UCL Université catholique de Louvain

École polytechnique de Louvain (EPL)

epl

Optimal interference nulling for large arrays of coupled antennas

Dissertation presented by
Valentin HAMAIDE

for obtaining the Master's degree in
Mathematical Engineering

Supervisor(s)
Christophe CRAEYE, François GLINEUR

Reader(s)
Bui Van HA, Eloy DE LERA ACEDO , Claude OESTGES

Academic year 2016-2017

Acknowledgements

I would like to thank my two advisors, Pr. Craeye and Pr. Glineur for giving me the opportunity to work on such a challenging project but also for their advice and the time they devoted to me during our numerous meetings.

I would also like to give a special thanks to Bui Van Ha for helping me with the code, for answering my many questions regarding antennas, for our long exchange of emails and for all the advice he gave me during the year.

I also thank Dr. Eloy de Lera Acedo from University of Cambridge, who followed this work from the beginning and provided advice in the early stage of this work.

Finally, I would like to thank every person that supported me during those 5 years at UCL, the professors, the students and the friends I have met here.

Contents

Acknowledgements **ii**

Abbreviations **v**

Symbols **vi**

1 Introduction **1**
1.1 Radiation pattern synthesis and interference nulling 1
1.2 Current techniques 2
1.3 SKA Project 2
1.4 Scope of this thesis 3

2 Background theory **4**
2.1 Radiation pattern 4
2.2 Array of antennas 5
2.2.1 Array factor 6
2.2.2 Array pattern 7
2.3 Data Model 8
2.4 Beamforming 9
2.4.1 Optimal beamforming 10
2.4.2 Conventional beamforming 10
2.4.3 Minimum variance beamformer (MV) 11
2.4.4 Other methods 12

3 Direct methods **13**
3.1 Need for a model adapted to radio astronomy 13
3.2 Orthogonal projection method 14
3.2.1 Application to the SKA telescope 16
3.3 Introduction of mutual coupling 21
3.3.1 Application to the SKA telescope 21
3.4 Two polarization antennas 22
3.4.1 Application to the SKA telescope 23
3.5 Taking into account the 2 arms of the SKALA element 23
3.6 Influence of the noise 24
3.7 New method for nulling 26
3.8 Comparison of the two methods 27
3.8.1 Isotropic case 28

3.8.2 Extension to mutual coupling and 2 polarizations 31

4 Convex optimization models **33**
4.1 Problem formulation 34
4.1.1 Convexity & second order cone programming (SOCP) 36
4.1.2 Main beam constraints 39
4.1.3 Weights constraints 41
4.1.4 Final model 41
4.2 Application to the SKA telescope 42
4.2.1 Test for different sets of parameters 43
4.2.1.1 Variation of Q 44
4.2.1.2 Variation of HB 45
4.2.1.3 Variation of NL 45
4.2.2 Comparison of the SLL between direct method and optimization 46
4.3 Problem of Robustness 46
4.4 Two arms antennas 47
4.5 Optimize both the position and weights 48
4.6 Real time optimization 49

5 Conclusion **51**

A Pseudo-inverse and least squares solution **53**
A.1 Rank of a matrix 53
A.2 Moore-Penrose Pseudoinverse 54
A.3 Over-determined System 55
A.3.1 **A** is full rank 55
A.3.2 **A** is rank deficient 55
A.4 Under-determined System 56
A.4.1 **A** is full rank 56
A.4.2 **A** is rank deficient 57

B Analytic comparison between the projection and pseudo-inverse method 58

Bibliography **61**

Abbreviations

SKA	**S**quare **K**ilometer **A**rray
SKALA	**SKA** **L**og-periodic **A**ntenna
IS	**I**solated **P**attern
EP	**E**lement **P**attern
EEP	**E**mbedded **E**lement **P**attern
AP	**A**rray **P**attern
DOA	**D**irection **O**f **A**rrival
SLL	**S**ide **L**obe **L**evel
NL	**N**ull **L**evel
SOCP	**S**econd **O**rder **C**one **P**programming

Symbols

$\mathbf{P}$	Matrix of EEPs at the center (multiplied by the phase factor)
$\mathrm{Re}\{\cdot\}$	Real part
$\mathrm{Im}\{\cdot\}$	Imaginary part
A^{+}	Pseudo-inverse of A
$(\cdot)^{H}$	Conjugate transpose
$(\cdot)^{*}$	conjugate
$(\cdot)^{T}$	transpose
$(\mathbf{V})^{\perp}$	Subspace orthogonal to $\mathbf{V}$

Chapter 1

Introduction

1.1 Radiation pattern synthesis and interference nulling

In this thesis, we will work with antenna arrays, which are sets of several antennas working together. They can combine and superpose to improve the performance over that of a single antenna. By applying different excitation currents on the elements, we are able to produce a wide variety of radiation pattern.

The objectives of this thesis is to design a numerical algorithm and/or an optimization model capable of matching a desired radiation pattern, by finding the proper excitation currents on the elements. Throughout this thesis, we will work with large antenna arrays and take into account mutual coupling between the antennas. By matching a desired radiation pattern, we mean having a radiation pattern that is highly sensitive in a desired direction (beam direction), less sensitive in every other direction and forming nulls in specific directions of associated interferences (i.e. interference cancellation). The positions that we want to null correspond to large astronomical sources in the sky such as the sun, the moon, Cassiopeia, etc. Indeed, those large sources emit electromagnetic signals that are strong in amplitude, which overwhelms other weaker radio sources in space. By spatially discriminating the sensitivity of the array of antennas in the direction of those strong sources, we are able to study regions in the sky with weaker radio signals.

Techniques and algorithms to match a desired radiation pattern are called beamforming, or pattern synthesis [1]. There are two types of beamforming : deterministic beamforming which is independent of the data received by the system or adaptive beamforming

for which we use the statistics of the signals received. In this thesis we will only be interested in deterministic beamforming with no prior knowledge of the statistics or the structure of the signals received.

Nulling is essentially a beamforming process. We want a zero in the radiation pattern of the array of antenna where the interference occurs so that it has no sensitivity at all in that direction. Of course, the location of the interferer needs to be known in order to cancel it. Fortunately, there exist images of the sky that are available to determine the positions of astronomical sources in real time from anywhere on earth.

1.2 Current techniques

Beamforming has been considerably studied in the literature. A thorough survey about beamforming can already be found in [1] back in 1988; it explains different techniques of beamforming for both the deterministic and adaptive cases. Conventional deterministic beamforming is detailed in the paper and some optimization problems are solved. However, interference cancellation techniques involving orthogonal projections, a type of method that we use in this thesis, were not used back then. In 1993, a first subspace separation technique involving orthogonal projection was derived in [2] and in more recent papers the method has been adapted for radio telescopes phased array in [3, 4]. This beamforming approach along with suppression of interference via subspace orthogonal projections will be partly the subject of chapter 3.

Another way to undertake beamforming is through a convex optimization model. Pattern synthesis via convex optimization has also been studied in the literature in [5–7]. Chapter 4 will be devoted to those models.

1.3 SKA Project

This thesis is in the continuity of research projects related to the Square Kilometer Array (SKA) telescope at UCL. A first UCL master thesis involving the optimization of SKA arrays was presented by Thibault Clavier in 2012 [8]. However, the present thesis is only weakly related to the work done by Clavier, which was not involving mutual coupling and was not focused on nulling.

The SKA project is an international effort to build the worlds largest radio telescope, with eventually over a square kilometer (one million square meters) of collecting area [9]. The scale of this radio telescope will be unprecedented with thousands of dishes and a million of low-frequency antennas. This will provide the telescope the ability to observe the sky with a level of details never achieved before and proceed much faster than any other telescope.

The project is being developed in two phases. For Phase 1, antennas are planed in two different countries. Australia is hosting the LOW frequency antennas that are placed in 500 stations of about 250 antennas in a remote location near Perth and South Africa is hosting 200 MID frequency antennas. This first phase has already started and is needed to test the whole system and obtain the first results. Once this is done, the second phase will complete the project by establishing 2000 MID and HIGH frequency antennas in South Africa and a million low frequency antennas in Australia.

1.4 Scope of this thesis

In this thesis, we will implement and extend current techniques of beamforming and nulling to an experimental station of 256 elements of the SKA LOW (low frequency antennas) array. We will also extend those existing methods to introduce the mutual coupling between antennas. We believe this is important, as the techniques found in the literature usually do not take into account these coupling effects.

The thesis is organized as follows. In chapter 2 we will remind the reader about some basics of antennas and introduce some vocabulary along with a mathematical framework needed for the rest of the thesis. In chapter 3, we will discuss about direct methods, i.e. projection methods that are used to find the set of weights to be applied to every antennas in order to obtain the desired radiation pattern. In chapter 4, we will discuss about convex optimization models used for nulling. Specific attention will be given to maintaining a good sensitivity. Finally, a conclusion about the techniques developed in this thesis and possible future work will be drawn.

Chapter 2

Background theory

This chapter reviews the basics about antennas that are necessary to understand the remainder of the thesis [1]. We will also explain some vocabulary and abbreviations that will be used in the next chapters and we will define a mathematical framework for the study of antennas.

2.1 Radiation pattern

First of all, before introducing the concept of radiation pattern, it is necessary to understand the main property of an antenna, that is the reciprocity principle. The principle of reciprocity states that the receive and transmit properties of an antenna are identical. Hence, the sensitivity of an antenna when used for receiving is the same as the power radiated by the antenna when used for transmitting, within a constant factor.

In this thesis, we will only consider the antenna as receiver. Therefore, when we talk about radiation pattern, we can refer to it as the sensitivity of the antenna as a function of the direction. For a single antenna, we call it the element pattern (EP). We can represent the radiation pattern in several ways i.e. as a polar plot (2D), a 3D plot in the uv plane (defined hereunder) or a cut (in 2D) along a certain angle. An example of a radiation pattern in a 2D plot is shown in Fig. 2.1. This allows us to introduce some vocabulary about radiation pattern. The main beam (or main lobe) is the beam with

[1] Most of the material from this chapter is based on the Antenna course given by Pr. Craeye at UCL and the corresponding lecture notes of A. Guissard and C. Craeye. [10]

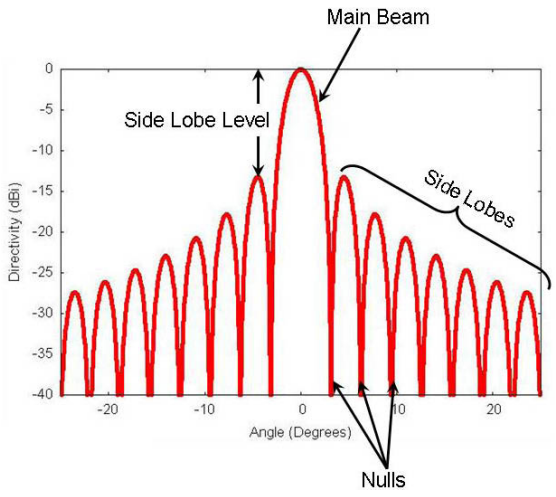


FIGURE 2.1: Example of radiation pattern

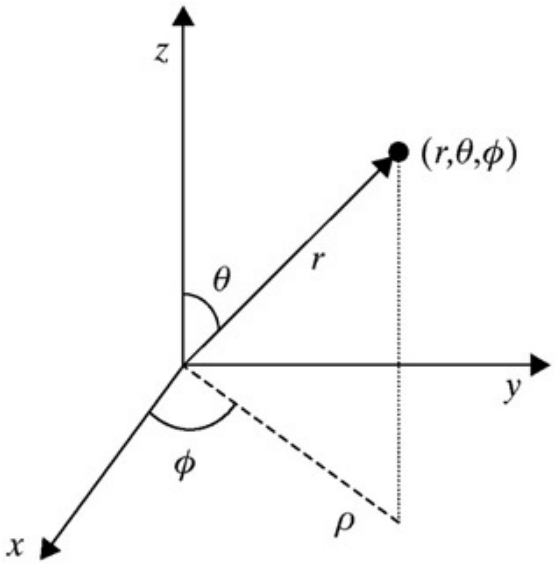


FIGURE 2.2: Spherical coordinate system

the highest gain, i.e. the direction in which we want to observe the radio waves. The sidelobes are all the other smaller beams that are parasitic. We usually want to avoid those as much as possible. The nulls are the zero in the patterns.

Let us now define a spherical coordinate system that will be used throughout this thesis. In Fig. 2.2, we define the coordinate system with θ the angle from the z-axis (inverse of elevation angle) and ϕ the azimuth angle from the x-axis. The angle θ vary from 0° to 90° and ϕ from 0° to 360°.

We denote u, v the cartesian coordinates corresponding to the xy plane as

$$\begin{aligned} u &= \sin(\theta)\cos(\phi) \\ v &= \sin(\theta)\sin(\phi) \end{aligned}$$

2.2 Array of antennas

This thesis does not consider the radiation pattern of a single antenna but rather concentrates on the radiation by an array of several antennas. All our numerical tests will be focused on an array of 256 elements, which represent one SKA station. An array of antenna is a set of antennas that are connected together and work as a single antenna. Indeed, based on the the principle of superposition, the field generated by N antennas

will be denoted

$$\mathbf{E} = \sum_{n=1}^{N} \mathbf{E}_n \tag{2.1}$$

This is due to the linearity of the Maxwell equations. One has to note that in this formula, each field $\mathbf{E}_n$ is computed in the presence of the other antennas while they are inactive. This field is actually different from the field that would be radiated by the antenna alone, if it was isolated. This is due to the mutual coupling between antennas, an important concept for this thesis that we will emphasize later on. Usually this coupling is assumed to be negligible for easier handling by mathematical and numerical methods but we will see that in the case of nulling, it is essential to take this effect into account.

The pattern of an antenna isolated from the other ones is called the isolated pattern (IP) while the pattern of an antenna in the presence of other (but inactive) antennas is called the embedded element pattern (EEP).

There are several advantages to work with an array of antennas rather than with a single element [11, Ch. 2]. First of all, the array can achieve a high gain and narrow beam in a given direction by combining the radiation pattern of each antenna. Secondly, it can lower the sidelobes of the radiation pattern. And thirdly, an array of antennas has more degrees of freedom which means that we can "play" with the antennas weights to meet a desired radiation pattern.

The antenna arrays that we will work with are called phased arrays. They are arrays of antennas controlled electronically by a computer and connected together with a phase shifter. They can steer the beam in a given direction without having to move the antennas.

2.2.1 Array factor

The array factor is a mathematical factor used to represent the geometry of an array of antennas. When multiplied by the directivity of each antenna (i.e. their element pattern), we get the radiation pattern of the whole array, called the array pattern (AP).

The array factor is defined as

$$AF(\theta,\phi) = \sum_n w_n \exp(jk\mathbf{u}\cdot\mathbf{r}_n) \tag{2.2}$$

$$= \sum_n w_n \exp\left(j\frac{2\pi}{\lambda}\left(r_{x_n}\sin(\theta)\cos(\phi) + r_{y_n}\sin(\theta)\sin(\phi)\right)\right) \tag{2.3}$$

where $k = \dfrac{2\pi}{\lambda}$ is the wavenumber, $\mathbf{r}_n = (r_{x_n}, r_{y_n})$ is the position of the n^{th} antennas relative to the center of the array and $\mathbf{u} = (u, v, w) = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ is the unit vector pointing in the desired direction (θ, ϕ). However we will assume that the all the elements have the same heights and belong to the same 2D plane. Thus, we can neglect the last component w.

The expression in the exponential represents the time delay in the reception of a radio wave coming from the direction (θ, ϕ) relative to the center of the array (can be chosen arbitrarily) for each antenna position $\mathbf{r}_n$. This time delay can be represented by a phase shift if we assume that the signal received is narrow-band. The factor w_n is a complex coefficient that represents the excitation current applied to antenna n. We will choose these coefficients to match a desired radiation pattern; this is called beamforming. Note that if the w_n are real and equal, this is just adding all the signals in phase.

2.2.2 Array pattern

The array pattern (AP) is the radiation pattern of all the elements of the array combined together. It is defined as $AP = EP \times AF$ with $EP = f(\theta, \phi)$ the element pattern. If we do not take into account the mutual coupling, the element pattern is the same for every elements, i.e. $f_n(\theta, \phi) = f(\theta, \phi) \quad \forall n$. Conversely, if we take into account the mutual coupling, we have that $f_{n1}(\theta, \phi) \neq f_{n2}(\theta, \phi)$ for $n_1 \neq n_2$. Therefore, we can write the general formula for the array pattern

$$AP(\theta,\phi) = \sum_n w_n f_n(\theta,\phi) \exp\left(j\frac{2\pi}{\lambda}\left(r_{x_n}\sin(\theta)\cos(\phi) + r_{y_n}\sin(\theta)\sin(\phi)\right)\right) \tag{2.4}$$

2.3 Data Model

Let us suppose we have a N-element antenna array of arbitrary geometry, the array output is given by

$$\mathbf{x}(t) = \mathbf{a}(\theta, \phi) \cdot s(t) \tag{2.5}$$

The vector $\mathbf{a}(\theta, \phi) = [a_1(\theta, \phi), ..., a_N(\theta, \phi)]^T$ is called a *steering vector* and it represents the array response in the direction (θ, ϕ). We assume that the signals arise from the far field (which is a good approximation for radio astronomy) propagating as plane wave. If we further assume that all the elements have the same directivity (same element pattern $f_1(\theta, \phi) = ... = f_N(\theta, \phi) = f(\theta, \phi)$), we have

$$\mathbf{a}(\theta, \phi) = f(\theta, \phi) \begin{bmatrix} e^{jk\mathbf{u}(\theta,\phi)\cdot\mathbf{r}_1} \\ \vdots \\ e^{jk\mathbf{u}(\theta,\phi)\cdot\mathbf{r}_N} \end{bmatrix} \tag{2.6}$$

The signal $s(t)$ in expression (2.5) represents the baseband signal.

