Optimization of antenna array layouts in planar and non-planar geometries for DoA estimation

JOHAN ZANICHELLI
MASTER´S THESIS
DEPARTMENT OF ELECTRICAL AND INFORMATION TECHNOLOGY
FACULTY OF ENGINEERING | LTH | LUND UNIVERSITY

Optimization of antenna array layouts in planar and non-planar geometries for DoA estimation

Johan Zanichelli
jo6843za-s@student.lu.se

Department of Electrical and Information Technology
Lund University

Supervisor: Daniel Sjöberg

Examiner: Mats Gustafsson

July 3, 2024

Printed in Sweden
Tryckeriet i E-huset, Lund

Abstract

The use of antenna arrays allows for direction-of-arrival (DoA) estimation in a radar system. The antenna elements in an array are conventionally positioned equidistantly on a grid, but they can also be placed relatively freely to form sparse antenna arrays. These non-conventional layouts could potentially offer improved DoA capabilities. In this work, a particle swarm optimization algorithm has been implemented to find optimal array layouts according to a set of requirements. These requirements have been based on the ambiguity function, which is related to the risk of error and to the resolution of the DoA estimations. The results indicate that the implemented method works as intended, and that it can be applied to two different geometries. The first geometry is planar and circular, while the second consists of multiple planar surfaces, positioned to cover 360° in azimuth. A comparison was made between the noise performance of the optimized arrays and that of a reference array, in which only small differences were seen. The chosen method was implemented to work for any given dimension of the specified geometries, and could potentially be used for restrictive dimensions and a varying number of antenna elements.

Keywords: Radar, DoA, MIMO, Particle Swarm Optimization, Sparse Arrays

Acknowledgements

First and foremost, I would like to express my gratitude to my supervisors Magnus and Daniel. You have both offered much support and insightful comments during the work. I would also like to extend a thank you to all other master thesis students in the team who all have made these last six months a real treat. Finally, thanks to my friends and family for their support during my studies.

Popular Science Summary

The existence of electromagnetic (EM) waves was first experimentally shown in the late 19th century, a discovery that would revolutionize the way we communicate. From the first radio to the wireless systems of today, our world has become increasingly dependent on EM waves in our everyday lives. Parallel to the communication applications, this discovery also led to the development of radar. A radar system works by sending out signals through an antenna, and listening for a reflected signal. Oftentimes, it is desired to determine in what direction you are receiving the reflected signal in order to know where your target is. To explain how this is done, it is worth taking a look at ourselves. Without much thought, humans are able to locate sounds that we hear, which is mainly thanks to us having two ears. The direction-of-arrival (DoA) can be estimated by comparing how loud the signal is at each ear, and by sensing the time difference between them. The latter is essentially what is used by a radar system to do a similar estimation. For the same reason that we need two ears to determine the DoA, a radar will also need more than one antenna to do the same thing. Something that we are unable to experience is that the ability to determine the DoA is dependent on how the antennas are placed. For instance, if the antennas are badly placed, it can "sound" like a single signal is coming from multiple directions. In order to help with this problem, a radar system often has more than two antennas and these antennas can be positioned in different ways. This raises a question, which is also the aim of this thesis: how do you place the antennas to get as good DoA estimation as possible? Unfortunately, there is often an infinite amount of possible solutions to this problem, so it is impossible to try all of them. Instead, we can once again be inspired by a natural phenomenon. This time, we look at the way that bird flocks collaborate when they hunt for insects. In such a flock, each individual bird wants to find a position in the air with the most amount of food. If they have once found such a place, they will tend to return there. At the same time, they will also look at where the other birds are flying, as there could be another spot with even more insects. While this might sound abstract, this same approach can be used to find good antenna positions. By allowing an algorithm to search for antenna positions, it was possible to find satisfactory layouts in different types of geometries. This method could prove useful when it is needed to place antennas in restrictive geometries, where it is difficult to find suitable placements.

Table of Contents

Abbreviations

DoA Direction of Arrival
ToF Time of Flight
CW Continuous Wave
FMCW Frequency Modulated Continuous Wave
DFT Discrete Fourier Tansform
PEC Perfect Electric Conductor
PMC Perfect Magnetic Conductor
RF Radio Frequency
AoA Angle of Arrival
Rx Receiving
Tx Transmitting
MIMO Multiple Input Multiple Output
SNR Signal to Noise Ratio
CBF Classical Beamformer
FFT Fast Fourier Transform
MUSIC Multiple Signal Classification
SLL Sidelobe Level
MLW Main Lobe Width
PSLL Peak Sidelobe Level
ULA Uniform Linear Array
CRLB Cramér-Rao Lower Bound
GA Genetic Algorithm
PSO Particle Swarm Optimization
RoI Region of Interest
SBR Shooting and Bouncing Rays
FoV Field of View

Chapter 1
Introduction

1.1 Background

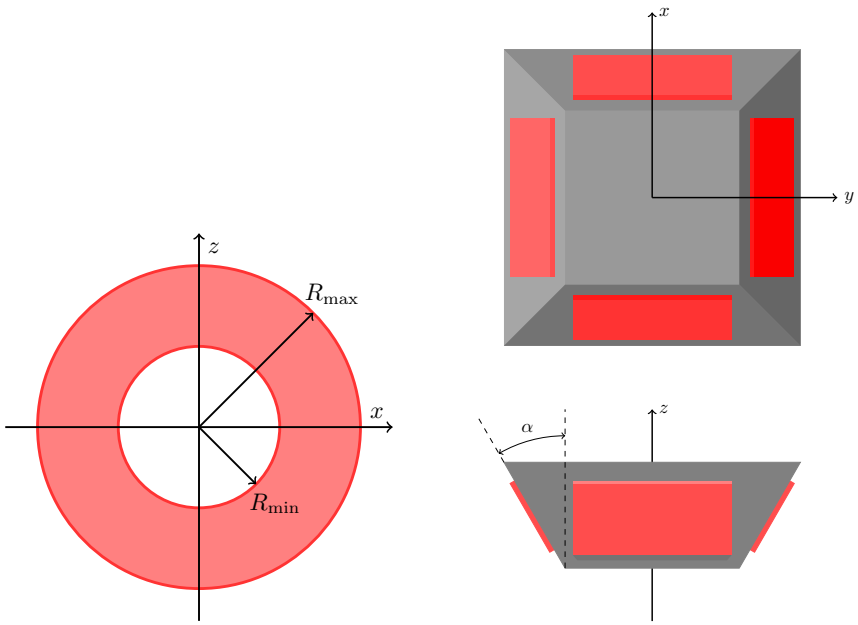
The history of radar can be traced back to the end of the 19th century, and to early electromagnetic experiments conducted by Heinrich Hertz. With these experiments, Hertz was able to experimentally show the existence of electromagnetic waves, including their ability to reflect of surfaces [1]. This discovery led to numerous inventions in the following years, of which the wireless telegraph by Guglielmo Marconi is perhaps the most famous [2]. Less famous is the invention of radar, which was first patented in 1904 by Christian Hülsmeyer as a way of detecting the presence of metal objects, such as ships and trains. While the principles of radars were studied in the coming decades, it was not widely researched until the 1930s, when the main focus was on military applications. During WW2, radars became an essential part in Britain's defense against German bombers. After WW2, the research was broadened to civilian applications, such as weather radars, which gained the capability of not only detecting the range to targets, but also the radial velocity of the target [3]. Thanks to improvements in semiconductor techniques and signal processing, radars have become frequently used in the automotive industry in recent years. Here, cameras can be complemented by radar modules that offer longer range and better visibility in bad weather conditions [4]. These advantages can also motivate the use of radars in conjunction with surveillance cameras.

Beyond range and velocity measurements, many radars can also detect the direction to the target. This is called the direction of arrival (DoA), and is commonly implemented by using several spatially separated receiving antennas. As the received signal is different between the antennas depending on the DoA, this direction can (potentially) be estimated. The quality of the DoA estimation is directly linked to the relative positions of the antennas [5].

1.2 Aim and scope

The aim of this thesis is to investigate how the antenna elements in an array should be placed in order to give the best possible direction of arrival estimation within a specified angular range. Furthermore, it is of special interest that the estimation is equally well for all angles within the angular range. The two geometries on which

the elements should be placed are illustrated in red in Figure 1.1. In this thesis, these geometries are called the circular and multiple plane geometry, respectively. The implementation of the method should be applicable to arbitrary dimensions of the geometries, and notably an arbitrary amount of sides in the multiple plane geometry. Furthermore, the multiple plane geometry should be designed to have invariant performance over azimuth, and should have the same antenna layout on each individual side. The two geometries are chosen as they can be easily integrated with surveillance cameras. The circular geometry could be placed around a camera lens, while the multiple plane geometry could surround a so-called dome camera which offers 360° coverage around the z axis.



(a) Circular geometry. **(b)** Multiple plane geometry.

Figure 1.1: Illustration of the two different geometries on which the antennas can be placed. The allowed positions are illustrated in red.

In order to achieve the aim, a method of assessing the performance of the arrays is presented. Based on this method, an optimization algorithm is implemented. The optimized arrays are simulated, and their noise performance is analyzed.

Chapter 2
Theory

This chapter is divided into four main parts, with the first being a quick introduction to the basic principles of radar. After this, some general antenna theory is presented and a specific type of antenna, the patch antenna, is described in further detail. The most substantial part of the theory is about antenna arrays, and how they are used for DoA estimation. A signal model for antenna arrays is presented, from which two DoA algorithms are introduced. Some ways of assessing the performance of an array are also described. Finally, two suitable optimization methods are outlined.

2.1 Radar

Radar, an acronym for Radio Detection and Ranging, is a system that uses electromagnetic waves to measure its surroundings. In its most simple form a radar system only detects the range to a target. This is achieved by transmitting and timing an electromagnetic pulse that is reflected by the target and then received. By measuring the time of flight (ToF) of the pulse, the range to the target can be determined. Modern radar systems have added capabilities beyond ranging, and are used in a wide range of applications. Some examples are military radars, weather radars and satellite mapping radars [6]. As the costs associated with radar systems have reduced greatly in the last decades, they are now used in a wider range of applications, such as for safety features in cars [4].

Radar systems are often divided into two main categories, pulsed and continuous wave (CW), based on the characteristics of the transmitted signals. Pulsed radars repeatedly transmit during short periods of time, and listen for the returned signals. This allows for both ToF measurement and target velocity measurements, the latter by using the Doppler shift of the received signal. In many cases, the receiver is blocked during transmission as to not overpower the receiver. If a signal is reflected and received during this time, it is not detected. Thus, pulsed radars usually have a lower limit for range detection. In contrast, the simplest CW radars transmit a signal with constant frequency over a relatively long period of time. There is no way of measuring the ToF with this type of CW radar, and they are thus only used for velocity measurements. Furthermore, as signals are received and transmitted in the same time, the transmitted power has to be relatively low compared to pulsed radar systems in order to prevent substantial crosstalk between

the transmitting and receiving sides [6]. By introducing frequency modulation of the transmitted signal, a CW radar can also determine range. This type of signal is called frequency modulated continuous wave (FMCW) and is widely used in for instance the automotive industry [7].

2.2 FMCW radar

An FMCW radar system frequency modulates the transmitted signal, resulting in a linear frequency chirp described by:

$$f(t) = f_c + \frac{B}{T}t, \quad -\frac{T}{2} \leq t \leq \frac{T}{2} \tag{2.1}$$

where f is the frequency, f_c the center frequency, B the bandwidth and T the chirp duration. The amplitude variations in each chirp is described by:

$$A(t) = A_0 \cos\left(2\pi \int_{-T/2}^{t} f(t)\mathrm{d}t\right) = A_0 \cos\left(2\pi\left(f_c t + \frac{1}{2}\frac{B}{T}t^2\right) - \varphi_{T_0}\right) \tag{2.2}$$

where φ_{T_0} is the initial phase of the chirp. The received signal is mixed with the transmitted signal, producing a constant intermediate frequency, f_{IF}, dependent on the velocity, v, and range, R, of the target as [7]:

$$f_{\mathrm{IF}} = \frac{2f_c v}{c} + \frac{2BR}{Tc} \tag{2.3}$$

where c is the speed of light. The velocity v only includes movements towards or away from the radar system, as only movements in these directions produces a Doppler shift. If the parameters are chosen beneficially, the effect of v on f_{IF} are negligible. Each received chirp is sampled, and a discrete Fourier ransform (DFT) is performed, resulting in a peak at the intermediate frequency [7]. Using (2.3), the range can then be estimated. This process is illustrated in Figure 2.1.

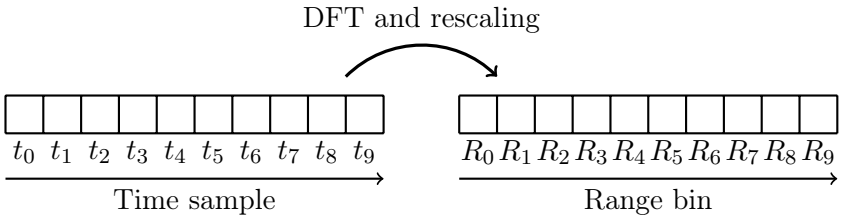


Figure 2.1: Illustration of the time samples and their transformation to range bins through a DFT.

In order to also estimate the velocity, v, multiple chirps have to be analyzed similarly to the range estimation. In order to visualize this process it can be useful to illustrate the samples of multiple chirps using a grid as in Figure 2.2. Performing a DFT through the time samples gives the range while a DFT through the chirps

gives the velocity. More precisely, a Doppler velocity of v gives a frequency peak at [7]:

$$f_{\text{Doppler}} = \frac{2f_c v}{c} \tag{2.4}$$

where the velocity is positive if the object is moving towards the radar.

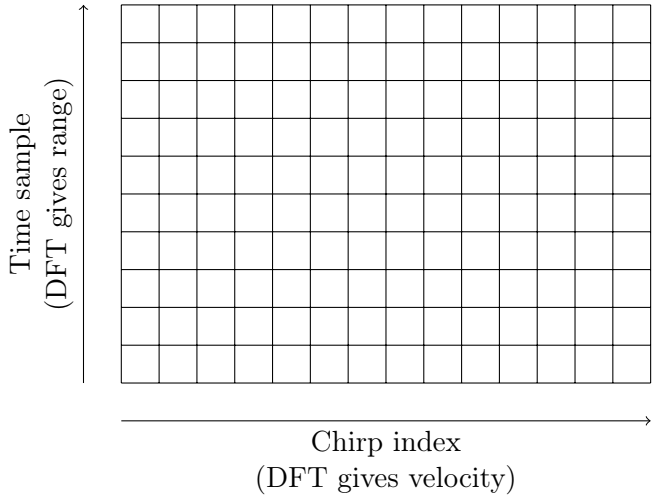


Figure 2.2: Illustration of the radar data grid.

