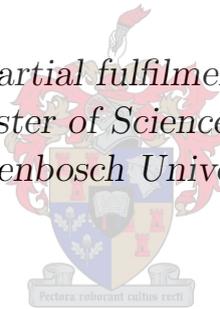


The design of a two-element correlation interferometer operating at L-band

by

Juan-Pierre Jansen van Rensburg

*Thesis presented in partial fulfilment of the requirements
for the degree Master of Science in Engineering at
Stellenbosch University*



Supervisor:

Prof. David B. Davidson
Department of Electrical and Electronic Engineering

December 2012

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the owner of the copyright thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: December 2012

Copyright © 2012 Stellenbosch University
All rights reserved.

Abstract

Keywords - Radio astronomy, interferometry, noise temperature, aperture synthesis, brightness, visibility, correlator, sensitivity, noise figure, radio-frequency interference.

This thesis describes the design and construction of a two-element digital correlation interferometer operating at 1.42 GHz. This instrument consists of two 92 cm offset-parabolic reflectors arranged in an east-west baseline, with a maximum length of 24 m. The reflectors are outfitted with helical beam antennas as primary feeds, and are further equipped with low-noise front-ends. A wideband dual-conversion superheterodyne receiver processes the astronomical signals before digitisation. The receiver is implemented as far as possible using affordable off-the-shelf technologies. The cross-correlation between the two antenna signals is measured using a 256 MHz bandwidth digital FX correlator, and is implemented on a Reconfigurable Open Architecture for Computing Hardware (ROACH) board. The preliminary observations made with this interferometer, suggests that it is possible to detect the sun, and some other much weaker sources such as Taurus A.

The design of an interferometer in general is presented, from the formulation of the underlying instrumental requirements, to making meridian drift scan observations. The interferometer developed may serve as a demonstrator for other engineering students to gain a working knowledge of radio interferometry, which should prove invaluable when addressing the challenges the Square Kilometre Array (SKA) is faced with.

Opsomming

Sleutelwoorde - Radio astronomie, interferometrie, ruistemperatuur, stralingsvlak sintese, helderheid, sigbaarheid, korreleerder, sensitiwiteit, ruissyfer, radio frekwensie steurings.

In hierdie tesis word die ontwerp en konstruksie van 'n twee-element digitale korrelasie interferometer wat by 1.42 GHz ontvang beskryf. Die instrument bestaan uit twee 92 cm afset-paraboliese reflektorantennas met 'n oos-wes basislyn, met maksimum lengte 24 m. Heliese antennes word gebruik as primêre voer vir die reflektors, en verder maak die voorent gebruik van laeruis versterkers. Die wyeband superheterodineontvanger verwerk die astronomiese seine in twee stappe voordat dit digitaliseer word. Bekostigbare kommersiële komponente word gebruik in die vervaardiging van die ontvanger. Die kruiskorrelasie tussen die twee antenne seine word digitaal gemeet deur 'n FX korreleerder met 'n 256 MHz bandwydte op 'n ROACH bord. Die resultate toon aan dat dit moontlik is om die son waar te neem, sowel as sommige ander swakke bronne, soos Taurus A.

Al die kwessies insluitende die instrumentale vereistes betrokke by die ontwerp van 'n interferometer word bespreek. Verder word die gebruik van interferometers om meridiaan meetings te doen bespreek. Die interferometer dien as 'n praktiese demonstrasie van radio interferometrie vir ingenieurstudente. Hierdie demonstrasie is voordelig vir studente om die uitdagings van die SKA te bestudeer.

Acknowledgements

I would like to express my sincere gratitude to the following people who assisted and supported me throughout this thesis, and if it was not for them, this thesis would not have been possible.

- First and foremost, Prof. D. Davidson for introducing me to the fascinating field of radio astronomy, and providing me with this opportunity. For all the guidance and advice throughout the project. And, for all the funding provided. I have certainly learnt so much during the course of this thesis, and I am truly grateful for that.
- Prof. J. De Swardt for all the advice and suggestions with regards to the design of the radio receiver front-end.
- Prof. K. Palmer, for the useful discussions regarding the antenna design. Furthermore, for donating the two offset-parabolic reflectors for this project.
- Prof. H. Reader for the helpful suggestions regarding the EMC shielding of the receiver.
- The team at the KAT office, especially Francois Kapp and Jason Manley for reviewing the instrument design, and the useful suggestions they made regarding the use of the ROACH board and the correlator design. Lastly, thanks for inviting me the CASPER Workshop in India, I certainly learnt a lot, and at the same time had a lot of fun.
- Rob Anderson for assisting me with my measurements, and for all the very useful discussions.
- Wessel Croukamp for your expertise in helping me realise my designs, and for the excellent suggestions you gave me in terms of the practical design aspects of my system.
- Reutech, especially Robert Engelbrecht for allowing me to use their NFA.
- To everyone in my research group, for their support and advice.
- My parents, for always providing me with so many opportunities and their constant support and love.
- My brother and sister for their encouragement.
- My friends for taking my mind off work and allowing me to de-stress.
- Lastly, to a very special person in my life, Melissa Colby for always giving me so much love and support.

Table of Contents

List of Figures	ix
List of Tables	xiii
Nomenclature	xiv
1 Introduction	1
1.1 Background	1
1.2 Objectives	2
1.3 Overview	3
1.4 Contribution	4
2 Fundamentals of radio astronomy	5
2.1 Introduction	5
2.2 Reception of astronomical signals	5
2.3 Connection between thermal noise and blackbody radiation	7
2.4 Emission mechanisms and types of cosmic sources	9
2.4.1 Types of radio emissions	9
2.4.2 Emission mechanisms	11
2.5 Identification of some intense radio sources	12
2.6 Choosing an operating band for the radio interferometer	14
2.7 Conclusion	20
3 Principles of radio interferometry and aperture synthesis	21
3.1 Introduction	21
3.2 The van Cittert-Zernike theorem	24
3.3 Coordinate system in synthesis imaging	27
3.4 Baseline coordinates and earth rotation synthesis	29
3.5 Computation of the interferometer baseline length	32
3.6 Response of a two-element interferometer	33
3.6.1 Finite bandwidth effects	35
3.6.2 Frequency conversion effects	36
3.7 Aperture synthesis and imaging	38
3.7.1 Spatial transfer function	38
3.7.2 Array configuration	39

3.7.3	Mapping by discrete transformation	42
3.8	Conclusion	43
4	Characterisation and design of the radio front-end	44
4.1	Principles of radio astronomy receivers	44
4.1.1	Sensitivity of an interferometer	47
4.2	Receiver performance metrics	53
4.2.1	Noise figure	53
4.2.2	Phase stability considerations	54
4.2.3	Sensitivity degradation due to the receiver frequency response	55
4.2.4	Signal transmission system	56
4.3	Linearity	56
4.4	Design details of the receiver front-end	58
4.4.1	Development of the down conversion scheme	59
4.4.2	Behavioural design of the dual-conversion superheterodyne receiver	62
4.4.3	Linear performance of the RF stage	65
4.4.4	Linear performance of the IF Stage	68
4.4.5	Linear performance of the output stage	69
4.4.6	Characterisation of the frequency converters	70
4.4.7	Receiver cascade design performance	78
4.4.7.1	The receiver gain	80
4.4.7.2	The receiver noise figure	81
4.4.7.3	The linearity of the receiver	82
4.4.8	Other receiver specifications	84
4.5	Conclusion	84
5	An antenna suitable for radio interferometry	87
5.1	Introduction	87
5.2	Antenna literature study	87
5.2.1	Overview	87
5.2.2	The pyramidal horn	89
5.2.3	The Yagi-Uda	90
5.2.4	The LPDA	90
5.2.5	The helix	92
5.2.6	The selected antenna	93
5.3	Geometrical features of offset-parabolic reflectors	94
5.4	Analysis and detailed design of the helical beam antenna feed	97
5.4.1	Optimising the pattern characteristics of the helix feed by shaping the ground plane	99
5.4.2	Mechanical support considerations	104
5.4.3	Matching the feed antenna to the offset-parabolic reflector	107
5.4.4	Input impedance matching for the optimised helical antenna	110
5.5	Integrating the helical feed antenna with the offset-parabolic reflector	114
5.5.1	Phase center of the helical beam antenna	115
5.5.2	Setting the feed pointing angle	117

5.5.3	Simulation of the helical beam antenna fed offset-parabolic reflector . . .	117
5.6	Measurement results of the helical beam antenna fed offset-parabolic reflector . .	121
5.6.1	The measured radiation patterns of the helical beam antennas	122
5.6.2	The measured return loss of the helical beam antennas	126
5.6.3	The gain measurement of the offset-parabolic reflector	126
5.7	Conclusion	128
6	Correlator digital back-end	130
6.1	Introduction	130
6.2	Quantisation noise of spectroscopic correlators	133
6.3	Spectroscopic correlator topologies	134
6.3.1	XF correlator	134
6.3.2	FX correlator	135
6.3.3	XF and FX correlator comparison	136
6.4	CASPER tools for developing FX correlators	138
6.4.1	The ROACH FPGA-based processing board	138
6.4.2	The KatADC digitiser	140
6.4.3	Software toolflow	141
6.5	Design of the 256 MHz, 1024-channel FX correlator	142
6.5.1	FX correlator pipeline synchronisation	145
6.5.2	Data logging and supervisory control software	145
6.6	FX correlator testing and validation	148
6.7	Conclusion	153
7	Interferometric observations	154
7.1	Introduction	154
7.2	A more precise RFI characterisation of the 1292-1548 MHz band	154
7.3	First light	158
7.4	Experiments and other observations	158
7.4.1	One-dimensional image of the sun	160
7.4.2	Observations of non-solar sources	161
7.5	Conclusion	162
8	Conclusion	164
8.1	Summary of the work	164
8.2	Observations and findings	165
8.3	Suggestions for future work	166
A	System of coordinates and geometric relationships	167
A.1	Horizon coordinate system	167
A.2	Equatorial coordinates	168
A.3	Conversion from hour angle-declination to azimuth-elevation	170
B	Noise as a random process	172
B.1	The Gaussian PDF	172

B.2	Specification and classification of a random process	173
B.3	PSD of a random process	174
B.4	Ergodic processes	174
C	Hardware details of the dual-conversion superheterodyne receiver	176
C.1	Enclosure	176
C.2	Power requirements and distribution	177
C.3	Details of the power limiting circuit	177
C.4	Coaxial cables	179
D	User guide for the FX correlator	182
D.1	Hardware setup	182
D.2	Software configuration	184
E	Photos	187
	List of References	190

List of Figures

1.1	A simplified functional block diagram of the interferometer system.	3
2.1	Radiation of brightness B from an element of solid angle $d\Omega$, incident on a flat area A	6
2.2	The noise temperature of an antenna.	8
2.3	Typical spectra of continuum sources, for both thermal and non-thermal radiation.	12
2.4	The transparency of the atmosphere at different altitude levels over the electromagnetic spectrum.	16
2.5	Example of an RFI scan performed of the local interference environment from 300 MHz to 2900 MHz.	18
2.6	The local interference environment spectrum measured from 1300 MHz to 1500 MHz.	19
3.1	A schematic of a simple (adding) interferometer.	22
3.2	A schematic of a phase-switching interferometer	23
3.3	Geometry for the van Cittert-Zernike theorem.	25
3.4	Baseline and source position vectors used to define the coordinate system for an interferometric array.	28
3.5	Geometrical relationship between the source under observation and the projected baseline of the interferometric array.	29
3.6	The spherical triangle used to compute the (u, v) -coordinates as a function of the baseline coordinates (h, d) , and the source coordinates (HA, δ)	30
3.7	The (u, v) loci for different baseline azimuth \mathcal{A} and source declination δ	31
3.8	A simplified schematic of a two-element interferometer.	34
3.9	The (u, v) -coverage of the WSRT.	40
3.10	The (u, v) -coverage of the VLA and GMRT.	41
4.1	Block diagram of a single-conversion superheterodyne receiver.	45
4.2	Block diagram of a total power receiver.	46
4.3	Block diagram of a Dicke receiver.	47
4.4	Block diagram of Graham's receiver.	48
4.5	Simplified schematic of a correlation interferometer.	49
4.6	Illustration of a noisy network used to determine the NF of a two-port network.	54
4.7	A simplified block diagram of the interferometer receiver system.	59
4.8	The linear spur plot normalised by the LO for the upconverting stage.	61
4.9	The linear spur plot normalised by the LO for the downconverting stage.	62

4.10	The frequency mapping scheme of the dual-conversion superheterodyne receiver. . .	63
4.11	A simplified block diagram of the dual-conversion superheterodyne receiver architecture.	64
4.12	A block diagram of the input RF stage, $\nu_{\text{RF}} = 1292 - 1548$ MHz	65
4.13	The measured S-parameters of the RF stage from 800 to 2200 MHz.	67
4.14	The low noise front-end topologies considered.	67
4.15	The measured noise figure and gain of the two alternate low-noise front-end topologies.	68
4.16	A block diagram of the intermediate IF stage, $\nu_{\text{IF}} = 3372 - 3628$ MHz.	69
4.17	The measured S-parameters of the IF stage from 2500 to 4500 MHz.	70
4.18	A block diagram of the Output stage, $\nu_0 = 94 - 350$ MHz.	70
4.19	The measured S-parameters of the output stage from 50 to 1000 MHz.	71
4.20	The measured conversion loss L_c of the Mini-Circuits ZX05-73L double-balanced mixer.	72
4.21	The measured return loss of the ZX05-73L upconverting mixer.	73
4.22	The measured return loss of the ZX05-73L downconverting mixer.	73
4.23	The block diagrams of the (a) 4920 MHz local oscillator, and the (b) 3278 MHz local oscillator.	75
4.24	The measured S-parameters of the fixed 4920 MHz LO topology.	76
4.25	The output spectrum of the 4920 MHz local oscillator.	76
4.26	The output spectrum of the 3278 MHz local oscillator.	77
4.27	The phase noise calculation of the local oscillators.	78
4.28	The block diagram for the complete dual-conversion superheterodyne receiver.	78
4.29	The simulated distributed receiver gain, computed using the analytical model.	80
4.30	The measured gain at the receiver output from 94 MHz to 350 MHz.	81
4.31	The simulated cascaded noise figure of the analytical receiver model.	82
4.32	The measured noise figure of the dual-conversion superheterodyne receiver from 94 MHz to 350 MHz.	83
4.33	The measured third-order two-tone intermodulation product of the receiver.	84
4.34	The measured 1 dB compression point of the dual-conversion superheterodyne receiver.	85
4.35	The spectrum of the receiver for an input signal of -80 dBm at 1548 MHz.	85
5.1	The geometry of a pyramidal horn antenna.	89
5.2	The geometry of an LPDA.	91
5.3	The geometry of an offset-parabolic reflector.	95
5.4	The geometry of a uniform (left-hand wound) helix antenna with dimensions.	98
5.5	Far-field patterns of the initial helix design.	100
5.6	A helical antenna above a square, a circular and a cylindrical ground plane.	100
5.7	Initial helix RHCP and LHCP patterns for different ground plane geometries.	101
5.8	The results of the cylindrical ground plane parametric study	103
5.9	Gain comparison between the air- and PVC-core helices.	105
5.10	Far-field pattern comparison between the air- and PVC-core helices.	106
5.11	Simulated radiation patterns of the optimised helical beam antenna for 1292 MHz, 1420 MHz and 1548 MHz.	109

5.12	Simulated axial ratio of the manufactured helical beam antenna from 1292 MHz – 1548 MHz.	110
5.13	Illustration of the triangular copper matching strip.	111
5.14	Simulated input impedance of the optimised helical beam antenna for the non-matched and matched designs.	112
5.15	The return loss variation as the triangular strip matching network dimensions are changed.	113
5.16	Simulated return loss [dB] of the optimised helical beam antenna with the designed matching network.	114
5.17	Geometry for determining the phase center of the helical beam antenna.	116
5.18	Simulated normalised E-field phase of the optimised helical feed antenna.	116
5.19	The secondary patterns of the offset-parabolic reflector at 1420 MHz for the different feed pointing angles.	118
5.20	Simulated radiation patterns of the offset-parabolic reflector fed by the optimised helical beam antenna for 1292 MHz, 1420 MHz and 1548 MHz.	120
5.21	The gain on boresight for the offset-parabolic reflector fed by the optimised helical beam antenna.	121
5.22	The measured and simulated RHCP- and LHCP-patterns in both the (a) $\phi = 0^\circ$ - and (b) 90° -planes, at 1292 MHz.	123
5.23	The measured and simulated RHCP- and LHCP-patterns in both the (a) $\phi = 0^\circ$ - and (b) 90° -planes, at 1420 MHz.	124
5.24	The measured and simulated RHCP- and LHCP-patterns in both the (a) $\phi = 0^\circ$ - and (b) 90° -planes, at 1548 MHz.	125
5.25	The measured and simulated axial ratio of the manufactured helical beam antennas.	126
5.26	The measured and simulated return loss of the manufactured helical beam antenna.	127
5.27	The gain of the measured offset-parabolic reflector fed by the helical beam antenna.	128
5.28	The measured radiation pattern of the offset-parabolic reflector fed by the helical beam antenna.	128
6.1	A simplified diagram of a lag (XF) digital correlator.	135
6.2	A simplified diagram of an FX digital correlator.	137
6.3	A simplified block diagram of the ROACH board.	139
6.4	A simplified schematic of the KatADC.	141
6.5	Comparison of a single channel frequency response of a PFB with a direct FFT.	142
6.6	The system model of the 256 MHz, 1024-channel FX correlator.	144
6.7	The flowchart for the generation of the accumulation sync pulse.	146
6.8	The I/O paths and interfaces between the control computer, PowerPC and FPGA.	147
6.9	The flowchart for the data-logging and supervisory control software.	148
6.10	The computed complex visibility of the software FX correlator model.	149
6.11	The synthesised image of the 1-D source, using the software (Simulink) FX correlator model.	150
6.12	The power spectrum measured for both input channels of the FX correlator from 0 – 400 MHz (Nyquist frequency).	151
6.13	The computed complex visibility of the FX correlator (FPGA).	152

6.14	The synthesised image of the source using the FX correlator (FPGA).	153
7.1	A more precise RFI characterisation of the 1292-1548 MHz band.	157
7.2	Interference fringes of the quiet sun at 1.42 GHz for a baseline of 15 m.	159
7.3	The fringes of the quiet sun at 1.42 GHz for the baselines, $b = 3$ m, $b = 12$ m and $b = 21$ m.	161
7.4	The synthesised image of the quiet sun at 1.42 GHz.	161
7.5	The measured fringes of Taurus A for a baseline of $b = 5$ m.	162
A.1	The horizon system of coordinates.	168
A.2	The equatorial system of coordinates.	169
A.3	Relationship between the celestial coordinates (δ , HA) and the azimuth and elevation (\mathcal{A} , \mathcal{E}) coordinates.	170
C.1	Mechanical drawing of the receiver chassis.	176
C.2	An illustration of the power distribution throughout the receiver chassis.	178
C.3	An illustration of the power distribution within the low-noise front-end.	178
C.4	Block diagram of the power limiting circuit.	179
C.5	The circuit schematic and PCB of the Schmitt trigger.	180
D.1	ROACH board with the labels.	183
E.1	Photo of the helical beam antenna.	187
E.2	Photo of the receiver board.	188
E.3	Photo of the low-noise front-end units.	188
E.4	Photo of the interferometer.	189

List of Tables

2.1	A catalogue of discrete intense sources at 408 MHz and 1420 MHz.	15
2.2	The ITU (WRC-2003) frequency band allocations for radio astronomy between 10 – 3000 MHz, for region 1 (South-Africa)	17
4.1	A summary of the required receiver specifications.	58
4.2	A bill of materials for the dual-conversion superheterodyne receiver.	79
5.1	A summary of the geometrical parameters of the 92 cm offset-parabolic reflector. . .	96
5.2	The dimensions of the manufactured (optimised) helical feed antenna.	107
6.1	Summary of the efficiency factor for different quantisation schemes.	134
6.2	The required specifications for the FX correlator design.	143
7.1	A summary of the interferometer’s system performance specifications.	155
7.2	Instrument parameters of the two-element digital correlation interferometer.	156
C.1	A list of the Schmitt trigger circuit components.	179
C.2	A list of the coaxial cables used throughout the interferometer system.	181

Nomenclature

List of Acronyms

ADC	Analogue-to-Digital Converter
ALMA	Atacama Large Millimeter Array
AR	Axial Ratio
ASIC	Application Specific Integrated Circuit
ASKAP	Australian Square Kilometer Array Pathfinder
ATA	Allen Telescope Array
BEE2	Berkeley Emulation Engine 2
BORPH	Berkeley Operating system for Re-Programmable Hardware
BPF	BandPass Filter
BRAM	Block Random Access Memory
CASPER	Collaboration for Signal Processing and Electronic Research
CMB	Cosmic Microwave Background
COPS	Complex Operations Per Second
COTS	Components Off The Shelf
DFT	Discrete Fourier Transform
DSP	Digital Signal Processing
EI	Edge Illumination
EMBRACE	Electronic Multi Beam Radio Astronomy ConcEpt
FET	Field Effect Transistor
FoV	Field-of-View
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
FPGA	Field Programmable Gate Array
FT	Fourier Transform
FX	DFT Cross-correlating correlator
GaAs	Gallium Arsenide
GbE	Gigabits per second Ethernet
GMRT	Giant Metrewave Radio Telescope
GMT	Greenwich Mean Time

GSM	Global System for Mobile communications
HEMT	High Electron Mobility Transistor
HFET	Heterostructure Field Effect Transistor
HPBW	Half Power BeamWidth
HSLO	High-Side injected Local Oscillator
IBOB	Internet Break-Out-Board
IF	Intermediate-Frequency
IFFT	Inverse Fast Fourier Transform
IIP3	Input third order intercept point
IL	Insertion Loss
IP3	Third order intercept point
ISM	Interstellar Medium
ITU	International Telecommunications Union
KATCP	Karoo Array Telescope Control Protocol
KSP	Key Science Project
LHCP	Left-Hand Circular Polarisation
LNA	Low Noise Amplifier
LO	Local Oscillator
LOFAR	LOW Frequency ARray
LPDA	Log-Periodic Dipole Array
LPF	LowPass Filter
LSLO	Low-Side injected Local Oscillator
LWA	Long Wavelength Array
MbE	Megabits per second Ethernet
MoM	Method of Moments
MSH	Mills, Slee and Hill
MWA	Murchison Wide-field Array
MWO	MicroWave Office
NCP	North Celestial Pole
NF	Noise Figure
NFA	Noise Figure Analyser
OIP3	Output third order intercept point
P1dB	1dB Compression point
PDF	Probability Density Function
PFB	Polyphase FilterBank
PKS	Parkes
PSD	Power Spectral Density

PSF	Point Spread Function
PVC	PolyVinyl Chloride
ROACH	Reconfigurable Open Architecture for Computing Hardware
RF	Radio-Frequency
RHCP	Right-Hand Circular Polarisation
RL	Return Loss
RV	Random Variable
SCP	South Celestial Pole
SMA	SubMiniature version A
RFI	Radio-Frequency Interference
SFU	Solar Flux Unit
SEFD	System Equivalent Flux Density
SIS	Superconductor-Insulator-Superconductor
SKA	Square Kilometre Array
SNR	Signal to Noise Ratio
SPL	Spherical Spreading Loss
SPST	Single Pull Single Throw
TCXO	Temperature Compensated Crystal Oscillator
UTC	Coordinated Universal Time
VHDL	Very high-speed integrated circuit Hardware Description Language
VLA	Very Large Array
VLBA	Very Long Baseline Array
VSWR	Voltage Standing Wave Ratio
WSRT	Westerbork Synthesis Radio Telescope
WSS	Wide Sense Stationary
XF	Cross-correlating DFT correlator
XMAC	Complex Multiplication and Accumulation

Constants

$1\text{Jy} =$	$10^{-26}\text{W} \cdot \text{m}^{-2}\text{Hz}^{-1}$
$h =$	$6.63 \times 10^{-34}\text{J} \cdot \text{s}$
$k =$	$1.38 \times 10^{-23}\text{J/K}$
$c =$	$3 \times 10^8\text{m/s}$
$\omega_e =$	$7.29115 \times 10^{-5}\text{rad/s}$
$T_0 =$	290K

Variables

B	Brightness $[\text{W} \cdot \text{m}^{-2}\text{Hz}^{-1}\text{rad}^{-2}]$
-----	--

P_A	Power received by antenna	[W]
S	Flux density	[W · m ⁻² Hz ⁻¹]
Ω_s	Solid angle subtended by source	[sr]
P_n	Normalised antenna power pattern	[unitless]
S_0	Observed flux density	[W · m ⁻² Hz ⁻¹]
Ω_A	Antenna pattern solid angle	[sr]
Ω_M	Antenna main-lobe solid angle	[sr]
A_{eff}	Antenna effective collecting area	[m ²]
$\Delta\nu$	Frequency bandwidth	[Hz]
w	Thermal noise power per unit bandwidth	[W · Hz ⁻¹]
λ	Free space wavelength	[m]
T_A	Antenna noise temperature	[K]
T_b	Brightness temperature	[K]
RA	Right ascension	[degrees, time]
HA	Hour angle	[degrees, time]
δ	Declination	[degrees]
b	Interferometer baseline	[m]
V	Visibility	[W · m ⁻² Hz ⁻¹ rad ⁻²]
I	Intensity	[W · m ⁻² Hz ⁻¹ rad ⁻²]
\mathcal{E}	Electric field of celestial source	[V · m ⁻¹]
E	Electric field at the observer	[V · m ⁻¹]
A_N	Normalised effective collecting area of the antenna	[unitless]
(u, v, w)	Projected antenna spacing (baseline) coordinates	[λ]
(l, m, n)	Direction cosines with respect to baseline components	[unitless]
s_0	Phase tracking center	[unitless]
s	Direction of a differential element $d\Omega$ of the source	[unitless]
σ	Vector normal to s_0	[degrees]
h	Hour angle for baseline intersection with celestial sphere	[time]
d	Declination for baseline intersection with celestial sphere	[degrees]
τ_g	Geometric delay	[sec]
\mathcal{A}	Baseline azimuth	[degrees]

NOMENCLATURE

\mathcal{E}	Baseline elevation	[degrees]
\mathcal{L}	Baseline latitude	[degrees]
$v_{c,n}(t)$	Voltage signal at correlator input from antenna n	[V]
$r(\tau)$	Correlator output signal for delay τ . .	[V ²]
n_a	Number of antennas	[unitless]
$W(u, v)$	Spatial transfer function	[V ² · m ⁻²]
$g(u, v)$	Weighting function for the spatial transfer function	[unitless]
$b_o(l, m)$	Synthesised beam	[V ² · m ⁻²]
E_a	Aperture electric field	[V · m ⁻¹]
ΔT_{\min}	Minimum detectable noise temperature	[K]
T_{sys}	System noise temperature	[K]
τ_a	Post-detection integration time	[sec]
$\Delta\nu_{\text{IF}}$	Pre-detection receiver bandwidth	[Hz]
T_{rec}	Receiver noise temperature	[K]
K_s	Sensitivity constant of receiver	[unitless]
F	Noise factor	[unitless]
T_e	Equivalent noise temperature	[K]
L_c	Mixer conversion loss	[dB]
Ψ_f	Feed angle	[degrees]
α_h	Helix pitch angle	[degrees]
η_q	Efficiency factor	[unitless]

Chapter 1

Introduction

1.1 Background

The first discovery of radio waves emitted by an extraterrestrial source from the direction of the Galactic center, was made by Jansky at Bell Telephone Laboratories in 1932 [1]. However, this made little scientific impact on the astronomy community at the time. Later, Reber extended Jansky's work by performing a systematic survey of the sky at several frequencies between 160 and 480 MHz [1]. These observations were published in the *Astrophysical Journal* and was therefore much more accessible by the astronomy community.

After, astronomers started to realise the significance of radio observations in their pursuit of studying the universe, instruments with larger collecting areas, finer angular resolutions and higher sensitivities were constructed. Traditionally, single-aperture radio telescopes were the workhorses of early radio astronomy. However, the angular resolution achievable with such an instrument is limited by practical constraints, as the dish size increases. This led to the development of radio interferometry, which enabled astronomers to obtain sufficiently fine angular resolution with the use of much smaller antennas. In, the last few decades the number of operational interferometric arrays have increased. Notably, some of the larger systems are the Very Large Array (VLA), the Westerbork Synthesis Radio Telescope (WSRT), the Australia Telescope, and the MERLIN array [2]. In the mm-band, the CARMA array is an example of a fully operational synthesis array, and the Atacama Large Millimeter Array (ALMA) operates in the sub-mm range [2, 3]. On the opposite side of the spectrum, the Giant Metrewave Radio Telescope (GMRT) operates at meter wavelengths [4].

The proposed Square Kilometre Array (SKA) is the next major step in probing the fundamental questions of the universe. The SKA is an international project to design and build a revolutionary, next-generation radio telescope. As the name suggests, this radio telescope will have a collecting area of 1 square kilometer, which operates over the range of 70 MHz to 30 GHz, or equivalently covers the wavelengths from 4.3 m to 1 cm [5, 6, 7]. This bandwidth cannot be managed by a single antenna technology, and the preliminary draft design for the SKA thus proposes three different frequency bands. For the low frequency regime ($\sim 70 - 300$ MHz), sparse aperture arrays, are likely the most practical solution. There are multiple antenna technologies

being considered for the mid-frequency band (300-1000 MHz). These include dense aperture arrays, parabolic reflectors outfitted with either phased-array feeds or single pixel feeds. The higher operating band ($\sim 1 - 10$ GHz) will almost certainly use reflector antennas with single pixel wideband feeds [5, 8]. The SKA is being designed to address questions of fundamental importance in physics and astrophysics. The SKA's international science working group, have formulated these unanswered questions into 5 Key Science Projects (KSPs). The KSPs range from the strong field tests of gravity; probing the dark ages; the origin and evolution of cosmic magnetism; the cradle of life and the evolution of galaxies [7]. In turn the technical design requirements are derived from these science goals, and it is required that the imaging capabilities of the SKA far exceed the performance of present-day interferometers [7].

Compared to existing radio telescopes, the SKA will have about a 50 times higher sensitivity, and 10,000 times faster survey speed [7]. As mentioned previously, to achieve these sensitivities a collecting area of 1 square kilometer is needed. This large aperture will be implemented with an array of coherently connected antennas that will have a baseline extending up to 3000 km [5]. The provisional distribution of the collecting area for the array configuration, sees 50% of the collecting area being located within a 2.5 km radius from the array center, 25% within the next 180 km and the remainder extends up to the maximum baseline [5]. A project of this scale has inherently, major engineering challenges, this naturally motivates the design of new technologies and novel concepts. This has sparked a lot of interest, and currently 19 countries and about 55 institutions around the world are involved. Thus, major developments are taking place in the fields of antennas, low-noise Radio-Frequency (RF) systems, long distance data transmission, real-time signal processing and highly complex computing systems [6]. This global interest regarding the SKA has spurred the development and construction of many radio telescopes worldwide. These so-called pathfinders are used to test some of the proposed technologies and to provide important technical information. This helps mitigating risk within the complex SKA design. For example, both the South-African and Australian SKA pathfinder projects, are demonstrating reflector antenna technologies for the SKA. However, the MeerKAT project of South-Africa is focused on the large number small diameter reflector antennas scenario, using wide-band single pixel feeds [9], whereas, the Australian SKA Pathfinder (ASKAP) utilises focal plane phased-array feeds. Furthermore ASKAP will demonstrate the FoV enhancements, and the computing/processing requirements for such a large-scale synthesis array [10]. Some of the other technology demonstrators taking place are focused on the longer wavelength regime of the SKA. These include, the LOw Frequency ARray (LOFAR), the Long Wavelength Array (LWA) and the Murchison Wide-field Array (MWA) [8].

1.2 Objectives

The problem addressed in this project is the design, implementation and verification of a two-element digital correlation interferometer as applied to the domain of radio astronomy. The aim was to go full circle, from the conceptualisation of a reference design complete with system specifications, through to the system integration phase, and finally, testing and verifying the instrument; simply put, the end-to-end design of an interferometer. It was decided to implement the interferometer as far as possible using affordable off-the-shelf technologies.

The major engineering challenge addressed in this project was the design of an instrument capable of making professional-type radio astronomy measurements, in a local interference environment with strong sources of radio frequency interference (RFI). In this regard, the investigative component for this project was to verify whether it is possible to achieve a high enough sensitivity to reliably detect cosmic sources other than the sun, where, the presence of man-made interference and cost adversely influences the system.

1.3 Overview

The subsystems that needed to be designed for the interferometer are shown in figure 1.1. The layout of this thesis is logically organised in accordance to the various major subsystems. To formulate a reference design, it is first necessary to cover the fundamentals of radio astronomy, and particularly interferometry. These topics are discussed respectively in chapters 2 and 3. Additionally, the results of the RFI survey performed to evaluate the impact the local interference environment potentially has on the receiver, are presented in chapter 2. Chapter 4 commences with a rigorous derivation for the sensitivity of an interferometer. From this the required sensitivity of the instrument is defined, along with the other global system specifications. The remainder of chapter 4, presents the details of the radio front-end design of the instrument, along with the simulated and measurement results.

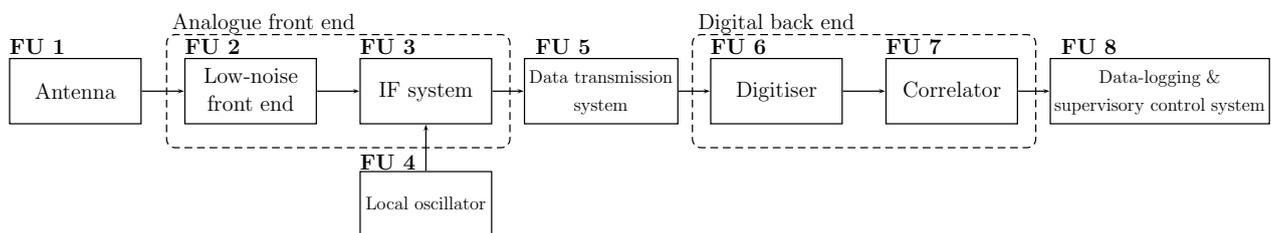


Figure 1.1: A simplified functional block diagram of the interferometer system.

After briefly discussing the potential antenna technologies usable, the remainder of chapter 5 describes the antennas designed for the interferometer; this includes discussing both the measured and simulated results. Chapter 6 then focuses on the digital back-end design: this includes discussing the digitiser; the FPGA-based processing board and software libraries used to realise the correlator design. After integrating the various subsystems, a series of observational experiments were conducted with the two-element digital correlation interferometer - the results are presented in chapter 7. Both the sun and Taurus A were successfully detected; this verifies the end-to-end performance of the interferometer. Finally, chapter 8 concludes this work by presenting a summary of the results and findings obtained during the operation of the instrument. Additionally, the limitations of the system are pointed out, and suggestions are made for future work regarding the improvement of the instrument.

1.4 Contribution

The work presented in this thesis is not novel and does not break new ground; various authors have published papers on their successes of implementing smaller amateur instruments [11, 12, 13]. This being said, the advent of the SKA and its many engineering challenges are rich in research opportunities. This has led to many engineers from all over the world to get involved, and help address the pressing issues facing the SKA. However, much of the research is focused on very specific areas, without regard for the system as a whole. Moreover, most engineers do not necessarily have a good working knowledge of radio astronomy, and more importantly radio interferometry. The intent of this thesis was therefore to develop an experimental interferometer, which can be used by engineering students to develop a deeper and more practical understanding of interferometry, on a system (instrument) level. This instrument further serves the purpose of being a platform for future academic research, where specific aspects of the system can be focused on, and improved.

The contribution unique to this thesis is the RFI characterisation of the local interference environment of the instrument. These results may assist future radio astronomy projects at this site, and again creates additional research opportunities, such as looking into strategies for mitigating RFI in wideband receiver instruments.

Chapter 2

Fundamentals of radio astronomy

2.1 Introduction

The purpose of this chapter is to introduce some of the important concepts found in radio astronomy. More specifically, topics such as the brightness, flux density and antenna temperature will be treated. This is followed by a discussion of the types of emission mechanisms found in radio astronomy, and how this influences the design of a radio telescope. A list of some of the stronger radio sources which are routinely observed by amateur instruments are also presented. Lastly, this chapter concludes with a discussion based around the selection of the operating frequency band for the interferometer.

2.2 Reception of astronomical signals

The front-end of any radio telescope is outfitted with an antenna. It is therefore possible to collect the power radiated by some distance cosmic source in a particular direction, see figure 2.1. In this regard, the most fundamental measure of power received per unit area, per unit solid angle, and per unit bandwidth [$\text{W} \cdot \text{m}^{-2} \text{Hz}^{-1} \text{rad}^{-2}$] in radio astronomy is called the brightness, $B(\theta, \phi)$ [14]. The infinitesimal measure of power dW incident on a surface of area dA from an element of solid angle $d\Omega$ is expressed as

$$dW = B(\theta, \phi) \cos \theta d\Omega dA d\nu, \quad (2.1)$$

where $d\nu$ is the frequency bandwidth. The power received by the antenna P_A is obtained by integrating (2.1) by the entire collecting area A , the receiver bandwidth $\Delta\nu$, from a solid angle Ω of the sky

$$P_A = A \int_{\nu}^{\nu+\Delta\nu} \iint_{\Omega} B(\theta, \phi) \cos \theta d\Omega d\nu. \quad (2.2)$$

Since, the sky brightness may vary with direction, it is useful to quantify the strength of the radio signal received from some portion of the sky. This quantity is known as the spectral flux density S , and for a discrete source it is expressed as [15]

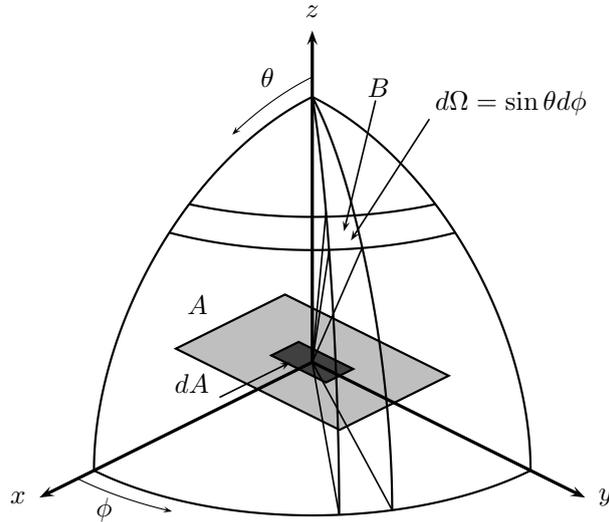


Figure 2.1: Radiation of brightness B from an element of solid angle $d\Omega$, incident on a flat area A .

$$S = \iint_{\Omega_s} B(\theta, \phi) d\Omega, \quad (2.3)$$

where Ω_s is the solid angle subtended by the source. Furthermore, the flux density is measured as $\text{W} \cdot \text{m}^{-2}\text{Hz}^{-1}$, and the unit is the jansky (Jy), where $1 \text{ Jy} = 10^{-26} \text{W} \cdot \text{m}^{-2}\text{Hz}^{-1}$. Practically, all sources have an extent and subtend some finite solid angle. It should thus be obvious that the antenna power pattern $P_n(\theta, \phi)$ acts as a weighting function, and the observed flux density S_0 is less than the true value, and is given by

$$S_0 = \iint_{\Omega_s} B(\theta, \phi) P_n(\theta, \phi) d\Omega. \quad (2.4)$$

Consider the following cases; if the source extent is sufficiently small compared to the antenna pattern solid angle Ω_A , then the observed flux density given by (2.4) is nearly equal to the true flux density S ; and if the source extent is large enough such that the brightness may be regarded as uniform over the antenna main-lobe solid angle, then from (2.4), $S_0 \approx B(\theta, \phi)\Omega_M$ [14]. The power received by the antenna can now be re-expressed in terms of the observed flux density

$$\begin{aligned} P_A &= \frac{1}{2} A_{\text{eff}} \Delta\nu \iint_{\Omega} B(\theta, \phi) P_n(\theta, \phi) d\Omega \\ &= \frac{1}{2} A_{\text{eff}} \Delta\nu S_0, \end{aligned} \quad (2.5)$$

where $\cos \theta$ has been replaced by the antenna power pattern, the area is now the effective collecting area A_{eff} of the antenna. Additionally, it has been assumed that both the brightness

and antenna pattern are uniform with respect to frequency over the bandwidth $\Delta\nu$. Since most cosmic sources are unpolarised, the factor of $1/2$ in (2.5) accounts for the antenna only being able to respond to one polarisation component.

2.3 Connection between thermal noise and blackbody radiation

In order to successfully design a radio telescope, it is important to characterise the signal that is expected to be received from the cosmic source under observation. In this regard, a connection between thermal (Johnson) noise and blackbody radiation will be made.

In his 1928 paper Nyquist presents the theoretical deduction of the electromotive force due to thermal agitation in a simple manner, basing it on the laws of thermodynamics and statistics [16]. Moreover, Nyquist also showed that the mean-square thermal noise voltage generated by a lossy device (resistance R), is indistinguishable from the thermal radiation of a black body. The brightness of a black body radiator is related to physical temperature T and frequency ν by Planck's radiation law

$$B = \frac{2kT\nu^2}{c^2} \left[\frac{\frac{h\nu}{kT}}{e^{h\nu/kT} - 1} \right], \quad (2.6)$$

where $h = 6.626 \times 10^{-34} \text{ J} \cdot \text{s}$ is Planck's constant, $k = 1.38 \times 10^{-23} \text{ J/K}$ is Boltzmann's constant, and $c = 3 \times 10^8 \text{ m/s}$ is the speed of light. The Nyquist formulation considers two conductors of resistance R connected by a non-dissipative transmission line of characteristic impedance R , and held at a temperature T [16]. A more instructive example is considered here, where a simple thermodynamic model of an antenna is used. Consider the antenna depicted in figure 2.2(b). The antenna is matched to the transmission line and is further terminated by a matched resistor R . The antenna is completely surrounded by a black body enclosure, which is held at the same temperature T as the terminating resistor. The thermal radiation emitted by the walls of the enclosure is intercepted by the antenna and transmitted via the transmission line to the resistor, where it is then absorbed. The power per unit bandwidth received by the antenna is given by (2.5) divided by $\Delta\nu$, where the brightness of the black body enclosure is defined by (2.6). Simultaneously, the thermal noise power generated by the resistor travels up the transmission line and is then radiated outwards by the antenna, which is then absorbed by the black body enclosure. The thermal noise power per unit bandwidth of the resistor is defined as [16]

$$w = kT. \quad (2.7)$$

In accordance with the second law of thermodynamics, the thermal noise power generated by the resistor must be equal to the power picked up by the antenna, when pointed to a black body of the same temperature T . This condition is met in the radio wavelength region when moderately high temperatures are considered; this is known as the Rayleigh-Jeans region

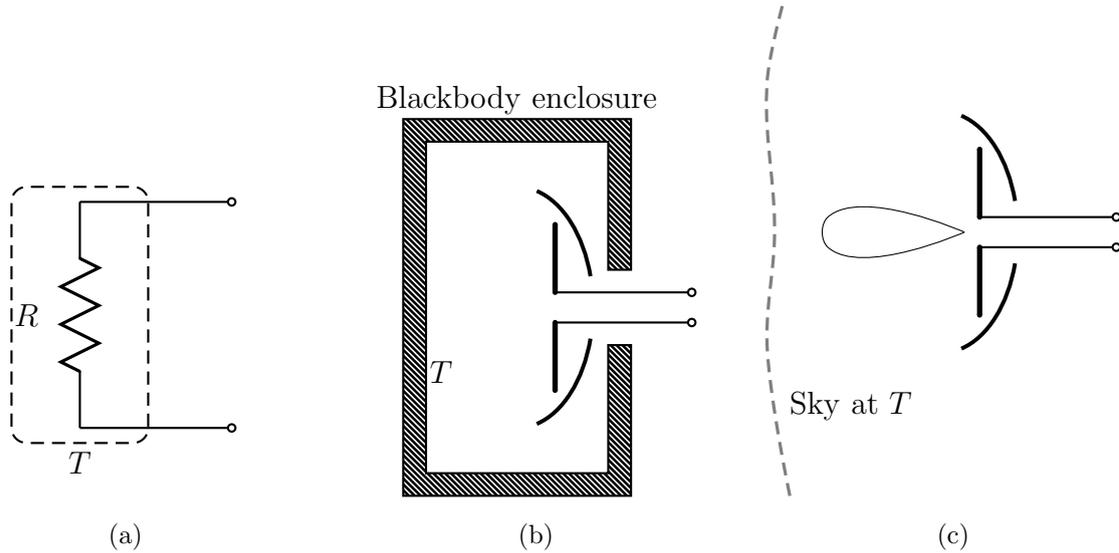


Figure 2.2: (a) Resistor at T ; (b) antenna in a blackbody enclosure at T ; and (c) an antenna observing the sky at a temperature T . In all three cases the noise power available at the terminals are equal.

($h\nu \ll kT$). In this case, the exponential denominator in (2.6) can be approximated by a Taylor series expansion. If only the first two terms of this expansion are significant then

$$e^{h\nu/kT} - 1 \approx \frac{h\nu}{kT}. \quad (2.8)$$

Substituting (2.8) into (2.6) gives

$$B = \frac{2kT}{\lambda^2}, \quad (2.9)$$

where λ is the free space wavelength. This is known as the Rayleigh-Jeans radiation law. Introducing (2.9) into (2.5) yields

$$w = \frac{kT}{\lambda^2} A_{\text{eff}} \iint_{4\pi} P_n(\theta, \phi) d\Omega, \quad (2.10)$$

but $\iint_{4\pi} P_n(\theta, \phi) d\Omega = \Omega_A$ [14] and $A_{\text{eff}}\Omega_A = \lambda^2$, thus (2.10) equals $w = kT$. This is the same as the noise power of the resistor. Thus, the antenna behaves like a resistor which terminates the transmission line. Equivalently, if the enclosure is removed, but the antenna beam observes a region of the sky at T , then all three situations illustrated in figure 2.2 result in the same noise power per unit bandwidth. It is also convenient to describe the radiation resistance of the antenna when observing a region of the sky at temperature T , in terms of the antenna temperature T_A .

If the source being observed does not extend over the entire antenna area, then T_A measured will be less than T of the source. In this case, (2.10) becomes

$$w = \frac{1}{2} A_{\text{eff}} \iint_{\Omega} B(\theta, \phi) P_n(\theta, \phi) d\Omega = kT_A. \quad (2.11)$$

By substituting (2.4) into (2.11), the observed flux density can be evaluated in terms of the antenna noise temperature

$$S_0 = \frac{2kT_A}{A_{\text{eff}}}. \quad (2.12)$$

To conclude, the noise power of a resistor R can be equated to the mean-square noise voltage (assuming a matched load of R) as $\overline{v_n^2} = 4kTR\Delta\nu$. It is obvious that the mean-square noise voltage is fundamentally only a function of the bandwidth, resistance and temperature. Thus, thermal noise is independent of frequency and has a flat power spectral density (white noise). Equivalently, the signals received from a cosmic source, also generally has the form of Gaussian (white) noise.

2.4 Emission mechanisms and types of cosmic sources

A successful observation generally starts by firstly choosing a science question. This is followed by the selection of a cosmic source to observe. Given that the type- and radiative mechanism- of the cosmic source dictate what observational sciences are possible, this directly influences the interferometer design, in particular the array configuration and receiver setup. For example, the instrumentation used for pulsar¹ studies differ from others receivers used in radio astronomy [19]. Pulsars, characteristically emit periodic pulses, and the receiver must therefore be sensitive to short time intensity variations. Additionally, the Interstellar Medium (ISM) causes pulse dispersion and the receiver must account (cancel) for this affect by a technique called dedispersion. Note that this thesis is not concerned with transient or pulsar observations, and will not be discussed further.

This section initially gives an overview of the radio emission types, namely continuum- and spectral line-radiation, and then highlights some of the science studies associated with each. This is followed with a brief discussion based on the emission mechanisms involved in radio astronomy.

2.4.1 Types of radio emissions

The radiation spectrum of cosmic sources can be categorised as either continuum or spectral line radiation [17]. In the former, apart from line emissions of atoms and molecules and some other celestial bodies, it is observed that all other forms of radio emission arise from free electrons. These free electrons have no definite energy jumps, and this results in a continuous spectrum

¹Pulsars are periodic radio sources that are the remnants of supernova explosions. Pulsars are very dense neutron stars, where the magnetic and spin axes are not aligned, this leads to a narrow beam of coherent radiation. Pulsars have emission periods ranging from milliseconds to many seconds. After the first pulsar was discovered in 1967, they have contributed to our understanding of the final stages of stellar evolution, properties of the ISM and the fundamental physical affects of general relativity [17, 18].

with no sharp absorption or emission lines [19].

Most cosmic sources such as quasars, radio galaxies and supernova remnants have continuous spectra. The three principle emission mechanisms responsible for continuous radiation are; black body radiation, free-free emission and non-thermal emission [17]. In [15], a qualitative explanation is given as to why most of the power of cosmic sources are in the form of continuous radiation. In their explanation, the radiation is generated by the synchrotron mechanism, where free high-energy electrons are accelerated in an orbital motion by a magnetic field. The observer therefore sees pulses of radiation from the electrons, where the power of these pulses are concentrated at harmonics of the orbital frequencies. The individual pulses are too numerous to be separable, and thus the electric field appear as a continuous random process with a zero mean. This confirms the initial definition, which stated that cosmic signals appear as Gaussian random noise with a wide, uniform power spectral density. Continuum radiation has contributed to the majority of the cosmic sources discovered, and our current knowledge about their nature [17]. Continuum observations consider the variation of flux density (intensity) as a function of frequency, with a power law distribution of $S \propto f^{-\alpha}$ [20]. The spectral index ² α , describes the spectral slope of a cosmic source, and is used to study the processes responsible for the emission mechanism of sources. Furthermore, the steepness of the spectrum is related to the energy of the electrons. In the case of synchrotron radiation, the spectral slope tends to increase as the most energetic electrons are lost, and astronomers can use this data to predict the age of sources [17].

More importantly, continuum observations place certain design requirements on the receiver. The power (flux) spectrum of continuum sources show a slow variation with frequency, where non-thermal sources generally have a spectral index of ~ 0.8 , but values as low as 0.17, have been derived [20, 17]. This highlights the fact that the flux density may be regarded as constant across the bandwidth of the receiver. Instruments designed for continuum observations exploit this, and utilise the entire instantaneous bandwidth available. This results in higher sensitivity [21].

In contrast, spectral line radiation is emitted at specific frequencies, due to the electrons in atoms and molecules changing state. This causes a photon of energy to be emitted at a wavelength which is characteristic to the atom or molecule [18]. One of the most important and widely observed spectral lines is the neutral (non-ionised) atomic hydrogen line, HI, at 1420 MHz. Numerous studies pertaining to the HI distribution in our galaxy are being made to investigate cold interstellar matter, the kinematics of gas in the galaxies, and to also learn about the gravitational potential of galaxies [17]. Moreover, the 21 cm HI line is being used to search for extra-terrestrial intelligence, since researchers believe this spectral line might be used by extra-terrestrial lifeforms for communication [18]. Other spectral lines of interest is the Deuterium line at 327.3 MHz, used for researching the origin of the universe, and the Hydroxyl (OH) lines at 1612, 1665, 1667 and 1720 MHz for studies focused on the low-density envelopes

²The flux density of a cosmic source is dependent on frequency, and the spectral index, α is a unitless value that quantifies this dependence. Usually, the total flux density versus frequency curves are plotted as log-log graphs of the form $\log S = -\alpha \log f + \text{constant}$.

of molecular clouds. A list of a few important spectral lines are given in [15].

Spectral line instrumentation (especially the back-end) differs substantially from its continuum counterpart. The simplest back-end representation, consists of a large bank of narrow band filters. Each filter defines a spectral channel, and every channel pair has a detector (correlator) [21]. Digital spectrometers are preferred over analogue methods for carrying out spectral-line observations, given their flexibility and stability [21]. It is also mentioned that spectral line correlators can be used for continuum observations. This offer advantages such as RFI mitigation, and to reduce the smearing effect caused by wide bandwidths [15].

2.4.2 Emission mechanisms

During the pioneering stages of radio astronomy, two broad classes of emission mechanisms were distinguished, namely thermal and non-thermal radiation. The physical processes responsible for the generation of these emission mechanisms, and how they influence the interferometer design is briefly discussed.

Any cosmic source that has a physical temperature above absolute zero (0 K) will radiate energy in the form of electromagnetic waves [14]. Thus, the radiative mechanism is due to a thermal process. The quiet sun, the moon, the cosmic microwave background (CMB)³ and the Orion nebula (M42), are all examples of sources that are thermal in nature [19]. One type of thermal emission is free-free radiation. Free-free radiation, which is also commonly known as bremsstrahlung, is a broadband emission due to the acceleration of a free electron in the field of an ion [21]. The Rayleigh-Jeans law given in (2.9), explicitly shows that the brightness, and the temperature of a black body radiator are related. This is an important result as it has become the custom in radio astronomy to measure the brightness temperature⁴, T_b and by using (2.9), the brightness can be computed [19].

Initially, there was no interest in observing the radio sky, since it was believed that thermal bodies were the principal sources of radiation. And, according to Planck's law, the brightness falls off steadily at longer wavelengths (i.e. radio) [21]. Jansky's discovery of radio emissions from the Milky Way was the catalyst of radio astronomy. However, Jansky's observation raised a lot controversy, because the observed intensity was too high to be of thermal origin. In 1947, Townes concluded that Jansky's observation had a brightness temperature of 150,000 K, which is clearly much higher than the 6000 K of the visible sun [21]. Soon after, the distinction was made that this intense temperature was generated by a non-thermal process. Many emission mechanisms are responsible for non-thermal radiation, but in practice synchrotron radiation is generally encountered. The supernova remnant Cassiopeia A and the radio galaxy Cygnus A, are examples of some of the most intense non-thermal sources that are routinely observed in the radio spectrum [21]. The flux density of non-thermal sources are more intense at lower

³The cosmic microwave background (CMB) is the relic radiation from the early universe and is nearly constant at 2.74 K [21].

⁴The brightness temperature can be summarised, as the temperature assigned to an emitter of radiation at a frequency ν , as if it were a black body [21].

frequencies, and tends to decrease as frequency increases. To contrast the spectral behaviour of both thermal and non-thermal radiation, an illustration of the flux density variation as a function of frequency is given in figure 2.3.

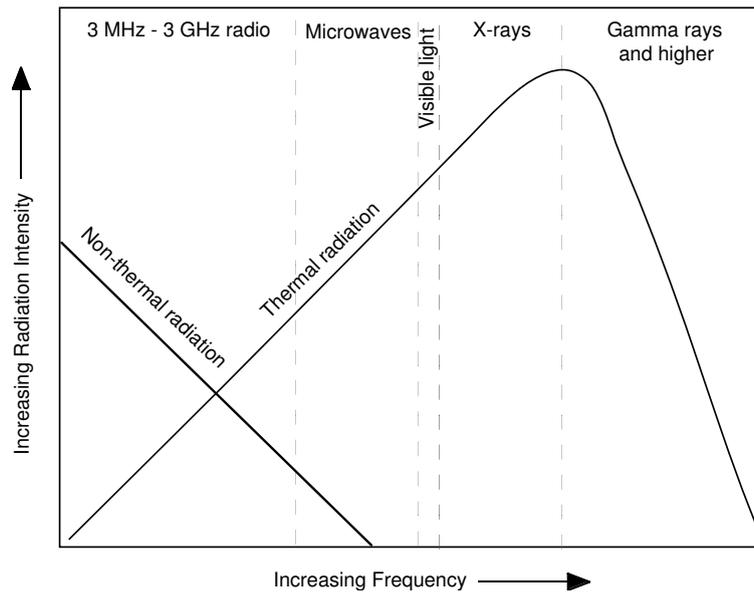


Figure 2.3: Typical spectra of continuum sources, for both thermal and non-thermal radiation, taken from [18].

These flux density trends can influence the system design. If a particular source is of interest, then an appropriate frequency can be chosen, where the flux density of this particular source is more intense. However, the selection of an operating frequency relies on numerous factors, such as the science cases, RFI and instrumentation complexity.

2.5 Identification of some intense radio sources

Large regions of the galaxy are inaccessible to optical wavelengths, since light is heavily obscured by clouds of interstellar gas. However, it is transparent to radio waves. Moreover, ionised-gas regions, such as the corona of the sun, are transparent at very short wavelengths ($< \text{mm}$), but optically thick at radio wavelengths. The picture of the sky is therefore different at radio wavelengths. This led to numerous all-sky surveys being carried out. Lists of intense, discrete radio sources were published as a result of these surveys.

A literature study was undertaken to identify some of the stronger radio sources that are observable from the southern hemisphere. Two important southern hemisphere catalogues are namely, MSH (Mills, Slee and Hill) and PKS (Parkes). They are arguably the southern hemisphere equivalent of the well known 3CR catalogue (Bennett, Cambridge northern sky survey

at 178 MHz, 1963). The PKS catalogues appears as a series of 4 papers (1964-1966), where the survey covered declinations of $-60^\circ \leq \delta \leq -20^\circ$, $-90^\circ \leq \delta \leq -60^\circ$, $0^\circ \leq \delta \leq 20^\circ$ and $-20^\circ \leq \delta \leq 0^\circ$. The survey was performed with the 210-ft reflector antenna at the frequencies of 85, 408, 1410 and 2650 MHz at Parkes, NSW Australia [22]. The intent of this survey was to perform detailed measurements of the position, flux density, spectrum and polarisation of radio sources. The MSH catalogue is also divided according to declination, where 3 papers were published covering declinations of $-20^\circ \leq \delta \leq 10^\circ$ (1958), $-50^\circ \leq \delta \leq -20^\circ$ (1960) and $-80^\circ \leq \delta \leq -50^\circ$ (1961). The Sydney cross-type telescope at a wavelength of 3.5m (~ 85 MHz), was used for this survey

Initially, the MSH and PKS catalogues were used to identify intense discrete sources. The radio sources Virgo A and Centaurus A are identified as being bright, with flux densities of 1050 Jy (178 MHz) and 8700 Jy (85 MHz), respectively [22, 23, 24]. However, some of the most intense radio sources have been observed in the northern sky ($\delta > 20^\circ$), where Cassiopeia A, Cygnus A and Taurus A are some of the brightest. These sources are discussed here for completeness, even though some of them are not observable from the southern sky. To follow is a brief description of each.

Cassiopeia A is the youngest supernova remnant (about 300 years old), and is the brightest radio source in the sky [21]. Cygnus A is radio galaxy with a double-lobe shape, and is one of the strongest radio sources, even at a distance of 170 Mpc [21, 19]. The Crab nebula radio source, Taurus A, is a remnant of a supernova explosion, and has an elliptical shape [19]. Centaurus A is a double radio source (radio galaxy) that has a power output of 10^{35} W. Lastly, Virgo A (M87) is an elliptical radio galaxy, that is famous for a jet that extends from its core. The radio sources discussed are all galactic and extragalactic, and non-thermal in origin. For a more comprehensive discussion of these sources, the interested reader is referred to the many excellent radio astronomy texts available, for instance see [21, 19].

The radio sources with their respective flux densities, as presently discussed, are given in table 2.1. The flux densities at 408 MHz and 1420 MHz are given, and the values are mostly from [20] and [25], respectively. The flux density values given in [25], are referred to outside the atmosphere. An expression is given to translate the galactic flux densities to an earth-bound value; however this is not applied to the densities provided table 2.1. In addition, the right ascension and declination (RA, δ) of the sources given in both the MSH and PKS catalogues, refer to epoch⁵ 1950. Refer to appendix A for a definition of the (RA, δ)-coordinates of a cosmic source. To calculate the position for a specific date, the following corrections can be applied [14]

⁵The equatorial coordinate system (see Appendix A) specifies the sky position (RA, δ) of a radio source, in a relatively fixed manner. However, the gradual procession of the earth's axis around the pole of the ecliptic, causes a slow change in the coordinate of a source. To precisely define a radio source's position, this procession has to be taken into account, by referring to an explicit date known as the epoch. Where, epoch 1950, refers to January 1, 1950.

$$\Delta\text{RA} = m + n \sin \text{RA} \tan \delta \quad \text{per year,} \quad (2.13)$$

$$\Delta\delta = n \cos \text{RA} \quad \text{per year,} \quad (2.14)$$

where (RA, δ) are for epoch 1950, and $m = 3.07420 \text{ sec}$, $n = 1.33589 \text{ sec}$ for (2.13) or $m = 20.0383 \text{ sec of arc}$ for (2.14), are for epoch 2000. The coordinates for a given date is then $\text{RA} + \Delta\text{RA}$ and $\delta + \Delta\delta$. The source positions given in table 2.1, are corrected for epoch 2000 (1 January 2000), using (2.13) and (2.14). The angular dimensions for the sources in table 2.1, are taken from [26].

To conclude this discussion, it is certainly important to investigate the strongest discrete source experienced on earth, the sun. The sun emits both thermal and non-thermal radiation. At optical wavelengths, the solar photosphere of the sun is visible, with a surface brightness of 5770 K ($\approx 6000 \text{ K}$) [21, 18]. The visible spectrum of the sun is approximately that of a black body radiator. But, observing the sun at longer wavelengths, the corona of the sun is optically thick and appears larger and brighter, with temperatures ranging well above several million K [19]. To get an estimation of the brightness temperature of the corona emission, the quiet sun (no disturbances) was measured in [27], between $1.4 \text{ GHz} \leq \nu \leq 18 \text{ GHz}$. From this data a brightness temperature model was developed, where the corona temperature has a wavelength dependence of $\nu^{-2.1}$ and is expressed as

$$T_B = A\nu^{-2.1} + T_{\text{chrom}}, \quad (2.15)$$

where the constant $A = 140,077 \text{ K}$, $T_{\text{chrom}} = 10,880 \text{ K}$ is a correction for the sun's chromosphere, and ν is expressed in GHz. The flux density values given for the sun in table 2.1 are computed using (2.15). The brightness temperatures at 408 MHz and 1420 MHz were computed as $9.31 \times 10^5 \text{ K}$ and $7.79 \times 10^4 \text{ K}$, respectively. The average brightness B_{avg} is then calculated using the Rayleigh-Jeans approximation of (2.9). To finally obtain the values given in table 2.1, it is assumed that the solid angle subtended by the sun, $\Omega_s = 5.981 \times 10^{-5} \text{ sr}$, is sufficiently small compared to Ω_A .

From the flux density data of the sources presented in table 2.1, a specification for the minimum sensitivity of the interferometer can be defined. It is seen that Virgo A has the lowest flux density of $\sim 200 \text{ Jy}$ at 1420 MHz . For a measurement to be discernible from the system noise, the sensitivity should at least be 5σ better [21]. In this case the minimum sensitivity must then be 40 Jy .

2.6 Choosing an operating band for the radio interferometer

Astronomy at radio frequencies on earth is possible because of the transparency of the earth's atmosphere at these wavelengths. This is known as the radio window and extends from a lower

Source	RA (2000.0)	Dec(2000.0)	Angular dimensions	S_{408} [Jy]	S_{1420} [Jy]	Catalogue Ref. number
Sun	-	-	$30'(0.5^\circ)$	284, 343 ⁽¹⁾	286, 320 ⁽¹⁾	N/A
Cassiopeia A *	$23^{\text{h}}23^{\text{m}}24^{\text{s}}$	$+58^\circ 48.9'$	$3' \times 4'$	5500	2397	3C-461
Cygnus A *	$19^{\text{h}}59^{\text{m}}28.3^{\text{s}}$	$+40^\circ 44'0.2''$	$2'10'' \times 35''$	5500	1621	3C-405
Taurus A *	$05^{\text{h}}34^{\text{m}}31.9^{\text{s}}$	$+22^\circ 00'52''$	$3\frac{1}{2}' \times 5\frac{1}{2}'$	1250	893	NGC 1952
Virgo A *	$12^{\text{h}}30^{\text{m}}49.4^{\text{s}}$	$+12^\circ 23'28''$	$2\frac{1}{2}' \times 5'$	494	208	PKS-1228+12
Centaurus A	$13^{\text{h}}22^{\text{m}}30^{\text{s}}$	$-42^\circ 46'$	$3' \times 6\frac{1}{2}'$	2740 ⁽²⁾	1000	PKS-1322-42, MHS-13-42

⁽¹⁾ H.Zürin (1990), ⁽²⁾ J.G. Robertson (1972).

Table 2.1: A catalogue of discrete sources at 408 MHz and 1420 MHz, that are observable from the southern sky. Note that any source with a * suffix, are northern sky sources and not necessarily observable from the southern sky, but included here for completeness.

frequency limit of ~ 15 MHz to about 1.5 THz [19]. The radio window is seen in figure 2.4, but it is also shown that the opacity of the atmosphere is dependent on frequency and altitude. It is apparent from figure 2.4 that there are regions within the radio window, where radio waves are absorbed. For instance, at the H_2O vapour band at 22 GHz and the O_2 lines at 60 GHz. At these less transparent atmospheric frequencies, the sky noise temperature T_{sky} that the antenna sees increases. In [14] it is shown that T_{sky} monotonically decreases as frequency increases and then at about 1 GHz, T_{sky} levels out to a constant value until about 10 GHz. The sky noise temperature at 1420 MHz when the antenna is pointed to zenith is about 15 K, this includes the radiation from the CMB and the atmosphere [28]. Whereas, at 408 MHz, the sky noise temperature is closer to 40 K [4]. Furthermore, it is seen that at zenith the sky noise temperature is a minimum and at the horizon it is a maximum. This suggests that T_{sky} decreases with elevation angle. It is evident from the above discussion that there is a wide frequency band that can be considered for radio astronomy. However, there is the unfortunate issue of RFI, and it is important to consider choosing an operating frequency band, that has been allocated to radio astronomy.

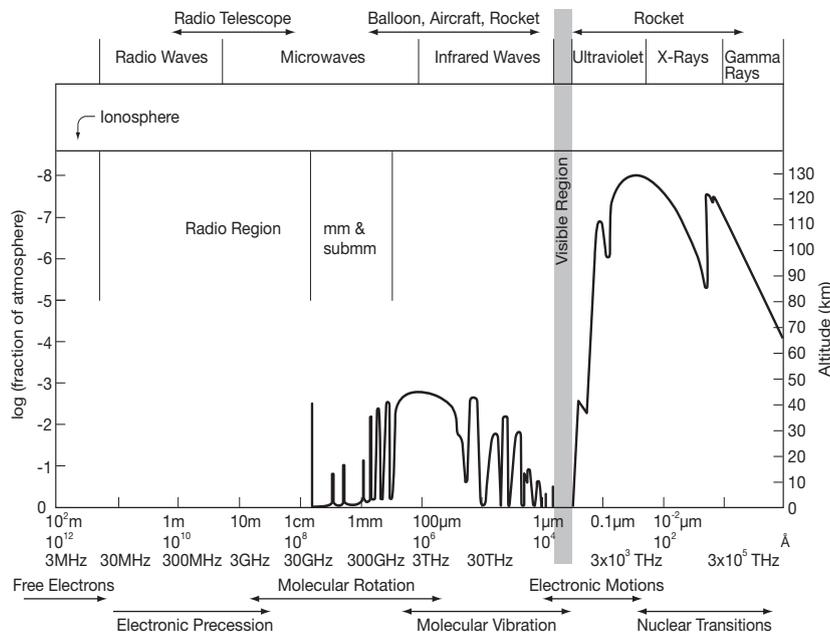


Figure 2.4: The transparency of the atmosphere at different altitude levels over the electromagnetic spectrum, taken from [2].