If we consider that there are M signals coming from M directions, we can extend the definition of (2.5) to

$$\mathbf{x}(t) = \mathbf{A}(\theta, \phi) \cdot \mathbf{s}(t) \tag{2.7}$$

with $\mathbf{A}(\theta, \phi) = [\mathbf{a}(\theta_1, \phi_1), ..., \mathbf{a}(\theta_M, \phi_M)]$ and $\mathbf{s}(t) = [s_1(t), ..., s_M(t)]^T$. We can also assume the presence of an additive white gaussian noise (AWGN) present, and extend the definition to

$$\mathbf{x}(t) = \mathbf{A}(\theta, \phi) \cdot \mathbf{s}(t) + \mathbf{n}(t) \tag{2.8}$$

However, most of the methods that will be developed throughout this thesis are data independent, and we will generally assume

$$\mathbf{x} = \begin{bmatrix} e^{jk\mathbf{u}(\theta,\phi)\cdot\mathbf{r}_1} \\ \vdots \\ e^{jk\mathbf{u}(\theta,\phi)\cdot\mathbf{r}_N} \end{bmatrix}$$

for the output of the antenna and omit the time dependence and the directivity, while when we take into account mutual coupling, the directivity is different for every antenna

and therefore $\mathbf{x}$ is rewritten as

$$\mathbf{x} = \begin{bmatrix} f_1(\theta,\phi)e^{jk\mathbf{u}(\theta,\phi)\cdot\mathbf{r}_1} \\ \vdots \\ f_N(\theta,\phi)e^{jk\mathbf{u}(\theta,\phi)\cdot\mathbf{r}_N} \end{bmatrix}$$

2.4 Beamforming

As explained in the introduction, beamforming is a type of spatial filtering process used to match a desired radiation pattern. There are essentially two types of beamforming : deterministic beamforming and adaptive beamforming. The information and explanation in this section are mostly based on papers in [1, 12, 13].

In adaptive beamforming, we use the statistics of the signal received at the array to optimize the array response so that the weights are statistically optimum. The interference components within the received signals are identified and combined in such a way that they cancel each other out. The direction of the interferers do not need to be known in advance as they can be found from the statistics of the signals. Adaptive beamformers are able to automatically adapt to a given situation optimally. The criteria for optimality includes maximizing the output power for a given direction (conventional beamforming) and minimizing the output power while maintaining a fixed gain (Minimum variance beamformer).

In deterministic beamforming, we only use information about the location of the receivers and the direction of the incident waves. There are advantages and disadvantages in using this method. An example of drawback is that the location of the interferers need to be known with high accuracy. And an advantage is that we do not need to worry about the strength of the interference compared to the noise. Indeed, even if the interferer is weaker than the noise, we will always be able to place a proper null in the radiation pattern while in adaptive beamforming, if the interference to noise ratio (INR) is too small, we will not be able to detect it.

Even though we are mainly interested in deterministic beamforming in this work, we will nevertheless mention the most common adaptive beamforming algorithm used in phased arrays. The general concept of beamforming is that we add the signals x_n coming from each element and multiply each of them by a complex weight w_n^* that we can interpret

as a current excitation where * represent the complex conjugate. The output of the beamformer y is given by

$$y = \sum_{n=1}^{N} w_n^* x_n \tag{2.9}$$

$$= \mathbf{w}^H \mathbf{x} \tag{2.10}$$

where ^{H} represents the conjugate transpose.

2.4.1 Optimal beamforming

There are different criteria for optimality in beamforming and we present hereunder the two most common beamformers used in the literature. The first one is called the conventional beamformer (or Bartlett beamformer, or again delay-and-sum beamformer) and is based on the maximization of the power output in a given direction. The second one is called the minimum variance beamformer (or Capon's beamformer) and is based on the minimization of the output power subject to a fixed gain constraint. We define the covariance matrix of the data $\mathbf{R} = \mathbb{E}\{\mathbf{x}\mathbf{x}^H\}$ and assume that the noise is AWGN (Additive White Gaussian Noise), therefore $\mathbb{E}\{\mathbf{n}\mathbf{n}^H\} = \sigma^2\mathbf{I}$.

2.4.2 Conventional beamforming

The conventional beamformer maximizes the output power of the beamforming output for a given direction (θ, ϕ). The power of the output is given by

$$P = \mathbb{E}\{|y|^2\} = \mathbb{E}\{|\mathbf{w}^H\mathbf{x}|^2\} \tag{2.11}$$

The beamformer output for a signal arriving from a direction of interest corrupted by noise is given by

$$\mathbf{x}(t) = \mathbf{a}(\theta, \phi)s(t) + \mathbf{n}(t) \tag{2.12}$$

Maximizing the output power in that direction of interest results in

$$\max_{\mathbf{w}} \quad \mathbb{E}\left\{\mathbf{w}^H\mathbf{x}\mathbf{x}^H\mathbf{w}\right\} \quad \text{s.t.} \quad |\mathbf{w}| = 1 \tag{2.13}$$

$$\Leftrightarrow \max_{\mathbf{w}} \quad \mathbb{E}\{s(t)^2\}|\mathbf{w}^H\mathbf{a}(\theta, \phi)|^2 + \sigma^2|\mathbf{w}|^2 \quad \text{s.t.} \quad |\mathbf{w}| = 1 \tag{2.14}$$

The resulting optimal weight is then given by

$$\mathbf{w} = \frac{\mathbf{a}(\theta, \phi)}{\sqrt{\mathbf{a}^H(\theta, \phi)\mathbf{a}(\theta, \phi)}} \tag{2.15}$$

This weight can be interpreted as the delay applied to each antenna so that it optimally combines with a maximum in the desired direction (θ, ϕ). One has to notice that no information about the statistics of the signals is needed. The expression is therefore purely deterministic. Note that the denominator can be omitted as it is a constant and because we are concerned about the shape of the pattern rather than its magnitude (which depends on the distance) in the far field.

2.4.3 Minimum variance beamformer (MV)

The optimality criterion that is chosen here is to minimize the total energy output while holding constant the gain in a desired direction. This is a slightly better formulation since it allows one to reduce the importance of all the undesired effects such as interferences. However, this time we need information about the covariance matrix of the data; hence, the formulation is not deterministic anymore. Mathematically the problem can be formulated as [13]

$$\arg \quad \min_{\mathbf{w}} \quad \mathbb{E}\{|y|^2\} \quad \text{s.t.} \quad \mathbf{w}^H\mathbf{a}_0 = 1 \tag{2.16}$$

$$\Leftrightarrow \arg \quad \min_{\mathbf{w}} \quad \mathbb{E}\{|\mathbf{w}^H\mathbf{x}|^2\} \quad \text{s.t.} \quad \mathbf{w}^H\mathbf{a}_0 = 1 \tag{2.17}$$

The minimization problem can be solved using the Lagrange multipliers. We minimize $\mathcal{L}(\mathbf{w}, \lambda)$ where

$$\mathcal{L}(\mathbf{w}, \lambda) = \mathbb{E}\{|\mathbf{w}^H\mathbf{x}|^2\} + \lambda\left(\mathbf{w}^H\mathbf{a}_0 - 1\right) \tag{2.18}$$

$$= \mathbf{w}^H\mathbf{R}\mathbf{w} + \lambda\left(\mathbf{w}^H\mathbf{a}_0 - 1\right) \tag{2.19}$$

Setting the partial derivative $\frac{\partial \mathcal{L}}{\partial \mathbf{w}^H}, \frac{\partial \mathcal{L}}{\partial \lambda}$ equal to zero, we have

$$\frac{\partial \mathcal{L}}{\partial \mathbf{w}^H} = \mathbf{R}\mathbf{w} + \lambda\mathbf{a}_0 \tag{2.20}$$

$$\Rightarrow \mathbf{w} = -\lambda\mathbf{R}^{-1}\mathbf{a}_0 \tag{2.21}$$

and

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \mathbf{w}^H \mathbf{a}_0 - 1 = 0 \tag{2.22}$$

Substituting (2.21) into (2.22), we get $\lambda = \frac{1}{\mathbf{a}_0^H \mathbf{R}^{-1} \mathbf{a}_0}$. Therefore, the optimal weights are given by

$$\mathbf{w} = \frac{\mathbf{R}^{-1}\mathbf{a}_0}{\mathbf{a}_0^H \mathbf{R}^{-1} \mathbf{a}_0} \tag{2.23}$$

This beamformer is also called Capon's beamformer.

2.4.4 Other methods

Other methods do exist and are not discussed here as they are out of the scope of this thesis. But for information, a well-known alternative method is the so-called MUSIC (MUltiple SIgnal Classification) algorithm which is based on the eigen-decomposition of the covariance matrix $\mathbf{R}$, that is furthered decomposed into a noise and signal eigenvectors.

Chapter 3

Direct methods

In this chapter, we will derive direct methods for both beamforming and nulling. By direct methods, we mean methods that are not based on an iterative algorithm, as opposed to an optimization problem. Moreover we assume that no *a priori* information about the signals are known. Those methods are purely based on matrix computations, projections and solving of a linear system of equations. This enables to solve the problem with fast computation times.

We will start by explaining the principles of the general method based on orthogonal projections. Interference nulling with orthogonal projections has been studied in the literature in [2–4]. We will also explain why this approach is best suited for phased-array based radio telescopes. We will extend the existing methods to the case where mutual coupling between antennas is included and the case with dual polarization antennas. Indeed, the antennas used for the SKA telescope [14] have two polarizations. An alternative technique using matrix computations will also be derived and a comparison between the two methods will finally be studied.

3.1 Need for a model adapted to radio astronomy

The motivation for new (adaptive or not) nulling techniques to deal with radio interference in radio astronomy has increased in the past decades. The usual approach for adaptive nulling using power minimization such as the MV beamformer is not suitable for astronomy applications for several reasons [3].

First, the power inversion property of the MV beamformer limits the null depth to be proportional to the interference to noise ratio (INR) [3]. This is a problem for radio astronomy as there can sometimes be interference with INR less than 1 and yet stronger than the desired signals. This is for example the case when observing a spectral line with really weak power density.

Second, MV beamformer is subject to the phenomenon of *weight jitter*. Indeed, the inversion of matrix $\mathbf{R}$ causes the noise eigenvalues to become large and hence make a large contribution to the weights. Therefore, there is a considerable variability between the updates [3].

Projection methods can be used as alternatives for preventing the problems explained above. Those methods first explicitly estimate the interference subspace and then compute the weights by projecting them onto the subspace orthogonal to the interferers.

There are two cases to distinguish. If we know the direction of arrival (DOA) of the interferer, we can directly compute the interferer subspace and we are therefore in the case of deterministic nulling. If instead, we do not know those DOAs, we need to estimate them explicitly. To estimate those DOAs, we could simply compute an estimate of matrix $\mathbf{R}$ and then find the interference subspace using an eigen-decomposition. The eigenvectors will then define the interference subspace. This actually works for a small number of antennas but it becomes computationally intensive for large arrays of antennas such as the SKA [3]. What is used instead is a subspace tracking approach. This approach is explained in details in [3] but we will not study it here as we will assume in the following sections that the direction of the interferer is known. Another more recent paper that also uses interference subspace projection is [15]. Instead of a subspace tracking approach, a polynomial-based model is used to track changes in the covariance matrix over time and is able to improve the cancellation performance.

3.2 Orthogonal projection method

In this section, we will derive a method for nulling by proposing an orthogonal projection in the simple case where we neglect mutual coupling and suppose the antenna has only one polarization. This section is strongly based on the paper by Cazemier & Ellingson [4] about subspace nulling.

First, we want to to maximize the gain in a certain direction (θ, ϕ). We therefore use the conventional beamformer developed in section 2.4.2 and obtain the following weights, that we will denote $\tilde{\mathbf{w}}$

$$\tilde{\mathbf{w}} = \frac{\mathbf{a}(\theta, \phi)}{\mathbf{a}^H(\theta, \phi)\mathbf{a}(\theta, \phi)} \tag{3.1}$$

We now want to null the array pattern in the directions of the interferers. We assume that those direction are known as previously stated. Imagine that we want to null k interferers from different locations. We can use the interference steering vectors $\mathbf{v}_1 \dots \mathbf{v}_k$ with

$$\mathbf{v}_i = \begin{bmatrix} e^{jk\mathbf{u}(\theta_i,\phi_i)\cdot\mathbf{r}_1} \\ \vdots \\ e^{jk\mathbf{u}(\theta_i,\phi_i)\cdot\mathbf{r}_N} \end{bmatrix} \tag{3.2}$$

to form the interference subspace. The matrix of the interference is then formed by concatenating the k interference steering vectors. $\mathbf{V} = [\mathbf{v}_1 \dots \mathbf{v}_k]$ of size $N \times k$ with N being the number of elements and k being the number of interferers. This way of forming the interference subspace with interference steering vectors from known DOAs guarantees the null to be as deep as possible in the desirable direction.

If the DOAs are not available, two solutions are possible. One solution is to perform an eigendecomposition of $\mathbf{R}$ and the estimated eigenvectors $\mathbf{u}_n$ are used to compute a nullspace projection. The other solution consists of estimating the DOAs with the MUSIC algorithm or a maximum likelihood method and computing the estimated steering vectors with the DOAs found by the algorithm.

Now that we have our interference subspace, we will project the weights found in (3.1) in a subspace orthogonal to the subspace occupied by the interferers, i.e. $\mathbf{V}^\perp$. The projection operator (on the orthogonal complement of $\mathbf{V}$) is defined as

$$\mathbf{P}_{\mathbf{V}}^{\perp} = \mathbf{I} - \mathbf{V}(\mathbf{V}^{\mathbf{H}}\mathbf{V})^{-\mathbf{1}}\mathbf{V}^{\mathbf{H}} \tag{3.3}$$

$$= \mathbf{I} - \mathbf{V}\mathbf{V}^{+} \tag{3.4}$$

where $\mathbf{V}^{+} = (\mathbf{V}^{\mathbf{H}}\mathbf{V})^{-\mathbf{1}}\mathbf{V}^{\mathbf{H}}$ is the MoorePenrose pseudo-inverse[1] of $\mathbf{V}$, which can also be computed with a singular value decomposition (SVD) if $\mathbf{V}^{\mathbf{H}}\mathbf{V}$ is singular. $\mathbf{V}\mathbf{V}^{+}$ is

[1]The Moore-Penrose pseudo-inverse is the generalization of the inverse of a matrix. A complete description of the Moore-Penrose pseudo-inverse, and its relation with least square system is done in Appendix A

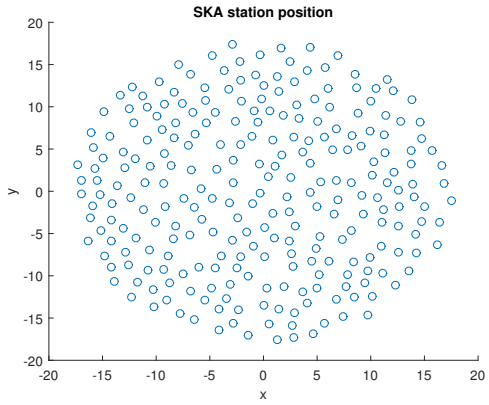


FIGURE 3.1: Position of the antennas on a typical SKA station

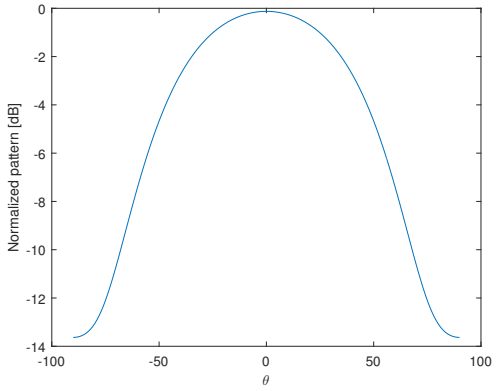


FIGURE 3.2: SKALA element pattern for cut at $\phi = 0$

an orthogonal projection operator on the image $\mathbf{V}$ and $\mathbf{I} - \mathbf{V}\mathbf{V}^{+}$ is the orthogonal projection operator on the image of the orthogonal complementary of $\mathbf{V}$, i.e. the subspace orthogonal to the interferers.

The final weights are given by

$$\mathbf{w} = \mathbf{P}_{\mathbf{V}}^{\perp}\tilde{\mathbf{w}} \tag{3.5}$$

and the array output is given by $y(t) = \mathbf{w}^{H}\mathbf{x}(t)$.

There are two important criteria for this method to be applicable. First, we must have $k \ll N$. This is not a problem for a radio telescope such as the SKA as a typical station would have about 256 antennas and we usually require fewer nulls.

The second condition is that no projection is allowed close to the main beam. Nevertheless, in that case we could use a frequency or time-domain technique to suppress the remaining interferences [4].

3.2.1 Application to the SKA telescope

Let us now show some results of this technique applied to a typical station of 256 elements of the SKA telescope. The disposition of the elements can be found in Fig. 3.1. As can be seen on the figure this is a purely non-regular array geometry.

The elements that compose the station are called SKALA, for SKA Log-periodic antenna. They are log-periodic antennas and have been designed to cover the SKA-low frequency band (50-350MHz). They have a maximum sensitivity in a wide field of view

FIGURE 3.3: SKA Low antenna [14]

(±45° from zenith). The antenna has four arms placed in two geometrically orthogonal planes (see Fig. 3.3). Thus, there are actually two pairs of arms orthogonal to each other. In the following, when we talk about an arm, we will actually mean a pair of arm. The two arms can be interpreted as two independent antennas in one SKALA element. Indeed, those arms are completely independent from each other as we can impose different excitation current (or weights $\mathbf{w}$) on both of them. However, even though the two arms are independent, they need to combine to get the full pattern.

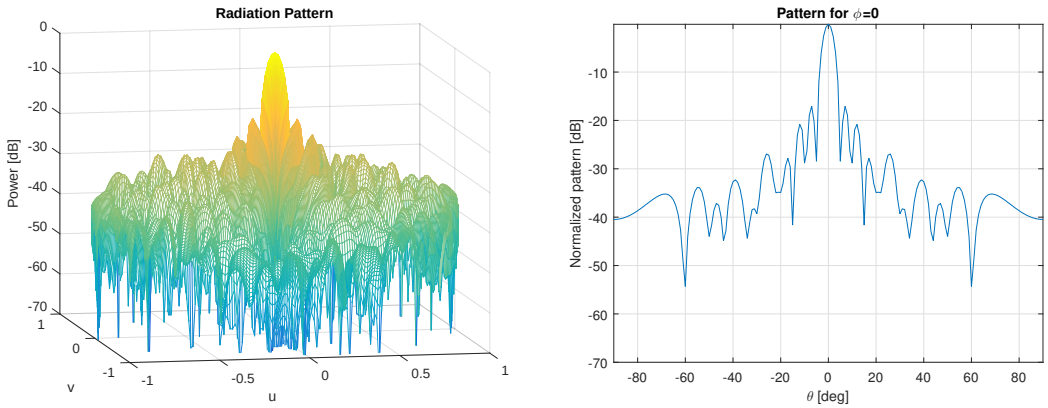
At first, and throughout all this section (unless otherwise stated) we will only consider 1 arm, and thus only one set of weights for each element. However, if one wants to compute the full pattern, the weights on both arms should be computed. More information about this antenna can be found in [14]. A picture of this antenna is shown in Fig. 3.3.