2.3 Antenna radiation

An antenna radiates electromagnetic fields as a result of currents on the structure of the antenna. If the currents are time-harmonic, the radiated electric field is given by $\boldsymbol{E}(\boldsymbol{r},t) = \text{Re}\big(\boldsymbol{E}(\boldsymbol{r},\omega)e^{-\text{i}\omega t}\big)$, where $\boldsymbol{E}(\boldsymbol{r},\omega)$ is the electric field in the frequency domain and ω is the angular frequency. In vacuum, $\boldsymbol{E}(\boldsymbol{r},\omega) = \boldsymbol{E}(\boldsymbol{r})$ is determined by [8]:

$$\boldsymbol{E}(\boldsymbol{r}) = \text{i}\omega\mu_0 \left(\mathbf{I} + \frac{1}{k^2}\nabla\nabla\right) \cdot \iiint_V \frac{e^{\text{i}k|\boldsymbol{r}-\boldsymbol{r'}|}}{4\pi|\boldsymbol{r}-\boldsymbol{r'}|}\boldsymbol{J}(\boldsymbol{r'})\ \text{d}v' \tag{2.5}$$

where μ_0 is the permeability of vacuum, $\boldsymbol{J}(\boldsymbol{r'})$ is the current density, $\mathbf{I}$ is the identity operator, $\boldsymbol{r}$ is the observation point, $\boldsymbol{r'}$ is the source point and k is the wavenumber, related to the wavelength λ as:

$$k = \frac{2\pi}{\lambda} \tag{2.6}$$

Equation (2.5) can be simplified when the electric field is observed in the far field, which intuitively is at a distance far from the radiating sources. The requirements on the observation point $\boldsymbol{r}$ for being in the far field can be formulated mathematically as [8]:

$$\begin{cases} r \gg d \\ r \gg kd^2 \\ r \gg \lambda \end{cases} \tag{2.7}$$

where d is the largest extent of the current density (in the case of this thesis, the largest dimension of the antenna). However, other definitions of the far field exist, for instance as $r > 2d^2/\lambda$ in [9]. When the requirements for far field are met, the distance between the observation points and source points, $|\boldsymbol{r} - \boldsymbol{r'}|$, can be approximated as:

$$|\boldsymbol{r} - \boldsymbol{r'}| \approx r - \hat{\boldsymbol{r}} \cdot \boldsymbol{r'} \tag{2.8}$$

Using this approximation gives the following far field formulation of the electric field [8]:

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{e^{\mathrm{i}kr}}{r} \boldsymbol{F}(\hat{\boldsymbol{r}}) + \mathcal{O}(r^{-2}) \tag{2.9}$$

where $\mathcal{O}(r^{-2})$ includes all parts that decrease as r^{-2} or faster and $\boldsymbol{F}(\hat{\boldsymbol{r}})$, the radiation vector, is given by:

$$\boldsymbol{F}(\hat{\boldsymbol{r}}) = \frac{\mathrm{i}k\eta_0}{4\pi} \hat{\boldsymbol{r}} \times \left(\iiint_V e^{-\mathrm{i}k\hat{\boldsymbol{r}} \cdot \boldsymbol{r'}} \boldsymbol{J}(\boldsymbol{r'}) \, \mathrm{d}v' \times \hat{\boldsymbol{r}} \right) \tag{2.10}$$

where η_0 is the impedance of vacuum and where it can be seen that $\boldsymbol{F}(\hat{\boldsymbol{r}})$ is orthogonal to $\hat{\boldsymbol{r}}$. The corresponding magnetic field, $\boldsymbol{H}(\boldsymbol{r})$, can be calculated using Maxwell's equations for a vacuum that is free of sources:

$$\begin{cases} \nabla \times \boldsymbol{E} = \mathrm{i}k\eta_0 \boldsymbol{H} \\ \nabla \times \boldsymbol{H} = -\mathrm{i}\frac{k}{\eta_0} \boldsymbol{E} \end{cases} \tag{2.11}$$

which gives:

$$\boldsymbol{H}(\boldsymbol{r}) = -\mathrm{i}\frac{1}{k\eta_0} \nabla \times \boldsymbol{E} = \frac{e^{\mathrm{i}kr}}{\eta_0 r} \hat{\boldsymbol{r}} \times \boldsymbol{F}(\hat{\boldsymbol{r}}) + \mathcal{O}(r^{-2}) \tag{2.12}$$

From the electric and magnetic fields, the Poynting vector, $\boldsymbol{S}(\hat{\boldsymbol{r}})$, is defined as $\boldsymbol{S} = \frac{1}{2}\boldsymbol{E} \times \boldsymbol{H}^*$ and gives the complex power density of an electromagnetic wave in the direction of $\boldsymbol{S}$. Furthermore, the average Poynting vector is $\boldsymbol{S}_{\mathrm{av}} = \frac{1}{2}\mathrm{Re}(\boldsymbol{E} \times \boldsymbol{H}^*)$ [9, 10]. Using $\boldsymbol{E}(\boldsymbol{r})$ from (2.9) and $\boldsymbol{H}(\boldsymbol{r})$ from (2.12) and ignoring the terms $\mathcal{O}(r^{-2})$ gives:

$$\boldsymbol{S}_{\mathrm{av}}(\hat{\boldsymbol{r}}) = \frac{1}{2}\mathrm{Re}(\boldsymbol{E}(\hat{\boldsymbol{r}}) \times \boldsymbol{H}(\hat{\boldsymbol{r}})^*) = \frac{1}{2\eta_0 r^2} \boldsymbol{F}(\hat{\boldsymbol{r}}) \times (\hat{\boldsymbol{r}} \times \boldsymbol{F}(\hat{\boldsymbol{r}})^*) = \frac{\hat{\boldsymbol{r}}}{2\eta_0 r^2} \left| \boldsymbol{F}(\hat{\boldsymbol{r}}) \right|^2 \tag{2.13}$$

A commonly used metric for an antenna is the radiation intensity, $U(\hat{\boldsymbol{r}})$, which states the radiated power from an antenna in the far field per unit solid angle. It can be defined using the average Poynting vector as [9]:

$$U(\hat{\boldsymbol{r}}) = r^2 \, \hat{\boldsymbol{r}} \cdot \boldsymbol{S}_{\mathrm{av}}(\hat{\boldsymbol{r}}) = r^2 \left| \boldsymbol{S}_{\mathrm{av}}(\hat{\boldsymbol{r}}) \right| = \frac{\left| \boldsymbol{F}(\hat{\boldsymbol{r}}) \right|^2}{2\eta_0} \tag{2.14}$$

From the radiation intensity, the antenna gain, $G(\hat{\boldsymbol{r}})$, can also be defined. This gives the ratio between an antenna's radiation intensity and the radiation intensity from a (hypothetical) isotropic antenna, both supplied with the same power. The definition is [9]:

$$G(\hat{\boldsymbol{r}}) = \frac{4\pi U(\hat{\boldsymbol{r}})}{P_{\text{in}}} \tag{2.15}$$

where P_{in} is the supplied power.

Another important antenna metric, apart from the total gain, is the polarization of the antenna. In spherical coordinates, $\boldsymbol{F}(\hat{\boldsymbol{r}})$ is given by two orthogonal components [9]:

$$\boldsymbol{F}(\hat{\boldsymbol{r}}) = F_\theta(\hat{\boldsymbol{r}})\hat{\boldsymbol{\theta}} + F_\phi(\hat{\boldsymbol{r}})\hat{\boldsymbol{\phi}} \tag{2.16}$$

where the directions of $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\phi}}$ are illustrated in Figure 2.3. Using the notation from (2.16), the gain can similarly be divided into two components:

$$G(\hat{\boldsymbol{r}}) = G_\theta(\hat{\boldsymbol{r}}) + G_\phi(\hat{\boldsymbol{r}}) \tag{2.17}$$

where $G_\theta(\hat{\boldsymbol{r}})$ and $G_\phi(\hat{\boldsymbol{r}})$ are sometimes called partial gains [9].

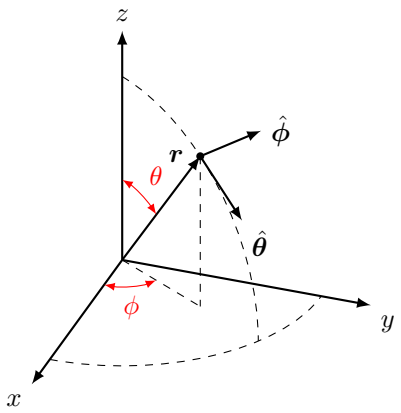


Figure 2.3: Illustration of the spherical coordinate system used in this work.

2.4 Patch antenna

A common antenna type is the patch antenna, also called the microstrip antenna. The patch antenna is a conformal antenna, meaning it can be placed on a flat or curved shape, making it especially suitable in low-profile products. It is also characterized by its low production cost as well as easy circuit integration. Depending on the desired radiation pattern, polarization and other factors, the patch antenna can take many forms. An example of a specific shape, the rectangular patch, is illustrated in Figure 2.4. The patch itself is made of a conducting material and is placed above a ground plane (both illustrated in red, where the ground plane is

seen as a thin red layer at the bottom). The sheet between these two is often made of a substrate material, which allows for easy feeding using a microstrip line. A microstrip line is a type of transmission line that, like the patch antenna, consists of a thin conducting material above a ground plane [9].

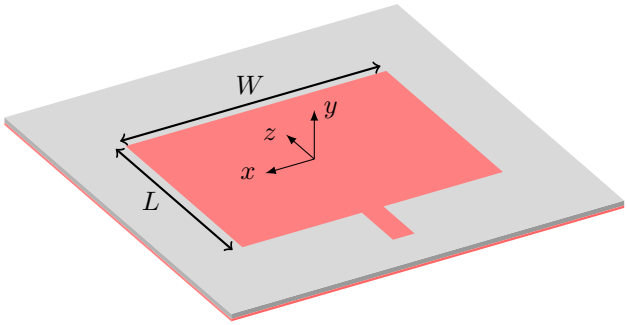


Figure 2.4: Example of a simple rectangular patch antenna.

The resonances and the radiation patterns of the patch antenna can be analytically analyzed using different approximate models, one of which is the cavity model. The cavity model approximates that the electromagnetic field in the antenna structure is confined to the volume below the patch, and above the ground plane. Both the ground plane and the patch are treated as perfect electric conductors (PEC) while the sides to the confined volumes are treated as perfect magnetic conductors (PMC). Solving Maxwell's equations for this geometry gives that the lowest resonant frequency, $f_{\text{res},0}$, is given by:

$$f_{\text{res},0} = \frac{c}{2L\sqrt{\epsilon_r}} \tag{2.18}$$

where ϵ_r is the relative permittivity of the substrate [9]. From (2.18) an approximate value of the length can thus be calculated based on the desired resonance frequency. In the cavity model, all radiation stems from the induced currents in the PMC surfaces. The main contribution to the far field radiation is from induced currents along the width of the patch. From (2.9) and (2.10), it can thus be seen that the $\boldsymbol{E}$-field is mainly polarized in the yz plane as given in Figure 2.4. The width of the antenna has an effect on its bandwidth, and a typical starting value in a design process is [9]:

$$W = \frac{c}{2f_{\text{res},0}}\sqrt{\frac{2}{\epsilon_r + 1}} \tag{2.19}$$

The patch has to be fed in some way in order to excite the fundamental mode. This can be done as in Figure 2.4, where a microstrip line extends from one of the sides. However, this does not allow for tuning of the input impedance of the antenna, which preferably should be matched to the microstrip line which in turn should be matched to the radio frequency (RF) front end [9]. A discontinuity in impedance between these parts leads to both power loss and distortion of transmitted and received signals. The distortion occurs as the signals are reflected multiple times at the boundaries [11]. One way of solving this is by using an inset fed patch as

illustrated in Figure 2.5, where i_l is the inset length, i_w is the inset width and f_w is the feed width. According to the cavity model, the $\boldsymbol{E}$-field under a patch varies sinusoidally in the length direction, which results in a varying input impedance depending on where the patch is fed [9]. This theoretical variation is illustrated in Figure 2.6, where it is assumed that the feed itself does not affect the variations of the electric field under the patch. The desired impedance can thus be achieved by iterating the inset length in the design process. Usually, an input impedance of 50 Ω is used. An appropriate width of the microstrip line can be calculated using an empirical formula (for instance given in [9]).

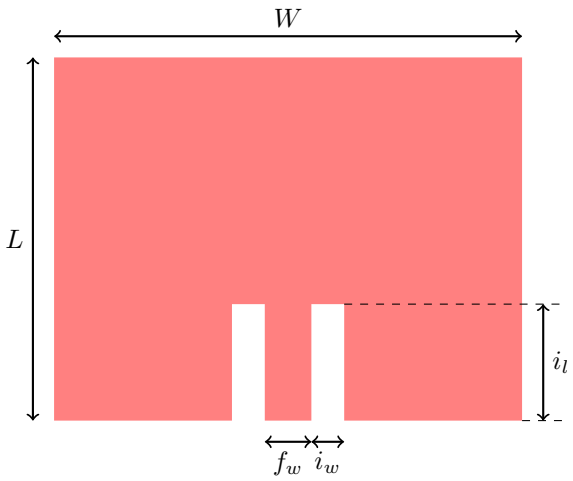


Figure 2.5: Rectangular patch antenna with inset feed.

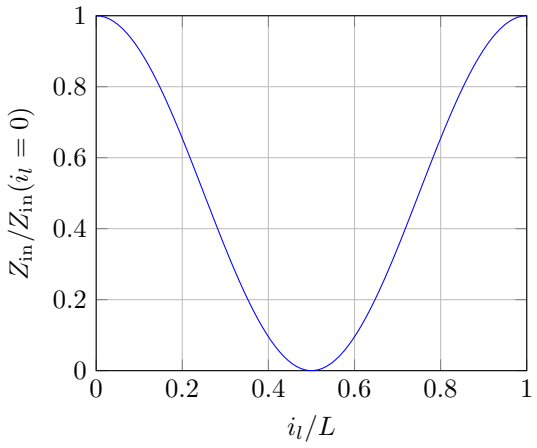


Figure 2.6: Variation of the input impedance as a function of feed position.

2.5 Antenna arrays

In many radar applications it is desired to detect the direction from which an incoming wave is arriving. This direction is often called the direction of arrival (DoA) or the angle of arrival (AoA). Determining the DoA can be done by sweeping a directive antenna over an angular range, either mechanically or electronically. By detecting in which direction the received signal amplitude is at a maximum, the DoA can be estimated. More commonly, a DoA can be estimated by measuring the phase of a reflected wave in spatially separated points, which requires the use of an antenna array [6].

A simple receiver antenna array is illustrated in Figure 2.7 together with an impinging wave. The distance to the source of the wave is assumed to be large, meaning that the wavefront (illustrated in red) is planar. As the wavefront reaches the elements in different times, the phase of the induced currents on the elements vary. The phase φ of an element placed at the position $\boldsymbol{d}$ with a wave impinging in the direction $\hat{\boldsymbol{r}}$ is $\varphi = k\hat{\boldsymbol{r}} \cdot \boldsymbol{d}$. Thus, with an array of N elements defined by the element positions $\boldsymbol{d}_n$, $n = 0, ..., N-1$ the phase variations on the array can be described by the array vector $\boldsymbol{a}(\hat{\boldsymbol{r}})$ [5]:

$$\boldsymbol{a}(\hat{\boldsymbol{r}}) = \begin{bmatrix} e^{\mathrm{i}k\hat{\boldsymbol{r}}\cdot\boldsymbol{d}_0} \\ \vdots \\ e^{\mathrm{i}k\hat{\boldsymbol{r}}\cdot\boldsymbol{d}_{N-1}} \end{bmatrix} \tag{2.20}$$

As is seen in section 2.8, only the relative phases between elements are of importance in the array vector. This means that the origin for the positions $\boldsymbol{d}_n$ can be set arbitrarily.

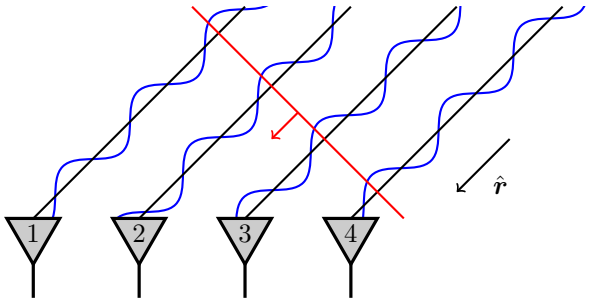


Figure 2.7: Antenna array with impinging wavefront.

2.6 MIMO systems

In many applications it can be beneficial to not only use an array of receiving (Rx) antennas, but also transmitting (Tx) antennas. This forms a multiple input multiple output (MIMO) radar system, assuming that the transmitted signals are orthogonal to each other. This orthogonality can be achieved by time-division of

the transmitted signals. By using a MIMO array, the amount of usable phase information for DoA estimation can be greatly increased. A MIMO system with N Rx antennas and M Tx antennas is equivalent to an array with one Tx antenna and $N \times M$ Rx antennas, assuming that the radar scene does not change between the transmitted signals. The resulting $N \times M$ array is often called a virtual array, with a virtual steering vector $\boldsymbol{a}_v(\hat{\boldsymbol{r}})$. A MIMO array is illustrated in Figure 2.8, where it is seen that for a specific direction $\hat{\boldsymbol{r}}$, the transmitted signals have different phases which can be calculated as $\varphi_{\mathrm{Tx},m} = k\hat{\boldsymbol{r}} \cdot \boldsymbol{d}_{\mathrm{Tx},m}$. These phases are added to the phases of the receiving antennas, meaning that when the first Tx antenna transmits, the resulting received phases are $\varphi_{m=0,n} = k\hat{\boldsymbol{r}} \cdot (\boldsymbol{d}_{\mathrm{Rx},n} + \boldsymbol{d}_{\mathrm{Tx},0})$. In total, the combination of all Tx and Rx antennas yields the virtual array vector $\boldsymbol{a}_V(\hat{\boldsymbol{r}})$ [12]:

$$\boldsymbol{a}_V(\hat{\boldsymbol{r}}) = \boldsymbol{a}_{\mathrm{Tx}}(\hat{\boldsymbol{r}}) \otimes \boldsymbol{a}_{\mathrm{Rx}}(\hat{\boldsymbol{r}}) = \begin{bmatrix} e^{\mathrm{i}k\hat{\boldsymbol{r}}\cdot(\boldsymbol{d}_{\mathrm{Rx},0}+\boldsymbol{d}_{\mathrm{Tx},0})} \\ \vdots \\ e^{\mathrm{i}k\hat{\boldsymbol{r}}\cdot(\boldsymbol{d}_{\mathrm{Rx},0}+\boldsymbol{d}_{\mathrm{Tx},M-1})} \\ \vdots \\ e^{\mathrm{i}k\hat{\boldsymbol{r}}\cdot(\boldsymbol{d}_{\mathrm{Rx},N-1}+\boldsymbol{d}_{\mathrm{Tx},0})} \\ \vdots \\ e^{\mathrm{i}k\hat{\boldsymbol{r}}\cdot(\boldsymbol{d}_{\mathrm{Rx},N-1}+\boldsymbol{d}_{\mathrm{Tx},M-1})} \end{bmatrix} \tag{2.21}$$

where $\otimes$ denotes the Kronecker product. Thus, the positions of the virtual antenna elements are simply $\boldsymbol{d}_{n,m} = \boldsymbol{d}_{\mathrm{Rx},n} + \boldsymbol{d}_{\mathrm{Tx},m}$.