The International Telecommunications Union (ITU) is responsible for the global management of the radio spectrum and also the protection of radio astronomy. According to the ITU, South-Africa falls into region 1. The radio astronomy bands allocated for region 1 was tabulated from a 2008 ICASA draft of radio frequency allocations in South-Africa [29], and the ITU summary given in [17]. The table is presented in figure 2.2. Also shown in this table is the scientific observations that are possible at the different bands. The 406.1-410 MHz and the

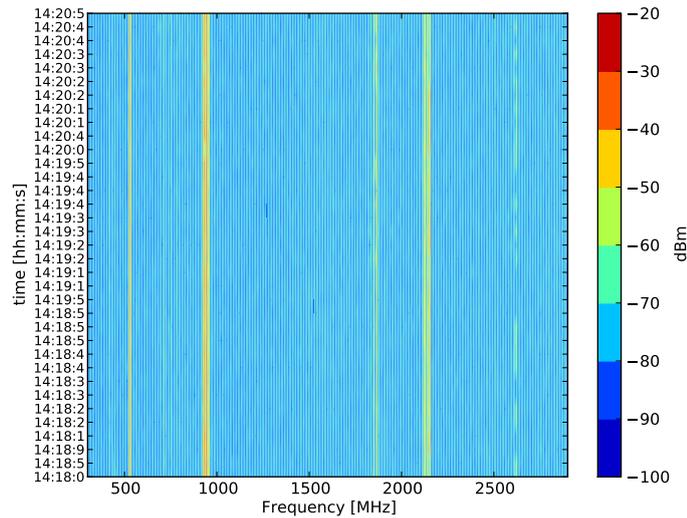
Band [MHz]	Services	Scientific observations
Frequency allocations below 100 MHz		
13.36 – 13.41	RAS ¹ , FS ²	Sun (Solar radio bursts), Jupiter (Radio bursts), ISM, Steep-spectrum continuum sources.
25.55 – 25.67	RAS ¹ , FS ²	Sun (Solar radio bursts), Jupiter (Radio bursts), ISM, Steep-spectrum continuum sources.
73.0 – 4.60	RAS, FS, MS ³	Sun (Solar radio bursts), Jupiter (Radio bursts), ISM, Steep-spectrum continuum sources
Frequency allocations between 100 MHz and 1000 MHz		
150.05 – 153.0	RAS (Region1), FS, MS	Sun, ISM, Steep-spectrum continuum sources, pulsars.
322.0 – 328.6	RAS, FS, MS	Deuterium (Spectral line, 327.384 MHz), Sun, ISM, pulsars.
406.1 – 410.0	RAS, FS, MS	Sun, ISM, Steep-spectrum continuum sources, pulsars.
608.0 – 614.0	RAS (Region 1 & 3), MSS ⁴	Sun, ISM, Steep-spectrum continuum sources, pulsars.
Frequency allocations between 1000 MHz and 3000 MHz		
1400.0 – 1427.0	RAS, EESS ⁵ (passive), SRS ⁶ (passive)	Sun, galactic and local extragalactic HI (Spectral line, 1420.406 MHz), continuum.
1610.6 – 1613.8	RAS, MSS, AeRNS ⁷ , RDSS ⁸	Sun, extragalactic OH (Hydroxyl), continuum.
1668.4 – 1670.0	RAS, MSS, AeRNS, RDSS	Sun, extragalactic OH (Hydroxyl), continuum.
2690.0 – 2700.0	RAS, EESS(passive), SRS(passive)	

¹RAS - Radio Astronomy Service, ²FS - Fixed Service, ³MS - Mobile Service, ⁴MSS - Mobile Satellite Service, ⁵EESS- Earth Exploration-Satellite Service, ⁶SRS - Space Research Service, ⁷AeRNS - Aeronautical Radionavigation Service, ⁸RDSS - Radio determination Satellite Service.

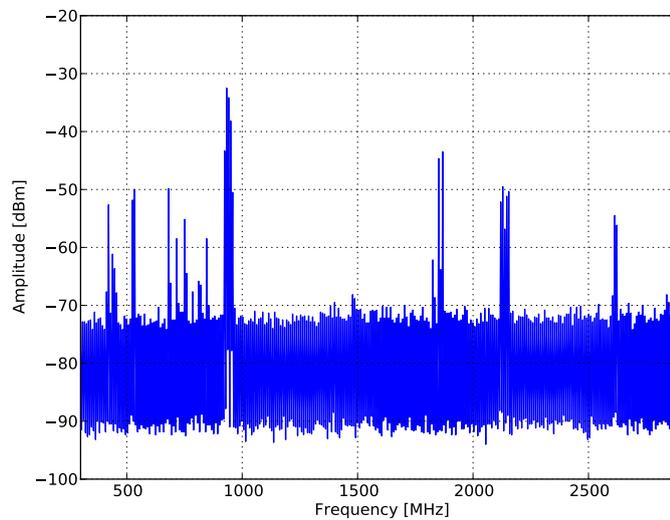
Table 2.2: The ITU (WRC-2003) frequency band allocations for radio astronomy between 10 – 3000 MHz, for region 1 (South-Africa). The table also elaborates on the scientific uses of each band. Also listed are the other radio services occupying this band.

1400-1427 MHz bands, are of interest. At the 408 MHz band, studies of continuum sources as routinely done. Such observations are also possible at 1420 MHz, given the broadband nature of most radio sources. In addition, hydrogen line studies are possible at 1420 MHz (and lower depending on red shift). It is evident from table 2.2 that continuum observations are possible at all the frequencies bands listed. Therefore, this table was accompanied by an RFI survey of the local interference environment, in order to decide on a operating frequency.

The RFI survey was performed using a log-periodic dipole array (LPDA), that operated over a band of 300 – 6000 MHz, and a spectrum analyser set to maximum hold. The measurement routine consisted of capturing the spectral data in all four quadrants (N,E,S and W) for both horizontal and vertical polarisations. An example of one measurement is shown in figure 2.5. This measurement was made with the LPDA directed North and horizontally orientated.



(a)



(b)

Figure 2.5: Example of an RFI scan performed of the local interference environment from 300 MHz to 2900 MHz, with the LPDA horizontally polarised and pointed to the North. Shown is the (a) spectra measured against time, and (b) a snapshot in time of one of the measurements.

From figure 2.5(b) it is clear that there is a significant amount of interference in the 300 to 2900 MHz frequency band. There are strong levels of RFI from 920 MHz to 960 MHz, the maximum power levels measured in this band were close to -30 dBm. According to [29], this is the lower Global System for Mobile communications (GSM) band. As expected there is also a lot of activity at the upper GSM band of 1805-1880 MHz, with power levels as high as -42 dBm. At 2100-2155 MHz, the RFI levels are close to -49 dBm. It is also no surprise that majority of the band from about 400 MHz to 800 MHz is cluttered with RFI, as this portion of the band is used for terrestrial TV broadcasting.

From this initial survey the operating bands that must be avoided could easily be identified. There was however, a portion of the band seen around 1400 MHz that showed very little activity. A more sensitive measurement was performed around this portion of the band, and the result is shown in figure 2.6. The noise floor of the spectrum analyser was pushed down to -110 dBm. As figure 2.6 shows, there seems to be no prominent spikes which could signify potential RFI. There does however seem to be a slight rise in the noise at about 1450 MHz, but it was not clear whether this was because of man-made interference or not. According to table 2.2, the 1400-1427 MHz band is allocated for radio astronomy.

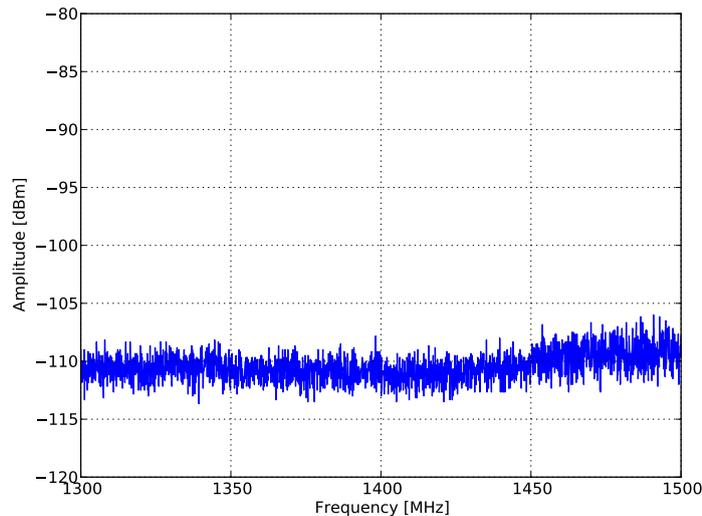


Figure 2.6: The local interference environment spectrum measured from 1300 MHz to 1500 MHz, with the LPDA vertically polarised and directed towards the North. This was a more sensitive measurement of the RFI environment around 1400 MHz, with a much finer resolution bandwidth.

From the information presented in this section, it was decided that the most suitable operating frequency is centered around 1420 MHz. Furthermore, it was decided early on in the thesis that the aim of the interferometer is to detect some of the more routinely observed continuum radio sources. It is then possible to use a wider bandwidth to improve the sensitivity, which makes it easier to detect weaker sources. In this regard, from the RFI survey it was seen that a band

of more than 250 MHz, centered around 1420 MHz, appeared free from any RFI. Thus, it was decided to use an instantaneous bandwidth of 256 MHz.

2.7 Conclusion

In conclusion, initially the emphasis of this chapter was to cover some of the fundamental quantities and equations used in radio astronomy. This was followed by a discussion on the types of cosmic sources and their emission mechanisms. To summarise, both sources with continuous- or line-spectra are found. Their emission mechanisms can either be thermal or non-thermal. In this thesis only broadband emitters are of interest. This has certain advantages. Firstly, a wider receiver passband can be utilised to improve the interferometer's sensitivity. Secondly, given the flat power spectral density of these sources, the observing frequency of the instrument is not restricted to a specific frequency. It is therefore possible to select an operating band that is free of any strong RFI. This was followed by an investigation to identify some of the stronger radio source that are observable with amateur instruments; these are listed in table 2.1. Finally, an RFI survey of the operating environment of the interferometer was performed. This was done in order to identify a radio-quiet portion of the band ($\text{cm} \leq \lambda \leq \text{m}$). A 256 MHz band centered around 1420 MHz was selected.

Chapter 3

Principles of radio interferometry and aperture synthesis

3.1 Introduction

The early radio interferometry observations made by Ryle, Bolton, Stanley and others were motivated by accurately determining the position and angular size of radio sources. This required instruments with high angular resolutions. According to the diffraction limit the angular resolution of a radio telescope is approximately λ/D radians. Considering the radio spectrum, λ is in the meter to millimeter range. Therefore, dishes with excessively large apertures are required to achieve sub-arcsecond resolution. Such dishes are impractical and this led to the use of interferometry, where the aperture D is increased by using smaller antennas spaced many wavelengths apart. The use of interferometry in astronomy was not a new concept. The Michelson optical interferometer was used in the 1890s to measure the diameters of some of the larger and nearer stars [15]. Ryle and Vonberg in 1946 made the first astronomical observation with a two-element interferometer, at 175 MHz with a variable baseline of 17 to 240 m [15]. However, it was the analogue Lloyd's mirror interferometer, an instrument formed by an antenna on a seaside cliff and its image below the water, that really highlighted the importance of radio interferometry in astronomy [30]. This was as a result of the observation of Cygnus A by Bolton and Stanley. It was the first positive evidence of the existence of discrete non-solar radio sources. Other applications of interferometers soon followed, such as full sky surveys, where the source catalogues composed using the Mills cross and the Cambridge interferometer are well known. Other uses include mapping of the sun, looking at the intensity profile of sources and spectral line interferometry. This naturally led to the development of new technology for more sensitive instruments.

The radio interferometer used by Ryle in 1946 was of the adding type, also known as a simple interferometer [15, 14]. As the name suggests, the signal voltages from the antennas are combined additively. This output voltage is then passed through a square-law detector, resulting in a power signal. An illustration of a simple adding interferometer is shown in figure 3.1. For a simple interferometer the detector output, assuming a monochromatic source is

$$F = v^2 \left[1 + \cos \left(\frac{2\pi\nu b \sin \theta}{c} \right) \right],$$

where v^2 is the power after the square-law detector, and b the baseline length. The cosine term represents the Fourier component of the source brightness to which the interferometer responds. The simple interferometer also has a constant additive term that consists of the thermal galactic background, receiver thermal noise and the noise picked-up from the ground through the antenna sidelobes. However, the problem with the simple interferometer is that any gain instability leads to random fluctuations on the recorder, and this can potentially make the detection of weak sources impossible. In 1951, Ryle proposed a new type of radio interferometer known as a phase-switched (multiplying) interferometer, which overcame this problem [31].

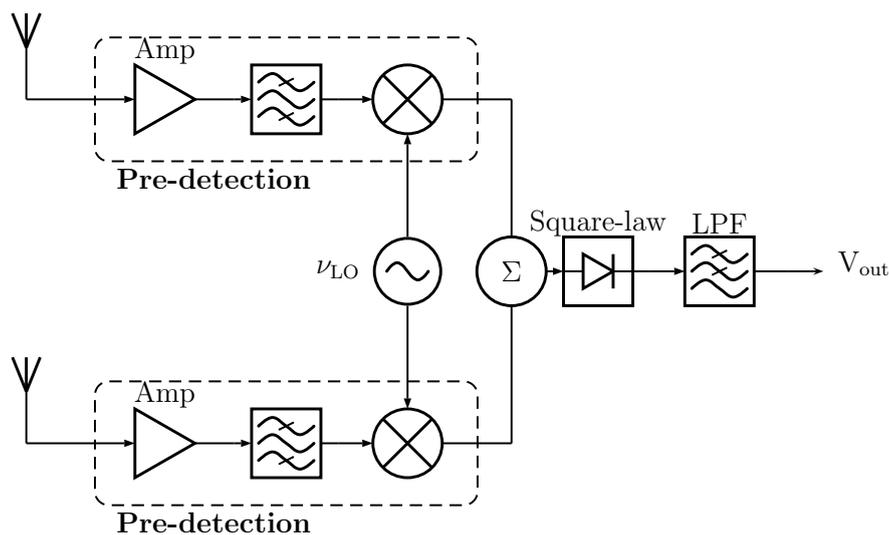


Figure 3.1: A schematic of a simple (adding) interferometer.

The phase-switched interferometer has the advantage of discriminating against both the background and receiver noise [31, 30]. The phase-switched interferometer is shown in figure 3.2. It is seen that the interferometer uses a method whereby an additional $\lambda/2$ length of transmission line is switched into the path of one antenna. This displaces the interference pattern, so that the maximum of the in-phase pattern corresponds to the minimum of the anti-phase pattern. This section of transmission line is rapidly switched in and out of the path, and a synchronous detector takes the difference of the two signals. The output is then seen to be the time average of the product of the two signals, with a zero average [31, 15, 14]. This is essentially a cross-correlation operation of the two signals. Another very important advantage of the phase-switched interferometer is that pre-amplification can be used [31]. The correlating action of the interferometer means that it is insensitive to any gain variation. Therefore, longer baselines can be achieved with the use of pre-amplification. This made it possible to detect weaker sources.

Most modern implementations of this method uses a multiplier-integrator circuit known as a correlator, and this is further implemented digitally. This is discussed in chapter 6.

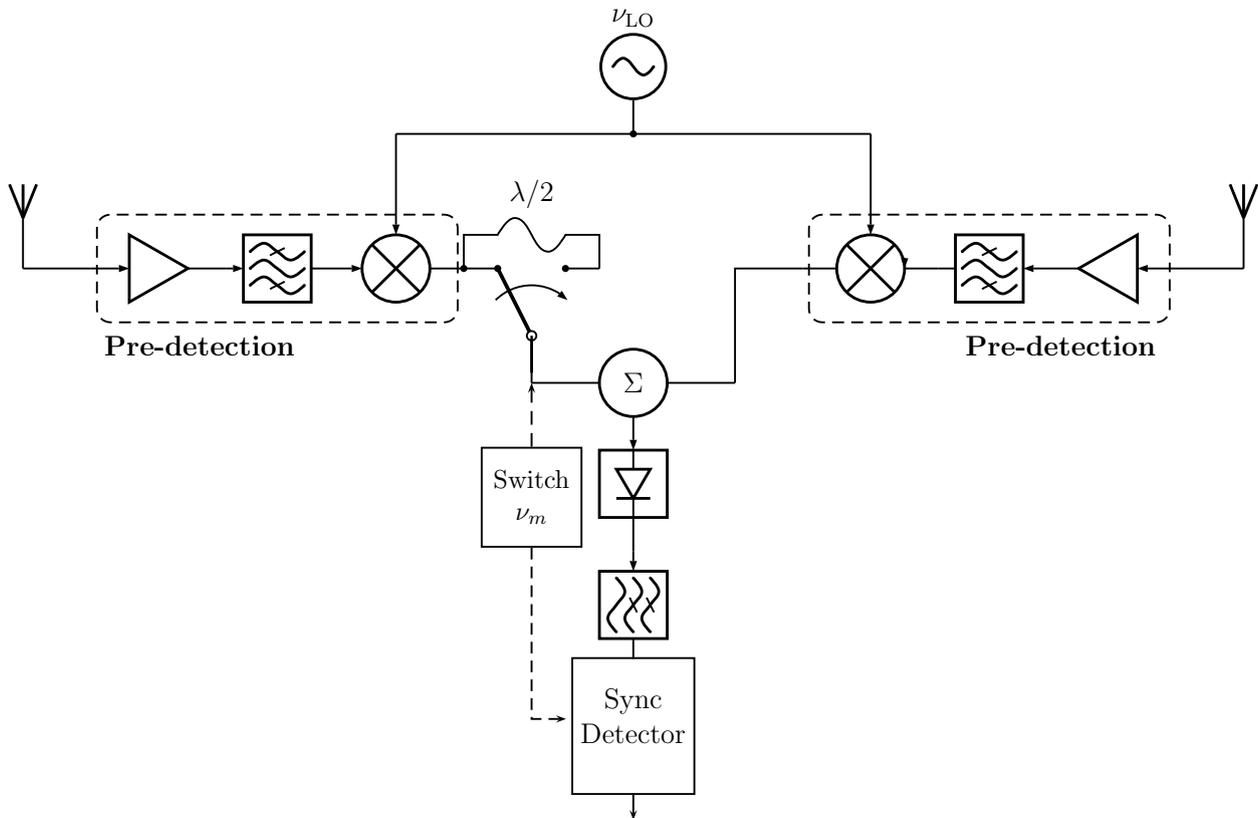


Figure 3.2: A schematic of a phase-switching interferometer. The input signal from one antenna is periodically reversed in phase, by switching in an additional $\lambda/2$ section of transmission line.

Radio interferometry is perfectly suited for mapping the sky, given the high angular resolutions achievable. This makes it possible to image the finer structural details of a source. To image a source with high fidelity requires measurements with many different baseline lengths. Each measurement for a given baseline corresponds to one Fourier component of the radio image [4]. Ryle showed that the rotation of the earth offered different projected baselines to the source. The combination of many such projected baselines effectively synthesise a large aperture [32, 4]. This process is known as earth rotation synthesis and this was a very important step in synthesis mapping. Ryle was awarded the Noble prize in 1974 for his pioneering work in this regard [32]. The Cambridge 1-mile telescope exploited the earth rotation synthesis technique, and the images made of Cassiopeia A and Cygnus A showed great structural detail [15, 14]. This led to the construction of many synthesis arrays.

Several modern synthesis arrays worth mentioning are the VLA, the WSRT, the GMRT, the ALMA and the SKA. The VLA is a synthesis array that operates in the wavelength range of

$\sim 1 - 20$ cm, and has been in existence since the 1980s [33, 32]. The VLA has a two-dimensional configuration with 27 antennas in a “Y”-configuration, each arm is 21 km in length with 9 antennas on each [33]. The maximum baseline spacing achievable is 36.4 km and this gives an angular resolution of $< 0.6''$. The GMRT came into operation in 1998 and is another synthesis array with a “Y”-configuration, and consists of 30 45 m antennas [32, 4]. There are 12 dishes randomly distributed in a central region of 1 km across and the remaining antennas are arranged in the shape of an irregular “Y”. The length of each arm is approximately 14 km, and the maximum baseline length between the extreme antenna arms is about 25 km [4]. Where, the GMRT really differs from the VLA is with respect to the operating wavelength range. The GMRT is a low frequency synthesis array, and covers a frequency range of 50 to 1450 MHz [32, 4]. In contrast, the WSRT is an example of a one-dimensional synthesis array where the antennas are arranged in an east-west baseline. This array consists of 40 25 m antennas, where 10 of the antennas are fixed [19, 34]. ALMA is an instrument that operates in the mm and sub-mm band. This is an international project among North America, Europe and East Asia [3]. The array is proposed to consist of 64 antennas, each 12 m in diameter. ALMA is expected operate over the range from 31 to 900 GHz [3]. For a brief summary of the SKA see chapter 1.

The importance of radio interferometry has now been highlighted. This chapter focuses on the fundamentals of interferometry, where the concept of aperture synthesis is introduced as a technique to achieve high angular resolutions.

The van Cittert-Zernike theorem is firstly presented with the intent of presenting a physical explanation of radio interferometry. This is followed by the analysis of a two-element interferometer, since large synthesis arrays are simply an ensemble of two-element interferometers. In conclusion, the theory of mapping the sky brightness distribution is discussed.

3.2 The van Cittert-Zernike theorem

Assume that the source is distant and that it may be approximated as a brightness (intensity) distribution on a conceived celestial sphere, see figure 3.3. The brightness distribution may further be viewed as a collection or ensemble of point sources. Therefore, any arbitrary point source at location \mathbf{R} ¹, causes a time-varying electric field $\mathcal{E}(t, \mathbf{R})$ to propagate. The receiving antenna pair of the interferometer is spatially separated by some baseline, and the field $\mathbf{E}(t, \mathbf{r})$ is received at two different locations, namely \mathbf{r}_1 and \mathbf{r}_2 . This suggests that the intensity distribution is somehow related to the fields at these two observation points. Stated more formally, the van Cittert-Zernike theorem relates the spatial coherence function $V(\mathbf{r}_1, \mathbf{r}_2)$ to the distribution of intensity $I(\mathbf{R})$, where \mathbf{R} is the vector in the direction of the source under observation [19]. The spatial coherence function, also known as the complex visibility, is simply the correlation of the field at the observation points (\mathbf{r}_1 and \mathbf{r}_2), and is defined as

$$V(\mathbf{r}_1, \mathbf{r}_2) = \langle \mathbf{E}(\mathbf{r}_1) \mathbf{E}^*(\mathbf{r}_2) \rangle, \quad (3.1)$$

¹A bold face variable \mathbf{x} , represents a vector.

where $\langle \cdot \rangle$ denotes the expectation, and the asterisk indicates the complex conjugate. An interferometer is thus a device that measures this spatial coherence function [35]. To follow is a simplified proof of the van Cittert-Zernike theorem.

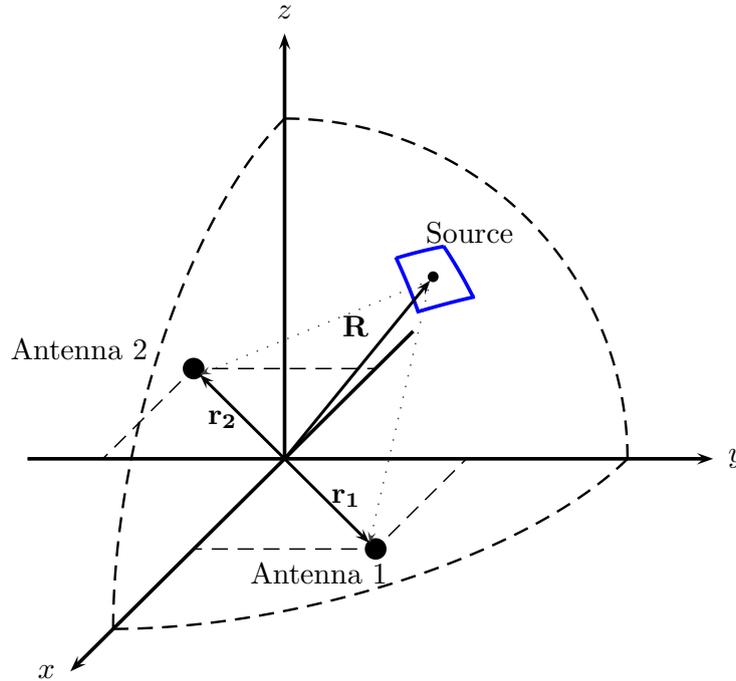


Figure 3.3: Geometry for the van Cittert-Zernike theorem.

Firstly, the field at the observation point is defined. As previously mentioned the brightness distribution may be viewed as a collection of point sources. The field at the observation point \mathbf{r} is then simply the superposition of the fields produced by these point sources. This is possible because of the linearity imposed by Maxwell's equations. Further assuming that the brightness distribution is sufficiently far away, and that there are no additional sources of radiation within the celestial sphere. Then using Huygen's principle the field at the observation point is [35]

$$E(\mathbf{r}) = \int \mathcal{E}(\mathbf{R}) \frac{e^{j2\pi\nu|\mathbf{R}-\mathbf{r}|/c}}{|\mathbf{R}-\mathbf{r}|} dS. \quad (3.2)$$

Note that (3.2) is simplified by ignoring the time dependence, since the quasi-monochromatic components of the electric field are considered [35]. Moreover, the electric field is treated as a scalar quantity. This simplifies the analysis without obscuring the physical meaning [35]. The field given by (3.2), can be substituted into (3.1), in order to compute the spatial coherence function. However, before this is done an additional assumption is made, where the emission from the source is assumed to be spatially incoherent. This can be explained by considering the superposition of two waves generated by the point sources at locations \mathbf{R}_1 and \mathbf{R}_2 . At the

observation points \mathbf{r}_1 and \mathbf{r}_2 , these waves can locally be described as plane waves and the total fields are then simply

$$E(\mathbf{r}) = \mathcal{E}(\mathbf{R}_1) + \mathcal{E}(\mathbf{R}_2). \quad (3.3)$$

Then the spatial coherence function of (3.3) is given as

$$\begin{aligned} \langle E(\mathbf{r}_1)E^*(\mathbf{r}_2) \rangle &= \langle [\mathcal{E}(\mathbf{R}_1) + \mathcal{E}(\mathbf{R}_2)] [\mathcal{E}^*(\mathbf{R}_1) + \mathcal{E}^*(\mathbf{R}_2)] \rangle \\ &= \langle \mathcal{E}(\mathbf{R}_1)\mathcal{E}^*(\mathbf{R}_1) + \mathcal{E}(\mathbf{R}_1)\mathcal{E}^*(\mathbf{R}_2) + \mathcal{E}^*(\mathbf{R}_1)\mathcal{E}(\mathbf{R}_2) + \mathcal{E}(\mathbf{R}_2)\mathcal{E}^*(\mathbf{R}_2) \rangle \\ &= \langle |\mathcal{E}(\mathbf{R}_1)|^2 \rangle + \langle |\mathcal{E}(\mathbf{R}_2)|^2 \rangle, \end{aligned} \quad (3.4)$$

where $\langle \mathcal{E}(\mathbf{R}_1)\mathcal{E}^*(\mathbf{R}_2) \rangle = \langle \mathcal{E}^*(\mathbf{R}_1) + \mathcal{E}(\mathbf{R}_2) \rangle = 0$. The spatial coherence function of the brightness distribution is thus

$$\begin{aligned} V(\mathbf{r}_1, \mathbf{r}_2) &= \int \langle \mathcal{E}(\mathbf{R}_1)\mathcal{E}^*(\mathbf{R}_2) \rangle \frac{e^{-j2\pi\nu|\mathbf{R}_1-\mathbf{r}_1|/c}}{|\mathbf{R}_1-\mathbf{r}_1|} \frac{e^{j2\pi\nu|\mathbf{R}_2-\mathbf{r}_2|/c}}{|\mathbf{R}_2-\mathbf{r}_2|} dS_1 dS_2 \\ &= \int \langle |\mathcal{E}(\mathbf{R})|^2 \rangle |\mathbf{R}|^2 \frac{e^{-j2\pi\nu|\mathbf{R}-\mathbf{r}_1|/c}}{|\mathbf{R}-\mathbf{r}_1|} \frac{e^{j2\pi\nu|\mathbf{R}-\mathbf{r}_2|/c}}{|\mathbf{R}-\mathbf{r}_2|} dS_1 dS_2 \\ &= \int I(\mathbf{R}) \frac{e^{-j2\pi\nu|\mathbf{R}-\mathbf{r}_1|/c}}{|\mathbf{R}-\mathbf{r}_1|} \frac{e^{j2\pi\nu|\mathbf{R}-\mathbf{r}_2|/c}}{|\mathbf{R}-\mathbf{r}_2|} dS \end{aligned} \quad (3.5)$$

where $I(\mathbf{R})$ is the intensity distribution at point $\mathbf{R} = \mathbf{R}_1 = \mathbf{R}_2$ and $dS = dS_1 = dS_2$. Given that the source is a great distance from the observation point, the following approximations apply; $|\mathbf{R}-\mathbf{r}_1| = |\mathbf{R}-\mathbf{r}_2| \approx |\mathbf{R}|$ and $|\mathbf{R}-\mathbf{s}\cdot\mathbf{r}_1|$, $|\mathbf{R}-\mathbf{s}\cdot\mathbf{r}_2|$ for the amplitude and phase respectively, where \mathbf{s} is a unit vector in the direction of the source. These approximations are seen to be very similar to the far-field approximations used in antenna theory. Furthermore, from the definition of a steradian the differential solid angle on the surface of a sphere is defined as $d\Omega = dS/R^2$. Using the above approximations and replacing dS , results in

$$V(\mathbf{r}_1, \mathbf{r}_2) = \int I(\mathbf{s}) e^{-j2\pi\nu\mathbf{s}\cdot(\mathbf{r}_1-\mathbf{r}_2)/c} d\Omega. \quad (3.6)$$

To cast (3.6) in a more useful form, the coordinate system introduced in section 3.3 is used. It is worth noting that (3.6) is dependent on the difference $\mathbf{r}_1 - \mathbf{r}_2$, which is the baseline vector \mathbf{b}_λ defined in section 3.3. This allows (3.14) to be inserted into (3.6), thus

$$V(u, v, w) = \iint I(l, m) e^{-j2\pi(u l + v m + w [\sqrt{1-l^2-m^2}-1])} \frac{dl dm}{\sqrt{1-l^2-m^2}}. \quad (3.7)$$

This is the fundamental relationship in radio interferometry, where the visibility is related to the intensity distribution, and this is known as the van Cittert-Zernike theorem in optics. The intensity distribution in (3.7) is not a function of n , since only the surface brightness can be measured. Practically the interferometer elements are finite antennas that are sensitive to radiation from some arrival direction. There is then an additional factor of $A_N(l, m)$ in (3.7). Where, $A_N(l, m)$ is the normalised effective collecting area of the antenna, and (3.7) becomes

$$V(u, v, w) = \iint A_N(l, m) I(l, m) e^{-j2\pi(ul+vm+w[\sqrt{1-l^2-m^2}-1])} \frac{dl dm}{\sqrt{1-l^2-m^2}}. \quad (3.8)$$

It is easily seen that (3.8) resembles a Fourier transform, where (u, v, w) are now spatial frequencies. There are two conditions under which this equation reduces to the exact definition of a two-dimensional Fourier transform.

Firstly, for a coplanar baseline-vector (east-west array). If the w -axis is chosen to lie in the direction of the celestial pole, then $w = 0$ [35]. In other words for an observation confined to the $u - v$ plane, equation (3.8) then reduces to

$$V(u, v, 0) = V(u, v) = \iint \frac{A_N(l, m) I(l, m)}{\sqrt{1-l^2-m^2}} e^{-j2\pi(ul+vm)} dl dm, \quad (3.9)$$

and this is seen to represent a two-dimensional Fourier transform. The second scenario that reduces (3.8) to a Fourier transform is when the baseline does not remain coplanar. This is the general case and the visibility function has a three-dimensional distribution (u, v, w) . However, if the angular extent of the synthesised field is confined to be small, then the term $w [\sqrt{1-l^2-m^2}-1] \approx -\frac{1}{2}(l^2+m^2)w \approx 0$. Thus, (3.7) becomes

$$V(u, v, 0) \approx V(u, v) = \iint A_N(l, m) I(l, m) e^{-j2\pi(ul+vm)} dl dm. \quad (3.10)$$

The approximation made in (3.10) is only valid for small field imaging, because this introduces a phase error. A common criterion used to restrict this phase error is that it may not exceed 0.1 rad [15], this requires that

$$\theta_f < \frac{1}{3} \sqrt{\theta_b}, \quad (3.11)$$

where θ_f is the width of the synthesised field, and θ_b is the synthesised beamwidth of the antenna array. In [15], a more thorough discussion regarding this phase error is given.

To summarise, if a measurement is confined to a plane, then according to the van Cittert-Zernike theorem the visibility function $V(u, v)$ is related to the distribution of intensity by

$$V(u, v) \rightleftharpoons I(l, m), \quad (3.12)$$

where \rightleftharpoons is a Fourier transformation pair.

3.3 Coordinate system in synthesis imaging

To image of an area of the sky, there is a convenient coordinate system that results in an expression for the correlator response, that is in the familiar form of a Fourier transform. Firstly, a center (origin) of the field being mapped is specified. This reference position is indicated by the unit vector \mathbf{s}_0 , and is aptly called the phase-tracking center, since most observations involve the antennas tracking the source, and this position serves as the phase reference. The direction

\mathbf{s} of a differential element $d\Omega$ of the source, relative to the phase-tracking center is then defined as

$$\mathbf{s} = \mathbf{s}_0 + \sigma, \quad (3.13)$$

where σ is a vector normal to \mathbf{s}_0 , see figure 3.4.

It is now possible to introduce a coordinate system that expresses the source position \mathbf{s} , in a form suitable for mapping the brightness distribution. Firstly, the visibility function introduced in section 3.2 is measured in baseline coordinates. The baseline vector \mathbf{b}_λ is measured in wavelengths, and may be expressed in a right-handed rectilinear coordinate system, with components (u, v, w) [21]. The w axis points towards the phase tracking center \mathbf{s}_0 , and in a plane normal to this direction, the u and v axes are projected towards the East and North, respectively. The u, v -plane thus represents the antenna spacing as seen from the phase reference position, and w is in effect the geometrical delay τ_g between the antennas [33], see figure 3.5. Next, the coordinates on the sky must be defined. It is assumed that the radio source is distributed on the surface of a celestial sphere, and that the phase-tracking center is at the origin of the intensity distribution. Then the coordinates of σ is defined as (l, m) , which are the direction cosines measured with respect to the u and v axes, this is shown in figure 3.5. Using these coordinates, the following expressions are derived

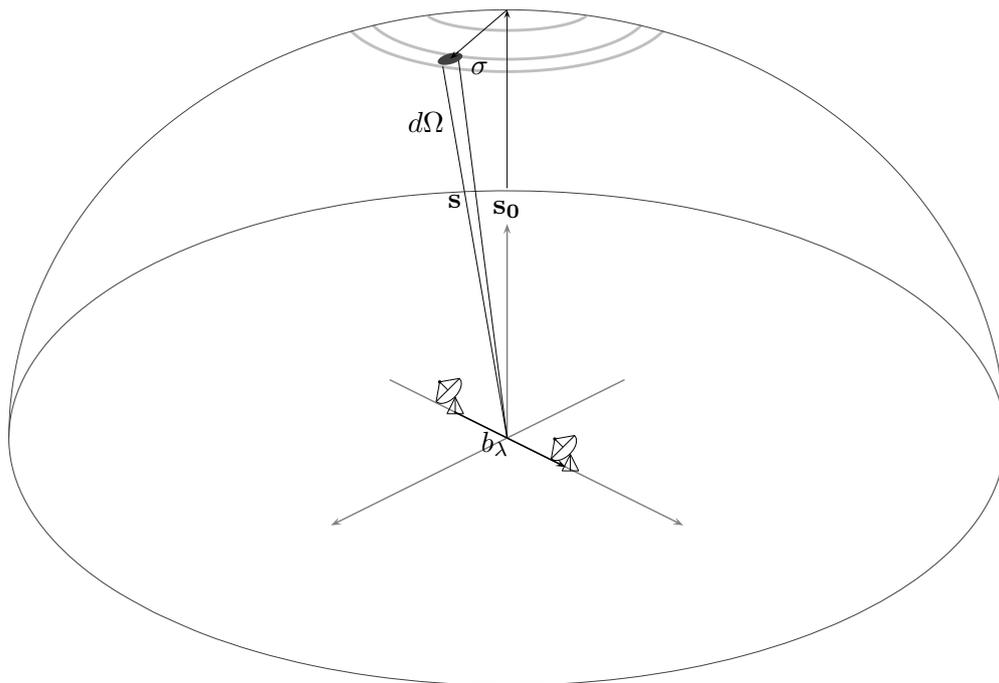


Figure 3.4: Baseline and source position vectors used to define the coordinate system for an interferometric array.

$$\mathbf{b}_\lambda \cdot \mathbf{s}_0 = w, \quad (3.14a)$$

$$\mathbf{b}_\lambda \cdot \mathbf{s} = \left(ul + vm + w\sqrt{1 - l^2 - m^2} \right), \quad n = \sqrt{1 - l^2 - m^2} \quad (3.14b)$$

$$d\Omega = \frac{dl dm}{\sqrt{1 - l^2 - m^2}}. \quad (3.14c)$$

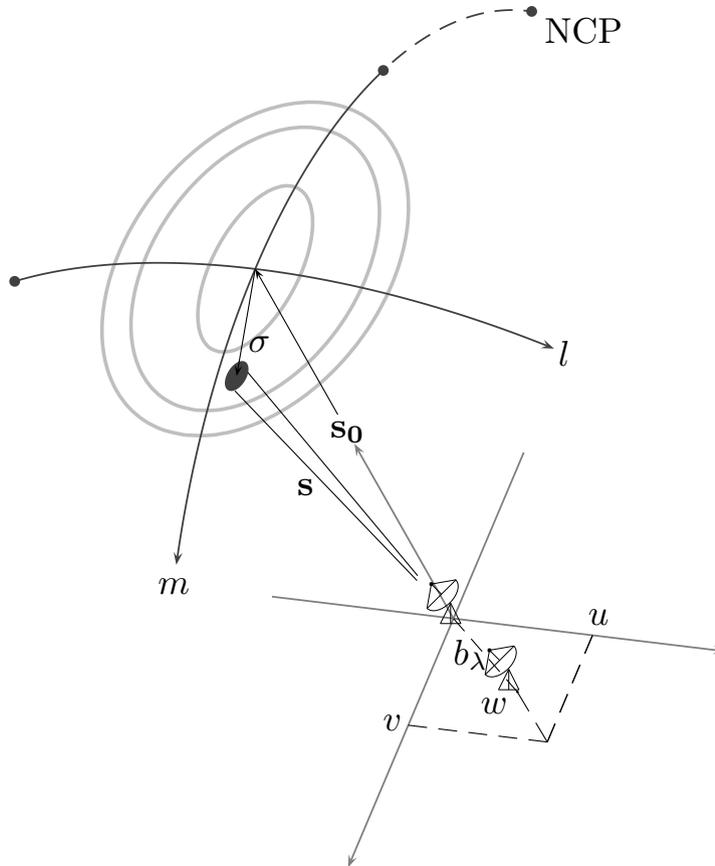


Figure 3.5: Geometrical relationship between the source under observation and the projected baseline of the interferometric array.

The coordinate system presented here was shown in section 3.2 to cast (3.6) in a convenient form, where the synthesised image of the source, is simply the inverse Fourier transform of the measured visibility.

3.4 Baseline coordinates and earth rotation synthesis

Apart from specifying the coordinates for pointing the aerial beam to a celestial object when making interferometric observations, it is also required that the relative position of the antennas

in the array be specified. Equatorial coordinates are mostly used to specify the position of an object. A useful method is to specify the direction of the baseline in this system of coordinates as well, as proposed in [36]. Refer to appendix A for a review of the coordinate systems commonly used in radio astronomy.

The baseline direction is extended to intersect the celestial sphere as shown in figure 3.6. The position of this intersection is then represented by the hour angle and declination (h, d) . The source position S in figure 3.6, is further given by (HA, δ) - refer to appendix A for a definition of hour angle (HA). Using the spherical triangle SNB , it can be shown that the (u, v) plane coordinates are given by [36]

$$u = b_\lambda \cos d \sin(HA - h), \quad (3.15a)$$

$$v = b_\lambda [\sin d \cos \delta - \cos d \sin \delta \cos(HA - h)]. \quad (3.15b)$$

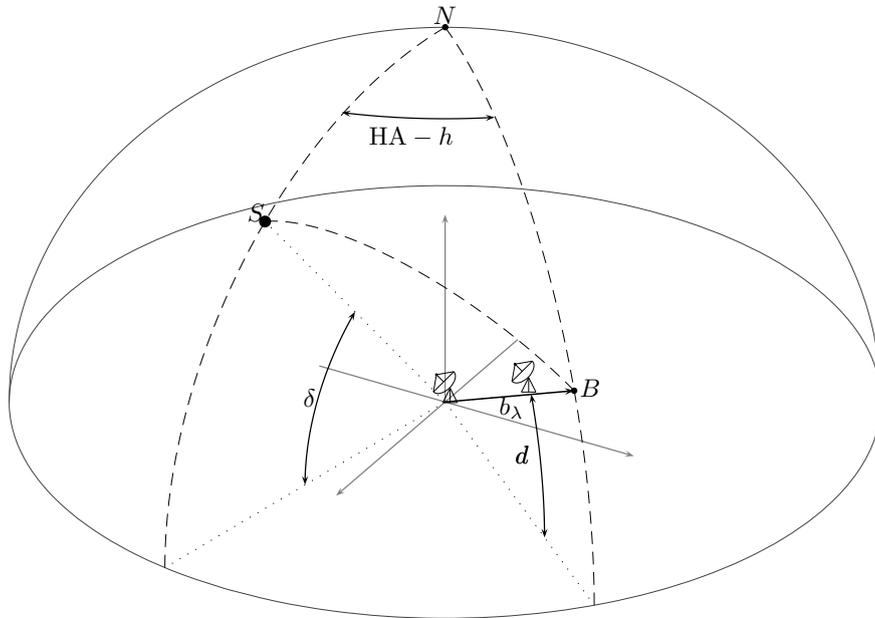


Figure 3.6: The spherical triangle used to compute the (u, v) -coordinates as a function of the baseline coordinates (h, d) , and the source coordinates (HA, δ) .

Two deductions can be made from (3.15). Firstly, the u (East-West) component of the projected baseline depends only on the hour angle of the source and not on the declination δ . And more importantly, during an observation, it can be shown that the locus of the projected antenna components (u, v) is an ellipse, with hour angle as the variable. In the (u, v) plane the ellipse can be expressed as a parametric equation

$$u = a \sin \phi, \quad (3.16)$$

$$v = b \cos \phi, \quad (3.17)$$

with semi-major axis $a = b_\lambda \cos d$, semi-minor axis $b = b_\lambda \cos d \sin \delta$, and parameter $\phi = (\text{HA} - h)$. Furthermore, the ellipse is centered at $(u, v) = (0, b_\lambda \sin d \cos \delta)$. Note that for a coplanar (East-West) baseline, $d = 0$. The ellipse is then centered at the origin. The arc of an ellipse is traced out during an observation, and this process fills the $u - v$ plane as the earth rotates. This is known as earth-rotation synthesis. The locus of the ellipse is dependent on the location (longitude and latitude) of each end of the interferometer baseline, the source declination and the range of HA covered [36, 15]. Figure 3.7 is a plot of the (u, v) loci for different baseline azimuth \mathcal{A} and source declination δ . In figure 3.7, it is seen that for $\mathcal{A} = 90^\circ$, the locus of the ellipse is centered at the origin. Furthermore, for higher declinations the $u - v$ locus has a much more circular ellipticity. It is therefore possible to make high fidelity images with an east-west (one-dimensional) array [33]. If the declination δ of the source is lower (closer to the equator), then the axis of the ellipse in the v -direction becomes narrower and eventually it becomes a straight line for sources on the equator. This results in poor (u, v) coverage and a two-dimensional antenna array with various azimuthal directions are needed for improving the (u, v) coverage. At any given instant a pair of antennas measure the visibility values at two points, $V(-u, -v) = V^*(u, v)$. This is because the brightness distribution is a real function [33]. To follow is a discussion of how to compute the equatorial coordinates (h, d) of the baseline.

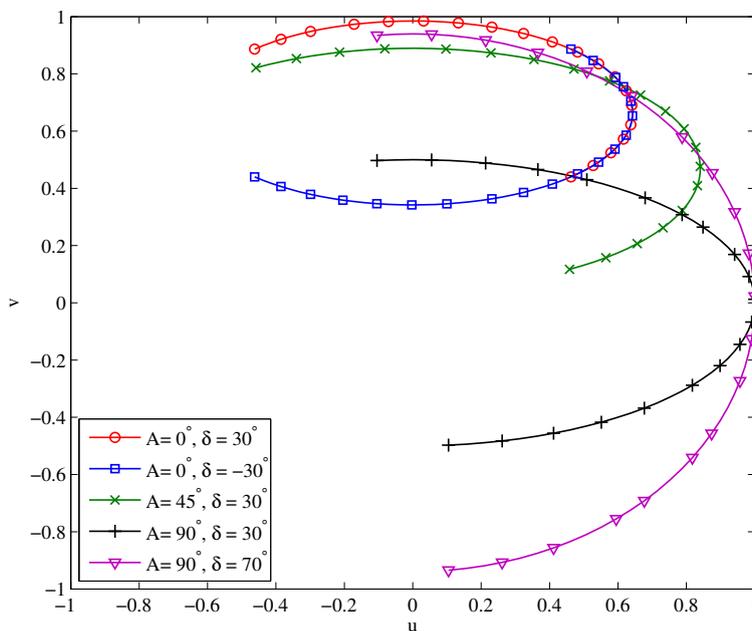


Figure 3.7: The (u, v) loci for different baseline azimuth \mathcal{A} and source declination δ . In all cases it is assumed that $\mathcal{E} = 0^\circ$, $\mathcal{L} = 40^\circ$. The HA range is further chosen to be from $-6h$ to $+6h$. The baseline length was normalised.

To find the coordinate values of the baseline (h, d) , the elevation \mathcal{E} and azimuth \mathcal{A} of the interferometer baseline must first be determined. This is usually established using field surveying

techniques. From $(\mathcal{E}, \mathcal{A})$ the baseline intersection coordinates with the celestial sphere (h, d) , can be determined using the following conversions [15]

$$\sin d = \sin \mathcal{L} \sin \mathcal{E} + \cos \mathcal{L} \cos \mathcal{E} \cos \mathcal{A}, \quad (3.18a)$$

$$\cos d \cos h = \cos \mathcal{L} \sin \mathcal{E} - \sin \mathcal{L} \cos \mathcal{E} \cos \mathcal{A}, \quad (3.18b)$$

$$\cos d \sin h = -\cos \mathcal{E} \sin \mathcal{A}, \quad (3.18c)$$

where \mathcal{L} is the latitude of the baseline. Figure A.3 in appendix A presents a graphical illustration of the relationship between (HA, δ) and $(\mathcal{E}, \mathcal{A})$. It is important to note that S in figure A.3, is now the end point of the baseline. The correctness of (3.18) is validated by computing (h, d) for an azimuth of 80.083° and elevation of -0.483° . From (3.18a) the declination is found to be 5.5167° , and the hour angle is computed as -98.267° , by solving (3.18b) and (3.18c) simultaneously. This is the same result as given in [36].

3.5 Computation of the interferometer baseline length

It should be mentioned early on that the intent of this thesis is not to form images of the sky. Mapping is a specialised subject on its own. However, interesting science is still possible by observing the fringes of a two-element interferometer. As pointed out in the introduction the early use of interferometry in radio astronomy was aimed at accurately determining the position of sources and their sizes. The principal method used to determine a source's position by Ryle and others in the late 1940s, consisted of measuring the time of transit of the central fringe, using an east-west baseline. This gave the RA of the source. The measurement of the source declination relied on the recording of the maximum fringe frequency, this frequency is given as

$$\frac{dw}{dt} = -\omega_e u \cos \delta, \quad (3.19)$$

where $\omega_e = \frac{dH}{dt} = 7.29115 \times 10^{-5}$ rad/s is the rotational velocity of earth. Equation (3.19) shows that if the baseline length u and the fringe frequency is known, then it is a simple matter to calculate the declination.

The simplest procedure used to compute the angular width of a source is based on measuring the fringe amplitude. If the source extent is comparable or larger than the fringe width then the fringe amplitude is attenuated. This phenomenon is easily explained by considering an extended source. A plane wave originating from a given point on the source is assumed to arrive in-phase at the two antennas, however the plane wave from an adjacent point will arrive slightly out of phase as compared to the initial plane wave. This clearly results in a decrease in the fringe amplitude recorded. The more extended the source is, the lower the fringe amplitude will be. If the fringe amplitude drops to zero then the source has been completely resolved out.

This criterion is used to specify the maximum baseline for the interferometer. To compute the approximate baseline needed to resolve the sources given in table 2.1, it is assumed that

these sources are symmetrical with uniform intensity. To further simplify the derivation, a one-dimensional source is assumed, with $l = \sin \theta$ and the projected baseline u . For a source small in extent, $\Delta\theta$, the small-angle approximations can be used, resulting in $l = \sin \theta \approx \theta$, and $u = \frac{b \cos \theta_0}{\lambda}$, with respect to the phase reference position θ_0 . It is well known from Fourier theory that a symmetrical function consists only of a cosine Fourier series, therefore the visibility expression in (3.10) becomes

$$V(u) = \int_{-\Delta\theta/2}^{\Delta\theta/2} A_N(\theta) I(\theta) \cos(2\pi u \theta) d\theta. \quad (3.20)$$

If the antenna pattern solid angle Ω_A is much wider than the solid angle of the source, Ω_s then the observed flux density is equal to $B(\theta, \phi)\Omega_s$. This simplifies to $I(\theta)\Delta\theta$, where $I = -B$ has been used. Equation 3.20 is then found to

$$V(u) = |V| \frac{\sin 2\pi u \Delta\theta/2}{2\pi u \Delta\theta/2}. \quad (3.21)$$

In (3.21) it is seen that the fringe amplitude approaches zero when $2\pi u \Delta\theta/2 \rightarrow \pi$. From table 2.1, the sun has an angular size $\Delta\theta$ of $0.5^\circ (30')$, so the minimum baseline needed to resolve the sun is $b = \frac{\lambda}{\Delta\theta} \approx 24$ m for $\nu = 1.42$ GHz. The baselines needed to resolve out Cassiopeia A and Cygnus A are 181.3 m and 362.7 m, respectively. These baselines are too long and impractical for a small instrument. So the maximum baseline designed for is 24 m.

3.6 Response of a two-element interferometer

It is necessary to analyse the two-element interferometer problem. Since, most synthesis arrays are essentially an ensemble of two-element interferometers. The analysis that follows will use a quasi-monochromatic approximation ($\Delta\nu/\nu \ll 1$). The reason for this has to do with the bandwidth pattern of the interferometer. The bandwidth pattern must ideally be much wider than the antenna pattern, else it is shown in [30] that the correlator output cannot be interpreted as a Fourier integral of the brightness distribution. This is because the integrand is then a function of both (l, m) and (u, v) . Moreover, it is assumed that the source is completely incoherent. In the analysis, it is also important to consider effects such as the bandwidth and frequency conversion of the receiver, in order to correctly interpret the interferometer output.

Consider the elementary two-element interferometer shown in figure 3.8. The antennas are separated by a baseline b , and the receivers are denoted by $H_1(\nu)$ and $H_2(\nu)$. The receivers are to provide any amplification, filtering and frequency conversion needed. The incident field $E(l, m)$ from a distant point source, induces a voltage at the output of the antenna terminals. The signals then pass through the respective receivers, and are denoted by $v_{c,1}(t)$ and $v_{c,2}(t)$ before the correlator. The multiplying- and averaging- circuits in figure 3.8, are collectively called the correlator. The output of the correlator is the cross-correlation of the voltages $v_{c,1}(t)$ and $v_{c,2}(t)$

$$r(\tau_g) = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T v_{c,1}(t) v_{c,2}^*(t - \tau_g) dt, \quad (3.22)$$

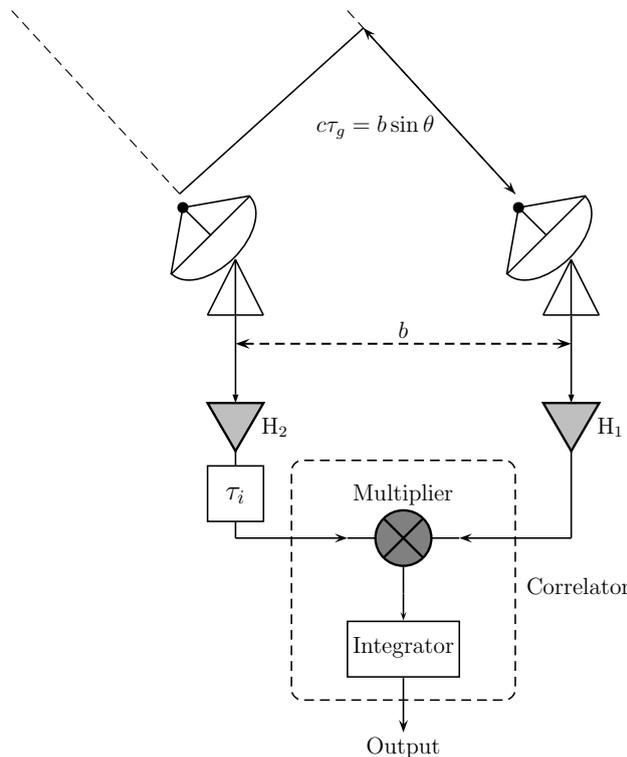


Figure 3.8: A simplified schematic of a two-element interferometer, where H_1 and H_2 are the receivers chains, τ_g the geometrical time delay and τ_i the compensating instrumental time delay.

where τ_g is the geometrical delay, which is the w -component of the projected baseline divided by the center frequency ν_0 of the observing band, $w/\nu_0 = \mathbf{b}_\lambda \cdot \mathbf{s}_0/\nu_0$. Note that the instrumental delay τ_i is assumed to be zero. The voltage signals at the input of the correlator can be expressed as

$$v_{c,1}(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E(l, m) \sqrt{A_{\text{eff}}(l, m)} H_1(\nu) e^{j2\pi\nu t} d\nu dl dm, \quad (3.23a)$$

$$v_{c,2}(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E(l, m) \sqrt{A_{\text{eff}}(l, m)} H_2(\nu) e^{j2\pi\nu(t-\tau_g)} d\nu dl dm, \quad (3.23b)$$

If the received signals are represented by quasi-monochromatic Fourier components of frequency ν , and assuming that identical receivers are used, with uniform passbands of bandwidth $\Delta\nu$, then (3.22) becomes

$$\begin{aligned} r(\tau_g) &= \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E(l, m) E^*(l, m) A_{\text{eff}}(l, m) H_1(\nu) H_2^*(\nu) e^{j2\pi\nu\tau_g} d\nu dl dm dt \\ &= \Delta\nu \cos(2\pi\nu\tau_g) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I(l, m) A_{\text{eff}}(l, m) e^{-j2\pi(u l + v m)} dl dm. \end{aligned} \quad (3.24)$$

The form of (3.24) is only valid if the bandwidth pattern is much wider than the antenna pattern (quasi-monochromatic approximation). The fringe term $\cos(2\pi\nu\tau_g)$ is responsible for the quasi-sinusoidal fringes as the earth rotates and τ_g changes. By observing (3.10) it is evident that the integral in (3.24) is the complex visibility, $V = |V| e^{j\phi_V}$ and the correlator output can be re-written as

$$r = A_0 \Delta\nu |V| \cos(2\pi\nu\tau_g - \phi_V), \quad (3.25)$$

where $A_0 = A_{\text{eff}}(\sigma)/A_N(\sigma)$ is the antenna collecting area in the direction of the phase reference position, \mathbf{s}_0 . Practical interferometry measurements consist of measuring the amplitude and phase of the fringe pattern given in (3.25), and then extracting the complex visibility after calibration. If the monochromatic constraint is relaxed then the bandwidth of the interferometer has the effect of reducing the fringe amplitude, and this is discussed next.

3.6.1 Finite bandwidth effects

The quasi-monochromatic assumption becomes invalid as the receiver bandwidth and/or baseline increases. This ultimately causes the bandwidth pattern to become narrower than the antenna pattern. This effect causes the angular range over which the fringes appear to decrease. This effect is intuitively described by considering a source that transits the beam. If the source is directly normal to the baseline then the geometrical delay is zero and constructive interference occurs at all frequencies. As τ_g increases the phase of the fringe pattern changes, and the phase is different, at the different frequencies across the band. This can cause destructive interference, and averaging over all the frequencies causes the fringe amplitude to decrease [21]. The frequency response of the interferometer is bandlimited by the receivers $H_1(\nu)$ and $H_2(\nu)$. For this analysis they are assumed to be identical. The interferometer response for a rectangular passband of $\Delta\nu$, centered at ν_0 is then

$$\begin{aligned} r &= \Re \left\{ A_0 |V| \int_{\nu_0 - \Delta\nu/2}^{\nu_0 + \Delta\nu/2} e^{j2\pi\nu(\tau_g - \phi_V)} \right\} \\ &= A_0 |V| \cos(2\pi\nu_0\tau_g - \phi_V) \left[\frac{\sin(\pi\Delta\nu\tau_g)}{\pi\Delta\nu\tau_g} \right]. \end{aligned} \quad (3.26)$$

The interferometer response is now amplitude modulated by a sinc function. It is evident from (3.26) that the sinc function decreases rapidly as $\Delta\nu$ and/or $\tau_g = \mathbf{b} \cdot \mathbf{s}/c$ increases. It is desirable to process as wide a bandwidth as possible, since this improves the sensitivity. It is necessary to find a solution to average over the bandwidth without losing fringe amplitude. The fringe term $\cos(2\pi\nu_0\tau_g)$ in (3.26) contains no useful information. This allows for τ_g to be compensated for by inserting an instrumental delay τ_i , as shown in figure 3.8. Ideally, $\Delta\tau = \tau_g - \tau_i = 0$, as this causes the bandwidth pattern (sinc function) to be at its maximum, and the fringe washing problem is solved. This procedure is complicated by the fact the τ_g changes as the earth rotates, and τ_i must therefore continuously be adjusted. This is known as delay tracking.

The first null in the bandwidth pattern for the designed interferometer is now computed, where the assumption of a rectangular passband is made. For the 24m baseline, the maximum geometrical delay is found to be $\tau_g = 80$ ns. This is when the source is at an elevation angle of 0° - along the horizon. From the bandwidth pattern term in (3.26), the first null occurs when the angle θ from zenith is 2.8° . A processing bandwidth of $\Delta\nu_{IF} = 256$ MHz was used in this calculation. This is clearly a narrow angle over which to observe the source. As suggested, delay tracking can be implemented, but this is not trivial and the easiest solution is to use a shorter baseline. To fully understand how delay tracking influences the interferometer response, it is required to look at the frequency conversion.

3.6.2 Frequency conversion effects

Practically all radio astronomy receivers mix the incoming RF signal, ν_{RF} to some intermediate frequency (IF), ν_{IF} . There are many reasons for this, but one in particular that is worth mentioning is the need for mixing the incoming signal down to a lower IF for digitisation. Frequency conversion is performed with the use of a mixer, which responds to both the upper- and lower-sidebands. This suggests that one or both the sidebands can be accepted. Double sideband systems are typically employed above frequencies of about 100 GHz [3]. This is because of the difficulty of making low-noise amplifiers at these high frequencies. Superconductor-insulator-superconductor (SIS) ² mixer receivers are normally used at these frequencies, with reported noise temperatures of approximately 20 K at 100 GHz for a double sideband receiver [3]. ALMA is an example of an instrument that uses SIS mixers at mm and sub-mm wavelengths. Double sideband systems were used in the 1960s and 1970s, but due to increasing levels of RFI, modern systems rely on single sideband operation. Similarly, for this thesis a superheterodyne receiver with single sideband operation was designed. The used of double sideband systems in interferometry are not discussed further and the interested reader can refer to [15].

For the single sideband response of an interferometer consider again the basic interferometer system in figure 3.8. The receivers $H_1(\nu)$ and $H_2(\nu)$ must contain filtering for sideband selectivity before the mixers. It is useful to consider the phase changes ϕ_1 and ϕ_2 imposed on the received signals, when analysing the effect the frequency conversion has on (3.26). These phases are different in the upper and lower sideband cases, because the input signal is an hermitian function. If the lower sideband case is considered, $\nu_{RF} = \nu_{LO} - \nu_{IF}$, then the channel phases of the interferometer can simply be written as

$$\phi_1 = 2\pi(\nu_{LO} - \nu_{IF})\tau_g + \theta_1, \quad (3.27a)$$

$$\phi_2 = 2\pi\nu_{IF}\tau_g + \theta_2, \quad (3.27b)$$

where θ_1 and θ_2 are the local oscillator phase differences. The signal at antenna 1 is delayed by τ_g at a frequency of ν_{RF} , and this causes a phase shift of $2\pi\nu_{RF}\tau_g$ in (3.27a). The instrumental

² A SIS mixer uses a superconducting tunnel junction as the non-linear mixing element. This suggests that for satisfactory operation the SIS mixer must be cooled, where modern SIS mixers use a Nb junction and are cooled to 4 K. New SIS mixers have been developed to be wideband, tunerless and single-sideband [3, 19]

delay, τ_i is inserted at the ν_{IF} frequency, and this results in a phase delay of $2\pi\nu_{IF}\tau_i$ in (3.27b). Inserting (3.27) into (3.26) results in

$$r = A_0 |V| \left[\frac{\sin(\pi \Delta \nu \tau_g)}{\pi \Delta \nu \tau_g} \right] \cos [2\pi\nu_{LO}\tau_g - 2\pi\nu_{IF}\Delta\tau - \phi_V + (\theta_1 - \theta_2)]. \quad (3.28)$$

Since, the delay tracking is implemented at a frequency different to ν_{RF} , the term $2\pi\nu_{LO}\tau_g$ in (3.28) results in a quasi-sinusoidal oscillation, even when $\tau_g = \tau_i$. The frequency of this output oscillation is $\nu_{LO}d\tau_g/dt$, when $\Delta\tau = 0$, and is referred to as the natural fringe frequency. The frequency can exceed 100 kHz in VLBI and the VLA has a fringe frequency of more than 150 Hz for the longest baseline [35]. This requires a sampling rate of at least twice the fringe rate at the correlator output to preserve the fringe information. This can cause a significant increase in the data rate. To avoid this there are numerous methods available to reduce the fringe frequency. One common method is to reduce the fringe frequency to zero by adding a phase of $2\pi\nu_{LO}d\tau_g/dt$ to the local oscillator signal. Alternatively, this can be achieved digitally by multiplying the signal from antenna 2 by a sinusoid with the appropriate instantaneous frequency, this is the approach taken at the GMRT [4]. This procedure is referred to as fringe stopping.

The designed interferometer uses a dual superheterodyne receiver, where the RF signal is first upconverted with a Low-Side injected Local Oscillator (LSLO) ($\nu_{LO,1}$) and then downconverted with a High-Side injected Local Oscillator (HSLO) ($\nu_{LO,2}$). The motivation for this conversion scheme and the design details for the receiver is presented in chapter 4. The approximate (quasi-monochromatic) interferometer response implementing such a frequency conversion scheme is

$$r = |V| |G(\Delta\tau)| \cos \left(2\pi\nu_{LO,1}\tau_g - 2\pi\nu_{LO,2}\tau_g - 2\pi\nu_{IF}\Delta\tau + (\theta_{11} - \theta_{21}) - (\theta_{12} - \theta_{22}) - \phi_V - \phi_G \right),$$

where $G(\Delta\tau)\angle\phi_G$ is the instrumental response of the receiver, θ_{mn} is the phase of local oscillator m for antenna n , $\nu_{LO,1} = 4920$ MHz and $\nu_{LO,2} = 3278$ MHz. The fringe frequency is computed using (3.19) and is found to be

$$\frac{d\tau_g}{dt} \times (\nu_{LO,1} - \nu_{LO,2} - \nu_{IF}) = \frac{1}{c} \omega_e u \cos \delta \times (\nu_{LO,1} - \nu_{LO,2} - \nu_{IF}) = 5.86 \text{ mHz},$$

if no delay tracking is used. The designed interferometer therefore has a fringe period of $t = 170.74$ s. The integration time for the instrument is chosen to be in the order of a few seconds, this suggests that the correlator output is sampled fast enough and no fringe rotation is needed.

In order to measure the phase of the visibility, the fringe oscillations and their phases must be measured. If fringe stopping is used to reduce the fringe frequency to zero than another method is needed to measure the complex amplitude of the visibility. The complex correlator is a scheme where two correlators are used to measure the cosine and sine terms of the complex visibility. The second correlator uses a quadrature phase shift (Hilbert transform) to obtain the sine term. These two terms are then the real and imaginary parts of the complex amplitude.

3.7 Aperture synthesis and imaging

In section 3.2, it is summarised that the measured visibility in the (u, v) plane is related to the source brightness distribution by a simple Fourier transform, see (3.12). It is also mentioned that at any measurement instant, a single Fourier component of the brightness distribution ($V(u, v); V(-u, -v)$) is measured. In order to image a complex source with a high fidelity, many measurements are required to satisfactorily fill the (u, v) plane. The map of the brightness distribution is then obtained by the inverse Fourier transform [4].

Fortunately, most radio sources in the sky are statistically independent, and it is therefore not necessary to measure the whole (u, v) plane simultaneously [4]. This allows for several different methods to be used to fill in the (u, v) plane. Earth rotation synthesis is one such technique where the vector of a tracking array traces out an arc of an ellipse in the (u, v) plane, due to the earth's rotation. Alternatively, pairs of antennas may be used with different baselines. It is possible to synthesise an image to any desired complexity with only two antennas. This is accomplished by moving one antenna to a new position for each observation [33]. If an array of n_a antennas are used then $n_a(n_a - 1)/2$ baseline pairs are simultaneously formed. The (u, v) coverage provided by these methods simulate a larger aperture. Thus, the spatial frequency coverage is the same as that achievable with a single antenna with equal aperture. This is known as aperture synthesis.

