq (C_0 + 1) \sqrt{ML} \sigma$ with a numerical constant C_0 , where σ is the standard deviation of the white noise sequence.*

The proof is given in Appendix E.

7.4 Robust Low Rank Beamforming

The recovered training data matrix $\hat{\mathbf{S}}$ contains errors which result in degradation of the performance of the traditional beamforming methods. Therefore, robust beamforming methods are required. In this paper, with the robust optimization techniques [104], we develop a beamformer that is robust against training data matrix mismatch, and it is based on worst-case performance optimization, along the lines of [108].

The training data matrix \mathbf{X} can be modeled as

$$\mathbf{X} = \hat{\mathbf{S}} + \mathbf{\Delta}. \quad (7.14)$$

Based on Lemma 3, the error matrix $\mathbf{\Delta}$ is bounded, i.e., $\|\mathbf{\Delta}\|_F \leq (C_0 + 1) \sqrt{ML}\sigma \triangleq \beta$.

In the Capon beamformer, the objective is to minimize

$$\begin{aligned} \mathbf{w}^H \hat{\mathbf{R}} \mathbf{w} &= \frac{1}{L} \mathbf{w}^H \mathbf{X} \mathbf{X}^H \mathbf{w} \\ &= \frac{1}{L} (\mathbf{X}^H \mathbf{w})^H (\mathbf{X}^H \mathbf{w}) \\ &= \frac{1}{L} \|\mathbf{X}^H \mathbf{w}\|^2. \end{aligned} \quad (7.15)$$

Therefore, minimizing $\mathbf{w}^H \hat{\mathbf{R}} \mathbf{w}$ is equivalent to minimizing $\|\mathbf{X}^H \mathbf{w}\|$. We want to obtain the weight vector by solving the following optimization problem

$$\min_{\mathbf{w}} \max_{\|\mathbf{\Delta}\|_F \leq \beta} \|\mathbf{X}^H \mathbf{w}\| \quad \text{s.t.} \quad |\mathbf{w}^H \mathbf{a}| \geq 1. \quad (7.16)$$

As in [108], we want to minimize the worst-case output power corresponding to the worst-case error matrix $\mathbf{\Delta}$ subject to the distortionless response for the desired signal.

The problem (7.16) can be equivalently written as

$$\min_{\mathbf{w}} \max_{\|\mathbf{\Delta}\|_F \leq \beta} \left\| \hat{\mathbf{S}}^H \mathbf{w} + \mathbf{\Delta}^H \mathbf{w} \right\| \quad \text{s.t.} \quad |\mathbf{w}^H \mathbf{a}| \geq 1. \quad (7.17)$$

Let $f(\mathbf{w}) = \max_{\|\mathbf{\Delta}\|_F \leq \beta} \left\| \hat{\mathbf{S}}^H \mathbf{w} + \mathbf{\Delta}^H \mathbf{w} \right\|$, then, it was shown in [108] that $f(\mathbf{w}) = \left\| \hat{\mathbf{S}}^H \mathbf{w} \right\| + \beta \|\mathbf{w}\|$.

Thus, we can rewrite the problem of (7.17) as

$$\min_{\mathbf{w}} \left\| \hat{\mathbf{S}}^H \mathbf{w} \right\| + \beta \|\mathbf{w}\| \quad \text{s.t.} \quad |\mathbf{w}^H \mathbf{a}| \geq 1. \quad (7.18)$$

The absolute operation in the constraint of problem (7.18) makes it nonconvex. Fortunately, when the weight vector \mathbf{w} undergoes any phase rotation, the cost function in (7.18) is unchanged. Therefore, the weight vector \mathbf{w} can be chosen to satisfy

$$\Re\{\mathbf{w}^H \mathbf{a}\} \geq 0, \quad \Im\{\mathbf{w}^H \mathbf{a}\} = 0. \quad (7.19)$$

Thus, the problem of (7.18) can be equivalently written as

$$\min_{\mathbf{w}} \left\| \hat{\mathbf{S}}^H \mathbf{w} \right\| + \beta \|\mathbf{w}\| \quad \text{s.t.} \quad \mathbf{w}^H \mathbf{a} \geq 1. \quad (7.20)$$

If the constraint of problem (7.20) is true, then the conditions in (7.19) would be satisfied. Also, the inequality constraint in (7.20) can be replaced by the equality $\mathbf{w}^H \mathbf{a} = 1$.

Alternative robust beamforming methods are the diagonal loading (DL) [109] and the eigenspace-based beamformer [102]. The idea of DL is to replace the sample covariance matrix $\hat{\mathbf{R}}_{\text{mc}}$ with $\hat{\mathbf{R}}_{\text{DL}} = \hat{\mathbf{R}}_{\text{mc}} + \xi \mathbf{I}$. The main difficulty in the DL method is in choosing the parameter ξ ; if ξ is too large, the beamformer fails to suppress the interference since most of the effort is used for white noise suppression. In practice, ξ is usually chosen as $10\sigma^2$ [105].

In the eigenspace-based beamformer, the basic idea is to use a projection of the steering vector \mathbf{a} onto the sample signal-plus-interference subspace. The eigenspace beamformer is known to be powerful but its performance degrades a lot at low SNR [105].

7.5 Numerical Results

In the simulations, we use a ULA with $M = 40$ sensors. The interspace distance between each sensor is set to $\lambda/2$. Assume there is one signal source and its direction of arrival (DOA) is -20° and supposed to be known. Two interference sources are in the plane space with DOA as -15° and -10° , respectively. Thus, $K = 3$. The frequencies of both signal and interference sources are set to $f = 1 \times 10^9 \text{Hz}$. The observation noise sequence is assumed to be white Gaussian with zero mean and standard deviation as σ . The sample data matrix recovery is done using the SVT algorithm [71]. The parameter

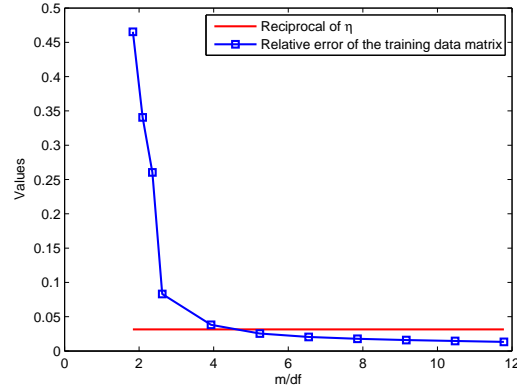


Figure 7.1: The relative recovery error. The x-axis is the number of samples per degree of freedom (df).

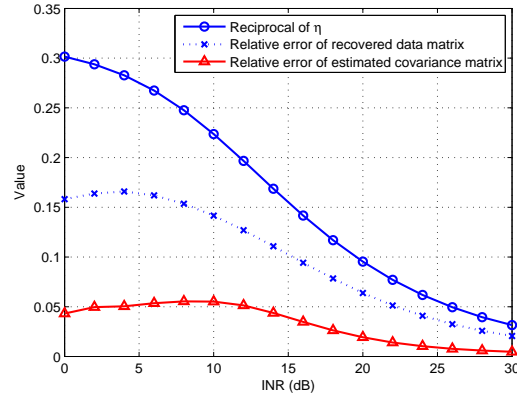


Figure 7.2: The relative errors of $\hat{\mathbf{S}}$ and $\hat{\mathbf{R}}_{\text{mc}}$ versus INR.

α is set to $\alpha = 20\sqrt{ML}$. Each sensor carries out a uniformly random sampling and forwards the sub-Nyquist samples to the fusion center. The DL and eigenspace beamformers were also implemented. In DL, the diagonal loading factor is chosen as $\xi = 10\sigma^2$. Define the interference-to-noise ratio (INR) as $\text{INR} = \|\mathbf{S}_{\text{int}}\|_F / \|\mathbf{Z}\|_F$, where \mathbf{S}_{int} is the sampled interference data matrix.

Fig. 7.1 shows the relative recovery error $\phi_{\mathbf{S}}$ of the training data matrix versus the number of samples per degree of freedom. The simulation parameters are set as $\text{SNR} = 10\text{dB}$, $\text{INR} = 30\text{dB}$ and $L = 2000$. The degree of freedom of training data matrix \mathbf{S} is $\text{df} = (M + L - K)K = 6111$. In total, 100 iterations are run and the relative errors are averaged. It can be seen that when m/df increases from 2 to 4, the

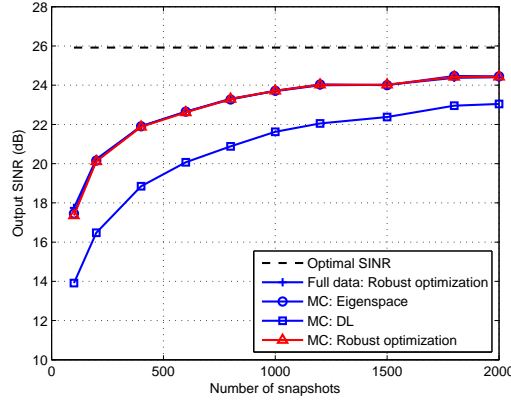


Figure 7.3: The output SINR versus the number of snapshots.

relative error $\phi_{\hat{\mathbf{S}}}$ drops sharply to the reciprocal of the observed signal-plus-interference-to-noise ratio level, i.e., the “phase transition” happens. In our following simulations, we set sampling ratio as $p = \frac{m}{ML} = 0.5$, i.e., $m/\text{df} \approx 6$.

