Algebraic and Adaptive MIMO Radar

Matthew W. Morency

School of Electrical Engineering

Thesis submitted for examination for the degree of Master of Science in Technology.
Espoo 21.11.2015

Thesis supervisor:

Prof. Sergiy A. Vorobyov

Aalto University
School of Electrical Engineering

AALTO UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING

ABSTRACT OF THE
MASTER'S THESIS

Author: Matthew W. Morency

Title: Algebraic and Adaptive MIMO Radar

Date: 21.11.2015 Language: English Number of pages: 5+50

Department of Signal Processing and Acoustics

Professorship: Digital Signal Processing

Supervisor and advisor: Prof. Sergiy A. Vorobyov

Breaking causality is the main distinction of the multiple-input multiple-output (MIMO) paradigm as used for active sensing/radar. This is because the transmitting side can be optimized in many ways to manipulate the capabilities of the system. Adaptive beamforming is a fundamental problem in array-processing, communications, and radar among other fields which has once again garnered significant research interest in recent years within the MIMO paradigm. In this work, transmit adaptive beamforming algorithms are developed. One class of algorithms allows search free DOA estimation in 1D and 2D MIMO radar with an arbitrary receive array geometry while allowing transmit power gain. The other uses polynomial ideals in order to recompose the rank-constrained beamforming problem from non-convex problem to a convex one. In the first case, modern algebra is used to analyze target identifiability in the radar system. In the second, algebra reshapes the problem formulation. In both cases, performance improvements are demonstrated compared to previous methods.

Keywords: Convex Optimization, Algebra, Beamforming, MIMO, Radar

Preface

I want to thank Professor Sergiy A. Vorobyov for everything he's done. He has believed in me, supported me, encouraged me, and provided opportunities without all of which I would not be where I am today.

Otaniemi, 21.11.2015

Matthew W. Morency

Contents

1 Introduction

Radar (RAdio Detection And Ranging) is one of the most widely used remote sensing technologies in existence today. A radar system transmits microwave energy from one or more antennas (called the transmitter) into the environment and collects the reflections of these waves (echos) with one or more antennas (called the receiver) [2], [1]. At the time of transmission, everything is known about the signal being transmitted. When the signal echos are collected, however, they have been altered by the environment. The manner in which the signal echos differ from the transmitted signal contain information about the environment, and thus provides us with a tool to learn about our environment. While radio communications and radar systems are similar, herein lies a key difference: in a radar system, the environment itself is the source of information, as opposed to the transmitted signal. Radar systems can be used to detect targets at great distances [3], or through obstructions [4]. It can also be used to estimate parameters of a target such as its position relative to a known station, or its velocity [5]-[8]. Synthetic Aperture Radar (SAR) can be used to provide detailed images of landbased features such as ice flows in cases where optical imaging would not be possible (e.g. through heavy cloud cover) [9]. All this is to say that radar is a tool which is ubiquitous in its applicability.

The most common type of radar in use today is Phased Array (PA) radar [10]. In a PA-radar system, a single known signal is transmitted from multiple antennas with a known phase relationship among one another. This phase relationship is used to coherently add waveforms transmitted from different antennas in certain directions, while incoherently adding in other directions. This is a process known as transmit beamforming.

In the early 1990's, a new paradigm emerged within communications systems, known as Multiple-input-multiple-output (MIMO) [11]. In this paradigm, multiple signals are transmitted (multiple input) and collected at multiple receivers (multiple output). It was shown that a MIMO communication system could exploit spatial diversity in order to drastically increase the throughput of a communication system [12], [13]. It was some time after that the concept of MIMO radar emerged [14]. A MIMO radar system transmits several signals as opposed to a PA-radar system which only transmits one. A MIMO radar system uses waveform diversity in order to increase the number of detectable targets by a factor of N, where N is the number of waveforms.

MIMO radar systems can be divided into two main groups: monostatic, and multistatic systems. Monostatic radar systems have both the transmitter and the

receiver in the same location, and thus view only a single aspect of the target [16]. Multistatic radar systems have transmitters and receivers in at least two different locations, and as such, view different aspects of the target [15]. We concern ourselves with the monostatic case in this thesis.

Within the monostatic case, there are two further dominant approaches: full waveform diversity [16], and transmit beamspace based radar [17]-[21]. In full waveform diversity MIMO radar (henceforth referred to simply as FWD-MIMO radar) one waveform is transmitted per transmit antenna element. In a transmit beamspace based MIMO radar system (henceforth referred to TB-MIMO radar) however, fewer waveforms than transmit antenna elements are transmitted. Since we wish to use all the antennas in a system, it implies that linear combinations of the waveforms are formed by the system before the waveforms are transmitted. This in turn implies transmit gain in some directions and suppression in others, much as in Phased Array radar [18]. This property can be exploited by correctly designing the linear combinations of waveforms to achieve desired properties, e.g., a desired angular spectrum ("shape") [17], [18], [22].

The transmit gain enjoyed by TB-MIMO radar systems comes at the price of reduced virtual aperture compared to FWD-MIMO radar. That is, fewer targets are theoretically detectable in a TB-MIMO radar system than an otherwise identical FWD-MIMO system. This result comes from an eigenanalysis of the observed signal cross-correlation matrix. Specifically, a system transmitting K waveforms, the echos of which are collected at R receive antenna elements will produce a signal cross-correlation matrix of dimension $KR \times KR$. Given an observed signal cross-correlation matrix $\mathbf{Y}$ of dimension $KR \times KR$ the eigendecomposition can be decomposed as

$$\mathbf{Y} = \mathbf{Q}_s \mathbf{\Lambda}_s \mathbf{Q}_s^H + \mathbf{Q}_n \mathbf{\Lambda}_n \mathbf{Q}_n^H$$

where $\mathbf{Q}_s$,$\mathbf{Q}_n$ are the signal and noise eigenvector matrices, and $\mathbf{\Lambda}_s$,$\mathbf{\Lambda}_n$ are the diagonal matrices containing the signal and noise eigenvectors, respectively. Since $\mathbf{Q}_s$ and $\mathbf{Q}_n$ are orthogonal matrices, the above decomposition only makes sense if the number of signals in question is at most $KR - 1$. However, this does not imply that the signal parameters will be identifiable. If the noise eigenvalues are larger than any of the signal eigenvalues, how shall we decide which eigenvalues/vectors corresponds to a signal, and which correspond to noise [23]? The condition that the product of waveforms and receive antennas be larger than the number of sources is therefore a necessary condition for parameter identifiability, but far from sufficient. The signal

eigenvalues not only dictate how well the signal parameters can be estimated, but how many can be correctly identified.

If one assumes that the locations of the targets are randomly distributed in angular space according to some prior distribution, then it makes sense to direct as much energy as possible into the directions in which the targets are most likely to be located. In the physical realm there exist several examples in which this assumption is valid. For example, in a commercial airport, depending on the terrain, aircraft may have to approach the airport from certain directions in order to land. In a military context, attacks may certainly come from one direction due to operational ranges of enemy aircraft, or simply geography. In this thesis, we consider uniform prior distributions of $\mathcal{U}(\theta_1, \theta_2)$ which has a probability density function of

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{if } a \leq x \leq b \\ 0, & \text{otherwise} \end{cases}$$

The sector $\Theta = [\theta_1, \theta_2]$ is referred to as the "sector of interest." More realistic priors can (and likely should) be investigated. However, this is left for future work.

2 Preliminaries

2.1 Array Processing Fundamentals

Each algorithm in this thesis relies on two assumptions at the most basic level. The first is the narrow-band approximation [24]. In radar and communications systems, information or probing signals are not simply transmitted from an array. They are typically modulated by another waveform called a carrier waveform in order to improve the signal propagation. The propagation of an electromagnetic waveform is greatly influenced by the waveform's frequency. The modulated signal has the form

$$s(t) = A(t)cos(\omega_0 t + \phi(t)) \tag{1}$$

where $A(t)$ and $\phi(t)$ are the "baseband" signal amplitude and phase as a function of time, respectively. The narrowband approximation states that during the propagation time across the array aperture, τ, the baseband signal amplitude and phase do not vary significantly. More formally

$$A(t+\tau) \approx A(t)$$
$$\phi(t+\tau) \approx \phi(t).$$

The narrowband approximation simplifies analysis of waveform propagation across the array aperture in stating that the observed phase difference between signals at different parts of array elements is a function of propagation of the carrier waveform in the direction of travel. This leads naturally to the plane-wave approximation.

Signals transmitted from a point source, or reflected by a point target will propagate in a sphere in 3 dimensional space. Much as how the curvature of the earth is difficult to discern with the naked eye in most scenarios, it would be difficult to discern the spherical nature of a wavefront transmitted from a source at the earth's radius away from us. The wavefront would appear to be approximately flat. This is the plane-wave approximation. If the array aperture is small enough relative to the distance between the array and the source, the wavefronts appear to the array as approximately planar. Fig. 1 shows a plane wave impinging on a uniform linear array (ULA) with N elements, and antenna element spacing of dx.

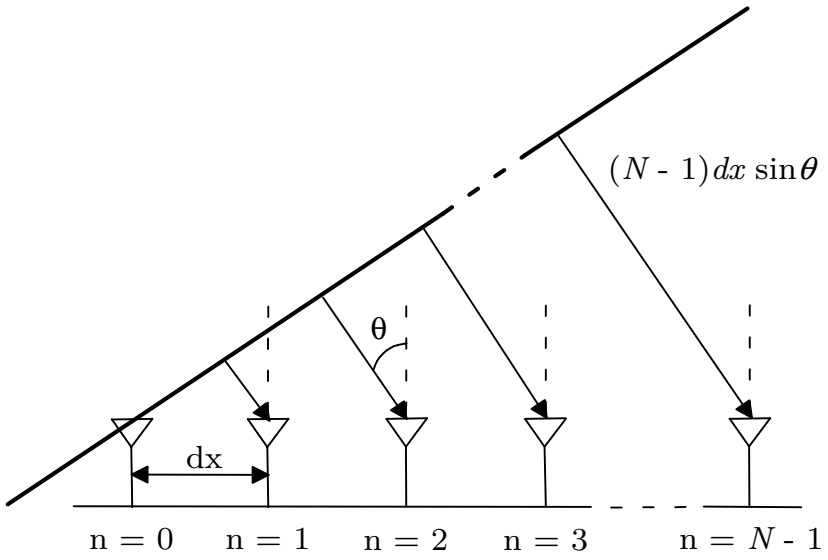

Figure 1: 1D ULA with plane wave impinging from direction θ.

The wavefront is a line of constant phase. That is, if the array in Fig. 1 were perfectly aligned with the wavefront, there would be no difference (ignoring noise) in phase between them. Phase is incurred by travel along the wave-vector. The arrows normal to the wavefront in Fig. 1 are the wavevector. With that in mind, it is simple to derive the phase difference between signals observed at two different antenna elements. The phase incurred by travelling along the wavevector a distance equal to one wavelength of the carrier waveform, λ_c, is $2\pi/\lambda_c$. Therefore, the phase incurred by travelling along the wavevector by any distance is some real multiple of this fraction. By examining Fig. 1, simple trigonometry shows that the phase difference between a signal arriving at a reference element, and one $n \cdot dx$ away is precisely $2\pi/\lambda_c \cdot n \cdot dx \cdot sin(\theta)$. Let s_0 be the signal observed at a reference element, and s_n be the signal observed at the n-th element. Then $s_n = e^{2\pi/\lambda_c ndxsin(\theta)} s_0$. We

compile these phase difference in an N dimensional vector, called the "array response vector," as follows,

$$\mathbf{a}(\theta) = \begin{bmatrix} 1 \\ e^{2\pi/\lambda_c dx sin(\theta)} \\ \vdots \\ e^{2\pi/\lambda_c (n-1) dx sin(\theta)} \end{bmatrix}. \tag{2}$$

This allows us to write the vector of observed signals from L sources at time t as

$$x(t) = \mathbf{A}(\theta)\mathbf{s}(t) + n(t) \tag{3}$$

where $\mathbf{A}(\theta)$ is an $N \times L$ matrix with columns $\mathbf{a}(\theta_l)$, $1 \leq l \leq L$, $\mathbf{s}(t)$ is an L dimensional vector containing signals from L sources at time t, and $n(t)$ is a random vector taken to represent the noise present at the sensors at time t.

Based on this signal model, we can identify up to two spherical coordinates of the source. If we transmit a signal from our array and measure its echo, the delay between transmission and reception allows us to determine the range of the target, assuming we roughly know the propagation speed of the carrier waveform in transmission media. The phase difference between array elements allows us to determine the elevation angle, depicted in Fig. 1. However, this angle lies in a plane which contains both the ULA and the target. Clearly, a ULA cannot distinguish the true plane from any other. This is the so-called "cone of ambiguity" [22]. It is so called because what we know is an angular ray, but we don't know its orientation in the ambient space. Thus, the source could lie in any ray of angle θ with respect to the array axis, which describes a cone aligned with the array. Note that in the case of a single antenna element, the cone would become a sphere.

In order to resolve this ambiguity, we need a sensor which does not lie on the array axis. A sensible choice is to consider multiple identical parallel ULAs, which we call a uniform rectangular array (URA). Fig. 2 depicts a source at direction (θ, ϕ) impinging upon a URA with antenna element spacing dx in the $\mathbf{x}$ direction, and dy in the $\mathbf{y}$ direction.

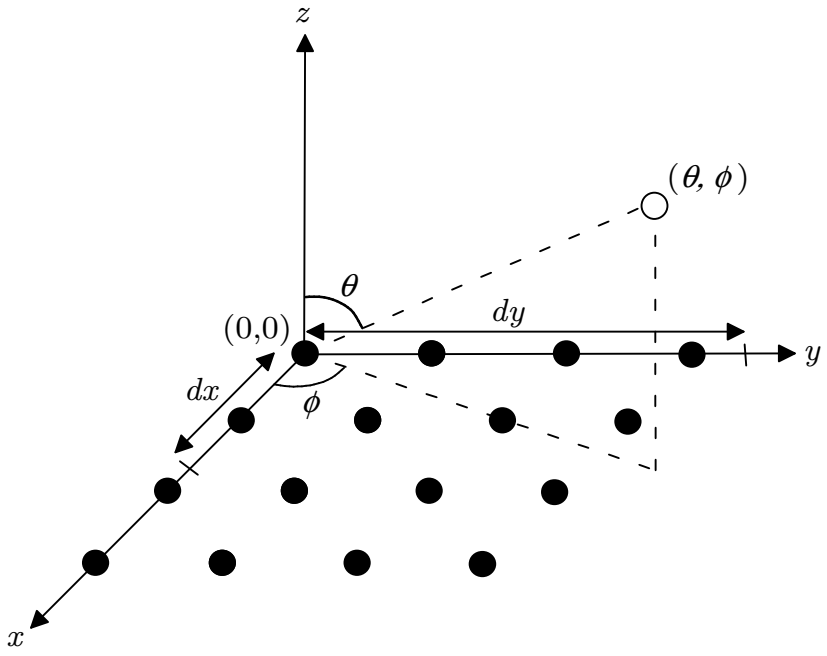

Figure 2: 2D URA with source impinging from direction (θ, ϕ).

The line connecting the source and the origin is the wave vector, and the wavefront is the plane to which the wave vector is normal. Unless the wave vector is $-\mathbf{z}$ then the signals observed at the array elements will differ by a phase which is a function of their separation in the $\mathbf{x}$ and $\mathbf{y}$ directions. This phase difference is again incurred by travel along the wave vector. As a matter of nomenclature, we say that the columns of the array are parallel with the $\mathbf{x}$ axis, and the rows are parallel with the $\mathbf{y}$ axis. Assume first that the target is located in the $\mathbf{x} - \mathbf{z}$ plane, that is, $(\theta, \phi) = (\theta, 0)$. The wavefronts will strike each row at the same time. Thus, the only phase difference between elements will be between different elements in the same column. The array response vector for each column is then exactly given by (2). Considering a target in the $\mathbf{y} - \mathbf{z}$ plane, that is $(\theta, \phi) = (\theta, \pi/2)$, yields the same steering vector, except dx becomes dy. So, the phase difference between antenna elements in the array is

so far $\psi = 2\pi/\lambda(c)sin(\theta)(f \cdot dx + g \cdot dy)$. The functions f and g must satisfy two properties. The first is that they must be continuous functions of ϕ. Secondly, they must satisfy the boundary conditions satisfy two boundary conditions.

$$f(0) = 1, f(\pi/2) = 0, \tag{4}$$

$$g(0) = 0, g(\pi/2) = 1. \tag{5}$$

A solution for this system of equations is $f = cos(\phi)$ and $g = sin(\phi)$. This solution is easily verified by considering a target in the $\mathbf{x} - \mathbf{y}$ plane. Thus, the phase difference between any given element, and one n rows and m columns away is $\psi = 2\pi/\lambda(c)sin(\theta)(mcos(\phi) \cdot dx + nsin(\phi) \cdot dy)$. Thus, the mn-th entry of the 2D array response vector is $[\mathbf{a}(\theta, \phi)]_{mn} = e^{j2\pi/\lambda_c sin(\theta)((m-1)cos(\phi)\cdot dx + (n-1)sin(\phi)\cdot dy)}$.

2.2 Algebraic Structures

This thesis makes frequent use of several algebraic structures with which the general engineering audience may not be familiar. It is therefore important to both clearly define these structures and their relationships.

Definition 2.1 *A group G is a set of elements with an operation $\bullet$ called addition with the following properties:*

Property 1. $a \bullet b \in G, \ \forall a, b \in G$

Property 2. $\exists e \in G \mid a \bullet e = e \bullet a = a, \ \forall a \in G$

Property 3. $\forall a \in G, \ \exists b \in G | a \bullet b = b \bullet a = e$

Property 4. $(a \bullet b) \bullet c = a \bullet (b \bullet c)$

A group is a set of elements with identity, and an operation $\bullet$ such that the group is closed under $\bullet$, there is an inverse element for every element in the set (again, with respect to $\bullet$), and the set is associative with respect to the operation $\bullet$. Further, if a group G is commutative with respect to $\bullet$ then G is said to be Abelian. Note, that although we refer to the operation $\bullet$ as addition, the operation $\bullet$ need not be the conventional notion of addition for the set of elements under consideration. For example, the integers are a group under multiplication. In this case $e = 1$. The integers under multiplication are also an Abelian group. Every vector space is an

Abelian group under vector addition, which is a property which will be used in this section. Building on this definition, we must next introduce the concept of a subgroup.