Mapping the intensity distribution from the measured visibility data, by direct Fourier transformation, results in an image that is the true brightness distribution convolved with the Point Spread Function (PSF) or synthesised beam

$$V_m(u, v) = W(u, v)g(u, v)V(u, v), \quad (3.29)$$

where $W(u, v)$ is the spatial transfer function, and $g(u, v)$ the weighting function. The Fourier transform of $W(u, v)g(u, v)$ is the synthesised beam, $b_o(l, m)$. This suggests that the shape of the synthesised beam is a function of the (u, v) coverage, which in turn is a function of the antenna array configuration [4]. Before looking at some different array topologies the weighting function and spatial transfer function are discussed. The weighting function provides a means of controlling the shape of the synthesised beam. Two widely used weighting functions are the natural and the uniform schemes. Natural weighting results in a lower angular resolution, but has a higher Signal to Noise Ratio (SNR) than uniform weighting [19].

3.7.1 Spatial transfer function

The spatial transfer function, $W(u, v)$ is equal to the autocorrelation of the electric field $E_a(x_\lambda, y_\lambda)$ across the aperture

$$W(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} E_a(x_\lambda, y_\lambda) E_a^*(x_\lambda - u, y_\lambda - v) dx_\lambda dy_\lambda, \quad (3.30)$$

where (x_λ, y_λ) are the coordinates in the aperture plane of the antenna [15, 14]. It can also be shown that $W(u, v)$ is the Fourier transform of the antenna power pattern. The spatial transfer function further defines the sensitivity of the interferometer to different spatial frequencies. Thus, the array (or antenna) acts as a spatial filter, and defines the (u, v) coverage over which measurements can be made [15]. For example, assume that a two-element interferometer has identical antennas with uniform aperture distribution $E_a(x_\lambda, y_\lambda)$. The spatial transfer function $W(u, v)$ can be shown to have a bandpass filter response. As previously mentioned, $W(u, v)$ forms part of the synthesised beam. If it is assumed that the (u, v) -plane is continuously sampled and that $g(u, v) = 1$, for all (u, v) -points, then the effect $W(u, v)$ has on the formed image is easily observed. Two cases of interest are the Gaussian and uniform (u, v) distributions. The (u, v) coverage has some finite cut-off at the maximum baseline. The Gaussian distribution tapers neatly to zero at the edge which results in a PSF with no sidelobes but a wider beam, resulting in a lower angular resolution. For a uniform distribution, the PSF has a more directive pattern but now with large inner sidelobes, like a J_0 Bessel function [34].

3.7.2 Array configuration

A linear array with the antennas orientated usually in an East-West line is probably the simplest topology. The WSRT is an example of a synthesis array with an East-West baseline. The spatial frequency coverage of the WSRT for different source declinations are shown in figure 3.10. These images were created with AntConfig³. Good images of sources at high declinations (near the celestial pole) can be made, see figures 3.9(b) and 3.9(c), for declinations of $\delta = 90^\circ$ and $\delta = 60^\circ$, respectively. The spacing vector of the array is then more nearly circular. A disadvantage of these one-dimensional arrays are that the beam becomes elongated near the equator as shown in figure 3.9(d), for $\delta = 5^\circ$. It is further seen in figure 3.9(a) that snapshot images are not possible with one-dimensional arrays. To overcome this shortcoming a two-dimensional array must be considered.

A simple two-dimensional array to consider is the “Y”-configuration. Both the GMRT and the VLA are examples of interferometers that use this configuration. The “Y”-configuration allows for reasonable two-dimensional snapshots to be taken, see figure 3.10(a). Furthermore, high fidelity images are possible at low declination angles, for example in figure 3.10(b) good (u, v) -coverage is achieved with the VLA (configuration A) for a declination of $\delta = 10^\circ$. However, there is a sort of grating lobe in the PSF which can only be overcome by several hours of earth rotation synthesis [34]. The (u, v) -coverage for both the VLA and GMRT, for a 12 hour observation is shown in figures 3.10(c) and 3.10(d), respectively. Lastly, circular arrays achieve the best uniform coverage possible, but this results in the synthesised beam having large sidelobes [34].

³AntConfig is a program written by Mattieu de Villiers. This tool reserves as an aid to design an interferometer layout in radio astronomy.

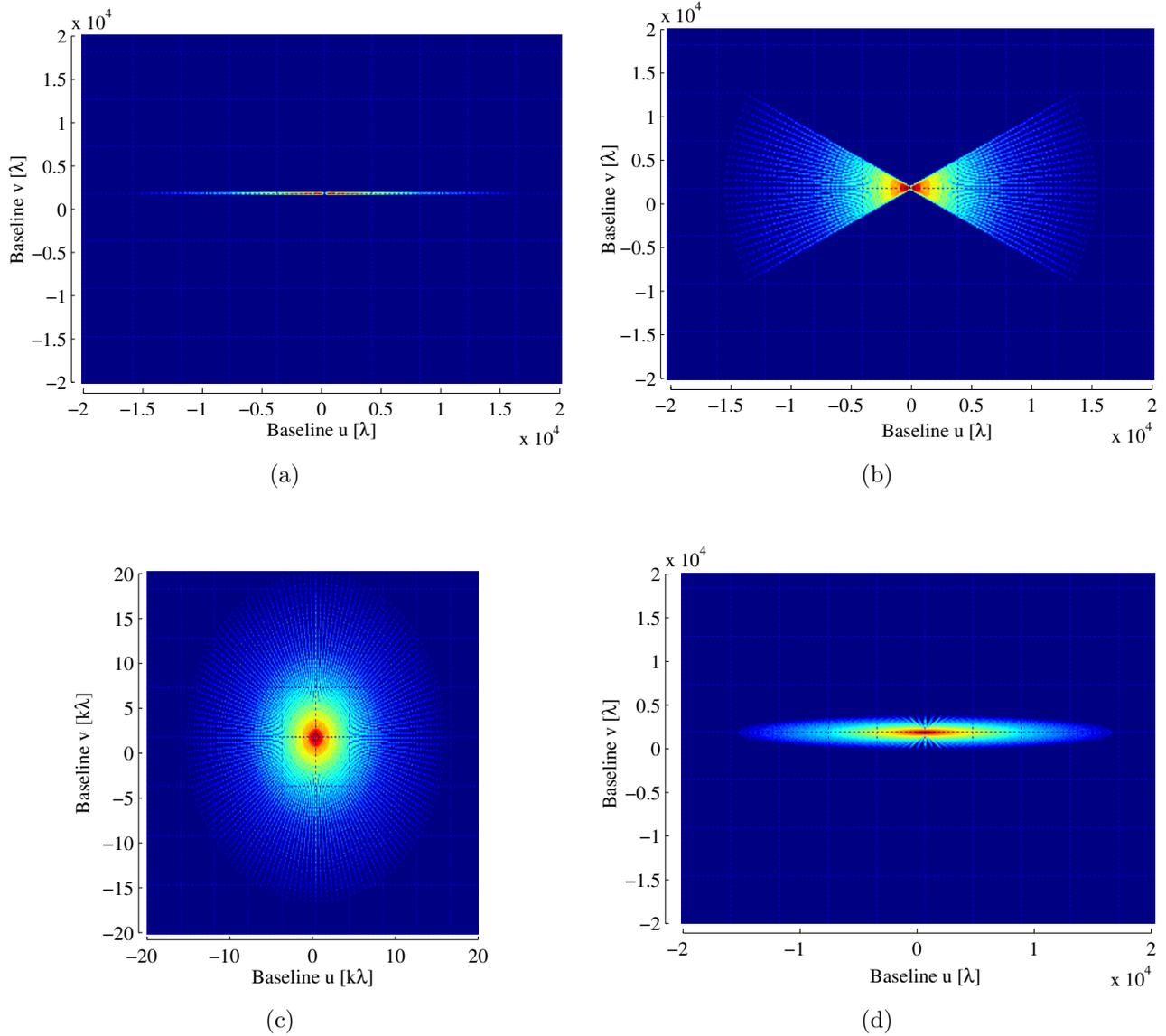


Figure 3.9: The (u, v) -coverage of the WSRT; (a) snapshot at $\delta = 90^\circ$; (b) $\delta = 90^\circ$ for 4 hour coverage; (c) $\delta = 60^\circ$ for 12 hour coverage; and (d) $\delta = 5^\circ$ for 12 hour coverage.

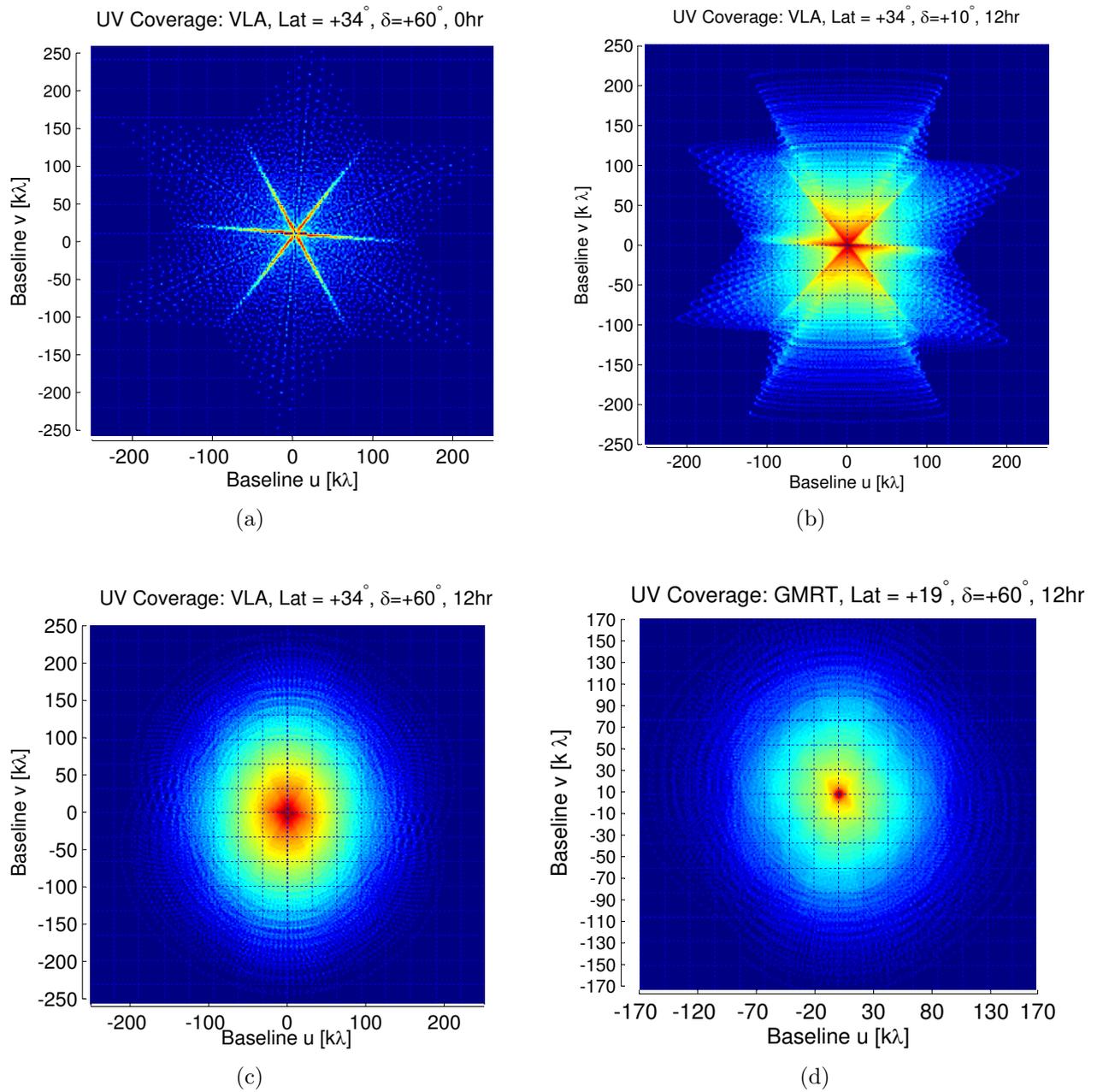


Figure 3.10: The (u, v) -coverage of the VLA (configuration A); (a) snapshot at $\delta = 60^\circ$; (b) $\delta = 10^\circ$ for 12 hour coverage; (c) $\delta = 60^\circ$ for 12 hour coverage; and the (u, v) -coverage of the GMRT (d) $\delta = 60^\circ$ for 12 hour coverage.

3.7.3 Mapping by discrete transformation

Applying the direct Fourier transformation on the measured visibility data, in order to compute the brightness distribution is computationally expensive. The number of multiplications increases as N^4 , for an image of size, $N \times N$ pixels and N^2 visibility measurements [33]. Using the Fast Fourier Transform (FFT), the number of multiplications are $2N^2 \log_2 N$, which provides a computational advantage as N increases. However, the FFT introduces some complications, firstly, the visibility must be evaluated at points on a rectangular grid, and secondly, aliasing can occur.

In order to apply the FFT it is first necessary to determine the visibility values at discrete points. This requires that a sampling function be defined

$$S(u, v) = \sum_{i,j} \delta(u - i\Delta u, v - j\Delta v), \quad (3.31)$$

where $\Delta u, \Delta v$ are uniformly spaced at the grid intervals. This suggests that in the (l, m) domain, the measured source distribution is convolved with the Fourier transform of the sampling function. Images of the source distribution will therefore appear at multiples of $(\Delta u^{-1}, \Delta v^{-1})$, and this can clearly result in aliasing. To avoid any aliasing it is required that the change in the baseline length between measurements, $(\Delta u)^{-1}$ and $(\Delta v)^{-1}$ be no greater than the reciprocal of the interval where $I(l, m)$ is non-zero [15].

Moreover, the measured visibility lie on elliptical loci, and it is therefore necessary to estimate the values at the required grid points. This process is known as gridding. The estimation of the grid points involve the convolution of the measured visibility, V_m (as in (3.29)) with a function $C(u, v)$, in order to produce a continuous visibility distribution [15, 33, 4]. This is followed by resampling at the grid points

$$V_g(u, v) = \Delta u \Delta v \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \delta^2(u - i\Delta u, v - j\Delta v) \{C(u, v) \otimes \otimes [W(u, v)V(u, v)]\}, \quad (3.32)$$

where $\otimes \otimes$ represents a two-dimensional convolution, and $W(u, v)$ is now measured at discrete points according to the sampling function. The FFT algorithm can then be applied to $V_g(u, v)$ to yield an estimate of the source intensity distribution. The emphasis of this thesis is not to map a region of the sky, but for completeness imaging was briefly discussed. See [35, 15] for a much more thorough treatment of imaging theory, especially the choice of convolution function, $C(u, v)$ and deconvolution.

3.8 Conclusion

To conclude, this chapter focused on the fundamentals of radio interferometry. In particular, the van Cittert-Zernike theorem was presented, which established the connection between the source intensity and the measured spatial coherence function (visibility), and the response of a two-element interferometer was discussed in full. The topic of aperture synthesis and mapping was also briefly discussed. It was also shown that a baseline of at least 24 m is needed to resolve the sun. In this regard, the maximum baseline for the interferometer was chosen as 24 m.

Chapter 4

Characterisation and design of the radio front-end

4.1 Principles of radio astronomy receivers

The radio emissions of a cosmic source that are concentrated at the focus of a radio telescope is in most cases too weak to be detected directly. Furthermore, these signals have noise-like characteristics, and are therefore indistinguishable from the inherent noise of the receiver. The function of the radio-telescope receiver is then to discriminate between the cosmic signal and the much stronger receiver noise. This implies that both the sensitivity and stability (due to the high gain) of the receiver, are important requirements. The sensitivity can simply be defined as a measure of the weakest source of radio emission that can be detected [35]. The minimum detectable noise temperature ΔT_{\min} of a receiver, is a stochastic signal, and it is therefore defined statistically equal to the root-mean-square (rms) deviation about the mean level

$$\Delta T_{\min} = K_s \frac{T_{\text{sys}}}{\sqrt{\Delta\nu_{\text{IF}}\tau_a}}, \quad (4.1)$$

where T_{sys} is the effective system noise temperature, $\Delta\nu_{\text{IF}}$ the pre-detection bandwidth, τ_a the post-detection integration time constant and K_s is the sensitivity constant of different radio-telescope receivers. The system noise temperature includes both the receiver noise T_{rec} and effective antenna noise T_A , thus $T_{\text{sys}} = T_{\text{rec}} + T_A$. Furthermore, T_A includes the cosmic signal noise temperature ΔT and the background noise from the sky, ground, etc. This expression was first derived by Dicke [37]. From (4.1) it is seen that the sensitivity can indefinitely be improved by increasing the pre-detector bandwidth and the integration time. However, there are limits to the extent to which these parameters may be increased. The bandwidth is ultimately limited by the operating bandwidth of the receiver components and by RFI [14, 38]. Too long an integration time can distort the true source structure (drift-scan), and it can further cause source displacement [38]. For a measurement to be discernible the sensitivity should at least be 5σ [21].

The type of receiver architectures used in radio astronomy are very similar to the receivers used in other branches of radio engineering, in particular radar and telecommunications engineering.

But, by far the most popular type is the superheterodyne receiver. A block diagram of a single-conversion superheterodyne receiver, consisting of an RF stage, a mixer/LO and an IF stage is given in figure 4.1. The RF stage is in most cases preceded with a low-noise amplifier (LNA), followed by a bandpass filter (BPF) to suppress any out of band interference. The majority of the signal processing is performed after coherently mixing the RF signal down to a lower IF signal. At lower frequencies the filters have lower (more practical) Q-values for good selectivity, a larger selection of amplifiers are available to the designer and the digitisation of the signals is more manageable. The majority of the signal amplification occurs in the IF stage, and this stage also usually determines the signal bandwidth $\Delta\nu_{\text{IF}}$ [38]. The quest for ultra low-noise ($T_{\text{sys}} \ll$) reception is especially important in radio astronomy, see (4.1). A significant contribution of the overall system noise temperature is due to the receiver. This has led to the development of low-noise receiver technologies. In the 1980s, low-noise receivers were built using GaAs FET technology combined with cryogenic cooling [3, 33]. The VLA made use of these technologies and a receiver noise performance of 30 K was achieved for a wavelength of 20 cm [33]. Modern radio astronomy instrumentation use Heterostructure Field-Effect Transistors (HFETs) below frequencies of about 120 GHz [3]. The noise performance achievable with a cooled $0.1 \mu\text{m}$ -long gate InP HFET is as low as 15 K at 43 GHz [3]. HFET technology has matured to the point where their use in large radio astronomy systems have become practical. Several instruments including the VLA, VLBA and the Greenbank telescope have successfully made use of HFETs. The VLBA has a state of the art receiver that exhibits a noise temperature of less than 10 K [3]. It is not feasible to cool the receiver front-end of an amateur instrument. This being said, new-generation amplifier technology has improved to the point where GaAs and HEMT devices, can easily achieve noise figures of less than 0.4 dB (~ 28 K) [28]. An example of a synthesis array that uses uncooled receivers is the GMRT [4]. The GMRT has a receiver noise temperature of only 45 K (28 K LNA) at 1390 MHz. This suggests that the receiver noise performance of an amateur instrument is almost comparable to that of a large professional system.

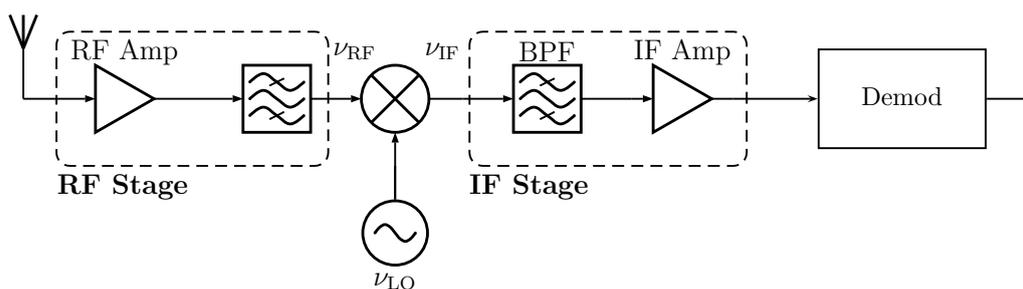


Figure 4.1: Block diagram of a single-conversion superheterodyne receiver.

There exists numerous methods for detecting weak astronomical signals. These have been culminated into various different receiver types, each with its own advantages and drawbacks. The total power receiver depicted in figure 4.2 is the simplest receiver topology, but this simplicity also makes it the least practical. The total power receiver is susceptible to gain variations, ground radiation, and removing any interference from the desired signal is also problematic [12].

The gain variation directly influences the sensitivity, such that a differentially small gain fluctuation, ΔG can easily exceed the smallest detectable change in signal temperature ΔT . The gain fluctuations and the thermal noise are both independent random processes. It is therefore possible to statistically add their powers. This changes the sensitivity of (4.1) accordingly

$$\Delta T_{\min} = K_s T_{\text{sys}} \left[\left(\frac{1}{\Delta \nu_{\text{IF}} \tau_a} \right) + \left(\frac{\Delta G}{G} \right)^2 \right]^{1/2}. \quad (4.2)$$

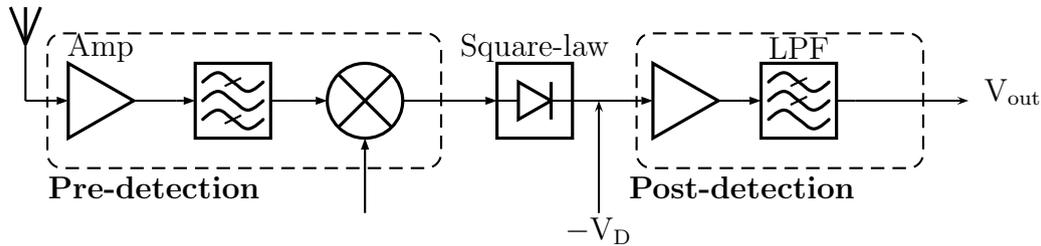


Figure 4.2: Block diagram of a total power receiver.

The gain instability in (4.2) is ultimately the limiting factor of the sensitivity in a practical system. The sensitivity constant K_s of the total power receiver is specified as 1. This provides a basis for comparing the sensitivities of the other receiver types against. Dicke was the first to introduce the use of modulation principles for eliminating low frequency receiver instability, including gain fluctuations [38, 37]. The Dicke receiver shown in figure 4.3 switches the input of the receiver between the antenna and some reference load. If the load noise power matches the noise power of the system, then the signal power is amplitude modulated at the switching frequency. Using a phase detector which is switched in synchronism with the receiver, then removes all the low frequency instabilities. The sensitivity of the Dicke receiver is one half the theoretical sensitivity of the total power receiver, $K_s = 2$, assuming a square wave modulating signal. The reason for this is that the Dicke receiver only connects to the signal half the time. However, the gain instabilities are removed which makes this a much more practical receiver.

The Dicke receiver only observes a radio source half of the time. This is inefficient and results in a loss of sensitivity. A method of improving the sensitivity, whilst maintaining the gain stability advantage is to switch the antenna between two Dicke receivers [38]. This is known as the Graham's receiver, and a simplified block diagram of this receiver is given in figure 4.4. Both the output signals in this receiver are added which results in a sensitivity improvement of $\sqrt{2}$. In contrast, interferometric receivers use two (or more) antennas. This means that more power is received from the source, which naturally leads to higher sensitivities. The maximum sensitivity of the adding (simple) interferometer depicted in figure 3.1, is shown to be twice that of a total power receiver, $K_s = 1/2$. But, the adding interferometer also suffers from the same drawbacks as the total power receiver. The correlation interferometer as shown in figure 4.5, has a slightly worse sensitivity by a factor of $1/\sqrt{2}$. The reason for this is that the adding interferometer first adds the input voltages, before passing the signal through a square-law detector, $(v_1 + v_2)^2$. The correlation interferometer is still more sensitive

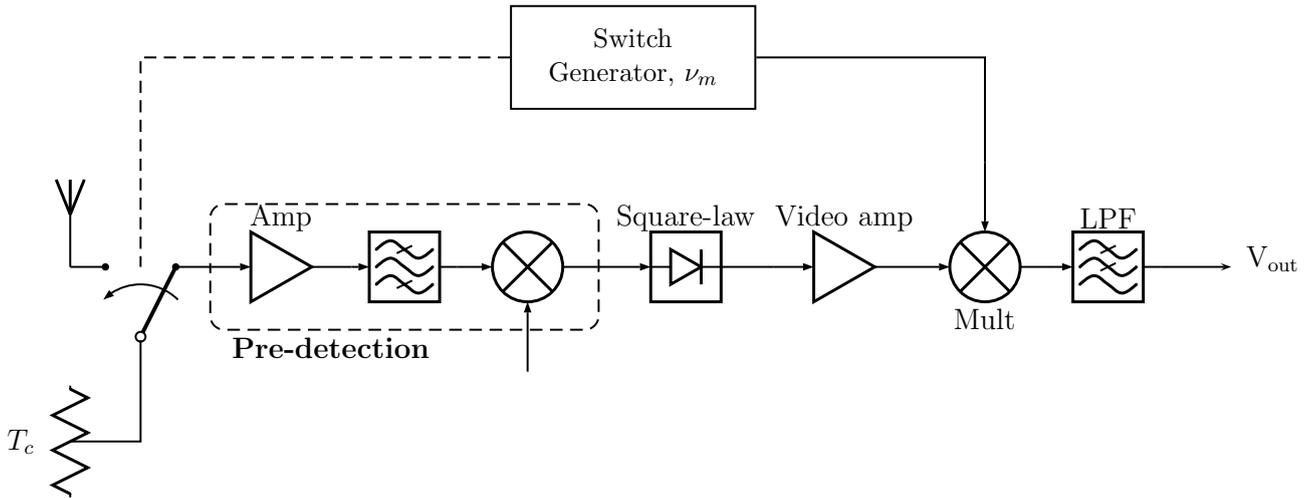


Figure 4.3: Block diagram of a Dicke receiver.

than the single antenna receivers, where it can be seen to be $2\sqrt{2}$ times better than a Dicke receiver [14]. Moreover, the correlation interferometer has superior performance and some of its advantages are highlighted in chapter 3. A rigorous derivation of the sensitivity of the correlation interferometer is presented in this chapter. The receiver characteristics are critical for achieving the desired accuracy and maintaining the sensitivity of the measured visibility. For these reasons, the performance metrics such as the receiver noise temperature, phase stability and linearity are also discussed. Lastly, the design details and performance results of the implemented receiver are presented and discussed.

4.1.1 Sensitivity of an interferometer

The sensitivity derivation presented here follows closely the approach of Goldstein, which makes use of the autocorrelation function [39].

Consider the elementary functional block diagram of the correlation interferometer given in figure 4.5. The signal voltages induced in the antennas by the cosmic source are highly correlated, $v_{s,1}(t) = v_{s,2}(t) = v_s(t)$. However, the receiver noise voltages, $n_1(t)$ and $n_2(t)$, are uncorrelated. The voltage waveforms at the correlator inputs are denoted by $v_{c,1}(t) = v_s(t) + n_1(t)$ and $v_{c,2}(t) = v_s(t) + n_2(t)$. For this derivation it is assumed that the receiver accounts for the geometric delay τ_g , with the use of delay tracking and fringe stopping. Hence, the identical signal voltages $v_s(t)$ at the correlator inputs. Furthermore, it is assumed that all the signals are independent, stationary ¹ random processes ², with a Gaussian amplitude distribution given by (B.1), and zero mean. The average signal powers are $\sigma_{v_s}^2$, $\sigma_{n_1}^2$ and $\sigma_{n_2}^2$. The output of the multiplier circuit in figure 4.5, is the cross-multiplication of the two signal channels

¹The random process must at least be Wide-Sense Stationary (WSS), see Appendix B for the definition of WSS.

²A Roman type (x) is used to denote a random variable or random process.

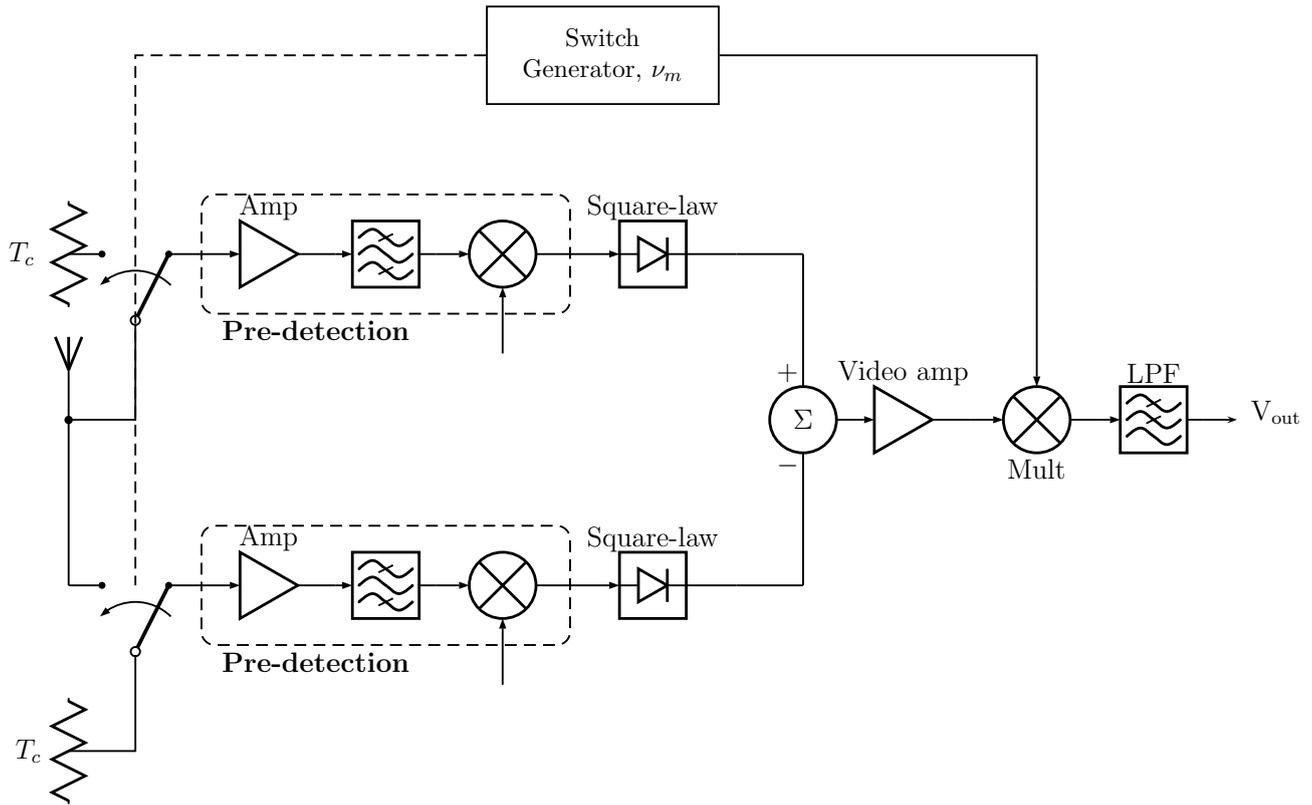


Figure 4.4: Block diagram of Graham's receiver.

$$r(t) = K (v_s^2(t) + v_s(t) [n_1(t) + n_2(t)] + n_1(t)n_2(t)), \quad (4.3)$$

where K is the receiver proportionality constant. The sensitivity is basically a measure of the SNR of the receiver. The method used to compute the signal- and noise-powers is based on the autocorrelation of $r(t)$. From the autocorrelation function, the power spectral densities (PSDs) can be computed using the Wiener-Khinchin theorem (see appendix B). Following this approach, the autocorrelation of $r(t)$ is obtained as

$$\rho_r(\tau) = K^2 \left\{ \overline{v_s^2(t)v_s^2(t+\tau)} + \overline{v_s(t)v_s(t+\tau)} \left[\overline{n_1(t)n_1(t+\tau)} + \overline{n_2(t)n_2(t+\tau)} \right] + \overline{n_1(t)n_1(t+\tau)} + \overline{n_2(t)n_2(t+\tau)} \right\}, \quad (4.4)$$

where the products of the uncorrelated terms (i.e. the products of the signal and noise voltages, and the noise voltages from the different receivers) are omitted, due to their expectations being zero. The overbar \bar{x} indicates the average or expected value of the random process x . It is shown in [40], that a random process of the form $y(t) = x^2(t)$, has an autocorrelation function of

$$\rho_y(\tau) = \rho_x^2(0) + 2\rho_x^2(\tau), \quad (4.5)$$

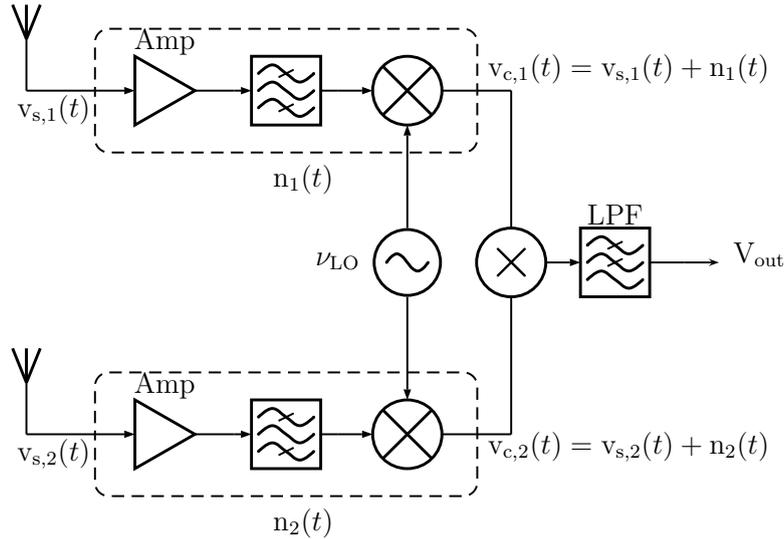


Figure 4.5: Simplified schematic of a correlation interferometer.

where $x(t)$ is a stationary, Gaussian random process with zero mean, and autocorrelation $\rho_x(\tau)$. Using (4.5), it is seen that

$$\overline{v_s^2(t)v_s^2(t+\tau)} = \sigma_{v_s}^4 + 2 \left[\overline{v_s(t)v_s(t+\tau)} \right]^2. \quad (4.6)$$

In order to solve (4.6) and (4.4), the autocorrelation functions, $\rho_{v_s}(\tau)$, $\rho_{n_1}(\tau)$ and $\rho_{n_2}(\tau)$ must first be computed. To simplify these computations, assume that the signals are quasi-monochromatic and that both channels have identical rectangular passbands $\Delta\nu_{IF}$, centered at ν_{IF} . The autocorrelation functions are then mathematically described by

$$\overline{v_s(t)v_s(t+\tau)} = \frac{\sigma_{v_s}^2 \sin(\pi\Delta\nu_{IF}\tau)}{\pi\Delta\nu_{IF}\tau} \cos 2\pi\nu_{IF}\tau, \quad (4.7a)$$

$$\overline{n_1(t)n_1(t+\tau)} = \frac{\sigma_{n_1}^2 \sin(\pi\Delta\nu_{IF}\tau)}{\pi\Delta\nu_{IF}\tau} \cos 2\pi\nu_{IF}\tau, \quad (4.7b)$$

$$\overline{n_2(t)n_2(t+\tau)} = \frac{\sigma_{n_2}^2 \sin(\pi\Delta\nu_{IF}\tau)}{\pi\Delta\nu_{IF}\tau} \cos 2\pi\nu_{IF}\tau. \quad (4.7c)$$

Substituting (4.7) and (4.6) into (4.4) yields

$$\begin{aligned} \rho_r(\tau) = K^2 & \left(\sigma_{v_s}^4 + \frac{1}{2} \left[2\sigma_{v_s}^4 + \sigma_{v_s}^2 (\sigma_{n_1}^2 + \sigma_{n_2}^2) \right. \right. \\ & \left. \left. + \sigma_{n_1}^2 \sigma_{n_2}^2 \right] \frac{\sin^2 \pi \Delta \nu_{IF} \tau}{\pi^2 \Delta \nu_{IF}^2 \tau^2} \right), \end{aligned} \quad (4.8)$$

where $K^2\sigma_{v_s}^4$ represent the signal and the remaining terms describe the noise. The integrator (averaging) after the multiplier circuit can be modelled as a low-pass filter, with a rectangular passband from 0 to $\Delta\nu_{LF}$. Assuming an integration time constant in the order of tens of seconds,

which is typical for radio astronomy applications [38]. The bandwidth of $\Delta\nu_{\text{LF}} \ll \Delta\nu_{\text{IF}}$, and the output noise power is approximately $P_n = S_n(0)\Delta\nu_{\text{LF}}$, where $S_n(\nu)$ is the PSD of the noise components in (4.8). Using the Wiener-Khinchin theorem, it is shown that the output noise power of the correlator is

$$P_n = 4\Delta\nu_{\text{IF}} \int_0^\infty \rho_n(\tau) d\tau, \quad (4.9)$$

where $\rho_n(\tau)$ denotes the noise term of (4.8), and (B.13) was used. Also, using (B.11), the signal power at the correlator output is given as $K^2\sigma_{\text{vs}}^4\delta(0)$. The signal to noise ratio is then found to be

$$\text{SNR} = \frac{\sigma_{\text{vs}}^4}{2\sigma_{\text{vs}}^4 + \sigma_{\text{vs}}^2(\sigma_{\text{n1}}^2 + \sigma_{\text{n2}}^2) + \sigma_{\text{n1}}^2\sigma_{\text{n2}}^2} \left(\frac{\Delta\nu_{\text{IF}}}{\Delta\nu_{\text{LF}}} \right). \quad (4.10)$$

In most circumstances, the case where the source is near the threshold of detectability, such that $\sigma_{\text{n}}^2 \gg \sigma_{\text{vs}}^2$ is of interest. The sensitivity measure given by (4.10) is then simplified to

$$\text{SNR} = \frac{\sigma_{\text{vs}}^4 \Delta\nu_{\text{IF}}}{\sigma_{\text{n}}^4 \Delta\nu_{\text{LF}}}, \quad (4.11)$$

where it is further assumed that the noise powers in both channels are identical, $\sigma_{\text{n1}}^2 = \sigma_{\text{n2}}^2 = \sigma_{\text{n}}^2$. It is convenient to express the SNR in terms of the detectable signal temperature ΔT and the system noise temperature T_{sys} , where

$$\sigma_{\text{vs}}^2 = k\Delta T\Delta\nu_{\text{IF}}, \quad (4.12a)$$

$$\sigma_{\text{n}}^2 = kT_{\text{sys}}\Delta\nu_{\text{IF}}, \quad (4.12b)$$

and k is Boltzmann's constant. A SNR of unity results in the least-detectable signal level, and using (4.12) this is given by

$$\Delta T = T_{\text{sys}} \sqrt{\frac{\Delta\nu_{\text{LF}}}{\Delta\nu_{\text{IF}}}}. \quad (4.13)$$

The $\frac{\sin^2 \pi \Delta\nu_{\text{IF}} \tau}{\pi^2 \Delta\nu_{\text{IF}}^2 \tau^2}$ factor in (4.8) is an artefact of the rectangular passband, and the equivalent bandwidth is determined by integrating this function

$$2\Delta\nu_{\text{LF}} = \int_{-\infty}^{\infty} \frac{\sin^2 \pi \Delta\nu_{\text{IF}} \tau}{\pi^2 \Delta\nu_{\text{IF}}^2 \tau^2} d\nu = \frac{1}{\tau_a}, \quad (4.14)$$

where τ_a is the time constant of the low-pass filter. Substituting (4.14) into (4.13), results in

$$\Delta T = \frac{T_{\text{sys}}}{\sqrt{2\Delta\nu_{\text{IF}}\tau_a}}. \quad (4.15)$$

Alternatively, the sensitivity can be expressed as the minimum detectable flux density, ΔS . The thermal power density $k\Delta T$ is related to both the flux density and the effective collecting area of the antenna $\eta_a A$, by (2.12), and (4.15) becomes

$$\Delta S = \frac{2kT_{\text{sys}}}{\eta_a A} \frac{1}{\sqrt{2\Delta\nu_{\text{IF}}\tau_a}}, \quad (4.16)$$

where η_a is the aperture efficiency of the antenna. The system equivalent flux density (SEFD) is an indicator that accounts for both the effective antenna area plus system noise. The SEFD is defined as

$$SEFD = \frac{2kT_{\text{sys}}}{\eta_a A}. \quad (4.17)$$

Expressing the sensitivity given by (4.16), in terms of the SEFD

$$\Delta S = \frac{SEFD}{\sqrt{2\Delta\nu_{\text{IF}}\tau_a}}. \quad (4.18)$$

Another factor that influences the sensitivity is the efficiency factor η_q (see chapter 6). If the signals are digitised before entering the correlator, the finite quantisation levels result in a loss of sensitivity. If η_q is included, (4.18) becomes

$$\Delta S = \frac{SEFD}{\eta_q \sqrt{2\Delta\nu_{\text{IF}}\tau_a}}. \quad (4.19)$$

Similar sensitivity derivations can be found in [15, 35, 14]. Note that this is not the imaging sensitivity. See [35] for a derivation of the rms noise level in a synthesised image. To get a better idea of how sensitive a large synthesis array is, the minimum observable flux density of some of these instruments are presented next. Note that only a single baseline (two antennas) is considered here, to determine the rms noise level in the synthesised image the number of baseline pairs must be accounted for, see [33, 35]. The VLA is capable of detecting sources with flux densities as low as $\Delta S = 0.264$ Jy for a 10 second integration time, at a wavelength of 20 cm [33]. This assumes an aperture efficiency of 0.52, using the full bandwidth of 46 MHz, 2-bit (3-level) quantisation, $\eta_q = 0.81$ and $T_{\text{sys}} = 60$ K [33]. In order to draw comparisons between the different systems, it is assumed that all the instruments operate at or near 1.42 GHz and $\tau_a = 10$ s. The GMRT has an L-band feed with $\eta_a = 0.422$, a processing bandwidth of 32 MHz (this assumes that both 16 MHz sidebands are used), 6-bits are used for signal quantisation resulting in $\eta_q \sim 1$ and a system noise temperature of 72 K is specified. From these performance figures, the minimum discernible flux density is 0.117 Jy [4]. The Meerkat is still under development, but it is worth looking at the sensitivity capabilities of a next-generation synthesis array. From the initial specifications, the Meerkat is able to detect down to an impressive sensitivity of $\Delta S = 0.1$ Jy. This is largely due to the massive instantaneous bandwidth being processed $\Delta\nu = 512$ MHz, and a low system noise temperature $T_{\text{sys}} = 30$ K, the other important performance parameters are $\eta_a = 0.7$ and $\eta_q \sim 1$ [9]. The sensitivities achievable with small amateur instruments are also very encouraging [13, 11, 12]. The details of these amateur instruments are more fully discussed in chapter 5, a brief summary of their predicted sensitivities are only given here. The adding interferometer in [13] is defined to have a sensitivity of $\Delta S = 5.5$ Jy, where $T_{\text{sys}} = 150$ K, $A_{\text{eff}} \approx 4.24$ m² and 12-bits are used for the digitisation. The 408 MHz interferometer is expected to achieved a sensitivity of $\Delta S = 171$ Jy, for $T_{\text{sys}} = 150$ K, $G = 14$ dBi and $\Delta\nu = 250$ kHz [12]. This instrument uses an analogue correlator and an initial

bandwidth of 1 MHz was designed for, but because of RFI issues, a much lower bandwidth was ultimately used. The 1.3 GHz correlation interferometer of [11] achieves a sensitivity of about $\Delta S = 163$ Jy, where $\Delta\nu = 4$ MHz, $G = 18$ dBi and an 8-bit ADC is used. The sensitivities achieved with these small interferometers are good enough to detect the radio sources given in table 2.1. These figures seem very promising, but they are probably too optimistic, given the RFI levels such instruments have to contend with and the poor phase tolerances of these instruments.

To define an initial set of receiver specifications the sensitivity equation of (4.19) can be used. The minimum sensitivity ΔS was specified as 40 Jy and the bandwidth of the system as $\Delta\nu = 256$ MHz. As a first order approximation of the system noise temperature T_{sys} , assume that the antenna temperature owing to the sky, T_{sky} is roughly ~ 48 K. This includes the brightness temperature of the CMB of ~ 2.7 K, the sky noise at 1.42 GHz which is approximately 15 K, it also includes the background noise received through the sidelobes of the antenna and the warm ground (290 – 300 K), this can potentially add an additional 30 K or so. The GMRT and VLA at 1.4 GHz have a T_{sys} of 28 K, and 20 K, respectively [4, 33]. It has to be stressed that this is only an approximation based on the T_{sys} figures given in the literature. To accurately measure the background noise temperature of the system the spatial distribution of the background temperature must be weighted by the pattern of the antenna as follows

$$T_A = \frac{1}{\Omega_M + \Omega_S + \Omega_G} \int \int_{4\pi} T_B(\theta, \phi) P_n(\theta, \phi) \sin \theta d\theta d\phi,$$

where Ω_S is the spillover solid angle, Ω_M the main beam solid angle, Ω_G the solid angle directed towards the ground and $T_B(\theta, \phi)$ is the temperature distribution of the background. This is however not a trivial computation to perform, and it is also dependent on the direction the antenna is pointing in. Therefore, this value varies as a source is being tracked. The receiver noise temperature T_{rec} is also approximated, but with far less error. Off-the-shelf LNAs with Noise Figures (NFs) in the order of ~ 0.8 dB (59 K), are readily available. A well designed receiver will thus have a noise temperature close to that of the LNA. Accounting for any losses before the LNA (say 0.3 dB), then $T_{\text{sys}} \approx 130$ K. Using $\tau_a = 10$ sec, it can be shown from (4.19) that an antenna with a gain of at least 15.5 dBi is needed to achieve a sensitivity of 40 Jy. This is however a theoretical prediction and most likely very under-estimated. Furthermore, a 256 MHz bandwidth is probably very ambitious because of the cluttered RFI environment the interferometer is expected to operate in. If for example only a 64 MHz bandwidth can be used, and assuming an integration time constant of 10s, then the gain of the antenna needed is 18.5 dBi. Thus, for a conservative design the specifications of the receiver noise temperature and the gain of the antenna are

- $T_{\text{rec}} \leq 85$ K (NF = 1.1 dB),
- $G \geq 19$ dBi.

This chapter deals with the design of the receiver, and the design details of the antenna are deferred to chapter 5.

4.2 Receiver performance metrics

The receiver performance is critical for maintaining a high sensitivity and for avoiding any inaccuracies in the measured visibility data. In this section some of the more critical performance parameters of the receiver are discussed in terms of the potential performance degradation they may cause, and the ways in which professional instruments try to avoid such problems. The performance metrics that are focused on is the receiver's noise figure, the linearity, its frequency response and the phase stability.

4.2.1 Noise figure

The NF of a receiver is a measure of the absolute sensitivity of the system. The NF metric makes it possible to easily analyse the noise content of the whole receiver, and the contribution of the individual components. It will be shown that the NF can simply be translated into the system noise temperature, T_{sys} . The NF is defined as the ratio of the available SNR at the input to the available SNR at the output of a noisy network (or component) [41], see figure 4.6

$$\text{NF} = 10 \log \left(\frac{P_{s_i}/P_{n_i}}{P_{s_0}/P_{n_0}} \right), \quad (4.20)$$

where

$$F = \frac{P_{s_i}/P_{n_i}}{P_{s_0}/P_{n_0}}, \quad (4.21)$$

is the noise factor, F . Alternatively, the noise factor may be expressed as

$$1 + \frac{T_e}{T_0} = \frac{P_{n_0}}{kT_0\Delta\nu G_a} = \frac{P_{n_0}}{P_{n_i} G_a} = \frac{P_{s_i}/P_{n_i}}{P_{s_0}/P_{n_0}}, \quad (4.22)$$

where G_a is the available gain, and T_e is the equivalent noise temperature of the network (or component). It is important to note that the noise figure is defined for a matched source impedance at room temperature, $T_0 = 290$ K. Another, useful expression is obtained by solving (4.22) for T_e

$$T_e = (F - 1)T_0. \quad (4.23)$$

This is an especially useful expression, since the NF of a network can be related to an equivalent temperature. The noise factor of a cascaded system of n -networks is given by

$$F_{\text{cas}} = F_1 + \frac{F_2 - 1}{G_1} + \frac{F_3 - 1}{G_1 G_2} + \dots + \frac{F_N - 1}{G_1 G_2 \dots G_{N-1}}, \quad (4.24)$$

where F_N is the noise factor-, and G_N is the available gain- of the N^{th} network. The NF is then simply $10 \log (F_{\text{cas}})$. Equation (4.24) indicates that if the gain of the first stage is large enough, then the NF of the cascade is effectively determined by the first stage. Thus, to achieve the lowest NF possible, the first stage should have a very low NF and reasonably high gain. A practical guideline is to design for $G_1 + \text{NF}_1 > \text{NF}_2 + 15$, this ensures that the second stage contributes less than 0.1 dB to the NF. The NF standard makes it simple to analyse the noise

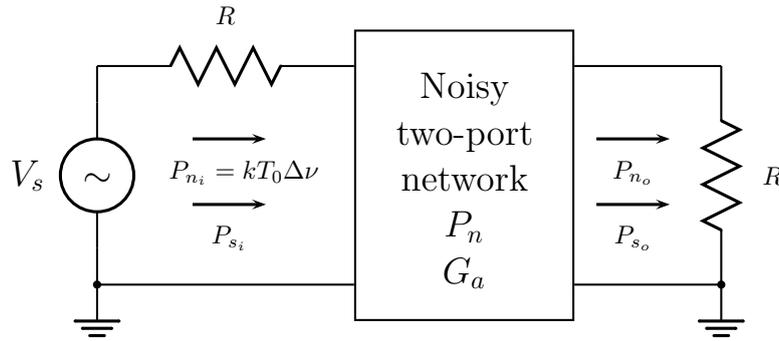


Figure 4.6: Illustration of a noisy network used to determine the NF of a two-port network.

content, and equivalently the noise temperature of the receiver. Additionally, the NF is important from a design point of view, since the NFs of the individual components can be identified. This enables a designer to improve the part of the system responsible for most of the noise.

Care must be exercised if a section of transmission line is placed between the antenna and the first LNA in the receiver. In [42], the noise factor of a passive, lossy component such as a transmission line is derived

$$F = 1 + (L - 1) \frac{T_p}{T_0}, \quad (4.25)$$

where $L = 1/G > 1$ is the loss factor and T_p is the physical temperature of the line. For a line at $T_p = T_0$, (4.25) shows that $F = L$. The transmission line clearly adds additional noise to the system if it is the first component in (4.24). For example, each 0.1 dB increase in attenuation (L), raises the system noise temperature by 6.7 K. Hence, it is important to keep the transmission line short and low-loss [38].

It is worth mentioning that the NF expressions given in this section, ignore any impedance mismatches that might exist between the networks. This can result in an under-estimation of the actual NF [42]. For the designed superheterodyne receiver the VSWR between the components are specified as better than 1.5:1 ($RL \approx -14$ dB). Thus, the expressions given here are certainly accurate for an initial estimation.

4.2.2 Phase stability considerations

Poor phase stability results in phase errors which adversely affects the sensitivity, and the measured visibility phase [15]. The phase error $\Delta\phi$ often has an even probability distribution with a zero mean. The rms phase error $\Delta\phi_{\text{rms}}$ at the output of the correlator for this case is approximately given by

$$r \approx 1 - \frac{1}{2} \Delta\phi_{\text{rms}}^2. \quad (4.26)$$

Equation (4.26) is only valid for $\Delta\phi_{\text{rms}}$ less than about 37° , and the maximum sensitivity loss is then 20.85% [15]. This highlights the importance of good phase stability for a professional

synthesis array. A phase-stable system requires a good LO distribution system where the phase coherency at all the antennas are maintained. Furthermore, a stable IF transmission path from the antennas to the correlators are also crucial [33]. There are three principal techniques used for obtaining this phase stability in radio astronomy. The method suggested by Swarup and Yang is based on measuring the phase variation between the master oscillator and a portion of the LO that is reflected back from the antennas [33]. The reflector at the end of the line is switched at a low frequency, in order to make it possible to distinguish the reflected LO from all the other unavoidable reflections. The round-trip-phase measurement method is another popular technique that is similar to the Swarup and Yang technique. Again a sample of the LO is returned from the antenna for comparison with the master oscillator. But, this time a phase-lock loop at the antenna is used to add a small frequency offset, which makes the reflected LO distinguishable from all the other reflections. The VLA, the Nobeyama mm-wave synthesis telescope and Jodrell bank have all successfully applied this technique [33]. The final, commonly used technique for distributing phase-stable LOs, consists of injecting two signals at opposite ends of the transmission line. After mixing these two signals, the sum signal must have the same phase at each antenna [33, 15].

A phase-stable LO distribution system forms a critical part of a large synthesis array, but this is not the case for amateur applications. Similarly, for this thesis all the frequency mixing is done remotely (not at the antennas), and a phase-coherent LO distribution system is not required. There will almost certainly be phase errors between the two receiver paths, resulting in a loss of sensitivity. The maximum allowable sensitivity degradation for the system is specified as 2.5%, this is in accordance to the specifications given in [15]. Using (4.26), the phase error must therefore be less than 12.8° rms.

4.2.3 Sensitivity degradation due to the receiver frequency response

The overall frequency response of the receiver can be used as a means of characterising how the receiver components influences the sensitivity. It is well known at this stage that the PSD of both the cosmic signal and the receiver noise is flat. This means that the signal spectrum delivered to the correlator is dictated by the shape of the receiver's frequency response. Ideally, the receiver must have a flat passband response. The loss in sensitivity caused by a non-ideal frequency response is quantified by the degradation factor, \mathcal{D} [15]. The degradation factor is specified as the ratio between the SNR for channel responses $H_1(\nu)$ and $H_2(\nu)$, as compared to the SNR obtained with an ideal rectangular passband of bandwidth $\Delta\nu$

$$\mathcal{D} = \frac{\text{Re} \left[\int_0^\infty H_1(\nu) H_2^*(\nu) d\nu \right]}{\sqrt{\Delta\nu \int_0^\infty |H_1(\nu)|^2 |H_2(\nu)|^2 d\nu}}. \quad (4.27)$$

In [15], numerous degradation expressions are derived by imposing different distortions such as passband ripple and amplitude slope variation, on (4.27). The expressions derived can be used to determine the maximum allowable passband distortion, for the various distortion effects. For

a maximum sensitivity loss of 2.5% ($\mathcal{D} = 0.975$), the maximum amplitude slope limit is 3.5 dB edge-to-edge, and the passband ripple limit is 2.9 dB peak-to-peak [15].

4.2.4 Signal transmission system

The purpose of the signal transmission system is to provide a low-loss and phase-stable path for the signals from the antennas to the correlator. The primary requirement for any signal transmission path is that it has sufficient bandwidth available to cope with the data rate, and suitably low attenuation over the length of the path, preferably without the need for repeaters [33]. For large synthesis arrays the design of a high performance and cost-effective signal transmission system is complex. In the past, most radio observatories made use of parallel wire transmission lines, waveguides, radio links and coaxial cable [33, 15]. The VLA for instance required signal paths of up to 22 km with bandwidths of 200 MHz. Because of the high attenuation losses of coaxial cable, the limited bandwidth of radio links and since optical communication technologies were not fully developed at that stage, they decided to use a circular waveguide transmission path [33]. Modern instruments seldom use anything other than optical fiber links. In comparison to the other technologies mentioned, optical fiber has unsurpassed bandwidth capabilities and the attenuation is as low as 0.2 dB/km near 1550 nm and 0.4 dB/km near 1300 nm [15]. The GMRT uses a fiber-optic link between the antennas and the central electronics building. The largest loss recorded is only ~ 11 dB, and this is for the most distant antenna (> 20 km) [4]. However, the design of optical fiber transmission systems is much more complicated than cable-based system. Furthermore, most amateur instruments have short baselines, and the use of low-loss coaxial cable is probably the easiest and most cost-effective solution [13, 11, 12].

For this thesis the signal transmission system is implemented with low-loss coaxial cable. After reviewing the coaxial products offered by the various vendors, it was decided to use the Times Microwave LMR-600 cable. This cable has exceptional performance with less than 10 dB attenuation per 100 m at 1.42 GHz, shielding of more than 90 dB and operates over a wide temperature range of -40°C to $+85^\circ\text{C}$.

4.3 Linearity

The receiver front-end is basically a cascade of quasi-linear two-port devices. The output voltage signal v_c can be expressed as a Taylor series expansion about the operating point, in terms of the input signal v_s

$$v_c = a_0 + a_1 v_s + a_2 v_s^2 + a_3 v_s^3 + \dots,$$

where a_0 is a bias (DC) term, a_1 is the linear voltage gain term, and all the remaining higher order components are distortion terms. Ideally, only the linear gain term, a_1 is desired. This is to ensure that the system remains linear over a wide range of noise temperatures, from the cold sky ($\sim 5 - 10$ K) to high antenna temperatures when observing strong sources such as the sun ($\sim 70\,000$ K). The linearity of the receiver directly impacts on the system calibration and accuracy. A useful approach for quantifying the linearity of a receiver, is to inject two closely

spaced tones. This situation is illustrative of a practical scenario of two adjacent channels in a multi-channel communication system. This then gives rise to harmonics and intermodulation products³. The second-order intermodulation components generally fall outside of the useful band. However, the third-order products are problematic, given that they are usually within the passband of the receiver. This is known as third-order intermodulation distortion. The third-order intercept point (IP3) is a useful metric for quantifying the linearity of a component or system in terms of third-order intermodulation.

The IP3 is defined as the fictitious power of one of the fundamental tones for which the sideband of a third-order spectral component has the same power [43]. It should be apparent that the concept of the IP3 is fundamentally based on monochromatic signals. For its applicability in radio astronomy it must be extended to address wideband (Gaussian) signals as well. In [43], a method is presented of linking the gain linearity performance of the receiver to the IP3. Thus, it is now possible to characterise the end-to-end receiver linearity using the IP3 for a Gaussian input signal. Fundamentally, this solution is based on the same expression used for characterising the reduction in sensitivity due to gain instability, $\Delta T/T_{\text{sys}} = \Delta G/G$. This is intuitive, since gain compression occurs once the receiver is driven nonlinearly, and this can cause differential gain errors ΔG , that cause inaccuracies (sensitivity reduction). After a rigorous statistical analysis a useful expression for the minimum output power P_{out} back-off with respect to the IP3, for a required gain linearity $\Delta G/G$, is given as

$$P_{\text{out}} \leq \frac{\Delta G}{G} \frac{\text{OIP3}}{4(1 - 10^{-DR/10})}, \quad (4.28)$$

where DR is the dynamic range of the antenna temperatures expected. For example from table 2.1, the sun is the strongest source with nominal flux density of 48 solar flux units (SFU)⁴, and Virgo A is the weakest source with a flux density of only 208 Jy at 1420 MHz. The DR is then found to ~ 1.8 dB for a bandwidth of $\Delta\nu = 256$ MHz, antenna gain of $G = 19$ dBi and effective system temperature of $T_{\text{sys}} = 130$ K. For an accuracy of 26.7 mK, the gain linearity is specified as 0.02%. Then from (4.28), it is seen that the output power P_{out} must be 38.3 dB less than the OIP3, to achieve this gain linearity. Alternatively, the equivalent 1 dB compression point expression, $P_{1\text{dB}} = \text{OIP3} - 10.6$, can be used to see what the output power back-off with respect to P1dB should be. In this case it is -27.7 dB. The IP3 of components are given by the manufacturer. IIP3 is usually specified for mixers and OIP3 for amplifiers [44]. The IP3 of an n -network cascade is easily computed using

$$\frac{1}{\text{OIP3}_{\text{cas}}} = \frac{1}{G_2 G_3 \dots G_n \text{OIP3}_{1,1}} + \frac{1}{G_3 G_4 \dots G_n \text{OIP3}_{2,2}} + \dots + \frac{1}{\text{OIP3}_{n,n}}, \quad (4.29a)$$

$$\frac{1}{\text{IIP3}_{\text{cas}}} = \frac{1}{\text{IIP3}_{1,1}} + \frac{1}{\text{IIP3}_{2,2}/G_1} + \frac{1}{\text{IIP3}_{3,3}/G_1 G_2} + \dots + \frac{1}{\text{IIP3}_{n,n}/G_1 G_2 \dots G_{n-1}}. \quad (4.29b)$$

This equation shows that to ensure a high degree of linearity the IP3 of the later stages in the cascade must be high and the excess gain low. It is also important to realise that the linearity

³Intermodulation products are the undesired spectral components that are produced when two input frequencies are mixed.

⁴1 SFU is equivalent to 10000 Jy, and this unit is primarily used to express the flux density of the sun.

Specification	Target	Explanation
Receiver gain, G_{rec}	67 dB	Section 4.4
Noise figure, NF	1.1 dB	Section 4.1.1
VSWR	$\leq 1 : 1.5$, (RL ≈ -14 dB)	-
Linearity, OIP3 _{rec}	+11.3 dBm	Section 4.3
Phase error, $\Delta\phi_{\text{rms}}$	$\leq 12.8^\circ$	Section 4.2.2
Gain slope edge-to-edge, ΔG_{rec}	3.5 dB	Section 4.2.3
Gain ripple peak-to-peak, G_{ripple}	2.9 dB	Section 4.2.3

Table 4.1: A summary of the required receiver specifications. This table also indicates the section in which an explanation is given for the respective specifications.

can not be increased by adding more components, it can only get worse. In a well designed system all the translated IP3s of the components should be equal and this will cause all the components to become non-linear at the same input power. It is seen that there is clearly a trade-off between NF which is determined by high gains early on in the cascade, and the linearity of the receiver.

4.4 Design details of the receiver front-end

This section presents the simulation and measurement details of the receiver front-end that was designed for the interferometer. A summary of the receiver target specifications are presented in table 4.1.

A simplified block diagram of the interferometer receiver system is illustrated in figure 4.7. Since the correlator is implemented digitally (see chapter 6), it is required that the signal noise temperature ΔT , including the system noise temperature T_{sys} , be amplified, filtered and additionally mixed such that the signal presented to the input of the Analogue-to-Digital Converter (ADC) can properly be digitised. For a T_{sys} of 130 K, the minimum detectable signal is $P_{\text{MDS}} = -94$ dBm, where $\Delta\nu = 256$ MHz. From table 2.1, the noise powers $P_{\Delta T}$ of the sun and Virgo A are -99.8 dBm and -131.2 dBm, respectively for $G = 19$ dBi. This shows that the signal of interest is buried well below the noise floor of the system, and this validates the need for a correlating interferometer. Before the required gain of the receiver can be defined it is necessary to decide how many bits will be used to digitise the astronomical signal. Theoretically digitisers with only 1 or 2 bits are needed for describing the weak sky noise temperature (see chapter 6) [45]. However, the presence of any strong RFI can cause the ADC to saturate, and this has to be accounted for. It is decided that only 3 bits ($\eta_q \rightarrow 1$) will be used for the astronomical signal (equivalent system noise), and the remaining bits are available for any RFI. The digitiser used is the KatADC (see chapter 6), which is an 8-bit ADC with a full-scale input voltage of 650 mV_{pp}. The resolution of the ADC is thus $\Delta v = 2.55$ mV_{pp}. Digitising a Gaussian input signal (CF = 4.42) with 3 bits, presents an input power of -45.5 dBm (100 Ω) to the ADC.

However, the KatADC has a 20 dB pre-amplifier and 31.5 dB variable attenuator that has been pre-configured to toggle 3- to 4-bits for an input power level of -27 dBm. The required power gain of the receiver is then found to be 67 dB. The KatADC uses the National Semiconductor ADC08D1520. The effective number of bits for this part is 7.4 at 748 MHz [46]. This gives an approximate dynamic range of 25 dB for any RFI before the ADC starts saturating.

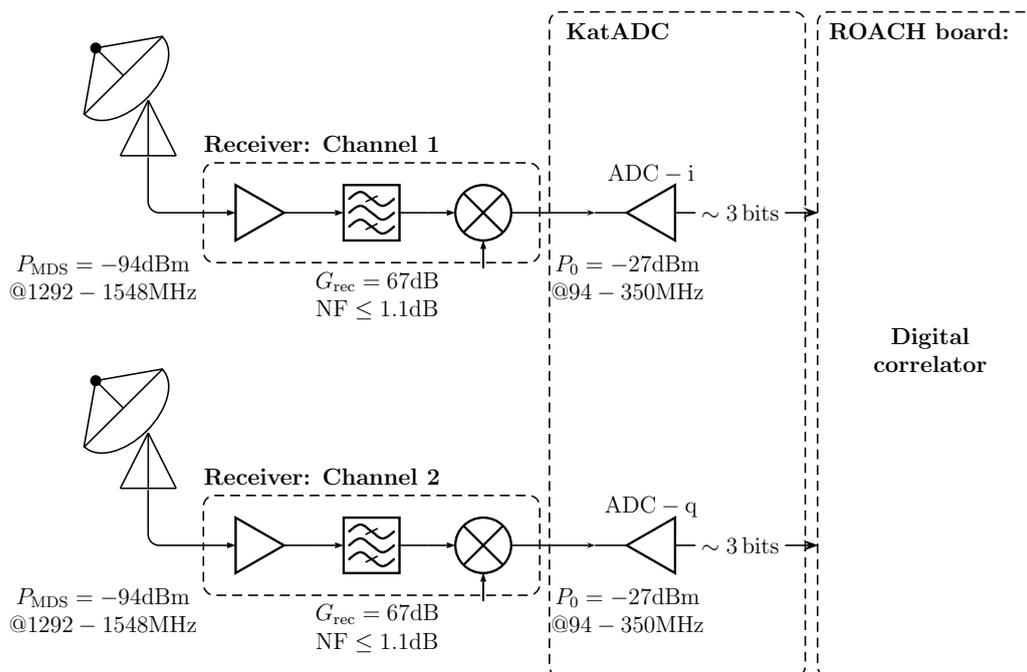


Figure 4.7: A simplified block diagram of the interferometer receiver system.

4.4.1 Development of the down conversion scheme

The RF band $\Delta\nu_{\text{RF}}$ for the interferometer is from 1292 MHz to 1548 MHz, and it is centered at 1420 MHz. The KatADC has a maximum dual channel sampling rate of 1.5GSamples/s, and is further capable of operating as a single converter with a sample rate of 3GSamples/s [46]. Since, only a single KatADC was available the option of digitising at 3GSamples/s was not an option. Unfortunately, the 1.5GSamples/s rate is too low for Nyquist sampling. There are two versions of the KatADC- one using the SBB-2089Z (50-850 MHz) pre-amplifier and the other using the much wider band SBB-5089Z (50-6000 MHz) amplifier [76]. If the later is utilised, then sub-Nyquist (bandpass) sampling may be considered, since the KatADC (ADC08D1520) has a 2 GHz full power bandwidth [46]. However, The KatADC with the 50 – 850 MHz pre-amplifier was used. It was then decided that the RF signal band would be mixed down to a lower more manageable frequency range. The receiver output band $\Delta\nu_0$ was chosen as 94 – 350 MHz, centered about 222 MHz. The lower output frequency limit was chosen to be above the 50 MHz restriction imposed by the KatADC pre-amplifier. Since, the KatADC was clocked at 800 MHz, the upper output frequency limit had to be ≤ 400 MHz to adhere to the Nyquist sample rate.

