AFRL-RY-WP-TR-2021-0162

CONVEX OPTIMIZATION FOR COGNITIVE RADAR (Preprint)

Muralidhar Rangaswamy
Multiband Multifunction Radio Frequency Sensing Branch
Multispectral Sensing & Detection Division

Bosung Kang
University of Dayton Research Institute

Khaled AlHujaili
Taibah University

Vishal Monga
Pennsylvania State University

JUNE 2021
Final Report

AIR FORCE RESEARCH LABORATORY
SENSORS DIRECTORATE
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-7320
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
June 2021	Book Chapter Preprint	11 June 2021 –11 June 2021

4. TITLE AND SUBTITLE
CONVEX OPTIMIZATION FOR COGNITIVE RADAR (Preprint)

5a. CONTRACT NUMBER
FA8650-18-C-1073

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
61102F/62212F/62204F

6. AUTHOR(S)
Muralidhar Rangaswamy (AFRL/RYMF)
Bosung Kang (University of Dayton Research Institute)
Khaled AlHujaili (Taibah University)
Vishal Monga (Pennsylvania State University)

5d. PROJECT NUMBER
3001/2030/4920

5e. TASK NUMBER
N/A

5f. WORK UNIT NUMBER
Y1Q0

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
Sensors Directorate (AFRL/RYMF)
Wright-Patterson Air Force Base, OH 45433-7320
Air Force Materiel Command
United States Air Force

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
Sensors Directorate
Wright-Patterson Air Force Base, OH 45433-7320
Air Force Materiel Command
United States Air Force

Air Force Office of Scientific Research Public Affairs (AFOSR)
875 North Randolph Street, Suite 325, Room 3112
Arlington, VA 22203

10. SPONSORING/MONITORING AGENCY ACRONYM(S)
AFRL/RYMF

11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)
AFRL-RY-WP-TR-2021-0162

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
PAO case number AFRL-2021-1819, Clearance Date 11 June 2021. Chapter 1 of a forthcoming book (book title to be determined). The U.S. Government is joint author of this work and has the right to use, modify, reproduce, release, perform, display, or disclose the work. Report contains color.

14. ABSTRACT
A confluence of factors continues to increase the complexity and challenges of modern high performance radars [1]. Cognitive radar was described and introduced as an advanced form of radar systems for the first time by Haykin [2] to meet the challenges of increasingly complex operating environments. On the contrary to a conventional radar, a cognitive radar includes an adaptive transmitter in addition to an adaptive receiver, which entails a number of new adaptation and knowledge-aided methods.

15. SUBJECT TERMS
cognitive radar, convex optimization

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT:	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON (Monitor)
a. REPORT	b. ABSTRACT	c. THIS PAGE			Muralidhar Rangaswamy
Unclassified	Unclassified	Unclassified	SAR	39	19b. TELEPHONE NUMBER *(Include Area Code)* N/A

Chapter 1

Convex optimization for cognitive radar

Bosung Kang[1], Khaled AlHujaili[2], Muralidhar Rangaswamy[3], and Vishal Monga[4]

1.1 Introduction

A confluence of factors continues to increase the complexity and challenges of modern high performance radars [1]. Cognitive radar was described and introduced as an advanced form of radar systems for the first time by Haykin [2] to meet the challenges of increasingly complex operating environments. On the contrary to a conventional radar, a cognitive radar includes an adaptive transmitter in addition to an adaptive receiver, which entails a number of new adaptation and knowledge-aided methods. A fundamental goal of a cognitive radar is to sense, learn and adapt (SLA) to a complex environment [1]. The cognitive radar continuously learns about the environment and updates the receiver with relevant information on the environment. Then the transmitter continually adjusts its signal in intelligent manner based on the sensed environment such as the size and range of the targets and clutter. This closed-loop dynamic system is the key aspect of the cognitive radar. The basic structure of cognitive radar is shown in Figure 1.1. The learning process starts when the receiver collects the returns from the target and scatterers. From these returns, the cognitive radar system acquires the required information about its external environment. The transmitter uses the obtained information to alter its transmission and hence compensates for the changes in the environment that are captured through the receiver's previous interactions.

For a cognitive radar, both transmit and receive functions are utilized in new ways to enhance channel estimation and the radar optimizes a spatio-temporal transmit and receive strategy. This optimization process involves solving mathematical optimization problems. For example, it has been shown that the optimum transmit and receive functions maximize output signal-to-interference ratio (SINR), which means that it is required to solve an SINR maximization problem to obtain the optimum transmit and receive function. Such mathematical optimization problems may be convex and the solution can be efficiently obtained by numerical approaches or

[1]University of Dayton Research Institute, Dayton, OH, USA
[2]Taibah University, Al-Madinah, Saudi Arabia
[3]Air Force Research Laboratory, WPAFB, OH, USA
[4]The Pennsylvania State University, University Park, PA, USA

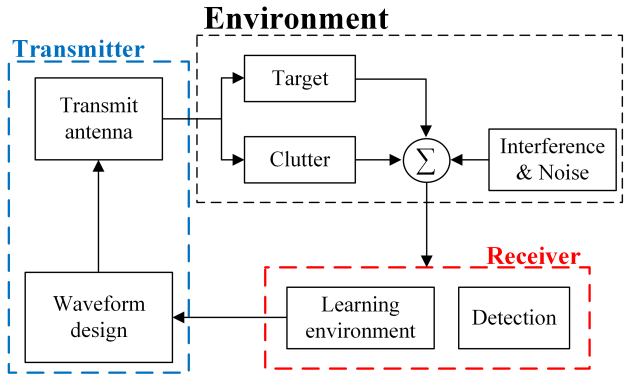


Figure 1.1 Basic structure of cognitive radar system

a closed form solution in some cases. However, the optimization problems are in general non-convex and it is hard to obtain a solution particularly when practical constraints are exploited in the optimization problems to ensure performance of the radar and satisfy the hardware requirements under complex environment.

In this chapter, we introduce several kinds of optimization problems that are widely and actively studied in cognitive radar applications and how the problems can be solved using mathematical skills and principles of convex optimization. We first provide importance and purposes of waveform design problems in cognitive radar and practical challenges in the waveform design problems, which is followed by the principles and fundamentals of convex optimization. Typical approaches to solve convex optimization problems and non-convex optimization problems are also introduced. Lastly, successful examples of waveform design algorithms that solve hard non-convex optimization problems using the principals of convex optimization are discussed.

1.1.1 Waveform Design Problems in Cognitive Radar

The basic principle in radar is to illuminate a certain region of interest by transmitting a radio frequency (RF) electromagnetic (EM) signal and receive its echo caused by an object of interest known as a *target* and other non of interest objects. Utilizing this echo along with the knowledge about the transmitted signal, the radar performs various functions such as detection, tracking, imaging, and classification. In addition to the target return, the received echo contains, as shown in Figure 1.2, signal-dependent returns from objects not of interest known as *clutters*, electromagnetic returns from other radiators known as *interference*, and *noise* [3].

1.1.1.1 Waveform Design: Background & Motivation

In the presence of those unwanted contributions, the ability to extract the target return or suppress the unwanted returns is encouraged to enhance the functionality of the radar systems. Towards this goal, during the last decades, many techniques

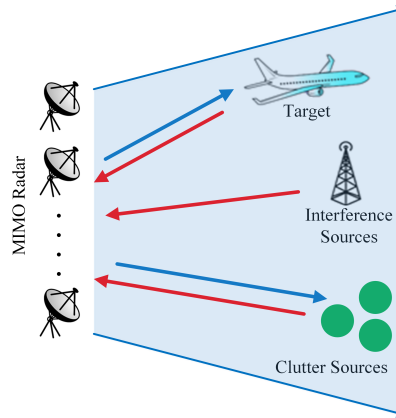


Figure 1.2 Crowded environment

have been proposed to adaptively suppress those aforementioned unwanted returns at the receiver side [4]. Examples include constant false alarm rate (CFAR) detectors, space-time adaptive processing (STAP), and rejecting range-ambiguous scatterer returns. On the other hand, the transmitter could be also incorporated into this task due to the dependence of the target and clutter returns on the transmit waveform. Accordingly, adjusting this waveform helps in reducing the effect of the unwanted returns, and hence leads to better extracting of the target return. This concept of adaptively adjusting the transmitted signal was proposed by H. Van Trees in 1965 [5] and is known in the literature of the radar signal processing as *the waveform design*. The concept of the waveform design can not be applied in a radar system without exploiting the knowledge about its surrounding *environment*. The surrounding environment includes all objects that contribute to the received signal and affect the performance of radar [4]. The knowledge about the environment is obtainable under the framework of the cognitive radar mentioned earlier.

1.1.1.2 Waveform Design via Optimization

The concept of waveform design has been enormously addressed in the literature of radar signal processing during the last decades for different applications, i.e., detection, estimation, and tracking. In general, the adaptive transmission technology can be divided into two main different categories according to the degree of the freedom of the transmitter. These categories are *selection* and *design* [4]. Under the selection category, we have the methods that either adaptively select pre-designed signals or select certain parameters of pre-defined signals such as pulse repetition frequency (PRF) and the signal pulse width. On the other hand, the design category includes the methods that either arbitrarily design the transmit waveform or design the aforementioned parameters of an existing waveform. The discussion presented in this chapter deals with the waveform design category. More specific, the proposed methods or algorithms here design suitable transmit waveforms to compensate for the changes in the radar's surrounding environment. Furthermore, it is assumed that no specific

structure for the transmit waveform. In other words, in the design process, each time sample in the desired waveform is considered to be freely available to be designed and constructed.

Mostly, this type design problems is conducted under the *numerical constrained optimization* paradigm where the radar performance metric of interest will be optimized as an *objective function* over the transmitted waveforms while considering some practical/hardware requirements that are imposed by the system as *constraints*. In general, the objective functions and the constraints depend on the task of the radar system, i.e., detection, estimation, and tracking tasks. In the following subsections, different common performance metrics and practical constraints will be discussed in more detail.

1.1.1.3 Radar Performance Metrics (Objective Functions)

As mentioned before, the transmit waveforms from the radar can be designed by solving optimization problems. The objective functions in these problems usually represent some performance measures as figure of merits. In this part, our goal is to highlight some of these merits.

Signal-to-interference-plus-noise ratio (SINR)

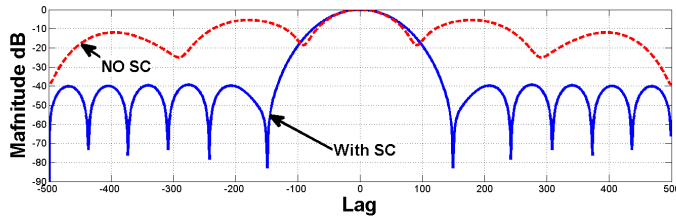
Most of the waveform design approaches aim to enhance the detection ability of radar systems. Enhancing this ability is equivalent to maximizing the radar signal-to-interference-plus-noise ratio (SINR). SINR has a proportional influence on the probability of detection, and maximizing SINR is equivalent to directly maximizing the probability of detection and hence the detection ability [6, 7, 8] and the references therein.

The transmit beampattern

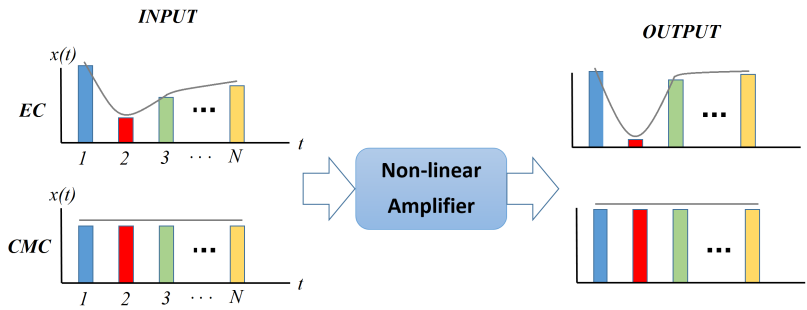
Under excessive clutter and/or interference disturbance the improvement in SINR can be indirectly achieved by considering different metrics. For instance, maximizing the energy of the target return (by focusing the transmit power to the expected target location) and reducing it for the unwanted returns has been conducted in the literature by controlling *the transmit beampattern* [9, 10, 11] and the references therein. The main idea here is to design a set of waveforms such that the transmitted beampattern matches certain specifications, e.g. a desired beampattern, by minimizing the deviation between the produced and the desired beampatterns.

The ambiguity function

Another indirect SINR improvement can be achieved by designing a transmit waveform that has a specific *ambiguity* shape [12, 11]. This design problem is known in literature as *the ambiguity function* (AF) shaping. The AF represents the range-Doppler response at the output of a filter matched to the transmitted signal when it arrives with a time delay (represents the range) and uncompensated Doppler shift. The transmit waveform has a significant impact on this response, and, hence, designing this waveform can be employed to control the ambiguity function for radar systems.



(a)



(b)

Figure 1.3 Necessity and importance of practical constraints on radar transmit waveform (a) the similarity constraint and (b) the constant modulus constraint

1.1.1.4 Practical Constraints in the Waveform Design Process

Furthermore, besides improving the performance metric using the optimal waveform, many hardware limitations imposed by system components must be considered during the design process. These limitations are reflected in the design process as constraints on the designed waveform. Many practical constraints are used in literature. Salient examples include peak-to-average-power-ratio (PAPR) and energy constraints (EC), similarity constraint (SC), spectral constraint (SpecC), and constant modulus constraint (CMC). In the following, we provide brief descriptions of some of the intensively used/studied constraints.

Peak-to-average-power-ratio (PAPR) and the energy constraint (EC)

Peak-to-average-power-ratio (PAPR) and energy constraints belong to the family of the modulus constraints [4] and usually are imposed on the transmitted waveform to maximize the efficiency of the transmitter hardware [13, 14]. Since the constant modulus constraint results in a hard non-convex optimization problem, PAPR and the EC are employed as a relaxation of the CMC in the literature.

Similarity constraint (SC)

The similarity constraint (SC) is used to produce a waveform that has some of the desirable properties of a reference signal. The advantage of imposing this constraint has been reported in different works in the literature such as in [15, 8]. In these

works, without enforcing SC the produced waveforms suffer from undesirable effects in pulse compression, and ambiguity function properties as shown in Figure 1.3a.

Spectral constraint (SpecC)
This constraint has been introduced in the literature of radar signal processing to ensure the co-existence among radar and communication systems, in a spectrally crowded environment [16, 17]. The need for this co-existence appears when both systems occupy the same frequency band. For an instant, in wideband transmission, radar systems occupy a large bandwidth and hence overlap with the spectrum for other radiators could arise. The term co-existence implies that the frequency signature of the transmit waveforms from the radar system exhibits nulls in the frequency bands of the communication systems.

Orthogonality constraint
Imposing orthogonality across antennas has been shown to be particularly meritorious. Orthogonal MIMO waveforms enable the radar system to achieve an increased virtual array and, hence leads to many practical benefits [18, 19]. A compelling practical challenge is that the "directional knowledge" of target and interference sources utilized in specifying the desired beampattern may not be perfect. In such scenarios, it has been shown in [18, 19] that the gain loss in the transmit-receive patterns for orthogonal waveform transmission is very small under target direction mismatch.

Constant modulus constraint (CMC)
The constant modulus constraint (CMC) aims to ensure that the waveform's envelope have a constant amplitude. The CMC is crucial in the design process due to the presence of non-linear amplifiers in radar systems [6]. These components are known to operate in saturation mode, and CMC is required to maximize their efficiency. Due to this saturation mode, radar transmitters have a peak power level that cannot be exceeded. Therefore, if the peak amplitude of the waveform exceeds this upper level, it will be clipped, and hence the transmitted power will not be fully utilized as shown in Figure 1.3b. Consequently, less than the expected power is carried to the target, and thus system performance will be degraded.