Definition 2.2 *A subgroup H in G is a subset of G for which all of the group properties hold with respect to $\bullet$.*

Thus, every subgroup is a group unto itself which is also contained in G. A subgroup of G which does not contain every element of G is said to be a proper subgroup of G. Abelian groups have the property that all of its subgroups are Abelian.

Definition 2.3 *A ring R is an Abelian group under an operation $\bullet$ (called addition), and a second operation $\diamond$ (called multiplication) with the following properties:*

Property 1. $\exists 1 \in R \mid 1 \diamond r = r \diamond 1 = r, \ \forall r \in R$

Property 2. $(a \diamond b) \diamond c = a \diamond (b \diamond c), \forall a, b, c \in R$

Property 3. $a \diamond (b \bullet c) = (a \diamond b) \ \bullet (a \diamond c), \forall a, b, c \in R$

Property 4. $(a \bullet b) \diamond c = (a \diamond c) \bullet (b \diamond c), \forall a, b, c \in R$

Properties 1 and 2 of Definition 2.3 taken together say that a ring R is a *monoid* under $\diamond$. This would be true even if the ring were not an Abelian group. A monoid is simply a set of elements closed under an associative operation $\diamond$ (called multiplication) with a multiplicative identity element. Properties 3 and 4 of Definition 2.3 simply state that in a ring, multiplication ($\diamond$) is left and right distributive over addition ($\bullet$). We will only consider commutative rings which have the addition property that $a \diamond b = b \diamond a, \forall a, b \in R$.

Definition 2.4 *An ideal $\mathcal{I}$ in a commutative ring R is a subgroup of R with the following property:*

Property 1. $\forall a \in \mathcal{I}, \ r \in R, a \diamond r \in \mathcal{I}, r \diamond a \in \mathcal{I}$

A relevant example of a commutative ring is the ring of univariate polynomials over a field $\mathbb{K}$, which will we denote as $\mathbb{K}[x]$ (read as "K adjoin x"). To show that this is a commutative ring we need to show the relevant properties of Definition 2.1 and Definition 2.3 hold. That the polynomials are a ring is obvious since $P(x) + Q(x) = Q(x) + P(x) = R(x)$, and $A(x)B(x) = B(x)A(x) = C(x)$ for any polynomials

A, B, C and P, Q, R with coefficients in $\mathbb{K}$. Polynomial multiplication is distributive over polynomial addition since $A(x)(B(x) + C(x)) = A(x)B(x) + A(x)C(x)$, and $(A(x) + B(x))C(x) = A(x)C(x) + B(x)C(x)$. The additive inverse of any polynomial $P(x)$ is trivially $-P(x)$. The additive and multiplicative identities of $\mathbb{K}[x]$ are 0 and 1. Lastly, $\mathbb{K}[x]$ is associative with respect to both multiplication and addition. The univariate polynomials $\mathbb{K}[x]$ with a root at x_1 form an ideal. To see this, consider a polynomial with a single root at x_1. By euclid's division algorithm, a univariate polynomial has a root at a point x_1 if and only if it can be written as $P(x) = Q(x)(x - x_1)$. Consider Definition 2.4 and let $P(x) = Q(x)(x - x_1)$. Then $P(x)R(x) = Q(x)R(x)(x-x_1) = S(x)(x-x_1)$. This shows that the set of polynomials with a root at a given point are an ideal in the ring of univariate polynomials. If we consider $\mathbb{K} = \mathbb{C}$, then this ideal is also an infinite dimensional vector space over $\mathbb{C}$. Let $A(x) = P(x)(x - x_1)$ and $B(x) = Q(x)(x - x_1)$.

$$\begin{aligned}\alpha(A(x) + B(x)) &= \alpha A(x) + \alpha B(x), \alpha \in \mathbb{C}, A(x), B(x) \in \mathcal{I} \\ &= \alpha(P(x) + Q(x))(x - x_1) \\ &= \alpha R(x)(x - x_1),\ R(x) = P(x) + Q(x)\end{aligned}$$

This vector space has a basis of $(x - x_1)\{x^i, 0 \leq i \leq \infty\}$. A finite dimensional vector space of polynomials with roots at a certain point can be constructed by restricting the degree of the polynomials to a finite number N. This vector space has a basis of $(x - x_1)\{x^i, 0 \leq i \leq N - 1\}$. We denote this space by $\mathbb{C}_{N-1}[x]Q(x)$ where $Q(x) = x - x_1$, and $\mathbb{C}_m[x]$ is the space of polynomials with degree strictly less than m. By induction, it is clear that the set of polynomials with roots at points $x_1, \ldots, x_m$ also forms an ideal in $\mathbb{C}[x]$, and a vector space over $\mathbb{C}$. In this case, $Q(x) \triangleq \prod_{i=1}^{m}(x - x_i)$.

2.3 Problem Formulation

In Section 1 the notion of energy concentration within a sector of interest in which targets are most likely to be located has been introduced. Given a target prior distribution of $\mathcal{U}(\theta_1, \theta_2)$, it is obvious that the ideal beampattern with which to illuminate these targets is one in which all the energy is concentrated in the sector of interest Θ, with no energy radiated anywhere else. Let this ideal beampattern be $G_d(\theta)$. As was said in Section 1, TB-MIMO transmits linear combinations of basis

waveforms in order to concentrate energy. Thus the transmitted signal at time t, and direction θ is

$$\mathbf{s}(t,\theta) = \mathbf{a}^H(\theta)\mathbf{W}\boldsymbol{\psi}(t) \tag{6}$$

where $(\cdot)^H$ is the Hermitian transpose operator, and $\mathbf{W} \in \mathbb{C}^{N\times K}$ is the transmit beamforming matrix to be designed.

Using (6), the magnitude of the beampattern in direction θ is given by the innerproduct of $\mathbf{s}(t,\theta)$ with itself

$$\begin{aligned} G(\theta) &= \int_T \mathbf{s}(t,\theta)\mathbf{s}^H(t,\theta)dt \\ &= \mathbf{a}^H(\theta)\mathbf{W}\left(\int_T \boldsymbol{\psi}(t)\boldsymbol{\psi}^*(t)dt\right)\mathbf{W}^H\mathbf{a}(\theta) \\ &= \mathbf{a}^H(\theta)\mathbf{W}\mathbf{W}^H\mathbf{a}(\theta) = \|\mathbf{W}^H\mathbf{a}(\theta,\phi)\|^2 \end{aligned} \tag{7}$$

where $\int_T \boldsymbol{\psi}(t)\boldsymbol{\psi}^*(t)dt = \mathbf{I}_k$ as $\psi(t)$ are chosen to be orthogonal, and T is the pulse length, $(\cdot)^*$ is the complex conjugate, and $\|\cdot\|$ denotes the Euclidean norm. The dominant approach to designing the transmit beampattern since [22] has been through convex programming, utilizing a paradigm called Semidefinite Relaxation (SDR), which we describe in this section. Given a desired beampattern $G_d(\theta)$, and an achievable one $G(\theta)$, a sensible objective function is given by

$$\min_{\mathbf{X}} \|G_d(\theta) - G(\theta)\|_* = \min_{\mathbf{X}} \|G_d(\theta) - \mathbf{a}^H(\theta)\mathbf{X}\mathbf{a}(\theta)\|_*$$

where $\|\cdot\|_*$ is some norm. Let us consider the Chebyshev norm whereby the objective function becomes

$$\min_{\mathbf{X}} \max_{\theta} \|G_d(\theta) - \mathbf{a}^H(\theta)\mathbf{X}\mathbf{a}(\theta)\|$$

where $-\pi/2 \le \theta \le \pi/2$. This, on its own, is an unconstrained minimization problem. However, the solution is likely to violate constraints of the physical system for which the beamforming matrix is being designed. For example, every physical system must have some limit on the total amount of energy used. Moreover, it may be important to constrain the amount of energy used by certain antenna elements, or constrain the amount of energy used by all antenna elements to be equal. Since $\mathbf{X} \triangleq \mathbf{W}\mathbf{W}^H$, the i-th diagonal element of $\mathbf{X}$ is clearly $\sum_{k=1}^{K}[w_k]_i^2$. If the transmitted signals are chosen

to have constant modulus 1, the i-th diagonal element of $\mathbf{X}$ is the amount of energy used by all signals at antenna element i. Let $\mathbf{B}_i$ be an $N \times N$ matrix of all zeros, except for in the i-th diagonal position. Then $\mathbf{X}_{i,i} = \text{tr}(\mathbf{B}_i\mathbf{X})$, where $\text{tr}(\cdot)$ is the matrix trace operator, which allows us to express constraints on the energy used by a specific antenna element as

$$\text{tr}(\mathbf{B}_i\mathbf{X}) = \delta, \ \delta \in \mathbb{R} \tag{8}$$

where delta is the desired desired amount of energy. Constraints of this form are linear in $\mathbf{X}$ and thus convex. Depending on the requirements of the physical system, $\mathbf{B}_i$ can take a variety of forms, as can the constant δ. For example, let the total amount of energy available to the system be E. If we want to constrain the total amount of energy used by the radar system, then $\mathbf{B} = \mathbf{I}_N$ and (8) becomes

$$\text{tr}(\mathbf{X}) = E.$$

A uniform power distribution among all antenna elements, where $\delta = E/N$ can be written as

$$\text{tr}(\mathbf{B}_i\mathbf{X}) = E/N \ 1 \leq i \leq N.$$

Lastly, given that we want to use fewer waveforms than transmit antennas, the rank of $\mathbf{X}$ must be constrained. Taken together, the rank-constrained beamforming problem, with uniform power constraint is written as

$$\min_{\mathbf{X}} \max_{\theta} |G_d(\theta) - \text{tr}\{\mathbf{C}\mathbf{X}\}| \tag{9}$$

$$\text{s.t. } \text{rk}\{\mathbf{X}\} = K \tag{10}$$

$$\text{tr}\{\mathbf{B}_i\mathbf{X}\} = \frac{E}{N}, \ i \in 1, \cdots, N \tag{11}$$

$$\mathbf{X} \succeq 0 \tag{12}$$

where $\text{rk}(\cdot)$ returns the rank of a matrix, and constraint (12) is required given that energy is always a positive quantity, by definition. Unfortunately, problem (9)-(12) is non-convex, and therefore, difficult to solve. This is entirely due to constraint (10), as the set of matrices of a given rank is non-convex. To wit, the sum of two matrices

of rank K is not necessarily rank K. Following the approach in [25], [22] adopted the SDR approach. That is, constraint (10) is ignored, and the resulting convex problem is solved. Once the problem is solved, the issue of extracting $\mathbf{W}$ from $\mathbf{X}$ remains. To do this, the authors in [19] adopted the approach of randomization. Random collections of K vectors are drawn from the columnspace of $\mathbf{X}^\star$, and among these the best is chosen according to overall fit to the beampattern corresponding to $\mathbf{X}^\star$. If the candidate solution chosen is not a feasible point of the original optimization problem, then it is mapped to a nearby feasible point.

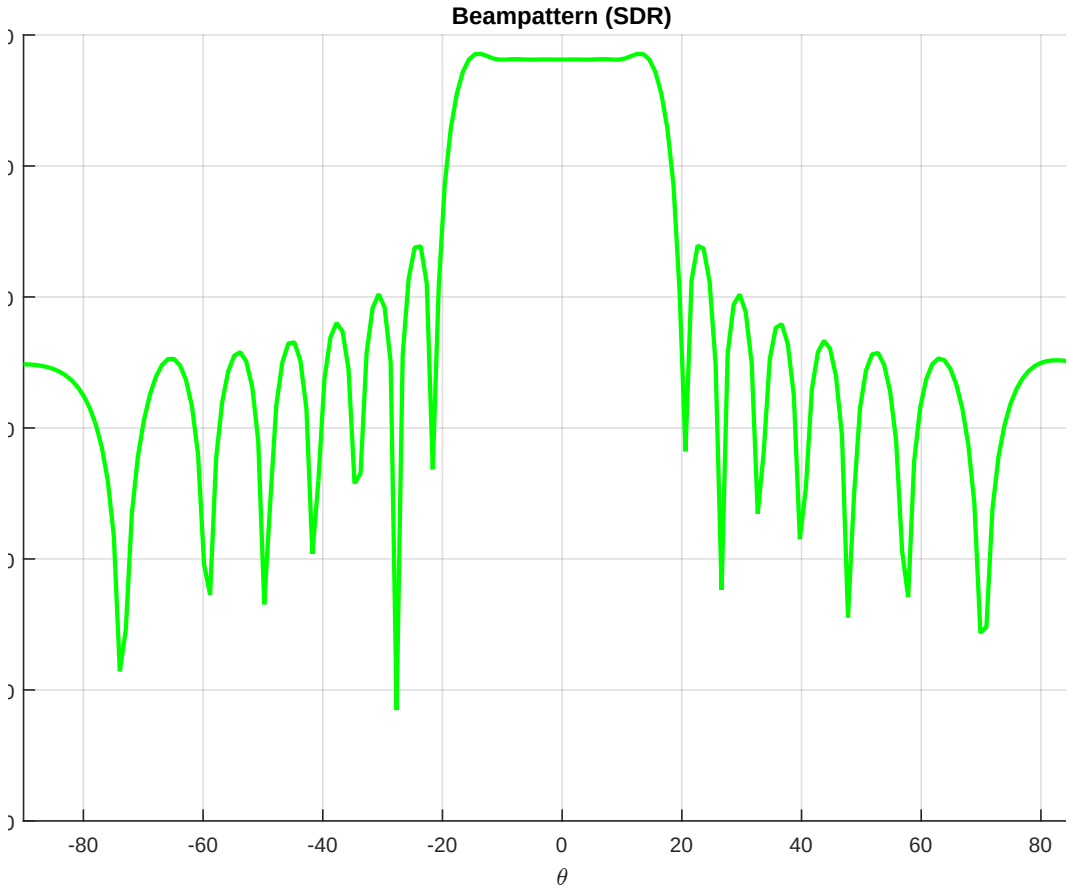

Figure 3: Beampattern designed using method of [22] for a sector of $[-15^o, 15^o]$.

Fig. 3 demonstrates the performance of the method proposed in [22] with a sector of width 30^o, centered at $\theta = 0$.

There are two key problems with the relaxation based approach. The first is that there are no guarantees on the optimality of the solution with respect to the solution of (9)-(12). In [25], statistical bounds on the optimality of the solution of the relaxed problem were given, however, these bounds were only shown to hold

for the rank-1 case. There is no reason to assume that the relaxation bound in the rank-K case should remain unchanged from the rank-1 case. The second problem is that the relaxation based approach disregards any and all underlying algebraic structures which could be used to solve, or more deeply understand the problem. In subsection 3.2, a new algorithm will be proposed which exploits the underlying algebraic structure of the rank constrained beamforming problem.

3 Algebraic Rank-Constrained Beamforming

3.1 System Model

With the rank constrained beamforming problem now formulated, consider a ULA consisting of N antenna elements with an inter-element spacing of $\lambda_c/2$ where λ_c is the wavelength of the associated carrier waveform. Thus,the elements of the array response vector become $[\mathbf{a}(\theta)]_n \triangleq e^{j2\pi/\lambda_c(n-1)d_x \sin(\theta)},\ \theta \in [-\pi/2, \pi/2],\ n \in 1, \cdots, N$. A linear combination of K orthogonal baseband waveforms $\boldsymbol{\psi}(t) \triangleq [\psi_1(t), \ldots, \psi_K(t)]^T$ is transmitted, where $K < N$, in order to concentrate energy over a desired sector $\Theta = [\theta_1,\ \theta_2]$. Our challenge is to circumvent constraint (10) in order to efficiently solve the transmit beamforming problem. However, we wish to avoid having to rely on randomization methods, as guarantees on the optimality of the solution provided by them have not been shown to hold. Further, as we show in the next section, they may not even be necessary, and perhaps obfuscate the true nature of the problem.

3.2 Restriction

Consider a matrix $\mathbf{A} \in \mathbb{V}^{N\times L}$, where $\mathbb{V}^{N\times L}$ is the set of $N \times L$ vandermonde matrices, whose columns are the powers from 0 to $N-1$ of L complex generators α_l, and $L < N$. Let $\alpha_l = e^{j2\pi/\lambda_c d_x \sin(\theta_l)}$. This matrix can be used to implicitly enforce constraints on the beampattern at locations $\theta_l,\ l \in 1, \cdots, L$. For example, the equations

$$\mathbf{a}^H(\theta_l)\mathbf{W}\mathbf{W}^H\mathbf{a}(\theta_l) = 0,\ l \in 1, \cdots, L. \tag{13}$$

can be written as

$$\text{diag}\left(\mathbf{A}^H\mathbf{W}\mathbf{W}^H\mathbf{A}\right) = \mathbf{0} \tag{14}$$

where $\mathbf{0}$ is the zero vector, and diag$(\cdot)$ is an operator which takes the elements of the diagonal of a matrix and puts them into a vector. Constraints of this type can be used to control the location of nulls in the beampattern. The equations in (13) are sum of squares polynomials by construction.

For some θ_l, let $\mathbf{y} \triangleq \mathbf{a}^H(\theta_l)\mathbf{W}$, then each of the equations in (13) can be rewritten as $\sum_{k=1}^{K} |\mathbf{y}_k|^2 = 0$, which has only the obvious solution that $\mathbf{y}_k = 0$ where $\mathbf{y}_k = \mathbf{a}^H(\theta_l)\mathbf{w}_k, \forall k \in 1, \cdots, K$. Equations (13) are therefore satisfied if and only if $\mathcal{C}(\mathbf{W}) \subset \mathcal{N}(\mathbf{A}^H)$, where $\mathcal{C}(\cdot)$ and $\mathcal{N}(\cdot)$ denote the column and nullspace of a matrix.