Fig. 7.2 compares the the reciprocal of the observed signal-plus-interference-to-noise ratio η and the relative error $\phi_{\hat{\mathbf{S}}}$ in $\hat{\mathbf{S}}$ as well as the relative error in $\hat{\mathbf{R}}_{\text{mc}}$. Set $\text{SNR} = 10\text{dB}$ and $L = 2000$. 100 iterations are run and relative errors are averaged. It can be seen that in the entire INR range, the $\phi_{\hat{\mathbf{S}}}$ are much smaller than the reciprocal of η , which confirms the conclusion in Lemma 2. In addition, the relative error in $\hat{\mathbf{R}}_{\text{mc}}$ is even smaller.

Fig. 7.3 compares the performances of several beamformers in terms of the number of snapshot L , i.e., the robust optimization, DL as well as eigenspace-based beamformers, applied on the covariance matrix estimate obtained via matrix completion (MC). The SNR and INR are set as 10dB and 30dB, respectively. 100 iterations have been run to calculate the average output SINR. The optimal SINR (7.6) is also plotted for reference. It can be seen from Fig. 7.3 that as L increases, the SINR output under both robust optimization and eigenspace-based low rank beamformers approaches the optimal SINR within about 3dB. The results also confirms that to achieve a high SINR output, large snapshots are required. Therefore, our matrix completion based low rank beamforming method is important.

The SINR performance of these beamformers versus the SNR is shown in Fig. 7.4

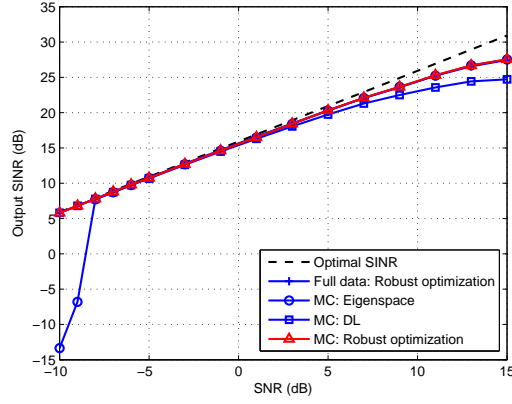


Figure 7.4: The output SINR versus SNR.

for $\text{INR} = 30\text{dB}$ and $L = 2000$. 100 iterations were run and the relative errors were averaged. It should be noted that at low SNR, the signal of interest is buried in the background noise and the training data matrix \mathbf{X} is approximately of rank 2, which is the number of interference sources. Via matrix completion, the recovered training data matrix $\hat{\mathbf{S}}$ is of rank 2, and contains the interference signal information, based on which, the suppression of interference can be achieved with the beamformers. It can be seen in Fig. 7.4 that among all matrix completion based robust beamformers, the robust optimization low rank adaptive beamformer has the best SINR output in the entire SNR range, and its SINR is identical to that of the robust optimization beamformer with full data. The DL beamformer loses some performance in the high SNR region, while the eigenspace-based beamformer does not work well at low SNR.

Last, the beampattern comparison is displayed in Fig. 7.5. The parameters are set as $L = 2000$, $\text{SNR} = 10\text{dB}$ and $\text{INR} = 30\text{dB}$. Compared with the DL beamformer, the proposed robust optimization low rank adaptive beamformer gives the best suppression to the interference sources at DOA -15° and -10° while keeping the signal source at DOA -20° distortionless.

7.6 Summary

In this chapter, beamforming problem in the large size sensor array with signal of interest present has been studied, in which huge training data needs to be collected to

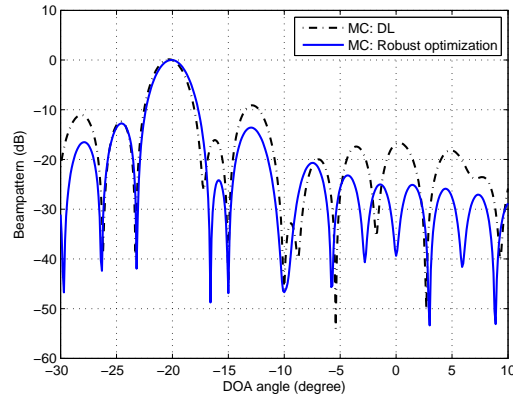


Figure 7.5: Beampatterns comparison.

generate a comparable SINR output. When the training data matrix is low rank, each sensor only needs to carry out a uniformly random sampling and forwards the sub-Nyquist samples to the fusion center. The full training data matrix is reconstructed by matrix completion. In the noisy observation case, the recovery is not exact. To deal with the errors in the reconstructed training data matrix introduced by the matrix completion, robust beamforming method was then proposed based on the worst-case performance optimization. Numerical results show that the proposed beamformer achieves near-optimal performance with half observations.

Chapter 8

Conclusions and Future Research Directions

8.1 Conclusions

The dissertation has exploited the sparsity in the context of MIMO radars.

First, we have proposed power allocation and waveform design algorithms to improve the target estimation performance in CS-based colocated and widely separated MIMO radars. In the context of compressive sensing, the uniform uncertain principle (UUP) states that the columns of the sensing matrix should be as orthogonal as possible. Based on the UUP, the power allocation and waveform design objective is to minimize the coherence among the columns of the sensing matrix arising in the CS-based MIMO radars. However, the design problem is not convex and we have proposed a semidefinite relaxation approach. The target estimation performance under the proposed algorithms has been improved greatly compared to the CS-based MIMO radars without power allocation or waveform design. However, the sparsity assumption of target space does not hold if there is clutter. In the context of CS-based MIMO radars, we have proposed a clutter suppression approach based on Capon beamforming. In the proposal, the Capon beamformer has been applied to the compressed receive data and consequently the block sparsity of the beamformed data has been exploited for target estimation using CS. The power allocation scheme based on the UUP has been proposed to further improve the target estimation performance in the CS-based MIMO radars with clutter suppression.

In the second part of the dissertation, we have proposed a colocated MIMO radars approach based on matrix completion, termed as MIMO-MC radar. It has been shown that the data matrix arising in the fusion center of the colocated MIMO radars with

large scale antennas is low-rank. Additionally, we have shown that the data matrices under both sampling schemes, i.e., with and without matched filtering satisfy the coherence conditions required by the matrix completion theory. Consequently, the data matrix can be completed with a small portion of samples collected in a uniformly random fashion from each receive antenna. The target estimation follows the recovered matrix using standard array signal processing. The proposed MIMO-MC radar approach does not require discretization of target space and thus avoids the basis mismatch issues inherent in CS. High resolution target estimation can be achieved by the proposed MIMO-MC approach with significantly reduced samples as compared to MIMO radars with the same number of elements.

We have further conducted an analysis of the coherence of the data matrix arising in MIMO-MC radars with ULA configurations and transmitting orthogonal waveforms. We have shown that, the data matrix attains its lowest possible coherence if the waveform snapshots across the transmit array have flat power spectra for all time instances. The waveform design problem has been approached as an optimization problem on the complex Stiefel manifold and has been solved via the modified steepest descent algorithm and the modified Newton algorithm. The numerical results have shown that as the number of antennas increases, the optimized waveforms result in optimal data matrix coherence, i.e., 1, and thus, only a small portion of samples are needed for the data matrix recovery. Although the optimal waveforms are designed based on the assumption that the targets fall in the same range bin, our analysis and simulations showed that they cause only small amount of performance degradation for relatively small delays, i.e., of the order of the symbol interval when used in a scenario in which the targets appear in different range bins.

Finally, we have investigated the transmit and receive beamforming problems in MIMO-MC radar, respectively. In the tracking mode, we have proposed a transmit beamforming scheme in MIMO-MC radars using sub-Nyquist sampling at the receiver end while achieving comparable target estimation performance to Nyquist sampling. The optimal waveform conditions has been derived for both the rank-1 and multi-rank transmit beamforming in MIMO-MC radars by conducting the coherence analysis of

the data matrix. A robust receive beamforming scheme based on matrix completion has been proposed for a large scale sensor array. The performance of receive beamforming techniques are based on the accurate estimation of the covariance matrix, which requires a large amount samples, especially when the targets of interest are in the presence of receive data. Our proposed robust beamforming scheme based on the matrix completion techniques at the receiver end could yield significantly improved performance with only a small portion of receive samples.

8.2 Future Research Directions

8.2.1 Waveform Design for Range Compression in MIMO-MC Radars

In Chapter 5, we have conducted coherence analysis of data matrix arising in the scenario of targets falling into different range bins, and shown in Corollary 1 that the matrix coherence achieves its lowest upper bound if the proposed waveforms additionally have good correlation properties. It is worth investigating of joint waveform design to meet both spatial spectra and correlation properties. It should be noted that it is impossible to achieve orthogonal waveforms with exactly zero correlations for the whole delay range [80] due to the lack of degree of freedom. Instead, we could design waveforms satisfying both spectra conditions and correlation properties only for a small portion delay or apply weights to correlation conditions.

8.2.2 Waveform Design for Transmit Beamforming in MIMO-MC Radars

In Chapter 6, we have derived optimal waveform conditions for multi-rank transmit beamforming (see Theorem 5). Additionally, it is shown that the waveform condition is independent of the multi-rank beamformer. Therefore, the results hold for all kinds of multi-rank beamformers obtained via different methods, e.g., multi-rank beamformer for the approximation of a desired beampattern, proposed in [101]. It is worth investigating of joint waveform design to meet both the waveform conditions stated in Theorem 5 for matrix completion as well as other design objectives, such as a desired beampattern. In the joint waveform design problem, the waveform conditions stated in

Theorem 5 can be formulated as a constraint in the new optimization problem.

8.2.3 Clutter Suppression in MIMO-MC Radars

When there is clutter, the sparsity assumption of CS-based MIMO radars about the target space does not hold. Similarly, clutter would invalidate the low-rank structure of the data matrix arising in MIMO-MC radars. In Chapter 3, we proposed a Capon beamforming method to suppress the clutter in CS-based MIMO radars assuming knowledge of the clutter covariance. The key step is applying the Capon beamformer to the compressed receive data and observing that the beamformer output is block sparse.