The 800 MHz sampling rate of the ADC was mainly defined by the clock rate of the FPGA. The FPGA can be clocked off the same 800 MHz source, which is then decimated by a factor of 4, resulting in a 200 MHz clock frequency. This is a compromise between meeting performance demands and making the timing constraints of the FPGA code easier to meet. The bandwidth of the receiver is relatively wide, and is certainly more than an octave at the output frequency. This makes preventing, or more realistically minimising spurious signals from falling in band more of a challenge. The first task for the receiver design was to consider the various alternative downconversion schemes.

The first scheme investigated was a direct downconversion superheterodyne receiver, where $\Delta\nu_{\text{RF}}$ was mixed down to $\Delta\nu_0$ with a single frequency conversion stage. This is the most economical solution in terms of component count. Considering a HSLO, the LO frequency is at $\nu_{\text{LO},1} = 1642$ MHz. An inherent problem of HSLO downconversion is that it has poor image and LO suppression, especially if the ratio between the RF and IF frequencies is large, $\nu_{\text{RF}} \gg \nu_{\text{IF}}$ [47]. Another limitation of a large ratio HSLO downconverter, is the increased susceptibility to single frequency spurs falling within the passband, $m\nu_{\text{LO}} + n\nu_{\text{RF}}$, where $m, n = \pm 1, \pm 2, \pm 3, \dots$ and simultaneously $|m| = |n| \neq 1$, where the $\pm 2 \times \pm 2$ spurs parallel the desired response [44]. The rejection of all even inter-modulation spurious signals can be increased by using a double-balance mixers [42, 44]. For instance the ZX05-73L double-balanced mixer from Mini-Circuits, has a spurious signal level of < -61 dBc for the $m \times n = 2 \times 2$ harmonic, this assumes that $P_{\text{LO}} = +4$ dBm and $P_{\text{RF}} = -14$ dBm. This spur will even be lower if P_{RF} is less, an estimation of the actual spurious signal level for the n^{th} harmonic, can be computed using [44]

$$\Delta R_{mn} = (|n| - 1) [\Delta P_{\text{RF}} - \Delta P_{\text{LO}}] \text{ dB}, \quad (4.30)$$

where ΔP_{RF} is the change in signal level, ΔP_{LO} the change in LO level. However, the biggest problem with this conversion scheme is that the image frequency ν_{IM} is situated around 1864 MHz, 1736 MHz – 1992 MHz. From the RFI survey results presented in figure 2.5, the upper GSM band is 1805-1880 MHz, and will thus be mixed into the IF band. At least 55 dB attenuation is needed to suppress the signal levels at 1850 MHz to below the noise floor of the receiver. Such a steep cutoff filter is difficult to achieve at L-band, and is furthermore not readily available off-the-shelf. Using a LSLO will not perform any better, since they both have similar drawbacks, and the image frequency will now rather be at the lower GSM band, which is even worse.

In an attempt to alleviate the RFI problem, the receiver design was forced to use a dual-conversion superheterodyne receiver. The idea is to upconvert first, where it is then easier to design high selectivity cavity filters. This will provide better attenuation for the RFI signals in the image band. Finally the higher IF frequency ν_{IF} , is downconverted to the output frequency, $\nu_0 = 222$ MHz. Care must be exercised when designing an upconverting system. This is to ensure that the RF harmonics are not located inside the passband. A commonly used design criterion is to ensure that the lower edge of the IF passband is at least 2.1 times above the upper edge of the RF passband, $\Delta\nu_{\text{IF,L}} \geq 2.1\nu_{\text{RF,H}}$ [47]. Since, $\nu_{\text{RF,H}} = 1548$ MHz, therefore $\Delta\nu_{\text{IF,L}} \geq 3250$ MHz. The intermediate band was chosen to be centered about 3500 MHz, because of the availability of good cavity filters at this frequency. This meant that the 256 MHz

bandwidth extended from 3372 MHz to 3628 MHz. This is above the 3250 MHz restriction. The fixed LO frequency needed to mix the RF band up from 1420 MHz to 3500 MHz, is $\nu_{LO,1} = 4920$ MHz. The receiver thus uses a HSLO for the upconversion, which causes frequency inversion. The image frequency ν_{IM} is 8292-8548 MHz, centered at 8420 MHz, and is easily removed using a lowpass filter (LPF). Another advantage of upconverting is that any LO leakage and LO re-radiation are effectively suppressed with a LPF. The linear spur plot normalised to $\nu_{LO,1} = 4920$ MHz, for the upconverting stage is shown in figure 4.8. Only the harmonics up to, and including the 6th-order are present, $m, n \leq 3$. It is seen that the 2nd-order RF harmonic is out of band, as well as the 3rd-order RF harmonic. The passband is clear of any low order harmonics ($|n| + |m| \leq 6$) which is good. There are higher-order harmonics (not shown in figure 4.8) that potentially will fall inside the 3372 MHz – 3628 MHz band, but will be of little consequence, given their low signal levels.

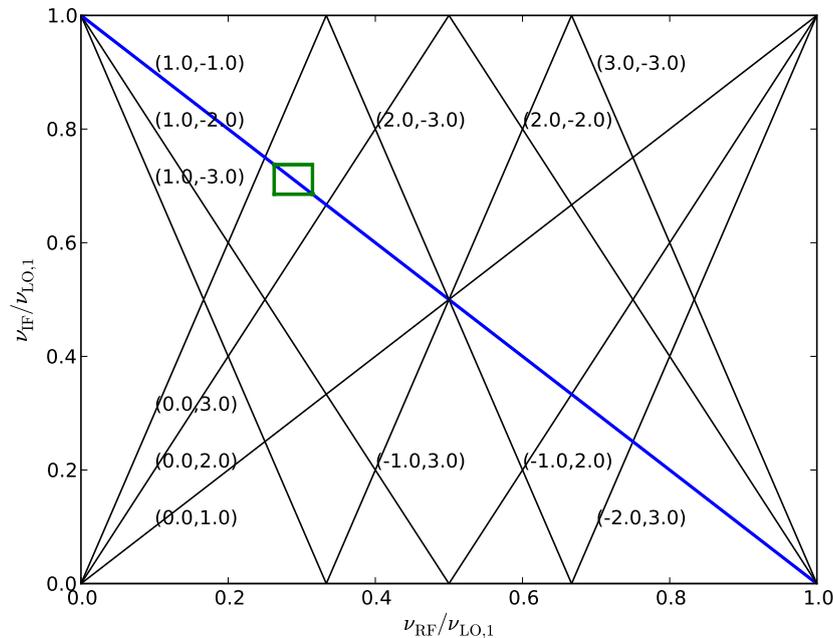


Figure 4.8: The linear spur plot normalised to $\nu_{LO,1} = 4920$ MHz, for the upconverting stage. Only harmonics up to and including the 6th-order are present, $m, n \leq 3$. The RF band extends from 1292 to 1548 MHz, and the IF band from 3372 to 3628 MHz.

Finally, the 3500 MHz centered band is downconverted to the required 222 MHz centered output band. This is accomplished by using a fixed LSLO $\nu_{LO,2}$ of 3278 MHz. The image frequency is now from 2928 MHz to 3184 MHz, centered at 3056 MHz. The upper GSM band is also mixed up, and still appears in the image band. However, at a much reduced power level given the filtering action applied before upconversion, and then again before mixing down to the output band. For the downconversion stage the same harmonics appear in the passband as for the direct downconverting case. This is shown in the spur plot of figure 4.9. The spur signal levels

are -61 dBc and -64 dBc for $-2\nu_{LO,2} \times 2\nu_{RF}$ and $-3\nu_{LO,2} \times 3\nu_{RF}$, respectively, for the ZX05-73L double balanced mixer. These levels are low enough to avoid any signal distortion (beating). Both harmonics are seen to be present only over a small region close to the lower-edge of the passband.

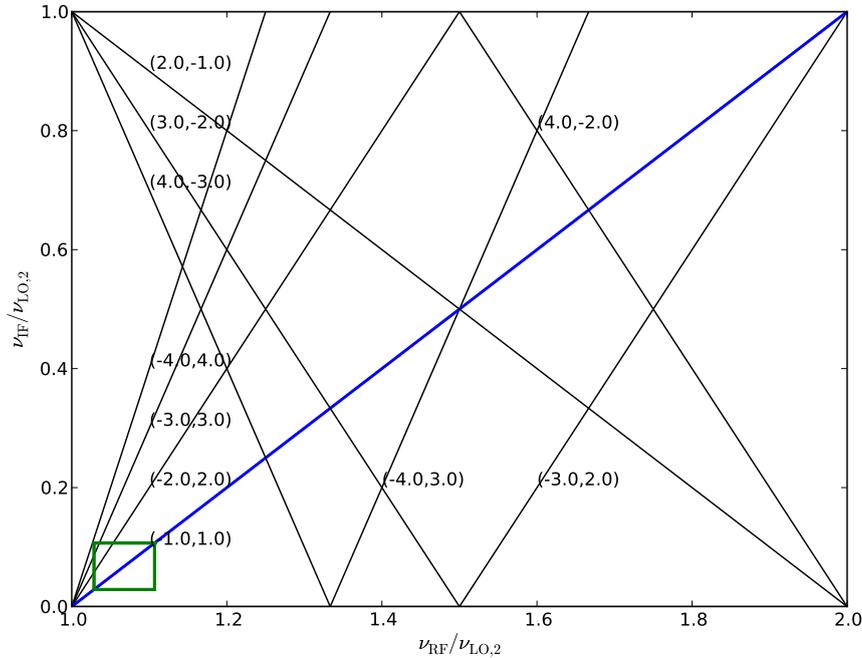


Figure 4.9: The linear spur plot normalised to $\nu_{LO,2} = 3278$ MHz, for the downconverting stage. Only harmonics up to and including the 8th-order are present, $m, n \leq 4$. The RF band extends from 3372 to 3628 MHz, and the IF band from 94 to 350 MHz.

This was the chosen frequency conversion scheme, and figure 4.10 summarises this frequency conversion process. The design of the dual-conversion superheterodyne receiver is discussed in the next section.

4.4.2 Behavioural design of the dual-conversion superheterodyne receiver

A simple receiver model was developed in Microwave Office (MWO ⁵) based on the cascaded network equations (4.24) and (4.29b). The general outline of the receiver architecture is shown in figure 4.11. The MWO model made it possible to iteratively populate the general architecture with different components and layout orders. This made it possible to investigate the

⁵Microwave Office is an IC, PCB and module design suite that includes linear as well as non-linear circuit simulators, EM analysis, integrated schematic and layout support and more. MWO is built on the AWR high-frequency design environment platform, <http://web.awrcorp.com/Usa/Products/Microwave-Office/>

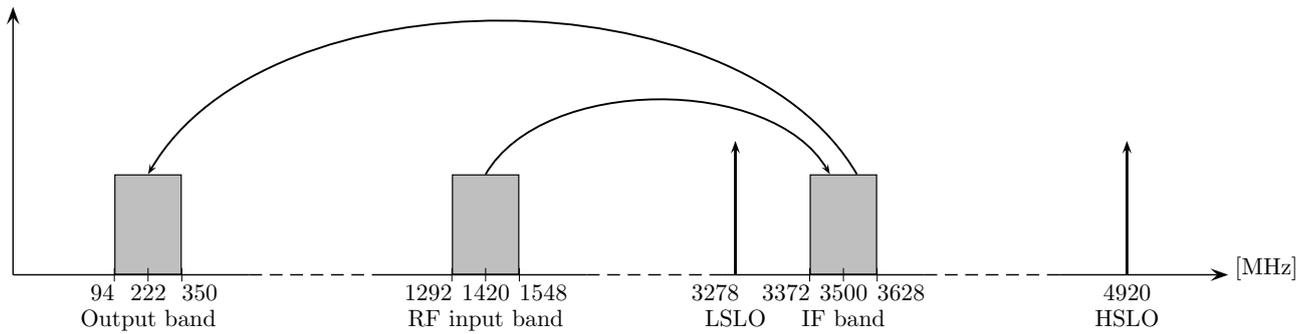


Figure 4.10: The frequency mapping scheme of the dual-conversion superheterodyne receiver.

performance of different implementations and to decide on a best solution.

Figure 4.11 shows that the line-up of some of the components are fixed. These components perform a very specific function, and are crucial in terms of receiver performance. Upfront, an LNA with enough gain is needed to characterise the NF of the whole receiver. This is followed by a BPF, which is used to reject some of the output of band RFI. This is to prevent any strong signals from saturating the components, especially in the later stages. The selectivity BPF ultimately defines the receiver processing bandwidth, and in this case is 256 MHz. The LPF at the receiver output is the anti-aliasing filter. Figure 4.11, further shows that the amplifiers are distributed throughout the receiver chain, in order to keep gain up and simultaneously the NF low. The design strategy followed consisted of

- ensuring that throughout the receiver chain $G_1 + NF_1 > NF_2 + 15$ [dB],
- further down the chain the higher the IP3 of the components must be,
- and that the matching between the components are better than a VSWR of 1:1.5.

Another important aspect is appropriate filtering near the mixers. If a broadband amplifier precedes the mixer then image noise could be a serious problem. It is mentioned that the NF of the receiver can increase by as much as 3 dB in this case [44]. The effective noise factor of a mixer with image noise is given as

$$f_{\text{eff,mixer}} = f_{\text{mixer}} + (f'_{n-1}g'_{n-1} - 1) \frac{g'_{\text{mixer}}}{g_{\text{mixer}}}, \quad (4.31)$$

where the primes are used to represent the parameter values at the image frequency. From (4.31) it is evident that an amplifier ($f'_{n-1} \ll g'_{n-1}$) will cause the effective NF of the mixer to increase. However, if a suitable filter is used before the mixer then the image noise can effectively be suppressed. A filter preceding the mixer can serve an additional purpose of attenuating the LO leakage from being re-radiated. Filtering after the mixer is also generally needed to suppress the LO and RF leakage into the IF stage. The LO leakage can cause compression in the components to follow if not properly suppressed, since P_{LO} is generally high.

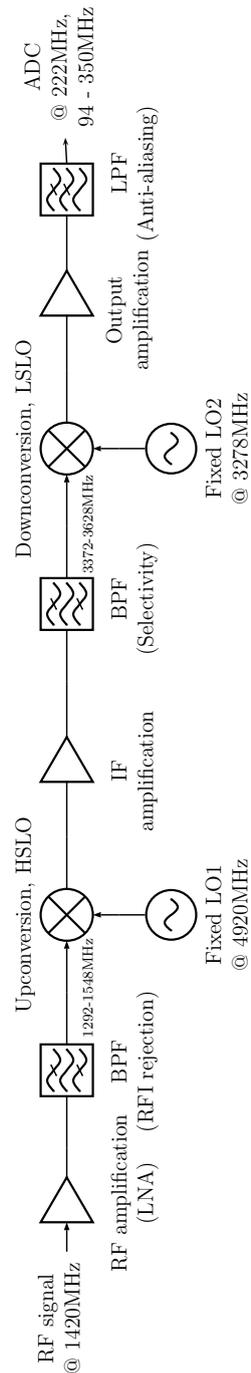


Figure 4.11: A simplified block diagram of the dual-conversion superheterodyne receiver architecture.

The receiver was designed and built using components off-the-shelf (COTS). Suitable components were thus sourced which would fulfil the specifications presented in table 4.1. The analytical design approach using the simplified receiver model, led to an initial receiver design with an estimated $G_{\text{rec}} = 67$ dB and $\text{NF} = 1$ dB. This was followed by simulating the non-linear

performance of the receiver, using the Harmonic-Balance (HB⁶) solver in MWO. This provided an estimation for the 1 dB compression point, $P_{1dB} = -50.9$ dBm, the IIP3 = -42.7 dBm, and the spurious free dynamic range, $SFDR = 34.1$ dB. The simulated and measured results for the RF, Intermediate and Output stages are presented in the sections to follow.

4.4.3 Linear performance of the RF stage

A block diagram of the input RF stage (1292 – 1548 MHz) is given in figure 4.12, and a summary of the electrical specifications for the component line-up is presented in table 4.2.

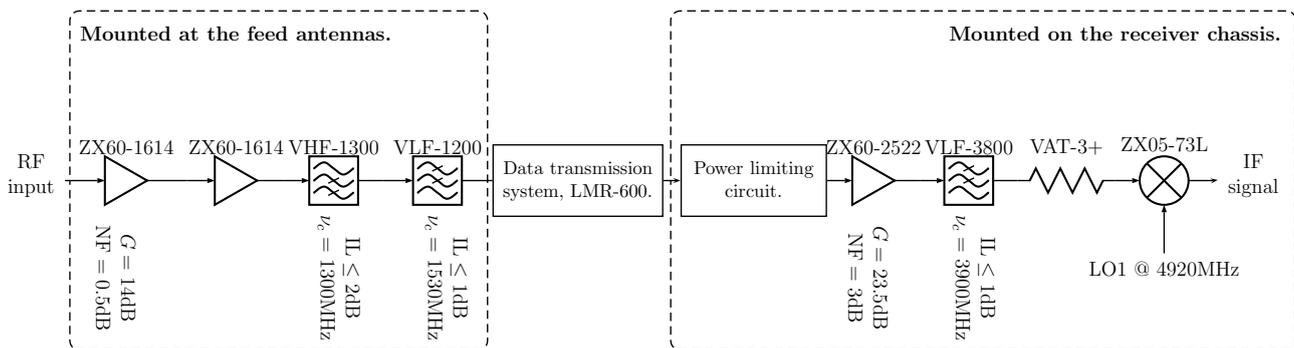


Figure 4.12: A block diagram of the input RF stage, $\nu_{RF} = 1292 - 1548$ MHz. The figure additionally shows that the low-noise front-end is situated at the output of the feed antenna, and the remainder of the RF stage circuitry is attached to the receiver chassis. The signal transmission system is implemented with low-loss LMR-600 coaxial cable. The power-limiting circuit protects the receiver against exceedingly high input power levels, and is more fully discussed in appendix C.

To achieve the best possible receiver noise temperature, the RF stage must have an LNA with a low NF right upfront at the antenna feed. The Mini-Circuits ZX60-1614LN LNA has an exceptionally low NF of typically 0.5 dB. However, this generally requires a trade-off between the NF and available gain. The ZX60-1614LN LNA is no exception, with a gain of only $14 \text{ dB} \pm 1.1 \text{ dB}(\text{typ.})$. In order to maintain $G_1 + NF_1 > NF_2 + 15$ [dB], it is required that at least two of these amplifiers be used. Generally, LNAs also have relatively poor linearity, the ZX60-1614LN has an OIP3 of +30 dBm, and a P1dB of 13.5 dB, referenced to the output. Since, these amplifiers are relatively broadband it is important to have some filtering upfront to prevent linearity degradation in the presence of strong RFI. To this end, the results of the RFI survey were used in selecting the appropriate filters. As previously mentioned, there is a limited selection of steep cut-off filters near 1-2 GHz. In this regard, a high-pass and a low-pass filter was cascaded to approximate a bandpass response. The Mini-Circuits VLF-1200 and VHF-1300 filters are used, with respective 20 dB attenuation stopbands at 1865 MHz, and 930 MHz. Both are 7th-order filters. An additional line amplifier is needed to compensate for

⁶The Harmonic-Balance solver in MWO is a non-linear frequency domain technique that provides a steady-state solution.

the losses in the signal-transmission system (LMR-600), which over a 35 m distance only has a loss of ~ 3.7 dB at 1420 MHz, and the conversion loss of $L_c \approx 6$ dB, of the ZX05-73L double balanced mixer. The Mini-Circuits ZX60-2522M (500-2500 MHz) amplifier is used and placed after the power-limiting circuit. The ZX60-2522M has a nominal gain of $G = 23.5$ dB. At this stage components with higher NFs can be chosen, where the ZX60-2522M amplifier has a typical NF of 3 dB.

The measured return loss (RL) and the gain for the RF stage are shown figure 4.13. The input RL (S_{11}) is seen to be well matched across the operating band, where the RL was measured as -37 dB, -19.9 dB and -18.4 dB at 1292 MHz, 1420 MHz and 1548 MHz, respectively. Similarly, the output RL (S_{22}) is well below -20 dB across the passband. The RL measured for S_{22} is better than S_{11} primarily because of the 3 dB attenuator at the output. The receiver gain (S_{21}) is seen to have a gradual slope across the passband. The gain at the lower- and upper-edge of the passband is measured as 41.5 dB and 36.9 dB, respectively. The maximum variation in the gain across the band is seen to be $\Delta G = 5.2$ dB. This is higher than the required 3.5 dB specification, but it will be shown that the IF passband of the receiver is less than the 256 MHz. As a result, the ΔG_{rec} requirement of the receiver is met. The gain roll-off of the amplifier cascade is only ~ 1.3 dB. Thus, this main contributor to the measured slope across the passband is the two cascaded filters. Unfortunately, the cut-off frequencies of the two filters are fairly close, and this causes the uneven frequency response observed in figure 4.13. To investigate the RFI rejection effectiveness of the input stage, the maximum RFI levels from figure 2.5 were considered. At ~ 930 MHz the maximum signal level was measured as -32 dBm. From figure 4.13 the attenuation is measured as 25.06 dB, relative to the normalised gain at 1420 MHz. At the upper GSM band the largest signal power is measured as -45 dB at 1850 MHz. The rejection measured at 1850 MHz is more than 32 dB. The amount of RFI rejection is high enough to avoid driving the RF stage into compression. However, these power levels are still much stronger than any astronomical signal and additional filtering is a necessity.

The two ZX60-1416LN amplifiers, the VLF-1200, and the VHF-1300 filters are combined to form the low-noise front-end of the RF stage. It is indicated in figure 4.12 that this part of the receiver is placed right upfront at the antennas. Since, both the NF and linearity of the receiver are of importance, two different configurations for this low-noise front-end was considered, see figure 4.14. The first design alternative places the filter cascade between the two amplifiers. This configuration should have better linearity for large signal levels. However, since the ZX60-1614LN has a typical gain of 14 dB, and the IL of the filter cascade is ~ 3 dB, the NF performance will be somewhat worse. The other configuration places both amplifiers upfront, this configuration is more susceptible to saturation, but has superior NF performance. To quantify the performance trade-off between the two topologies, the NF for both configurations were measured, and the results are shown in figure 4.15. It is evident from figure 4.15(a) that the NF of the second configuration is better. At the center frequency, the respective NFs for topologies 1 and 2 are measured as 0.625 dB and 0.839 dB. It is also seen that the NF in both cases tend to increase near the upper edge (1548 MHz), where the NF of topology 1 is as high as 1.06 dB. The higher NFs is to be expected because of the gain slop across the passband. Figure 4.15(b) shows that the gains at the upper frequency edge are 18.44 dB and 17.8 dB for

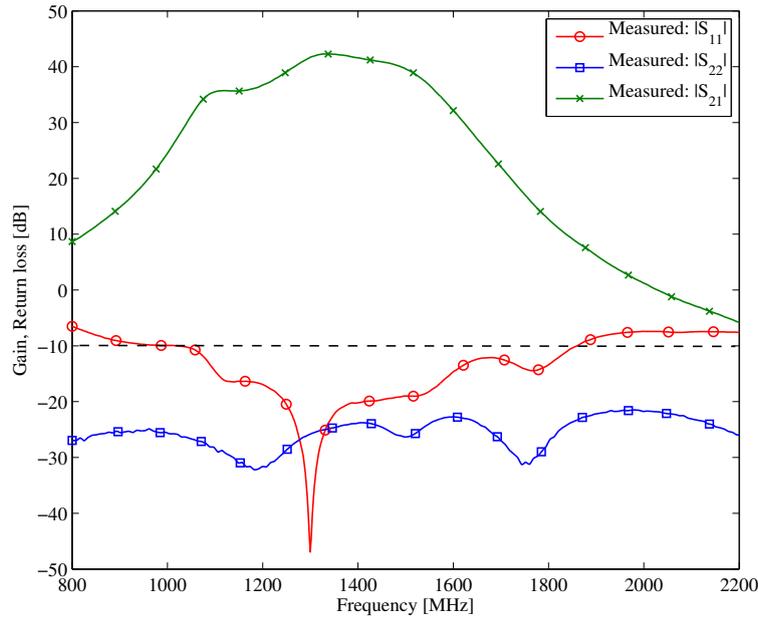
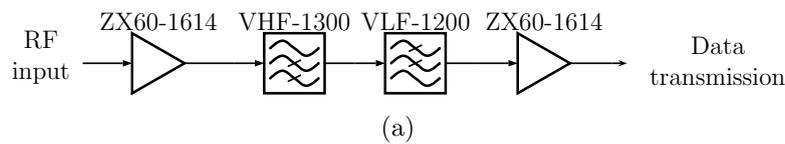


Figure 4.13: The measured S-parameters of the RF stage from 800 to 2200 MHz. The insertion loss of the power limiting circuit is accounted for, but not for the LMR-600 cable.

Topology 1:



Topology 2:

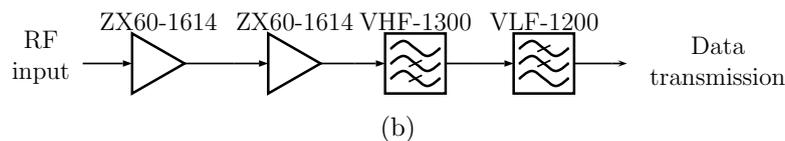
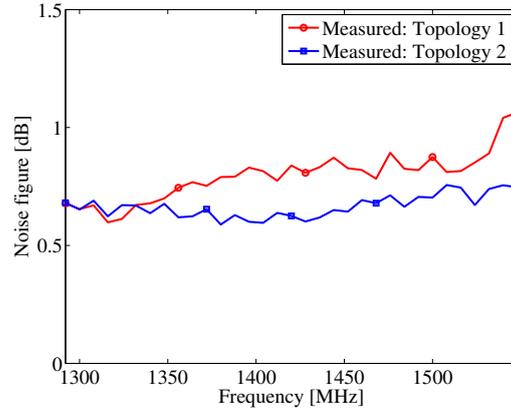
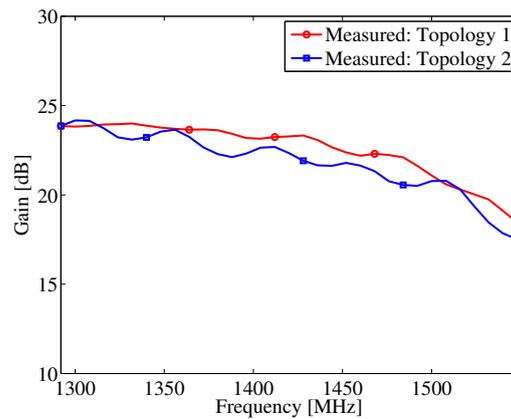


Figure 4.14: The low noise front-end topologies considered. The first option (a) places the filters between the two LNAs (topology 1), and the (b) second design cascades both LNAs upfront, followed by the filters.

topologies 1 and 2, respectively. This is lower than the gains observed at 1420 MHz, of 22.34 dB and 23.27 dB. A similar gain roll-off is shown in figure 4.13. In terms of linearity, it was mentioned that there is more than 30 dB attenuation for the out of band RFI, and this ensures that both configurations are operated in the linear region. Thus, it was decided to implement the low-noise front-ends using the second configuration.



(a)



(b)

Figure 4.15: The measured (a) noise figure and (b) gain of the two low-noise front-end topologies shown in figure 4.14.

4.4.4 Linear performance of the IF Stage

After upconverting, the layout details of the IF stage (3372-3628 MHz) is shown in figure 4.16, and the measured S-parameters are further presented in figure 4.17. The Mini-Circuits VLF-3400 LPF is positioned after the upconverting mixer to suppress the LO1 leakage signal. The upconverting mixer is driven by $P_{LO1} = +7$ dBm and the LO/RF isolation was measured as 28.5 dB, see section 4.4.6. The VLF-3400 provides a further rejection of ~ 25 dB, resulting in an LO1 leakage level of -47 dBm. The intermediate amplifier after the LPF is the ZX60-3800LN (Mini-Circuits). This amplifier operates over a frequency range from 3300-3800 MHz, with a flat gain response of $G = 24$ dB, ± 0.1 dB. The gain response of the IF stage in figure 4.17 is shown to be very flat, with only about 0.18 dB ripple across the required 3372-3628 MHz band. The maximum gain is measured as nominally 17 dB across the band. The ZX60-3800LN has both a higher IP3 and P1dB as compared to the LNA used in the RF stage, the values are +36 dBm and +18 dBm, respectively. Referring to figure 4.16, it is seen that a bandpass filter follows the intermediate amplifier. The BF3500M200 (Dytech) BPF defines the selectivity

of the interferometer. The 3 dB-bandwidth of the BF3500M200 filter, and hence of the receiver is measured as 250 MHz (3380-3630 MHz). Moreover, the steep skirts of this cavity filter provides the required attenuation needed to suppress the out-of-band RFI. It was mentioned that the upper GSM band is situated in the image band (2928-3184 MHz) of the receiver. Upon downconversion these RFI signals are mixed into the output passband of the receiver. The image rejection of the IF stage is better than -74 dBc. This suggests that the RFI levels are suppressed to below -90 dBm. Moreover, the 2nd-order harmonic of the RF signal, appearing at 3096 MHz, can be as high as -56 dBc. The BF3500M200 filter also effectively suppresses this spurious signal. The input return loss S_{11} is below -13 dB across the band, and is measured as -16 dB at center frequency of 3500 MHz. The output return loss S_{22} is also seen to be matched better than -18 dB over the entire passband.

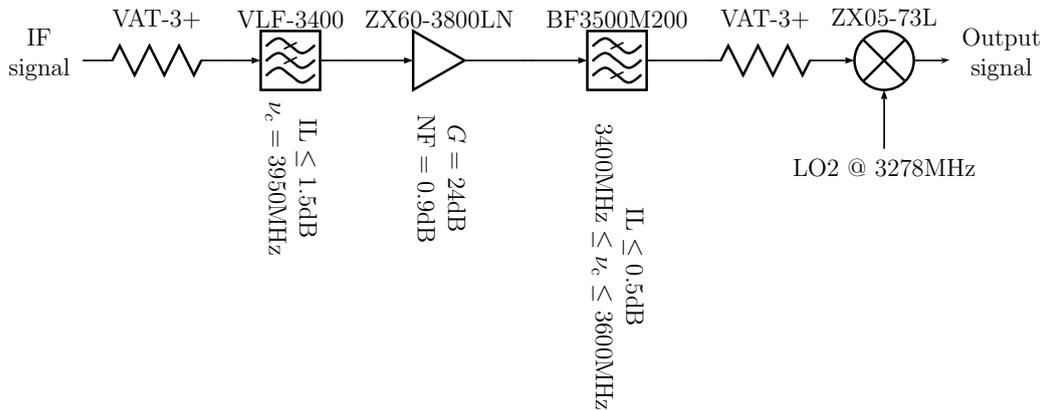


Figure 4.16: A block diagram of the IF stage, $\nu_{\text{IF}} = 3372 - 3628$ MHz. A summary of the electrical specifications of the IF stage components are given in table 4.2.

4.4.5 Linear performance of the output stage

The component line-up for the output stage (94-350 MHz) is shown in figure 4.18. Furthermore, the measured S-parameters associated with this stage is given in figure 4.19. In figure 4.18 it is shown that a LPF is needed after the downconverting mixer. The VLF-400 (Mini-Circuits) filter serves the same purpose as the LPF filter used in the IF stage. The LO2 leakage in this case is suppressed by 60 dB, this includes the LO/IF isolation of 25 dB and the filter attenuation of ~ 35 dB. The ZFL-500HLN (Mini-Circuits) is the final amplifier in the receiver chain. It has a gain of 19 dB, ± 0.4 dB and a high IP3 of +30 dBm. From figure 4.19, the gain is measured as 18.7 dB, 18.07 dB and 15.6 dB at 94 MHz, 222 MHz and 350 MHz, respectively. This indicates the gain variation across the band is only 2.48 dB, which is in compliance with the $\Delta G_{\text{rec}} \leq 3.5$ dB specification. It is necessary at this stage to apply filtering which will prevent aliasing during the signal digitisation. The VLFX-225 (Mini-Circuits) LPF is used in this regard. This filter must be able to attenuate the signals above the Nyquist frequency (400 MHz) to less than the quantisation noise level ($< 1/2\text{LSB}$). The minimum required stopband attenuation must typically be

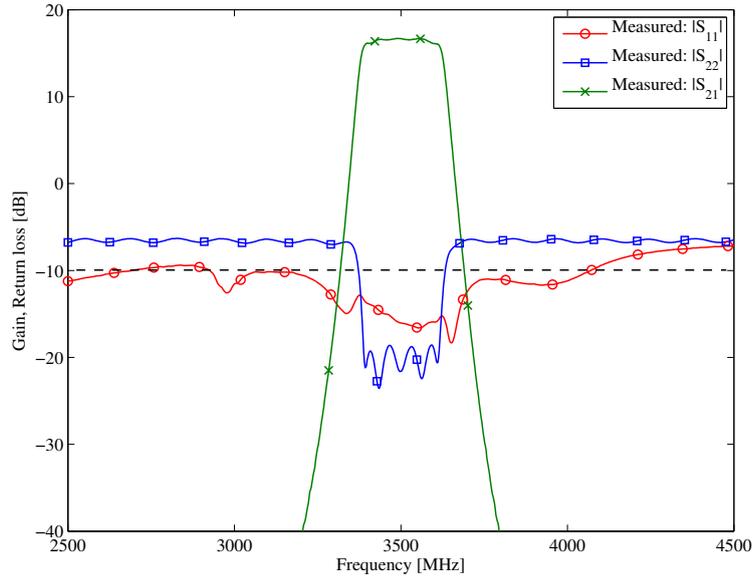


Figure 4.17: The measured S-parameters of the IF stage from 2500 to 4500 MHz.

$$A = 20 \log(\sqrt{1.5} \times 2^B), \quad (4.32)$$

where B is the number of bits in the ADC. The KatADC uses 8-bits and according to (4.32) the attenuation must at least be 50 dB at 400 MHz, assuming a continuous wave signal. From figure 4.19, the attenuation at 400 MHz is shown to be 10 dB. This is somewhat low, but the signal level into the ADC is only -27 dBm, which suggests that only 20 dB attenuation is needed. The rejection at 450 MHz is 36 dB, which is much better. The input return loss is measured as below -25 dB from 94-350 MHz. The output return loss is below -20 dB across most of the band, but as shown in figure 4.19, the return loss decreases to -12 dB at the upper-band edge.

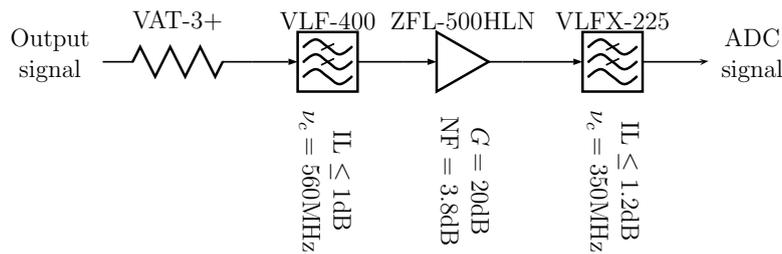


Figure 4.18: A block diagram of the Output stage, $\nu_0 = 94 - 350$ MHz. A summary of the electrical specifications of the Output stage components are given in table 4.2.

4.4.6 Characterisation of the frequency converters

The design details for the LOs are discussed in this section. According to figure 4.10, two local oscillators must be generated at 4920 MHz, and at 3278 MHz, for the upconversion- and

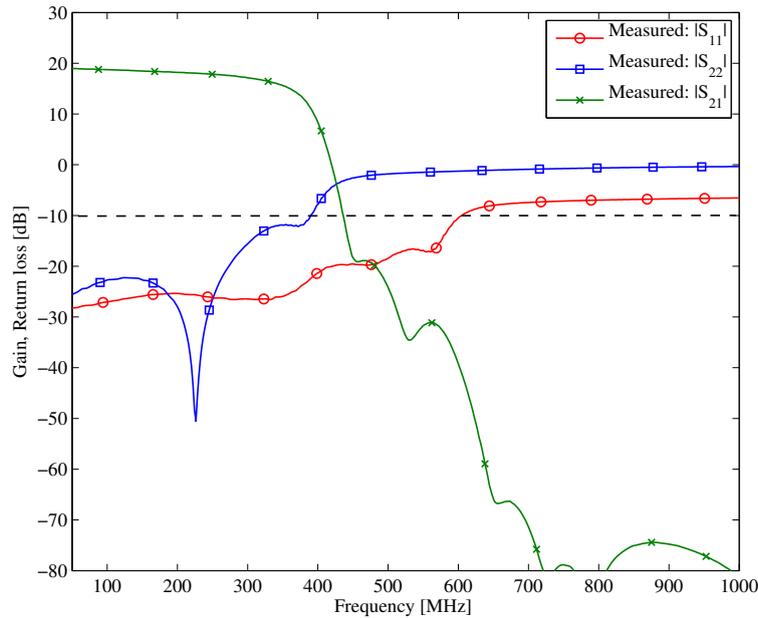


Figure 4.19: The measured S-parameters of the output stage from 50 to 1000 MHz.

downconversion-stages, respectively. Key specifications of the LO signals are the power levels and frequency stability (phase noise). However, before the LOs can be designed, it was required that the mixers be fully characterised.

The decision of what mixers to use was primarily based on the expected RF input power P_{RF} . Thus, the lowest LO power P_{LO} mixers, were chosen for which the 1 dB compression point exceeded the anticipated maximum RF signal level. This in turn also ensures a minimum LO leakage signal. In this case, the Mini-Circuits ZX05-73L double-balanced mixer, with $P_{\text{LO}} = +4$ dBm, was chosen for both the downconverting and upconverting mixers. For the upconverting case, the RF port of the mixer is the output, and the input signal is applied to the IF port. A summary of the electrical specifications of the ZX05-73L mixer is presented below:

- The RF/LO bandwidth is from 2400-7000 MHz.
- The IF bandwidth extends from DC-3000 MHz.
- The conversion loss is $L_c \leq 6.2$ dB.
- The LO/RF Isolation is specified as > 26 dB (30 dB typ.).
- The LO/IF Isolation is better than > 21 dB and typically 26 dB.
- See table 4.2 for some of the other specifications, such as VSWR, P1dB, etc.

These specifications are given for an IF of 30 MHz, and at the midrange frequencies. To meet the receiver's design specifications it was required to measure the performance parameters of

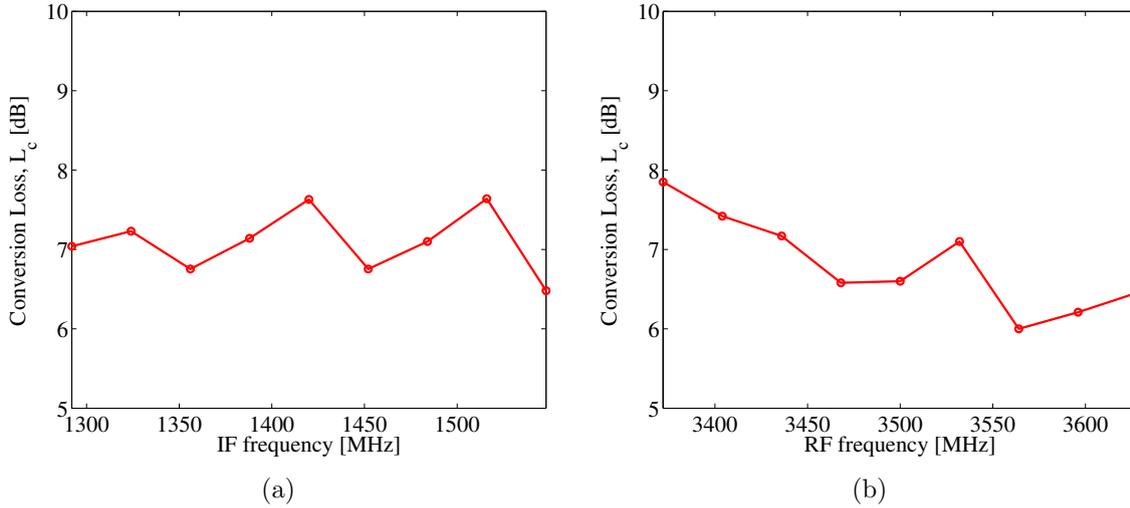


Figure 4.20: The measured conversion loss L_c of the Mini-Circuits ZX05-73L double-balanced mixer for both the (a) upconverting and (b) downconverting cases. The P_{LO} for the upconverting and downconverting cases are +7 dBm and +4 dBm, respectively. The input power in both cases were fixed at -14 dBm. To improve the VSWR at the mixer ports 10 dB attenuators were added.

the ZX05-73L mixer under the operating conditions it will be used for. This made it possible to more accurately account for the conversion loss and LO leakage. The measured conversion loss, L_c for both the upconverting and downconverting mixers are shown in figure 4.20. It is seen that the L_c is higher than the specified value of 6.2 dB. However, in both cases the conversion loss is nominally about 7 dB across the band. The isolation between LO/RF and LO/IF were measured as 28.5 dB and 17.8 dB for the upconverting mixer. Whereas, the downconverting mixer has an LO/RF isolation of 30.22 dB, and an LO/IF isolation of more than 25 dB. These values are important for selecting the appropriate LO rejection filters.

A mixer inherently has poor matching at its ports, where the VSWR can easily exceed 3 ($RL < -6$ dB). It is common practice to place attenuators between the mixer ports and the adjoining components. This effectively masks the impedance mismatches, and a lower VSWR can be achieved. The improved return loss RL' is given as

$$RL' = RL + 2\text{Atten [dB]}, \quad (4.33)$$

where Atten is the value of the attenuator. It is easily seen from (4.33) that the impedance mismatch can indefinitely be improved by increasing the attenuator value. However, in a receiver there is an upper limit on the attenuator size that can be used, since the gain decreases, and this can potentially result in a degradation in the NF. The return loss for both the up- and down-converting mixers were measured, and the results are presented in figures 4.21 and 4.22.

For the upconverting mixer in figure 4.21, it is shown that without the 3 dB attenuators the ports are poorly matched and the return loss is mostly above -10 dB. For the IF port the return loss is even as high as -6 dB. By placing a 3 dB attenuator before the mixer ports, a significant

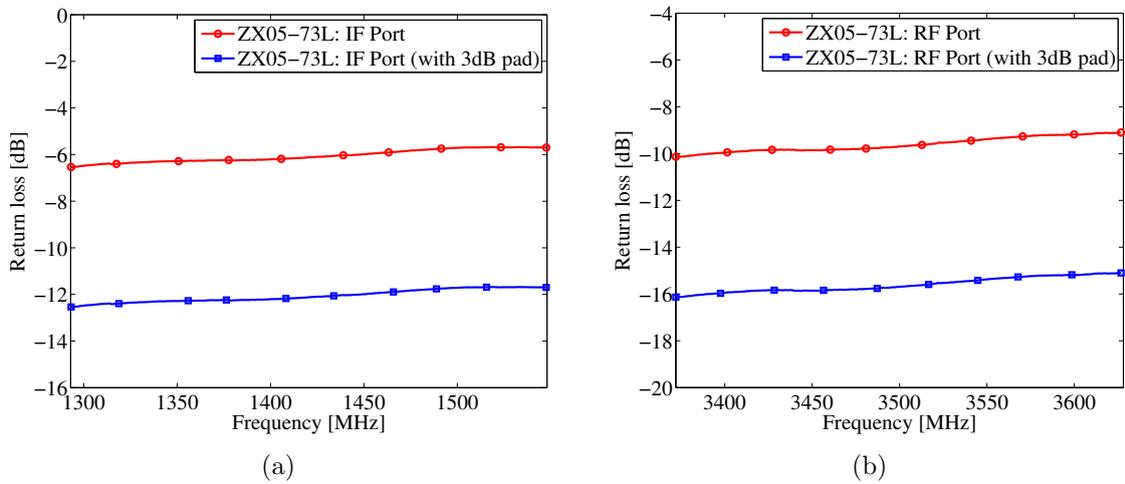


Figure 4.21: The measured return loss of the ZX05-73L upconverting mixer at the (a) IF and (b) RF ports. It is also shown that a 6 dB return loss improvement is achieved by placing 3 dB pads before the mixer ports.

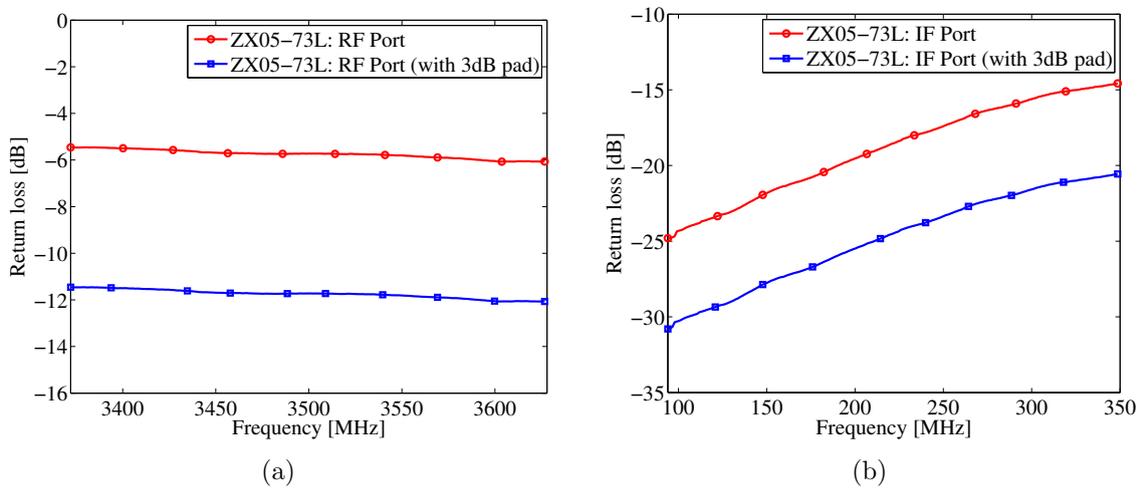


Figure 4.22: The measured return loss of the ZX05-73L downconverting mixer at the (a) RF and (b) IF ports. It is also shown that a 6 dB return loss improvement is achieved by placing 3 dB pads before the mixer ports.

improvement in the return loss is seen. For the RF port in figure 4.21(a), the return loss is seen to be near -16 dB across the band, and thus meets the VSWR 1.5:1 requirement. The return loss of the IF port is better than -12 dB, which is marginally above the -14 dB requirement. However, it is still well matched, and using a larger attenuator will negatively impact the NF of the receiver. Similarly, the return loss for the downconverting mixer is poorly matched at the RF port, and the 3 dB attenuator improves the return loss to better than -12 dB. The IF port of the downconverting mixer is surprisingly well matched without a 3 dB attenuator, but an attenuator was added since the return loss without a pad approached -14 dB at the upper-edge of the band. The design details of the fixed local oscillators are discussed next.

The Valon 5007 Dual synthesiser was used to generate the two local oscillator signals. The Valon synthesiser is capable of sourcing two independent LOs, at any frequency between 138 to 4400 MHz, with a minimum step size of 2.5 kHz. Both LOs are referenced to the same internal 10 MHz temperature compensated crystal oscillator (TCXO). To adequately drive the ZX05-73L mixers, the LO power levels must at least be +4 dBm. The output power levels of the Valon synthesiser is programmable up to +7 dBm. Two of the Mini-Circuits ZX10-20-71 2-way 0° power splitters are used to coherently distribute the LO signals to both receiver chains. These are 3 dB power splitters that operate from 2950-7100 MHz, with a total IL of close to 3.25 dB. Another important specification is the maximum phase imbalance of these power splitters. This factors into the correlator output equation (3.27), where these phase imbalances are represented in the LO phase differences θ_1 and θ_2 . A 3° phase imbalance is specified for this power splitter, and it is expected that this low phase variance, will have little impact on the correlator. Additional components were needed for the fixed 4920 MHz HSLO design, since the maximum synthesis frequency of the Valon is 4400 MHz.

The signal path of the 4920 MHz LO is shown in figure 4.23(a), and the frequency response in figure 4.24. The LO signal generated by the Valon synthesiser has a power level of +7 dBm, and an output frequency of 2460 MHz. The Mini-Circuits ZX90-2-36 frequency multiplier (x2) is used to double the frequency of the initial LO signal to 4920 MHz. The ZX90-20-36 specifies a conversion loss of 10 dB for a minimum input power of +8 dBm. Fortunately, the actual output power level of the Valon synthesiser is typically 1 dB higher. This is to compensate for some cable loss and gain roll-off. The newly generated LO at 4920 MHz was measured to have a power level of -1.5 dBm, corresponding to a conversion loss of ~ 10 dB. In order to drive the upconverting mixer hard (and to keep L_c low), some additional amplification after the frequency multiplier is required. The broadband (1500-5900 MHz) Mini-Circuits ZX60-5916M amplifier was used. The ZX60-5916M has a nominal gain of 17.5 dB and a high P1dB of +14.5 dBm. The amplifier is followed by a bandpass filter. The purpose of this filter is to attenuate the harmonics that were generated during the frequency multiplication process. The harmonics measured after the amplifier in figure 4.25, are as follows:

- the leaked input signal at 2460 MHz is -33.43 dBc,
- the 3rd harmonic at 7380 MHz is -33.38 dBc,
- the 4th harmonic at 9840 MHz is only -30.86 dBc.

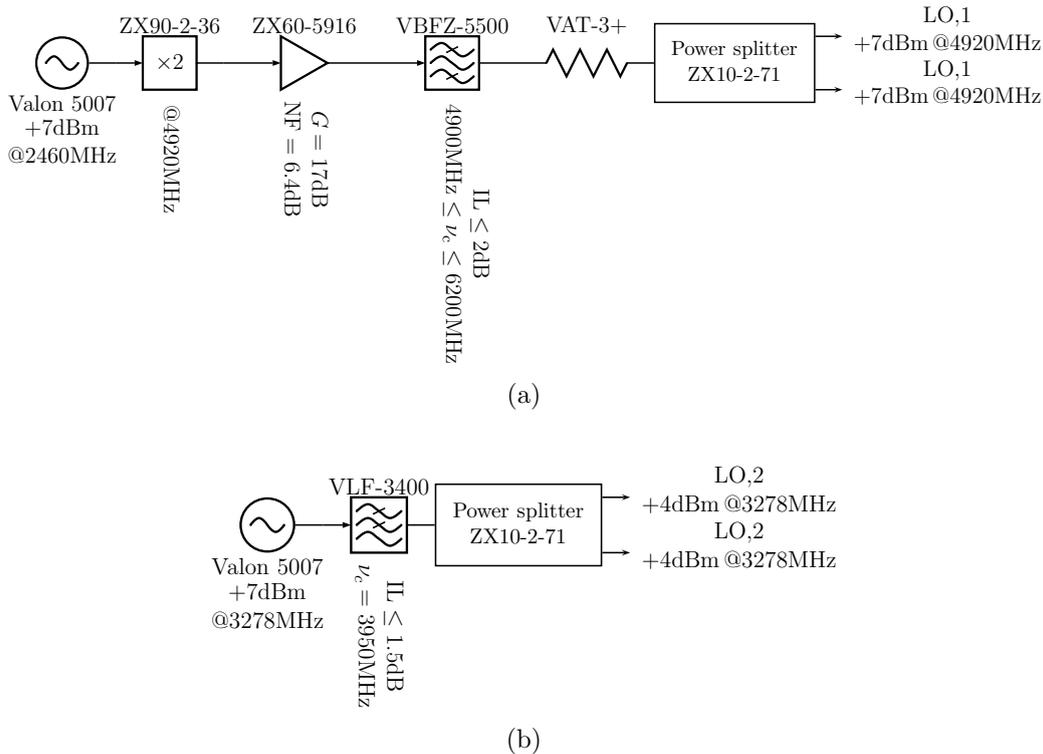


Figure 4.23: The block diagrams of the (a) 4920 MHz local oscillator, and the (b) 3278 MHz local oscillator.

These levels are still too high, and must further be suppressed to prevent them from reaching the mixer. The Mini-Circuits VBFZ-5500 BPF fulfils this purpose. Figure 4.24 shows that the additional attenuation provided by this filter is -58.71 dBc, -26.25 dBc and > -35 dBc at 2460 MHz, 7380 MHz and 9840 MHz, respectively. The spectrum of LO1 after the filter is shown in figure 4.25. This further confirms the effectiveness of the VBFZ-5500 filter. It is shown in figure 4.24, that the S_{21} roll-off is not smooth and abruptly drops by more than 3 dB at 5100 MHz. This is as a result of the poor gain flatness of the ZX60-5916M amplifier, after about 5000 MHz. This is however not a problem because the LO is a single tone at 4920 MHz. The gain at 4920 MHz is measured as $+9.5$ dBm, from figure 4.24. This suggests that P_{LO1} is close to $+7$ dBm. Furthermore, figure 4.24 shows that the LO1 signal path has an input and output return loss of better than -14 dB at 4920 MHz.

The 3278 MHz LO is directly synthesised from the Valon source at a power level of $+7$ dBm. The harmonic content of the Valon synthesiser was measured in figure 4.26 as follows:

- the 2nd-harmonic at 6556 MHz is -15.82 dBc,
- and the 3rd-harmonic at 9843 MHz is above -21 dBc.

This indicates that additional filtering was required. The Mini-Circuits VLF-3400 LPF was used in this case. This filter has an attenuation of more than 20 dB at the 2nd-harmonic frequency. The harmonic suppression is shown in figure 4.26, where it is seen that the both

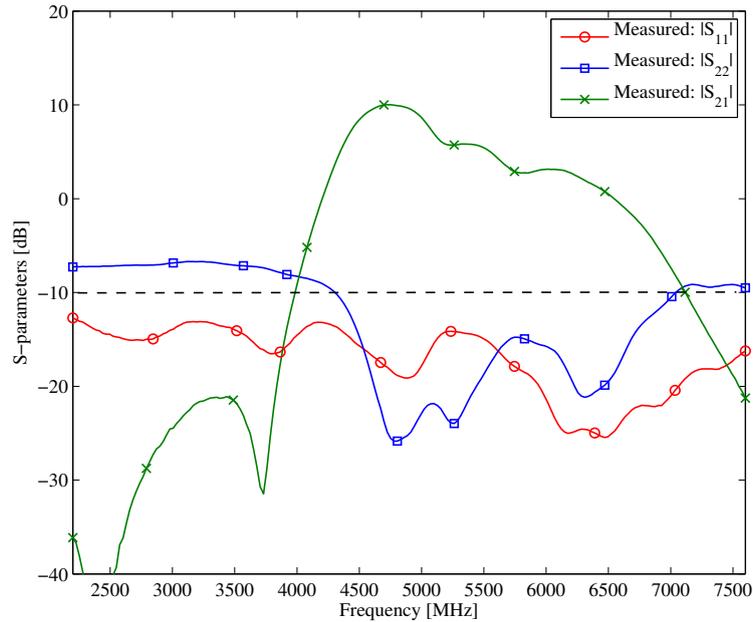


Figure 4.24: The measured S-parameters of the fixed 4920 MHz LO topology from 2200 to 7600 MHz. This figure excludes the ZX10-2-71 3 dB power divider, however its IL is accounted for.

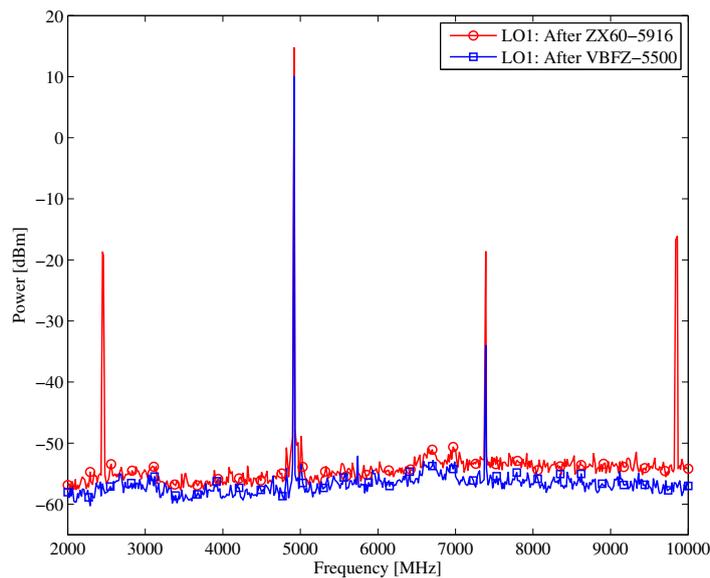


Figure 4.25: The output spectrum of the 4920 MHz local oscillator.

harmonics at 6556 MHz and 9843 MHz are effectively suppressed to -40.25 dBc and -33 dBc, respectively. The block diagram of the LO2 signal path is shown in figure 4.23(b). After the ZX10-2-71 power splitter, $P_{LO2} \approx +4$ dBm.

Frequency stability is another key specification for a receiver design, and is often characterised

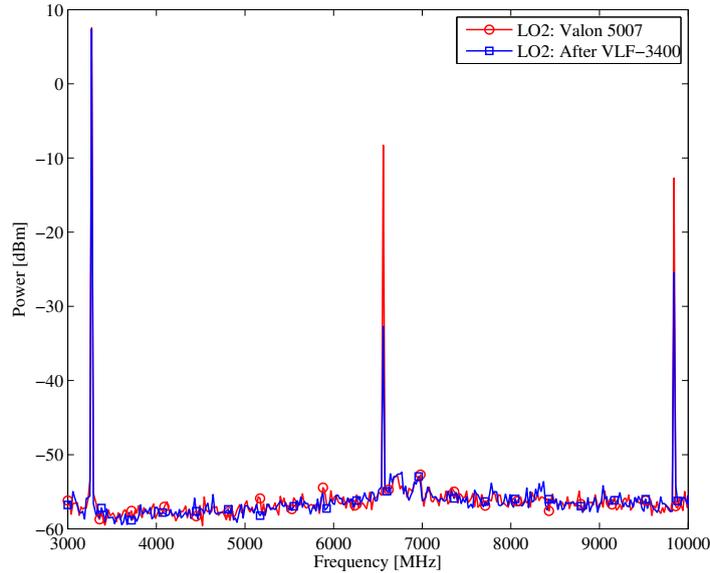


Figure 4.26: The output spectrum of the 3278 MHz local oscillator.

by the short-term phase noise stability. This was however not a primary design specification for the system, but the phase noise was none the less computed. Phase noise can simply be described as an undesired phase modulation of a signal, and can more specifically be defined as the ratio of power in a single phase-modulation sideband to the total signal power per unit bandwidth, at a particular offset frequency f_m , and is denoted as $\mathcal{L}(f_m)$ [44, 42]. To simplify the phase noise calculation it was assumed that the phase contribution of the components are negligible. The phase noise of the receiver was translated to the output center frequency of 222 MHz. The translation of the phase noise to a different frequency is computed using

$$\frac{\mathcal{L}_{\text{trans}}(f_m)}{\mathcal{L}_{\text{ref}}(f_m)} = 20 \log(n), \quad (4.34)$$

where $\mathcal{L}_{\text{ref}}(f_m)$ is the phase noise at the reference frequency, $\mathcal{L}_{\text{trans}}(f_m)$ is the phase noise of the translated frequency, and n is the ratio of the translated- to the reference-frequencies. The phase noise of the LOs are added throughout the receiver chain according to figure 4.27.

At 2460 MHz, a phase noise of -96 dBc/Hz at $f_m = 10$ kHz is specified for the Valon synthesiser. After the frequency multiplier the phase noise is worsened by 6 dB. After translating the phase noise to 3500 MHz and then to 222 MHz, using (4.34), the phase noise at the output for LO1 is found to be -117 dBc/Hz. Similarly, the phase noise at the output for LO2 is found as -113.8 dBc/Hz. Thus, the phase noise of the receiver is close to -112 dBc/Hz at an offset of 10 kHz. It is expected that this value is good enough to prevent reciprocal mixing and receiver desensitisation. Another parameter often specified is the Allan variance, which characterises the long-term frequency stability of a receiver. This was however not computed.

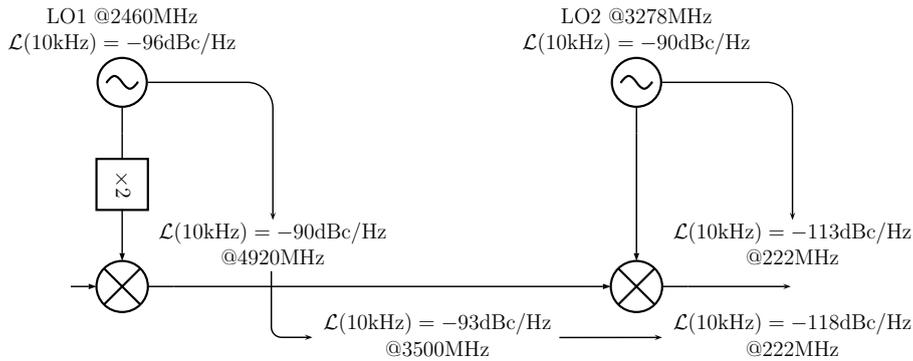


Figure 4.27: The phase noise calculation of the local oscillators.

4.4.7 Receiver cascade design performance

The integration of the RF stage, the IF stage and the output stage forms the dual-conversion superheterodyne receiver. The architectures of the RF-, IF- and output-stages are shown in figures 4.12, 4.16 and 4.18, and the complete receiver chain is shown in figure 4.28. Furthermore, a list of the receiver components is given in table 4.2. This section presents the simulated and measured results of the fully integrated superheterodyne receiver.

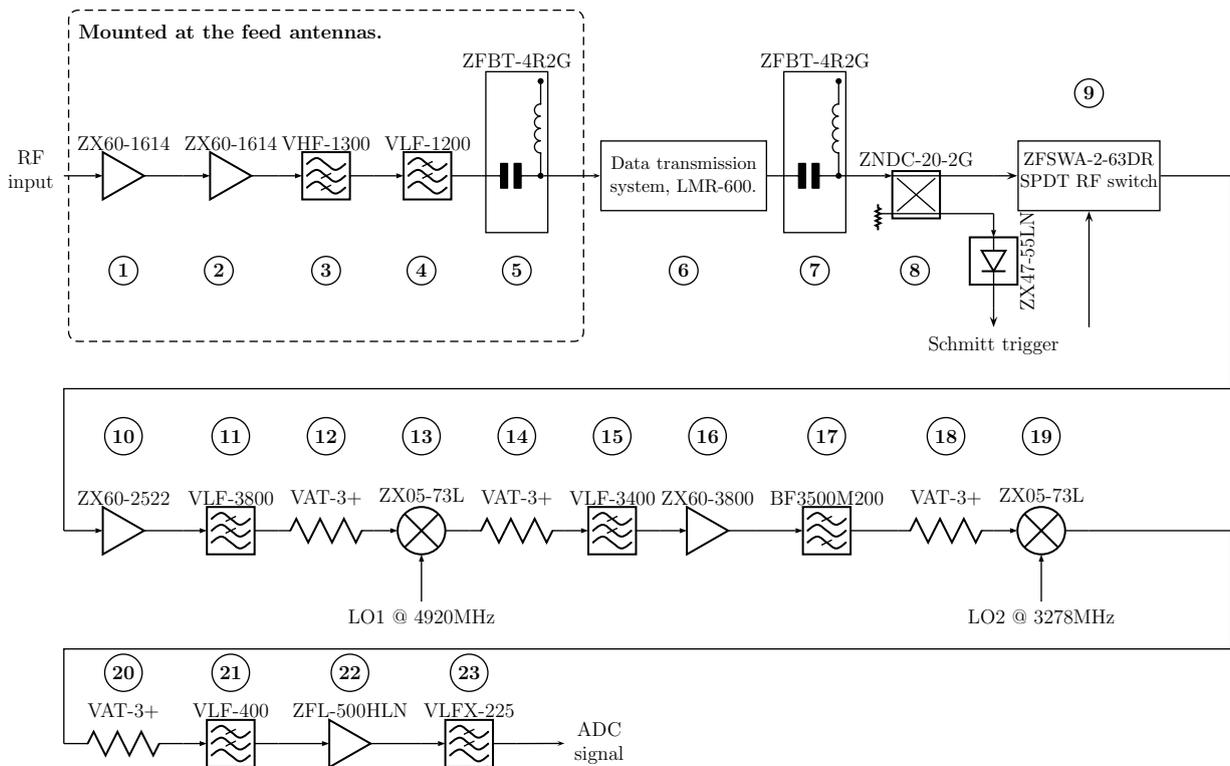


Figure 4.28: The block diagram for the complete dual-conversion superheterodyne receiver. The components for the power limiting circuit are also shown in this figure, refer to Appendix C for more information. Also, shown are the bias-tees (ZFBT-4R2G) which distribute power the low-noise front-end units, again refer to Appendix C for more information.

Model number	Units	Price (\$)	Gain/IL (dB)	Input (dBm)	P1dB (dBm)	NF (dB)	IP3	VSWR (:1)		DC power		Connector	
								In	Out	V (V)	I (mA)	Input	Output
Input stage													
VLF-1200	2	21.95	≤ 1	10W	-	~ 1	-	1.2	1.2	-	500	SMA F	SMA M
VHF-1300	2	24.95	≤ 2	7W	-	~ 2	-	1.5	1.5	-	-	SMA F	SMA M
VLF-3800	2	21.95	≤ 1	8W	-	~ 1	-	1.3	1.3	-	500	SMA F	SMA M
ZX60-1614LN	4	149.95	14, ±1.1	+13	+13.5	0.5	+30	1.3	1.3	12	42	SMA F	SMA F
ZX60-2522M+	2	59.95	23.5	+10	+15.8	3	+30.6	1.5	1.7	5	95	SMA F	SMA F
Intermediate stage													
VLF-3400	2	21.95	≤ 1.5	8W	-	~ 1.5	-	1.2	1.2	-	500	SMA M	SMA F
ZX60-3800LN	2	119.95	19, ±1.0	+1	+18	0.9	+36	1.5	1.4	5	110	SMA F	SMA F
BF3500M200	2	120.00	0.5	50W	-	~ 0.5	-	1.3	1.3	-	-	SMA F	SMA F
Output stage													
VLFX-225	2	39.95	≤ 1.2	10W	-	~ 1.2	-	1.15	1.15	-	-	SMA F	SMA M
VLF-400	2	21.95	≤ 1	8.5W	-	~ 1	-	1.2	1.2	-	500	SMA F	SMA M
ZFL-500HLN	2	99.95	20	+15	+16	3.8	+30	2.0	2.0	12	110	SMA F	SMA F
Mixers and Miscellaneous													
ZX05-73L-S+	4	48.95	6.6	50mW	+1	6.2	+12	3.38	2.52	-	-	SMA F	SMA F
VAT-3+	10	11.95	3	1W	-	~ 3	-	1.4	1.4	-	-	SMA F	SMA M
SM-SM50+	26	5.95	0.05	-	-	~ 0.05	-	1.12	1.12	-	-	SMA M	SMA M
ZFBT-4R2G	4	59.95	0.6	30	-	~ 0.6	-	1.13	1.13	30	500	SMA F	SMA M
Local Oscillators													
Valon 5007	1	295.00	-	-	-	-	-	-	1.22	5	340	SMA F	-
ZX60-5916M-S+	1	59.95	18	+10	+15.7	6.4	+28.9	2.2	1.2	5	96	SMA F	SMA F
ZX10-2-71	2	34.95	≤ 0.3	1W	-	~ 0.3	-	1.14	1.06	-	-	SMA F	SMA F
VBFZ-5500-S+	1	39.95	≤ 2	7W	-	~ 2	-	2.1	2.1	-	-	SMA F	SMA M
ZX90-2-36	1	36.95	≤ 10	+13	-	~ 10	-	-	-	-	-	SMA F	SMA F
Power limiting circuit													
ZNDC-20-2G	2	49.95	0.5	3W	-	~ 0.5	-	1.15	3 ports	-	-	SMA F	3 ports
ZFSWA2-63DR	2	69.95	1.15	1W	+24	~ 1.15	+49	1.2	3 ports	5	0.018	SMA F	4 ports
ZX47-55LN	2	89.95	-	+15	-	-	-	1.8	-	5	120	SMA M	SMA F

Table 4.2: A bill of materials for the dual-conversion superheterodyne receiver. Included in this table is a summary of the electrical performance specifications of the components.