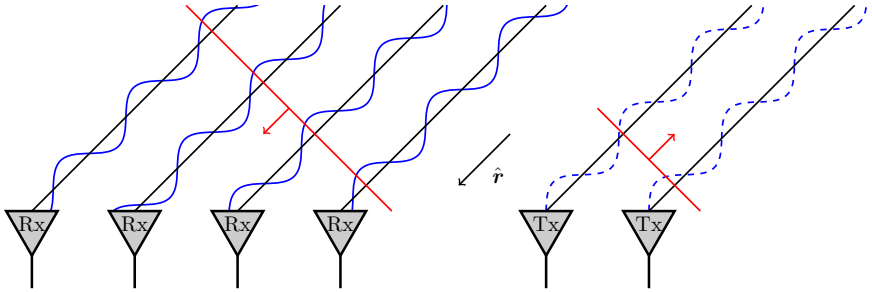


Figure 2.8: MIMO antenna array with an outgoing and incoming wavefront. The waves from the Tx antennas are dashed to illustrate that they do not transmit simultaneously.

2.7 Signal model

The output from a receiver antenna array is, as described in the previous sections, dependent on the positions of the antenna elements and the direction of the impinging waves. Beyond this, the amplitude and phase of the waves also impact the

output, as does added noise. Based on multiple signals $s_d(t)$ from the directions $\hat{\boldsymbol{r}}_d$, $, d = 1, ..., D$ and an array vector $\boldsymbol{a}(\hat{\boldsymbol{r}})$, the received signal $\boldsymbol{x}(t)$ is:

$$\boldsymbol{x}(t) = \begin{bmatrix} \boldsymbol{a}(\hat{\boldsymbol{r}}_1) & \boldsymbol{a}(\hat{\boldsymbol{r}}_2) & ... & \boldsymbol{a}(\hat{\boldsymbol{r}}_D) \end{bmatrix} \begin{bmatrix} s_1(t) \\ s_2(t) \\ \vdots \\ s_D(t) \end{bmatrix} + \boldsymbol{\eta}(t) \tag{2.22}$$

or when sampled:

$$\boldsymbol{x}[k] = \begin{bmatrix} \boldsymbol{a}(\hat{\boldsymbol{r}}_1) & \boldsymbol{a}(\hat{\boldsymbol{r}}_2) & ... & \boldsymbol{a}(\hat{\boldsymbol{r}}_D) \end{bmatrix} \begin{bmatrix} s_1[k] \\ s_2[k] \\ \vdots \\ s_D[k] \end{bmatrix} + \boldsymbol{\eta}[k] = \boldsymbol{A}\boldsymbol{s}[k] + \boldsymbol{\eta}[k] \tag{2.23}$$

where $\boldsymbol{\eta}$ is the noise vector, that could originate from for instance RF components and the environment [13]. The noise is often assumed to be complex zero-mean Gaussian and spatially uncorrelated, meaning that the noise at each antenna is independent of each other [14]. A useful metric in array signal processing is the signal to noise ratio (SNR) which can be defined as:

$$\text{SNR} = \frac{|s|^2}{\sigma^2} \tag{2.24}$$

where $|s|$ is the peak signal amplitude and σ^2 is the variance of the noise.

The signal model above assumes that there is perfect isolation between the receiving antennas, which is a noteworthy simplification. In reality, each antenna produces an electromagnetic field as a result of the impinging waves, and this field in turn affects the phase and amplitude of the other antennas. This effect is called mutual coupling, and similarly affect the transmitting antennas. In many cases, mutual coupling is difficult to model and instead requires simulations to be characterized. Qualitatively, mutual coupling can be reduced by increasing the distance between the antenna elements [9]. However, there is a limitation in how much the spacing between elements can be increased, as this ultimately negatively affects the DoA estimation abilities, further described in section 2.8.3. Mutual coupling also affects the noise and introduce some spatial correlation.

2.8 Direction of arrival algorithms

Direction of arrival algorithms are used to determine the direction to one or several sources based on a signal $\boldsymbol{x}$, which can be acquired from a specific range-Doppler bin [15]. For this estimation, the array correlation matrix $\boldsymbol{R}_{xx}$ is often used. This can be defined as [5]:

$$\boldsymbol{R}_{xx} = \text{E}[\boldsymbol{x}\boldsymbol{x}^{\text{H}}] = \text{E}[(\boldsymbol{A}\boldsymbol{s} + \boldsymbol{\eta})(\boldsymbol{s}^{\text{H}}\boldsymbol{A}^{\text{H}} + \boldsymbol{\eta}^{\text{H}})] = \boldsymbol{A}\boldsymbol{R}_{ss}\boldsymbol{A}^{\text{H}} + \boldsymbol{R}_{\eta\eta} \tag{2.25}$$

where $\boldsymbol{R}_{ss} = \text{E}[\boldsymbol{s}\boldsymbol{s}^{\text{H}}]$ and $\boldsymbol{R}_{nn} = \text{E}[\boldsymbol{\eta}\boldsymbol{\eta}^{\text{H}}]$ are the source correlation and noise correlation matrices, respectively. These can not be calculated without a priori

knowledge of the source or the noise, and in real world scenarios the array correlation matrix has to be approximated as a time average of the signal $\boldsymbol{x}$:

$$\boldsymbol{R}_{xx} \approx \frac{1}{K}\sum_{k=1}^{K} \boldsymbol{x}[k]\boldsymbol{x}^{\mathrm{H}}[k] \tag{2.26}$$

Some common algorithms used for DoA estimation are Classical Beamformer, Capon, MUSIC and ESPRIT which are described in [5]. Shorter descriptions of two popular ones, Classical Beamformer and MUSIC, are given below.

2.8.1 Classical Beamformer

The Classical Beamformer (CBF) method, also known as the Bartlett method, is a simple DoA method given by:

$$P_{\mathrm{CBF}}(\hat{\boldsymbol{r}}_{\mathrm{eval}}) = \left|\boldsymbol{a}^{\mathrm{H}}(\hat{\boldsymbol{r}}_{\mathrm{eval}})\boldsymbol{R}_{xx}\boldsymbol{a}(\hat{\boldsymbol{r}}_{\mathrm{eval}})\right| \tag{2.27}$$

where the peaks of the pseudospectrum $P_{\mathrm{CBF}}(\hat{\boldsymbol{r}})$ correspond to the DoAs. A simple case of CBF is when only a single sample is taken, meaning that the pseudospectrum takes the form:

$$\begin{aligned}
P_{\mathrm{CBF}}(\hat{\boldsymbol{r}}_{\mathrm{eval}}) &= \left|\boldsymbol{a}^{\mathrm{H}}(\hat{\boldsymbol{r}}_{\mathrm{eval}})\boldsymbol{x}\boldsymbol{x}^{\mathrm{H}}\boldsymbol{a}(\hat{\boldsymbol{r}}_{\mathrm{eval}})\right| = \|\boldsymbol{a}^{\mathrm{H}}(\hat{\boldsymbol{r}}_{\mathrm{eval}})\boldsymbol{x}\|^2 \\
&= \|\boldsymbol{a}^{\mathrm{H}}(\hat{\boldsymbol{r}}_{\mathrm{eval}})\left(\boldsymbol{A}\boldsymbol{s}+\boldsymbol{\eta}\right)\|^2 \\
&= \left\| \begin{bmatrix} e^{-\mathrm{i}k\hat{\boldsymbol{r}}_{\mathrm{eval}}\cdot\boldsymbol{d}_0} \\ \vdots \\ e^{-\mathrm{i}k\hat{\boldsymbol{r}}_{\mathrm{eval}}\cdot\boldsymbol{d}_{N-1}} \end{bmatrix}^T \left(\sum_{d=1}^{D} \left(s_d \begin{bmatrix} e^{\mathrm{i}k\hat{\boldsymbol{r}}_d\cdot\boldsymbol{d}_0} \\ \vdots \\ e^{\mathrm{i}k\hat{\boldsymbol{r}}_d\cdot\boldsymbol{d}_{N-1}} \end{bmatrix} \right) + \begin{bmatrix} \eta_0 \\ \vdots \\ \eta_{N-1} \end{bmatrix} \right) \right\|^2 \\
&= \left| \sum_{n=0}^{N-1} e^{-\mathrm{i}k\hat{\boldsymbol{r}}_{\mathrm{eval}}\cdot\boldsymbol{d}_n} \left(\sum_{d=1}^{D} \left(s_d e^{\mathrm{i}k\hat{\boldsymbol{r}}_d\cdot\boldsymbol{d}_n} \right) + \eta_n \right) \right|^2
\end{aligned} \tag{2.28}$$

A benefit of the CBF is that this form resembles a DFT:

$$\hat{x}_l = \sum_{n=0}^{N-1} x_n e^{-\mathrm{i}2\pi\frac{l}{N}n} \tag{2.29}$$

In order for the DFT to be applicable for CBF, and have only non-zero values x_n, the antenna array has to fill a cartesian equidistant grid. A simple such array is an equidistant $\hat{z}$-directed 1D array. Setting the spacing to d and placing the first antenna in the origin gives the element phases $\varphi_n = k\hat{\boldsymbol{r}}\cdot\boldsymbol{d}_n = kdn\cos\theta$. A common choice for the spacing is $d = \lambda/2$, as this fulfills the Nyquist sampling criteria for all incident angles, making it possible to determine angles $0° < \theta < 180°$ if noise is neglected. The pseudospectrum of such an array with 4 elements is given in Figure 2.9. Using a fast Fourier transform (FFT) to perform the DFT makes the CBF computationally efficient, which is a notable advantage of CBF [5].

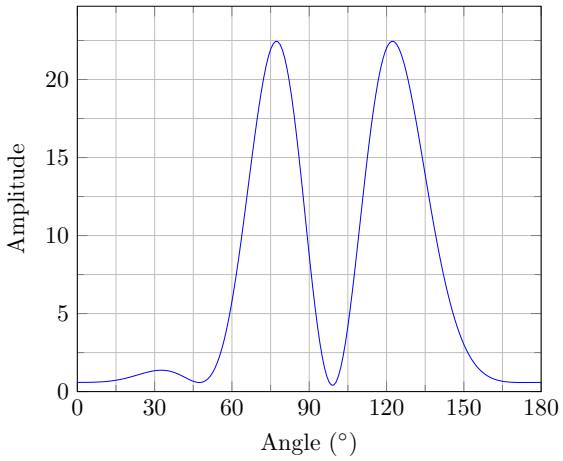


Figure 2.9: Classical Beamformer Estimation using a $\hat{z}$-directed 1D array with 4 elements and with incident directions 75° and 125°, both with equal amplitude.

2.8.2 MUSIC

Another popular DoA method which produces higher resolution pseudospectrums compared to CBF is Multiple Signal Classification (MUSIC). This method uses the eigenstructure of the array correlation matrix $\boldsymbol{R}_{xx}$, meaning that an eigenvalue decomposition has to be performed on $\boldsymbol{R}_{xx}$. This decomposition makes MUSIC more computationally intensive than the simpler CBF. Furthermore, the number of sources D must be known in advance or the eigenvalues have to be searched to determine this [5].

Assuming D sources impinging on an array with N elements, with $D < N$, $\boldsymbol{R}_{xx}$ has D signal eigenvectors and $N - D$ noise eigenvectors. In the MUSIC algorithm, the corresponding eigenvalues are sorted from smallest to largest and the $N - D$ smallest $(\boldsymbol{e}_1, \ldots, \boldsymbol{e}_{N-D})$ are assumed to correspond with the noise eigenvectors. These eigenvectors build up the noise subspace matrix $\boldsymbol{E}_{\text{noise}}$:

$$\boldsymbol{E}_{\text{noise}} = \begin{bmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \ldots & \boldsymbol{e}_{N-D} \end{bmatrix} \tag{2.30}$$

Ideally, the noise eigenvectors are orthogonal to the signal eigenvectors at the directions of arrival. This is utilized in the definition of the pseudospectrum [5]:

$$P_{\text{MUSIC}}(\hat{\boldsymbol{r}}_{\text{eval}}) = \frac{1}{\left| \boldsymbol{a}^{\text{H}}(\hat{\boldsymbol{r}}_{\text{eval}}) \boldsymbol{E}_{\text{noise}} \boldsymbol{E}_{\text{noise}}^{\text{H}} \boldsymbol{a}(\hat{\boldsymbol{r}}_{\text{eval}}) \right|} \tag{2.31}$$

The pseudospectrum for the same scenario as in Figure 2.9 is given in Figure 2.10.

2.8.3 Ambiguity function

A valuable metric in the design of an antenna array is the ambiguity function, which is related to the certainty and resolution of DoA estimation. The ambiguity

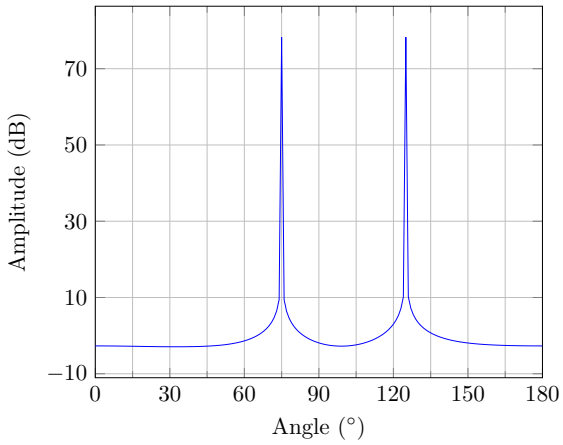


Figure 2.10: MUSIC estimation using a $\hat{z}$-directed 1D array with 4 elements and with incident directions $75°$ and $125°$, both with equal amplitude. The SNR is 10 dB.

function gives the normalized correlation between array vectors corresponding to two incident directions $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{r}}_{\text{eval}}$. It is defined as [12]:

$$\chi(\hat{\boldsymbol{r}}, \hat{\boldsymbol{r}}_{\text{eval}}) = \frac{\left|\boldsymbol{a}(\hat{\boldsymbol{r}})^{\text{H}} \boldsymbol{a}(\hat{\boldsymbol{r}}_{\text{eval}})\right|}{\|\boldsymbol{a}(\hat{\boldsymbol{r}})\| \, \|\boldsymbol{a}(\hat{\boldsymbol{r}}_{\text{eval}})\|} \tag{2.32}$$

There is an obvious similarity between the ambiguity function and the pseudospectrum of the CBF as given in (2.27), meaning that the ambiguity function gives a direct metric on the noiseless performance of the CBF. The ambiguity function always has a maximum of 1 when $\hat{\boldsymbol{r}} = \hat{\boldsymbol{r}}_{\text{eval}}$ and should preferably have low values elsewhere. Two important metrics for the ambiguity function are its sidelobe levels (SLL) and its main lobe width (MLW). The main lobe in the ambiguity function is the values surrounding the peak value, while the side lobes are all the local maximas excluding the main peak. The risk of ambiguities in DoA estimates is determined by the height of the SLLs, and they should therefore be as low as possible to minimize the risk of faulty DoA estimates. It is possible that $\chi(\hat{\boldsymbol{r}}, \hat{\boldsymbol{r}}_{\text{eval}}) = 1$ for several directions $\hat{\boldsymbol{r}}_{\text{eval}}$ if the distances between array elements are too large. In this case, it is impossible to determine the DoA with certainty. At low SNRs, the DoA performance of an array is typically limited by the SLLs [16]. When the SNR instead is high, the performance is largely determined by the shape of the main lobe, which gives an indication of the resolution of the DoA estimations [17]. When multiple signals are received with different DoAs, it is beneficial to have a small MLW as this makes it more likely that the signals can be differentiated. What SNR levels that are considered high or low is however difficult to determine, and is dependent on the specific array parameters.

As first mentioned in section 2.8.1, a common choice of antenna array is an equidistant 1D array with half-a-wavelength distance between elements. This is a type of uniform linear array (ULA). Figure 2.11 shows the ambiguity function

of such an ULA with 4 elements, with an incident angle $\theta = 90°$. Some important definitions are also included in the figure, and the width of the main lobe is measured at half of the peak value, corresponding to -3 dB of peak value, if the spectrum is interpreted as given in power.

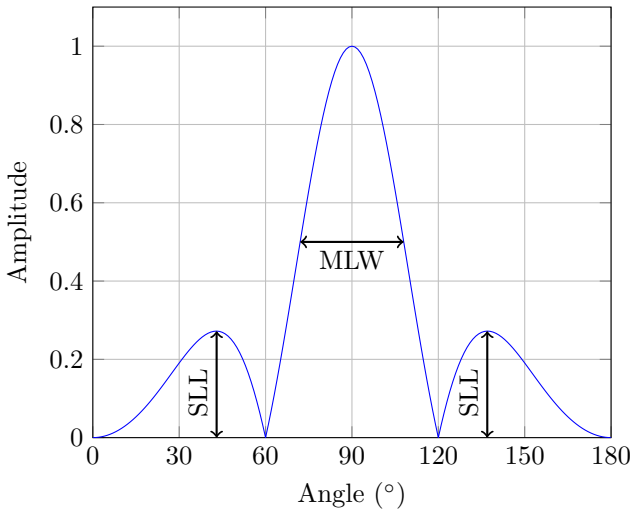


Figure 2.11: Ambiguity function of a $\hat{z}$-directed half-a-wavelength ULA with 4 elements. The incident angle is $\theta = 90°$.

