Improving the Direction-Dependent Gain Calibration of Reflector Antenna Radio Telescopes

by

André Young

Dissertation presented for the degree of Doctor of Philosophy in Electronic Engineering in the Faculty of Engineering at Stellenbosch University

Promoters:

Prof. David B. Davidson
Department of Electrical and Electronic Engineering
University of Stellenbosch
Stellenbosch, South Africa

Prof. Rob Maaskant
Department of Signals and Systems
Chalmers University of Technology
Gothenburg, Sweden

Prof. Marianna V. Ivashina
Department of Signals and Systems
Chalmers University of Technology
Gothenburg, Sweden

December 2013

Declaration

By submitting this dissertation electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: December 2013 .....

Abstract

Utilising future radio interferometer arrays, such as the Square Kilometre Array (SKA), to their full potential will require calibrating for various direction-dependent effects, including the radiation pattern (or *primary beam* in the parlance of radio astronomers) of each of the antennas in such an array. This requires an accurate characterisation of the radiation patterns at the time of observation, as changing operating conditions may cause substantial variation in these patterns. Furthermore, fundamental imaging limits, as well as practical time constraints, limit the amount of measurement data that can be used to perform such characterisation. Herein three techniques are presented which aim to address this requirement by providing pattern models that use the least amount of measurement data for an accurate characterisation of the radiation pattern. These methods are demonstrated through application to the MeerKAT Offset Gregorian (OG) dual-reflector antenna.

The first technique is based on a novel application of the Jacobi-Bessel series in which the expansion coefficients are solved directly from the secondary pattern. Improving the efficiency of this model in the desired application leads to the development of a different set of basis functions, as well as two constrained solution approaches which reduce the number of pattern measurements required to yield an accurate and unique solution.

The second approach extends the application of the recently proposed Characteristic Basis Function Patterns (CBFPs) to compensate for non-linear pattern variations resulting from mechanical deformations in a reflector antenna system. The superior modelling capabilities of these numerical basis functions, which contain most of the pattern features of the given antenna design in a single term, over that of analytic basis functions are demonstrated.

The final method focusses on an antenna employing a Phased Array Feed (PAF) in which multiple beam patterns are created through the use of a beamformer. Calibration of such systems poses a difficult problem as the radiation pattern shape is susceptible to gain variations. Here we propose a solution which is based on using a Linearly Constrained Minimum Variance (LCMV) beamformer to conform the realised beam pattern to a physics-based analytic function. Results show that the LCMV beamformer successfully produces circularly symmetric beams that are accurately characterised with a single-term analytic function over a wide FoV.

Opsomming

Die volle benutting van toekomstige radio interferometersamestellings, soos die *Square Kilometre Array* (SKA), benodig die kalibrering van verskeie rigting-afhanklike effekte, insluitend die stralingspatroon (bekend as die *primêre bundel* onder radio astronome) van elke antenne in só 'n samestelling. Hierdie benodig 'n akkurate karakterisering van die stralingspatrone op die waarnemingstydstip, aangesien veranderende bedryfskarakteristieke 'n beduidende afwyking in hierdie patrone veroorsaak. Verder, weens fundamentele perke in beeldverwerking, asook praktiese tydbeperkinge, bestaan daar 'n limiet op die hoeveelheid gemeetde data wat benut kan word om die nodige karakterisering mee te doen. Hierin word drie tegnieke ten toon gestel wat gemik is daarop om aan hierdie behoefte te voorsien deur die gebruik van modelle wat 'n minimum hoeveelheid metingdata benodig om 'n akkurate beskrywing van die stralingspatroon te lewer. Die verskeie metodes word aangebied aan die hand van die MeerKAT afset-Gregorian dubbelreflektorantenne.

Die eerste tegniek is gebasseer op 'n nuwe toepassing van die Jacobi-Besselreeks waarin die sekondêre stralingspatroon direk gebruik word om die uitsettingskoëffisiënte op te los. Die doelmatigheidsverbetering van hierdie model in die huidige toepassing lei na die ontwikkeling van 'n nuwe versameling van basisfunksies, asook twee voorwaardelike oplossings wat die nodige aantal metings vir 'n akkurate, unieke oplossing verminder.

In die tweede tegniek word die toepassing van die onlangs voorgestelde Karakteristieke Basisfunksie Patrone uitgebrei om te vergoed vir die nie-lineêre stralingspatroonafwykings wat teweeggebring word deur meganiese vervormings in die reflektorantenne. Die superieure modelleringsvermoëns van hierdie numeriese basisfunksies, wat meeste van die patroonkenmerke vasvang in 'n enkele term, bo dié van analitiese basisfunksies word gedemonstreer.

Die laaste metode fokus op die gebruik van 'n gefaseerde samestellingvoer waarin veelvoudige bundelpatrone geskep word deur die gebruik van 'n bundelvormer. Die kalibrering van sulke instrumente word bemoeilik daardeur dat die patroonvorm gevoelig is vir aanwinsafwykings. Hier stel ons 'n oplossing voor waarin 'n lineêrbegrensde minimumstrooiing bundelvormer gebruik word om die stralingspatroon te pas op 'n fisika-gebasseerde analitiese funksie. Resultate toon dat hierdie bundelvormer sirkelsimmetriese bundels kan skep wat akkuraat beskryf word deur 'n een-term analitiese funksie oor 'n wye gesigsveld.

Acknowledgements

I would like to express my sincerest gratitude to the following people and institutions without whom this work would not have realised.

Firstly, I would like to thank Professor David B. Davidson, my supervisor at the University of Stellenbosch, for his continued support, encouragement, and patience throughout the past eight years since I did my undergraduate final year project under his supervision. It has been a privilege to work under his leadership — an experience that has surely made a significant contribution to my development as an independent researcher — and I am especially grateful to him for allowing me the freedom to wander into unknown territories in search of a topic for this work. He has also managed to find time out of an extremely busy schedule to review this thesis and to ensure that the examination process was completed in good time.

In the same breath, I would like to thank Professors Rob Maaskant and Marianna V. Ivashina, my co-supervisors at the Chalmers University of Technology, for introducing me to this very interesting research area. Their in-depth knowledge and many years of experience in the field of radio astronomy have proven invaluable in the completion of this work. Collaborating with them has been a great pleasure, and I look forward to continue working with them on various projects in the future. I am also grateful to have had the opportunity to visit Chalmers twice on their invitation, and for the many hours that they have spent discussing my work, which in some cases extended well beyond office hours.

Thanks to Doctor Dirk I. L. de Villiers at the University of Stellenbosch, especially for providing me with the various shaped geometries based on the MeerKAT optical design, and for the many conversations on various aspects of reflector antennas.

Thanks to Oleg A. Iupikov for providing me with the GRASP toolbox interface, without which many of the simulations performed as part of this work would have been more difficult by at least an order of magnitude. I am also grateful for his contributions to the work on beamforming, for very useful comments, and for the opportunity to collaborate on projects that he is working on.

Thanks to the Antenna Group at Chalmers, and especially to Professor Per-Simon Kildal for the opportunity to visit Chalmers on two occasions. These

visits have proven to be most productive and have enriched me professionally. The warmth and friendliness of this group have also managed to transform a cold Christmas far from home into a most joyous occasion.

I would like to thank Doctor Isak P. Theron at EMSS Antennas for information on MeerKAT and very useful feedback at the early stages of this work.

Thanks to Professor Oleg. M. Smirnov at Rhodes University for various discussions and insightful comments on the application of this work. I am also grateful to him for inviting me to participate in the *Third Generation Calibration Workshop III* held in Port Alfred during February 2013, and for Professor Rob Maaskant for extending his invitation to me.

Thanks to Doctor Ludwig Schwardt and the organisers of the *Calibration and Imaging Workshop 2012* held in Cape Town during December 2012 for inviting me to participate in this workshop.

I would also like to thank Professor Marco A. B. Terada at the University of Brasilia for the opportunity to collaborate on the sensitivity assessment of the MeerKAT antenna. Although this work did not find its way in this thesis, it has surely contributed to a greater understanding of the calibration challenges that face future radio telescopes.

Thanks to all my Colleagues, past and present, in the Computational Electromagnetics Group (CEMAGG). Although working on vastly different projects meant that there was little chance for the CEMAGG members to discuss our work in detail, it has been a tremendous help to know that I was not working alone. I would like to extend a special word of thanks to two former CEMAGG members, Doctor Evan Lezar and Doctor Renier G. Marchand, for their support during difficult stages of this work.

Thanks to Doctor Danie N. J. Els at the University of Stellenbosch for creating the various LaTeXpackages that have been used to compile this thesis in the correct format.

I am indebted to the National Research Foundation and the South African Research Chairs Initiative of the Department of Science and Technology for providing financial support for my pursuit of a doctoral degree.

I am also grateful for the financial support of the Swedish SIDA grant which contributed to my visits to Chalmers.

To my parents, André Young and Petro J. Young, and to my brothers, Timothy B. Young and Donovan Young, thank you for the education that you have given me. I am very grateful for the ways in which you have shaped, and continue to shape me.

To Suleen, thank you for your words of encouragement during difficult times, for your understanding when work continued into evenings and weekends, for always expressing sincere interest in my work, and most of all, for your unfailing love.

Contents

List of Figures

List of Tables

Nomenclature

Abbreviations

AA	Aperture Array
APERTIF	APERture Tile-In-Focus
ASKAP	Australian Square Kilometre Array Pathfinder
CBFP	Characteristic Basis Function Pattern
CP	Co-Polarised
DoF	Degree of Freedom
DS	Direct Solution (see Section 3.2.2)
EEP	Embedded Element Pattern
FoM	Figure of Merit
FoV	Field of View
HPBW	Half-Power Beam Width
JB	Jacobi-Bessel (as in JB-model, see Section 3.1.2)
LCMV	Linearly Constrained Minimum Variance
LMS	Lagrange Multiplier Solution (see Section 3.2.2)
MaxDir	Maximum Directivity (as in MaxDir beamformer)
MaxSNR	Maximum Signal-to-Noise Ratio (as in MaxSNR beamformer)
MeerKAT	Meer Karoo Array Telescope
MoM	Method of Moments
OG	Offset Gregorian
PAF	Phased Array Feed
PO	Physical Optics
PTD	Physical Theory of Diffraction
QPS	Quadratic Penalty Solution (see Section 3.2.2)
RIME	Radio Interferometer Measurement Equation
SKA	Square Kilometre Array
SLL	Side Lobe Level (maximum first sidelobe relative to pattern maximum, unless stated otherwise)

SNR	Signal-to-Noise Ratio
SVD	Singular Value Decomposition
TSA	Tapered Slot Antenna
VLA	Very Large Array
WRS	Weighted Residual Solution (see Section 3.2.1)
WSRT	Westerbork Synthesis Radio Telescope
XP	Cross-Polarised

Conventions

a, A	Scalar variables
$\mathbf{a}$, $\mathbf{A}$	Physical vector quantities
$\hat{\mathbf{a}}$	Unit vectors
$\mathbf{a}$, $\mathbf{A}$	Matrices and vectors in systems of linear equations
$(\mathbf{A})_{ij}$, A_{ij}	Element on the ith row and jth column in matrix $\mathbf{A}$

Mathematical Symbols

$\lvert x\rvert$	Magnitude or absolute value of x
$\overline{a}$	Complex conjugate of a
$n\,!$	Factorial, $n \times (n-1) \times (n-2) \times \cdots \times (2) \times (1)$
$\mathbf{a} \cdot \mathbf{b}$	Vector dot-product, $\sum_i a_i b_i$
$\mathbf{a} \times \mathbf{b}$	Vector cross-product
$\mathbf{a} \perp \mathbf{b}$	$\mathbf{a}$ is perpendicular to $\mathbf{b}$
$\mathbf{a} \parallel \mathbf{b}$	$\mathbf{a}$ is parallel to $\mathbf{b}$
$\lVert\mathbf{x}\rVert$	Euclidean norm, $\sqrt{\sum_i x_i^2}$
$\mathbf{A}^\dagger$	Hermitian transpose, $\left(\mathbf{A}^\dagger\right)_{ij} = \overline{A_{ji}}$
$\mathbf{A}^T$	Transpose, $\left(\mathbf{A}^T\right)_{ij} = A_{ji}$
$E\left[f(x)\right]$	Expectation of f

$$E\left[f(x)\right] = \int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x$$

$f(t) \star g(t)$	Correlation of f and g

$$f(t) \star g(t) = E\left[f(t_1) g(t_2)\right]$$

$\langle f(t)\rangle$	Time-average of f

$$\langle f(t)\rangle = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} f(t)\,\mathrm{d}t$$

$\mathbf{C} = \mathbf{A} \circ \mathbf{B}$	Hadamard, or element-wise product $C_{ij} = A_{ij}B_{ij}$
$\operatorname{diag} \mathbf{A}$	Returns the main diagonal elements of $\mathbf{A}$ as a column vector
$\min f$, $\min \mathbf{f}$	Minimum value of the function f and minimum element in $\mathbf{f}$, respectively
$\max f$, $\max \mathbf{f}$	Maximum value of the function f and maximum element in $\mathbf{f}$, respectively
$\arg\min_x f(x)$	Value of x which minimises f
$\arg\max_x f(x)$	Value of x which maximises f
$\nabla_{\overline{\mathbf{x}}} f(\mathbf{x})$	Complex gradient operator with respect to $\overline{\mathbf{x}}$ as defined in [1]

$$\nabla_{\overline{\mathbf{x}}} f(\mathbf{x}) = \left[\frac{\partial f}{\partial \overline{x_1}}, \frac{\partial f}{\partial \overline{x_2}}, \cdots, \frac{\partial f}{\partial \overline{x_N}} \right]^T$$

$a \in A$	a is an element of the set A
$\emptyset$	Empty set
$\mathbb{C}$	Set of all complex numbers
$\{x_i\}_{i=1}^N$	Set consisting of the elements $x_1, x_2, x_3, \ldots, x_N$ Equivalent to $\{x_1\,, x_2\,, \ldots, x_N\}$
$C = A \cup B$	Set union, so that C contains all elements from A and all elements from B
$C = A \cap B$	Set intersection, so that C contains only elements that are both in A and in B
$\triangleq$	Defined as

Functions

e^x, $\exp(x)$	Exponential function
$J_\nu(x)$	Bessel function of the first kind of order ν
$\operatorname{jinc}(x)$	Jinc-function

$$\operatorname{jinc}(x) = \begin{cases} \dfrac{J_1(x)}{x} & x > 0 \\ \dfrac{1}{2} & x = 0 \end{cases}$$

$U^o_{\mu,\nu}(\xi,\psi)$, $U^e_{\mu,\nu}(\xi,\psi)$	Zernike polynomials
$P_l^{\alpha,\beta}(x)$	Jacobi polynomial
$\Gamma(x)$	Gamma function

$\delta(x)$	Dirac-delta function which has the property $$f(a) = \int_{-\infty}^{\infty} f(x)\delta(x-a)\,\mathrm{d}x$$
$\delta(x_1, x_2, \ldots, x_N)$	Multi-dimensional Dirac-delta $\delta(x_1, x_2, \ldots, x_N) = \delta(x_1)\delta(x_2)\cdots\delta(x_N)$
$\delta_{m,n}$	Kronecker delta $$\delta_{m,n} = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases}$$

Constants

π	Ratio of a circle's circumference to its diameter, $\pi = 3.1415926535\ldots$
e	Euler's number, $e = 2.7182818284\ldots$
j	Imaginary unit, $j = \sqrt{-1}$
$\mathbf{1}$	Identity matrix
$\mathbf{0}$	Zero vector
$\mathbf{0}$	Zero matrix

Frequently used variables

The following is a list of symbols that are frequently used to represent specific entities. It shall be noted in the text where a symbol represents a different entity as listed below and where its meaning is not clear from the context in which it is used.

x, y, z	Cartesian coordinates
ξ, ψ	Radial and azimuthal coordinates, respectively, in a polar coordinate system
r, θ, ϕ	Radial, zenith angle, and azimuthal coordinates, respectively, in a spherical coordinate system
a	Reflector antenna projected aperture radius
$\mathbf{E}$	Electric field
f	Frequency
F, $\mathbf{F}$	Scalar and vector pattern functions, respectively
$\tilde{F}$, $\tilde{\mathbf{F}}$	Scalar and vector pattern function models, respectively
$\mathbf{H}$	Magnetic field

k	Wavenumber $k = \frac{2\pi}{\lambda}$
t	Time
$\boldsymbol{\delta}$	Subreflector displacement vector in OG antenna
$\kappa(\mathbf{A})$	Condition number of matrix $\mathbf{A}$, $\kappa(\mathbf{A}) = \sigma_{\max}/\sigma_{\min}$
λ	Wavelength (also penalty factor in QPS, see Section 3.2.2)
$\boldsymbol{\lambda}$	Lagrange multipliers vector
σ	Singular value of a matrix
$\mathrm{d}\sigma$	Surface differential
Σ	Domain for surface integral

Chapter 1

Introduction

Radio astronomy was born with the accidental discovery of radio emission from sources outside the solar system by Karl G. Jansky when measuring radio frequency interference in the 1930s [2]. It was only about a decade later that the first radio map of the sky was produced by Reber using a parabolic reflector antenna [3], and the following years would see a burst of radio astronomical activities and research [4].

Since the wavelengths at radio frequencies are several orders of magnitude longer than those at optical frequencies, interferometry techniques have from the outset played an integral role in the observation of the radio sky, starting with the first two-element radio interferometer used by Ryle and Vonberg in 1946 [5]. Here a much higher angular resolution than what could be achieved with a single antenna was obtained by correlating the signals received by two antenna elements that are separated by a distance (called a *baseline*) of several wavelengths. This led to the development of aperture synthesis techniques, which is based on the Van Cittert-Zernike theorem and compiles measurements on baselines of different lengths and orientations into a single image [6; 7]. Following the success of these earlier instruments a drive for faster observations, higher resolution, and increased sensitivity eventually led to the development of larger synthesis arrays in the 1970s, including systems such as the Westerbork Synthesis Radio Telescope (WSRT) in The Netherlands [8], and the Very Large Array (VLA) in New Mexico [9].

Continuing on this line of development the idea of the Square Kilometre Array (SKA) was conceived — an ultra-sensitive radio telescope array with a total collecting area in the order of a million square metres and baselines of up to three thousand kilometres in length [10; 11]. The proposed instrument will cover an extremely wide frequency band from around 70 MHz up to 10 GHz, and will be divided into Aperture Arrays (AAs) for operation at the lower frequencies (up to about 800 MHz), and dish arrays for operation at the higher frequencies (from about 500 MHz). The dish array alone is expected to consist of between two and three thousand antennas.

It is no surprise that ambitious projects such as these have been a driving

force in the advancement of antenna and other related technologies [12], and as the SKA is nearing a final detailed design before the first phase of construction commences in a few years [13], various pathfinder projects are also being developed and constructed as demonstrators of cutting edge technology which is expected to be used in the SKA. Perhaps the most notable of these are the telescope arrays that are being constructed at the two core sites which will host the SKA; these are MeerKAT (*Meer*[1] Karoo Array Telescope) in the Karoo, South Africa [14], and ASKAP (Australian Square Kilometre Array Pathfinder) in Western Australia [15].

The capabilities of these next generation radio telescopes will surpass by far that of any existing systems and consequently the factors that will ultimately limit the performance of these future systems may be quite different from those for existing systems [16]. One such aspect that needs to be addressed specifically, is the direction-dependent calibration of the instrument which compensates for the various manners in which the incoming signal of a distant radio source is affected as it propagates towards and through the different receiving elements in the interferometer array [17]. Strategies to address this problem are currently a much discussed and publicised topic within the radio astronomy community [18; 19; 20; 21]

1.1 The Purpose of this Study

Calibration in radio astronomy is a multi-faceted problem involving the modelling of calibrator radio sources in the sky, compensating for variable atmospheric conditions, and correcting for the response characteristics of the observational instrument itself [22]. This includes calibrating for the radiation pattern of the antenna, which in turn requires an accurate description of that radiation pattern at the time of an observation. For the extremely sensitive future instruments being developed this poses a difficult challenge as the temporal and station to station variation of radiation patterns resulting from varying operating conditions may impact significantly on the calibration accuracy, and will have to be accounted for [23]. Depending on the length of an observation, calibration may need to be performed a number of times during the course of that observation, which precludes time consuming measurements of the radiation pattern at many positions over a wide Field of View (FoV). This motivates the use of pattern models that accurately characterise the radiation pattern after solving for a small number of model parameters [24]. Furthermore, the number of calibration parameters should be kept at a minimum, as the noise in the image produced from radio interferometric data increases with the number of calibration parameters [25; 26].

It is this problem — the accurate characterisation of the antenna radiation pattern through models that contain as few as possible solvable parameters —

[1] *Meer* is the Afrikaans word for *more*.

that is addressed in this work.

1.1.1 Novel Contributions of this Work

Three different proposed solutions to this problem, and the novel contributions of each, are as follows.

The first solution is the use of the series expansions in [27; 28] as an analytic pattern model. These expansions were originally developed as an efficient interpolation technique to compute the radiation pattern of reflector antennas, and the expansion coefficients are usually solved from the knowledge of the primary (feed) pattern[2]. However, in the present application the coefficients are solved directly from sparse sampling of the secondary pattern [29]. Obtaining high accuracy through this approach poses some challenges, which are overcome by developing a slightly different form of the expansion functions and through the use of constrained solutions to solve for the expansion coefficients. The latter approach also results in a significant reduction in the number of pattern samples required to solve for the expansion coefficients.

The second solution uses a recently proposed technique which employs numerical expansion functions called Characteristic Basis Function Patterns (CBFPs), and which was demonstrated to yield highly accurate models of the element patterns in a strongly-coupled array [30; 31]. Herein this technique is applied to compensate for the non-linear pattern variations that result from mechanical deformations in a single-pixel dual-reflector antenna system [32]. Specifically, the method is shown to accurately predict the patterns resulting from displacement of the feed and subreflector as well as deformations of the main and subreflectors.

Finally, the third solution is based on reducing the calibration complexity of a Phased Array Feed (PAF)[3] based system through constrained beamforming [34; 35]. It is shown here how the use of a physics-based analytic pattern function, which is fitted to patterns obtained with a beamforming scheme that maximises directivity, can be employed to produce patterns with a constrained beamformer that are accurately characterised by a single-term analytic function [36; 37].

[2] Within the radio astronomy community the term *primary beam* is often used to refer to the radiation pattern of a single telescope antenna (e.g. dish or aperture array station) on the sky, whereas *Point Spread Function (PSF)* or *synthesised beam* is used to refer to the interference pattern. Amongst antenna engineers the term *primary pattern* refers to the feed pattern illuminating the reflector, and the *secondary pattern* refers to the reflector antenna pattern, after reflection, on the sky. Herein the latter terminology is adopted.

[3] Again, the possible confusion that may exist surrounding the use of the term *phased array feed* is pointed out. The term *phased array* often refers to a large array with uniform amplitude and linearly varying phase excitation [33, § 3.8]. However, within the radio astronomy community the term *phased array feed* is used to refer to a reflector antenna feed which consists of densely packed antennas, which is also how this term is used herein.

The various approaches stated above are presented here within the context of a proposed optical design of the MeerKAT reflector antenna, which features a circular aperture offset Gregorian dual-reflector system, although many of the developed techniques may generally be applied to other reflector antenna designs, or even aperture array stations [38; 39]. Furthermore, the focus is on characterising the spatial dependence of radiation patterns; temporal variation is addressed implicitly in that few calibration parameters need to be solved in order to estimate the radiation pattern so that the calibration intervals may be reduced, whereas pattern variation over frequency is mostly considered to fall outside the scope of this work and is recommended as a topic for future research.

1.2 Document Outline

In Chapter 2 the problem of direction-dependent calibration for the radiation pattern of a radio telescope is formulated through an overview of the Radio Interferometer Measurement Equation (RIME). The need for an accurate model with a few solvable coefficients is also motivated and the desired form for such a model is stated.

The ensuing Chapter 3 uses this desired pattern model as a starting point to develop two sets of analytic basis functions with which accurate pattern models are constructed. Following the derivation of these pattern models, various methods to solve for the model coefficients are presented and compared.

In Chapter 4 the CBFP method is presented. Here the generation of basis functions to compensate for specific pattern variations is discussed and the excellent modelling capabilities of such basis functions are demonstrated. In conclusion a comparison of these numerical basis functions and the analytic basis functions of the previous chapter is presented.

Chapter 5 starts with an overview of beamforming techniques for a PAF based reflector antenna. A physics-based analytic pattern model is then derived and used to define directional constraints in a Linearly Constrained Minimum Variance (LCMV) beamformer to produce secondary patterns that are accurately modelled with the analytic pattern model. The performance of this beamforming scheme is compared to that of a Maximum Directivity (MaxDir) beamformer.

Finally, in Chapter 6 some concluding remarks are presented. There the novel contributions of this work are reviewed and some topics for future research are proposed.

Chapter 2

Calibration of Reflector Antenna Radio Telescopes

The calibration challenges of future radio telescopes were briefly mentioned in the previous chapter, and here a more detailed discussion is presented in order to provide the necessary context for the work that comprises the remainder of this thesis. The objective here is not to discuss in detail any of the number of calibration algorithms of varying levels of sophistication and maturity that exist, but rather to illustrate some of the general underlying principles of calibration.

The starting point of this discussion will be an introduction to the Radio Interferometer Measurement Equation (RIME) [40] which will serve to illustrate the need for an accurate radiation pattern model of the antennas in an interferometer array. Following this a study of the variability of the radiation pattern of a reflector antenna under the range of expected operating conditions is considered. These results will motivate the need for routine pattern calibrations during the course of an observation. Thereafter the preferred form of the pattern model will be stated to set the scene for the following chapters in which various such pattern models are presented.

The work herein is presented within the context of the MeerKAT reflector antenna, and before we proceed with the above mentioned outline, a brief overview of this antenna and the use thereof in this work is discussed.

2.1 MeerKAT Antenna Overview

The MeerKAT antenna optical design features an Offset Gregorian (OG), dual-reflector system with an unblocked, circular aperture [41], and which is designed to meet the Mizugutch criteria for low-cross polarisation [42]. The parameters that define the geometry [43] are illustrated in Figure 2.1 and the value of each parameter is listed in Table 2.1.

The defining parameters are: the projected aperture diameter D_m, the main

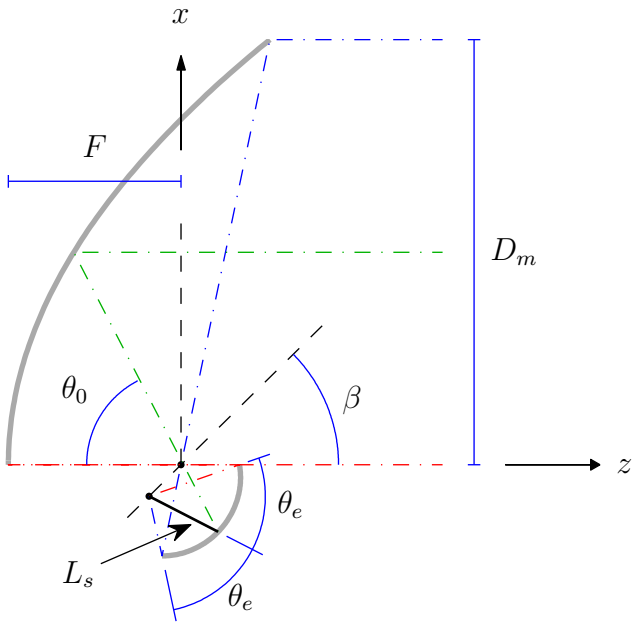


Figure 2.1: Defining parameters for offset Gregorian geometry.

Parameter designation	Parameter value
D_m	13.5 m
θ_0	63.20°
θ_e	48.89°
β	45.47°
L_s	2.419 m
F	5.486 m

Table 2.1: Geometric parameter values for MeerKAT design.

reflector offset angle θ_0, the half-angle θ_e subtended by the subreflector at the secondary focus, the tilt angle β between the axis of the main reflector and the major axis of the elliptical subreflector, the distance L_s between the secondary focus and the point of incidence of the central ray on the subreflector, and the focal length F of the main reflector.

Herein the above reflector system is analysed through simulation to calculate the secondary far-field patterns at various points within the frequency band ranging from 580 MHz to 1.75 GHz. The analysis is performed using the Physical Optics and Physical Theory of Diffraction (PO + PTD) engine available in GRASP [44]. For the purposes of this study two feed types are employed. The first represents the feed for a single-beam system and uses a Gaussian beam primary pattern (built-in feed available in GRASP [45, § 2.3.1.3]) with a 12 dB taper towards the subreflector half-angle.

The second feeding technique employs a Phased Array Feed (PAF), in

which case the analysis of the reflector antenna uses the tabulated feed functionality in GRASP [45, § 2.3.1.1] through a toolbox interface [46] implemented in Matlab [47]. The PAF consists of 121 Tapered Slot Antenna (TSA) elements arranged in a dual-polarised array [48], and radiation patterns for the elements in this feed array were obtained through simulation with a Method of Moments (MoM) code called CAESAR [49; 50]. The exact illumination of the reflector, and consequently the secondary pattern on the sky, depends on the excitation weights of the elements in the feed array, which is left as a topic for discussion in Chapter 5.

The above reflector design is used throughout and the applicable feed type will be noted in the numerical results that are presented in the remainder of this document.

2.2 The Radio Interferometer Measurement Equation

The measurement equation for radio interferometers provides a simple yet powerful mathematical framework within which the full propagation path of a radio signal from a distant source to the output of the interferometer can be described and analysed. Although the original equation was formulated using Mueller matrices [40], herein the more intuitive formulation in terms of Jones matrices is used [51], and the derivation presented here follows mostly that in [52].

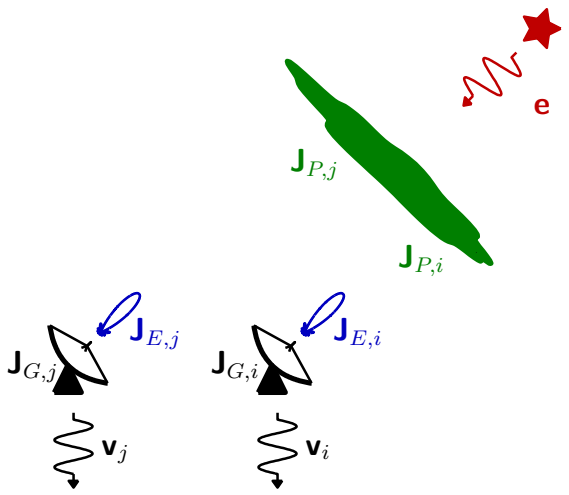


Figure 2.2: The radio interferometer measurement equation.

Consider the two-element interferometer shown in Figure 2.2. The two antennas each contain a dual-polarised feed and are pointed towards a common source which is in the far-field of the interferometer. The source radiates an electromagnetic field in all directions and the signal in the direction of the interferometer is denoted by the 2×1 vector

$$\mathbf{e} = \begin{bmatrix} e_p \\ e_q \end{bmatrix} \tag{2.2.1}$$

which represents two orthogonal components of the electric field in some polarisation system. As this signal propagates towards the interferometer it is subject to various transformations which, if it is assumed that all transformations are linear, may each be represented as 2×2 Jones matrix. The signal $\mathbf{e}_o$ after one such transformation is then related to the signal $\mathbf{e}_i$ before that transformation through

$$\begin{bmatrix} e_{o,p} \\ e_{o,q} \end{bmatrix} = \begin{bmatrix} J_{pp} & J_{pq} \\ J_{qp} & J_{qq} \end{bmatrix} \begin{bmatrix} e_{i,p} \\ e_{i,q} \end{bmatrix}. \tag{2.2.2}$$

Two or more consecutive transformations may also be combined into a single equivalent transformation

$$\mathbf{e}_o = \mathbf{J}_n \mathbf{J}_{n-1} \dots \mathbf{J}_2 \mathbf{J}_1 \mathbf{e}_i = \mathbf{J} \mathbf{e}_i \tag{2.2.3}$$

as long as the order of transformations is preserved[1]. In the above Jones chain the signal is affected by the transformations in the order $1, 2, \dots, n$.

For the purpose at hand all the transformations along the propagation path are grouped into three Jones matrices. The first of these (left-most in a three-matrix Jones chain as above) includes all effects occurring after conversion of the incident electromagnetic wave into an electrical signal at the terminals of the antenna, such as low-noise amplifier gains, cable losses, etc. These effects are unaffected by the position of the source relative to the pointing of the antenna, and the representative Jones matrix is termed the direction-*independent* gain $\mathbf{J}_G$. The actual conversion from the incident electromagnetic wave to an electrical signal at the antenna terminals is defined by the radiation pattern of the antenna, and comprises the second Jones matrix $\mathbf{J}_E$ which is the direction-*dependent* gain. In this Jones matrix the diagonal elements are the co-polarised far-field patterns of the dual-polarised antenna, and the off-diagonal elements are the cross-polarised far-field patterns. Finally, all the remaining signal transformations between the source and the antenna are subsumed in the third matrix $\mathbf{J}_P$. Included in this term are effects such as the phase delay along the path from the source to the antenna, ionospheric effects such as Faraday rotation, and parallactic rotation.

[1] A more complete summary of Jones matrix properties regarding algebraic manipulation can be found in [52]. Here only the combination property is required.

In this formalism, and assuming that the nature of the source and that of all relevant Jones matrices along the propagation path are known, the dual-polarised signal $\mathbf{v}_i$ output by antenna i on the right in Figure 2.2 is simply calculated as

$$\mathbf{v}_i = \begin{bmatrix} v_{i,p} \\ v_{i,q} \end{bmatrix} = \mathbf{J}_{G,i}\mathbf{J}_{E,i}\mathbf{J}_{P,i}\mathbf{e} \tag{2.2.4}$$

where each of the transformations is applied to the signal in the same order as that in which they occur. The signal $\mathbf{v}_j$ output by antenna j on the left in Figure 2.2 is similarly determined. However, since the propagation path towards antenna j is different to that towards antenna i the actual Jones matrices applicable to the signal at antenna j may be different from those for antenna i. For example, for a relatively long baseline separating the antennas the ionospheric effects may be quite different as the propagation paths of the signal towards the two antennas is along different paths in the atmosphere [22].

In fact, all instrument related signal transformations are prone to be different for each telescope in an interferometer array since each instrument will generally exhibit different deviations from the actual design in practice, be it a result of manufacturing tolerances or varying operating conditions. This fact is the central motivation for the work presented herein and will be considered in more detail later in this chapter.

To produce an interferometer the signals output by each of the antennas need to be correlated to form four correlation pairs, which can be arranged in the *visibility matrix*[2] $\mathbf{V}_{i,j}$ as follows

$$\mathbf{V}_{i,j} = \begin{bmatrix} v_{p,i} \star v_{p,j} & v_{p,i} \star v_{q,j} \\ v_{q,i} \star v_{p,j} & v_{q,i} \star v_{q,j} \end{bmatrix}. \tag{2.2.5}$$

Here $\star$ denotes the cross-correlation operator, which for ergodic processes allows the visibility matrix to be written as

$$\mathbf{V}_{i,j} = \begin{bmatrix} \langle v_{p,i}\overline{v_{p,j}} \rangle & \langle v_{p,i}\overline{v_{q,j}} \rangle \\ \langle v_{q,i}\overline{v_{p,j}} \rangle & \langle v_{q,i}\overline{v_{q,j}} \rangle \end{bmatrix} = \left\langle \mathbf{v}_i \mathbf{v}_j^{\dagger} \right\rangle \tag{2.2.6}$$

where $\langle . \rangle$ computes the time-average of its argument [53]. Substituting the Jones chain in (2.2.4) into the expression for the visibility matrix, using $\mathbf{J}_i = \mathbf{J}_{G,i}\mathbf{J}_{E,i}\mathbf{J}_{P,i}$, and assuming that the Jones matrices are relatively constant over the integration period of the correlator, gives

$$\mathbf{V}_{i,j} = \mathbf{J}_i \left\langle \mathbf{e}\mathbf{e}^{\dagger} \right\rangle \mathbf{J}_j^{\dagger} = \mathbf{J}_i \mathbf{B} \mathbf{J}_j^{\dagger} \tag{2.2.7}$$

where $\mathbf{B}$ is called the *brightness matrix*.

The above relation (2.2.7) between the visibility matrix as measured by an interferometer and the radiation produced by a distant radio source was

[2] The omission of constant gain factors such as the factor 2 appearing in (4) of [52] is done without consequence for the present discussion.

derived for a single point source. By superposition the interferometer response to a number of sources is simply the sum of the responses to each source individually, and in the limiting case where the intensity distribution on the sky is a continuous function the summation becomes integration over the portion of the sky S which is visible to the interferometer. The visibility can then be stated as (cf. Equation (18) in [52])

$$\mathbf{V}_{i,j} = \mathbf{J}_{G,i} \left(\int_S \mathbf{J}_{E,i} \mathbf{J}_{P,i} \mathbf{B} \mathbf{J}_{P,j}^{\dagger} \mathbf{J}_{E,j}^{\dagger} \, \mathrm{d}\Omega \right) \mathbf{J}_{G,j}^{\dagger} \tag{2.2.8}$$

where all direction-independent terms have been removed from under the integral. The RIME is now in a form that will suffice for the discussion on calibration which follows.

2.2.1 Calibration

In the derivation of (2.2.8) we were mostly concerned with the forward calculation of the RIME, that is, to determine the output of the interferometer given a certain brightness distribution on the sky and fully known propagation transformations towards the receiver. However, the primary task of the radio astronomer is to determine the nature of the radio sky, and in this regard the backward calculation is of more importance: given the data output by the interferometer, what is the brightness distribution on the sky? Answering this question lies at the heart of calibration (and imaging) in radio astronomy [17].

From (2.2.8) it is apparent that the correction for direction-independent transformations $\mathbf{J}_G$ is relatively simple, whereas the same for the direction-dependent transformations $\mathbf{J}_E$ and $\mathbf{J}_P$ is a much more complicated problem. Many approaches to solving this have been proposed and an overview of many of these to date can be found in [17, § 2]. What should be clear is that whichever strategy is used to recover $\mathbf{B}$ from $\mathbf{V}_{i,j}$ knowledge of the various Jones terms $\mathbf{J}_G$, $\mathbf{J}_E$ and $\mathbf{J}_P$ (or at least their combined effect which may be represented as a single Jones matrix $\mathbf{J}$) is required.

Depending on the calibration algorithm which is used some direction-dependent effects are modelled with a predetermined Jones term based on a priori knowledge, whereas the Jones terms for other effects may be parametrised and solved using interferometry data. In the next section the implications of assuming a predetermined direction-dependent gain term $\mathbf{J}_E$ for a reflector antenna radio telescope are explored.

2.2.2 Pattern Variability in Changing Operating Conditions

The direction-dependent gain consists of the four far-field components

$$\mathbf{J}_E = \begin{bmatrix} F_p^{\mathrm{CP}}(\theta, \phi) & F_p^{\mathrm{XP}}(\theta, \phi) \\ F_q^{\mathrm{XP}}(\theta, \phi) & F_q^{\mathrm{CP}}(\theta, \phi) \end{bmatrix} \tag{2.2.9}$$

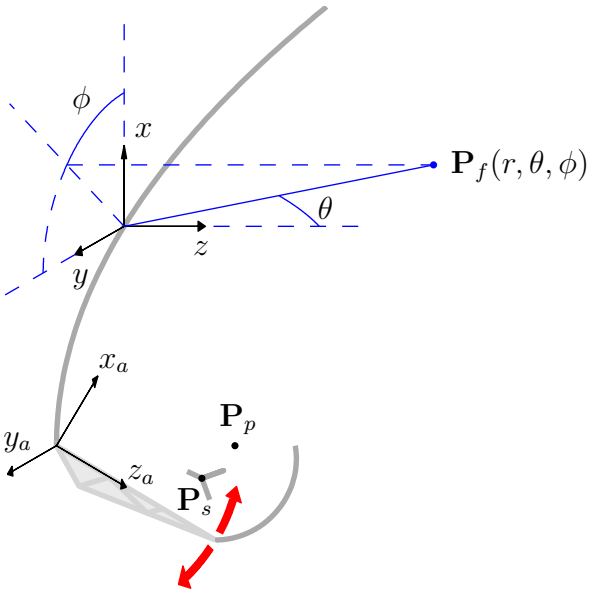


Figure 2.3: Mechanical deformations in the MeerKAT OG antenna.

where the superscripts CP and XP indicate co- and cross-polarisation components, respectively, and p and q represent two orthogonal polarisation components. Herein the polarisation convention corresponding to the third definition in [54] is used. It is assumed here that $\mathbf{J}_E$ is determined through measurement under ideal operating conditions using the correlation measurement technique in [55; 23].

An illustration of some of the expected mechanical deformations in the MeerKAT OG system is shown in Figure 2.3. Due to changes in temperature, orientation of the antenna (and therefore gravitational loading), and wind loading the structure is subject to various mechanical deformations, which affect the accuracy of both reflector surfaces, as well as the position of the feed and the subreflector. Assume here that the deformations are limited to those affecting the support arm, as indicated in the figure. Then the tolerances on the position of the feed and subreflector are as provided in Table 2.2 [56], and defined in the coordinate system (x_a, y_a, z_a) where $\hat{\mathbf{z}}_a$ is directed along the length of the support arm from the bottom of the main reflector to the bottom of the subreflector. Note also the position of the global coordinate system (x, y, z) which is located on the main reflector and in the centre of the projected aperture. Herein all secondary patterns are referenced to this point, with the spherical far-field coordinates (θ, ϕ) as defined in the figure.

The variations in the co-polarised pattern resulting from the deformations are shown in Figure 2.4 for a linear-x polarised feed at 1.42 GHz. Here various feed and subreflector displacements in the (x_a, y_a)-plane were applied to the single-beam antenna and the resulting far-field patterns computed. The shaded area in Figure 2.4 indicates the range of resulting patterns as a function of θ

Direction	Subreflector	Feed
$\hat{\mathbf{x}}_a$	± 10.0 mm	± 7.7 mm
$\hat{\mathbf{y}}_a$	± 5.0 mm	± 3.9 mm
$\hat{\mathbf{z}}_a$	± 20.0 mm	± 15.4 mm

Table 2.2: Estimated displacement tolerances for MeerKAT subreflector and feed.

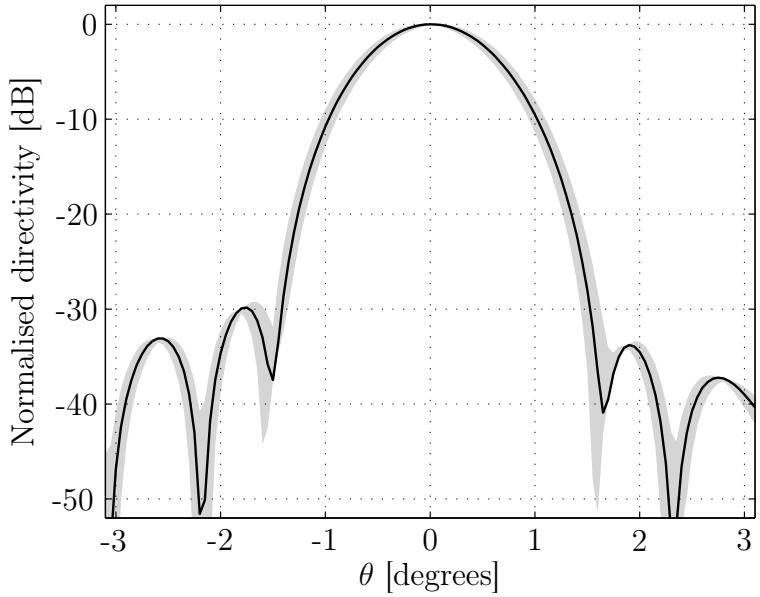


Figure 2.4: Co-polarised pattern variation due to feed and subreflector displacement, indicated as the shaded area. The pattern corresponding to the ideal geometry is shown as the solid line.

in the $\phi = 0°$ plane, whereas the solid line shows the pattern corresponding to the ideal geometry. As the results show this particular deformation mainly results in a pointing error, which can be expected for displacement of the feed or subreflector.