4.4.7.1 The receiver gain

The receiver gain requirement given in table 4.1, is $G_{\text{rec}} = 67$ dB. The analytical model of the superheterodyne receiver was initially used to simulate the gain distribution throughout the receiver. This model accounts for the losses in all the passive components, such as the mixers, filters, and attenuators. Moreover, the losses in the signal transmission system is included. The LMR-600 cables are exposed to wide temperature variations, since they are placed above ground. This was also taken into account during the simulation, and a maximum temperature of 85°C was defined. The simplicity of the model restricts it to a single frequency computation, and in this regard the specifications of the components were chosen at the mid-band frequencies of the receiver stages. The simulated gain distribution of the receiver is given in figure 4.29. The receiver gain is computed as 61 dB, assuming the typical component specifications. This is 6 dB below the required specification, but the KatADC can provide up to 20 dB of additional amplification, see section 6.4.2. Also shown in figure 4.29 is the overall receiver gain, using the minimum component values. For this case the gain is shown to be 46 dB. Figure 4.29 further illustrates that the difference in the gain is minimal, for the LMR-600 cable at room temperature and at 85°C . Lastly, it is clear from figure 4.29 that after the two ZX60-1614LN amplifiers the gain is > 15 dB, and the NF is therefore defined by the low-noise front-end.

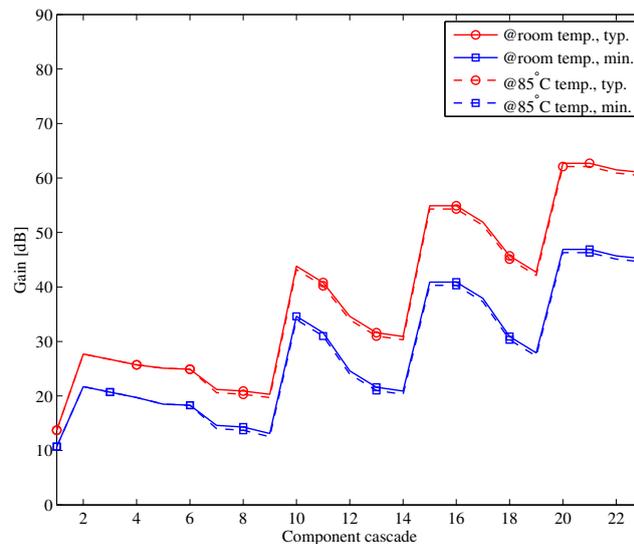


Figure 4.29: The simulated distributed gain of the receiver, computed using the analytical model. The distributed gain was computed for both the typical and minimum gain values of the components. Moreover, the gain variation for the LMR-600 coaxial cable at room temperature and at 85°C is shown. The horizontal scale refers to the component numbering used in figure 4.28.

The measured gain at the output of the receiver is shown in figure 4.30. The gain is measured as 52.57 dB, 58.22 dB and 56.6 dB at the lower-edge frequency of 94 MHz, the center frequency 222 MHz and at the upper-edge frequency of 350 MHz. The maximum passband gain of 59.28 dB occurs at 326 MHz. The gain is shown to be on average about 3 dB lower than the simulated

gain. This can possibly be attributed to mismatches between some of the components, and the below typical component values. The gain deficit can easily be corrected for by the KatADC.

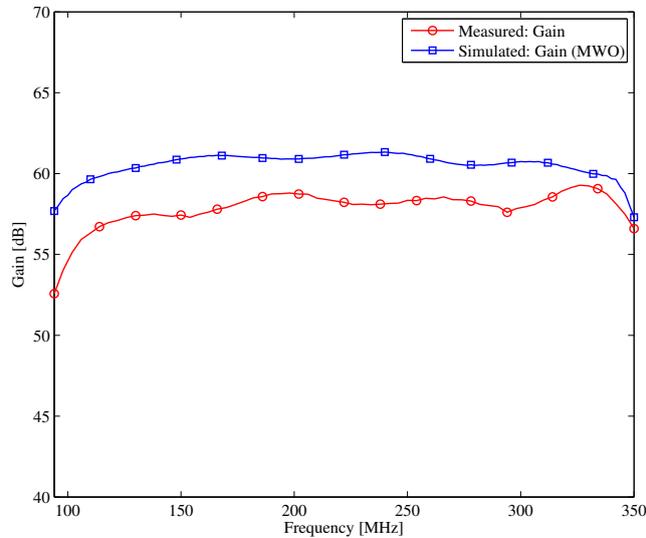


Figure 4.30: The measured gain at the receiver output from 94 MHz to 350 MHz. The gain measurement accounts for the loss in the LMR-600 coaxial cable as well.

In order to meet the $\Delta G_{\text{rec}} \leq 3.5$ dB requirement, a slightly smaller passband must be defined. From figure 4.30 it is seen that this condition is met from 102-350 MHz. This is then effectively the final receiver passband. The gain ripple in the passband is measured to be less than 1.6 dB, and the $G_{\text{ripple}} \leq 2.9$ dB specification is met.

4.4.7.2 The receiver noise figure

In table 4.1, it is specified that the noise figure of the receiver should be less than 1.1 dB, or equivalently 83.6 K. The simulated noise figure for the cascaded analytical model is shown in figure 4.31.

A section of transmission line connects the antenna to the first LNA. According to (4.25) it is important to select a low-loss cable, since any losses at this point, sets the lower NF limit of the receiver. In this regard, a 1 m length of low-loss LMR-240 coaxial cable is used. This cable has 0.3 dB attenuation at 1420 MHz, assuming room temperature. This explains the initial 0.3 dB NF in figure 4.31. For a complete list of the coaxial cables used in the interferometer system see table C.2. It is shown in figure 4.31 that after the low-noise front-end, the NF of the overall receiver has been characterised. The impact of the later stages are insignificant. It is however noted that the upconverting mixers slightly increases the NF from about 0.94 dB to 0.98 dB. The final receiver NF is found to be 0.987 dB, assuming typical component values. The simulation based on the minimum component values, indicates a severe degradation in the achieved NF. From figure 4.31 it is seen that the NF is now 1.78 dB, this is a 0.8 dB increase,

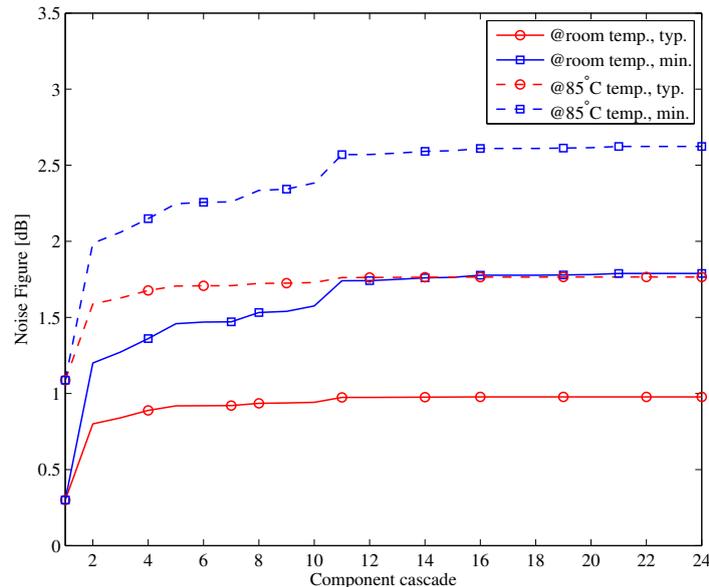


Figure 4.31: The simulated cascaded noise figure of the analytical receiver model. The antenna feed line (LMR-240) was also included in this simulation. At 1420 MHz the antenna feed line has a 0.3 dB loss. The horizontal scale refers to the component numbering used in figure 4.28.

and exceeds the 1.1 dB requirement. At elevated temperatures, the NF is even worse. This is due to the antenna feed line also being exposed to wide temperature variations. According to (4.25), as the physical temperature T_p of the line increases so does the NF. For the minimum component value scenario, at 85°C, the NF is shown to be 2.62 dB.

The noise figure measurement of the dual-conversion superheterodyne receiver consisted of making two separate measurements, since the Agilent N8975A noise figure analyser (NFA) used, does not directly support double frequency conversion measurements. Firstly, the NF and the gain after the upconverting stage was measured. This was followed by measuring the NF of the rest of the receiver chain. From these two measurements the overall NF of the receiver was determined using (4.24). The measured NF of the receiver is shown in figure 4.32. At the band center of 222 MHz the NF was measured as 0.617 dB. A slightly higher NF of 0.83 dB was measured at the lower-edge frequency of 94 MHz. The NF at the upper-edge frequency of 350 MHz was measured as 0.63 dB. It is important to note that these measurements excluded the losses of the antenna feed line, which is about 0.3 dB. If this is included it is noted that the NF on average is about 0.92 dB, which is close the expected value. It is clear that across the entire operating band the NF is below the required 1.1 dB specification.

4.4.7.3 The linearity of the receiver

The linearity of the receiver is characterised by both the third-order intermodulation distortion and the 1 dB compression point of the receiver. The specification given in table 4.1, for the OIP3 is 11.3 dBm. The OIP3 of the analytical receiver model was determined using (4.29b), which was then translated to the output. The estimated IIP3 is found to be -42.7 dBm. Assum-

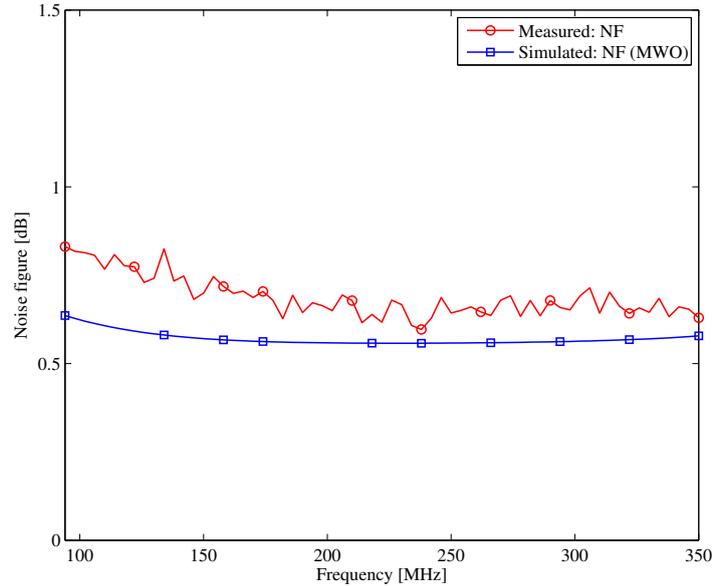


Figure 4.32: The measured noise figure of the dual-conversion superheterodyne receiver from 94 MHz to 350 MHz. The NF measurement does not account for the losses in the LMR-240 antenna feed line.

ing a gain of 61 dB, the final OIP3 is then computed as +18.3 dBm. The OIP3 of the receiver was measured and the result is presented in figure 4.33. The measurement setup consisted of combining two tones at $\nu_1 = 1395$ MHz and $\nu_2 = 1445$ MHz. The third-order harmonics at the output are thus situated at $2\nu_1 - \nu_2 = 279$ MHz and $2\nu_2 - \nu_1 = 147$ MHz.

From figure 4.33, the OIP3 is found to be $\sim +19$ dBm. The requirement of at least +11.3 dBm is met. The spurious free dynamic range of the receiver is then $\text{SFDR} = 2/3 (\text{IIP3} - \text{P}_{\text{MDS}}) = 34.1$ dB. It is apparent from figure 4.33 that the simulated and measured results are not identical, but in agreement. The differences can be ascribed to the accuracy of the simulation, and it was also noted that the gain of the actual receiver was measured to be about 3 dB less. Furthermore, the test setup could also be a contributing factor. The measurement could be repeated with better isolation between the two generators.

The 1 dB compression point of the receiver was not formally defined, since it was pointed out earlier that even the signal noise power $P_{\Delta T}$ of the sun, is below the noise floor of the receiver. So there is essentially no linear dynamic range specification. However, ideally a high P1dB is desired, since the presence of strong RFI signals can affect the linearity of the receiver. The compression point in a cascade does not easily lend itself to analysis. In this regard, the input power for each component with a P1dB value was monitored, and the receiver input power at which the limiting component reached its compression point was used as the P1dB value of the receiver. The compression point was determined to be -50.9 dBm. The limiting component is the ZX05-73L downconverting mixer, followed by the output stage ZFL-500HLN amplifier. The measured P1dB of the dual-conversion superheterodyne receiver is shown in figure 4.34. The actual 1 dB compression point is measured to be -45.4 dBm. This is almost 5 dB higher than the simulated result. This can be as a result of the lower receiver gain measured. The

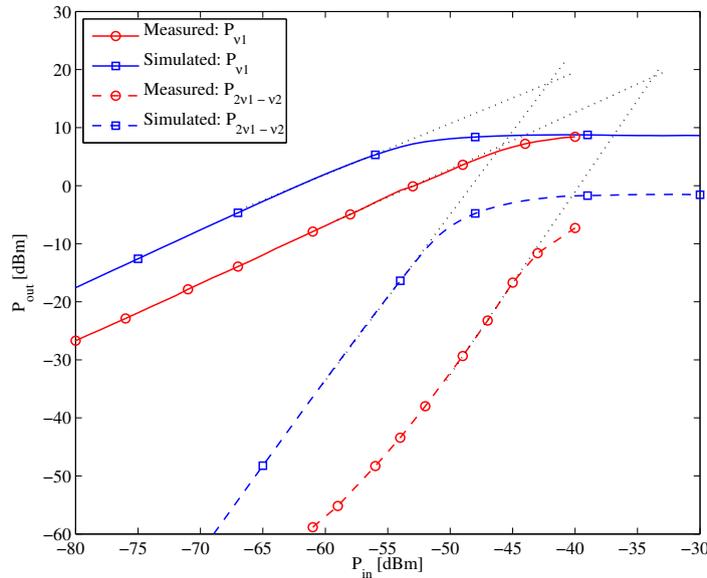


Figure 4.33: The measured output power of the fundamental tone and the third-order intermodulation product. The linear and third-order responses are extended, and where these two curves intersect is the OIP3 of the receiver. The test setup consisted of two input tones at $\nu_1 = 1395$ MHz and 1445 MHz, of equal power.

linear dynamic range of the receiver is 48.5 dBm.

In section 4.4.1 it was mentioned that the $-2\nu_{LO,2} \times 2\nu_{RF}$ and $-3\nu_{LO,2} \times 3\nu_{RF}$ harmonics appear in the receiver passband after downconverting. In figure 4.9, it is shown that if the input signal is converted to the lower-passband edge (94 MHz) then the two spurious signals are situated at 188 MHz and 282 MHz. In figure 4.35, the spectral content of the receiver is shown for an input signal of -80 dBm at 1548 MHz. It is clearly evident from the spectrum that these harmonics are below the noise floor of the receiver, and are thus not a problem.

4.4.8 Other receiver specifications

Apart from the performance specifications outlined for the receiver, there are other receiver parameters that need specifying, such as the maximum allowable input power that the receiver can handle, and the receiver's power requirements. These topics among others are discussed in appendix C.

4.5 Conclusion

Firstly, the concept of sensitivity was introduced as a means of quantifying the minimum detectable flux density that a radiometer is able to detect. Next, a rigorous derivation of the sensitivity equation for an interferometer was presented. This was followed by a discussion

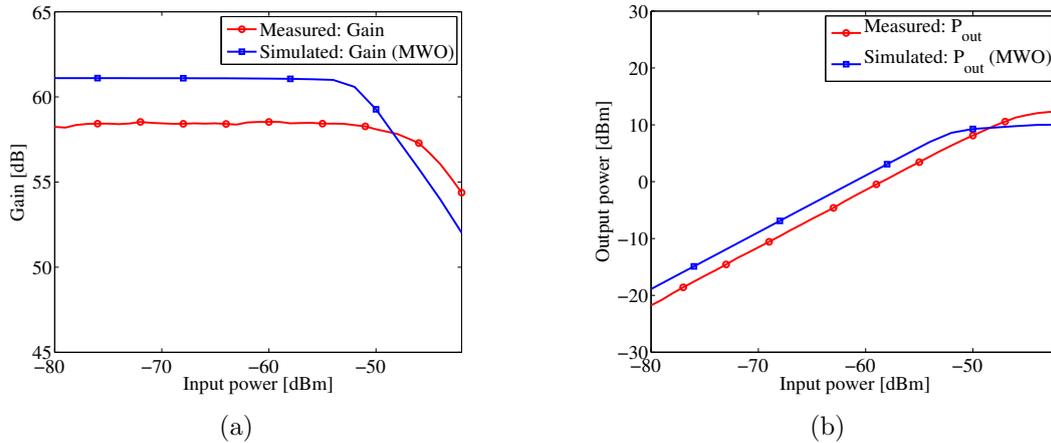


Figure 4.34: The measured (a) gain and (b) output power of the receiver at 222 MHz as a function of input power. The 1 dB compression point of the receiver can be determined from this figure. The input power is limited to -40 dBm, after this the power limiting circuits switches the signal path out after the low-noise front-end.

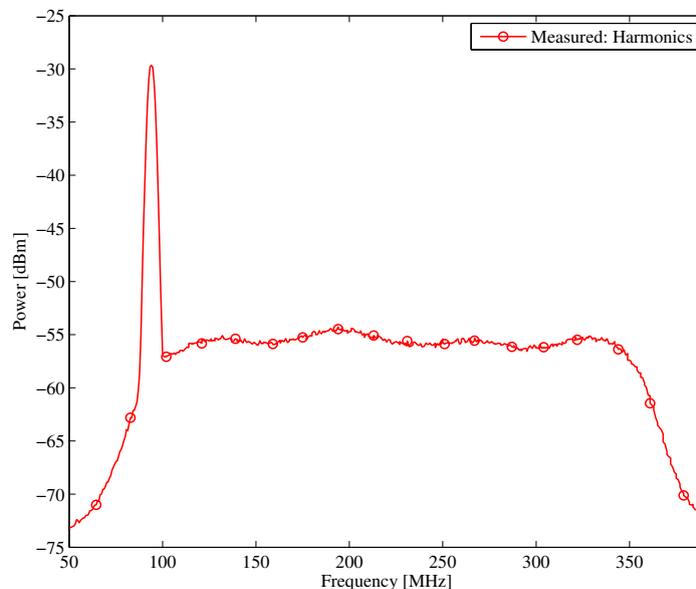


Figure 4.35: The spectrum of the dual-conversion superheterodyne receiver for an input signal of -80 dBm at 1548 MHz. This figure shows that the $-2\nu_{LO,2} \times 2\nu_{RF}$ and $-3\nu_{LO,2} \times 3\nu_{RF}$ spurious products are below the receiver noise floor.

of how the common performance metrics of radio receivers used in other branches of radio science and engineering - such as gain flatness, passband ripple and IP3 can be related to the performance of an interferometer. Specifically, the sensitivity degradation caused by these receiver specifications. This led to a set of design specifications which is summarised in table 4.1. A dual-conversion superheterodyne receiver topology was chosen, since this approach provided the best solution in terms of RFI rejection. The simulation and measurement details

of this receiver was presented and discussed in full. The complete block diagram of the receiver system is shown in figure 4.28. It was shown that the receiver met the design specifications outlined in table 4.1, and a summary of the achieved specifications are given in table 7.1. The only specification which could not be verified was the rms phase error $\Delta\phi_{\text{rms}}$ at the output of the correlator. This requires a very precise, and concurrent phase measurement of the two receiver channels. All the possible steps were taken to reduce the phase difference between the two receiver channels, such as keeping both paths of equal length and minimising the LO distribution distances.

Chapter 5

An antenna suitable for radio interferometry

5.1 Introduction

For the SKA to meet some its key science goals it is estimated that a sensitivity, $A_{\text{eff}}/T_{\text{sys}}$ of $2 \times 10^4 \text{ m}^2/\text{K}$ is needed [6, 5]. Assuming a conservative system temperature T_{sys} of 50 K, the required effective collecting area, A_{eff} of the synthesis array is 10^6 m^2 or 1 km^2 . This has major cost implications on the antennas, and it has been estimated that the antennas will account for approximately 40% of the cost [6]. This has initiated many SKA demonstrator projects such as the Allen Telescope Array (ATA), LOFAR and Electronic Multi Beam Radio Astronomy ConcEpt (EMBRACE), which all focuses on different antenna design solutions [6]. The ATA focuses on the feasibility of a small reflector solution, incorporating cryogenic-cooling techniques for wideband feeds [6]. LOFAR and EMBRACE are both SKA pathfinder projects focused on aperture phased arrays, where LOFAR more specifically is investigating sparse arrays operating around 20 - 240 MHz, and EMBRACE is focused on dense arrays at 0.4 - 1.55 GHz [6]. These pathfinder projects highlight the cost/performance trade-offs offered, and applicability for the SKA project. Even for a small instrument such as the interferometer built for this thesis, the same criteria can be used to motivate the suitability of a particular antenna design. This chapter commences with a literature review of plausible antenna types.

5.2 Antenna literature study

5.2.1 Overview

The approach followed to narrow down the possible antenna types, considered the bandwidth requirements of the system; optimising the sensitivity $A_{\text{eff}}/T_{\text{sys}}$, where a gain of at least 19 dBi is required; looking at the manufacturing costs involved, and investigating what antenna types are commonly used for radio astronomy. The polarisation of cosmic sources is mostly unpolarised. This is because of the direction variation of the magnetic fields within the sources, and Faraday rotation [15, 14]. The type of antenna chosen is therefore not restricted by the

type of polarisation, and both linearly- and circularly- polarised antennas are considered. This being said polarisation measurements are still very important in radio astronomy. In [15], an overview of the uses of polarimetry in radio astronomy is given. The processing bandwidth of the instrument was chosen as 256 MHz, and for a center frequency of 1420 MHz the fractional bandwidth is 18%. This immediately excludes any resonant antenna types, such as dipoles and Yagi-Uda arrays. These antennas are capable of only narrow bandwidths, which range from about 3% to 8%, considering thin dipoles [48]. It is possible to increase the operational bandwidth of these antennas by decreasing the length to diameter ratio ($l/2a$), basically this means using thicker dipoles. In [48], it is shown that for a $l/2a$ ratio of 50, a bandwidth of 16% is achievable for a single half-wave length dipole. The collecting area A_{eff} has a more profound effect on improving the sensitivity, as compared to using a wider bandwidth, since the latter only increases the sensitivity in proportion to $\sqrt{\Delta\nu}$. The effective collecting area is related to the gain G of the antenna, by

$$G = \frac{4\pi}{\lambda^2} A_{\text{eff}}. \quad (5.1)$$

This shows that the sensitivity can be improved by maximising the gain. Most of the large radio telescopes achieve high gains by using parabolic- or spherical-reflectors, or arrays of a large number of elemental antenna types, such as dipoles and helical antennas [14]. To follow is a concise review of the variety of antenna types used in radio interferometry.

Early radio astronomy antennas operated in the meter wavelength range, and the instruments often consisted of dipole arrays and parabolic-cylindrical reflectors [15]. The aperture-synthesis array at Cambridge, England had a 442 m long by 20 m wide cylindrical parabolic reflector, and it operated at 178 MHz. In 1952, Kraus constructed a 96 element helical array antenna at Ohio State University, which operated at wavelengths from 1 to 3 meters [14]. As a final example of early radio telescope antennas, the 85 MHz Mills cross in Sydney, Australia consisted of two arms, E-W and N-S, which were about 457 meters long [14]. Each arm had 500 half-wave dipole antennas and could produce a fan-beam ($50^\circ \times 0.6^\circ$), or a pencil beam (0.8°) when operated as a phase-switching instrument [14]. Modern synthesis arrays operate at much higher frequencies (cm range), making it possible to use parabolic reflector antennas, which are now very well developed for radio astronomy. The VLA consist of 25 m shaped parabolic reflectors, where the feed system has a Cassegrain geometry with an off-axis secondary focal point [33]. Each main reflector is outfitted with four different feeds that are instrumented for wavelengths of 17-22, 6, 2, and 1.3 cm. The GMRT also uses parabolic reflectors but of the prime-focus type, since it is not practical to have a Cassegrain design at meter wavelengths [4]. The longer wavelengths allowed the reflector surface to be constructed out of mesh panels, which significantly reduced the cost. The GMRT feeds are placed on a cubical turret at the focus and can operate over 6 frequency bands, 1000-1450 MHz, 610 MHz, 327 MHz, 233 MHz, 150 MHz and 50 MHz [4]. Even the SKA is provisionally planning to use parabolic reflectors with wideband single pixel feeds at the frequency band ($\sim 1000 \text{ MHz} - 10 \text{ GHz}$) [5]. These professional instruments achieve sensitivities that are just not possible to obtain with small amateur arrays. However, many amateur instruments have been constructed using significantly smaller antennas, and the results achieved are very encouraging [49, 13, 12, 11].

5.2.2 The pyramidal horn

The University of California at Berkeley have built a radio telescope that is optimised for observing the HI line at 1420 MHz [49]. The antenna used is a pyramidal horn, and attempts are being made to observe the 1665 MHz OH maser line. Pyramidal horn antennas find wide application at microwave frequencies given their high gains (in excess of 20 dBi), low VSWR and moderate bandwidths (50%) [48, 50]. Their low sidelobes are also an attraction. In the E- and H-planes of a standard gain pyramidal horn, the side lobes are below -13 dB and -20 dB, respectively [48, 51]. Moreover, their mechanical simplicity makes manufacturing straightforward, but it can nonetheless be costly. The geometry of a pyramidal horn antenna is given in figure 5.1. The aperture dimensions are specified by A and B , the throat-to-aperture lengths are given by R_E and R_H , and the broad and narrow dimensions of the connecting waveguide are $a \times b$, respectively. The analytical expressions derived in [50], computes the dimensions for an optimum gain pyramidal horn, where the phase lag in the H-plane is 135° , and 90° in the E-plane. The expressions are repeated here

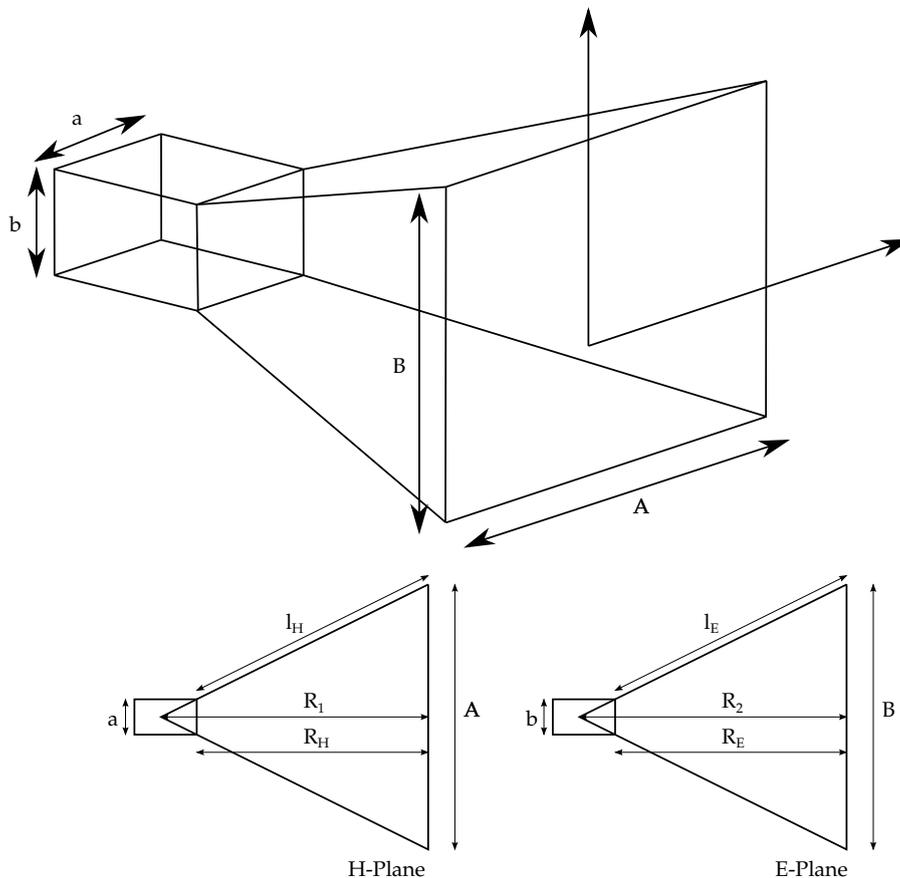


Figure 5.1: The geometry of a pyramidal horn antenna. The cross-sections through the E- and H-planes are also shown.

$$\frac{3G^2\lambda^2}{32\pi^2\eta_a^2} = A^4 - aA^3 + \frac{3bG\lambda^2}{8\pi\eta_a}A, \quad (5.2a)$$

$$B = \frac{G\lambda^2}{4\pi A\eta_a}, \quad (5.2b)$$

$$R_E = R_H = \frac{A(A - a)}{3\lambda}, \quad (5.2c)$$

where η_a is the aperture efficiency, and it is normally chosen as 51%, since it has been found that this value results in the most accurate prediction for the gain level [50]. The size of the pyramidal horn used in [49] can be computed using (5.2). The waveguide dimensions are chosen such that it supports the propagation of the dominant TE_{10} mode at 1.42 GHz. The standard WR-650 waveguide with dimensions, $a = 16.51$ cm and $b = 8.225$ cm is suitable [42]. Solving (5.2a) analytically, followed by (5.2b) and (5.2c), it is found that the dimensions are, $A = 101.26$ cm, $B = 78.8$ cm and $R_E = R_H = 135.41$ cm. The dimensions for a pyramidal horn operating at L-band are large, and the manufacturing costs can be high if this antenna needs to be accurately manufactured.

5.2.3 The Yagi-Uda

A popular antenna choice for amateur radio astronomy is the Yagi-Uda array [12, 11]. It was already mentioned that the bandwidth of this antenna is inherently narrow ($\sim 3\%$). It is however worth discussing their characteristics, which make them such a popular choice. In the literature, two correlation interferometers operating at 408 MHz and 1.3 GHz, have successfully been designed using Yagi-Uda arrays [12, 11]. In both cases, strong radio sources other than the sun could be detected, such as Cassiopeia A, Taurus A and Cygnus A. In [12], each interferometer element consisted of an array of two smaller Yagi-Uda antennas. Because of the higher operating frequency, electrically larger Yagi-Uda antennas were used in [11], with gains of 18 dBi and 2% bandwidth. Apart from the high gains, these antennas are simple and inexpensive to manufacture.

5.2.4 The LPDA

An alternative wire dipole antenna with excellent broadband performance is the log-periodic dipole array (LPDA). This antenna is discussed next.

The LPDA was invented by Isbell at the University of Illinois in 1958 [52, 53]. This was as a result of the pioneering work on broadband logarithmically periodic antenna structures by DuHamel [53]. The LPDA structure is defined such that the pattern and impedance repeat periodically with frequency, and the period is controlled by the design parameter τ . If the variation of the pattern and impedance is negligible over one frequency period, then the antenna is essentially frequency independent. The geometry of the LPDA is described in figure 5.2. The LPDA is a series-fed coplanar array of linear dipoles. The dipole lengths, and their spacing from

the apex are scaled in such a way that it forms a geometric progression with a constant ratio, $\tau < 1$. The spacing factor σ is further defined as the distance between successive elements. These parameters are mathematically defined as

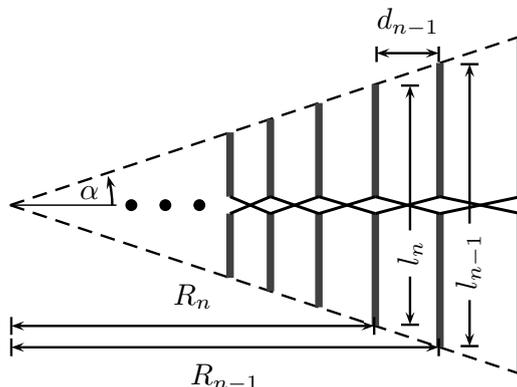


Figure 5.2: The geometry of a log-periodic dipole array (LPDA).

$$\tau = \frac{l_n}{l_{n-1}} = \frac{d_n}{d_{n-1}}, \quad (5.3a)$$

$$\sigma = \frac{d_n}{2l_n} = \frac{1 - \tau}{4 \tan(\alpha)}, \quad (5.3b)$$

where α is half the angle of the wedge enclosing the dipole lengths. Adjacent elements are connected to a balanced transmission line in an alternating fashion. This is shown in figure 5.2. This ensures an end-fire radiation pattern towards the apex of the LPDA [52]. Theoretically, LPDAs can have an infinite bandwidth, but a practical antenna must have a finite size. Truncating the antenna then ultimately limits the operating bandwidth. The LPDA has an active region where the elements are about $\lambda/2$ long, this suggests that the lower frequency limit is set by the longest element, similarly the upper frequency limit is set by the shortest element. However, LPDAs are still broadband, and 45% bandwidths are easily achieved, and log-periodic antennas are further capable of 1:10 bandwidths [52]. This is well above the 18% bandwidth requirement. Carrel introduced a straightforward design procedure for LPDAs that allows for control over the input impedance, bandwidth, gain and antenna size [53]. This procedure was accompanied by contour plots of the directivity as a function of τ and σ . These curves predict that gains in excess of 11 dBi are possible. However, experimental results from other researchers, showed disagreement between their gains measured and those predicted by Carrel's work. It is suggested, that Carrel's gain curves must be uniformly lower by 2 to 3 dB. The gain of an LPDA is therefore unsatisfactorily low for radio astronomy applications. DuHamel proposed arraying a number of log-periodic antennas together in order to achieve a higher gain [54]. To maintain the frequency independence, the array element spacings must be defined by angles and not distances. It is reported that gains of 15 dBi were achieved, and it is expected that 20 dBi gains are possible [54]. This is a plausible solution for increasing the gain, but this

is an overcomplicated design, which is not fully mechanically self-supported.

For radio astronomy it is also desirable to have low backlobe levels. This is because the radiation from the backlobe is normally pointed towards the warm ground, and this causes the system noise temperature to increase. The F/B ratio of an LPDA is usually better than 15 dB [52].

The N²I² adding interferometer is an amateur interferometry project between the NRAO and New Mexico Tech [13]. This instrument makes use of a more conventional radio astronomy antenna, the parabolic reflector. The 3 m reflectors used are outfitted with cylindrical feed horns and operate at 1.4 GHz. The gain achieved is approximately 30.7 dBi, and this is a representation of the high gains achievable with reflector antennas. Out of all the amateur instruments discussed in this section, the N²I² interferometer has the highest sensitivities, owing to the high gain antennas used. This made it possible to observe even weaker sources such as Virgo A [13]. The prices of commercial parabolic reflectors have decreased dramatically, due to the rapid development in direct broadcast satellite TV technology. The use of an inexpensive reflector antenna promises to be a good solution. These antennas are however designed to operate over portions of the X- and Ku-bands (10.7-12.75 GHz). This requires that a suitable feed antenna be instrumented.

The LPDA could very well be used as a feed for a reflector antenna. It is capable of operating over the specified 18% bandwidth, and its beam characteristics can be controlled to give optimum illumination of the reflector. Another important requirement for a feed antenna is to have a constant phase center with frequency. The LPDA unfortunately suffers from phase center variation as the frequency changes [55]. The reduced efficiency can be quantified as the defocusing loss [55]. For the example considered in this paper, the defocusing loss approached 1.9 dB for a 90 cm reflector, covering the range from 1 to 12.4 GHz.

5.2.5 The helix

An interesting alternate feed antenna to consider, is the axial mode helical antenna. The helix was invented by Kraus, and is one of the simplest antennas with outstanding broadband characteristics [56]. A helix with a circumference C , on the order of one wavelength radiates in the direction of the helix axis, and is circularly polarised [56, 48]. Over a wide frequency band the pattern shape, circularity of the polarisation and the impedance are all relatively stable. More specifically, the helical beam antenna performs satisfactorily over a bandwidth of

$$\frac{3}{4}\lambda \leq C \leq \frac{4}{3}\lambda,$$

which corresponds to a bandwidth ratio $\sim 1 : 1.78$, and compares favourably to the required 1:1.198 bandwidth [57, 48, 56]. In [57], it is shown that the bandwidth decreases as the axial length or equivalently the gain of the helix increases. From the parametric study performed in this article, an empirical expression was derived for the bandwidth

$$\text{BW} \approx 1.07 \left(\frac{0.91}{G/G_p} \right)^{4/(3\sqrt{N})}, \quad (5.4)$$

where G/G_p is the gain versus peak gain ratio, and specified as either -2 dB or -3 dB. In their study the bandwidth reduction as a function of axial length was summarised as follows, for a 5 and 35 turn helix the bandwidths were found to be $\sim 42\%$ and 21% , respectively [57]. The persistence of the axial-mode radiation pattern and the stable terminal impedance of a helix, over a wide frequency range, is owed to the current distribution and the phase velocity properties of the helix. Firstly, the current distribution of a helix antenna can be resolved into a current for an outward travelling wave and a much attenuated inward travelling wave. The large attenuation of the reflected wave results in a nearly uniform current distribution along the center of the helix. The reflected wave at the input of the helix is insignificant, which ensures a flat VSWR over a wide frequency range [56]. Secondly, the phase velocity of the propagating wave is such that the electric fields at every half turn ($\lambda/2$) adds in-phase. The phase velocity naturally adjusts itself to produce this result [56]. This ensures an end-fire radiation pattern in the direction of the helix axis, over a wide frequency range.

The circularly polarised radiation of the helical beam antenna is not a problem, since the radiation from most cosmic sources is unpolarised. In fact, using a helical beam antenna feed has the added advantage of being insensitive to reflections from nearby metallic surfaces. Upon reflection the handedness of the radiation changes, for example, a right-hand circularly polarised (RHCP) helix antenna is insensitive to left-hand circular polarisation (LHCP). The axial ratio is generally better than 1.5 dB for $0.8 \leq C/\lambda \leq 1.2$ [57], and improves as the number of turns N increases, approximately by

$$\text{AR} = \frac{2N + 1}{2N}. \quad (5.5)$$

The axial ratio can be improved even further by tapering the last two turns, the improvement is most noticeable at the higher-band end [57]. A poorly designed helix can have high sidelobe levels (> -10 dB) and back radiation. This can be corrected for, by shaping and re-sizing the ground conductor of the helix [58]. This is seen to be especially true for cylindrical and conical ground planes. The decaying current of the first few helix turns are responsible for the higher sidelobe levels at low elevation angles [58]. The sidelobe levels are suppressed by redirecting this radiation favourably upwards. The improvement is controlled by the finite height of the cylindrical and conical ground conductors [58]. The size of the ground plane is also critical, and it is usually chosen to be between $0.5 - 0.75\lambda$. If the ground plane is made to be small ($< -0.3\lambda$), then the helical antenna operates in the backfire mode. Lastly, the shaping can also optimise the gain of the antenna [58].

5.2.6 The selected antenna

In conclusion, based on the literature study the proposed antenna solution consists of a parabolic reflector antenna fed by a helical beam antenna. More specifically, two 92 cm commercial DBS offset parabolic reflectors are used. According to (5.1), a gain of at least 20.5 dBi can be achieved

at 1420 MHz, assuming an aperture efficiency of 0.6. Firstly, the important geometrical features of this reflector is discussed. This is followed by the analysis and detailed design of the helical feed antenna. Finally, the simulated and measured results are presented and discussed.

5.3 Geometrical features of offset-parabolic reflectors

Offset-parabolic reflectors have found widespread use, particularly in satellite communication systems. The main principal advantage of an offset-parabolic reflector is the reduced aperture blockage [59]. This is a significant advantage over axisymmetric reflectors, which suffers from degradation in the gain, sidelobes and polarisation as a result of aperture blockage. Another advantage of the offset-parabolic reflector is the good isolation between the primary feed and reflector [59]. The advantage of this is that the VSWR of the feed is independent of the reflector. Probably the biggest disadvantage of the offset-reflector is the higher cross-polarisation levels - this is as result of the asymmetrical geometry [59]. For circular polarisation, the reflector does not depolarise the radiated field, so there is no cross-polarisation. However, the beam is squinted from the boresight direction. It is important to study some of the geometrical features of an offset-parabolic reflector, in order to calculate the electrical performance of the antenna.

The geometry of an offset-parabolic reflector with a focal length F , diameter D and offset height H is shown in figure 5.3. An offset-reflector can be constructed by carving out a portion of the symmetrical parent reflector. This is typically achieved by intersecting the parent reflector surface with a circular cone with its apex at the focal point, or by using a circular cylinder that is parallel to the reflector axis [60]. A thorough mathematical treatment of the offset-reflector generated from these conic-sections are given in [60]. Only the important results are summarised here. The projection of the intersection curve (rim) on the $x' - y'$ plane is elliptical, with its major axis along the y' -axis, while its projection on the $x - y$ plane is a circular aperture. This assumes that the reflector surface is paraboloidal (eccentricity = 1). It is of some interest to note that the point of intersection of the cone axis with the intersection plane (point c') is not coincident with the center of the intersection curve (o'). This results in two angles; the angle subtended in the bisect direction β (point c'); and the angle subtended to the center of the projected aperture β' (o'). The primary feed is often directed towards β or β' , resulting in different far-field patterns. It is shown that the far-field pattern in the plane of symmetry ($y - z$ plane), shows some improvement for the feed angle β' , such as lower sidelobe levels [60]. The far-field pattern differences in the perpendicular plane, is less pronounced for the two feed angles. The effect the feed angle Ψ_f has on the far-field pattern was investigated for the designed reflector system, and the results are presented in section 5.5.2. Some of the other important geometrical parameters that describe the offset-parabolic reflector, is defined below:

- The angle subtended at the upper tip of the reflector rim,

$$\Psi_U = 2 \tan^{-1} [(D + H)/2F].$$
- The angle subtended at the lower tip of the reflector rim,

$$\Psi_L = 2 \tan^{-1} [H/2F].$$

- The angle subtended to the center of the projected aperture, $\beta' = 2 \tan^{-1} [(D/2 + H)/2F]$.
- The angle subtended in the bisect direction, $\beta = (\Psi_U + \Psi_L)/2$.
- The half angle subtended to the upper and lower tips of the reflector rim, $\alpha/2 = (\Psi_U - \Psi_L)/2$.

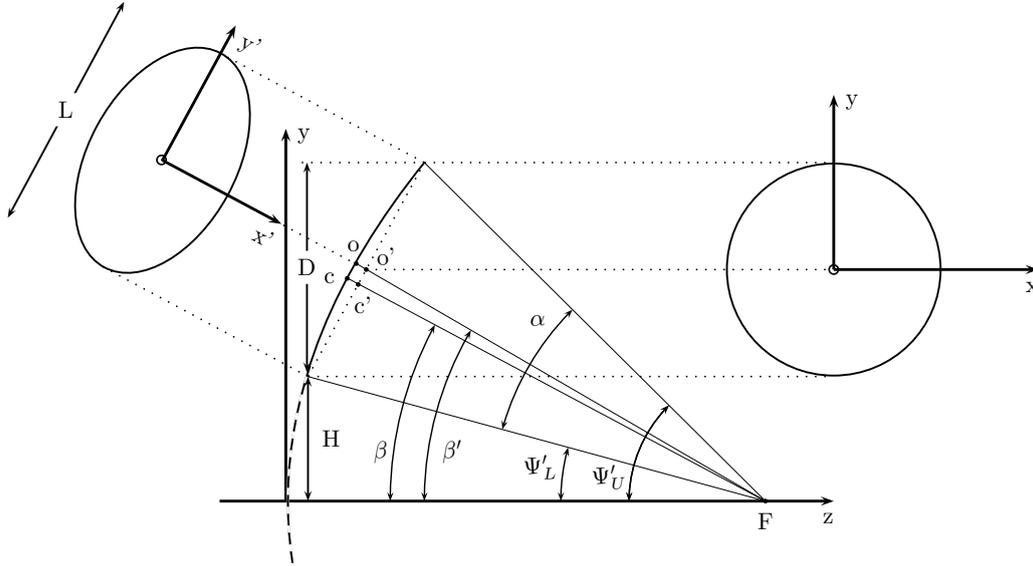


Figure 5.3: The geometry of an offset-parabolic reflector, with focal length F , diameter D and offset height H . The offset-parabolic reflector has an elliptical contour (rim), while the projection of this contour onto the $x - y$ plane will produce a circular aperture.

The geometrical parameters of a reflector antenna are usually related to the focal length, F . It is however useful to convert these geometrical parameters, to dimensions relative to the plane of the rim [61]. Situations where this is useful, is during manufacturing, and to compute the focal length if it is unknown from the easily measured reflector parameters, such as the lengths of the major axis (L), minor axis (D) and depth (B), see figure 5.3, where B is measured as $c - c'$. The simple expression derived in [61] to compute the focal length from L , D and B is

$$F = \frac{D^3}{16LB}. \quad (5.6)$$

The last parameter needed to completely specify the geometry of an offset-parabolic reflector is the offset height, H . This parameter can also be computed using the measured parameters L and D

$$H = -\frac{D}{2} + 2F\sqrt{\frac{L^2}{D^2} - 1}. \quad (5.7)$$

The two identical offset-parabolic reflectors were donated to this project and the focal length was thus not known. The dimensions of the dishes were measured as $L = 102$ cm, $D = 92$ cm

Geometrical Parameter	Value
Dish diameter, D	92 cm
Parent parabola diameter, $D_p = 2(D + H)$	194.6 cm
F/D_p	0.275
Offset height, H (5.7)	5.33 cm
$\alpha/2$	39.38°
β	45.08°
β'	51.16°
Ψ_L	5.69°
Ψ_U	84.46°

Table 5.1: A summary of the geometrical parameters of the 92 cm offset-parabolic reflector.

and $B = 8.9$ cm. The focal length was computed as $F = 53.61$ cm, using (5.6). Some of the other important parameters of the offset-parabolic reflectors are summarised in table 5.1.

The performance of a reflector antenna is only as good as the matching of the primary feed pattern that illuminates it. The aperture efficiency $0 \leq \eta_a \leq 1$ is a metric which indicates how well a reflector is illuminated. A high aperture efficiency results in good reflector antenna performance. The aperture efficiency can further be separated into various subefficiencies, such as amplitude taper efficiency η_t , spillover efficiency η_s , crosspolar efficiency η_x , phase efficiency η_p , etc. For this study only the amplitude taper- and spillover- efficiencies are of concern. The amplitude taper efficiency η_t , is a measure of the non-uniformity of the field amplitude distribution across the aperture of the reflector, this is caused by the tapered primary pattern [4]. For $\eta_t \ll 1$, the edges of the reflector are under-illuminated. This gives better sidelobe performance, but a smaller area of the reflector is illuminated causing a reduction in the achievable gain. The spillover efficiency η_s , quantifies the portion of the power that is radiated from the feed that gets intercepted by the reflector, relative to the total feed power. It is seen that η_t and η_s are complementary. If the edges of the reflector are strongly illuminated (with a broad primary beam), then η_t increase, but η_s decreases and there is more spillover. In contrast, using a narrow primary beam results in a better $\eta_s \rightarrow 1$, but the η_t decreases. There clearly exists a trade-off between these two efficiencies. The trade-off between η_t and η_s does however have an optimum solution. A maximum aperture efficiency, $\eta_a = \eta_t \eta_s$ occurs for an edge illumination (EI) of -11 dB, this is about an 82% efficiency [4, 48]. In practice, a value of -10 dB is commonly quoted, and the aperture efficiency achieved is also lower, about 60% – 75%. The effect of spherical spreading loss (SPL), must also be accounted for at the rim of the reflector. This loss is due to the fact that the power density of the spherical wave leaving the feed falls off as $1/R^2$, where R is the distance from the focus to the edge of the reflector rim. The SPL is defined as

$$\text{SPL} = \frac{1 + \cos \Psi_R}{2}, \quad (5.8)$$

for a parabolic reflector, and $\Psi_R = \Psi_L$ or Ψ_U . The EI is then more accurately defined as

$$EI = -FT - SPL \text{ [dB]}, \quad (5.9)$$

where FT is the feed taper in the direction of the reflector rim. The goal of an EI of -11 dB was specified for the primary feed pattern. This includes the ~ -0.02 dB and -5.22 dB SPL at the lower and upper tip of the reflector rim, respectively. The analysis and design of the helical beam antenna is presented in the next section.

5.4 Analysis and detailed design of the helical beam antenna feed

The geometry of a uniform monofilar helix antenna is shown in figure 5.4. If one turn of the helix is unrolled on a flat surface, then the circumference C , the spacing S and turn length L are related by the triangle shown in figure 5.4. It is then a simple matter to define the geometrical parameters of the helix:

- The diameter of the helix, D_h . The diameter is measured between the centers of the conductor used ($D_h = D'_h + 2a$), where a is the radius of the conductor.
- The circumference of the helix, $C = \pi D_h$. This has to be on the order of about a wavelength ($\sim 1\lambda$) to ensure the axial mode of operation.
- The spacing between the helix turns, $S = C \tan \alpha_h$.
- The pitch angle of the windings, $\alpha_h = \tan^{-1}(S/C)$.
- The length of one turn, $L = \sqrt{(C^2 + S^2)}$.
- The axial length of the helix, $h = NS$.

To obtain a first-order prediction of the radiation pattern of the helical beam antenna, an approximation can be made, where it is assumed that the current distribution of the helix is a single uniform travelling wave [56]. The axial-mode helix can then be modelled as a uniformly excited array, each turn being an element spaced S apart from the next adjacent element. The radiation pattern can then be approximated by taking the product of the array factor and the element (single turn) pattern. In [56], it is shown that if the number of turns are large ($N > 7$) then the radiation pattern of the helix is largely dominated by the array factor. The normalised array factor for a uniform array of N elements is given by

$$Y_N = \frac{1}{N} \frac{\sin \frac{N\psi}{2}}{\sin \frac{\psi}{2}}, \quad (5.10)$$

where ψ is the auxiliary phase factor. The phase factor of an axial-mode helix antenna can be approximated as [56]

$$\psi = 2\pi \left(S_\lambda \cos \phi - \frac{L_\lambda}{p} - \frac{1}{2N} \right), \quad (5.11)$$

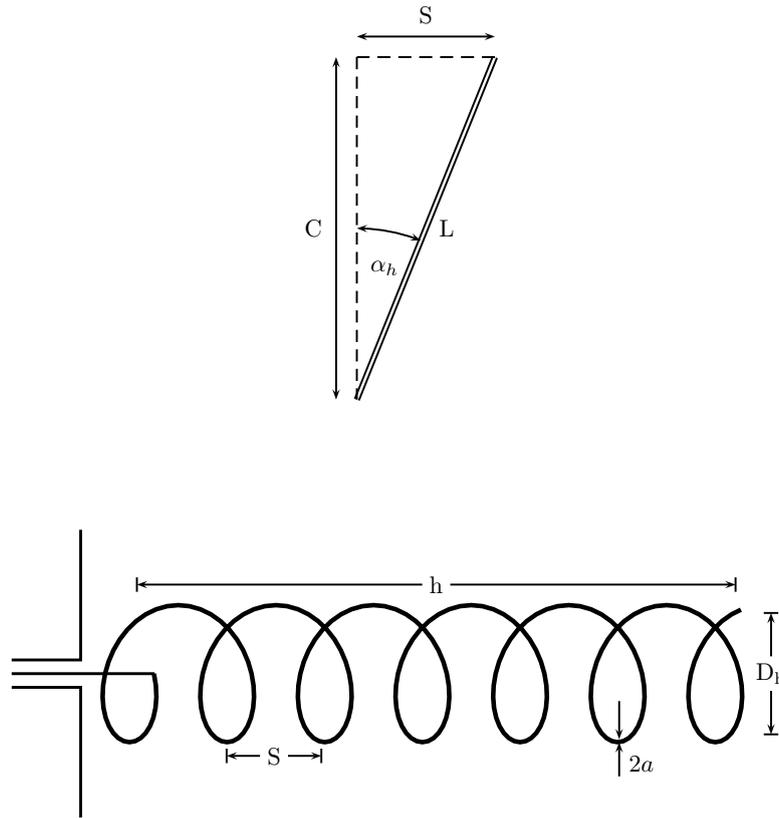


Figure 5.4: The geometry of a uniform (left-hand wound) helix antenna with dimensions.

where S_λ and L_λ are measured in wavelengths, p is the phase velocity factor and ϕ is the angle measured relative to the helix axis. The phase factor can be understood as follows, the phase difference at some distant point between two adjacent elements is $2\pi S_\lambda \cos \phi$. But, there is a phase retardation between these two elements of $2\pi L_\lambda/p$, because the wave has to travel around one turn. The $\frac{\pi}{N}$ factor in (5.11), is a natural phenomenon in a helical beam antenna, and causes the helix pattern to be sharper. This corresponds to the Hansen-Woodyard end-fire array [48, 56]. For the case where the maximum radiated field is in the direction of the helix axis ($\phi = 0$ and $\psi = -2\pi$) it is seen that (5.10) reduces to

$$\frac{(-1)^{N+1}}{N \sin(\pi/2N)}. \quad (5.12)$$

This factor can be used to normalise the radiation pattern of the array. This theoretical model of a helical beam antenna is easily implemented as a computer routine. Furthermore, the computational demands for such a model is very low, and is quickly solved. This provides an analytical method for determining an initial set of helix parameters (i.e. N , S , and α). To reduce the number of variables and their adjustment range, certain design constraints are used. As mentioned in the introduction, the helix antenna performs satisfactorily as a beam antenna over the circumference range of $\frac{3}{4}\lambda \leq C \leq \frac{4}{3}\lambda$ and this corresponds to $15.84 \text{ cm} \leq C \leq 28.17 \text{ cm}$, at the center frequency of 1420 MHz. Where, the peak gain occurs at nearly $\pi D_h/\lambda \sim 1.135$ ($D_h = 7.63 \text{ cm}$) for a short helix $N < 10$ [57]. Based on extensive experimental work, it is

suggested that the optimum pitch angle for a helix antenna is between $12^\circ \leq \alpha_h \leq 14^\circ$. For the initial design a pitch angle of 13° was chosen. The number of helix turns N , were iteratively adjusted until a -10 dB beamwidth of 78.7° was achieved. This exercise led to an initial right-handed helix design with parameters; $N = 7$; $\alpha_h = 13^\circ$; $D_h = 7.63$ cm; $S = 5.53$ cm and $h = 38.73$ cm. The field pattern of this model is shown in figure 5.5, and it is seen that the -10 dB beamwidth is close to the desired value. To verify the accuracy of this simple model a helix antenna with exactly the same specifications was simulated in FEKO¹. The helix antenna can be approximated as a thin wire problem and this can efficiently be solved with the Method of Moments (MoM) technique. The diameter of the helix conductor is chosen to be $2a = 3$ mm, as this is a standard thickness. An infinite ground plane was assumed for this simulation, of which the far-field pattern results are shown in figure 5.5. The principal pattern in the $\phi = 0^\circ$ and $\phi = 90^\circ$ are in good agreement, resulting in an almost ideally symmetrical pattern. A comparison between the analytical and MoM solutions clearly show some variances. The most noticeable difference being the more directive pattern predicted by the MoM solution. The -10 dB beamwidth is now only 69.6° , this is a 13.07% difference. From figure 5.5, the maximum sidelobe levels are close to -7.3 dB and -9.5 dB for the analytical and the numerical (MoM) solutions, respectively. Both these levels are unsatisfactorily high. In the literature study it was mentioned that sidelobe levels can be improved by shaping the ground plane. This is investigated in the section to follow.

5.4.1 Optimising the pattern characteristics of the helix feed by shaping the ground plane

The size and shape of the ground conductor has a significant impact on the pattern characteristics of an axial-mode helical antenna [58]. The gain can be improved by as much as 4 dB, and the use of a conical ground plane has shown to suppress the sidelobe levels and improve the axial ratio. An alternative approach would be to perform an in-depth parametric study, based on the helix parameters. The dependency between the parameters could make this a complicated task. It is decided that the preferred approach for improving the pattern characteristics is to shape the ground plane.

Only square, circular and cylindrical ground conductors are considered, as shown in figure 5.6, since the manufacturing of a conical ground plane is more complicated. It is recommended that the ground plane size should be between $0.5\lambda - 0.75\lambda$ [58]. As an initial choice, the side lengths l_{gp} for the square ground plane of figure 5.6, was chosen as 0.75λ . For this study, the size of the ground planes were fixed to $0.75\lambda \times 0.75\lambda$. This insures that the ground plane size does not influence the patterns, and only the shape dependency can be investigated. In this regard, the cylindrical and circular ground conductors both have a diameter d_{gp} of 0.846λ . Furthermore, the height h_{gp} of the cylindrical cup is chosen as 0.25λ , this is suggested to be optimum [58]. The initial helix parameters given in section 5.4, are kept for consistency. The simulated results for both the RHCP and LHCP patterns at 1420 MHz in both principal planes ($\phi = 0^\circ, 90^\circ$),

¹FEKO is a comprehensive electromagnetic simulation software tool, which uses multiple CEM techniques, provided by EMSS <http://www.feko.info/>.

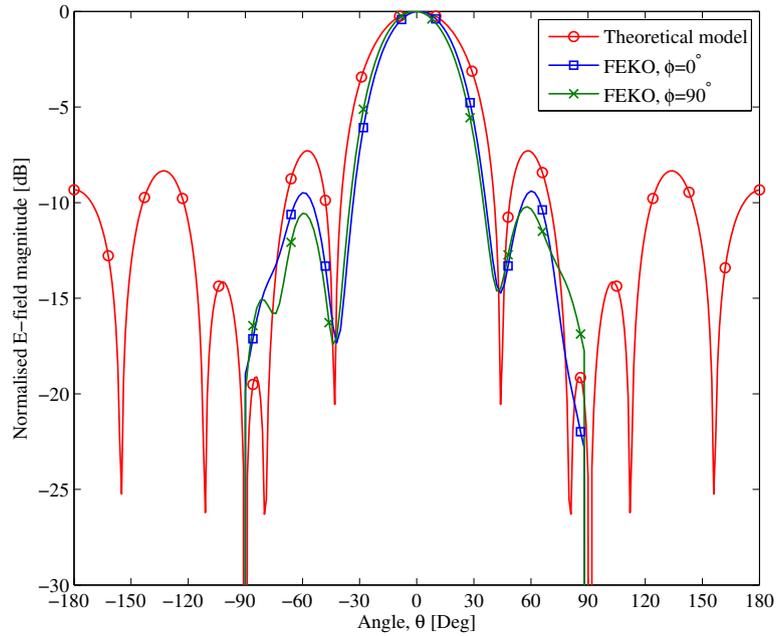


Figure 5.5: The far-field patterns for the initial right-handed helical beam antenna, with parameters, $N = 7$, $\alpha_h = 13^\circ$, $D_h = 7.63$ cm, $S = 5.53$ cm and $h = 38.73$ cm. The simulated field patterns at 1420 MHz for both the analytical- and the FEKO-models are shown in this figure. The FEKO solution assumes an infinite ground plane, which restricts the pattern calculation to $-90^\circ \leq \theta \leq 90^\circ$ (above the ground plane).

are shown in figure 5.7, for the three different ground plane geometries.

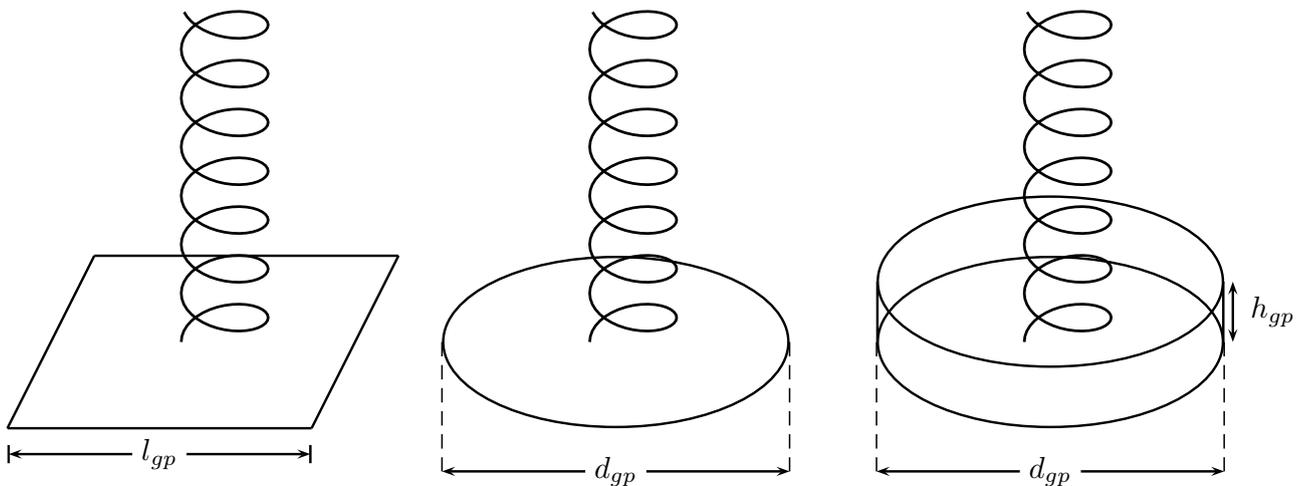
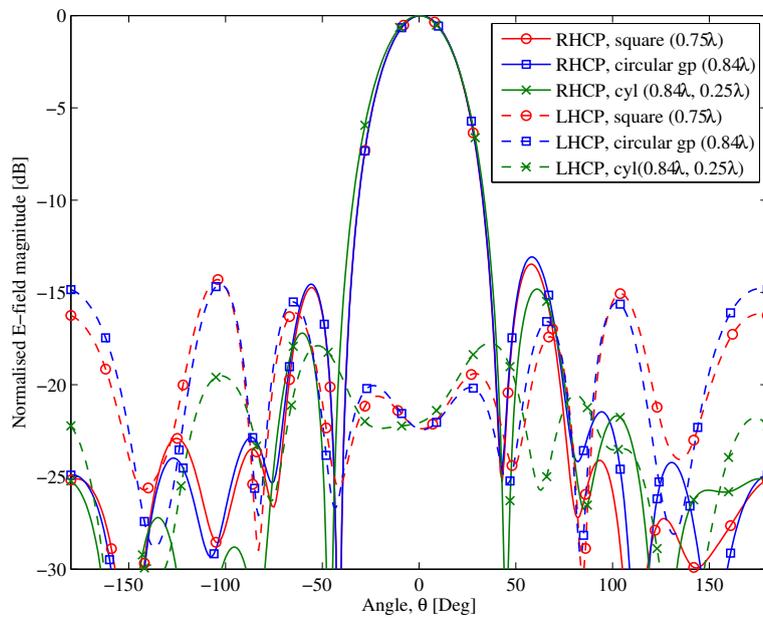
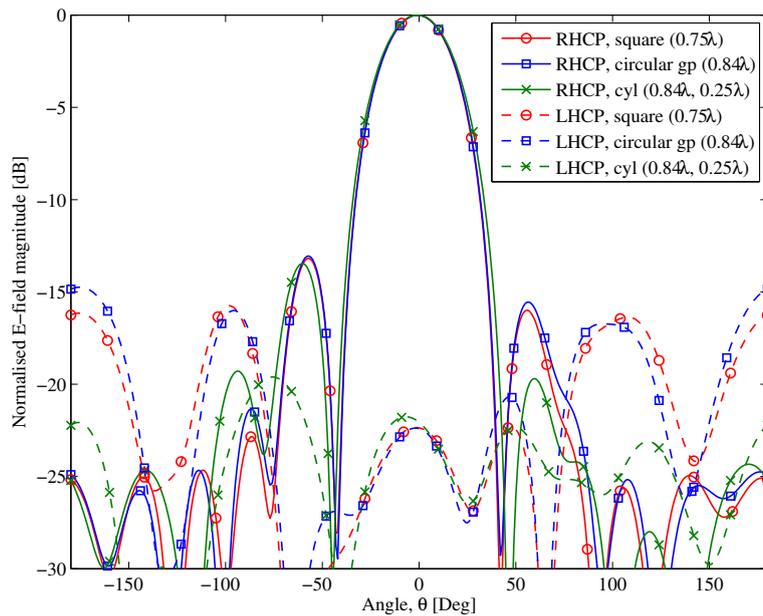


Figure 5.6: A helical antenna above a square, a circular and a cylindrical ground plane.