In the case of a multidimensional array, the ambiguity function has to be evaluated in both θ and ϕ to assess the full DoA performance of the array. As an example of this, the layout of a 2D array, together with its ambiguity function is given in Figure 2.12.

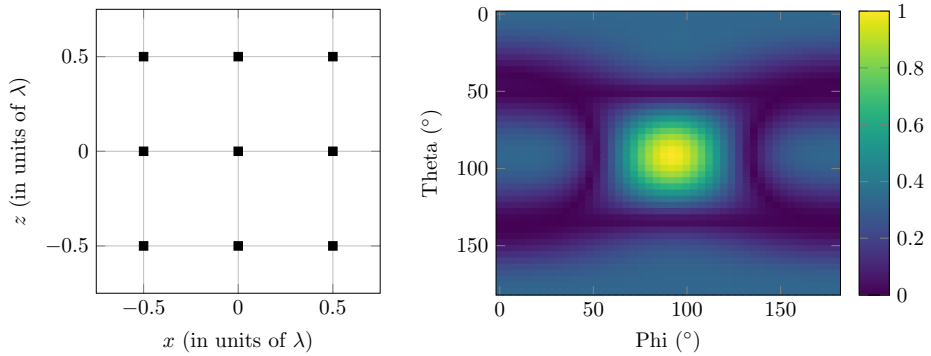


Figure 2.12: The left figure shows the layout of a 2D 9 element array and the right figure shows its resulting ambiguity function for an incident angle $\theta = 90°$, $\phi = 90°$.

2.8.4 Cramér-Rao Bound

One common tool in the assessment of antenna arrays for DoA estimation is the Cramér-Rao Lower Bound (CRLB). The CRLB gives the lower bound of variance for an unbiased estimator, with unbiased meaning that the expected value of the estimator is equal to the real value of the estimated parameter. In the case of DoA estimation, this variance is calculated from the array vector and noise characteristics, and is given in degrees^2 or radians^2 [13]. It is mathematically intricate to calculate the CRLB for arbitrary arrays, especially for non-isotropic receivers, and previous research has mainly focused on derivations for specific geometries, gain profiles and noise assumptions [14, 18]. It is worth noting that the CRLB does not consider ambiguities in the DoA estimation, and it is thus only related to the shape of the main lobe in the ambiguity function.

2.9 Optimization algorithms for antenna arrays

In many cases, the problem of finding an optimal antenna layout for DoA estimation becomes extensive. If the number of antennas and/or size of the bounding geometry is large, it is impractical or even impossible to test all possible placements. In these cases, an optimization algorithm has to be used. Due to the often complex nature of the optimization problem, and the risk of finding local minimas, different stochastic methods have been used in the optimization of antenna arrays. To the author's knowledge, the two most used classes of algorithms for optimizing antenna arrays are genetic algorithms (GA) and particle swarm optimization (PSO) algorithms. Especially genetic algorithms are often used in beamforming problems for antenna arrays.

2.9.1 Genetic algorithms

Genetic algorithms are inspired by the evolutionary process in order to optimize an objective function. Specifically, the processes of natural selection, mating and mutation are used. This has proven to be a successful optimization method in many applications, and has specifically been used for antenna array layouts previously [12, 19].

In order to use a genetic algorithm, the solution space has to be discretized. In this way, all potential solutions can be represented by a binary sequence. If the solution space is built up of several elements, such as several antennas, the value of each element is represented by a binary sequence called a gene. In a beamforming problem, the genes can for instance contain the amplitude and phase of each element. In the case of this report, it is instead the positions of the antenna elements. The genes build up what is called a chromosome, which has an associated cost that states its fitness [19].

The first step of a genetic algorithm is to randomly set the values of the genes in a number of different chromosomes. The associated cost of each chromosome is then evaluated, and the chromosomes are ranked from best to worst. A certain number of the worst chromosomes are discarded (for instance the worst half) and are replaced by new chromosomes. These are produced by mating of the genes

from the saved chromosomes. In this process, there is a certain probability for mutation, meaning that one or several genes are altered randomly. The cost of the new chromosomes are then calculated, and the process is repeated. The number of iterations can be set in advance, or the algorithm can be set to stop when all chromosomes have a cost below a certain threshold [19].

2.9.2 Particle Swarm Optimization

Particle swarm optimization is inspired by social models that simulate flock behavior of birds and other animals. Like GA, its connection to nature makes it intuitive to understand. PSO has previously been used for optimization of radiation patterns for 1D arrays [20]. Unlike GA, the solution space does not have to be discretized for PSO. Each possible solution is instead represented by a vector, $\boldsymbol{d}$, that define what is called a particle. Thus, there is an infinite number of positions in the solution space [21].

The first step of PSO is to randomly initialize a number of particles, and each particle is also given a random velocity within the solution space. The cost of the particles are evaluated, and they are moved according to their velocities. After the associated cost of the particles have been calculated, the next velocity of each particle, $\boldsymbol{v}_n$, is calculated from the last velocity, $\boldsymbol{v}_{n-1}$, and position, $\boldsymbol{d}_{n-1}$, according to the following formula:

$$\begin{aligned}\boldsymbol{v}_n = w \cdot \boldsymbol{v}_{n-1} + b \cdot r_1 \cdot (\boldsymbol{d}_{\text{local best}} - \boldsymbol{d}_{n-1}) \\ + r_2 \cdot c \cdot (\boldsymbol{d}_{\text{global best}} - \boldsymbol{d}_{n-1})\end{aligned} \tag{2.33}$$

where $\boldsymbol{d}_{\text{local best}}$ is the best position achieved by the current particle and $\boldsymbol{d}_{\text{global best}}$ is the best position achieved by any of the particles. r_1 and r_2 are random variables uniformly distributed between 0 and 1. w is an inertia factor while a and b are learning factors. In the original paper introducing PSO, these factors were set to $w = 1$, $a = 2$, $b = 2$. After the velocity $\boldsymbol{v}_n$ has been calculated, the positions are updated as $\boldsymbol{d}_n = \boldsymbol{d}_{n-1} + \boldsymbol{v}_n$ [21].

PSO has been compared to GA in for instance [20] where it was seen that they offered comparable performance for a phased array optimization problem. It was also noted that the PSO is easier to implement, as the positional updates are less complicated to calculate compared to the mutation and mating processes of the GA.

Chapter 3
Previous work

The problem of assessing and optimizing an array for DoA estimation has been studied extensively, for instance in [12, 13, 17, 22–24]. The approaches to the problem differ between papers, and some are focused on specific geometries. The approaches and results of these papers are summarized below.

In [13], Jackson et al. derived the CRLB for uniform circular arrays, meaning arrays where the elements are evenly distributed on a circle. Furthermore, they also used analytical models for patch antenna gain profiles and analyzed how the directivity of the antennas and the radius of the circle affected the CRLB. They showed that the CRLB was improved by using directive instead of isotropic elements.

In [23], Birinci and Tanık optimized a 2D array based on merits derived from the CRLB and the ambiguity function. In the paper, it is shown that for evaluated directions, $\hat{\boldsymbol{r}}_{\text{eval}}$, that are close to the "real" direction, $\hat{\boldsymbol{r}}$, the value of the ambiguity function is directly related to the CRLB. Thus, only the ambiguity function was used. A genetic algorithm was used for the optimization, however the discretization of the geometry as well as the mating and mutation processes were not described in any detail. The performance of the array was compared to a conventional uniform circular array, and showed better DoA estimation abilities for a range of SNRs.

In [17] and [22], a similar approach was used. An analysis using the ambiguity function in conjunction with the CRLB was made. In contrast to Birinci and Tanık, the CRLB was explicitly used in the optimization by calculating a simple expression based on the positions of the antenna elements. In [17], this was done for a 1D array and in [22] extended to 2D. In both of these papers, a special region of interest (RoI) is defined, which is the angular range in which the array should be able to perform DoA estimation without ambiguity. Thus, only directions $\hat{\boldsymbol{r}}$ that were inside the RoI were evaluated in the ambiguity function. As with [23], the details of the optimization algorithm were limited.

In [12], Di Serio et al. extended the optimization to also handle 2D MIMO arrays. In the introduction of the article, the approach of [17, 22] is discussed and some potential improvements are found. It is noted that the interpretation of the RoI given in these papers neglects interference, meaning that if a wave impinges from outside the RoI, it can still cause sidelobes within this region. Because of this, Di Serio et al. evaluated all directions $\hat{\boldsymbol{r}}$ in the ambiguity function. As with all the papers above, a genetic algorithm was used but not described in detail. The CRLB was not used in this paper, and instead the resolution of the array was

evaluated by using the following approximate formulas:

$$\Delta\phi = 1.22\frac{\lambda}{d_x}, \quad \Delta\theta = 1.22\frac{\lambda}{d_z} \tag{3.1}$$

where $\Delta\phi$ and $\Delta\theta$ are the resolutions in ϕ and θ, and d_x and d_z are the maximum extents of the virtual array in x and z direction. The resolution of the array is not included in the cost function of the optimization algorithm, but is evaluated in the end.

To the author's knowledge, the particle swarm optimization method has only been used for array layout optimization in [24]. Here, Chen et al. analyzed 1D arrays, and defined the cost to be the peak sidelobe level (PSLL) in the DoA estimation. The analysis is limited compared to the previously mentioned papers, and it is unclear if more than one incident direction has been evaluated.

Chapter 4
Method

The general outline of the work can be summarized in 7 steps as given in the list below.

1. Implementation of Python code used for calculating the array response for an arbitrary MIMO array, see section 4.1.
2. Definition of a cost function for the DoA estimation capability of an array, see section 4.2.
3. Implementation of optimization algorithm, see section 4.3.
4. Design of a patch antenna used for optimization algorithm, see section 4.4.
5. Choice of array parameters for optimization, see section 4.5.
6. Simulation of the optimized arrays, see section 4.6.
7. Assessment of noise performance of the optimized arrays, see section 4.7.

4.1 Array response calculations

In the Python code, the array vector was calculated using the formulation of the virtual array vector as described in section 2.6. Any effects of mutual coupling between elements were ignored, and it was assumed that all elements had identical gain profiles. For the circular geometry, all elements were placed in a plane and with the same orientation. Thus, the received amplitude was modelled to be identical between elements for a specific DoA. This was not the case for the multiple plane geometry, as the sides point in different directions. In order to calculate the signal amplitude for a specific DoA, both the θ and ϕ components of the gain profile had to be used. Assuming that the n:th Tx antenna transmitted and the m:th Rx antenna received, the total voltage gain (ignoring phase) was calculated as:

$$g(\hat{\boldsymbol{r}}) = \sqrt{G_{\theta,\mathrm{Tx},n}(\hat{\boldsymbol{r}})G_{\theta,\mathrm{Rx},m}(\hat{\boldsymbol{r}}) + G_{\phi,\mathrm{Tx},n}(\hat{\boldsymbol{r}})G_{\phi,\mathrm{Rx},m}(\hat{\boldsymbol{r}})} \tag{4.1}$$

where $G_{\theta,\mathrm{Tx},n}(\hat{\boldsymbol{r}})$ and $G_{\phi,\mathrm{Tx},n}(\hat{\boldsymbol{r}})$ are the θ and ϕ components of the n:th Tx, respectively, and where $G_{\theta,\mathrm{Rx},m}(\hat{\boldsymbol{r}})$ and $G_{\phi,\mathrm{Rx},m}(\hat{\boldsymbol{r}})$ are the θ and ϕ components of the m:th Rx. The gain profile used for the amplitude calculations were taken from a simulated patch antenna, described in section 4.4.

4.2 Assessment of array performance

Two ways of assessing the DoA estimation performance of an array are described in section 2.8.3 and 2.8.4. The Cramér-Rao bound should be used if the goal is to maximize the theoretical DoA estimation capability of an array for a given SNR. However, the CRLB does not give any guarantee that the lower bound can be reached using a practical DoA method. In this work, only the ambiguity function is used as it considers ambiguities as well as DoA resolution through the MLW. The two geometries, as given in Figure 1.1, are to a large extent handled in the same way, with a few exceptions. Below, the assessment of arrays in the circular geometry is described first followed by the multiple plane geometry.

4.2.1 Circular geometry

To assess the full performance of an array, all combinations of $\hat{\boldsymbol{r}}$ and $\hat{\boldsymbol{r}}_{\text{eval}}$ have to be evaluated in the ambiguity function. In order to do this numerically, the vectors have to be discretized into $N \times M$ directions $\hat{\boldsymbol{r}}(\theta_i, \phi_j)$ and $\hat{\boldsymbol{r}}_{\text{eval}}(\theta_k, \phi_l)$, where $i, k = 1, ..., N$ and $j, l = 1, ..., M$. This means that the total number of combinations of $\hat{\boldsymbol{r}}(\theta_i, \phi_j)$ and $\hat{\boldsymbol{r}}_{\text{eval}}(\theta_k, \phi_l)$ that have to be evaluated in the ambiguity function is $N^2 \times M^2$. If the array is placed in the xz-plane and radiates in the half-space where y is positive, θ_i, ϕ_j, θ_k and ϕ_l can be uniformly distributed in the interval $0 \leq \theta_i, \phi_j, \theta_k, \phi_l \leq 180°$. It was determined that setting $N = M = 46$ offered reasonable computational times and angular resolution of 4°.

Following the same notation as in [12], the ambiguity matrix $\boldsymbol{X}_{ij}$ is defined as:

$$\boldsymbol{X}_{ij} = \begin{bmatrix} \chi(\hat{\boldsymbol{r}}(\theta_i, \phi_j), \hat{\boldsymbol{r}}_{\text{eval}}(\theta_1, \phi_1)) & \dots & \chi(\hat{\boldsymbol{r}}(\theta_i, \phi_j), \hat{\boldsymbol{r}}_{\text{eval}}(\theta_1, \phi_M)) \\ \vdots & \ddots & \vdots \\ \chi(\hat{\boldsymbol{r}}(\theta_i, \phi_j), \hat{\boldsymbol{r}}_{\text{eval}}(\theta_N, \phi_1)) & \dots & \chi(\hat{\boldsymbol{r}}(\theta_i, \phi_j), \hat{\boldsymbol{r}}_{\text{eval}}(\theta_N, \phi_M)) \end{bmatrix} \tag{4.2}$$

Each $\boldsymbol{X}_{ij}$ has a main lobe and (in most cases) sidelobes, as illustrated in Figure 2.12. As the highest SLL (denoted PSLL) determines the risk of ambiguity in DoA estimation, it is of special interest. The maximum half-width of the main lobe (denoted $\text{MLW}_{\max}$) is similarly important for the resolution of the DoA estimation. Therefore, two matrices are introduced to evaluate these two qualities. The first is the ambiguity indicator matrix, $\boldsymbol{Q}$, inspired by [12]:

$$\boldsymbol{Q} = [Q_{ij}], \quad Q_{ij} = \begin{cases} 0, & \text{PSLL}(\boldsymbol{X}_{ij}) \leq t \\ \text{PSLL}(\boldsymbol{X}_{ij}), & \text{PSLL}(\boldsymbol{X}_{ij}) > t \end{cases} \tag{4.3}$$

where t is the threshold value for the maximum SLL. It is chosen as $t = 0.5$ in the optimization algorithm as this is slightly better than the maximum PSLL of the reference array, which is described in section 4.5. Similarly, a width matrix $\boldsymbol{W}$ is defined as:

$$\boldsymbol{W} = [W_{ij}], \quad W_{ij} = \text{MLW}_{\max}(\boldsymbol{X}_{ij}) \tag{4.4}$$

where $\text{MLW}_{\max}(\boldsymbol{X}_{ij})$ is the largest possible angle, Δ, between $\hat{\boldsymbol{r}}(\theta_i, \phi_j)$ and a vector $\hat{\boldsymbol{r}}_w(\theta_k, \phi_l)$ that is part of the main lobe such that $\chi(\hat{\boldsymbol{r}}(\theta_i, \phi_j), \hat{\boldsymbol{r}}_w(\theta_k, \phi_l)) \geq$

w. The angle is calculated using $\Delta = \cos^{-1}(\hat{\boldsymbol{r}} \cdot \hat{\boldsymbol{r}}_w)$. Here, w defines at what amplitude the main lobe width should be measured, and it was chosen as $w = 0.5$. This way of assessing the DoA resolution is more complicated than what is presented in (3.1), but is also more complete as it considers the resolution in all possible directions instead of θ and ϕ separately. The retrieval of the values Q_{ij} and W_{ij} are illustrated in Figure 4.1 for a specific ambiguity matrix.

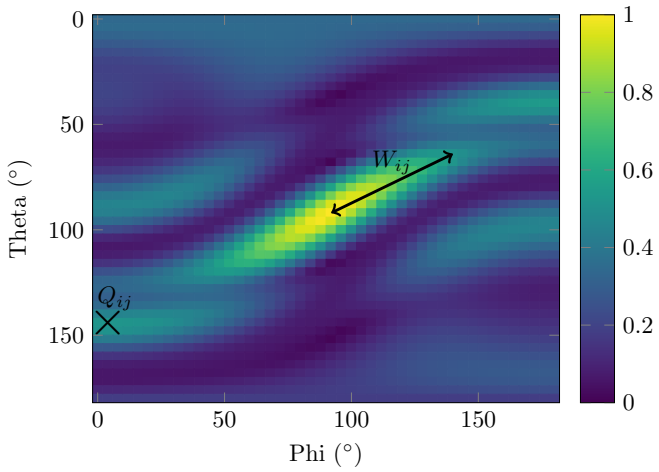


Figure 4.1: Illustration of the values Q_{ij} (4.3) and W_{ij} (4.4) from the ambiguity matrix.