Moreover each arm has two polarizations, a vertical and horizontal polarization (or TE/TM fields). The element pattern (EP) of the antenna (only 1 arm) with two polarizations is thus decomposed as follows :

$$EP = \sqrt{|EP_h|^2 + |EP_v|^2} \tag{3.6}$$

with EP_h and EP_v being the horizontal and vertical polarization respectively.

The projection method presented earlier only works for a single polarization but will be adapted for the case with two polarizations. In a first instance though, we will only



(A) 3D pattern (B) cut along $\phi = 0$

FIGURE 3.4: Radiation pattern of the station for $w_i = 1/N$

work with one polarization (the vertical one), the extension will be straightforward. The radiation pattern of the isolated element in the vertical polarization is shown in Fig. 3.2. As we can see on the graph, the pattern has a half beamwidth (at -3dB) of 45°.

The element pattern of the SKALA antenna is computed using the method of moments (MoM) with macro basis functions (MBFs) with a software developed at UCL called HARP (for HARmonic Polynomial) [16, 17]. The software can compute the EEPs of the elements in the presence of mutual coupling. However, we will not explain the details of this procedure as this is not the purpose of this thesis.

The array pattern of the station is computed by multiplying the element pattern of the antennas (all identical as we suppose there is no coupling) by the array factor for a given frequency of interest as described in equation (2.4). The radiation pattern of the station when all the weights are equal is shown in Fig. 3.4a. A cut along the axis $\phi = 0$, shown in Fig. 3.4b and is provided for better visibility. To obtain the plots, we discretized AP for $\theta = 0 : 90$ with a step equal to 1 and $\phi = 0 : 360$ also with a step equal to 1. This results in $91 \times 360 = 32,760$ directions computed.

Now let us apply the procedure of orthogonal projection described in section 3.2. We will apply this technique to an array of SKALA elements (shown in Fig. 3.1) in the simplest case where we only take into account the vertical polarization and mutual coupling is neglected. The isolated pattern of the element in the vertical polarization have been computed via the HARP software. The computations are done for a frequency $f = 110MHz$.

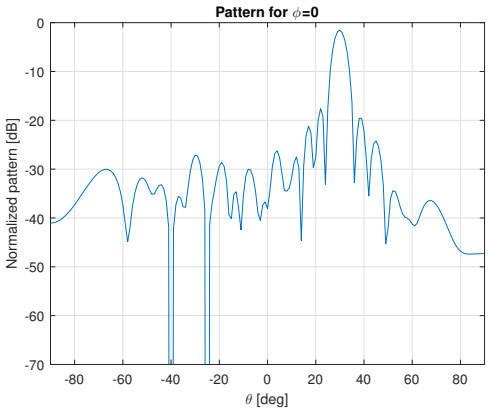


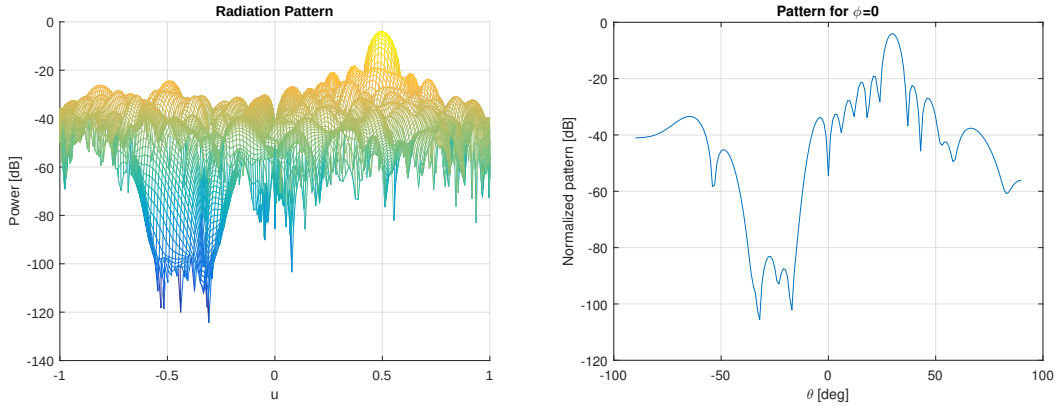
FIGURE 3.5: Array pattern of vertical polarization - no coupling

Let us imagine that we want to steer in the direction $(\theta_0, \phi_0) = (30, 0)$ and null in the directions $(\theta_1, \phi_1) = (25, 180)$ and $(\theta_2, \phi_2) = (40, 180)$. We find the set of weights to apply to each element according to the method of projection so that it matches our desired pattern. The resulting array pattern is given in Fig. 3.5.

As can be seen on Fig. 3.5, we have the two directions of interference that exhibit very deep nulls (-300 dB) and a steering direction at $\theta = 30°$ like we wanted it to be.

We could also null a region instead of just a specific direction. For example, imagine we want to spatially discriminate the region defined by $\theta = [20\ 30]$ and $\phi = [170\ 180]$. One way to deal with this is to force nulls in every directions in that region with a step equal to 1 for both angles, i.e. $(\theta, \phi) = [(20, 170), (21, 170), \ldots, (30, 170), (20, 171), (21, 171), \ldots, (30, 180)]$. This means forcing $k = 121$ nulls in the pattern. Of course, the smaller the step, the better the null region but we have to keep in mind that the number of nulls should always remain smaller than the number of antennas ($k \ll N$). If we null the region we just mentioned with a spatial step equal to 1, we obtain the following graphs 3.6. As can be seen, all the region we want to null is at least below -80 dB. This is still acceptable even though the nulls are not as deep as in the previous example.

If we choose a step equal to 2 for both angles, we obtain a null region in which there is a lot of variation as the nulls are not close enough for the null region to remain stable as it is shown in Fig. 3.7. But the region still remains below -80 dB. However if we continue to increase the step at which we null, we will arrive at a point where only points near the discrete directions only will be deep nulls (around -300 dB) and the other part of the region will not be small enough.



(A) 3D pattern (B) cut along $\phi = 0$

FIGURE 3.6: Nulling region in between [20 170] and [30 180] for a step of 1°.

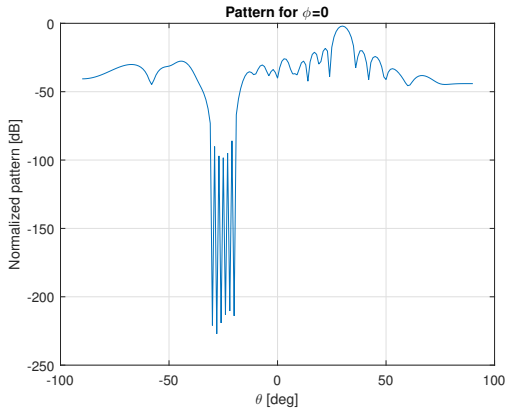


FIGURE 3.7: Nulling region in between [20 170] and [30 180] for a step of 2°.

Nulling a wide region in the pattern by discretizing it with points close enough, so that we can force them to be zero is not very elegant even though it works as long as the nulling region is not too big. Indeed, if we want to null a region of size $20° \times 20°$ with a step of size 1, we need to force 441 nulls while there are only 256 antennas in the station and there are not enough degrees of freedom to achieve this.

A more elegant way of doing this would have been to use a constraint of the type

$$(AP(\theta, \phi))_{\text{dB}} \leq c \quad \forall(\theta, \phi) \in NR \tag{3.7}$$

where NR = null region and c is a constant, for example $c = -70$ dB. However this is not achievable with this type of method but could be done in an optimization problem. This will be discussed in Chapter 4.

3.3 Introduction of mutual coupling

The projection method described in section 3.2 can easily be extended to the case where we take into account the mutual coupling between antennas. This however has not been addressed in the literature to the best of my knowledge.

The idea is to include the embedded element pattern (EEP) of each antenna in the projection matrix $\mathbf{P}_{\mathbf{V}}^{\perp}$. To do that, we multiply every element in the steering vectors (previously defined in eq. 3.2) associated with the direction of the interferer by EEP. Therefore the new steering vector associated with interferer i is defined as

$$\mathbf{v}_i = \begin{bmatrix} EEP_1(\theta_i, \phi_i) \cdot e^{jk\mathbf{u}(\theta_i,\phi_i)\cdot\mathbf{r}_1} \\ \vdots \\ EEP_N(\theta_i, \phi_i) \cdot e^{jk\mathbf{u}(\theta_i,\phi_i)\cdot\mathbf{r}_N} \end{bmatrix} \tag{3.8}$$

The matrix $\mathbf{V}$ is the concatenation of the k interference steering vectors : $\mathbf{V} = [\mathbf{v}_1 \dots \mathbf{v}_k]$. And the projection matrix is defined in the same way as the usual method in equation (3.3).

3.3.1 Application to the SKA telescope

Let us now apply this technique to the SKA station of Fig. 3.1. We still only work with one polarization (the vertical polarization). First of all we will compare the difference when neglecting or integrating mutual coupling between the elements in the station. The comparison of the radiation pattern of the station with and without mutual coupling is shown in Fig. 3.8 for a cut along the angle $\phi = 0$. The embedded pattern of the elements have been computed with HARP.

As we can observe, the main beam is not affected by the coupling but the sidelobes further apart from that region are quite different. However, it is generally in that region further away from the main lobe that we want to cancel out interferences. Therefore we can now understand why it is not suitable use nulling without taking into account mutual coupling.

Let us now apply the technique that we just developed for the same direction than that of the the previous section, i.e. a steering direction $(\theta_0, \phi_0) = (30, 0)$ and two nulls at

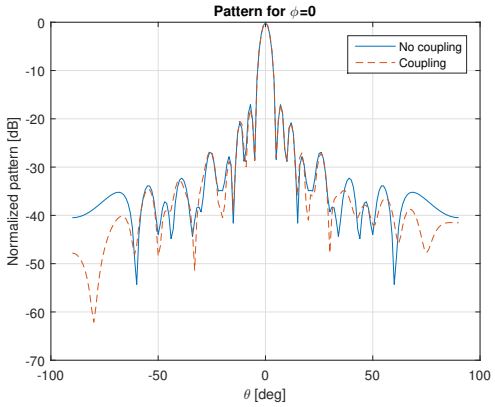


FIGURE 3.8: Comparaison of AP with and without coupling

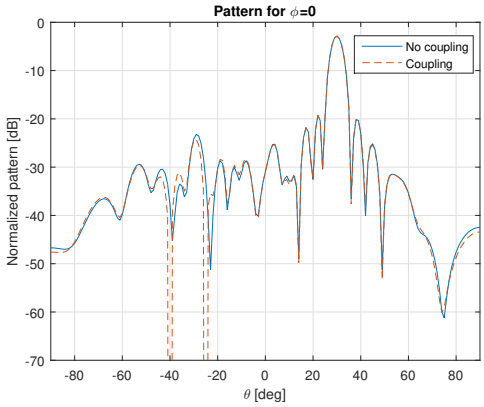


FIGURE 3.9: Comparison between the projection methods with coupling (orange) and without (blue)

$(\theta_1, \phi_1) = (25, 180)$ and $(\theta_2, \phi_2) = (40, 180)$. The comparison between the new technique and the naive orthogonal projection (without taking into account the coupling) is shown in Fig. 3.9.

Now that we have computed the array pattern with the embedded element pattern, we observe that the simple nulling technique does not work, so it is essential to take the EEPs into account in (3.8).

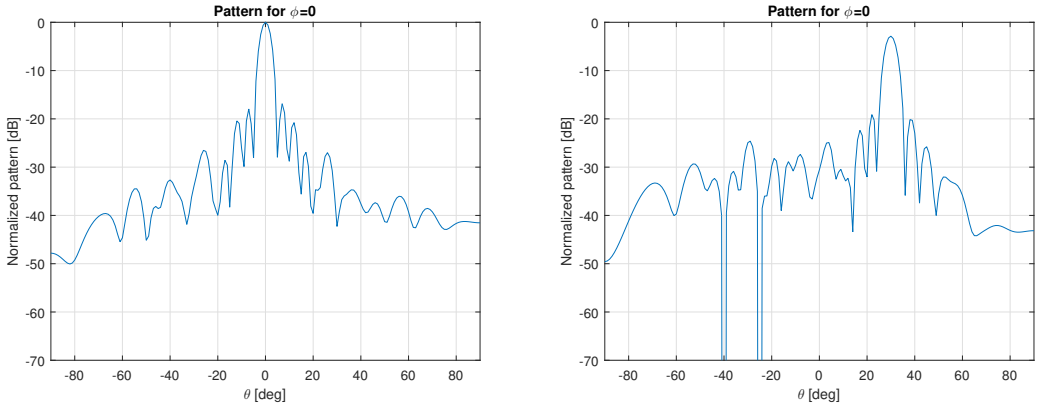
3.4 Two polarization antennas

The concept of sections 3.2 and 3.3 cannot be applied directly to the case where the antenna has two polarizations, i.e. its radiation pattern is given by taking

$$AP = \sqrt{AP_h^2 + AP_h^2}$$

Indeed, in that case the EEPs for the vertical and horizontal polarizations are different. A simple way to deal with this is to null in both polarizations. Instead of having k interference steering vectors, we now have $2k$ steering vectors. For each interferer i, we now have a $N \times 2$ matrix

$$\mathbf{v}_i = \begin{bmatrix} EEP_1^h(\theta_i, \phi_i) \cdot e^{jk\mathbf{u}(\theta_i,\phi_i)\cdot\mathbf{r}_1} & EEP_1^v(\theta_i, \phi_i) \cdot e^{jk\mathbf{u}(\theta_i,\phi_i)\cdot\mathbf{r}_1} \\ \vdots & \\ EEP_N^h(\theta_i, \phi_i) \cdot e^{jk\mathbf{u}(\theta_i,\phi_i)\cdot\mathbf{r}_N} & EEP_N^v(\theta_i, \phi_i) \cdot e^{jk\mathbf{u}(\theta_i,\phi_i)\cdot\mathbf{r}_N} \end{bmatrix} \tag{3.9}$$



(A) Initial pattern (B) Pattern after weights application

FIGURE 3.10: Array pattern for 2 polarization with mutual coupling

where EEP_n^h and EEP_n^v are the embedded element patterns of element n in the horizontal polarization and vertical polarization respectively. The matrix $\mathbf{V}$ is obtained by concatenating the $\mathbf{v}_i$.

3.4.1 Application to the SKA telescope

We will now try this extension on the SKA station where the array pattern takes into account both polarizations and where mutual coupling is included. The array pattern when all the weights are equal is given in Fig. 3.10a and Fig. 3.10b illustrates the array pattern when the weights are applied.

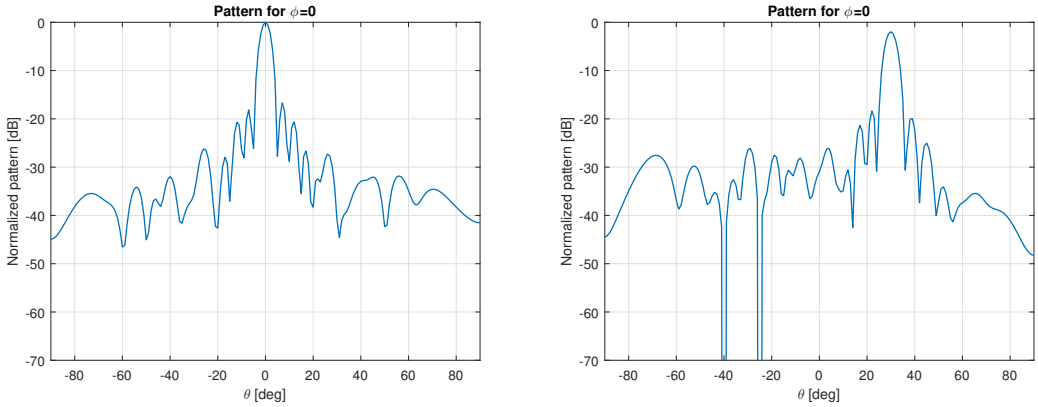
3.5 Taking into account the 2 arms of the SKALA element

Until now, we have only worked with 1 of the arm of the SKA antenna. However, to get the full radiation pattern, we need to compute the weights for both arms. Instead of having a set of N weights, where N is the number of antennas, we now have $2N$ weights.

The weights can be computed independently with the same nulls and main beam direction :

$$\mathbf{w}_1 = \mathbf{P}^{\perp}_{\mathbf{V_1}}\mathbf{a}_0 \tag{3.10}$$

$$\mathbf{w}_2 = \mathbf{P}^{\perp}_{\mathbf{V_2}}\mathbf{a}_0 \tag{3.11}$$



(A) Initial pattern (B) Pattern after weights application

FIGURE 3.11: Full array pattern with two arms

with $\mathbf{a}_0$, the steering vector in the main beam direction and $\mathbf{P_{V_1}}, \mathbf{P_{V_2}}$ the projection matrix on subspace orthogonal to $\mathbf{V}_1$ and $\mathbf{V}_2$. The matrix $\mathbf{V}_1$ is defined as a concatenation of $N \times 2$ matrices $\mathbf{v}_i$ as described in (3.9) with the EEPs of the first arm and matrix $\mathbf{V}_2$ is defined in the same way with the EEPs of the second arm.

The full radiation pattern is then computed as follows

$$AP = \sqrt{|\mathbf{P}_1^H \cdot \mathbf{w_1} + \mathbf{P}_2^H \cdot \mathbf{w}_2|^2 + |\mathbf{P}_1^V \cdot \mathbf{w_1} + \mathbf{P}_2^V \cdot \mathbf{w}_2|^2} \tag{3.12}$$

where $\mathbf{P}_1^H$ stands for embedded element pattern of the first arm in the horizontal polarization at the center (multiplied by the phase factor), $\mathbf{P}_2^H$ stands for the second arm in the horizontal polarization centered and so on.

A zero is thus enforced in both polarization of both arms. The main beams will add up in both arm to get the final pattern within a constant factor.

An application to the SKA telescope is shown on Fig. 3.11 where the nulls and main beam direction is chosen the same as in the previous sections.