Clutter suppression in MIMO-MC radars is worth investigating. Under the stationary clutter assumption, we can optimally design the transmit waveforms to suppress the clutter. The objective of the optimization problem would be to minimize the effective clutter power, which can be calculated assuming that the random matrix completion sampling pattern as well as the clutter covariance matrix are known prior, subject to the total transmit power constraint as well as the optimal transmit waveforms conditions stated in Chapter 5.

Appendix A

Proof of $\mu_0 = \mu_1 = 1$ for rank-1 matrix \mathbf{Z}_q^{MF}

Proof. Suppose that there are $K, K \geq 2$ targets in the search space, all with the same DOA, say θ_1 . The transmit and receive steering matrices are given by

$$\mathbf{A} = [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_1)], \quad (\text{A.1})$$

$$\mathbf{B} = [\mathbf{b}(\theta_1), \dots, \mathbf{b}(\theta_1)], \quad (\text{A.2})$$

where the transmit and receive steering vectors $\mathbf{a}(\theta_1)$ and $\mathbf{b}(\theta_1)$ are defined in equations (4.2) and (4.5), respectively. The noise-free receive data matrix \mathbf{Z}_q^{MF} can be written as

$$\begin{aligned} \mathbf{Z}_q^{MF} &= \mathbf{B} \mathbf{\Sigma} \mathbf{D}_q \mathbf{A}^T \\ &= [\mathbf{b}(\theta_1), \dots, \mathbf{b}(\theta_1)] \begin{bmatrix} \beta_1 & & \\ & \ddots & \\ & & \beta_K \end{bmatrix} \times \\ &\quad \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_K \end{bmatrix} [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_1)]^T \\ &= \left(\sum_{k=1}^K \beta_k d_k \right) \mathbf{b}(\theta_1) \mathbf{a}^T(\theta_1), \end{aligned} \quad (\text{A.3})$$

where d_k is the Doppler shift of the k -th target. Its compact SVD is

$$\mathbf{Z}_q^{MF} = \mathbf{u} \sigma \mathbf{v}^H, \quad (\text{A.4})$$

where $\mathbf{u}^H \mathbf{u} = 1, \mathbf{v}^H \mathbf{v} = 1$, and σ is the singular value.

By applying the QR decomposition to the receive steering vector $\mathbf{b}(\theta_1)$, we have $\mathbf{b}(\theta_1) = \mathbf{q}_r r_r$, where $\mathbf{q}_r^H \mathbf{q}_r = 1$ and $r_r = \sqrt{M_r}$. The expression of \mathbf{q}_r is given by

$$\mathbf{q}_r = \frac{1}{\sqrt{M_r}} \left[1, e^{j \frac{2\pi}{\lambda} d_r \sin(\theta_1)}, \dots, e^{j \frac{2\pi}{\lambda} (M_r-1) d_r \sin(\theta_1)} \right]^T. \quad (\text{A.5})$$

Similarly, applying the QR decomposition to the transmit steering vector $\mathbf{a}(\theta_1)$, we have $\mathbf{a}(\theta_1) = \mathbf{q}_t r_t$, where $\mathbf{q}_t^H \mathbf{q}_t = 1$ and $r_t = \sqrt{M_t}$. The expression of \mathbf{q}_t is given by

$$\mathbf{q}_t = \frac{1}{\sqrt{M_t}} \left[1, e^{j \frac{2\pi}{\lambda} d_t \sin(\theta_1)}, \dots, e^{j \frac{2\pi}{\lambda} (M_t-1) d_t \sin(\theta_1)} \right]^T. \quad (\text{A.6})$$

Therefore, it holds that

$$\mathbf{Z}_q^{MF} = \mathbf{q}_r \underbrace{r_r \left(\sum_{k=1}^K \beta_k d_k \right)}_{\eta} r_t \mathbf{q}_t^T, \quad (\text{A.7})$$

where η is a complex number. Its SVD can be written as $\eta = q_1 \rho q_2^*$, where $|q_1| = |q_2| = 1$, and ρ is a real number. Thus,

$$\mathbf{Z}_q^{MF} = \mathbf{q}_r q_1 \rho q_2^* \mathbf{q}_t^T = \mathbf{q}_r q_1 \rho (\mathbf{q}_t^* q_2)^H, \quad (\text{A.8})$$

where $(\mathbf{q}_r q_1)^H \mathbf{q}_r q_1 = |q_1|^2 \mathbf{q}_r^H \mathbf{q}_r = 1$ and $(\mathbf{q}_t^* q_2)^H \mathbf{q}_t^* q_2 = |q_2|^2 (\mathbf{q}_t^H \mathbf{q}_t)^* = 1$. By the uniqueness of the singular value, it holds that $\rho = \sigma$. Therefore, we can set $\mathbf{u} = \mathbf{q}_r q_1$ and $\mathbf{v} = \mathbf{q}_t^* q_2$.

Let $\mathbf{q}_r^{(i)}$ denote the i -th element of vector \mathbf{q}_r . The coherence $\mu(U)$ is given by

$$\begin{aligned} \mu(U) &= \frac{M_r}{1} \sup_{i \in \mathbb{N}_{M_r}^+} \left\| \mathbf{q}_r^{(i)} q_1 \right\|_2^2 \\ &= M_r \sup_{i \in \mathbb{N}_{M_r}^+} \left\| \mathbf{q}_r^{(i)} \right\|_2^2 \\ &= 1. \end{aligned} \quad (\text{A.9})$$

Let $\mathbf{q}_t^{*(i)}$ denote the i -th element of vector \mathbf{q}_t^* . The coherence $\mu(V)$ is given by

$$\begin{aligned} \mu(V) &= \frac{M_t}{1} \sup_{i \in \mathbb{N}_{M_t}^+} \left\| \mathbf{q}_t^{*(i)} q_2 \right\|_2^2 \\ &= M_t \sup_{i \in \mathbb{N}_{M_t}^+} \left\| \mathbf{q}_t^{*(i)} \right\|_2^2 \\ &= 1. \end{aligned} \quad (\text{A.10})$$

Consequently, we have $\mu_0 = \max(\mu(U), \mu(V)) = 1$. In addition, we have $\mu_1 \leq \mu_0 \sqrt{K} = 1$ [41]. It always holds that $\mu_1 \geq 1$. Thus, $\mu_1 = 1$. Therefore, we have $\mu_0 = \mu_1 = 1$. \square

Appendix B

Proof of Lemma 1

Proof. Assume the MIMO radars systems are configured with ULA transmit array of size M_t and inter-element spacing as d_t . There are K targets in the far-field at DOAs $\{\theta_k \in [-\frac{\pi}{2}, \frac{\pi}{2}]\}_{k \in \mathbb{N}_K^+}$, corresponding to spatial frequencies $\{\alpha_k^t \in [-\frac{1}{2}, \frac{1}{2}]\}_{k \in \mathbb{N}_K^+}$. Then the transmit steering matrix \mathbf{A} has the Vandermonde form. As a result, it holds that $\text{tr}(\mathbf{A}\mathbf{A}^H) = KM_t$.

Suppose that orthogonal waveforms are transmitted so that $\mathbf{S} \in \mathbb{C}^{N \times M_t}$. Since $\mathbf{S}^H \mathbf{S} = \mathbf{I}_{M_t}$, it holds that

$$\sum_{i=1}^N s_{m'}^*(i) s_m(i) = \begin{cases} 1, m = m' \\ 0, m \neq m' \end{cases}, \quad m, m' \in \mathbb{N}_{M_t}^+. \quad (\text{B.1})$$

Consequently,

$$\sum_{i=1}^N (\mathbf{S}^{(i)})^H \mathbf{S}^{(i)} = \mathbf{I}_{M_t}, \quad (\text{B.2})$$

where $\mathbf{S}^{(i)}$ denotes the i -th row of \mathbf{S} . Following the equation (5.17), it holds that

$$\begin{aligned} \sum_{i=1}^N \sum_{k=1}^K |S_i(\alpha_k^t)|^2 &= \sum_{i=1}^N \left\| \mathbf{S}^{*(i)} \mathbf{A}^* \right\|_2^2 = \sum_{i=1}^N \text{tr} \left(\mathbf{A}^T (\mathbf{S}^{*(i)})^H \mathbf{S}^{*(i)} \mathbf{A}^* \right) \\ &= \sum_{i=1}^N \text{tr} \left((\mathbf{S}^{*(i)})^H \mathbf{S}^{*(i)} \mathbf{A}^* \mathbf{A}^T \right) = \sum_{i=1}^N \text{tr} \left((\mathbf{S}^{(i)})^H \mathbf{S}^{(i)} \mathbf{A} \mathbf{A}^H \right) \\ &= \text{tr} \left(\left(\sum_{i=1}^N (\mathbf{S}^{(i)})^H \mathbf{S}^{(i)} \right) \mathbf{A} \mathbf{A}^H \right) \\ &= \text{tr}(\mathbf{A} \mathbf{A}^H) \\ &= KM_t. \end{aligned} \quad (\text{B.3})$$

Thus the statements in Lemma 1 follow. \square

Appendix C

Derivative and Hessian of $f(\mathbf{S})$

First, we give the following two Lemmas, which will be applied in finding the the derivation.