In some cases, it might not be necessary to optimize an array for all possible DoAs as a radar system often operates within an angular range of less than 180° in both elevation and azimuth. While the useful range of a radar system could take many forms, it is in this work given as $\theta_{\min} \leq \theta \leq \theta_{\max}$ and $\phi_{\min} \leq \phi \leq \phi_{\max}$ for simplicity. Furthermore, it is denoted the region of interest (RoI). By defining such a RoI, the ambiguity indicator matrix and the width matrix were evaluated slightly differently from (4.3) and (4.4). It is important to note that reflected waves that arrive from outside the RoI can give rise to sidelobes within this region, and result in false detections. Thus, the ambiguity matrix $\boldsymbol{X}_{ij}$ can not simply be evaluated for $\theta_{\min} \leq \theta_k \leq \theta_{\max}$, $\phi_{\min} \leq \phi_l \leq \phi_{\max}$. Instead, if the real direction $\hat{\boldsymbol{r}}$ lies within the given range, the whole ambiguity matrix has to be evaluated. However, if $\hat{\boldsymbol{r}}$ lies outside the range, only values $\theta_{\min} \leq \theta_k \leq \theta_{\max}$ and $\phi_{\min} \leq \phi_l \leq \phi_{\max}$ are evaluated. The width matrix is only evaluated within the specified range. The ranges used were $45° \leq \theta, \phi \leq 135°$ for the circular geometry, as this covers a reasonable angular range when compared to the coverage of surveillance cameras.

From the definition of $\boldsymbol{Q}$ and $\boldsymbol{W}$, a cost function C was defined. As it is beneficial to have low SLLs and low MLWs, the average SLL and MLW was included in the cost. In order to ensure that the performance was relatively uniform over the angular range, the worst-case values were also specifically included. The total

cost function was thus given as:

$$C(\boldsymbol{Q}, \boldsymbol{W}, c_1, c_2, c_3, c_4) = \frac{c_1}{NM} \sum_{i,j} Q_{ij} + c_2 \max(\boldsymbol{Q}) + \frac{c_3}{NM} \sum_{i,j} W_{ij} + c_4 \max(\boldsymbol{W}) \tag{4.5}$$

where $\max(\boldsymbol{A})$ denotes the maximum value contained in a matrix $\boldsymbol{A}$. The main objective of the optimization was to reduce all SLLs below the threshold t, while the secondary goal was to minimize the MLWs. Thus, the parameters were chosen such that $c_1 > c_3$ and $c_2 > c_4$. The values used were $c_1 = 1$, $c_2 = 0.4$, $c_3 = 0.2$ and $c_4 = 0.05$, as they were seen to give reasonable results.

4.2.2 Multiple plane geometry

For the multiple plane geometry (illustrated in Figure 1.1b) all sides were used for the DoA estimation. As the total number of antennas is large, it becomes time-consuming to use each Tx and Rx for a single DoA estimation. Thus, the array was simulated such that only Tx antennas on one of the sides transmitted, while all Rx antennas received. This resulted in a virtual array of size $N_s \times n_{\text{Tx}} \times n_{\text{Rx}}$, where N_s is the number of sides, and n_{Tx} and n_{Rx} is the number of Tx and Rx elements per side, respectively. One factor that needed special care in the optimization of the array was that the distance between the four different arrays could be large in terms of wavelengths. This results in many ambiguities if the total array vector for all elements is used in the ambiguity function (2.32). To solve this, a DoA estimation was performed for each individual side instead. As the sides point in different directions, thus giving different gain profiles, the amplitude information could be used together with the phase information. A signal model of only the output amplitudes of the n:th side, $\boldsymbol{A}_n$, is introduced as:

$$\boldsymbol{A}_n[k] = \begin{bmatrix} \boldsymbol{g}_n(\hat{\boldsymbol{r}}_1) & \boldsymbol{g}_n(\hat{\boldsymbol{r}}_2) & \dots & \boldsymbol{g}_n(\hat{\boldsymbol{r}}_D) \end{bmatrix} \begin{bmatrix} s_1[k] \\ s_2[k] \\ \vdots \\ s_D[k] \end{bmatrix} + \boldsymbol{\eta}[k] \tag{4.6}$$

where $\boldsymbol{g}_n(\hat{\boldsymbol{r}})$ consists of the combined Tx and Rx gains for the receiving elements of the n:th side and s is real valued. It is assumed that the gain profile is identical between elements on the same side and that the gains are voltage gains. Thus, $\boldsymbol{g}_n(\hat{\boldsymbol{r}})$ is defined as:

$$\boldsymbol{g}_n(\hat{\boldsymbol{r}}) = \begin{bmatrix} g_n(\hat{\boldsymbol{r}}) \\ g_n(\hat{\boldsymbol{r}}) \\ \dots \\ g_n(\hat{\boldsymbol{r}}) \end{bmatrix}, \quad g_n(\hat{\boldsymbol{r}}) = \sqrt{G_{\theta,\text{Tx}}(\hat{\boldsymbol{r}}) G_{\theta,\text{Rx},n}(\hat{\boldsymbol{r}}) + G_{\phi,\text{Tx}}(\hat{\boldsymbol{r}}) G_{\phi,\text{Rx},n}(\hat{\boldsymbol{r}})} \tag{4.7}$$

where all the gains are given in the global coordinate system given in Figure 1.1b. The mean over the elements of $\boldsymbol{A}_n[k]$ is denoted $\bar{\boldsymbol{A}}_n[k]$ and can be valuable to

calculate in order to reduce the effects of noise in the signals. From this, the vector $\boldsymbol{A}[k]$ containing the mean amplitude of each side is defined as:

$$\boldsymbol{A}[k] = \begin{bmatrix} \bar{\boldsymbol{A}}_1[k] \\ \bar{\boldsymbol{A}}_2[k] \\ \vdots \\ \bar{\boldsymbol{A}}_{N_s}[k] \end{bmatrix} \tag{4.8}$$

where N_s is the number of sides.

Based on this signal model and the signal model of (2.23), a new DoA method is defined. This DoA estimator uses the gain variations of the different sides to roughly assess in what angular range the DoA is likely to be. With this knowledge, the phase information is used on one or two sides to determine a more precise DoA. Assuming that the transmitting side has side index $n = n_{\mathrm{Tx}}$ and that the side that receives the maximum mean amplitude $\bar{\boldsymbol{A}}$ has index $n = n_{\max}$, a pseudospectrum similar to the CBF is defined as:

$$P_{\mathrm{new}}(\hat{\boldsymbol{r}}_{\mathrm{eval}}) = \frac{\boldsymbol{A}^{\mathrm{T}} R_{AA} \boldsymbol{A}}{\boldsymbol{A}^{\mathrm{T}} \boldsymbol{A}} \times \begin{cases} \left| \boldsymbol{a}_{n_{\mathrm{Tx}}}(\hat{\boldsymbol{r}}_{\mathrm{eval}})^{\mathrm{H}} R_{xx,\, n_{\mathrm{Tx}}} \boldsymbol{a}_{n_{\mathrm{Tx}}}(\hat{\boldsymbol{r}}_{\mathrm{eval}}) \right|, & \text{if } n_{\mathrm{Tx}} = n_{\max} \\ 0.5 \sum\limits_{n=n_{\mathrm{Tx}}, n_{\max}} \left| \boldsymbol{a}_n(\hat{\boldsymbol{r}}_{\mathrm{eval}})^{\mathrm{H}} R_{xx,\, n} \boldsymbol{a}_n(\hat{\boldsymbol{r}}_{\mathrm{eval}}) \right|, & \text{if } n_{\mathrm{Tx}} \neq n_{\max} \end{cases} \tag{4.9}$$

where $\boldsymbol{a}_n$ denotes the array vector of the n:th side and R_{AA} is the correlation matrix of $\boldsymbol{A}$, calculated in the same way as in (2.26). Based on the DoA method given by P_{new}, a corresponding ambiguity function is also defined as:

$$\chi(\hat{\boldsymbol{r}}, \hat{\boldsymbol{r}}_{\mathrm{eval}}) = \frac{\boldsymbol{g}(\hat{\boldsymbol{r}})^{\mathrm{T}} \boldsymbol{g}(\hat{\boldsymbol{r}}_{\mathrm{eval}})}{\|\boldsymbol{g}(\hat{\boldsymbol{r}})\| \, \|\boldsymbol{g}(\hat{\boldsymbol{r}}_{\mathrm{eval}})\|} \times \begin{cases} \dfrac{\left| \boldsymbol{a}_{n_{\mathrm{Tx}}}(\hat{\boldsymbol{r}})^{\mathrm{H}} \boldsymbol{a}_{n_{\mathrm{Tx}}}(\hat{\boldsymbol{r}}_{\mathrm{eval}}) \right|}{\|\boldsymbol{a}_{n_{\mathrm{Tx}}}(\hat{\boldsymbol{r}})\| \, \|\boldsymbol{a}_{n_{\mathrm{Tx}}}(\hat{\boldsymbol{r}}_{\mathrm{eval}})\|}, & \text{if } n_{\mathrm{Tx}} = n_{\max} \\ 0.5 \sum\limits_{n=n_{\mathrm{Tx}}, n_{\max}} \dfrac{\left| \boldsymbol{a}_n(\hat{\boldsymbol{r}})^{\mathrm{H}} \boldsymbol{a}_n(\hat{\boldsymbol{r}}_{\mathrm{eval}}) \right|}{\|\boldsymbol{a}_n(\hat{\boldsymbol{r}})\| \, \|\boldsymbol{a}_n(\hat{\boldsymbol{r}}_{\mathrm{eval}})\|}, & \text{if } n_{\mathrm{Tx}} \neq n_{\max} \end{cases} \tag{4.10}$$

where $\boldsymbol{g}$ is given by:

$$\boldsymbol{g}(\hat{\boldsymbol{r}}) = \begin{bmatrix} g_1(\hat{\boldsymbol{r}}) \\ g_2(\hat{\boldsymbol{r}}) \\ \vdots \\ g_{N_s}(\hat{\boldsymbol{r}}) \end{bmatrix} \tag{4.11}$$

As the multiple plane array can receive signals from a larger angular range compared to the circular array, the ranges were increased to cover the full unit sphere. However, only angles covered by the Tx side were evaluated. These useful angles are plotted in Figure 4.2, where it is assumed that the tilt of the sides is $\alpha = 20°$ and that the Tx side points in the positive y direction.

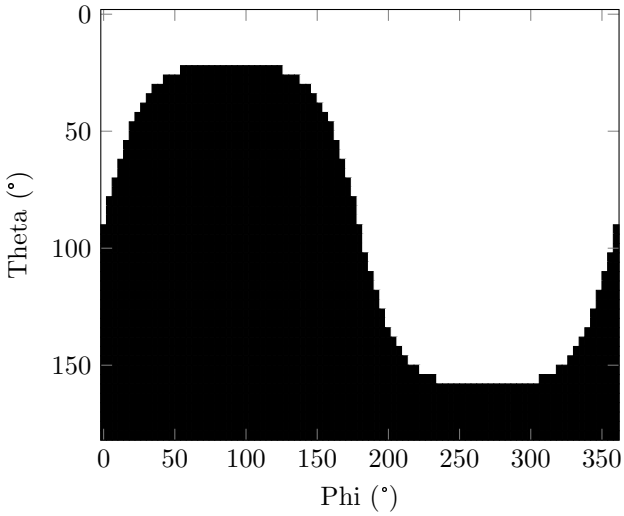


Figure 4.2: The useful angles for the multiple plane layout (see Figure 1.1b), plotted in black.

4.3 Implementation of optimization algorithm

The two types of optimization algorithms considered for this work were GA and PSO, as described in sections 2.9.1 and 2.9.2. While GA is the most common choice for this type of problem, PSO seems to offer simpler implementation. Furthermore, the latter has to the author's knowledge not been used for 2D array layout optimization before. It was therefore deemed interesting to investigate the effectiveness of PSO for this specific problem.

The outline of the optimization method is given below in Algorithm 1, where N_{part} is the number of particles, N_{it} is the number of iterations and w, b, c are the constants used for updating the particle velocities according to (2.33). For the circular geometry, the positions of the Tx and Rx elements in the i:th particle are stored in $\boldsymbol{d}_{i,\text{Tx}}$ and $\boldsymbol{d}_{i,\text{Rx}}$ as:

$$\boldsymbol{d}_{i,\text{Tx}} = \begin{bmatrix} x_{i,\text{Tx},1} & z_{i,\text{Tx},1} \\ x_{i,\text{Tx},2} & z_{i,\text{Tx},2} \\ \dots & \dots \\ x_{i,\text{Tx},N_{\text{Tx}}} & z_{i,\text{Tx},N_{\text{Tx}}} \end{bmatrix}, \quad \boldsymbol{d}_{i,\text{Rx}} = \begin{bmatrix} x_{i,\text{Rx},1} & z_{i,\text{Rx},1} \\ x_{i,\text{Rx},2} & z_{i,\text{Rx},2} \\ \dots & \dots \\ x_{i,\text{Rx},N_{\text{Rx}}} & z_{i,\text{Rx},N_{\text{Rx}}} \end{bmatrix} \tag{4.12}$$

Similarly, the velocities of the Tx and Rx elements in the i:th particle are stored in $\boldsymbol{v}_{i,\text{Tx}}$ and $\boldsymbol{v}_{i,\text{Rx}}$ as:

$$\boldsymbol{v}_{i,\text{Tx}} = \begin{bmatrix} v_{x,i,\text{Tx},1} & v_{z,i,\text{Tx},1} \\ v_{x,i,\text{Tx},2} & v_{z,i,\text{Tx},2} \\ \dots & \dots \\ v_{x,i,\text{Tx},N_{\text{Tx}}} & v_{z,i,\text{Tx},N_{\text{Tx}}} \end{bmatrix}, \quad \boldsymbol{v}_{i,\text{Rx}} = \begin{bmatrix} v_{x,i,\text{Rx},1} & v_{z,i,\text{Rx},1} \\ v_{x,i,\text{Rx},2} & v_{z,i,\text{Rx},2} \\ \dots & \dots \\ v_{x,i,\text{Rx},N_{\text{Rx}}} & v_{z,i,\text{Rx},N_{\text{Rx}}} \end{bmatrix} \tag{4.13}$$

In the case of the multiple plane geometry, x and z are replaced by coordinates in the local coordinate system of a side. These coordinates are called $\tilde{x}$ and $\tilde{z}$, in

order to distinguish them from the global coordinate system given in Figure 1.1b. As the layouts should be identical between sides, it is sufficient to only consider one planar layout in the optimization. The resulting layout is then copied to all sides.

Some extra steps were added to the implemented PSO compared to the general description in section 2.9.2. Firstly, the values of the constants w, b, c are reduced towards the end of the optimization loop. This was added as it was difficult to choose a set of constant parameters that did not give either too high or too low average element velocities. Too high velocities discouraged any fine-tuning of the positions, while too low velocities instead made the algorithm get stuck in local minimas. Thus, two sets of parameters were found that gave suitable randomness in the initial search as well as some fine-tuning during the end of the optimization. Secondly, a maximum allowed speed of individual elements, $v_{\max}$, was introduced to reduce the risk of instability in the element movements. The number of particles was set to $N_{\text{part}} = 20$.

When the element positions are updated in line 36 and 37 of the pseudocode, it is possible that the new positions are invalid. This could be either because at least one element is outside the defined geometry, or it could be because two elements are closer to each other than the minimum allowed spacing of Δd. These cases are handled in line 38 and 39, respectively. In line 38, any element that is outside the boundary is moved to the first intersect point of the line of movement and the boundary. Also, the element bounces on the boundary, meaning that the component of the velocity vector that is orthogonal to the boundary is negated. In line 39, collisions between elements is handled similarly to bounces on the boundary. If an element is placed closer than Δd to another element, the last element to be moved is placed at a distance Δd from the other element. As the elements can collide in several ways, and with multiple other elements, it is difficult to handle all edge cases. Thus, line 38 and 39 do not always succeed in placing the elements in an allowed way, which necessitates the validity check in line 11.

The number of iterations that was used in the optimization was largely limited by the runtime of the algorithm. As to not exceed 24 hour optimization runs on the available hardware, the number of iterations were limited to 500 for the circular geometry. Only 300 iterations were used for the multiple plane geometry, due to the additional computations associated with a larger array.

4.4 Design of antenna elements

The antenna type used in the work is the inset-fed patch antenna as described in section 2.4. Through manual tuning of the dimensions given in Figure 2.5, the patch was designed to have a center frequency at 24.125 GHz with a 250 MHz bandwidth at -10 dB of the reflection coefficient, Γ. The frequency range was chosen as it covers one of the frequency bands that can be used for commercial radar applications [25]. The patch and ground plane were simulated as PEC and the substrate as Rogers RO3003 with a relative permittivity $\epsilon = 3$ and thickness of 0.254 mm. The size of the ground plane was 20 mm $\times$ 20 mm.