Since we want $\mathbf{W}$ to be tall and full rank, the equations (13) constrain K to be

less than or equal to the dimension of the nullspace of $\mathbf{A}^H$. Thus, a concise expression for $\mathcal{N}(\mathbf{A}^H)$ becomes important. Using the definition of the nullspace

$$\mathcal{N}(\mathbf{A}^H) \triangleq \{\mathbf{w} \in \mathbb{C}^N | \mathbf{A}^H \mathbf{w} = \mathbf{0}\} \tag{15}$$

it is easy to show that every vector in $\mathcal{N}(\mathbf{A}^H)$ describes the coefficients of a polynomial of degree $N-1$ with roots at $\alpha_1^*, \cdots, \alpha_l^*$, that is,

$$\mathbf{A}^H \mathbf{w} = \mathbf{0} \iff \sum_{i=0}^{N-1} (\alpha_l^*)^i \mathbf{w}_i = 0, \ \forall l \in 1, \cdots, L. \tag{16}$$

A polynomial $P(x)$ has a root at some point α if and only if $(x-\alpha)$ is a factor of $P(x)$ [26]. By induction, it can be seen that a polynomial $P(x)$ has roots at points $\alpha_1^*, \cdots, \alpha_l^*$ if and only if $P(x) = Q(x)B(x)$ where

$$Q(x) \triangleq \prod_{l=1}^{L} (x - \alpha_l^*). \tag{17}$$

From (16) and (17), $\mathcal{N}(\mathbf{A}^H)$ can be expressed as

$$\mathcal{N}(\mathbf{A}^H) = Q(x)\mathbb{C}_{N-L}[x]. \tag{18}$$

As was discussed in the previous section, every member of this vector space is also a member of the polynomial ideal generated by points $V = \{\alpha_1^*, \cdots, \alpha_l^*\}$. Let $\mathbf{q} \triangleq [(-1)^{L-1} s_{L-1}, (-1)^{L-2} s_{L-2}, \cdots, (-1)s_1, 1]^T$, where $s_1, \cdots, s_{L-1}$ are the elementary symmetric functions of $\alpha_1^*, \cdots, \alpha_l^*$. The k-th elementary symmetric function in L variables (in this case, $\alpha_1^*, \cdots, \alpha_l^*$) is the sum of the products of the k subsets of those L variables. For example, if $L = 3$ then

$$\begin{aligned} s_3 &= \alpha_1^* + \alpha_2^* + \alpha_3^* \\ s_2 &= (\alpha_1\alpha_2)^* + (\alpha_2\alpha_3)^* \\ s_1 &= (\alpha_1\alpha_2\alpha_3)^* \end{aligned}$$

Let $\mathbf{q}' \triangleq [\mathbf{q}, 0, \cdots, 0]^T \in \mathbb{C}^N$. Then a basis of $\mathcal{N}(\mathbf{A}^H)$ has the matrix representation

$$\mathbf{Q} = [\mathbf{q}', \mathbf{q}_1', \cdots, \mathbf{q}_{N-L-1}'] \tag{19}$$

where $\mathbf{q}_i'$ is the i-th cyclical shift of $\mathbf{q}'$. For a polynomial $Q(x) = a_0 + a_1 x + \cdots + a_L x^L$,

with roots $\alpha_1^*, \cdots, \alpha_l^*$, Viète's formulas yield the coefficiens $a_0, \cdots, a_{L-1}$ as

$$\begin{aligned} s_1(\alpha_1^*, \cdots, \alpha_l^*) &= -\frac{a_{L-1}}{a_L} \\ s_2(\alpha_1^*, \cdots, \alpha_l^*) &= \frac{a_{L-2}}{a_L} \\ &\vdots \\ s_L(\alpha_1^*, \cdots, \alpha_l^*) &= (-1)^L \frac{a_0}{a_L}. \end{aligned}$$

Thus, the elements of the vector $\mathbf{q}$ are the coefficients of $Q(x)$, which are given as a function of the roots of $Q(x)$ by Viète's formulas, with $a_L = 1$. Armed now with both the structure of $\mathcal{N}(\mathbf{A}^H)$ and a basis for this vector space in the columns of $\mathbf{Q}$, the following lemma becomes clear.

Lemma 1 *The columns of any matrix of the form* $\mathbf{QR}$*, where* $\mathbf{R}$ *is some matrix conformable to* $\mathbf{Q}$*, will both remain in* $\mathcal{N}(\mathbf{A}^H)$ *and the associated polynomial ideal* $\mathcal{I}(V)$*.*

That is, the polynomials described by the columns of $\mathbf{QR}$ will continue to have roots at V. Therefore, now that we have an exact procedure for constructing $\mathbf{Q}$, we can use this property to manipulate the beampattern by choosing the set V. Using the cyclic property of the matrix trace operator, an optimization problem whose feasible set is a polynomial ideal generated by the set V is realized by rewriting (9)–(12) as

$$\min_{\mathbf{X}} \max_{\theta} \left| G_d(\theta) - \mathrm{tr}\{\mathbf{a}^H(\theta)\mathbf{Q}\mathbf{X}\mathbf{Q}^H\mathbf{a}(\theta)\} \right| \tag{20}$$

$$\mathrm{tr}\{\mathbf{B}_i\mathbf{Q}\mathbf{X}\mathbf{Q}^H\} = \frac{E}{N}, \ i \in 1, \cdots, N \tag{21}$$

$$\mathbf{X} \succeq 0 \tag{22}$$

By letting $\mathbf{D} \triangleq \mathbf{Q}^H\mathbf{a}(\theta)\mathbf{a}^H(\theta)\mathbf{Q}$ and $\mathbf{H}_i \triangleq \mathbf{Q}^H\mathbf{B}_i\mathbf{Q}$, and using the cyclical property of the matrix trace operator again, (20)–(22) becomes the following problem in $\mathbf{X}$

$$\min_{\mathbf{X}} \max_{\theta} |G_d(\theta) - \mathrm{tr}\{\mathbf{D}\mathbf{X}\}| \tag{23}$$

$$\mathrm{tr}\{\mathbf{H}_i\mathbf{X}\} = \frac{E}{N}, \ i \in 1, \cdots, N \tag{24}$$

$$\mathbf{X} - \gamma\mathbf{I}_{N-L} \succeq 0 \tag{25}$$

where γ is a real-valued tuning parameter, which serves to constrain the minimum eigenvalue of $\mathbf{X}$. Constraint (10) has been replaced by (25). Since this problem is

convex in $\mathbf{X}$, it can be solved optimally in polynomial time by using, for example, interior point methods [27]. Note that instead of looking for rank K solutions for $\mathbf{X}$ in the cone of PSD matrices, we search for full-rank solutions in a space of dimension K parameterized by V. However, as will be examined in the next section, a judicious selection of K will ensure that $X^\star$ will be positive definite and well-conditioned enough to use, for example, the Cholesky decomposition to recover $\mathbf{W}$ as

$$\mathbf{W} = \mathbf{QR} \tag{26}$$

$$\mathbf{X}^\star = \mathbf{RR}^H. \tag{27}$$

Lastly, the formulation of $\mathbf{Q}$ allows for exact control over the nulls in the beampattern. The set V can be chosen arbitrarily. For example, V can contain multiple-roots which would correspond to a particularly deep null, or roots placed in a uniform, but tightly packed spread, which would correspond to a deep and wide null. The only thing that must be changed in either situation is a recalculation of the symmetric functions.

The proposed algorithm relies on the algebraic structure imposed on the rank-constrained beamforming problem by the structure of the transmit array. The polynomial structure extends directly from the definition of the steering vector $[\mathbf{a}(\theta)]_i = \alpha^{i-1}$ which, with α as defined in this section, corresponds to a ULA with array spacing $\lambda_c/2$. However, other array structures also enforce this polynomial structure. For example, consider an array with elements occupying points along a 1-D lattice, and assume the lattice points have the form $q \cdot \lambda_c,\ q \in \mathbb{Q}$. If this array is fully populated, then clearly, the steering vector is a polynomial since

$$\begin{aligned}
[\mathbf{a}(\theta)]_i &= e^{j2\pi/\lambda_c q \cdot \lambda_c (i-1)\sin(\theta)} \\
&= e^{j2\pi q(i-1)\sin(\theta)} = \gamma^{i-1} \\
\gamma &= e^{j2\pi q \sin(\theta)}.
\end{aligned}$$

Further, if the i-th position of the array is not occupied by an antenna element, the steering vector still describes a polynomial, only with a coefficient of zero for the $i-1$-th power of γ. Therefore, any array whose elements occupy arbitrary positions on a lattice with vertices of the form $q \cdot \lambda_c,\ q \in \mathbb{Q}$ will have a steering vector with a polynomial structure. This is significant since the only points in space any linear array can occupy are of the form $r \cdot \lambda_c, r \in \mathbb{R}$, and $\mathbb{Q}$ is dense in $\mathbb{R}$. That is, for any

two numbers $x, y \in \mathbb{R}$ there is a rational number q between them. It is clear that the steering vector of an arbitrary linear array can thus be approximated by one with a polynomial structure, thus enabling the method proposed in this section.

3.3 Selection of Optimal Number of Waveforms

The rank constraint in previous section raises a further question: what should K be? Depending on the problem, the selection of K can either be obvious, or require careful consideration. In the case where we wish to design a single null, K can be chosen in an iterative fashion. That is, the problem can be executed several times (if required) with one more multiple root added each time to reach the desired level of null depth/precision, and overall beampattern fit. The algorithm would be started with $K = N - 1$ and proceed until some criterion is violated (e.g. overall beampattern fit). With each addition root, $K_{i+1} = K_i - 1$, and thus the number of degrees of freedom will decrease by one in each iteration. As a consequence, the overall beampattern fit will worsen. Once the desired tradeoff between null depth and beampattern fit is reached, the algorithm terminates. However, in the case where we are designing a main lobe, the scenario becomes more complicated. This is because we are no longer setting multiple roots at the same point, but rather setting many nulls in several different locations. Moreover, the location, and number of the roots affect the condition number of the designed signal cross-correlation matrix $\mathbf{X}^\star$, which can prohibit the use of the Cholesky decomposition to recover $\mathbf{W}$. Some general findings are useful in selecting K. A wider sector requires the use of more waveforms and hence, fewer nulls. This fact is demonstrated by the following two figures.

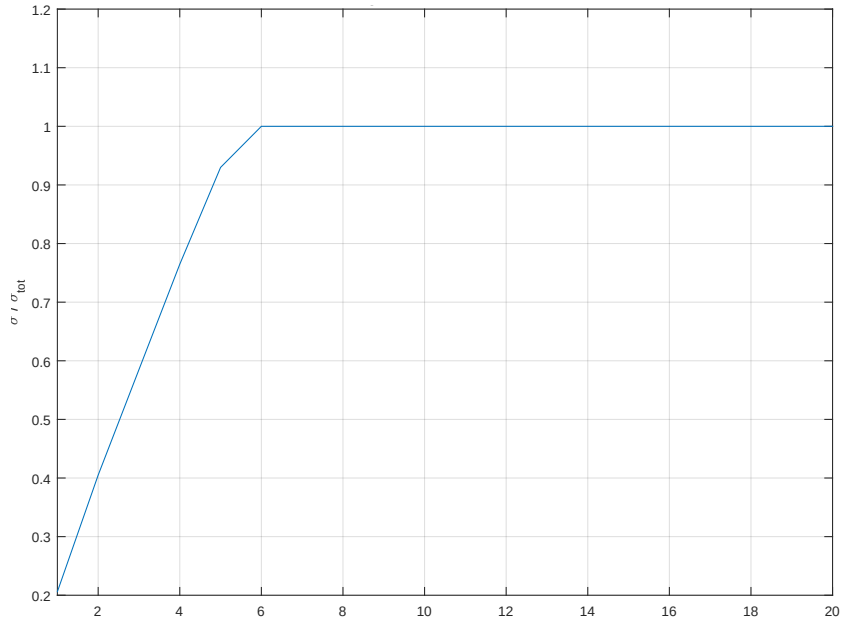

Figure 4: Distribution of energy among eigenvalues of signal cross-correlation matrix for sector width of 30 degrees.

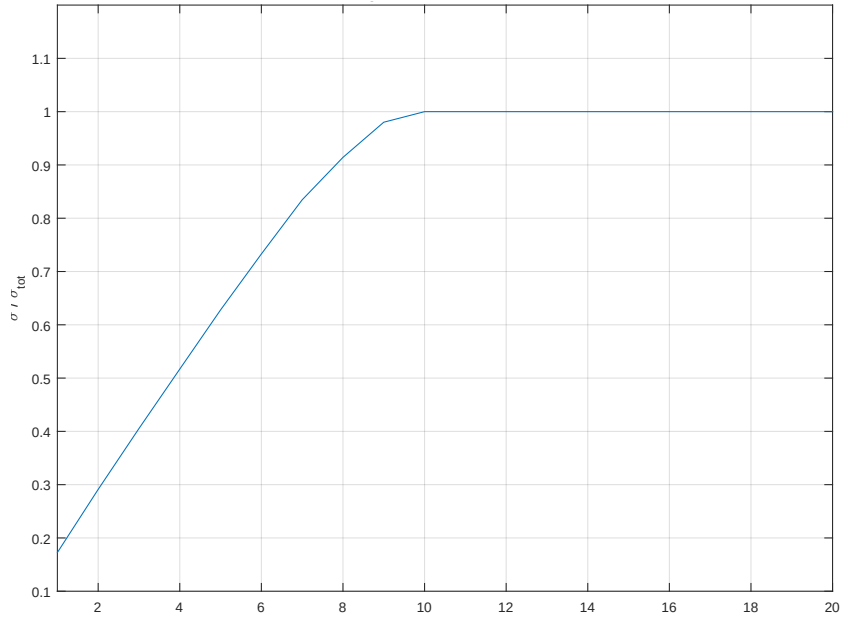

Figure 5: Distribution of energy among eigenvalues of signal cross-correlation matrix for sector width of 60 degrees.

It can be seen by the distribution of the eigenvalues of the two signal cross-correlation matrices that they could both be approximated with matrices of lower rank. However, while the cross-correlation matrix designed for a sector of width 30

degrees has 6 dominant eigenvalues, the cross-correlation matrix designed fora sector of width 60 degrees has 10 dominant eigenvalues. Both matrices were designed for use with the same ULA consisting of 20 elements. The first matrix was designed to transmit 12 waveforms, while the second was designed to transmit 14 waveforms. In both cases, it's clear that both matrices were designed for more waveforms than was entirely necessary.

It would be better, therefore, to understand before designing the matrix how many waveforms should be used in the design. One sensible criteria in beamforming would be to maximize the ratio of energy within the desired sector, to the total amount of energy radiated in the whole angular space. This figure has been considered in [17]. To wit, the ratio

$$\Gamma_k = \frac{\int_T \int_\Theta |\mathbf{w}_k^H \mathbf{a}(\theta)|^2 \mathrm{d}\theta \mathrm{d}t}{\int_T \int_{-\pi/2}^{\pi/2} |\mathbf{w}_k^H \mathbf{a}(\theta)|^2 \mathrm{d}\theta \mathrm{d}t} \tag{28}$$

should be maximized for each $\mathbf{w}_k$. Neither the numerator or denominator in (28) are functions of time, so the outside integral can be ignored. Noting the fact that $\mathbf{w}_k$ is not a function of θ allows us to represent (28) as

$$\Gamma_k = \frac{\mathbf{w}_k^H \mathbf{A} \mathbf{w}_k}{\mathbf{w}_k^H \mathbf{D} \mathbf{w}_k} \tag{29}$$

where $\mathbf{A} \triangleq \int_\Theta \mathbf{a}(\theta)\mathbf{a}^H(\theta)\mathrm{d}\theta$, and $\mathbf{D} \triangleq \int_{-\pi/2}^{\pi/2} \mathbf{a}(\theta)\mathbf{a}^H(\theta)\mathrm{d}\theta$. In the case of a vandermonde structure for $\mathbf{a}(\theta)$ (which is presumed by the algorithm in this section) $\mathbf{D} = \pi \mathbf{I}_N$. To simply analysis, let $\|\mathbf{w}_k\| = 1$, and $\mathbf{w}_i \neq \mathbf{w}_j,\ i \neq j,\ i, j \in 1, \ldots, K$. Thus, maximizing (28) is equivalent to maximizing the numerator in (29). With the preceeding assumptions, the numerator will achieve its maximum when w_k are aligned with the eigenvectors of $\mathbf{A}$. Further, if $\mathbf{A}$ has dominant eigenvalues, then most of the energy represented by the numerator of (28) will be the product of w_k which are aligned with the corresponding eigenvectors.

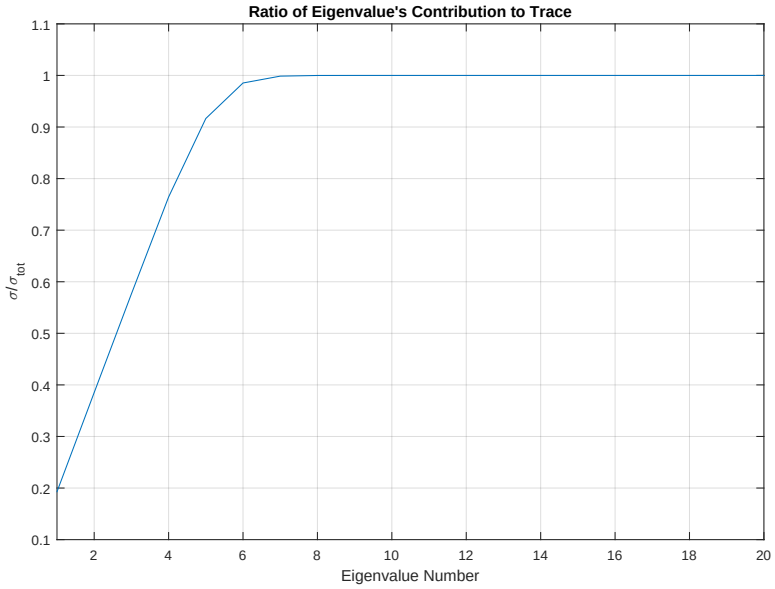

Figure 6: Distribution of energy among eigenvalues of **A** matrix for sector width of 30 degrees.