3.6 Influence of the noise

We now want to study the effect on the signal to noise ratio (SNR) after applying a set of weights on the antennas where

$$\text{SNR} = \frac{P_s}{P_n} \tag{3.13}$$

P_s being the signal power and P_n the noise power. The signal power P_s (in a given direction $\mathbf{u}_s$) is computed as the time-average of the square of the total field [2] (in the same direction), i.e.

$$P_s(\mathbf{u}_s) = \left\langle \left| \sum_n w_n A P_n(\mathbf{u}_s) \right|^2 \right\rangle \tag{3.14}$$

$$= \left| \sum_n w_n f_n(\mathbf{u}_s) e^{jk\mathbf{u}_s \cdot \mathbf{r}_n} \right|^2 \tag{3.15}$$

If we suppose that the direction we want to compute the SNR is $(\theta, \phi) = (0, 0)$, that the antenna is isotropic and there is no coupling, the signal power is simply

$$P_s = \left| \sum_n w_n \right|^2 \tag{3.16}$$

The noise power is given by the time-average of the square of the sum of the noise voltage on each antennas.

$$P_n = \left\langle \left| \sum_i w_i n_i \right|^2 \right\rangle \tag{3.17}$$

where n_i is the amplitude of the noise on antenna i and $N_i = n_i^2$, used hereunder, is the noise power of antenna i prior to the weight application. We can safely [3] assume that the noise on every antenna is uncorrelated, therefore, we can express P_n as

$$P_n = \sum_i \sum_j w_i w_j^* < n_i n_j^* > \tag{3.18}$$

$$= \sum_i \sum_j w_i w_j^* N_i \delta_{ij} \tag{3.19}$$

$$= \sum_i |N_i w_i|^2 \tag{3.20}$$

If we suppose that $N_i = 1 \; \forall i$, we finally have that the SNR is computed as

$$\text{SNR} = \frac{\left|\sum_n w_n\right|^2}{\sum_n |w_n|^2} \tag{3.21}$$

[2] This directly comes from Poynting theorem where the complex power is given by $\mathbf{S} = \mathbf{E} \times \mathbf{H}^*$ where $\mathbf{E}$ is the electric field and $\mathbf{H}$ the magnetic field. Since we assume that we work with plane wave, the time-average magnitude of the Poynting vector is given by $< S >= \frac{1}{2\eta}|E|^2$ because $H = \frac{E}{\eta}$ for a plane wave, with η the impedance of free space. We will omit the denominator in the computation, since we are computing a ratio of power and it will therefore cancel out.

[3] More precisely, there can be a correlation through the "noise coupling" phenomenon. It makes the global noise level slightly dependent on the direction of observation, but that does not changes the global conclusions.

This expression is maximized if all the weights are equal. In that case, if there are N elements, the maximum we can obtain will be SNR $= N$. Therefore for any weighting scheme, the SNR will be smaller or equal to N. What we want is to remain as close as possible to this value. Of course this is only valid for isotropic antennas and no coupling. Nevertheless, it remains a good indicator of the noise contribution.

3.7 New method for nulling

Let us now investigate a new method and return to the basic problem. We want the pattern to steer in a given direction, i.e. we want it to be maximal in that direction. Now we can impose this maximal value to match any constant C, for example $C = 1$. We also want to have zeros in the pattern in the direction of the interferences. This can be formulated as a linear system of equations (3.22)

$$\underbrace{\begin{bmatrix} \mathbf{P}(\theta_s, \phi_s) \\ \mathbf{P}(\theta_{\text{null}_1}, \phi_{\text{null}_1}) \\ \vdots \\ \mathbf{P}(\theta_{\text{null}_M}, \phi_{\text{null}_M}) \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} w_1 \\ \vdots \\ \vdots \\ w_N \end{bmatrix}}_{\mathbf{w}} = \underbrace{\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{\mathbf{b}} \tag{3.22}$$

where $\mathbf{P}(\theta, \phi) = EEP(\theta, \phi) \cdot e^{jk\mathbf{u}(\theta,\phi)\cdot\mathbf{r}}$. The $\mathbf{A}$ is of size $(M+1) \times N$ with M the number of nulls. There are three cases to distinguish here.

If $M + 1 < N$, there are fewer equations than unknowns, i.e. the number of nulls is smaller than the number of antennas which is most of the time the case. The system of equations is in that case underdetermined. Therefore, there exists an infinite number of solutions that satisfy the system. We have many degrees of freedom here, actually we have $N - k - 1$ degrees of freedom. A question that then naturally arises is which particular solution should we choose from this set of solutions that is the best from an antenna perspective. We would also like a solution that maximizes the SNR derived in the last section. However this imposes solving an optimization problem while we want a direct method computable via matrix operations only. A good alternative to this would be finding the best solution that minimizes the noise power, which is defined by $\|\mathbf{w}\|_2^2$ as shown in (3.21). This is actually solvable via matrix computations thanks to the

Moore-Penrose pseudo-inverse with

$$\mathbf{w} = \mathbf{A}^{+}\mathbf{b} \tag{3.23}$$

Indeed, the pseudo-inverse gives the least square solution with minimal 2-norm $\|\mathbf{w}\|_2$ [18, p. 28]. The pseudo-inverse when $\mathbf{A}$ has less rows than columns ($M+1 < N$) and is full-rank, can be computed with a simple algebraic relation $\mathbf{A}^{+} = \mathbf{A}^{H}(\mathbf{A}\mathbf{A}^{H})^{-1}$. The solution with minimal norm is then given by $\mathbf{w} = \mathbf{A}^{+}\mathbf{b}$. If $\mathbf{A}$ is not full rank, the pseudo-inverse still exists but is not given by the previous formula, it is computed via SVD. The reader may refer to Appendix A for a full reminder of the theory on the Moore-penrose pseudo-inverse and the relation between solving linear systems of equations in a least square sense.

If $M+1 = N$ and $\mathbf{A}$ is full rank, there exists an exact solution and the pseudoinverse is equal to the usual inverse $\mathbf{A}^{+} = \mathbf{A}^{-1}$ and $\mathbf{w} = \mathbf{A}^{-1}\mathbf{b}$. If $\mathbf{A}$ is not full rank however, the solution is not unique and we cannot compute $\mathbf{A}^{-1}$. The pseudoinverse provides the solution with the smallest norm.

If $M+1 > N$, there are more equations than unknowns, i.e. the number of nulls is greater than the number of antennas. The system of equations is in that case overdetermined and no solution exists. We wish to compute the least-squares solution which minimizes $\|\mathbf{A}\mathbf{w} - \mathbf{b}\|_2^2$. The minimal least squares solution can again be computed with the pseudoinverse, i.e. $\mathbf{w} = \mathbf{A}^{+}\mathbf{b}$ where $\mathbf{A}^{+} = (\mathbf{A}^{H}\mathbf{A})^{-1}\mathbf{A}^{H}$ if $\mathbf{A}$ is full rank. Otherwise, it is computed via SVD.

3.8 Comparison of the two methods

The two methods presented look similar but they are actually not the same. In the first method, we project the steering vector (normalized) associated with the direction we want to observe onto the subspace orthogonal to the the interferers; that is we find the vector $\mathbf{w}$ that minimizes the distance from vector $\mathbf{a}_0$ to $\mathbf{V}^{\perp}$. Therefore, the projection method does not necessarily enforce the nulls but finds a solution in the least square sense. It minimizes $\|\mathbf{a}_0 - \mathbf{p}\|$ where $\mathbf{p}$ is the projection of $\mathbf{a}_0$ on $\mathbf{V}^{\perp}$. The advantage of this method is that it keeps the structure of the main beam as similar as possible to the original main beam.

While for the second method, we find a solution with minimal norm that minimizes the distance between $\mathbf{b}$ and $\mathbf{Aw}$, which is the orthogonal projection of $\mathbf{b}$ onto the row space of $\mathbf{A}$ [19, p. 442]. If the system is underdetermined it will always enforce the nulls, which is not necessarily the case for the former method. The drawback of this method is that it does necessarily keep the structure of the original main lobe.

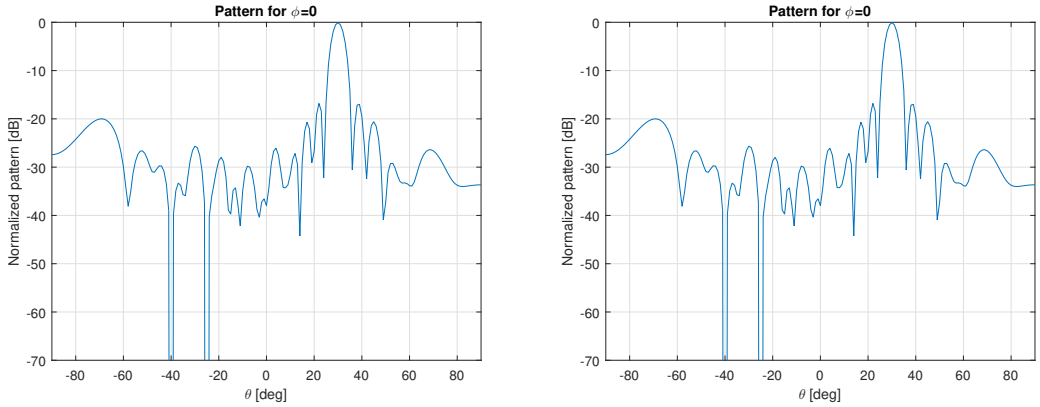
We will first analyze the two methods for the simplest case were all the antennas are isotropic, when there is no mutual coupling and when only 1 polarization is taken into account. An extension that include mutual coupling and 2 polarizations will be undertaken afterwards.

An analytical comparison of the two methods is done in Appendix B. However, this is more of a theoretical exercise than a truly useful comparison. Nevertheless it is interesting to take a look at the similitude between the two methods.

3.8.1 Isotropic case

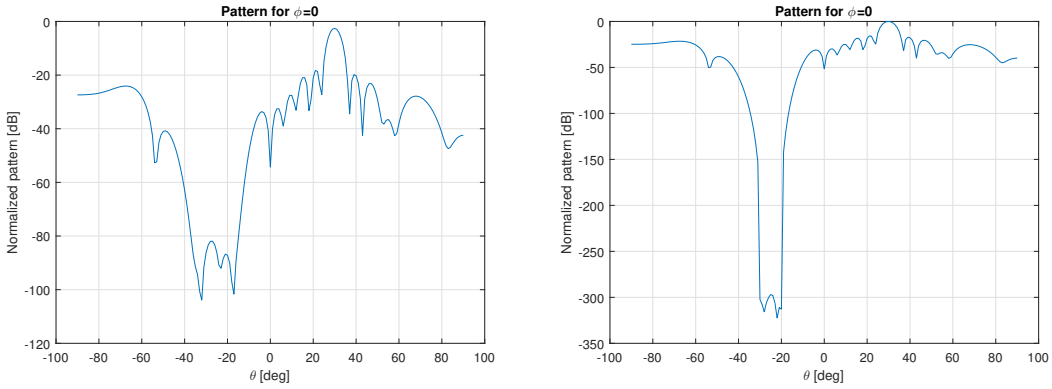
In this section, we suppose that the antennas are isotropic. We will first compare the two methods when only a few nulls are computed. We will keep the same position of interferers as in the previous examples, i.e. a beam at $(\theta_0, \phi_0) = (30, 0)$ and two nulls at $(\theta_1, \phi_1) = (25, 180)$ and $(\theta_1, \phi_1) = (40, 180)$. The results are given in Fig. 3.12 where we can observe that the two pattern are the same. Actually they slighlty differ from each other but it is not visible. To give an idea, the difference in norm is $\|\mathbf{w}_1 - \mathbf{w}_2^*\|_2 = 5.3 \times 10^{-4}$ where $\mathbf{w}_1$ are the weights obtained by the projection method and $\mathbf{w}_2$ the weights obtained by the pseudoinverse method. We try to give an explanation to this in appendix B.

We will now compare the difference when we null a wide region. We will keep the same wide region as in the previous example, i.e. $(\theta, \phi) = [(20, 170), (21, 170), \ldots, (30, 170), (20, 171), (21, 171), \ldots, (30, 180)]$. The results are given in Fig. 3.13 where we can observe that the two patterns (given for a cut for a better visibility) are very different. We can observe that the new method is much better when a greater number of nulls is required. The direction of interferences are true zeros in this case while it is not so the case for the projection method, where the pattern in the direction of interferences simply consists of small values. Moreover, the pattern found with the new method



(A) Array pattern with projection method (B) Array pattern with pseudoinverse method

FIGURE 3.12: Array pattern comparison for 2 nulls



(A) Array pattern with projection method (B) Array pattern with pseudoinverse method

FIGURE 3.13: Array pattern comparison for 141 nulls

scrupulously respects the region of interferences while for the former method the region nulled is wider than expected.

This is actually not strange that the two methods behave quite differently, this is because they are not optimal in the same sense. Let us show some numerical results about different criteria of optimality for those two sets of weights. We define $\mathbf{w}_1$ as the conjugate of the set of weights obtained via projection and $\mathbf{w}_2$, the set of weights obtained via pseudoinverse (i.e. the new method). Four criteria of optimality will be studied : the residual of system (3.22) $\|\mathbf{A}\mathbf{w} - \mathbf{b}\|_2$, the norm of the difference of $\frac{\mathbf{a}_0}{\mathbf{a}_0^H \mathbf{a}_0}$ and its projection on $\mathbf{V}^\perp$, i.e. $\|\frac{\mathbf{a}_0}{\mathbf{a}_0^H \mathbf{a}_0} - \mathbf{w}^*\|_2$, the norm of the weights $\mathbf{w}_2$ and the SNR. The results are shown in Tab. 3.1.

From those results, we can see that the projection method is far from being optimal in the sense of the problem posed in the beginning, i.e. the one satisfying (3.22) and we

	$\|\mathbf{Aw}-\mathbf{b}\|_2$	$\left\|\frac{\mathbf{a}_0}{\mathbf{a}_0^H\mathbf{a}_0}-\mathbf{w}^*\right\|_2$	$\|\mathbf{w}\|_2$	SNR/N
Projection	0.2555	0.0316	0.0539	0.8641
Pseudoinverse	2.2×10^{-14}	0.0366	0.0724	0.8641

TABLE 3.1: Comparison between projection and pseudoinverse methods for 141 nulls

understand why the method does not give the results expected. We could then wonder if a better method could potentially decrease another optimal criterion such as SNR, but we can see from those results that it is not the case which is reassuring.

Let us now turn our attention to the case where there are more interferers than antennas ($k > N$). For example let us null the complete region defined by $\theta = 20:40$ and $\phi = 170:190$ with a step angle of 1, that is $21\times21 = 441$ interferers. Those results are shown in Fig. 3.14 for a cut, to have a better visibility.

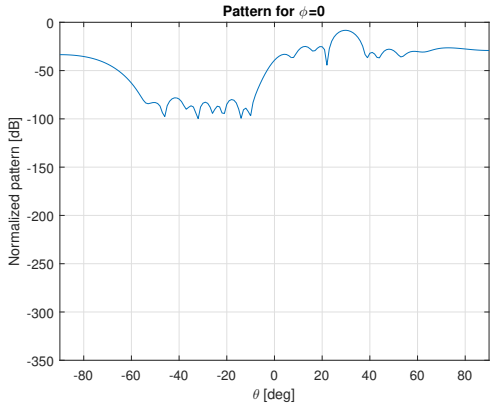
We can see that the pseudoinverse method does a much better job at nulling, with effectively nulling just the region required while the projection method only reaches -100dB and nulls a much wider region than needed (it is even worse when we look at the 3D plot). Moreover, the mainbeam decreases to -8dB. The sidelobe level are quite similar for the two methods however.

As in the previous example, we will look at the numerical results for different optimal criteria in Tab. 3.2. We can observe that the residual of the linear system for the pseudoinverse method is really small and we may wonder how we can have such a small value while we should be solving a least square system since there are more equations than unknowns. This comes from the fact that, in this case $\text{rank}(V) = 189 < N = 256$. Therefore $\text{rank}(A) = 190 < N$ and there exists a solution that satisfies every equations.

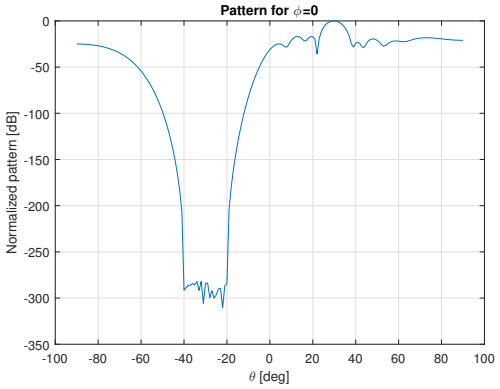
	$\|\mathbf{Aw}-\mathbf{b}\|_2$	$\left\|\frac{\mathbf{a}_0}{\mathbf{a}_0^H\mathbf{a}_0}-\mathbf{w}^*\right\|_2$	$\|\mathbf{w}\|_2$	SNR/N
Projection	0.6146	0.0490	0.0388	0.6099
Pseudoinverse	1.1×10^{-13}	0.0787	0.1005	0.6112

TABLE 3.2: Comparison between projection and pseudoinverse methods for 441 nulls

A last comparison will be done for the case where the number of interferers $k \gg N$ so there exists no exact solution. If we impose a null on a ball of radius 20° along $(\theta, \phi) = (30, 180)$ with a step 1, we impose 1773 null directions, almost 7 times the number of antennas. The results obtained are shown in Fig. 3.15. We can observe that the projection method completely fails. The pseudoinverse method however is able to

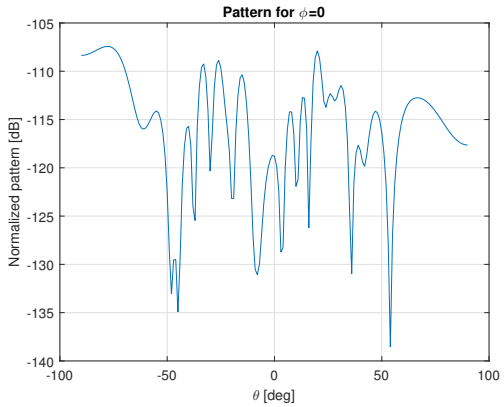


(A) Array pattern with projection method

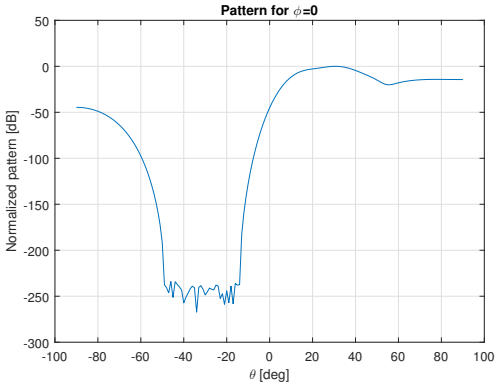


(B) Array pattern with pseudoinverse method

FIGURE 3.14: Array pattern comparison for 441 nulls



(A) Array pattern with projection method



(B) Array pattern with pseudoinverse method

FIGURE 3.15: Array pattern comparison for 1173 nulls

impose the nulls quite efficiently but the main beam becomes too wide even though it is maximal in the desired direction. This is one of the major drawback of this method that we have to keep in mind. We do not impose any constraint on the shape of the main beam with this method which can result in unexpected outcomes.