A greater appreciation of these results may be gained through a more quantitative analysis of the pattern variations, and to this end the difference (or *error*) between the pattern of the ideal geometry and that corresponding to a deformed geometry is calculated. Herein two error metrics are used. The first error used is the power in the error pattern normalised to the power in the total ideal pattern and computed over a certain angular region

$$\epsilon_N(\theta_M) = \frac{\int_{S_M} |F - F_\epsilon|^2 \, \mathrm{d}\Omega}{\int_{S_M} |F|^2 \, \mathrm{d}\Omega} \tag{2.2.10}$$

where F and F_ϵ are complex-valued voltage pattern functions (either the co- or cross-polarised component) corresponding to the ideal and erroneous ge-

ometries, respectively. The angular region over which the error is computed is

$$S_M = \{(\theta, \phi) : (\theta, \phi) \in S,\, \theta \leq \theta_M\}\,. \tag{2.2.11}$$

The second error is computed as the average relative error magnitude in the complex-valued voltage pattern

$$\epsilon_A(L) = \frac{1}{N_L} \sum_{\substack{i=1 \\ (\theta_i,\phi_i)\in S_L}}^{N_S} \left| \frac{F(\theta_i, \phi_i) - F_\epsilon(\theta_i, \phi_i)}{F(\theta_i, \phi_i)} \right| \tag{2.2.12}$$

where F and F_ϵ are the same quantities as in (2.2.10), (θ_i, ϕ_i) are the positions where the patterns were computed, and N_S is the total number of sampling points. In order to prevent this error from exploding as a result of small values of the denominator near pattern nulls the computation of this error will be limited to those points where the pattern function F is above or equal to a certain level L relative to maximum and expressed in decibels

$$S_L = \left\{ (\theta_i, \phi_i) : (\theta_i, \phi_i) \in S,\, 20 \log_{10} \left(\frac{F(\theta_i, \phi_i)}{\max F(\theta, \phi)} \right) \geq L \right\}. \tag{2.2.13}$$

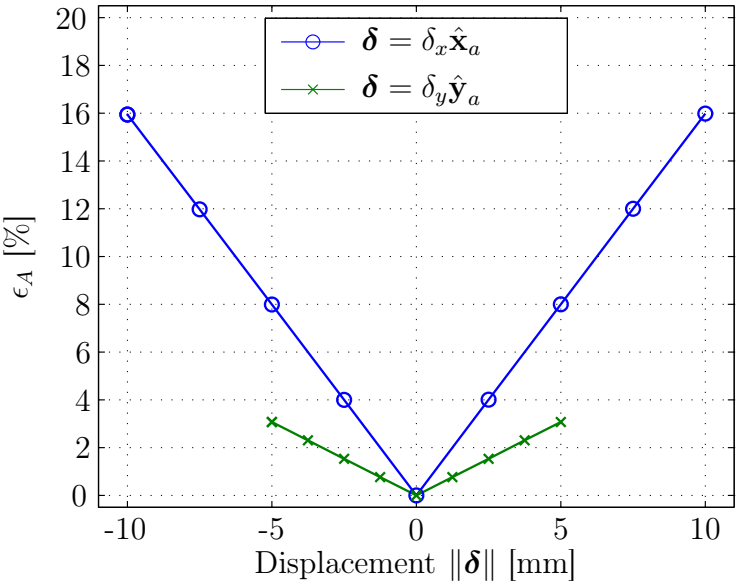
N_L is the number of sampling points contained in S_L. This second error metric is similar to the Figure of Merit (FoM) defined in [23, § 2.E] as (using the present notation)

$$\epsilon_R = \sqrt{\frac{1}{N} \sum_{i}^{N} \left| \frac{F(\theta_i, \phi_i) - F_\epsilon(\theta_i, \phi_i)}{F(\theta_i, \phi_i)} \right|^2}. \tag{2.2.14}$$

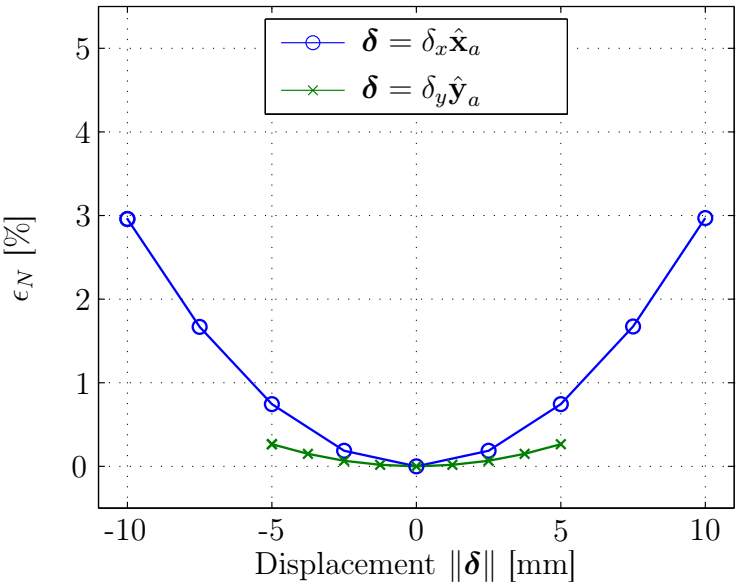
The errors between the ideally expected pattern and the actual patterns (co-polarised components) as a function of displacement towards different directions are shown in Figure 2.5. As expected the error steadily increases as the magnitude of the displacement increases, and the rate of increase is higher for displacement along $\hat{\mathbf{x}}_a$ than for displacement along $\hat{\mathbf{y}}_a$. In addition to this the range of expected displacement along $\hat{\mathbf{x}}_a$ is also larger than along $\hat{\mathbf{y}}_a$ so that the largest errors can be expected for displacements in the symmetry plane (xz-plane). The results obtained using either error metric may be compared. The average relative error ϵ_A shown in Figure 2.5 (a) varies nearly linearly as the displacement increases. On the other hand, the error ϵ_N in Figure 2.5 (b) is seen to increase more or less quadratically since this error is proportional to power and the average error ϵ_A is proportional to voltage.

To put the results of Figure 2.5 (a) in perspective, it is useful to consider the relation between the FoM in (2.2.14) and image fidelity, which is a measure of the relative error in the image of the radio sky produced from interferometric data [57]. In [23] the image fidelity is *optimistically* related to the FoM as

$$\text{Image Fidelity} = \frac{\epsilon_R}{\sqrt{N_A}} \tag{2.2.15}$$



(a) Average relative error ϵ_A



(b) Normalised error power ϵ_N

Figure 2.5: Error in co-polarisation pattern of deformed geometry relative to pattern for ideal geometry.

where N_A is the number of antennas in the interferometer array. Achieving an image fidelity of 10^{-4} with the SKA [23] and using $N_A = 3000$ for the dish array [10] requires that $\epsilon_R < 0.55\%$. This means that even the smallest displacement along $\hat{\mathbf{y}}_a$ in Figure 2.5 (a) results in an unacceptably high error. Note that the use of the root-mean-square in the definition of ϵ_R as opposed to the absolute-mean used to compute ϵ_A means that the restriction on ϵ_A should

be even more stringent.

Despite the fact that the relation between errors in the final image and errors in the beam model is not as simple, and the fact that the required accuracy depends also on the science case for a particular observation, the above calculated result still serves as an indication that compensating for pattern variations such as those in Figure 2.4 is essential. In order to achieve this the direction-dependent gain Jones term should be parametrised in some way so as to allow solution using interferometric data. A general preferred form for such a pattern model is considered in the next section.

2.3 Modelling the Direction-Dependent Antenna Gain

Performing the calculations required to invert (2.2.8) requires a function which accurately describes the direction-dependent antenna gain pattern. As was shown in the previous section the radiation pattern of even a single-beam reflector antenna may exhibit significant variation with changing operating conditions. For PAF based systems such variation is exacerbated by the fact that the feed array (primary) pattern also varies with electronic drift, causing the illumination of the reflector, and therefore also the secondary pattern to vary.

To compensate for such variations the function describing the radiation pattern (hereafter, referred to as the *pattern model*) should contain a number of solvable parameters which are determined at the time of observation. The preferred form of the pattern model $\mathbf{J}_E$ can be expressed as [24]

$$\mathbf{J}_E(\theta, \phi, t, f) = \sum_{k=1}^{K} \mathbf{x}_k(t, f) \circ \mathbf{f}_k(\theta, \phi, t, f), \tag{2.3.1}$$

where $\circ$ is used to denote the Hadamard (element-wise) matrix product, $\{\mathbf{f}_k\}_{k=1}^{K}$ is an appropriate set of basis functions, and $\{\mathbf{x}_k\}_{k=1}^{K}$ are the unknown model parameters that need to be solved[3]. Note that the basis functions and weighting coefficients in the above model are 2×2 matrices as is required to model the direction-dependent gain Jones matrix in (2.2.9).

Ideally the basis functions $\mathbf{f}$ in (2.3.1) would be selected such that the error between the pattern model and the actual radiation pattern of the antenna is minimised for a given number of terms. This emphasises the use of basis functions that contain information related to the physics of the antenna system. In addition it is preferred that a trade-off exists between the number of terms K

[3] In the referenced source the spatial dependence of the pattern models are defined using direction-cosines. For convenience, herein the pattern models are obtained in the spherical coordinate system shown in Figure 2.3. A simple conversion to direction-cosines [5, § 3.1] then transforms the model to the appropriate expression.

and the model accuracy. This would allow the model to be adjusted depending on whether accuracy or processing time is of more importance.

The form expressed in (2.3.1) is quite different from some of the models that have been used to approximate the radiation patterns of existing radio telescopes. For example, the co-polarisation voltage pattern main beam of the Westerbork Synthesis Radio Telescope (WSRT) is often approximated at GHz frequencies as [58]

$$F^{\mathrm{CP}}(\theta,\phi) = \cos^3\left(Cf\theta\right), \tag{2.3.2}$$

where C is slowly varying over frequency[4]. This model is only accurate down to about 10% of the main lobe. Another example is a pattern model for the VLA, which approximates the antenna pattern (above 1 GHz) down to the 5% level by the jinc-function [60; 61]

$$F^{\mathrm{CP}}(\theta,\phi) = \frac{J_1(C\theta)}{C\theta}, \tag{2.3.3}$$

where J_1 is the Bessel function of the first kind of order one and C is a constant. The main advantage of (2.3.1) is that with an appropriate choice of basis functions, pattern features such as beam asymmetry and complex sidelobe structures may be modelled accurately. This is not the case for the models in (2.3.2) and (2.3.3); these functions are by definition constant in ϕ (circularly symmetric). Furthermore, the sidelobe structures that are contained in these models are very simplistic and may even not be realistic at all!

Of course, the benefits of the model in (2.3.1) comes at the cost of having a number of terms, and therefore a number of unknowns that need to be solved. This emphasises the need for basis functions that are able to achieve a highly accurate pattern model with as few terms as possible.

The next two chapters are devoted to the development of such basis functions. Throughout it will be assumed that the model coefficients are solved using the correlation measurement method [55] to determine the actual radiation pattern at a few positions and at the time of an observation. This method is only used to illustrate that an accurate characterisation of the antenna radiation pattern can be obtained through the solution of only a few parameters, and the incorporation of the presented pattern models in calibration algorithms such as those discussed in [17, § 2] is left for future work.

Furthermore, it will be assumed that the different models are required to characterise the radiation pattern over an angular region that extends beyond the main lobe and which includes the first few sidelobes. Such characterisation may be necessary to meet the dynamic range limitations of future systems [16]. For example, specifications on the first sidelobe level of the MeerKAT antenna were relaxed, the idea being that sources observed in this region would be

[4] The WSRT power beam is also modelled in practice (e.g. in the NEWSTAR interferometry software package) as $|F^{\mathrm{CP}}(\theta,\phi)|^2 = \max\left[\cos^6\left(Cf\theta\right), 0.01\right]$ out beyond the main lobe [59].

taken into account during calibration/imaging [62]. This obviously requires a pattern model that includes the first sidelobe.

2.4 Conclusion

Herein the general calibration problem was illustrated through a brief introduction to the RIME, and the need for an accurate model that describes the radiation pattern of the antennas in an interferometer array was motivated. It was also shown that pattern variability resulting from changes within the range of expected operating conditions may result in unacceptably high errors in an assumed fixed pattern model, and necessitates the use of a solvable pattern model. The preferred form of such a model was stated as a weighted sum of basis functions in which the weighting coefficients are the unknown model parameters that need to be solved.

In such a model an appropriate set of basis functions is required to minimise the number of terms necessary to achieve the desired level of accuracy, and in the next two chapters different such bases are considered. In Chapter 3 the use of analytic basis functions is presented, followed by the use of numerical basis functions in Chapter 4.

Chapter 3

Analytic Pattern Models

The need for an accurate description of the radiation pattern of a radio telescope antenna was discussed in the previous chapter, and here analytical pattern models are developed for that purpose.

Obtaining the radiation pattern of an antenna in the form of an analytic function is very seldom possible, and even then simplistic models of the antenna are used in many instances in order to solve the pertaining equations in closed form. For aperture antennas such a solution usually employs Huygens' principle [63, § 12] and calculates the far-field pattern from a simplified description of the field distribution over the antenna aperture. This procedure is also applicable to reflector antennas and will be used here to derive an analytic pattern model.

In order to be able to represent a general far-field pattern the models presented here are derived from an expansion of the aperture fields in a series of functions which form a complete and orthogonal set over the antenna aperture. Once the pattern models are defined, various approaches toward solving the model parameters for a particular pattern are developed. Numerical results are presented to evaluate the efficiency of the pattern models and the various model parameters solutions, followed by some conclusions.

3.1 Analytic Pattern Model Derivation

A widely used basis for expanding the aperture distribution over circular apertures are the Zernike polynomials [64, § 9.2.1] which model azimuthal and radial dependence using trigonometric Fourier and Jacobi polynomial series, respectively. Using such an aperture field expansion yields a far-field pattern model which contains a sum of Bessel functions of increasing order, which will be referred to herein as the *Jacobi-Bessel pattern model* or JB-model [27; 28]. Historically this result has been used widely as an efficient method by which the radiation pattern of large reflector antennas could be calculated, and even recent implementations appear in literature [65; 66].

Within the radio interferometry calibration community this analytic result has also gained attention recently, and a comparison with other similar models may be found in [38; 39]. Therein the JB-model was shown to exhibit superior convergence when compared to a model based on a Fourier-Bessel expansion of the aperture fields, and the present study is restricted to models based on a Zernike polynomial expansion.

In what follows it will be shown that JB-model far-field model also has certain shortcomings, and to counter these a second analytic far-field pattern is derived by modifying the definition of the JB-model. This second analytic far-field pattern will be referred to herein as the *Neumann pattern model* or N-model.

3.1.1 Aperture Field to Far-Field Transformation

Consider the reflector antenna shown in Figure 3.1.

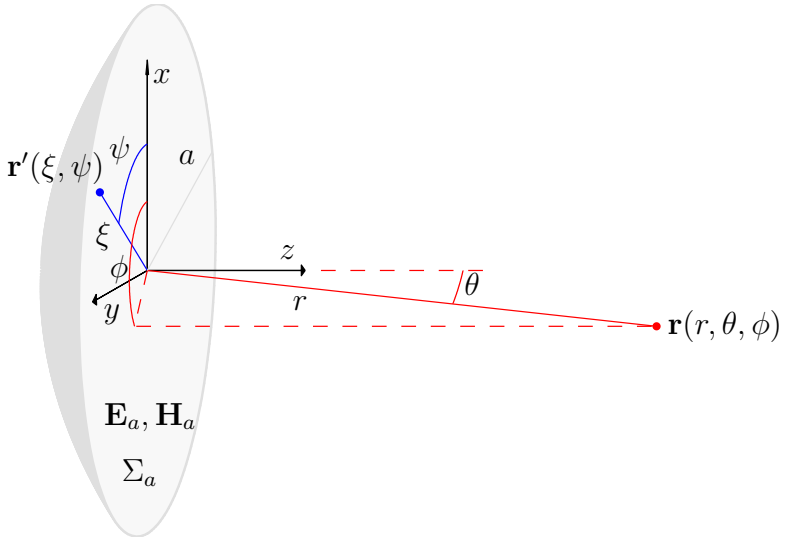


Figure 3.1: Aperture field to far-field transformation.

Let the electric field $\mathbf{E}_a$ radiated by the antenna be known on the projected aperture Σ_a. In the far-field approximation the radiated electric field $\mathbf{E}$ at a location $\mathbf{r}(r,\theta,\phi)$ can then be calculated to be [63, § 12.3] (see Appendix A)

$$\mathbf{E}(\mathbf{r}) \approx -jk\frac{e^{-jkr}}{2\pi r}\int_{\Sigma_a} \hat{\mathbf{r}} \times \hat{\mathbf{z}} \times \mathbf{E}_a(\mathbf{r}')e^{j\mathbf{k}\cdot\mathbf{r}'}\,\mathrm{d}\sigma \tag{3.1.1}$$

where $\hat{\mathbf{z}}$ is the unit vector normal to the aperture plane, $\hat{\mathbf{r}}$ is a unit vector towards the far-field point, $\mathbf{r}'(\xi,\psi)$ is a point in the aperture plane

$$\mathbf{r}' = \xi\left(\cos\psi\hat{\mathbf{x}} + \sin\psi\hat{\mathbf{y}}\right) \tag{3.1.2a}$$

expressed in the polar coordinate system (ξ, ψ) and $\mathbf{k}$ is the wave vector

$$\mathbf{k} = k\hat{\mathbf{r}} = k\left(\cos\phi\sin\theta\hat{\mathbf{x}} + \sin\phi\sin\theta\hat{\mathbf{y}} + \cos\theta\hat{\mathbf{z}}\right). \tag{3.1.2b}$$

Letting $\mathbf{F}_a(\xi, \psi) = \hat{\mathbf{r}} \times \hat{\mathbf{z}} \times \mathbf{E}_a(\mathbf{r}')$ and substituting (3.1.2) in (3.1.1) gives

$$\mathbf{E}(\mathbf{r}) \approx -jk\frac{e^{-jkr}}{2\pi r}\mathbf{F}(\theta, \phi) \tag{3.1.3a}$$

where the far-field pattern function $\mathbf{F}(\theta, \phi)$ is given by

$$\mathbf{F}(\theta, \phi) = \int_0^{2\pi}\int_0^a \mathbf{F}_a(\xi, \psi)e^{jk\xi\sin\theta\cos(\psi-\phi)}\xi\,\mathrm{d}\xi\,\mathrm{d}\psi. \tag{3.1.3b}$$

The objective in this case is to approximate the far-field pattern function $\mathbf{F}$ as a sum of analytic functions by expanding the function $\mathbf{F}_a$ in a manner which allows analytic evaluation of (3.1.3b). The procedure is carried out in the next section using the Zernike polynomials to yield the JB-model.

3.1.2 Jacobi-Bessel Pattern Model

The Zernike polynomials [64, § 9.2.1] are defined as

$$\left\{\begin{array}{c} U^o_{\mu,\nu}(\rho, \psi) \\ U^e_{\mu,\nu}(\rho, \psi) \end{array}\right\} = R_{\mu,\nu}(\xi)\left\{\begin{array}{c} \sin\nu\psi \\ \cos\nu\psi \end{array}\right\} \tag{3.1.4a}$$

$$R_{\mu,\nu}(\rho) = \begin{cases} (-1)^{(\mu-\nu)/2}\rho^\nu P^{(\nu,0)}_{(\mu-\nu)/2}\left(1-2\rho^2\right) & \mu \geq \nu \geq 0 \text{ and } (\mu-\nu) \text{ even}, \\ 0 & \text{otherwise} \end{cases} \tag{3.1.4b}$$

where $P^{(\alpha,\beta)}_l$ is a Jacobi polynomial and the polar coordinate system (ρ, ψ) is defined such that $\xi = a\rho$. The Zernike polynomials are orthogonal on, and form a complete set over the unit circle $\rho \leq 1$ which coincides with the circular antenna aperture $\xi \leq a$. A well-behaved function $\mathbf{F}_a(\xi, \psi)$ then allows expansion as a sum of Zernike polynomials to give

$$\mathbf{F}_a(\xi, \psi) = \sum_{\nu=0}^{\infty}\sum_{\substack{\mu=\nu \\ \mu-\nu \text{ even}}}^{\infty} \mathbf{A}_{\mu,\nu}U^o_{\mu,\nu}(\tfrac{\xi}{a}, \psi) + \mathbf{B}_{\mu,\nu}U^e_{\mu,\nu}(\tfrac{\xi}{a}, \psi) \tag{3.1.5a}$$

for which the coefficients may be calculated using

$$\left\{\begin{array}{c} \mathbf{A}_{\mu,\nu} \\ \mathbf{B}_{\mu,\nu} \end{array}\right\} = \frac{\mu+1}{\epsilon^2_{\mu,\nu}\pi}\int_0^1\int_0^{2\pi}\mathbf{F}_a(a\rho, \psi)\left\{\begin{array}{c} U^o_{\mu,\nu}(\rho, \psi) \\ U^e_{\mu,\nu}(\rho, \psi) \end{array}\right\}\rho\,\mathrm{d}\rho\,\mathrm{d}\psi \tag{3.1.5b}$$

with

$$\epsilon_{\mu,\nu} = \begin{cases} \frac{1}{\sqrt{2}} & \nu = 0, \mu \neq 0 \\ 1 & \text{otherwise.} \end{cases}$$

In the above the vector nature of $\mathbf{F}_a$ is expressed through the vector valued coefficients $\mathbf{A}_{\mu,\nu}, \mathbf{B}_{\mu,\nu}$. Given the restrictions on μ, ν in (3.1.4b) it is useful to introduce the indexing variables

$$m = \frac{\mu - \nu}{2} \tag{3.1.6a}$$

$$n = \nu \tag{3.1.6b}$$

which permits simplification of the summation variables in (3.1.5a). Substitution of this sum in (3.1.3b) gives the desired far-field pattern model [27; 28]

$$\begin{aligned}\mathbf{F}(\theta,\phi) &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\left(\mathbf{A}_{2m+n,n}\sin n\phi + \mathbf{B}_{2m+n,n}\cos n\phi\right)\frac{J_{n+2m+1}(ka\sin\theta)}{ka\sin\theta} \\ &\triangleq \tilde{\mathbf{F}}^{(\mathrm{JB})}(\theta,\phi).\end{aligned} \tag{3.1.7}$$

Note that since $\sin n\phi = 0$ for $n = 0$ we can immediately state that $\mathbf{B}_{m,0} = \mathbf{0}$. For practical evaluation of the function on the right hand side the summation needs to be terminated at $n = N, m = M$. The special case of the far-field pattern for a uniform aperture field distribution is obtained with $M, N = 0$, in which case (3.1.7) reduces to the jinc-function.

Note that the model is undefined for $\theta = 0$ since the denominator is then equal to zero. However, for Bessel functions of the first kind it can be shown that [67, § 3.1] (see Appendix B.1)

$$\lim_{u\to 0}\frac{J_q(u)}{u} = \begin{cases}\frac{1}{2} & q = 1 \\ 0 & q > 1\end{cases} \tag{3.1.8}$$

and herein expressions of the form $J_\nu(u)/u$ will be understood as assuming the limiting value stated above for $\theta = 0$.

3.1.3 Neumann Model

Towards developing the second analytic model we note that the JB-model expands the far-field function into azimuthal (n-modes) which can each be separated as

$$\tilde{\mathbf{F}}_n^{(\mathrm{JB})}(\theta,\phi) = \Phi_n(\phi)\boldsymbol{\Theta}_n(ka\sin\theta) \tag{3.1.9}$$

with Φ_n either a sine or cosine function, and the zenith function $\boldsymbol{\Theta}_n$ which can be expressed as

$$\mathbf{U}_n(u) = u\boldsymbol{\Theta}_n(u) = \sum_{m=0}^{\infty}\mathbf{a}_{m,n}J_{n+2m+1}(u) \tag{3.1.10}$$

after multiplying with $u = ka\sin\theta$. The expansion on the right-hand side of (3.1.10) is a particular case of the Neumann Series [67, § 16.1] which has

the more general form

$$\mathbf{U}_n(u) = \sum_{k=0}^{\infty} \mathbf{a}_k J_k(u). \tag{3.1.11}$$

The requirements for $\mathbf{U}_n$ to have an expansion of the form (3.1.10) are more restrictive as compared to the expansion in (3.1.11) [67, § 16.4]. Indeed, the only requirement to ensure that $\mathbf{U}_n$ admits an expansion of the form in (3.1.11) within the region $u \leq R$, is that $\mathbf{U}_n$ be analytic within this region (see Appendix B.2). Therefore it makes sense that the far-field function $\mathbf{F}$ should also allow expansion employing the more general form of the Neumann series[1].

Noting from the left-hand side of (3.1.10) that $\mathbf{U}_n(0) = 0$ we may immediately determine that the coefficient of the first term on the right-hand side of (3.1.11) is equal to zero, $\mathbf{a}_0 = \mathbf{0}$ since $J_0(0) = 1$. For the second analytic far-field pattern model we therefore have

$$\begin{aligned}\mathbf{F}(\theta,\phi) &= \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} \left(\mathbf{A}_{m,n}\sin n\phi + \mathbf{B}_{m,n}\cos n\phi\right) \frac{J_{m+1}(ka\sin\theta)}{ka\sin\theta} \\ &\triangleq \tilde{\mathbf{F}}^{(\mathrm{N})}(\theta,\phi).\end{aligned} \tag{3.1.12}$$

Note that the above summation will result in a $J_1(u)/u$ term for each n-mode. From (3.1.8) we realise that this will result in a higher order azimuthal Fourier expansion for the pattern at $\theta = 0$. However, since the pattern is single-valued at this point only the $n = 0$ term is required, and the coefficients $\mathbf{A}_{0,n}, \mathbf{B}_{0,n} = \mathbf{0}$ for $n > 0$. Also, as for the JB-model we have $\mathbf{B}_{m,0} = \mathbf{0}$.

Comparing this expression to that of $\tilde{\mathbf{F}}^{(\mathrm{JB})}$ we note two differences, specifically in the order of the Bessel functions. Firstly, for a particular azimuthal mode the zenith variation in $\tilde{\mathbf{F}}^{(\mathrm{N})}$ is modelled using functions from order one, whereas in $\tilde{\mathbf{F}}^{(\mathrm{JB})}$ only functions of order $n+1$ and higher are used. This means that the N-model is better adapted than the JB-model to express higher-order azimuthal variation close to $\theta = 0$. Secondly, since the function order increases as $2m$ in the case of the JB-model fewer functions are able to contribute significantly to the pattern shape within a certain angular region, as compared to the N-model in which the function order only increases as m. These differences are illustrated in Figure 3.2 which shows the zenith angle dependence of the basis functions for various values of m, n.

[1] That is not to say the expansion in (3.1.10) is invalid. In fact, for a given aperture field distribution that allows expansion as in (3.1.5) the right-hand side of (3.1.7) does converge to $\mathbf{F}$.

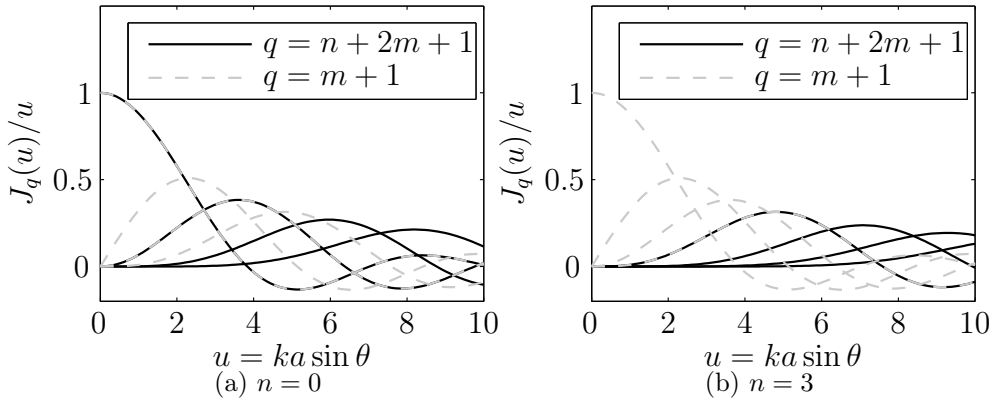


Figure 3.2: Comparison of JB-model ($q = n+2m+1$) and N-model ($q = m+1$) basis functions zenith angle dependence. In each figure plots are shown for $m = 0, 1, 2$.

3.2 Calculating Model Parameters

The formulation for determining the JB-model coefficients in Section 3.1.2 follows a typical solution to the far-field computation problem, which is to calculate the secondary pattern of a reflector antenna given the primary pattern. In the present application it is assumed that the secondary pattern is already known (to a certain extent) and that the coefficients for the various models have to be determined from this knowledge alone. One possible solution is to inverse transform (3.1.1) to determine the aperture field distribution, and from there calculate the coefficients. However, for cases where the secondary pattern is only known over a limited angular region this may lead to an inaccurate result for the aperture field distribution [23], and a subsequent error in the model coefficients. Moreover, for the N-model it is not possible to determine the coefficients from the aperture fields directly.

Instead the model coefficients are calculated directly from the far-field patterns using a weighted residual method.

3.2.1 Method of Weighted Residuals

Since the different polarization components of the far-field pattern are modelled separately and in order to simplify notation the formulation presented here is shown for a single polarization, although the extension to full-polarization is trivial. Also, the solution for the pattern model that follows in this section will be referred to as the *Weighted Residual Solution* (WRS).

Assuming that each of the components F of the far-field pattern $\mathbf{F}$ can be expressed as in (3.1.7) or (3.1.12), we multiply on both sides by one of the yet to be defined weighting functions $\{w_s(\theta, \phi)\}_{s=0}^{S}$ and then integrate to give S

equations of the form

$$\int_0^{\pi}\int_0^{2\pi} F(\theta,\phi)w_s(\theta,\phi)\,\mathrm{d}\phi\,\mathrm{d}\theta = \int_0^{\pi}\int_0^{2\pi}\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\left(A_{m,n}\cos n\phi + B_{m,n}\sin n\phi\right)\frac{J_{q(m,n)}(ka\sin\theta)}{ka\sin\theta}w_s(\theta,\phi)\,\mathrm{d}\phi\,\mathrm{d}\theta. \tag{3.2.1}$$

Note that the order of the Bessel function in this expression is defined as the function $q(m,n)$ which is $q(m,n) = n + 2m + 1$ for the JB-model, and $q(m,n) = m + 1$ for the N-model. After changing the order of integration and summation we have

$$\begin{aligned}\int_0^{\pi}\int_0^{2\pi} F(\theta,\phi)w_s(\theta,\phi)\,\mathrm{d}\phi\,\mathrm{d}\theta = &\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}A_{m,n}\int_0^{\pi}\int_0^{2\pi}\cos n\phi\frac{J_{q(m,n)}(ka\sin\theta)}{ka\sin\theta}w_s(\theta,\phi)\,\mathrm{d}\phi\,\mathrm{d}\theta \\ &+\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}B_{m,n}\int_0^{\pi}\int_0^{2\pi}\sin n\phi\frac{J_{q(m,n)}(ka\sin\theta)}{ka\sin\theta}w_s(\theta,\phi)\,\mathrm{d}\phi\,\mathrm{d}\theta.\end{aligned} \tag{3.2.2}$$

An obvious choice for w_s is to use Galerkin weighting[2]

$$w_s(\theta,\phi) = \left\{\begin{array}{c}\cos n_s\phi\\ \sin n_s\phi\end{array}\right\}\frac{J_{q(m_s,n_s)}(ka\sin\theta)}{ka\sin\theta} \tag{3.2.3}$$

and then using the orthogonality properties of the sine and cosine functions for $n_s = 0, 1, 2, 3, \ldots$

$$\int_0^{2\pi}\sin n_s\phi\sin n\phi\,\mathrm{d}\phi = \pi\left(1-\delta_{n_s,0}\right)\delta_{n_s,n} \tag{3.2.4a}$$

$$\int_0^{2\pi}\cos n_s\phi\sin n\phi\,\mathrm{d}\phi = 0 \tag{3.2.4b}$$

$$\int_0^{2\pi}\cos n_s\phi\cos n\phi\,\mathrm{d}\phi = \pi\left(1+\delta_{n_s,0}\right)\delta_{n_s,n}, \tag{3.2.4c}$$

where $\delta_{n_s,n}$ is the Kronecker delta

$$\delta_{n_s,n} = \begin{cases}1 & n_s = n\\ 0 & n_s \neq n,\end{cases} \tag{3.2.5}$$

[2]This weighting exploits only the orthogonality of the sine and cosine functions. For the JB-model specifically an orthogonality property of the Bessel functions may also be used by dropping the $ka\sin\theta$ factor in the denominator (see Appendix B.3). However, this does not apply to the N-model and Galerkin weighting is used for both models.

to reduce (3.2.1) to

$$\int_0^{\pi}\int_0^{2\pi} F(\theta,\phi)\cos n_s\phi \frac{J_{q(m_s,n_s)}(ka\sin\theta)}{ka\sin\theta}\,\mathrm{d}\phi\,\mathrm{d}\theta = \sum_{m=0}^{\infty} A_{m,n_s}(1+\delta_{n_s,0})\pi\int_0^{\pi}\frac{J_{q(m,n_s)}(ka\sin\theta)}{ka\sin\theta}\frac{J_{q(m_s,n_s)}(ka\sin\theta)}{ka\sin\theta}\,\mathrm{d}\theta \quad (3.2.6a)$$

for the cosine weighting functions ($n_s = 0, 1, 2, 3, \ldots$) and

$$\int_0^{\pi}\int_0^{2\pi} F(\theta,\phi)\sin n_s\phi \frac{J_{q(m_s,n_s)}(ka\sin\theta)}{ka\sin\theta}\,\mathrm{d}\phi\,\mathrm{d}\theta = \sum_{m=0}^{\infty} B_{m,n_s}\pi\int_0^{\pi}\frac{J_{q(m,n_s)}(ka\sin\theta)}{ka\sin\theta}\frac{J_{q(m_s,n_s)}(ka\sin\theta)}{ka\sin\theta}\,\mathrm{d}\theta \quad (3.2.6b)$$

for the sine weighting functions ($n_s = 1, 2, 3, \ldots$).

Therefore we have for the n_sth azimuthal mode, after truncating the summation at $m = M$, a set of equations

$$\mathbf{Z}^{\cos}_{n_s}\mathbf{A}_{n_s} = \mathbf{V}^{\cos}_{n_s} \quad (3.2.7a)$$

$$\mathbf{Z}^{\sin}_{n_s}\mathbf{B}_{n_s} = \mathbf{V}^{\sin}_{n_s} \quad (3.2.7b)$$

that can be solved separately from the equations similarly obtained for the other azimuthal modes. In this set of equations we have

$$\left(\mathbf{Z}^{\cos}_{n_s}\right)_{ij} = (1+\delta_{n_s,0})\pi\int_0^{\pi}\frac{J_{q(j,n_s)}(ka\sin\theta)}{ka\sin\theta}\frac{J_{q(i,n_s)}(ka\sin\theta)}{ka\sin\theta}\,\mathrm{d}\theta \quad (3.2.8a)$$

$$\left(\mathbf{Z}^{\sin}_{n_s}\right)_{ij} = \pi\int_0^{\pi}\frac{J_{q(j,n_s)}(ka\sin\theta)}{ka\sin\theta}\frac{J_{q(i,n_s)}(ka\sin\theta)}{ka\sin\theta}\,\mathrm{d}\theta \quad (3.2.8b)$$

$$\left(\mathbf{V}^{\cos}_{n_s}\right)_{i} = \int_0^{\pi}\int_0^{2\pi} F(\theta,\phi)\cos n_s\phi\frac{J_{q(i,n_s)}(ka\sin\theta)}{ka\sin\theta}\,\mathrm{d}\phi\,\mathrm{d}\theta \quad (3.2.8c)$$

$$\left(\mathbf{V}^{\sin}_{n_s}\right)_{i} = \int_0^{\pi}\int_0^{2\pi} F(\theta,\phi)\sin n_s\phi\frac{J_{q(i,n_s)}(ka\sin\theta)}{ka\sin\theta}\,\mathrm{d}\phi\,\mathrm{d}\theta, \quad (3.2.8d)$$

where the cosine equations are valid for $n_s \geq 0$ and the sine equations are only valid for $n_s > 0$.

In order to yield a determined system in (3.2.7) we produce equations for the weighting functions w_s with $m_s = 0, 1, 2, \ldots, M$. Then $\mathbf{A}_{n_s}, \mathbf{B}_{n_s}$ is each an unknown $(M+1)\times 1$ vector that needs to be solved. The exception is the N-model for the case $n > 0$ where we already have $\mathbf{A}_{0,n} = \mathbf{B}_{0,n} = \mathbf{0}$ and only M equations are produced.

Combining the sets of equations for $n_s = 0, 1, 2, \ldots, N$ then yields the block-diagonal system

$$\begin{bmatrix} \mathbf{Z}_0^{\cos} & \mathbf{0} & \cdots & & & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_1^{\cos} & \mathbf{0} & \cdots & & \cdots & \mathbf{0} \\ \vdots & \mathbf{0} & \ddots & \ddots & & & \vdots \\ & \vdots & \ddots & \mathbf{Z}_N^{\cos} & & & \\ & & & & \mathbf{Z}_1^{\sin} & & \\ & & & & & \mathbf{Z}_2^{\sin} & \ddots & \vdots \\ \vdots & \vdots & & & & \ddots & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \cdots & & & \mathbf{0} & & \mathbf{Z}_N^{\sin} \end{bmatrix} \begin{bmatrix} \mathbf{A}_0 \\ \mathbf{A}_1 \\ \vdots \\ \mathbf{A}_N \\ \mathbf{B}_1 \\ \mathbf{B}_2 \\ \vdots \\ \mathbf{B}_N \end{bmatrix} = \begin{bmatrix} \mathbf{V}_0^{\cos} \\ \mathbf{V}_1^{\cos} \\ \vdots \\ \mathbf{V}_N^{\cos} \\ \mathbf{V}_1^{\sin} \\ \mathbf{V}_2^{\sin} \\ \vdots \\ \mathbf{V}_N^{\sin} \end{bmatrix}$$

$$\mathbf{Z}\mathbf{x} = \mathbf{V}. \tag{3.2.9}$$

Since we already have $\mathbf{B}_{m,0} = \mathbf{0}$ in both (3.1.7) and (3.1.12) there are only equations $n_s = 1, 2, \ldots, N$ for the $\mathbf{B}_{n_s}$ vectors. Therefore in the determined system above the matrix is $\mathbf{Z}$ is $S \times S$ where $S = (2N + 1)(M + 1)$ for the JB-model and $S = (2N + 1)M + 3$ for the N-model.

In practice the far-field pattern may not be known to the extent that is required to calculate model coefficients as above. Firstly, the pattern is generally only known over a limited angular region $\theta \leq \theta_R$. In this case the limits of integration in (3.2.1) may be changed such that $\theta \in [0, \theta_R]$ to yield a pattern model that is valid over this limited region.

Secondly, and more importantly, practical estimation of the antenna pattern for calibration purposes requires that only a small number of measurements be necessary. In fact, if the knowledge of the pattern is sufficient to allow calculating the pattern model as above, this would eliminate the need for a model in the first place. However, since the pattern models are only approximate the method presented above is useful to determine how accurate these models are in the *ideal* case where the exact pattern is already known.

In the next section methods that require only a few calibration measurements to calculate the pattern model coefficients are presented.

3.2.2 Sparse Pattern Sampling

The simplest method to estimate the model parameters practically is by sampling the actual pattern at S number of points and then requiring that the model exactly fits the pattern at these points, i.e. point-matching. This is equivalent to the weighted residual method in (3.2.2) but with the weighting function chosen as the Dirac-delta function $w_s(\theta, \phi) = \delta(\theta_s, \phi_s)$. Following this approach results in the linear system

$$\mathbf{Z}_R\mathbf{x} = \mathbf{V}_R \tag{3.2.10}$$

where

$$(\mathbf{Z}_R)_{ij} = \tilde{f}_{m(j),n(j)}(\theta_i, \phi_i) \tag{3.2.11a}$$

$$(\mathbf{V}_R)_i = F(\theta_i, \phi_i) \tag{3.2.11b}$$

and $\tilde{f}_{m(j),n(j)}$ is a model basis function of which the order (m, n) and type (sine/cosine) depends on the value of j.

Since the objective here is to determine the model parameters with as few possible measurements it is assumed that only a sufficient number of pattern samples to yield a determined system in (3.2.10) are used. This requires a total of $S = (2N+1)(M+1)$ measurements for the JB-model and $S = (2N+1)M+3$ measurements for the N-model when the highest order terms in each of the models are $m = M$ and $n = N$, see (3.2.9).

Solution of the model parameters through (3.2.10) will hereafter be referred to as the *Direct Solution* (DS). As will be shown later this method may not be preferred since a large number of calibration measurements within a limited angular region are required to yield an accurate pattern model, which causes the system in (3.2.10) to become ill-conditioned. In the following two alternative approaches are presented to address this problem.

Generally the secondary pattern is at least *approximately* known to the extent required for the use of the WR solution in Section 3.2.1, whether through careful direct measurement, sophisticated simulation models, or a combination of both. In this case a set of coefficients $\mathbf{x}_0$ for the model of the *approximate* pattern can be computed by solving (3.2.9). Now if the *actual* pattern is similar to the approximate pattern, then the actual pattern model coefficients $\mathbf{x}$ may also expected to be similar to the approximate pattern model coefficients $\mathbf{x}_0$. In this case a constrained solution may be used to reduce the number of measurements required to find a unique solution to the system (3.2.10). In general the problem can be stated as finding a solution to

$$\arg\min_{\mathbf{x}} \|\mathbf{x} - \mathbf{x}_0\| \tag{3.2.12a}$$

$$\text{subject to } \mathbf{Z}_R\mathbf{x} = \mathbf{V}_R \tag{3.2.12b}$$

where $\arg\min_x f(x)$ finds the solution for x which minimises the function $f(x)$.

Here two possible solutions to the above problem are investigated. The first solution formulates the above as a quadratic programming problem with equality constraints [68, § 16.1]. Using the method of Lagrange multipliers a solution is sought by minimising the functional

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = (\mathbf{x} - \mathbf{x}_0)^\dagger \mathbf{1} (\mathbf{x} - \mathbf{x}_0) + \boldsymbol{\lambda}^\dagger (\mathbf{Z}_R\mathbf{x} - \mathbf{V}_R) + (\mathbf{Z}_R\mathbf{x} - \mathbf{V}_R)^\dagger \boldsymbol{\lambda}. \tag{3.2.13}$$

To this end the complex gradient [1] of $\mathcal{L}$ is set equal to zero which yields the system of equations

$$\nabla_{\mathbf{x}^\dagger}\mathcal{L} = \mathbf{1}\mathbf{x} - \mathbf{1}\mathbf{x}_0 + \mathbf{Z}_R^\dagger\boldsymbol{\lambda} = \mathbf{0} \tag{3.2.14a}$$

$$\nabla_{\boldsymbol{\lambda}^\dagger}\mathcal{L} = \mathbf{Z}_R\mathbf{x} - \mathbf{V}_R = \mathbf{0}. \tag{3.2.14b}$$

Finally the model coefficients are determined by solving

$$\begin{bmatrix} \mathbf{1} & \mathbf{Z}_R^\dagger \\ \mathbf{Z}_R & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_0 \\ \mathbf{V}_R \end{bmatrix}. \tag{3.2.15}$$

The model coefficient solution obtained with (3.2.15) will hereafter be referred to as the *Lagrange Multiplier Solution* (LMS).

The second approach uses a quadratic penalty method [68, § 17.1] to find a solution to (3.2.12). The objective here is to minimise the penalty function

$$\mathcal{L}(\mathbf{x}) = (\mathbf{Z}_R\mathbf{x} - \mathbf{V}_R)^\dagger (\mathbf{Z}_R\mathbf{x} - \mathbf{V}_R) + \lambda (\mathbf{x} - \mathbf{x}_0)^\dagger (\mathbf{x} - \mathbf{x}_0) \tag{3.2.16}$$

to which the solution is again obtained by setting the gradient equal to zero to find a stationary point of $\mathcal{L}$

$$\nabla_{\mathbf{x}^\dagger}\mathcal{L} = \mathbf{Z}_R^\dagger\mathbf{Z}_R\mathbf{x} - \mathbf{Z}_R^\dagger\mathbf{V}_R + \lambda\mathbf{x} - \lambda\mathbf{x}_0 = 0. \tag{3.2.17}$$

Finally, the vector $\mathbf{x}$ is found by solving

$$\left(\mathbf{Z}_R^\dagger\mathbf{Z}_R + \mathbf{1}\lambda\right)\mathbf{x} = \left(\mathbf{Z}_R^\dagger\mathbf{V}_R + \lambda\mathbf{x}_0\right) \tag{3.2.18}$$

where $\mathbf{1}$ is the identity matrix. The model coefficients solution thus obtained will hereafter be referred to as the *Quadratic Penalty Solution* (QPS).

In order to solve (3.2.18) an appropriate value should be chosen for the penalty factor λ which affects the relative weighting between satisfying (3.2.12a) and satisfying (3.2.12b). If $\lambda = 0$ the solution is unaffected by $\mathbf{x}_0$ and reduces to the DS. Conversely, if $\lambda \to \infty$ then the actual pattern samples have no effect and the solution is trivially $\mathbf{x} = \mathbf{x}_0$. For intermediate values of λ it is noted that the obtained stationary point of $\mathcal{L}$ may not necessarily satisfy (3.2.12b) exactly, and λ may be adjusted to reflect the level of confidence in the pattern samples.