Lemma 4. *Let $\mathbf{A} \in \mathbb{C}^{M_t \times L}$, $\mathbf{Z} \in \mathbb{C}^{N \times M_t}$, $\mathbf{Y} \in \mathbb{C}^{L \times N}$, $\mathbf{H} \in \mathbb{R}^{N \times L}$ be arbitrary matrices. It can be shown that*

$$\text{tr} \{ \mathbf{H} [(\mathbf{A}^H \mathbf{Z}^H) \odot \mathbf{Y}] \} = \text{tr} \{ \mathbf{Z}^H [(\mathbf{H} \odot \mathbf{Y}^T) \mathbf{A}^H] \}. \quad (\text{C.1})$$

Proof. See the Appendix F. □

Lemma 5. *Let $\tilde{\mathbf{H}} \in \mathbb{C}^{N \times L}$ and $\mathbf{G}, \mathbf{M} \in \mathbb{C}^{L \times N}$ be general matrices with arbitrary elements. It holds that*

$$\text{tr} \left((\mathbf{G} \odot \mathbf{M}) \tilde{\mathbf{H}} \right) = [\text{vec}(\mathbf{G})]^T \text{vec} \left(\mathbf{M} \odot \tilde{\mathbf{H}}^T \right). \quad (\text{C.2})$$

Proof. See the Appendix G. □

Since $\mathbf{F} \odot \mathbf{F}^*$ and \mathbf{N} are real-valued matrices, the objective function $f(\mathbf{S})$ defined in (5.29) can be written as

$$f(\mathbf{S}) = \text{tr} \left\{ (\mathbf{F} \odot \mathbf{F}^* - \mathbf{N}) (\mathbf{F} \odot \mathbf{F}^* - \mathbf{N})^T \right\}. \quad (\text{C.3})$$

In order to find the derivative and Hessian of $f(\mathbf{S})$, we do the following expansion:

$$\begin{aligned} f(\mathbf{S} + \delta \mathbf{Z}) &= \text{tr} \left\{ [((\mathbf{S} + \delta \mathbf{Z})^* \mathbf{A}^*) \odot ((\mathbf{S} + \delta \mathbf{Z}) \mathbf{A}) - \mathbf{N}] [((\mathbf{S} + \delta \mathbf{Z})^* \mathbf{A}^*) \odot ((\mathbf{S} + \delta \mathbf{Z}) \mathbf{A}) - \mathbf{N}]^T \right\} \\ &= f(\mathbf{S}) + \delta \text{tr} \{ \mathbf{T} \} + \delta^2 \text{tr} \{ \mathbf{T}' \} + \mathcal{O}(\delta^3). \end{aligned} \quad (\text{C.4})$$

Here,

$$\begin{aligned} \mathbf{T} = & \{[(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{SA})] - \mathbf{N}\} \left\{ [(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{ZA})]^T + [(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{ZA})]^H \right\} \\ & + \{[(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{ZA})] + [(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{ZA})]^*\} \left\{ [(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{SA})]^T - \mathbf{N}^T \right\}, \end{aligned} \quad (\text{C.5})$$

$$\begin{aligned} \mathbf{T}' = & [(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{SA})] [(\mathbf{Z}^* \mathbf{A}^*) \odot (\mathbf{ZA})]^T + \left\{ [(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{SA})] [(\mathbf{Z}^* \mathbf{A}^*) \odot (\mathbf{ZA})]^T \right\}^T \\ & + [(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{ZA})] [(\mathbf{Z}^* \mathbf{A}^*) \odot (\mathbf{SA})]^T + \left\{ [(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{ZA})] [(\mathbf{Z}^* \mathbf{A}^*) \odot (\mathbf{SA})]^T \right\}^T \\ & + [(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{ZA})] [(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{ZA})]^T + \left\{ [(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{ZA})] [(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{ZA})]^T \right\}^H \\ & - [(\mathbf{Z}^* \mathbf{A}^*) \odot (\mathbf{ZA})] \mathbf{N}^T - \left\{ [(\mathbf{Z}^* \mathbf{A}^*) \odot (\mathbf{ZA})] \mathbf{N}^T \right\}^T. \end{aligned} \quad (\text{C.6})$$

Thus, it holds that

$$\begin{aligned} \text{tr}(\mathbf{T}) = & \text{tr} \left(2 \{[(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{SA})] - \mathbf{N}\} \left\{ [(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{ZA})]^T + [(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{ZA})]^H \right\} \right) \\ = & \Re \left\{ \text{tr} \left(\mathbf{H} [(\mathbf{A}^H \mathbf{Z}^H) \odot (\mathbf{A}^T \mathbf{S}^T)] \right) \right\}, \end{aligned} \quad (\text{C.7})$$

where $\mathbf{H} = 2[(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{SA}) - \mathbf{N}] \in \mathbb{R}^{N \times L}$. Let $\mathbf{Y} = \mathbf{A}^T \mathbf{S}^T$. Following Lemma 4, it holds that

$$\text{tr}(\mathbf{T}) = \Re \left\{ \text{tr} \left(\mathbf{Z}^H [(\mathbf{H} \odot \mathbf{Y}^T) \mathbf{A}^H] \right) \right\}. \quad (\text{C.8})$$

In addition, it holds that

$$\begin{aligned} \text{tr}(\mathbf{T}') = & 2\text{tr} \left([(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{SA})] [(\mathbf{Z}^* \mathbf{A}^*) \odot (\mathbf{ZA})]^T \right) \\ & + 2\text{tr} \left([(\mathbf{SA}) \odot (\mathbf{Z}^* \mathbf{A}^*)] [(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{ZA})]^T \right) \\ & + 2\Re \left\{ \text{tr} \left([(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{ZA})] [(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{ZA})]^T \right) \right\} \\ & - 2\text{tr} \left(\mathbf{N} [(\mathbf{A}^H \mathbf{Z}^H) \odot (\mathbf{A}^T \mathbf{Z}^T)] \right). \end{aligned} \quad (\text{C.9})$$

Now, we focus on the first term on the right side of equation (C.9). Following Lemma 5, it holds that

$$\begin{aligned} & \text{tr} \left([(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{SA})] [(\mathbf{Z}^* \mathbf{A}^*) \odot (\mathbf{ZA})]^T \right) \\ = & \text{tr} \left([(\mathbf{A}^H \mathbf{Z}^H) \odot (\mathbf{A}^T \mathbf{Z}^T)] \tilde{\mathbf{H}} \right) \\ = & [\text{vec}(\mathbf{A}^H \mathbf{Z}^H)]^T \text{vec} \left((\mathbf{A}^T \mathbf{Z}^T) \odot \tilde{\mathbf{H}}^T \right), \end{aligned} \quad (\text{C.10})$$

where $\tilde{\mathbf{H}} = (\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{S} \mathbf{A}) \in \mathbb{R}^{N \times L}$. Further, via equations (5.38) and (H.9), it holds that

$$\begin{aligned}
 [\text{vec}(\mathbf{A}^H \mathbf{Z}^H)]^T &= [(\mathbf{I}_N \otimes \mathbf{A}^H) \text{vec}(\mathbf{Z}^H)]^T \\
 &= \left\{ (\mathbf{I}_N \otimes \mathbf{A}^H) [\text{vec}(\mathbf{Z}^T)]^* \right\}^T \\
 &= \left\{ (\mathbf{I}_N \otimes \mathbf{A}^H) [\mathbf{P}_{N \times M_t} \text{vec}(\mathbf{Z})]^* \right\}^T \\
 &= [\text{vec}(\mathbf{Z})]^H \mathbf{P}_{M_t \times N} (\mathbf{I}_N \otimes \mathbf{A}^*), \tag{C.11}
 \end{aligned}$$

as well as

$$\begin{aligned}
 \text{vec} \left((\mathbf{A}^T \mathbf{Z}^T) \odot \tilde{\mathbf{H}}^T \right) &= \text{diag} \left(\text{vec}(\tilde{\mathbf{H}}^T) \right) \text{vec}(\mathbf{A}^T \mathbf{Z}^T) \\
 &= \text{diag} \left(\text{vec}(\tilde{\mathbf{H}}^T) \right) (\mathbf{I}_N \otimes \mathbf{A}^T) \text{vec}(\mathbf{Z}^T) \\
 &= \text{diag} \left(\mathbf{P}_{N \times L} \text{vec}(\tilde{\mathbf{H}}) \right) (\mathbf{I}_N \otimes \mathbf{A}^T) \mathbf{P}_{N \times M_t} \text{vec}(\mathbf{Z}). \tag{C.12}
 \end{aligned}$$

Consequently, it holds that

$$\begin{aligned}
 &\text{tr} \left([(\mathbf{A}^H \mathbf{Z}^H) \odot (\mathbf{A}^T \mathbf{Z}^T)] \tilde{\mathbf{H}} \right) \\
 &= [\text{vec}(\mathbf{Z})]^H \mathbf{P}_{M_t \times N} (\mathbf{I}_N \otimes \mathbf{A}^*) \text{diag} \left(\mathbf{P}_{N \times L} \text{vec}(\tilde{\mathbf{H}}) \right) (\mathbf{I}_N \otimes \mathbf{A}^T) \mathbf{P}_{N \times M_t} \text{vec}(\mathbf{Z}). \tag{C.13}
 \end{aligned}$$