1.2 Background and Motivation

1.2.1 Principles of Convex Optimization

1.2.1.1 Convex Optimization Problem

Constrained optimization problem
Many estimation and design problems in signal processing, particularly, radar applications, can be posed as a constrained optimization problem which has the form

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & f(\boldsymbol{x}) \\ \text{subject to} & g_i(\boldsymbol{x}) \leq 0 \\ & h_i(\boldsymbol{x}) = 0 \end{array} \tag{1.1}$$

where the vector $\boldsymbol{x}$ is the optimization variable of the problem, the function f is the objective function or the cost function we desire to minimize, the functions g_i are the inequality constraint functions, and the functions h_i are the equality constraint functions. The optimal solution of the optimization problem, $\boldsymbol{x}^\star$, is a vector that achieves the smallest objective value of the cost function among all vectors that satisfy the constraints. $\boldsymbol{x}^\star$ can be expressed by

$$\boldsymbol{x}^\star = \arg\min_{\boldsymbol{x}\in\left\{\{\boldsymbol{x}|g_i(\boldsymbol{x})\leq 0\}\cap\{\boldsymbol{x}|h_i(\boldsymbol{x})=0\}\right\}} f(\boldsymbol{x}) \tag{1.2}$$

where the set $\left\{\{\boldsymbol{x}|g_i(\boldsymbol{x}) \leq 0\} \cap \{\boldsymbol{x}|h_i(\boldsymbol{x}) = 0\}\right\}$ is called the constraint set or the domain of the problem. In general, the objective function f is a non-convex function and the problem may have many local optima. Therefore, it is challenging to obtain the global optimal solution of the problem using numerical algorithms or even impossible to find the solution. A specific class of optimization problems is called convex optimization problems if the cost function f and the inequality constraint functions g_i are convex and the equality constraint functions h_i are affine. For convex optimization problems, any local optimum of the problem is the global minimum for convex optimization problems.

Convex sets

A set S is called an affine set if it contains the line through any two points in the set. If S is an affine set, for any two points $\boldsymbol{x}$ and $\boldsymbol{y}$ in the set S, every linear combination of them is in the set S.

$$\boldsymbol{x},\boldsymbol{y} \in S \;\Rightarrow\; \lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y} \in S \;\text{ for }\; \lambda \in \mathbb{R} \tag{1.3}$$

In addition, a set S is convex if it contains every line segment between any two points in the set. In other words, if S is a convex set, for any two points $\boldsymbol{x}$ and $\boldsymbol{y}$ in the set S,

$$\boldsymbol{x},\boldsymbol{y} \in S \;\Rightarrow\; \lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y} \in S \;\text{ for }\; 0 \leq \lambda \leq 1 \tag{1.4}$$

The weighted average shown in (1.4) can be generalized to more than two points. We refer to a point of the form $\lambda_1\boldsymbol{x}_1 + \lambda_2\boldsymbol{x}_2 + \cdots + \lambda_k\boldsymbol{x}_k$, where $\lambda_1 + \lambda_2 + \cdots + \lambda_k = 1$ and $\lambda_i \geq 0$ for all i, to a convex combination. Then it can be also shown that a set is convex if and only if it contains every convex combination of points in the set. We can easily know that an affine set is a subset of a convex set and every affine set is a convex set.

The convex hull is defined on a set S as the set of all convex combinations of points in the set S.

$$\textbf{conv}\, S = \{\lambda_1\boldsymbol{x}_1 + \lambda_2\boldsymbol{x}_2 + \cdots + \lambda_k\boldsymbol{x}_k | x_i \in S, \lambda_i \geq 0, \lambda_1 + \lambda_2 + \cdots + \lambda_k = 1\} \tag{1.5}$$

The convex hull is always convex and it is the smallest convex set that contains S. For example, if S is a set of two points, **conv** S is a line segment between the two points. If S is a set of three points, **conv** S is a triangle that is formed by the three points. If a set C that contains S is given, then **conv** $S \subseteq C$.

There are some more examples of convex sets that are widely used in the optimization problems. A hyperplane defined by $\{\boldsymbol{x}|\boldsymbol{a}^T\boldsymbol{x} = b\}$ is affine and convex. A euclidean ball with the center at $\boldsymbol{x}_c$ and the radius r given by

$$B(\boldsymbol{x}_c, r) = \{\|\boldsymbol{x} - \boldsymbol{x}_c\|_2 \leq r\} \tag{1.6}$$

is a special case of a convex ellipsoid which has the form $\{\boldsymbol{x}|(\boldsymbol{x} - \boldsymbol{x}_c)^T\boldsymbol{A}^{-1}(\boldsymbol{x} - \boldsymbol{x}_c) \leq 1\}$ where $\boldsymbol{A}$ is symmetric and positive definite. A norm ball is one of the mostly used convex set in the optimization problem, which is defined by

$$\{\boldsymbol{x}|\|\boldsymbol{x} - \boldsymbol{x}_c\| \leq r\} \tag{1.7}$$

Note that a kind of norm is not specified. Any function that is positive definite and absolutely homogeneous and satisfies the triangle inequality is a norm.

Convexity is preserved under intersection and an affine transformation, namely, the intersection of convex sets is convex and the image and inverse image of a convex set on an affine function is convex. For example, a polyhedron which is defined as the solution set of a finite number of linear equalities and inequalities, in other words, the intersection of a finite number of halfspaces and hyperplanes, $\{\boldsymbol{x}|\boldsymbol{A}\boldsymbol{x} \preceq \boldsymbol{b}, \boldsymbol{C}\boldsymbol{x} = \boldsymbol{d}\}$, is a convex set.

Convex functions

A function f is convex if the domain of f is a convex set and

$$f\big(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}\big) \leq \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) \tag{1.8}$$

for all $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{dom}\ f$ and $0 \leq \lambda \leq 1$. This inequality geometrically means that the line segment between $\big(\boldsymbol{x}, f(\boldsymbol{x})\big)$ and $\big(\boldsymbol{y}, f(\boldsymbol{y})\big)$ lies above the graph of f. For example, an affine function always hold the equality in (1.8) and therefore all affine functions are convex. If $-f$ is convex, f is called a concave function. An affine function is also a concave function and any function that is convex and concave is affine.

A convex function can be defined using the α-sublevel set of a function f which is defined as

$$S_\alpha \triangleq \{\boldsymbol{x} \in \mathbf{dom}\ f \mid f(\boldsymbol{x}) \leq \alpha\} \tag{1.9}$$

If f is convex, then its sublevel sets S_α are convex sets for any value of α. Note that the converse is not true. For example, a non-convex monotonically decreasing function has convex sublevel sets. In addition, if a function is differentiable or twice differentiable, in other words, its gradient or Hessian exists at each point in $\mathbf{dom}\, f$, then f is convex if and only if $\mathbf{dom}\, f$ is convex and

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T(\boldsymbol{y} - \boldsymbol{x}) \qquad \text{for all } \boldsymbol{x}, \boldsymbol{y} \in \mathbf{dom}\, f \tag{1.10}$$

$$\nabla^2 f(\boldsymbol{x}) \succeq 0 \qquad \text{for all } \boldsymbol{x} \in \mathbf{dom}\, f \tag{1.11}$$

Note that the right term of (1.10) represents the first-order Taylor approximation of f at $\boldsymbol{x}$ and the inequality (1.10) implies that the first-order Taylor approximation of a convex function is a global underestimator of the function. The inequality (1.11) shows that the Hessian of a convex function is positive semidefinite.

Convex optimization problem

For the standard form of the optimization problem in (1.1), an optimization problem is a convex optimization problem if the objective function f and the inequality constraint functions g_i are convex and the equality constraint function h_i are affine. Therefore, the standard form of the convex optimization problem can be given as

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & f(\boldsymbol{x}) \\ \text{subject to} & g_i(\boldsymbol{x}) \leq 0 \\ & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{array} \tag{1.12}$$

Since the equality constraint functions are affine, the equality constraint can be expressed by $h_i(\boldsymbol{x}) = \boldsymbol{a}_i^T\boldsymbol{x} - b_i$ and rewritten as $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ by stacking all the equality constraints. Convex optimization problems facilitate solving optimization problems since any local optimum of the convex optimization problem is necessarily a global optimum. Therefore, there are many efficient numerical solution methods available that can handle very large problems with high dimensional variables and a number of constraints.

Some examples of canonical convex optimization problems are useful to introduce. When the objective function and constraint functions are all affine, the problem is called a linear programming (LP) which has the form

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & \boldsymbol{c}^T\boldsymbol{x} \\ \text{subject to} & \boldsymbol{G}\boldsymbol{x} \preceq \boldsymbol{h} \\ & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{array} \tag{1.13}$$

where $\boldsymbol{G} \in \mathbb{R}^{m\times n}$ and $\boldsymbol{A} \in \mathbb{R}^{p\times n}$, m is the number of the inequality constraints, p is the number of the equality constraints, and n is the number of the optimization variables. The feasible set of the LP forms a polyhedron. Linear programmings are applicable to a number of fields and applications such as finding Chebyshev center of a polyhedron and dynamic activity planning.

A quadratically constrained quadratic program (QCQP) is also a kind of convex optimization problems that is widely used in applications. It has a convex quadratic objective function and convex quadratic inequality constraint functions.

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & \boldsymbol{x}^T\boldsymbol{R}\boldsymbol{x} + \boldsymbol{c}^T\boldsymbol{x} + d \\ \text{subject to} & \boldsymbol{x}^T\boldsymbol{P}_i\boldsymbol{x} + \boldsymbol{q}_i^T\boldsymbol{x} + r_i \leq \boldsymbol{0} \qquad i = 1,\ldots,m \\ & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{array} \tag{1.14}$$

where $\boldsymbol{R}$ and $\boldsymbol{P}_i$ are all positive semidefinite matrices. In QCQP, each inequality constraint forms an ellipsoid. In the case that $\boldsymbol{P}_i = 0$, the optimization problem becomes a quadratic program (QP) which has a convex quadratic objective function and affine constraint functions. It also includes linear programs when $\boldsymbol{R} =$ and $\boldsymbol{P}_i = 0$. The examples of the quadratic program include least squares approximation and Markowitz portfolio optimization which is a classical portfolio problem.

A semidefinite programming (SDP) has the form

$$\begin{array}{ll} \min\limits_{\boldsymbol{X}} & \mathrm{tr}(\boldsymbol{R}\boldsymbol{X}) \\ \text{subject to} & \mathrm{tr}(\boldsymbol{P}_i\boldsymbol{X}) \leq 0 \qquad i = 1,\ldots,m \\ & \mathrm{tr}(\boldsymbol{Q}_i\boldsymbol{X}) = 0 \qquad i = 1,\ldots,p \\ & \boldsymbol{X} \succeq 0 \end{array} \tag{1.15}$$

where $\boldsymbol{X} \succeq 0$ represents that $\boldsymbol{X}$ is a positive semidefinte matrix. Note that it has a non-negativity constraint in addition to a linear objective function and linear constraint functions of $\boldsymbol{X}$. Many optimization problems including a matrix norm minimization problem, moment problems, and a fastest mixing Markov chain problem can be casted to a semidefinite programming.

1.2.1.2 Solving Convex Optimization Problems

There are two main categories for solving convex optimization problems, *i.e.,* obtaining the optimal solution of optimization problems, analytical approaches and numerical approaches. Analytical approaches enable a closed form solution but are not available to all the optimization problems. For those problems where an analytical approach is not available, a solution can be found via numerical approaches.

Analytical approaches

Lagrangian associated with the problem (1.1) is defined as

$$L(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\nu}) = f(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i g_i(\boldsymbol{x}) + \sum_{i=1}^{p} \nu_i h_i(\boldsymbol{x}) \tag{1.16}$$

where $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$ are the Lagrange multipliers. The Lagrange dual function is the minimum value of the Lagrangian over $\boldsymbol{x}$

$$g(\boldsymbol{\lambda},\boldsymbol{\nu}) = \inf_{\boldsymbol{x}\in D} L(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\nu}) \tag{1.17}$$

where D is the domain of the optimization problem. The dual function gives the information about the optimal value $p^\star$ of the problem. For any $\boldsymbol{\lambda} \succeq 0$ and any $\boldsymbol{\nu}$,

$$g(\boldsymbol{\lambda},\boldsymbol{\nu}) \leq p^\star \tag{1.18}$$

The inequality implies that the Lagrange dual function is a lower bound on the optimal value $p^\star$. Then the best lower bound of the optimum value can be obtained by maximizing the Lagrangian dual function.

$$\begin{array}{ll} \max & g(\boldsymbol{\lambda},\boldsymbol{\nu}) \\ \text{subject to} & \boldsymbol{\lambda} \succeq 0 \end{array} \tag{1.19}$$

This is the Lagrange dual problem of (1.1). It is always a convex optimization problem regardless of convexity of (1.1) since the dual function is concave and the constraint is convex. Let $d^\star$ denote the optimal value of (1.19). The following weak duality inequality always hold even if the original problem is not convex.

$$d^\star \leq p^\star \tag{1.20}$$

In the case that the problem is convex and there exists a strictly feasible point, strong duality holds.

$$d^\star = p^\star \tag{1.21}$$

This means if the dual optimal points $\boldsymbol{\lambda}^\star$ and $\boldsymbol{\nu}^\star$ can be obtained the optimality conditions of the optimization problem can be derived. Let $\boldsymbol{x}^\star$ be the optimal point of the optimization problem (1.1). If strong duality holds,

$$\begin{aligned} f(\boldsymbol{x}^\star) &= g(\boldsymbol{\lambda}^\star, \boldsymbol{\nu}^\star) && (1.22)\\ &\leq f(\boldsymbol{x}^\star) + \sum_{i=1}^{m} \lambda_i^\star g_i(\boldsymbol{x}^\star) + \sum_{i=1}^{p} \nu_i^\star h_i(\boldsymbol{x}^\star) && (1.23)\\ &\leq f(\boldsymbol{x}^\star) && (1.24) \end{aligned}$$

From the equality and inqualities above, the Karush-Kuhn-Tucker (KKT) conditions are derived as following.

$$g_i(\boldsymbol{x}^\star) \leq 0 \qquad \text{Primal feasibility} \tag{1.25}$$

$$h_i(\boldsymbol{x}^\star) = 0 \qquad \text{Primal feasibility} \tag{1.26}$$

$$\lambda_i^\star \geq 0 \qquad \text{Dual feasibility} \tag{1.27}$$

$$\lambda_i^\star g_i(\boldsymbol{x}^\star) = 0 \qquad \text{Complementary slackness} \tag{1.28}$$

$$\nabla f(\boldsymbol{x}^\star) + \sum_{i=1}^{m} \lambda_i^\star \nabla g_i(\boldsymbol{x}^\star) + \sum_{i=1}^{p} \nu_i^\star \nabla h_i(\boldsymbol{x}^\star) = 0 \qquad \text{Gradient condition} \tag{1.29}$$

If the objective function and constraint functions of the problem are differentiable and strong duality holds, the optimal and the dual optimal points must satisfy the KKT conditions. Moreover, the KKT conditions are sufficient for convex optimization problems, which means any points $\boldsymbol{x}$, $\boldsymbol{\lambda}$, and $\boldsymbol{\nu}$ that satisfy the KKT conditions are primal and dual optimal. One simple example of the optimization problem that can be solved by the KKT conditions is an equality constrained convex quadratic minimization problem which is given by

$$\begin{array}{ll} \text{min} & (1/2)\boldsymbol{x}^T\boldsymbol{P}\boldsymbol{x} + \boldsymbol{q}^T\boldsymbol{x} + r \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{array} \tag{1.30}$$

The KKT conditions for this problem are

$$\boldsymbol{A}\boldsymbol{x}^\star = \boldsymbol{b}, \qquad \boldsymbol{P}\boldsymbol{x}^\star + \boldsymbol{q} + \boldsymbol{A}^T\boldsymbol{\nu}^\star = 0 \tag{1.31}$$

$\boldsymbol{x}^\star$ and $\boldsymbol{\nu}^\star$ can be obtained by solving the following linear equation

$$\begin{bmatrix} \boldsymbol{P} & \boldsymbol{A}^T \\ \boldsymbol{A} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}^\star \\ \boldsymbol{\nu}^\star \end{bmatrix} = \begin{bmatrix} -\boldsymbol{q} \\ \boldsymbol{b} \end{bmatrix} \tag{1.32}$$

Numerical Approaches

The KKT conditions enable an analytical solution of optimization problems, however, they are applicable to the limited number of convex optimization problems. In

most cases, it is required to find the solution by numerical algorithms. First consider a simple unconstrained convex optimization problem,

$$\min_{\boldsymbol{x}} f(\boldsymbol{x}) \tag{1.33}$$

A necessary and sufficient condition for the optimal point $\boldsymbol{x}^\star$ is $\nabla f(\boldsymbol{x}^\star) = 0$. Starting from an initial point $\boldsymbol{x}_0$, a sequence of $\boldsymbol{x}_n$ is selected by

$$\boldsymbol{x}_{n+1} = \boldsymbol{x}_n + t\boldsymbol{d} \tag{1.34}$$

where t and $\boldsymbol{d}$ are the step size and the search direction, respectively. This is called a descent method if $f(\boldsymbol{x}_{n+1}) < f(\boldsymbol{x}_n)$ holds for every pair of $\boldsymbol{x}_n$ and $\boldsymbol{x}_{n+1}$ in the sequence. For any decent methods for the convex objective function, the search direction $\boldsymbol{d}$ must be a descent direction which satisfies

$$\nabla f(\boldsymbol{x})^T \boldsymbol{d} < 0 \tag{1.35}$$

The step size t can be determined by the exact line search or the inexact line search methods, for example, the backtracking line search. Different choice of the search direction for the descent methods results in various kinds of the descent methods. A descent method which takes the negative gradient as the search direction is called the gradient descent method. For the gradient method, the search direction is chosen by

$$\boldsymbol{d} = -\nabla f(\boldsymbol{x}) \tag{1.36}$$

The gradient descent method shows approximately linear convergence and the convergence rate highly depends on the condition number of the Hessian. Though it is such a simple descent method, the convergence is very slow with the high condition number of the Hessian.