3.3 Numerical Results

With two analytical pattern models and various methods by which to solve the parameters within these models, numerical results will now be presented. All results in this section pertain to the MeerKAT OG system (see Section 2.1 for an overview of the system) and are for the operating frequency of 1.42 GHz. In all cases it is assumed that the antenna patterns are defined in a reference frame where the pattern maximum is approximately located at $(\theta = 0, \phi = 0)$. Finally, the error metric ϵ_N defined in (2.2.10) is used below to measure the accuracy of the various pattern models.

3.3.1 Weighted Residual Models

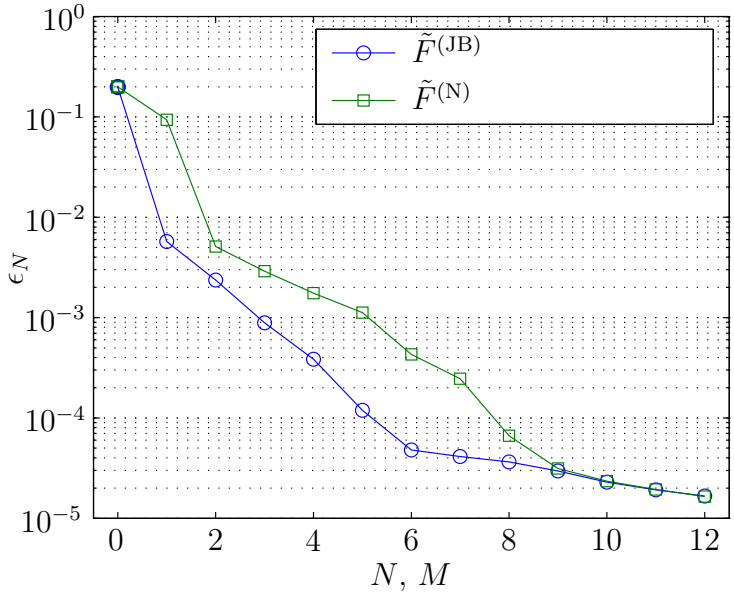
In order to determine how accurately the JB- and N-models can be made to approximate the far-field pattern WRS models were obtained for various orders of expansion M and N. The models were computed over the region $\theta \leq 5°$. This means that the pattern model extends far enough from the beam centre to include roughly the first three sidelobes (the half-power and null beamwidths are about 1° and 3°, respectively). Figure 3.3 (a) and (b) show how the error in each of the models decreases as the model expansion orders are increased and as more terms are added, respectively.

When $M = N = 0$ each of the models has a single term and is equal to the jinc-function, and the model accuracies are equal with an error of around $\epsilon_N = 0.2$. As more terms are added the JB-model error rapidly decreases whereas the N-model error is seen to decrease at a slower rate. This is attributed to the fact that low order models are already a relatively accurate approximation to the main beam, so that after only a few terms the dominant contribution to the model error is due to the sidelobes. As the Bessel function order increases more rapidly for the JB-model than for the N-model, higher accuracy is attainable at these wider regions for a low-order JB-model than an N-model of the same order. Consequently the overall error is smaller for the JB-model than for the N-model. For the same reason we observe that from around $M, N = 6$ the rate at which the JB-model error decreases slows down as additional terms are of such high order that their dominant contribution falls outside the region $\theta \leq \theta_R$ within which the model is calculated. On the other hand the error in the N-model still decreases steadily until around $M, N = 9$ and from there on the error for the two models are approximately equal.

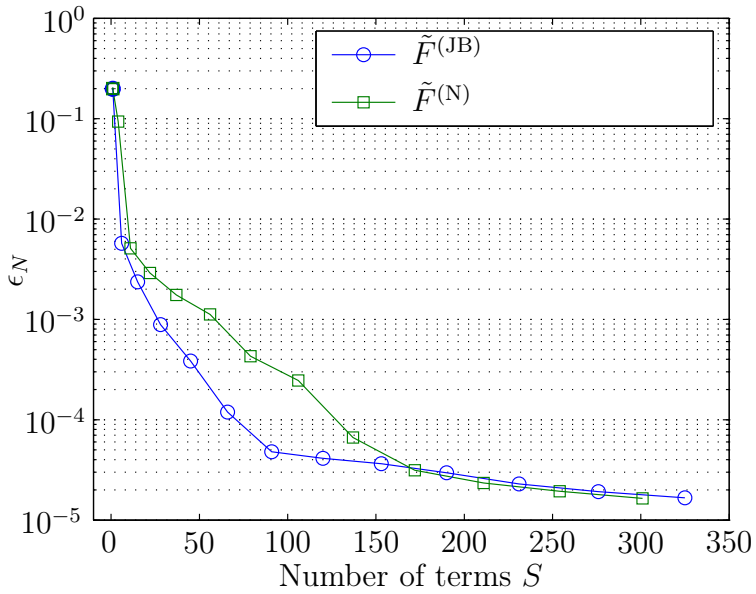
A number of $J_q(u)/u$ functions of various orders q are plotted in Figure 3.4. In order to ensure that all the terms in the model have a dominant contribution to the model within the region of interest, all terms that have their first peaks outside that region are discarded. This limits the order of expansion in both the azimuth and the zenith for the JB-model, whereas only the order of expansion in the zenith is limited in the N-model. For $\theta \leq 5°$ we see that all terms with $q \geq 16$ are discarded.

By limiting the number of terms in the model as such improves the conditioning of the system in (3.2.9) as most of the higher order terms add little independent components to the model. The maximum condition number $\kappa_{\max} = \max(\kappa(\mathbf{Z}_n))$ of all the sub-matrices on the diagonal of $\mathbf{Z}$ is shown as a function of expansion order M, N in Figure 3.5. As expected for low-order models the condition number remains relatively low as more terms are added, and from around $N = M = 5$ the conditioning of the system for the JB-model deteriorates rapidly. For the N-model the system conditioning also degrades as more terms are added, however this is at a much slower rate.

Setting $\theta = 5°$ and discarding the corresponding terms from the model is seen to flatten the condition number for the JB-model from $N, M = 5$, since



(a) Error as a function of M, N.



(b) Error as a function of number of terms.

Figure 3.3: Comparison of JB-model and N-model accuracy as computed with the WRS. The number of terms S in each model is defined below (3.2.9).

from here on higher order terms have $q = n + 2m + 1 \geq 16$ and are discarded. The highest value for $\kappa_{\max}$ is reached at $N = M = 7$ which corresponds to the largest n-mode added. For right up to $N = M = 12$ the results for the N-model are exactly the same as where no limit was placed on the term order. This result is not surprising since even for $M = 12$ we have $q = M + 1 < 16$.

Finally, with $\theta_R = 2.5°$ the highest Bessel function order allowed is $q = 7$

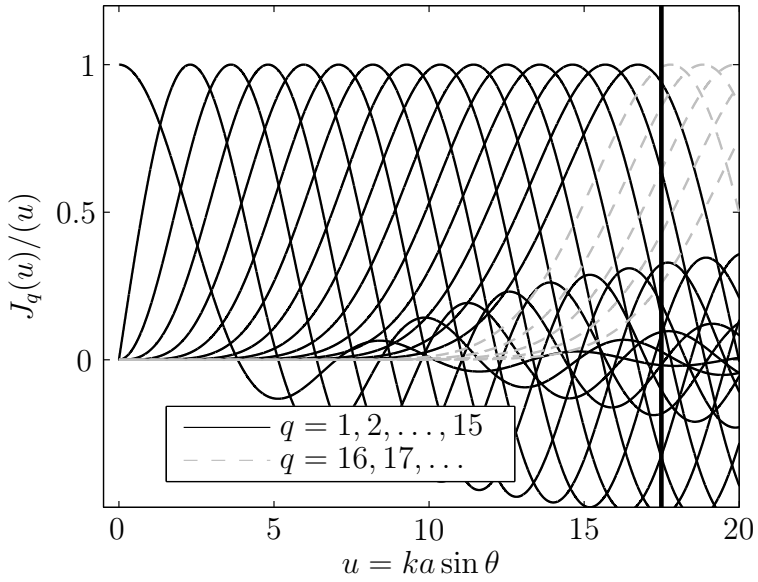


Figure 3.4: Zenith angle dependence of the $J_q(u)/u$ functions. Functions are normalised to their respective maxima for clarity, and the solid vertical line ($\theta = 5°$) indicates the extent of the region of interest. Terms that have peaks outside this region are discarded.

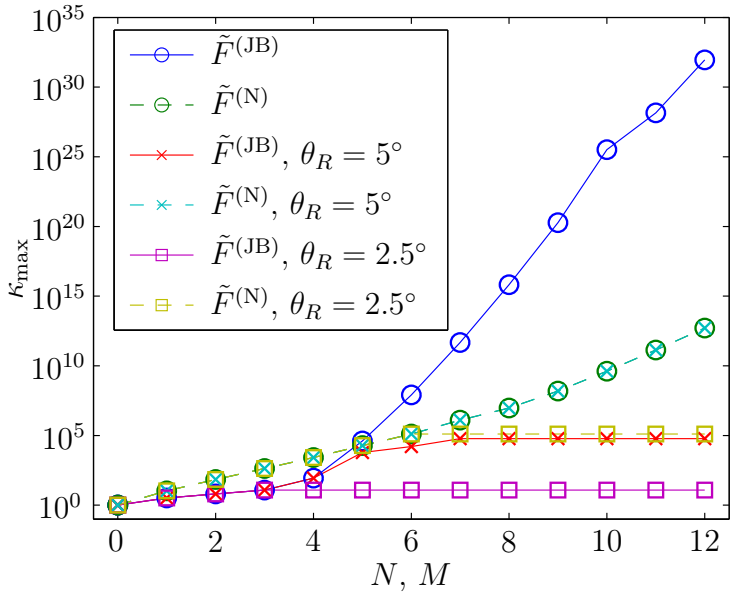
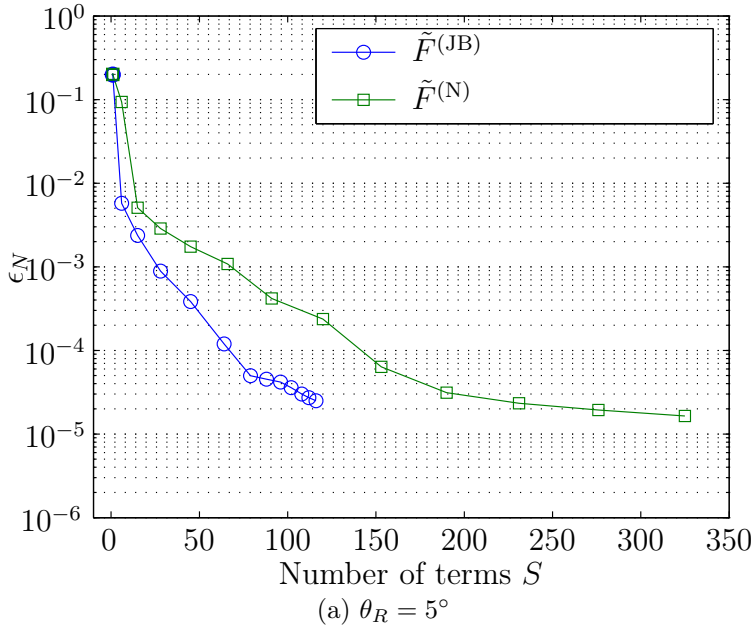
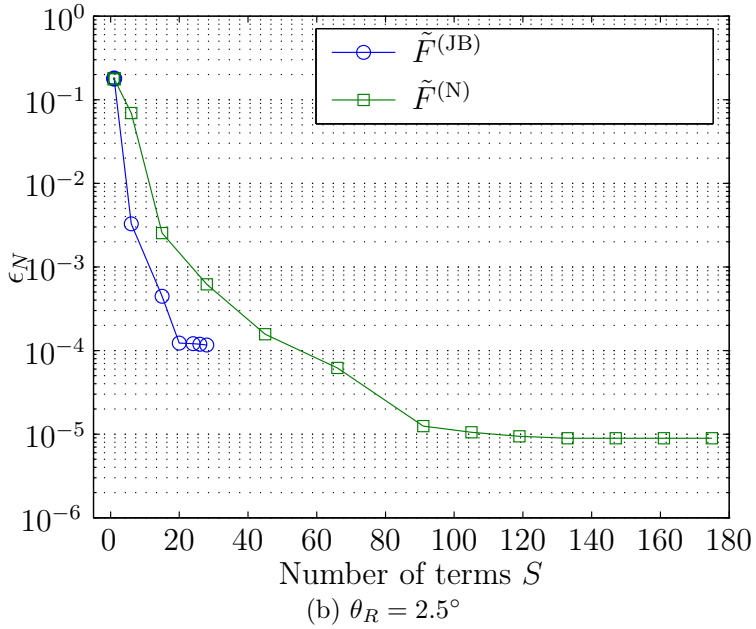


Figure 3.5: Maximum condition number of the sub-matrices in $\mathbf{Z}$ as a function of N, M. A comparison is shown between JB- and N-models containing all terms up to a given order, and JB- and N-models where non-dominant terms have been discarded for $\theta_R = 5°$ and $\theta_R = 2.5°$.

and above $N = M = 6$ no more terms are added to the JB-model. Consequently the system for this model remains well-conditioned. Similarly for the N-model $\kappa_{\max}$ stabilises at $N = M = 6$ after which no more m-modes are added, although more n-modes are allowed.



(a) $\theta_R = 5°$



(b) $\theta_R = 2.5°$

Figure 3.6: Pattern model accuracy as a function of number of terms. A comparison is shown between JB-model and N-model for different values of θ_R.

Figure 3.6 shows how the error decreases with number of terms when terms with too high an order are discarded form the model. The results in Figure 3.6 (a) can be compared directly with that in Figure 3.3 (b). Since no terms are discarded from the N-model the results for this model are exactly as before. On the other hand, for the JB-model the decrease in error from around 70 terms is more rapid and the same level of accuracy is achieved here with 116 terms as compared to that achieved before with 220 terms.

When limiting the region of interest further down to $\theta_R = 2.5°$ an important difference between the JB-model and N-model becomes apparent, as shown in Figure 3.6 (a). Whereas the JB-model is able to rapidly decrease the error down to $\epsilon_N \approx 10^{-4}$ with only 28 terms, the N-model requires roughly twice as many terms to achieve this same level of accuracy. However, the N-model allows many more terms to be added without severely degrading the system conditioning, and the error can be reduced to $\epsilon_N \approx 10^{-5}$ by doubling again the number of terms in the model.

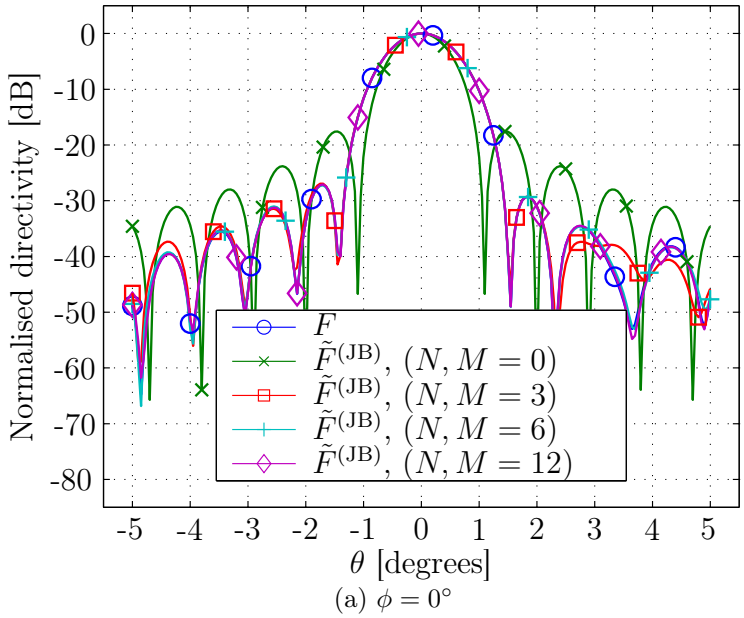
The actual pattern that is modelled along with JB-models of various orders is shown in Figure 3.7 (a) and (b) in the planes $\phi = 0°$ and $90°$, respectively. With $M, N = 0$ the model is the circularly symmetric jinc-function with a narrower main beam than the actual pattern and much higher first sidelobes at more or less -17.5 dB. Increasing the model order to $M, N = 3$ results in a model that is relatively accurate out to the first / second sidelobe, and after $M, N = 6$ the difference between the model and the pattern is almost indistinguishable over the region shown.

Figure 3.8 compares the actual pattern to N-models of various orders. As for the JB-model it is observed that the angular region over which an accurate model is obtained increases as the model order increases, albeit at a slower rate for the N-model. Here an $M, N = 3$ order model is only accurate to the first sidelobe, and an $M, N = 6$ order model only accurate to the second sidelobe.

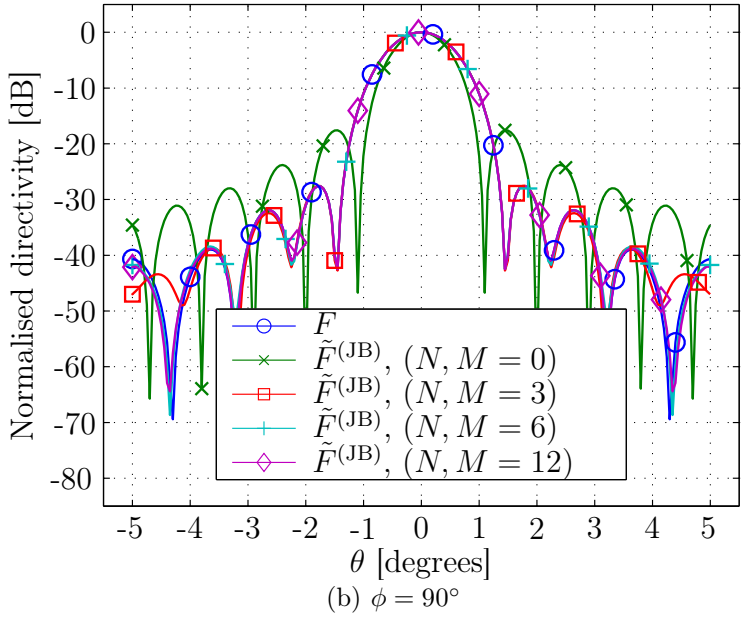
Until now results were presented for models computed through the weighted residuals method as an indication of the accuracy that can be achieved with the various analytical models. Next results will be presented for models estimated through more practical means.

3.3.2 Sparse Pattern Sampling - Direct Solution

First we consider JB-models obtained through the direct solution of (3.2.10). The system is exactly determined as the number of pattern samples equals the number of basis functions in the model $S = (2N + 1)(M + 1)$. Pattern samples are taken within the region $\theta \leq 1°$ which roughly coincides with the 10 dB beamwidth. The sampling points used to produce the system (3.2.10) are shown in Figure 3.9 (a) for models of various orders. For $M, N = 0$ only one model coefficient needs to be solved and the pattern is sampled on-axis; for higher order models more sampling points are required. Starting with the first sampling point on-axis for models of all orders, additional sampling



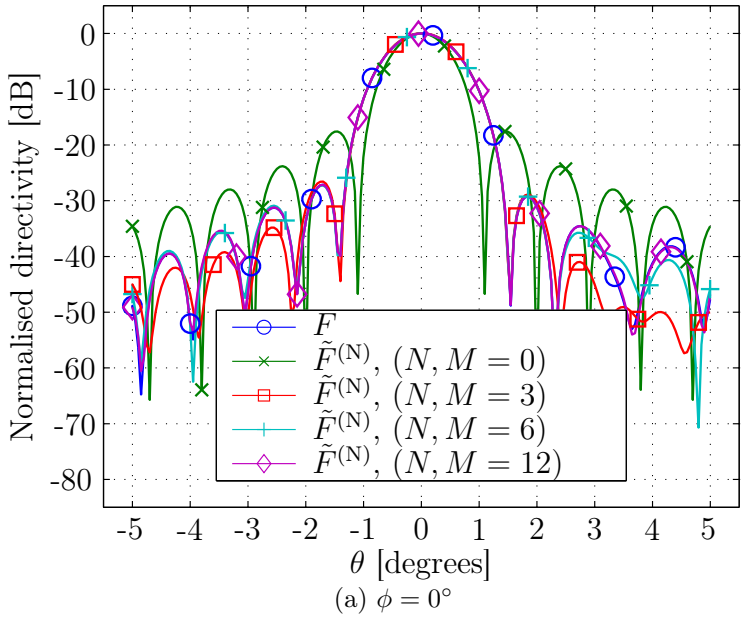
(a) $\phi = 0°$



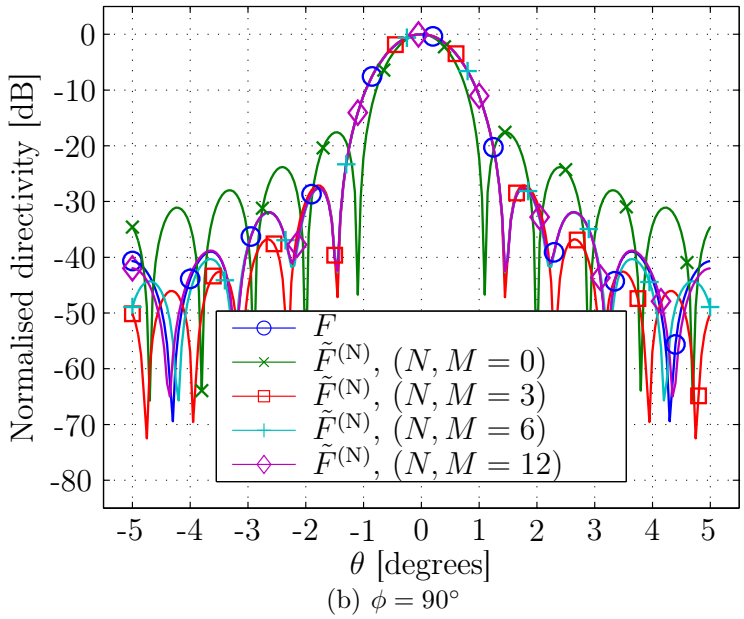
(b) $\phi = 90°$

Figure 3.7: Reference pattern and JB-models of various order.

points are then selected such that the increase in the condition number $\kappa(\mathbf{Z}_R)$ is minimised. This is done in an attempt to improve the conditioning of the linear system from which the model parameters are solved, and as a result the accuracy of the obtained solution. This procedure results in the seemingly random scattering of points seen in the figure. Although for $M, N = 1$ it is seen that the sampling points reside mostly near the $\phi = 0°, 90°, 180°$ axes which maximises independence of the different azimuthal modes. For higher order



(a) $\phi = 0°$

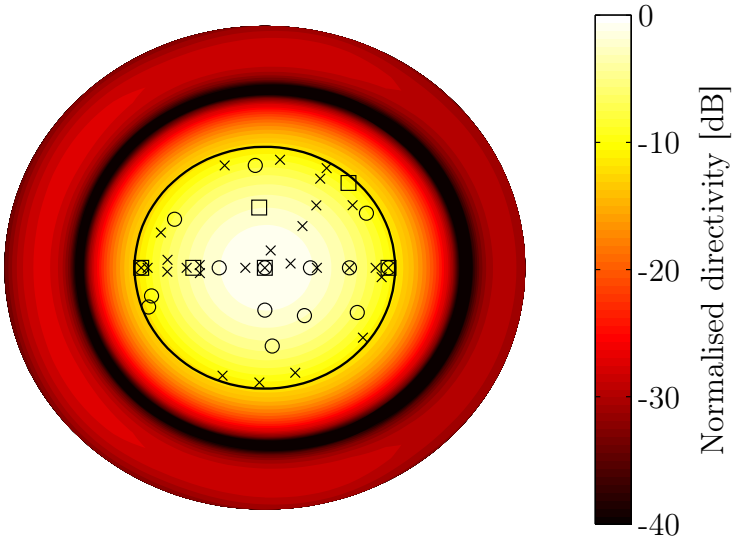


(b) $\phi = 90°$

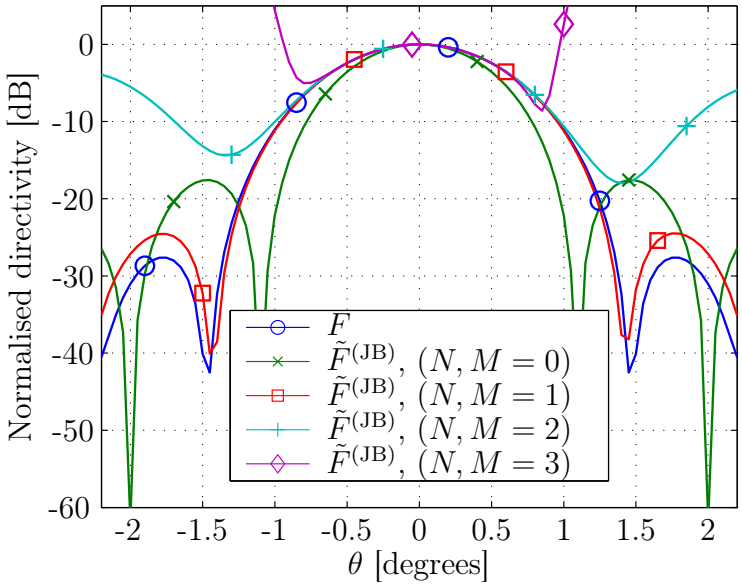
Figure 3.8: Reference pattern and N-models of various order.

modes many points are also seen to be positioned near the $\theta = 1°$ boundary, which maximises the independence of higher order $J_q(u)/u$ functions which are zero at and very slowly increasing from $\theta = 0$.

The JB-models of various orders for the pattern are shown in Figure 3.9 (b). For the first two orders we observe that the model accuracy improves as more terms are added, specifically over the main beam and within the region where the pattern is sampled. The first sidelobe level is also relatively accurate for



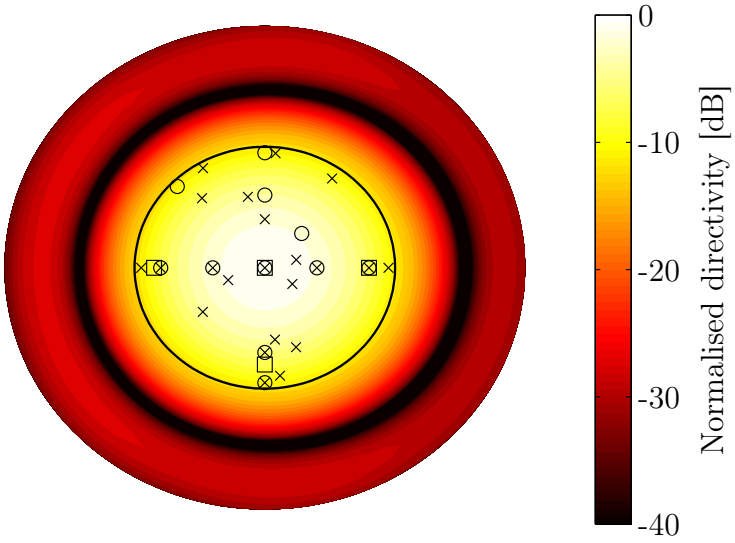
(a) Pattern sampling points



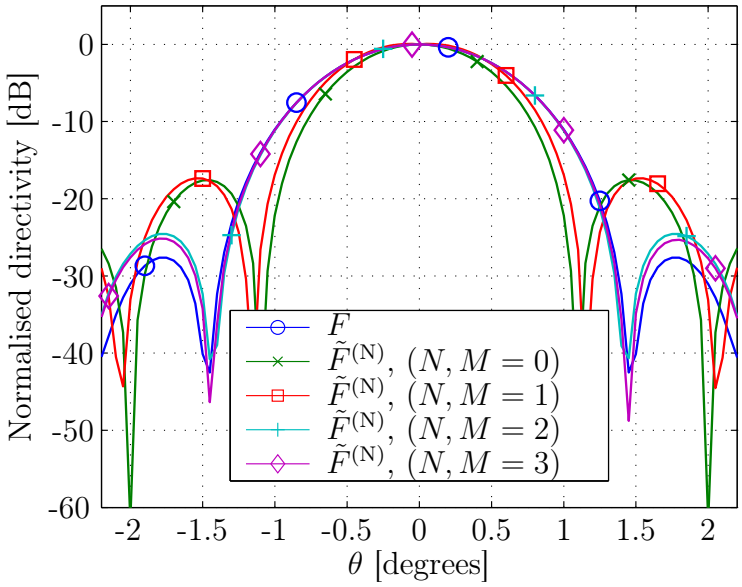
(b) Reference and JB-model patterns in $\phi = 0°$ plane

Figure 3.9: Direct solution of JB-models. In (a) the sampling points of the models of various order are indicated as '□' $M, N = 1$, '○' $M, N = 2$, and '×' $M, N = 3$. The solid black line in (a) indicates the $\theta \leq 1°$ region within which pattern samples are taken.

$M, N = 1$ although this is well away from the sampling region. As the model order is increased further the model is seen to diverge significantly from the pattern, at first in the sidelobe region for $M, N = 2$ but also over the main beam within the sampling region for $M, N = 3$. This is attributed to the fact that the higher order terms degrade the conditioning of the system if the



(a) Pattern sampling points



(b) Reference and N-model patterns in $\phi = 0°$ plane

Figure 3.10: Direct solution of N-models. In (a) the sampling points of the models of various order are indicated as '□' $M, N = 1$, '∘' $M, N = 2$, and '×' $M, N = 3$. The solid black line in (a) indicates the $\theta \leq 1°$ region within which pattern samples are taken.

sampling is limited to a relatively small region.

We now consider the same results but for the N-model shown in Figure 3.10. The pattern sampling points in Figure 3.10 (a) are chosen using the same procedure as for the JB-model and the results are — as is expected — somewhat different, with a cluster of points near on-axis and a more uniform spread of

points for higher order models. Once again this is due to the inclusion of more lower-order Bessel functions in the N-model as compared to the JB-model. Consequently a greater independence exist among the basis functions over the limited sampling region. For this same reason the higher order N-models obtained with direct solution do not diverge as rapidly as was the case for the JB-models, as shown in Figure 3.10 (b).

In order to quantify the accuracy of the different models obtained with the direct solution the error was computed over the region within which the pattern was sampled. This was done for JB-models and N-models of orders ranging from $M, N = 0$ up to $M, N = 6$ and a sampling region limited first to $\theta \leq 1^\circ$ and then $\theta \leq 0.5^\circ$. The results are shown in Figure 3.11. As before the low-order JB-models are relatively accurate, but as the model order is increased the model diverges from the pattern so that the smallest error within the 1° region is $\epsilon_N \approx 10^{-5}$. Low-order N-models exhibit a somewhat larger error, but by sufficiently increasing the order of expansion the error can be reduced to much lower levels than that of the JB-models, reaching an error of $\epsilon_N \approx 10^{-7}$ within the 1° region. Also, the N-model does not seem to diverge as much from the pattern as the order is increased as opposed to the case for the JB-model.

The results for a sampling region of $\theta \leq 0.5^\circ$ are very similar to that of the larger sampling region, except that the errors are generally much smaller for the lower order models. This is expected since the same number of terms are used to model the pattern over a smaller region.

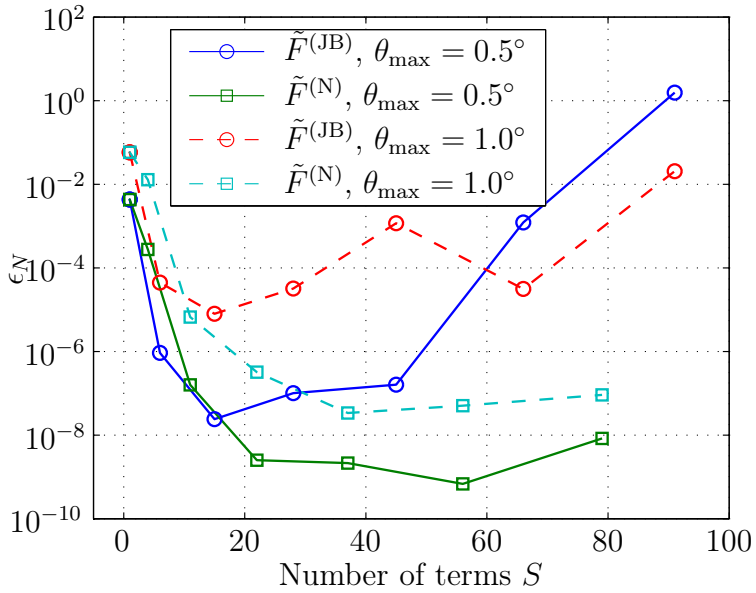


Figure 3.11: Direct solution JB-model and N-model accuracy as a function of number of terms. A comparison is shown for different sampling regions. Markers indicate models of order $M, N = 0, 1, 2, , \ldots, 6$.

We now focus on solving for the model coefficients using one of the constrained solutions in an attempt to improve the accuracy versus number of measurements trade-off.

3.3.3 Sparse Pattern Sampling - Constrained Solutions

For this section two antenna patterns are used to construct the various models using one of the constrained solutions in Section 3.2.2. The first pattern F_0 is the ideally expected on-axis pattern and is assumed to be known; the second pattern is the actual pattern F at the time of operation and is obtained by introducing a small error in the positioning of the feed and subreflector. These patterns are shown in Figure 3.12 where the difference between the patterns is seen to be mainly a pointing error. The error in using F as an approximation of F_0 is $\epsilon_N \approx 10^{-2}$ when computed over the region $\theta \leq 1°$.

Now models will be estimated for this perturbed pattern F using the least possible number of measurements. All pattern samples taken for F are within the region $\theta \leq 1°$ and the initial model coefficients vector $\mathbf{x}_0$ as in (3.2.12a) is obtained using the WRS to compute a model for F_0 over a $\theta \leq 5°$ region.

Figure 3.13 shows the accuracy of the models as obtained using the LMS. The errors are shown as functions of the number of pattern samples taken, ranging from a single sample up to the number of terms in each of the models. In Figure 3.13 (a) the results for the JB-model indicate that if the number of samples approximately equals half the number of terms in the model at least as much accuracy is achieved as when the direct solution is used to compute the model coefficients. In addition to reducing the number of pattern

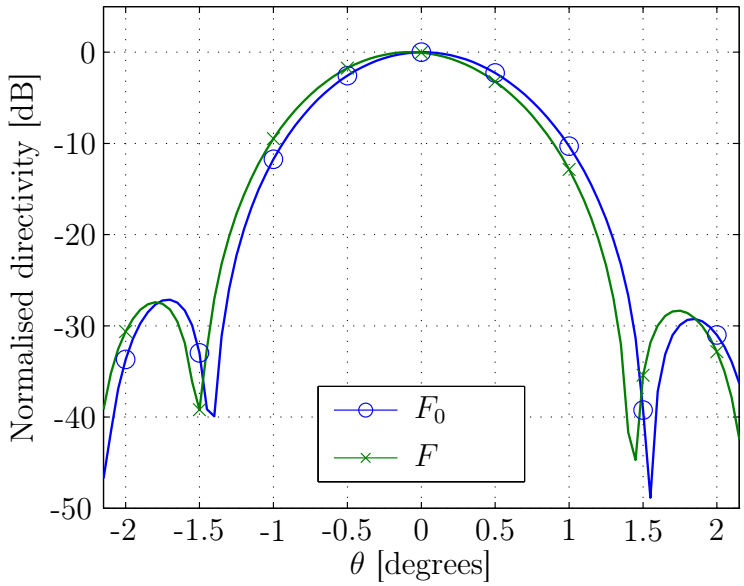
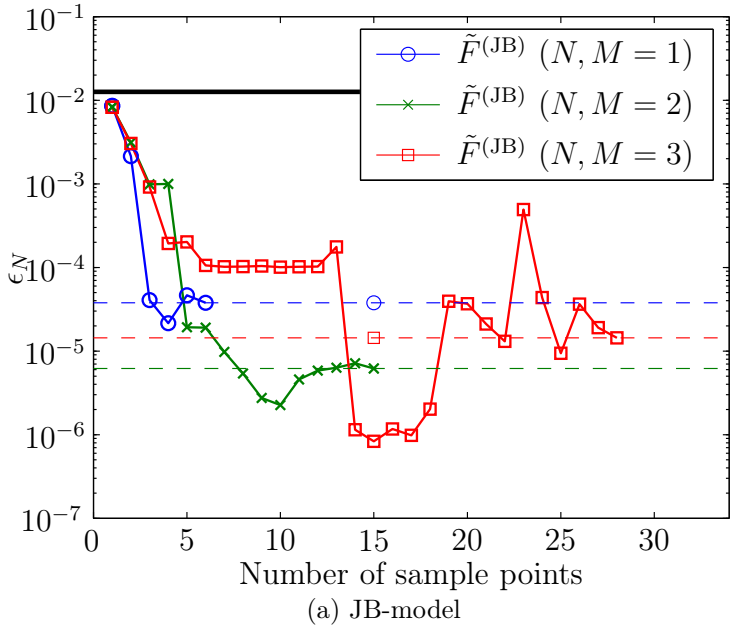
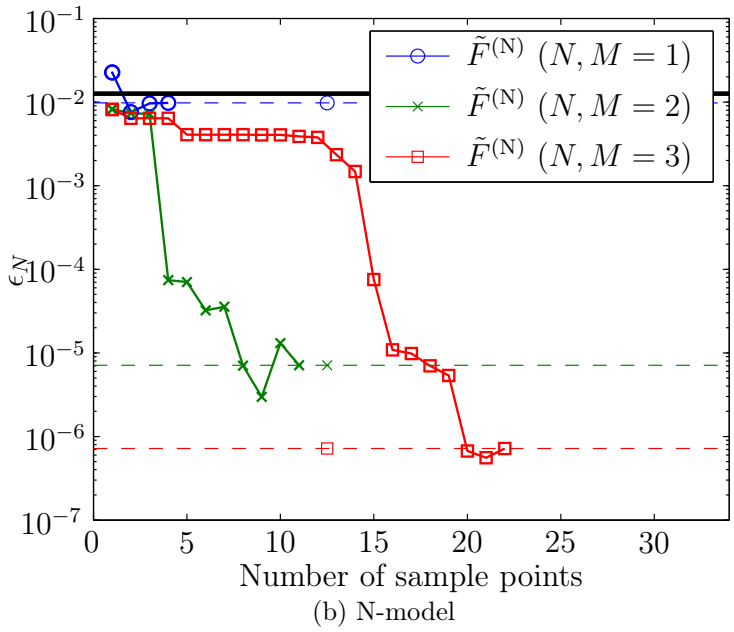


Figure 3.12: Ideally expected pattern $\mathbf{F}_0$ and perturbed pattern $\mathbf{F}$.



(a) JB-model



(b) N-model

Figure 3.13: Accuracy of models obtained with LMS as a function of number of pattern samples. Dashed lines indicate error obtained with direct solution for the same order model. Solid line at the top indicates error in assuming $F \approx F_0$.

samples required for a unique solution, the constrained solution approach may also improve the solution accuracy in some cases; note that the error for the constrained solution models reduce to below the error in the direct solution models. However, as the number of samples approaches the number of terms

in the model the error approaches that of the direct solution. Finally, note that with only a single sample the error in the model is approximately equal to that assuming $F \approx F_0$. This is expected since the solution is then dominated by minimising the difference between the expected pattern coefficients vector and the estimated pattern coefficients $\|\mathbf{x} - \mathbf{x}_0\|$.

The same results but for the N-model are shown in Figure 3.13 (b). Overall it is observed that almost as many pattern samples as terms in each model are required to reduce the model accuracy to the level of the error in the direct solution models. That is, the constrained solution does not improve the performance of the N-models as much as was found for the JB-models.

Next the QPS was used to determine pattern models with a reduced number of measurements. Since such a solution requires selecting the penalty factor λ in (3.2.18) a parametric study was first used to determine the effect of this parameter on the model accuracy. Here the value of λ is normalised such that

$$\lambda = \frac{\varepsilon \left\| \mathbf{Z}_R^{\dagger} \mathbf{V}_R \right\|}{\|\mathbf{x}_0\|} \tag{3.3.1}$$

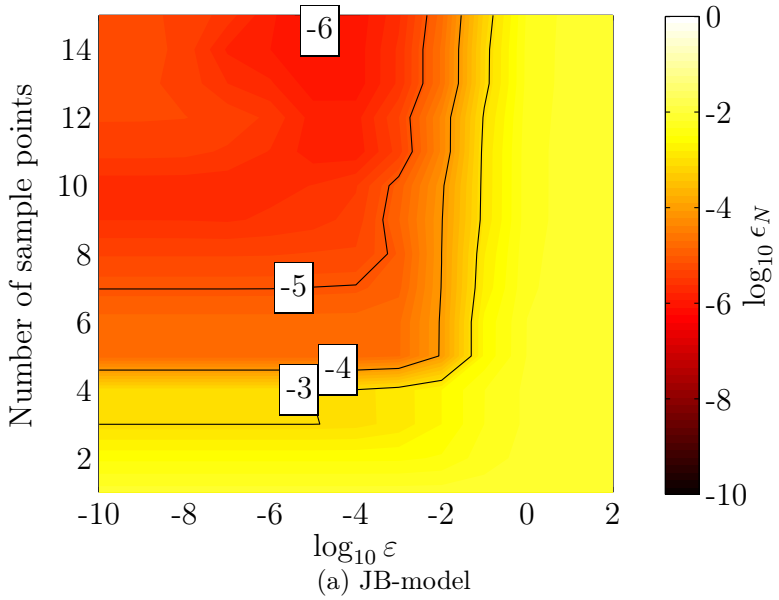
where ε is selected to control the relative weighting of the terms on the right-hand side of (3.2.18).

For models of various orders the QPS was used to estimate a model of the perturbed pattern F using the initial coefficients vector $\mathbf{x}_0$ for the ideally expected pattern F_0, and different values of ε. For each model the number of pattern samples ranged from one up to the number of terms in that model, and the error computed as a function of the number of sampling points.

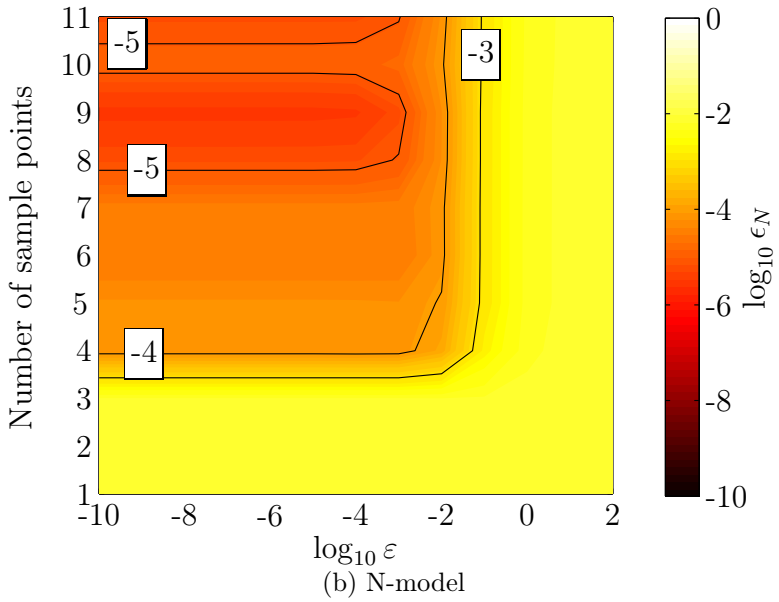
The error in the JB-model of order $M, N = 2$ is shown as a function of ε and the number of pattern samples S in Figure 3.14 (a). In general the error in the model is seen to decrease as $\varepsilon \to 0$ (towards the left) and as S increases (towards the top). This result is expected to some extent since a small ε and a large number of pattern samples means that the solution is dominated by the samples of the actual pattern that is being modelled. Over a wide range of values $10^{-10} < \varepsilon < 10^{-3}$ the variation in error seems very small for a given number of pattern samples. As the value of ε is increased above 10^{-3} the relative weighting of the penalty for $\|\mathbf{x} - \mathbf{x}_0\|$ starts to dominate the solution, resulting in a model accuracy comparable to the error in assuming $F \approx F_0$ when $\varepsilon \approx 1$ and for any number of pattern samples.

Similar results are observed for the N-model in Figure 3.14 (b) as were found for the JB-model. However, the N-model seems somewhat less accurate (almost an order of magnitude for larger numbers of pattern samples) than the JB-model for the same penalty factor, which shows an improvement in the JB-model accuracy as compared to the direct solution models and $M, N = 2$ in Figure 3.11.