Let us focus on the second term on the right side of equation (C.9). It holds that

$$\begin{aligned}
 &\text{tr} \left([(\mathbf{S} \mathbf{A}) \odot (\mathbf{Z}^* \mathbf{A}^*)] [(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{Z} \mathbf{A})]^T \right) \\
 &= \text{tr} \left([(\mathbf{A}^H \mathbf{S}^H) \odot (\mathbf{A}^T \mathbf{Z}^T)]^H [(\mathbf{A}^H \mathbf{S}^H) \odot (\mathbf{A}^T \mathbf{Z}^T)] \right) \\
 &= \text{tr} \left([\mathbf{Y}^* \odot (\mathbf{A}^T \mathbf{Z}^T)]^H [\mathbf{Y}^* \odot (\mathbf{A}^T \mathbf{Z}^T)] \right) \\
 &= [\text{vec}(\mathbf{Y}^* \odot (\mathbf{A}^T \mathbf{Z}^T))]^H \text{vec}(\mathbf{Y}^* \odot (\mathbf{A}^T \mathbf{Z}^T)), \tag{C.14}
 \end{aligned}$$

as well as

$$\begin{aligned}
 \text{vec}(\mathbf{Y}^* \odot (\mathbf{A}^T \mathbf{Z}^T)) &= \text{diag}(\text{vec}(\mathbf{Y}^*)) \text{vec}(\mathbf{A}^T \mathbf{Z}^T) \\
 &= \text{diag}(\text{vec}(\mathbf{Y}^*)) (\mathbf{I}_N \otimes \mathbf{A}^T) \text{vec}(\mathbf{Z}^T) \\
 &= \text{diag}(\text{vec}(\mathbf{Y}^*)) (\mathbf{I}_N \otimes \mathbf{A}^T) \mathbf{P}_{N \times M_t} \text{vec}(\mathbf{Z}). \tag{C.15}
 \end{aligned}$$

Consequently, it holds that

$$\begin{aligned}
& \text{tr} \left([\mathbf{Y}^* \odot (\mathbf{A}^T \mathbf{Z}^T)]^H [\mathbf{Y}^* \odot (\mathbf{A}^T \mathbf{Z}^T)] \right) \\
&= [\text{vec}(\mathbf{Z})]^H \mathbf{P}_{M_t \times N} (\mathbf{I}_N \otimes \mathbf{A}^*) [\text{diag}(\text{vec}(\mathbf{Y}^*))]^H \\
&\quad \times \text{diag}(\text{vec}(\mathbf{Y}^*)) (\mathbf{I}_N \otimes \mathbf{A}^T) \mathbf{P}_{N \times M_t} \text{vec}(\mathbf{Z}) \\
&= [\text{vec}(\mathbf{Z})]^H \mathbf{P}_{M_t \times N} (\mathbf{I}_N \otimes \mathbf{A}^*) \text{diag}(\text{vec}(\mathbf{Y} \odot \mathbf{Y}^*)) (\mathbf{I}_N \otimes \mathbf{A}^T) \mathbf{P}_{N \times M_t} \text{vec}(\mathbf{Z}). \quad (\text{C.16})
\end{aligned}$$

Next, let us focus on the third term on the right side of equation (C.9). With equations (H.9) and (C.15), it holds that

$$\begin{aligned}
& \text{tr} \left([(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{Z} \mathbf{A})] [(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{Z} \mathbf{A})]^T \right) \\
&= \left[\text{vec} \left([(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{Z} \mathbf{A})]^T \right) \right]^T \text{vec} \left([(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{Z} \mathbf{A})]^T \right) \\
&= [\text{vec}(\mathbf{Y}^* \odot (\mathbf{A}^T \mathbf{Z}^T))]^T \text{vec}(\mathbf{Y}^* \odot (\mathbf{A}^T \mathbf{Z}^T)) \\
&= [\text{vec}(\mathbf{Z})]^T \mathbf{P}_{M_t \times N} (\mathbf{I}_N \otimes \mathbf{A}) [\text{diag}(\text{vec}(\mathbf{Y}^*))]^T \\
&\quad \times \text{diag}(\text{vec}(\mathbf{Y}^*)) (\mathbf{I}_N \otimes \mathbf{A}^T) \mathbf{P}_{N \times M_t} \text{vec}(\mathbf{Z}) \\
&= [\text{vec}(\mathbf{Z})]^T \mathbf{P}_{M_t \times N} (\mathbf{I}_N \otimes \mathbf{A}) \text{diag}(\text{vec}(\mathbf{Y}^* \odot \mathbf{Y}^*)) (\mathbf{I}_N \otimes \mathbf{A}^T) \mathbf{P}_{N \times M_t} \text{vec}(\mathbf{Z}). \quad (\text{C.17})
\end{aligned}$$

Finally, let us focus on the forth term on the right side of equation (C.9). Via equations (H.2) (C.11) (C.12) and Lemma 5, it holds that

$$\begin{aligned}
& \text{tr}(\mathbf{N} [(\mathbf{A}^H \mathbf{Z}^H) \odot (\mathbf{A}^T \mathbf{Z}^T)]) = \text{tr}([(\mathbf{A}^H \mathbf{Z}^H) \odot (\mathbf{A}^T \mathbf{Z}^T)] \mathbf{N}) \\
&= [\text{vec}(\mathbf{A}^H \mathbf{Z}^H)]^T \text{vec}((\mathbf{A}^T \mathbf{Z}^T) \odot \mathbf{N}^T) \\
&= [\text{vec}(\mathbf{Z})]^H \mathbf{P}_{M_t \times N} (\mathbf{I}_N \otimes \mathbf{A}^*) \text{diag}(\mathbf{P}_{N \times L} \text{vec}(\mathbf{N})) (\mathbf{I}_N \otimes \mathbf{A}^T) \mathbf{P}_{N \times M_t} \text{vec}(\mathbf{Z}). \quad (\text{C.18})
\end{aligned}$$

Therefore, it holds that

$$\begin{aligned}
f(\mathbf{S} + \delta \mathbf{Z}) &= f(\mathbf{S}) + \delta \Re \left\{ \text{tr}(\mathbf{Z}^H [(\mathbf{H} \odot \mathbf{Y}^T) \mathbf{A}^H]) \right\} \\
&\quad + 2\delta^2 [\text{vec}(\mathbf{Z})]^H \mathbf{P}_{M_t \times N} (\mathbf{I}_N \otimes \mathbf{A}^*) \text{diag} \left(\text{vec} \left(\tilde{\mathbf{H}}^T + \mathbf{Y} \odot \mathbf{Y}^* - \mathbf{N}^T \right) \right) \\
&\quad \times (\mathbf{I}_N \otimes \mathbf{A}^T) \mathbf{P}_{N \times M_t} \text{vec}(\mathbf{Z}) \\
&\quad + 2\delta^2 \Re \left\{ [\text{vec}(\mathbf{Z})]^T \mathbf{P}_{M_t \times N} (\mathbf{I}_N \otimes \mathbf{A}) \text{diag}(\text{vec}(\mathbf{Y}^* \odot \mathbf{Y}^*)) (\mathbf{I}_N \otimes \mathbf{A}^T) \mathbf{P}_{N \times M_t} \text{vec}(\mathbf{Z}) \right\} \\
&\quad + \mathcal{O}(\delta^3). \quad (\text{C.19})
\end{aligned}$$

By coefficient comparison between (C.19) and the matrix form of the second-order Taylor series (5.32), we finally obtain

$$\begin{aligned} \mathbf{D}_{\mathbf{S}} &= (\mathbf{H} \odot \mathbf{Y}^T) \mathbf{A}^H, \\ \mathbf{H}_{\mathbf{S}} &= 4\mathbf{P}_{M_t \times N} (\mathbf{I}_N \otimes \mathbf{A}^*) \text{diag} \left(\text{vec} \left(2\tilde{\mathbf{H}}^T - \mathbf{N}^T \right) \right) (\mathbf{I}_N \otimes \mathbf{A}^T) \mathbf{P}_{N \times M_t}, \\ \mathbf{C}_{\mathbf{S}} &= 4\mathbf{P}_{M_t \times N} (\mathbf{I}_N \otimes \mathbf{A}) \text{diag} \left(\text{vec} (\mathbf{Y}^* \odot \mathbf{Y}^*) \right) (\mathbf{I}_N \otimes \mathbf{A}^T) \mathbf{P}_{N \times M_t}, \end{aligned}$$

where $\mathbf{H} = 2[(\mathbf{S}^* \mathbf{A}^*) \odot (\mathbf{S} \mathbf{A}) - \mathbf{N}]$, $\mathbf{Y} = \mathbf{A}^T \mathbf{S}^T$ and the fact $\tilde{\mathbf{H}}^T = \mathbf{Y} \odot \mathbf{Y}^* = (\mathbf{A}^T \mathbf{S}^T) \odot (\mathbf{A}^H \mathbf{S}^H)$ is applied. It is easy to verify that $\mathbf{H}_{\mathbf{S}} = \mathbf{H}_{\mathbf{S}}^H$, $\mathbf{C}_{\mathbf{S}} = \mathbf{C}_{\mathbf{S}}^T$.

Appendix D

Proof of Lemma 2

Proof. The proof of Lemma 2 is based on the assumption that matrix \mathbf{S} satisfies the *restricted isometry property* (RIP) [36]. The RIP asserts that the sampling operator obeys

$$(1 - \epsilon) p \|\mathbf{S}\|_F^2 \leq \|\mathcal{P}_\Omega(\mathbf{S})\|_F^2 \leq (1 + \epsilon) p \|\mathbf{S}\|_F^2, \quad (\text{D.1})$$

provided that \mathbf{S} is low rank. Here, ϵ is a smaller constant. Since the training data matrix \mathbf{S} satisfies the *incoherence properties* defined in Section 1.4 of Chapter 1, then with high probability, on the observation set Ω whose elements are uniformly random sampled entries, the *restricted isometry property* holds [71]. The RIP property means that the ‘energy’ of \mathbf{S} on the observation set Ω is about proportional to the size of Ω .