Algorithm 1 Particle swarm optimization pseudocode

for $i = 1$ to N_{part} **do**
    randomly set initial positions $\boldsymbol{d}_{i,\text{Tx}}$ and $\boldsymbol{d}_{i,\text{Rx}}$ for *particle*(i)
    calculate corresponding virtual array
    randomly set initial velocities $\boldsymbol{v}_{i,\text{Tx}}$ and $\boldsymbol{v}_{i,\text{Rx}}$ for *particle*(i)
end for

$w \leftarrow 1$, $b \leftarrow 2$, $c \leftarrow 2$
define vector *costs* of size N_{part} with current costs
for $i = 1$ to N_{it} **do**
    for $j = 1$ to N_{part} **do**
        if $\boldsymbol{d}_{j,\text{Tx}}$, $\boldsymbol{d}_{j,\text{Rx}}$ not valid **then**
            costs(j) $\leftarrow \infty$
        else
            costs(j) $\leftarrow$ cost of *particle*(j)
        end if
    end for
    for $j = 1$ to N_{part} **do**
        costs(j) $\leftarrow$ cost of *particle*(j)
        if *costs*(j) $< cost_{\text{local best, j}}$ **then**
            $\boldsymbol{d}_{\text{local best, j, Tx}} \leftarrow \boldsymbol{d}_{j,\text{Tx}}$
            $\boldsymbol{d}_{\text{local best, j, Rx}} \leftarrow \boldsymbol{d}_{j,\text{Rx}}$
            $cost_{\text{local best, j}} \leftarrow$ *costs*(j)
        end if
        if *costs*(j) $< cost_{\text{global best}}$ **then**
            $\boldsymbol{d}_{\text{global best, Tx}} \leftarrow \boldsymbol{d}_{j,\text{Tx}}$
            $\boldsymbol{d}_{\text{global best, Rx}} \leftarrow \boldsymbol{d}_{j,\text{Rx}}$
            $cost_{\text{global best}} \leftarrow$ *costs*(j)
        end if
    end for
    if $i \approx 0.75 N_{\text{it}}$ **then**
        $w \leftarrow 0.5$, $b \leftarrow 1.5$, $c \leftarrow 1.5$
    end if
    for $j = 1$ to N_{part} **do**
        $\boldsymbol{v}_{i,\text{Tx}}$, $\boldsymbol{v}_{i,\text{Rx}} \leftarrow$ new velocities according to (2.33) using w, b, c
        reduce element velocities exceeding $v_{\max}$
        $\boldsymbol{d}_{j,\text{Tx}} \leftarrow \boldsymbol{d}_{j,\text{Tx}} + \boldsymbol{v}_{i,\text{Tx}}$
        $\boldsymbol{d}_{j,\text{Rx}} \leftarrow \boldsymbol{d}_{j,\text{Rx}} + \boldsymbol{v}_{i,\text{Rx}}$
        relocate elements outside of boundary
        handle collisions between elements
        calculate virtual array
    end for
end for
return $\boldsymbol{d}_{\text{global best, Tx}}$, $\boldsymbol{d}_{\text{global best, Rx}}$

4.5 Choice of array parameters

The optimization algorithm was implemented in such a way that it could handle any number of Tx and Rx antennas, as well as arbitrary dimensions of the geometries given in Figure 1.1a and Figure 1.1b. In order to limit the scope of the report, only a few specific choices of parameters were evaluated. For the circular array, 3 Tx and 4 Rx antennas were used as it is a common choice for commercial radar products, meaning that the optimized array could be compared to existing layouts. A conventional array layout using 3 Tx and 4 Rx antennas is given in Figure 4.3 and 4.4. The dimensions of the circular array as given in Figure 1.1a were set to $R_{\text{min}} = 15\,\text{mm}$ and $R_{\text{max}} = 30\,\text{mm}$. The minimum distance between antenna elements was set to $\Delta d = \lambda/2$ as this is also the closest distance in the reference array. Thus, mutual coupling effects should be manageable.

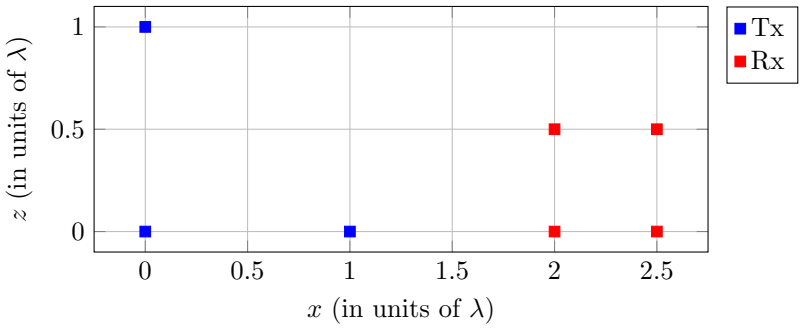


Figure 4.3: Rx and Tx positions of reference ULA.

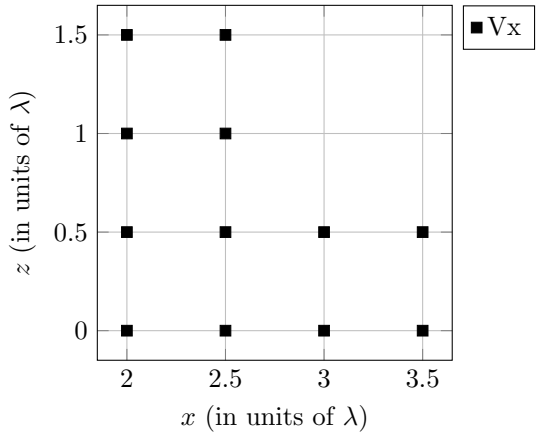


Figure 4.4: Virtual array of reference ULA.

A setup with four sides was chosen for the multiple plane array, where each side consisted of 3 Tx and 4 Rx antennas. As mentioned earlier, each side was set to have the same array layout. The dimensions of each side was set to 30 mm ×

40 mm, with a distance of 80 mm from the z-axis to the center point of the sides and a tilt of $\alpha = 20°$. The ULA layout given in Figure 4.3 was used as a reference layout.

4.6 Simulation of arrays

The optimized arrays were simulated in HFSS to ensure that the array vectors for different DoAs corresponded to those of the analytical model. This was done by simulating the responses from the Rx antennas using a PEC sphere with a radius of 50 mm as a target, placed with a radial distance of 10 m from the arrays. The dimensions of the substrate and ground plane had to be increased slightly in order to ensure that the antennas did not extend beyond the boundaries. Specifically, $R_{\min}$ was decreased to 10 mm and $R_{\max}$ was increased to 35 mm. Furthermore, all sides in the multiple plane geometry were widened to 40 mm $\times$ 50 mm. As the distance from the array to the target is very large in terms of wavelength, a full wave simulation of the setup was unviable. Thus, the far field characteristics of only the array itself was simulated separately and used in a shooting and bouncing rays (SBR) simulator together with the target. The SBR solver in HFSS utilizes that the target is in the far field of the array, and thus the waves are treated as optical rays. This means that the angle of reflection is the same as the angle of incidence onto the reflector. By performing the simulation, mutual coupling effects as well as effects due to element placement on the geometry are taken into account.

A special FMCW module in the SBR solver was used in the simulation and resulted in a range-Doppler map for each Tx and Rx pair. From these maps, only the bins corresponding to a Doppler velocity of 0 m/s and a range of 10 m were extracted. For each position of the sphere, a response vector $\boldsymbol{x}$ (as given in section 2.7) was thus gathered. In the case of the circular geometry, a CBF DoA estimation was performed using (2.27), where $\hat{\boldsymbol{r}}_{\text{eval}}$ was set as the analytically calculated array vector. As there was only one "sample" per position of the sphere, the covariance matrix was reduced to $\boldsymbol{x}\boldsymbol{x}^H$. The deviation of the DoA estimation from the actual DoA was then calculated. Similarly, the DoA method defined in (4.9) was used for the multiple plane geometry. Due to the substantial amount of time needed for the simulations to run, the angular spacing of simulated DoAs could not be as small as for the optimization. Thus, the angular spacing was set to 8° for both θ and ϕ. Time constraints also limited the analysis to DoAs within the RoI, and the effects of interfering signals was thus not evaluated.

4.7 Noise performance of arrays

The final step in the evaluation was to investigate the performance of the arrays with added noise $\boldsymbol{\eta}$, as given in the signal models (2.23) and (4.6). This was done by generating complex zero-mean Gaussian noise with variance given by (2.24), and by seeing how this affected the average angular estimation deviation in the RoI. The CBF was once again used for the DoA estimation with only a single sample (i.e. $K = 1$). The deviation was evaluated using both the analytical and simulated array responses. Furthermore, the deviation was calculated from 500

instances of signal noise for each DoA, and an average deviation for each DoA was calculated from this. In the case of the multiple plane geometry, each side has its own SNR as the received signal amplitudes are different. The signal amplitude to calculate the SNR from can thus be chosen in different ways. For the results, it was defined as the signal amplitude of the received signal on the transmitting side. Furthermore, the noise variance was calculated for each incident angle to match the SNR, instead of adding a constant noise level based on one incident angle. The noise performance was evaluated using both the analytical and simulated array responses.

Chapter 5
Results

5.1 Patch antenna

The dimensions of the patch antenna are given in Table 5.1 and its resulting reflection coefficient is plotted in Figure 5.1. The far field radiation pattern of the antenna is given in Figure 5.2 and 5.3, using the spherical coordinate system defined in Figure 2.3. Please note that the dB scale is different between plots. The antenna is mainly polarized in the θ direction, as expected, and its gain profiles are symmetrical in all plots except for the ϕ component in the yz plane. This could potentially be caused by the feed which introduces an asymmetry in the patch antenna.

Parameter	L	W	f_w	i_w	i_l
Value (mm)	3.43	4.42	0.44	0.31	1.1

Table 5.1: Dimensions of the patch antenna using the parameters defined in Figure 2.5.

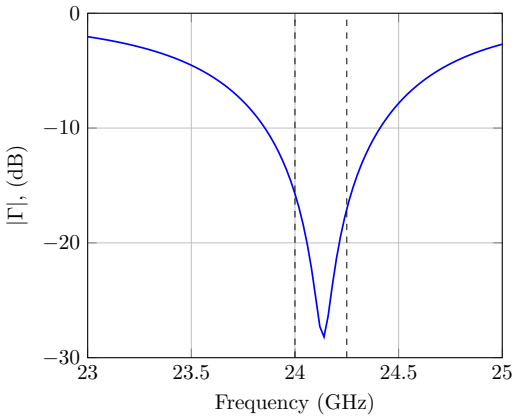
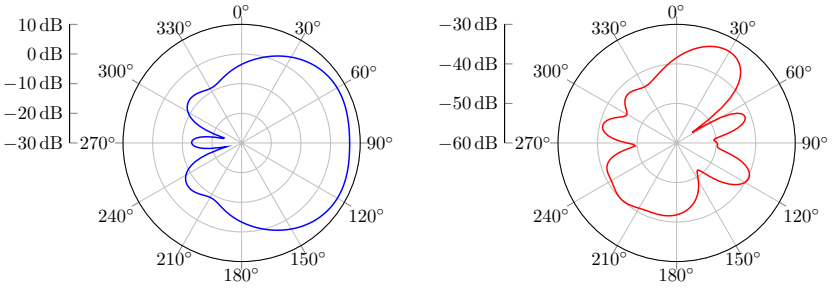
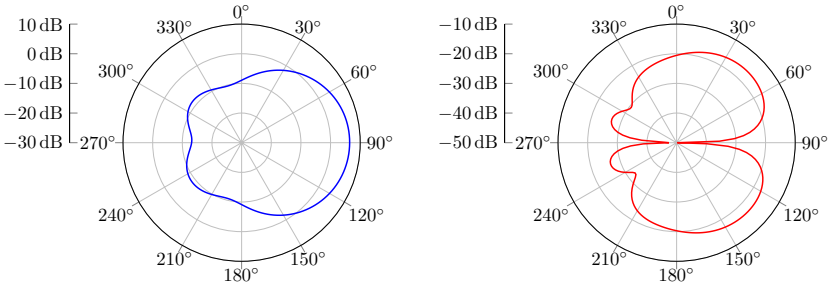


Figure 5.1: Reflection coefficient of the patch antenna. The frequency range of interest is marked by the dashed lines.



(a) θ component. **(b)** ϕ component.

Figure 5.2: Gain in yz plane, swept over θ.



(a) θ component. **(b)** ϕ component.

Figure 5.3: Gain in xy plane, swept over ϕ.

5.2 Optimization results

5.2.1 Circular layout

The element positions of the optimized array are given in Figure 5.4, and the corresponding virtual array in Figure 5.5. The optimization algorithm managed to suppress the relevant SLLs so that $\boldsymbol{Q} = \mathbf{0}_{46\times46}$, compared to $\boldsymbol{Q}$ of the reference array which is given in Figure 5.6. Figure 5.7 and 5.8 show the evolution of the cost and the average movement of particles during the iteration. It can be seen that the cost reduces greatly during iteration 35 as $\boldsymbol{Q} = \mathbf{0}_{46\times46}$ at this point. Only smaller improvements occur in the later iterations. The average movement between iterations is relatively high during the first 375 iterations (in the order of a wavelength) and reduces greatly as the speed parameters are reduced.

The resulting array is further compared to the reference array in Figure 5.9, where the PSLLs are plotted within the region of interest. Due to the discretization, this region has extended slightly from $45° \leq \theta \leq 135°$, $45° \leq \phi \leq 135°$ to $44° \leq \theta \leq 136°$, $44° \leq \phi \leq 136°$. The values of the two plots are summarized in Table 5.2. It is seen that the optimized array has a lower maximum PSLL, with the downside of having a higher average PSLL.

Similarly, the MLWs of the optimized and reference arrays are plotted in Figure 5.10 and summarized in Table 5.3. The performance is similar, but the optimized array has slightly lower MLWs.

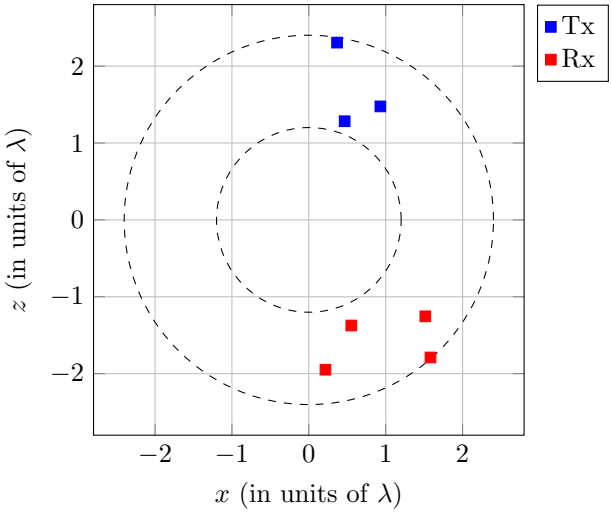


Figure 5.4: Rx and Tx positions of optimized array.

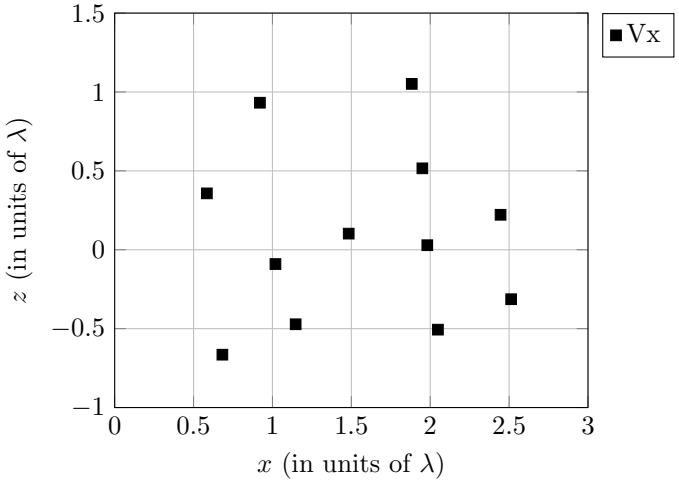


Figure 5.5: Virtual element positions of optimized array.

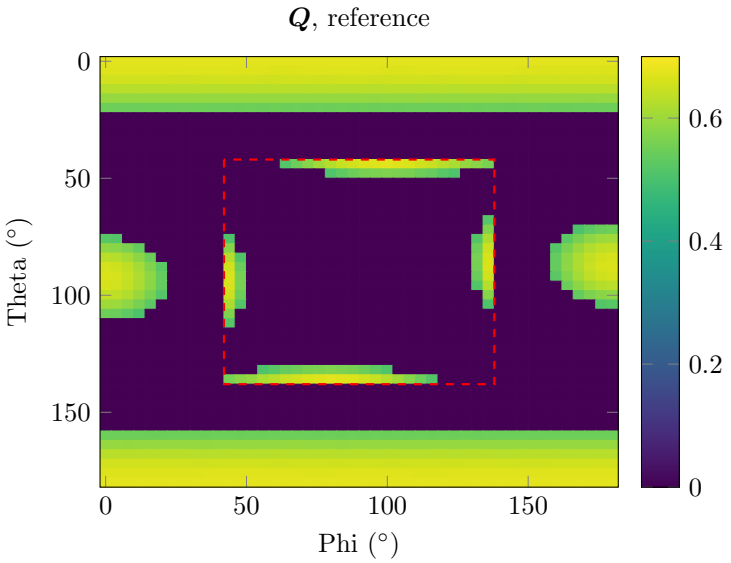


Figure 5.6: Values of $\boldsymbol{Q}$ for the reference array in the full angular region. Region of interest is marked with the dashed red box.