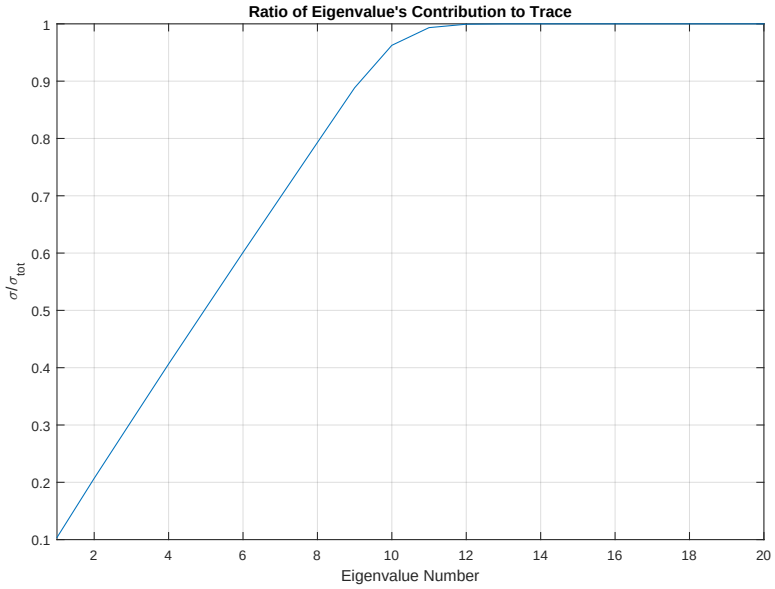

Figure 7: Distribution of energy among eigenvalues of **A** matrix for sector width of 60 degrees.

As we see from Fig. 6 and 7, the first 6 and 10 eigenvalues account for close to 100% of the total energy in the numerator of (29) for sectors of 30 and 60 degrees respectively. This corresponds exactly with what we know of the distribution of

eigenvalues of $\mathbf{Q}\mathbf{X}^{\star}\mathbf{Q}^{H}$. However, we know our sector of interest a-priori, and thus, the number of dominant eigenvalues of $\mathbf{A}$ gives us a useful proxy to determine K. Indeed, from the preceeding analysis, it is obvious that only as many $\mathbf{w}_k$ as there are dominant eigenvalues of $\mathbf{A}$ can make any significant contribution to the distribution of energy in the desired sector, as the columns of $\mathbf{W}$ are linearly independent by construction, and in order to maximize energy within the sector, they will, at least somewhat, align themselves with the eigenvectors of $\mathbf{A}$. This fact is borne out by the norm of the individual columns of $\mathbf{W}$. In both cases there are exactly 6 and 10 beamforming vectors in $\mathbf{W}$ with $\|\mathbf{w}_k\| \geq 1$, and the columns of $\mathbf{W}$ are close to orthogonal.

3.4 Generalized Sidelobe Canceller (GSC)

In [28], Griffiths and Jim proposed an algorithm that has become the standard approach to linearly constrained adaptive beamforming. Their approach, referred to within [28] as the generalized sidelobe canceller (GSC), uses a two step procedure in order to produce a beampattern with a fixed mainlobe and suppressed sidelobes. In the first step, a beampattern with a fixed response in the "look" direction is produced by convolving a vector of constraints with a normalized beamforming vector with the desired mainlobe response. In the second step, the signals in the "look" direction are blocked out, while the output power is minimized. If $y_w(k)$ is the signal corresponding to the first part of the algorithm, and $y_n(k)$ is the signal corresponding to the second, then the overall beamformed signal is $y(k) = y_w(k) - y_n(k)$, where k in this case is some discrete time index. In order to block the signal in the "look" direction, the authors use the assumption of ideal steering. To wit, they assume that the signal impinges on the broadside of the array. If we assume a ULA of M antennas, the signal at the m-th antenna is $x_m(k) = s(k) + n_m(k)$. The assumption of ideal steering allows us to state that the desired signal $s(k)$ will be identical at each antenna (differing only by noise), and thus a sufficient condition for blocking of the desired signal is $\mathbf{w}_i^T\mathbf{1} = 0$, where $\mathbf{w}_i$ is the blocking beamforming vector.

Using the definition of (6), it is clear that $\mathbf{a}(0) = \mathbf{1}$, and thus, any beamforming vector satisfying $\mathbf{w}^T\mathbf{1} = 0$ will have a null at $\theta = 0$. Equivalently, $\mathbf{w}$ contains the coefficients of a polynomial with at least one root at $\alpha(0) = e^{j2\pi/\lambda_c d_x \sin(0)} = 1$. The $M-1$ vectors $\mathbf{w}_m$ are compiled into an $(M-1) \times M$ matrix $\mathbf{W}_B$ with rows $\mathbf{w}_m^T$. It is clear that all $\mathbf{w}_m \in (x-1)\mathbb{C}_{M-1}[x]$, and thus, lie in the polynomial ideal $\mathcal{I}(1)$. The underlying algebraic structure allows several generalizing statements to be made. Instead of requiring the assumption of ideal steering, instead require $\mathbf{w}^T a(\theta_0) = 0$.

That $\mathbf{w} \in (x - \alpha(\theta_0))\mathbb{C}_{M-1}[x]$ is a necessary and sufficient condition for $s(k)$ to be blocked. Requiring that $\mathbf{w}_i$ be linearly independent for all $1 \leq i \leq M-1$ implies that the polynomials all of the polynomials only share a single root at $\alpha(\theta_0)$. If multiple signals impinge upon the array from directions $\theta_l, 1 \leq l \leq L$, and we wish to simultaneously block each of them in order to implement the GSC, the row-space of $\mathbf{W}_B \subset Q(x)\mathbb{C}_{M-2}[x]$ where $Q(x) = \prod_{l=1}^{L}(x - \alpha(\theta_l))$. In this case, we can have as many as $M - L$ vectors $\mathbf{w}_m$. In [28], the authors give an example of the matrix $\mathbf{W}_B$, for $M = 4$.

$$\mathbf{W}_B = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \tag{30}$$

The rows of (30) are in fact a basis for $(x-1)\mathbb{C}_{M-1}[x]$, the entries of which are predicted exactly by Viète's formulas. It is exactly $\mathbf{Q}^T$ for the case of $M = 4$, $L = 1$, and $\theta_l = 0$. However, as we've seen, the GSC can be interpreted as a special case of the method presented in this section.

3.5 Simulation Results

Throughout the simulations, we assume a ULA of 20 elements acting as a transmitter. As the focus of this paper is transmit beamforming, noise has not yet entered the model. We assume that our transmitter has elements spaced at exactly $\lambda_c/2$. The SDR approach to the rank-constrained beamforming problem will be adopted as a comparison to the proposed method in all examples.

3.5.1 Example 1: Mainlobe

In this example, the sector of interest is $\Theta = [-15^o, 15^o]$. We choose $K = 4$, and thus must set $N - K = 16$ roots. The only question is what set V will lead to the best beampattern. It makes sense to select α_l^* such that they correspond to nulls which are uniformly distributed outside of the passband and transition regions. In order to suppress the sidelobes, we set 8 roots per side of the mainlobe at $V = \{\pm 75^o, \pm 60^o, \pm 50^o, \pm 43^o, \pm 34^o, \pm 33^o, \pm 26^o, \pm 22^o\}$. A transition region of 5^o is allowed on either side of the mainlobe in order to allow a smooth transition from the mainlobe, to the stop-band.

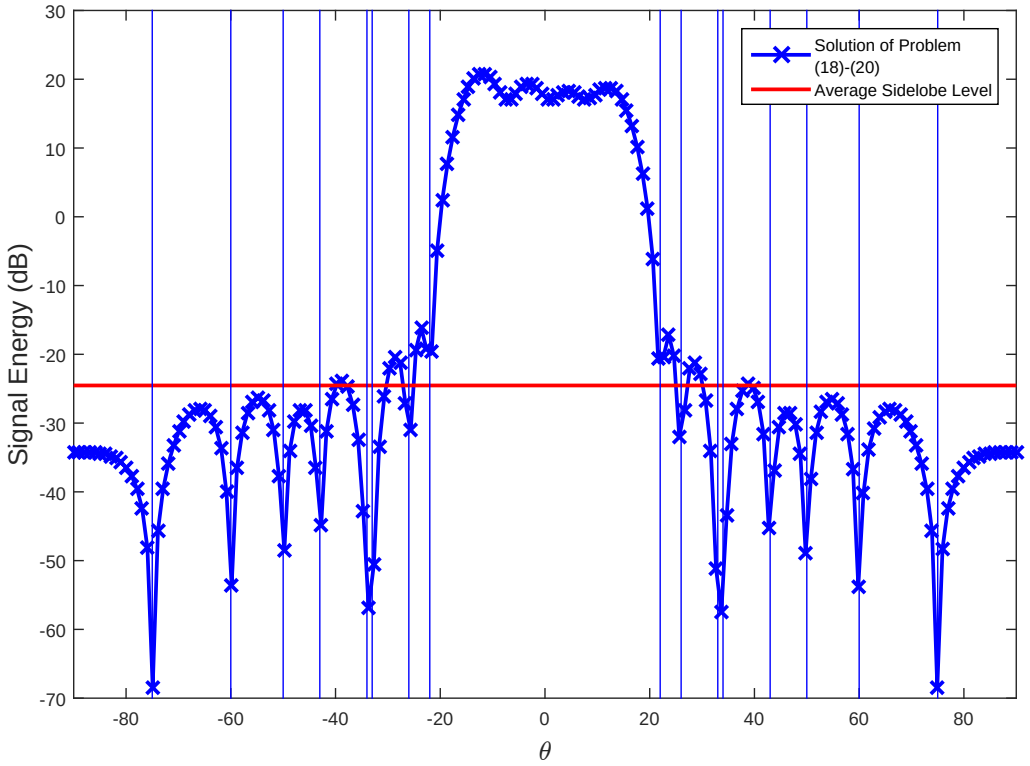

Figure 8: Solution of problem (23)-(25). A clear agreement between algebraic structure and result is observed.

Fig. 8 depicts the result of solving the optimization problem (23)-(25) using a total power constraint, i.e. $\mathbf{B} = \mathbf{I}$. A clear agreement between the vanishing set V and the nulls of the beampattern is shown by the vertical lines. The average sidelobe level (ASL) of the beampattern is -24.53 dB.

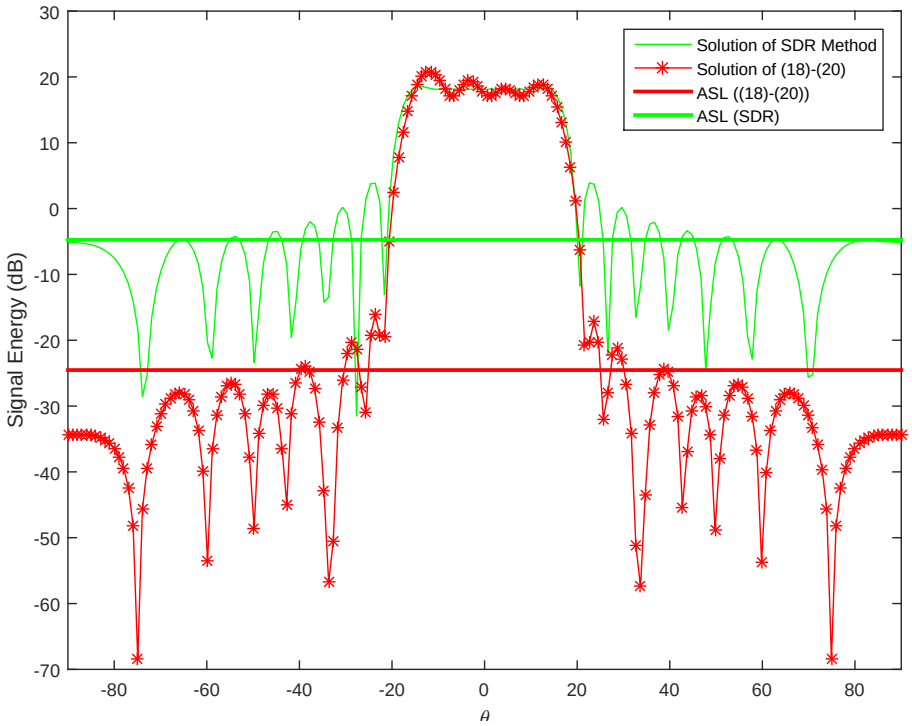

Figure 9: Comparison between proposed and SDR methods. A gap in ASL of 20dB is observed.

A comparison between the proposed method and the solution obtained by the SDR approach to problem (9)–(12) is shown in Fig. 9. The problem relaxed via SDR was designed with the exact same system specification, passband, and transition regions as the proposed method. The ASL corresponding to the solution of the problem relaxed via SDR is -4.73 dB. A gap of 19.8 dB in terms of ASL while a very close agreement between the passbands is observed. Indeed, the peak sidelobe level of the proposed method is 10dB below the ASL of the SDR result, while the mean-squared error (MSE) between the solution of the proposed method and the desired beampattern is 229, compared to 247 with the SDR method.

3.5.2 Example: Single Null

In this example, the goal is to transmit as much energy in all directions as possible, while transmitting no energy in the direction $\theta = -13^o$. To do so, the vanishing set V only takes one value, however, the cardinality of the set can be any integer less than N. We allow a transition region of 5 degrees on either side of the null.

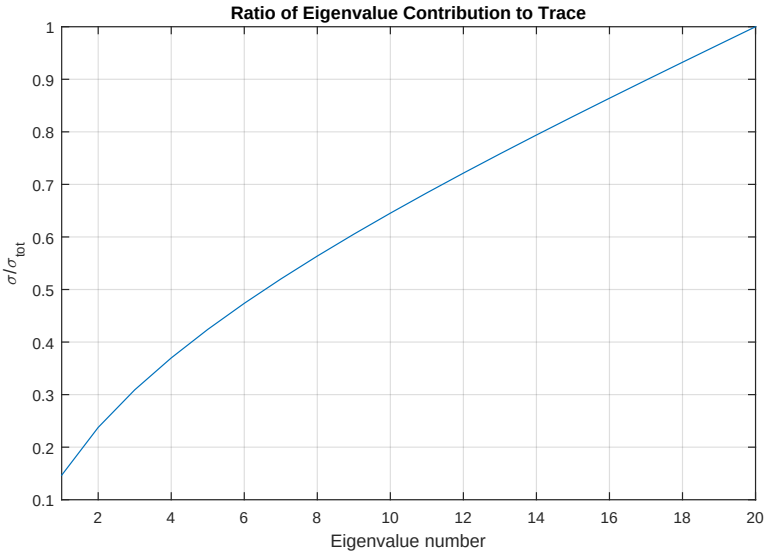

Figure 10: Distribution of energy among eigenvalues of $\mathbf{A}$ matrix for sector width of 170 degrees.

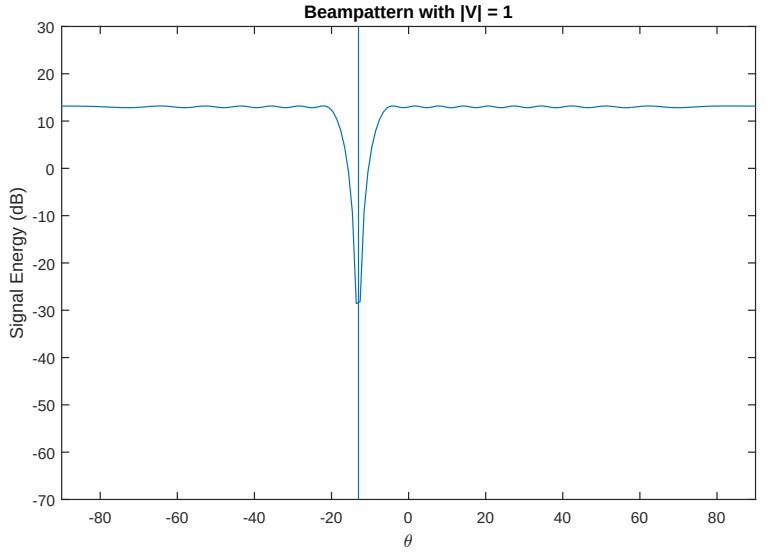

Figure 11: Solution of (23)-(25) with a single root placed at $\theta = -13$.

As we can see from figure 10, the eigenvalues of $\mathbf{A} = \int_{\Theta} \mathbf{a}(\theta)\mathbf{a}^H(\theta)\mathrm{d}\theta$ are all relatively equal. There are no dominant eigenvalues which would guide our choice of K. Each additional null that we set should worsen the fit of the beampattern to G_d.

Fig. 11 demonstrates the performance of the proposed algorithm with a single null placed at $\theta = -13$. Again, we see an exact agreement between the proposed algebraic structure and the solution of problem (23)-(25). To investigate the performance of the algorithm when multiple roots are placed in the same location we set $|V| = 3$ and $V = \{-13 - 13 - 13\}$.

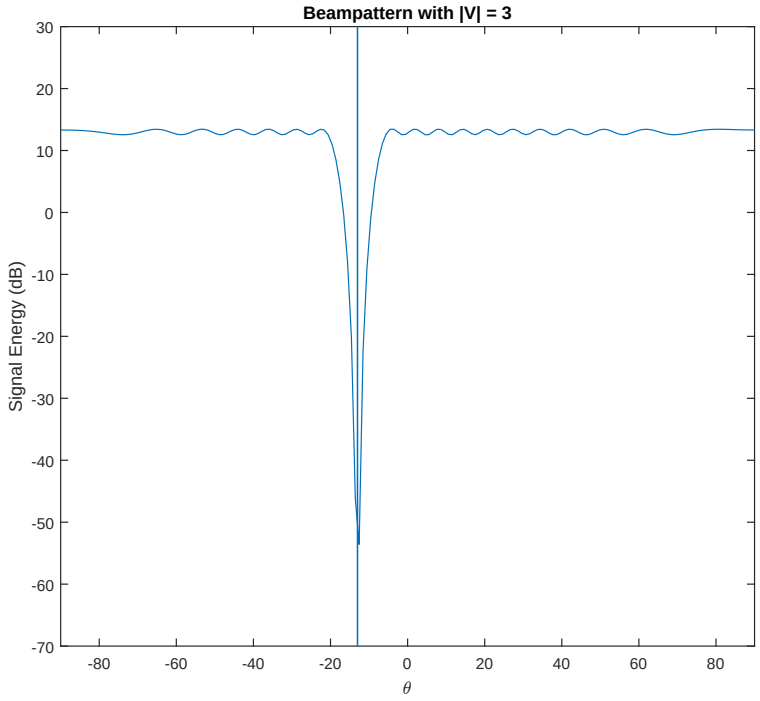

Figure 12: Solution of (23)-(25) with a triple root placed at $\theta = -13$.

Comparing Fig. 12 to 11, we observe that the passband ripple has increased slightly from approximately 0.5dB to approximately $1dB$. However, the null depth at $\theta = -13$ has gone from -28.13dB to -53.6dB: a difference of more than two orders of magnitude. Since there is little point to pushing a null lower than $-53.6dB$ and 1dB passband ripple is tolerable, we settle at placing a triple root at $\theta = -13$ for the purposes of comparison with the SDR method.