Based on RIP, the results in [103] would yield that the recovered matrix $\hat{\mathbf{S}}$ by solving the convex optimization problem (7.11) satisfies

$$\|\hat{\mathbf{S}} - \mathbf{S}\|_F \leq \frac{C_0 \delta}{\sqrt{p}}, \quad (\text{D.2})$$

where C_0 is a numerical constant [43]. This means that the right hand of (D.2) is on the order of $\frac{\delta}{\sqrt{p}}$, i.e., $\|\hat{\mathbf{S}} - \mathbf{S}\|_F \leq O\left(\frac{\delta}{\sqrt{p}}\right)$. Thus, we have

$$\phi_{\hat{\mathbf{S}}} = \frac{\|\hat{\mathbf{S}} - \mathbf{S}\|_F}{\|\mathbf{S}\|_F} \leq \frac{O\left(\frac{\delta}{\sqrt{p}}\right)}{\|\mathbf{S}\|_F}. \quad (\text{D.3})$$

From the RIP (D.1), we have $\|\mathcal{P}_\Omega(\mathbf{S})\|_F \asymp \sqrt{p} \|\mathbf{S}\|_F$, where \asymp means there is a smaller number ϵ in (D.1). Since $\|\mathcal{P}_\Omega(\mathbf{Z})\|_F \leq \delta$, we have

$$\frac{1}{\eta} = \frac{\|\mathcal{P}_\Omega(\mathbf{Z})\|_F}{\|\mathcal{P}_\Omega(\mathbf{S})\|_F} \leq \frac{\delta}{\sqrt{p} \|\mathbf{S}\|_F}. \quad (\text{D.4})$$

By comparing the right hands of inequalities (D.3) and (D.4), the conclusion in Lemma 2 is proved. \square

Appendix E

Proof of Lemma 3

Proof. We have

$$\left\| \hat{\mathbf{S}} - \mathbf{X} \right\|_F = \left\| \hat{\mathbf{S}} - \mathbf{S} + \mathbf{S} - \mathbf{X} \right\|_F \quad (\text{E.1})$$

$$\leq \left\| \hat{\mathbf{S}} - \mathbf{S} \right\|_F + \left\| \mathbf{S} - \mathbf{X} \right\|_F \quad (\text{E.2})$$

$$\leq \frac{C_0 \delta}{\sqrt{p}} + \sqrt{ML} \sigma \quad (\text{E.3})$$

$$\leq C_0 \sqrt{\frac{m + \sqrt{8m}}{p}} \sigma + \sqrt{ML} \sigma \quad (\text{E.4})$$

$$\asymp (C_0 + 1) \sqrt{ML} \sigma. \quad (\text{E.5})$$

Here, in (E.2) the triangle inequality of matrix norm is applied. In (E.3), the bound (D.2) is applied. \square

Appendix F

Proof of Lemma 4

Proof. We use $a_{ij}, z_{ij}, y_{ij}, h_{ij}$ to denote the ij -th element of the corresponding matrices $\mathbf{A} \in \mathbb{C}^{M_t \times L}, \mathbf{Z} \in \mathbb{C}^{N \times M_t}, \mathbf{Y} \in \mathbb{C}^{L \times N}, \mathbf{H} \in \mathbb{R}^{N \times L}$, respectively. It holds that

$$\begin{aligned} \text{tr} \{ \mathbf{H} [(\mathbf{A}^H \mathbf{Z}^H) \odot \mathbf{Y}] \} &= \sum_{m=1}^N \sum_{n=1}^L h_{mn} \left(\sum_{i=1}^{M_t} a_{in}^* z_{mi}^* \right) y_{nm} \\ &= \sum_{i=1}^{M_t} \sum_{m=1}^N z_{mi}^* \left(\sum_{n=1}^L h_{mn} y_{nm} a_{in}^* \right). \end{aligned} \quad (\text{F.1})$$

Let $\mathbf{D}_{\mathbf{S}} \in \mathbb{C}^{N \times M_t}$ such that $\text{tr} \{ \mathbf{H} [(\mathbf{A}^H \mathbf{Z}^H) \odot \mathbf{Y}] \} = \text{tr} \{ \mathbf{Z}^H \mathbf{D}_{\mathbf{S}} \}$. We use d_{ij} to denote the ij -th element of $\mathbf{D}_{\mathbf{S}}$ and it holds that

$$\text{tr} \{ \mathbf{Z}^H \mathbf{D}_{\mathbf{S}} \} = \sum_{i=1}^{M_t} \sum_{m=1}^N z_{mi}^* d_{mi}. \quad (\text{F.2})$$

By comparing equations (F.1) and (F.2), we have

$$d_{mi} = \sum_{n=1}^L h_{mn} y_{nm} a_{in}^*. \quad (\text{F.3})$$

As a result, the matrix $\mathbf{D}_{\mathbf{S}}$ has the form as

$$\mathbf{D}_{\mathbf{S}} = (\mathbf{H} \odot \mathbf{Y}^T) \mathbf{A}^H. \quad (\text{F.4})$$

Consequently, it holds that $\text{tr} \{ \mathbf{H} [(\mathbf{A}^H \mathbf{Z}^H) \odot \mathbf{Y}] \} = \text{tr} \{ \mathbf{Z}^H [(\mathbf{H} \odot \mathbf{Y}^T) \mathbf{A}^H] \}$, which completes the proof. \square

Appendix G

Proof of Lemma 5

Proof. We use $g_{ij}, m_{ij}, \tilde{h}_{ij}$ to denote the ij -th element of the corresponding matrices $\mathbf{G}, \mathbf{M} \in \mathbf{C}^{L \times N}$ and $\tilde{\mathbf{H}} \in \mathbf{C}^{N \times L}$. Then, it holds that

$$\text{tr} \left\{ (\mathbf{G} \odot \mathbf{M}) \tilde{\mathbf{H}} \right\} = \sum_{j=1}^L \sum_{i=1}^N g_{ji} m_{ji} \tilde{h}_{ij}. \quad (\text{G.1})$$

On the other hand, it holds that

$$\text{vec} \left(\mathbf{M} \odot \tilde{\mathbf{H}}^T \right) = \left[m_{11} \tilde{h}_{11}, \dots, m_{L1} \tilde{h}_{1L}, \dots, m_{1N} \tilde{h}_{N1}, \dots, m_{LN} \tilde{h}_{NL} \right]^T, \quad (\text{G.2})$$

$$\text{vec}(\mathbf{G}) = [g_{11}, \dots, g_{L1}, \dots, g_{1N}, \dots, g_{LN}]^T. \quad (\text{G.3})$$

Consequently, it holds that

$$\begin{aligned} \text{vec}(\mathbf{G})^T \text{vec} \left(\mathbf{M} \odot \tilde{\mathbf{H}}^T \right) &= \sum_{i=1}^N \sum_{j=1}^L g_{ji} m_{ji} \tilde{h}_{ij} \\ &= \sum_{j=1}^L \sum_{i=1}^N g_{ji} m_{ji} \tilde{h}_{ij}. \end{aligned} \quad (\text{G.4})$$

By comparing equations (G.1) and (G.4), we have

$$\text{tr} \left\{ (\mathbf{G} \odot \mathbf{M}) \tilde{\mathbf{H}} \right\} = \text{vec}(\mathbf{G})^T \text{vec} \left(\mathbf{M} \odot \tilde{\mathbf{H}}^T \right), \quad (\text{G.5})$$

which completes the proof. \square

Appendix H

Useful Equations

Here, we list some useful equations for deriving the derivative and Hessian of a matrix-valued cost function with $\mathbf{Z} \in \mathbb{C}^{N \times M_t}$. They are

$$\|\mathbf{A}\|_F^2 = \text{tr}(\mathbf{A}\mathbf{A}^H), \quad (\text{H.1})$$

$$\text{tr}(\mathbf{Z}\mathbf{H}) = \text{tr}(\mathbf{H}\mathbf{Z}), \quad (\text{H.2})$$

$$\text{tr}(\mathbf{A}^T) = \text{tr}(\mathbf{A}), \quad (\text{H.3})$$

$$\text{tr}(\mathbf{A}^H) = \text{tr}(\mathbf{A}^*) = (\text{tr}(\mathbf{A}))^*, \quad (\text{H.4})$$

$$(\mathbf{H} \otimes \mathbf{Z})^T = \mathbf{H}^T \otimes \mathbf{Z}^T, \quad (\text{H.5})$$

$$(\mathbf{H} \odot \mathbf{Z})^T = \mathbf{H}^T \odot \mathbf{Z}^T, \quad (\text{H.6})$$

$$\text{tr}(\mathbf{H}\mathbf{Z}) = [\text{vec}(\mathbf{H}^T)]^T \text{vec}(\mathbf{Z}), \quad (\text{H.7})$$

$$\text{tr}(\mathbf{H}^H \mathbf{Z}) = [\text{vec}(\mathbf{H})]^H \text{vec}(\mathbf{Z}), \quad (\text{H.8})$$

$$\text{vec}(\mathbf{H}\mathbf{Z}\mathbf{G}) = (\mathbf{G}^T \otimes \mathbf{H}) \text{vec}(\mathbf{Z}). \quad (\text{H.9})$$