A descent method with a fixed step size $t = 1$ and the search direction

$$\boldsymbol{d} = -\nabla^2 f(\boldsymbol{x})^{-1} \nabla f(\boldsymbol{x}) \tag{1.37}$$

is called Newton's method. It is motivated by the fact that $\boldsymbol{d}$ in (1.37) is the minimizer of the second-order approximation of f at $\boldsymbol{x}$, *i.e.*,

$$\boldsymbol{d} = \arg\min \hat{f}(\boldsymbol{x}+\boldsymbol{d}) = \arg\min f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T \boldsymbol{d} + \frac{1}{2}\boldsymbol{d}^T \nabla^2 f(\boldsymbol{x}) \boldsymbol{d} \tag{1.38}$$

Newton's method shows faster convergence than the gradient descent method and it converges quadratically near $\boldsymbol{x}^\star$. However, it involves computation of the Hessian and the Newton step at every iteration, which requires solving a set of linear equations.

An equality constrained minimization problem can be solved by the decent methods for unconstrained minimization problems described above after eliminating the equality constraint and simplified to an unconstrained optimization problem. Consider the following equality constrained minimization problem.

$$\begin{array}{ll} \min & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \end{array} \tag{1.39}$$

After finding a particular solution $\hat{\boldsymbol{x}}$ of the equality constraint $\boldsymbol{Ax} = \boldsymbol{b}$, the equivalent unconstrained optimization problem can be written by

$$\min g(\boldsymbol{z}) = f(\boldsymbol{Fz} + \hat{\boldsymbol{x}}) \tag{1.40}$$

where $\boldsymbol{F} \in \mathbb{R}$ is a matrix whose range is the nullspace of $\boldsymbol{A}$, *i.e.*, $\boldsymbol{AF} = \boldsymbol{0}$. This problem can be solved by the descent methods for unconstrained optimization problems. After obtaining $\boldsymbol{z}^\star$, the solution of (1.39) can be easily found by

$$\boldsymbol{x}^\star = \boldsymbol{Fz}^\star + \hat{\boldsymbol{x}} \tag{1.41}$$

Newton's method for the inequality constrained minimization problem is also available. Taking the second-order Taylor approximation of (1.39) yields

$$\begin{array}{ll} \min & \hat{f}(\boldsymbol{x}+\boldsymbol{d}) = f(\boldsymbol{x}) + \nabla f(\boldsymbol{x})^T \boldsymbol{d} + \frac{1}{2}\boldsymbol{d}^T \nabla^2 f(\boldsymbol{x}) \boldsymbol{d} \\ \text{subject to} & \boldsymbol{A}(\boldsymbol{x}+\boldsymbol{d}) = \boldsymbol{b} \end{array} \tag{1.42}$$

Recall that the search direction is the minimizer of the above problem and since it is an equality constrained quadratic minimization problem as (1.30), $\boldsymbol{d}$ can be obtained by the KKT condition,

$$\begin{bmatrix} \nabla^2 f(\boldsymbol{x}) & \boldsymbol{A}^T \\ \boldsymbol{A} & \boldsymbol{0} \end{bmatrix} \begin{bmatrix} \boldsymbol{d} \\ \boldsymbol{v} \end{bmatrix} = \begin{bmatrix} -\nabla f(\boldsymbol{x}) \\ \boldsymbol{0} \end{bmatrix} \tag{1.43}$$

Lastly, a key idea for solving inequality constrained minimization problems is approximately to formulate an inequality constrained problem which can be solved by the methods described above. Consider the following equality constrained optimization problem

$$\begin{array}{ll} \min & f(\boldsymbol{x}) + \sum_{i=1}^{m} I_-(g_i(\boldsymbol{x})) \\ \text{subject to} & \boldsymbol{Ax} = \boldsymbol{b} \end{array} \tag{1.44}$$

where I_- is the indicator function of $\mathbb{R}_-$ such that $I_-(u) = 0$ if $u \leq 0$ and $I_-(u) = \infty$ otherwise. Though (1.44) has no equality constraint and it is equivalent to (1.12), Newton's method can not be applied since $I_-(u)$ is not differentiable. The indicator function can also be approximated by the logarithmic barrier function which is differentiable and given by

$$\hat{I}(u) = -(1/t)\log(-u) \tag{1.45}$$

Then the inequality constrained problem can be approximated by the following equality constrained problem

$$\begin{array}{ll} \min & f(\boldsymbol{x}) - (1/t)\sum_{i=1}^{m} \log(-g_i(\boldsymbol{x})) \\ \text{subject to} & \boldsymbol{Ax} = \boldsymbol{b} \end{array} \tag{1.46}$$

This problem is convex and the objective function is differentiable. Therefore, it can be solved by Newton's method.

1.2.1.3 Approaches for Non-convex Optimization Problems

Convex problems can be solved by the analytical and numerical approaches described in the previous sections, however, many of the optimization problems in practice are non-convex. In general, it is difficult to find a solution of the non-convex

optimization problem since it has a lot of local minima and maxima and saddle points. We introduce mathematical skills that help to find optimal or sub-optimal solution of the non-convex problem, convex relaxation/approximation and transform of variables with examples.

A compressive sensing signal recovery problem is one of the most actively studied area in signal processing applications. The optimization problem is given by an $l-0$ norm minimization problem,

$$\hat{\boldsymbol{x}} = \min_{\boldsymbol{x}} \|\boldsymbol{x}\|_0 \qquad \text{subject to} \qquad \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} \tag{1.47}$$

Since the objective function $\|\boldsymbol{x}\|_0$ is non-convex, it can be relaxed as $l-1$ norm which is convex

$$\hat{\boldsymbol{x}} = \min_{\boldsymbol{x}} \|\boldsymbol{x}\|_1 \qquad \text{subject to} \qquad \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} \tag{1.48}$$

This problem is a convex optimization problem. Though the objective function is still not differentiable at $\boldsymbol{x} = \boldsymbol{0}$, it can be solved by efficient solvers such as the interior point method and least angle regression (LARS). In radar applications, the CMC is used in many waveform design problems. The SINR maximization subject to the CMC is given by

$$\begin{array}{ll} \min & \boldsymbol{x}^H \boldsymbol{\Phi} \boldsymbol{x} \\ \text{subject to} & |\boldsymbol{x}| = \boldsymbol{1}/n \end{array} \tag{1.49}$$

The constant modulus constraint is a non-convex constraint and very hard to exploit in the optimization problem. The energy constraint is used as a relaxation of the CMC in many problems

$$\begin{array}{ll} \min & \boldsymbol{x}^H \boldsymbol{\Phi} \boldsymbol{x} \\ \text{subject to} & \|\boldsymbol{x}\|_2 = 1 \end{array} \tag{1.50}$$

A non-convex problem can be converted to a convex problem by transformation variables. Consider a geometric programming which is given by

$$\begin{array}{ll} \min & \sum_{k=1}^{K_0} e^{\boldsymbol{a}_{0_k}^T \boldsymbol{x} + b_{0_k}} \\ \text{subject to} & \sum_{k=1}^{K_i} e^{\boldsymbol{a}_{i_k}^T \boldsymbol{x} + b_{i_k}} \leq 1, \qquad i = 1, \ldots, m \\ & e^{\boldsymbol{g}_i^T \boldsymbol{x} + h_i} = 1, \qquad i = 1, \ldots, p \end{array} \tag{1.51}$$

This problem is non-convex because of non-convexity of the equality constraint. Taking the logarithm on the objective and constraint functions gives

$$\begin{array}{ll} \min & \log\left(\sum_{k=1}^{K_0} e^{\boldsymbol{a}_{0_k}^T \boldsymbol{x} + b_{0_k}}\right) \\ \text{subject to} & \log\left(\sum_{k=1}^{K_i} e^{\boldsymbol{a}_{i_k}^T \boldsymbol{x} + b_{i_k}}\right) \leq 0, \qquad i = 1, \ldots, m \\ & \boldsymbol{g}_i^T \boldsymbol{x} + h_i = 0, \qquad i = 1, \ldots, p \end{array} \tag{1.52}$$

The objective and equality constraint functions are convex and the equality constraint function is affine, and therefore the problem is convex.

1.2.2 Challenges of Optimization Problems for Cognitive Radar

Convex optimization has a long history in signal processing applications including control, circuit design, economics and finance, statistics and machine learning, and radar applications [20] and has emerged as a major signal processing tool that has made a significant impact on numerous problems previously considered intractable [21]. The classical simplex algorithm was developed during the World War II by Dantzig to solve linear programming problems. It begins at a starting extreme point and moves along the edges of the feasible region until it reaches the vertex of the optimal solution. It showed a great improvement over earlier methods, however, it takes a long time to converge and the computational complexity is exponential time at the worst case [22]. In 1980s, Karmarka developed Karmarka's algorithm [23] which has been proven to be four times faster than the simplex method, particularly in polynomial time, by reinventing the interior-point method. It reaches a best solution by traversing the interior of the feasible region contrary to the simplex method. This recognition of the interior-point methods stimulated huge interest in new classes of convex optimization problems such as semidefinite programs and second-order cone programs and enabled to solve the problems as easily as linear programs [24]. The recent advances in processor power dramatically reduce solution time and even accelerated usage of convex optimization in numerous applications.

Convex optimization also benefits cognitive radar in many ways and has been applied to essential problems for cognitive radar, including but not limited to target detection and estimation, channel estimation, waveform design, beampattern design, and resource allocation. In practice, however, the optimization problems become more challenging when practical constraints induced by the environment and radar physics are exploited in the optimization problems. For example, the waveform design problem that maximizes a radar output SINR under the EC, $\boldsymbol{x}^H\boldsymbol{x} = 1$, can be easily solved by the eigenvalue problem and the optimal waveform is given as the eigenvector corresponding to the greatest eigenvalue of the SINR matrix [25]. However, the waveform optimization problem becomes onerous and the solution is difficult to obtain by conventional numerical approaches when the CMC which is practically essential is employed in the optimization problem. Additional constraints such as the SC, the SpecC, and the interference constraint make the optimization problem even more challenging.

1.3 Constrained Optimization for Cognitive Radar

In this section, we introduce three successful waveform design algorithms that solve a non-convex optimization problems that optimize SINR, the beampattern and the AF under the challenging constraints. They exploit the CMC in common and have different additional constraint depending on the purpose of waveform design. Though the CMC can not be achievable by conventional numerical approaches, the solutions of these algorithms approach the constant modulus in iterative ways where a convex optimization problem is solved at each iteration step. It has been proven that the solutions converge and achieve the CMC at convergence.

1.3.1 SINR Maximization

We consider a collocated narrow band MIMO radar system with N_T transmit antennas and N_R receive antennas. The received signal in this model is given by

$$\boldsymbol{r} = \alpha_0 \boldsymbol{U}(\theta_0)\boldsymbol{x} + \sum_{k=1}^{K} \alpha_k \boldsymbol{U}(\theta_k)\boldsymbol{x} + \boldsymbol{n} \tag{1.53}$$

where n is a circular complex Gaussian noise vector with zero mean and covariance matrix $\sigma_2 \boldsymbol{I}$, α_0, α_k, θ_0 and θ_k denote the complex amplitudes and the angle of the target and the k-th clutter source, respectively and $\boldsymbol{U}(\theta)$ is the steering matrix of a uniform linear array (ULA) antenna with half-wavelength separation between the antennas. The most common criterion in waveform design involves SINR maximization, which involves joint optimization of the transmit waveform and the receive filter. In particular, the receive filter is assumed to be a linear finite impulse response filter $\boldsymbol{w}$. The output is given by

$$r_f = \boldsymbol{w}^H \boldsymbol{r} = \alpha_0 \boldsymbol{w}^H \boldsymbol{U}(\theta_0)\boldsymbol{x} + \sum_{k=1}^{K} \alpha_k \boldsymbol{w}^H \boldsymbol{U}(\theta_k)\boldsymbol{x} + \boldsymbol{w}^H \boldsymbol{n} \tag{1.54}$$

Then, the output SINR can be expressed as

$$\text{SINR} = \frac{\sigma |\boldsymbol{w}^H \boldsymbol{U}(\theta_0)\boldsymbol{x}|^2}{\boldsymbol{w}^H \boldsymbol{\Sigma}(\boldsymbol{x})\boldsymbol{w} + \boldsymbol{w}^H \boldsymbol{w}} \tag{1.55}$$

where $\sigma = E[|\alpha_0|^2]/\sigma_n^2$, $\boldsymbol{\Sigma}(\boldsymbol{x}) = \sum_{k=1}^{K} I_k \boldsymbol{U}(\theta_k)\boldsymbol{x}\boldsymbol{x}^H \boldsymbol{U}^H(\theta_k)$ and $I_k = E[|\alpha_k|^2]/\sigma_n^2$.