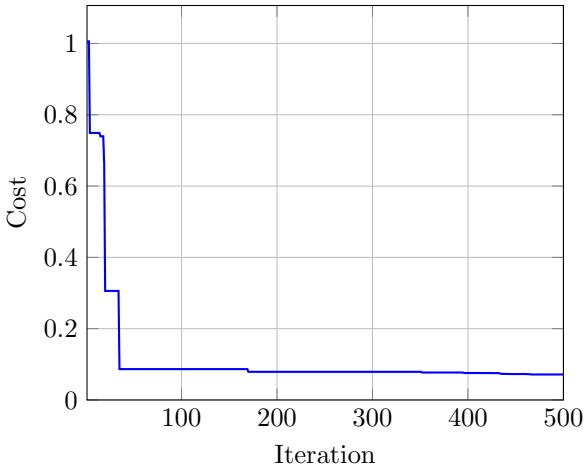


Figure 5.7: Best cost (including historical best) per iteration.

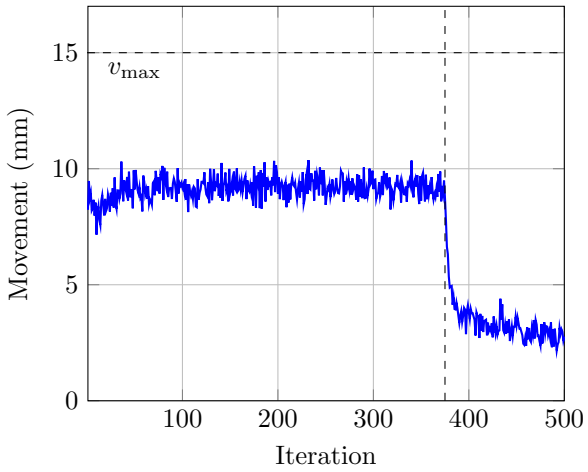


Figure 5.8: Average movement of elements per iteration. Vertical dashed line marks the iteration where the movement parameters are changed.

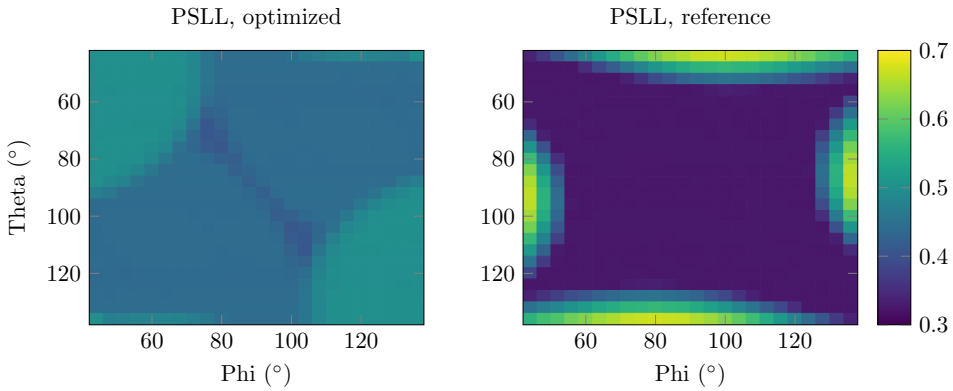


Figure 5.9: Peak side lobe levels in the region of interest.

Array	Average PSLL	Minimum PSLL	Maximum PSLL
Optimized	0.456	0.408	0.500
Reference	0.381	0.323	0.671

Table 5.2: Side lobe values of optimized array compared to reference array within the region of interest.

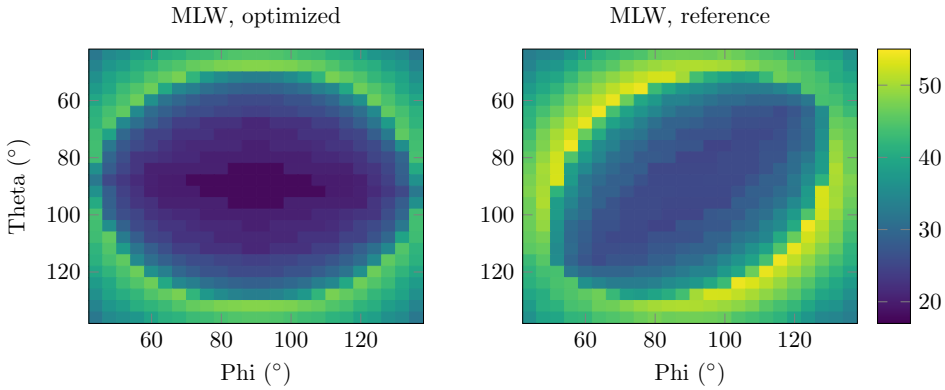


Figure 5.10: Main lobe widths in the region of interest. Values given in degrees.

Array	Average MLW	Minimum MLW	Maximum MLW
Optimized	31	18	48
Reference	38	25	54

Table 5.3: Main lobe width values of optimized array compared to reference array within region of interest. All values given in degrees.

5.2.2 Multiple plane layout

The final antenna layout after an optimization loop consisting of 300 iterations is given in Figure 5.11. Figure 5.13 and 5.14 show the evolution of the cost and the average movement of particles during the iterations. The optimization algorithm managed to suppress the relevant SLLs so that $\boldsymbol{Q} = \mathbf{0}_{46\times46}$, compared to $\boldsymbol{Q}$ of the reference array which is given in Figure 5.12. In this optimization run, it took 175 iterations before this was achieved. The PSLLs of the optimized array and the reference array are given in Figure 5.15. The variations of PSLLs are seen to be relatively smooth, except for the angles with the smallest and largest values of ϕ. These discontinuities are introduced by the shift in the calculation of the ambiguity function as given by the two cases in (4.10). The PSLL values are summarized in Table 5.4, which shows the same general tendencies as for the circular geometry. The maximum PSLL of the optimized array is reduced compared to the reference array, with the cost of higher average and slightly higher minimum PSLL. The improvements in MLW because of the optimization are more discernible for this geometry than for the circular geometry. The MLWs of the optimized and reference array are given in 5.16. These values are summarized in Table 5.5.

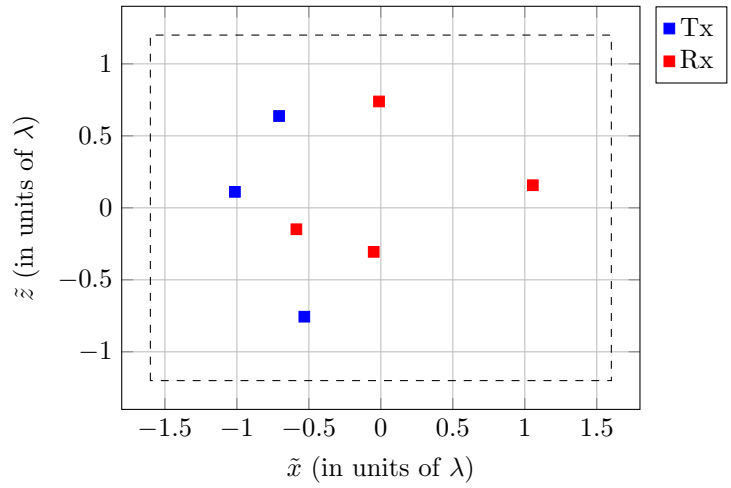


Figure 5.11: Rx and Tx positions on each side of optimized array.

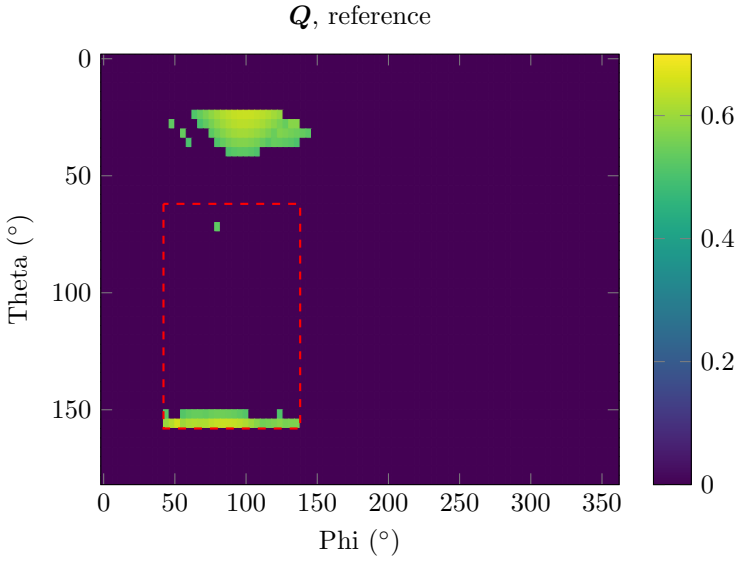


Figure 5.12: Values of $\boldsymbol{Q}$ for the reference array in the full angular region. Region of interest is marked with the dashed red box.

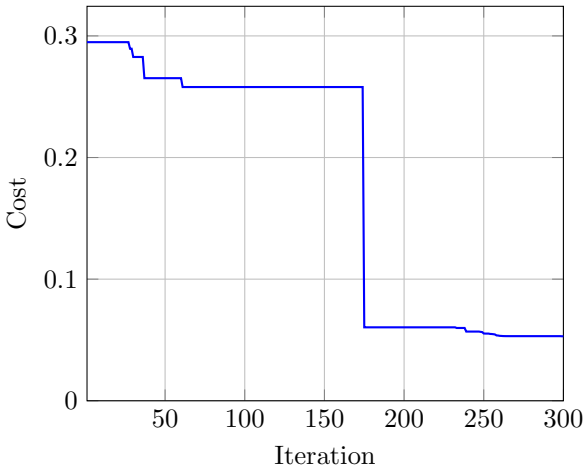


Figure 5.13: Best cost (including historical best) per iteration.

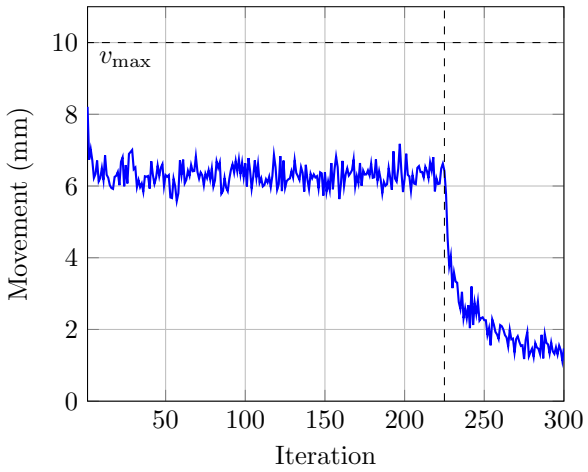


Figure 5.14: Average movement of elements per iteration. Vertical dashed line marks the iteration where the movement parameters are changed.

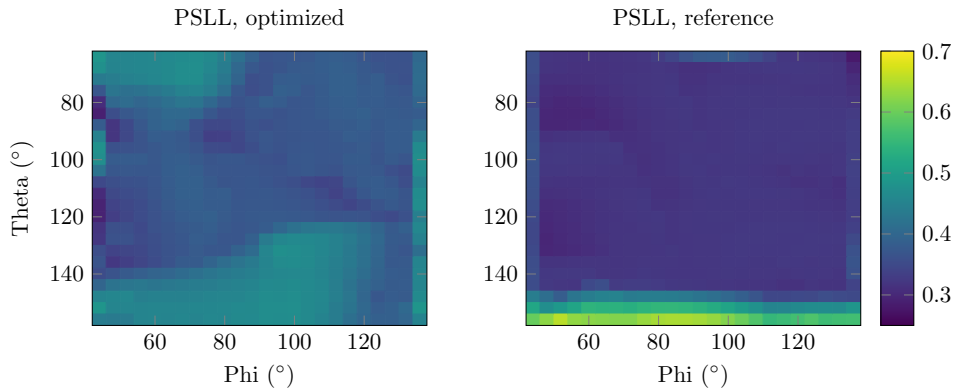


Figure 5.15: Peak side lobe levels in the region of interest.

Array	Average PSLL	Minimum PSLL	Maximum PSLL
Optimized	0.396	0.289	0.484
Reference	0.344	0.280	0.653

Table 5.4: Side lobe values of optimized array compared to reference array within the region of interest.

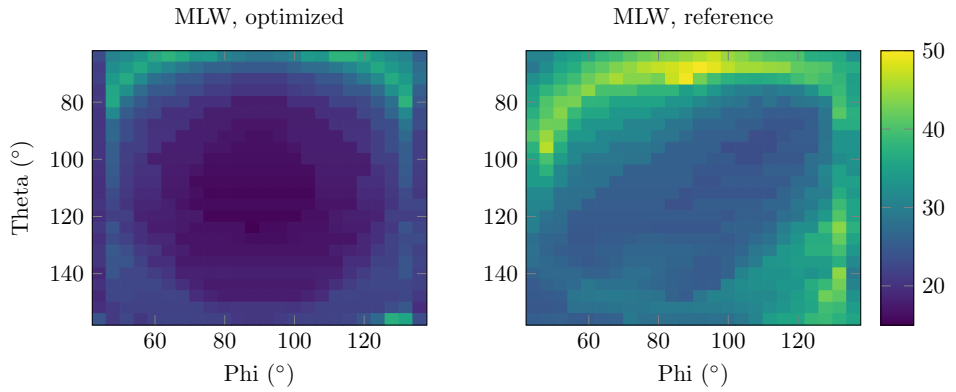


Figure 5.16: Main lobe widths in the region of interest. Values given in degrees.

Array	Average MLW	Minimum MLW	Maximum MLW
Optimized	21	15	38
Reference	31	23	50

Table 5.5: Main lobe width values of optimized array compared to reference array within region of interest. All values given in degrees.

5.3 Simulation results

5.3.1 Circular layout

As described in section 4.6, the response of the optimized array was simulated for a number of DoAs. By using the analytical model and estimating the DoA from this, a deviation was found between the actual and estimated DoA. This deviation is given in Figure 5.17. It can be seen that the deviation is generally small, with a maximum of 6°. However, it should be noted that there is no guarantee that the simulated responses are more accurate than the analytically calculated responses.

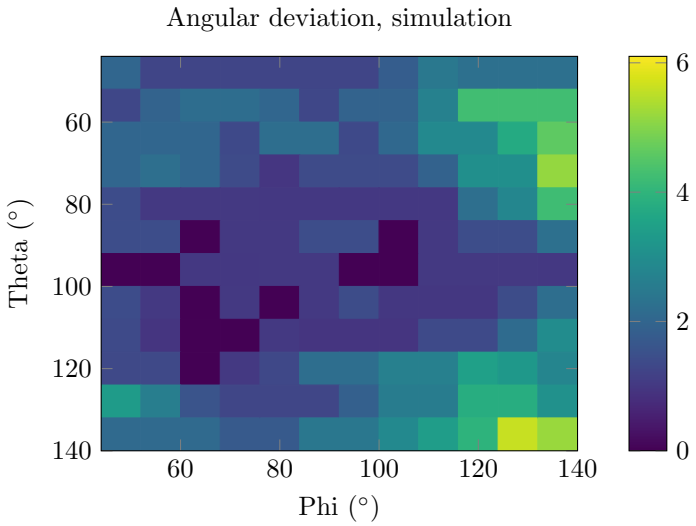


Figure 5.17: Angular estimation error of simulated responses.

5.3.2 Multiple plane layout

In the same way as for the circular geometry, the angular deviations in the RoI is given in Figure 5.18. While the average deviation is small (1.7°), there are two angles with deviations around 10°.

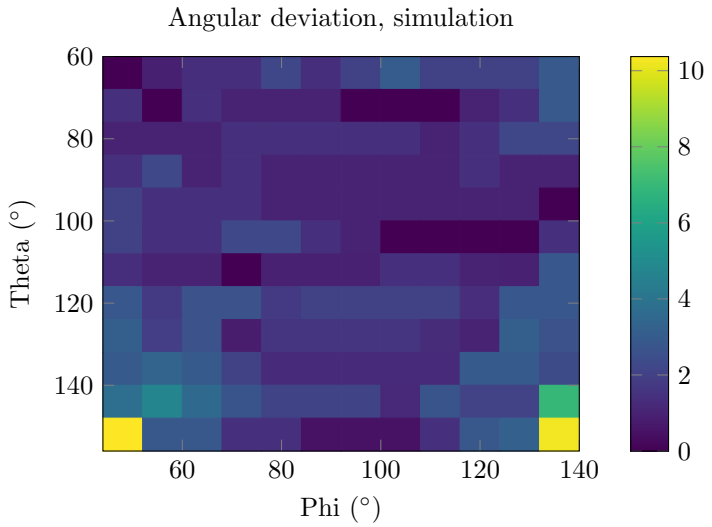


Figure 5.18: Angular estimation error of simulated responses.