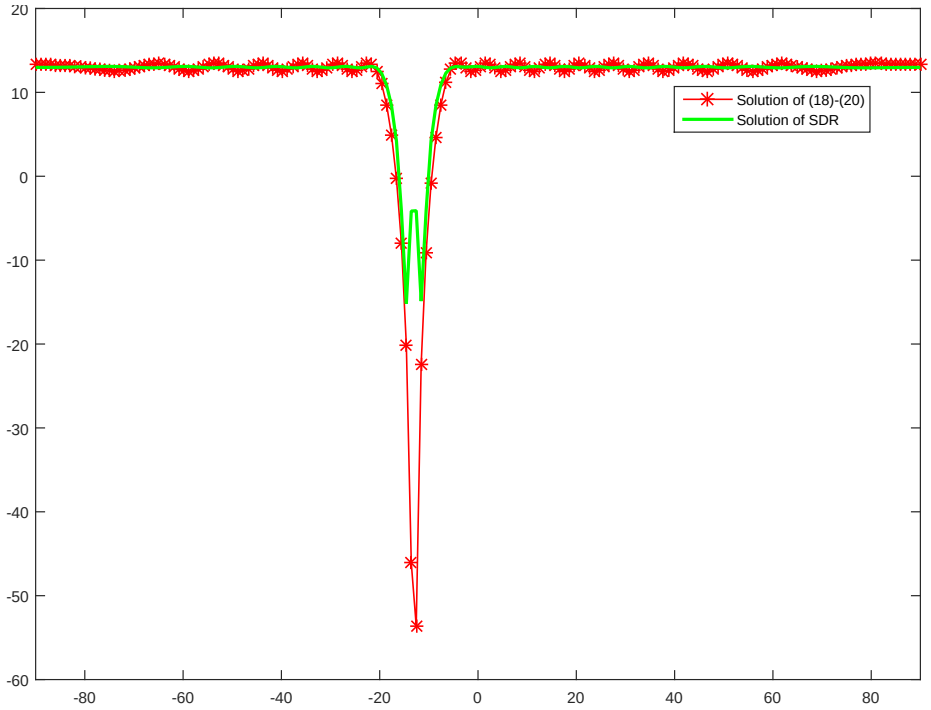

Figure 13: Comparison between proposed and SDR methods.

As we can see, the proposed method greatly outperforms the SDR method in terms of null depth, while closely matching it in terms of pass-band performance.

4 Partially Adaptive 2D Beamforming

4.1 System Model

Consider a monostatic MIMO radar system consisting of a uniform rectangular array (URA) with $M \times N$ antenna elements acting as a transmitter, and a planar receive array with R antenna elements in an arbitrary configuration. Pursuant to the findings in 2.1, we write the $MN \times 1$ transmit array response vector is defined as

$$\mathbf{a}(\theta, \phi) = \text{vec}\left(\mathbf{u}(\theta, \phi)\mathbf{v}^T(\theta, \phi)\right) \tag{31}$$

where

$$[\mathbf{u}(\theta, \phi)]_m = e^{j2\pi m \cdot d_x \sin\theta \cos\phi}, m \in \{0, 1, \cdots, M-1\}$$
$$[\mathbf{v}(\theta, \phi)]_n = e^{j2\pi n \cdot d_y \sin\theta \sin\phi}, n \in \{0, 1, \cdots, N-1\}$$

correspond to the antenna response coefficients for displacements $m \cdot d_x$ and $n \cdot d_y$ from a reference element respectively. A linear combination of K orthogonal baseband waveforms $\boldsymbol{\psi}(t) = [\psi_1(t), \ldots, \psi_K(t)]^T$is transmitted, where $K \ll MN$, which permits energy focussing over a desired sector $\Theta = [\theta_1 \ \theta_2]$ in the elevation domain and $\Phi = [\phi_1 \ \phi_2]$ in the azimuthal domain. This corresponds to a prior distribution of $\mathcal{U}_r(\Theta, \Phi)$ for the target locations.

The signal at the transmitter within a single slow-time pulse, for a given angular direction (θ, ϕ), $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ and $\phi \in [0, 2\pi]$, at time t is expressed as

$$\mathbf{s}(t) = \mathbf{a}(\theta, \phi)^H \mathbf{W} \boldsymbol{\psi}(t) \tag{32}$$

where $\mathbf{W}$ is a $MN \times K$ complex valued beamforming weight matrix with $w_{p,k}$ corresponding to the weighting coefficient of waveform k at antenna element $p = mn$(that is the element in the m-th row and n-th column).

The magnitude of the beampattern in the direction (θ, ϕ) is given by

$$\begin{aligned} G(\theta, \phi) &= \int_T \mathbf{s}(t)\mathbf{s}^H(t)dt \\ &= \mathbf{a}^H(\theta, \phi)\mathbf{W}\left(\int_T \boldsymbol{\psi}(t)\boldsymbol{\psi}^*(t)dt\right)\mathbf{W}^H\mathbf{a}(\theta, \phi) \\ &= \mathbf{a}^H(\theta, \phi)\mathbf{W}\mathbf{W}^H\mathbf{a}(\theta, \phi) = \|\mathbf{W}^H\mathbf{a}(\theta, \phi)\|^2 \end{aligned} \tag{33}$$

where $\int_T \boldsymbol{\psi}(t)\boldsymbol{\psi}^*(t)dt = \mathbf{I}_k$, as $\boldsymbol{\psi}(t)$ are chosen to be orthogonal, T_s is the period of a

slow-time pulse, $\|\cdot\|$ is the Euclidean norm, and $(\cdot)^*$ is the complex conjugate.

The presence of L targets in a Doppler-range bin following a Swerling II model results in a noisy $R \times 1$ receive array observation vector at a time t and pulse τ which can be expressed as

$$\mathbf{x}(t,\tau) = \mathbf{B}\boldsymbol{\Sigma}(\tau)\mathbf{A}^H\mathbf{W}\boldsymbol{\psi}(t) + \mathbf{z}(t,\tau) \tag{34}$$

where $\mathbf{B} \triangleq [\mathbf{b}(\theta_1,\phi_1),\ldots,\mathbf{b}(\theta_L,\phi_L)]$, $\mathbf{A} \triangleq [\mathbf{a}(\theta_1,\phi_1),\cdots,\mathbf{a}(\theta_L,\phi_L)]$, $\boldsymbol{\Sigma}(\tau) \triangleq \text{diag}([\beta_1(\tau),\ldots,\beta_L(\tau)])$, $\mathbf{z}(t,\tau)$ is an $R \times 1$ zero mean Gaussian random vector with covariance $\mathbf{Q} = \sigma^2\mathbf{I}_R$, and $\beta_l(\tau)$ is the complex radar reflection coefficient corresponding to the l-th target. The operator diag$(\cdot)$ creates a diagonal matrix with entries equal to the elements of a vector. The receive array response vectors are $\mathbf{b}(\theta_l,\phi_l) \triangleq [e^{j2\pi\boldsymbol{\xi}[\gamma_l,\zeta_l]^T}]$ where $\boldsymbol{\xi}$ is an $R \times 2$ matrix containing the x and y coordinates of a receiver element r, $\gamma_l \triangleq \sin\theta_l\cos\phi_l$, and $\zeta_l \triangleq \sin\theta_l\sin\phi_l$. The receive antenna element coordinates in $\boldsymbol{\xi}$ are defined relative to a reference element in terms of the carrier wavelength λ_c, but are otherwise arbitrary. Therefore, the first row of $\boldsymbol{\xi}$ is the zero row vector. The columns of the matrices $\mathbf{A}$ and $\mathbf{B}$ are the array response vectors for a target located at direction (θ_l,ϕ_l) for the transmitter and receiver respectively.

Defining the $K \times N$ matrix $\boldsymbol{\Psi} \triangleq [\boldsymbol{\psi}(1),\ldots,\boldsymbol{\psi}(T_f)]$, where T_f is the number of fast-time samples of the K orthogonal waveforms, the result of the matched filter operation at the receiver over a slow-time pulse τ is expressed as

$$\begin{aligned}\mathbf{y}(\tau) &= \text{vec}\left(\frac{1}{N}\mathbf{B}\boldsymbol{\Sigma}(\tau)\mathbf{A}^H\mathbf{W}\boldsymbol{\Psi}\boldsymbol{\Psi}^H + \frac{1}{T_f}\mathbf{z}(t,\tau)\boldsymbol{\Psi}^H\right)\\ &= \text{vec}\Big(\mathbf{B}\boldsymbol{\Sigma}(\tau)\mathbf{A}^H\mathbf{W} + \mathbf{V}(\tau)\Big) \qquad (35)\\ &= \Big((\mathbf{W}^H\mathbf{A}) \odot \mathbf{B}\Big)\mathbf{c}(\boldsymbol{\Sigma}(\tau)) + \mathbf{v}(\tau) \qquad (36)\end{aligned}$$

where vec$(\cdot)$ is the vectorization operator (which stacks the columns of a matrix on top of one another into a vector), $\odot$ is the column-wise Khatri-Rao product, $\mathbf{c}(\boldsymbol{\Sigma}(\tau))$ is a column vector consisting of the diagonal entries of $\boldsymbol{\Sigma}(\tau)$, and $\mathbf{v}(\tau) \triangleq T_f^{-1}\text{vec}(\mathbf{z}(t,\tau)\boldsymbol{\Psi}^H)$.

The noisy virtual data vectors (35) have dimension of $KR \times 1$ and form a $KR \times I$ matrix

$$\mathbf{Y} \triangleq [\mathbf{y}(1),\ldots,\mathbf{y}(I)] = \Big((\mathbf{W}^H\mathbf{A}) \odot \mathbf{B}\Big)\mathbf{C}(\boldsymbol{\Sigma}) + \mathbf{Z} \tag{37}$$

where I is the number of slow-time pulses in a scan, $\mathbf{C}(\mathbf{\Sigma})$ has columns $\mathbf{c}(\mathbf{\Sigma}(\tau))$, and $\mathbf{V}$ has columns $\mathbf{v}(\tau)$.

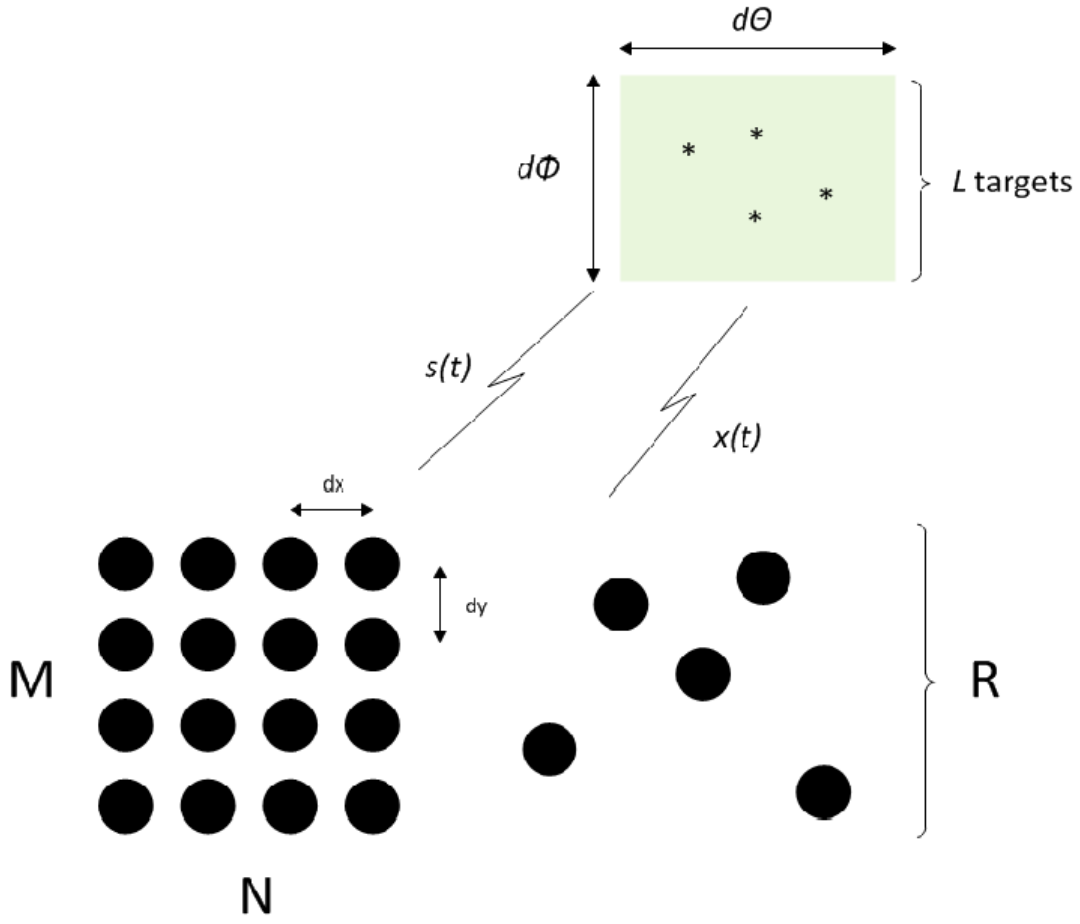

Figure 14: Visual representation of the 2D TB-MIMO radar system considered in this section. The receiver is in an arbitrary configuration. The green patch represents the sector of interest.

In order to enable search-free direction finding methods, it is necessary to enforce a structure on (35). This is normally done by controlling the geometry of the receive array. By requiring the receiver to be translationally symmetric, datasets which are identical in all but a phase rotation. This phase rotation will depend on the location of the target, and hence, can be used for direction finding. However, the structure of (35) can be controlled from the transmitter as well as was shown in [19], [20], [21]. In order to unamibuously determine the direction of arrival (DOA) of a target in two dimensions we need to generate a data set which has multiple invariances (at least two, one to resolve each angle). We enforce these by requiring

that the waveforms $\boldsymbol{\psi}(t)$ are transmitted in groups, each radiated in an identical beampattern. A sufficient condition for two groups of waveforms to be radiated with identical beampatterns is that each waveform in one group be radiated with the exact same beampattern as a waveform in the other group, which we will now prove.

Consider two beamforming matrices $\mathbf{W} \in \mathbb{C}^{N\times K}$ and $\mathbf{V} \in \mathbb{C}^{N\times K}$. The beampatterns associated to these matrices can be written as $\mathbf{a}^H(\theta,\phi)\mathbf{W}\mathbf{W}^H\mathbf{a}(\theta,\phi)$, and $\mathbf{a}^H(\theta,\phi)\mathbf{V}\mathbf{V}^H\mathbf{a}(\theta,\phi)$. We are interested in the conditions for when

$$\mathbf{a}^H(\theta,\phi)\mathbf{W}\mathbf{W}^H\mathbf{a}(\theta,\phi) = \mathbf{a}^H(\theta,\phi)\mathbf{V}\mathbf{V}^H\mathbf{a}(\theta,\phi),\ \forall\ \theta,\phi \tag{38}$$

which can be rewritten as

$$\text{trace}\Big(\mathbf{a}^H(\theta,\phi)\mathbf{W}\mathbf{W}^H\mathbf{a}(\theta,\phi)\Big) = \text{trace}\Big(\mathbf{a}^H(\theta,\phi)\mathbf{V}\mathbf{V}^H\mathbf{a}(\theta,\phi)\Big). \tag{39}$$

Then, using the cyclical property of the matrix trace operator, (39) is easily rewritten as

$$\text{trace}\Big(\mathbf{W}^H\mathbf{D}(\theta,\phi)\mathbf{W}\Big) = \text{trace}\Big(\mathbf{V}^H\mathbf{D}(\theta,\phi)\mathbf{V}\Big) \tag{40}$$

where $\mathbf{D}(\theta,\phi) \triangleq \mathbf{a}(\theta,\phi)a^H(\theta,\phi)$. By expanding (40) it is easily shown that the diagonal elements of $\mathbf{W}^H\mathbf{D}(\theta,\phi)\mathbf{W}$ are exactly the beampatterns corresponding to the individual vectors of $\mathbf{W}$.

$$\begin{aligned}
\text{trace}\left(\begin{bmatrix}\mathbf{w}_1^H\mathbf{D}(\theta,\phi)\\ \vdots \\ \mathbf{w}_k^H\mathbf{D}(\theta,\phi)\end{bmatrix}\mathbf{W}\right) &= \text{trace}\left(\begin{bmatrix}\mathbf{v}_1^H\mathbf{D}(\theta,\phi)\\ \vdots \\ \mathbf{v}_k^H\mathbf{D}(\theta,\phi)\end{bmatrix}\mathbf{V}\right) \\
&= \text{trace}\left(\begin{bmatrix}\mathbf{v}_1^H\mathbf{D}(\theta,\phi)\mathbf{v}_1 & \cdots & \mathbf{v}_1^H\mathbf{D}(\theta,\phi)\mathbf{v}_k \\ \vdots & \ddots & \vdots \\ \mathbf{v}_k^H\mathbf{D}(\theta,\phi)\mathbf{v}_1 & \cdots & \mathbf{v}_k^H\mathbf{D}(\theta,\phi)\mathbf{v}_k\end{bmatrix}\right) \\
\sum_{k=1}^{K}\mathbf{w}_k^H\mathbf{D}(\theta,\phi)\mathbf{w}_k &= \sum_{k=1}^{K}\mathbf{v}_k^H\mathbf{D}(\theta,\phi)\mathbf{v}_k.
\end{aligned} \tag{41}$$

Equation (41) completes the proof. It is perhaps possible that this condition is not necessary, given that there are an infinite collection of K non-negative real numbers which sum up to any given real number, however, we do not investigate this further

here.

Now, consider a beamforming matrix $\mathbf{W} \triangleq [\mathbf{U}_0 \cdots \mathbf{U}_{Q-1}]$ where Q is the number of invariant waveform groups. Then equation (35) can be rewritten as

$$\left(\Pi \odot (\mathbf{U}_0^H \mathbf{A}) \odot \mathbf{B}\right)\mathbf{c}(\mathbf{\Sigma}(\tau)) + \mathbf{v}(\tau) \tag{42}$$

where Π is a $Q \times K$ matrix of unity modulo phase arguments which relate the waveform groups to one another. So, Π_q, k is the complex exponential relating the waveform group p-th waveform group to waveform group 1 for the k-th target. The first row is then, obviously, all ones. Equation (42) clearly shows the enforced multiple invariances as a result of proper beamspace matrix design which is also completely independent of the receive array geometry. Indeed, if there is only one antenna element in the receive array, we will still have the Q invariant data sets required to unambiguously locate targets in 2 dimensions. What changes, then, is the number of detectable targets, and the variance of the estimator. The next section explains exactly how the multiple data invariances are enforced at the transmitter such that the phase information is useable for DOA estimation.