1.3.1.1 Problem Formulation

The objective is to design the optimal waveform which maximizes the SINR subject to the CMC and the SC, *i.e.*, to solve the following optimization problem

$$\begin{array}{ll} \max\limits_{\boldsymbol{w},\boldsymbol{x}} & \frac{\sigma |\boldsymbol{w}^H \boldsymbol{U}(\theta_0)\boldsymbol{x}|^2}{\boldsymbol{w}^H \boldsymbol{\Sigma}(\boldsymbol{x})\boldsymbol{w} + \boldsymbol{w}^H \boldsymbol{w}} \\ \text{subject to} & \|\boldsymbol{x} - \boldsymbol{x}_0\|_\infty \leq \varepsilon \\ & |\boldsymbol{x}(k)| = 1/\sqrt{N_T N} \end{array} \tag{1.56}$$

It is a joint problem with respect to $\boldsymbol{x}$ and $\boldsymbol{w}$, however, it is separable and an unconstrained optimization problem of $\boldsymbol{w}$. The optimal $\boldsymbol{w}$ can be obtained from the well-known MVDR problem [26]. By substituting the optimal $\boldsymbol{w}$ into (1.50), we obtain an equivalent problem

$$\begin{array}{ll} \max\limits_{\boldsymbol{x}} & \boldsymbol{x}^H \boldsymbol{\Phi}(\boldsymbol{x})\boldsymbol{x} \\ \text{subject to} & \|\boldsymbol{x} - \boldsymbol{x}_0\|_\infty \leq \varepsilon \\ & |\boldsymbol{x}(k)| = 1/\sqrt{N_T N} \end{array} \tag{1.57}$$

where $\boldsymbol{\Phi}(\boldsymbol{x}) = \boldsymbol{U}^H(\theta_0)[\boldsymbol{\Sigma}(\boldsymbol{x}) + \boldsymbol{I}]^{-1}\boldsymbol{U}(\theta_0)$. Since the dependence of $\boldsymbol{\Phi}(\boldsymbol{x})$ on the waveform $\boldsymbol{x}$ makes the optimization problem onerous, it has been solved iteratively assuming $\boldsymbol{\Phi}(\boldsymbol{x}) = \boldsymbol{\Phi}$ for a fixed $\boldsymbol{x}$ and repeatedly optimizing $\boldsymbol{x}$ with a new $\boldsymbol{\Phi}$ till convergence [27, 15, 8]. However, even for a fixed $\boldsymbol{\Phi}$, the optimization of $\boldsymbol{x}$ is a hard non-convex problem for which one of approaches to solve the problem involves SDR with randomization [28, 15].

1.3.1.2 Successive QCQP Refinement

The optimization problem (1.56) with signal independent clutter (*i.e.*, $\boldsymbol{\Phi}(\boldsymbol{x}) = \boldsymbol{\Phi}$) is equivalent to the following non-convex problem

$$\begin{array}{ll} \max\limits_{\boldsymbol{x}} & \boldsymbol{x}^H(\boldsymbol{\Phi} - \lambda \boldsymbol{I})\boldsymbol{x} \\ \text{subject to} & \arg \boldsymbol{x}(k) \in [\gamma_k, \gamma_k + \delta] \\ & |\boldsymbol{x}(k)| = 1/\sqrt{N_T N} \end{array} \tag{1.58}$$

where λ is a constant greater than the largest eigenvalue of $\boldsymbol{\Phi}$ so that $\boldsymbol{\Phi} - \lambda \boldsymbol{I}$ is negative semidefinite. Due to the CMC, the SC can be expressed by the phase only where $\gamma_k = \arg \boldsymbol{x}_0(k) - \arccos(1-\varepsilon^2/2)$ and $\delta = 2\arccos(1-\varepsilon^2/2)$. Though it maximizes a concave quadratic function with a negative semidefinite matrix, it is still non-convex due to the feasible set. The key idea of the successive QCQP refinement is to solve the non-convex optimization problem (1.58) by solving a sequence of convex QCQP problems such that in each iteration of the sequence the designed waveform satisfies the SC and the constant modulus is successively achieved at convergence. The CMC enforces the modulus of every element of $\boldsymbol{x}$ to be a constant $(1/\sqrt{N_T N})$, in other words, every element of $\boldsymbol{x}$ should be located on a scaled unit circle in the complex plane. The SC restricts the phase of each element as shown in (1.58). Based on this observation, the successive QCQP refinement method finds the solution by solving the sequence of convex QCQP problems for which the feasible set gets smaller and closer to the constant modulus circle.

Consider the following convex optimization problem which is a relaxation of the non-convex problem (1.58).

$$\begin{array}{ll} \max\limits_{\boldsymbol{x}} & \boldsymbol{x}^H \boldsymbol{Q} \boldsymbol{x} \\ \text{subject to} & a_k \Re\{\boldsymbol{x}(k)\} + b_k \Im\{\boldsymbol{x}(k)\} \geq c_k \\ & |\boldsymbol{x}(k)|^2 \leq 1/\sqrt{N_T N} \end{array} \tag{1.59}$$

where $\boldsymbol{Q} = \boldsymbol{\Phi} - \lambda \boldsymbol{I}$ is a negative semidefinite matrix, $\Re\{\boldsymbol{x}\}$ and $\Im\{\boldsymbol{x}\}$ represent the real and imaginary part of a complex vector $\boldsymbol{x}$, respectively. Note that the CMC and the SC are relaxed to the convex quadratic inequality constraint and the affine inequality constraint, respectively. Therefore, (1.59) is a convex optimization problem. The parameters a_k, b_k, and c_k represent the line that intersects with the constant modulus at the interval $[\gamma_k, \gamma_k + \delta_k]$. The tighter the SC of (1.59) (which implies the smaller δ_k), the closer (1.59) to (1.58). For instance, if $\delta = \pi/2$, the feasible value of $|\boldsymbol{x}(k)|$ lies between $1/\sqrt{2}$ and 1 and $|\boldsymbol{x}(k)|$ approaches 1 as δ_k reduces as shown in Figure 1.4.

Figure 1.4 illustrates of the successive QCQP refinement algorithm. At the first iteration, a_k, b_k, and c_k are calculated from the SC parameters, γ_k and δ_k. The feasible set forms a circular segment of the unit circle with a radius $1/N_T N$ as shown in Figure 1.4a. Then $\boldsymbol{x}^\star(k)$ is obtained by solving the problem (1.59). Denote the solution at the first iteration by $\boldsymbol{x}^{(0)}$. The basic idea of updating the feasible set such that a new feasible set becomes closer to the constant modulus circle at every iteration is to choose a half of the current feasible set which the solution $\boldsymbol{x}^{(0)}$ belongs to. Specifically, if $\arg \boldsymbol{x}^{(0)}(k) \geq \gamma_k + \delta/2$, we set a new SC as $[\gamma_k + \delta/2, \gamma_k + \delta]$, *i.e.*, the new constraint angles becomes $\gamma_k^{(1)} = \gamma_k + \delta/2$ and $\delta^{(1)} = \delta/2$. If $\arg \boldsymbol{x}^{(0)}(k) <$

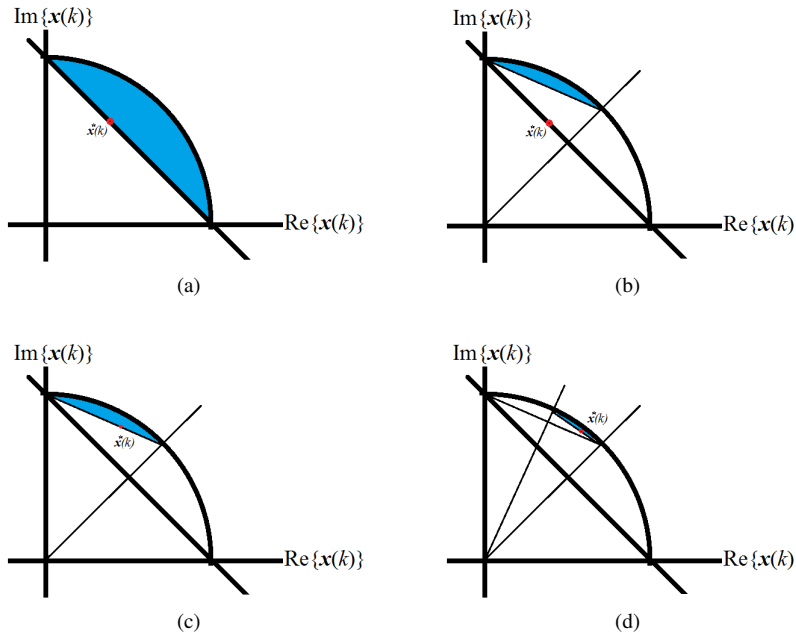


Figure 1.4 Illustration of the successive approximation of problem (1.58). *(a) The convex hull of the feasible set of* (1.58) *is the blue area. (b) The solution point of the convex problem in red. Now we consider only the upper half of the SC and solve again. (c) Second refinement (d) Third refinement, here solution in the third refinement is very close to unity.*

$\gamma_k + \delta/2$, $\gamma_k^{(1)} = \gamma_k$ and $\delta^{(1)} = \delta/2$. In the same way, the problem (1.59) is solved in the next refinement with the updated γ_k and δ. Repeating the refinements for $n = 2,3,\ldots$, the interval $[\gamma_k^{(n)}, \gamma_k^{(n)} + \delta/2^n]$ gets smaller and smaller and eventually the modulus of $\boldsymbol{x}^{(n)}(k)$ will converge to the constant modulus circle as shown in Figure 1.4.

1.3.1.3 Analytical and Experimental Results

Complexity analysis

Based on the computational complexity of a QCQP [29] in each refinement, the overall computational complexity of SQR is $\mathcal{O}(FN_T^{3.5}N^{3.5})$ where F is the total number of refinements. In comparison, SDR with randomization has a computational complexity of $\mathcal{O}(N_T^{3.5}N^{3.5}) + \mathcal{O}(LN_T^2N^2)$ [30]. It is shown that the number of required refinements F is independent of N_TN and in fact $F \ll N_TN$ [8] and the SQR algorithm typically has much lower complexity. The SDR with randomization invariably needs a large number of randomization trials L [15] which makes the term $\mathcal{O}(LN_T^2N^2)$ much larger.

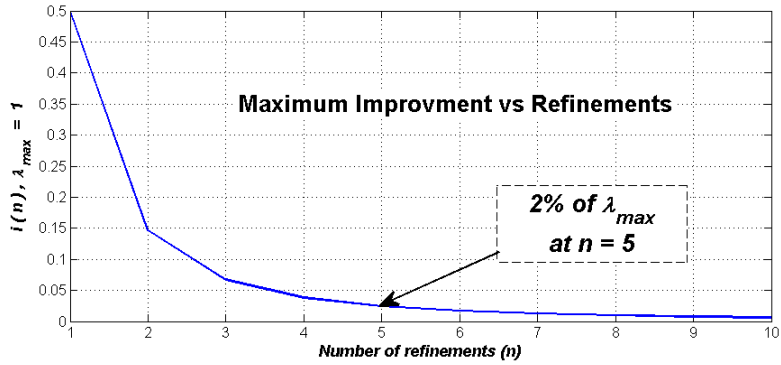


Figure 1.5 $i(n)$ *versus the number of refinements for* $\lambda_{\max} = 1$.

Convergence Analysis

It is shown that the SINR of the SQR algorithm is non-decreasing with each refinement, which means the following inequality always holds at each refinement.

$$\text{SINR}^{n-1} = \boldsymbol{x}^{(n-1)^H} \boldsymbol{\Phi} \boldsymbol{x}^{(n-1)} \leq \boldsymbol{x}^{(n)^H} \boldsymbol{\Phi} \boldsymbol{x}^{(n)} = \text{SINR}^{n} \tag{1.60}$$

The inequality implies that the SQR algorithm improves the SINR after each refinement. Furthermore, it is also shown that the sequence SINR^n is bounded and therefore it converges to a finite value $\text{SINR}^\star$ since a function which is monotonically non-decreasing and bounded converges to a finite value from the monotone convergence criterion [31].

The convergence rate of the algorithm highly depends on the number of refinements. The total number of refinements F can be determined from $i(n)$, the maximum improvement of the new iteration $n+1$ over n, which is given by

$$i(n) = \max\{\text{SINR}^{n+1} - \text{SINR}^{n}\} = \lambda_{\max}(1 - \beta^2(n)) \tag{1.61}$$

where $\beta(n) = \cos(\frac{\pi}{2^{n+1}})$ and $\lambda_{\max}$ is the largest eigenvalue of $\boldsymbol{\Phi}$. Figure 1.5 plots $i(n)$ when $\lambda_{\max} = 1$. It is shown that at most only 2% of $\lambda_{\max}$ of improvement on the SINR value is expected after 5 refinements, which means the SINR nearly converges after the small number of refinements.

Experimental results

The numerical simulations are provided to show the performance of the SQR binary search (SQR-BS) method and the sequential optimization algorithm 1 (SOA1) [15]. For experimental setup, transmit and receive antennas have $N_T = 4$ and $N_R = 8$ elements, respectively, and the orthogonal linear frequency modulation (LFM) waveform is considered as the reference waveform $\boldsymbol{x}_0$. The number of randomization trials used in SOA1 is 20,000 and the SQR method involves four refinement steps, *i.e.*, $F = 4$. The target is located at an angle $\theta_0 = 15^\circ$ and three interference sources is at $\theta_1 = -50^\circ$, $theta_2 = -10^\circ$ and $\theta_3 = 40^\circ$.

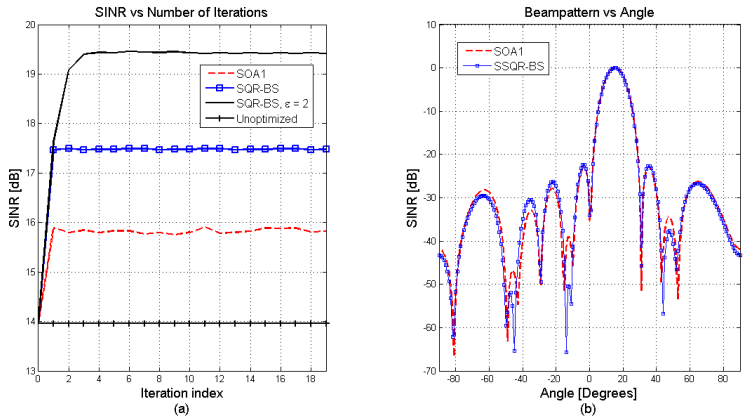


Figure 1.6 (a) The SINR values at each iteration and (b) the beampattern for SOA1 and SQR-BS algorithms with $\varepsilon = 0.7$.

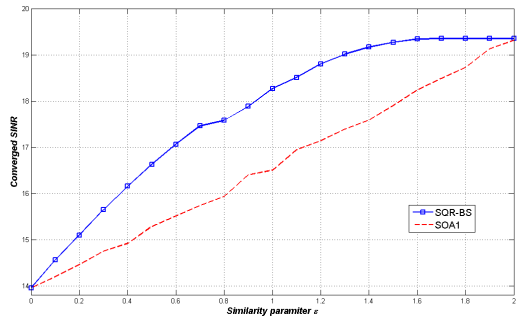


Figure 1.7 The SINR values at convergence of SQR-BS and SOA1 vs. ε.

Figure 1.6 shows the SINR improvement in each iteration in (a) and the beampattern in (b) for the SOA1 and SQR-BS. It is shown that SQR-BS achieves an SINR 1.59 dB higher than SOA1 for the same SC and exhibits much better suppression performance at $\theta = -10^\circ$ and $\theta = 40^\circ$ when compared to SOA1. A plot of the converged SINR value versus the SC parameter ε is shown in Figure 1.7. the SOA1 increases approximately linearly with ε while the SQR-BS exhibits a superlinear increase.

1.3.2 Spatio-Spectral Radar Beampattern Design

In practice, the transmit beampattern design is more challenging for two reasons. The first reason is the requirement of the CMC on the radar transmit waveform, i.e. a constant envelope transmit signal [6]. The second reason is the requirement of spectral compatibility of radar and telecommunication systems, which demands

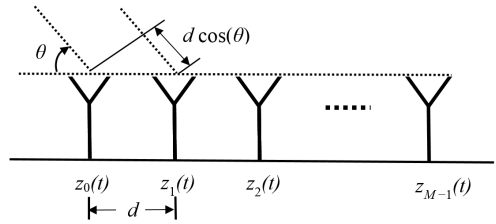


Figure 1.8 Configuration of ULA antenna

a spectral constraint on the radar waveform spectral shape. Designing the MIMO radar beampattern in the simultaneous presence of constant modulus and spectral constraints remains a stiff open challenge.