5.4 Noise results

5.4.1 Circular layout

The noise performance of the optimized circular array can be seen in Figure 5.19, together with the noise performance of the reference array. The deviation given in the figure is the average deviation within the RoI. It can be seen that the arrays have equal estimation performance for low SNRs where the noise causes several ambiguities. As the SNR increases, small variations can be seen between the different instances. The simulated arrays consistently have larger errors than the analytical arrays. This should be expected, as the simulated arrays exhibit deviations even without noise. The difference between the optimized and reference array is small throughout the SNR range.

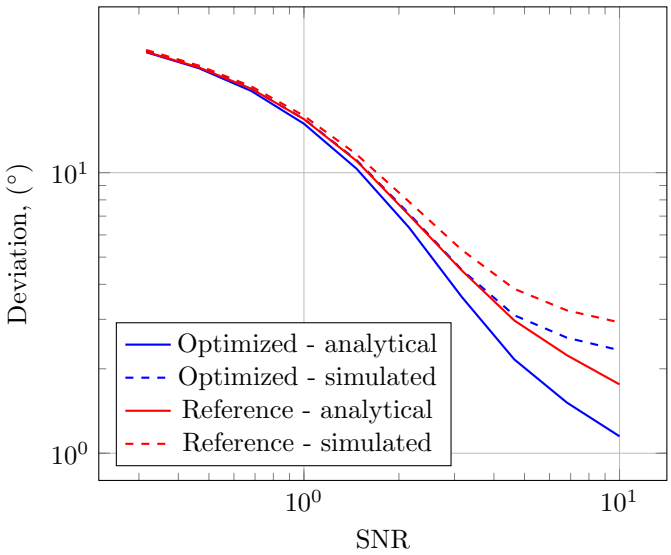


Figure 5.19: Angular deviation error for a range of SNRs.

5.4.2 Multiple plane layout

The deviation-to-SNR plot for the multiple plane geometry is given in Figure 5.20, where the minimum SNR has been reduced to 0.1. Comparing this to Figure 5.19, it can be seen the deviations at low SNRs are smaller for the multiple plane geometry compared to the circular geometry. This should be expected, as the former utilises gain information that is averaged over several signals, making it less sensitive to noise. For higher SNRs, the differences between Figure 5.19 and 5.20 are less pronounced. Furthermore, it can be seen that the reference array performs slightly better for SNRs below around 0.8. This might be explained by the PSLLs of the arrays, as the side lobes should affect the risk of ambiguities. Observing Figure 5.15, it can be seen that the number of angles at which the PSLL is higher for the reference array compared to the optimized array is low. Thus,

as the deviation is calculated as the mean over all angles in the RoI, the effect of reducing these sidelobes below the threshold value might not have a significant effect on the average deviation. For higher SNRs, the optimized array performs slightly better than the reference array.

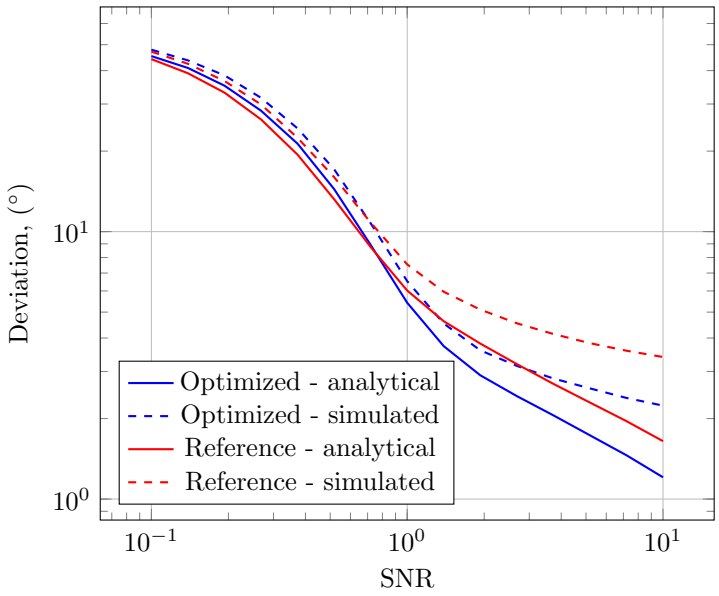


Figure 5.20: Angular deviation error for a range of SNRs.

Chapter 6
Discussion

6.1 Discussion on results

Based on the results presented in section 5.2, the optimization method seems to work as intended. For both the circular and multiple plane geometry, the PSLLs were reduced to below the threshold value of $t = 0.5$ while also maintaining slightly lower MLWs than the reference array. From the plots of the evolution of the cost, it can be seen that the main objective of reducing all sidelobe levels below the threshold was achieved with the fewest iterations for the circular geometry. The few other optimisation runs that were made (but not presented in the results) showed the same tendency. This is noteworthy, as the main objective should be easier to accomplish for the multiple plane geometry, due the use of more antenna elements. A reason for this behaviour was not found, but could be because of a more complex relation between element placements and the resulting ambiguity function.

Another result that is worth further discussion is the simulation of the multiple plane layout, as given in Figure 5.18. As pointed out, there are two directions with a rather high deviation of around 10° at the highest θ value. These deviations are caused by a mismatch of how the signal amplitudes are calculated analytically compared to the values given by the simulation. In the analytical model, it is assumed that the sides only receives signals that are covered by their field of view (FoV). This means that if the normal direction of the side is $\hat{\boldsymbol{n}}$ and a signal arrives from the direction $\hat{\boldsymbol{r}}$, the analytical model assumes that no signal is received if $\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{r}} \leq 0$. This is an approximation, and some signal amplitude is still received out of the FoV. The two values with higher deviation in the figure comes from directions just outside of the FoV of the opposing side to the Tx side, resulting in an estimation error. This problem could probably be addressed by adjusting the gain profile used in the analytical model to include gain slightly outside of the FoV.

Lastly, the noise results indicate that the optimized arrays perform similarly to the reference array. This might mean that it is not valuable to implement the optimized arrays practically, however the method is still useful as it can be used for arbitrary dimensions of the geometries and for different RoIs. In order to illustrate this, two additional arrays were optimized in the circular geometry using different sets of parameters. The first array consisted of 4 Tx and 6 Rx antennas, while

the second one was optimized for a much thinner geometry with $R_{\min} = 20\,\text{mm}$ and $R_{\max} = 25\,\text{mm}$. The optimization algorithm managed to suppress all SLLs to below $t = 0.5$ for these geometries. The resulting layouts are given in Figure 6.1 for reference. It is worth noting that the reference array would not be able to fit within the thin geometry.

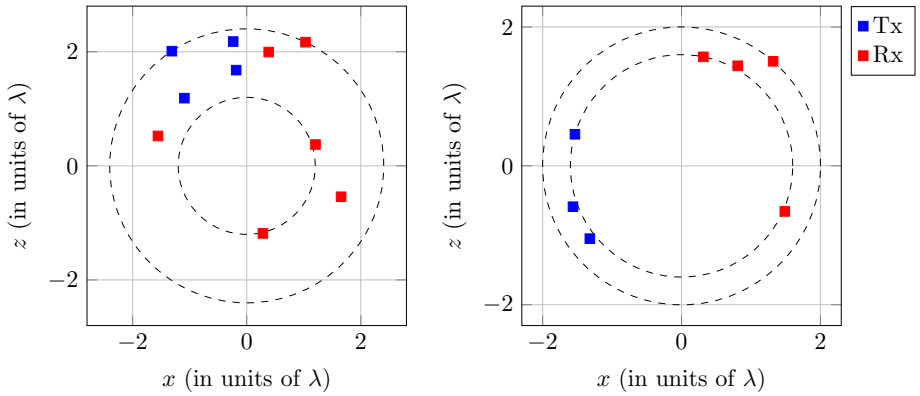


Figure 6.1: Additional optimized arrays

6.2 Practical implementation

There are notable challenges that would have to be overcome in order to implement the optimized arrays instead of the reference array in practice, and which reduces the potential gain of using them. Firstly, the complexity of the signal processing could increase significantly, at least if CBF is used as the DoA estimation method. As described in section 2.8.1, the FFT can be applied on arrays with elements that are placed on a grid, such as the reference array, in order to make the DoA estimation efficient. As the optimized array layouts are gridless, extra processing time would have to be taken into account in a practical implementation. One way to alleviate this problem would be to instead use the GA, as this naturally places the elements on a specified grid. While most of the grid points would be empty, the FFT can still be applied on the resulting sparse array.

Another important aspect to consider is the feeding of the patch antennas. If the inset-fed patch antenna is used, microstrip lines would have to extend from each Rx and Tx element to the receiver and transmitter. Finding suitable feeding routes for the optimized array could potentially be more intricate than for the reference array depending on the optimized layout.

6.3 Possible improvements

There are several possible improvements to the method that could have been implemented to improve the effectiveness of the optimization algorithm, as well as

the accuracy of the quality measures. These improvements are discussed in the sections below.

6.3.1 Improvements to assessment

The assessment of the arrays in this report has been based on the conventional ambiguity function as used in several papers for optimization of 1D and 2D arrays, as well as a newly defined ambiguity function that takes the gain profiles of the antennas into account. One limitation of the assessment in this report is that only the single signal case is considered, as it is deemed unlikely that two or more signals should appear within the same range-Doppler bin. This limitation is also imposed in the optimization approaches presented in Chapter 3. By considering the case of multiple signals, the optimization problem would grow substantially, and it is not clear to the author how it could be efficiently studied. It might prove sufficient to reduce the SLL threshold t in the single signal case to also ensure that multiple signals can be detected without ambiguity. Furthermore, the MLW should be directly linked to the separability of multiple signals. However, this has to be studied in more detail.

Secondly, the optimization method builds on the assumption that the PSLLs are related to the risk of DoA ambiguities, and that the MLWs are related to the resolution. As described in section 2.8.3, these values are relevant in different SNR ranges. In a real world application, a desired SNR range within which the radar should operate would probably be presented. Based on this, the requirements on the PSLLs and MLW could be altered in order to achieve the best performance within this range. For instance, if the SNR range was high, the risk of ambiguities would be low and the threshold t could thus be increased.

6.3.2 Improvements to optimization algorithm

While the optimization seems to perform as intended, there are always improvements that could be implemented. One of the limiting factors in the optimization was the speed. As mentioned previously, the optimization runs took around a full day on the available hardware. While this is not unmanageable for a single optimization, it limits the possibility of fine-tuning the parameters to give faster convergence and lower final cost. There are a few improvements that could be implemented that might alleviate these problems.

Firstly, the positions of the Tx and Rx antennas are, as groups, translationally invariant with respect to the performance of the array. This means that if all Rx or all Tx antennas in an array are equally displaced within the geometry, the resulting array is identical in performance, at least analytically. This concept can be seen by observing the formulation of the virtual array vector in (2.21). When the Tx or Rx antennas are moved, all elements in the vector are phase shifted by the same amount, which does not affect the DoA estimation. This translational invariance is not utilized in the optimization algorithm as the best absolute positions of the Tx and Rx are saved. Secondly, the array performance is also unchanged when a permutation is performed on the Tx or Rx antennas, respectively. This is also not taken into account in the optimization, and would probably lead to faster

convergence.

Beyond speed, other improvements could be implemented to discourage the optimization from getting stuck in local minimas. This has been the focus of much research on PSO since the introduction in [21]. In for instance [26], Kao and Zahara combined the mating and mutation processess of the genetic algorithm with PSO and showed that this lead to improved solution quality and faster convergence. However, implementing such an algorithm would further complicate the fine tuning of the algorithm as more parameters would be introduced.

Chapter 7
Conclusions

In this report, two array layouts have been optimized using the particle swarm optimization method. The first geometry was planar and circular, while the second one consisted of four separate planar surfaces that covered 360° in azimuth. In the optimization of the arrays, they were assessed based on their ambiguity functions. It was deemed of interest to focus the analysis on two parameters, the PSLL and the MLW of the ambiguity function, as these are related to the accuracy of the DoA estimation. The optimization method achieved the primary objective of reducing the PSLLs to below the threshold value, and could reduce the MLWs to smaller values compared to the reference array. The resulting arrays were simulated and the responses were seen to be close to the analytical model used for the optimization. Lastly, the noise performance of the arrays was investigated. Only a smaller difference between the optimized and reference arrays were found, which indicate that a practical implementation of the optimized arrays could be hard to justify. However, the implemented method can handle variations in geometry sizes and an arbitrary number of antenna elements, and is therefore applicable to larger arrays or more limited geometries, where conventional arrays either do not exist or are not usable.

Bibliography

[1] Merrill I. Skolnik. *radar*. `https://www.britannica.com/technology/radar`. Accessed: 2024-06-03. 2024.

[2] Nobel Foundation. *Nobel Lectures, Physics 1901-1921*. Elsevier Publishing Company, 1967.

[3] Gaspare Galati. *100 Years of Radar*. Springer, 2016.

[4] Sujeet Milind Patole et al. "Automotive radars: A review of signal processing techniques". In: *IEEE Signal Processing Magazine* 34.2 (2017).

[5] Frank B. Gross. *Smart Antennas with MATLAB, Second Edition*. McGraw-Hill Education, 2015.

[6] Mark A. Richards, James A. Scheer, and William A. Holm. *Principles of Modern Radar - Basic Principles*. SciTech Publishing, 2010.

[7] Volker Winkler. "Range Doppler detection for automotive FMCW radars". In: *Proceedings to the 37th European Microwave Conference*. 2007, pp. 1445–1448.

[8] Gerhard Kristensson. *Spridningsteori med antenntillämpningar*. LTH, 2008.

[9] Constantine A. Balanis. *Antenna Theory - Analysis and design, Fourth edition*. John Wiley & Sons, 2016.

[10] John D. Kraus and Ronald J. Marhefka. *Antennas For All Applications*. McGraw-Hill, 2002.

[11] Donald J. Lanzinger. "Group delay caused by impedance mismatch". In: *29th ARFTG Conference Digest*. 1987, pp. 247–264.

[12] Adolfo Di Serio et al. "2-D MIMO radar: A method for array performance assessment and design of a planar antenna array". In: *IEEE Transactions on Antennas and Propagation* 68.6 (2020).

[13] Brad Jackson et al. *Theory, design, and measurement of novel uniform circular antenna arrays for direction of arrival estimation.* Tech. rep. DRDC-RDDC-2015-R010. Defence Research and Development Canada, 2015.

[14] Yibao Liang et al. "A review of closed-form Cramér-Rao bounds for DOA estimation in the presence of gaussian noise under a unified framework". In: *IEEE Access* 8 (2020).

[15] Cesar Iovescu and Sandeep Rao. *The fundamentals of millimeter wave radar sensors.* Tech. rep. SPYY005A. Texas Instruments, 2020.

[16] Fredrik Athley. "Threshold region performance of maximum likelihood direction of arrival estimators". In: *IEEE Transactions on Signal Processing* 53.4 (2005).

[17] Oliver Lange and Bin Yang. "Array geometry optimization for direction-of-arrival estimation including subarrays and tapering". In: *2010 International ITG Workshop on Smart Antennas (WSA).* 2010, pp. 135–142.

[18] Harry L. Van Trees. *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory.* John Wiley & Sons, 2002, pp. 938–984.

[19] Randy L. Haupt. "An introduction to genetic algorithms for electromagnetics". In: *IEEE Antennas and Propagation Magazine* 37.2 (1995).

[20] Daniel W. Boeringer and Douglas H. Werner. "Particle swarm optimization versus genetic algorithms for phased array synthesis". In: *IEEE Transactions on Antennas and Propagation* 52.3 (2004).

[21] James Kennedy and Russell Eberhart. "Particle swarm optimization". In: *Proceedings of ICNN'95 - International Conference on Neural Networks.* 1995, pp. 1942–1948.

[22] Oliver Lange and Bin Yang. "Antenna geometry optimization for 2D direction-of-arrival estimation for radar imaging". In: *2011 International ITG Workshop on Smart Antennas.* 2011.

[23] Toygar Birinci and Yalcın Tanık. "Optimization of nonuniform array geometry for DOA estimation with the constraint on gross error probability". In: *Signal Processing* 87.10 (2007).

[24] Xiaoyan Chen et al. "Optimization of virtual array element position for sparse array based on particle swarm algorithm". In: *2019 IEEE International Conference on Signal, Information and Data Processing (ICSIDP).* 2019.

[25] International Telecommunication Union. *Radio Regulations - Articles*. 2020.

[26] Yi-Tung Kao and Erwie Zahara. "A hybrid genetic algorithm and particle swarm optimization for multimodal functions". In: *Applied Soft Computing* 8.2 (2008).

Printed by Tryckeriet i E-huset, Lund 2024

LUND
UNIVERSITY

Series of Master´s theses
Department of Electrical and Information Technology
LU/LTH-EIT 2024-1010
http://www.eit.lth.se