4.2 Beamspace Matrix Design

The design of $\mathbf{W}$ is performed in two stages. First, a beamspace matrix $\mathbf{U}_0 = [\mathbf{u}_1, \ldots, \mathbf{u}_k]$, with full column rank K, is designed over a spatial sector $\Theta = [\theta_1\ \theta_2]$ and $\Phi = [\phi_1\ \phi_2]$ using only the first $(M-1)$ rows and $(N-1)$ columns of the transmit array. Then, a simple transformation is performed on $\mathbf{U}_0$ to produce beamforming matrices with identical beampatterns, but which correspond to different subarrays. The number of orthogonal waveforms K is selected in line with the findings of 3.3. In 2D, however, we have a two dimensional sector of interest, and thus the number of waveforms is selected to be the number of dominant eigenvalues of the matrix

$$\mathbf{D}(\theta, \phi) \triangleq \int_{\Theta} \int_{\Phi} \mathbf{a}(\theta, \phi)\mathbf{a}^H(\theta, \phi) d\theta d\phi. \tag{43}$$

As we saw in 3.3, K depends heavily on the area of the given sector $\{\Theta, \Phi\}$. In the case where we have no prior information about the location of the targets, the prior distribution is $\mathcal{U}_r([-\pi/2, \pi/2], [0, \pi])$. The eigenvalues corresponding to matrix $\mathbf{D}(\theta, \phi)$ will be equal. As the sector shrinks, however, fewer and fewer waveforms are necessary to concentrate energy within the desired sector.

The initial design of $\mathbf{U}_0$ can be stated as the following optimization problem

$$\min_{\mathbf{u}_1,\ldots,\mathbf{u}_K} \max_{\theta,\phi} \left| G_d(\theta,\phi) - \sum_{k=1}^{K} \mathbf{u}_k^H \mathbf{a}(\theta,\phi)\mathbf{a}^H(\theta,\phi)\mathbf{u}_k \right| \tag{44}$$

$$\text{s.t.} \sum_{k=1}^{K} \left| \mathbf{U}_{[jk]} \right|^2 = \frac{E}{4KMN}, j \in \{1, \cdots, (P-1)(Q-1)\} \tag{45}$$

where $G_d(\theta,\phi)$ is an ideal beampattern over the desired sector $\{\Theta,\Phi\}$. (44) is a non-convex quadratically-constrained quadratic optimization problem. It can be solved via a semi-definite programming relaxation approach, [21]–[20] by first introducing the new variables $\mathbf{X}_k \triangleq \mathbf{u}_k\mathbf{u}_k^H$, $k = 1,\ldots,K$, and relaxing the rank constraint by requiring only that $\mathbf{X}_k$ be positive semidefinite. Then, using the technique of randomization a rank 1 solution can be extracted from the signal cross-correlation matrix $\mathbf{X}_k$. This will yield a beamforming matrix of dimension $(M-1)(N-1) \times K$.

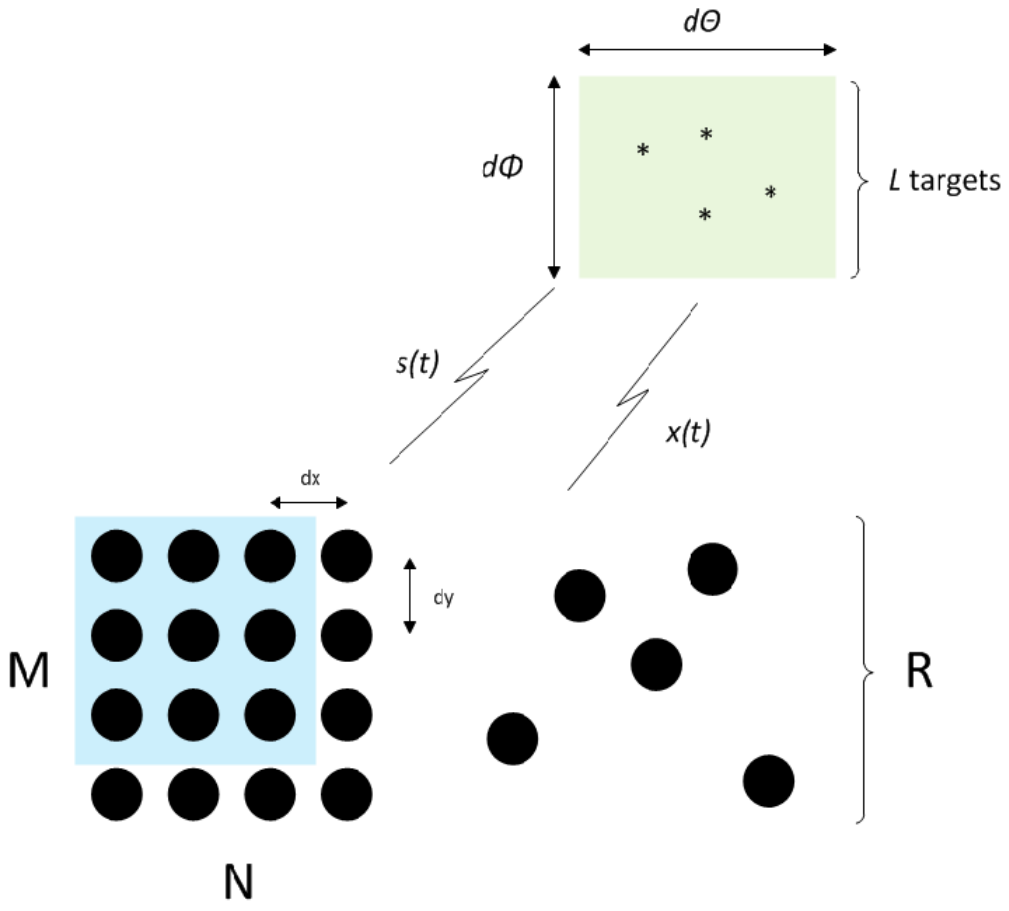

Figure 15: Visual representation of the initial beamforming process in section 4.2. The transmitter elements highlighted in blue show the elements used in the beamforming problem (44)-(45).

It is trivial to show that a matrix $\mathbf{U}_0'$ of dimension $MN \times K$, with an identical beampattern to that of $\mathbf{U}_0$, can be constructed by placing zeros in the spots corresponding to the antenna elements which were omitted from the original design of $\mathbf{U}_0$. The matrix $\mathbf{U}_0'$ then denotes a beamforming matrix where K beams are transmitted from the first $M-1$ rows and $N-1$ columns of a transmit array of dimension $M \times N$. Given the shape of the transmit array, it is simple to show that by shifting the positions of the zeros in $\mathbf{U}_0'$, the exact same beampattern can be achieved by subarrays containing the first $M-1$ rows and last $N-1$ columns, the last $M-1$ rows and the first $N-1$ columns, and finally the last $M-1$ and last $N-1$ columns of the transmit array. These three matrices are denoted as $\mathbf{U}_1'$, $\mathbf{U}_2'$, and $\mathbf{U}_3'$, respectively. With these matrices defined, it is easy to show that the following is true

$$\begin{aligned} \mathbf{a}^H(\theta,\phi)\mathbf{U}_0' &= e^{j2\pi d_x \sin\theta\cos\phi}\left(\mathbf{a}^H(\theta,\phi)\mathbf{U}_1'\right) \\ &= e^{j2\pi d_y \sin\theta\sin\phi}\left(\mathbf{a}^H(\theta,\phi)\mathbf{U}_2'\right) \\ &= e^{j2\pi(d_x \sin\theta\cos\phi + d_y \sin\theta\sin\phi)}\left(\mathbf{a}^H(\theta,\phi)\mathbf{U}_3'\right). \end{aligned} \tag{46}$$

The beamforming matrix $\mathbf{W}$ is then defined as $\mathbf{W} \triangleq [\mathbf{U}_0', \mathbf{U}_1', \mathbf{U}_2', \mathbf{U}_3']$ with an overall dimension of $MN \times 4K$. Clearly, in the original design problem, K must be no larger than $MN/4$. It should be stated that the partially adaptive approach described in this section implies that only $(M-1)(N-1)$ elements operate at full power. This power loss, however, is a strictly decreasing function of transmit array size. With the assumption of constant modulo 1 signals, the total radiated energy is $\mathrm{tr}(\mathbf{W}\mathbf{W}^H)$. Now we assume, without loss of generality that the amount of power available to a transmitter is 1 unit of power. The total amount of power available to a MIMO system will be the number of antennas MN. In the case of the proposed algorithm, we will have $((M-1)(N-1)/4)\cdot 4 = (M-1)(N-1)$. The fraction of power used by an array of dimension $P \times Q$ by the proposed method is then $F = MN/((M-1)(N-1))$. Let $E = MN$, then so long as $EF \cdot 1/4K \geq 1$, this method will yield superior resolution and DOA estimation performance relative to FWD-MIMO provided the targets exist within our sector of interest. This is because we will transmit more power per unit waveform, thereby increasing the SNR of the received signal.

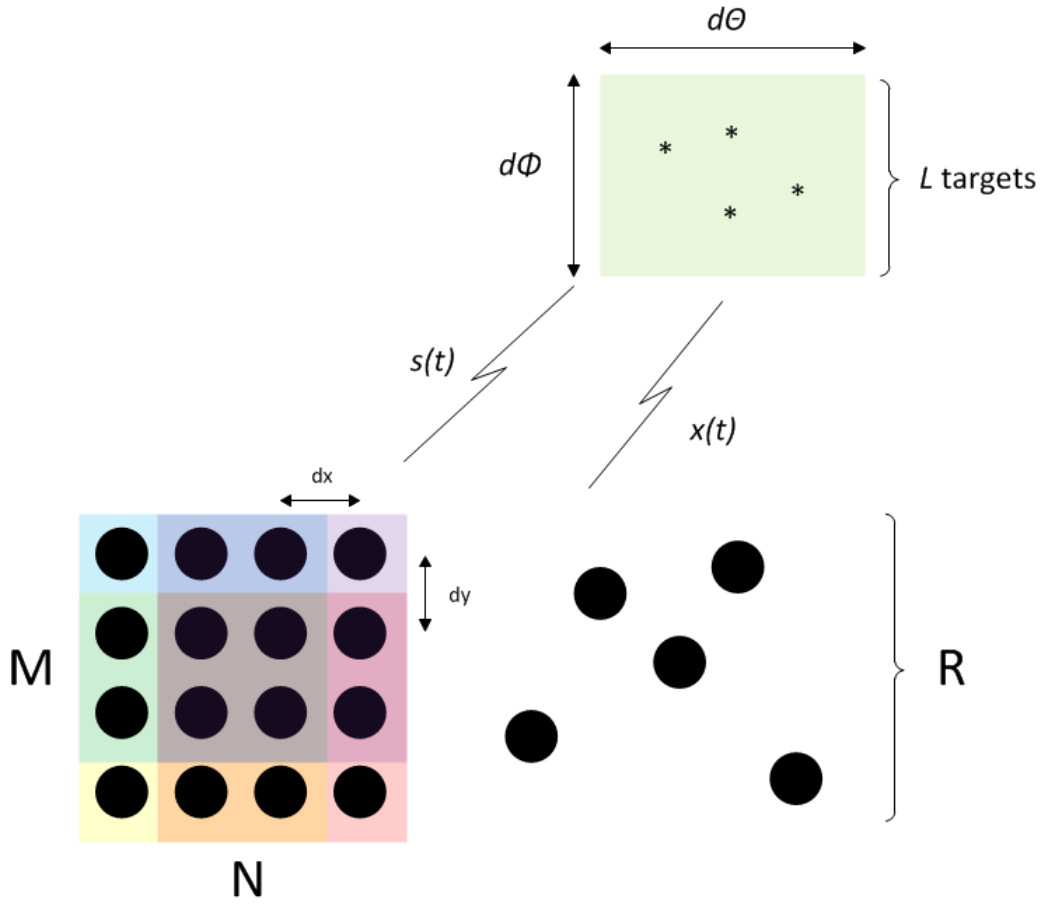

Figure 16: Visual representation of the 2D TB-MIMO radar system after the beamspace transformation. Only the center 4 elements are operating at full power while the corner elements are operating at 1/4 power. The elements on the sides are operating at half power.

4.3 DOA Estimation

Given the structure (46) imposed on the beamspace matrix $\mathbf{W}$, let us turn our attention to (35). Rewriting the noiseless matrix before vectorization allows a clear illustration of the effect of the proposed structure of $\mathbf{W}$ on DOA estimation. Specifically, we can write that

$$\mathbf{B}\boldsymbol{\Sigma}(\tau)\mathbf{A}^H\mathbf{W} = \mathbf{B}\boldsymbol{\Gamma} \tag{47}$$

where $\boldsymbol{\Gamma} \triangleq \boldsymbol{\Sigma}(\tau)\mathbf{A}^H\mathbf{W}$. The matrix $\boldsymbol{\Gamma}$ is the source signal matrix, and has dimension $L \times 4K$. In the following, $\boldsymbol{\Gamma}_0 = \boldsymbol{\Sigma}(\tau)\mathbf{A}^H\mathbf{U}_0'$ is the source signal matrix corresponding to K beams emanated from the first $(M-1)$ rows and $(N-1)$ columns of the transmit array. Using the relations (46) we define matrices $\boldsymbol{\Omega}_i, i \in \{0,1,2,3\}$ as the $L \times L$

diagonal matrices with the l-th diagonal entry of $\mathbf{\Omega}_i$ being the complex exponential in (46) which relates $\mathbf{a}^H(\theta,\phi)\mathbf{U}_0'$ to $\mathbf{a}^H(\theta,\phi)\mathbf{U}_i'$. The matrix $\mathbf{\Omega}_0$ is obviously the identity matrix. Then (47) can be expressed as the following block partitioned matrix

$$\mathbf{B\Gamma} = \mathbf{B}\left[\begin{array}{c|c|c|c} \mathbf{\Omega}_0\mathbf{\Gamma}_0 & \mathbf{\Omega}_1\mathbf{\Gamma}_0 & \mathbf{\Omega}_2\mathbf{\Gamma}_0 & \mathbf{\Omega}_3\mathbf{\Gamma}_0 \end{array}\right] \tag{48}$$

$$= \left[\begin{array}{c|c|c|c} \mathbf{B\Omega}_0 & \mathbf{B\Omega}_1 & \cdots & \mathbf{B\Omega}_3 \end{array}\right] \mathrm{bdiag}_4(\mathbf{\Gamma}_0) \tag{49}$$

where $\mathrm{bdiag}_m(\cdot)$ takes a single matrix as an argument, and creates a block diagonal matrix whose m blocks are equal to its argument. The matrix $\mathbf{B\Omega}_0$ is simply the receiver response matrix to L targets. The virtual receiver response matrices $\mathbf{B\Omega}_1$, $\mathbf{B\Omega}_2$, and $\mathbf{B\Omega}_3$ are exactly the receiver response matrices to L targets, for identical receive arrays that are linearly displaced from our actual receiver by $[d_x, 0]$, $[0, d_y]$, and $[d_x, d_y]$, respectively. The source signal matrix $\mathbf{\Gamma}_0$ is a common factor for each. From (48) it is visible that the proposed structure for $\mathbf{W}$ enforces an algebraic structure on $\mathbf{Y}$ which can be exploited by search-free algorithms for DOA estimation, including, but not limited to, ESPRIT (Estimation of Signal Parameters through Rotational Invariance Techniques) [29], [30]. Further, it should be noted from (37) that the matrix $\mathbf{Y}$ has rank L, with probability 1, if the targets are incoherent. The proof of this statement can be found in 4.4.

Matrix $\mathbf{Y}$ has dimension $4RK \times I$. After defining the matrix selection operator $\mathbf{F}_j(\cdot)$ which selects the $(jM/4+1)$–$M/4(j+1)$ rows from an arbitrary matrix with M rows, where $j \in \{0,1,2,3\}$, $\mathbf{Y}$ and a new matrix $\mathbf{Y}'$ can be expressed as

$$\mathbf{Y} = \begin{bmatrix} \mathbf{F}_0(\mathbf{Y}) \\ \mathbf{F}_1(\mathbf{Y}) \\ \mathbf{F}_2(\mathbf{Y}) \\ \mathbf{F}_3(\mathbf{Y}) \end{bmatrix}, \mathbf{Y}' = \begin{bmatrix} \mathbf{F}_0(\mathbf{Y}) \\ \mathbf{F}_2(\mathbf{Y}) \\ \mathbf{F}_1(\mathbf{Y}) \\ \mathbf{F}_3(\mathbf{Y}) \end{bmatrix}. \tag{50}$$

Forming the cross correlation matrices $\mathbf{R}_{\mathbf{Y}} = I^{-1}\mathbf{Y}\mathbf{Y}^H$ and $\mathbf{R}_{\mathbf{Y}'} = I^{-1}\mathbf{Y}'\mathbf{Y}'^H$, and performing ESPRIT on both will yield a vector of L phase arguments which are directly proportional to ζ_l and γ_l. Defining a complex number $z_l = \gamma_l + j\zeta_l$ the angle estimates are given by $\phi_l = \arctan(\zeta_l/\gamma_l)$, and $\theta_l = |z_l|$.

4.4 Identifiability Results

In this section, we prove a few key results relating to the identifiability of targets using this method, as well as others. The DOA estimation methods described in Section 4.3 rely on the use of an eigendecomposition on the cross-correlation matrix

$$\mathbf{X} = \mathbf{V}\mathbf{P}\mathbf{V}^H + \mathbf{U} \tag{51}$$

where $\mathbf{V} = \left(\Pi \odot (\mathbf{U}_1^H \mathbf{A}) \odot \mathbf{B}\right)$, $\mathbf{P} = (1/T_s)\mathbf{C}(\mathbf{\Sigma})\mathbf{C}(\mathbf{\Sigma})^H$, and $\mathbf{U}$ contains all noise and cross-terms from the product $\mathbf{Y}\mathbf{Y}^H$. With sufficent SNR, up to L targets can be identified with these methods if the matrix V has rank L. Because of the matrix product $\mathbf{W}^H\mathbf{A}$, the matrix V is not obviously tall and full rank. Identifiability with these methods therefore requires an indepth analysis, which follows. First, it will be proven that the Khatri-Rao product of an injective matrix with a conformable non-zero matrix is injective. Second, it will be proven that, with probability 1, the beamforming matrix will not cause the rank of the signal-only matrix to drop. And finally, these two results put together will show that, under moderate conditions, $4KR$ targets will be identifiable with probability 1.