1.3.2.1 Problem Formulation

Consider a wideband MIMO radar with a uniform linear array (ULA) of M antennas and equal spacing distance of d as shown in Figure 1.8. The signal transmitted from the m-th element is denoted by $z_m(t)$. Let $z_m(t) = x_m(t)e^{j2\pi f_c t}$ where $x_m(t)$ is the baseband signal and f_c is the carrier frequency. We assume that the spectral support of $x_m(t)$ is within the interval $[-B/2, B/2]$ where B is the bandwidth in Hz. The sampled baseband signal transmitted by the m-th element is denoted by $x_m(n) \triangleq x_m(t = nT_s)$, $n = 0, ..., N-1$ with N being the number of time samples and $T_s = 1/B$ is the sampling rate. The discrete Fourier transform (DFT) of $x_m(n)$ is denoted by $y_m(p)$ and it is given by

$$y_m(p) = \sum_{n=0}^{N-1} x_m(n)e^{-j2\pi\frac{np}{N}}, \quad p = -\frac{N}{2}, \ldots, 0, \ldots, \frac{N}{2} - 1 \tag{1.62}$$

where N is assumed to be even in Eq. (1.62). If N is odd, then $p = -(N-1)/2, \ldots, 0, \ldots, (N-1)/2$.

The beampattern can be given by the following discrete angle-frequency grid [14]

$$P_{kp} = |\boldsymbol{a}_{kp}^H \boldsymbol{y}_p|^2 = |\boldsymbol{a}_{kp}^H \boldsymbol{W}_p \boldsymbol{x}|^2 \tag{1.63}$$

where $\boldsymbol{x} \in \mathbb{C}^{MN}$ is the concatenated vector i.e. $\boldsymbol{x} = [\boldsymbol{x}_0^T \quad \boldsymbol{x}_1^T \quad \ldots \quad \boldsymbol{x}_{M-1}^T]^T$, $\boldsymbol{a}_{kp} = [1 \quad e^{j2\pi(\frac{p}{NT_s}+f_c)\frac{d\cos\theta_k}{c}} \quad \ldots \quad e^{j2\pi(\frac{p}{NT_s}+f_c)\frac{(M-1)d\cos\theta_k}{c}}]^T$, and $\boldsymbol{W}_p \in \mathbb{C}^{M\times MN}$ is given by

$$\boldsymbol{W}_p = \boldsymbol{I}_M \otimes \boldsymbol{e}_p^H \tag{1.64}$$

where $\boldsymbol{e}_p^H = [1 \quad e^{-j2\pi\frac{p}{N}} \quad \ldots \quad e^{-j2\pi\frac{(N-1)p}{N}}] \in \mathbb{C}^N$ and $\boldsymbol{I}_M$ is an $M \times M$ identity matrix.

The problem of spectral co-existence has been of great interest recently [16, 32, 33] and involves minimization of interference caused by radar transmission at victim communication receivers operating in the same frequency band. In this case, the beampattern of the transmit waveform is required to have nulls in these bands to prevent interference. For J communication receivers, we suppose that the j-th

communication receiver operating on a frequency band $B_j = [p_l^j, p_u^j]$, where p_l^j and p_u^j are the lower and upper normalized frequency, respectively. We denote the desired (discrete) spectrum shape by $\hat{\boldsymbol{y}} = [\hat{y}_{-\frac{N}{2}}, \hat{y}_{-\frac{N}{2}+1}, ..., \hat{y}_{\frac{N}{2}-1}] \in \mathbb{C}^{N\times 1}$ defined as

$$\hat{y}_p = \begin{cases} 0 & \text{for } p \in B_j = [p_l^j, p_u^j], \qquad j = 1,2,...,J \\ \gamma & \text{otherwise.} \end{cases} \tag{1.65}$$

where γ is a scalar such that $\hat{\boldsymbol{y}}^H \boldsymbol{F}\boldsymbol{F}^H \hat{\boldsymbol{y}} = N$ and $\boldsymbol{F}$ is the DFT matrix. In SHAPE algorithm proposed by Rowe *et al.* [34], a least-squares fitting approach for the spectral shaping problem for SISO has been formulated by minimizing the following cost function

$$\|\boldsymbol{F}^H \boldsymbol{x} - \hat{\boldsymbol{y}}\|_2^2 \tag{1.66}$$

We extend (1.66) for MIMO radar and employ it as a constraint in the optimization problem as follows.

$$\|(\boldsymbol{I}_M \otimes \boldsymbol{F}^H)(\boldsymbol{1}_M \otimes \hat{\boldsymbol{y}}) - \boldsymbol{x}\|_2^2 = \|\bar{\boldsymbol{F}}^H \bar{\boldsymbol{y}} - \boldsymbol{x}\|_2^2 \leq E_R \tag{1.67}$$

where $\boldsymbol{1}_M = [1,1,\ldots,1] \in \mathbb{R}^{M\times 1}$, $\bar{\boldsymbol{F}} = \boldsymbol{I}_M \otimes \boldsymbol{F}^H$, and $\bar{\boldsymbol{y}} = \boldsymbol{1}_M \otimes \hat{\boldsymbol{y}}$, and E_R is the maximum tolerable spectral error.

1.3.2.2 Beampattern Design with Interference Control (BIC) under Constant Modulus

The optimization problem for beampattern design under the CMC and the spectral constraint can be formulated as the following matching problem.

$$\begin{cases} \min\limits_{\boldsymbol{x}} & \sum_{k=1}^{K} \sum_{p=-\frac{N}{2}}^{\frac{N}{2}-1} [d_{kp} - |\boldsymbol{a}_{kp}^H \boldsymbol{W}_p \boldsymbol{x}|]^2 \\ \text{s.t.:} & |x_m(n)| = 1, \text{ for } m = 1,2,\ldots,M \text{ and} \\ & \qquad\qquad n = 0,1,\ldots,N-1 \\ & \|\bar{\boldsymbol{F}}^H \bar{\boldsymbol{y}} - \boldsymbol{x}\|_2^2 \leq E_R \end{cases} \tag{1.68}$$

where $d_{kp} \in \mathbb{R}$ is the desired beampattern. These constraints are neither convex nor linear and it is well known in the literature that (1.68) is a hard non-convex problem even without the spectral constraint. The objective function can be rewritten as [14]

$$\sum_{k=1}^{K} \sum_{p=-\frac{N}{2}}^{\frac{N}{2}-1} |d_{kp} e^{j\phi_{kp}} - \boldsymbol{a}_{kp}^H \boldsymbol{W}_p \boldsymbol{x}|^2 \tag{1.69}$$

where $\phi_{kp} = \arg\{\boldsymbol{a}_{kp}^H \boldsymbol{W}_p \boldsymbol{x}\}$. This objective function can be optimized by an iterative method [14, 35, 36] which first optimizes $\boldsymbol{x}$ for a fixed ϕ_{kp} and then finds the optimal ϕ_{kp} for the fixed $\boldsymbol{x}$. For a fixed ϕ_{kp}, the objective function can be further simplified in a quadratic form as following [10].

$$\sum_{k=1}^{K} \sum_{p=-\frac{N}{2}}^{\frac{N}{2}-1} |d_{kp} e^{j\phi_{kp}} - \boldsymbol{a}_{kp}^H \boldsymbol{W}_p \boldsymbol{x}|^2 = \sum_p \|\boldsymbol{d}_p - \boldsymbol{A}_p \boldsymbol{W}_p \boldsymbol{x}\|_2^2 \tag{1.70}$$

$$= \boldsymbol{x}^H \boldsymbol{P}\boldsymbol{x} - \boldsymbol{q}^H \boldsymbol{x} - \boldsymbol{x}^H \boldsymbol{q} + r \tag{1.71}$$

Moreover, using $\boldsymbol{x}^H\boldsymbol{x} = 2L$, the spectral constraint can also be simplified as

$$\Re\{\bar{\boldsymbol{y}}^H\bar{\boldsymbol{F}}\boldsymbol{x}\} \geq (1 - E_R/2)L \tag{1.72}$$

The optimization problem (1.68) for a fixed ϕ_{kp} is equivalent to the following problem in real variables.

$$\begin{cases} \min\limits_{\boldsymbol{s}} & \boldsymbol{s}^T(\boldsymbol{R} + \lambda\boldsymbol{I})\boldsymbol{s} \\ \text{s.t.:} & \boldsymbol{s}^T\boldsymbol{E}_l\boldsymbol{s} = 1, \quad l = 1,2,\ldots,L \\ & \bar{\boldsymbol{s}}^T\boldsymbol{s} \geq (1 - E_R/2)L \end{cases} \tag{1.73}$$

where λ is an arbitrary positive number, $\bar{\boldsymbol{s}} = [\Re\{\bar{\boldsymbol{F}}^H\bar{\boldsymbol{y}}\}^T \; \Im\{\bar{\boldsymbol{F}}^H\bar{\boldsymbol{y}}\}^T \; 0]^T$, $\boldsymbol{R} = \begin{bmatrix} \boldsymbol{G} & -\boldsymbol{t} \\ -\boldsymbol{t}^T & r \end{bmatrix}$, $\boldsymbol{s} = [\Re\{\boldsymbol{x}^T\} \; \Im\{\boldsymbol{x}^T\} \; 1]^T$, and $\boldsymbol{t} = [\Re\{\boldsymbol{q}^T\} \; \Im\{\boldsymbol{q}^T\}]^T$.

Sequence of Closed Form Solutions

Though (1.73) is a minimization problem with a convex objective function since $\boldsymbol{R}$ is positive semi-definite, it is still non-convex because of the constant modulus constraint, $\boldsymbol{s}^T\boldsymbol{E}_l\boldsymbol{s} = 1$. It can also be solved by a sequential approach which involves solving a sequence of convex problems. Let us consider the following sequence of constrained quadratic programming (QP) where the n-th QP is given by

$$(CP)^{(n)} \begin{cases} \min\limits_{\boldsymbol{s}} & \boldsymbol{s}^T(\boldsymbol{R} + \lambda\boldsymbol{I})\boldsymbol{s} \\ \text{s.t.:} & \boldsymbol{B}^{(n)}\boldsymbol{s} = \boldsymbol{1} \\ & \bar{\boldsymbol{s}}^{(n)T}\boldsymbol{s} \geq (1 - E_R/2)L \end{cases} \tag{1.74}$$

where $\bar{\boldsymbol{s}}^{(n)}$ is given by:

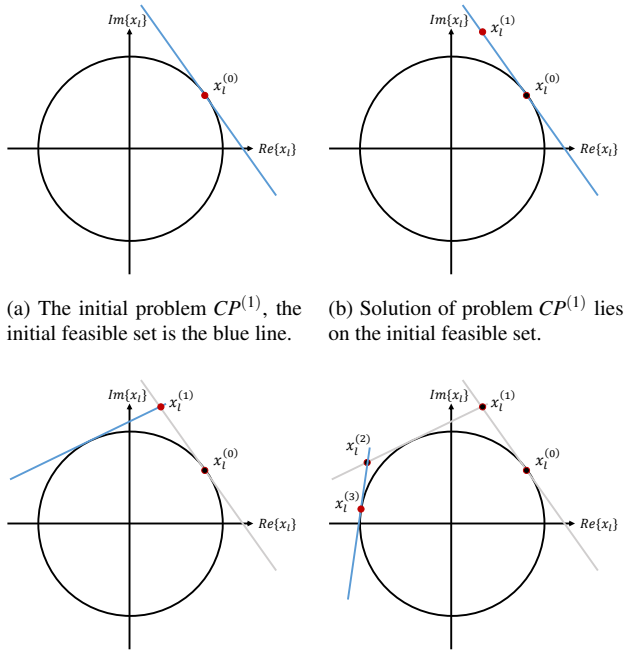
$$\bar{\boldsymbol{s}}^{(n)} = \begin{bmatrix} \Re\{(\bar{\boldsymbol{F}}^H\bar{\boldsymbol{y}}) \odot e^{\{j\arg(\boldsymbol{x}^{(n-1)}) - \arg(\bar{\boldsymbol{F}}^H\bar{\boldsymbol{y}})\}}\} \\ \Im\{(\bar{\boldsymbol{F}}^H\bar{\boldsymbol{y}}) \odot e^{\{j\arg(\boldsymbol{x}^{(n-1)}) - \arg(\bar{\boldsymbol{F}}^H\bar{\boldsymbol{y}})\}}\} \\ 0 \end{bmatrix} \tag{1.75}$$

and $\boldsymbol{B}^{(n)} = [\boldsymbol{b}_1^{(n)}, \boldsymbol{b}_2^{(n)}, ..., \boldsymbol{b}_{L+1}^{(n)}]^T \in \mathbb{R}^{(L+1)\times(2L+1)}$ such that the line defined by $\boldsymbol{b}_l^{(n)T}\boldsymbol{s} = 1$ is a tangent to the circle $\boldsymbol{s}^T\boldsymbol{E}_l\boldsymbol{s} = 1$ for $l = 1,2,\ldots,L$. Specifically, $\boldsymbol{b}_l$ is given by

$$\boldsymbol{b}_l^{(n)}(i) = \begin{cases} \cos(\gamma_l^{(n)}) & \text{if } i = l \\ \sin(\gamma_l^{(n)}) & \text{if } i = l + L \\ 0 & \text{otherwise.} \end{cases} \tag{1.76}$$

for $l = 1,\ldots,L$ and $\boldsymbol{b}_{L+1}^{(n)} = [0,\ldots,0,1]^T$ where $\gamma_l^{(n)} = 2\arg(x_l^{(n-1)}) - \gamma_l^{(n-1)}$ and $x_l^{(n)}$ is the l-th elements of $\boldsymbol{x}^{(n)}$ which is the complex version of the optimal solution of (1.74), $\boldsymbol{s}^{(n)}$, that is, $x_l^{(n)} = s_l^{(n)} + js_{l+L}^{(n)}$ and conversely $\boldsymbol{s}^{(n)} = [\Re\{\boldsymbol{x}^{(n)}\}^T \Im\{\boldsymbol{x}^{(n)}\}^T \; 1]^T$. Note that, the term $e^{\{j\arg(\boldsymbol{x}^{(n-1)}) - \arg(\bar{\boldsymbol{F}}^H\bar{\boldsymbol{y}})\}}$ in (1.75) depends on the argument $\boldsymbol{x}^{(n-1)}$, which changes $\bar{\boldsymbol{s}}^{(n)}$ in each iteration.

Although the problem (1.74) does not result in a constant modulus solution, a sequence of such problems (in the index n) ensures a non-increasing sequence of cost



(a) The initial problem $CP^{(1)}$, the initial feasible set is the blue line.

(b) Solution of problem $CP^{(1)}$ lies on the initial feasible set.

(c) The new adjusted feasible set (Contains $x_l^{(1)}$) in blue, the previous feasible set in gray.

(d) The converged solution now lies on the constant modulus.

Figure 1.9 Illustration of the successive solutions of (1.74) *for the l-th element of the vector* $\boldsymbol{x}^{(n)}$ *i.e.* $x_l^{(n)}$*. The current feasible set is shown via a blue line.*

function values, such that the sequence of the corresponding optimal solutions converges to constant modulus for large enough λ [37]. It is shown that the constraints of $CP^{(n)}$ in (1.74) are adjusted so that the feasible set of $CP^{(n)}$ includes $\boldsymbol{x}^{(n-1)}$ [17]. This means that the feasible set of each iteration is updated such that it contains the optimal solution of the optimization problem at the previous iteration step. If $|\boldsymbol{x}^{(n)}| = \mathbf{1}$, then the constraints of the next problem $CP^{(n+1)}$ are the same as problem $CP^{(n)}$, which means $\boldsymbol{x}^{(n+1)} = \boldsymbol{x}^{(n)}$ and, hence, the algorithm converges. It has been also shown that the cost function sequence is in fact non-increasing and converges. This procedure is visually illustrated in Figure 1.9.