Increasing the model order to $M, N = 6$ yields the results shown in Figure 3.15, where the accuracy of both the JB-models and N-models are seen to



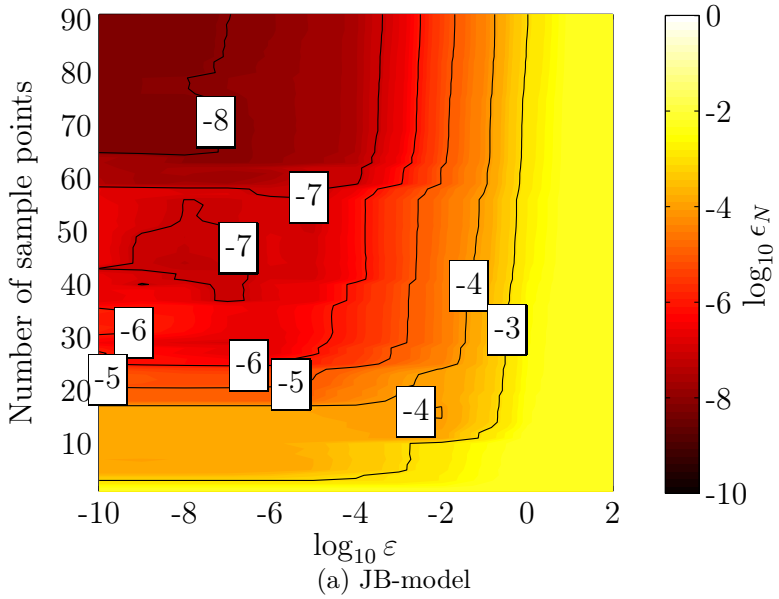
(a) JB-model



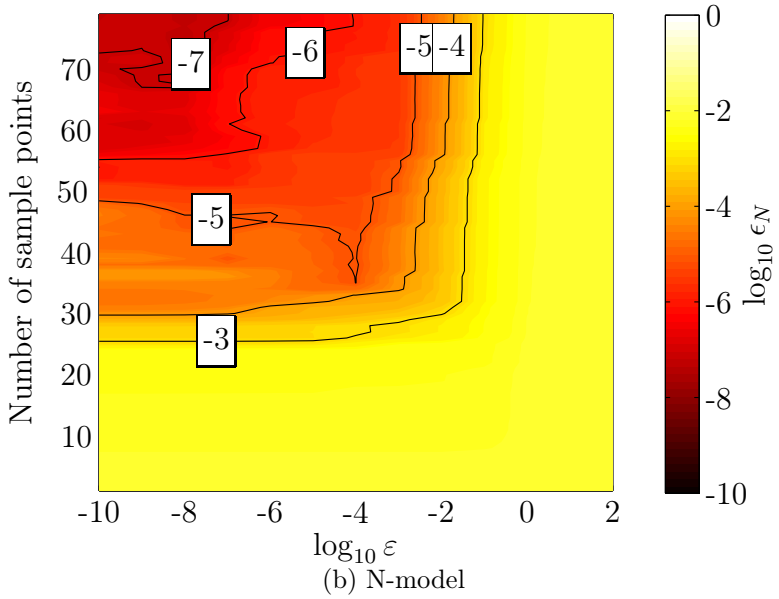
(b) N-model

Figure 3.14: Accuracy of JB- and N-models of order $M, N = 2$ obtained with QPS. Error is shown as a function of the normalised penalty factor ε and the number of pattern samples S. Contour lines are drawn at integer intervals of $\log_{10}\epsilon_N$.

have improved significantly. Compared with the lowest error attainable with the JB-model and using the DS in Figure 3.11, the error may be reduced by up to three orders of magnitude through using the QPS. For the DS of high-order models the ill-conditioning of $\mathbf{Z}_R$ resulted in the coefficients vector diverging rapidly as the number of pattern samples was increased, whereas for the QPS



(a) JB-model



(b) N-model

Figure 3.15: Accuracy of JB- and N-models of order $M, N = 6$ obtained with QPS. Error is shown as a function of the normalised penalty factor ε and the number of pattern samples S. Contour lines are drawn at integer intervals of $\log_{10}\epsilon_N$.

the penalty term in (3.2.16) ensures that the coefficients vector does not diverge too far from the initial vector $\mathbf{x}_0$. Consequently a much larger number of terms may be included in the JB-model, reducing the error to much lower levels. Once again the results for the N-models using the QPS are not much different to that using the DS.

The error as a function of model order is shown in Figure 3.16 for JB- and N-models obtained with the QPS, and using $\varepsilon = 10^{-8}, 10^{-6}, 10^{-4}$. For each model the number of sample points is equal to the number of terms in the model. As shown the error in the JB-model decreases steadily as the model order increases, except for high-order models with $\varepsilon = 10^{-4}$ where the error increases somewhat at $M, N = 4$. Even in this case the error may be reduced to below 10^{-7} by increasing the model order to $M, N = 6$. Similar results are observed for the N-models, although the error in these models are again somewhat higher than those in the JB-models.

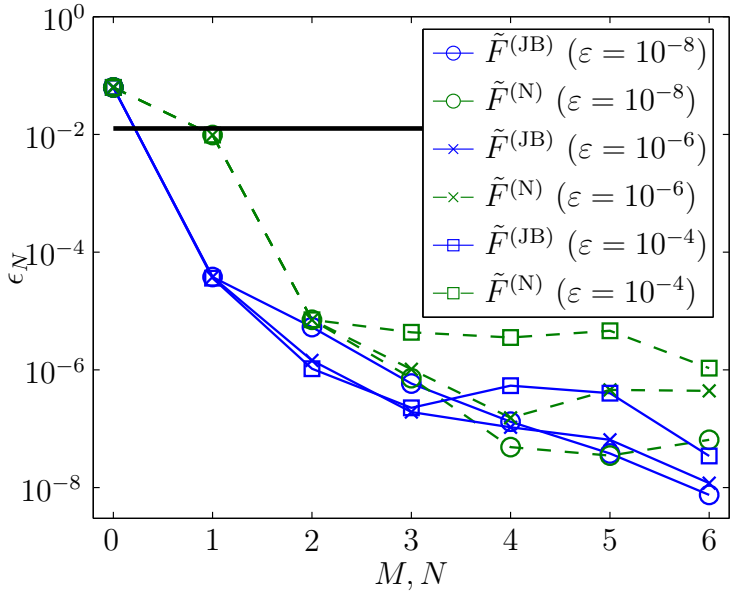


Figure 3.16: Error in QPS JB- and N-models of increasing order. Solid line at the top indicates error in assuming $F \approx F_0$.

Finally, errors in the models obtained with the QPS are shown as a function of the number of pattern samples in Figure 3.17. Similar to the results for JB-models obtained with the LMS approximately half as many pattern samples as number of model parameters are required to yield a model which is as accurate as the model obtained with the DS. Here the only difference is that as the number of measurements approaches the number of model parameters the error decreases steadily beyond that in the DS model. Compared to the error in the JB-model that in the N-model decreases at a slower rate as was also the case for the LMS models.

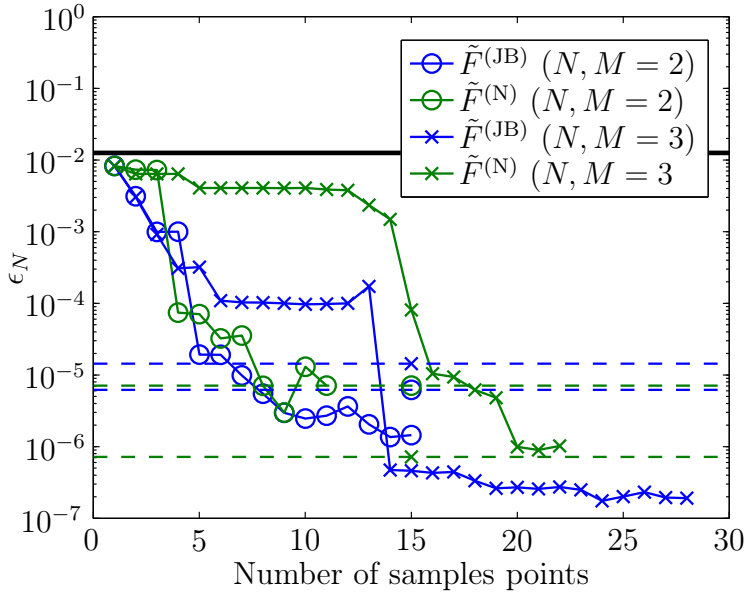


Figure 3.17: Error in QPS JB- and N-models as a function of number of pattern samples. For all models $\varepsilon = 10^{-6}$. Dashed lines indicate error obtained with direct solution for the same order model. Solid line at the top indicates error in assuming $F \approx F_0$.

3.4 Conclusion

In this chapter two analytic far-field pattern models were developed with which to approximate the far-field patterns of a circular aperture reflector antenna. The Jacobi-Bessel model was derived by expanding the aperture field distribution into a series of Zernike polynomials and calculating analytically the corresponding far-field function. This yielded a pattern model that expands the far-field into a trigonometric Fourier series in the azimuth and a special case of the Neumann series in the zenith. From this result the Neumann model was derived by using a more general form of the Neumann series that models pattern variation along the zenith.

Various techniques were presented by which the model parameters could be solved. For a pattern that is accurately known over a certain angular region the method of weighted residuals was used to determine the model parameters. Using models of sufficiently high order the approximation error could be reduced to below 10^{-4} over a region extending as much as 5° (approximately five half-power beamwidths for MeerKAT) off-axis. In general the JB-models of lower order was shown to be more accurate than the same order N-models.

To address the problem of calculating model coefficients for an unknown pattern and using only a few measurements a point-matching solution was presented. Here the JB-model accuracy was limited by the highest model order that could be solved before the system matrix became ill-conditioned. On

the other hand N-models of much higher order could be solved to yield pattern models with higher accuracy. Over a region extending 1° (approximately one half-power beamwidth) off-axis the most accurate JB-model and N-model contained an error of 10^{-5} and 10^{-7}, respectively. Solving for these models required 15 pattern samples for the JB-model and 37 pattern samples for the N-model.

Finally two constrained solution approaches were considered in an attempt to reduce the number of pattern samples required to solve the parameters in a model of a given order. Here it was assumed that the unknown pattern to be modelled is relatively similar to an already known pattern, as can be expected if certain deterministic errors (e.g. feed / subreflector displacement) caused the antenna pattern to vary. For the N-models, neither of the constrained solutions was shown to reduce significantly the number of pattern samples required for an accurate solution. On the other hand, for the JB-model the Lagrange multiplier solution was shown to reduce by approximately a half the number of pattern samples required to achieve a model of the same accuracy as would be obtained using the point-matching solution. Increasing the number of pattern samples beyond this point again resulted in an ill-conditioned system and the accuracy of the JB-model was limited as before. However, this problem could be alleviated by using the second constrained solution, that is the quadratic penalty method. The same reduction in number of required pattern samples was observed in this case, and furthermore, the JB-model accuracy could be improved to contain a much smaller error than that achieved with the point-matching solution for a model of the same order.

In the next chapter a novel beam modelling technique is presented which uses discrete pattern basis functions. That method is based on the same assumption made above in applying the constrained solutions, that is, that the unknown pattern may be approximated relatively accurately by a known and ideally expected pattern. It will be shown how this novel approach yields pattern models that are much more efficient, providing very high accuracy and requiring only a few pattern samples to solve for the model parameters.

Chapter 4

Characteristic Basis Function Patterns

Recently a novel technique has been proposed for the accurate prediction of antenna radiation patterns that requires very few calibration measurements [30; 31; 32]. The success of this method is due to the fact that it models the far-field pattern as an expansion of basis functions, called *Characteristic Basis Function Patterns* (CBFPs) that contain as much physics-based information of the antenna as possible.

The CBFPs are obtained through numerical calculation using sophisticated simulation models, from direct and careful measurement where this is feasible, or a combination of both. Consequently the basis functions are discrete[1]. If desired, each of the CBFPs may further be interpolated using the analytic models presented in the previous chapter, or any other suitable analytic functions to produce a hierarchical set of basis functions [30; 70][2].

The set of basis functions comprise a *primary* CBFP which is the radiation pattern associated with the antenna under ideal operating conditions, and a number of *secondary* CBFPs that are generated to model perturbations in the antenna pattern. Since many perturbations from the ideally expected pattern can be anticipated and attributed to certain kinds of errors that are known to have a dominant effect, secondary CBFPs are then specifically constructed to compensate for such errors.

The procedure used to generate the CBFPs and to solve for the pattern model parameters is presented below. Thereafter the CBFP method is used to compensate for pattern variations resulting from various geometrical de-

[1] A related beam modelling approach which is aimed specifically at producing accurate low-order models of the patterns of Aperture Arrays (AA) can be found in [69]. In that approach the pattern basis functions are related to macro-domain current basis functions that are obtained from a method of moments solution of the AA.

[2] It should be noted that in this case only the expansion coefficients for the CBFPs need to be solved during calibration, since the CBFPs and the analytical expansions of these basis functions are assumed to be known.

formations in a reflector antenna system, where the highly efficient modelling capabilities of the method will be demonstrated by numerical results. Finally, some concluding remarks are presented.

4.1 CBFP Generation

The first step in applying the CBFP method is to generate an appropriate set of basis function patterns. As already mentioned the primary CBFP corresponds to the ideally expected radiation pattern, whereas the generation of secondary CBFPs depends on which errors are to be compensated for. Here two kinds of geometrical errors in an OG reflector antenna will be considered; these are deformation of the support arm on which the subreflector and feed are hosted, and deformation of the subreflector and main reflector surfaces.

Certain simplifications will be made regarding the effect of these errors in the antenna system. Firstly, it is assumed that deformation of the support arm only results in *positional* errors of the feed and subreflector, i.e. *rotation* of these components are ignored. Also, since the same support arm hosts both these components, the error in their positions are affected in proportion to their relative positions on the arm.

Secondly, deformation of the reflector surfaces will be simulated by using different shaped OG geometries which are controlled by a single parameter and based on an unshaped OG design. The reasons for modelling the surface deformation as such is that the single parameter model for deformation allows a systematic approach to generating CBFPs and investigating their modelling capabilities. Furthermore, in what follows it will be shown that the pattern variations resulting from this deformation are distinctly different from those resulting from deformation of the support arm, as described above. Consequently the efficiency of the CBFP method when compensating for very different sources of error in the radiation pattern can be studied.

Nevertheless, the pattern variations resulting from these simplified deformations are significant and will suffice to demonstrate the potential of the CBFP method.

4.1.1 The primary CBFP

The primary CBFP is obtained by sampling in S different positions $\Omega_S = \{(\theta_i, \phi_i)\}_{i=1}^{S}$ the radiation pattern $\mathbf{F}_0$ corresponding to the antenna operating under ideal conditions. As this operation is typically only required once (or rarely at the most), the directions are densely arranged over a wide angular region $\theta_j \leq \theta_{\max}$ for $j = 1, 2, \ldots, S$. The idea here is to sample the CBFP in as many possible directions, or as many as would be required to satisfactorily model the actual radiation pattern of the antenna operating under non-ideal

conditions. Sampling obtains both polarizations of the complex voltage pattern to yield the primary CBFP $\mathbf{f}_p$ represented as a $2S \times 1$ complex vector.

4.1.2 Secondary CBFPs for Support Arm Deformation

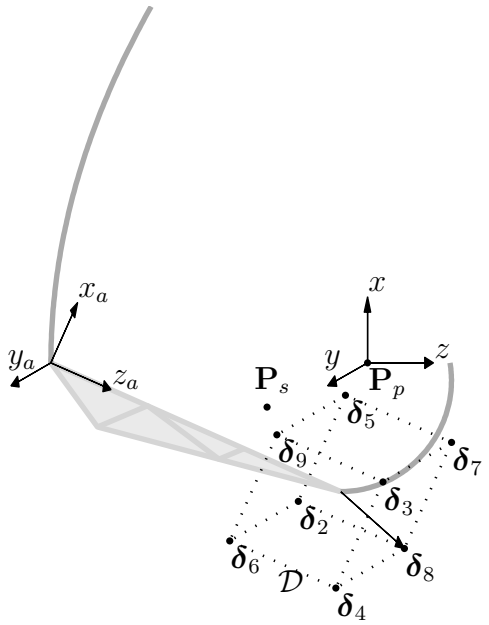


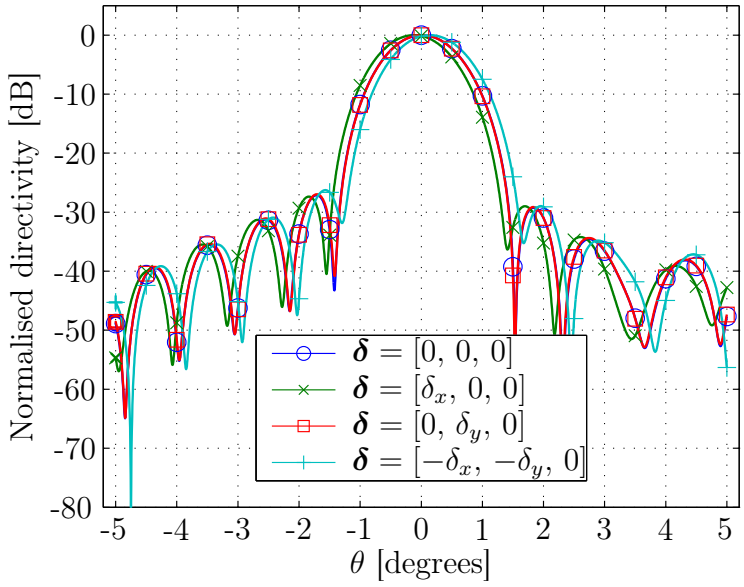
Figure 4.1: Range of possible subreflector positions resulting from deformation of the support arm, shown on an exaggerated scale for clarity. The possible feed positions (excluded in the figure) are within a similar, but smaller region centred at the ideal feed position $\mathbf{P}_s$.

Tolerances for the positional error of the feed and subreflector are tabulated in Chapter 2, and the range of possible positions of the subreflector are illustrated in Figure 4.1. Assuming that the only source of error in the antenna system is the displacement of the subreflector (and the proportional displacement of the feed), the resulting geometries and therefore the resulting radiation patterns can be defined exactly by the subreflector positional error three-vector $\boldsymbol{\delta} = [\delta_x;\, \delta_y;\, \delta_z] \in \mathcal{D}$.

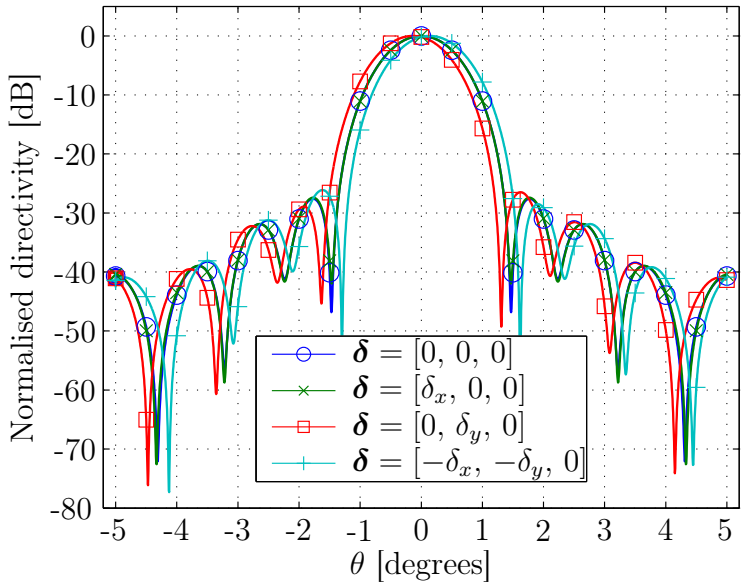
Figure 4.2 shows a number of patterns associated with the geometrical errors depicted in Figure 4.1. As expected the most prominent effect on the patterns is a pointing error, the direction of which is dependent on that of the positional error. In addition some variation in the sidelobes is also visible.

If the (dual-polarised) patterns are sampled in S directions as for the primary CBFP, then for the purpose at hand the erroneous antenna system can be viewed as a mapping

$$Y^{(\mathrm{D})} : \mathcal{D} \to \mathcal{P}^{(\mathrm{D})} \tag{4.1.1}$$



(a) $\phi = 0°$



(b) $\phi = 90°$

Figure 4.2: Perturbed patterns resulting from mechanical deformation of the support arm. $\delta_x = \delta_y = 16.5$ mm (0.0781 wavelengths).

where $\mathcal{D} \subseteq \mathbb{R}^3$, $\mathcal{P}^{(\mathrm{D})} \subseteq \mathbb{C}^{2S}$, and the parenthesised superscript 'D' indicates the mapping pertaining to the support arm deformation. Note that for the case where $\boldsymbol{\delta} = \mathbf{0}$ the resulting pattern is the ideally expected pattern or primary CBFP, that is $Y^{(\mathrm{D})}(\mathbf{0}) = \mathbf{f}_p$. Now, in generating the secondary CBFPs the objective is to obtain a set of patterns (comprising the primary CBFP and all secondary CBFPs) that spans the space of patterns resulting from the allowed

support arm deformations. Let the set of CBFPs be

$$C^{(\mathrm{D})} = C_p \cup C_s^{(\mathrm{D})} = \{\mathbf{f}_p\} \cup \left\{\mathbf{f}_{s,1}^{(\mathrm{D})}, \mathbf{f}_{s,2}^{(\mathrm{D})}, \cdots, \mathbf{f}_{s,(N-1)}^{(\mathrm{D})}\right\}, \tag{4.1.2}$$

then ideally we would have

$$\mathrm{span}\left(C^{(\mathrm{D})}\right) = \mathcal{P}^{(\mathrm{D})}. \tag{4.1.3}$$

However, this may require a total of $N = 2S$ CBFPs. Producing so many CBFPs for typically useful values of S is impractical, and instead we aim to obtain the set $C_s^{(\mathrm{D})}$ such that any pattern $\mathbf{F} \in \mathcal{P}^{(\mathrm{D})}$ may be approximated with sufficient accuracy through a linear combination of the elements in $C^{(\mathrm{D})}$.

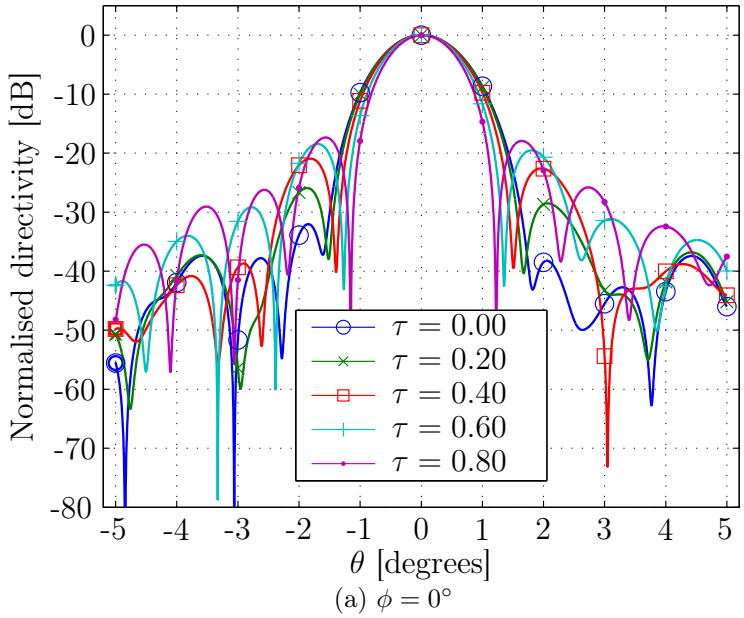
A first effort towards obtaining such a set uses each of the radiation patterns obtained with the positional error vector located at each of the corners of the domain $\mathcal{D}$ in Figure 4.1 so that $\mathbf{f}_{s,i} = Y^{(\mathrm{D})}(\boldsymbol{\delta}_i)$ for $i = 2, 3, \ldots, 9$. This approach for generating the CBFPs assumes that $Y^{(\mathrm{D})}$ can be linearised with sufficient accuracy over relatively large regions of the domain $\mathcal{D}$ and may produce inaccurate models. If higher accuracy is desired patterns may also be generated by using intermediate positions for $\boldsymbol{\delta}$. However, this accuracy comes at the cost of an increase in the number of CBFPs and consequently an increase in the number of model parameters that need to be solved.

4.1.3 Secondary CBFPs for Reflector Surface Deformation

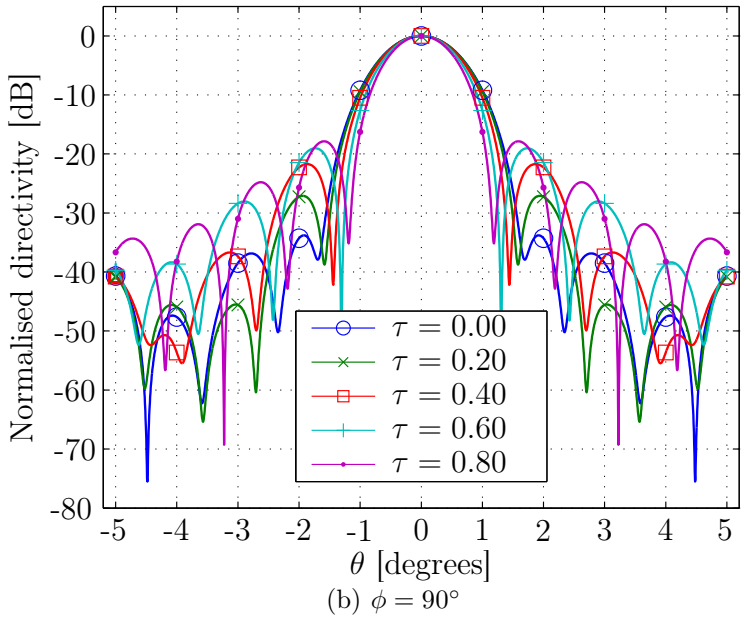
In this study a simplistic surface deformation model, which contains a single parameter $\tau \in \mathcal{T}$ that controls the aperture field distribution, is used to further illustrate the CBFP method. Using this model the geometry for a particular value of τ is obtained by shaping both the main reflector and subreflector to yield the corresponding aperture field distribution [71]. Specifically, the aperture field amplitude distribution is given by

$$E_A = \begin{cases} 1 & \text{for } 0 \leq \xi \leq \tau a \\ \exp\left[-b\left(\dfrac{\xi - \tau a}{a(1-\tau)}\right)^2\right] & \text{for } \tau a \leq \xi \leq a \end{cases} \tag{4.1.4}$$

in the polar coordinate system (ξ, ψ) of which the origin is located in the centre of the aperture. The circular aperture has a radius equal to a and the parameter b controls the aperture edge taper level. Figure 4.3 shows a number of patterns resulting from the various reflector geometries. Opposed to the pattern variations observed for deformation of the feed/subreflector support arm where the dominant effect was a pointing error, here the pattern variations mostly pertain to the main beam width and sidelobe levels. According to the aperture amplitude function in (4.1.4) for $\tau = 0$ the resulting pattern has a



(a) $\phi = 0°$



(b) $\phi = 90°$

Figure 4.3: Perturbed patterns resulting from the surface deformations obtained by varying the parameter τ.

relatively wide main beam and very low sidelobe levels, and as $\tau \to 1$ the main beam narrows as the sidelobe levels increase.

Producing secondary CBFPs to compensate for deformation of the reflector surfaces considered here follows very much the same approach as that outlined previously for the support arm deformation. Since the various deformations in this case are uniquely defined by a single parameter τ the mapping from the

erroneous geometry to a sampled radiation pattern is

$$Y^{(\mathrm{T})} : \mathcal{T} \to \mathcal{P}^{(\mathrm{T})} \tag{4.1.5}$$

where $\mathcal{T} \subseteq \mathbb{R}$, $\mathcal{P}^{(\mathrm{T})} \subseteq \mathbb{C}^{2S}$, and the parenthesised superscript 'T' indicates the mapping pertaining to the reflector surface deformation. Here the objective is to obtain the set of CBFPs

$$C^{(\mathrm{T})} = C_p \cup C_s^{(\mathrm{T})} = \{\mathbf{f}_p\} \cup \left\{\mathbf{f}_{s,1}^{(\mathrm{T})}, \mathbf{f}_{s,2}^{(\mathrm{T})}, \cdots, \mathbf{f}_{s,(N-1)}^{(\mathrm{T})}\right\} \tag{4.1.6}$$

such that

$$\mathrm{span}\left(C^{(\mathrm{T})}\right) = \mathcal{P}^{(\mathrm{T})}. \tag{4.1.7}$$

As before and for practical purposes only a few CBFPs are be generated to form $C_s^{(\mathrm{T})}$ such that every $\mathbf{F} \in P^{(\mathrm{T})}$ can be approximated with sufficient accuracy through a linear combination of the patterns in $C^{(\mathrm{T})}$.

Although the specific deformation of the reflector surfaces considered here is somewhat simplistic, it should be noted that the outlined procedure by which secondary CBFPs may be created is generally applicable. The objective should be to produce a number of linearly independent patterns that are representative of the variations that are expected to occur due to a certain kind of system error.

For example, the effects of solar heating on reflector antennas in the Allen Telescope Array are studied in [23, § II-J]. Using radio holography the aperture field distribution of an antenna was determined at different times of the day, and exhibited effects which could be related to the position of the sun. Applying the CBFP method to compensate for these pattern variations may generate secondary CBFPs by measuring the antenna pattern at different times of the day. In this case a one-parameter model for surface deformation as above applies where τ is related to the time of day.

Another cause of surface deformation (and generally other deformations of the antenna structure) is the variable gravitational loading which is dependent on the pointing of the antenna [72]. In order to compensate for the resulting pattern variations secondary CBFPs may be generated by measuring radiation patterns at various positions in the sky[3]. Depending on the telescope mount a one-parameter deformation model as above or a two-parameter model may be used.

4.1.4 Compensating for Multiple Simultaneous Errors

In the previous two sections secondary CBFP sets were generated to compensate separately for different geometrical errors, i.e. deformation of the support

[3]If the CBFPs are generated here by measurement the support of the basis functions may be limited to a region over which the deformations may be approximated as being constant.

arm, and deformation of the reflector surfaces. Here we briefly discuss the case where secondary CBFPs are used to compensate simultaneously for different sources of error.

Supposing that we wish to compensate for both deformation of the support arm as well as deformation of the reflector surfaces, it may seem sufficient to construct the set

$$C^{(\mathrm{D,T})} = C_p \cup C_s^{(\mathrm{D})} \cup C_s^{(\mathrm{T})} \tag{4.1.8}$$

by simply combining the primary CBFP and secondary sets of CBFPs obtained separately for the different errors. However, for the mapping from combined geometrical errors to sampled radiation patterns we have

$$Y^{(\mathrm{D,T})} : \mathcal{T} \times \mathcal{D} \to \mathcal{P}^{(\mathrm{T,D})} \tag{4.1.9}$$

where $\mathcal{T} \times \mathcal{D} \subseteq \mathbb{R}^4$ and $\mathcal{P}^{(\mathrm{T,D})} \in \mathbb{C}^{2S}$. Note that $\mathcal{P}^{(\mathrm{T})} \cup \mathcal{P}^{(\mathrm{D})} \subseteq \mathcal{P}^{(\mathrm{T,D})}$ so that all patterns resulting from combined reflector and support arm deformations may not be a linear combination of the patterns obtained from separate deformations of these components. In order to properly compensate for simultaneous multiple errors a more appropriate set of CBFPs would therefore be

$$C^{(\mathrm{D,T})} = C_p \cup C_s^{(\mathrm{Z})} \tag{4.1.10}$$

where elements in $C_s^{(\mathrm{Z})}$ are obtained from $\mathbf{f}_{s,i}^{(\mathrm{Z})} = Y^{(\mathrm{D,T})}(\boldsymbol{\zeta}_i)$ and $\boldsymbol{\zeta} = [\boldsymbol{\delta}; \tau] \in \mathcal{Z} = \mathcal{T} \times \mathcal{D}$.

Since $Y^{(\mathrm{D,T})}(\boldsymbol{\zeta}_i\,|_{\tau=0}) \in \mathcal{P}^{(\mathrm{D})}$ and $Y^{(\mathrm{D,T})}(\boldsymbol{\zeta}_i\,|_{\boldsymbol{\delta}=\mathbf{0}}) \in \mathcal{P}^{(\mathrm{T})}$, it is noted that the set of CBFPs in (4.1.10) may be generated by using the set in (4.1.8) and simply adding CBFPs that result from errors $\boldsymbol{\zeta}\,|_{\boldsymbol{\delta}\neq\mathbf{0},\tau\neq 0}$. In addition to possibly reducing the effort in constructing the appropriate set of CBFPs, this may also improve the modelling capabilities of the CBFP set for situations where errors expected to be dominant in the antenna system are in fact *not* dominant.

With sets of CBFPs constructed to compensate for certain errors in the antenna system, the next section discusses how these sets are applied to produce an accurate pattern model.

4.2 Calculating Model Parameters

In this section we develop the procedure by which the CBFP model parameters are determined. Since the CBFPs are discrete functions the objective here is to approximate the vector $\mathbf{F}$ which is formed by sampling the far-field pattern $\mathbf{F}$ in S directions Ω_S. The set of CBFPs is assumed to be $C = \{\mathbf{f}_1, \mathbf{f}_2, \cdots, \mathbf{f}_N\}$ and produces the model

$$\tilde{\mathbf{F}} = \mathbf{Z}\mathbf{x} \tag{4.2.1a}$$

where $\mathbf{Z}$ is the $2S \times N$ matix

$$\mathbf{Z} = \begin{bmatrix} \mathbf{f}_1 & \mathbf{f}_2 & \cdots & \mathbf{f}_N \end{bmatrix}. \tag{4.2.1b}$$

The CBFPs are sampled such that

$$(\mathbf{f}_i)_j = \begin{cases} F_i^{(\mathrm{CP})}(\theta_j, \phi_j) & \text{for } j = 1, 2, \ldots, S \\ F_i^{(\mathrm{XP})}(\theta_{S-j}, \phi_{S-j}) & \text{for } j = S+1, S+2, \ldots, 2S \end{cases} \tag{4.2.2}$$

where $F_i^{(\mathrm{CP})}$ and $F_j^{(\mathrm{XP})}$ are the co- and cross-polarised components, respectively, of the far-field pattern $\mathbf{F}_i$. The sampling points Ω_S satisfy $\theta_i \leq \theta_j$ for $i < j$ so that the support of the CBFPs can be restricted to $\theta \leq \theta_R$ by only keeping the rows $1, 2, \ldots, S_R$ and $S+1, S+2, \ldots, S+S_R$ in each $\mathbf{f}_i$ and where $\theta_1, \theta_2, \ldots, \theta_{S_R} \leq \theta_R$. (The second set of rows may be dropped if only the co-polarised component is required.)

Now the goal is to find a solution which minimises the error between the actual far-field pattern and the model

$$\arg\min_{\mathbf{x}} \left\| \mathbf{F} - \tilde{\mathbf{F}} \right\|. \tag{4.2.3}$$

For the case where $\mathbf{F}$ is known the solution is simply

$$\mathbf{x} = \left(\mathbf{Z}^\dagger \mathbf{Z}\right)^{-1} \mathbf{Z}^\dagger \mathbf{F}, \tag{4.2.4}$$

and in the following this solution will be used as a benchmark, similar to the weighted residual solution for the analytic pattern models. Note that typically $2S > N$ so that the system in (4.2.4) is over-determined.

Of course, in practice the model coefficients $\mathbf{x}$ have to be determined from only a few pattern measurements. Here it is also assumed that these measurements are restricted to the co-polarised component of the far-field. In this case the pattern $F^{(\mathrm{CP})}$ is only measured at the positions $\Omega_M = \{(\theta_{m,i}, \phi_{m,i})\}_{i=1}^{i=M}$ and the M corresponding rows in $\mathbf{Z}$ are used to produce the linear system

$$\mathbf{Z}_M \mathbf{x} = \mathbf{V}_M \tag{4.2.5}$$

where $\mathbf{Z}_M$ is an $M \times N$ matrix ($M \geq N$ for a unique solution), $\mathbf{V}_M$ is an $M \times 1$ column vector, and

$$(\mathbf{Z}_M)_{ij} = F_j^{(\mathrm{CP})}(\theta_{m,i}, \phi_{m,i}) \tag{4.2.6a}$$

$$(\mathbf{V}_M)_i = F^{(\mathrm{CP})}(\theta_{m,i}, \phi_{m,i}). \tag{4.2.6b}$$

Since the CBFPs are sampled radiation patterns of antennas that are very similar (in fact, the radiation patterns are generally also very similar), it is reasonable to assume that there would be a certain degree of redundancy in the basis C. As more patterns are added to the set of CBFPs, this degree of redundancy increases and may eventually yield a rank deficient matrix in (4.2.5), especially due to the sparse sampling used to produce $\mathbf{Z}_M$. In that case it is necessary to remove any redundant CBFPs from the set, and to improve the conditioning of (4.2.5). Such a procedure will now be presented.

4.2.1 Removing Redundancy in Basis Function Set

Suppose that the pattern measurements used to produce the system in (4.2.5) are limited to a certain angular region Ω_R and that only co-polarised measurements are performed. It is then desirable to orthogonalise the basis functions over this region; we start by forming the matrix[4]

$$\mathbf{Z}_R = \left[\begin{array}{cccc} \mathbf{f}_{1,R} & \mathbf{f}_{2,R} & \cdots & \mathbf{f}_{N,R} \end{array} \right] \tag{4.2.7}$$

where

$$\left(\mathbf{f}_{i,R}\right)_j = F_i^{(\mathrm{CP})}(\theta_j, \phi_j) \qquad \text{for } j = 1, 2, \ldots, S_R. \tag{4.2.8}$$

Next we compute Singular Value Decomposition (SVD) of $\mathbf{Z}_R$ to give

$$\mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^\dagger = \mathbf{Z}_R. \tag{4.2.9}$$

The left-singular vectors $\mathbf{u}_i$ with $i = 1, 2, \ldots, N$ (the first N columns of $\mathbf{U}$) form an orthonormal basis for the column space of $\mathbf{Z}_R$, and a reduced set C_R of CBFPs may now be obtained by keeping the first $N_\sigma \leq N$ left-singular vectors for which the singular values are above some tolerance $\sigma_i > \varepsilon$ to give

$$C_R = \left\{ \mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_{N_\sigma} \right\}. \tag{4.2.10}$$

This basis is orthonormal over its support Ω_R and may be used to construct a CBFP model over Ω_R as

$$\tilde{\mathbf{F}}_R = \mathbf{R}_R \mathbf{y} \tag{4.2.11}$$

where $\mathbf{R}_R$ is formed by stacking the elements in C_R as columns, and $\mathbf{y}$ is the model coefficients vector. This model may be used to approximate the far-field $\mathbf{F}$ over the region Ω_R by solving $\mathbf{y}$ in

$$\mathbf{R}_M \mathbf{y} = \mathbf{V}_M \tag{4.2.12}$$

where $\mathbf{V}_M$ may be the same as in (4.2.5) and $\mathbf{R}_M$ is formed by keeping the rows in $\mathbf{R}_R$ corresponding to the positions Ω_M where $\mathbf{F}$ is sampled to form $\mathbf{V}_M$.

If so desired the model $\tilde{\mathbf{F}}$ with support over Ω_S (and containing both polarisation components) can be recovered by evaluating

$$\tilde{\mathbf{F}} = \mathbf{R}\mathbf{y} \tag{4.2.13}$$

where $\mathbf{R}$ is the matrix formed by stacking the columns

$$\mathbf{r}_i = \frac{1}{\sigma_i} \mathbf{Z}\mathbf{v}_i. \tag{4.2.14}$$

Note that it follows from the definition of the SVD that the first S_R elements in $\mathbf{u}_i$ and $\mathbf{r}_i$ are equal.

[4] The procedure outlined in this section is illustrated by use of a numerical example in Appendix C.

Finally, it is noted that there is merit in keeping all left-singular vectors in $\mathbf{U}$ as basis functions, i.e. the case where $N_\sigma = N$. By using this set of basis vectors instead of the generated CBFPs $\mathbf{f}_i$ directly at least generally yields a better conditioned system in (4.2.12) than the one in (4.2.5), since all the singular values of $\mathbf{U}$ are exactly equal to one.

4.3 Numerical Results

Using the above procedures to generate basis function sets and to solve for the model coefficients, the CBFP method was applied to model far-field patterns corresponding to various deformed geometries of the MeerKAT OG system. Unless specifically noted otherwise, all results are at the operating frequency of 1.42 GHz. Once again the same error metric as used in Chapter 3 will be used to measure the accuracy of the various obtained pattern models.

4.3.1 Reflector Surface Deformations

As a first example the CBFP method will be used to model pattern variations resulting from the reflecting surface deformations in Section 4.1.3. Here it is assumed that the actual pattern to be modelled is the far-field pattern of the shaped offset Gregorian design with $\tau = 0.25$ in (4.1.4), that under ideal conditions the pattern is that which corresponds to $\tau = 0.4$, and that $\mathcal{T} = 0 \leq \tau \leq 0.8$. The primary CBFP is therefore $\mathbf{f}_p = Y^{(\mathrm{T})}(0.4)$, and secondary CBFP sets of various sizes were used to construct pattern models. To determine the effect of the sample density in $\mathcal{T}$ on the model accuracy the various CBFP sets used were as in (4.1.6) with

$$\mathbf{f}_{s,i} = Y^{(\mathrm{T})}(\tau_i) \tag{4.3.1}$$

and the various values for τ_i used in each set as tabulated in Table 4.1. The patterns were sampled over a region $\theta \leq 5°$.

CBFP set	Values of τ_i for secondary CBFPs	Number of CBFPs N
C_1	$\tau_i \in \emptyset$	1
C_2	$\tau_i \in \{0, 0.8\}$	3
C_3	$\tau_i \in \{0, 0.2, 0.6, 0.8\}$	5
C_4	$\tau_i \in \{0, 0.1, 0.2, 0.3, 0.5, 0.6, 0.7, 0.8\}$	9

Table 4.1: Secondary CBFP generation for reflector surface deformation.

For each set of CBFPs two sets of model coefficients were solved, one by solving the overdetermined system in (4.2.4), and the other by solving the determined system in (4.2.5). To produce the determined system the

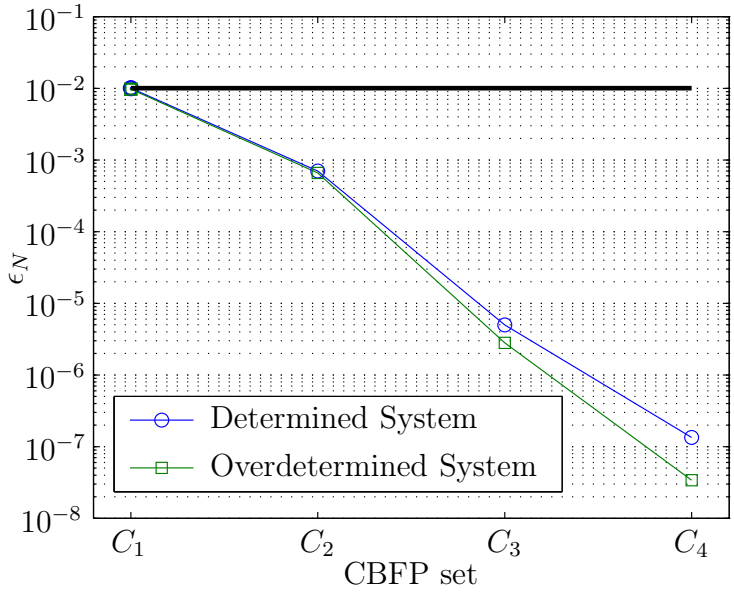


Figure 4.4: Error in CBFP models using the various sets in Table 4.1. Solid line indicates error in assuming pattern corresponds to ideal operating conditions.

pattern measurement positions were unrestricted within the $\theta \leq 5^\circ$ region, and the positions were selected to minimise the condition number $\kappa(\mathbf{Z}_M)$. Correspondingly the error in the model pattern ϵ_N was computed over the same region.