3.8.2 Extension to mutual coupling and 2 polarizations

The pseudoinverse method can easily be extended to the case where we take into account mutual coupling. In that case the EEPs are simply included in matrix **A** of the array pattern.

To take care of both polarizations however, the extension is not that obvious. We want the contribution of both polarizations to be equal to 1 in the desired direction, that is

$$\sqrt{|\mathbf{P}_V \cdot w|^2 + |\mathbf{P}_H \cdot w|^2} = 1 \tag{3.24}$$

where $\mathbf{P}_V$ and $\mathbf{P}_H$ stand for the embedded patterns at the centers (so multiplied by a phase factor) in the vertical polarization and horizontal polarization respectively. However this is not a linear equation anymore and hence, it cannot be computed using matrix computations. A trick would be to express the array pattern in term of its *co-polarization* and *cross-polarization* instead. Since the SKA antennas are designed to have low cross polarization [14], expressing the pattern to be equal to the co-polarization in the main beam region is a good approximation. The co and cross polarizations are expressed as follows [4]

$$\mathbf{P}_{\text{co}} = \mathbf{P}_V \cos(\phi + \alpha) - \mathbf{P}_H \sin(\phi + \alpha) \tag{3.25}$$

$$\mathbf{P}_{\text{cross}} = -\mathbf{P}_V \sin(\phi + \alpha) - \mathbf{P}_H \cos(\phi + \alpha) \tag{3.26}$$

where $\phi = [0\ 360]$ is the azimuth angle, α is the rotation angle where the combination TE/TM have the maximum co-polarization and minimun cross polarization and $\mathbf{P}$ stands for the EEPs at the center. Equation (3.24) can therefore be approximated by

$$\mathbf{P}_{\text{co}}(\theta_0, \phi_0) \cdot w = 1 \tag{3.27}$$

To give an idea about the goodness of the approximation, we compute

$$|\mathbf{P}_{\text{co}}(\theta_0, \phi_0)| - \sqrt{|\mathbf{P}_V(\theta_0, \phi_0)|^2 + |\mathbf{P}_H(\theta_0, \phi_0)|^2} = 2.7746 \times 10^{-5}$$

And we can see this is indeed a good approximation.

[4] The formula for co/cross-polarization are obtained by a projection of the vertical and horizontal polarizations on the co/cross-plane.

Chapter 4

Convex optimization models

In the previous chapter, we worked with direct methods based on matrix computations. The advantage of such methods is that they are fast and do not involve an iterative algorithm. The main drawback of those methods though is that there is no constraint on the region of the pattern external to the null region and the main beam. However, we would like to limit as much as possible the sidelobes everywhere and have some kind of constraints on the width of the main beam, while at the same time imposing the pattern in the interference directions to be zero. We will see that this can be formulated as an optimization problem. Moreover, the formulation of the problem can be made convex and therefore be solved efficiently using well-known interior point methods.

Pattern synthesis via convex optimization has been studied in the literature for quite some time now. The first work related to this problem goes back to the late 90's by Lebret and Boyd in [5, 6] where they derive a model to minimize the beam pattern level over a given region for linear arrays of antennas. Other models followed but most of them are based on adaptive beamforming, rather than deterministic beamforming, where the data statistics are required. They usually consists in some kind of variation of the minimum variance beamforming. A recent survey about those convex optimization-based adaptive methods can be found in [20]. More recent papers about deterministic beamforming tried to optimize the weights to meet a desired pattern as well as the position of the antennas but only for specific types of regular array geometries and without truly minimizing the sidelobes level [21, 22]. But all those methods found in the literature do not take into account mutual coupling, to the best of my knowledge.

In summary the problem that we are interested in is an extension of the model presented by Lebret and Boyd in [5] for non-regular arrays in the presence of mutual coupling and where we also include nulling.

4.1 Problem formulation

Let us first define some notations. Let

- Ω = set of all angles of the semi-sphere defined by the spherical coordinates $\theta = [0\ 90]$ and $\phi = [0\ 360]$
- $\mathcal{B}$ = set of angles which define the main beam. For example, if we steer in a direction (θ_0, ϕ_0) we will define a ball of fixed size around that direction and all angles in that ball define the region $\mathcal{B}$
- $\mathcal{N}$ = set of angles where we want to null the interferences. This can either be defined by several discrete directions or by balls around specific directions.
- $\mathbf{P}(\theta, \phi)$ is a row vector of size $1 \times N$ that defines the radiation pattern of the array prior to the weights summation, for the direction (θ, ϕ), i.e. the element pattern multiplied by the phase factor.

We want to minimize the level of the sidelobes subject to a specific main beam direction and several null directions. This can be formulated as the following optimization problem:

$$\begin{aligned} \min_{\mathbf{w}} \max_{(\theta,\phi)\in\Omega\setminus\mathcal{B}} \quad & |\mathbf{P}(\theta,\phi)\cdot\mathbf{w}| \\ \text{s.t.} \quad & \mathbf{P}(\theta_0,\phi_0)\cdot\mathbf{w} = 1 \\ & |\mathbf{P}(\theta,\phi)\cdot\mathbf{w}| = 0,\ \forall(\theta,\phi)\in\mathcal{N} \end{aligned} \tag{4.1}$$

We minimize the maximal sidelobe on the region external to the main beam while constraining the beampattern to unity[1] at the steering direction and enforcing zeros in the pattern at the null directions. Note that this last constraint can be relaxed to be

[1] This could actually be any constant since we are not interested in the magnitude of the pattern but rather in its shape.

smaller than a given maximum sidelobe level, hence we could express this last constraint as

$$|\mathbf{P}(\theta,\phi)\cdot\mathbf{w}|\leq \mathrm{NL},\ \forall(\theta,\phi)\in\mathcal{N} \tag{4.2}$$

where NL is the desired nulling level.

The objective function involves a maximum over an infinite set of parameters, we therefore chose to discretize[2] it with a small enough step so that the approximation is good. In our example, we chose a step of 1° for both the θ and ϕ angles, resulting in a discretization of $\mathrm{NA} = 91 \times 360 = 32,760$ points. A step of 1° will provide good accuracy but could lead to long computation time for real time application, some trade-off between accuracy and speed should be chosen by the engineer. Throughout this chapter, we will maintain a step of 1° though. The matrix

$$\mathbf{P} = \begin{bmatrix} f_1(\mathbf{u}_1)e^{(jk\mathbf{u}_1\cdot\mathbf{r}_1)} & \dots & f_N(\mathbf{u}_1)e^{(jk\mathbf{u}_1\cdot\mathbf{r}_N)} \\ \vdots & & \vdots \\ f_1(\mathbf{u}_{\mathrm{NA}})e^{(jk\mathbf{u}_{\mathrm{NA}}\cdot\mathbf{r}_1)} & \dots & f_N(\mathbf{u}_{\mathrm{NA}})e^{(jk\mathbf{u}_{\mathrm{NA}}\cdot\mathbf{r}_N)} \end{bmatrix}$$

is therefore of size $\mathrm{NA}\times N$.

The optimization problem defined in (4.1) assumes that the antenna is defined by only one component (vertical or horizontal component), while the SKALA antenna has two : $\mathbf{P}_V$ and $\mathbf{P}_H$. Therefore, when we take into account the contribution from those two components, we get a new optimization problem in (4.3) which is already a bit more complicated.

$$\begin{aligned} \min_{\mathbf{w}} \max_{(\theta,\phi)\in\Omega\setminus\mathcal{B}} \quad & \sqrt{|\mathbf{P}_V(\theta,\phi)\cdot\mathbf{w}|^2+|\mathbf{P}_H(\theta,\phi)\cdot\mathbf{w}|^2} \\ \text{s.t.} \quad & \sqrt{|\mathbf{P}_V(\theta_0,\phi_0)\cdot\mathbf{w}|^2+|\mathbf{P}_H(\theta_0,\phi_0)\cdot\mathbf{w}|^2}=1 \\ & \sqrt{|\mathbf{P}_V(\theta,\phi)\cdot\mathbf{w}|^2+|\mathbf{P}_H(\theta,\phi)\cdot\mathbf{w}|^2}=0,\ \forall(\theta,\phi)\in\mathcal{N} \end{aligned} \tag{4.3}$$

A last problem remains with this formulation. The main beam region $\mathcal{B}$ is not constrained by any equation except the fact that $AP(\theta_0,\phi_0)=1$ in the steering direction. This mean that the pattern could vary a lot, there could be sidelobes greater than the sidelobe level outside $\mathcal{B}$, the pattern could even go over 1. We therefore must add some

[2] In some cases this could be handled exactly using continuous optimization, but we chose here the simpler approach of discretization.

kind of constraints in this region to avoid those effects. This will be the subject of section 4.1.2 as some concepts about convexity need to be posed first.

4.1.1 Convexity & second order cone programming (SOCP)

For an optimization model to be solved efficiently, a very common technique is to make it convex. Indeed, in that case there exists well-known interior point algorithms that can solve any convex problems in polynomial time [23]. Moreover, in convex optimization every local minimum is a global minimum. Let us investigate the convexity[3] of the problem formulated in (4.3).

A model is convex if (1) The objective function is convex and (2) The set X defined by the constraints is convex. First let us use a simple trick to make the objective function linear. We can rewrite the objective function

$$\min_{\mathbf{w}} \max_{(\theta,\phi)\in\Omega\setminus\mathcal{B}} \sqrt{|\mathbf{P}_V(\theta,\phi)\cdot\mathbf{w}|^2 + |\mathbf{P}_H(\theta,\phi)\cdot\mathbf{w}|^2}$$

as

$$\begin{aligned} \min_{\mathbf{w}} \quad & t \\ \text{s.t.} \quad & \sqrt{|\mathbf{P}_V(\theta,\phi)\cdot\mathbf{w}|^2 + |\mathbf{P}_H(\theta,\phi)\cdot\mathbf{w}|^2} \leq t,\ \forall(\theta,\phi)\in\Omega\setminus\mathcal{B} \end{aligned} \tag{4.4}$$

which makes the objective function linear (and hence convex) but adds up new constraints. The number of new constraints added is equal to $\dim(\Omega \setminus \mathcal{B})$. Let us now investigate the convexity of the set of constraints. A set X is convex iff

$$x, y \in X \Rightarrow \lambda x + (1-\lambda y) \in X,\ 0 \leq \lambda \leq 1$$

However this is quite hard to show in practice. But since the model looks like a second order cone program (SOCP), we will use its broad theory to show that the model is convex. Let us first give some definitions regarding SOCP.

Definition 4.1. *Second order cone program* refers to an optimization model where a linear function is minimized over the intersection of an affine set and the product of second order (quadratic) cones [7].

[3] The analysis in this section is based the course of convex optimization (LINMA2471) given by Pr. Glineur at UCL [24]

Definition 4.2. A *convex cone* $\mathcal{K}$ is a special type of convex set that is closed under positive scaling, i.e. $\alpha x \in \mathcal{K}\ \forall x \in \mathcal{K}$ and $\forall \alpha \geq 0$.

Definition 4.3. A *second-order cone* is a type of convex cone and is defined as $\mathcal{Q}^n = \left\{(t, x) \in \mathbb{R}^{n+1} \mid t \geq \sqrt{x_1^2 + \ldots + x_n^2}\right\}$.

Definition 4.4. For any convex cone $\mathcal{K}$, we can define a *dual cone* $\mathcal{K}^* = \{x | \langle x, y\rangle \geq 0,\ \forall y \in \mathcal{K}\}$ where $\langle \cdot, \cdot \rangle$ denotes the inner product.

Definition 4.5. A *self-dual cone* is a cone that satisfies $\mathcal{K} = \mathcal{K}^*$.

It can be shown that the second-order cone is self-dual. Thanks to this property, it is possible to design efficient algorithm for solving a SOCP based on the primal-dual interior point method.

Since second-order cones (SOC) are defined for real variables, we have to split the complex valued variables and parameters into their real and imaginary parts. We will now denote $\mathbf{w}^r = \mathrm{Re}\{\mathbf{w}\}$, $\mathbf{w}^{im} = \mathrm{Im}\{\mathbf{w}\}$, $\mathbf{P}_V^r = \mathrm{Re}\{\mathbf{P}_V\}$, $\mathbf{P}_V^{im} = \mathrm{Im}\{\mathbf{P}_V\}$, $\mathbf{P}_H^r = \mathrm{Re}\{\mathbf{P}_H\}$ and $\mathbf{P}_H^{im} = \mathrm{Im}\{\mathbf{P}_H\}$. The general pattern expression

$$AP(\theta, \phi) = \sqrt{|\mathbf{P}_V(\theta, \phi) \cdot \mathbf{w}|^2 + |\mathbf{P}_H(\theta, \phi) \cdot \mathbf{w}|^2} \tag{4.5}$$

is then replaced by its real and imaginary parts as follows. First, we have that

$$\begin{aligned} |\mathbf{P} \cdot \mathbf{w}|^2 &= \left|\left(\mathbf{P}^r + j\mathbf{P}^{im}\right) \cdot \left(\mathbf{w}^r + j\mathbf{w}^{im}\right)\right|^2 && (4.6)\\ &= \left|\left(\mathbf{P}^r \cdot \mathbf{w}^r - \mathbf{P}^{im} \cdot \mathbf{w}^{im}\right) + j\left(\mathbf{P}^r \cdot \mathbf{w}^{im} + \mathbf{P}^{im} \cdot \mathbf{w}^r\right)\right|^2 && (4.7)\\ &= \left(\mathbf{P}^r \cdot \mathbf{w}^r - \mathbf{P}^{im} \cdot \mathbf{w}^{im}\right)^2 + \left(\mathbf{P}^r \cdot \mathbf{w}^{im} + \mathbf{P}^{im} \cdot \mathbf{w}^r\right)^2 && (4.8) \end{aligned}$$

where the (θ, ϕ) dependence has been omitted for clarity. By substituting (4.8) into (4.5), we obtain

$$AP(\theta, \phi) = \sqrt{\begin{aligned}&(\mathbf{P}_V^r \cdot \mathbf{w}^r - \mathbf{P}_V^{im} \cdot \mathbf{w}^{im})^2 + \left(\mathbf{P}_V^r \cdot \mathbf{w}^{im} + \mathbf{P}_V^{im} \cdot \mathbf{w}^r\right)^2 \\ &+ \left(\mathbf{P}_H^r \cdot \mathbf{w}^r - \mathbf{P}_H^{im} \cdot \mathbf{w}^{im}\right)^2 + \left(\mathbf{P}_H^r \cdot \mathbf{w}^{im} + \mathbf{P}_H^{im} \cdot \mathbf{w}^r\right)^2\end{aligned}} \tag{4.9}$$

We now define four new variables for each pair (θ, ϕ) :

$$
\begin{aligned}
Z_{V1}^{\theta,\phi} &= \mathbf{P}_V^r \cdot \mathbf{w}^r - \mathbf{P}_V^{im} \cdot \mathbf{w}^{im} && (4.10)\\
Z_{V2}^{\theta,\phi} &= \mathbf{P}_V^r \cdot \mathbf{w}^{im} + \mathbf{P}_V^{im} \cdot \mathbf{w}^r && (4.11)\\
Z_{H1}^{\theta,\phi} &= \mathbf{P}_H^r \cdot \mathbf{w}^r - \mathbf{P}_H^{im} \cdot \mathbf{w}^{im} && (4.12)\\
Z_{H2}^{\theta,\phi} &= \mathbf{P}_H^r \cdot \mathbf{w}^{im} + \mathbf{P}_H^{im} \cdot \mathbf{w}^r && (4.13)
\end{aligned}
$$

The problem can then finally be formulated as a SOCP :

$$
\begin{aligned}
&\min_{\mathbf{w}} \quad t \\
&\text{s.t.} \quad \sqrt{(Z_{V1}^{\theta,\phi})^2 + (Z_{V2}^{\theta,\phi})^2 + (Z_{H1}^{\theta,\phi})^2 + (Z_{H2}^{\theta,\phi})^2} \leq t, \ \forall (\theta, \phi) \in \Omega \setminus \mathcal{B}
\end{aligned} \tag{4.14}
$$

where the variables Z are defined by the linear equations (4.10 - 4.13) and this is convex by the definition of a second order cone. We still have to check if the last two constraints of our initial model (4.1) are convex.

The null constraint $AP(\theta, \phi) = 0, \ \forall (\theta, \phi) \in \mathcal{N}$ is convex because it is equivalent to imposing the imaginary and real part to be zero for those directions, i.e. $Z_{V1}^{\theta,\phi} = Z_{V2}^{\theta,\phi} = Z_{H1}^{\theta,\phi} = Z_{H2}^{\theta,\phi} = 0$ and this is a linear equality constraint. If we choose the relaxation of the null constraint (4.2), i.e. imposing the pattern to be smaller than a fixed value, this would also be convex since it is then a second order cone where t is a constant and equal to some small number NL.