The optimization problem (1.74) is a convex quadratic minimization with linear equality constraints. Using the optimality conditions for problem (1.74), the sufficient and necessary Karush-Kuhn-Tucker (KKT) conditions [24] of (1.74) give the following.

$$2(\boldsymbol{R}+\lambda\boldsymbol{I})\boldsymbol{s}^{(n)} + \boldsymbol{B}^{(n)T}\boldsymbol{v}^{(n)} - \mu^{(n)}\bar{\boldsymbol{s}} = 0 \tag{1.77}$$

$$\boldsymbol{B}^{(n)}\boldsymbol{s}^{(n)} = \mathbf{1} \tag{1.78}$$

$$\mu^{(n)}\left(\bar{\boldsymbol{s}}^{(n)T}\boldsymbol{s}^{(n)} - (1-E_R/2)L\right) = 0 \tag{1.79}$$

$$\bar{\boldsymbol{s}}^{(n)T}\boldsymbol{s}^{(n)} - (1-E_R/2)L \geq 0 \tag{1.80}$$

$$\mu^{(n)} \geq 0 \tag{1.81}$$

Solving the equations and the inequality above gives a closed form solution as following.

$$\boldsymbol{s}^{(n)} = \begin{cases} \bar{\boldsymbol{R}}^{-1}\boldsymbol{B}^{(n)T}\left(\boldsymbol{B}^{(n)}\bar{\boldsymbol{R}}^{-1}\boldsymbol{B}^{(n)T}\right)^{-1}\mathbf{1} & \text{if } \bar{\boldsymbol{s}}^{(n)T}\boldsymbol{s}^{(n)} - (1-E_R/2)L \geq 0 \\ \mu^{(n)}\bar{\boldsymbol{R}}^{-1}(\boldsymbol{I}-\boldsymbol{B}^{(n)T}\hat{\boldsymbol{R}}\boldsymbol{B}^{(n)}\bar{\boldsymbol{R}}^{-1})\bar{\boldsymbol{s}}^{(n)} + \hat{\boldsymbol{s}}^{(n)} & \text{otherwise} \end{cases} \tag{1.82}$$

where $\hat{\boldsymbol{R}} = \left(\boldsymbol{B}^{(n)}\bar{\boldsymbol{R}}^{-1}\boldsymbol{B}^{(n)T}\right)^{-1}$, $\mu^{(n)} = \frac{1}{\alpha^{(n)}}\left(\bar{\boldsymbol{s}}^{(n)T}\hat{\boldsymbol{s}}^{(n)} - (1-E_R/2)L\right)$, and

$$\alpha^{(n)} = -\begin{bmatrix}\bar{\boldsymbol{s}}^{(n)} \\ \mathbf{0}\end{bmatrix}^T\begin{bmatrix}\bar{\boldsymbol{R}} & \boldsymbol{B}^{(n)T} \\ \boldsymbol{B}^{(n)} & \mathbf{0}\end{bmatrix}^{-1}\begin{bmatrix}\bar{\boldsymbol{s}}^{(n)} \\ \mathbf{0}\end{bmatrix} \tag{1.83}$$

Nullforming Beampattern Design

The beampattern design method described above can also be applied the nullforming beampattern design problem which is a special case of the full beampattern design. The goal of nullforming beampattern design is to form a beampattern with nulls in desired directions and the optimization problem is given by

$$\begin{cases} \min\limits_{\boldsymbol{x}} & \boldsymbol{x}^H\boldsymbol{V}\boldsymbol{x} \\ \text{subject to:} & |\boldsymbol{x}(k)|^2 \leq 1/(MN) \\ \text{s.t.:} & |x_m(n)| = 1, \text{ for } m = 1,2,\ldots,M \text{ and } n = 0,1,\ldots,N-1 \\ & \|\bar{\boldsymbol{F}}^H\bar{\boldsymbol{y}} - \boldsymbol{x}\|_2^2 \leq E_R \end{cases} \tag{1.84}$$

where $\boldsymbol{V} = \sum_{p=-\frac{N}{2}}^{\frac{N}{2}-1}\boldsymbol{W}_p^H\boldsymbol{A}_p^H\boldsymbol{A}_p\boldsymbol{W}_p$. Since $\boldsymbol{V}$ is positive semidefinite and there are no linear terms in the objective function, the solution can be obtained by the algorithm.

1.3.2.3 Analytical and Experimental Results

Complexity and convergence analysis

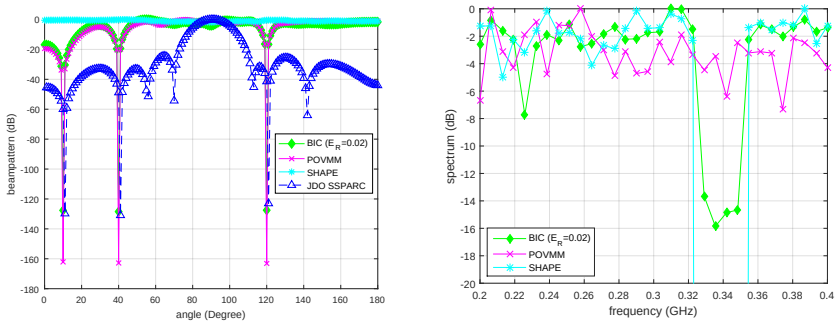
Based on the computational cost of each iteration, the overall computational complexity is $\mathcal{O}(FL^{2.373}) - \mathcal{O}(FL^3)$ [38] where F is the total number of iterations. It also converges to a finite value $\boldsymbol{s}^\star$ since the sequence $\{g(\boldsymbol{s}^{(n)})\}_{n=0}^{\infty}$ is non-increasing and bounded $g(\boldsymbol{s}) \geq 0$ where $g(\boldsymbol{s}) = \boldsymbol{s}^T(\boldsymbol{R}+\lambda\boldsymbol{I})\boldsymbol{s}$. It can be easily shown by the following inequality,

$$\boldsymbol{s}^{(n)T}(\boldsymbol{R}+\lambda\boldsymbol{I})\boldsymbol{s}^{(n)} \leq \boldsymbol{s}^{(n-1)T}(\boldsymbol{R}+\lambda\boldsymbol{I})\boldsymbol{s}^{(n-1)} \tag{1.85}$$

The inequality always holds since $\boldsymbol{s}^{(n-1)}$ belongs to the feasible set of $CP^{(n)}$ and $\boldsymbol{s}^{(n)}$ is the optimal solution of $CP^{(n)}$ which is a convex problem.

Numerical results

Figure 1.10 shows the results for nullforming beampattern of BIC, POVMM [39], SHAPE [34], and JDO SSPARC [40]. POVMM performs nullforming beampattern



(a) beampattern vs. angle

(b) spectrum vs. frequency

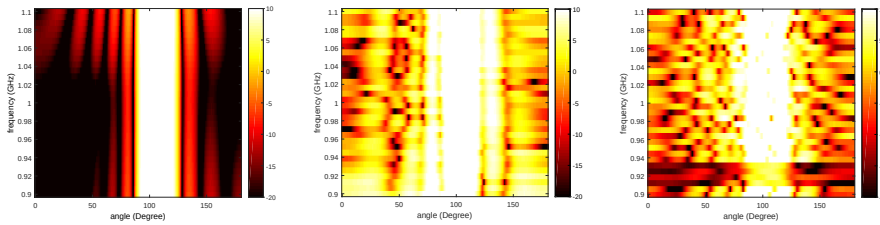
Figure 1.10 Nullforming beampattern design

design by optimizing phases of the waveform under the constant modulus constraint but no spectral constraint is involved. The SHAPE algorithm is a computationally efficient method of designing sequences with desired spectrum shapes. In particular, the spectral shape is optimized as a cost function subject to the constant modulus constraint but the resulting beampattern is an outcome (not explicitly controlled). JDO SSPARC is an approach for beamforming that maximizes the signal power through the forward channels while simultaneously minimizes the response at the co-channels. JDO SSPARC does not control the spectral shape of the waveform in the frequency domain. It is shown that BIC, POVMM, and JDO SSPARC achieve nulls in the desired angles, 10°, 40° and 120°, *i.e.*, desired spatial control in Figure 1.10a. SHAPE lacks a spatial control component by virtue of its design. Note that the forward channel for JDO SSPARC is set to be $\theta = [80^\circ \text{ to } 100^\circ]$, however, unlike the other methods, the resulting waveform is non-constant modulus. On the other hand, Figure 1.10b plots the spectrum versus the frequency. Here, BIC and SHAPE effectively suppress the energy in the frequency bands where the transmission should be mitigated. Unsurprisingly, POVMM do not provide the desired suppression in the frequency bands of interest because it is not designed for the same. In summary, only the BIC enables the desired spatio-spectral control.

For wideband beampattern design, we place a notch in the band 910-932 MHz and consider the following desired transmit beampattern

$$d(\theta, f) = \begin{cases} 1 & \theta = [95^\circ, 120^\circ] \\ 0 & \text{Otherwise.} \end{cases} \tag{1.86}$$

in Figure 1.11. It shows the angle-frequency plot of the beampattern for WBFIT method [14] (no spectral constraint) and BIC with the spectral constraint ($E_R = 0.01$). The BIC method is able to keep the energy of the waveform in particular frequency band low enough as well as achieve higher suppression at the undesired angles compared to WBFIT.



(a) (b) (c)

Figure 1.11 Plot of the beampattern. (a) unconstrained (b) WBFIT method (c) BIC

1.3.3 Quartic Gradient Descent for Tractable Radar Ambiguity Function Shaping

The problem of minimizing the disturbance power at the output of the matched filter in a single antenna cognitive radar set-up improves the SINR for the radar. The aforementioned disturbance power can be shown to be an expectation of the slow-time ambiguity function (STAF) of the transmitted waveform over range-Doppler bins of interest. The design problem is known to be a non-convex quartic function of the transmit radar waveform. This STAF shaping problem becomes even more challenging in the presence of practical constraints on the transmit waveform such as the CMC. Most existing approaches address the aforementioned challenges by suitably modifying or relaxing the design cost function and/or the CMC.

In this part, we discuss a solution that involves direct optimization over the non-convex complex circle manifold, i.e. the CMC set. This solution uses a new update strategy (quartic-gradient-descent (QGD)) that computes an *exact* gradient of the quartic cost and invokes principles of optimization over manifolds towards an iterative procedure with guarantees of monotonic cost function decrease and convergence [41]. Experimentally, QGD outperforms state of the art approaches for shaping the ambiguity function under the CMC while being computationally less expensive.

1.3.3.1 System Model and Problem Formulation

We consider a monostatic single-input-single-output (SISO) radar system which transmits a coherent burst of slow-time[5] coded pulses dented by

$$\boldsymbol{x} = [x(0), x(1), \ldots, x(N-1)]^T \in \mathbb{C}^N \tag{1.87}$$

The radar system illuminates the environment by sending N coherent burst of slow-time coded pulses $\boldsymbol{x}$. The signal at the receiver is down-converted to baseband, undergoes a pulse matched filtering operation, and then is sampled. The received vector $\boldsymbol{v} = [v(0), v(1), \ldots, v(N-1)]^T \in \mathbb{C}^N$ of observations from the range-azimuth cell under consideration is given by

$$\boldsymbol{v} = \alpha_T \boldsymbol{x} \odot \boldsymbol{p}(\boldsymbol{v}_{d_T}) + \boldsymbol{d}(\boldsymbol{x}) + \boldsymbol{n} \tag{1.88}$$

[5]For more details about the slow-time and fast-time coding, the reader is advise to see [19, 42, 11]

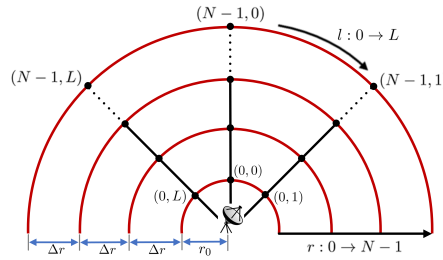


Figure 1.12 Range-azimuth bins, where the target is located at the point $(0,0)$. *The distance* r_0 *is assumed to be* $r_0 \leq r_{max} = \frac{cT_r}{2}$, *where* r_{max} *is the maximum unambiguous range defines the maximum distance to locate a target, and* $\Delta r = r_{max}$.

where α_T is a complex parameter accounting for channel propagation and back-scattering effects from the target within the range-azimuth bin of interest, $\boldsymbol{p}(v_{d_T}) = [1, e^{j2\pi v_{d_T}}, ..., e^{j2\pi(N-1)v_{d_T}}]^T$, v_{d_T} is the normalized target Doppler frequency, $\boldsymbol{d}(\boldsymbol{x})$ is the vector of interfering echo samples, and $\boldsymbol{n}$ is the filtered noise vector with $\mathbb{E}[\boldsymbol{n}] = 0$ and $\mathbb{E}[\boldsymbol{n}\boldsymbol{n}^H] = \sigma_n^2 \boldsymbol{I}$. According to [12], the vector $\boldsymbol{d}(\boldsymbol{x})$ captures the returns from different N_t interfering scatterers located at different range-azimuth bins[6] (r,l), where $r \in \{0,1,\ldots,N-1\}$, $l \in \{0,1,\ldots,L\}$ (as illustrated in Figure 1.12) and $L+1$ is the number of discrete azimuth sectors. This vector $\boldsymbol{d}(\boldsymbol{x})$ can be expressed as

$$\boldsymbol{d}(\boldsymbol{x}) = \sum_{i=1}^{N_t} \rho_i \boldsymbol{J}^{r_i} \underbrace{\left(\boldsymbol{x} \odot \boldsymbol{p}(v_{d_i})\right)}_{\boldsymbol{c}_{v_{d_i}}} = \sum_{i=1}^{N_t} \rho_i \boldsymbol{J}^{r_i} \boldsymbol{c}_{v_{d_i}} \tag{1.89}$$

where $r_i \in \{0,1,...,N-1\}$ is the range position, ρ_i is the echo complex amplitude, v_{d_i} and $\boldsymbol{c}_{v_{d_i}} = \boldsymbol{x} \odot \boldsymbol{p}(v_{d_i})$ are the normalized Doppler frequency and the signature[7] of the i^{th} scatterers, respectively. $\boldsymbol{J}^r$ is an $N \times N$ shift matrix and $\forall\ r \in \{-N+1,...,0,...,N-1\}$ it is denoted as:

$$\boldsymbol{J}^r(l_1,l_2) = \begin{cases} 1, & \text{if } l_1 - l_2 = r \\ 0, & \text{otherwise} \end{cases} \quad (l_1,l_2) \in \{1,...,N\}^2 \tag{1.90}$$

with $\boldsymbol{J}^{-r} = (\boldsymbol{J}^r)^T$. Combining Eq. (1.88) and Eq. (1.89), the output of the matched filter to the target signature $\boldsymbol{c}_{v_{d_T}} = \boldsymbol{x} \odot \boldsymbol{p}(v_{d_T})$ is given by

$$\boldsymbol{c}_{v_{d_T}}^H \boldsymbol{v} = \alpha_T \|\boldsymbol{x}\|_2^2 + \text{Dist}\big(\boldsymbol{x}, v, r\big) \tag{1.91}$$

[6]Without loss of generality, the target of interest can be assumed to be located at the range-azimuth bin $(0,0)$ and scatterers are located at further range bins [43, 44].