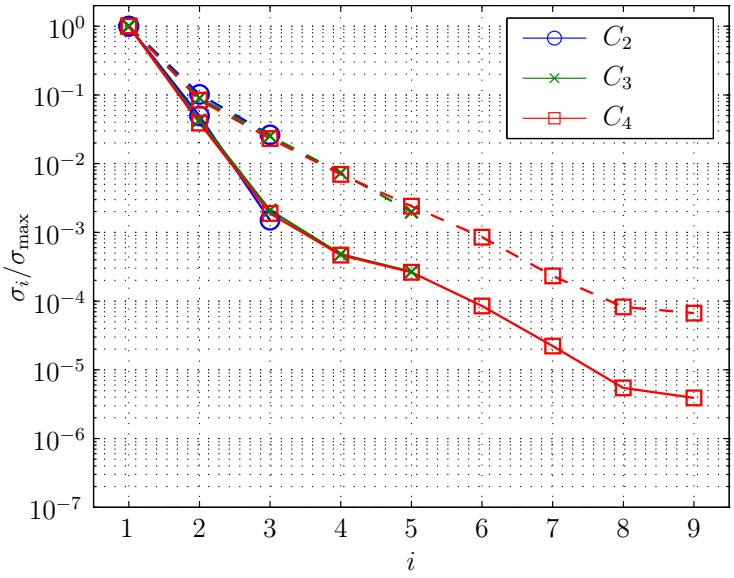
Figure 4.4 shows how the model accuracy decreases as CBFP sets of increasing number of patterns are used. C_1 contains only the primary CBFP, which is exactly the ideally expected pattern and the corresponding pattern model achieves the same accuracy $\epsilon_N \approx 10^{-2}$ as assuming the ideal pattern. With two additional CBFPs the error reduces by just over an order of magnitude, and reduces much further as more CBFPs are added until $\epsilon_N \approx 10^{-7}$ for set C_4 which contains nine CBFPs. As expected, for all sets the model obtained from the overdetermined system is slightly more accurate, especially for the larger sets, although high accuracy is still achieved using the determined system.

Next the support of the models was reduced to

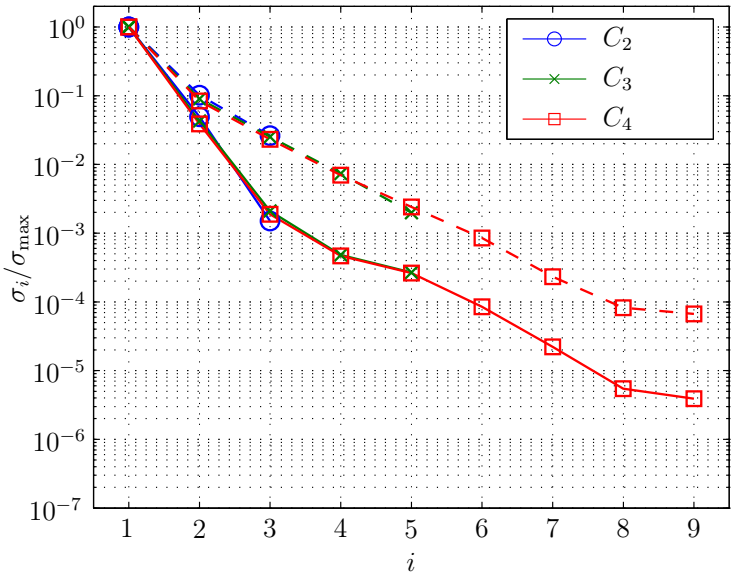
$$\Omega_R = \{(\theta_i, \phi_i) \,:\, (\theta_i, \phi_i) \in \Omega_S,\, \theta_i \leq \theta_R\}$$

and the SVD used to construct orthonormal bases over these regions from the various CBFP sets. The singular value spectra are shown in Figure 4.5 (a) and (b) for $\theta_R = 0.5^\circ$ and $\theta_R = 1.0^\circ$, respectively. As a comparison the results as computed for $\theta_R = 5.0^\circ$, i.e. orthonormalising over the entire Ω_S are also shown. As expected the degree of redundancy increases as θ_R is decreased.

Using the orthonormalised bases CBFP models were solved by sampling the far-field pattern within the respective angular regions and using the determined



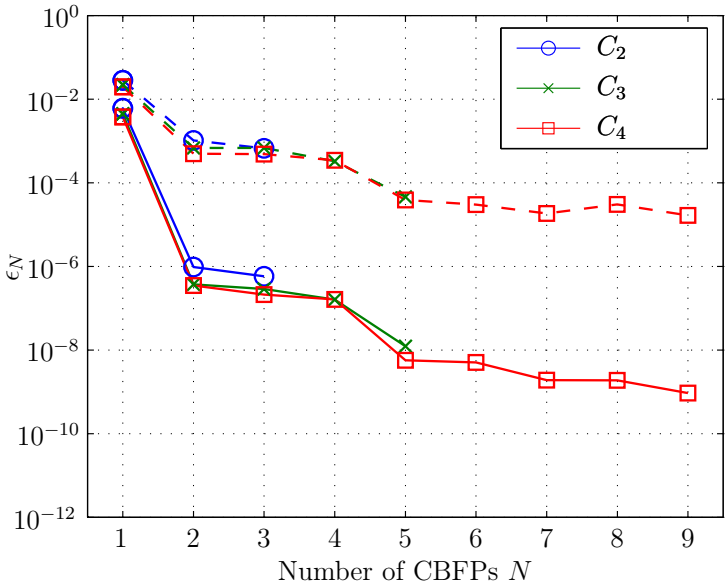
(a) $\theta_R = 1.0°$



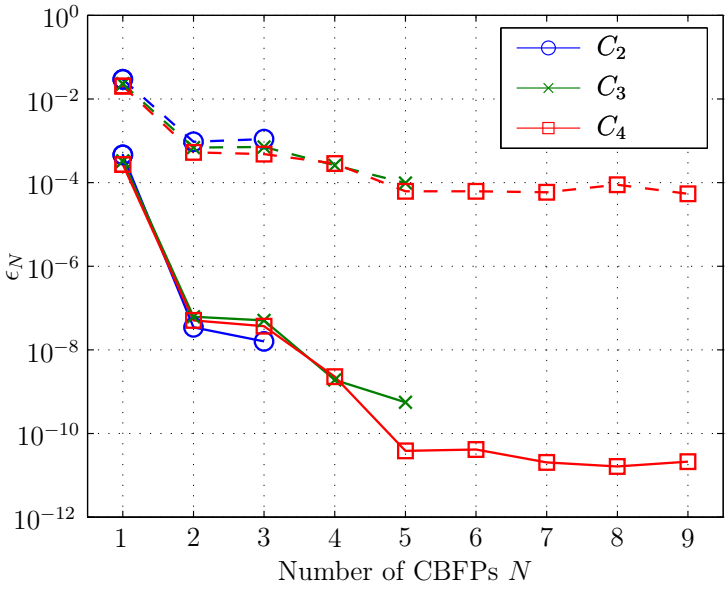
(b) $\theta_R = 0.5°$

Figure 4.5: Singular value spectra for various sets of CBFPs computed over different regions Ω_R. The singular value spectra for $\theta_R = 5°$ are shown on each plot as dashed lines.

system in (4.2.5). The support of the various models were then increased to Ω_S using (4.2.13) and the error in each pattern model was computed over two regions; these are the region Ω_R over which the basis used to construct the model was orthonormalised, and the entire region Ω_S over which the original CBFPs were defined. For each basis the error was also computed as a function



(a) $\theta_R = 1.0°$



(b) $\theta_R = 0.5°$

Figure 4.6: Error as a function of the number N of CBFPs used in the model. Solid lines and dashed lines indicate the error computed over Ω_R and Ω_S, respectively.

of how many elements from the basis were used to solve and construct a model. The results shown in Figure 4.5 (a) and (b) for $\theta_R = 0.5°$ and $\theta_R = 1.0°$, respectively.

Overall it is observed that the model error when computed over the respective Ω_R regions is much smaller than the same model error when computed

over Ω_S. This is due to the fact that pattern measurements are only taken within the region Ω_R so that the actual pattern shape outside this region does not affect the model solution at all. The results shown in Figure 4.4, especially those pertaining to the determined system solution serve as an indication of the highest accuracy that may be achieved. For example, using C_4 and measuring the far-field pattern anywhere within Ω_S the error is in the order of 10^{-7}. On the other hand, by limiting measurements to be within Ω_R as was done here results in an error in the order 10^{-5}.

Comparing the model errors computed over Ω_R for $\theta_R = 0.5°$ and that for $\theta_R = 1.0°$ it is also observed that the former is much lower (approximately two orders of magnitude) than the latter. This could be anticipated since in case of either the smaller or larger region the pattern model is exactly equal to the pattern at the same number of points. The only difference is that for the smaller region these points are in closer proximity than what they are for the larger region. However, when comparing the errors computed over the entire Ω_S for the models solved over the smaller and larger regions, the model solved over the larger region is more accurate.

Finally, for the set C_4 it is seen that as the number of terms in the pattern model is increased beyond five the model error does not decrease significantly. From the results in Figure 4.5 we note that for $i > 5$ the singular values are already very small $\sigma_i < 10^{-4}$, indicating that the additional basis functions may not be contributing dominant pattern components. Still, for the models solved over the larger region the inclusion of all the terms as opposed to using only five terms reduces the error by roughly a factor six; for the models solved over the smaller region the same reduces the error by less than a factor two.

Figure 4.7 shows the ideally expected pattern $\mathbf{f}_p$ (primary CBFP), the actual pattern to be modelled $\mathbf{F}$, and the CBFP model $\tilde{\mathbf{F}}$ as a visual indication of the accuracy that is achieved. Also shown are the secondary CBFPs (before orthonormalisation over the region $\theta \leq 1°$) and the set C_3 was used to solve and construct the model. Considering the significant difference between the ideally expected pattern and the actual pattern, especially in the sidelobe regions the CBFP model is remarkably accurate and the model error is only distinguishable from around -40 dB below the pattern maximum.

A more quantitative representation of the error in the pattern model is shown in Figure 4.8. For the model solved over $\theta \leq 1.0°$ in Figure 4.8 (a) the maximum error everywhere within that region is -74 dB relative to the pattern maximum, whereas the same error everywhere within Ω_S (not shown in figure) is -54 dB relative to the pattern maximum. Similarly for the model solved over $\theta \leq 0.5°$ in Figure 4.8 (b) the maximum errors are -86 dB and -52 dB relative to the pattern maximum within the regions Ω_R and Ω_S, respectively. Furthermore, the model obtained for $\theta \leq 0.5°$ is overall less accurate outside Ω_R than the model for $\theta \leq 1.0°$. This is in agreement with the results in Figure 4.6 where it was observed that higher accuracy within Ω_R resulted in lower accuracy outside Ω_R.

4.3.2 Support Arm Deformations

Next the CBFP method is applied to compensate for pattern variations resulting from deformation of the support arm as in Section 4.1.2. Here the actual pattern to be modelled is the secondary pattern of the unshaped MeerKAT optical design with an arbitrary positional error $\boldsymbol{\delta}_\varepsilon$ applied to the subreflector (and a proportional error applied to the feed). CBFPs are orthonormalised over the region Ω_R with $\theta_R = 1.0°$, and the pattern sampling points used to construct (4.2.12) are limited to within that same region. The primary CBFP is in all cases the ideally expected pattern $\mathbf{f}_p = Y^{(\mathrm{D})}(\mathbf{0})$.

First positional errors in the x and y directions are considered separately. The sets of CBFPs used to model the resulting patterns are shown in Figure 4.9 (a) for $\boldsymbol{\delta}_\varepsilon = [\epsilon_x; 0; 0]$ and (b) for $\boldsymbol{\delta}_\varepsilon = [0; \varepsilon_y; 0]$. After applying the SVD in (4.2.9) to each of the sets of CBFPs only the left-singular vectors with $\sigma_i/\sigma_{\max} > 10^{-6}$ were used.

Figure 4.10 (a) and (b) show the model accuracy as a function of the number of CBFPs for separate x- and y-directed errors, respectively, and compares the results for the various sets shown in Figure 4.9. As before the CBFP pattern models are observed to be extremely accurate and the error over Ω_S is reduced to $\epsilon_N \approx 10^{-10}$ by using only three or four terms. In general the model errors (as well as the error in assuming the ideal pattern) for y-displacement are smaller than those for x-displacement since the tolerances for position-

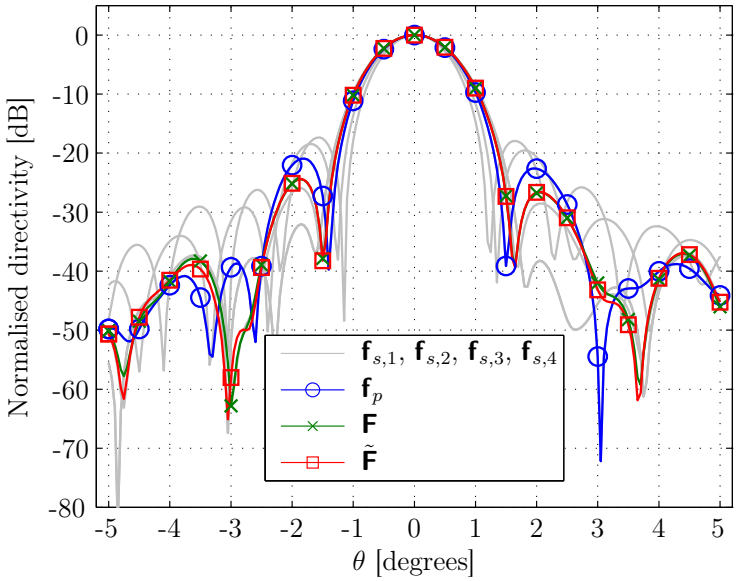
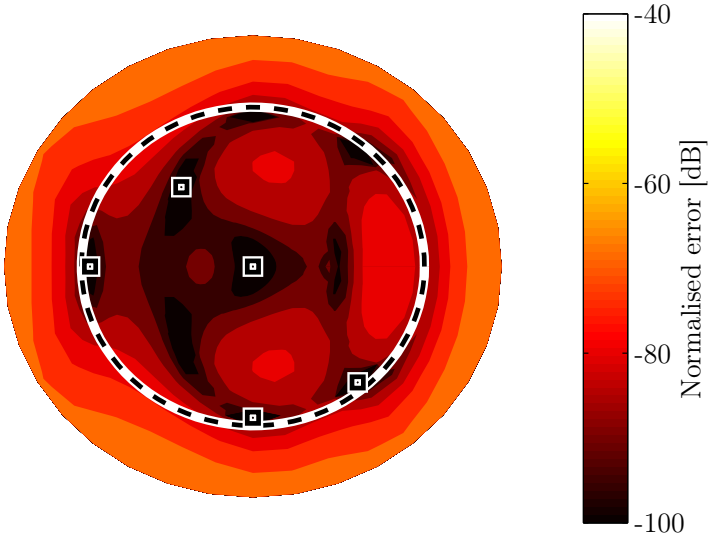


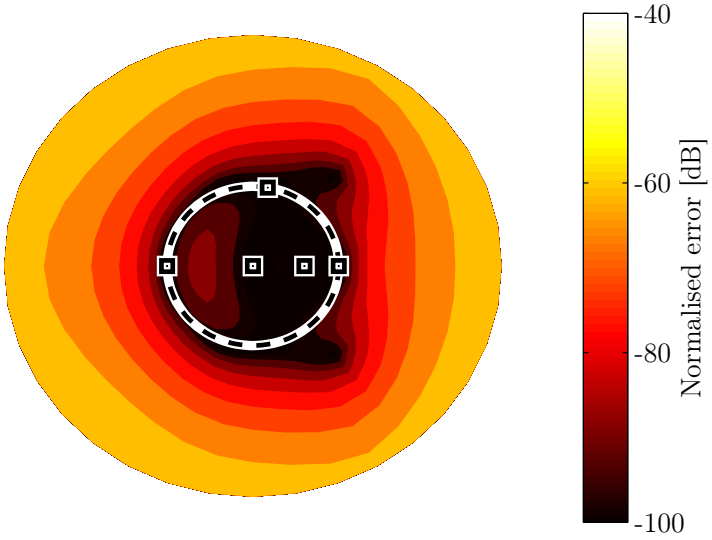
Figure 4.7: CBFP model $\tilde{\mathbf{F}}$ using all five terms from C_3 and orthonormalising over $\theta \leq 1°$. Also shown are the ideally expected pattern (or primary CBFP $\mathbf{f}_p$), the actual pattern $\mathbf{F}$, and secondary CBFPs $\mathbf{f}_{s,i}$. Patterns are shown in the plane $\phi = 0°$.

ing errors of the feed/subreflector are smaller in the y-direction than in the x-direction.

Next CBFPs were used to model a pattern resulting from a combined (x, y) error $\boldsymbol{\delta}_\varepsilon = [\varepsilon_y; \varepsilon_y; 0]$. For this purpose three different sets of CBFPs were constructed, as shown in Figure 4.11. Set C_1 is simply a combination of the sets $C_{x,3}$ and $C_{y,3}$ above, set C_2 uses as secondary CBFPs patterns obtained by applying displacements diagonally in the (x, y)-plane, and set C_3 is created

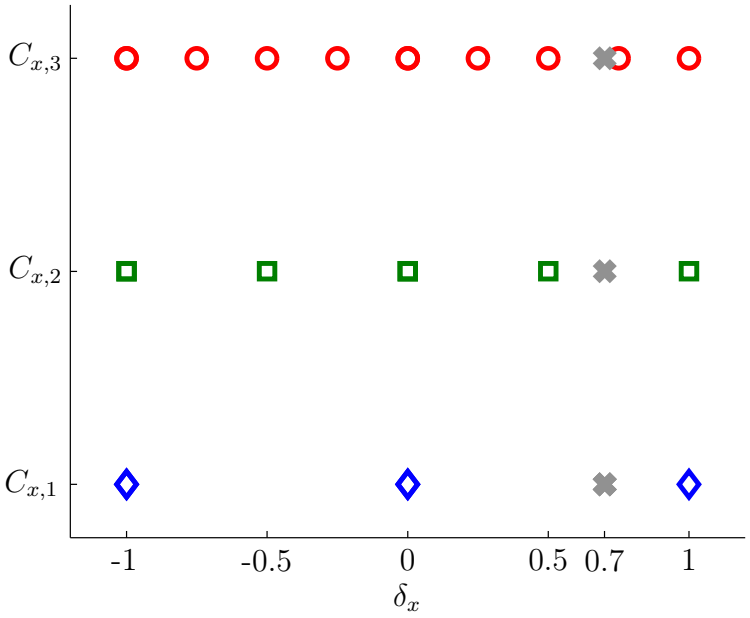


(a) $\theta_R = 1.0°$

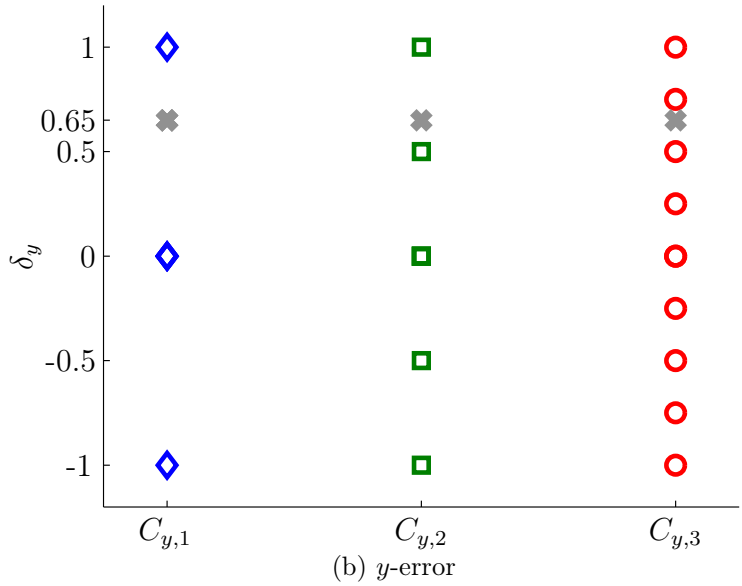


(b) $\theta_R = 0.5°$

Figure 4.8: Model error relative to pattern maximum shown over the region $\theta \leq 1.5°$. The boundary of Ω_R in each figure is shown as a dashed black line and pattern sampling positions to construct (4.2.12) are indicated with '□'s.



(a) x-error

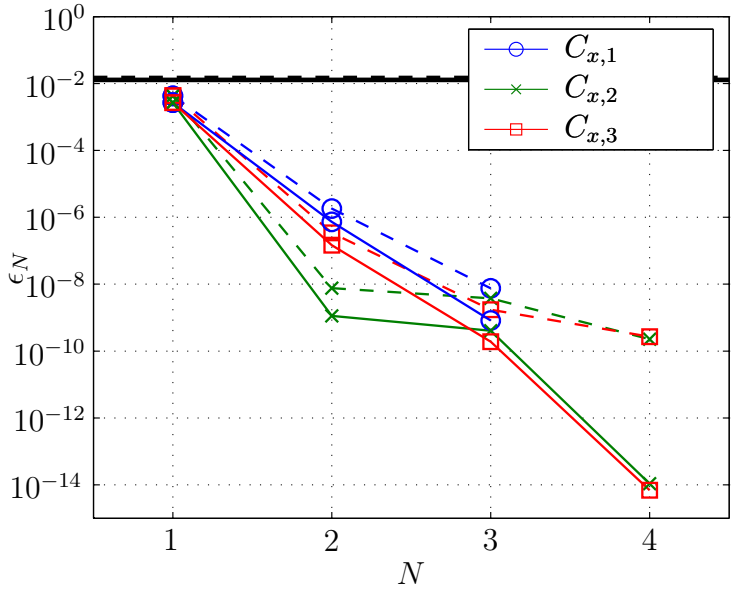


(b) y-error

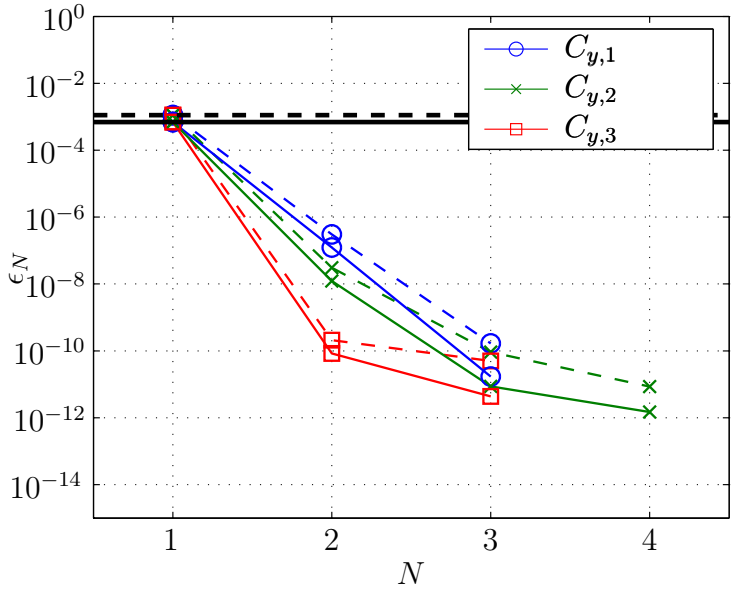
Figure 4.9: Displacement vectors for CBFP generation for separate x- and y-errors. Displacements corresponding to the actual pattern **F** to model are indicated with '×' markers. δ_x and δ_y axes are normalised to the maximum possible displacement along each axis.

from the patterns obtained by sampling the displacement vector over a grid in the (x, y)-plane.

Figure 4.12 shows the singular value spectra for sets C_1 and C_2. The singular values obtained after orthonormalising the various CBFP sets are



(a) x-error



(b) y-error

Figure 4.10: Error as a function of the number N of CBFPs used in each model. Dashed lines indicate error computed over Ω_S, solid lines indicate error over Ω_R, and solid black line indicates error in assuming ideal pattern.

shown in Figure 4.12. Left-singular vectors for which $\sigma_i/\sigma_{\max} < 10^{-6}$ were discarded, and resulted in a total of six basis functions from C_1, seven basis functions from C_2, and nine basis functions from C_3. It seems plausible that, as evidenced by the slightly slower decay of singular values for C_2 than those of C_1, the degree of linear dependence in C_2 be less than that in C_1, due

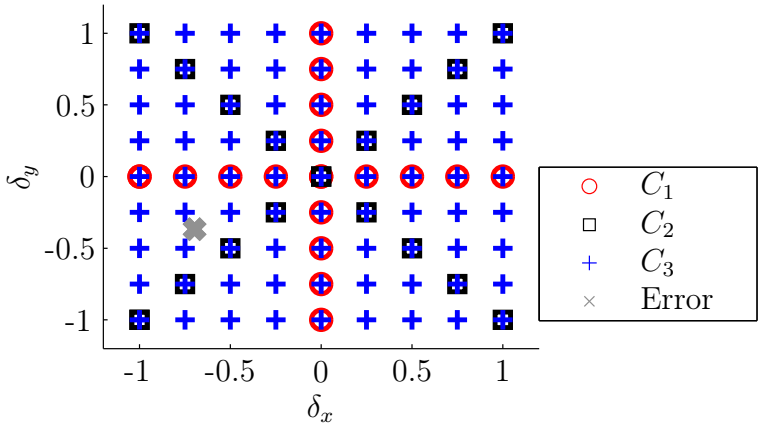


Figure 4.11: Displacements applied for CBFP generation for (x, y)-plane errors.

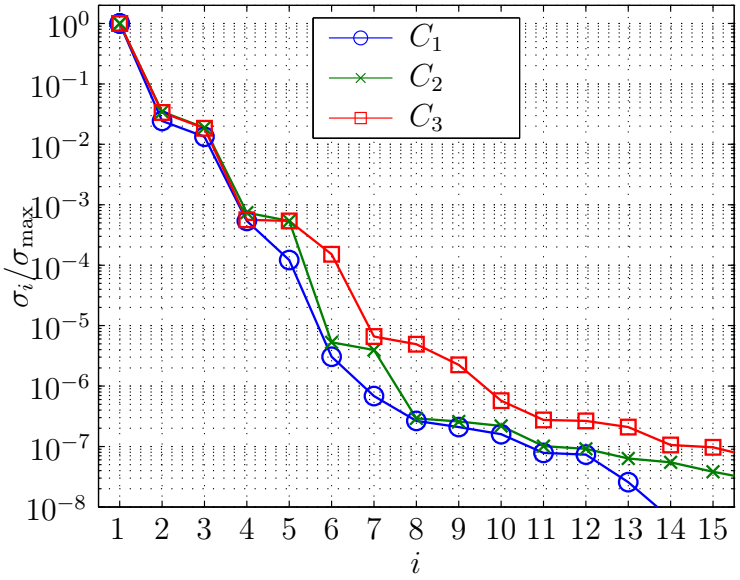
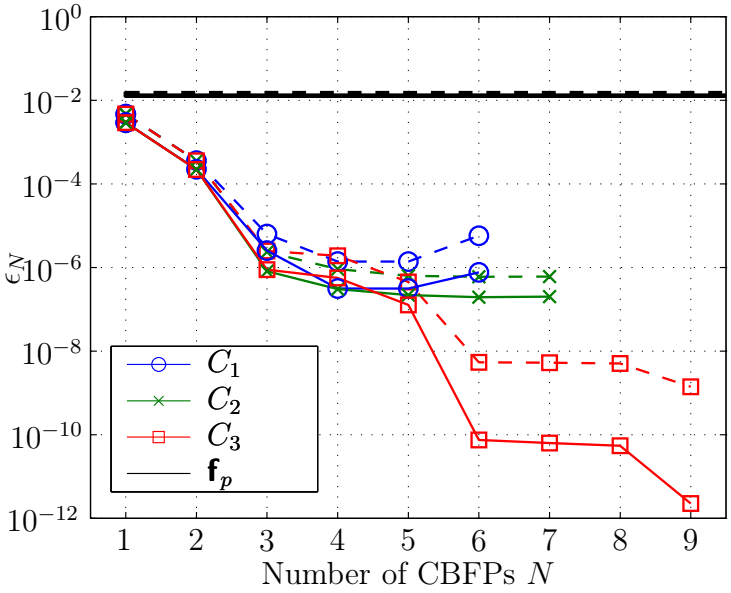


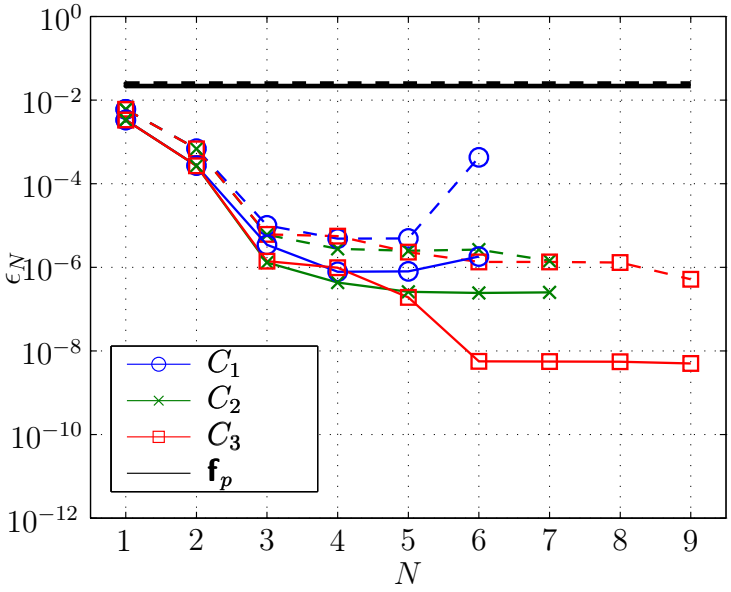
Figure 4.12: Singular value spectra for C_1 and C_2.

to the fact that the farthest displacement used to produce CBFPs in C_2 is physically larger than the same in C_1. However, it is clear that C_3 should be the preferred method of producing CBFPs to compensate for the assumed displacement error.

After solving for the coefficients vectors of the various pattern models the support of these models was extended to Ω_S and to include the cross-polarised component using (4.2.13). Figure 4.13 (a) shows the error in various models of the co-polarised field component as a function of the number of terms in each model. Up to $N = 5$ the models constructed with any of the sets exhibit approximately the same accuracy, although beyond that higher accuracy is



(a) Co-polarised pattern



(b) Cross-polarised pattern

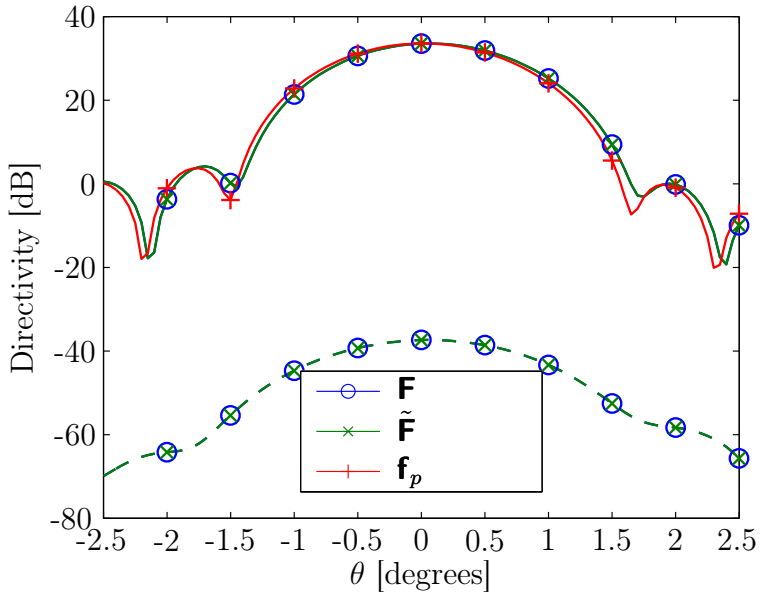
Figure 4.13: Error as a function of the number of CBFPs used in each model. Dashed lines indicate error computed over Ω_S, solid lines indicate error over Ω_R, and solid black line indicates error in assuming ideal pattern.

only obtained by using C_3. This confirms that the method used to produce secondary CBFPs for C_3 is the most appropriate.

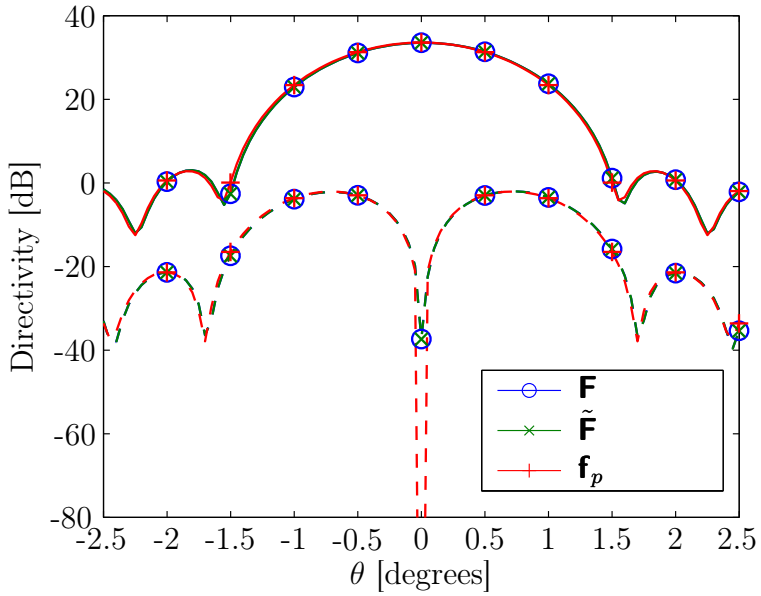
As compared to the results for only x- or y-directed displacements in Figure 4.10, here the error is seen to decrease at a slower rate. For example, achieving an error of less than 10^{-6} over Ω_S required about two or three terms

for the x- or y-displacement, whereas achieving the same error here for the (x, y)-displacement requires roughly double that.

The results for the cross-polarised field component models are shown in



(a) $\phi = 0°$



(b) $\phi = 90°$

Figure 4.14: Ideally expected, actual, and CBFP model patterns for (x, y)-displacement. Solid lines show co-polarised patterns and dashed lines show cross-polarised patterns. In (a) the cross-polarised component of $\mathbf{f}_p$ is around -180 dB and does not appear within the shown vertical axis range.

Figure 4.13 (b). As expected the error in these models are somewhat larger than those for the co-polarised field component, as no samples of the actual cross-polarised field pattern were used. Nevertheless, the error as calculated over Ω_S may be reduced to less than 10^{-6} compared to the error 10^{-2} in assuming the cross-polarised pattern corresponding to the antenna operating under ideal conditions.

The ideally expected pattern $\mathbf{f}_p$, actual pattern $\mathbf{F}$ and pattern model $\tilde{\mathbf{F}}$ for (x, y)-displacements are shown in Figure 4.14. For clarity the horizontal range is limited to $\theta \leq 2.5°$. As was noted before the geometrical deformations here mainly result in pointing errors, as can be seen in the difference between the expected and actual co-polarised patterns in the $\phi = 0°$ plane. However, due to the loss of symmetry caused by displacing the feed/subreflector the cross-polarised performance of the antenna is adversely affected and cross-polarisation in the $\phi = 0°$ plane (symmetry plane for the ideal antenna geometry) increases significantly. It is observed that these pattern variations that are present in both the co- and cross-polarisation components are compensated for in the CBFP model with high accuracy.

Finally, the CBFP method was applied to compensate for pattern variations resulting from displacement of the feed/subreflector in three dimensions $\boldsymbol{\delta}_\varepsilon = [\varepsilon_y; \varepsilon_y; \varepsilon_z]$. This was done at two frequencies; these are 580 MHz and 1.750 GHz, the lower bound of MeerKAT Band-1 and upper bound of MeerKAT Band-2, respectively. The CBFPs were created by evaluating $Y^{(\mathrm{D})}(\boldsymbol{\delta})$ over a three-dimensional grid in $\mathcal{D}$, and two separate sets of CBFPs were created at the two frequencies. After applying the SVD the basis functions cor-

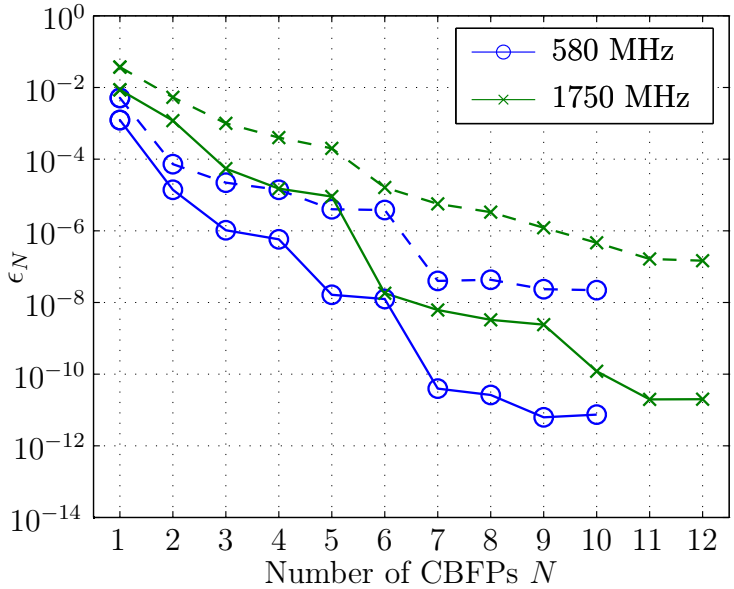


Figure 4.15: Error as a function of the number N of CBFPs used in the model. Solid and dashed lines indicate error as computed over Ω_R and Ω_S, respectively.

responding to normalised singular values below 10^{-6} were discarded. Next a number of erroneous patterns were created by obtaining the patterns $Y^{(\mathrm{D})}(\boldsymbol{\delta}_\varepsilon$ corresponding to each of fifty randomly selected displacement vectors $\boldsymbol{\delta}_\varepsilon \in \mathcal{D}$ and at each of the two frequencies. For each of these patterns a CBFP model was constructed and the error calculated. The error in the CBFP models for the patterns resulting in the least accurate models at each of the frequencies are shown in Figure 4.15 as a function of the number of terms in each model.

Note that the CBFP model for the pattern at 580 MHz only contains up to ten terms, whereas the model for the pattern at 1.750 GHz contains up to twelve terms. This is due to the fact that the same physical displacement of the feed/subreflector is electrically larger at higher frequencies, and therefore results in greater pattern variations at these frequencies. For this same reason it is also seen that a larger number of terms is required to achieve the same accuracy level at the higher frequency that at the lower frequency. With a seven term model at 580 MHz the error over the region Ω_R is reduced to less than 10^{-10}, whereas that over Ω_S is below 10^{-6}; at 1750 MHz a ten term model achieves approximately the same accuracy.

4.4 Conclusion

In this chapter the Characteristic Basis Function Patterns method was presented as a second approach towards an efficient pattern model. Using this method a number of discrete basis functions are constructed through numerical modelling or direct measurement of the far-field pattern of the complete antenna system under ideal operating conditions (primary CBFP), and various non-ideal operating conditions (secondary CBFPs) that are chosen specifically to be representative of the conditions under which the actual antenna pattern is to be modelled.

The generation of the basis functions to compensate for pattern variations resulting from mechanical deformations in a reflector antenna was discussed. Specifically, deformation of the feed/subreflector support arm and deformation of the reflector surfaces in an Offset Gregorian design were considered. This method was demonstrated to yield very efficient pattern models, and could estimate patterns with high accuracy using only very few pattern measurements. Possible redundancy in the bases and application of the Singular Value Decomposition to improve the model performance were also studied.

4.4.1 Comparison of Analytic and CBFP Models

In comparison with the analytic pattern functions presented in the previous chapter, the CBFP method is seen to be much more efficient in terms of the accuracy achieved for a given number of calibration measurements. This is to be expected since the entire antenna structure is incorporated into generating

the different CBFPs, and many of the finer pattern features are therefore already incorporated in the primary CBFP.

On the other hand, for the analytic pattern models terms up to an arbitrarily high order may be added to increase the degrees of freedom and improve the model accuracy. However, it was also shown how only a limited number of terms were effective at modelling the pattern over a given angular region, so that even then the accuracy of the analytic models were limited. In the case of the CBFP method the addition of higher order terms is determined by how well dominating errors in the antenna system can be identified, and how accurately these errors can be introduced in a controlled manner so as to allow the resulting patterns to be measured or calculated.

Chapter 5

Constrained Beamforming in a Phased Array Feed Based System

Phased Array Feed technology has in recent years experienced a growing interest from the radio astronomy community due to its ability to vastly increase the survey speed of radio telescopes by producing multiple simultaneous beams over an enlarged FoV [73; 74; 75; 15; 76]. In a PAF based system the feed array consists of a number of closely spaced antenna elements ($\leq 0.5\lambda$ spacing) and the signals of these elements are combined in a beamformer to produce the total beamformed pattern, as opposed to a cluster of horn feeds in which each beam on the sky corresponds to a single element in the feed array [74]. An important advantage of PAFs in this regard is that the small spacing between elements allows a sufficiently fine field sampling over the FoV [77; 78]. However, with this technology there are also a number of challenges that need to be addressed, specifically pertaining to the calibration of the instrument. Since the beamformed radiation pattern of a PAF system is dependent on the relative excitations in the feed array, variations in the pattern shape occur as a result of gain drifts in the system and have to be compensated for. In comparison with a single beam system where gain variations only impact on the total received power and are relatively easily calibrated for, the equivalent problem in a PAF based instrument is much more challenging [79; 80]. However, PAFs also offer a large degree of flexibility in terms of optimising the excitation of the feed to suit a particular application, and herein this property is utilised to improve the calibration of such systems.

Recently a number of beamforming schemes have been proposed within the context of improving the calibration of PAF based antennas, and some of these are briefly discussed here. One such approach aims to remove the paralactic rotation Jones-term out of the RIME for an alt-az mount telescope by de-rotating the antenna beam pattern on the sky as a source is tracked [81]. This is achieved by applying time-varying beamformer weights that constrain the beam pattern to remain constant in the sky coordinates. The method was demonstrated to reduce the rotation of the beam pattern over a significant

scan range at the cost of a loss in sensitivity.

Another beamforming scheme reduces the sensitivity ripple over the FoV of a PAF based telescope by applying directional constraints at the crossover points between adjacent beams [34]. This reduces the complexity of the calibration of the instrument and simplifies reconstruction of the original image since the intensity distortion of sources distributed over the FoV is lessened. A related approach constrains both the gain and noise response to be uniform over the FoV which aims to also improve the detection of weaker sources [82].

Finally, certain beamforming techniques are aimed at improving the beamformed radiation pattern stability [83]. In this context a directional constraint beamformer may also be used to conform the beam pattern to a specific analytic function which exhibits desirable properties, e.g. circular symmetry [35]. A great advantage in such an approach is that the beam pattern may then also be approximated with the same analytic function, thus providing an accurate beam model. In this chapter this approach, coupled with the use of a physics-based analytic pattern function is used to develop a beamforming strategy which creates beam patterns that are accurately characterised by a single-term model [36; 37].

In the following section an overview of the terminology pertaining to PAF based reflector antennas and the derivation of two beamformers that are used herein are presented. Thereafter the analytic pattern function is derived and the proposed beamforming strategy is developed. Next some numerical results for an example beamformer design based on this strategy are presented, followed by some concluding remarks.

5.1 Beamforming in a PAF Based System

A general PAF based reflector antenna is shown in Figure 5.1. The feed consists of an array of N antenna elements with a spacing of $\lesssim 0.5\lambda$ at the highest operating frequency and which is positioned near the geometrical focus of the reflector. When the ith element is excited with a unit voltage and zero phase while all other elements in the array are terminated in matching loads the primary Embedded Element Pattern (EEP) $\mathbf{f}_i$ is radiated, which in turn produces the secondary EEP $\mathbf{F}_i$ on the sky (mainly through inducing currents on the main reflector surface, but also including effects such as edge diffraction, support strut diffraction, back-radiation from the feed itself, etc.)[1]. In the following it is assumed that all primary EEPs are referenced to the geometric focus of the reflector (or the secondary focus for a dual-reflector antenna), and that all secondary EEPs are referenced to a point in the centre of the reflector antenna projected aperture.

[1] Note once again that the qualifier *primary* means the pattern of the feed illuminating the reflector, and *secondary* means the pattern radiated on the sky.

If each element in the PAF is excited with a complex voltage $\overline{w_i}$ then through superposition the total primary and secondary patterns $\mathbf{f}$ and $\mathbf{F}$, respectively, may be expressed as [84]

$$\mathbf{f} = \sum_{i=1}^{N} \overline{w_i}\mathbf{f}_i \tag{5.1.1a}$$

$$\mathbf{F} = \sum_{i=1}^{N} \overline{w_i}\mathbf{F}_i. \tag{5.1.1b}$$

Equivalently for the antenna in receive mode, an electromagnetic field incident upon the reflector antenna induces a voltage v_i at the terminals of the ith element in the PAF. Beamforming then produces a signal through the weighted summation of all these voltages to produce a single signal

$$b = \mathbf{w}^{\dagger}\mathbf{v} \tag{5.1.2}$$

where $\mathbf{v}$ is the $N \times 1$ vector of voltages that appear at the antenna terminals and $\mathbf{w}$ is the $N \times 1$ vector of excitation weights in (5.1.1). This system is illustrated in Figure 5.2 where all components in the signal chain between each antenna and the beamformer (e.g. low-noise amplifiers) are subsumed in the antenna symbols.