For the steering direction constraint $AP(\theta_0, \phi_0) = 1$, the situation obvious. Actually this constraint is not convex since we impose an equality on a quadratic function. Thus, we need to find an approximation that is convex. We can reuse the same trick as in Chap. 3 by imposing the real part of the copolarization pattern to be equal to 1 at the steering direction and the imaginary part to be equal to 0:

$$
\begin{aligned}
\mathrm{Re}\{\mathbf{P}_{\mathrm{co}}(\theta_0, \phi_0) \cdot \mathbf{w}\} &= \mathbf{P}_{\mathrm{co}}^r(\theta_0, \phi_0) \cdot \mathbf{w}^r - \mathbf{P}_{\mathrm{co}}^{im}(\theta_0, \phi_0) \cdot \mathbf{w}^{im} = 1 && (4.15)\\
\mathrm{Im}\{\mathbf{P}_{\mathrm{co}}(\theta_0, \phi_0) \cdot \mathbf{w}\} &= \mathbf{P}_{\mathrm{co}}^r(\theta_0, \phi_0) \cdot \mathbf{w}^{im} + \mathbf{P}_{\mathrm{co}}^{im}(\theta_0, \phi_0) \cdot \mathbf{w}^r = 0 && (4.16)
\end{aligned}
$$

where $\mathbf{P}_{\mathrm{co}}^r = \mathrm{Re}\{\mathbf{P}_{\mathrm{co}}\}$ and $\mathbf{P}_{\mathrm{co}}^{im} = \mathrm{Im}\{\mathbf{P}_{\mathrm{co}}\}$. This constraint is now linear and hence convex. This approximation is appropriate because the SKA antenna have been designed to have a really low cross polarization [14].

4.1.2 Main beam constraints

Now that we have a convex problem, we can go back to the question of the main beam shape. The only constraint that we have for now is that the pattern must be equal to 1 in the direction of the steering direction. We saw that this was not a convex constraint and we used the approximation of the co-polarization of the pattern that must be equal to 1 in the steering direction. But the pattern in every other direction is not constrained and can therefore vary a lot. The other thing that we have to fix is the width of that region. We would like this width to match with the size of the actual main beam (not just the region).

We could think of two ways to add new constraints to this region. The first is to impose that the gradient in the region be negative throughout the region so that no secondary lobe can coexist within the region. However we will see the gradient is quite painful to work with and is also not convex (in its components). Another idea is to impose a window within which the beam must remain. We will see that this solution is feasible and this is the solution we choose to work with.

First, let us take a look at the gradient method proposed. To impose a negative slope for the 3D beam throughout all the region, we must impose

$$\hat{n}_u \cdot \nabla_{\mathbf{u}} AP(\mathbf{u}, \mathbf{v}) + \hat{n}_{\mathbf{v}} \cdot \nabla_{\mathbf{v}} AP(\mathbf{u}, \mathbf{v}) < 0$$

Remember that the array pattern is defined in (4.9). For simplicity, we will work with the square of the pattern, so we get rid of the square root. AP is then just a sum of quadratic terms. But those quadratic terms are themselves a sum of sines and cosines of the variables of interests u, v. Indeed, remember that the phase factor of an element is defined as $\exp(jk(ux_i + vy_i))$ in the array pattern. Taking their real and imaginary parts (for the expression in $\mathbf{P}^r$,$\mathbf{P}^{im}$) results in a sum of sine and cosine functions. After differentiation, this gives a long and painful sum of product of sine and cosine that we are unable to work with and that is clearly not convex.

Our second idea was to impose a window within which the main beam must remain. An example of the type of window is shown in Fig. 4.1. Note that the actual window is in 3D, so it actually has a cylindrical shape. We can impose an upper bound in the pattern, so that it remains smaller than 1 for a region of width HB around (θ_0, ϕ_0) and smaller

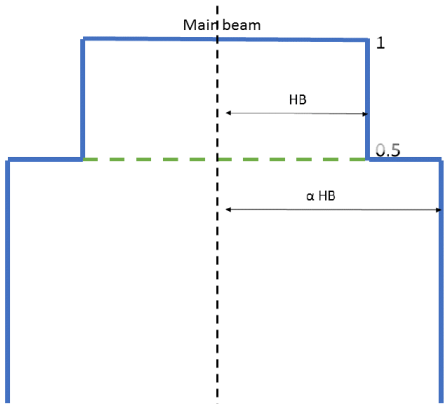


FIGURE 4.1: Main beam window

than 0.5 for a region starting at HB and finishing at $\alpha \cdot HB$ where $\alpha > 1$. These are of course some arbitrary value and they can be adapted by the engineer. Additional step functions may be included in order the shape that we want. Upper bounds are convex since they can be expressed as SOC like the other constraints. The upper bounds have the role of limiting the size of the sidelobes that could be generated in the main beam region; however the main beam can still have arbitrary width (of course smaller than the actual size $\mathcal{B}$) within the region. We would also like to impose a lower bound constraint for the main beam so that we would be able to control the true width of the beam. We could impose the pattern level to be greater than 0.5 (half power) in a ball of size HB centered at (θ_0, ϕ_0) (doted green line on the figure). The bad news is that a lower bound constraint is concave. However, we can use the same approximation as previously with the copolarization since we are close to the steering direction. We will therefore impose

$$\mathrm{Re}\{\mathbf{P}_{\mathrm{co}}(\theta, \phi) \cdot \mathbf{w}\} \geq 0.5, \ \forall (\theta, \phi) \in B[(\theta_0, \phi_0), \mathrm{HB}]$$

where $B[(\theta_0, \phi_0), \mathrm{HB}]$ stands for a ball centered at (θ_0, ϕ_0) of radius HB.

Finally the constraints imposed in the main beam region are the following

$$\mathrm{Re}\{\mathbf{P}_{\mathrm{co}}(\theta_0,\phi_0)\cdot\mathbf{w}\}=1 \tag{4.17}$$

$$\mathrm{Im}\{\mathbf{P}_{\mathrm{co}}(\theta_0,\phi_0)\cdot\mathbf{w}\}=0 \tag{4.18}$$

$$\sqrt{(Z_{V1}^{\theta,\phi})^2+(Z_{V2}^{\theta,\phi})^2+(Z_{H1}^{\theta,\phi})^2+(Z_{H2}^{\theta,\phi})^2}\leq 1,\ \forall(\theta,\phi)\in B[(\theta_0,\phi_0),\mathrm{HB}] \tag{4.19}$$

$$\sqrt{(Z_{V1}^{\theta,\phi})^2+(Z_{V2}^{\theta,\phi})^2+(Z_{H1}^{\theta,\phi})^2+(Z_{H2}^{\theta,\phi})^2}\leq 0.5,\ \forall(\theta,\phi)\in \mathcal{B}\setminus B[(\theta_0,\phi_0),\mathrm{HB}] \tag{4.20}$$

$$\mathrm{Re}\{\mathbf{P}_{\mathrm{co}}(\theta,\phi)\cdot\mathbf{w}\}\geq 0.5,\forall(\theta,\phi)\in B[(\theta_0,\phi_0),\mathrm{HB}] \tag{4.21}$$

where

$$\mathrm{Re}\{\mathbf{P}_{\mathrm{co}}\cdot\mathbf{w}\} = \mathbf{P}_{co}^{r}\cdot\mathbf{w}^{r}-\mathbf{P}_{co}^{im}\cdot\mathbf{w}^{im} \tag{4.22}$$

$$\mathrm{Im}\{\mathbf{P}_{\mathrm{co}}\} = \mathbf{P}_{co}^{r}\cdot\mathbf{w}^{im}+\mathbf{P}_{co}^{im}\cdot\mathbf{w}^{r} \tag{4.23}$$

4.1.3 Weights constraints

We would like to add a constraint on the weights so that the SNR remains as high as possible. We know that the SNR is maximized when all the weights are equal. Thus we would like the weights to remain within a tight interval by imposing a lower bound and an upper bound on the absolute value of the weights. However only the upper bound is convex. Indeed imposing an upper bound on each complex weights w_i defines a second order cone :

$$|w_i|=\sqrt{(w_i^r)^2+(w_i^{im})^2}\leq C,\ \forall i=1,\ldots,N \tag{4.24}$$

where C is a constant, while imposing a lower bound on the w_i makes the cone concave. But even only imposing an upper-bound will help to increase the SNR as we will see in section (4.2).

4.1.4 Final model

The final formulation of the problem is given below

$$
\begin{array}{rll}
\min\limits_{\mathbf{w}} & t & \text{subject to} \\
\text{(Sidelobe 1)} & \sqrt{(Z_{V1}^{\theta,\phi})^2 + (Z_{V2}^{\theta,\phi})^2 + (Z_{H1}^{\theta,\phi})^2 + (Z_{H2}^{\theta,\phi})^2} \leq t, & \forall (\theta,\phi) \in \Omega \setminus \mathcal{B} \\
\text{(Sidelobe 2)} & Z_{V1}^{\theta,\phi} = \mathbf{P}_V^r(\theta,\phi) \cdot \mathbf{w}^r - \mathbf{P}_V^{im}(\theta,\phi) \cdot \mathbf{w}^{im}, & \forall (\theta,\phi) \in \Omega \\
\text{(Sidelobe 3)} & Z_{V2}^{\theta,\phi} = \mathbf{P}_V^r(\theta,\phi) \cdot \mathbf{w}^{im} + \mathbf{P}_V^{im}(\theta,\phi) \cdot \mathbf{w}^r, & \forall (\theta,\phi) \in \Omega \\
\text{(Sidelobe 4)} & Z_{H1}^{\theta,\phi} = \mathbf{P}_H^r(\theta,\phi) \cdot \mathbf{w}^r - \mathbf{P}_H^{im}(\theta,\phi) \cdot \mathbf{w}^{im}, & \forall (\theta,\phi) \in \Omega \\
\text{(Sidelobe 5)} & Z_{H2}^{\theta,\phi} = \mathbf{P}_H^r(\theta,\phi) \cdot \mathbf{w}^{im} + \mathbf{P}_H^{im}(\theta,\phi) \cdot \mathbf{w}^r, & \forall (\theta,\phi) \in \Omega \\
\text{(Null 1)} & Z_{V1}^{\theta,\phi} = 0, & \forall (\theta,\phi) \in \mathcal{N} \\
\text{(Null 2)} & Z_{V2}^{\theta,\phi} = 0, & \forall (\theta,\phi) \in \mathcal{N} \\
\text{(Null 3)} & Z_{H1}^{\theta,\phi} = 0, & \forall (\theta,\phi) \in \mathcal{N} \\
\text{(Null 4)} & Z_{H2}^{\theta,\phi} = 0, & \forall (\theta,\phi) \in \mathcal{N} \\
\text{(Null relaxed)} & \sqrt{(Z_{V1}^{\theta,\phi})^2 + (Z_{V2}^{\theta,\phi})^2 + (Z_{H1}^{\theta,\phi})^2 + (Z_{H2}^{\theta,\phi})^2} \leq \text{NL}, & \forall (\theta,\phi) \in \mathcal{N} \\
\text{(Steer 1)} & \mathbf{P}_{co}^r(\theta_0,\phi_0) \cdot \mathbf{w}^r - \mathbf{P}_{co}^{im}(\theta_0,\phi_0) \cdot \mathbf{w}^{im} = 1 & \\
\text{(Steer 2)} & \mathbf{P}_{co}^r(\theta_0,\phi_0) \cdot \mathbf{w}^{im} + \mathbf{P}_{co}^{im}(\theta_0,\phi_0) \cdot \mathbf{w}^r = 0 & \\
\text{(Beam LB1)} & \sqrt{(Z_{V1}^{\theta,\phi})^2 + (Z_{V2}^{\theta,\phi})^2 + (Z_{H1}^{\theta,\phi})^2 + (Z_{H2}^{\theta,\phi})^2} \leq 1, & \forall (\theta,\phi) \in B[(\theta_0,\phi_0), \text{HB}] \\
\text{(Beam LB2)} & \sqrt{(Z_{V1}^{\theta,\phi})^2 + (Z_{V2}^{\theta,\phi})^2 + (Z_{H1}^{\theta,\phi})^2 + (Z_{H2}^{\theta,\phi})^2} \leq 0.5, & \forall (\theta,\phi) \in \mathcal{B} \setminus B[(\theta_0,\phi_0), \text{HB}] \\
\text{(Beam UB)} & \mathbf{P}_{co}^r \cdot \mathbf{w}^r - \mathbf{P}_{co}^{im} \cdot \mathbf{w}^{im} \geq 0.5, & \forall (\theta,\phi) \in B[(\theta_0,\phi_0), \text{HB}] \\
\text{(Weights UB)} & \sqrt{(w_i^r)^2 + (w_i^{im})^2} \leq C & \forall i = 1, \ldots, N
\end{array}
$$

For the null constraints, we either impose (Null 1-4) or (Null relaxed). Of course the feasibility of (Null 1-4) depends on the number of elements, i.e. the size of the weight vector. Therefore, we will usually work with the relaxed constraint.

4.2 Application to the SKA telescope

Let us now apply our optimization model to the SKA telescope. We will compute the radiation pattern of the same station as before with the embedded element pattern calculated via the HARP software [16] developed at UCL. Several parameters need to be fixed : the null level (NL), the half beamwidth (HB), and the upper bound of the weight C.

For our first test, we will choose a null level of -60dB. The half beamwidth, which is defined as the radius of the main lobe at half power (-6dB[4] in the power pattern), will be chosen to match the one when all weights are unity ($w_i = 1\ \forall i$), i.e. HB = 3°.

The upper bound on the weight will be chosen according to

$$\sqrt{(w_i^r)^2 + (w_i^{im})^2} \leq Q \cdot M \quad \forall i = 1, \ldots, N \tag{4.25}$$

where $M = \max_{\theta,\phi} \sqrt{|\mathbf{P}_H \cdot \mathbf{1}|^2 + |\mathbf{P}_V \cdot \mathbf{1}|^2}$ is the maximum value of the array pattern when all the weights are chosen equal to 1 and Q is some scaling factor. In fact, what we will do in the implementation of the model is to divide every EEPs by M, so that the constraints (Steer 1) and (Steer 2) will impose the main lobe to match the value M at the steering direction, and the constraint on the weights will just be $|w_i| \leq Q$.

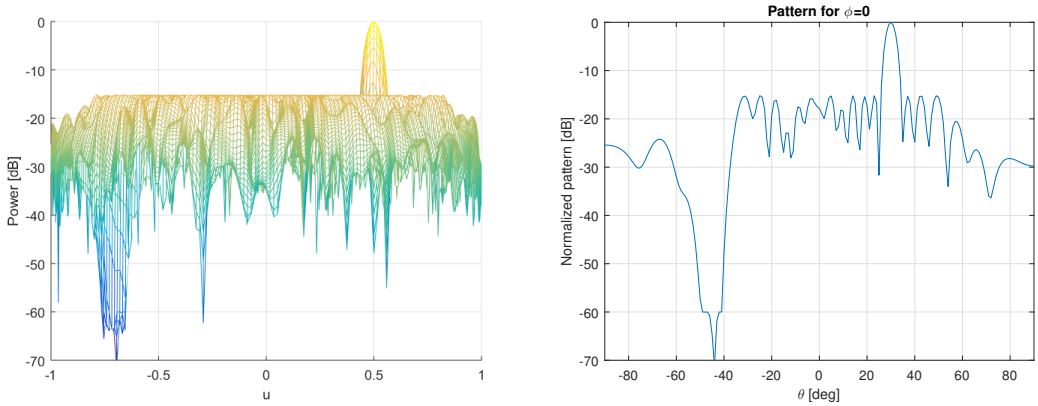
For our example, we will choose a main beam at $(\theta_0, \phi_0) = (30, 0)$ of half beamwidth HB = 3° and for the interference direction, we will choose a ball around $(\theta, \phi) = (45, 180)$ of radius 5°. The optimization is done on MATLAB with the second order cone solver of MOSEK [25]. The model is composed of 165,024 scalar variables, 164,549 constraints and 33,107 cones and takes about 300s to be solved. The resulting pattern is shown in Fig. 4.2 both in 3D and with a 2D cut along $\phi = 0$. We can see that the resulting pattern is quite smooth. The sidelobes are a little below -15dB. The pattern in the direction of the null is below -60dB as expected. Finally, on Fig. 4.1 we zoom at the mainbeam region and superimpose the window defined by the constraints, and we can see that the main lobe respects the constraints.

One can notice that the value of the sidelobes are not so low. This is in fact due to both the main beam width and the upper bound constraint. Indeed, if we choose HB = 5° instead and don't impose a lower bound, the SLL goes to −23dB. Moreover, if the main beam location is chosen close to $(\theta, \phi) = (0, 0)$, we are able to achieve SLL as low as −30dB. There are therefore lots of parameters to take into account.

4.2.1 Test for different sets of parameters

We will now investigate the behaviour of the optimization problem for different sets of parameters. We will vary the scaling factor Q, the half beamwidth (HB) and the Null

[4] Indeed, $AP_{dB} = 20\log_{10}(0.5) = -6dB$



(A) 3D pattern (B) cut along $\phi = 0$

FIGURE 4.2: Radiation pattern example 1

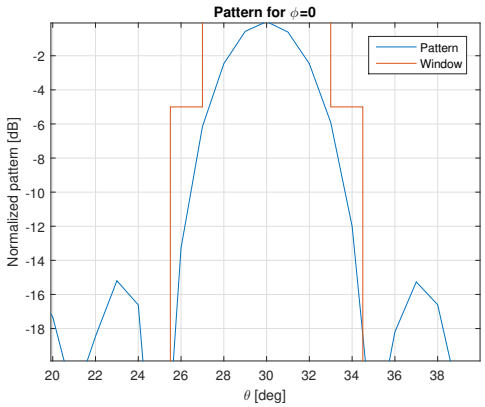


FIGURE 4.3: Zoom on main beam with window superimposed in orange

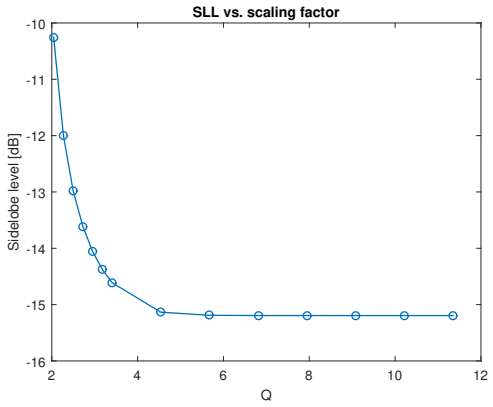


FIGURE 4.4: Variation of Q and impact on SLL

level (NL).

4.2.1.1 Variation of Q

Here we will choose to vary the scaling factor Q by a range of values from 1 to 10. We will fix NL $= -60$dB, HB $= 3°$. On Fig. 4.5 and 4.6, we show the dependence of Q on the level of the sidelobes and the SNR. We can observe that a small value of Q increases the SNR as expected but at the same time it also increases the sidelobe level. Some kind of trade-off has to be chosen by the engineer. A Pareto curve between the SLL and SNR is shown on Fig. 4.6.