[7]This vector $\boldsymbol{c}_v$ will be used in this chapter to represent the signature of any object with a Doppler frequency v, i.e., $\boldsymbol{c}_v = \boldsymbol{x} \odot \boldsymbol{p}(v)$.

where $\text{Dist}(\boldsymbol{x}, \nu, r)$ represents the disturbance at the output of the match filter, i.e.,

$$\text{Dist}(\boldsymbol{x}, \nu, r) = \underbrace{\boldsymbol{c}_{\nu_{d_T}}^H \boldsymbol{n}}_{\text{noise}} + \underbrace{\sum_{i=1}^{N_t} \rho_i \boldsymbol{c}_{\nu_{d_T}}^H \boldsymbol{J}^{r_i} \boldsymbol{c}_{\nu_{d_i}}}_{\text{interference}} \tag{1.92}$$

Assuming that $\boldsymbol{n}$ is uncorrelated with $\boldsymbol{d}(\boldsymbol{x})$, the energy of the disturbances in the match filter output can be expressed as:

$$\mathbb{E}[|\text{Dist}(\boldsymbol{x}, \nu, r)|^2] = \mathbb{E}\left[\left|\boldsymbol{c}_{\nu_{d_T}}^H \boldsymbol{n}\right|^2\right] + \mathbb{E}\left[\left|\sum_{i=1}^{N_t} \rho_i \boldsymbol{c}_{\nu_{d_T}}^H \boldsymbol{J}^{r_i} \boldsymbol{c}_{\nu_{d_i}}\right|^2\right] \tag{1.93}$$

Problem formulation

In [12], the normalized Doppler frequencies ν_{d_i} are expressed in terms of the difference with respect to ν_{d_T}, and the normalized Doppler frequency range $[-\frac{1}{2}, \frac{1}{2}]$ is divided into N_ν bins. Consequently, the normalized frequencies ν_{d_i} can be represented by the discrete frequencies $\nu_h = -\frac{1}{2} + \frac{h}{N_\nu},\ h = 0, ..., N_\nu - 1$. Using this representation and approximating the statistical expectation in Eq. (1.93) by the sample mean, the disturbances in the match filter output will be[8]

$$\mathbb{E}\left[\left|\text{Dist}(\boldsymbol{x}, \nu, r)\right|^2\right] = \sum_{r=0}^{N-1} \sum_{h=0}^{N_\nu - 1} p(r,h) \|\boldsymbol{x}\|_2^2 g_{\boldsymbol{x}}(r, \nu_h) + \sigma_n^2 \|\boldsymbol{x}\|_2^2 \tag{1.94}$$

where $p(r,h)$ is interference map for the range-Doppler bin (r, ν_h) and $g_{\boldsymbol{x}}(r, \nu_h)$ is the STAF of the transmitted code $\boldsymbol{x}$ defined as:

$$g_{\boldsymbol{x}}(r, \nu_h) = \frac{1}{\|\boldsymbol{x}\|_2^2} \left|\boldsymbol{x}^H \boldsymbol{J}^r \boldsymbol{c}_\nu\right|^2 \tag{1.95}$$

with $r \in \{0, 1, ..., N-1\}$ and $\nu_h = -\frac{1}{2} + \frac{h}{N_\nu},\ h = 0, ..., N_\nu - 1$. Given a (r, ν_h) pair, the STAF $g_{\boldsymbol{x}}(r, \nu_h)$ gives the range-Doppler response from an interfering patch corresponding with a Doppler frequency of ν_h located r time-lag away. As mentioned before, the goal is to design a suitable radar waveform $\boldsymbol{x}$ in order to shape the STAF to match a desired range-Doppler response (*shaping the STAF is equivalent to minimizing the disturbances in the match filter output in Eq. (1.94)* [12]) under the CMC, i.e., $|x(n)| = 1,\ n = 1, 2, \ldots, N$. With this constraint, the quantity $\|\boldsymbol{x}\|_2^2$ in Eq. (1.94) is constant and hence the disturbance in the output of the matched filter will be minimized using the following cost function [11]:

$$\phi(\boldsymbol{x}) = \sum_{i=1}^{M} \boldsymbol{x}^H \boldsymbol{C}_i \boldsymbol{x} \boldsymbol{x}^H \boldsymbol{C}_i^H \boldsymbol{x}. \tag{1.96}$$

where $M = N \times N_\nu$, and i is a one-to-one mapping index with (r,h), i.e., for each pair $(r,h) \in \{0, \ldots, N-1\} \times \{0, \ldots, N_\nu - 1\}$, we have $i = rN_\nu + h \in \{1, \ldots, M = N \times N_\nu\}$ and the matrix $\boldsymbol{C}_i$ is defined as $\boldsymbol{C}_i = \boldsymbol{C}_{(r,h)} = \sqrt{p(r,h)} \boldsymbol{J}^r \text{diag}(\boldsymbol{p}(\nu_h))$. Then

[8] A detailed derivation for Eq. (1.94) can be found in [12].

the optimization problem for shaping the STAF under the CMC will be the following complex quartic minimization problem.

$$\begin{cases} \min\limits_{\boldsymbol{x}} & \phi(\boldsymbol{x}) = \sum_{i=1}^{M} \boldsymbol{x}^H \boldsymbol{C}_i \boldsymbol{x}\boldsymbol{x}^H \boldsymbol{C}_i^H \boldsymbol{x}. \\ \text{s.t.} & \boldsymbol{x} \in \mathcal{S}^N \end{cases} \tag{1.97}$$

where $\mathcal{S}^N$ is the complex circle manifold (formal manifold terminology for the CMC set) defined as $\mathcal{S}^N \stackrel{\triangle}{=} \{\boldsymbol{x} \in \mathbb{C}^N : |x(n)| = 1,\ n = 1,2,\ldots,N\}$.

It has been shown in [12] that the optimization problem (1.97) is NP-hard. The authors in [12] approach this problem via a polynomial time waveform optimization procedure based on the maximum block improvement (MBI) method. In their work, the CMC is enforced by employing a randomization strategy [45] which leads to an effective solution but one that has high computational complexity. In [46], a combination of majorization-minimization (MM) (to majorize the quartic cost by a quadratic) and coordinate descent methods is used. The CMC is extracted in different parts of the optimization but a direct optimization over the non-convex CMC remains elusive.

We invoke principles of optimization over non-convex manifolds to address this open challenge. Our focus is on developing a gradient-based method, which can enable descent on the complex circle manifold while maintaining feasibility.

First, the cost function in (1.97) can be altered by adding the term $\gamma \boldsymbol{x}^H \boldsymbol{x}\boldsymbol{x}^H \boldsymbol{x}$, i.e.,

$$\begin{cases} \min\limits_{\boldsymbol{x}} & \bar{\phi}(\boldsymbol{x}) = \sum_{i=1}^{M} \boldsymbol{x}^H \boldsymbol{C}_i \boldsymbol{x}\boldsymbol{x}^H \boldsymbol{C}_i^H \boldsymbol{x} + \gamma \boldsymbol{x}^H \boldsymbol{x}\boldsymbol{x}^H \boldsymbol{x} \\ \text{s.t.} & \boldsymbol{x} \in \mathcal{S}^N \end{cases} \tag{1.98}$$

where $\gamma \geq 0$ (it will be used later in Lemma 1 to control convergence). Since the problem (1.98) enforces the CMC, the term $\gamma \boldsymbol{x}^H \boldsymbol{x}\boldsymbol{x}^H \boldsymbol{x}$ is constant (i.e. $\gamma \boldsymbol{x}^H \boldsymbol{x}\boldsymbol{x}^H \boldsymbol{x} = \gamma N^2$). Hence, the optimal solution of the problem (1.97) and the optimal solution of the problem (1.98) are identical for any $\gamma \geq 0$.

1.3.3.2 Quartic Gradient Descent (QGD) algorithm

The goal is to find an efficient method to deal with the non-convex feasible set of the problem (1.97) (or (1.98)), i.e., the complex circle manifold. Many classical line-search algorithms from unconstrained nonlinear optimization in $\mathbb{C}^N$ such as gradient descent can be used in optimization over manifolds but with some modifications. In general, line-search methods in $\mathbb{C}^N$ are based on the following update formula [47]:

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \beta_k \boldsymbol{\eta}_k \tag{1.99}$$

where $\boldsymbol{\eta}_k \in \mathbb{C}^N$ is the *search direction* and $\beta_k \in \mathbb{R}$ is the *step size*. The most obvious choice for the search direction is the steepest descent direction which is the negative gradient of $\bar{\phi}(\boldsymbol{x})$ at the point $\boldsymbol{x}_k$, i.e., $\boldsymbol{\eta}_k = -\nabla_{\boldsymbol{x}} \bar{\phi}(\boldsymbol{x}_k)$ [48, 47]. In the literature [49, 50], the following high level structure is suggested:

- The descent will be performed on the manifold itself rather than in the Euclidean space by means of the *intrinsic* search direction. The *intrinsic* search direction is a vector in the tangent space $\mathcal{T}_{\boldsymbol{x}_k}\mathcal{M}$ to the manifold $\mathcal{M}$ at the point

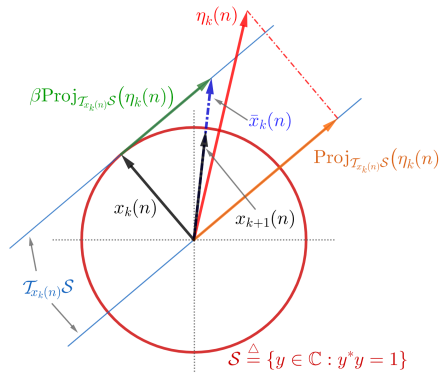


Figure 1.13 Illustration of the update of $x_{k+1}(n)$ starting from $x_k(n)$, where $x_k(n)$ and $\eta_k(n)$ are the n-th elements of the vectors $\boldsymbol{x}_k$ and $\boldsymbol{\eta}_k$, respectively.

$\boldsymbol{x}_k \in \mathcal{M}$. This *intrinsic* search direction can be obtained by projecting the standard search direction $\boldsymbol{\eta}_k = -\nabla_{\boldsymbol{x}}\bar{\phi}(\boldsymbol{x}_k)$ onto $\mathcal{T}_{\boldsymbol{x}_k}\mathcal{M}$ by means of a projection operator $\text{Proj}_{\mathcal{T}_{\boldsymbol{x}_k}\mathcal{M}}(\boldsymbol{\eta}_k)$.

- The update is performed on the tangent space along the direction of $\text{Proj}_{\mathcal{T}_{\boldsymbol{x}_k}\mathcal{M}}(\boldsymbol{\eta}_k)$ with a step β, i.e.,

$$\bar{\boldsymbol{x}}_k = \boldsymbol{x}_k + \beta\text{Proj}_{\mathcal{T}_{\boldsymbol{x}_k}\mathcal{M}}(\boldsymbol{\eta}_k) \in \mathcal{T}_{\boldsymbol{x}_k}\mathcal{M}$$

- Since $\bar{\boldsymbol{x}}_k \notin \mathcal{M}$, it will be mapped back to the manifold by the means of a *retraction* operator, $\boldsymbol{x}_{k+1} = \text{Ret}(\bar{\boldsymbol{x}}_k)$.

For many manifolds, the projection $\text{Proj}_{\mathcal{T}_{\boldsymbol{x}_k}\mathcal{M}}(.)$ and retraction $\text{Ret}(.)$ operators admit a closed form. Interested readers may refer to [49] for more details.

For the manifold under interest, i.e., the complex circle manifold, [11] developed a framework for the optimization over this manifold. Consequently, the problem of shaping the STAF over CMC defined in (1.98) can be solved by utilizing the aforementioned framework. Precisely, at the kth iteration, (1.98) will be solved iteratively using the following steps (illustrated visually in Fig. 1.13)

1. A projection of the search direction $\boldsymbol{\eta}_k = -\nabla_{\boldsymbol{x}}\bar{\phi}(\boldsymbol{x}_k)$ onto the tangent space of the manifold at the point $\boldsymbol{x}_k$, $\mathcal{T}_{\boldsymbol{x}_k}\mathcal{S}^N$, using

$$\boldsymbol{P}_{\mathcal{T}_{\boldsymbol{x}_k}\mathcal{S}^N}(\boldsymbol{\eta}_k) = \boldsymbol{\eta}_k - \text{Re}\{\boldsymbol{\eta}_k^* \odot \boldsymbol{x}_k\} \odot \boldsymbol{x}_k \tag{1.100}$$

2. A Descent on this tangent space to update the current value of $\boldsymbol{x}_k$ on the tangent space $\mathcal{T}_{\boldsymbol{x}_k}\mathcal{S}^N$ as

$$\bar{\boldsymbol{x}}_k = \boldsymbol{x}_k + \beta\boldsymbol{P}_{\mathcal{T}_{\boldsymbol{x}_k}\mathcal{S}^N}(\boldsymbol{\eta}_k) \tag{1.101}$$

3. A Retraction of this update to $\mathcal{S}^N$ by using $\boldsymbol{R}(\boldsymbol{w}) = \boldsymbol{w} \odot \frac{1}{|\boldsymbol{w}|}$ as

$$\boldsymbol{x}_{k+1} = \boldsymbol{R}(\bar{\boldsymbol{x}}_k) \tag{1.102}$$

where $\odot$ is the element wise product and $|\boldsymbol{x}_k|$ is a vector of element wise absolute values of $\boldsymbol{x}_k$, i.e., $|\boldsymbol{x}_k| = \big[|x_k(1)|\ |x_k(2)|\ \ldots\ |x_k(N)|\big]^T$.

The algorithm utilizing these steps to solve P_2 is named as quartic-gradient-descent (QGD). It is shown [11] that the gradient of the quartic cost in Eq. (1.98) is

$$\nabla(\bar{\phi}(\boldsymbol{x})) = 2\Big(\sum_{i=1}^{M} \boldsymbol{C}_i \boldsymbol{x}\boldsymbol{x}^H \boldsymbol{C}_i^H \boldsymbol{x} + \boldsymbol{C}_i^H \boldsymbol{x}\boldsymbol{x}^H \boldsymbol{C}_i \boldsymbol{x}\Big) + 4\gamma N \boldsymbol{x} \tag{1.103}$$

Using optimization over the complex circle manifold along with the gradient in Eq. (1.103), the QGD algorithm with Armijo line search method is formally described in Algorithm 1.

Algorithm 1 Quartic-Gradient-Descent (QGD) with Armijo line search

1: **Inputs:** The interference map $p(r,h)$, $\boldsymbol{x}_0 \in \mathcal{S}^N$, scalars $\tau > 0, \beta \in (0,1), \sigma \in (0,1)$ and a pre-defined threshold value ε.
2: **Output:** A solution $\boldsymbol{x}^\star$ for optimizing $\bar{\phi}(\boldsymbol{x})$ over the complex circle manifold $\mathcal{S}^N$.
3: $i = 1$
4: **for** each $(r,h) \in \{0,...,N-1\} \times \{0,...,N_v - 1\}$ **do**
5: Compute $\boldsymbol{J}^r$ as Eq. (1.90)
6: Compute v_h and $\boldsymbol{p}(v_h)$
7: $\boldsymbol{C}_{(r,h)} = \sqrt{p(r,h)}\boldsymbol{J}^r \mathrm{diag}(\boldsymbol{p}(v_h))$
8: $i \leftarrow i+1$
9: **end for**
10: Set $k = 0$.
11: Compute the search direction using the gradient in Eq. (1.103) as $\boldsymbol{\eta}_k = -\nabla(\bar{\phi}(\boldsymbol{x}_k))$.
12: Compute the projection of the $\boldsymbol{\eta}_k$ onto the tangent space according to Eq. (1.100), and let $\boldsymbol{z} = \boldsymbol{P}_{\mathcal{T}_{\boldsymbol{x}_k}\mathcal{S}^N}(\boldsymbol{\eta}_k)$.
13: (**Armijo line search**) Find the smallest integer $m \geq 0$ such that

$$\bar{\phi}(\boldsymbol{x}_k) - \bar{\phi}(\boldsymbol{x}_k + \tau\beta^m \boldsymbol{\eta}_k) \geq \sigma\tau\beta^m \boldsymbol{z}^H \boldsymbol{z}$$

14: Compute the update of $\boldsymbol{s}_k$ onto $\mathcal{T}_{\boldsymbol{s}_k}\mathcal{S}^N$ as

$$\bar{\boldsymbol{x}}_k = \boldsymbol{x}_k + \tau\beta^m \boldsymbol{\eta}_k \tag{1.104}$$

15: Compute the next iterate $\boldsymbol{x}_{k+1}$ by retracting $\bar{\boldsymbol{x}}_k$ to the complex circle manifold by using the retraction formula $\boldsymbol{x}_{k+1} = \boldsymbol{R}(\bar{\boldsymbol{x}}_k)$.
16: **if** $\|\boldsymbol{x}_{k+1} - \boldsymbol{x}_k\|_2^2 < \varepsilon$ **then**
17: STOP.
18: **else**
19: k = k+1.
20: GOTO step (11).
21: **end if**
22: **Output:** $\boldsymbol{x}^\star = \boldsymbol{x}_k$

1.3.3.3 Analytical and Experimental Results

Convergence analysis

The cost function in Eq. (1.98) is quartic w.r.t. $\boldsymbol{x}$ and hence finding a condition on the step size μ to ensure the monotonic decrease in the cost function during the descent step on the tangent space $\mathcal{T}_{\boldsymbol{x}_k}\mathcal{S}^N$ of $\mathcal{S}^N$ at $\boldsymbol{x}_k$ is a challenging task. Empirically, a small step size shows good results (monotone decrease in the cost) during this step. Instead of working with a fixed step size, we can employ well-known backtracking line search methods that produce a variable step size that ensures the reduction in the cost. One of these methods is Armijo line search [51, 49]. In [51], Proposition 1.2.1 states that for a gradient method (such as steepest descent method) with a step size chosen by the Armijo method, every limit point of the generated sequence is a stationary point. In the setup, the Armijo line search will be used to ensure the reduction on the tangent space (an affine set) and hence the result from the aforementioned proposition can be utilized here. In other words, using Armijo line search method will produce a point on the tangent space of $\mathcal{S}^N$ at $\boldsymbol{x}_k$ with an improvement on the cost, i.e., $\bar{\phi}(\boldsymbol{x}_k) \geq \bar{\phi}(\bar{\boldsymbol{x}}_k)$. Now, the point $\bar{\boldsymbol{x}}_k$ will be on the tangent space and it will be retracted to the complex circle manifold, hence we need to investigate the effect of this operator on the change in the cost. The following lemma establishes that the cost function $\bar{\phi}(\boldsymbol{x})$ will be non-increasing through the retraction step given that the positive scalar γ satisfies a certain condition.