In practice the voltage vector $\mathbf{v}$ typically contains a desired signal component $\mathbf{v}_s$ and an undesired noise component $\mathbf{v}_n$. The noise may be due to external sources (e.g. cosmic sources, communication links) or internal sources

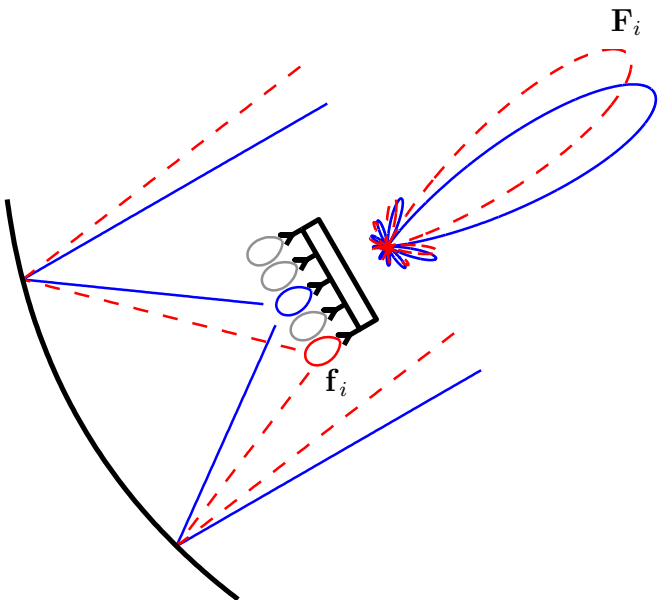


Figure 5.1: General PAF based reflector antenna.

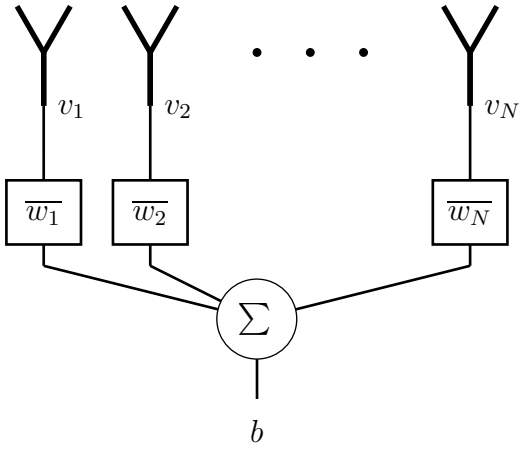


Figure 5.2: General beamforming array.

(e.g. low-noise amplifiers). Herein it is assumed that the noise is a wide-sense stationary ergodic random process [53, § 6.2] so that the noise correlation matrix $\mathbf{R}_n$, which is the correlation of the noise signals in each of the receiver channels, can be obtained through

$$(\mathbf{R}_n)_{ij} = \left\langle (\mathbf{v}_n)_i \, \overline{(\mathbf{v}_n)_j} \right\rangle \tag{5.1.3a}$$

where $\langle . \rangle$ denotes the time-average. If the signal itself is a similarly defined random process then the signal correlation matrix $\mathbf{R}_s$ may also be calculated as above. However, here it is assumed that the signal component is due to a plane wave incident upon the antenna in which case

$$\mathbf{R}_s = \mathbf{v}_s \mathbf{v}_s^\dagger. \tag{5.1.3b}$$

The signal power P_s and noise power P_n in the output of the beamformer are

$$P_s = \mathbf{w}^\dagger \mathbf{R}_s \mathbf{w} \tag{5.1.4a}$$

$$P_n = \mathbf{w}^\dagger \mathbf{R}_n \mathbf{w}. \tag{5.1.4b}$$

It is also assumed here that both $\mathbf{R}_s$ and $\mathbf{R}_n$ are known. A method by which these correlation matrices may be determined through measurement can be found in [85].

Two different beamformer designs are implemented and compared here; these are a Maximum Signal-to-Noise Ratio (MaxSNR) beamformer, and a Linearly Constrained Minimum Variance (LCMV) beamformer with directional constraints. Expressions for the excitation weights of these beamformers will now be derived.

5.1.1 Maximum Signal-to-Noise-Ratio Beamformer

One typically used beamformer maximises the SNR [86] in the beamformer output. That is, the ratio β is maximised in

$$\beta = \frac{P_s}{P_n} = \frac{\mathbf{w}^\dagger \mathbf{R}_s \mathbf{w}}{\mathbf{w}^\dagger \mathbf{R}_n \mathbf{w}}. \tag{5.1.5}$$

To this end the gradient of (5.1.5) is set equal to zero

$$\nabla_{\overline{\mathbf{w}}} \beta = \frac{\mathbf{R}_s \mathbf{w} P_n - P_s \mathbf{R}_n \mathbf{w}}{(\mathbf{w}^\dagger \mathbf{R}_n \mathbf{w})^2} = \mathbf{0} \tag{5.1.6}$$

to find a stationary point of β. Satisfying (5.1.6) requires the numerator to equal zero and a solution is obtained by solving the generalised eigenvalue problem

$$\mathbf{R}_s \mathbf{w} = \frac{P_s}{P_n} \mathbf{R}_n \mathbf{w} = \beta \mathbf{R}_n \mathbf{w}. \tag{5.1.7}$$

By setting $\mathbf{w}$ equal to the eigenvector in (5.1.7) corresponding to the largest eigenvalue the SNR is maximised. For the case where the desired signal is a plane wave the solution to (5.1.7) can be stated in closed form as [86, § 6.2.3]

$$\mathbf{w}_{\text{MaxSNR}} = \mathbf{R}_n^{-1} \mathbf{v}_s. \tag{5.1.8}$$

For uncorrelated noise of equal power in each of the receiver channels the noise correlation matrix is proportional to the identity matrix $\mathbf{R}_n = \sigma_n \mathbf{1}$ and the beamformer weights in (5.1.8) correspond to maximising the ratio of received signal power to the square of the norm of the weights vector. In order to make a clear distinction the weights vector for this case will be indicated as $\mathbf{w}_{\text{MaxSig}}$ (Maximum Signal, or MaxSig beamformer)[2].

5.1.2 Linearly Constrained Minimum Variance Beamformer

In some applications it is required that certain constraints are enforced on the beamformer, and here specifically directional constraints on the beamformed pattern are of interest. A Linearly Constrained Minimum Variance (LCMV) beamformer presents one approach through which this may be achieved, and minimises the noise output subject to a number of constraints. The derivation of the excitation weights for such a beamformer follows [86].

The objective is to minimise P_n in (5.1.4b) subject to

$$\mathbf{w}^\dagger \mathbf{C} = \mathbf{g}^\dagger. \tag{5.1.9}$$

[2] At the time of writing the correct term to use for such a beamformer is not yet clear; herein the term *Maximum Signal* beamformer is used.

For directional constraints each of the columns of $\mathbf{C}$ is the PAF voltage vector $\mathbf{v}_s(\theta_i, \phi_i)$ resulting from an incident plane wave signal from a direction (θ_i, ϕ_i) for which a constraint is defined, and the corresponding constraint value g_i is an element in the constraints vector $\mathbf{g}$. For N_Ω constraint directions and N elements in the beamforming array, the matrix $\mathbf{C}$ is $N \times N_\Omega$ and the vector $\mathbf{g}$ is $N_\Omega \times 1$. Usually the distortionless constraint $g_1 = 1$ is applied in the direction in which the beam is steered, with additional constraints used to specify relative responses towards certain other directions.

Through the method of Lagrange multipliers the above constrained minimisation problem can be reformulated as minimising the function [86, § 6.7.2]

$$\mathcal{L}(\mathbf{w}, \boldsymbol{\lambda}) = \mathbf{w}^\dagger \mathbf{R}_n \mathbf{w} + \left(\mathbf{w}^\dagger \mathbf{C} - \mathbf{g}^\dagger\right) \boldsymbol{\lambda} + \boldsymbol{\lambda}^\dagger \left(\mathbf{w}^\dagger \mathbf{C} - \mathbf{g}^\dagger\right)^\dagger . \tag{5.1.10}$$

By setting the complex gradient of $\mathcal{L}$ to zero gives

$$\nabla_{\overline{\mathbf{w}}} \mathcal{L} = \mathbf{R}_n \mathbf{w} + \mathbf{C}\boldsymbol{\lambda} = \mathbf{0} \tag{5.1.11a}$$

$$\nabla_{\overline{\boldsymbol{\lambda}}} \mathcal{L} = \mathbf{C}^\dagger \mathbf{w} - \mathbf{g} = \mathbf{0} \tag{5.1.11b}$$

in which the second equation is simply (5.1.9) rearranged. Solving for $\mathbf{w}$ in the first equation gives

$$\mathbf{w} = -\mathbf{R}_n^{-1} \mathbf{C} \boldsymbol{\lambda}. \tag{5.1.12}$$

Substitution of this expression in (5.1.11b) gives

$$\boldsymbol{\lambda} = -\left(\mathbf{C}^\dagger \mathbf{R}_n^{-1} \mathbf{C}\right)^{-1} \mathbf{g} \tag{5.1.13}$$

which after substitution in (5.1.11a) gives the desired result

$$\mathbf{w}_{\text{LCMV}} = \mathbf{R}_n^{-1} \mathbf{C} \left(\mathbf{C}^\dagger \mathbf{R}_n^{-1} \mathbf{C}\right)^{-1} \mathbf{g}. \tag{5.1.14}$$

The above derived beamformers will now be used to develop the constrained beamforming strategy used to improve calibration of the direction-dependent gain of a reflector telescope.

5.2 Beamforming Strategy

The goal here is to conform the co-polarised component of the beamformed pattern of a reflector antenna to a known analytic function (hereafter called the *reference pattern*) such that the analytic function may be used as an accurate approximation of the realised beam pattern over a relatively large angular region around the beam centre. Furthermore, this is to be achieved without too high a penalty in the sensitivity of the antenna, and using as few calibration measurements as possible. The derivation of an appropriate reference pattern and how the above beamformers are employed are the focus of this section.

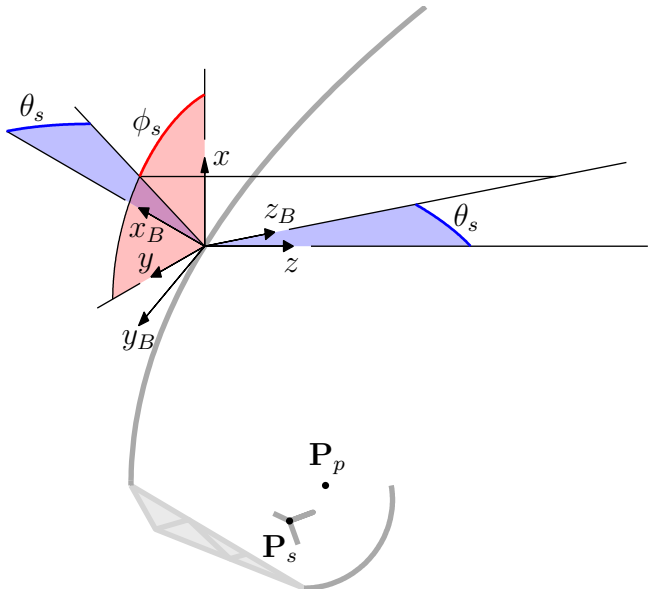


Figure 5.3: Reference pattern coordinate system.

Let the coordinate system (x_B, y_B, z_B) be defined such that $\hat{\mathbf{z}}_B$ points towards the beamformer steering direction (θ_s, ϕ_s) and that $\hat{\mathbf{x}}_B$ lies in the scan plane, as indicated in Figure 5.3. Herein the reference pattern and pattern models are then all defined in terms of the spherical coordinates (r_B, θ_B, ϕ_B) for which

$$x_B = r_B \sin\theta_B \cos\phi_B \tag{5.2.1a}$$
$$y_B = r_B \sin\theta_B \sin\phi_B \tag{5.2.1b}$$
$$z_B = r_B \cos\theta_B. \tag{5.2.1c}$$

5.2.1 Reference Pattern

A uniform amplitude and phase field distribution over a circular aperture yields the maximum directivity [63, § 12.6]

$$D_{\max} = \left(\frac{2\pi a}{\lambda}\right)^2 = (ka)^2 \tag{5.2.2}$$

which may be achieved with an aperture radius a at the frequency where the wavelength is λ (and the wavenumber k). The corresponding co-polarisation far-field function is accurately described in the region near the main beam by the jinc-function

$$F^{(\mathrm{ref})}(\theta_B, \phi_B) = \frac{J_1(ka\sin\theta_B)}{ka\sin\theta_B}. \tag{5.2.3}$$

However, for any practical feed, achieving the required uniform illumination of a reflector aperture that will yield such a pattern is not possible without a considerable amount of spillover loss. The directivity typically obtained is somewhat less than (5.2.2) and may be related to an *effective* aperture radius $a_{\text{eff}} \lesssim a$ through

$$D = (ka_{\text{eff}})^2 < D_{\max} \tag{5.2.4}$$

which corresponds to uniform illumination of an aperture with a radius smaller than a. Note that the effective aperture radius may be related to the physical aperture radius through

$$a_{\text{eff}} = sa \qquad \text{where } s \lesssim 1 \tag{5.2.5}$$

and the corresponding co-polarisation far-field function may be approximated as

$$F^{(\text{ref})}(\theta_B, \phi_B; s) = \frac{J_1(kas\sin\theta_B)}{kas\sin\theta_B}. \tag{5.2.6}$$

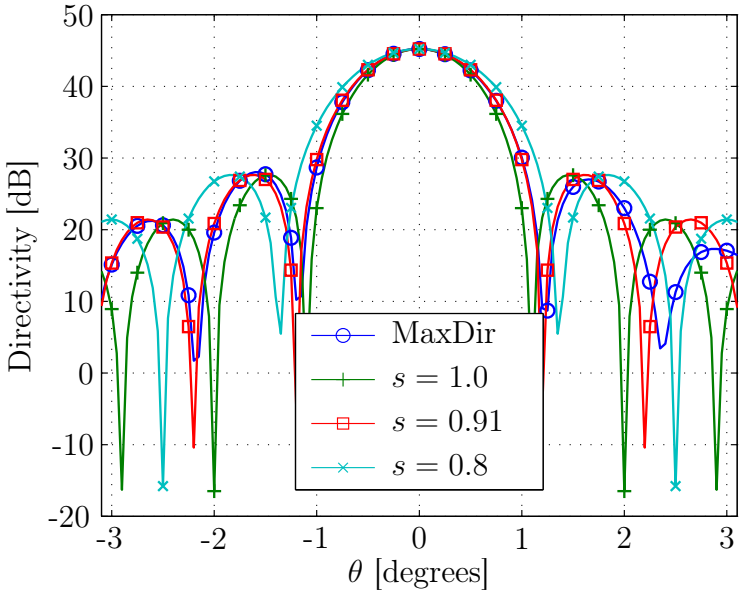
Including the scaling parameter s in the reference pattern results in a widening of the pattern in $\sin\theta_B$ space, and corresponds to the well-known relation between directivity and beamwidth.

The co-polarised component of two far-field patterns, as well as jinc-function approximations of these patterns using various values of s are shown in Figure 5.4. The patterns correspond to MaxSig beamformed patterns obtained with the MeerKAT optical design and a PAF, and steered towards two different directions. The jinc-function with $s = 1$ is seen to have a narrower main beam than either pattern and $s \approx 0.9$ is seen to yield a closer approximation. Note also that the best fit for the on-axis pattern has a slightly higher value of s than that for the off-axis pattern, and this is due to widening of the main beam as the antenna is scanned.

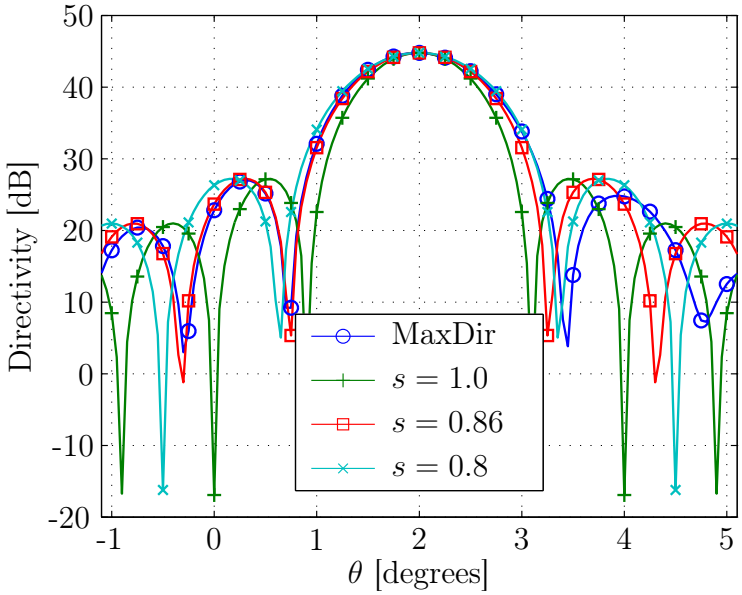
Since the function in (5.2.6) is real-valued it does not account for any phase variation of the radiation pattern, except for a sign inversion at the pattern nulls. In general the phase pattern is flat over a certain angular region if the phase reference of the pattern and the phase centre of the antenna coincide[3]. In this case a real-valued reference pattern may be accurate within a constant phase factor. However, if the phase centre and phase reference do not coincide a phase gradient may be present over the main beam of the radiation pattern and the reference pattern will only be accurate within a constant over a relatively small region.

The effect of non-coinciding phase reference and antenna phase centre on the phase pattern is illustrated in Figure 5.5 [87]. The phase reference and

[3] Here the term *phase centre* is used in the local sense, meaning that the requirements for the antenna to have a phase centre are limited to a certain angular region of interest. That is, the equiphase surfaces within a certain angular region of the radiation pattern reside on concentric spherical surfaces. The centre of these spherical surfaces is the phase centre of the antenna.



(a) On-axis ($\theta_s = 0°$, $\phi_s = 0°$) and shown in $\phi = 0°$ plane



(b) Off-axis ($\theta_s = 2°$, $\phi_s = 60°$) and shown in $\phi = 60°$ plane

Figure 5.4: Co-polarisation far-field amplitude patterns for the MeerKAT optical system with a PAF and beamforming for MaxSig patterns towards (θ_s, ϕ_s), and jinc-function approximations using various values of s.

phase centre are indicated as $\mathbf{p}_r$ and $\mathbf{p}_c$, respectively, and they are separated by a distance $\mathbf{d}$. The vectors $\mathbf{k}_1$ and $\mathbf{k}_2$ are the propagation vectors towards the far-field points $\mathbf{p}_1$ and $\mathbf{p}_2$, respectively, which are separated by an angular distance θ and reside on an equiphase surface referenced to $\mathbf{p}_r$. In the far-field approximation the vectors from $\mathbf{p}_c$ to $\mathbf{p}_1$ and from $\mathbf{p}_r$ to $\mathbf{p}_1$ are parallel, $\mathbf{k}_1^{(c)} \parallel$

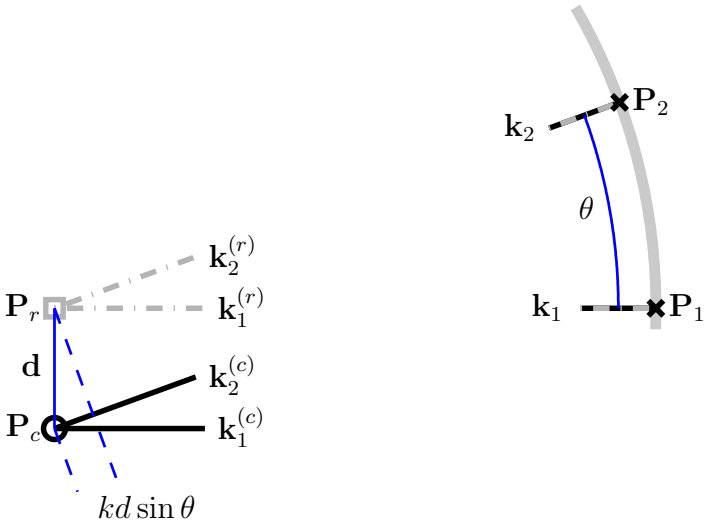


Figure 5.5: Effect of non-coinciding phase centre and phase reference on the far-field phase pattern.

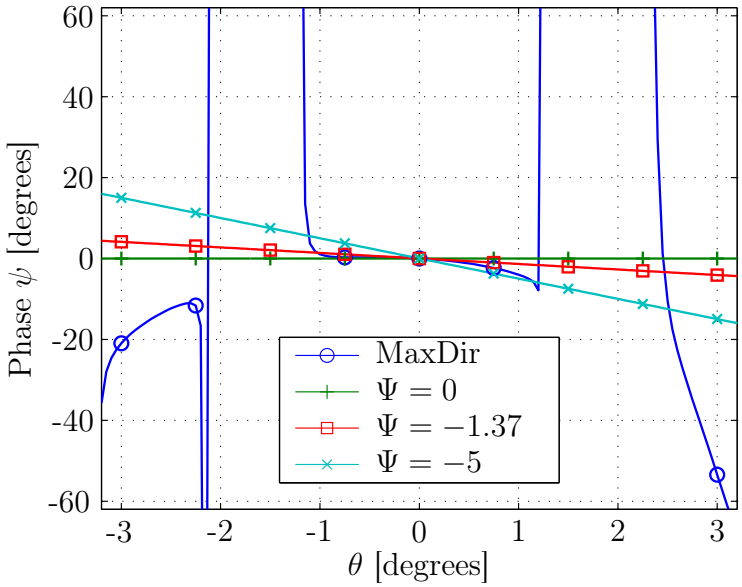
$\mathbf{k}_1^{(r)} \parallel \mathbf{k}_1$. Similarly we have $\mathbf{k}_2^{(c)} \parallel \mathbf{k}_2^{(r)} \parallel \mathbf{k}_2$. Assuming that $\mathbf{d}$ is orthogonal to $\mathbf{k}_1$ then spherical phase fronts originating from $\mathbf{p}_c$ and $\mathbf{p}_r$ will arrive at $\mathbf{p}_1$ in-phase. However, towards p_2 the relative phase difference between these phase fronts is $kd\sin\theta$. That is, for a point source at $\mathbf{p}_c$ the far-field pattern referenced at $\mathbf{p}_r$ will contain a phase variation as $\exp(jkd\sin\theta)$.

For the on-axis pattern of a reflector antenna the constraint of coinciding phase reference and phase centre is easily satisfied by choosing the phase reference at the centre of the projected aperture. However, for a scanned reflector the proper selection of a phase reference point is not as simple. In this case it is more convenient to keep the phase reference fixed at the centre of the projected aperture (as in Figure 2.3) and account for the shifting phase centre by including a phase variation in the reference pattern. Due to symmetry the shift of the phase centre is assumed to be in the scan plane $\phi = \phi_s$, and the reference pattern may be extended to account for this shift by adding a complex exponential factor as follows

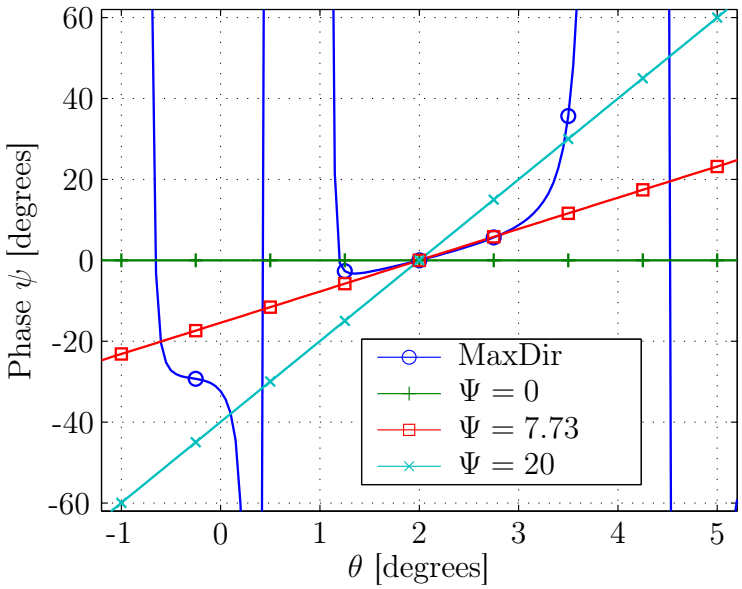
$$F^{(\text{ref})}(\theta_B, \phi_B; s, \Psi) = \frac{J_1\left(kas\sin\theta_B\right)}{kas\sin\theta_B} \exp\left(j\Psi\sin\theta_B\cos\phi_B\right). \tag{5.2.7}$$

Herein Ψ is a constant which is related to the product kd in Figure 5.5.

The phase patterns corresponding to the amplitude patterns in Figure 5.4 are shown in Figure 5.6, along with the phase of the reference pattern in (5.2.7) for various values of Ψ. For the on-axis pattern in Figure 5.6 (a) the phase is almost flat over the entire main beam, whereas for an off-axis pattern as in Figure 5.6 (b) the phase varies nearly linearly over the main beam.



(a) On-axis ($\theta_s = 0°$, $\phi_s = 0°$) and shown in $\phi = 0°$ plane



(b) Off-axis ($\theta_s = 2°$, $\phi_s = 60°$) and shown in $\phi = 60°$ plane

Figure 5.6: Co-polarisation far-field phase patterns for the MeerKAT optical system with a PAF and beamforming for MaxSig patterns towards (θ_s, ϕ_s), and complex exponential approximations using various values of Ψ.

The function in (5.2.7) is the desired form for the reference pattern and is used throughout to define the directional constraints for the LCMV beamformer. Evaluating (5.2.7) to obtain directional constraint values requires selecting the function parameters s and Ψ, as well as choosing the points (θ_B, ϕ_B) where the constraints are enforced. Selection of these parameters is discussed

in the following section.

5.2.2 Directional Constraints

As the secondary patterns for each feed array element have to be measured at each of the constraint positions in order to produce $\mathbf{C}$ in (5.1.9), the number of constraints that can practically be enforced is limited. This is especially true for PAF based systems in radio astronomy applications where such measurements need to be performed routinely during a long observation to compensate for electronic drift [80]. Note that the reflector antenna has to be steered mechanically[4] to place a calibration point source at each of the locations where constraints are enforced [88]. In addition to this each of the beamformer excitation weights represents a complex Degree of Freedom (DoF) which may be used to minimise the noise power in (5.1.14), and increasing the number of constraints reduces these DoFs. Therefore it is desired to limit the number of constraints to as few as possible while achieving the desired pattern shape. Herein a total of six directional constraints are enforced per LCMV beamformed pattern, in addition to the distortionless constraint.

Simultaneous design of the directional constraints and scanned beams over the FoV may reduce the number of required constraint measurements by enabling reuse of certain directions [34]. An example design is shown in Figure 5.7 where beams are positioned on a hexagonal grid such that the beam centres adjacent to a certain beam coincide with the directional constraints for that particular beam. The six directional constraints are positioned an angular distance θ_c away from the beam centre, and are arranged in a circularly symmetric fashion around the beam centre. In the arrangement shown a total of 37 measurements are required to realise 19 constrained beams; only constraint positions outside the FoV for the beams on the edge of the FoV require measurements additional to those required in the beam centres which are necessary in any case when employing a typically used MaxSNR or MaxSig beamformer.

The constraints vector $\mathbf{g}$ is formed by evaluating the reference pattern in (5.2.7) at each of the constraint positions for a particular beam, and in the reference frame in which $\theta_B = 0^\circ$ corresponds to the centre of that particular beam. The constraint vector elements are then

$$g_i = \begin{cases} F^{(\mathrm{ref})}(0, 0; s, \Psi) & i = 1 \\ F^{(\mathrm{ref})}(\theta_c, \phi_i; s, \Psi) & i = 2, 3, \ldots, 7. \end{cases} \tag{5.2.8}$$

Evaluating (5.2.8) requires selection of the reference pattern parameters s and Ψ. In the previous section it was shown how the reference pattern could

[4] This is required under the assumption used in this thesis that calibration of the patterns involves directly measuring each pattern in a (small) number of determined directions using a single celestial source. The use of so-called "in-beam" calibration techniques where a number of bright celestial calibration sources visible within the field of interest are used for such calibration is considered as a topic of future research in Chapter 6.

be fit to MaxSig beamformed patterns by choosing appropriate values for these parameters, and here the choice of parameter values is made using a similar approach. Attempting to reduce the sensitivity penalty for constraining the beamformed pattern shape the parameter values are chosen such that the reference pattern is fitted to a MaxSig beamformed pattern. That is, for each beam position over the FoV a MaxSig pattern is produced from which equivalent values for s and Ψ may be extracted and used to define directional constraints for an LCMV pattern at the same beam position in the FoV. Using the closed form relationship between pattern directivity and aperture size for a uniform circular aperture allows calculating an initial value for the beamwidth scaling parameter for each of the beams from

$$s_0 = \frac{a_{\text{eff,MaxSig}}}{a} = \frac{\lambda}{2\pi a}\sqrt{D_{\text{MaxSig}}} \tag{5.2.9a}$$

where D_{MaxSig} is the directivity of the MaxSig pattern. Similarly an initial value for the phase gradient parameter may be determined for each beam by direct computation from the corresponding MaxSig pattern

$$\Psi_0 = \left.\frac{\partial \psi_{\text{MaxSig}}}{\partial \theta_B}\right|_{\theta_B=0,\phi_B=0} \tag{5.2.9b}$$

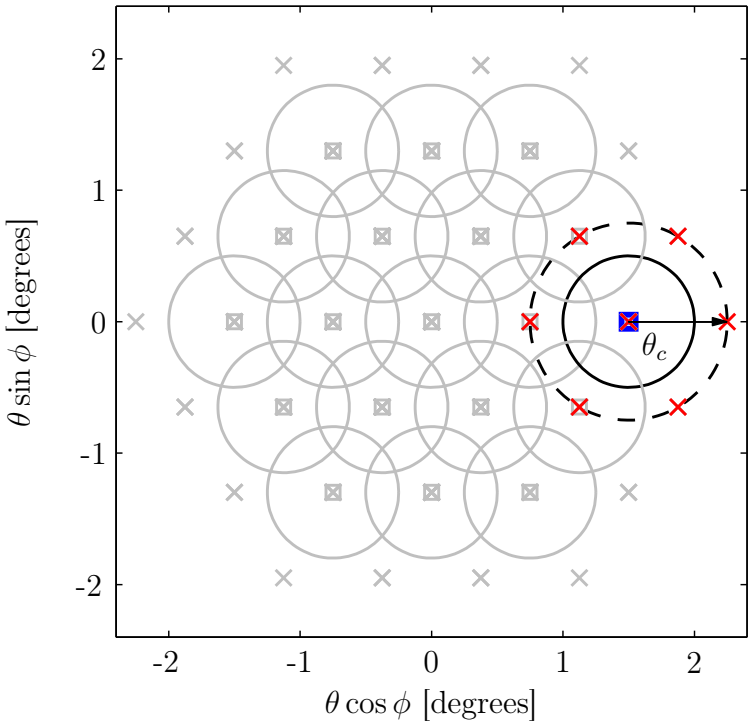


Figure 5.7: Layout of beams for measurement reuse. Nominal HPBW for each beam indicated in solid line, beam centres indicated as '□', and constraint positions indicated as '×'.

where ψ_{MaxSig} is the phase pattern for the MaxSig beamformer, and using the small angle approximation $\sin\theta_B \approx \theta_B$.

Taking full advantage of the flexibility offered by the parametrised reference pattern, the values obtained from the created MaxSig patterns need not be final, and the parameters may be further optimised to yield the desired performance.

Finally, it should be noted that these parameters are chosen to define the *reference pattern* that fixes the directional constraints; the same analytic function (i.e. using the same parameter values) is then used as a *pattern model* with which the realized beamformed pattern may be approximated.

5.3 Numerical Results

Numerical results for the above proposed beamforming strategy will now be presented. In this section the reflector telescope considered consists of the MeerKAT optical design, combined with a dual-polarised PAF based on the APERTIF array which consists of 121 Tapered Slot Antennas (TSAs) [48]. Although the beamformers herein are used to control only the co-polarised pattern, signals from elements of both polarisations in the feed array are used (as opposed to a bi-scalar beamformer in which only signals from one polarisation is used) [89; 90]. Beamforming is assumed to be performed off-line in software so that multiple beams can be produced with a single dataset [34].

For beam steering towards on-axis and relatively small scan angles off-axis, the PAF is centred on the secondary focus $\mathbf{P}_s$ of the elliptical subreflector, as indicated in Figure 5.8. Note the feed coordinates (x_f, y_f, z_f) in which $\hat{\mathbf{z}}_f$ is directed along the feed axis (indicated as L_s in Figure 2.1) and $\hat{\mathbf{x}}_f$ lies in the symmetry plane (xz-plane). In order to enlarge the FoV of the antenna the feed array was displaced by two element-widths (± 20 cm) in the directions $\phi_f = \{0°, 60°, \ldots, 300°\}$ depending on the direction of scan ϕ_s, and when scanning beyond a certain angular distance $\theta_s \geq \theta_M$ off-axis[5].

The results below are presented in two stages. Results of a parametric study to determine the impact of the reference pattern parameters and constraint positions on the beamformed pattern performance are presented first. Thereafter the performance of an LCMV beamformer designed using the results of the parametric study is compared to the performance of a MaxSig beamformer.

5.3.1 Parametric Study

For the parametric study a number of beam directions were defined over a FoV extending 3° off-axis. A MaxSig pattern was formed towards each beam direction and the initial reference pattern parameter values s_0 and Ψ_0 were

[5] This is simulates using a larger feed array.

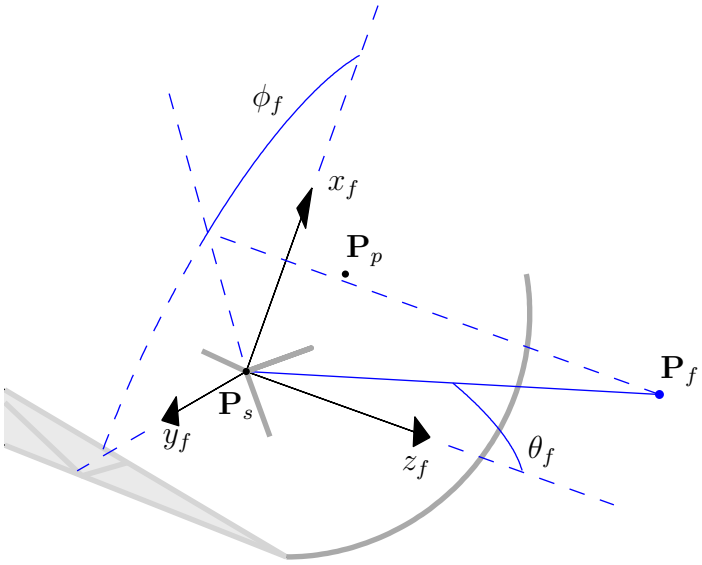


Figure 5.8: Feed coordinates.

calculated from (5.2.9). Next a number of LCMV patterns were formed for each beam direction using various values for s, Ψ and θ_c to determine the impact of these parameters on the resulting pattern at that beam direction. The figures of merit of interest were the directivity, first sidelobe level, and the average relative error ϵ_A calculated as in (2.2.12) when approximating the complex-valued co-polarisation voltage pattern as the reference pattern in (5.2.7), computed over all regions where the pattern is within 10 dB of the maximum.

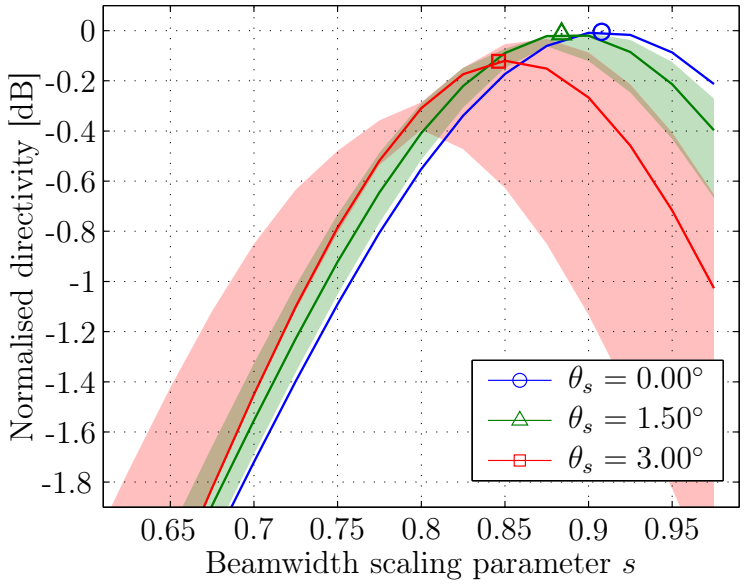
The parametric study was performed at the frequencies 1.20, 1.42, 1.60, and 1.75 GHz. In the following three sections Section 5.3.1.1 through Section 5.3.1.3 only results for a frequency of 1.42 GHz are presented, followed by a brief summary of all the results at other frequencies across the band in Section 5.3.1.4.

5.3.1.1 Beamwidth Scaling Parameter

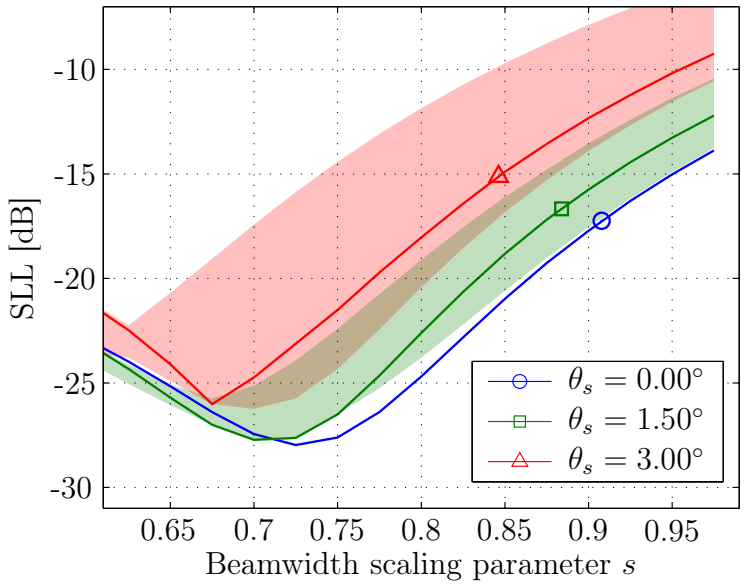
The effects of the beamwidth scaling parameter s on the pattern directivity and sidelobe level are shown in Figure 5.9. In this figure the solid lines indicate the results for the on-axis pattern, and two off-axis patterns scanned towards $(\theta_s = 1.5°, \phi_s = 120°)$ and $(\theta_s = 3.0°, \phi_s = 120°)$, respectively. Markers indicate the results for using the initial value $s = s_0$. The coloured regions around these lines indicate the variation of the results with azimuth (i.e. the minimum and maximum values obtained for all directions scanned towards the same θ_s off-axis). The directivities in Figure 5.9 (a) are normalised to that of the MaxSig beamformed patterns for the same scan directions. In all cases

$\Psi = \Psi_0$ and $\theta_c = 0.75°$ were used.

From these results the trade-off between directivity and sidelobe levels is recognised immediately; lowering the value of s widens the main beam by



(a) Directivity



(b) SLL

Figure 5.9: Impact of beamwidth scaling parameter s on LCMV pattern performance. Each line indicates the result for scanning in the plane $\phi_s = 120°$, and the shaded regions indicate the variation of that result over all values of ϕ_s.

tapering the reflector edge illumination, which generally results in a lower directivity and reduced sidelobe levels. As indicated by the proximity of the markers to the peaks of the curves in Figure 5.9 (a) the choice $s = s_0$ typically yields close to the optimum value for the beamwidth scaling parameter when directivity is to be maximised. If s is increased above s_0 the directivity is seen to decrease. This is attributed to an increased illumination of the reflector edge, which again results in higher spillover loss.

In Figure 5.9 (b) the sidelobe levels are seen to increase with scan angle, as expected due to coma aberration. For the on-axis pattern and $s = s_0$ the first sidelobe level is around the theoretical value of -17.6 dB for a uniform field distribution over a circular aperture. As the scan angle increases the sidelobe levels degrade rapidly and typical restrictions on sidelobe performance dictates using smaller values for s at wider scan angles. As was the case for spillover loss the increase in sidelobe levels is attributed to increased edge illumination of the reflector for $s \gtrsim s_0$. It is also observed that sidelobe levels cannot be reduced to arbitrarily low levels by reducing s and for $s \lesssim 0.75$ the sidelobe is seen to increase with decreasing s.

A number of LCMV beamformed on-axis primary and secondary patterns are shown in Figure 5.10 for various values of s. From the results in Figure 5.10 (a) it is clear that smaller values of s produce primary patterns that are tapered towards the reflector edge, whereas larger values of s produce flatter primary patterns, or even negative taper towards the reflector edge. Increased edge illumination for the MeerKAT optical system does not only result in a sharper transition in amplitude at the aperture edge, but also higher spillover and diffraction at the subreflector edges, which also affect the sidelobes of the total secondary pattern. The half-angle from the secondary focus towards the subreflector edge is approximately 48.9° and is indicated on the plot of the primary patterns.

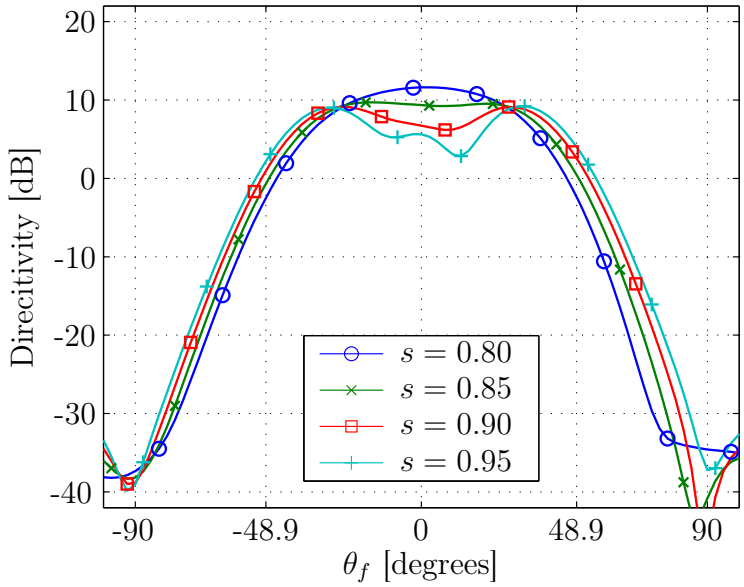
The corresponding secondary patterns are shown in Figure 5.10 (b) where the effect of s on beamwidth and sidelobe level is clearly visible.

For each LCMV beamformed pattern the error in approximating that pattern as the reference pattern was computed, using the same values of s and Ψ for both defining the directional constraints in (5.1.9), as well as computing the pattern modelling error. The average relative error over the region where the pattern is above -10 dB relative to maximum is shown as a function of s in Figure 5.11. Again markers indicate the results for $s = s_0$.

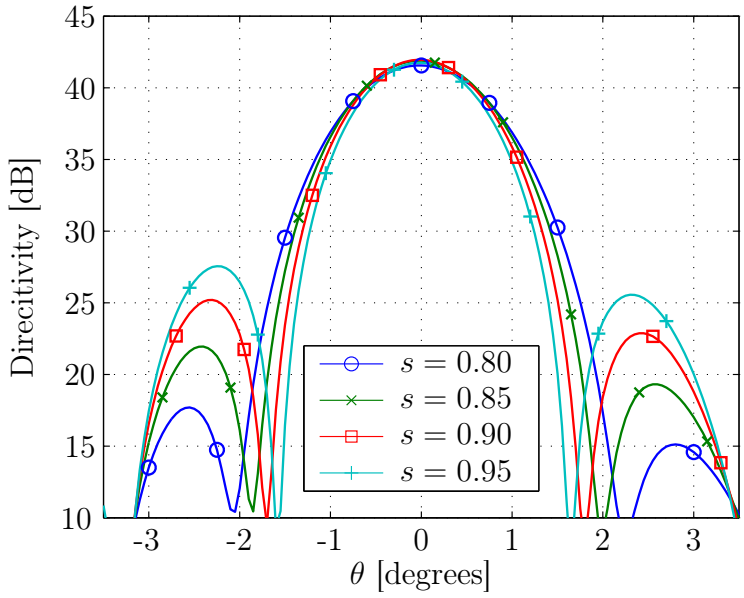
For all beam directions it is observed that the error is approximately at a minimum when using $s \approx s_0$. Furthermore, the error is slowly increasing as s is decreased, and remains below 1.5% for the on-axis patterns and below 5% for the off-axis patterns over a wide range of $s \lesssim s_0$. For wide scan directions the error is seen to increase significantly for $s > s_0$ due to the fact that the first sidelobe increases above -10 dB relative to the pattern maximum threshold (see Figure 5.9 (b)) and is therefore included in the error computation.