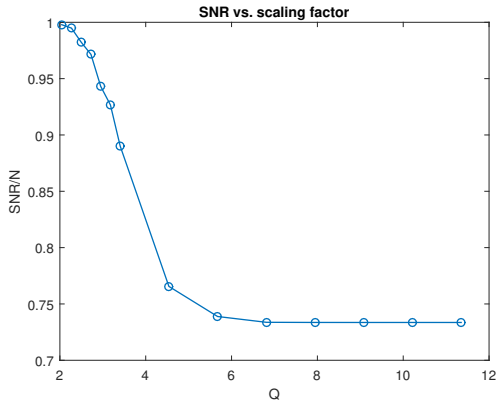


FIGURE 4.5: Variation of Q and impact on SNR

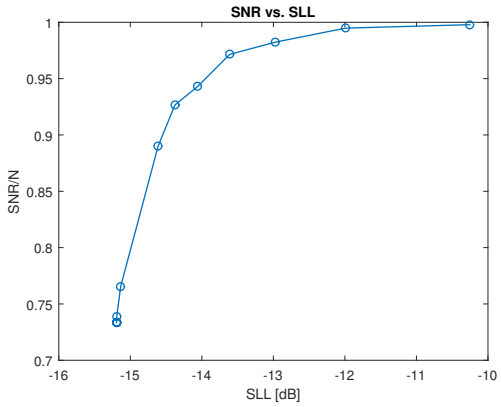


FIGURE 4.6: Trade-off between SNR and SLL for different values of Q

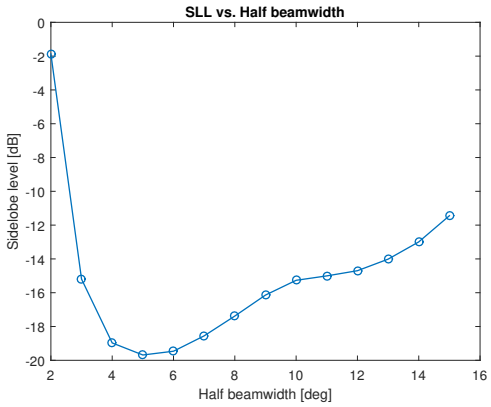


FIGURE 4.7: Variation of HB and impact on SLL

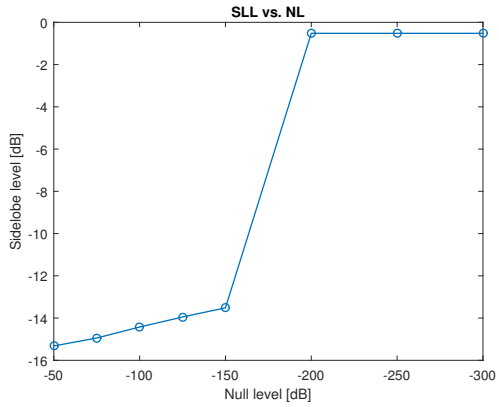


FIGURE 4.8: Variation of null level and impact on SLL

4.2.1.2 Variation of HB

We will now vary the half beamwidth from a range of 1° to 15°. We will fix NL = −60dB and give some high value to Q (so that the weight constraint is inactive). The result of the comparison of the sidelobes level as a function of the half-beamwidth is shown in Fig. 4.7. We observe that the SLL decreases up to a −20dB when HB decreases and then, it slowly increases again as we impose the beam to be wider. The best value is obtained when HB = 5°.

4.2.1.3 Variation of NL

We will finally vary NL, the maximal value of the pattern in the region where we null. We will fix HB = 3° and we will again give some high value to Q (so that the weight constraint is inactive). The result is shown in Fig. 4.8. We can see that NL does not

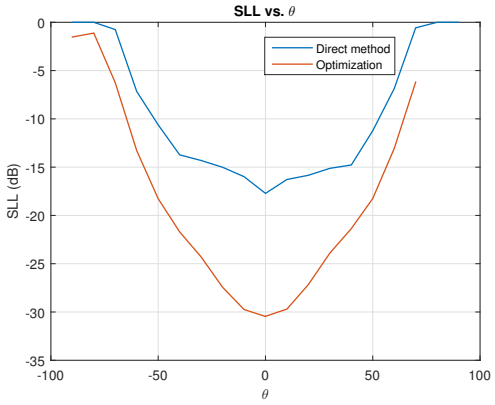


FIGURE 4.9: Comparison between SLL for direct method and optimization

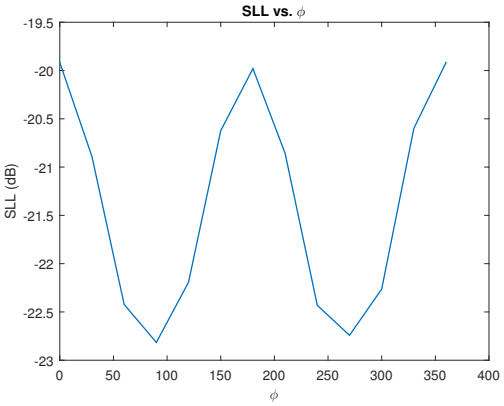


FIGURE 4.10: Variation of SLL in terms of ϕ for $\theta = 45°$ for optimization

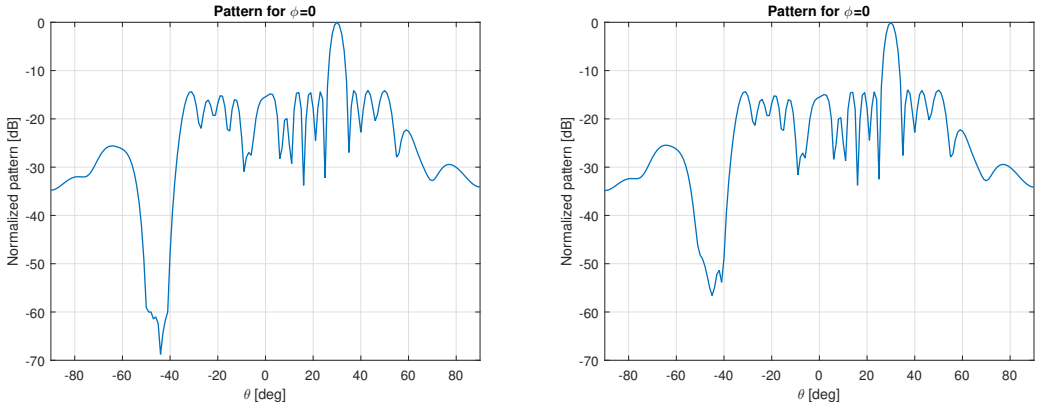
affect the SLL much, at least up to a certain point where the constraint cannot be respected anymore. Of course this will greatly depend on the size of the region we want to null. The bigger the region, the less we will be able to impose a high value for NL.

4.2.2 Comparison of the SLL between direct method and optimization

We will compare the level of the sidelobes when we use the projection method of Chapter 3 and the optimization. We will compare this when there is no nulling and for different angle θ when we fix $\phi = 0$. For the optimization problem, we use the following set of parameters : HB $= 5°$ and Q has some high value. The comparison is shown in Fig. 4.9 We can see that for all direction θ, the optimization problem have lower sidelobes than the direct method. Moreover, the directivity decreases as we move away from the center. This is due to the fact that the directivity varies in term of $\cos(\theta)$. Moreover, if we vary the azimuth angle ϕ for a fixed $\theta = 45°$, we will also have some attenuation in the directivity but not as much as for θ. The dependence of the SLL on ϕ for the optimization is shown in Fig. 4.10.

4.3 Problem of Robustness

Let us now investigate the robustness of our model. By robustness, we mean the behaviour of the model when a slight change occurs in one of the variables or parameters. A slight change could happen if the direction of the interferer is not known precisely



(A) Initial pattern (B) Pattern after perturbation

FIGURE 4.11: Comparison of the pattern before and after perturbation

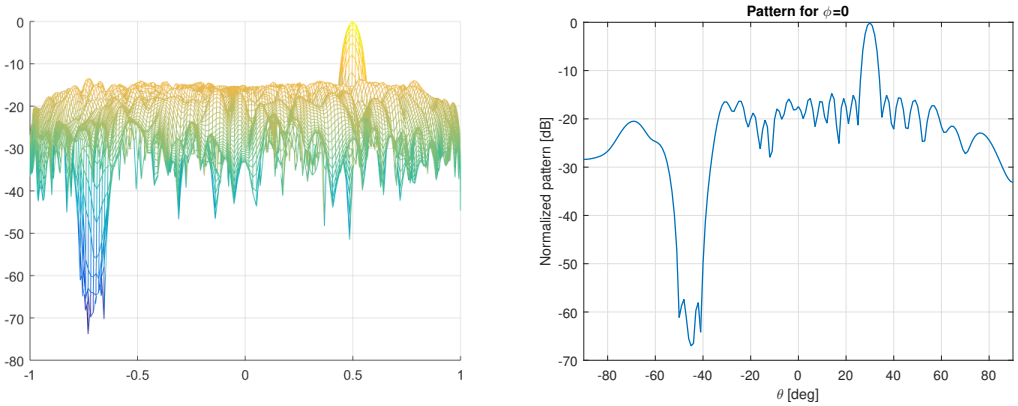
for example. In that case we can impose a ball of a certain radius around the interferer that we null. We would also like to know if the radiation pattern remains similar when the weights are perturbed by a random variation δ, e.g. for implementation precision.

Let us compute the weights when we fix the value $Q = 3$ so that the absolute value of the weights are at most 3 and compare this when we add a small perturbation uniformly distributed $\delta \in [-0.1\ 0.1] + j[-0.1\ 0.1]$ on those weights. Each δ is chosen independently on every weight and on the real and imaginary parts. We obtain the radiation pattern in Fig. 4.11. We observe that the optimization produces rather robust weights and this is actually the case for all the simulations I performed.

4.4 Two arms antennas

Until now, we have only worked with 1 arm of the antenna (but two polarization). The SKA antenna has however two (pair of) arms with two distinct and independent EEP. We could compute the weights separately for each arm (with the same null constraints and beam direction) by two separate optimization problems. We would hence find the set of weights $\mathbf{w}_1$, $\mathbf{w}_2$ satisfying both (pair of) arm. Then we would compute the full pattern according to

$$AP = \sqrt{\left|\mathbf{P}_V^1 \cdot w_1 + \mathbf{P}_V^2 \cdot w_2\right|^2 + \left|\mathbf{P}_H^1 \cdot w_1 + \mathbf{P}_H^2 \cdot w_2\right|^2} \tag{4.26}$$



(A) All pattern (B) cut at $\phi = 0$

FIGURE 4.12: Full pattern with the 2 polarizations

where $\mathbf{P}^1$ and $\mathbf{P}^2$ are the first and second arm respectively. The result of such a process gives the radiation pattern shown in Fig. 4.12. As we can observe the pattern is not as smooth as before but the sidelobes did not move too much, nor did the null region. Moreover, the power level of the main lobe, which is normalized to the maximum value of the full pattern given by replacing the weights by a vector of ones in (4.26), is the same. This is therefore a decent approximation to minimizing the true full pattern (4.26) in a single optimization problem involving both weights $\mathbf{w}_1, \mathbf{w}_2$.

The extension of the optimization problem with 1 arm to the optimization of the full pattern with both arm is not trivial. Indeed, equation (4.26) involves a sum including $\mathbf{w}_1 \cdot \mathbf{w}_2$ terms, hence this cannot be represented by a cone anymore. Semidefinite programming could be an idea to extend the method but we would have to investigate if the matrix of the objective is positive semi-definite. If this is not the case, this would not be a convex problem anymore and would therefore be really hard to handle.

4.5 Optimize both the position and weights

Throughout this thesis, we only tackled the problem of finding the best set of weights to achieve a desired radiation pattern. Indeed, the geometry of the station of the SKA telescope is already fixed and hence we were only interested in finding a proper set of weights. However, another optimization problem is to define the structure of the array to meet a desired pattern shape. This was the subject of a master thesis by Clavier at UCL which lead to a publication in [8]. But the optimization problem was performed by

assuming constant magnitude for the excitation of the elements. We may then wonder if it would be possible to optimize both the position of the antennas and the set of weights at the same time. Indeed, we could wonder how to fix the positions of the antennas so that it ensures the nulling problem is solved in the best way possible for example.

Some paper in the literature have tried to tackle the problem of optimizing the weights applied to the antennas at the same time as the position of the antennas. However, the methods developed in those papers do not usually effectively optimize both criteria at the same time and they have been developed for regular arrays and without mutual coupling.

For example, in [21], an initial rectangular array is given and the authors solve an optimization problem to find the sparsest array (i.e. select the least number of antennas) that achieves a desired SLL (and a main beam constraint). The optimization problem is based on a compressed sensing approach of minimizing the ℓ^0 norm of the weights (then approximated by a ℓ^1 norm for convexity) so that the solution gives the least number of nonzero weights and hence is able to select the fewest antennas. However, with this approach we do not minimize the SLL but essentially fix a SLL that we want to achieve with the sparsest geometry.

Another similar method is presented for linear arrays in [22] where a proper array size is fixed at first step according to the feasibility of the desired SLL and then the array positions are optimized according to a modified ℓ^0 norm.

In conclusion, there are actually no methods, to the best of my knowledge, that truly optimize both criteria. The best we can do right now if we had to design a new antenna array of fixed size is to find the proper geometry with the method developed by Clavier in [8] and then, based on this geometry, find a set of weights to achieve our desired radiation pattern with the model developed in this thesis.

4.6 Real time optimization

The SKA telescope will necessitate to find sets of weights every few minutes as the earth is moving and the position of the interferers is also moving. Therefore the optimization problem will be needed to be solved in real time on a strict time scale. This can be one

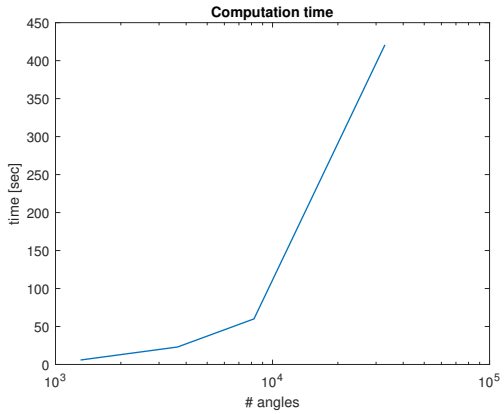


FIGURE 4.13: Computation time in terms of the number of angles

of the drawbacks about the optimization approach where it takes about 300s to solve (on a laptop) while the direct method takes less than a second. Although the time to solve the problem has been significantly reduced from the early beginning (going from 5 hours when used with CVX[5] to less than 5 minutes when directly used in MOSEK [25]) it nevertheless remains significant for a real time application. However, there exists solutions to tackle this problem, we could reoptimize by taking into account previous results in the algorithm. We could also improve the solver or the method. A third possibility would be to decrease the number of angles in the discretization. This would result in a smaller accuracy but it will increase the speed. If we use a step 2° instead of 1°, we decrease the number of angles by a factor 4, and the optimization problem is able to find the solution in 61s. If we choose an step of 3°, then we decrease the number of angles by a factor 9 and so on. A graph showing the computation time in terms of the number of angles is shown in Fig. 4.13. From those result, it seems that one would be able to use this optimization problem for application in real time, provided a good machine to solve the optimization problems.

[5]CVX [26] is a Matlab-based modeling system for convex optimization that parses the optimization problem written in simple mathematical expression into a structure readable by a solver.

Chapter 5

Conclusion

Throughout this thesis, we tackled the problem of designing efficient algorithms capable of matching a desired radiation pattern (i.e. steering in a desired direction, nulling interference regions and keeping the sidelobes as low as possible) by finding the proper weights (excitation current) to apply on each element. Moreover, we have done this while taking into account mutual coupling. Two types of methods have been developed.

The first type of method is based on matrix computations and does not need to be solved iteratively. The advantage of those methods is that they are fast but the main drawback is that we do not minimze the sidelobes. Two direct methods were derived : the first one is based on orthogonal projection of a steering vector on a subspace orthogonal to the interferers. This has the advantage of maintaining the main beam structure of the initial radiation pattern but it does not get nulls as deep as the second method when nulling wide regions. The second method is based on solving a least square system with minimal norm. The advantage of this method is that it ensures really deep nulls even for wide regions but the drawback is that it does not necessarily maintain the initial main beam structure.

The second type of method developed is an optimization problem. We showed that beamforming and nulling can be modeled as a convex optimization problem where the objective function minimizes the sidelobes level. The advantages of the optimization problem is the sidelobes minimization and the control of many parameters. However, this comes with the disadvantage that the it requires a greater amount of computation time. The model derived is quite polyvalent and allows the engineer to choose a desired

null level and the main beam width. Several set of parameters have to be fine-tuned and traded off to meet the needs of the engineer such as the speed/accuracy, SNR/SLL, etc. The robustness of the problem was also investigated, showing a good behavior to perturbations. We also saw that with some more improvements, the model could be suited for real time applications.

All the methods developed in this thesis have been tested on a SKA telescope station of 256 SKALA elements, showing promising results that could be applied in real life.

A possible extension of this thesis could be treating the case of the two pairs of arms of the SKALA element together in a single optimization problem rather than independently. Making the optimization faster to be sure it is suited for real time applications is also another challenge.

Appendix A

Pseudo-inverse and least squares solution

Most of the material from this appendix comes from [19, 27]. In this appendix, we will be interested to the solution to the linear system of equation

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{A.1}$$

when $\mathbf{A}$ is not a square matrix. The matrix $\mathbf{A}$ has m rows and n columns where $m \neq n$. We will investigate the cases where $m > n$, that is there is more equations than unknowns (overdetermined system) and when $m < n$, that is there are less equations than unknowns (underdetermined system). Moreover we will also investigate the cases where $\mathbf{A}$ is rank deficient. But first let us introduce the concept of rank of a matrix.

A.1 Rank of a matrix

Definition A.1. The column rank of a matrix is the dimension of its column space, i.e. the number of linearly independent columns of the matrix

Definition A.2. The row rank of a matrix is the dimension of its row space, i.e. the number of linearly independent rows of the matrix

Lemma A.3. *The row rank always equal the column rank. Thus we will generally use the term* rank *of a matrix.*

Definition A.4. A matrix $\mathbf{A} \in \mathbb{C}^{m \times n}$ is said to be *full rank* if $\text{rank}(\mathbf{A}) = \min(m, n)$. If the matrix is not full rank it is said to be rank deficient. The matrix is said to be *full column rank* if $\text{rank}(\mathbf{A}) = n$ and *full row rank* if $\text{rank}(\mathbf{A}) = m$.