Lemma 2 *The Khatri-Rao product of an injective matrix with any other conformable matrix which has at least one non-zero entry in every column is injective.*

The Khatri-Rao product $\mathbf{A} \odot \mathbf{B}$ with $\mathbf{A} \in \mathbb{C}^{N\times L}$, $\mathbf{B} \in \mathbb{C}^{M\times L}$ is injective if one of $\mathbf{A}$ or $\mathbf{B}$ is injective. A matrix of dimension $M \times N$ is injective if and only if $M \geq N$ and the matrix is full rank. The fact that this is true if both matrices are injective is yielded directly by the definition of the tensor product on vector spaces. Let matrix $\mathbf{A}$ define a linear transformation $S : \mathbb{C}^L \to \mathcal{R}(\mathbf{A})$, while the matrix $\mathbf{B}$ defines a linear transformation $T : \mathbb{C}^L \to \mathcal{R}(\mathbf{B})$. Then, the Kronecker product $\mathbf{A} \otimes \mathbf{B}$ is the tensor product of the linear maps S and T defined as $S \otimes T : V \otimes W \to \ \mathcal{R}(\mathbf{A}) \otimes \mathcal{R}(\mathbf{B})$. The tensor product of vector spaces $\mathcal{R}(\mathbf{A}) \otimes \mathcal{R}(\mathbf{B})$ has a few interesting properties. One is that the pairwise kronecker products of the bases of $\mathcal{R}(\mathbf{A})$ and $\mathcal{R}(\mathbf{B})$ form a basis represented by $\mathcal{R}(\mathbf{A}) \otimes \mathcal{R}(\mathbf{B})$. Therefore, $\mathrm{rk}(\mathbf{A} \otimes \mathbf{B}) = \mathrm{rk}(\mathbf{A}) \cdot \mathrm{rk}(\mathbf{B})$. Clearly $\mathcal{C}(\mathbf{A} \odot \mathbf{B}) \subset \mathcal{C}(\mathbf{A} \otimes \mathbf{B})$. Let both $\mathbf{A}$ and $\mathbf{B}$ be injective. Because $\mathcal{R}(\mathbf{A} \odot \mathbf{B}) \subset \mathcal{R}(\mathbf{A} \otimes \mathbf{B})$, then the basic properties of the tensor product of vector spaces tells us that $\dim(\mathcal{R}(\mathbf{A}) \otimes \mathcal{R}(\mathbf{B})) = \mathrm{L}^2$, and clearly, the Khatri-Rao product must have full rank.

Now, allow one of the matrices $\mathbf{A}$ or $\mathbf{B}$ to no longer be injective. Assume the case where $\sum_{l=1}^{L} \alpha_l \mathbf{a}_l = 0 \Rightarrow \alpha_l = 0,\ l = 1, \cdots, L$ and $\sum_{l=1}^{L} \beta_l \mathbf{b}_l = 0$ for some $\beta_l \neq 0$, where $\mathbf{a}_l$, and $\mathbf{b}_l$ are the columns of $\mathbf{A}$ and $\mathbf{B}$, and $\alpha_l, \beta_l \in \mathbb{C}$. $M \geq L$, and $N \geq 1$. Expanding the definition of the columnwise Khatri-Rao product $\mathbf{B} \odot \mathbf{A}$ yields

$$\begin{aligned} \mathbf{B} \odot \mathbf{A} &= [\mathbf{b}_1 \otimes \mathbf{a}_1, \mathbf{b}_2 \otimes \mathbf{a}_2, \cdots, \mathbf{b}_l \otimes \mathbf{a}_l] \\ &= \begin{bmatrix} [\mathbf{b}_1]_1 \mathbf{a}_1 & [\mathbf{b}_2]_1 \mathbf{a}_2 & \cdots & [\mathbf{b}_l]_1 \mathbf{a}_l \\ [\mathbf{b}_1]_2 \mathbf{a}_1 & [\mathbf{b}_2]_2 \mathbf{a}_2 & \cdots & [\mathbf{b}_l]_2 \mathbf{a}_l \\ \vdots & \vdots & \ddots & \vdots \\ [\mathbf{b}_1]_M \mathbf{a}_1 & [\mathbf{b}_2]_M \mathbf{a}_2 & \cdots & [\mathbf{b}_l]_M \mathbf{a}_l \end{bmatrix}. \end{aligned} \tag{52}$$

Clearly $MN \geq L$ in this case, so, if the matrix $\mathbf{B} \odot \mathbf{A}$ is full rank, it will also be injective. Let the columns of $\mathbf{B} \odot \mathbf{A}$ be denoted as $\mathbf{c}_l, l = 1, \cdots, L$. Then, $\mathbf{B} \odot \mathbf{A}$ will drop rank if and only if $\sum_{l=1}^{L} \gamma_l \mathbf{c}_l = 0$ with at least one non-zero $\gamma_l \in \mathbb{C}$. Using (52), and defining $\zeta_{m,l} = \gamma_l [\mathbf{b}_l]_m$, an equivalent condition is $\sum_{l=1}^{L} \zeta_{m,l} \mathbf{a}_l = 0, \forall m = 1, \cdots, M$. However, by assumption, these equations being satisfied implies that $\zeta_{m,l} = 0, \forall m, l$, while we assumed that at least one $\gamma_l \neq 0$, and for each $l = 1, \cdots, L$ there is some $[\mathbf{b}_l]_m \neq 0$. Therefore, at least one ζ_l must be non-zero. This is a contradiction, and thus proves that $\mathbf{B} \odot \mathbf{A}$ is injective. To show that $\mathbf{A} \odot \mathbf{B}$ is injective, we observe that $\mathbf{A} \odot \mathbf{B} = \mathbf{P}(\mathbf{B} \odot \mathbf{A})$ where $\mathbf{P}$ is a permutation matrix. As permutation matrices are bijective, multiplication by a permutation matrix preserves rank. ■

Lemma 3 *Let* $\mathbf{A} \triangleq [\mathbf{a}(\theta_1, \phi_1), \cdots, \mathbf{a}(\theta_L, \phi_L)]$ *for some fixed set of distinct target locations, where* $\mathbf{a}(\theta_L, \phi_L)$ *is defined in* (31)*. Then, if* $\mathbf{W} \in \mathbb{C}^{MN \times K}$ *is selected randomly from any continuous distribution in that space,* $(\mathbf{A}^H \mathbf{W})^T$ *will be injective with probability 1.*

It is necessary to introduce the concept of an algebraic variety in order to prove this Lemma. An algebraic variety is here defined as the set of solutions to a system of polynomial equations. The polynomial equations in this case are the L minors of the matrix $\mathbf{A}^H \mathbf{W}$. The matrix $\mathbf{A}^H \mathbf{W}$ will drop rank if and only if each L minor is 0. The set of solutions to these equations is called a determinantal variety. For a fixed $\mathbf{A}$ as defined above, each equation will be a multivariate polynomial of degree L with the individual entries of $\mathbf{W}$ as variables. Polynomials are analytic functions, and as such their roots occupy a set of measure zero within the MNK dimensional ambient affine space of complex matrices. Therefore, the probability of drawing a

matrix $\mathbf{W}$ from any continuous distribution on the ambient space which would result in a full-rank matrix product $\mathbf{A}^H\mathbf{W}$ is 1. ■

Combining these two statements results in the following strong result for identifiability. Consider a system with only a single receive antenna element. Then, with probability 1, it is possible to identify up to $4K$ targets.

To illustrate the argument, we consider a system with a uniform linear array consisting of $N = 4$ antenna elements, $K = 3$ waveforms, and $L = 2$ targets with distinct locations. In this case $[\mathbf{a}(\theta)]_n = e^{j2\pi/\lambda_c d_x(m-1)sin(\theta)}$ which corresponds to the definition in (31) where the targets are coplanar with the antenna array, and the number of rows $M = 1$. Then, the determinantal variety is given by the L-minors of $\mathbf{A}^H\mathbf{W}$, of which there are 3. Defining $\alpha_l = e^{-j2\pi/\lambda_c d_x sin(\theta_l)},\ l = 0, 1$, the minors can be expressed as

$$\begin{aligned}
&w_{2,1}w_{3,0}(\alpha_0^3\alpha_1^2 - \alpha_1^3\alpha_0^2) - w_{2,0}w_{3,1}(\alpha_0^3\alpha_1^2 - \alpha_1^3\alpha_0^2) - w_{1,1}w_{3,0}(\alpha_0\alpha_1^3 - \alpha_1\alpha_0^3) - \\
&w_{1,0}w_{3,1}(\alpha_0^3\alpha_1 - \alpha_1^3\alpha_0) - w_{1,1}w_{2,0}(\alpha_0\alpha_1^2 - \alpha_1\alpha_0^2) - w_{1,0}w_{2,1}(\alpha_0^2\alpha_1 - \alpha_1^2\alpha_0) - \\
&w_{0,1}w_{3,0}(\alpha_1^3 - \alpha_0^3) - w_{0,0}w_{3,1}(\alpha_0^3 - \alpha_1^3) - w_{0,1}w_{2,0}(\alpha_1^2 - \alpha_0^2) - \\
&w_{0,0}w_{2,1}(\alpha_0^2 - \alpha_1^2) - w_{0,1}w_{1,0}(\alpha_1 - \alpha_0) - w_{0,0}w_{1,1}(\alpha_0 - \alpha_1) = 0
\end{aligned} \tag{53}$$

$$\begin{aligned}
&w_{2,2}w_{3,0}(\alpha_0^3\alpha_1^2 - \alpha_1^3\alpha_0^2) - w_{2,0}w_{3,2}(\alpha_0^3\alpha_1^2 - \alpha_1^3\alpha_0^2) - w_{1,2}w_{3,0}(\alpha_0\alpha_1^3 - \alpha_1\alpha_0^3) - \\
&w_{1,0}w_{3,2}(\alpha_0^3\alpha_1 - \alpha_1^3\alpha_0) - w_{1,2}w_{2,0}(\alpha_0\alpha_1^2 - \alpha_1\alpha_0^2) - w_{1,0}w_{2,2}(\alpha_0^2\alpha_1 - \alpha_1^2\alpha_0) - \\
&w_{0,2}w_{3,0}(\alpha_1^3 - \alpha_0^3) - w_{0,0}w_{3,2}(\alpha_0^3 - \alpha_1^3) - w_{0,2}w_{2,0}(\alpha_1^2 - \alpha_0^2) - \\
&w_{0,0}w_{2,2}(\alpha_0^2 - \alpha_1^2) - w_{0,2}w_{1,0}(\alpha_1 - \alpha_0) - w_{0,0}w_{1,2}(\alpha_0 - \alpha_1) = 0
\end{aligned} \tag{54}$$

$$\begin{aligned}
&w_{2,2}w_{3,1}(\alpha_0^3\alpha_1^2 - \alpha_1^3\alpha_0^2) - w_{2,1}w_{3,2}(\alpha_0^3\alpha_1^2 - \alpha_1^3\alpha_0^2) - w_{1,2}w_{3,1}(\alpha_0\alpha_1^3 - \alpha_1\alpha_0^3) - \\
&w_{1,1}w_{3,2}(\alpha_0^3\alpha_1 - \alpha_1^3\alpha_0) - w_{1,2}w_{2,1}(\alpha_0\alpha_1^2 - \alpha_1\alpha_0^2) - w_{1,1}w_{2,2}(\alpha_0^2\alpha_1 - \alpha_1^2\alpha_0) - \\
&w_{0,2}w_{3,1}(\alpha_1^3 - \alpha_0^3) - w_{0,1}w_{3,2}(\alpha_0^3 - \alpha_1^3) - w_{0,2}w_{2,1}(\alpha_1^2 - \alpha_0^2) - \\
&w_{0,0}w_{2,2}(\alpha_0^2 - \alpha_1^2) - w_{0,2}w_{1,1}(\alpha_1 - \alpha_0) - w_{0,0}w_{1,2}(\alpha_0 - \alpha_1) = 0
\end{aligned} \tag{55}$$

There is a high degree of symmetry among the minors. In fact, each is the exact same polynomial, just in different variables $w_{i,j}$. For example, comparing (53) and (54) we observe that they are the exact same polynomial, except for the fact that $w_{i,1} \mapsto w_{i,2}$. Similarly, comparing (54) and (55), we observe that $w_{i,0} \mapsto w_{i,1}$. This is, of course, not true in general. It is a direct consequence of the vandermonde structure

of $\mathbf{A}$. It is likely that there is an underlying symmetric group action relating each of these polynomials to one another, which would allow for a more indepth analysis of the algebraic variety. However, that is beyond the scope of this thesis. The point which equations (53)–(55) illustrate is that these are indeed polynomials of degree 2 in the entries of $\mathbf{W}$, thus exemplifying Lemma 3.

However, the argument raised in support of Lemma 3 does not stop there. The two lemmas in this section prove that up to $4K$ targets are identifiable with probability 1, even with only a single receive element. However, additional receive elements should enable the detection of more targets. Further, if the number of targets L supercedes $4K$, the previous results no longer hold, since $(\mathbf{W}^H\mathbf{A})^T$ can no longer be injective, by definition (the product is no longer tall). Now consider a receiver of R antenna elements such that the position of each antenna element from a reference element in the array is some integer multiple of the smallest antenna element spacing in the x direction dx_r, and the smallest antenna element spacing in the y direction dy_r. The R antenna elements can otherwise occupy arbitrary positions on this perpendicular lattice. Similar to the array response vectors of the transmitter, the response of every element to a signal impinging on the array from (θ_l, ϕ_l) can be expressed by products of the powers of $\mu = e^{-j2\pi \cdot dx_r \sin\theta\cos\phi}$ and $\xi = e^{-j2\pi \cdot dy_r \sin\theta\sin\phi}$. Therefore the L minors of the matrix $\left((\mathbf{W}^H\mathbf{A}) \odot \mathbf{B}\right)$ will be polynomials of degree L in the variables α, β, μ, and ξ. Using Lemma 3 then proves that $4KR$ targets will be identifiable with probability 1 using such an aribtarily populated lattice array. This is a practical result as such a receiver could be a URA which has one or more broken or repurposed elements, or one which is purposely sparsely populated in order to reduce effects of mutual coupling.

4.5 Simulation Results

The simulations in this section test two separate functionalities of the algorithm proposed in this section. In the first the beamforming algorithm is tested, and the capability of generating multiple invariant beampatterns is demonstrated. In the second, the performance of the beamforming algorithm is tested with respect to two figures of merit. The first is root mean square error as a function of SNR. The second is the probability of resolving two closely located targets as a function of SNR.

4.5.1 Beampattern and RIP

In this example we assume a fully populated URA consisting of 11 rows and 11 columns, with interelement spacing $\lambda_c/2$. K is chosen to be 8 in line with the findings of Section 3.3. Therefore, the number of waveforms actually transmitted is $4K = 32$. The desired sector for this example is $\Theta = [30, 50]$ and $\Phi = [100, 120]$. A transition region of 10 degrees was allowed on either side of the passband.

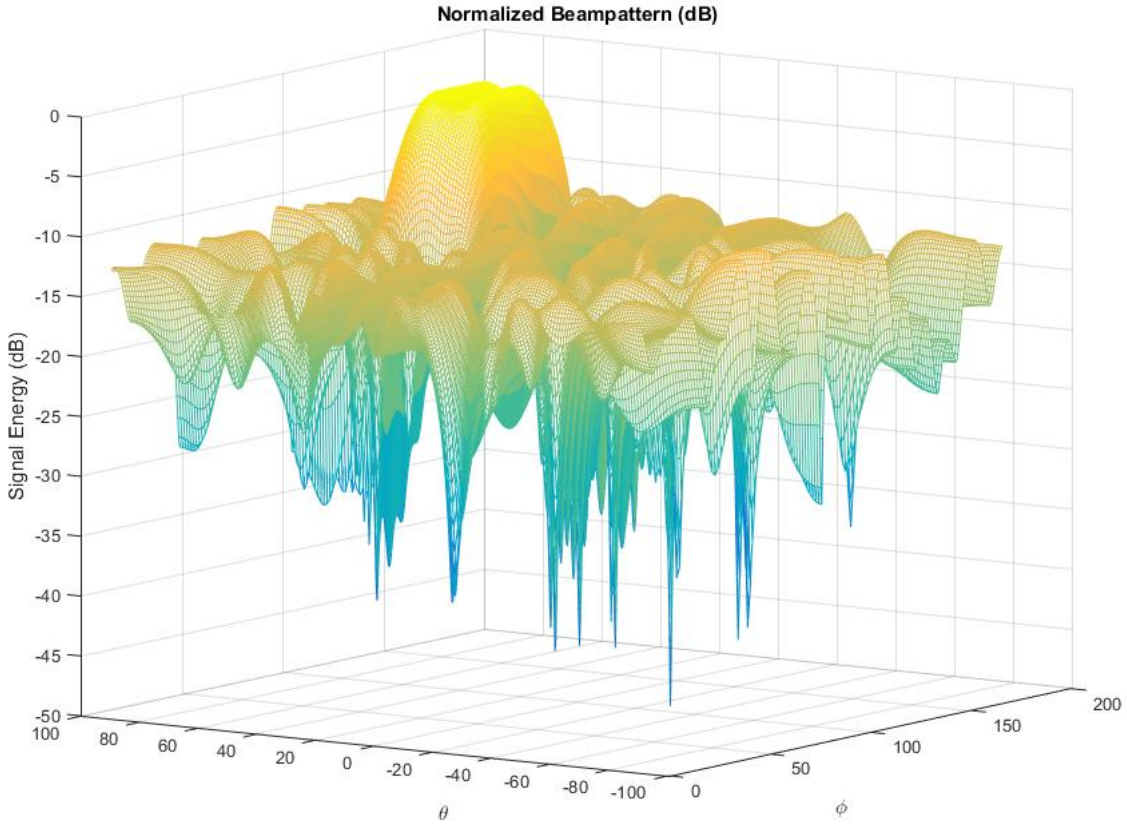

Figure 17: Normalized beampattern corresponding to 11×11 URA and 32 waveforms.