These results indicate that the value of s may be optimised over the region

$s \lesssim s_0$ for the desired trade-off between directivity and sidelobe levels without degrading the pattern model accuracy significantly. Note that the region $s > s_0$ is of little practical interest since there directivity is low and sidelobe levels are high.



(a) Primary patterns



(b) Secondary patterns

Figure 5.10: Primary and secondary patterns for various values of beamwidth scaling parameter s. Patterns are shown in the plane $\phi_f = 0°$.

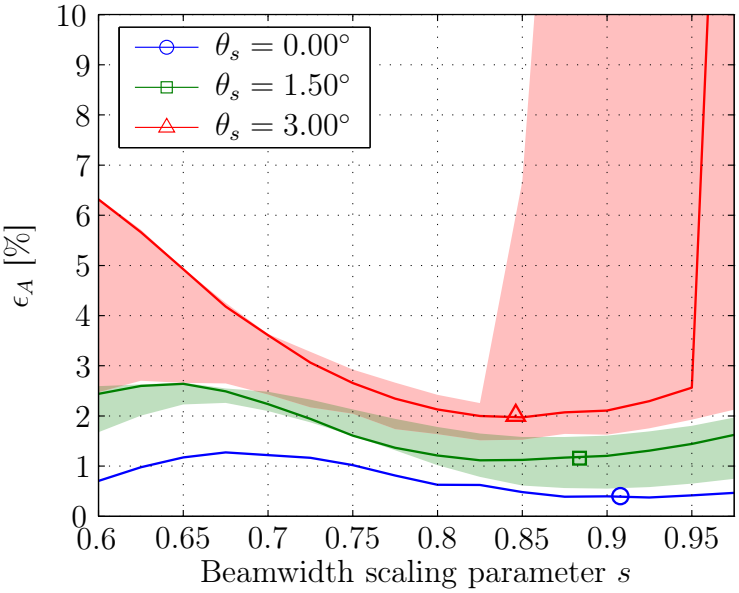


Figure 5.11: Pattern model error as a function of s. Each line indicates the result for scanning in the plane $\phi_s = 120°$, and the shaded regions indicate the variation of that result over all values of ϕ_s.

5.3.1.2 Phase Gradient Parameter

In order to appreciate the importance of using the reference pattern in (5.2.7) which includes the phase variation in the main beam region over the real-valued reference pattern in (5.2.6) the variation of Ψ_0 is shown as a function

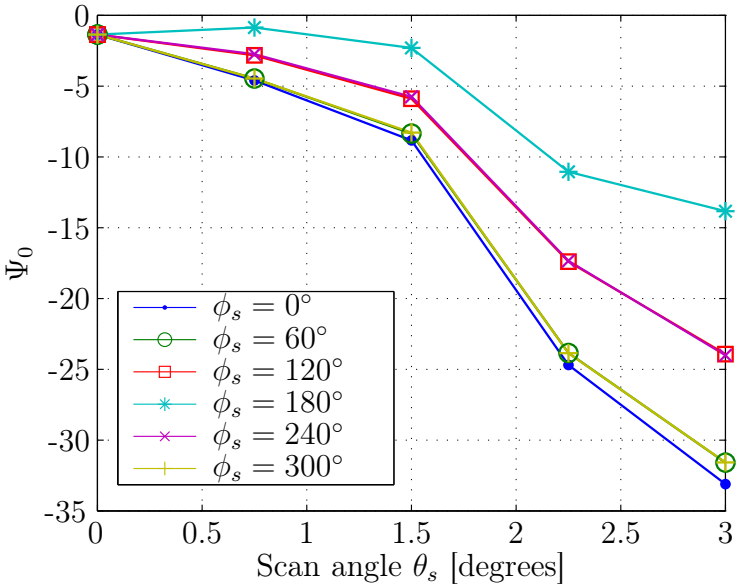
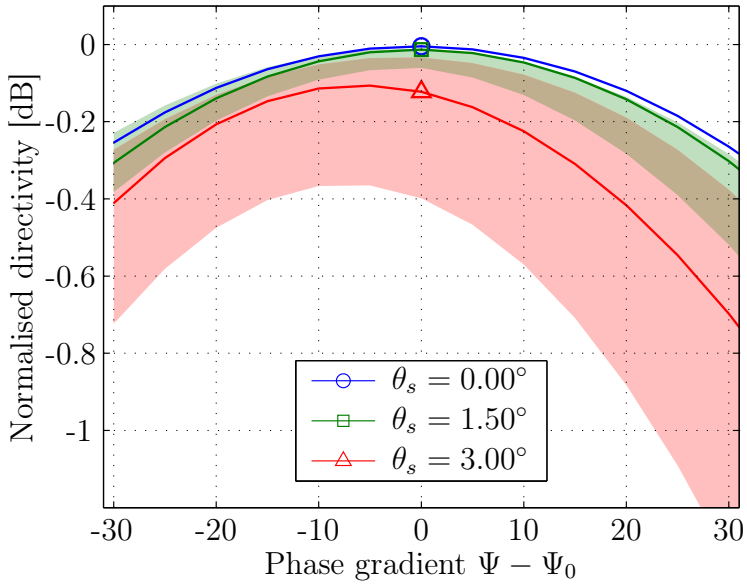


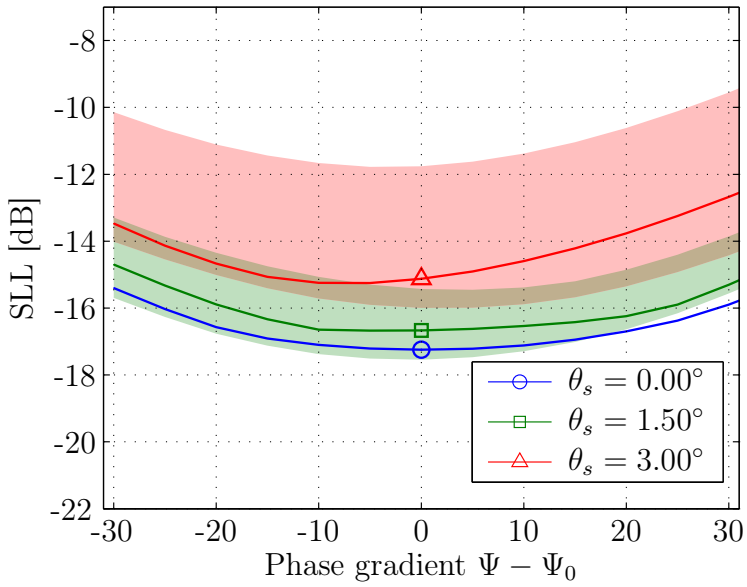
Figure 5.12: Variation of Ψ_0 with scan.

of scan angle θ_s and for various scan directions ϕ_s in Figure 5.12.

The value of Ψ_0 is seen to vary from nearly zero to between approximately -15 and -35, depending on the scan direction. Due to the asymmetry of the offset Gregorian geometry the results for scanning towards $\phi_s = 0°$ and towards



(a) Directivity



(b) SLL

Figure 5.13: Impact of phase gradient parameter Ψ on LCMV pattern performance. Each line indicates the result for scanning in the plane $\phi_s = 120°$, and the shaded regions indicate the variation of that result over all values of ϕ_s.

$\phi_s = 180°$ (in the offset plane) are different. As an example of the impact of this variation on the model accuracy, consider the relative error between the reference patterns $F^{(\text{ref})}(\theta_B, \phi_B; s)$ in (5.2.6) and $F^{(\text{ref})}(\theta_B, \phi_B; s, \Psi)$ in (5.2.7) evaluated at $\theta_B = 0.5°$ (approximate half-power point for the on-axis MaxSig pattern) with $\Psi = -20$ and $s = 0.9$. In the plane of scan the relative error is then approximately 17%

$$\begin{aligned}\text{Error} &= \left| \frac{F^{(\text{ref})}(0.5°, 0°; s = 0.9) - F^{(\text{ref})}(0.5°, 0°; s = 0.9, \Psi = -20)}{F^{(\text{ref})}(0.5°, 0°; s = 0.9, \Psi = -20)} \right| \\ &= |1 - \exp(-j20 \sin 0.5°)| = 0.174\end{aligned}$$

which is much higher than the modelling errors shown in e.g. Figure 5.11.

The effects of the phase gradient parameter Ψ on the directivity and sidelobe level are shown in Figure 5.13. Markers are used to indicate $\Psi = \Psi_0$, in all cases $s = s_0$ and $\theta_c = 0.75°$, and the directivity in Figure 5.13 (a) is normalised to that of the MaxSig patterns for each of the scan directions. Note that the abscissae indicate the difference $\Psi - \Psi_0$.

For both the directivity and sidelobe level it is concluded that the best performance is typically achieved when $\Psi = \Psi_0$. As Ψ is increased or decreased a gradual degradation in performance is observed, resulting in a decrease in directivity and an increase in the first sidelobe level. This confirms the importance of using the reference pattern in (5.2.7) over that in (5.2.6); the latter is equivalent to using $\Psi = 0$ and the corresponding pattern performance can be determined by noting the value of Ψ_0 for a particular scan direction in

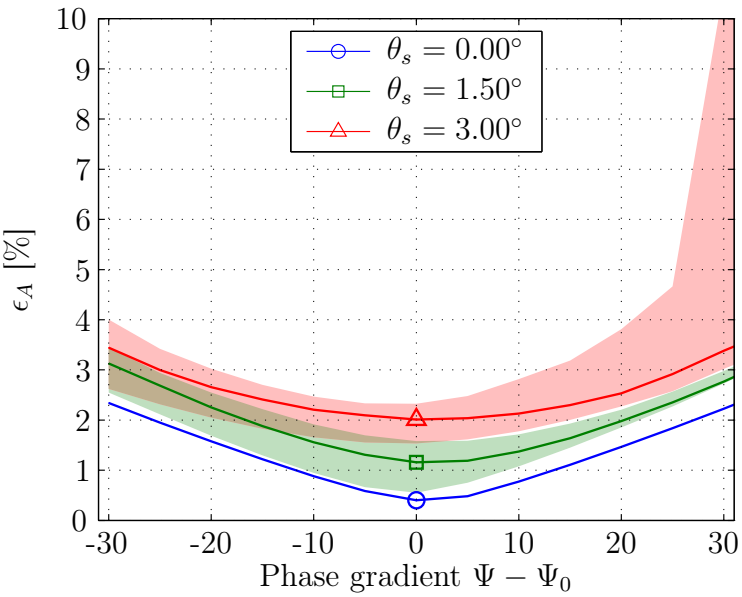


Figure 5.14: Pattern model error as a function of Ψ. Each line indicates the result for scanning in the plane $\phi_s = 120°$, and the shaded regions indicate the variation of that result over all values of ϕ_s.

Figure 5.12. For example, scanning towards ($\theta_s = 3°, \phi_s = 120°$) we have $\Psi_0 \gtrsim -25$, for which using $\Psi = 0$ corresponds to a loss in directivity of 0.6 dB and a sidelobe level increase of 2 dB compared to using the optimal value $\Psi = \Psi_0$.

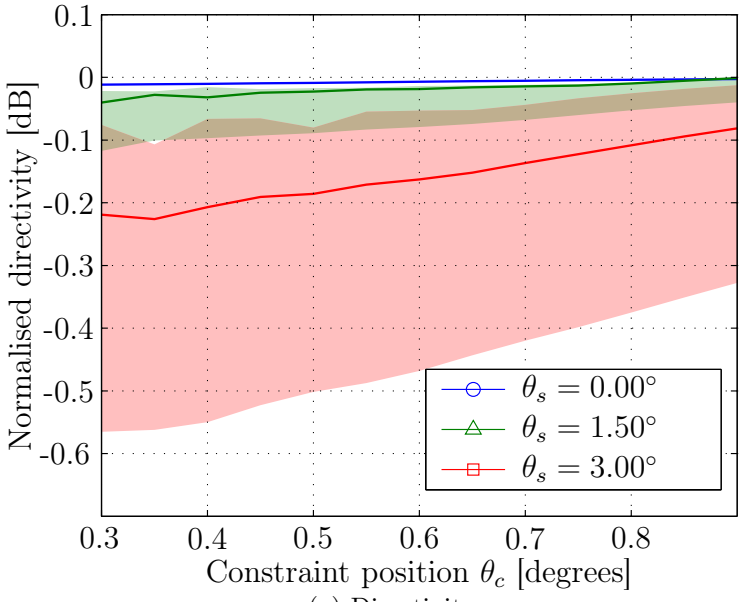
The pattern model error is shown in Figure 5.14 as a function of Ψ, where the best performance is obtained once again by choosing $\Psi = \Psi_0$. Similar to the increase in error that was observed when increasing $s > s_0$ here a significant increase in error also occurs when the value of Ψ is varied as much as is necessary to increase the first sidelobe above the -10 dB threshold.

5.3.1.3 Constraint Positions

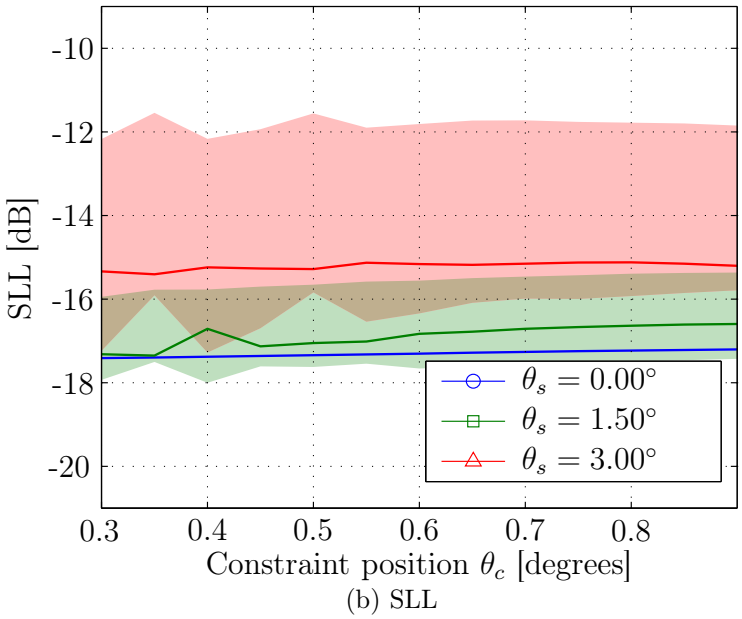
The effect on the various figures of merit of the angular distance θ_c from the beam centre to where directional constraints are enforced was investigated. Herein the reference pattern parameters used were $s = s_0$ and $\Psi = \Psi_0$.

Figure 5.15 (a) and (b) show how the directivity and sidelobe performance, respectively, of the beamformed patterns are affected by the position of directional constraints. For the on-axis and narrow scanned beams the directivity is weakly dependent on the constraint positions, and increases slightly as the constraints are enforced further from the beam centre; at wider scan angles this effect is somewhat more pronounced. On the other hand, the impact on the sidelobe performance is much weaker and the first sidelobe levels for all scan directions remain nearly constant as the constraint positions are varied.

In Figure 5.16 the pattern model error is shown as a function of θ_c. The results indicate that the average relative model error is generally minimised by placing constraints approximately between 0.7° and 0.8° from the beam centre, dependent on the scan direction. Using the marker positions in the figure, this constraint position range is seen to correspond to around the -6 dB or -7 dB beamwidth. This result can be expected since the realised pattern matches the reference pattern exactly at the positions where directional constraints are enforced, and may diverge from the reference pattern further from these positions. Enforcing constraints too close to the beam centre then results in a larger error at the edges of the beam, whereas constraints too far from the beam centre results in a larger error near the beam centre. As the error computation is limited within the 10 dB beamwidth (assuming that no sidelobes are above the -10 dB threshold), this error is at a minimum roughly where the constraints are enforced near the middle of the region of interest.



(a) Directivity



(b) SLL

Figure 5.15: Impact of constraint positions θ_c on LCMV pattern performance. Each line indicates the result for scanning in the plane $\phi_s = 120°$, and the shaded regions indicate the variation of that result over all values of ϕ_s.

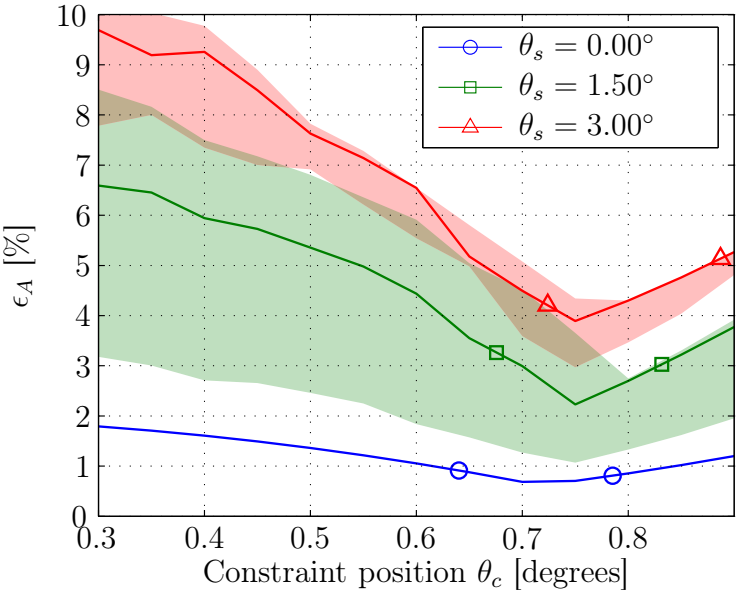


Figure 5.16: Effect of constraint position θ_c on pattern model error. Markers indicate roughly the -5 dB (left) and -8 dB (right) beamwidths for the realised LCMV patterns. Each line indicates the result for scanning in the plane $\phi_s = 120°$, and the shaded regions indicate the variation of that result over all values of ϕ_s.

5.3.1.4 Frequency Dependence

Results similar to those presented above were obtained at various frequencies within the band 1.20 GHz to 1.75 GHz. Where necessary frequency dependent quantities were simply scaled to values equivalent to that used above at 1.42 GHz, e.g. the scan angle for a particular beam at each frequency f_i is simply

$$\theta_s|_{f_i} = \theta_s|_{f_0} \times \frac{f_0}{f_i}$$

where $f_0 = 1.42$ GHz. Similarly the directional constraint positions in all cases were set to the equivalent of $\theta_c = 0.75°$ at 1.42 GHz.

Figure 5.17 shows the performance of LCMV beamformed patterns using $s = s_0$ and $\Psi = \Psi_0$ as a function of scan angle (in the direction of $\phi_s = 120°$) and at various frequencies. Across the frequency band the results are similar both in terms of directivity and first sidelobe level, with the exception of the widest scan angle at the lowest frequency where a drastic reduction in directivity is observed. This is due to the fact that the focal region for such a wide scan angle is beyond the extents of the feed array.

For the remainder of this section results are presented for the second off-axis beam direction in Figure 5.17, that is $\phi_s = 120°$ and θ_s ranges from 1.77° to 1.22° from the lowest to the highest frequency. The variation of results over

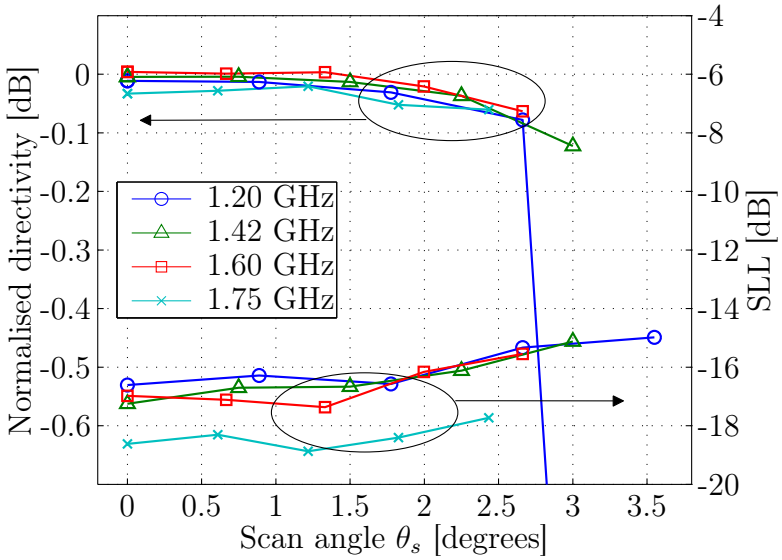
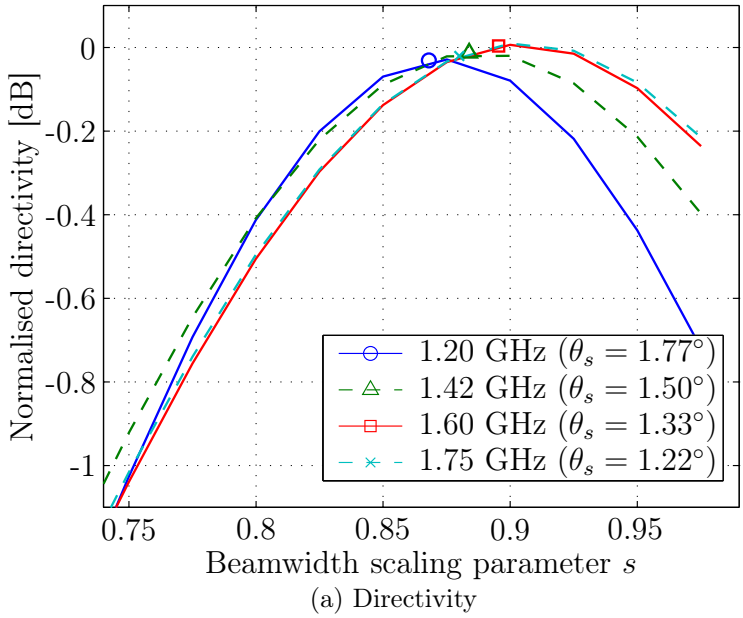


Figure 5.17: Directivity and first sidelobe level for LCMV patterns using s_0 and Ψ_0 at various frequencies.

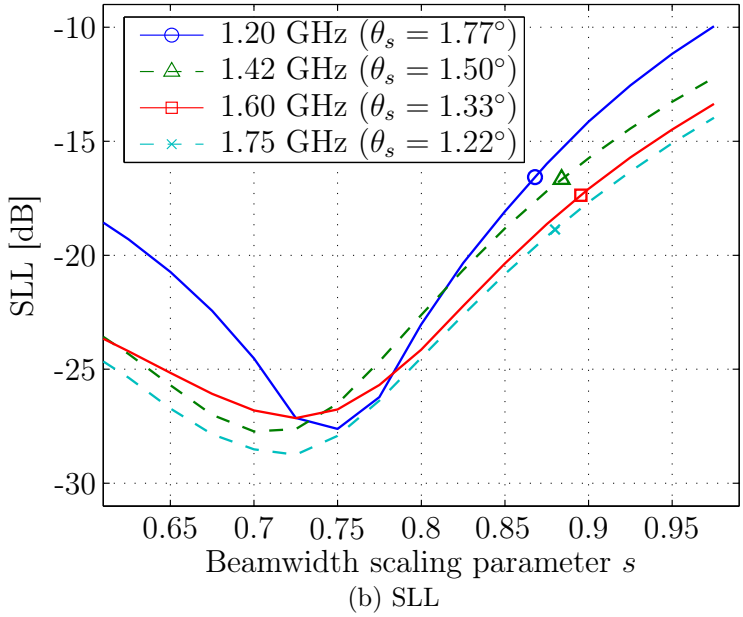
the various beam directions follows similar traits as that shown in previous sections and are not presented here.

The effect of s on the directivity and first sidelobe level of the LCMV beamformed patterns is shown in Figure 5.18 (a) and (b), respectively, for various frequencies. Note that by definition the reference pattern in (5.2.7) scales with frequency due to the factor $k = \frac{2\pi}{\lambda}$ in the argument of the jinc-function. It is therefore expected that the value of s_0 should remain relatively constant for a particular beam over frequency, as is seen in the figure. Once again the trade-off between beamwidth and sidelobe level is observed up until $s \approx s_0$ where directivity is nearly at a maximum, and beyond which directivity rapidly decreases.

Figure 5.19 shows how the value of Ψ impacts on the pattern figures of metric across the frequency band. Similar to previous results, using $\Psi = \Psi_0$ is seen to yield near-optimal directivity and first sidelobe performance. In Figure 5.19 (a) it is observed that the directivity is affected more severely by a certain variation in Ψ at lower frequencies than at higher frequencies; the results in Figure 5.19 (b) for the sidelobe level exhibit similar behaviour. Recall from Figure 5.5 that the parameter Ψ is related to a shift in the phase centre of the antenna relative to the phase reference, and that the difference in phase observed at two far-field points is proportional to the product kd, which is inherently frequency dependent. Therefore it is anticipated that the value of Ψ for a particular scan direction increases with frequency, so that at higher frequencies the same *absolute* change in this parameter corresponds to a smaller *relative* change than at lower frequencies. Consequently the impact of variation in the value Ψ on the pattern performance is more pronounced at



(a) Directivity

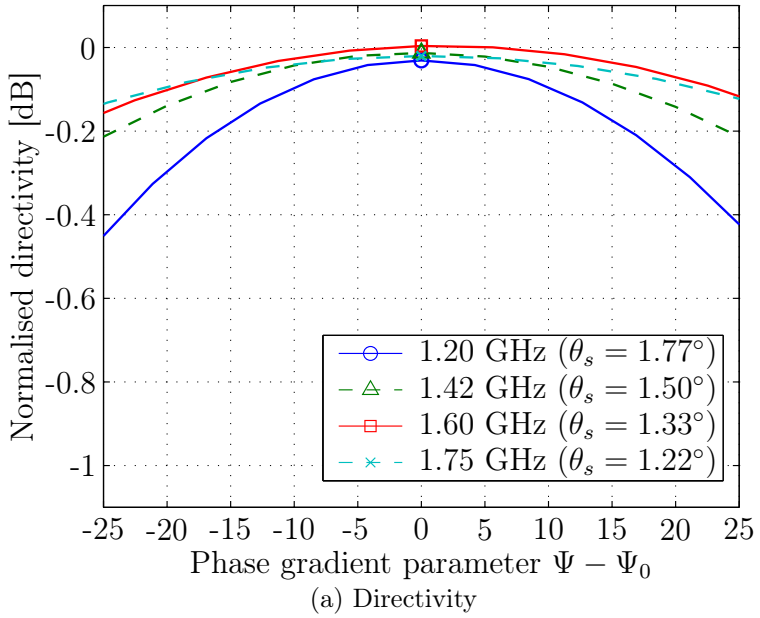


(b) SLL

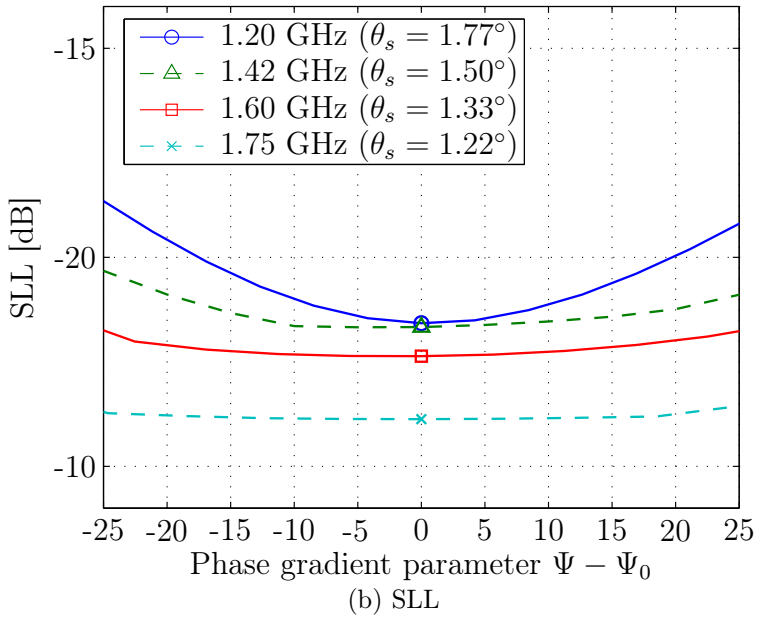
Figure 5.18: Impact of s on directivity and sidelobe performance. Markers indicate the values $s = s_0$ and in all cases $\Psi = \Psi_0$.

lower frequencies.

The pattern model error is shown as a function of s and Ψ in Figure 5.20 (a) and (b), respectively. With the exception of a relatively high error for small values of s at the lowest frequency, the error behaviour is very similar at the various frequencies. The large error for $s < 0.75$ at 1.20 GHz is due to asymmetry in the main beam, and as before the rapid increase in error for



(a) Directivity

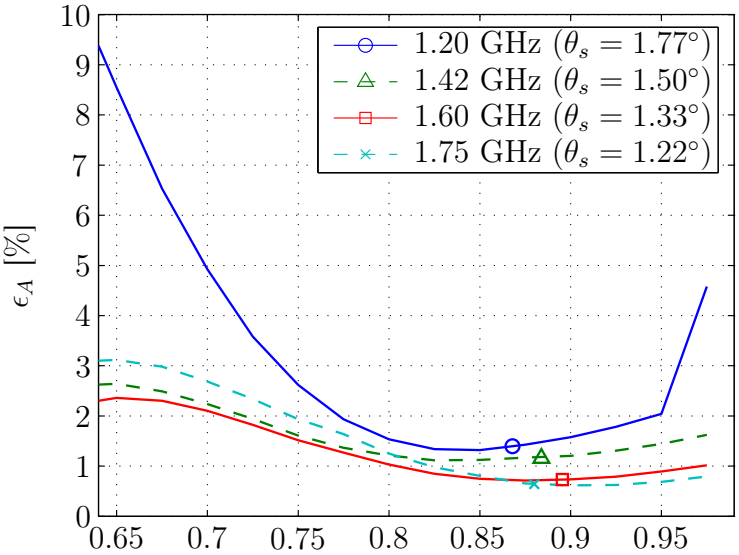


(b) SLL

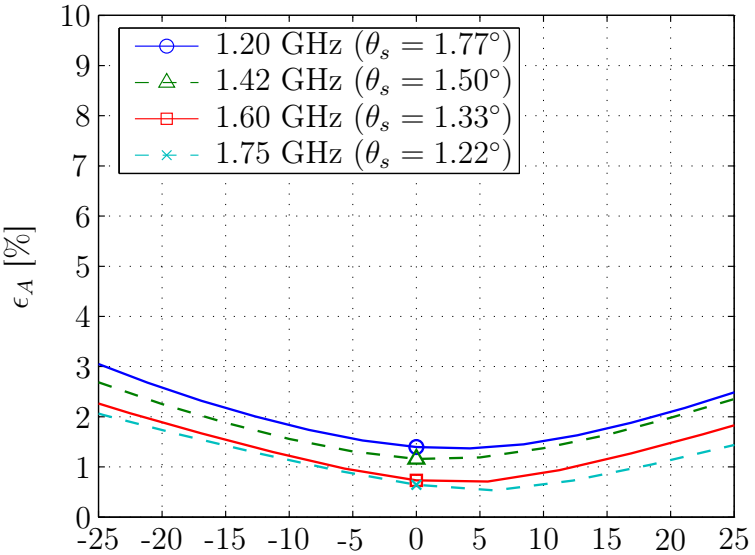
Figure 5.19: Impact of Ψ on directivity and sidelobe performance. Markers indicate values $\Psi = \Psi_0$ and in all cases $s = s_0$.

$s > s_0$ is due to the increase of the first sidelobe above the -10 dB threshold. Nevertheless, even at this lowest frequency s may be decreased to about 0.75 without incurring an error above 3%.

The results in this section indicate that the reference pattern and the LCMV beamforming strategy developed here may be used over a range of frequencies across the band from 1.20 GHz to 1.75 GHz. For the remainder of



(a) s



(b) Ψ

Figure 5.20: Impact of s and Ψ on the pattern model error.

this chapter results will only be presented for the single frequency 1.42 GHz.

5.3.2 Beamformer Design

Based on the results of the parametric study presented above the following constrained beamformer design procedure is derived:

1. For each beam direction a MaxSig pattern is realised and used to calculate initial model parameters s_0 and Ψ_0 from (5.2.9).

2. The reference pattern parameters are then adjusted to yield the desired directivity / sidelobe level trade-off

   a) The beamwidth scaling parameter s may be varied over a range $s \lesssim s_0$.

   b) The phase gradient parameter is fixed at the initial value $\Psi = \Psi_0$.

   c) Directional constraint positions are selected to conform the realised pattern to the reference pattern at about the -7 dB level.

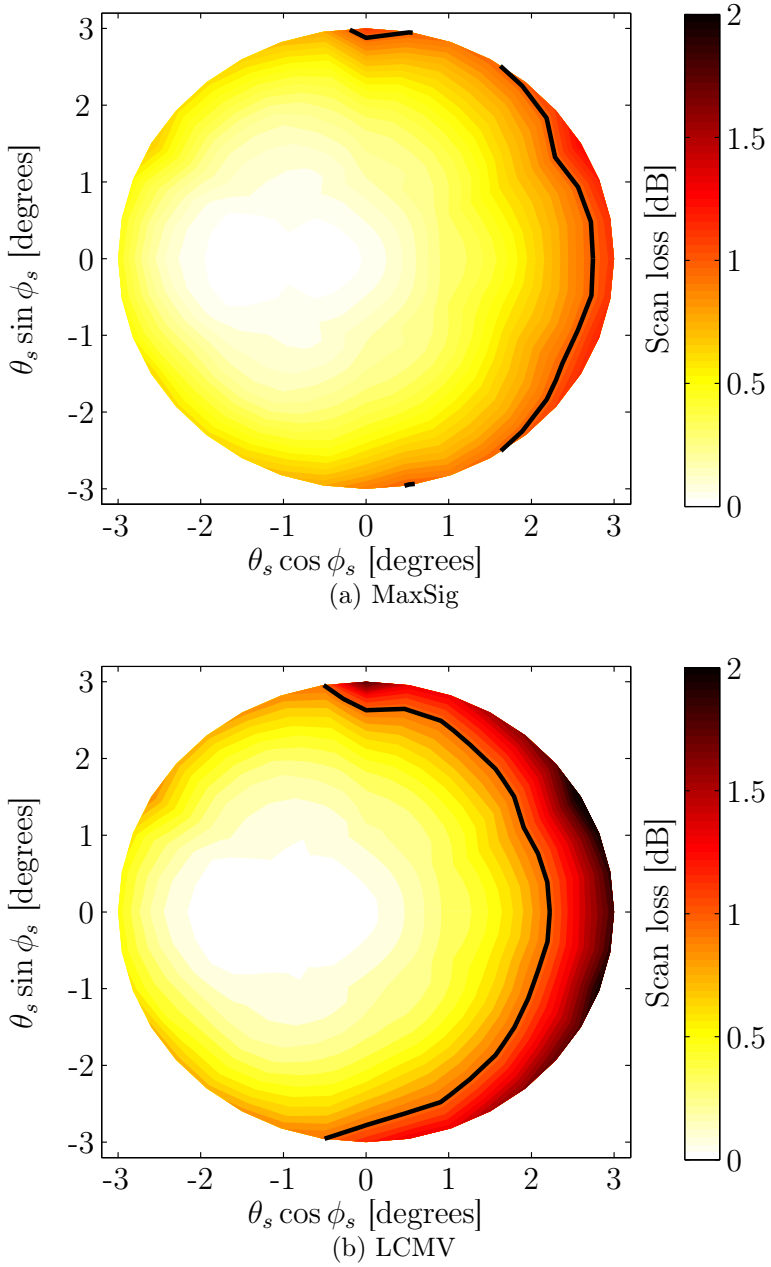
3. For each LCMV pattern realised using a set of parameters (s, Ψ), a pattern model is constructed using that *same* set of parameters.

Using the above design procedure an LCMV beamformer was implemented to produce beams over a FoV extending 3° off-axis. The performance of this beamformer was compared to that of a MaxSig beamformer which was also implemented to produce beams over the same FoV. The comparison was based on different FoMs, which were computed as functions of scan direction towards 592 different directions within the FoV. These FoMs used were the pattern model accuracy and scan loss over the FoV

The model accuracy is determined by computing the average relative error ϵ_A as in (2.2.12) between the model $\tilde{F}$ and the actual pattern F, where for the LCMV beamformer $\tilde{F} = F^{(\text{ref})}(\theta_B, \phi_B; s, \Psi_0)$, and for the MaxSig beamfomer $\tilde{F} = F^{(\text{ref})}(\theta_B, \phi_B; s_0, \Psi_0)$. The value of s in the LCMV beamformer reference pattern for a particular scan direction is chosen to yield an SLL within the range $(-0.2, 0)$ dB relative to the SLL of the MaxSig beamformed pattern towards that same direction. The scan loss for a particular scan direction is computed as the amount by which the directivity of the on-axis pattern of the beamformer is higher than the directivity of the pattern produced by the same beamformer for that scan direction.

The scan loss for each of the beamformers is shown over the FoV in Figure 5.21, and a loss of 1 dB is indicated with a solid black line. As expected the MaxSig beamformer exhibits a smaller loss over the FoV, and the scan loss is below 1 dB over a region of 26.98 sq. deg. In comparison the corresponding region for the LCMV beamformer is 23.15 sq. deg, and is about 14% smaller than that for the MaxSig beamformer.

For each of the beams over the FoV the average relative error in the pattern model was calculated over the region within 10 dB relative to the beam maximum. This error is shown for the MaxSig and LCMV beamformers in Figure 5.22 where the improvement in pattern model accuracy through using constrained beamforming is clearly seen. Patterns created through LCMV beamforming can be approximated with the single-term model with an error



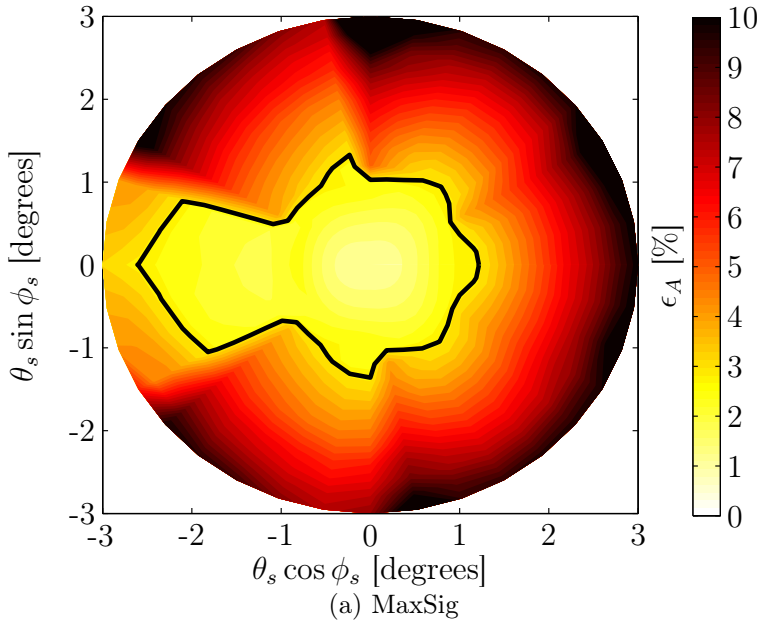
(a) MaxSig

(b) LCMV

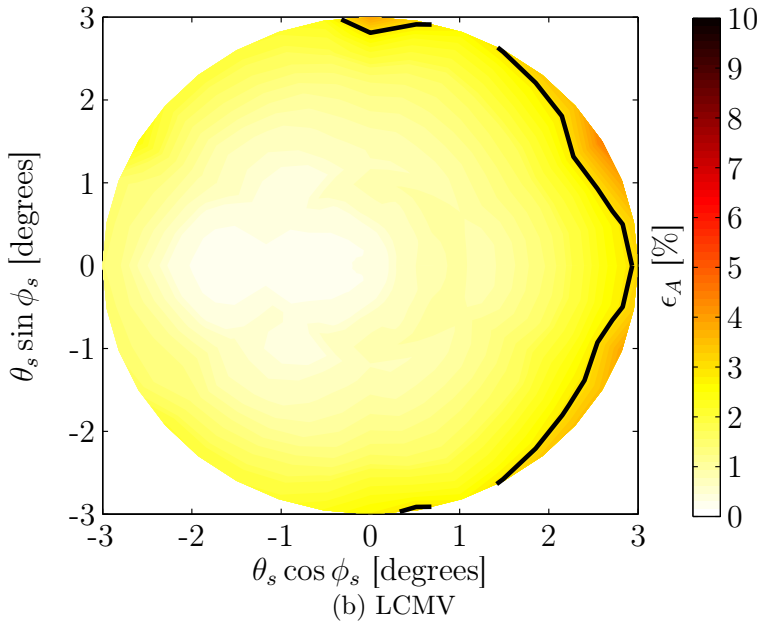
Figure 5.21: Scan loss over the FoV for each beamformer. Solid lines indicate a scan loss of 1 dB.

of less than 3% over almost the entire FoV. On the other hand, patterns created with the MaxSig beamformer are approximated with a similar degree of accuracy over a much smaller region which extends between roughly 1° and 2.5° off-axis, depending on the scan direction.

The dominant factor contributing to the large model error for MaxSig patterns is that the off-axis patterns become highly asymmetric, as is usually the



(a) MaxSig



(b) LCMV

Figure 5.22: Pattern model error for each beam over the FoV. Solid lines indicate an error of 3%.

case for reflector antennas [91; 92]. For each of the beamformed patterns over the FoV the half-power contour was determined and the aspect ratio of the contour was calculated. This ratio was found to be as high as 1.15:1 for wide scan directions and using the MaxSig beamformer. In contrast, the aspect ratio for the half-power contours of the LCMV patterns was found to be less than 1.01:1 over the entire FoV, and consequently these patterns may result

in somewhat simpler direction-dependent calibration.

5.4 Conclusion

In this chapter a novel beamforming strategy was proposed that can be used to conform the co-polarised component of the far-field pattern of a PAF based radio telescope to a known analytic reference pattern. Through the use of an LCVM beamformer and an appropriate physics-based reference pattern to define directional constraints, it was shown how beam patterns could be realised over a wide FoV that could be accurately modelled by this reference pattern. Combined design of the beam positions over the FoV and directional constraints positions for each beam was also shown to reduce the number of calibration measurements that would be required to perform constrained beamforming.

The reference pattern contains two parameters and the effects of these parameters on various performance figures of merit were studied. It was also shown how these parameters could be optimised for a particular beam direction in the FoV by first deriving initial values from a MaxSig pattern, and then adjusting the beamwidth scaling parameter to yield the desired trade-off between directivity and sidelobe level. Repeating the study at a number of frequencies, it was shown that the results exhibited a relatively weak dependence on frequency due to the inherent frequency scaling of the reference pattern.

Finally, a constrained beamformer design strategy was derived from the results of the parametric study. This strategy was then used to design an LCMV beamformer to produce patterns over a wide FoV, and the performance of the realised patterns compared to that for a MaxSig beamformer over the same FoV. The LCMV beamformer was shown to yield circularly symmetric beam patterns that could be approximated more accurately at the cost of a higher scan loss for wide off-axis beams. Thus a possible trade-off between calibration efficiency and sensitivity was shown to exist.

Chapter 6

Conclusion

The main focus of this work was the development of efficient pattern models — meaning high accuracy and few unknown parameters — for the direction-dependent gain calibration of reflector antenna radio telescopes.

In Chapter 2 the necessity for an accurate description of the radiation patterns of the antennas in an interferometer array was motivated. It was demonstrated that a *static* pattern model, such as that which would be obtained through measurement or simulation in a particular instance, is insufficient as the radiation pattern variation that can be expected under typical operating conditions may render that model inaccurate to the point of limiting the imaging performance of the instrument. Thus the need for a *dynamic* pattern model was stated, which can adapt to changes in the actual radiation pattern at the time of observation through solution of a few model parameters. The rest of this thesis was devoted to the development and application of such pattern models.

Chapter 3 cast the long-existing Jacobi-Bessel pattern model [27; 28] in a new light by solving for its expansion coefficients directly in the secondary pattern domain, as would be the case when using the model in a calibration algorithm. Direct solution through sparse sampling of the radiation pattern resulted in ill-conditioning so that a constrained solution, based on the assumption that the actual pattern at the time of observation is relatively close to a prior determined pattern, proved to yield higher accuracy models, along with the benefit of requiring less pattern samples to obtain a unique solution. Alternatively, the newly developed Neumann pattern model was demonstrated to yield higher accuracy through direct solution, although requiring somewhat more unknowns to be solved.

The use of the recently proposed Characteristic Basis Function Patterns (CBFPs) [30] was extended in Chapter 4 to compensate for the non-linear pattern variations that result from the mechanical deformations in a dual-reflector antenna. Here the excellent pattern prediction capability of these numerical basis functions was demonstrated, and although results were shown for simulation only, a measurement strategy for developing CBFP models for

an actual system was proposed.