A.2 Moore-Penrose Pseudoinverse

The Moore-Penrose pseudo-inverse of a matrix $A \in {}^{m \times n}$, denoted $A^+ \in {}^{n \times m}$ is a generalization of the inverse matrix which satisfies the 4 following properties :

1. $AA^+A = A$
2. $A^+AA^+ = A^+$
3. $(AA^+)^H = AA^+$
4. $(A^+A)^H = A^+A$

Moreover A^+ exist for any matrix and is unique. When $\mathbf{A}$ is full rank, A^+ can be expressed explicitly as an algebraic formula.

Indeed, when A is full column rank (A^HA is invertible), it can be computed as $A^+ = (A^HA)^{-1}A^H$, this is the left inverse since $A^+A = I$. While if $\mathbf{A}$ is full row rank (AA^H is invertible), it is computed as $A^+ = A(AA^H)^{-1}$, this is the right inverse $AA^+ = I$.

If $\mathbf{A}$ is rank deficient, the pseudo-inverse still exist but no simple algebraic solution exist. The pseudo-inverse is in that case computed via Singular value decompostion (SVD).

We will prove in the following sections that the pseudoinverse can be used to find the minimal least squares solution $x = A^+b$ of the linear system of equations $Ax = b$.

Definition A.5. If $\mathbf{A} \in \mathbb{C}^{m \times n}$ and $\mathbf{b} \in \mathbb{C}^m$. Then a vector $\mathbf{u} \in \mathbb{C}^n$ is called *least squares solution* to $\mathbf{Ax} = \mathbf{b}$ if $\|\mathbf{Au} - \mathbf{b}\| \leq \|\mathbf{Av} - \mathbf{b}\|$ $\forall \mathbf{v}$. A vector $\mathbf{u}$ is called *minimal least squares solution (or solution with minimal 2-norm)* to $\mathbf{Ax} = \mathbf{b}$ if $\mathbf{u}$ is a least square solution to $\mathbf{Ax} = \mathbf{b}$ and $\|\mathbf{u}\| \leq \|\mathbf{v}\|$ for all other least square solutions $\mathbf{w}$

A.3 Over-determined System

A.3.1 A is full rank

We suppose here that $\mathbf{A}$ has more equations than unknowns ($m > n$) and that $\mathbf{A}$ is full column rank.

We want to find the vector $\hat{\mathbf{x}}$ such that the residual $\mathbf{r} = \mathbf{b} - \mathbf{A}\hat{\mathbf{x}}$ is as small as possible. One can use the sum of squares of the residual to be as small as possible, i.e.

$$\begin{aligned}
\mathcal{S} &= \min_{\mathbf{x}} (\mathbf{r}^{\mathbf{T}} \cdot \mathbf{r}) && \text{(A.2)}\\
&= \min_{\mathbf{x}} (\mathbf{b} - \mathbf{A}\mathbf{x})^{\mathbf{H}} \cdot (\mathbf{b} - \mathbf{A}\mathbf{x}) && \text{(A.3)}\\
&= \min_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{b}\|_2 && \text{(A.4)}\\
&= \min_{\mathbf{x}} \mathbf{b}^{\mathbf{H}}\mathbf{b} - \mathbf{x}^{\mathbf{H}}(\mathbf{A}^{\mathbf{H}}\mathbf{b}) - (\mathbf{b}^{\mathbf{H}}\mathbf{A})\mathbf{x} + \mathbf{x}^{\mathbf{H}}(\mathbf{A}^{\mathbf{H}}\mathbf{A})\mathbf{x} && \text{(A.5)}\\
\Leftrightarrow (\mathbf{A}^{\mathbf{H}} \cdot \mathbf{A}) \cdot \mathbf{x} &= \mathbf{A}^{\mathbf{H}} \cdot \mathbf{b} && \text{(A.6)}\\
\mathbf{x} &= (\mathbf{A}^{\mathbf{H}}\mathbf{A})^{-\mathbf{1}}\mathbf{A}^{\mathbf{H}}\mathbf{b} && \text{(A.7)}
\end{aligned}$$

where $\mathbf{A}^{+} = (A^H A)^{-1} A^H$.

A.3.2 A is rank deficient

If $\mathbf{A}$ is rank deficient, $(A^H A)$ is not invertible in eq. (A.5) and hence cannot be computed in this manner. Moreover the least square solution is in this case not unique. There exist an equivalent matrix $\mathbf{A}^{+}$ that solves $\mathbf{A}\mathbf{x} = \mathbf{b}$ in a least square sense which is based on the singular value decomposition of $\mathbf{A}$. Moreover this least square solution will give the least square solution with minimal norm $\|\mathbf{x}\|_2$.

The SVD of $\mathbf{A}$ is given by

$$A = VDU^H \tag{A.8}$$

The pseudo-inverse of $\mathbf{A}$ is defined as

$$A^{+} = UD^{-1}V^H \tag{A.9}$$

The solution $x = A^+ b$ will give the least squares solution of smallest norm of $Ax = b$. The proof of this statement can be found in [19, p. 446].

A.4 Under-determined System

A.4.1 A is full rank

We suppose here that **A** has less equations than unknowns ($m < n$) and that **A** is full row rank.

In this case, there are an infinite number of solutions satisfying $\mathbf{Ax} = \mathbf{b}$. We want to pick the solution that has minimum 2-norm. Therefore, we want to find a solution that satisfies

$$\min x^H x \qquad \text{subject to} \qquad Ax = b \tag{A.10}$$

This can be formulated with Lagrange multipliers :

$$\mathcal{L} = x^H x + \lambda^H (Ax - b) \tag{A.11}$$

Setting the partial derivatives to zero, we get

$$\frac{\partial \mathcal{L}}{\partial x} = 2x^H + \lambda^H A = 0 \Rightarrow x = 1/2 A^H \lambda \tag{A.12}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = Ax - b = 0 \tag{A.13}$$

Substituting (A.12) into (A.13) and solving for λ, we get

$$\lambda = 2(AA^H)^{-1} b \tag{A.14}$$

The final solution is given by

$$x = A^H (AA^H)^{-1} b \tag{A.15}$$

where $A^+ = A^H (AA^H)^{-1}$.

A.4.2 A is rank deficient

If A is rank deficient, (AA^H) is not invertible in (A.15) and hence cannot be computed in this manner. The SVD will provide the pseudo-inverse in the same way as eq (A.9). Then $x = A^+b$ will provide the minimal norm solution.

Appendix B

Analytic comparison between the projection and pseudo-inverse method

Let us compare the two directs methods developed in Chap. 3 analytically and explain why they give similar results when only a few nulls are enforced. We will study this for the case where there is only 1 polarization and no mutual coupling as it is easier for notations.

When there is no mutual coupling and the antenna is isotropic, we have that

$$\mathbf{P}(\theta_s, \phi_s) = [e^{(jk\mathbf{u}_s \cdot \mathbf{r}_1)} \dots e^{(jk\mathbf{u}_s \cdot \mathbf{r}_N)}] = \mathbf{a}_0^T$$

and

$$\begin{bmatrix} \mathbf{P}(\theta_{\text{null}_1}, \phi_{\text{null}_1}) \\ \vdots \\ \mathbf{P}(\theta_{\text{null}_M}, \phi_{\text{null}_2}) \end{bmatrix} = \begin{bmatrix} e^{(jk\mathbf{u}_{\text{null}_1} \cdot \mathbf{r}_1)} & \dots & e^{(jk\mathbf{u}_{\text{null}_1} \cdot \mathbf{r}_N)} \\ \vdots & & \vdots \\ e^{(jk\mathbf{u}_{\text{null}_M} \cdot \mathbf{r}_1)} & \dots & e^{(jk\mathbf{u}_{\text{null}_M} \cdot \mathbf{r}_N)} \end{bmatrix} = \mathbf{V}^T$$

Therefore, the system (3.22) can be rewritten as

$$\begin{bmatrix} \mathbf{a}_0^T \\ \mathbf{V}^T \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_N \end{bmatrix} = \begin{bmatrix} 1 \\ \mathbf{0} \end{bmatrix} \tag{B.1}$$

We want to show that

$$\mathbf{w} = \begin{bmatrix} \mathbf{a}_0^T \\ \mathbf{V}^T \end{bmatrix}^+ \begin{bmatrix} 1 \\ \mathbf{0} \end{bmatrix} \tag{B.2}$$

is in some sense similar to [1]

$$\mathbf{w} = \left(\mathbf{P}_{\mathbf{V}}^{\perp} \frac{\mathbf{a}_0}{\mathbf{a}_0^H \mathbf{a}_0} \right)^* \tag{B.3}$$

with $\mathbf{P}_{\mathbf{V}}^{\perp} = \mathbf{I} - \mathbf{V}\mathbf{V}^+$, or more easily written $\mathbf{w} = \begin{bmatrix} \mathbf{a}_0^H \\ \mathbf{V}^H \end{bmatrix}^+ \begin{bmatrix} 1 \\ \mathbf{0} \end{bmatrix}$ is quite similar to $\mathbf{w} = \mathbf{P}_{\mathbf{V}}^{\perp} \frac{\mathbf{a}_0}{\mathbf{a}_0^H \mathbf{a}_0}$.

First, we need to compute the Moore-Penrose pseudo-inverse of a row-wise partitioned matrix. A particular formula for the MoorePenrose inverse of a columnwise partitioned matrix has been derived in [28]. The theorem is stated as follows :

Theorem B.1. *Let* $\mathbf{A} \in \mathbb{C}_{m,n}$ *be partitioned in* $\mathbf{A} = [\mathbf{A}_1 \ \mathbf{A}_2]$ *where* $\mathbf{A}_1, \mathbf{A}_2$ *are full rank and let* $\mathbf{P}_{\mathbf{1}}^{\perp}, \mathbf{P}_{\mathbf{2}}^{\perp} \in \mathbb{C}_{m,m}$ *be the orthogonal projectors defined as* $\mathbf{P}_{\mathbf{i}}^{\perp} = \mathbf{I}_m - \mathbf{A}_i \mathbf{A}_i^+$. *Moreover, let*

$$\mathbf{G} = \begin{bmatrix} (\mathbf{P}_{\mathbf{2}}^{\perp} \mathbf{A}_{\mathbf{1}})^+ \\ (\mathbf{P}_{\mathbf{1}}^{\perp} \mathbf{A}_{\mathbf{2}})^+ \end{bmatrix}.$$

Then $\mathbf{G} = \mathbf{A}^+$

The proof of this statement is given in [28]. This can easily be adapted to the case of a row-wise partitioned matrix. In that case, we have that $\mathbf{A}$ is partitioned as $\mathbf{A} = \begin{bmatrix} \mathbf{A}_1^H \\ \mathbf{A}_2^H \end{bmatrix}$ and $\mathbf{A}^+ = \begin{bmatrix} (\mathbf{A}_1^H \mathbf{P}_{\mathbf{2}}^{\perp})^+ & (\mathbf{A}_2^H \mathbf{P}_{\mathbf{1}}^{\perp})^+ \end{bmatrix}$.

In our initial problem, we have that $\mathbf{A}_1 = \mathbf{a}_0$, $\mathbf{A}_2 = \mathbf{V}$, $\mathbf{P}_{\mathbf{1}}^{\perp} = \mathbf{P}_{\mathbf{a_0}}^{\perp}$ and $\mathbf{P}_{\mathbf{2}}^{\perp} = \mathbf{P}_{\mathbf{V}}^{\perp}$. Therefore, equation (B.2) can be rewritten as

$$\begin{aligned} \mathbf{w} &= \begin{bmatrix} (\mathbf{a}_0^H \mathbf{P}_{\mathbf{V}}^{\perp})^+ & (\mathbf{V}^H \mathbf{P}_{\mathbf{a_0}}^{\perp})^+ \end{bmatrix} \begin{bmatrix} 1 \\ \mathbf{0} \end{bmatrix} && \text{(B.4)} \\ &= (\mathbf{a}_0^H \mathbf{P}_{\mathbf{V}}^{\perp})^+ && \text{(B.5)} \\ &\stackrel{?}{=} (\mathbf{P}_{\mathbf{V}}^{\perp})^+ (\mathbf{a}_0^H)^+ && \text{(B.6)} \\ &= \mathbf{P}_{\mathbf{V}}^{\perp} \frac{\mathbf{a}_0}{\mathbf{a}_0^H \mathbf{a}_0} && \text{(B.7)} \end{aligned}$$

[1] We take the conjugate $(\cdot)^*$ because the array pattern is computed as $AP = \mathbf{w}^H e^{jk\mathbf{u}\cdot\mathbf{r}}$

To go from line 3 to line 4 we use the fact that the pseudo-inverse of an orthogonal projector is the orthogonal projector and that the pseudo-inverse of a row vector is the conjugate transpose divided by its square magnitude. If $(AB)^+ = B^+A^+$, the two methods are equivalent. However, this is not in general the case and it is not the case for all of our examples unless there are only a few interferers in which case the two solutions (B.5) and (B.6) are almost equal. Unfortunately I was not able to come with an intuitive explanation for this. Looking in the literature about the pseudoinverse of a product and the relation $(AB)^+ = B^+A^+$, I realized that there is no real explanation to this, and I concluded this section like the authors of a recent and complete book [18] about the Moore-Penrose pseudoinverse who came up with this answer : *The authors feel that a good answer to the question 'What is* $(AB)^+$*' does not, and probably will not exist.* [18, p. 19].

Bibliography

[1] B. D. Van Veen and K. M. Buckley, "Beamforming: A versatile approach to spatial filtering," *IEEE assp magazine*, vol. 5, no. 2, pp. 4–24, 1988.

[2] H. Subbaram and K. Abend, "Interference suppression via orthogonal projections: A performance analysis," *IEEE Transactions on Antennas and Propagation*, vol. 41, no. 9, pp. 1187–1194, 1993.

[3] S. W. Ellingson and G. A. Hampson, "A subspace-tracking approach to interference nulling for phased array-based radio telescopes," *IEEE Transactions on Antennas and Propagation*, vol. 50, no. 1, pp. 25–30, 2002.

[4] S. W. Ellingson and W. Cazemier, "Efficient multibeam synthesis with interference nulling for large arrays," *IEEE Transactions on antennas and propagation*, vol. 51, no. 3, pp. 503–511, 2003.

[5] H. Lebret and S. Boyd, "Antenna array pattern synthesis via convex optimization," *IEEE transactions on signal processing*, vol. 45, no. 3, pp. 526–532, 1997.

[6] H. Lebret, "Antenna pattern synthesis through convex optimization," in *SPIE's 1995 International Symposium on Optical Science, Engineering, and Instrumentation*. International Society for Optics and Photonics, 1995, pp. 182–192.

[7] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear algebra and its applications*, vol. 284, no. 1-3, pp. 193–228, 1998.

[8] T. Clavier, "Ska telescope optimization," Master's thesis, Universit Catholique de Louvain, 2012.

[9] (2017, Feb) Ska project. Square Kilometer Array. [Online]. Available: https://www.skatelescope.org/project/

[10] C. Craeye, "Lelec2910 : Antenna and propagations," University Lecture, 2016.

[11] H. L. Van Trees, *Detection, estimation, and modulation theory, optimum array processing*. John Wiley & Sons, 2004.

[12] A. Thompson, "Spatial nulling for attenuation of interfering signals," *SKA Memo Ser*, vol. 34, 2003.

[13] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE signal processing magazine*, vol. 13, no. 4, pp. 67–94, 1996.

[14] E. de Lera Acedo, N. Razavi-Ghods, N. Troop, N. Drought, and A. Faulkner, "Skala, a log-periodic array antenna for the ska-low instrument: design, simulations, tests and system considerations," *Experimental Astronomy*, vol. 39, no. 3, pp. 567–594, 2015.

[15] J. Landon, B. D. Jeffs, and K. F. Warnick, "Model-based subspace projection beamforming for deep interference nulling," *IEEE Transactions on Signal Processing*, vol. 60, no. 3, pp. 1215–1228, 2012.

[16] Q. Gueuning, C. Raucy, C. Craeye, E. Colin-Beltran, and E. de Lera Acedo, "Mutual coupling analysis in non-regular arrays of skala antennas with the harp approach," in *Antennas and Propagation & USNC/URSI National Radio Science Meeting, 2015 IEEE International Symposium on*. IEEE, 2015, pp. 1528–1529.

[17] D. Gonzalez-Ovejero and C. Craeye, "Interpolatory macro basis functions analysis of non-periodic arrays," *IEEE Transactions on Antennas and Propagation*, vol. 59, no. 8, pp. 3117–3122, 2011.

[18] S. L. Campbell and C. D. Meyer, *Generalized inverses of linear transformations*. SIAM, 2009.

[19] G. Gerig, "Cs6640 : Image processing," University Lecture, 2012. [Online]. Available: http://www.sci.utah.edu/~gerig/CS6640-F2012/Materials/pseudoinverse-cis61009sl10.pdf

[20] A. B. Gershman, N. D. Sidiropoulos, S. Shahbazpanahi, M. Bengtsson, and B. Ottersten, "Convex optimization-based beamforming," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 62–75, 2010.

[21] S. E. Nai, W. Ser, Z. L. Yu, and H. Chen, "Beampattern synthesis for linear and planar arrays with antenna selection by convex optimization," *IEEE Transactions on Antennas and Propagation*, vol. 58, no. 12, pp. 3923–3930, 2010.

[22] G. Prisco and M. D'Urso, "Maximally sparse arrays via sequential convex optimizations," *IEEE Antennas and Wireless Propagation Letters*, vol. 11, pp. 192–195, 2012.

[23] Y. Nesterov and A. Nemirovskii, *Interior-point polynomial algorithms in convex programming*. SIAM, 1994.

[24] F. Glineur, "Linma2471 : Optimization models and method ii," University Lecture, 2016.

[25] M. ApS, *The MOSEK optimization toolbox for MATLAB manual. Version 7.1 (Revision 28).*, 2015. [Online]. Available: http://docs.mosek.com/7.1/toolbox/index.html

[26] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[27] B. G. Higgins, "Ech256 : Least squares solution and pseudo-inverse," University Lecture, 2012.

[28] J. K. Baksalary and O. M. Baksalary, "Particular formulae for the moore–penrose inverse of a columnwise partitioned matrix," *Linear algebra and its applications*, vol. 421, no. 1, pp. 16–23, 2007.

Rue Archimède, 1 bte L6.11.01, 1348 Louvain-la-Neuve **www.uclouvain.be/epl**