Fig. 17 shows the normalized beampattern over the entire 2D angular space. As we can see, there is a close correspondence between the desired beampattern and the achieved one. Energy is concentrated tightly within the desired sector, and the average transmit energy is approximately 10dB higher within the passband than in stopband.

To investigate the capability to enforce the RIP at the transmitter, we plot the normalized beampatterns of the 4 beamspace groups corresponding to $\mathbf{U}'_0$,$\mathbf{U}'_1$,$\mathbf{U}'_2$, and $\mathbf{U}'_3$, viewed from the elevation aspect. As we can see from Fig. 18, the beampatterns

are identical, and thus the rotational invariance property is enforced by the proposed structure.

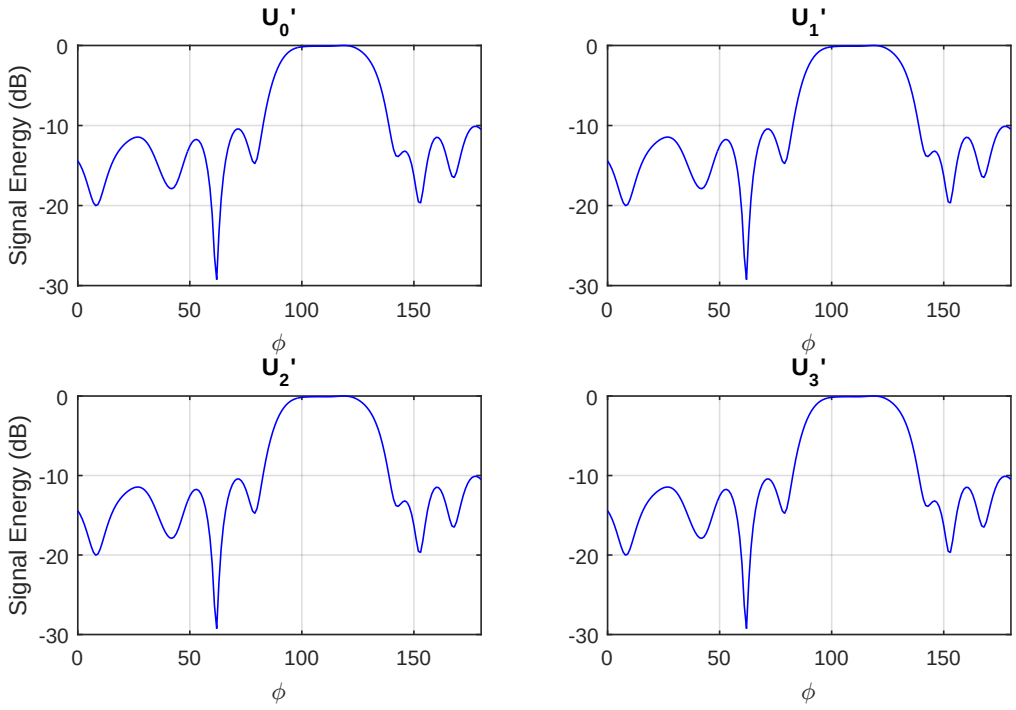

Figure 18: Elevation aspect of beampatterns corresponding to $\mathbf{U}'_0$ through $\mathbf{U}'_3$.

4.5.2 Parameter Estimation

In this example we assume the same system parameters as the previous example. To test the performance of the DOA estimation algorithm proposed in Section 4.3, we first consider RMSE performance. RMSE is calculated as

$$RMSE = \sqrt{\frac{1}{M}\sum_{m=1}^{M}\left(\frac{1}{L}\sum_{l=1}^{L}\|\hat{\rho}_l - \rho_l\|^2\right)} \tag{56}$$

where $\rho_l \triangleq [\theta_l, \phi_l]^T$. Thus, the figure of merit is actually the root mean square of the total euclidean error in 2D angular space. Fig. 19 shows the RMSE when estimating the location of a single target at $\theta = 47$, $\phi = 106$.

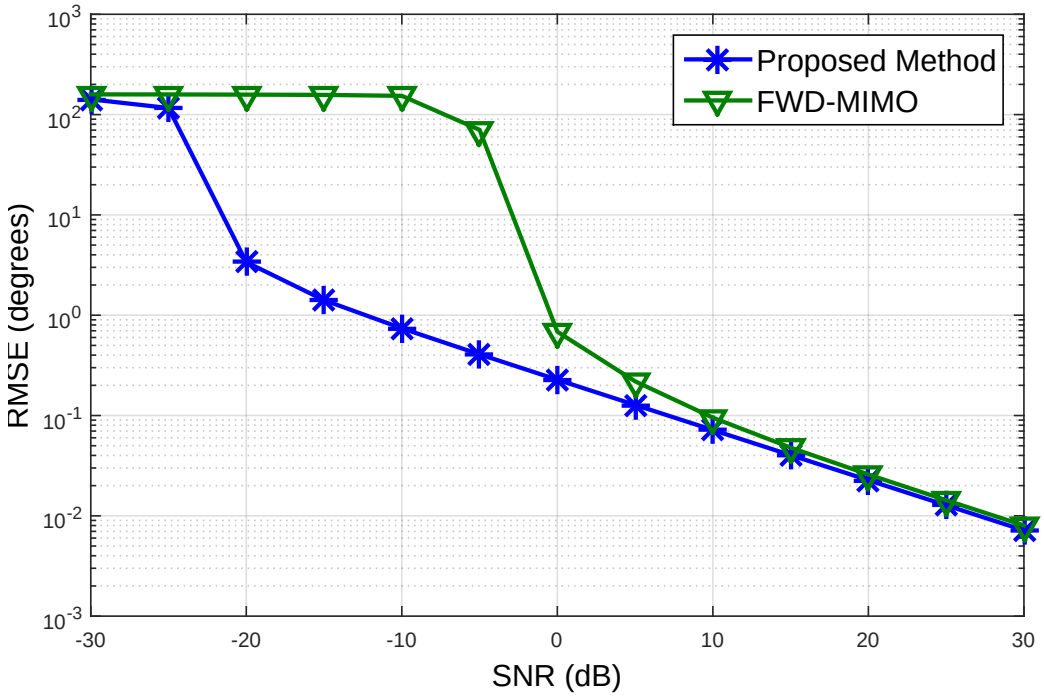

Figure 19: RMSE as a function of SNR. Single target located at $\theta = 47$, $\phi = 106$.

As we can see from the figure, the proposed method recovers from threshold behaviour due to eigenvalue swap at a much lower SNR than FWD-MIMO. At 0dB SNR the proposed method exhibits an RMSE of approximately 0.25^o. As SNR increases, the performance of FWD-MIMO approaches that of the proposed method. This performance characteristic is predictable as the benefits of the increased virtual aperture of FWD-MIMO are most noticeable at in the high SNR region. The reason for this was discussed in Section 1.

Fig. 20 compares the ability of the proposed method to FWD-MIMO with respect to the ability to resolve two closely located targets. Two targets were placed at $\theta_1 = 47^o$,$\theta_2 = 48^o$ and $\phi_1 = 106$,$\phi_2 = 107$. An approximately 10dB gap in SNR is observed between the proposed method and FWD-MIMO.

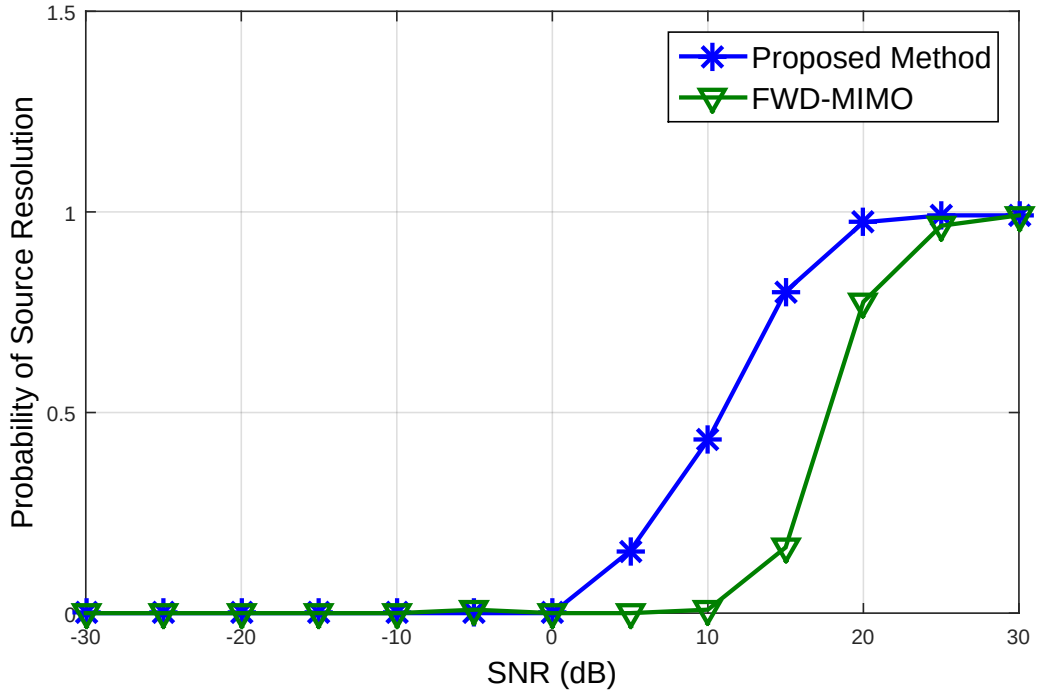

Figure 20: Probability of resolving two closely located targets as a function of SNR. Targets are located at $\theta_1 = 47^o$,$\theta_2 = 48^o$, and $\phi_1 = 106$,$\phi_2 = 107$.

It should be noted that even in the scenarios in which the performance of FWD-MIMO is close to that of the proposed method, the proposed method still benefits from a significantly lower computational complexity. The calculations required per iteration for the conventional MIMO radar are $\mathcal{O}((MNR)^3)$ while the calculations required for the proposed method are $\mathcal{O}((4KR)^3)$. With the configuration used in this example, FWD-MIMO radar requires approximately 50 times the calculations of the proposed method.

5 Conclusions

MIMO radar is an emerging remote sensing technology, ubiquitous in its potential applications. Within the paradigm of MIMO radar, a fundamental problem is that of transmit beamforming, in which energy is concentrated within a desired sector by means of forming linear combinations of basis signals. In this thesis, the problem of transmit beamforming for MIMO radar has been investigated in both 1 and 2 dimensions. Two novel algorithms have been presented which enhance both the capability and understanding of TB-MIMO radar.

The first algorithm exploits algebraic structure inherent in the TB-MIMO setting, as a consequence of array geometry. A wide class of array geometries are shown to enforce a polynomial structure on the transmitted signal. This polynomial structure is exploited by restricting the resulting beamforming problem to a polynomial ideal. This affords the designer complete control over the location of nulls in the beampattern. The performance of the algorithm was verified in comparison with the SDR method of [22] in several different scenarios. The proposed algorithm exhibited a significant advantage over the SDR in several respects. A gap of 19.8 dB in ASL was observed between the proposed and SDR methods. Also, an exact correspondence between the proposed algebraic structure and the simulation results is verified.

The second algorithm enforces the rotational invariance property (RIP) in the transmitted signal. This enables search free DOA estimation at the receive array without any constraint on the receive array structure. This is accomplished by designing a beampattern using only a subset of transmit array elements. Then, a novel beamspace transformation has been introduced which yields three other beamspace matrices with identical associated beampatterns. The algorithm has been verified by two sets of simulations. The first verified that the RIP is enforced by the proposed transformation. The second demonstrated a significant performance advantage in DOA estimation performance, and probability of source resolution over FWD-MIMO. Additionally, novel and strong results regarding parameter identifiability in the TB-MIMO regime have been shown.

The avenues for further work on this subject are numerous, and as promising as they are challenging. Modern algebraic techniques may hold the key which unlocks several yet unanswered questions about transmit beamforming. Further work must include an extension of the restriction method to 2D and 3D arrays. The extension will draw on the field of Algebraic Geometry, as the polynomials in question will no longer be univariate, but multivariate. Further, the wide range of array geometries which imply a polynomial structure raise the spectre of underlying group structure.

6 Bibliography

References

[1] G. R. Curry, *Radar System Performance Modeling*, 2nd ed. Boston: Artech House, 2005.

[2] M. I. Skolnik, *Introduction to Radar Systems*, 3rd ed. New York: McGraw-Hill, 2001.

[3] J.M. Headrick, M. I. Skolnik, "Over-the-Horizon radar in the HF band," *Proceedings of the IEEE*, vol. 62, no.6, pp. 664–673, June 1974.

[4] M. G. Amin and F. Ahmad,"Through-the-wall radar imaging: theory and applications," in R. Chellappa and S. Theodoridis (Eds), *Academic Press' Library in Signal Processing*, Academic Press, 2013.

[5] L. Xu, J. Li, and P. Stoica, "Target detection and parameter estimation for MIMO radar systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 44, no. 3, pp. 927–939, July 2008.

[6] I. Bekkerman and J. Tabrikian, "Target detection and localization using MIMO radars and sonars," *IEEE Trans. Signal Processing*, vol. 54, pp. 3873–3883, Oct. 2006.

[7] Q. He, R. S. Blum, H. Godrich, and A. M. Haimovich, "Cramer-Rao bound for target velocity estimation in MIMO radar with widely separated antennas," in *Proc. 42nd Annual Conf. Information Sciences and Systems (CISS'08)*, Princeton, NJ, Mar. 2008, pp. 123–127.

[8] Q. He, R. S. Blum, and A. M. Haimovich, "Noncoherent MIMO radar for location and velocity estimation: More antennas means better performance," *IEEE Trans. Signal Processing*, vol. 58, no. 7, pp. 3661–3680, Jul. 2010.

[9] C. W. Sherwin, J.P. Ruina, "Some Early Developments in Synthetic Aperture Radar Systems," *IRE Trans. Military Electronics*, vol. MIL-6, no. 2, pp. 111-115, April 1962.

[10] F. Daum and J. Huang, "MIMO radar: Snake oil or good idea," *IEEE Aerosp. Electron. Syst. Magazine*, pp. 8–12, May 2009.

[11] A.J. Paulraj, C.B. Papadias, "Space-time processing for wireless communications," *IEEE Signal Process. Magazine*, vol. 14, pp. 49-83, Nov. 1997.

[12] E. Telatar, "Capacity of multi-antenna Gaussian channels," *Eur. Trans. Telecommun.*, vol. 10, pp. 585–595, Nov. 1999.

[13] G.J. Foschini and M.J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Pers. Commun.*, vol. 6, pp. 311–335, 1998.

[14] E. Fishler, A. Haimovich, R. Blum, D. Chizhik, L. Cimini, and R. Valenzuela, "MIMO radar: An idea whose time has come," in *Proc. IEEE Radar Conf.*, Honolulu, Hawaii, USA, Apr. 2004, vol. 2, pp. 71–78.

[15] A. Haimovich, R. Blum, and L. Cimini, "MIMO radar with widely separated antennas," *IEEE Signal Processing Mag.*, vol. 25, pp. 116–129, Jan. 2008.

[16] J. Li and P. Stoica, "MIMO radar with colocated antennas," *IEEE Signal Processing Mag.*, vol. 24, pp. 106–114, Sept. 2007.

[17] A. Hassanien and S. A. Vorobyov, "Transmit energy focusing for DOA estimation in MIMO radar with colocated antennas," *IEEE Trans. Signal Process.*, vol. 59, pp. 2669–2682, June 2011.

[18] A. Hassanien and S. A. Vorobyov, "Phased-MIMO radar: A tradeoff between phased-array and MIMO radars," *IEEE Trans. Signal Process.*, vol. 58, pp. 3137–3151, June 2010.

[19] A. Khabbazibasmenj, A. Hassanien, S. A. Vorobyov, and M. W. Morency, "Efficient transmit beamspace design for search-free based DOA estimation in MIMO radar," *IEEE Trans. Signal Process.*, vol. 62, no. 6, pp. 1490–1500, Mar. 2014.

[20] A. Khabbazibasmenj, S. A. Vorobyov, A. Hassanien, and M. W. Morency, "Transmit beamspace design for direction finding in colocated MIMO radar with arbitrary receive array and even number of waveforms," in *Proc 46th Annual Asilomar Conf.*, Pacific Grove, California, USA, Nov. 2012, pp. 1307–1311.

[21] A. Hassanien, M. W. Morency, A. Khabbazibasmenj, S. A. Vorobyov, J.-Y. Park, and S.-J. Kim, "Two-dimentional transmit beamforming for MIMO radar with sparse symmetric arrays," in *Proc. IEEE Radar Conf.*, Ottawa, ON, Canada, Apr.-May 2013, pp. 1–5.

[22] D.R. Fuhrmann and G. San Antonio, "Transmit beamforming for MIMO radar systems using signal cross-correlation," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 44, no.1, pp. 171–186, Jan 2008.

[23] M. Shaghaghi and S.A. Vorobyov, "Subspace leakage analysis and improved DOA estimation with small sample size," *IEEE Trans. Signal Process.*, vol. 63, no. 12, pp. 3251–3265, June 2015.

[24] H.L. van Trees, *Optimum Array Processing: Part IV*. New Jersey: Wiley, 2002.

[25] M.X. Goemans, and D.P. Williamson, "Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming," *Journal of the Association for Computing Machinery*, vol. 42, no. 6, pp. 1115–1145, Nov. 1995.

[26] E.B. Vinberg, *A Course in Algebra*. Moscow: Factorial Press, 2001.

[27] Y. Nesterov, A. Nemirovsky, *Interior Point Polynomial Algorithms in Convex Programming*. Philadelphia, PA: SIAM, 1994.

[28] L.J. Griffiths, C.W. Jim, "An Alternative Approach to Linearly Constrained Adaptive Beamforming," *IEEE Trans. on Antennas and Propagation*, vol. AP-30, no.1, Jan 1982.

[29] A. Paulraj, R. Roy, and T. Khalaith, "Estimation of signal parameters via rotational invariance techniques- Esprit," in *Proc. 19th Annual Asilomar Conf.*, Pacific Grove, California, USA, Nov. 1985, pp. 83–89.

[30] N.D. Sidiropoulos, R. Bro, and G.B. Giannakis, "Parallel factor analysis in sensor array processing," *IEEE Trans. Signal Process.*, vol. 48, pp. 2377–2388, Aug. 2000.