Finally, in Chapter 5 the focus turned towards Phased Array Feed (PAF) based antennas, where the flexibility offered through beamforming was utilised to improve the calibration efficiency. Here a two-stage strategy was proposed which first fits a physics-based analytic model to beam patterns optimised for maximum directivity, and then constrains the beamformer to realise patterns that conform to that model. Through a performance comparison over a wide Field of View (FoV) of a constrained beamformer employing this strategy to a maximum directivity beamformer, a trade-off between directivity and calibration model accuracy was demonstrated.

6.1 Novelty and Impact of this Work

The work herein dealt specifically with the calibration of reflector antenna radio telescopes, and as such the results will be of interest mainly to the radio astronomy community, although it may also prove useful in other areas where the calibration of antenna radiation patterns is required.

The most significant contribution of this work is that it presents a number of pattern modelling techniques to the calibration specialist in the radio interferometry field. Drawing knowledge from antenna engineering, the emphasis was on using physics-based basis functions which improves the efficiency of the various models, while through the use of the Radio Interferometer Measurement Equation the appropriate context was created for the presentation of these models.

Additional to this, the application of the different pattern modelling techniques to the MeerKAT antenna through simulation served as an indication of the level of accuracy that could be achieved with each of these methods when used in the calibration algorithm. Specifically, the superior modelling ability of numerical basis functions, such as the CBFPs which encompass most of the pattern features in a single term, over that of analytic basis functions was demonstrated.

In the arena of the emerging PAF based technology where calibration of the radiation patterns remains a difficult challenge, a general beamforming strategy was devised through which a compromise between sensitivity and calibration efficiency may be obtained.

The work herein resulted in two journal publications [32; 37], two international conference publications [36; 56], and participation in international workshops focussed on addressing future challenges in the calibration of radio interferometers [29; 93; 94].

6.2 Topics for Future Research

Extension of this work in various directions is possible, and a few proposals for future research on this topic follows.

In general concerning the various models, there are a few points that still need to be addressed. One such aspect is the far sidelobe modelling, as the noise contribution from this region may be a performance limiting factor on smaller dishes [16]. Furthermore, modelling the frequency dependence of the radiation pattern is also of importance, especially frequency ripple effects as observed in prime focus systems [58], as well as in clear aperture systems [95].

Throughout this thesis it was assumed that a number of pattern measurements in predetermined directions are performed on a single celestial radio source in order to solve for the various model parameters. The impact of noise in such measurements on the resulting model accuracy needs to be considered in evaluating the performance of a particular modelling approach [96]. Furthermore, the use of the different pattern models in so-called "in-beam" calibration methods, where the pattern model parameters are solved using a number of celestial calibrator sources within a particular FoV of interest, also presents an important consideration. Such calibration offers the advantage of not having to reposition the antenna to perform a number of pattern measurements. On the other hand, in such a scenario we no longer have control over the positions in the pattern where such measured data is available, and depending on the brightness of each calibrator source within the FoV, the SNR associated with each point may vary [59].

Another general aspect concerning all beam modelling approaches which deserves attention is the derivation of beam model accuracy requirements based on the desired quality of the final image. Although this has been done to some extent for AA beams [97], as well as for dish arrays [23], the impact of the beam modelling error on the imaging performance is not yet that well understood, and quoting a specific accuracy requirement is in many cases still difficult.

In terms of analytic pattern models, herein only two related forms of basis functions were considered, and many other analytic expansions exist which may be explored. As suggested in [24], shapelets may prove useful due to their versatility and simplicity. Furthermore, these functions are already used within the astronomy community for image analysis [98; 99].

The next important step for CBFPs would be to demonstrate the method for an actual system, in order to determine how effective this method is at compensation for pattern variation in practice. The expansion of each CBFP into a fixed sum of analytical pattern functions may also be useful as a means to interpolate between discrete points sampled when constructing the basis function, as noted in [30], and the optimal interpolation scheme in [70] may be useful in this regard.

For the constrained beamforming strategy the inclusion of realistic internal and external noise is essential for a full evaluation its performance, as this will

allow the impact on the antenna *sensitivity* to be established, which is an important figure of merit in radio astronomy applications. Another aspect of this technique which may be explored further is the use of asymmetric pattern functions. Such models may provide a more accurate characterisation of off-axis unconstrained beamformed patterns, as well as improve the performance of beamformed patterns that are constrained to conform to these functions.

Finally, perhaps the most important next step is the integration of the presented pattern models in the calibration and imaging pipeline of existing radio interferometry tool sets, such as the MeqTrees software [18]. This will serve to test the capabilities of the various models within their intended application, and enable a more complete assessment of their performance.

Appendices

Appendix A

Aperture Field to Far-Field Transformation

The aperture field to far-field transformation used to derive the analytic pattern model in Section 3.1.1 is derived here following the derivation in [63, § 12.6].

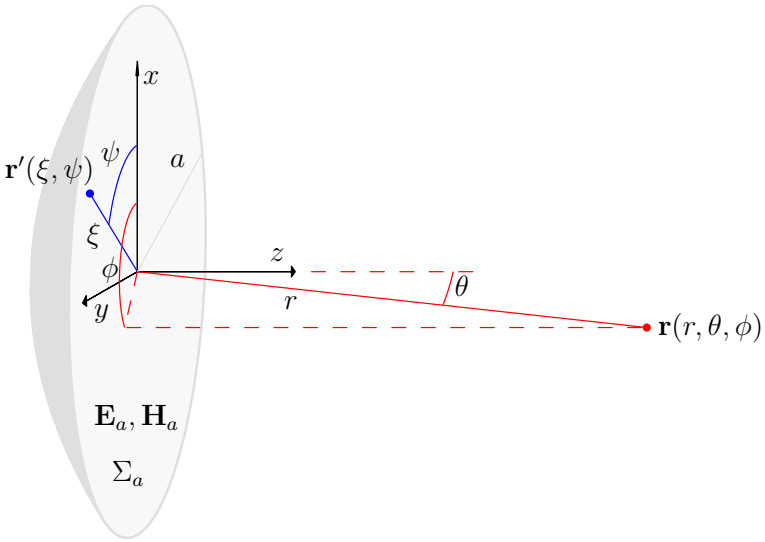


Figure A.1: Aperture field to far-field transformation.

Consider the reflector antenna shown in Figure A.1. Assume that the electric field $\mathbf{E}_a$ and the magnetic field $\mathbf{H}_a$ on the antenna aperture Σ_a are known. Select now a surface that extends to infinity in the aperture plane and placed such that the rim of the reflector lies in this surface, and assume that the fields on this surface outside the aperture are zero. Using Love's principle the aperture fields may be replaced by equivalent current sources on

the aperture as [63, Equations (12.3) and (12.4)]

$$\mathbf{J}_a = \hat{\mathbf{n}} \times \mathbf{H}_a \tag{A.0.1a}$$

$$\mathbf{M}_a = -\hat{\mathbf{n}} \times \mathbf{E}_a. \tag{A.0.1b}$$

Replacing the infinite surface with a perfectly electrical conductor results in $\mathbf{J}_a = \mathbf{0}$ and by image theory the magnetic current becomes

$$\mathbf{M}_a = -2\hat{\mathbf{n}} \times \mathbf{E}_a. \tag{A.0.2}$$

The magnetic vector potential $\mathbf{A}_M$ may now be determined as [63, Equation (3.52)]

$$\mathbf{A}_M = \frac{\epsilon}{4\pi} \int_{\Sigma_a} \mathbf{M}_a \frac{e^{-jkR}}{R} \, \mathrm{d}\sigma \tag{A.0.3}$$

where R is the distance from a point on the aperture to the point where $\mathbf{A}_M$ is calculated.

In the usual far-field approximation we have

$$R \approx r \tag{A.0.4a}$$

for the amplitude term and may be removed from under the integral, and

$$R \approx r - \mathbf{k} \cdot \mathbf{r}' \tag{A.0.4b}$$

for the phase term. Then the expression for the magnetic vector potential reduces to

$$\mathbf{A}_M \approx \frac{\epsilon}{4\pi} \frac{e^{-jkr}}{r} \int_{\Sigma_a} \mathbf{M}_a e^{j\mathbf{k}\cdot\mathbf{r}'} \, \mathrm{d}\sigma \tag{A.0.5}$$

Using $\mathbf{E} = j\omega\eta\hat{\mathbf{r}} \times \mathbf{A}_M$ [63, Equation (3.59b)] the electric far-field is then found to be

$$\mathbf{E} = \frac{j\omega\eta\epsilon}{4\pi} \frac{e^{-jkr}}{r} \int_{\Sigma_a} \hat{\mathbf{r}} \times \mathbf{M}_a e^{-j\mathbf{k}\cdot\mathbf{r}'} \, \mathrm{d}\sigma \tag{A.0.6}$$

which after manipulation reduces to the form of (3.1.1)

$$\mathbf{E}(\mathbf{r}) \approx -jk \frac{e^{-jkr}}{2\pi r} \int_{\Sigma_a} \hat{\mathbf{r}} \times \hat{\mathbf{n}} \times \mathbf{E}_a(\mathbf{r}') e^{j\mathbf{k}\cdot\mathbf{r}'} \, \mathrm{d}\sigma. \tag{A.0.7}$$

Appendix B

Miscellaneous Relations Pertaining to Bessel Functions

Herein various properties and relations pertaining to Bessel functions that are used in the thesis are shown. The primary reference in this regard is [67].

B.1 Limiting Value for Small Argument

The Bessel function of the first kind of order ν is defined by the equation [67, Equation (8) in § 3.1]

$$J_\nu(z) = \sum_{m=0}^{\infty} \frac{(-)^m (\frac{1}{2}z)^{\nu+2m}}{m\,!\,\Gamma(\nu+m+1)} \tag{B.1.1}$$

where $\Gamma(x)$ is the Gamma-function. Direct substitution of this expression into

$$\frac{J_\nu(z)}{z} \tag{B.1.2}$$

gives

$$\frac{J_\nu(z)}{z} = \sum_{m=0}^{\infty} \frac{(-)^m (\frac{1}{2}z)^{\nu-1+2m}}{m\,!\,\Gamma(\nu+m+1)}. \tag{B.1.3}$$

The expression on the right-hand side may be evaluated directly at $z = 0$ for $\nu \geq 1$, in which case all terms $m > 0$ are exactly equal to zero and

$$\left.\frac{(-)^m (\frac{1}{2}z)^{\nu-1+2m}}{m\,!\,\Gamma(\nu+m+1)}\right|_{m=0} = \frac{(\frac{1}{2}z)^{\nu-1}}{\Gamma(\nu+1)}. \tag{B.1.4}$$

Evaluating this expression for various values of $\nu \geq 1$ and $z = 0$ gives the desired result

$$\lim_{z\to 0} \frac{J_\nu(z)}{z} = \begin{cases} \frac{1}{2} & \nu = 1 \\ 0 & \nu > 1. \end{cases} \tag{B.1.5}$$

B.2 The Neumann Series

An arbitrary function $f(z)$ of the complex variable z which is analytic inside a circle of radius R centred at the origin $z = 0$ admits an expansion of the form [67, § 16.1]

$$f(z) = \sum_{n=0}^{\infty} a_n J_n(z) \tag{B.2.1}$$

which is called the *Neumann series* of f. The coefficients a_n may be calculated using

$$a_n = \frac{\epsilon_n}{j2\pi} \int_C f(t) O_n(t)\, \mathrm{d}t \tag{B.2.2}$$

where C is a contour in the complex plane defined by the circle with radius R, O_n is a Neumann polynomial [67, § 9.1]

$$O_n(t) = \begin{cases} \frac{1}{4} \sum_{m=0}^{m \leq \frac{1}{2}n} \frac{n(n-m-1)!}{m!\,\left(\frac{1}{2}t\right)^{n-2m+1}} & n > 0 \\ \frac{1}{t} & n = 0, \end{cases} \tag{B.2.3}$$

and

$$\epsilon_n = \begin{cases} 2 & n > 0 \\ 1 & n = 0. \end{cases} \tag{B.2.4}$$

B.3 An Orthogonality Property of Bessel Functions

Bessel functions of the first kind satisfy [67, § 16.4]

$$\int_0^{\infty} J_{2m+1}(t) J_{2n+1}(t) \frac{1}{t}\, \mathrm{d}t = \begin{cases} 0 & m \neq n \\ \frac{1}{4n+2} & m = n. \end{cases} \tag{B.3.1}$$

Appendix C

Removing Redundancy in the CBFP Basis

In Section 4.2.1 the SVD is used to remove redundancy in the set of CBFPs that are constructed to compensate for a particular error in the antenna system. The procedure outlined there is illustrated below for a numerical example.

The starting point for this example is the linear system in (4.2.1)

$$\tilde{\mathbf{F}} = \mathbf{Z}\mathbf{x} = \begin{bmatrix} \mathbf{f}_1 & \mathbf{f}_2 & \mathbf{f}_3 \end{bmatrix} \mathbf{x} \tag{C.0.1}$$

wherein the columns $\mathbf{f}_i$ of $\mathbf{Z}$ are the CBFPs (basis functions), $\tilde{\mathbf{F}}$ is the CBFP model, and $\mathbf{x}$ is the model parameters (coefficients) vector . Let the CBFPs be

$$\mathbf{f}_1 = \begin{bmatrix} 1.0000 & 0.5000 & 0.5000 & 0.1000 & 0.1000 \end{bmatrix}^T \tag{C.0.2a}$$

$$\mathbf{f}_2 = \begin{bmatrix} 1.0000 & 0.4500 & 0.5700 & 0.1200 & 0.1300 \end{bmatrix}^T \tag{C.0.2b}$$

$$\mathbf{f}_3 = \begin{bmatrix} 1.0000 & 0.5300 & 0.4200 & 0.1100 & 0.0900 \end{bmatrix}^T \tag{C.0.2c}$$

and the pattern to be modelled equal

$$\mathbf{F} = \begin{bmatrix} 1.0000 & 0.5800 & 0.5600 & 0.1100 & 0.1400 \end{bmatrix}^T. \tag{C.0.3}$$

Using the overdetermined system in (4.2.4) to obtain a benchmark solution yields

$$\mathbf{x} = \left(\mathbf{Z}^\dagger \mathbf{Z}\right)^{-1} \mathbf{Z}^\dagger \mathbf{F} = \begin{bmatrix} 1.9125 & -0.3341 & -0.5346 \end{bmatrix}^T \tag{C.0.4}$$

and

$$\tilde{\mathbf{F}} = \mathbf{Z}\mathbf{x} = \begin{bmatrix} 1.0438 & 0.5226 & 0.5413 & 0.0924 & 0.0997 \end{bmatrix}^T. \tag{C.0.5}$$

Since $\mathbf{F}$ is not in the column space of $\mathbf{Z}$ the above solution corresponds to the least-squares error between the model and the actual pattern, and the model may not fit the pattern exactly in any of the directions.

Assume the determined system in (4.2.5) is constructed to point-match the model to the pattern in the directions corresponding to the first three elements in $\mathbf{F}$. Then

$$\mathbf{Z}_M = \begin{bmatrix} 1.0000 & 1.0000 & 1.0000 \\ 0.5000 & 0.4500 & 0.5300 \\ 0.5000 & 0.5700 & 0.4200 \end{bmatrix} \tag{C.0.6a}$$

$$\mathbf{V}_M = \begin{bmatrix} 1.0000 \\ 0.5800 \\ 0.5600 \end{bmatrix} \tag{C.0.6b}$$

and the solution vector is found to be

$$\mathbf{x} = \mathbf{Z_M}^{-1}\mathbf{V}_M = \begin{bmatrix} 9.8421 & -4.3158 & -4.5263 \end{bmatrix}^T, \tag{C.0.7}$$

which yields the model

$$\tilde{\mathbf{F}} = \mathbf{Z}\mathbf{x} = \begin{bmatrix} 1.0000 & 0.5800 & 0.5600 & -0.0316 & 0.0158 \end{bmatrix}^T. \tag{C.0.8}$$

The columns in $\mathbf{Z}_M$ are linearly independent (they span $\mathbb{R}^3$) which means that the model *exactly* fits the pattern in the first three directions. However, the error in the fourth and fifth directions are observed to be significantly larger than previously obtained with the overdetermined system.

Consider now the case where we want to remove any CBFPs that are redundant over the region which contains the first four directions in $\mathbf{F}$. Form the matrix $\mathbf{Z}_R$ by inserting the first four elements in each basis function as a column

$$\mathbf{Z}_R = \begin{bmatrix} 1.0000 & 1.0000 & 1.0000 \\ 0.5000 & 0.4500 & 0.5300 \\ 0.5000 & 0.5700 & 0.4200 \\ 0.1000 & 0.1200 & 0.1100 \end{bmatrix}. \tag{C.0.9}$$

The condition number of $\mathbf{Z}_R$ is calculated to be equal to $\kappa(\mathbf{Z}_R) = 144.5111$, indicating a possibly large degree of redundancy in the basis over the limited region of interest. It is now desirable to orthogonalise the set of basis functions over that region. This may be done by computing the SVD[1]

$$\mathbf{U}\mathbf{\Sigma}\mathbf{V}^\dagger = \mathbf{Z}_R \tag{C.0.10}$$

[1] In *Matlab* the command `[U,S,V] = svd(Z_R,'econ');` can be used to compute the SVD. By passing the optional string parameter `'econ'` the returned matrices `U` and `V` contain only the left- and right-singular vectors of `Z_R`, respectively. That is, the columns of `U` form an orthonormal basis of the column space of `Z_R`. If the optional parameter `'econ'` is not passed, the returned `U` contains a number of additional columns which form an orthonormal basis for the orthogonal complement of the column space of `Z_R`. For an $m \times n$ matrix `Z_R` the columns of `U` then span $\mathbb{R}^m$.

wherein

$$\mathbf{U} = \begin{bmatrix} -0.8158 & 0.1435 & 0.2297 \\ -0.4022 & 0.5486 & -0.3881 \\ -0.4057 & -0.8224 & -0.2657 \\ -0.0898 & -0.0453 & 0.8521 \end{bmatrix} \tag{C.0.11a}$$

$$\mathbf{\Sigma} = \begin{bmatrix} 2.1230 & 0 & 0 \\ 0 & 0.1186 & 0 \\ 0 & 0 & 0.0147 \end{bmatrix} \tag{C.0.11b}$$

$$\mathbf{V} = \begin{bmatrix} -0.5788 & 0.0172 & -0.8153 \\ -0.5835 & -0.7071 & 0.3993 \\ -0.5696 & 0.7069 & 0.4193 \end{bmatrix}. \tag{C.0.11c}$$

Normalising the singular values $\boldsymbol{\sigma}$ (diagonal elements of $\mathbf{\Sigma}$) to the maximum singular value obtains the vector

$$\boldsymbol{\sigma}_{\text{norm}} = \frac{\boldsymbol{\sigma}}{\sigma_1} = \frac{\operatorname{diag}\mathbf{\Sigma}}{(\mathbf{\Sigma})_{11}} = \begin{bmatrix} 1.0000 & 0.0559 & 0.0069 \end{bmatrix}^T. \tag{C.0.12}$$

We now form a new set of basis functions using the columns of $\mathbf{U}$ corresponding to singular values for which $\boldsymbol{\sigma}_{\text{norm}} > 0.01$, that is, only the first two columns of $\mathbf{U}$ are selected as basis functions. The CBFP model using this basis is now expressed as in (4.2.11) as

$$\tilde{\mathbf{F}}_R = \mathbf{R}_R\mathbf{y} \tag{C.0.13}$$

where

$$\mathbf{R}_R = \begin{bmatrix} -0.8158 & 0.1435 \\ -0.4022 & 0.5486 \\ -0.4057 & -0.8224 \\ -0.0898 & -0.0453 \end{bmatrix} \tag{C.0.14}$$

and $\mathbf{y}$ is the coefficients vector for this model. A solution for this vector can also be obtained by matching the model to the pattern in the directions corresponding to the first three elements of $\mathbf{F}$. This requires producing the system in (4.2.12) with

$$\mathbf{R}_M = \begin{bmatrix} -0.8158 & 0.1435 \\ -0.4022 & 0.5486 \\ -0.4057 & -0.8224 \end{bmatrix} \tag{C.0.15}$$

and $\mathbf{V}_M$ as defined in (C.0.6b). The coefficients vector is obtained as

$$\mathbf{y} = \mathbf{R}_M^{-1}\mathbf{V}_M = \begin{bmatrix} -1.2867 & -0.0041 \end{bmatrix}^T \tag{C.0.16}$$

and the model produced with this solution is

$$\tilde{\mathbf{F}}_R = \mathbf{R}_R\mathbf{y} = \begin{bmatrix} 1.0491 & 0.5153 & 0.5254 & 0.1157 \end{bmatrix}^T. \tag{C.0.17}$$

Note that the model $\tilde{\mathbf{F}}_R$ only contains four directions, whereas the original pattern that needs to be modelled contains five directions. The model using the orthonormalised basis functions may now be extended to include the fifth direction by use of the right-singular vectors (columns of $\mathbf{V}$). Towards this end we form the matrix $\mathbf{R}$ by using the vectors from (4.2.14) as columns, that is

$$\mathbf{R} = \begin{bmatrix} \mathbf{r}_1 & \mathbf{r}_2 \end{bmatrix} \tag{C.0.18}$$

where

$$\mathbf{r}_1 = \frac{1}{\sigma_1}\mathbf{Z}\mathbf{v_1} = \begin{bmatrix} -0.8158 & -0.4022 & -0.4057 & -0.0898 & -0.0871 \end{bmatrix}^T \tag{C.0.19a}$$

$$\mathbf{r}_2 = \frac{1}{\sigma_2}\mathbf{Z}\mathbf{v_2} = \begin{bmatrix} 0.1435 & 0.5486 & -0.8224 & -0.0453 & -0.2241 \end{bmatrix}^T. \tag{C.0.19b}$$

Herein the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are the first and second columns of $\mathbf{V}$, respectively, as obtained from the SVD

$$\mathbf{v}_1 = \begin{bmatrix} -0.5788 & -0.5835 & -0.5696 \end{bmatrix}^T \tag{C.0.20a}$$

$$\mathbf{v}_2 = \begin{bmatrix} 0.0172 & -0.7071 & 0.7069 \end{bmatrix}^T, \tag{C.0.20b}$$

the singular values are

$$\sigma_1 = 2.1230 \tag{C.0.21a}$$

$$\sigma_2 = 0.1186, \tag{C.0.21b}$$

and $\mathbf{Z}$ is as defined in (C.0.1). By using the solution vector obtained from (C.0.16), a model over all five directions is now obtained using (4.2.13)

$$\tilde{\mathbf{F}} = \mathbf{R}\mathbf{y} = \begin{bmatrix} 1.0491 & 0.5153 & 0.5254 & 0.1157 & 0.1131 \end{bmatrix}^T. \tag{C.0.22}$$

Bibliography

[1] Brandwood, D.: A complex gradient operator and its application in adaptive array theory. *Microwaves, Optics and Antennas, IEE Proceedings H*, vol. 130, no. 1, pp. 11–16, February 1983.

[2] Jansky, K.G.: Electrical Disturbances Apparently of Extraterrestrial Origin. *Proceedings of the Institute of Radio Engineers*, vol. 21, no. 10, pp. 1387–1398, October 1933.

[3] Reber, G.: Cosmic Static. *The Astrophysical Journal*, vol. 100, pp. 279–287, 1944.

[4] Burke, B.F. and Graham-Smith, F.: *An Introduction to Radio Astronomy*. 1st edn. Cambridge University Press, 1997.

[5] Thompson, A.R., Moran, J.M. and Swenson, G.: *Interferometry and Synthesis in Radio Astronomy*. 2nd edn. Wiley, 2004.

[6] Ryle, M., Hewish, A. and Shakeshaft, J.: The synthesis of large radio telescopes by the use of radio interferometers. *IRE Transactions on Antennas and Propagation*, vol. 7, no. 5, pp. 120–124, December.

[7] Scott, P.F., Ryle, M. and Hewish, A.: First results of radio star observations using the method of aperture synthesis. *Monthly Notices of the Royal Astronomical Society*, vol. 122, no. 2, pp. 95–111, 1961.

[8] Baars, J. W. M. and van der Brugge, J. F. and Casse, J. L. and Hamaker, J. P. and Sondaar, L. H. and Visser, J. J. and Wellington, K. J.: The synthesis radio telescope at Westerbork. *Proceedings of the IEEE*, vol. 61, no. 9, pp. 1258–1266, 1973.

[9] Thompson, A.R., Clark, B.G., Wade, C.M. and Napier, P.J.: The Very Large Array. *The Astrophysical Journal Supplement Series*, vol. 44, pp. 151–167, October 1980.

[10] Dewdney, P.E., Hall, P.J., Schilizzi, R.T. and Lazio, T.J.L.W.: The Square Kilometre Array. *Proceedings of the IEEE*, vol. 97, no. 8, pp. 1482–1496, August 2009.

[11] The Square Kilometre Array. March 2013.
Available at: `http://www.skatelescope.org/`

[12] Warnick, K., Ivashina, M. and Hay, S.G.: Guest Editorial for the Special Issue on Antennas for Next Generation Radio Telescopes. *IEEE Transactions on Antennas and Propagation*, vol. 59, no. 6, pp. 1786–1789, June 2011.

[13] Taylor, A.R.: The Square Kilometre Array. In: *Proceedings of the International Astronomical Union Symposium*. 2012.

[14] Jonas, J.L.: MeerKAT - The South African Array With Composite Dishes and Wide-Band Single Pixel Feeds. *Proceedings of the IEEE*, vol. 97, no. 8, pp. 1522–1530, August 2009.

[15] DeBoer, D.R., Gough, R.G., Bunton, J.D., Cornwell, T.J., Beresford, R.J., Johnston, S., Feain, I.J., Schinckel, A.E., Jackson, C.A., Kesteven, M.J. *et al.*: Australian SKA pathfinder: A high-dynamic range wide-field of view survey telescope. *Proceedings of the IEEE*, vol. 97, no. 8, pp. 1507–1521, August 2009.

[16] Braun, R.: Understanding Synthesis Imaging Dynamic Range. *arXiv preprint arXiv:1208.6352*, 2013.

[17] Smirnov, O.M.: Revisiting the radio interferometer measurement equation. II. Calibration and direction-dependent effects. *Astronomy & Astrophysics*, vol. 527, p. A107, 2011.

[18] Noordam, J.E. and Smirnov, O.M.: The MeqTrees software system and its use for third-generation calibration of radio interferometers. *Astronomy & Astrophysics*, vol. 524, p. A61, 2010.

[19] Kazemi, S., Yatawatta, S., Zaroubi, S., Lampropoulos, P., de Bruyn, A.G., Koopmans, L.V.E. and Noordam, J.: Radio interferometric calibration using the SAGE algorithm. *Monthly Notices of the Royal Astronomical Society*, vol. 414, no. 2, pp. 1656–1666, 2011.

[20] Smirnov, O.M.: Revisiting the radio interferometer measurement equation. III. Addressing direction-dependent effects in 21 cm WSRT observations of 3C 147. *Astronomy & Astrophysics*, vol. 527, p. A108, March 2011.

[21] Intema, H.T., Van der Tol, S., Cotton, W.D., Cohen, A.S., Van Bemmel, I.M. and Röttgering, H.J.A.: Ionospheric calibration of low frequency radio interferometric observations using the peeling scheme. *Astronomy & Astrophysics*, vol. 501, no. 3, pp. 1185–1205, 2009.

[22] Wijnholds, S.J., van der Tol, S., Nijboer, R. and van der Veen, A.: Calibration challenges for future radio telescopes. *IEEE Signal Processing Magazine*, vol. 27, no. 1, pp. 30–42, January 2010.

[23] Harp, G. R. *et al.*: Primary Beam and Dish Surface Characterization at the Allen Telescope Array by Radio Holography. *IEEE Transactions on Antennas and Propagation*, vol. 59, no. 6, pp. 2004–2021, June 2011.

[24] Noordam, J.E.: Measuring Station Beamshapes as a function of time and frequency and in full polarization. In: *Third Generation Calibration Workshop II (3GC-II)*. Albufeira (Portugal), September 2011.

[25] Wijnholds, S.J. and van der Veen, A.: Fundamental imaging limits of radio telescope arrays. *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 5, pp. 613–623, 2008.

[26] Wijnholds, S.J., Grainge, K.J.B. and Nijboer, R.: Report on Calibration and Imaging Working Group Activities. Tech. Rep., Calibration and Imaging Working Group, 2012.

[27] Galindo-Israel, V. and Mittra, R.: A new series representation for the radiation integral with application to reflector antennas. *IEEE Transactions on Antennas and Propagation*, vol. 25, no. 5, pp. 631–641, September 1977.

[28] Rahmat-Samii, Y. and Galindo-Israel, V.: Shaped reflector antenna analysis using the Jacobi-Bessel series. *IEEE Transactions on Antennas and Propagation*, vol. 28, no. 4, pp. 425–435, July 1980.

[29] Young, A., Maaskant, R., Ivashina, M.V., Davidson, D.B., de Villiers, D.I.L. and Iupikov, O.A.: Improving Beam Pattern Estimation for Future Radio Telescopes. In: *Calibration and Imaging Workshop (CALIM)*. Cape Town, South Africa, December 2012.

[30] Maaskant, R., Ivashina, M.V., Wijnholds, S.J. and Warnick, K.F.: Efficient Prediction of Array Element Patterns Using Physics-Based Expansions and a Single Far-Field Measurement. *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 8, pp. 3614–3621, August 2012.

[31] Maaskant, R. and Ivashina, M.V.: Characteristic Basis Function Patterns – A Novel Expansion Method for the Fast and Accurate Prediction of Antenna Array Beams. In: *Proceedings of the International Conference on Electromagnetics in Advanced Applications (ICEAA)*, pp. 796–799. Cape Town, September 2012.

[32] Young, A., Maaskant, R., Ivashina, M.V., de Villiers, D.I.L. and Davidson, D.B.: Accurate Beam Prediction Through Characteristic Basis Function Patterns for the MeerKAT/SKA Radio Telescope Antenna. *IEEE Transactions on Antennas and Propagation*, Scheduled for publication in 2013.

[33] Stutzman, W.L. and Thiele, G.A.: *Antenna Theory and Design*. 2nd edn. Wiley, New York, N.Y., 1998.

[34] Ivashina, M.V., Iupikov, O.A., Maaskant, R., van Cappellen, W.A. and Oosterloo, T.: An Optimal Beamforming Strategy for Wide-Field Surveys With Phased-Array-Fed Reflector Antennas. *IEEE Transactions on Antennas and Propagation*, vol. 59, no. 6, pp. 1864–1875, June 2011.

[35] Iupikov, O.A., Ivashina, M.V. and Smirnov, O.M.: Reducing the complexity of the beam calibration models of phased-array radio telescopes. In: *Proceedings of the European Conference on Antennas and Propagation (EuCAP)*, pp. 930–933. April 2011.

[36] Young, A., Maaskant, R., Ivashina, M.V. and Davidson, D.B.: Performance Evaluation of Far Field Paterns for Radio Astronomy Application through the Use of the Jacobi-Bessel Series. In: *Proceedings of the International Conference on Electromagnetics in Advanced Applications (ICEAA)*, pp. 884–887. Cape Town, September 2012.

[37] Young, A., Ivashina, M.V., Maaskant, R., Iupikov, O.A. and Davidson, D.B.: Improving the Calibration Efficiency of an Array Fed Reflector Antenna through Constrained Beamforming. *IEEE Transactions on Antennas and Propagation*, 2013. Accepted for publication.

[38] Craeye, C. and Alexander, P.: Compact representation of the main beam and first few sidelobes for aperture-type antennas, August 2011. Unpublished.

[39] Craeye, C., Gonzalez-Ovejero, D., de Lera Acedo, E., Ghods, N.R. and Alexander, P.: Main beam representation in non-regular arrays. In: *Calibration and Imaging Workshop (CALIM)*. Manchester (UK), July 2011.

[40] Hamaker, J.P., Bregman, J.D. and Sault, R.J.: Understanding radio polarimetry. I. Mathematical foundations. *Astronomy & Astrophysics Supplement Series*, vol. 117, no. 1, pp. 137–147, 1996.

[41] Theron, I., Lehmensiek, R. and de Villiers, D.: The design of the MeerKAT dish optics. In: *Proceedings of the International Conference on Electromagnetics in Advanced Applications (ICEAA)*, pp. 539–542. September 2012.

[42] Mizugutch, Y., Akagawa, M. and Yokoi, H.: Offset dual reflector antenna. In: *Antennas and Propagation Society International Symposium*, vol. 14, pp. 2–5. October 1976.

[43] Granet, C.: Designing classical offset Cassegrain or Gregorian dual-reflector antennas from combinations of prescribed geometric parameters. *IEEE Antennas and Propagation Magazine*, vol. 44, no. 3, pp. 114–123, June 2002.

[44] TICRA. March 2013.
Available at: `http://www.ticra.com/`

[45] Pontoppidan, K.: *GRASP9 Technical Description*. Copenhagen, Denmark, September 2005.

[46] Ivashina, M.V., Iupikov, O. and Van Cappellen, W.: Extending the capabilities of the GRASP and CAESAR software to analyze and optimize active beamforming array feeds for reflector systems. In: *Proceedings of the International Conference on Electromagnetics in Advanced Applications (ICEAA)*, pp. 197–200. September 2010.

[47] The MathWorks, Inc. March 2013.
Available at: `http://www.mathworks.com/products/matlab`

[48] Verheijen, M.A.W., Oosterloo, T.A., van Cappellen, W.A., Bakker, L., Ivashina, M.V. and van der Hulst, J.M.: APERTIF, a focal plane array for the WSRT. In: *American Institute of Physics Conference Proceedings*. 2008.

[49] Maaskant, R., Tijhuis, A.G., Mittra, R., Ivashina, M.V., Van Cappellen, W.A. and Arts, M.J.: Hybridization of Efficient Modeling Techniques for Fast Analysis of Large-Scale Antenna Structures in the Context of the Square Kilometre Array Project. In: *38th European Microwave Conference (EuMC)*, pp. 837–840. October 2008.

[50] Maaskant, R.: *Analysis of Large Antenna Systems*. Ph.D. thesis, Eindhoven University of Technology, Eindhoven, 2010.

[51] Hamaker, J.: Understanding radio polarimetry. IV. The full-coherency analogue of scalar self-calibration: Self-alignment, dynamic range and polarimetric fidelity. *Astronomy & Astrophysics Supplement Series*, vol. 143, no. 3, pp. 515–534, 2000.

[52] Smirnov, O.M.: Revisiting the radio interferometer measurement equation. I. A full-sky Jones formalism. *Astronomy & Astrophysics*, vol. 527, p. A106, 2011.

[53] Peebles, P. Z., Jr.,: *Probability, Random Variables, and Random Signal Principles*. 4th edn. McGraw-Hill, 2001.

[54] Ludwig, A.: The definition of cross polarization. *IEEE Transactions on Antennas and Propagation*, vol. 21, no. 1, pp. 116–119, January 1973.

[55] Smith, P.G.: Measurement of the Complete Far-Field Pattern of Large Antennas by Radio-Star Sources. *IEEE Transactions on Antennas and Propagation*, vol. 14, no. 1, pp. 6–16, January 1966.

[56] Young, A., Terada, M.A.B., de Villiers, D.I.L. and Davidson, D.B.: Assessment of the Sensitivity of the South African KAT-7 and MeerKAT/SKA Radio Telescope Reflector Antennas. In: *Proceedings of the International Conference on Electromagnetics in Advanced Applications (ICEAA)*, pp. 486–489. Cape Town, September 2012.

[57] Cornwell, T.J., Holdaway, M.A. and Uson, J.M.: Radio-interferometric imaging of very large objects: implications for array design. *Astronomy & Astrophysics*, vol. 271, pp. 697–713, 1993.

[58] Popping, A. and Braun, R.: The standing wave phenomenon in radio telescopes. *Astronomy & Astrophysics*, vol. 479, no. 3, pp. 903–913, 2008.

[59] Smirnov, O.M.: 2013. Private communication.

[60] Uson, J.M. and Cotton, W.D.: Beam squint and Stokes V with off-axis feeds. *Astronomy & Astrophysics*, vol. 486, no. 2, pp. 647–654, 2008.

[61] Condon, J.J., Cotton, W.D., Greisen, E.W., Yin, Q.F., Perley, R.A., Taylor, G.B. and Broderick, J.J.: The NRAO VLA Sky Survey. *The Astronomical Journal*, vol. 115, no. 5, pp. 1693–1716, 1998.

[62] Theron, I.: 2012. Private communication.

[63] Balanis, C.A.: *Antenna Theory: Analysis and Design*. 2nd edn. Wiley, 1997.

[64] Born, M. and Wolf, E.: *Principles of Optics*. Third (revised) edn. Pergamon Press, 1965.

[65] Ismatullah and Rana, I.E.: Implementation of the Jacobi-bessel series method for radiation-pattern equations of an offset parabolic reflector antenna using MathCAD®. *IEEE Antennas and Propagation Magazine*, vol. 47, no. 3, pp. 99–104, June 2005.

[66] Rabelo, R., Terada, M.A.B. and Stutzman, W.: Analysis of reflector antennas through the World Wide Web [EM Programmer's Notebook]. *IEEE Antennas and Propagation Magazine*, vol. 49, no. 2, pp. 113–116, 2007.

[67] Watson, G.N.: *A Treatise on the Theory of Bessel Functions*. 2nd edn. Cambridge University Press, 1966.

[68] Nocedal, J. and Wright, S.J.: *Numerical Optimization*. 2nd edn. Springer, 2006.

[69] Craeye, C., Gonzalez-Ovejero, D., Ghods, N.R. and de Lera Acedo, E.: A Projection Approach to Model the Main Beam of Non-Regular Arrays in Presence of Mutual Coupling. In: *Proceedings of the European Conference on Antennas and Propagation (EUCAP)*, pp. 2609–2612. March 2012.

[70] Bucci, O.M., Gennarelli, C. and Savarese, C.: Optimal interpolation of radiated fields over a sphere. *IEEE Transactions on Antennas and Propagation*, vol. 39, no. 11, pp. 1633–1643, November 1991.

[71] Galindo-Israel, V., Veruttipong, W., Norrod, R.D. and Imbriale, W.A.: Scanning properties of large dual-shaped offset and symmetric reflector antennas. *IEEE Transactions on Antennas and Propagation*, vol. 40, no. 4, pp. 422–432, April 1992.

[72] von Hoerner, S. and Wong, W.-Y.: Gravitational deformation and astigmatism of tiltable radio telescopes. *IEEE Transactions on Antennas and Propagation*, vol. 23, no. 5, pp. 689–695, September 1975.

[73] Fisher, J.R.: Phased Array Feeds for Low Noise Reflector Antennas. Tech. Rep. 307, NRAO, 1996.

[74] Veidt, B.: Focal-Plane Array Architectures: Horn Clusters vs. Phased-Array Techniques. In: *SKA Memo Series*, Memo 71. February 2006.

[75] Van Ardenne, A., Bregman, J., van Cappellen, W., Kant, G. and de Vaate, J.: Extending the Field of View With Phased Array Techniques: Results of European SKA Research. *Proceedings of the IEEE*, vol. 97, no. 8, pp. 1531–1542, August 2009.

[76] Jeffs, B.D., Warnick, K.F., Landon, J., Waldron, J., Jones, D., Fisher, J.R. and Norrod, R.D.: Signal Processing for Phased Array Feeds in Radio Astronomical Telescopes. *IEEE Journal of Selected Topics in Signal Processing*, vol. 2, no. 5, pp. 635–646, October 2008.

[77] Johansson, J.F.: Fundamental limits for focal-plane array efficiency. In: *Multi-Feed Systems for Radio Telescopes*, vol. 75 of *ASP Conference Series*, pp. 34–41. 1995.

[78] Padman, R.: Optical fundamentals of array feeds. In: *Astronomical Society of the Pacific Conference Series: Multi-feed Systems for Radio Telescopes*, pp. 3–26. 1995.

[79] Van Cappellen, W.A., de Vaate, J.G.B., Ivashina, M.V., Bakker, L. and Oosterloo, T.: Focal plane arrays evolve. In: *Proceedings of the XXIXth General Assembly of the International Union of Radio Science (URSI)*. Chicago,IL, August 2008.

[80] Smirnov, O.M. and Ivashina, M.V.: Element gain drifts as an imaging dynamic range limitation in PAF-based interferometers. In: *Proceedings of the XXXth General Assembly and Scientific Symposium of the International Union of Radio Science (URSI)*, pp. 1–4. August 2011.

[81] Hay, S.G.: SKA field of view de-rotation using connected array beam scanning with constant beam shape. In: *Asia-Pacific Microwave Conference Proceedings (APMC)*, pp. 1174–1177. December 2011.

[82] Elmer, M., Jeffs, B., Warnick, K., Roshi, D. and Fisher, J.: Calibration and field flattening for single dish astronomical phased array feed radio camera imaging. In: *Proceedings of the International Conference on Electromagnetics in Advanced Applications (ICEAA)*, pp. 958–961. September 2012.

[83] Elmer, M., Jeffs, B.D., Warnick, K.F., Fisher, J.R. and Norrod, R.D.: Beamformer Design Methods for Radio Astronomical Phased Array Feeds. *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 2, pp. 903–914, February 2012.

[84] Aumann, H., Fenn, A. and Willwerth, F.: Phased array antenna calibration and pattern prediction using mutual coupling measurements. *IEEE Transactions on Antennas and Propagation*, vol. 37, no. 7, pp. 844–850, July 1989.

[85] Warnick, K.F., Jeffs, B.D., Landon, J., Waldron, J., Jones, D., Fisher, J.R. and Norrod, R.: Beamforming and imaging with the BYU/NRAO L-band 19-element phased array feed. In: *13th International Symposium on Antenna Technology and Applied Electromagnetics and the Canadian Radio Science Meeting (ANTEM/URSI)*, pp. 1–4. February 2009.

[86] Van Trees, H.L.: *Optimum Array Processing*. Wiley, New York, N.Y., 2002.

[87] Hu, Y.Y.: A method of determining phase centers and its application to electromagnetic horns. Tech. Rep. 1, Franklin Institute, 1961.

[88] Landon, J., Elmer, M., Waldron, J., Jones, D., A., Jeffs, B.D., Warnick, K.F., Fisher, J.R. and Norrod, R.D.: Phased Array Feed Calibration, Beamforming, and Imaging. *The Astronomical Journal*, vol. 139, no. 3, pp. 1154–1167, 2010.

[89] van Cappellen, W., Wijnholds, S. and Bakker, L.: Experimental evaluation of polarimetric beamformers for an L-band phased array feed. In: *Proceedings of the European Conference on Antennas and Propagation (EUCAP)*, pp. 634–637. March 2012.

[90] Warnick, K., Ivashina, M., Wijnholds, S. and Maaskant, R.: Polarimetry With Phased Array Antennas: Theoretical Framework and Definitions. *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 1, pp. 184–196, January 2012.

[91] Ruze, J.: Lateral-feed displacement in a paraboloid. *IEEE Transactions on Antennas and Propagation*, vol. 13, no. 5, pp. 660–665, September 1965.

[92] Imbriale, W.A., Ingerson, P.G. and Wong, W.C.: Large Lateral Feed Displacements in a Parabolic Reflector. *IEEE Transactions on Antennas and Propagation*, vol. 22, no. 6, pp. 742–745, November 1974.

[93] Young, A., Maaskant, R., Ivashina, M.V., Davidson, D.B., de Villiers, D.I.L. and Iupikov, O.A.: Characteristic Basis Function Patterns Method and other beam modelling techniques. In: *Third Generation Calibration Workshop III (3GC-III)*. Port Alfred, South Africa, February 2013.

[94] Young, A., Maaskant, R., Ivashina, M.V., Davidson, D.B., de Villiers, D.I.L. and Iupikov, O.A.: Improving PAF Based Radio Telescope Calibration through Constrained Beamforming. In: *Third Generation Calibration Workshop III (3GC-III)*. Port Alfred, South Africa, February 2013.

[95] De Villiers, D.I.L.: Prediction of Aperture Efficiency Ripple in Clear Aperture Offset Gregorian Antennas. *IEEE Transactions on Antennas and Propagation*, 2013. Accepted for publication.

[96] Warnick, K.F.: 2013. Private communication.

[97] Wijnholds, S.J.: Aperture Array Station Beam Specifications, March 2012. Unpublished.

[98] Refregier, A.: Shapelets — I. A method for image analysis. *Monthly Notices of the Royal Astronomical Society*, vol. 338, no. 1, pp. 35–47, 2002.

[99] Yatawatta, S.: Shapelets and related techniques in radio-astronomical imaging. In: *Proceedings of the XXXth General Assembly and Scientific Symposium of the International Union of Radio Science (URSI)*, pp. 1–4. August 2011.