From figure 5.7, it is evident that the shape of the ground plane influences the pattern characteristics of the axial-mode helical antenna, especially the sidelobes and back radiation. The



(a)



(b)

Figure 5.7: The RHCP and LHCP patterns in the (a) $\phi = 0^\circ$ - and (b) $\phi = 90^\circ$ -planes for the helical beam antenna with parameters, $N = 7$, $\alpha_h = 13^\circ$, $D_h = 7.63$ cm, $S = 5.53$ cm and $h = 38.73$ cm, at 1420 MHz. This figure compares the patter performance for three different ground plane geometries of equal size. The square- with side lengths 0.75λ , circular- with diameter $d_{gp} = 0.846\lambda$ and cylindrical- ground plane with a diameter and height of $d_{gp} = 0.846\lambda$, $h_{gp} = 0.25\lambda$.

main lobes for the square and circular ground planes are very close to identical. Both have a HPBW of 40.18° . In both the $\phi = 0^\circ$ - and $\phi = 90^\circ$ -planes the main lobe for the cylindrical ground plane is slightly broader with a HPBW of 42.23° . Of particular importance is the lower side lobe levels achieved with the cylindrical ground plane, where the sidelobes are below -14.82 dB in figure 5.7(a), and below -13.5 dB in figure 5.7(b). The sidelobes for the square and circular ground planes are -13.46 dB and -13.08 dB, respectively in the $\phi = 0^\circ$ plane, and the levels are slightly better in the $\phi = 90^\circ$ plane. The cylindrical ground plane has an almost 2 dB better sidelobe performance, as compared to the circular ground conductor. Figure 5.7 also shows the LHCP patterns, since this is essentially the cross-polarisation of the RHCP helix antenna. The radiation off the back of the helix is predominately LHCP, as expected. This is where the cylindrical ground plane is far superior. The LHCP levels in the backwards direction ($\theta = \pm 180^\circ$) for the cylindrical ground plane is only -22.23 dB, and this is much lower than the -16.2 dB and -14.8 dB levels for the square- and circular ground conductors, respectively. This much lower back radiation is especially important in radio astronomy, since this portion of the pattern is directed to the warm ground. In conclusion, the cylindrical ground conductor provides superior performance, and because of this a parametric study was performed to find the optimal dimensions for this ground plane.

The parametric study consisted of adjusting the cylindrical cup diameter d_{gp} and height h_{gp} . The initial dimensions of $d_{gp} = 0.423\lambda$ and $h_{gp} = 0.25\lambda$ were used as reference values, around which d_{gp} and h_{gp} , were adjusted in turn. For this study the following pattern characteristics were assessed; the HPBW, the maximum sidelobe levels and the cross-polarisation levels (LHCP) in the backwards direction. The tuning ranges for these parameters are; $0.4\lambda \leq d_{gp} \leq 1.5\lambda$; and $0.05\lambda \leq h_{gp} \leq 0.5\lambda$. The range over which d_{gp} was adjusted is limited by the diameter of the helix D_h , at the lower end, and the upper limit by a consideration for aperture blockage. For the height h_{gp} , the lower limit is where the cylindrical cup is simply a circular ground conductor ($h_{gp} \rightarrow 0$), and the upper limit is set arbitrarily to 0.5λ . The results of this parametric study is shown in figure 5.8.

The trends observed in figure 5.8, clearly show that the dimensions of the cylindrical ground plane significantly influences the pattern shape of the helical beam antenna. By properly choosing the diameter and cup height of the ground conductor, an optimum pattern response can be achieved. In terms of the HPBW, it is shown in figure 5.8(a), that as the diameter d_{gp} increases, the HPBW tends to slowly decrease, until about $d_{gp} = 1.2\lambda$, where it then decreases more steadily. This shows that the radiation pattern can be made more directive (within limits), by increasing the size of the ground plane. Ideally, d_{gp} should be kept as small as possible, in order to minimise the aperture blockage caused by the feed. Figure 5.8(b), illustrates that increasing the height h_{gp} , also monotonically increases the HPBW of the helix. The effective length of the helix is the portion outside of the cylindrical cavity. Thus, as h_{gp} increase the helix effectively becomes shorter, and this then results in a decrease in directivity. This is easily compensated for by lengthening the helix.

The sidelobe levels can also be improved by adjusting the ground plane size, see figures 5.8(c) and 5.8(d). The diameter of the ground plane has a more noticeable influence on the sidelobe

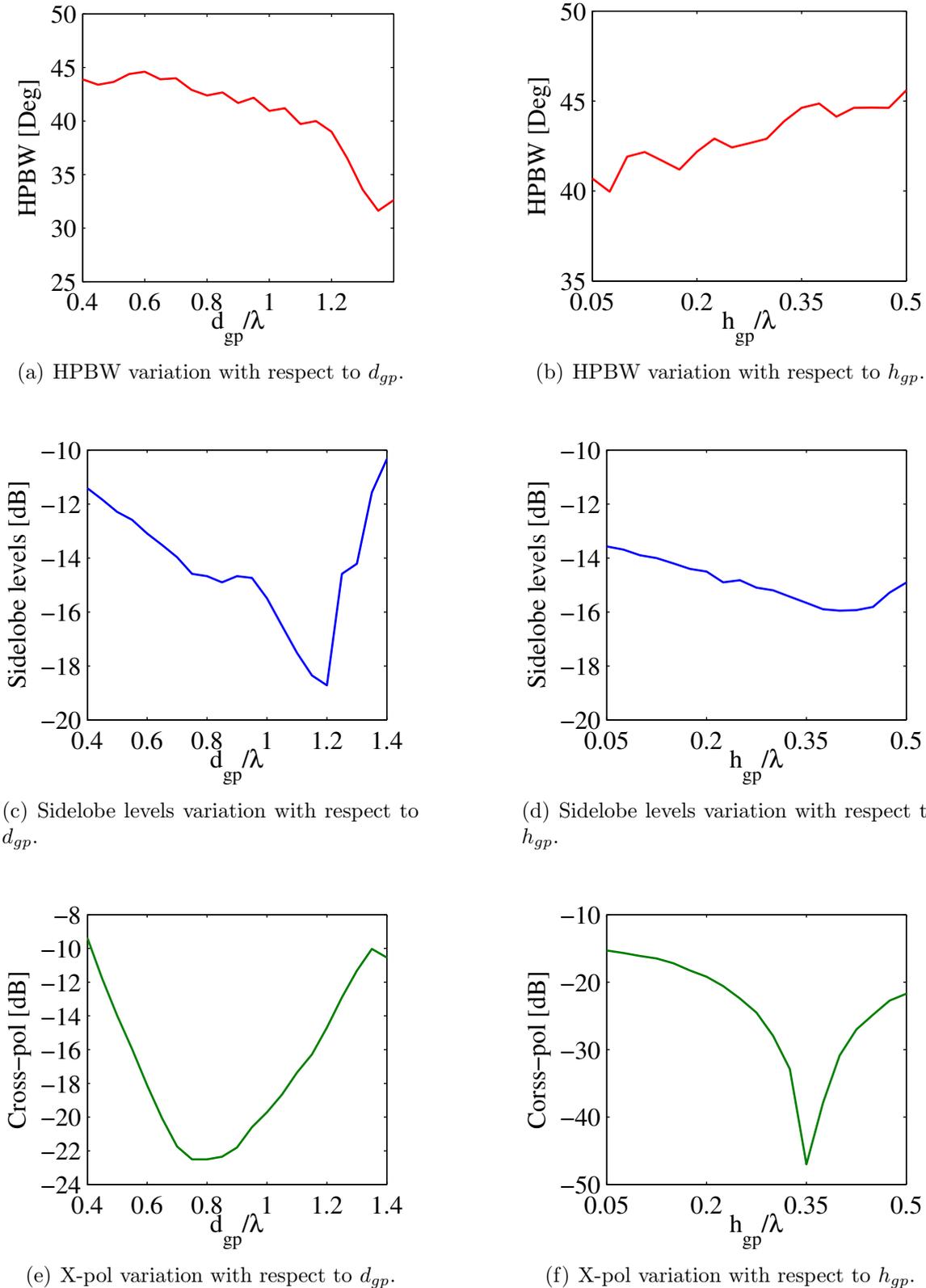


Figure 5.8: Cylindrical ground plane parametric study results. The height h_{gp} remained fixed at a value of 0.25λ for variations in d_{gp} , and $d_{gp} = 0.84\lambda$ when h_{gp} was adjusted. All results were computed for 1420 MHz, and only the $\phi = 0^\circ$ -plane was considered.

levels. There is almost a 10 dB variation over the d_{gp} range considered. It is also worth noting that there seems to be a singular optimum d_{gp} value, where the sidelobe levels are at their lowest. In this case, it is at about $d_{gp} = 1.2\lambda$, and the maximum sidelobe is below -18.7 dB. Increasing the cavity height also causes a steady decrease in the sidelobe levels. This was expected since it was pointed out in [58], that the cup intercepts the radiation of the first few helix turns. This results in sidelobe suppression.

The cross-polarisation levels is shown in figures 5.8(e) and 5.8(f) to decrease when both d_{gp} and h_{gp} increases. Until some threshold is reached, and then there is a degradation in the cross-polarisation performance. Thus, the cylindrical ground plane dimensions must carefully be selected to avoid poor cross-polarisation performance. Figure 5.8(e) shows that the optimal d_{gp} value is around 0.8λ , and it is clear from figure 5.8(f) that the optimal h_{gp} value is around 0.35λ . It can be concluded that by adjusting the size of the cylindrical ground plane, the pattern characteristics of a helical beam antenna can be improved. This study was conducted at a single frequency of 1420 MHz, and for a specific set of helix parameters. Therefore, to have a more complete view of how the ground plane dimensions influence the overall helix antenna performance, other aspects such as the operating bandwidth, axial ratio, and pattern symmetry must also be investigated, for a large sample of different helix parameters. From the curves presented in figure 5.8 the cylindrical ground plane dimensions were chosen as $d_{gp} = 1\lambda = 21.1$ cm and $h_{gp} = 0.35\lambda = 7.39$ cm, as this seemed to give the best cross-polarisation performance, while still having low sidelobe levels. The decrease in directivity can be compensated for by using a slightly longer helix antenna. These values are close the optimal values specified in [58], $d_{gp} = 1\lambda$ and $h_{gp} = 0.25\lambda$.

5.4.2 Mechanical support considerations

Since the helical feed antenna is made out of 3 mm diameter ($2a$) copper wire, it is not mechanically self-supportive. There are various methods proposed for improving the mechanical rigidity of a helix antenna. One popular choice is to use a thin metallic post in the center of the helix, and then to use either periodic supports or insert a low-loss styrofoam ($\epsilon_r = 1$) core [57]. Another variant is to use a hollow cylindrical tube to strengthen the antenna structure [62, 63]. Polyvinyl Chloride (PVC) tubing is readily available at very low cost, and it is probably the easiest method for supporting the helix antenna. However, PVC is a dielectric material with a relative permittivity ϵ_r of approximately ~ 2.7 F/m. The performance of the helix antenna will almost certainly be affected, and it is to be expected that the onset of the axial-mode radiation will be shifted down to a smaller circumference C_λ , accordingly by

$$\lambda_g = \frac{\lambda_0}{\sqrt{\epsilon_{\text{eff}}}},$$

where λ_g is the guided wavelength, and $1 \leq \epsilon_{\text{eff}} \leq \epsilon_r$ is the effective dielectric constant, whereby the air and dielectric regions have been replaced by a homogeneous medium. The range of C_λ over which the PVC-core helix optimally radiate (in the axial-mode), is from 0.7 to 1, assuming the free space wavelength [63]. This differs from the range given in section 5.4, for an

air-core helical antenna. Even though the thickness of most PVC cylinders are practically thin (1 mm – 5 mm), it was not apparent how this would influence the helical beam antenna. If the fields are confined along the axis of the helix antenna, then the dielectric loading of the PVC can potentially have a significant effect on the performance. This was best investigated by performing a full-wave analysis of the problem. In order to draw a comparison between the PVC- and air-core helix antennas, the same dimensions were considered for both. The parameters used are given in section 5.4. The shift in the peak gain is used to determine the influence the PVC-core has. The boresight ($\theta = 0^\circ, \phi = 0^\circ$) gain curves for the air- and PVC-core helices are given in figure 5.9.

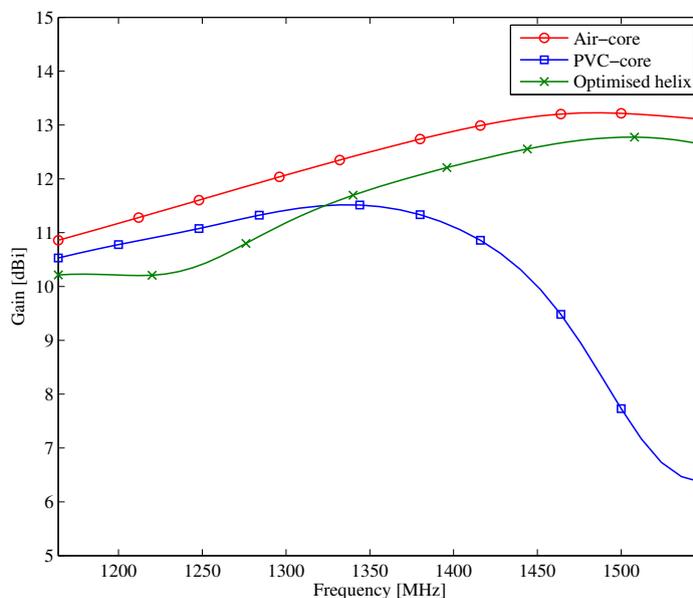


Figure 5.9: The gain on boresight for the air- and PVC-core helix antennas, with parameters, $N = 7$, $\alpha_h = 13^\circ$, $D_h = 7.63$ cm, $S = 5.53$ cm, $h = 38.73$ cm, $d_{gp} = 21.1$ cm and $h_{gp} = 7.39$ cm. The gain is compared over the frequency range from 1164 MHz to 1548 MHz. The gain for the optimised (manufactured) helix is also shown, and the dimensions are presented in table 5.2.

From figure 5.9 it is seen that the peak gain of the air-core helix antenna is 13.23 dBi, this is somewhat higher than the 11.65 dBi gain computed using the approximate formula

$$G = 6.2 \left(\frac{C}{\lambda} \right)^2 N \frac{S}{\lambda}. \quad (5.13)$$

In [57], it is shown that the empirical gain formula given by (5.13) is dependent on the helix parameters and the gain factor of 6.2 is not fixed, but can vary quite substantially. In comparison, the PVC-core antenna has a maximum gain of 11.65 dBi, this is 1.71 dB lower than the air-core helix, but yields a gain which is almost identical to the gain computed using (5.13). Apart from the lower gain observed for the PVC-core helix, there is also a clear shift between the peak gains. The peak gain of the air-core helix occurs at about 1.48 GHz, this is close to

the desired 1.42 GHz value, with only a 4.22% error. The peak gain of the PVC-core helix is shifted to the left by approximately 140 MHz, and occurs at 1.34 GHz. From this it can be computed that the effective permittivity, ϵ_{eff} is about 1.219, which is much lower than the relative permittivity of the PVC tubing. To fully appreciate the influence the PVC-core has on the antenna performance, the pattern responses must also be compared. The far-field patterns at 1.34 GHz and 1.84 GHz, for the PVC- and air-core helices, respectively, are plotted in figure 5.10. As expected from the gain curves, the PVC-core helix has a slightly broader main beam, because of its lower gain. The PVC-core helix in fact has superior sidelobe and back radiation performance as compared to the air-core helix. Figure 5.10 shows that the maximum side lobe levels of the PVC-core helix is only -16.28 dB, this is more than 2.2 dB better than the air-core helix. There is also just over a 10 dB improvement in the back radiation (LHCP) levels. It can be concluded that using a PVC-core fulfils its purpose of supporting the helix antenna and the pattern response remains satisfactory.

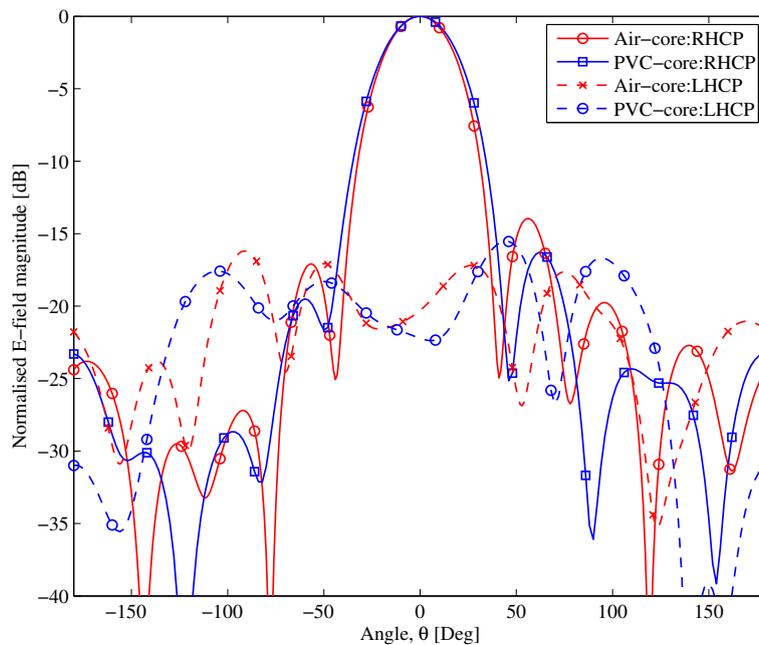


Figure 5.10: The RHCP and LHCP patterns in the $\phi = 0^\circ$ -plane for the air- and PVC-core helical beam antennas with parameters, $N = 7$, $\alpha_h = 13^\circ$, $D_h = 7.63$ cm, $S = 5.53$ cm, $h = 38.73$ cm, $d_{gp} = 21.1$ cm and $h_{gp} = 7.39$ cm. This figure compares the air-core helix patterns at 1480 MHz against the PVC-core patterns at 1340 MHz. These frequency values coincide with the peak gains of the respective helix antennas.

According to the estimated permittivity ϵ_{eff} , the guided wavelength at 1420 MHz is approximately $\lambda_g \approx 19.13$ cm, and the helix diameter required to ensure $C/\lambda \sim 1.135$ is then $D_h = 6.91$ cm. Due to availability, the selected PVC tube has an outer diameter D'_h of 6.3 cm, which corresponds to $D_h = D'_h + 2a = 6.6$ cm. The normalised helix circumference is then

Geometrical Parameter	Value
D_h	6.6 cm
$2a$	0.3 cm
α_h	11.76°
N	8
h	34.52 cm
Height of PVC tube, h_{PVC}	35.52 cm
Total helix length, L_h	162.05 cm
d_{gp}	21.2 cm
h_{gp}	7.4 cm
l_m	8.6 cm
w_m	2.4 cm

Table 5.2: The dimensions of the manufactured (optimised) helical feed antenna.

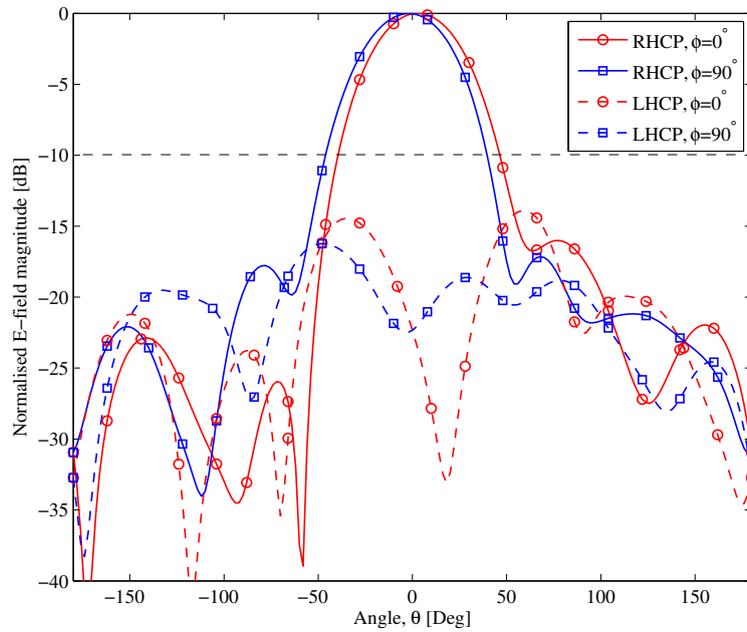
$C_\lambda \approx 0.98$, which is within the range specified in [63]. The thickness, t of the PVC tubing is 3 mm.

5.4.3 Matching the feed antenna to the offset-parabolic reflector

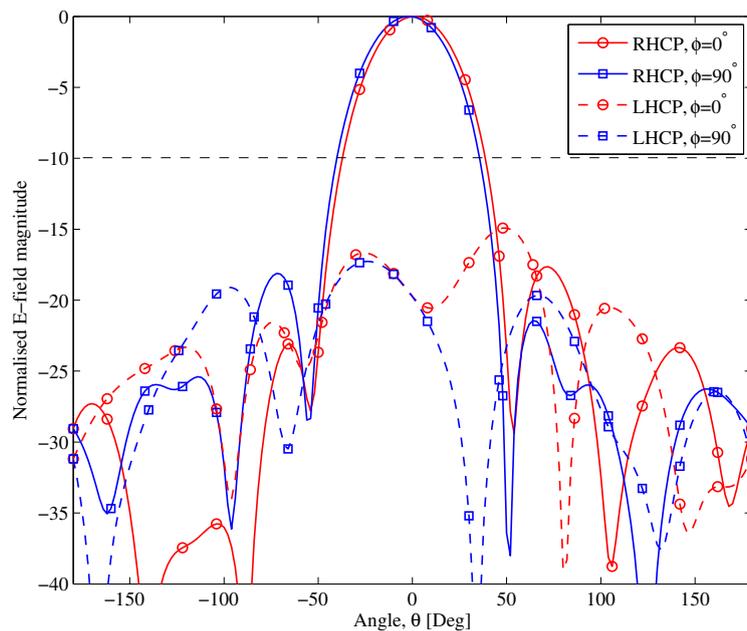
The initial helical beam antenna with parameters, $N = 7$, $\alpha_h = 13^\circ$, $D_h = 7.63$ cm, $S = 5.53$ cm, $h = 38.73$ cm, $d_{gp} = 21.1$ cm and $h_{gp} = 7.39$ cm can now be optimised to properly feed the offset-reflector antenna. The computed F/D_p of the offset reflector is ≈ 0.275 . However, the feed only needs to illuminate the offset section with a smaller subtended angle of $\alpha = 78.76^\circ$. This subtended angle is equivalent to a prime-fed reflector with an $F/D \approx 0.583$. It was mentioned in section 5.3 that an EI of -11 dB is desired. Because of the SPL and the offset-parabolic geometry there exists a 5.2 dB EI difference between the upper and lower tip of the reflector. In [64], a method is presented to compute the feed pointing angle Ψ_f , needed to achieved a balanced EI. In this case it corresponds to an EI of -12.6 dB for a -10 dB FT. This slightly under-illuminated reflector will only cause a 0.26° degradation in the HPBW, as compared to a -10 dB EI. This is acceptable, and the lower sidelobes are also appealing, since any RFI seen through the sidelobes are a bigger concern than a slightly less directive pattern. Thus, it was ultimately decided to design for a -10 dB FT at the half-subtended angle of $\alpha/2 = 39.38^\circ$.

The geometrical parameters of the optimised helical feed antenna is presented in table 5.2. These are the exact dimensions used during the manufacturing of the helical beam antennas.

The radiation patterns at 1292 MHz, 1420 MHz and 1548 MHz are shown in figure 5.11, this corresponds the lower, center and upper operating frequencies of the interferometer. The HPBWs are 50.8° , 45.5° and 41.2° from the lower to upper edge frequencies. This indicates that the main beam becomes more directive as the frequency increases. A more important design specification is the -10 dB beamwidth. At the center frequency the -10 dB beamwidth is 77.2° , this is in close agreement with the desired subtended angle of 78.7° . At the lower- and upper-band frequencies, the -10 dB beamwidth is found to be 85.9° and 71.5° . This suggests that the am-

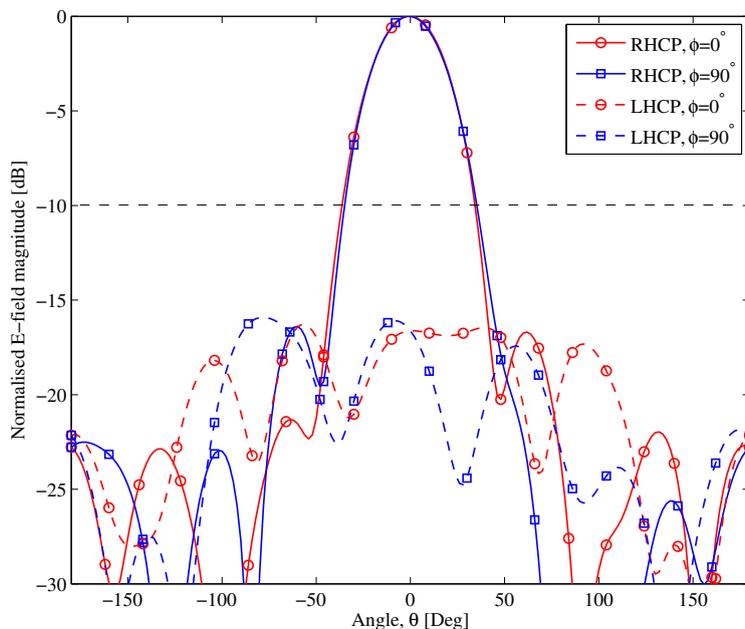


(a)



(b)

plitude tapering efficiency increases at 1292 MHz, because the reflector will be more strongly illuminated, whereas the reflector is somewhat under illuminated at 1548 MHz, and the spillover efficiency increases.



(c)

Figure 5.11: The simulated RHCP- and LHCP-patterns in both the $\phi = 0^\circ$ - and 90° -planes for the optimised helical beam antenna, with parameters $N = 8$, $\alpha_h = 11.76^\circ$, $D_h = 6.6$ cm, $S = 4.3$ cm, $h = 34.53$ cm, $d_{gp} = 21.1$ cm and $h_{gp} = 7.39$ cm. The patterns are shown (a) at 1292 MHz, (b) 1420 MHz and (c) 1548 MHz.

In figure 5.11 it is shown that the sidelobe levels for all three the frequencies considered are well below -15 dB. The cross-polarisation (LHCP) performance in the backwards ($\theta = -180^\circ$) direction is also excellent, where the levels never exceed -20 dB. It is observed that as the frequency increases so does the cross-polarisation levels. The only noticeable flaw of the helical feed antenna is the beam squint at the lower-edge frequency, where the patterns in the two principal planes are offset by $\sim 6^\circ$ from each other.

The gain of the optimised helical beam antenna is shown in figure 5.9. The maximum gain achieved is 12.77 dBi and occurs at about ~ 1500 MHz. There is only a 20 MHz (1.35%) difference between the peak gains of the air-core helix and the optimised helix. This confirms the accuracy of the 0.905 circumference scaling factor previously determined. A useful metric to define the pattern operating bandwidth of a helix antenna, is to use the 2 dB peak gain decrease criterion [57]. In this case, the gain drops to 10.77 dBi at 1274 MHz and 1673 MHz. This corresponds to a bandwidth ratio of 1:1.31. Equation (5.4), yields a bandwidth ratio of 1:1.27, this is almost exactly the same as the simulated result.

To fully characterise the performance of the helical feed antenna, it is also important to evaluate the axial ratio (AR)². The AR is computed using FEKO, and the on-axis ($\theta, \phi = 0^\circ, 0^\circ$) result is given in figure 5.12. The axial ratio is seen to be below 3 dB across the entire operating band,

²The axial ratio is defined as the ratio of the major to minor axes, and is ideally 1 for circular polarisation.

and figure 5.12 illustrates that the AR deteriorates as the frequency increases. More specifically the AR is 1.33 dB, 1.8 dB and 2.58 dB at 1292 MHz, 1420 MHz and 1548 MHz, respectively. The lowest AR is 1.3 dB and occurs at ~ 1.28 GHz. The AR is predicted as 0.53 dB, using (5.5). This differs significantly from the more accurate FEKO result. The AR achieved across the band is acceptable, but there are means of improving it. It is suggested that a larger pitch angle α_h , more helix turns N and a reduced ground plane size can improved the AR of a helical beam antenna [65].

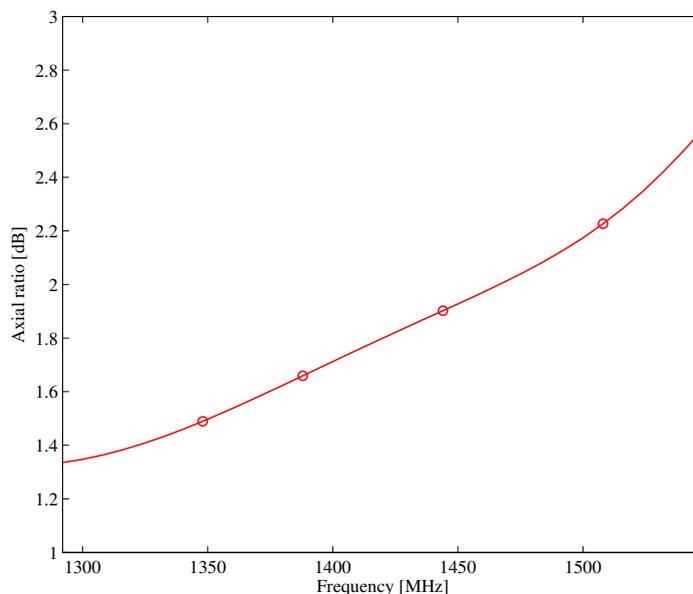


Figure 5.12: The simulated axial ratio (AR) of the optimised helical beam antenna in the on-axis ($\theta, \phi = 0^\circ, 0^\circ$) direction, for the frequency range 1292 MHz – 1548 MHz.

5.4.4 Input impedance matching for the optimised helical antenna

It is well known that a helical beam antenna with uniform conductor size has a nominal impedance of about 140Ω , across most of the frequency band [66]. An approximate formula for the input impedance of an axial-mode helix antenna is

$$R_{\text{in}} = \frac{140C}{\lambda}. \quad (5.14)$$

At 1420 MHz, using (5.14) gives a value of $R_{\text{in}} = 137.4\Omega$ for the optimised helix. This is not a convenient value, since the input of the low-noise front-end is matched to 50Ω . Consequently, an impedance matching technique is necessary to minimise the mismatch. In the literature, various impedance matching techniques have been proposed [66, 62]. Some of these methods include: increasing the conductor size close to the feed point at the ground plane; tapering the helical turns at the end; and using a triangular copper strip near the feed point [66, 62]. Using the first technique, the flat conductor and ground plane acts as a transmission line which transforms the high impedance to a lower value, but it reduces the bandwidth [66]. The impedance

bandwidth reduces to about 40% from 70% ($VSWR < 2$) for a $50\ \Omega$ match [51, 66]. Furthermore, this technique is better suited for a self-supported helix antenna, it becomes challenging to bend the flattened conductor around the periphery of the PVC tube. Using the technique of tapering the helix turns is not plausible, since the PVC support is of constant radius, and again this technique is more suited for an air-core helix. The use of a triangular copper strip is an attractive solution in comparison to the other two techniques discussed. It is simple and easily to implement, and this technique has successfully been used before to match a PVC-core helical beam antenna [62, 63]. Another advantage of this input matching technique is that the strip length l_m is not critical. It is shown in [63], that as long as the strip length is not too far away from some optimal length, varying the length does not cause a serious degradation in the input impedance match. Therefore, the matching process is somewhat simplified. Figure 5.13, shows a detailed sketch of this matching network. The copper triangular strip is shaped around the periphery of the helix antenna, and follows the conductor. The corner of the strip is soldered to the inner conductor of the connector.

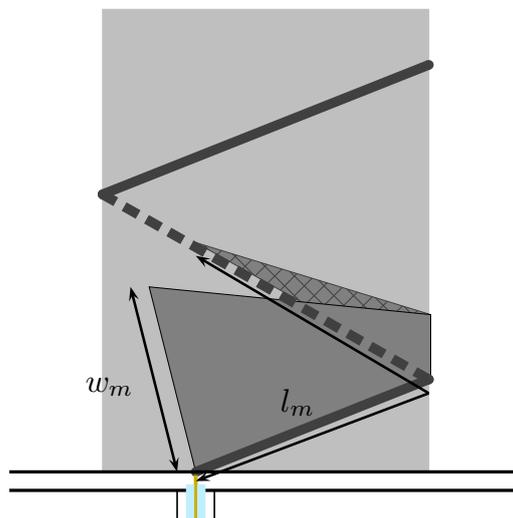


Figure 5.13: Illustration of the triangular copper matching strip.

At the center frequency the impedance of the optimised helical beam antenna is found to be $Z_{in} = 80.82 - j24.77\ \Omega$, as shown in figure 5.14, whereas the approximate input resistance was calculated as $137.4\ \Omega$. Secondly, the input impedance is not purely real as there is some capacitive reactance. Equation (5.14) is only an approximation, and the input impedance of a helical antenna depends on various parameters, such as the wire radius, number of turns, helix pitch, helix radius, the location of feeding point with respect to the ground plane, etc. This can explain the deviation between the impedance values. What is of interest in figure 5.14, it that the resistance and reactance does not vary greatly over a wide frequency range, where $60\ \Omega \leq R_{in} \leq 87\ \Omega$ and $-42\ \Omega \leq X_{in} \leq -22\ \Omega$ for $1292\ \text{MHz} \rightarrow 1548\ \text{MHz}$. This is characteristic of a helical beam antenna, and this shows the broadband impedance nature of this type of antenna. If an inductive reactance can be added to cancel the capacitance of the antenna, and if the resistance can be lowered than an ideal match can be achieved over the desired operating

band.

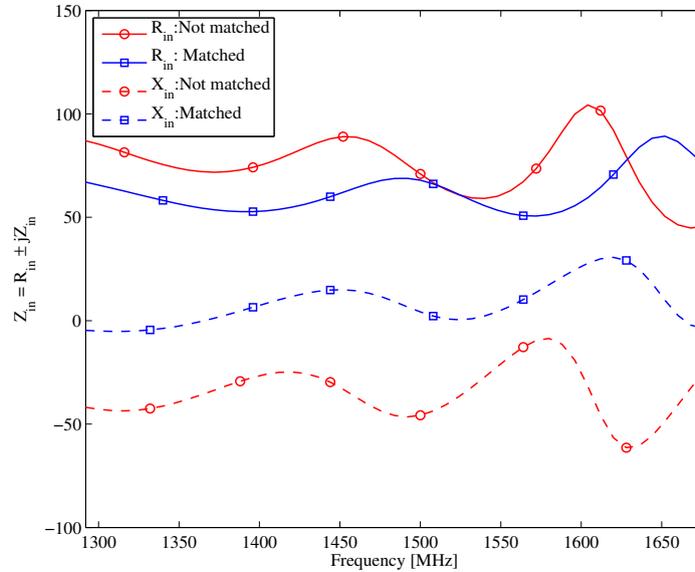
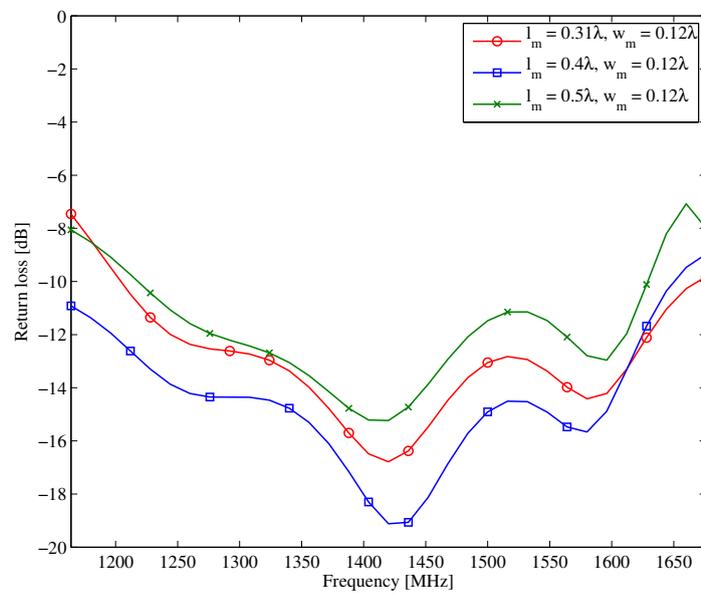


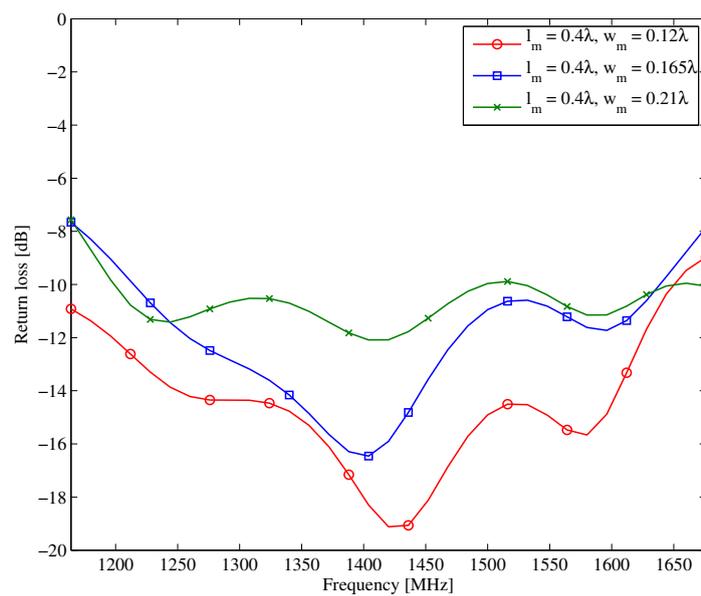
Figure 5.14: The input impedance of the optimised helical beam antenna, where the resistance and reactance for the non-matched and matched cases are shown as a function of frequency. The matching technique employed consists of a triangular copper strip with length $l_m = 8.6$ cm and width $w_m = 2.4$ cm.

By inserting the triangular copper strip close to the ground plane, the helical beam antenna can effectively be modelled as a parallel RLC resonant circuit [62]. The resonant frequency of a parallel RLC circuit is given by $\omega_0 = 1/\sqrt{LC}$. This implies that by adjusting the length l_m and width w_m of the triangular copper strip, the impedance of the helical beam antenna can be controlled. There seems to be no theoretical equivalent circuit model for the triangular copper strip. Thus, as a starting point for the strip design, the optimum dimensions given in [62] were normalised with respect to wavelength and then scaled using the center frequency of 1420 MHz. It was determined that the strip dimensions are $l_w = 0.4\lambda$ and $h_w = 0.12\lambda$. To investigate the individual influences of l_m and w_m on the impedance match of the helical beam antenna, each dimension was adjusted in turn whilst keeping the other dimension fixed. In figure 5.15 the RL results from this investigation is shown. As the length is varied from $l_m = 0.3\lambda - 0.5\lambda$ for $w_m = 0.12\lambda$, it seen from figure 5.15(a), that the resonance shifts towards a lower frequency. This indicates that the optimum match can be positioned at the desired frequency by adjusting the length of l_m . It is also seen that l_m influences the absolute match at the resonance frequency, where $l_m = 0.5\lambda$ has a maximum (lowest) RL of -15.23 dB, and $l_m = 0.4\lambda$ a return loss of -19.2 dB.

In contrast, adjusting w_m , while l_m is fixed at 0.4λ , has a much more profound effect on the overall match. It is shown in figure 5.15(b), that for $w_m = 0.21\lambda$ the RL is only -12 dB, but for $w_m = 0.12\lambda$ the RL is -19.2 dB. This is a 7 dB difference, and this suggests that w_m can



(a)



(b)

Figure 5.15: A study to investigate the change in the simulated return loss of the optimised helical beam antenna, for different triangular copper strip dimensions, where w_m is fixed at 0.2λ while (a) l_m is changed, and $l_m = 0.4\lambda$ while (b) w_m is varied.

be adjusted to improve the overall match of the helical beam antenna. The procedure thus followed to optimise the helix antenna consisted of adjusting l_m such that the maximum RL was achieved at the center frequency (1420 MHz), and then w_m was adjusted to improve the absolute RL over the required passband. Following this procedure, the optimum dimensions were found to be $l_m = 8.6$ cm and $w_m = 2.4$ cm.

The return loss of the helical feed antenna with the optimised matching network is compared to the same helix antenna without a matching network, in figure 5.16. Where the helix antenna without the matching network is poorly matched over majority of the band, and only at a few frequencies are matched better than $\text{VSWR} < 2$. The helix with the matching network is well matched over the entire passband (1292-1548 MHz). The RL is seen to be below -15 dB over this portion of the band, and even goes down to -28 dB at parts. Assuming a minimum RL of -10 dB ($\text{VSWR} < 2$), the helix antenna with the matching network has a fractional impedance bandwidth of more than 36%, and the specification of at least 18% bandwidth is met. The input impedance of the matched helical beam antenna is also shown in figure 5.14. The resistance is seen to be close to 50Ω over the entire band, and the reactance is ideally close to 0Ω as well.

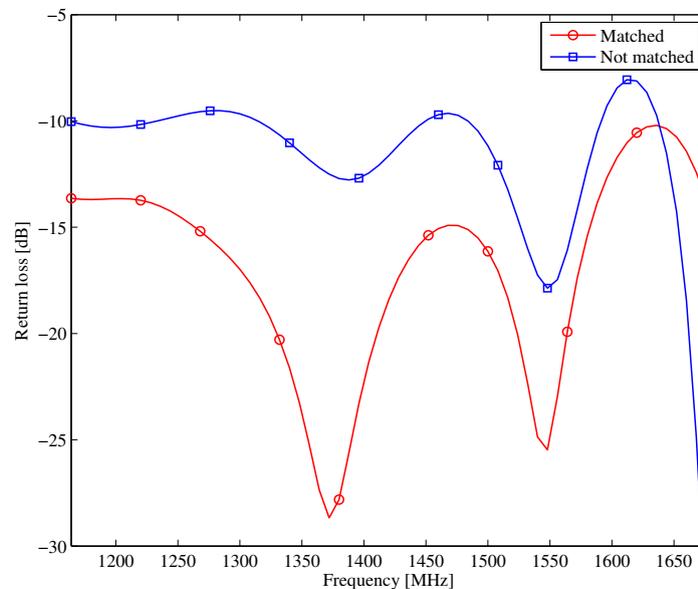


Figure 5.16: A comparison between the return loss [dB] of the optimised helical beam antenna with the designed matching network of dimensions $l_m = 8.6$ cm and $w_m = 2.4$ cm, and the same helix antenna without a matching network.

5.5 Integrating the helical feed antenna with the offset-parabolic reflector

In the previous section the design details of the helical feed antenna was presented, in this section the feed antenna is used as the primary beam that illuminates the offset-parabolic re-

flector. In this regard, it is important to specify the relative position of the feed antenna with respect to the reflector. Firstly, the phase center of the helical beam antenna is computed. Then, the feed pointing angle, Ψ_f is defined. Once, the optimum feed position is known, then the performance of the secondary pattern can be evaluated.

5.5.1 Phase center of the helical beam antenna

It is known that a uniform phase distribution over the aperture of the reflector gives the most desirable radiation pattern in terms of low sidelobes and high directivity. From the geometry of a paraboloidal reflector, equiphase aperture illumination can be achieved by locating a point source at the focus [65]. However, the helix or any other antenna for that matter is not a point source, but if the radiated fields in some angular region can be characterised as spherical equiphase surfaces, then a phase center can be defined. This phase center is then placed at the focus of the parabolic reflector. In [65] a method is proposed for determining the phase-center of a helical beam antenna. This method assumes a helix centered at the origin with its axis of symmetry along the z-axis, as shown in figure 5.17. If the point PC is the phase center which is located d from the helix center. Then the radiation in the $\phi = 90^\circ$ plane will have a locus of constant phase, described by a point $p(R, \theta, \pi/2)$, that is approximately a circular arc centered at PC , as θ varies. This spherical radiation from point PC can be described by

$$E_p = |M| e^{-jk[R-d\cos\theta]},$$

where $|M|$ is the amplitude of the field, $k = 2\pi/\lambda$ is the propagation constant, and the far-field approximation has been used. If the phase factor of this spherical source is compared to the phase factor of the radiated field of the helix antenna, $E_p = |M(\theta)| e^{-jk[R-h(\theta)]}$, where $h(\theta)$ is some phase term of the helix, then d can be computed as

$$d = \frac{h(\theta)}{k \cos \theta}. \quad (5.15)$$

The phase center is most easily found by plotting $h(\theta)$ against $\cos \theta$, and d is then the slope of the linear portion of the curve. For the offset-parabolic reflector the phase center only has to be constant in the subtended half angle $\alpha/2$, (5.15) can then be re-written as

$$d = \frac{(h(0) - h(\alpha/2))}{k(1 - \cos \alpha/2)}. \quad (5.16)$$

For the optimised helical feed antenna with parameters as given in table 5.2, the phase center was determined using (5.16). According to (5.16), $h(\theta)$ must be computed for $\theta = 0^\circ$ and $\alpha/2 = 39.38^\circ$. This is done using FEKO, where the phase of the E_θ -field was plotted for $0^\circ \leq \theta \leq 39.3^\circ$, with the helix centered along the z-axis. The phase of the E_θ -field for $d = 0$ cm is shown in figure 5.18. It is seen than in both planes, the phase errors exceed the maximum allowable phase variation of 22.5° . This indicates that the center of the helix does not coincide with its phase center. The averaged distance d from the helix origin was computed as $d = -5.14$ cm. The negative sign indicates that the helix must move up (along the z-axis). The phase of the E_θ -field for the helix shifted by $d = -5.14$ cm is shown in figure 5.18. It is clear

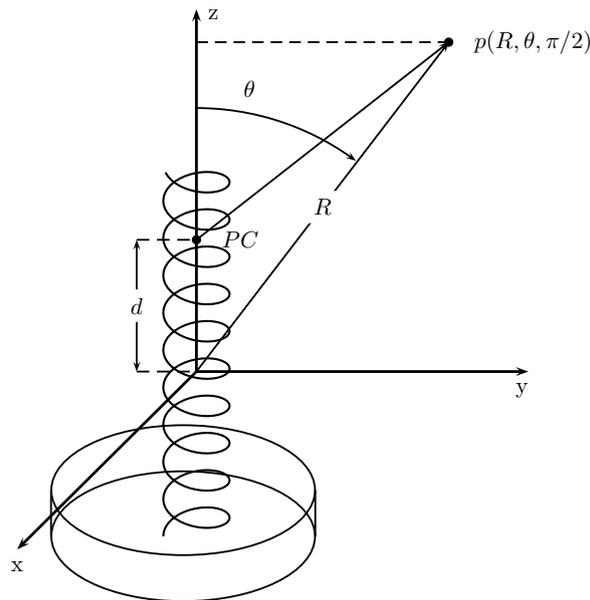


Figure 5.17: Geometry for determining the phase center of the helical beam antenna.

that the phase error across the subtended angle of the reflector is much less, where for the $\phi = 0^\circ$ -plane the phase varies by less than 10° , and less than 8° in the $\phi = 90^\circ$ -plane.

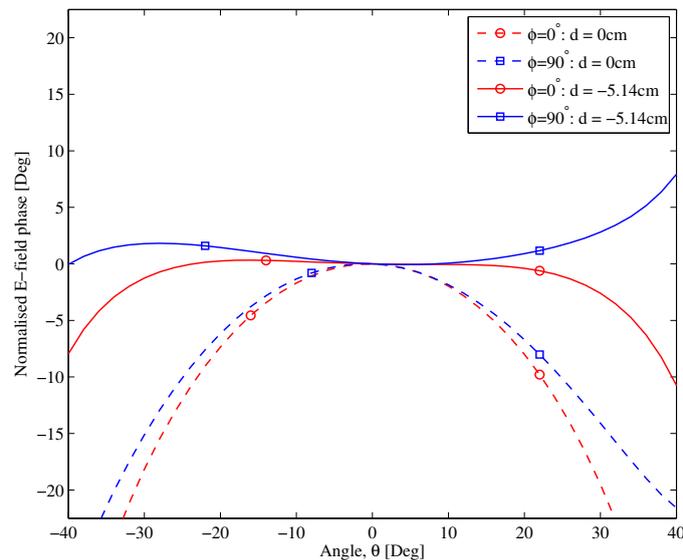


Figure 5.18: The normalised E-field phase of the optimised helical feed antenna across the half-subtended angle $\alpha/2$ of the offset-parabolic reflector, before ($d=0$ cm) and after ($d=-5.14$ cm) adjusting the helix to its phase center location. This simulation was performed at the center frequency of 1420 MHz. The phase curves have been shifted to coincide with 0° at $\theta = 0^\circ$.

5.5.2 Setting the feed pointing angle

It was mentioned early on in this chapter that the feed pointing angle Ψ_f , can potentially improve the secondary pattern. The different pointing angles considered for the primary feed; was towards the angle subtended in the bisect direction β ; the angle subtended to the center of the projected aperture β' ; and the subtended angle which results in equal (balanced) EI, Ψ_E . The pointing angles β and β' are given in table 5.1. A procedure is presented in [64], for determining Ψ_E . Following their procedure the equal EI angle was computed as $\Psi_E = 48.7^\circ$. The secondary patterns of the offset-parabolic reflector fed by the helical beam antenna, for the different pointing angles are shown in figure 5.19. To reduce the computation time needed to compute the secondary patterns, the helical beam antenna and the parabolic reflector were decoupled from each other, and simulated separately. The helical feed was solved using the MoM and the far-field pattern data was exported, and later recalled as the primary source for the reflector. This is known as the spherical mode excitation method in FEKO. The reflector used the PO technique. In both principal planes there does not seem to be any noticeable difference in the main beam for the three different feed pointing angles considered. In both principal planes the HPBW is measured to be about 16.4° . There are slight differences in the sidelobe levels, but these levels are below -25 dB. The only slight advantage of the β' feed angle is the better F/B ratio. The F/B ratio is lower by 2.3 dB in comparison to the β feed angle. In conclusion, there is no real advantage in picking one of these feed angles over the other. However, because of the slight F/B improvement it is decided that the β' angle will be used.

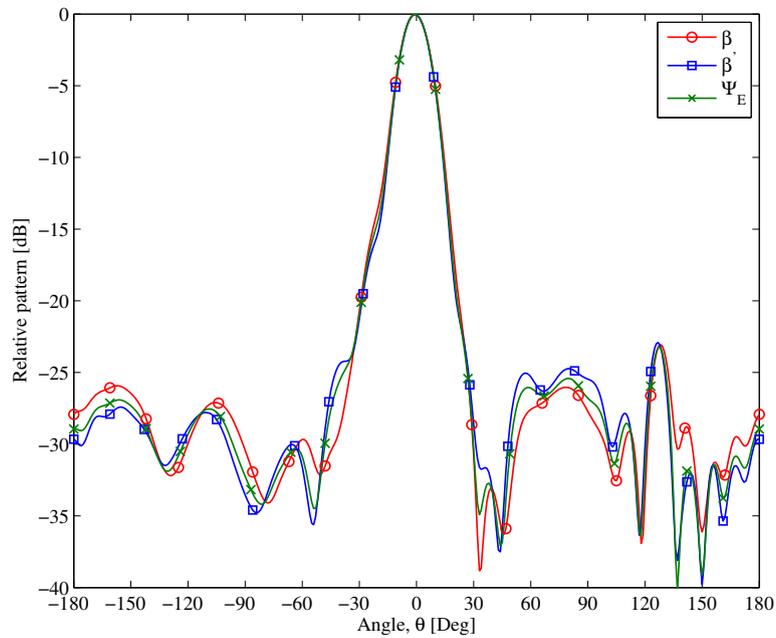
5.5.3 Simulation of the helical beam antenna fed offset-parabolic reflector

The details of the simulation results for the 92 cm offset-parabolic reflector fed by the optimised helical beam antenna, are presented in this section. It is important to note that the sense of circular polarisation changes upon reflection from the dish. More specifically, the helical feed antenna is RHCP and the field radiated off the reflector is then LHCP. The field patterns of the offset-fed reflector at 1292 MHz, 1420 MHz and 1548 MHz are shown in figure 5.20. For higher accuracy, the entire antenna system was coupled and solved using the MoM technique.

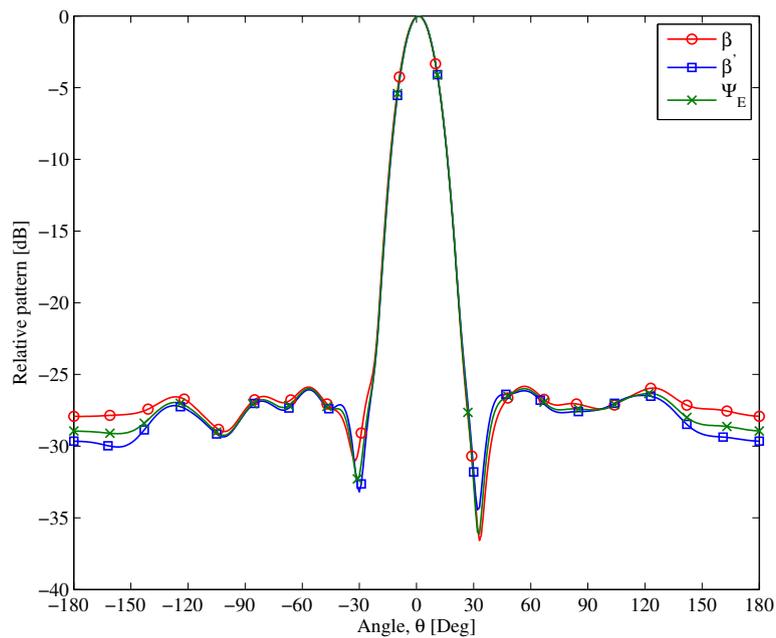
From figure 5.20, it is seen that the pattern shape remains satisfactory across the entire operating band. The pattern does however improve in terms of sidelobe levels, cross-polarisation (RHCP) and directivity as the frequency increases. The HPBWs are 16.63° , 15.77° and 14.98° at the lower-, center- and upper-frequencies. This is in close agreement to the values that were expected for an EI of -12 dB. Using the approximate HPBW formula in [48]

$$\text{HPBW} = 1.16 \frac{\lambda}{D}, \quad (5.17)$$

the expected HPBW values are 16.7° , 15.26° and 14° . It is apparent in all three figures that the beam is squinted in the asymmetric ($\phi = 90^\circ$)-plane. This result is expected, since it is mentioned in [59], that an offset-parabolic reflector fed by a circularly polarised primary pattern, experiences some squinting from its boresight axis. Furthermore, the direction of squinting is

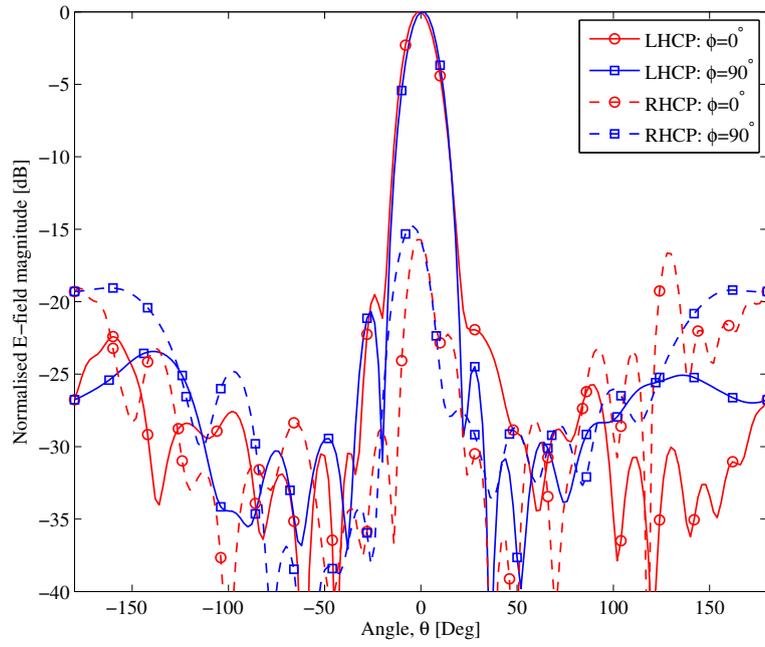


(a)

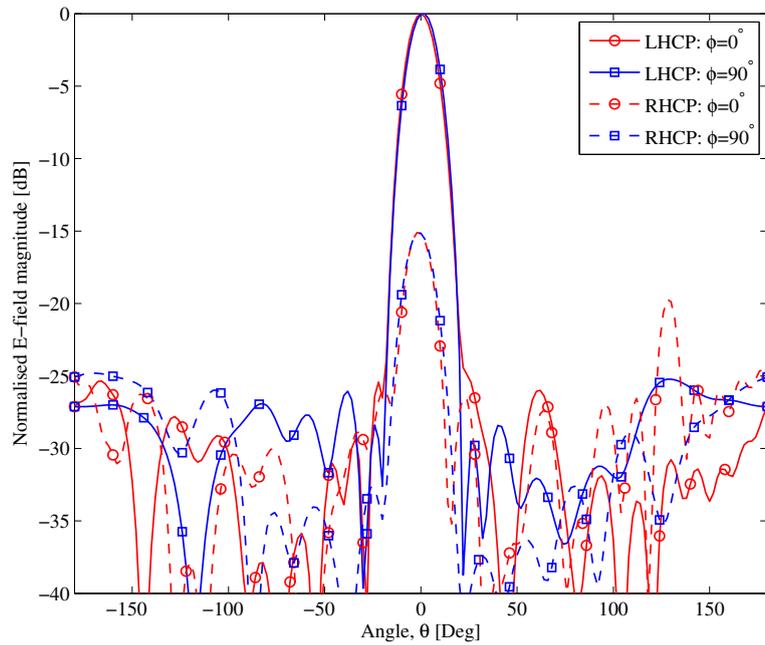


(b)

Figure 5.19: The secondary radiation patterns of the offset-parabolic reflector fed by the optimised helical beam antenna at 1420 MHz. The patterns in the (a) $\phi = 0^\circ$ - and (b) $\phi = 90^\circ$ -planes are shown for the three different feed pointing angles, namely, $\beta = 45.08^\circ$, $\beta = 51.16^\circ$ and $\Psi_E = 48.7^\circ$.



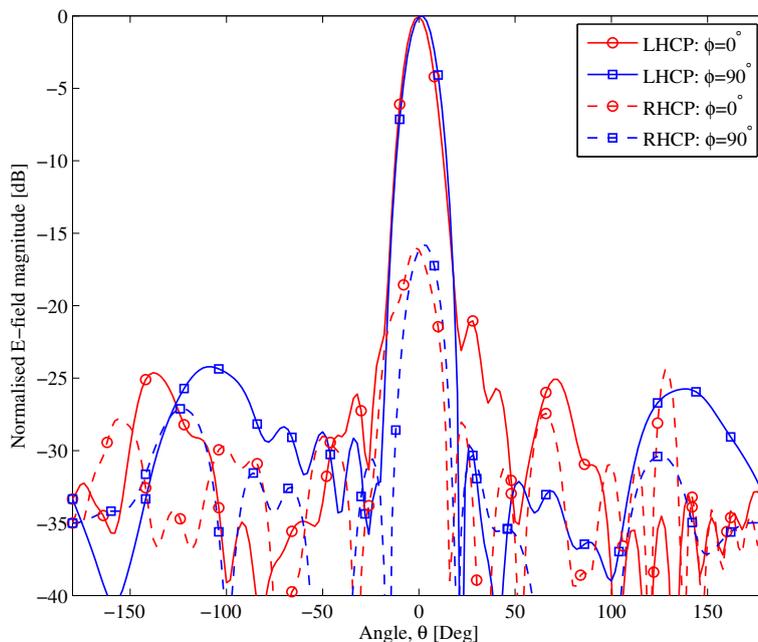
(a)



(b)

dependent on the handedness of the polarisation. An approximate formula for predicting the amount of squint to expect is given as

$$\theta_s = \pm \sin^{-1} \left(\frac{\lambda \sin \Psi_f}{4\pi F} \right). \quad (5.18)$$



(c)

Figure 5.20: The simulated LHCP- and RHCP-patterns in both the $\phi = 0^\circ$ - and 90° -planes for the offset-parabolic reflector fed by the optimised helical beam antenna, with parameters $N = 8$, $\alpha_h = 11.76^\circ$, $D_h = 6.6$ cm, $S = 4.3$ cm, $h = 34.53$ cm, $d_{gp} = 21.1$ cm and $h_{gp} = 7.39$ cm. The patterns are shown (a) for 1292 MHz, (b) 1420 MHz and (c) 1548 MHz.

From (5.18) the expected beam squint is $\sim 1.4^\circ$, for $\Psi_f = \beta' = 51.16^\circ$ and $F = 53.61$ cm. The beam squint in figure 5.20 is seen to be closer to 2° , but this approximation is still fairly close. For an EI of -12 dB the sidelobe levels should be below -22.9 dB [48]. Interestingly, the maximum sidelobe levels are even further down. The maximum sidelobes at 1292 MHz is -19.51 dB, -26 dB at 1420 MHz and -21.05 dB at 1548 MHz. The slightly higher sidelobe levels at the lower-edge frequency is expected since the feed has a broader pattern at 1292 MHz, and it therefore more strongly illuminates the edges of the reflector. The F/B ratio is also seen to improve as the frequency increases. The RHCP back radiation is -20 dB, -25 dB and -35 dB in figures 5.20(a), 5.20(b) and 5.19(c), respectively. The cross-polarisation (RHCP) on boresight is below -15 dB across the band, this seems slightly high and indicates that the radiation is not purely circularly polarised.

The gain on boresight for the offset-fed reflector is shown in figure 5.21. The gain expected at the center frequency is 20.5 dBi, this assumes an aperture efficiency of $\eta_a = 0.6$. From figure 5.21, the gain is more accurately seen to be 21.09 dBi, which meets the design requirement of 19 dBi. This suggests that the aperture efficiencies is more likely to be 0.687, at the wavelength of 21.1 cm. The gain at 1292 MHz and 1548 MHz is further seen to be 19.62 dBi and 21.48 dBi, respectively. This confirms the increased directivity seen in figure 5.20 at the upper-edge frequency. The gain variation across the passband is ~ 1.85 dB, this is fairly flat. The increase in the gain with frequency initially seems strange, given that the primary beam also becomes

more directive as frequency increases. However, the reflector becomes electrically larger with frequency. Furthermore, even though the primary beam becomes narrower and starts to under illuminate the reflector, for a -16 dB EI the HPBW is found to be 14.36° at 1548 MHz, which is still more directive than the -12 dB EI case at 1292 MHz. However, at about 1532 MHz the reflector becomes seriously under illuminated, and the gain is seen to decrease at this point.

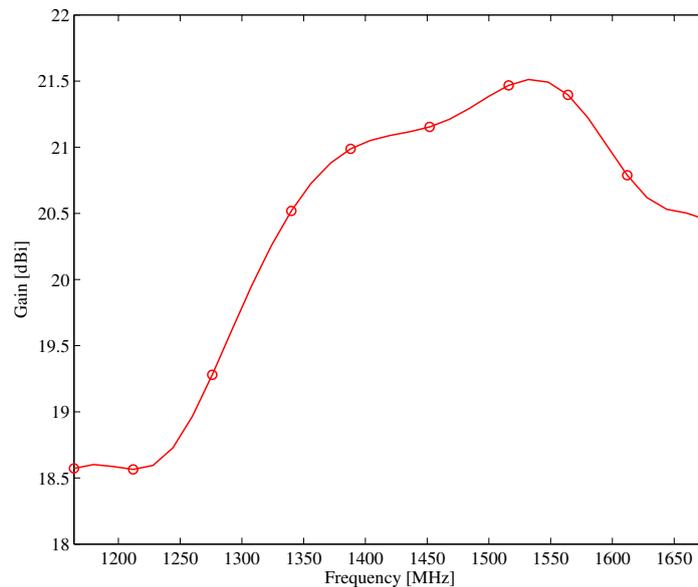


Figure 5.21: The gain on boresight for the offset-parabolic reflector fed by the optimised helical beam antenna. The gain is shown over the frequency range from 1164 MHz to 1548 MHz.

The performance of the offset-parabolic reflector fed by the helical beam antenna is acceptable. The principal plane patterns are highly directive with gains in excess of 20 dBi across most of the band. This means that the antenna gain requirement is met, and exceeded. The sidelobe levels are found to be at least -20 dB down. The F/B ratio is also better than -20 dB. This concludes the design of the offset-fed reflector antenna, and the measured results for the physical antenna is presented in the next section.

5.6 Measurement results of the helical beam antenna fed offset-parabolic reflector

Two helical beam antennas with dimensions as given in table 5.2 were fabricated and measured. Furthermore, the gain of the offset-parabolic reflector was measured. These measurement results are presented and discussed in this section.

5.6.1 The measured radiation patterns of the helical beam antennas

The principal pattern cuts of the helical beam antennas were measured in an anechoic chamber. The measurement setup was only calibrated for measuring the relative field patterns, since it was not necessary to measure the gain of the antennas as only the pattern shape was of interest. The patterns measured at 1292 MHz, 1420 MHz, and 1548 MHz are shown in figures 5.22, 5.23 and 5.24, respectively. In all the figures the measured and simulated patterns are plotted together for ease of comparison.

Figure 5.23 shows that the simulated and measured patterns at the center frequency are in good agreement, except for the slightly more directive patterns measured. The -10 dB beamwidth in this instance was measured as 69° , as opposed to the required 78.7° . The EI at the half-subtended angle $\alpha/2$ of the reflector will now be closer to -14 dB. This has been shown to reduce the maximum gain by a very small fraction (less than 0.4 dB). It is expected that this increase in directivity can be attributed to a lower dielectric constant for the PVC support, since this will result in the PVC-core helix having a higher gain, see figure 5.9. The maximum sidelobe levels have shifted up to about -15.9 dB, in the $\phi = 0^\circ$ -plane, but is now lower in the $\phi = 90^\circ$ -plane. The front-to-back ratio measured is very similar to the simulated result. The cross-polarisation (LHCP) is improved for the measured helix antenna. There are some discrepancies in the lower sidelobes for the simulated and measured results. A possible reason for this could be inaccuracies in the measurement. The anechoic chamber shows decreasing absorptivity below 2 GHz.

The patterns at the lower-passband edge is shown in figure 5.22. Again, it is seen that the measured pattern is more directive, and the -10 dB beamwidth is found to be exactly 78° . The sidelobes are below -15 dB and -20 dB in the $\phi = 0^\circ$ - and 90° -planes, respectively. The measured LHCP levels are also seen to be lower than the simulated results. The patterns at 1548 MHz are shown in figure 5.24. The main beams of the measured and simulated patterns at this frequency are nearly the same, there is only a 2° difference in the -10 dB beamwidth. There is however a much higher sidelobe in the $\phi = 0^\circ$ -plane. Where, the maximum sidelobe measured is close to -13 dB, which is much higher than the simulated -16.4 dB value. The backwards radiation is also seen to have increased by about 10 dB, and is now only -22 dB down.

The axial ratio measurement is shown in figure 5.25. The axial ratio is measured on-axis as a function of frequency. It is shown that the manufactured helical beam antenna achieved a much better axial ratio. The axial ratio is measured as 0.9 dB, 0.8 dB and 1.28 dB at 1292 MHz, 1420 MHz and 1548 MHz, respectively. It is not apparent what caused this improvement, but the simulated air-core helix had a better axial ratio than the PVC-core helix, and it is suspected that the effective permittivity of the actual PVC support is thus lower, and closer to $\epsilon_{\text{eff}} \approx \epsilon_0$. The low axial ratio measured, indicates that the radiated field on boresight is ideally circularly polarised.