References

- [1] M. I. Skolnik, *Introduction to Radar Systems, 2nd edition*, New York, McGraw-Hill, 1980.
- [2] M. A. Richards, *Fundamentals of Radar Signal Processing, 2nd edition*, McGraw-Hill, New York, 2014.
- [3] P. Stoica and R. Moses, *Spectral Analysis of Signals*, New Jersey, Pearson Prentice Hall, 2005.
- [4] H. L. Van Trees, *Detection, Estimation, and Modulation Theory, Part III: Radar-Sonar Signal Processing and Gaussian Signals in Noise*, New York: Wiley, 2001.
- [5] H. L. Van Trees, *Optimum Array Processing*, New York: Wiley, 2002.
- [6] Woodward P.M., *Probability and Information Theory with Application to Radar*, McGraw-Hill, New York, 1953.
- [7] B. R. Mahafza, *Radar Systems Analysis and Design Using MATLAB*, New York, Chapman & Hall/CRC, 2000.
- [8] R. H. Barker, "Group synchronizing of binary digital sequences," *Communication Theory*, pp. 273–287, W. Jackson, Ed., London, U.K.: Butterworth, 1953.
- [9] M. H. Ackroyd and F. Ghani, "Optimum mismatched filters for sidelobe suppression," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-9, no. 2, pp. 214–218, 1973.
- [10] E. Fishler, A. Haimovich, R. Blum, L. Cimini, D. Chizhik, and R. Valenziela, "MIMO radar: An idea whose time has come," in *Proc. of IEEE Int. Conf. on Radar*, Philadelphia, PA, April 26-29, 2004.
- [11] J. Li and P. Stoica, *MIMO Radar Signal Processing*, Hoboken, NJ, Wiley, 2009.
- [12] M. I. Skolnik, *Radar Handbook, 3rd Ed*, New York, McGraw-Hill, 2008.
- [13] V. S. Chernyak, "On the concept of MIMO radar," in *Proc. IEEE International Radar Conference*, Arlington, VA, May 10-14, 2010.
- [14] E. Fishler, A. Haimovich, R. Blum, L. Cimini, D. Chizhik, and R. Valenziela, "Performance of MIMO radar systems: advantages of angular diversity," in *Proc. 38th Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, Nov. 7-10, 2004.
- [15] N. H. Lehmann, A. M. Haimovich, R. S. Blum, and L. J. Cimini, "High resolution capabilities of MIMO radar," in *Proc. 40th Asilomar Conf. Signals, Systems and Computers*, Pacific Grove, CA, Oct. 29-Nov.1, 2006.

- [16] E. Fishler, R. Blum A. M. Haimovich, L. Cimini, D. Chizhik, and R. Valenzuela, "Spatial diversity in radars-models and detection performance," *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 823–838, 2006.
- [17] A. M. Haimovich, R. S. Blum, and L. J. Cimini, "MIMO radar with widely separated antennas," *IEEE Signal Process. Mag.*, vol. 25, no. 1, pp. 116–129, 2008.
- [18] J. Li and P. Stoica, "MIMO radar with colocated antennas," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 106–114, 2007.
- [19] J. Li, P. Stoica, L. Xu, and W. Roberts, "On parameter identifiability of MIMO radar," *IEEE Signal Process. Lett.*, vol. 14, no. 12, pp. 968–971, 2007.
- [20] T. Aittomaki and V. Koivunen, "MIMO radar filterbank design for interference mitigation," in *Proc. of IEEE 39th Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Florence, Italy, May 4-9, 2014.
- [21] G. S. Antonio, D. R. Fuhrmann, and F. C. Robey, "MIMO radar ambiguity functions," *IEEE Trans. Signal Process.*, vol. 1, no. 1, pp. 167–177, 2007.
- [22] C. Y. Chen and P. P. Vaidyanathan, "MIMO radar ambiguity properties and optimization using frequency-hopping waveforms," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 5926–5936, 2008.
- [23] D. V. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [24] E. J. Candès, "Compressive sampling," in *Proc. of International Congress of Mathematicians*, Madrid, Spain, Aug. 22-30, 2006.
- [25] E. J. Candès and M. B. Wakin, "An introduction to compressive sampling: A sensing/sampling paradigm that goes against the common knowledge in data acquisition," *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, 2008.
- [26] C. Y. Chen and P. P. Vaidyanathan, "Compressed sensing in MIMO radar," in *42th Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Oct. 26-29, 2008.
- [27] T. Strohmer and B. Friedlander, "Compressed sensing for MIMO radar - algorithms and performance," in *43th Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 1-4, 2009.
- [28] M. Herman and T. Strohmer, "High resolution radar via compressed sensing," *IEEE Trans. Signal Process.*, vol. 57, no. 6, pp. 2275–2284, 2009.
- [29] Y. Yu, A. P. Petropulu, and H. V. Poor, "MIMO radar using compressive sampling," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 1, pp. 146–163, 2010.
- [30] Y. Yu, S. Sun, R. N. Madan, and A. P. Petropulu, "Power allocation and waveform design for the compressive sensing based MIMO radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 2, pp. 898–909, 2014.

- [31] T. T. Cai and L. Wang, “Orthogonal matching pursuit for sparse signal recovery with noise,” *IEEE Trans. Inf. Theory*, vol. 57, no. 7, pp. 4680–4688, 2011.
- [32] S. S. Chen, D. L. Donoho, and M. A. Saunders, “Atomic decomposition by basis pursuit,” *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1999.
- [33] B. Efron, T. Hstie, I. Johnstone, and R. Tibshirani, “Least angle regression,” *Annals of Statistics*, vol. 32, no. 2, pp. 407–499, 2004.
- [34] E. J. Candès, J. Romberg, and T. Tao, “Stable signal recovery from incomplete and inaccurate measurements,” *Communications on Pure and Applied Mathematics*, vol. 59, no. 8, pp. 1207–1223, 2006.
- [35] E. J. Candès and T. Tao, “Near-optimal signal recovery from random projections: Universal encoding strategies,” *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, 2006.
- [36] E. J. Candès and T. Tao, “Decoding by linear programming,” *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4203–4215, 2005.
- [37] D. L. Donoho and X. Huo, “Uncertainty principles and ideal atomic decomposition,” *IEEE Trans. Inf. Theory*, vol. 47, no. 7, pp. 2845–2862, 2001.
- [38] J. A. Tropp, “Greed is good: Algorithmic results for sparse approximation,” *IEEE Trans. Inf. Theory*, vol. 50, no. 10, pp. 2231–2242, 2004.
- [39] Y. Chi, L. L. Scharf, A. Pezeshki, and A. R. Calderbank, “Sensitivity to basis mismatch in compressed sensing,” *IEEE Trans. Signal Process.*, vol. 59, no. 5, pp. 2182–2195, 2011.
- [40] W. U. Bajwa, K. Gedalyahu, and Y. C. Eldar, “Identification of parametric underspread linear systems and super-resolution radar,” *IEEE Trans. Signal Process.*, vol. 59, no. 6, pp. 2548–2561, 2011.
- [41] E. J. Candès and B. Recht, “Exact matrix completion via convex optimization,” *Foundations of Computational Mathematics*, vol. 9, no. 6, pp. 717–772, 2009.
- [42] E. J. Candès and T. Tao, “The power of convex relaxation: Near-optimal matrix completion,” *IEEE Trans. Inf. Theory*, vol. 56, no. 5, pp. 2053–2080, 2010.
- [43] E. J. Candès and Y. Plan, “Matrix completion with noise,” *Proc. IEEE*, vol. 98, no. 6, pp. 925–936, 2010.
- [44] R. H. Keshavan, A. Montanari, and S. Oh, “Matrix completion from a few entries,” *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2980–2998, 2010.
- [45] W. Dai, E. Kerman, and O. Milenkovic, “A geometric approach to low-rank matrix completion,” *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 237–247, 2012.
- [46] B. Vandereycken, “Low-rank matrix completion by Riemannian optimization,” *SIAM Journal on Optimization*, vol. 23, no. 2, pp. 1214–1236, 2013.

- [47] A. P. Petropulu, Y. Yao, and H. V. Poor, "Distributed MIMO radar using compressive sampling," in *42th Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Oct. 26-29, 2008.
- [48] S. Wang, Q. He, and Z. He, "Compressed sensing moving target detection for MIMO radar with widely spaced antennas," in *2010 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, Chengdu, China, Dec., 2010.
- [49] S. Gogineni and A. Nehorai, "Target estimation using sparse modeling for distributed MIMO radar," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5315–5325, 2011.
- [50] Y. Yao, A. P. Petropulu, and J. Huang, "Exploring sparsity in widely separated MIMO radar," in *45th Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 6-9, 2011.
- [51] Y. Yu, A. P. Petropulu, and H. V. Poor, "CSSF MIMO radar: Low complexity compressive sensing based MIMO radar that uses step frequency," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 2, pp. 1490–1504, 2012.
- [52] A. C. Gurbuz, J. H. McClellan, and W. R. Scott, "A compressive sensing data acquisition and imaging method for stepped frequency GPRs," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2460–2650, 2009.
- [53] Y.-S. Yoon and M. G. Amin, "Imaging of behind the wall targets using wide-band beamforming with compressive sensing," in *Proc. of IEEE Workshop on Statistical Signal Processing (SSP)*, Cardiff, Wales, Aug. 31- Sept. 3, 2009.
- [54] Y. Yao, A. P. Petropulu, and H. V. Poor, "The measurement matrix design for compressive sensing based MIMO radar," *IEEE Trans. Signal Process.*, vol. 59, no. 11, pp. 5338–5352, 2011.
- [55] H. Godrich, A. P. Petropulu, and H. V. Poor, "Power allocation strategies for target localization in distributed multiple radar architectures," *IEEE Trans. Signal Process.*, vol. 59, no. 7, pp. 3226–3240, 2011.
- [56] J. Akhtar, "Energy allocation for correlated MIMO radar antennas with ricean targets," in *Proc. 11th International Radar Symposium (IRS)*, Vilnius, Lithuania, June 16-18, 2010.
- [57] J. Huang and T. Zhang, "The benefit of group sparsity," *Annals of Statistics*, vol. 38, no. 4, pp. 1978–2004, 2010.
- [58] J. Friedman, T. Hastie, and R. Tibshirani, "A note on the group lasso and a sparse group lasso," in *[online] statweb.stanford.edu/tibs/ftp/sparse-grlasso.pdf*, 2010.
- [59] X. Chen, Q. Lin, S. Kim, J. Carbonell, and E. Xing, "Smoothing proximal gradient method for general structured sparse learning," *Annals of Applied Statistics*, vol. 6, no. 2, pp. 719–752, 2012.