Lemma 1: *Let $\lambda_{\boldsymbol{B}}$ denote the largest eigenvalue of the matrix*

$$\boldsymbol{B} = \sum_{i=1}^{M} \left(vec(\boldsymbol{C}_i) vec(\boldsymbol{C}_i)^H \right)$$

If $\gamma \geq \frac{N^2}{8}\lambda_{\boldsymbol{B}}$ then $\bar{\phi}(\bar{\boldsymbol{x}}_k) \geq \bar{\phi}(\boldsymbol{x}_{k+1})$.

Proof. See [11]. □

Enabled by monotonic cost function decrease in both the Armijo line search and retraction steps (Lemma above), we have $\bar{\phi}(\boldsymbol{x}_k) \geq \bar{\phi}(\boldsymbol{x}_{k+1}) \Rightarrow \bar{\phi}(\boldsymbol{x}_k) - \bar{\phi}(\boldsymbol{x}_{k+1}) \geq 0 \Rightarrow \phi(\boldsymbol{x}_k) - \phi(\boldsymbol{x}_{k+1}) \geq 0$. Then the sequence $\{\phi(\boldsymbol{x}_k)\}_{k=0}^{\infty}$ is non-increasing and since $\phi(\boldsymbol{x}_k) \geq 0$, $\forall\, \boldsymbol{x}$ (bounded below), converges to a finite value ϕ^* is guaranteed.

Experimental results

We show performance of ambiguity function shaping algorithms, the QGD algorithm, the MBI method with a quadratic improvement (MBIQ) [12], and the coordinate iteration for ambiguity function iterative shaping (CIAFIS) [46]. Consistent with existing work [12], the number of bins N_v in the normalized Doppler frequency axis is set to 50 which produces the discrete frequencies $v_h = -\frac{1}{2} + \frac{h}{N_v}$, $h = 0, ..., 49$. The desired response $p(r,h) = 1$ for $(r,h) \in \{2,3,4\} \times \{35,36,37,38\}$, $(r,h) \in \{3,4\} \times \{18,19,20\}$, and 0 otherwise (see Figure 1.14). The signal-to-interference-ratio (SIR) provides numerical assessment of all algorithms and is defined as

$$\text{SIR} = \frac{N^2}{\sum_{r=1}^{N}\sum_{h=1}^{N_v} p(r,h)\|\boldsymbol{x}\|_2^2 g_{\boldsymbol{x}}(r, v_h)} \tag{1.105}$$

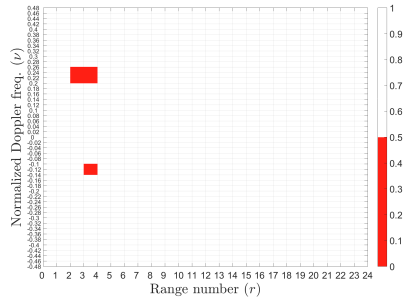
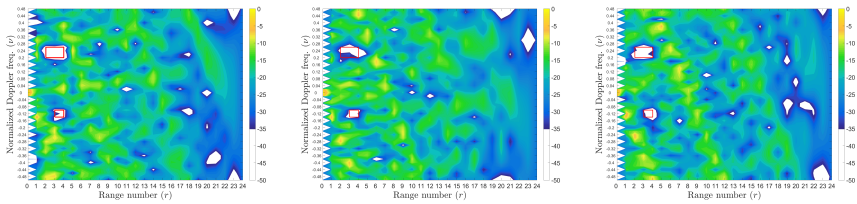


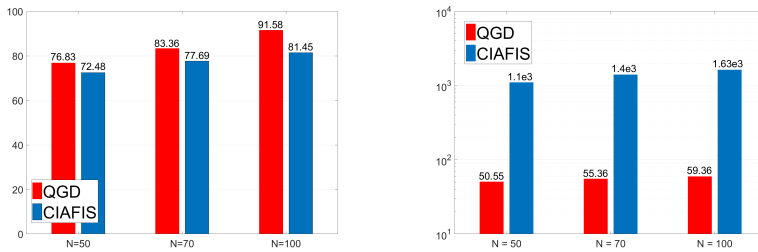
Figure 1.14 The desired STAF ($p(r,h)$).



(a) QGD (b) MBIQ (c) CIAFIS

Figure 1.15 STAF for (a) QGD, (b) MBIQ, and (c) CIAFIS for $N = 25$.

In Figures 1.15 (a) - (c), 2D plots for the STAFs are shown for QGD, MBIQ, and CIAFIS for $N = 25$. From these figures, it is evident that the response for the QGD waveform is the closest one to the desired one (the unwanted range-Doppler responses in the red rectangles are suppressed with average values around -45 dB). In



(a) Average SIR in dB (b) Average simulation time (sec)

Figure 1.16 (a) SIR average values, and (b) Average Simulation times for QGD and CIAFIS for $N = 50,\ 70,$ and 100. Each value is averaged over 100 random trials.

Figures 1.16 (a) and (b), QGD is compared against the CIAFIS method for different large values values for N varying from $50 - 100$. We focus on comparisons only against CIAFIS because [46], it has been reported that the results for MBIQ [12] for

relatively large values of N, i.e., beyond $N = 25$ take prohibitively long to generate. Figure 1.16a shows that both QGD and CIAFIS exhibit expected average SIR gains with increasing N but QGD can still outperform CIAFIS by 4-10 dB as N varies from $50 - 100$. On the other hand as Figure 1.16b reveals, the complexity of QGD increases more gracefully (slowly) with increasing N as compared to that of CIAFIS.

References

[1] Guerci JR, Guerci RM, Ranagaswamy M, et al. CoFAR: Cognitive Fully Adaptive Radar. In: IEEE Radar Conference; 2014. p. 984–989.

[2] Haykin S. Cognitive Dynamic Systems. Proceedings of the IEEE. 2006 November;94(11):1910–1911.

[3] Richards M, Scheer J, Holm W, et al. Principles of modern radar. Citeseer; 2010.

[4] Patton LK. On the Satisfaction of Modulus and Ambiguity Function Constraints in Radar Waveform Optimization for Detection. Wright State University; 2009.

[5] Trees HLV. Optimum signal design and processing for reverberation-limited environments. IEEE Transactions on Military Electronics. 1965;9(3):212–229.

[6] Patton L, Rigling BD. Modulus Constraints in Adaptive Radar Waveform Design. In: IEEE Radar Conference; 2008. p. 1–6.

[7] Maio AD, Nicola SD, Huang Y, et al. Design of phase codes for radar performance optimization with a similarity constraint. IEEE Trans Signal Processing. 2008;57(2):610–621.

[8] Aldayel O, Monga V, Rangaswamy M. Successive QCQP Refinement for MIMO Radar Waveform Design under Practical Constraints. IEEE Trans Signal Processing. 2016 July;64(14):3760–3773.

[9] San Antonio G, Fuhrmann DR. Beampattern synthesis for wideband MIMO radar systems. In: 1st IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing; 2005. p. 105–108.

[10] Aldayel O, Monga V, Rangaswmay M. Tractable Transmit MIMO Beampattern Design under a Constant Modulus Constraint. IEEE Trans Signal Processing. 2017;35(2):237–246.

[11] Alhujaili KA, Monga V, Rangaswamy M. Quartic Gradient Descent for Tractable Radar Slow-Time Ambiguity Function (STAF) Shaping. IEEE Trans Aerosp Electron Syst. 2019 August;56(2):1474–1489.

[12] Aubry A, De Maio A, Jiang B, et al. Ambiguity function shaping for cognitive radar via complex quartic optimization. Signal Process, IEEE Trans on. 2013 Nov;61(22):5603–5619.

[13] Stoica P, Li J, Zhu X. Waveform synthesis for diversity-based transmit beampattern design. IEEE Trans Signal Processing. 2008 Jun;56(6):2593–2598.

[14] He H, Stoica P, Li J. Wideband MIMO Systems: Signal Design for Transmit Beampattern Synthesis. IEEE Trans Signal Processing. 2011 February;59(2):618–628.

[15] Cui G, Li H, Rangaswamy M. MIMO Radar Waveform Design with Constant Modulus and Similarity Constraints. IEEE Trans Signal Processing. 2014 January;62(2):343–353.
[16] Aubry A, Maio AD, Farina A. Radar Waveform Design in a Spectrally Crowded Environment Via Nonconvex Quadratic Optimization. IEEE Trans Aerosp Electron Syst. 2014 April;50(2):1138–1152.
[17] Kang B, Aldayel O, Monga V, et al. Spatio-Spectral Radar Beampattern Design for Coexistence With Wireless Communication Systems. IEEE Trans Aerosp Electron Syst. 2019 April;55(2):644–657.
[18] Bekkerman I, Tabrikian J. Target detection and localization using MIMO radars and sonars. IEEE Trans Signal Processing. 2006;54(10):3873–3883.
[19] Li J, Stoica P. MIMO radar signal processing. Wiley Online Library; 2009.
[20] Mattingley J, Boyd S. Real-Time Convex Optimization in Signal Processing. IEEE Signal Processing Mag. 2010 May;27(3):50–61.
[21] Eldar YC, Luo Z, Ma W, et al. Convex Optimization in Signal Processing. IEEE Signal Processing Mag. 2010 May;27(3):19,145.
[22] Klee V, Minty G. How Good Is the Simplex Algorithm; 1970. .
[23] Karmarka N. A New Polynomial-Time Algorithm for Linear Programming. Combinatorica. 1984 December;4(4):373–395.
[24] Boyd S, Vandenberghe L. Convex Optimization. 2nd ed. Cambridge University Press; 2004.
[25] Guerci JR, Bergin JS, Guerci RJ, et al. A New MIMO Clutter Model for Cognitive Radar. In: IEEE Radar Conference; 2016. .
[26] Capon J. High Resolution Frequency-Wavenumber Spectrum Analysis. Proceedings of the IEEE. 1969 August;57(8):1408–1418.
[27] Friedlander B. Waveform Design for MIMO Radars. IEEE Trans Aerosp Electron Syst. 2007 July;43(3):1227–1238.
[28] Luo Z, Ma W, So A, et al. Semidefinite Relaxation of Quadratic Optimization Problems. IEEE Signal Processing Mag. 2010;27(3):20–34.
[29] Lobo MS, Vandenberghe L, Boyd S, et al. Applications of Second-Order Cone Programming. Linear Algebra and its Applications. 1998 Novemeber;284(1):193–228.
[30] Aubry A, Maio AD, Piezzo M, et al. Cognitive Design of the Receive Filter and Transmitted Phase Code in Reverberating Environment. IET Radar Sonar Navig. 2012;6(9):822–833.
[31] Royden H, Fitzpatrick P. Real Analysis. 4th ed. Prentice Hall; 2010.
[32] Huleihel W, Tabrikian J, Shavit R. Optimal Adaptive Waveform Design for Cognitive MIMO Radar. IEEE Trans Signal Processing. 2013 October;61(20):5075–5089.
[33] Tang B, Tang J, Peng Y. MIMO Radar Waveform Design in Colored Noise Based on Information Theory. IEEE Trans Signal Processing. 2010 September;58(9):4684–4697.
[34] Rowe W, Stoica P, Li J. Spectrally Constrained Waveform Design. IEEE Signal Processing Mag. 2014 May;31(3):157–162.

[35] Sussman SM. Least-Square Synthesis of Radar Ambiguity Functions. IRE Transactions on Information Theory. 1962 April;8(3):246–254.

[36] Gerchberg RW, Saxton WO. A Practical Algorithm for the Determination of Phase from Image and Diffraction Plane Pictures. Optik. 1972;35(2):237–246.

[37] Aldayel O, Kang B, Monga V, et al. Technical Report: Spatio-Spectral Radar Beampattern Design for Co-existence with Wireless Communication Systems. The Pennsylvania State University; 2017. Can be found at website : http://www.personal.psu.edu/osa105/BICTechReport.pdf.

[38] Williams VV. Multiplying Matrices Faster than Coppersmith-Winograd. In: Proceedings of the forty-fourth annual ACM symposium on Theory of computing; 2012. p. 887–898.

[39] Guo L, Deng H, Himed B, et al. Waveform Optimization for Transmit Beamforming with MIMO Radar Antenna Arrays. IEEE Trans Antennas Propagat. 2015 February;63(2):543–552.

[40] Guerci JR, Guerci RM, Lackpour A, et al. Joint design and operation of shared spectrum access for radar and communications. In: IEEE Radar Conference (RadarCon); 2015. .

[41] Alhujaili K, Monga V, Rangaswamy M. Transmit MIMO Radar Beampattern Design via Optimization on the Complex Circle Manifold. IEEE Trans Signal Processing. 2019;67(13):3561–3575.

[42] Zheng L, Lops M, Eldar YC, et al. Radar and communication coexistence: An overview: A review of recent methods. IEEE Signal Processing Mag. 2019;36(5):85–99.

[43] DeLong D, Hofstetter E. The design of clutter-resistant radar waveforms with limited dynamic range. IEEE Trans on Information Theory. 1969;15(3):376–385.

[44] Gregers-Hansen V. Clutter suppression using amplitude weighted waveforms. IET. 1997;.

[45] Zhang S, Huang Y. Complex quadratic optimization and semidefinite programming. SIAM Journal on Optimization. 2006;16(3):871–890.

[46] Wu L, Babu P, Palomar DP. Cognitive radar-based sequence design via SINR maximization. IEEE Trans on Signal Process. 2017;65(3):779–793.

[47] Sayed AH. Adaptive filters. John Wiley & Sons; 2011.

[48] Nocedal J, Wright SJ. Numerical optimization. 2nd ed. New York: Springer; 2006.

[49] Absil PA, Mahony R, Sepulchre R. Optimization algorithms on matrix manifolds. Princeton University Press; 2009.

[50] Kovnatsky A, Glashoff K, Bronstein M. MADMM: a generic algorithm for non-smooth optimization on manifolds. In: European Conf. on Computer Vision. Springer; 2016. p. 680–696.

[51] Bertsekas DP. Nonlinear programming. Athena Scientific Belmont; 1999.