Overall, the measured patterns of the helical beam antennas show stability across the operating

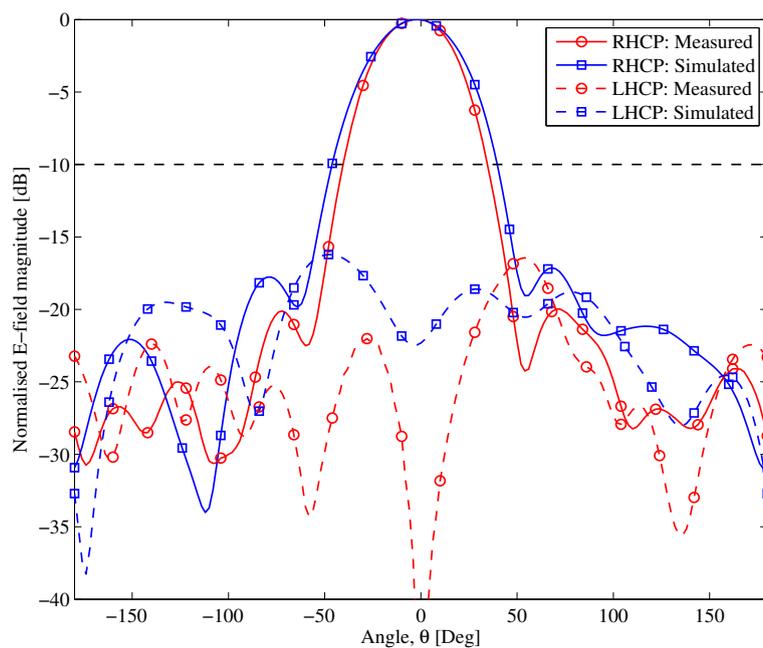
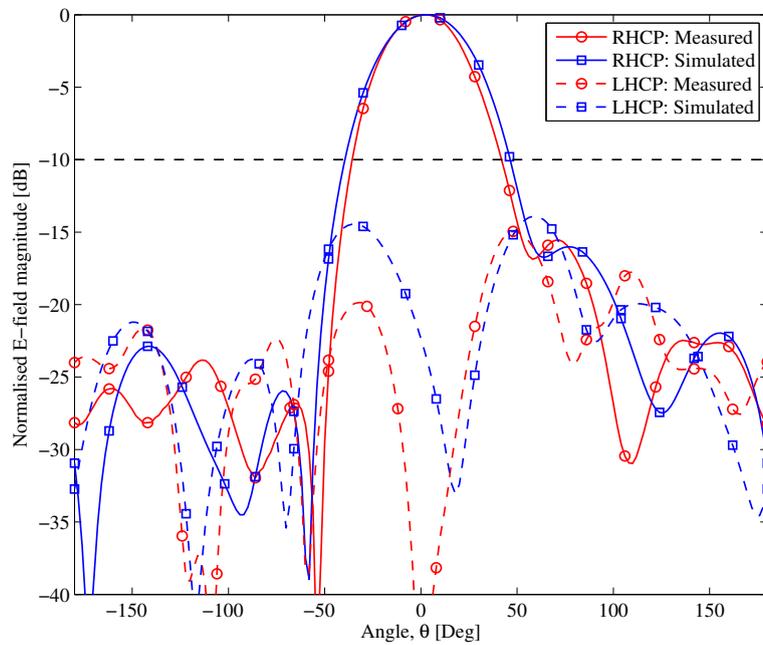
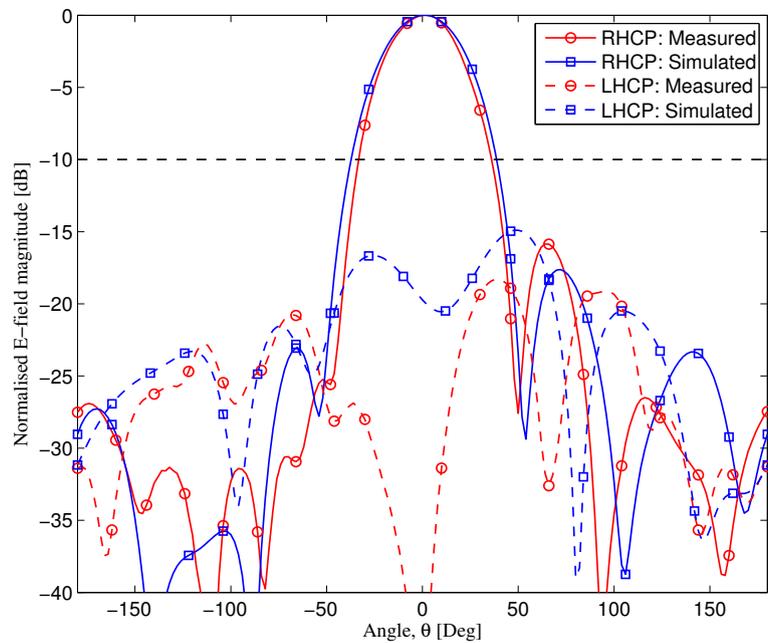
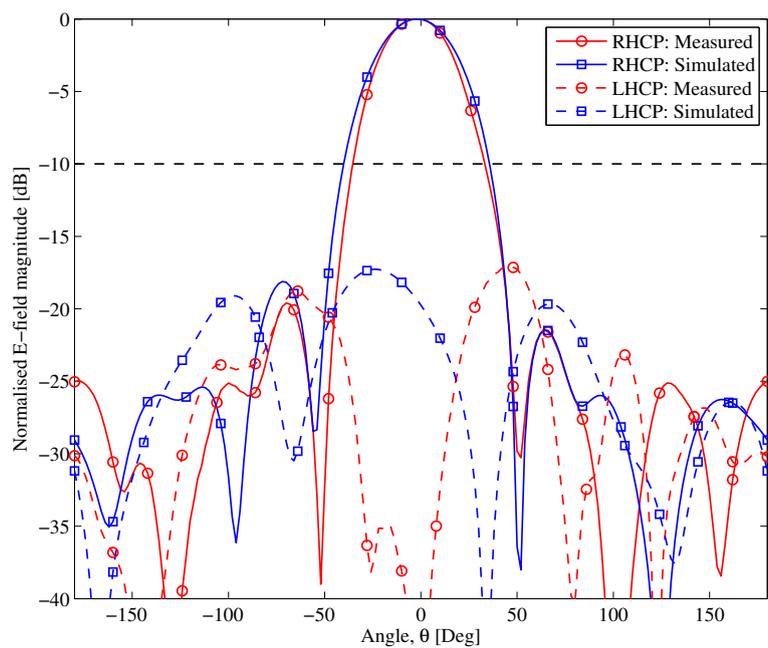


Figure 5.22: The measured and simulated RHCP- and LHCP-patterns in both the (a) $\phi = 0^\circ$ - and (b) 90° -planes, at 1292 MHz.

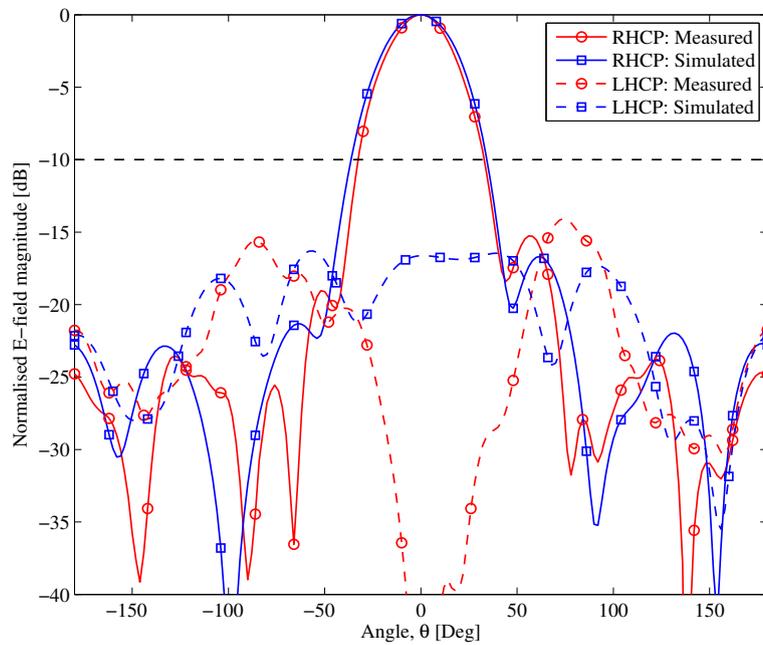


(a)

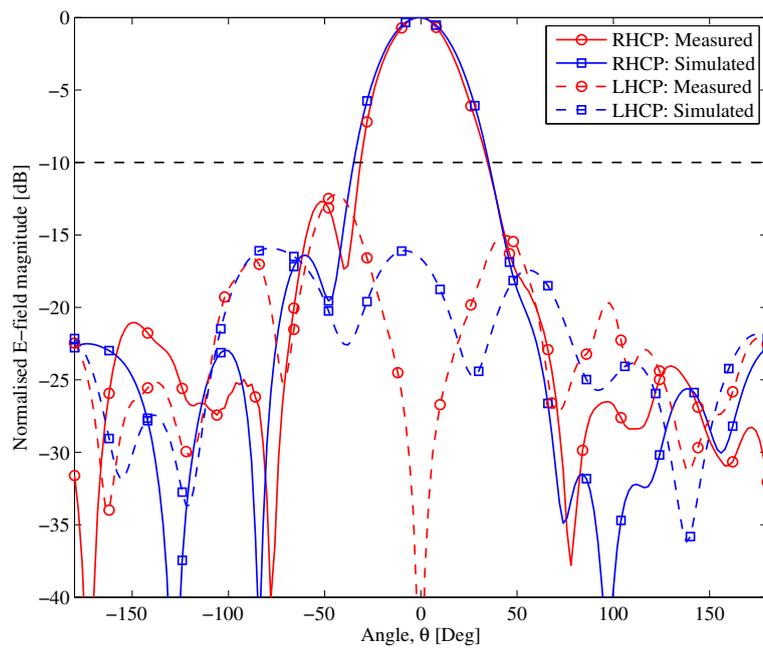


(b)

Figure 5.23: The measured and simulated RHCP- and LHCP-patterns in both the (a) $\phi = 0^\circ$ - and (b) 90° -planes, at 1420 MHz.



(a)



(b)

Figure 5.24: The measured and simulated RHCP- and LHCP-patterns in both the (a) $\phi = 0^\circ$ - and (b) 90° -planes, at 1548 MHz.

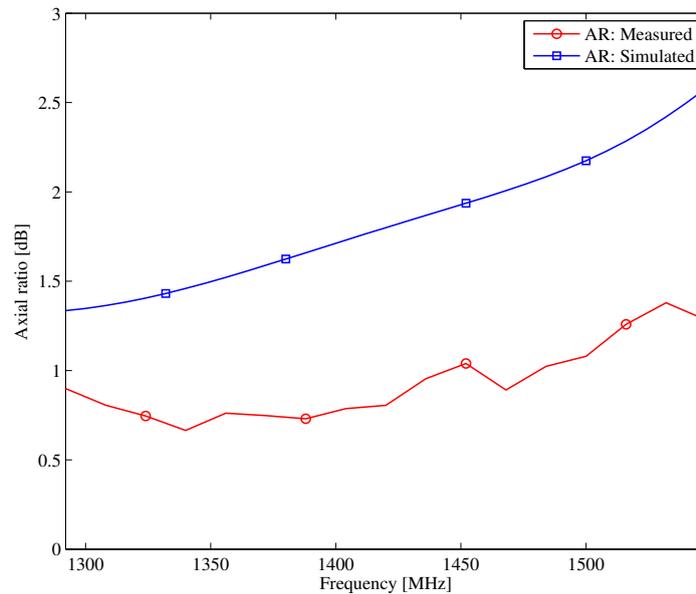


Figure 5.25: The measured and simulated axial ratio of the manufactured helical beam antennas.

band, and the pattern characteristics are in close agreement with the simulated results. A photograph of the built helix antennas are shown in appendix E.

5.6.2 The measured return loss of the helical beam antennas

The return loss of the helical beam antennas were also measured and the results are shown in figure 5.26. Both the antennas are properly matched across the entire passband, where the return loss was recorded as better than -13 dB. More specifically, at 1292 MHz, the return loss for the helices are measured as -15.3 dB and -17.31 dB, at 1420 MHz the return loss is -16.71 dB and -19.37 dB, and at the upper-edge frequency the return loss is -13.44 dB and -15.75 dB. Figure 5.26 shows that the measured return loss response is shifted to a higher frequency. In this regard, the maximum simulated return loss is -28.68 dB at 1372 MHz, and the measured result is -24.06 dB at 1454 MHz. This is a 6% shift in the resonant frequency. This indicates that the triangular matching strip is a bit too short.

5.6.3 The gain measurement of the offset-parabolic reflector

The boresight gain of the offset-parabolic reflector was measured using an over the air experiment. The anechoic chamber was too small for the reflector to be in the far-field. Furthermore, performing a near-field measurement was also not plausible, since a custom mount and an L-band probe would have had to be designed. This was outside the scope of the project. The far-field of the reflector at 1420 MHz is $2D^2/\lambda \approx 8$ m. The source antenna used was an LPDA with a relatively flat gain of about 5.5 dBi across the band. This is a linearly polarised antenna and the total gain was thus obtained by measuring the partial gains in the vertical and horizontal planes, and then combining the results [48]. The partial gains were measured using the Friis transmission formula

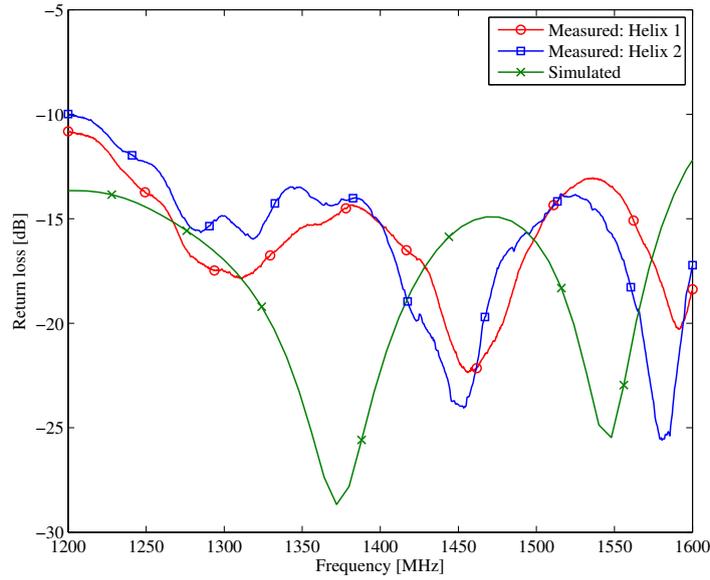


Figure 5.26: The measured and simulated return loss of the manufactured helical beam antenna.

$$P_r = P_t + G_t + G_r - 20 \log R - 20 \log \nu - 32.44 \text{ [dBm]}, \quad (5.19)$$

where P_r is the power received, P_t [dBm] is the power transmitted, G_t is the gain of the transmitting antenna, G_r is the gain of the receiving antenna, R [km] is the distance between the two antennas and ν [MHz] is the transmitting and receiving frequency. The power received P_r was measured on a spectrum analyser, and compensation was made for the losses in the coaxial cables used. The gain of the helical beam antenna fed offset-reflector is shown in figure 5.27. The gain is measured as 17.89 dBi at 1292 MHz, 18.1 dBi at 1420 MHz and 19.8 dBi at 1548 MHz. This is somewhat lower than the simulated gain. This is possibly caused by various measurement errors, such as alignment errors between the boresight direction of the reflector and the source antenna, ground reflections, and the relative feed positioning. It could in part also be due to the more directive feed patterns that were measured. This measurement was extremely difficult to perform and several attempts were made. There were just too many sources of error involved, especially the reflections, off buildings for instance. The effect of the reflections on the gain can be observed in figure 5.27, where it is seen that the gain uncharacteristically will increase or decrease abruptly.

The secondary pattern of the offset-parabolic reflector was measured using the interferometer and the sun as the source, the result at 1420 MHz is shown in figure 5.28. The HPBW is measured as 14.76° . This is in good agreement with the simulated value of 15.7° . This result indicates that the pattern of the manufactured helical beam antenna fed offset-parabolic reflector is satisfactory. The measurement dynamic range was limited to about 8 dB, this is because of the sensitivity of the interferometer.

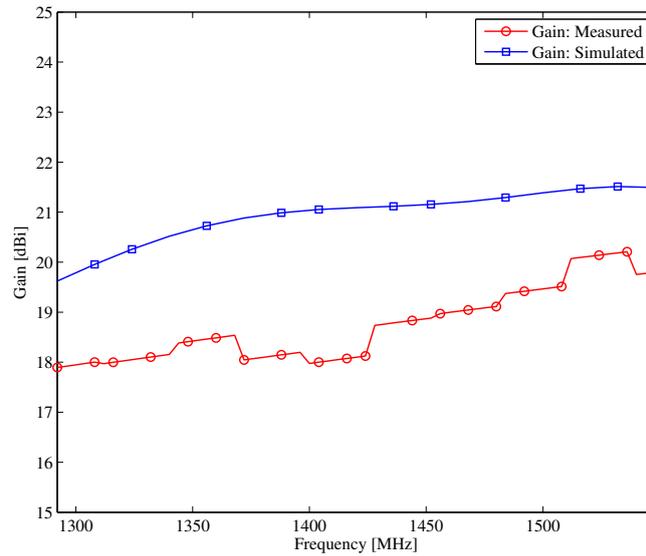


Figure 5.27: The gain of the measured offset-parabolic reflector fed by the helical beam antenna.

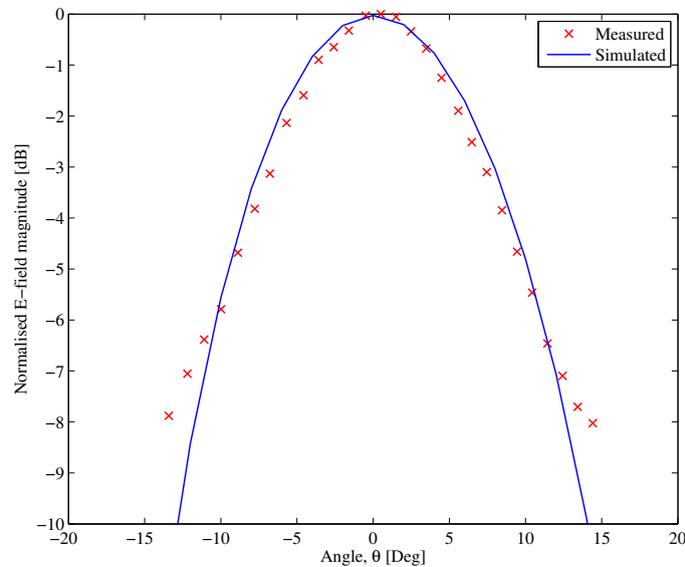


Figure 5.28: The measured radiation pattern of the offset-parabolic reflector fed by the helical beam antenna. This measurement was performed using the interferometer, and the sun as the source with a baseline of $b = 15$ m. The measurement data was integrated for 1 second over a bandwidth of 2 MHz. The measured pattern is compared to the LHCP-pattern in the $\phi = 0^\circ$ plane, at 1420 MHz.

5.7 Conclusion

In conclusion, this chapter presented the analysis and design details of the interferometer antennas. Ultimately it was decided that the most suitable antenna type was an offset-fed parabolic reflector that is fed by a helical beam antenna. The performance of the antenna is summarised as follows: a gain of more than 18 dBi is maintained across the 1292-1548 MHz operating band,

and the antennas are further matched to $50\ \Omega$ over this frequency band. The sidelobe levels are better than $-15\ \text{dB}$ and the front-to-back radiation is below $-20\ \text{dB}$. The feed is RHCP and the radiated fields off the dish are then LHCP.

Chapter 6

Correlator digital back-end

6.1 Introduction

It was in 1963 when Weinreb successfully demonstrated the use of a digital autocorrelation technique in radio astronomy [67]. The instrument built was a first of its kind and its intended use was for the measurement of spectral lines, but more specifically the galactic Deuterium line [67, 15]. The advantage of using digital instrumentation in radio astronomy became apparent. Some of these advantages include; repeatability; stability against temperature and other environmental conditions; and delay lines are more accurately implemented digitally if long baselines and bandwidths are needed [68, 69]. Because of these and the numerous other advantages, modern synthesis arrays nowadays always implement the correlators digitally. Following the modern trends in the radio astronomy community, and given all the benefits of a digital approach, it was decided early on in this thesis to make use of a digital correlator.

The cross correlation of a quasi-monochromatic ($\Delta\nu/\nu \ll 1$) signal can be written as an analytical signal

$$r(\tau_g) = V_R \cos 2\pi\nu_{IF} (\tau_g - \tau_i) + V_I \sin 2\pi\nu_{IF} (\tau_g - \tau_i), \quad (6.1)$$

where V_R and V_I are the real and imaginary parts of the visibility, respectively. It is clear from (6.1) that the single complex visibility function, $V = V_R + jV_I$ is specified over τ_g . It is therefore only necessary to measure $r(\tau_g)$ for $\Delta\tau = 0$ and $\frac{1}{4\nu_{IF}}$. This is exactly what a complex correlator with delay tracking and fringe stopping does. However, the aim of this thesis is to observe strong continuum sources as opposed to quasi-monochromatic signals. Fortunately, the idea of a complex correlator can be extended for continuum observations. The $\frac{1}{4\nu_{IF}}$ delay cycle corresponds to a phase shift of $\pi/2$ at ν_{IF} . If all the frequencies in the passband, $\Delta\nu$ are shifted by $\pi/2$ then a wideband complex correlator is constructed. Mathematically, this $\pi/2$ phase shift is equivalent to a Hilbert transform. The cross correlation function is then expressed as

$$r(\tau_g) = \langle v_{c,1}(t)v_{c,2}(t + \tau_g) \rangle + j\langle \{\mathcal{H}\{v_{c,1}(t)\}\}^* v_{c,2}(t + \tau_g) \rangle, \quad (6.2)$$

where $\mathcal{H}\{\cdot\}$ represents the Hilbert transformation. Implementing (6.2) digitally is probably the most straightforward approach for constructing a complex correlator for continuum observations. Alternatively, spectroscopic correlators are often used for making continuum observations

[15]. This offers the advantages of rejecting narrow-band RFI, and reducing the smearing effect if the passband is divided into narrower sub-bands [15]. In addition, spectroscopic correlators require only a single-multiplying correlator to measure the complex visibility, and the Hilbert transformation filter (which could limit the instrument's performance) is not needed. At this point it is apparent that there are two proposed alternatives for the digital correlator, namely the continuum complex correlator and the spectroscopic correlator. Apart from the simplicity of the complex correlator, the spectroscopic correlator is the preferred solution given the advantages just discussed. Furthermore, this design choice makes for a more flexible system with less project-specific constraints, given that the instrument may be used for spectral-line studies in the future.

A spectroscopic correlator channelises a measurement across the signal frequency band. This is done by measuring the cross-correlation for a large number of delays $2N$, near $\Delta\tau$ [68]. For a digital correlator the delays are specified as integer multiples of the sampling interval, t_s . The $2N$ real samples are then transformed to N discrete complex values at frequency harmonics, $(2Nt_s)^{-1}$, using a Discrete Fourier Transformation (DFT). An FFT is used to perform the DFT, this gives an appreciable computational saving which is critical for real-time processing. The output of the DFT is then the cross PSD according to the Wiener-Khinchin theorem, see Appendix B. After astronomical calibration the complex PSD becomes the complex visibility at N frequency intervals across the signal band. The operation of a digital spectroscopic correlator, can be described mathematically by the discretisation of (B.11)

$$S(k) = \sum_{n=0}^{2N-1} r(n) e^{-j\pi kn/N}, \quad (6.3)$$

where n is an integer from 0 to $2N - 1$ which represents the n -th cross-correlation sample, and k is an integer from 0 to $N - 1$ and represents the k -th frequency channel. The discrete cross-correlation function $r(n)$ is further given as

$$r(n) = \frac{1}{2N} \sum_{i=0}^{2N-1} \hat{v}_{c,1}(i) \hat{v}_{c,2}(i+n), \quad (6.4)$$

and the circumflex $\hat{\cdot}$ denotes a quantised value. Equivalently, the cross-power spectrum may also be expressed as $S(\nu) = V_{c,1}(\nu) V_{c,2}^*(\nu)$, according to the convolution/multiplication Fourier property. For a digital implementation this becomes

$$S(k) = \frac{1}{(2N)^2} \sum_{n=0}^{2N-1} \sum_{m=0}^{2N-1} \hat{v}_{c,1}(n) \hat{v}_{c,2}^*(m) e^{-j\pi k(n-m)/N}. \quad (6.5)$$

Equations (6.3) and (6.5) provide two alternative means for implementing a spectroscopic correlator. Firstly, (6.3) suggests that the cross-correlation is measured before the DFT is applied, whereas (6.5) is based upon measuring the cross PSD directly. In radio astronomy, the traditional method used, was based on measuring the cross-correlation and then performing the frequency analysis [70]. The VLA, Haystack, the Ryle telescope and the first version of ALMA all made use of this method [70, 69]. This type of correlator later became known as an XF

correlator, where the “X” represent the cross correlation and this is followed by an FFT, “F” for the frequency transformation. There was however a demand for processing larger bandwidths, and for finer frequency resolutions. Chikada (1987) proposed and built the first FX correlator. The FX correlator addressed these shortcomings, which are inherent in the XF design, but at the expense of correlator efficiency and complexity [71]. As the name suggests, the FFT, “F” precedes the cross-correlation “X”, and is therefore based on (6.5). In this chapter both methods are investigated, and the advantages as well as the disadvantages of each are pointed out.

What Digital Signal Processing (DSP) hardware platform to use, is another very important consideration. Generally, custom-hardware (ASIC) boards are built to meet the specialised signal processing needs for a radio astronomy instrument. This means developing custom interconnects, communication protocols, control software and using current DSP processor technology [72]. This leads to long development times that are superseded by the rapid growth of DSP computing technology. This ultimately results in these custom-built back-ends to become obsolete fairly quickly. Moreover, this approach is not modular or flexible and complete redesigns are often needed to include new science applications. The Collaboration for Signal Processing and Electronic Research (CASPER) group at the University California, Berkeley, proposes a different approach for developing digital instrumentation for radio astronomy [72]. Their aim is to shorten the time-to-science by relying on modular, connectible and upgradeable hardware components and open-source, platform-independent DSP software libraries [73]. This general-purpose approach is most likely not the most efficient in terms of resources and power consumption, but it certainly shortens the development time. Given the time constraints of this thesis, it is firstly not plausible to design and manufacture a custom DSP processing board. Secondly, this will also not be cost-effective. Moreover, it is more desirable to have a flexible system where the back-end can easily be changed for other science studies, such as spectrometry, pulsar timing, etc.

The digital correlator is implemented on the Reconfigurable Open Architecture for Computing Hardware (ROACH) board. This is a CASPER Field Programmable Gate Array (FPGA)-based processing board. The FPGA gateway for the correlator is developed using the CASPER open-source libraries. These libraries target the Simulink/Xilinx System Generator FPGA programming language. This is a high-level programming language that abstracts the user from the low-level VHDL ¹ code. This has the benefit of ease-of-use, and thus allows the user to rapidly develop FPGA-based DSP applications. The library further provides a huge collection of radio astronomy signal processing functions. The CASPER hardware and software toolflow is further discussed later in the chapter.

The digitisation (quantisation) of the analogue input signals must also be considered. As this process causes some loss in sensitivity, this topic is discussed first in this chapter. This is followed by a comparison between the XF and FX correlators. Then an overview of the CASPER toolflow is given. Lastly, the details of the digital correlator designed is presented.

¹VHDL stands for Very high-speed integrated circuit Hardware Description Language.

6.2 Quantisation noise of spectroscopic correlators

Apart from the numerous advantages of DSP in radio astronomy, there is one associated disadvantage that needs addressing, and that is the degradation in the SNR as a result of signal quantisation. The Analogue-to-Digital Converter (ADC) samples the continuous analogue input signals at discrete time instances, and further quantises the signal to several discrete levels. If the input signals (lowpass) are sampled at least by the Nyquist rate ($t_s = 1/2\Delta\nu$), then there is no loss of information, and the sensitivity remains unchanged. However, the input signals have a stair-case approximation after quantisation, this results in the signals not being bandlimited to $2\Delta\nu$ any more. This is intuitively explained by noting that the rapid change in signal level in the time-domain corresponds to higher order harmonics in the frequency-domain. This results in quantisation noise, which may be viewed as adding an error signal to the true input signal. If the input signal is random noise, as is generally the case in radio astronomy - then the error signal is also noise-like, and the digitiser therefore effectively adds noise to the signal [68]. Sampling at the Nyquist rate is not fast enough any more to fully represent the input signals. This causes a degradation in the SNR [15, 68]. The efficiency factor, η_q , is a figure of merit used to quantify this degradation, and it is specified as a ratio, $SNR_{\text{quantised}}/SNR_{\text{continuous}}$. Researchers have investigated ways of improving the efficiency, and the obvious solutions are to use more bits for the digitisation (finer quantisation intervals) and oversampling [15, 68]. However, this is at the cost of more complex quantisers and electronics. In this section a brief review of the efficiency factor for various quantisations schemes and oversampling rates is presented.

For a Gaussian input signal, very coarse quantisation can be used with surprisingly little loss of information [68]. At the onset of DSP in radio astronomy in the late 1950s, each input sample was represented by only 1 or 2 bits [33, 15]. This was a reflection of the ADC and DSP technology at the time. Even Weinreb's original correlator made use of only 1 bit (2 level) sampling, where only the sign of the signal was retained. Quantisation is a highly nonlinear operation, and the quantised correlation coefficient ρ_q , is therefore non-linearly related to the true correlation coefficient ρ . For the 1 bit case, the well known Van Vleck relationship is used to derive ρ from ρ_q

$$\rho_q = \left(\frac{2}{\pi}\right) \sin^{-1} \rho. \quad (6.6)$$

In the case of observing weak sources ($\rho \ll 1$), the quantised correlation coefficient is then approximated as $\rho_q \approx (2/\pi)\rho = 0.64\rho$, where (6.6) is used. Remarkably, this is only a 36% degradation in the SNR [68]. Increasing the number of bits greatly improves the sensitivity. If one additional bit is used (2 bits), then the efficiency factor is shown to be 0.81 for 3 levels and 0.88 for 4 levels [68, 33, 15]. For modern correlators the norm is 4 bits minimum, and 8 bits are used whenever possible [74]. Using 8 bits is especially important in high RFI environments, as the additional bits are used to prevent saturation [45]. In [15], it is shown that for 4 bits (16 levels) the η_q is 0.996, and it is even closer to 1 for 8 bits. As previously mentioned, oversampling (above the Nyquist rate) can also improve the sensitivity. For example, in a 2-level system, the quantization efficiency is improved from 0.64 to 0.74 by sampling at twice the Nyquist rate

Number of quantisation levels	Efficiency factor, η_q (Nyquist rate)	Efficiency factor, η_q (2xNyquist rate)
2	0.64	0.74
3	0.81	0.89
4	0.88	0.94
16	0.996	-
∞	1	-

Table 6.1: Summary of the efficiency factor, η_q for different quantisation schemes. The schemes considered are 1, 2, 3 and 4bit(s). The improvement by sampling at twice the Nyquist rate is also indicated in the table.

[33, 15]. A summary of the degradation in sensitivity for various quantisation schemes and oversampling rates, is given in table 6.1. It is important to note that the results in table 6.1 assumes a rectangular passband and small correlation coefficients (linear approximation).

6.3 Spectroscopic correlator topologies

In the introduction it was pointed out that there are two main implementations for a spectroscopic correlator, namely the XF and FX designs. In this section the operating principles of both designs are presented. Furthermore the advantages, and disadvantages of each are discussed. In conclusion, a comparison between the two correlators is made.

6.3.1 XF correlator

In the past all the digital correlators used in radio astronomy were of the XF type. A major contributing factor for this was the ability to use a lower number of bits per sample, whereas an FX correlator has to process multi-bit data after the FFT. The architecture of the XF correlator is realised by implementing (6.3), and a simplified schematic is shown in figure 6.1. The delay line is shown to be implemented in a shift register, with each delay being one sampling interval, t_s . The delay tracking block shown in figure 6.1, is also implemented as a multiple of the sampling interval. This can result in a fractional delay error that must be corrected. To compensate for the fractional delay error in an XF system is a complicated task, and it is handled much easier in an FX correlator [15, 35]. In an FX design, this correction is done after the FFT by incrementing the phase values across the band. Additionally, fringe rotation is also normally poorly implemented in an XF correlator, where primitive phase rotation functions consisting of 3 or 4 levels are possible. This makes station-based corrections difficult, and fringe stopping is implemented on a baseline basis [35, 69]. This is as a result of the low number of bits commonly used in an XF design, whereas the FX correlator is a multi-bit system and the fringe rotation function can be represented with a higher resolution, which maintains the required accuracy [15, 35, 69]. The self-multiplier terms in figure 6.1, are used to compute the average signal power (given that at least three levels of quantisation is used) [68]. From

the signal powers and the measured correlation coefficients, the non-linear correction can be applied to obtain the true (unquantised) correlation coefficients [68]. In the FX case, applying the quantisation correction is more problematic. An inverse DFT must be applied to the cross-power spectrum to produce the cross-correlation data. Then the quantisation correction can be applied, and finally a DFT must again be applied to obtain the corrected cross PSD [15, 68, 69].

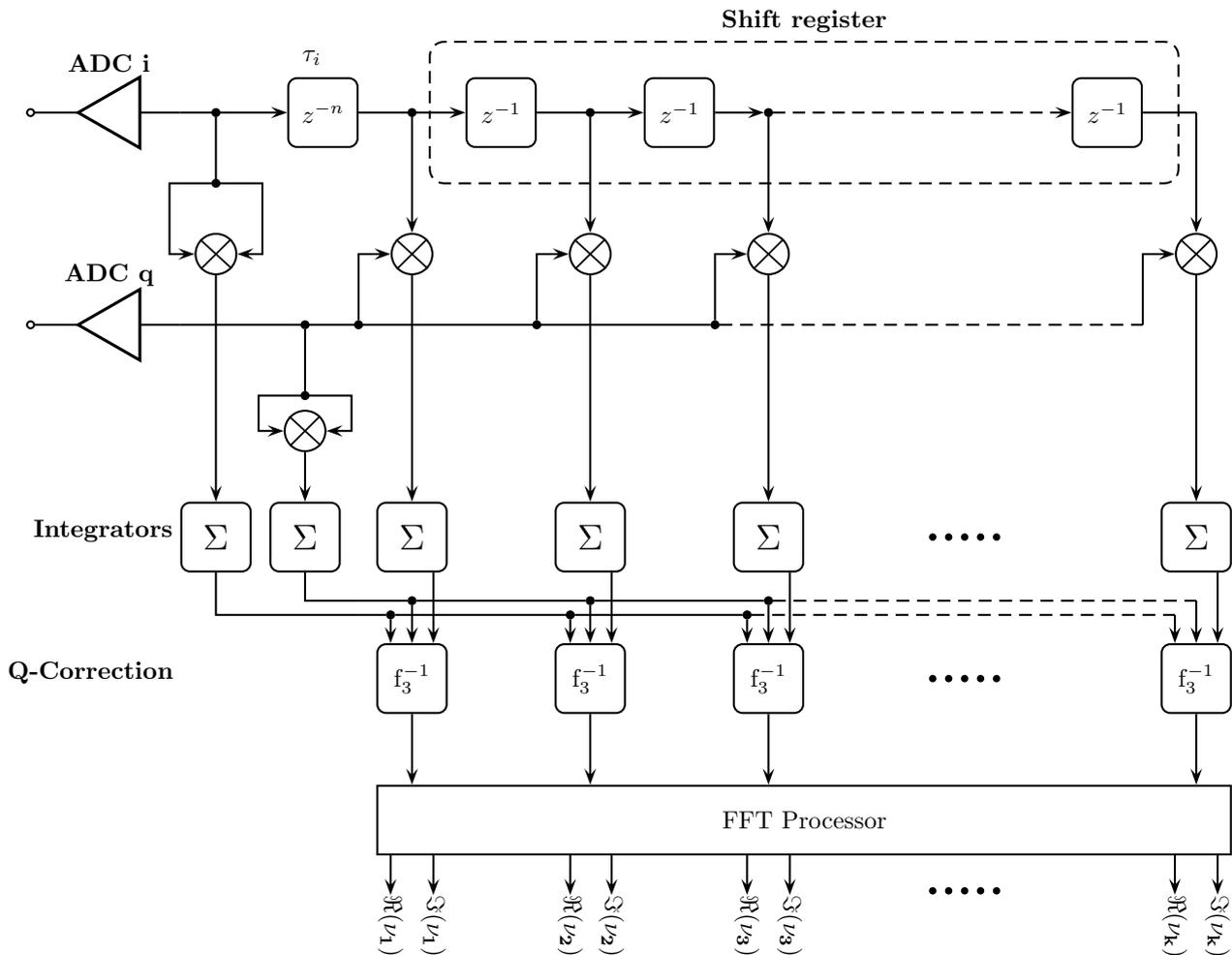


Figure 6.1: A simplified diagram of a lag (XF) digital correlator.

6.3.2 FX correlator

The FX correlator is based on (6.5), and a possible architecture is given in figure 6.2. Traditional FX correlators retain the coarse 2-bit quantisation scheme of the XF correlator [69, 70]. The reason for this is that the hardware interconnection costs are minimised. However, after the DFT the N complex values must be represented with more bits, for precision and because there is now real and imaginary data. This increases the complexity and overall hardware

costs. The FX correlator further suffers from reduced correlation efficiency [71]. Since the FFT processes a truncated, $2N$ contiguous dataset, it becomes apparent that the cross-correlation coefficients in the lag domain, are weighted by a triangular function [35, 68, 15]. This triangular tapering causes the cross-correlation data at larger lags to be poorly represented, and this results in a SNR degradation. It has been shown that the SNR decreases by a factor of 1.22, which is equivalent to an efficiency factor of 0.82 [70]. A possible solution is to overlap data segments and then perform the FFT [68]. However, this is at the expense of a higher data rate and overall implementation cost. A compromise is to use a 50% overlap, but this still nearly doubles the correlator cost for a given bandwidth [70].

The most popular correlator choice in radio astronomy was the XF correlator, because of the high interconnect costs, high bit- and data-rates and degradation in SNR of the FX topology. Probably the most important advantage of the FX correlator is that the number of complex operations per second scale in proportion to the number of antennas used, whereas the number of operations of an XF correlator is predominantly determined by the number of baselines. Additionally, $2N$ multipliers are needed for the N spectral channels. This has a serious implication on modern synthesis arrays with a large numbers of antennas, n_a . To put this into perspective, for the VLA with 27 antennas the number of baselines are 368, for the GMRT with 30 antennas there are 425 baselines, and for ALMA which has about twice as many antennas, this results in 2016 baseline pairs. It is seen that the number of baselines increase according to $\sim n_a^2/2$. The computational cost and power consumption can be reduced by orders of magnitude, by simply using an FX correlator design. This is even more relevant for the SKA which is estimated to use between 2000-3000 antennas [5]. The advances in DSP computing technology, especially the current generation of FPGAs, have made the FX correlator a viable option [71]. To follow is a comparison between the two correlator designs.

6.3.3 XF and FX correlator comparison

Cost is always a big driving factor that influences the final choice in any design. A typical approach followed to quantify a correlator's cost, is based on determining the number of complex operations performed per second [35, 68, 15]. This calculation is a function of the number of antennas used n_a , the number of spectral channels N and the processing bandwidth $\Delta\nu$. In [68], expressions are given for the number of complex operations performed per second for both the XF and FX type correlators

$$c_{XF} = \Delta\nu N n_a (n_a - 1) / 2, \quad (6.7a)$$

$$c_{FX} = \Delta\nu [n_a \log_2 N + n_a (n_a - 1/4)]. \quad (6.7b)$$

The derivation of (6.7b) assumes that an FFT algorithm is used. Furthermore, N complex multipliers and accumulators are used, where each complex multiplication and accumulation (XMAC) operation occurs every $2N$ samples, only. This implementation is similar to the architecture shown in figure 6.2. Equations (6.7), only gives an approximation of the actual number of operations performed, because the implementation details of practical correlators may differ. Using (6.7), the number of operations are computed as $c_{XF} = 2.62 \times 10^{11}$ COPS

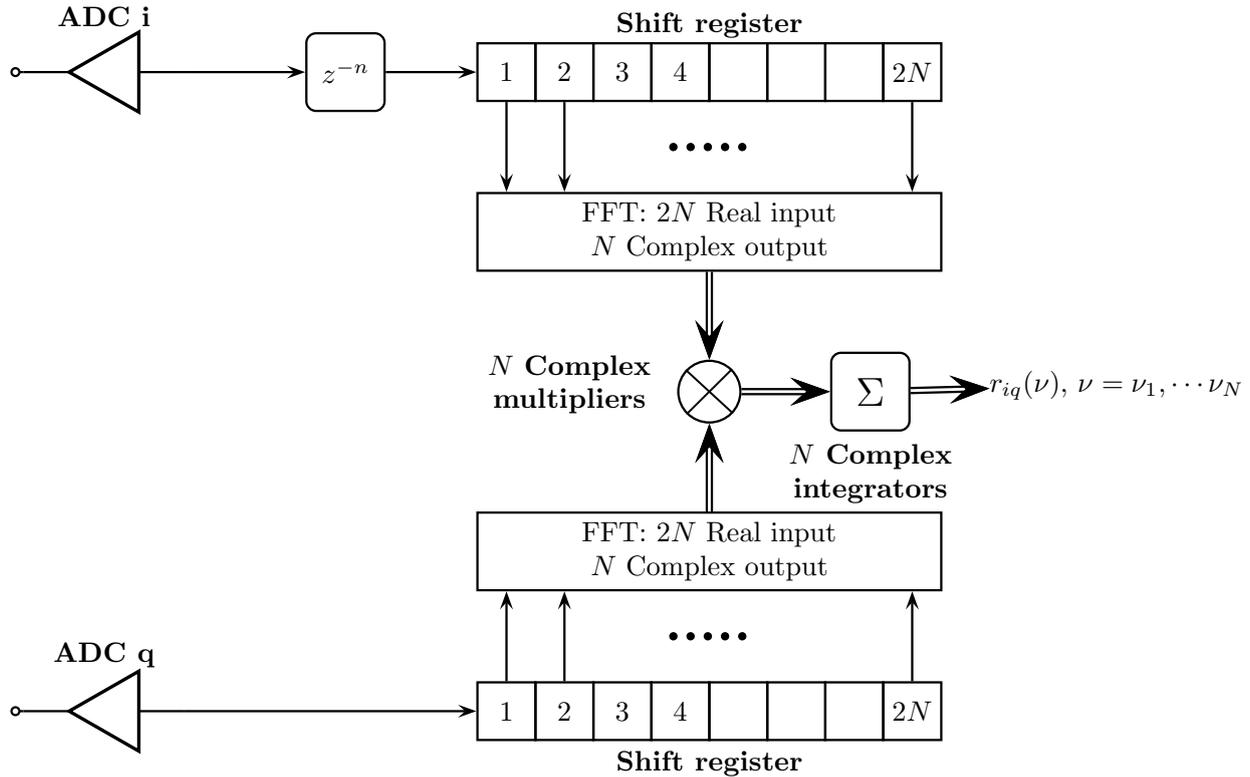


Figure 6.2: A simplified diagram of an FX digital correlator.

and $c_{FX} = 5.24 \times 10^9$ COPS, where $\Delta\nu = 256$ MHz, $n_a = 2$ and $N = 1024$. The XF correlator is seen to be computationally more expensive by a factor of $c_{XF}/c_{FX} \approx 50$. This illustrates the fact that the number of complex operations per second of an FX correlator increases slowly with respect to n_a and N , and this is advantageous.

In the previous sub-section some of the disadvantages of the FX correlator were outlined, such as the reduced sensitivity (triangular tapering), and the increased difficulty of implementing the quantisation correction. It will now be shown that these problems can effectively be solved for an FX correlator. For continuum (broadband) observations, it is evident that the power beyond the first few lags relative to lag zero, is negligible. The cross-correlation data is then near the peak of the triangular weighting function. This suggests that the triangle tapering has little effect on continuum observations and no significant loss in sensitivity is observed [35, 69]. Ideally, the interferometer should not be confined to one specific science application, in this case continuum observations. For instance, the possibility of performing spectral line studies of the hydrogen (HI) line in the future could be another application. The reduced sensitivity at the longer delays, then becomes a problem. It has been mentioned that a commonly employed method of improving the efficiency is to overlap the data segments. This method has the unfortunate drawback of where the data rate increases. A more efficient approach is possible due to the rapid advances in DSP technology. This technique consists of implementing a real-time Polyphase FilterBank (PFB) on a high-performance digital processor, such as an FPGA

[70, 71]. The PFB is used to augment the FFT, and an almost ideal channel response can be achieved, but many of these channels are redundant [71]. The removal of the redundant channels is equivalent to decimating the FFT output. By selecting a suitable windowing function and decimating factor, the data rate can be made to be the same as that of a non-overlapping FFT correlator [71]. The polyphase filterbank topic is further discussed in the section to follow. The need to apply quantisation correction can also be avoided [74]. The criterion for this is to use at least 4 bits, and small correlation values ($\rho \ll 1$), this is generally the case when observing weak sources. This was shown in [74], where the normalised correlation output values after quantisation (using 4 bits) are very close to the input (unquantised) correlation values, over a modest range of correlation coefficients. In this thesis 3-4 bits are used for digitising the astronomical signal, it is therefore not necessary to implement some nonlinear quantisation correction function. Lastly, the FX correlator is a more modular implementation and adapting the system when more antennas are added is easier [15]. This is because an FX correlator is expandable per antenna, rather than per baseline as in the XF correlator case.

After evaluating the advantages and caveats of both correlator designs, it was decided that an FX correlator would be used. The main reason for this decision is that a more modular and flexible instrument can be designed. This makes it easier to adapt the system for different science applications in the future. Most radio observatories are using FX correlator designs nowadays. Thus, the back-end designed in this thesis, mimics the current practices in the radio astronomy community more closely.

6.4 CASPER tools for developing FX correlators

A portion of the research being done by the CASPER group is aimed towards the design of scalable correlator architectures, that are based on modular FPGA hardware, reusable gateway and data packetisation [75]. They have decided to exclusively develop FX correlators, this is to maintain the scalability as the number of antennas grow large. The modular FPGA-based processing hardware and specialised gateway developed by the CASPER group, is ideal for implementing the FX correlator in this thesis.

6.4.1 The ROACH FPGA-based processing board

One of the primary goals of the CASPER group is to develop multipurpose processing boards that use industry-standard interconnection protocols [75]. FPGAs are well suited in this regard, given their flexibility and performance capabilities. For this thesis a ROACH board was used for implementing the FX correlator. The ROACH board is CASPER's new-generation processing board, and this platform is the successor of the Internet Break-Out Board (IBOB) and the Berkeley Emulation Engine 2 (BEE2) platforms [75]. The development of the ROACH board was in collaboration with MeerKAT and NRAO. A simplified block diagram of the ROACH board is given in figure 6.3. The DSP core of the ROACH board is a single Xilinx Virtex-5 XC5V95T FPGA. The Virtex-5 contains 94000 logic cells and 640 multiplier/accumulators and is capable of providing 400Gops/s [75]. Figure 6.3, shows that there are multiple interfaces to

the FPGA, and the four CX-4 and two Z-DOK connectors are worth discussing further. The Z-DOK (40-differential pair) connectors are primarily used for attaching ADCs to the ROACH board, however other interface cards may also be connected [76]. The four CX-4 connectors serve the purpose of connecting multiple ROACH boards together, as point-to-point XAUI links or standard 10Gbps Ethernet (10GbE) links in large systems [76, 75]. This allows for the computational load for a large antenna array, with a wide instantaneous bandwidth to be effectively managed. The frequency channel independence of an FX correlator allows for the spectral decomposition (“F” operation) and the cross-multiplication (“X” operation) to be implemented on separate hardware. The so-called F-engine channelises the antenna bandwidth and the X-engine cross-multiplies all the antenna pairs for as many frequency channels as possible. Commercial Ethernet switches are used for transporting the packetised channel data to the X-engines.

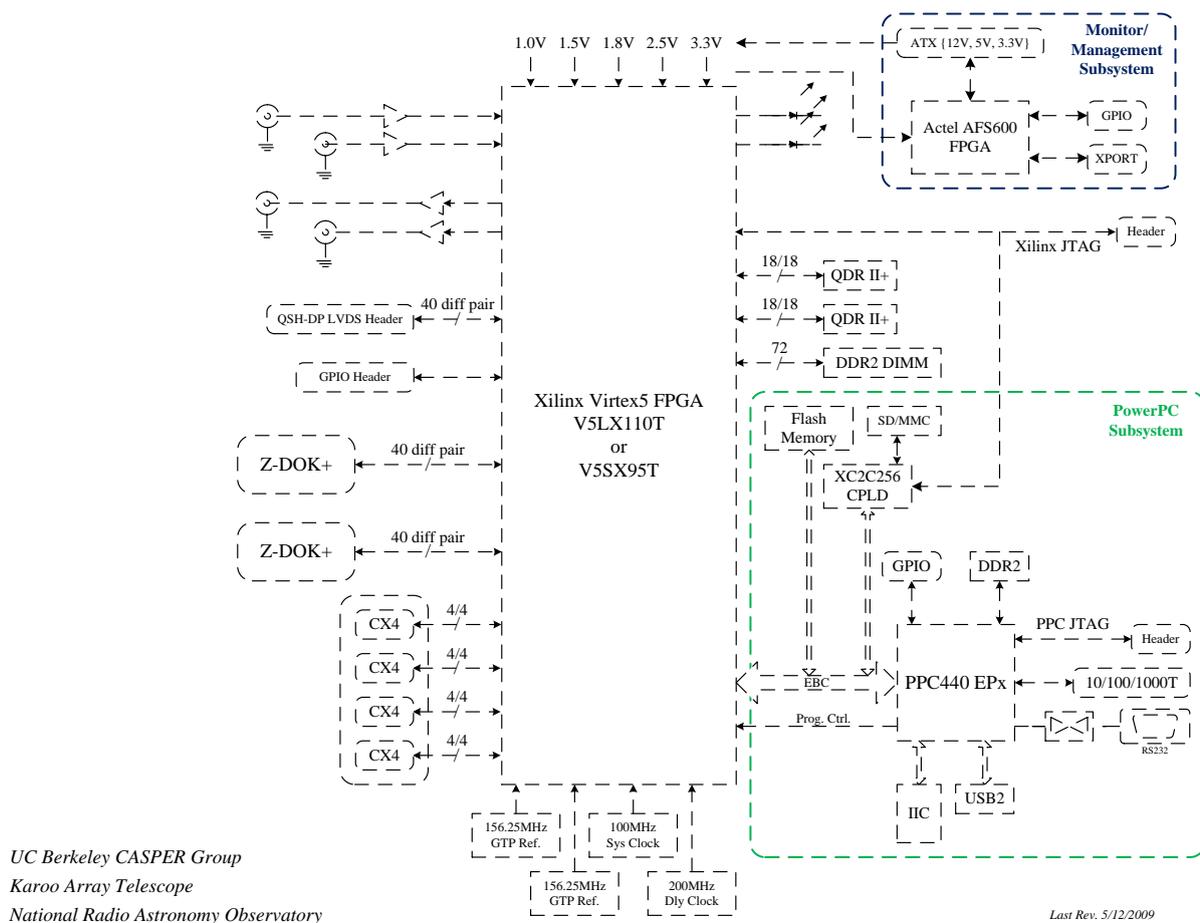


Figure 6.3: A simplified block diagram of the ROACH board, taken from [76].

The instrument developed for this thesis is a small two-element interferometer with a single

polarisation. The processing demands and FPGA resource usage are low enough that both the F- and X-engines can be implemented on a single ROACH board. The CX-4 ports are thus not used. To expand the memory capacity of the ROACH board, there are two fast, but low-capacity quad data rate SRAM interfaces- and one slower, but larger capacity DRAM interface- to the FPGA. The FPGA is further connected to a 440EPx PowerPC, via a 66 MHz bus. The PowerPC provides a simplified interface for controlling and communicating with the FPGA, along with a moderate I/O bandwidth [76].

6.4.2 The KatADC digitiser

The digitiser used is the KatADC, this is a dual 8-bit National Semiconductor ADC08D1520 ADC. This ADC is capable of digitising two channels up to a sample rate of 1.5Gsamples/s. It also has the added feature of interleaving the two converters to effectively create a single 3Gsample/s ADC [46, 76]. The instrument built in this thesis ignores the source's polarisation (Stokes parameters) and this means that only one KatADC is needed. Each channel of the KatADC has some additional circuitry that precedes the ADC08D1520. A simplified schematic of the KatADC is given in figure 6.4. It is shown in figure 6.4, that a Peregrine PE4246 SPST non-reflective RF switch is placed right at the input of the KatADC. The PE4246 covers a wide frequency range from 1 to 5000 MHz, with a typical IL = 1.1 dB, and high isolation of more than 44 dB. This is followed by a Hittite HMC624LP4 6-bit digital variable attenuator and Sirenza SBB-2089Z amplifier in cascade. The HMC624LP4 operates from DC up to 6 GHz, and the attenuation can be adjusted in steps of 0.5 dB up to a maximum of 31.5 dB. The IL = 1.8 dB (DC-3 GHz), typically. The SSB-2089Z amplifier is well matched from 50-850 MHz with a fairly flat gain response of 20 dB, the NF = 2.7 dB. The amplifier and attenuator can be used to control the power level into the ADC. The ADC08D1520 has no provision for single-ended input signals. The CX2156 is a 1:1 balun that converts the single-ended input to a differential signal, and it further provides isolation. The midband IL of the CX2156 is only 0.2 dB. The ADC08D1520 has a $650 \text{ mV}_{\text{pp}}$ differential full-scale voltage range, assuming pin 14 is low [46]. This corresponds to a power level of -2.77 dBm for a continuous wave tone. Then to compute the actual input power into the KatADC, requires that the losses of the preceding circuitry be taken into account. The 100Ω matching resistor in parallel to the ADC08D1520's input immediately causes a power loss of 3 dB, see figure 6.4. The input power into the KatADC is found to be $\sim +3.3 \text{ dBm}$ for a full scale input voltage, this assumes a -20 dB attenuator setting. To prevent any permanent damage, the differential input power must be less than $2.15 \text{ V}_{\text{pp}}$ or equivalent +7.6 dBm. The ADC08D1520 is also driven by a differential clock signal, and a typical sinusoidal voltage of $0.6 \text{ V}_{\text{pp}}$ ($\sim -3.4 \text{ dBm}$) is specified [46]. Again, the CX2156 balun is used for the clock input. The parallel 100Ω resistor causes a 3 dB loss in power, and a 0 dBm input signal is typically used. The KatADC is connected to the ROACH board via a Z-DOK connector.

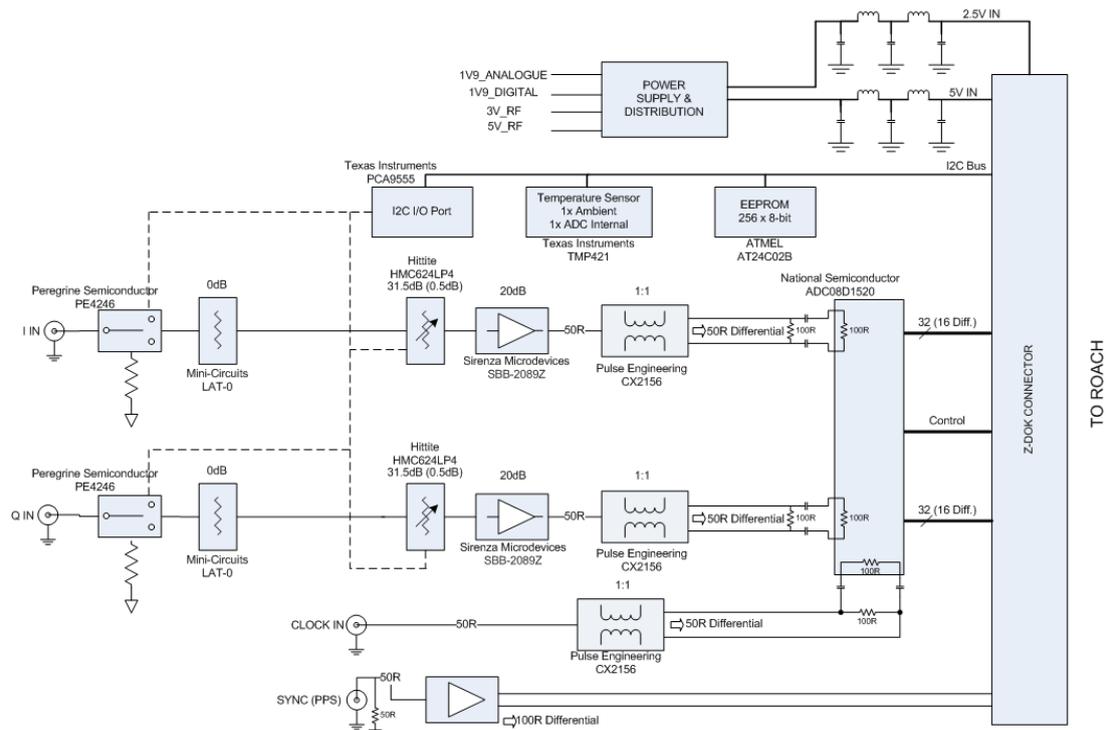


Figure 6.4: A simplified schematic of the KatADC, taken from [76].

6.4.3 Software toolflow

In the introduction it was mentioned that the CASPER group have also developed open-source libraries, for the Simulink/Xilinx System Generator FPGA programming language. The Simulink toolflow firstly provides hardware support for the ROACH, IBOB and BEE2 processing boards, and the various ADCs. Additionally, the CASPER DSP library consists of functions that are specifically tailored for building digital radio astronomy instrumentation. The core library components used for designing the FX correlator are briefly discussed.

The Polyphase FilterBank (PFB) is an efficient implementation of a bank of evenly spaced decimating Finite Impulse Response (FIR) filters, followed by a DFT [75]. The PFB can be used to great effect in radio astronomy signal processing. An inherent problem of the DFT is the leakage of the signal power into adjacent frequency channels. The consequence of this is that strong RFI signals can drown out any faint astronomical signals in the nearby bands [77]. It is shown in figure 6.5 that the PFB produces a flat channel response which closely resembles a rectangular function. This gives excellent out of band suppression, with satisfactorily low sidelobe levels [77]. The use of a PFB is especially beneficial for this thesis, given the RFI environment in which the interferometer is expected to operate in. The advantage of the PFB comes at the expense of additional sample buffering and complex multipliers (dependent on the number of taps used) [75].

The CASPER FFT core is implemented as a radix-2 biphase FFT [75]. To handle higher bandwidths the core is extended to combine multiple biphase cores with additional butterfly cores.

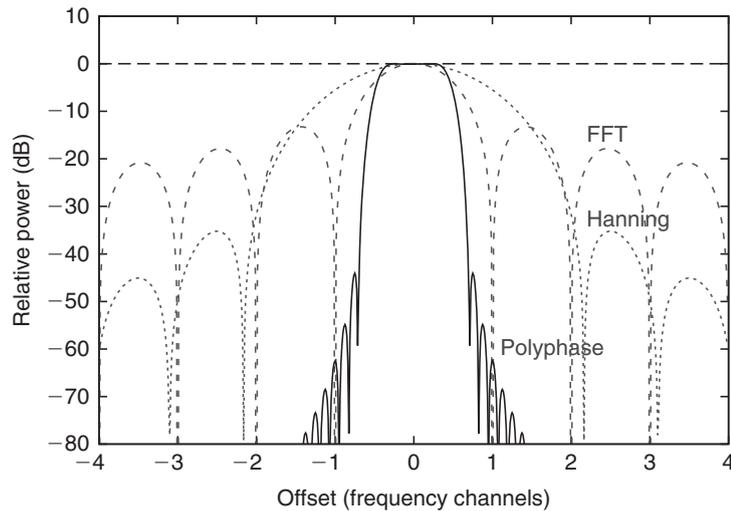


Figure 6.5: Comparison of a single channel frequency response of a PFB with a direct FFT, taken from C. Harris and K. Haines, “A mathematical review of polyphase filterbank implementations for radio astronomy,” *Publications of the Astronomical Society of Australia*, 2011, 28, 317-322.

This creates an FFT that can handle 2^P samples in parallel, where P is the number of simultaneous inputs. This core is capable of outputting valid data on every clock cycle and is 100% efficient [75]. A real wideband FFT is used in the design of the FX correlator, this is half a biplex FFT using Hermitian conjugation. This FFT requires at least 4 ($P = 2$) parallel data streams. The ADC sampling rate must therefore be higher than the FPGA clock rate. This is typically the case for most CASPER instruments, and the designed correlator uses an ADC clock rate of 800 MHz. The FPGA is then clocked at 200 MHz to ensure 4 samples are presented to the FFT every clock cycle.

An extended Linux kernel was developed by the CASPER group to simplify interfacing and controlling the FPGA. This modified version of Linux is called BORPH (Berkeley Operating system for Re-Programmable Hardware). BORPH runs on the PowerPC and maps the shared registers and block RAM spaces, from the user’s FPGA design, into files under the Linux system. This provides a filesystem abstraction for the FPGA resources.

6.5 Design of the 256 MHz, 1024-channel FX correlator

The target specifications for the FX correlator design are listed in table 6.2. At this point both antenna signals have been downconverted to 94-350 MHz, been amplified to a nominal power level of -27 dBm and finally been digitised by the KatADC, at a sample rate of 800MSamples/s. This gives a 400 MHz Nyquist sampled spectrum. It was further decided to channelise the spectrum using 1024 bins. This gives a frequency resolution of 390.625 kHz. This choice was based on a trade-off between sufficient frequency resolution and manageable data rates. The higher the frequency channel count, the less spectral data gets lost, if any channels con-

Specification	Target	Explanation
Instantaneous bandwidth	256 MHz	Section 6.5
Frequency channels	1024	Section 6.5
Onboard integration time	≤ 1 second	Section 6.5
Data transfer rate	32 MB/sec	Section 6.5.2

Table 6.2: The required specifications for the FX correlator design.

taining narrowband RFI gets excluded from further processing. However, the more frequency channels there are, the more data needs to be dumped, and this results in an increased data rate.

The system model of the designed FX correlator showing the signal flow paths of the two digital data streams is shown in figure 6.6. It is depicted in figure 6.6 that each of the digitised input signals have a word length of 8 bits (signed), and the antenna signals are represented by 3 of these bits. According to table 6.1 the efficiency factor $\eta_q \rightarrow 1$. In addition, four parallel time samples are available every FPGA clock cycle, per input channel. Next, the PFB FIR blocks divide the input signals into parallel taps and applies some filtering (FIR). Each PFB FIR is implemented with 8 taps and uses a Hamming window function. The real-sampled wideband FFT then channelises the data streams into 1024 frequency channels. As previously mentioned the PFB FIR and the wideband FFT together constitutes a polyphase filter bank. Each real and imaginary part of an FFT output sample has a bit width of 18, and 17 of these bits are used for precision. At this point, it is common practice in larger designs to equalise the FFT data. The purpose of this is to prepare the FFT output data for re-quantisation, where in many cases the resolution is reduced from 18 to 4 bits. The purpose of this equalisation process is to shift the data bits up to sit within the 4 most significant bits. This is needed to handle the high data rates in larger synthesis arrays. But, for a single two-element interferometer, this is not needed, and the full precision of the FFT output is thus maintained.

After the FFT, the even channels for both input streams are grouped together. Similarly, the odd channels are concurrently processed. The X-engines shown in figure 6.6, performs the XMAC operation on the channelised data. Firstly, the complex auto- and cross-terms are multiplied, point-by-point. This computes both the PSD of each input signal, as well as the cross PSD between the two signals. Both the real and imaginary parts of the data at this stage have a bit width of 37, with the binary point placed at bit 34. A set number of spectra are then accumulated, with each accumulator register containing 512 addresses. It is important to point out here that the bit growth for a 1 second integration period, at an FPGA clock rate of 200 MHz is ~ 28 bits. The bit width of the accumulators are thus set to 64 bits. If a longer accumulation time period is desired, it has to be done in post-processing. An accumulation sync pulse is sent once the full accumulation time has elapsed, see section 6.5.1. The averaged spectral data is then re-interpreted to 64 bits with the binary position shifted to bit zero (no fractional part). Additionally, this data is then sliced into two 32 bit partitions. The reason for this is that the BRAM buffers at the output are fixed to 32 bits, with a zero data binary point. These BRAM registers are then accessible from the PowerPC, and the spectral data can be grabbed from here, using the control computer. The control computer runs the data-logging

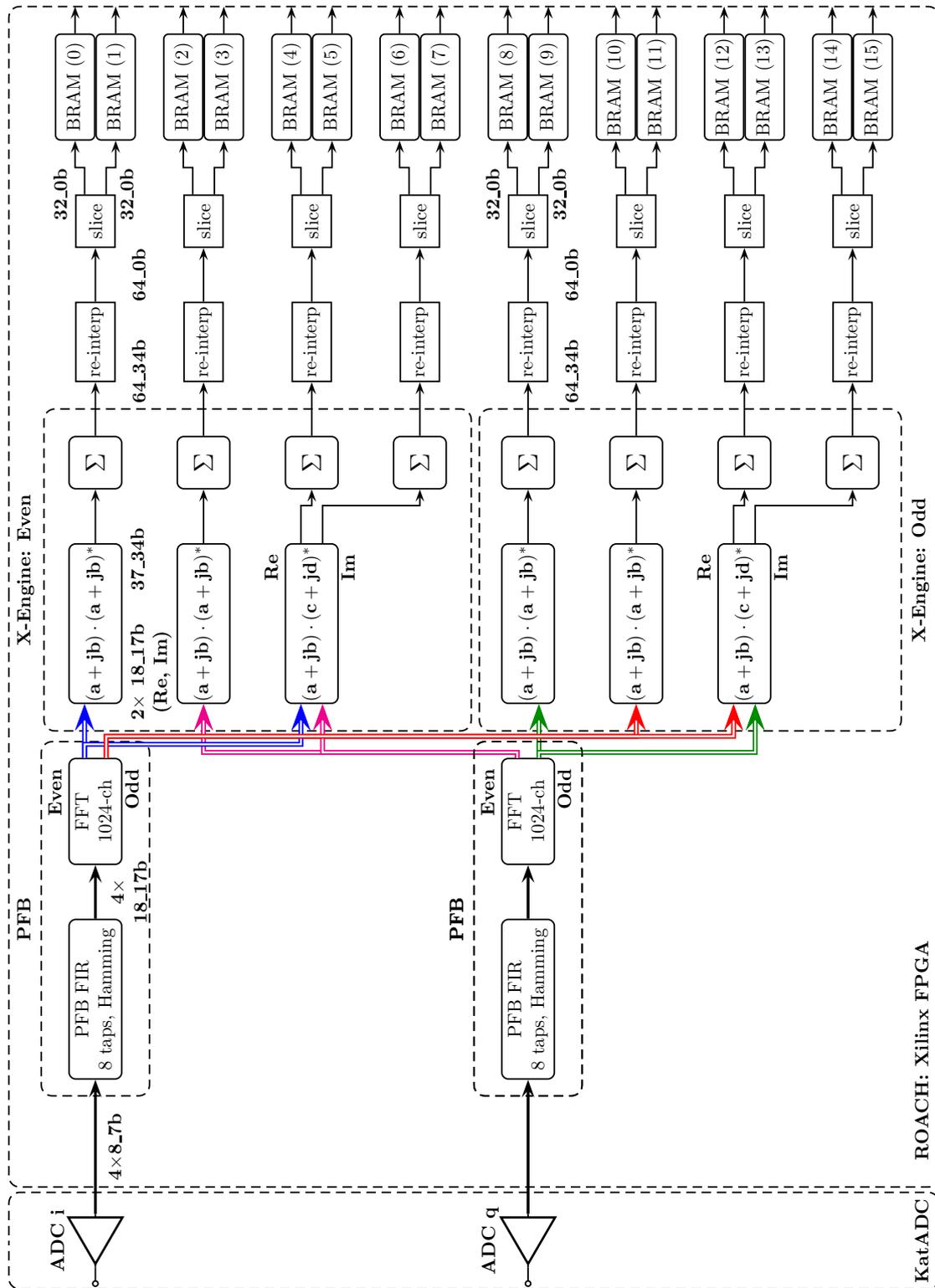


Figure 6.6: The system model of the 256 MHz, 1024-channel FX correlator. The bit width throughout the pipeline is also indicated in the figure.

and supervisory control software that is responsible for acquiring and logging the spectral data. The data-logging and supervisory control software is more fully discussed in section 6.5.2.