- [60] P. L. Combettes and J. C. Pesqust, "Proximal splitting methods in signal processing," *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, vol. 49, pp. 185–212, 2011.
- [61] A. R. Hunt, "Use of a frequency-hopping radar for imaging and motion detection through walls," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1402–11408, 2009.
- [62] N. Maaref, P. Millot, C. Pichot, and O. Picon, "A study of UWB FM-CW radar for the detection of human beings in motion inside a building," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1297–1300, 2009.
- [63] F. Ahmad and M. G. Amin, "Through-the-wall human motion indication using sparsity-driven change detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 2, pp. 881–890, 2013.
- [64] S. Kraut, K. Harmanci, and J. Krolik, "Space-time adaptive processing for over-the-horizon spread Doppler clutter estimation," in *Proc. of IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Cambridge, MA, Mar. 16–17, 2000.
- [65] L. J. Nickisch, "Non-uniform motion and extended media effects on the mutual coherence function: an analytic solution for spaced frequency, position, and time," *Radio Science*, vol. 27, no. 1, pp. 9–22, 1992.
- [66] C. Y. Chen and P. P. Vaidyanathan, "MIMO radar space-time adaptive processing using prolate spheroidal wave functions," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 623–635, 2008.
- [67] D. Nion and N. D. Sidiropoulos, "Tensor algebra and multi-dimensional harmonic retrieval in signal processing for MIMO radar," *IEEE Trans. Signal Process.*, vol. 58, no. 11, pp. 5693–5705, 2010.
- [68] D. S. Kalogerias and A. P. Petropulu, "Matrix completion in colocated MIMO radar: Recoverability, bounds & theoretical guarantees," *IEEE Trans. Signal Process.*, vol. 62, no. 2, pp. 309–321, 2014.
- [69] P. Dutta, A. Aroraand, and S. Bibyk, "Towards radar-enabled sensor networks," in *Proc. ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, Nashville, Tennessee, USA, April 19–21, 2006.
- [70] Q. Liang, "Radar sensor networks: Algorithms for waveform design and diversity with application to ATR with delay-Doppler uncertainty," *EURASIP Journal on Wireless Communications and Networking*, pp. 1–9, 2007.
- [71] J. F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 2, pp. 1956–1982, 2010.
- [72] S. Becker, E. J. Candès, and M. Grant, "Templates for convex cone problems with applications to sparse signal recovery," *Mathematical and Programming Computation*, vol. 3, no. 3, pp. 165–218, 2011.

- [73] P. Stoica, J. Li, and Y. Xie, "On probing signal design for MIMO radar," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4151–4161, 2007.
- [74] N. Levanon and E. Mozeson, *Radar Signals*, Hoboken, NJ, Wiley, 2004.
- [75] S. Sun, A. P. Petropulu, and W. U. Bajwa, "Target estimation in colocated MIMO radar via matrix completion," in *Proc. of IEEE 38th Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Vancouver, Canada, May 26–31, 2013.
- [76] S. Sun, W. U. Bajwa, and A. P. Petropulu, "MIMO-MC radar: A MIMO radar approach based on matrix completion," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 51, no. 3, pp. 1839–1852, 2015.
- [77] D. S. Kalogerias, S. Sun, and A. P. Petropulu, "Sparse sensing in colocated MIMO radar: A matrix completion approach," in *Proc. of IEEE 13th International Symposium on Signal Processing and Information Technology (ISSPIT)*, Athens, Greece, Dec. 12–15, 2013.
- [78] Y. Yang and R. S. Blum, "MIMO radar waveform design based on mutual information and minimum mean-square error estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 43, no. 1, pp. 330–343, 2007.
- [79] Y. Yang and R. S. Blum, "Minimax robust MIMO radar waveform design," *IEEE J. Sel. Topics Signal Process.*, vol. 1, no. 1, pp. 147–155, 2007.
- [80] H. He, P. Stoica, and J. Li, "Designing unimodular sequence sets with good correlations-including an application to MIMO radar," *IEEE Trans. Signal Process.*, vol. 57, no. 11, pp. 4391–4405, 2009.
- [81] K. W. Forsythe and D. W. Bliss, "MIMO radar waveform constraints for GMTI," *IEEE J. Sel. Topics Signal Process.*, vol. 4, no. 1, pp. 21–32, 2010.
- [82] H. Deng, "Polyphase code design for orthogonal netted radar systems," *IEEE Trans. Signal Process.*, vol. 52, no. 10, pp. 3126–3135, 2004.
- [83] S. Gogineni and A. Nehorai, "Frequency-hopping code design for MIMO radar estimation using sparse modeling," *IEEE Trans. Signal Process.*, vol. 60, no. 6, pp. 3022–3035, 2012.
- [84] A. Edelman, T. A. Arias, and S. T. Smith, "The geometry of algorithms with orthogonality constraints," *SIAM J. Matrix Anal. Appl.*, vol. 20, no. 2, pp. 303–353, 1998.
- [85] J. H. Manton, "Optimization algorithms exploiting unitary constraints," *IEEE Trans. Signal Process.*, vol. 50, no. 3, pp. 635–650, 2002.
- [86] P. A. Absil, R. Mathony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*, Princeton, Princeton University Press, 2008.
- [87] H. Zhang and W. W. Hager, "A nonmonotone line search technique and its application to unconstrained optimization," *SIAM Journal on Optimization*, vol. 14, no. 4, pp. 1043–1056, 2004.

- [88] A. Antoniou and W.-S. Lu, *Practical Optimization: Algorithms and Engineering Applications*, New York, Springer, 2007.
- [89] M. Joho and K. Rahbar, "Joint diagonalization of correlation matrices by using Newton methods with application to blind signal separation," in *Proc. of IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Rosslyn, VA, Aug. 4-6, 2002.
- [90] J. R. Magnus and H. Neudecker, "Symmetry, 0-1 matrices and Jacobians: A review," *Econometric Theory*, vol. 2, no. 2, pp. 157–190, 1986.
- [91] D. P. Bertsekas, *Nonlinear Programming*, 2nd ed., Athena Scientific, 1999.
- [92] M. Joho, "Newton method for joint approximate diagonalization of positive definite Hermitian matrices," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, pp. 1205–1218, 2008.
- [93] H. Wolkowicz and G. P. H. Styan, "Bounds for eigenvalues using traces," *Linear Algebra and Its Applications*, vol. 29, no. 2, pp. 471–506, 1980.
- [94] J. Li, L. Xu, P. Stoica, K. W. Forsythe, and D. W. Bliss, "Range compression and waveform optimization for MIMO radar: A Cramer-Rao bound based study," *IEEE Trans. Signal Process.*, vol. 56, no. 1, pp. 218–232, 2008.
- [95] S. Sun and A. P. Petropulu, "Waveform design for MIMO radar with matrix completion," *Technical Report, Rutgers University*, Jan., 2015 (<http://eceweb1.rutgers.edu/~cspl/publications/Waveform-TR.pdf>).
- [96] S. Sun and A. P. Petropulu, "On the applicability of matrix completion on MIMO radars," in *48th Annual Asilomar Conference on Signals, Systems, and Computers*, Pacific Grove, CA, Nov. 2-5, 2014.
- [97] S. Sun and A. P. Petropulu, "On waveform design for MIMO radar with matrix completion," in *Proc. of IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Atlanta, GA, USA, Dec. 3-5, 2014.
- [98] D. R. Fuhrmann and G. S. Antonio, "Transmit beamforming for MIMO radar systems using signal cross-correlation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, no. 1, pp. 171–186, 2008.
- [99] A. Hassanien and S. A. Vorobyov, "Phased-MIMO radar: A tradeoff between phased-array and MIMO radars," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3137–3151, 2010.
- [100] B. Friedlander, "On transmit beamforming for MIMO radar," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 48, no. 4, pp. 3376–3388, 2012.
- [101] A. Hassanien and S. A. Vorobyov, "Transmit energy focusing for DOA estimation in MIMO radar with colocated antennas," *IEEE Trans. Signal Process.*, vol. 59, no. 6, pp. 2669–2682, 2011.
- [102] D. D. Feldman and L. J. Griffiths, "A projection approach to robust adaptive beamforming," *IEEE Trans. Signal Process.*, vol. 42, pp. 867–876, 1994.

- [103] M. Fazel, E. Candès, B. Recht, and P. Parrilo, “Compressed sensing and robust recovery of low rank matrices,” in *Proc. of 42nd Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, Oct. 26-29, 2008.
- [104] A. Ben-Tal, L. E. Ghaoui, and A. Nemirovski, *Robust Optimization*, Princeton University Press, 2009.
- [105] 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 Trans. Signal Process.*, vol. 51, no. 2, pp. 313–324, 2003.
- [106] J. Capon, “High-resolution frequency-wavenumber spectrum analysis,” *Proc. IEEE*, vol. 57, no. 8, pp. 1408–1418, 1969.
- [107] A. Waters and V. Cevher, “Distributed bearing estimation via matrix completion,” in *Proc. of IEEE International Conference on Acoustic, Speech and Signal Processing (ICASSP)*, Dallas, TX, Mar. 14-19, 2010.
- [108] S. A. Vorobyov, A. B. Gershman, Z. Q. Luo, and N. Ma, “Adaptive beamforming with joint robustness against mismatched signal steering vector and interference nonstationarity,” *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 108–111, 2004.
- [109] J. Li, P. Stocia, and Z. Wang, “On robust capon beamforming and diagonal loading,” *IEEE Trans. Signal Process.*, vol. 51, no. 7, pp. 1702–1715, 2003.