This FX correlator design easily fitted on the FPGA fabric. The FPGA resource usage according to FPGA utilisation summary, showed that only 35% of the FPGA look-up tables, 29% of the DSP48Es, 45% of the Block RAM and 39% of the slice registers are used.

6.5.1 FX correlator pipeline synchronisation

Most of the CASPER DSP blocks have a streaming architecture, which outputs valid data after some initialisation period [78]. In this regard, a sync pulse is used in most cases to manage the propagation of the data stream throughout the architecture. Firstly, the sync pulse is used for controlling the timing throughout the data path, and can act as a reset if the sync comes out of alignment. The sync pulse is usually periodic with a pulse width of one clock cycle. The required sync period is dependent on the particular design, and is affected by the processing blocks used. The minimum pulse period required can be determined, using [78]

$$T_{\text{sync}} = N_{\text{acc}} \cdot \text{PFB FIR taps} \cdot \text{LCM}(\text{reorder orders}) \frac{\text{FFT size}}{\text{Simultaneous inputs}}, \quad (6.8)$$

where N_{acc} is the number of spectra accumulated, the number of PFB FIR taps used, the lowest common multiple of the orders of the various reorder blocks used in the FFT, the FFT channel size and the number of simultaneous inputs. For the FX correlator designed, the minimum sync period was determined as 73,726 clock cycles, using (6.8). The sync pulse may also be used to align the initialisation condition of a processing block with the arrival of valid data. This ensures that the block has been reset to some initial state and is ready to accept the new data frame. This usage of the sync pulse is included in the FX correlator design for signalling when a new accumulation period must start. The generation of the accumulation sync pulse is depicted in the flow chart shown in figure 6.7. The first accumulation sync pulse is generated once the sync out of the FFT indicates that a valid data frame will become available the next clock cycle. This synchronises the accumulation with the output of the FFT. The next accumulation sync pulse is sent only once the specified number of accumulations have been reached. The accumulation length is a user definable parameter. It is also possible to reset the accumulator. In this case, the next accumulation sync pulse is sent only once the FFT sync out goes high again.

6.5.2 Data logging and supervisory control software

The data throughput requirements for the FX correlator is low, and this makes it possible to transfer the data to the control computer via the 100MbE port on the ROACH board. An illustration of the I/O paths between the control computer and the FPGA resources are shown in figure 6.8. As previously mentioned, BORPH maps the FPGA resources to entries in the *proc* filesystem. The EBC bus is the interface between the FPGA and the PowerPC. This bus is 16 bits wide at 66 MHz, the theoretical maximum data rate is then 132 MB/sec. However, the bus has a transfer bus efficiency of about 20% due to the transaction overhead between

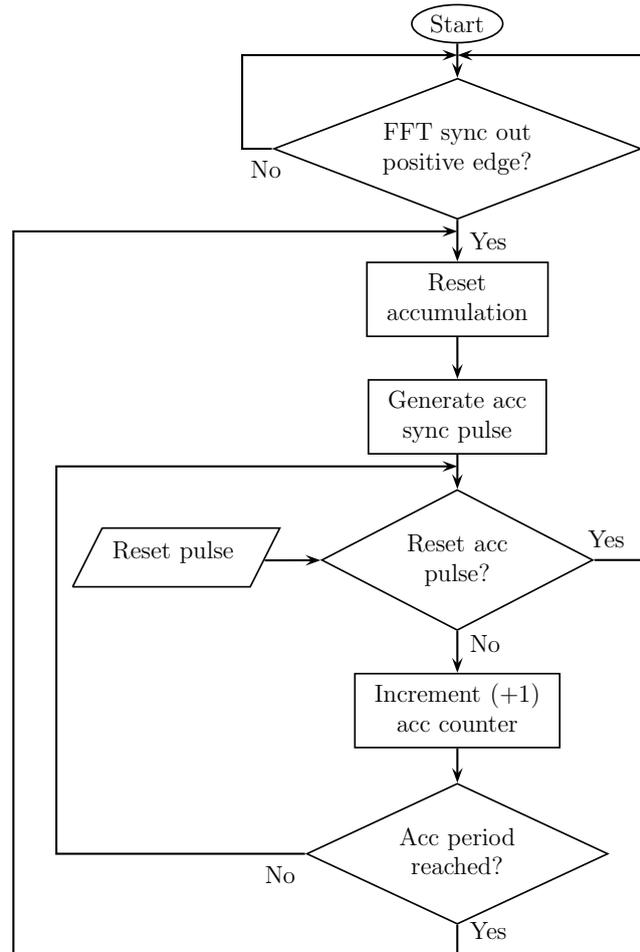


Figure 6.7: The flowchart for the generation of the accumulation sync pulse.

the FPGA and PowerPC. This means that the achievable data rate is closer to 26 MB/sec. Any additional CPU access overhead can cause this data rate to be even lower. The data is then pumped through the 100MbE port to the control computer. The data rate required for the FX correlator is only 32 MB/sec. This assumes a 32 bit wordlength, 1024 channels, the auto-correlation as well as the cross-correlation data for both input signals, and additionally a 1 second integration time. This creates a slight bottleneck between the FPGA and the PowerPC. This said, the data rate is still low enough to validate the usage of the PowerPC for transferring the correlator data. The data logging interval must therefore be slightly longer than a second.

Interfacing to the ROACH board is handled by the KATCP protocol. A standalone KATCP server runs under BORPH, and the user can interact with the FPGA's resources by opening a telnet connection to the ROACH board using port 7147. The FPGA resources are then queried over the network using the KATCP commands. Interacting with the ROACH board is simplified even further by the *corr* python library. This library is a python wrapper for KATCP. It then becomes a simple task to script a control routine for the ROACH board.

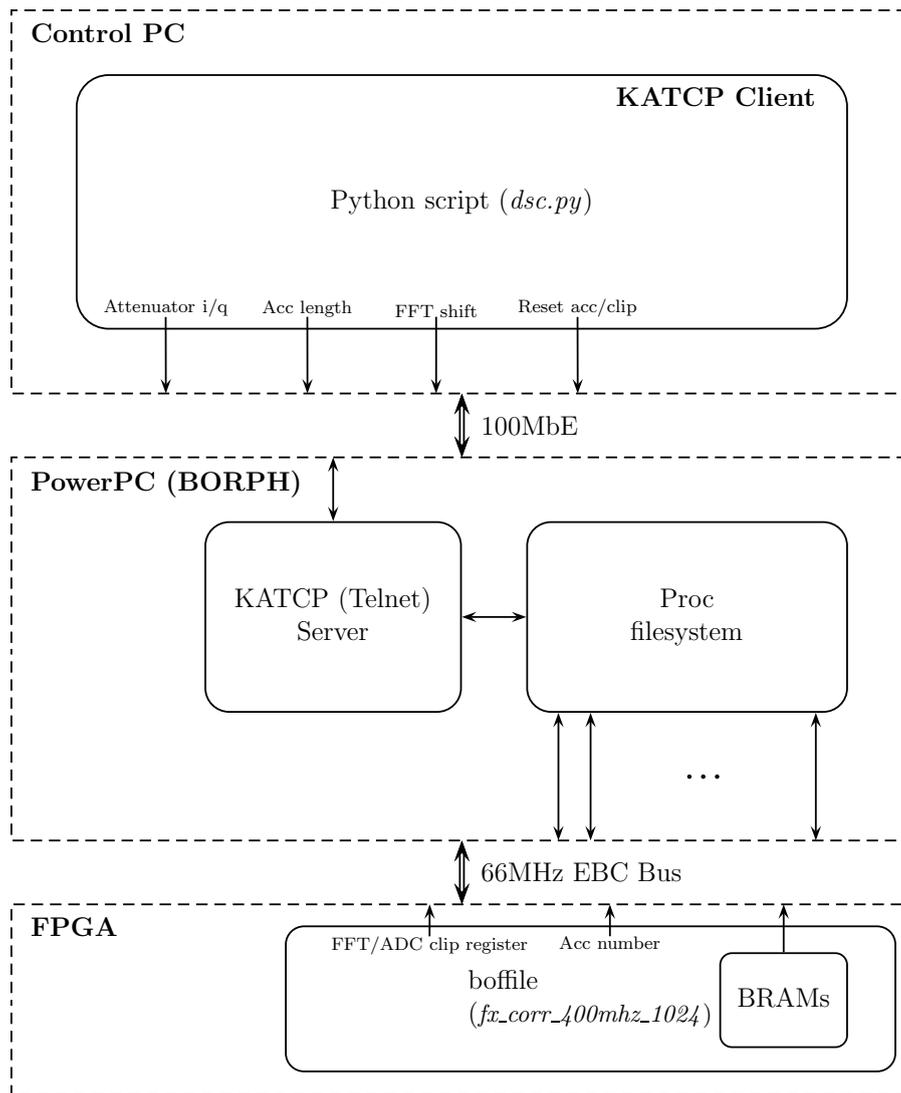


Figure 6.8: The I/O paths and interfaces between the control computer, PowerPC and FPGA.

A python script was written to program the FPGA, to configure the FX correlator, and to automate the data recoding and monitoring process. A flowchart of the data-logging and supervisory control software routine is depicted in figure 6.9. Firstly, an FPGA client object is instantiated, and a connection to the ROACH board is established. After verifying the connection between the client (control computer) and the server (ROACH)- the FPGA gets programmed with the FX correlator boffile. The FX correlator is then configured via the pre-defined control registers. The accumulation length, the KatADC attenuator levels and the FFT shift schedule are some of the configuration parameters. The routine then manages the observation, until the user specified stop time is reached. A user guide for the data-logging and supervisory control software is given in appendix D. Additionally, a status register is implemented into the FX correlator design. This register sends ADC clipping and FFT overflow information to the control computer. This informs the operator and corrective action can be taken.

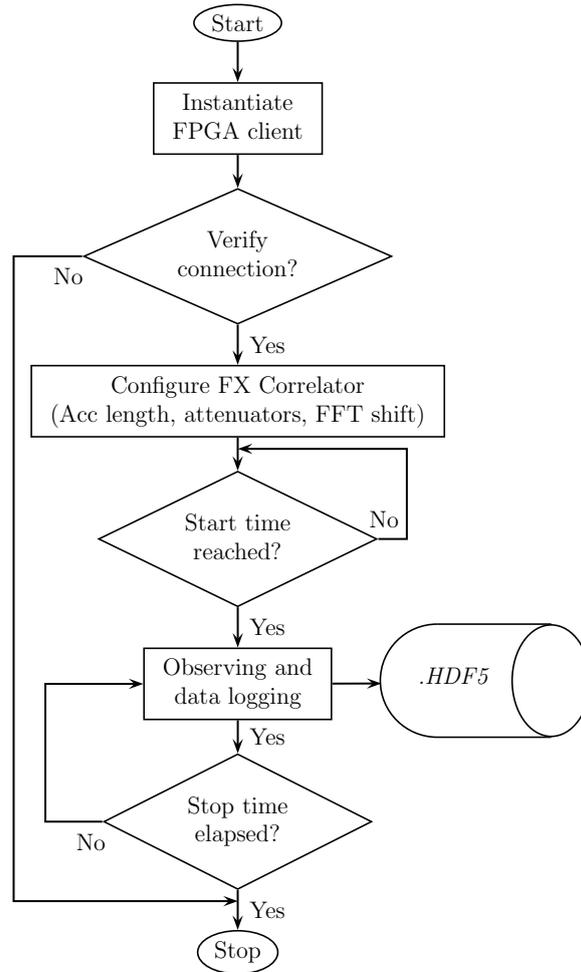
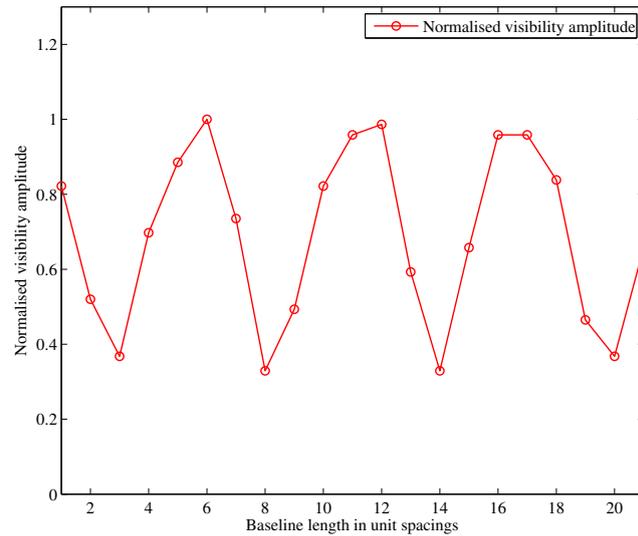


Figure 6.9: The flowchart for the data-logging and supervisory control software. This software routine runs on the control computer and is responsible for automating an astronomical observation, and to control- and communicate with- the ROACH board.

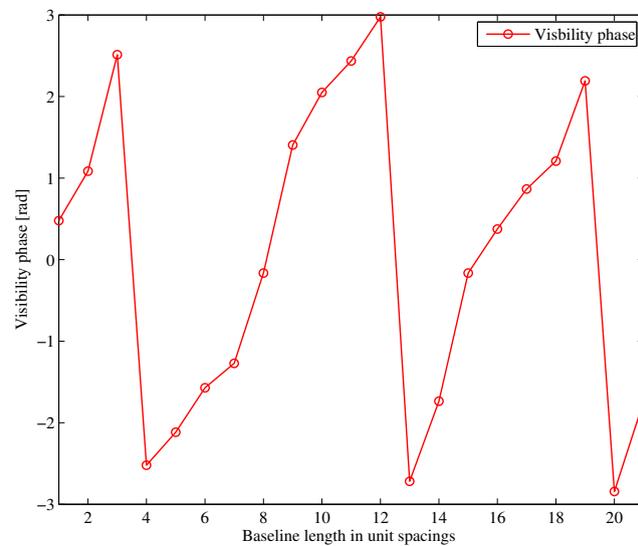
6.6 FX correlator testing and validation

The one-dimensional source synthesis example given in [15], was used as a test for verifying the correctness of the FX correlator, in both the software- and FPGA implemented- models. The source consists of two point source components at -0.24° and 0.1° with respect to the phase reference position θ_0 . The respective flux densities are in a ratio of 2:1. The visibility measurements are made with a pair of antennas positioned along an E-W baseline. According to section 3.5, the projected baseline is then $u = \frac{b \cos \theta_0}{\lambda}$. To avoid aliasing a baseline spacing of less than $\Delta u < 84.2\lambda$ is needed. The baseline spacing is chosen as $\Delta u = 30\lambda$, this is the same as in [15]. The number of measurements made are 23. The input signals are assumed to be monochromatic. A phase difference between the two input signals were induced to simulate a change in the baseline length. To follow are the results for the software (Simulink) FX correlator design.

The output of the Simulink model is the cross-power spectrum of the two input signals. This



(a)



(b)

Figure 6.10: The (a) amplitude and (b) phase of the complex visibility as measured by the software (Simulink) FX correlator model, plotted as a function of baseline length which has been normalised by the unit baseline spacing, $\Delta u = 30\lambda$.

is essentially the channelised complex visibility of the source. For a continuum observation all the frequency channels can simply be added together. This is however not necessary since a monochromatic input signal was considered. The computed visibility amplitude and phase are shown in figures 6.10(a) and 6.10(b), respectively.

These complex visibility values are samples of the Fourier transform of the source distribution.

Recalling that the inverse Fourier transform of the complex visibility yields the synthesised image of the source. The source obtained after performing the IFFT is shown in figure 6.11. The two point sources are clearly visible and are situated at the correct angular positions. The intensity is normalised, where it is seen that the point source at -0.24° is twice as large as the source at 0.1° . The sidelobes seen in figure 6.11 is as a result of the finite extent of the visibility measurements. These measurements are uniformly weighted, resulting in a rectangular windowing function. Upon transformation the channel response is that of a sinc function.

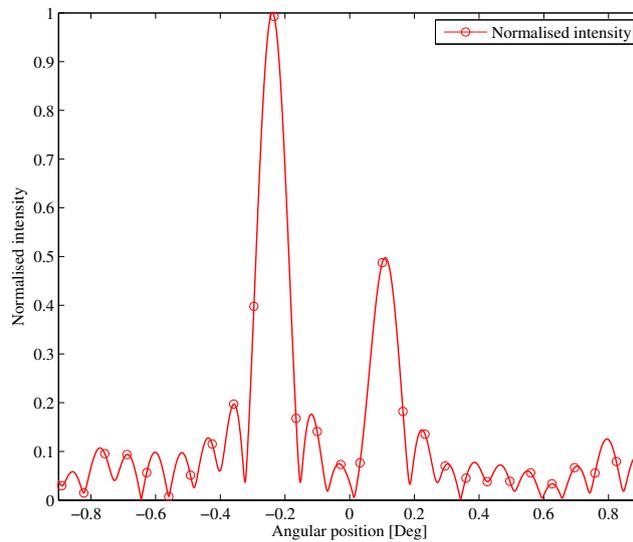
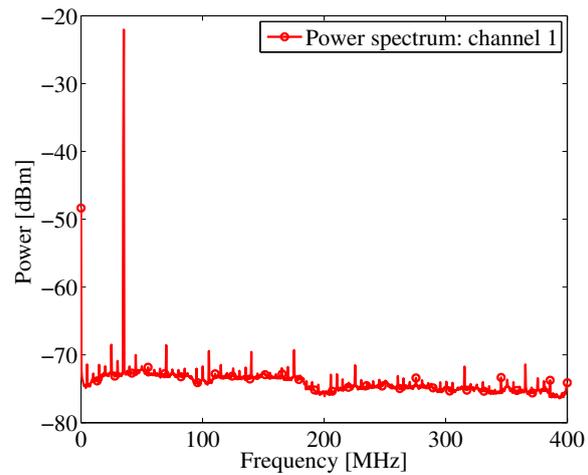
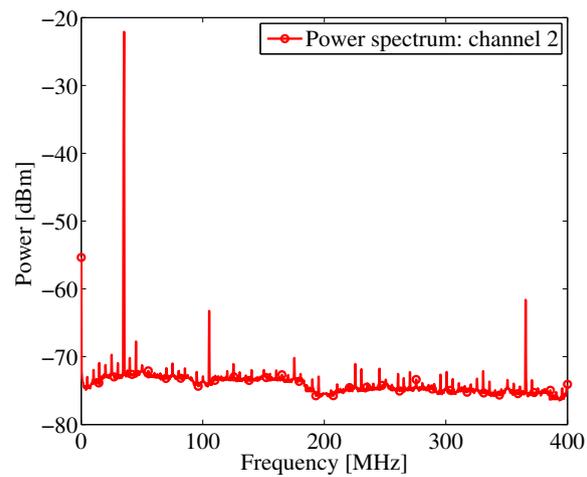


Figure 6.11: The synthesised image of the 1-D source, using the complex visibility values computed with the software (Simulink) FX correlator model. The inverse Fourier transform was zero-padded to 512 samples.

A similar approach was followed to validate the FX correlator implemented on the ROACH board. Unfortunately, due to the measurement equipment limitations it was possible to only simulate a single point source. It was decided that the source is positioned at -0.24° from the phase reference direction θ_0 . The test procedure consisted of generating two signals that are phase-locked, and then adjusting the relative phase between these two signals in order to simulate different baseline lengths. The AM300 arbitrary waveform generator was used for this experiment, since this instrument is capable of dialling in a phase delay between its two output channels. The test signals were generated at 35 MHz with a nominal power level of -20 dBm. Again a unit baseline spacing of 30λ was assumed, and the resulting phase delay is 45.24° . The experiment further consisted of 23 different measurements. The power spectrum measured for both input channels for one accumulation period is shown in figure 6.12. The input tone for both channels are seen to be at 35 MHz. The harmonics of the input signal can also be seen in figure 6.12, but they are at least -40 dBc. These harmonics are most probably generated by the ADC. Moreover, the input signals have not been filtered.

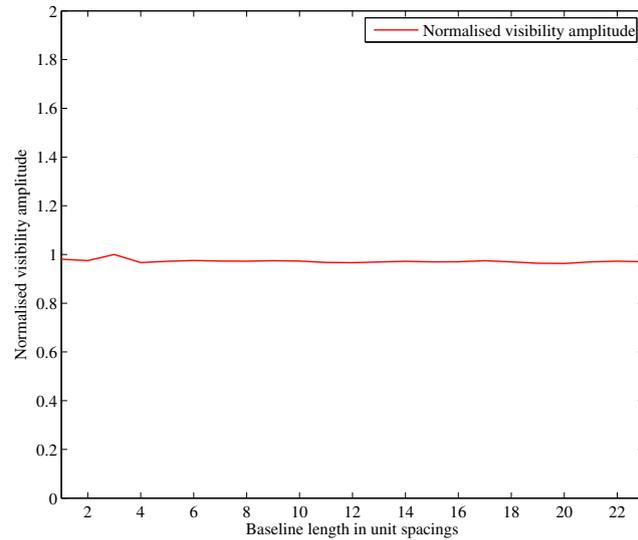


(a)

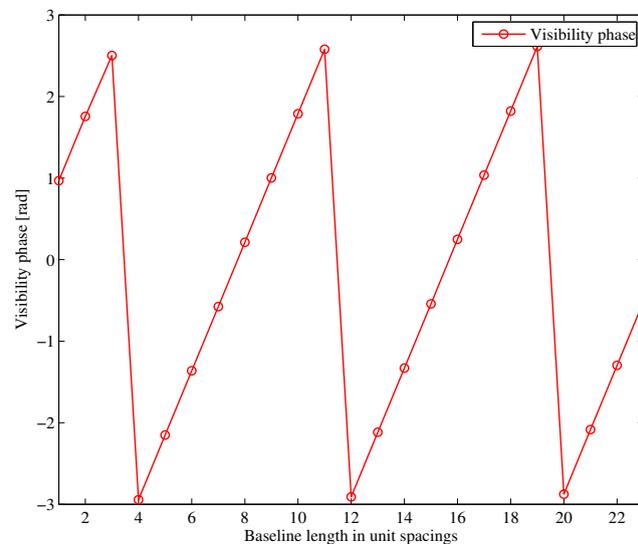


(b)

Figure 6.12: The power spectrum measured for both input channels of the FX correlator from 0 to 400 MHz (Nyquist frequency). Both input signals are generated at 35 MHz with a nominal input power of -20 dBm.



(a)



(b)

Figure 6.13: The (a) amplitude and (b) phase of complex visibility as measured by the FX Correlator (FPGA), plotted as a function of baseline length which has been normalised by the unit baseline spacing, $\Delta u = 30\lambda$.

The amplitude and phase of the measured visibility is shown in figure 6.13. Taking the inverse Fourier transform of the visibility results in the synthesised source in figure 6.14. The point source is shown to be positioned at -0.24° from the phase reference position. This is the expected result. The finite measurement extent again results in extraneous sidelobes. The sidelobe levels are measured to be -13.1 dB, which is characteristic of a sinc response. The results presented here verifies that the FX correlator is successfully implemented.

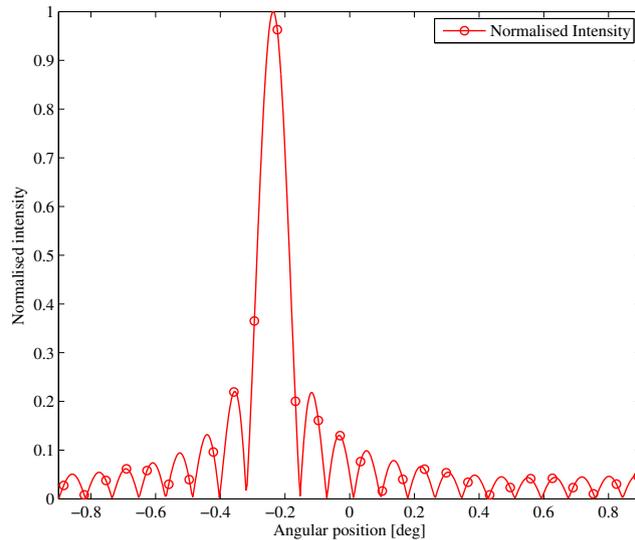


Figure 6.14: The synthesised image of the source using the complex visibility values computed by the FX correlator (FPGA). The inverse Fourier transform was zero-padded to 512 samples.

6.7 Conclusion

It was pointed out early on in this chapter that a spectroscopic correlator is the best choice in terms of RFI rejection. In this regard, the XF and FX correlators were compared. It was found that the advantages offered by the FX scheme, such as its modularity and computational advantages as the number of antennas increase, made it an ideal solution for future expansion of the interferometer. Thus, it was decided to implement an FX correlator. The topic of SNR degradation as a result of signal quantisation was also discussed. The efficiency factor η_q quantifies this degradation. It was decided that 3 bits will be used for quantising the sky noise temperature (astronomical signal), and it is seen that the efficiency factor tends to 1 (no sensitivity loss). The CASPER approach towards developing digital instrumentation for radio astronomy was discussed. Their aim is to use modular general-purpose DSP hardware coupled with open-source libraries, in order to rapidly develop instruments and reduce the time-to-science. This approach was very appealing and their software and hardware framework was used to implement the FX correlator. Ultimately, a 256 MHz, 1024-channel FX correlator was implemented on a ROACH board. The results presented in section 6.6 experimentally validates the design.

Chapter 7

Interferometric observations

7.1 Introduction

According to the function block diagram of the interferometer as presented in figure 1.1 - the three main subsystems that had to be designed are the antennas, the radio front-end and the digital back-end. The design details of the dual-conversion superheterodyne receiver, the helical beam antenna fed offset-parabolic reflectors and the digital FX correlator have been presented and discussed in chapters 4, 5 and 6, respectively. Presented below in table 7.1, is a comparative summary of the intended and the measured specifications for these sub-systems. It is seen that the measured results are in compliance with the target specifications.

The various sub-systems were integrated - forming the two-element digital correlation interferometer. The expected instrument performance specifications are listed in table 7.2. This chapter presents and discusses the observational results obtained with this instrument.

7.2 A more precise RFI characterisation of the 1292-1548 MHz band

The seemingly radio quiet band of 1300-1500 MHz, shown in figure 2.5 was found to be cluttered with RFI. Typically, the noise figure of a spectrum analyser is more than 20 dB, and as a result it has poor sensitivity. This makes it difficult to detect weaker signals that might be present. The low-noise front-end of the interferometer has a much better sensitivity, ~ 173.5 dBm/Hz. The signals that were initially buried below the noise floor of the spectrum analyser could now be detected. The spectrum measured with only the dish, the low-noise front-end and a spectrum analyser is shown in figure 7.1. From this measurement it is clear that there is a significant amount of activity within the passband of the interferometer. A review of [29], indicated that certain portions of the 1292-1548 MHz band are allocated to other applications. From figure 7.1(a), the interfering signals are seen to be both narrowband and broadband, and a summary of the stronger RFI signals are given below:

- -77.83 dBm at 1297 MHz,

Specification	Target	Measured	Explanation
Helical beam antenna fed offset-parabolic reflector			
Impedance bandwidth (VSWR better than 2:1)	18%	> 28.1%	Section 5.6.2
Gain across the 1292-1548 MHz band	> 19 dBi	~ 18 dBi	Section 5.6.3
-10 dB BW (helical feed) at 1420 MHz	78.7°	69°	Section 5.6.1
Axial ratio across the 1292-1548 MHz band	–	< 1.3 dB	Section 5.6.1
HPBW at 1420 MHz	–	~ 14.7°	Section 5.6.3
Maximum sidelobe levels across the 1292-1548 MHz band*	–	< -19 dB	Section 5.5.3
F/B ratio across the 1292-1548 MHz band*	–	< -20 dB	Section 5.5.3
Dual-conversion superheterodyne receiver			
Receiver gain	67 dB	58.2 dB	Section 4.4.7.1
Noise figure	1.1 dB	0.92 dB	Section 4.4.7.2
Gain roll-off, edge-to-edge	< 3.5 dB	~ 5.6 dB	Section 4.4.7.1
Gain ripple	< 2.9 dB	< 1.6 dB	Section 4.4.7.1
Output IP3	+11.3 dBm	+19 dBm	Section 4.4.7.3
input 1 dB compression point	–	-45.4 dBm	Section 4.4.7.3
Linear dynamic range	–	+48.5 dB	Section 4.4.7.3
Spurious free dynamic range	–	+34.1 dB	Section 4.4.7.3
FX Correlator			
Nyquist bandwidth	256 MHz	400 MHz	Section 6.5
Frequency resolution	390.625 kHz	390.625 kHz	Section 6.5
On-board integration time	≤ 1 sec	~ 1 sec	Section 6.5
Data transfer rate*	32 MB/sec	~ 20 MB/sec	Section 6.5.2

Table 7.1: A summary of the interferometer's system performance specifications. The specifications indicated with an * are the expected values, but have not been verified experimentally.

Instrument parameter	Value
Operating frequency range	1292-1548 MHz
Effective collecting area, A_{eff}	0.281 m ²
System temperature, T_{sys}	~ 130 K
Total instantaneous bandwidth	256 MHz
Sensitivity, ($\tau_a = 10$ s)	17.84 Jy
Polarisation	Circular (single)
Maximum baseline	24 m
Maximum fringe frequency	5.9 mHz
Bandwidth pattern null	$\sim 2.8^\circ$

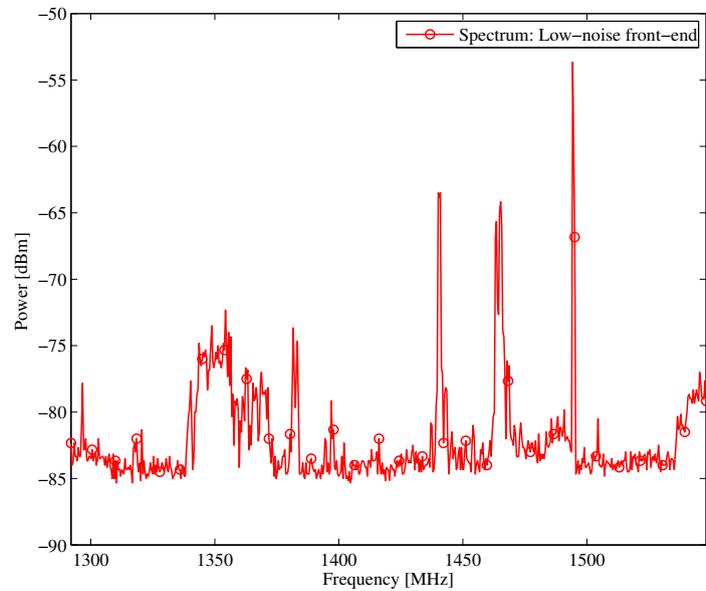
Table 7.2: Instrument parameters of the two-element digital correlation interferometer.

- -73.5 dBm at 1349 MHz,
- -73.66 dBm at 1382 MHz,
- -63.5 dBm at 1441 MHz,
- -64.1 dBm at 1465 MHz
- and -53.66 dBm at 1494 MHz.

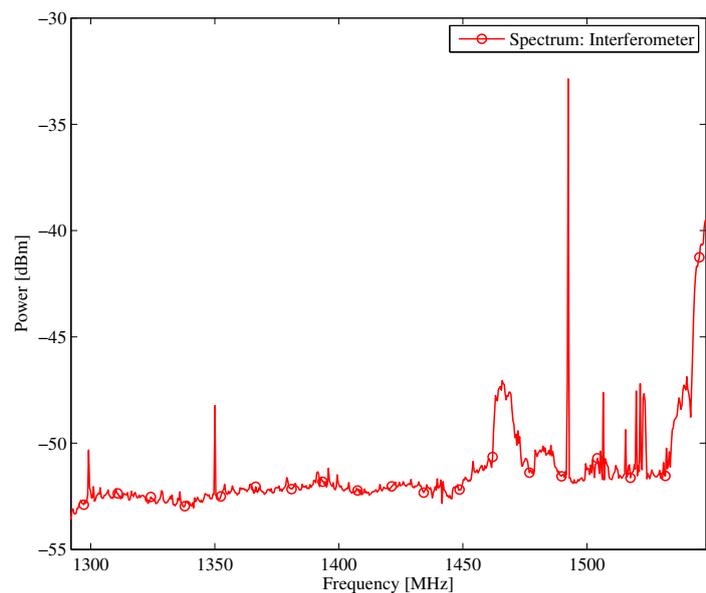
It is indicated in [29] that the frequency bands 1350-1400 MHz, 1429-1452 MHz and 1492-1518 MHz are allocated to fixed point-to-point link applications. Furthermore, the band from 1452-1492 MHz is used for digital audio broadcasting. The RFI measurement made with the interferometer is shown in figure 7.1(b). It is noticeable that much of the RFI that appeared in the middle of the band from the measurement using only the low-noise front-end, has dropped off significantly. This can be explained by the different pointing elevations used during these two measurements. The reflector in this measurement was pointed towards an elevation of 40° , whereas a pointing angle along the horizon was used during the initial measurements. Since most of the interfering signals according to [29], are used for point-to-point (line-of-sight) applications, it could explain the lower signals levels observed at increased elevations. However, in both measurements, the same strong RFI signals appear. It should be mentioned that the signal at 1441 MHz in figure 7.1(a) is not seen in figure 7.1(b). This was an intermittent (short-time span) signal, and it only sometimes appeared during a measurement.

Fortunately, even though there is a lot of RFI in the band, their power levels are still low enough to avoid the ADC from saturating. In figure 7.1(b) the maximum signal level is measured as -32 dBm. If this was not the case then a narrowband BPF was required upfront.

These results suggest that only a small segment of the available 256 MHz bandwidth can be used. This is unfortunate, but it was expected that there would be some level of interference. The RFI signals are furthermore coherent and will thus not be eliminated by the correlator. However, the FX correlator has the advantage of channelising the spectrum, and it is therefore possible to select an RFI-free portion of the band. The use of a PFB is also validated, since



(a)



(b)

Figure 7.1: RFI characterisation of the 1292-1548 MHz band using (a) the low-noise front-end with the reflector pointed towards the horizon, and (b) the spectrum measured using the interferometer, with the reflector pointed towards an elevation of 40° .

it provides excellent out-of-band rejection. This to some extent should ensure that the much stronger interfering signals, will not swamp out the much weaker astronomical signal. The results in figure 7.1 show that the band around 1400 MHz is fairly clean. This is expected, given that the 1400-1427 MHz band is assigned for radio astronomy [29]. All the observational results presented here after was made using this band. The achievable sensitivity in this case, for $\Delta\nu_{\text{IF}} = 27$ MHz and $\tau_a = 10$ sec is 55 Jy. This will make detecting Virgo A much more challenging, but the other sources should still be detectable.

7.3 First light

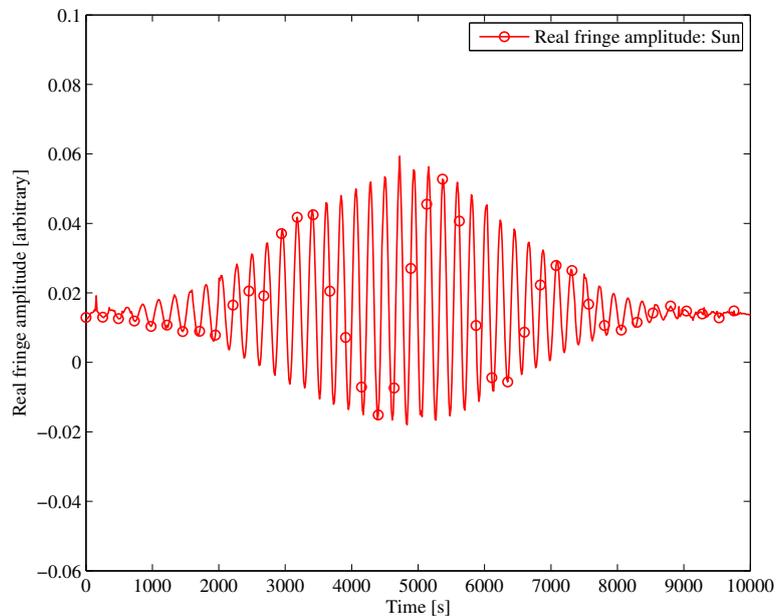
The interferometer operates in the meridian transit mode, with the antennas directed towards the northern meridian. By knowing the elevation \mathcal{E} of the source as it crosses the meridian, the antennas may then be pointed in this direction some time before the source transits. The source is then allowed to drift through the beam. A procedure is outlined in appendix A to compute the $(\mathcal{A}, \mathcal{E})$ coordinates- given the (δ, RA) - of the source. This procedure was used to determine the local transit time of the source.

The first observation of the sun was made on 2012/08/23, starting at 10:00 (08:00 UTC, GMT+2) and ending at 14:00 (12:00 UTC, GMT+2). The baseline used was $b = 15$ m. The interference fringes of the sun measured is shown in figure 7.2. In figure 7.2(b), a shorter time frame for this measurement is shown, which more clearly shows the fringes. The fringe frequency was computed using 3.19. The declination of the sun was $\sim 11.55^\circ$ and the E-W baseline was determined to be $u \approx 13.76$ m. The baseline was not precisely east-west, which caused a foreshortening in the baseline. It was estimated that the orientation error was about 23.4° relative to true north. This was determined using the time occurrence of the maximum fringe amplitude and computing what the azimuthal position of the sun was at that instance. The final pointing error was then determined using the average over a number of measurements. The theoretical fringe frequency is then found to be 4.62 mHz. From figure 7.2 the actual fringe period measured is 220 s, resulting in a fringe frequency of 4.54 mHz. This is only a 1.73% error.

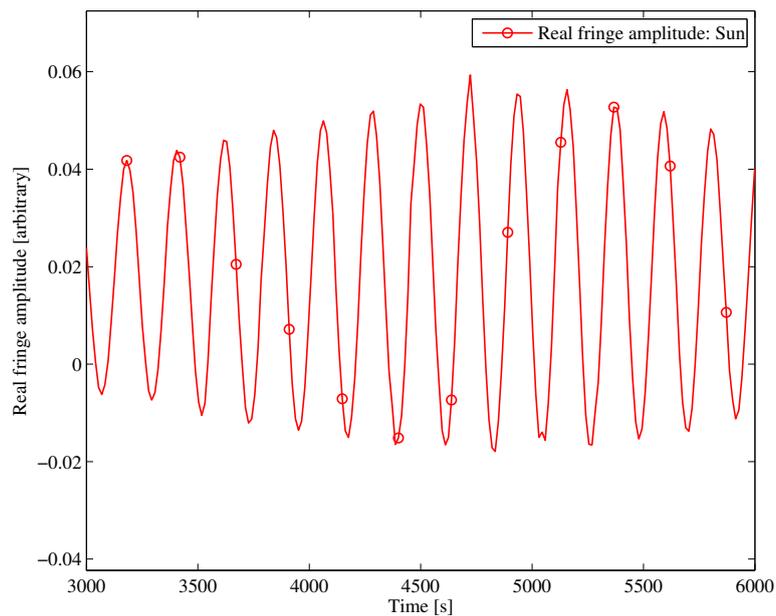
Given the much smaller bandwidth used $\Delta\nu_{\text{IF}} = 27$ MHz, the fringe washing pattern is not of a concern, since the first null only occurs at 38.8° from zenith. This is much wider than the HPBW of the antenna. The modulation of the fringes by the antenna pattern is however noticeable. The pattern of the main lobe as shown in figure 5.28, was determined from the fringe pattern in figure 7.2. The logarithm of the normalised fringe pattern was taken and the time between the -3 dB points was measured. Given that the sun moves about $\sim 15^\circ$ per hour the HPBW was computed as about 14.7° .

7.4 Experiments and other observations

From the first light measurement it was confirmed that the interferometer was sensitive enough to perform solar observations. Other science experiments were then undertaken using the two-



(a)



(b)

Figure 7.2: The (a) interference fringes of the quiet sun at 1.42 GHz for a baseline of 15 m, and (b) a close-up of the fringes. This was the first light observation and was made on 2012/08/23. The interferometer was not calibrated and the fringe amplitude is in an arbitrary scale.

element correlation interferometer. Specifically, the investigative component of this project was to determine whether it is possible to detect galactic sources, such as those specified in table 2.1, with affordable off-the-shelf technologies in the presence of man-made interference. This section discusses the observational results of both Virgo A and Taurus A. Furthermore, the interferometer had a variable baseline up to a maximum of 24 m. A series of solar observations were made in an attempt to synthesise a 1-dimensional image (profile) of the sun. These results are also discussed.

7.4.1 One-dimensional image of the sun

In section 3.7 it is mentioned that it is possible to synthesise an image of a source, with only two antennas. This is accomplished by making a series of observations over several days. Each day the antennas are moved to a new position, resulting in a different baseline. Following this approach an attempt was made to synthesise a one-dimensional image of the sun. The experiment consisted of 8 separate measurements, where the minimum baseline was 3 m and the longest 24 m. Since, the angular (optical) size of the sun is $\sim 30'$, the baseline sampling interval Δu had to be small enough to avoid aliasing in the image. This meant that Δu must be less than 114.6λ . At $\lambda = 21.1$ cm the baseline spacing must be $\Delta b \leq 24$ m, the spacing was then chosen as 3 m.

The measured fringes of the sun for a baseline of 3, 12 and 21 meters are shown in figure 7.3. It is shown that as the baseline length increases so does the fringe frequency. The maximum fringe frequencies for the three different baselines shown in figure 7.3, are measured as:

- 1.03 mHz for $b = 3$ m, theoretically it should be 0.932 mHz (10.5% error),
- 3.76 mHz for $b = 12$ m, theoretically it should be 3.73 mHz (8% error) and
- 7.14 mHz for $b = 21$ m, theoretically it should be 6.6 mHz (8.1% error).

As expected, the fringe amplitude tends to decrease as the baseline increases, this is because the source is being resolved out. The peak-to-peak fringe amplitude for the three cases in figure 7.3 was measured as 0.404, 0.183 and 0.021 from the shortest to longest baseline, respectively. The very strong spikes observed in figure 7.3 are caused by interference. This highlights the fact that even with the much reduced bandwidth, there is still some interference present.

The synthesised image of the sun made with the 8 measured Fourier components, is presented in figure 7.4. The image of the sun is assumed to be a disc of angular extent 0.5° , with a uniform brightness. The synthesised image is shown to be nearly Gaussian, which is not the expected result. The reasons for this is not entirely known. Possibly, too few Fourier components were used to reconstruct the source distribution, and perhaps the maximum baseline is not long enough.

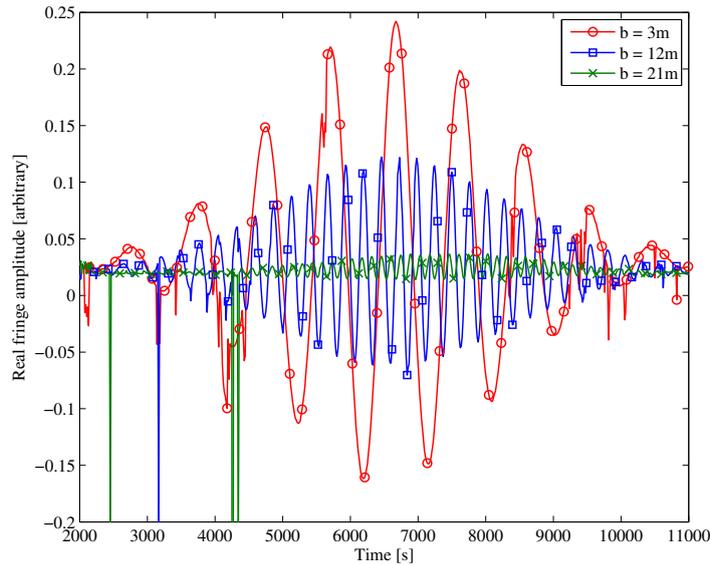


Figure 7.3: The fringes of the quiet sun at 1.42 GHz for the baselines, $b = 3$ m, $b = 12$ m and $b = 21$ m.

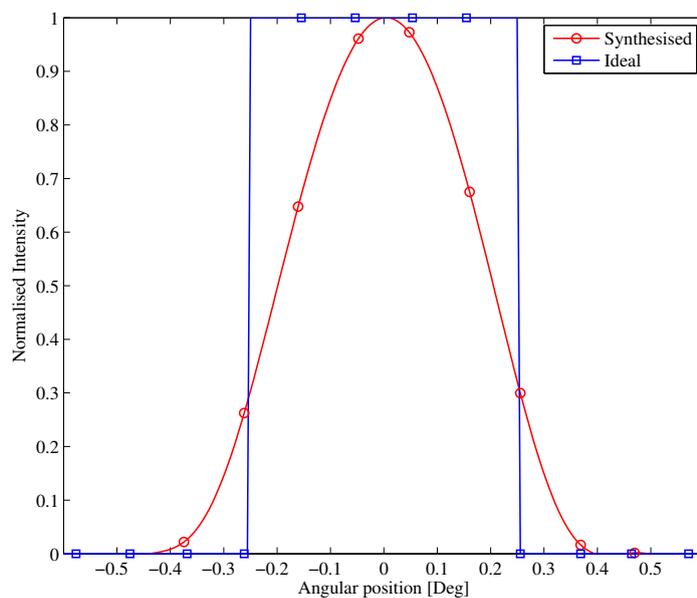


Figure 7.4: The synthesised image of the quiet sun at 1.42GHz. This image was constructed using 8 Fourier components ($b = 3 \dots 24$ m, $\Delta u = 3$ m), and the Inverse FFT was zero-padded to 512.

7.4.2 Observations of non-solar sources

Due to the geographical location of the interferometer, the altazimuth mounts ($\mathcal{E} \leq 55^\circ$) used and the time of year, the only two non-solar sources that could be observed was Virgo A and Taurus A. According to table 2.1, the flux density of Taurus A at 1420 MHz is 893 Jy, and should therefore be detectable. Using the same approach as before, the elevation of Taurus A

was computed as 34.8° at the time of transit. The fringes produced by Taurus A are shown in figure 7.5. The fringes are easily discernible from the noise, even though the noise corrupts much of the signal. The theoretical fringe frequency for this observation is 1.46 mHz, for a baseline of $b = 5$ m and $\delta \approx 22^\circ$. The fringe frequency was measured as 1.43 mHz which is in good agreement with the theoretical value.

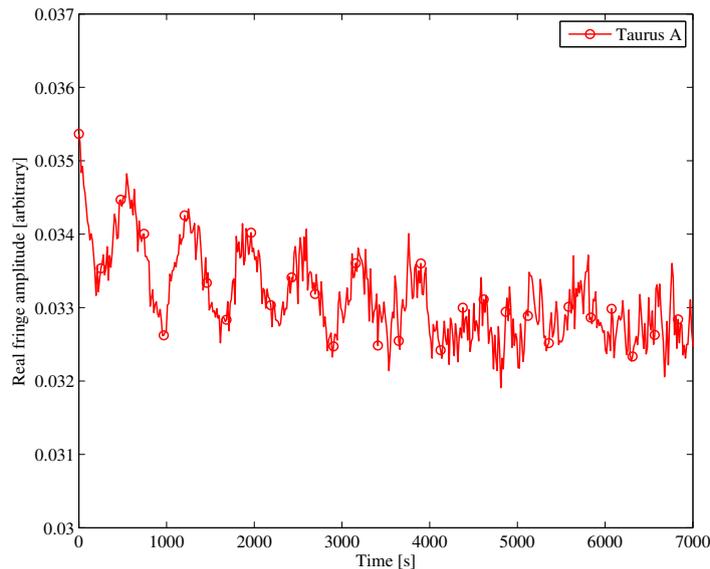


Figure 7.5: The measured fringes of Taurus A for a baseline of $b = 5$ m. This observation was made on 2012/09/06, and the time of transit was 06:32:00. The integration time τ_a used was 10s and $\Delta\nu_{IF} = 27$ MHz.

Several attempts were made to detect Virgo A, but there was too much uncertainty in the results obtained to conclusively state whether there are fringes or not. It is not to say that the sensitivity of the instrument is too low. It could also very well be ascribed to pointing inaccuracies. It is mentioned in the literature that a pointing error of less than one tenth of the beamwidth is required to make a proper measurement [12, 11].

The results presented in this section indicates that the sensitivity of the interferometer is most likely less than predicted. However, the instrument is still sensitive enough to detect much weaker sources other than the sun. The sources - Cassiopeia A; Cygnus A; and Centaurus A should also be detectable, given that their flux densities are higher than that of Taurus A at 1420 MHz.

7.5 Conclusion

This chapter presented the observational results obtained from the measurements made with the two-element digital correlation interferometer. It was noted that the RFI environment

at 1292-1548 MHz was in fact much more cluttered than initially anticipated. From the RFI measurements made with the much more sensitive low-noise front-end of the instrument, it was evident that the band from 1400-1427 MHz was free from any sources of RFI, this limited the usable bandwidth of the interferometer to 27 MHz. The achievable sensitivity is thus lower, and expected to be around 55 Jy. From the observations made it was verified that the instrument is capable of detecting the sun and other much weaker galactic (non-solar) sources, in this case Taurus A. Several unsuccessful attempts were made to detect Virgo A. Finally, the synthesis of a 1-dimensional image of the sun was also attempted, by making 8 measurements of the sun at different baselines. The image of the sun constructed was Gaussian in shape, and not the expected uniform profile of 0.5° in extent.

Chapter 8

Conclusion

8.1 Summary of the work

A two-element digital correlation interferometer was investigated and designed specifically for detecting some of the stronger routinely observed sources in the sky. The main objective was to demonstrate whether it is possible to make professional radio astronomy measurements in the presence of strong RFI, using off-the-shelf technologies. To follow is a summary of the work performed in this thesis to realise the interferometer.

- Initially, the primary focus of the thesis was to develop a working knowledge of the fundamentals of radio astronomy and interferometry. Concepts such as the brightness distribution, flux density, antenna noise temperature, the Van Cittert-Zernike theorem and aperture synthesis proved to be of importance. This theoretical review was necessary in order to establish a reference design for the interferometer.
- The focus then shifted towards quantifying the system specifications. This required, identifying a list of routinely observed sources and their minimum flux densities. Table 2.1 lists some of the stronger sources that are easily observed with amateur instruments.
- These sources are all broadband radiators, and the decision of what operating band to use, was thus primarily driven by RFI considerations. In this regard, an RFI survey of the local interference environment was performed, and from this a radio quiet band was identified.
- Next, the minimum allowable sensitivity of the interferometer was defined, such that the source with the lowest flux density in table 2.1 could still be detected. From this the global system specifications, such as the system noise temperature, the antenna gain, the receiver bandwidth and the integration time of the interferometer was quantified, see table 7.2.
- Firstly, the receiver for the interferometer was designed. However, this required that the sensitivity degradation owing the receiver parameters, such as passband ripple, gain roll-off and linearity be defined. From this the receiver performance parameters were derived.

Ultimately, a dual-conversion superheterodyne receiver was designed, constructed and measured.

- This was followed by designing a suitable antenna for the interferometer. The chosen antenna type was a helical beam antenna fed offset-parabolic reflector. The helical beam antennas were designed, manufactured and measured to verify their performance. The offset-parabolic reflectors used, were commercially available direct-broadcast satellite dishes.
- An FX correlator was designed and implemented on a ROACH board. The CASPER hardware and software tools were used throughout the correlator design.
- Finally, the various subsystems were integrated, and several observations (solar and galactic) and experiments were made to verify the performance of the interferometer.

8.2 Observations and findings

The astronomical observations made with the interferometer verified the functionality of the instrument, for which it was designed for. Presented below is a summary of the results obtained.

- The fringes of the quiet sun was successfully measured. This was done for multiple baselines in an attempt to make a one-dimensional image of the sun. The reconstructed image did not resemble the expected uniform profile of the sun.
- The sensitivity of the interferometer also proved to be good enough to detect Taurus A (~ 893 Jy at 1420 MHz).
- Several attempts were made to measure Virgo A, but since the presence of fringes in these measurements could not be verified with certainty. It was decided the interferometer was unsuccessful in detecting this source.

These measurements verified the end-to-end functionality of the instrument. Furthermore, it has been confirmed that it is possible to detect the sun and other much weaker sources with a small instrument in the presence of RFI.

Some of the important findings made during the operation of the interferometer are presented below.

- It was found that there was in fact a lot of RFI in the passband. These signals only appeared after the much more sensitive receiver was used. This resulted in the usable passband only being about 27 MHz. Correspondingly, the sensitivity of the interferometer decreased. This was a contributing factor to the unsuccessful detection of Virgo A.
- It was found early on that the baseline was not truly east-west by about 23.4° . Taking this into account the fringe frequencies measured for the various observations were in close agreement with the theoretical values.

8.3 Suggestions for future work

The prototype developed in this thesis demonstrates the feasibility of performing radio interferometry using commercially available hardware. However, as a system there is still much work which has to be done. This suggests that there is an abundance of potential research opportunities. On a system level, there are a couple of weak points that have been identified. Firstly, the electrically small antennas are possibly the biggest limitation on the interferometer's sensitivity. This provides an opportunity to investigate other more suitable antenna technologies. Another aspect of the system that should be improved is the filter used upfront, for rejecting the out-of-band RFI. Moreover, it was shown that a much narrower passband had to be used in the end. This provides for interesting opportunities in researching RFI mitigation strategies, to protect wideband receivers from strong sources of RFI. Lastly, the interferometer must be calibrated. Given the low sensitivity of the instrument, using sources in the sky to calibrate the system is not plausible, this provides an opportunity to investigate other calibration solutions. On a larger scale, more antennas can be used to construct a small synthesis array, allowing research into imaging. In conclusion, the author hopes the interferometer developed in this thesis will serve as a platform for other engineers to gain a working knowledge of the field of radio interferometry, and to develop the skill sets needed to start addressing the major engineering challenges the SKA is faced with.

Appendix A

System of coordinates and geometric relationships

Several coordinate systems have been defined to describe the position of celestial bodies relative to an observer on earth. These coordinate systems are the horizon-, equatorial-, ecliptic- and the galactic-coordinate systems. The choice of which coordinate system to use depends on the observation being performed, and the celestial object being studied [18]. For instance, the ecliptic system takes the plane of the ecliptic as its reference, and is most appropriate for studying the solar system, whereas the galactic coordinate system is better suited for observing more distant sources in our galaxy [14].

Two identical altazimuth (altitude-azimuth) mounts are used for the antennas in this thesis. The horizon system of coordinates is then a convenient coordinate system to use, since the position of the object is given by the azimuth and altitude (or elevation). However, due to the earth's rotation the coordinates of the object, if specified in the horizon system changes throughout the day. This is clearly not ideal if a fixed position must be defined for a celestial object. A more useful coordinate system is then the equatorial coordinate system, where the position of the object is fixed and furthermore not dependent on the geographical location of the observer. The equatorial system defines the coordinates of the object by the right ascension and declination. Since, the position of cosmic sources are mostly specified in the equatorial system, a transformation is thus needed to convert this position to the azimuth and elevation coordinates, for purposes of observations made with altazimuth mounts.

This appendix presents a brief overview of the horizon and equatorial coordinate systems and how to perform the transformation between these coordinate systems.

A.1 Horizon coordinate system

The local horizon of an observer at longitude and latitude (λ, \mathcal{L}) serves as the plane of reference. The coordinate system is further specified by the zenith, which is the point 90° above the local horizon and nadir, 90° below the horizon. Lastly, the object circle is specified as the circle through the object and the zenith, see figure A.1 for an illustration. The position of the source

is specified by the azimuthal and elevation angles (\mathcal{A}, \mathcal{E}). More specifically, the azimuth is the angle measured horizontally from the north in a clockwise direction, and the elevation is the vertical angle measured from the horizon upward to the object. The coordinates of the object is dependent on the observer's position and is only meaningful when combined with the time of the observation. It is also clear that the position of the object changes as the earth rotates.

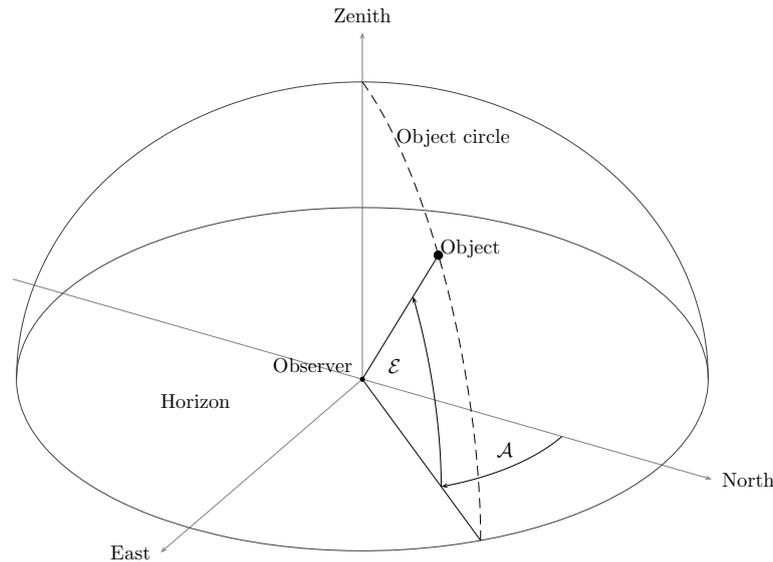


Figure A.1: The horizon system of coordinates.

A.2 Equatorial coordinates

The equatorial coordinate system overcomes the variable position problem of the horizon coordinate system. Moreover, the position is not dependent on the observer's location on earth. The coordinate system is conceived by assuming an imaginary celestial sphere that has the earth at its center. The reference plane is then specified as the earth's equator, which is projected onto the celestial sphere and is known as the celestial equator. The earth's axis extends to intersect the celestial sphere, and the poles are known as the north celestial pole (NCP) and south celestial pole (SCP). In figure A.2, the coordinates of the object is specified by constructing a great circle which passes through the object and connected to the NCP and SCP. This circle is known as the object's hour circle, and is similar to the object circle defined in the horizon system of coordinates. Next the position of the celestial object is given by the declination and right ascension (δ, RA). Where, the declination is simply the angle between the celestial equator and the object. The declination is expressed as positive if north of the celestial equator and negative below the equator. Any longitudinal lines from the earth projected onto the celestial sphere will change as the earth rotates, and is therefore not a suitable solution for specifying the horizontal position of the object. The earth's axis is inclined by 23.5° and this results in the intersection of the plane of the ecliptic and the earth's equator. One of these intersection

points is the Vernal Equinox ($RA = 0h$) and this is the reference position from where the RA of the source is measured from. The RA is measured eastwards from the Vernal Equinox and is commonly expressed in hours, minutes, seconds [14]. The right ascension and declination allows the position of the object to be defined in a relatively fixed manner that is independent of the diurnal motion of the earth.

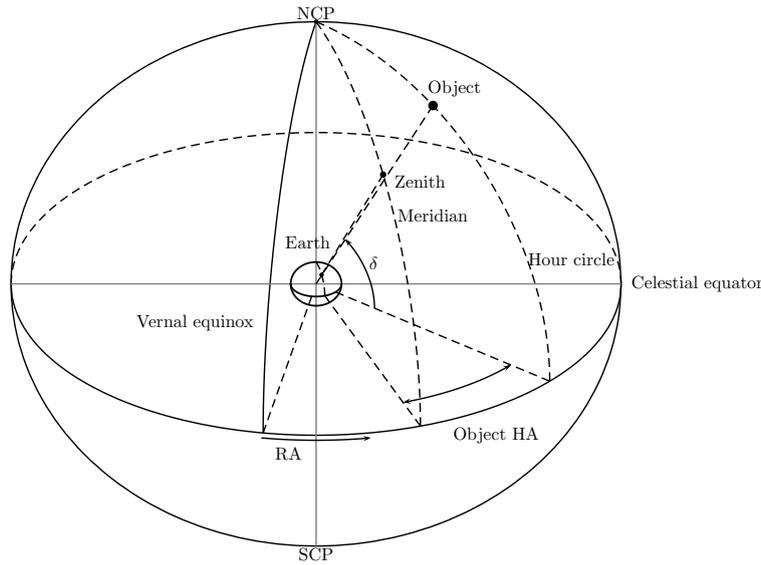


Figure A.2: The equatorial system of coordinates.

The equatorial coordinate system is relatively fixed, but the cyclic precession of the earth causes a gradual change in the coordinates of the object. The precession of the earth's axis around the pole of the ecliptic is cyclic with a period of 26000 years [14, 18]. To account for this shift in the equatorial coordinates, the declination and right ascension refers to a specific date known as an epoch. The epochs are defined in 50 year intervals, and epoch 1900 for example refers to January, 1, 1900. The change in (δ, RA) between epochs was discussed in detail in section 2.5.

To define the equatorial coordinates of the object relative to the observer's location for a given time, then the hour angle (HA) of the object must be specified [18]. The HA is defined by firstly drawing a great circle passing through the zenith of the observer and the celestial poles. This is known as the meridian circle. Then the angle between between the hour circle and meridian circle is the object's HA

$$HA = RA \text{ of meridian} - RA \text{ of object.} \quad (\text{A.1})$$

The HA is negative before the object transits the meridian and positive thereafter, and it is also normally specified in hours, minutes, seconds.

For example, the first light measurement performed in section 7.3 was conducted on 2012/08/23. Assuming, a time of 11:25 (09:25 UTC, GMT+2). With the aid of *PyEphem*, the declination and right ascension of the sun was computed as $(\delta, \text{RA}) = 11.23^\circ, 152.42^\circ$. Before (A.2) can be used to compute the local $(\mathcal{A}, \mathcal{E})$ position of the sun, the HA must first be determined. This requires that the longitude and latitude (λ, \mathcal{L}) of the observer be known, for Stellenbosch this is $(18.51^\circ\text{E}, 33.56^\circ\text{S})$. The HA was found to be -20.87° , this suggests that the sun has not yet passed the meridian of the observer. Substituting (δ, HA) into (A.2), results in $(\mathcal{A}, \mathcal{E}) = (27.5^\circ, 41^\circ)$.

Appendix B

Noise as a random process

The radio emissions received from a radio source, can simply be described as a series of voltage pulses, that vary randomly on a very short time scale. The ensemble of these pulses have a Gaussian amplitude distribution [15]. Hence, the non-deterministic nature of these signals validate the need for a statistical approach to analysing a radio telescope receiver.

The intent of this appendix is to provide an overview of some of the important statistical concepts, that are related to radio astronomy.

B.1 The Gaussian PDF

The random, short-duration pulses that describe a cosmic source can be viewed as a summation of independent random variables (RV), $x = x_1 + x_2 + \dots + x_n$. By the central limit theorem - the RV x tends to a Gaussian distribution for large n [40], with a probability density function (PDF)

$$f_x(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)/2\sigma^2}, \quad (\text{B.1})$$

with mean μ and variance σ^2 . The second central moment of a RV x , is known as the variance σ_x^2 , and is defined as

$$\sigma_x^2 = \overline{x^2} - \bar{x}^2, \quad (\text{B.2})$$

where \bar{x} denotes the expected value of the RV x . For thermal noise, μ is assumed zero. For a correlation receiver two RVs x and y are needed to describe the signals of the two separate receiver chains. The bivariate (joint) Gaussian density of x and y is given by

$$f_{xy}(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left(\frac{1}{2(1-\rho^2)} \left[\frac{(x-\mu_x)^2}{\sigma_x^2} - 2\rho\frac{(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} + \frac{(y-\mu_y)^2}{\sigma_y^2} \right]\right). \quad (\text{B.3})$$

The correlation coefficient ρ , between the RVs x and y , is by definition given by the ratio

$$\rho = \frac{C}{\sigma_x \sigma_y}, \quad (\text{B.4})$$

where C is the covariance between the RVs x and y ,

$$C = \overline{xy} - \bar{x}\bar{y}. \quad (\text{B.5})$$

B.2 Specification and classification of a random process

Spectrometers are widely used in radio astronomy, and the function of a spectrometer is to measure the Power Spectral Density (PSD) of the source under observation. This suggests that the time dependency of the cosmic signal needs to be considered. Thus, the RV x (defined in section B.1) is now a function of time, $x(t)$, and is formally known as a random (or stochastic) process.

The immediate goal is to find some quantitative measure, that will completely specify the random process $x(t)$. In this regard, a random process may be seen as a collection of an infinite number of RVs, that are usually dependent [80]. In other words, let x_i represent the RV $x(t_i)$, that has a sample space equal to the amplitudes of the random process at time t_i . It then follows that $x(t)$ is completely described by the joint PDF $f_{x_1 x_2 \dots x_n}(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n)$, where $n \rightarrow \infty$. However, only the first- and second-order statistics are of importance for the topics discussed in this thesis.

The first-order statistic can be derived from $f_{x_1 x_2}(x_1, x_2)$, by simple integration

$$f_{x_1}(x_1) = \int_{-\infty}^{\infty} f_{x_1 x_2}(x_1, x_2) dx_2, \quad (\text{B.6})$$

this represents the marginal density, which can then be used to determine the mean $\overline{x(t)}$ of the random process $x(t)$,

$$\overline{x(t)} = \int_{-\infty}^{\infty} x f_x(x; t) dx. \quad (\text{B.7})$$

It is clear from (B.7), that the mean is a function of time, thus $\overline{x_1}$ will most likely differ from $\overline{x_2}$. However, it will be shown that if the random process is stationary, then the mean $\overline{x(t)}$ is constant. The second-order statistic is arguably the most important characteristic of a random process, especially regarding radio astronomy. The ensemble average from the joint PDF of x_1 and x_2 , is the autocorrelation of $x(t)$ [80]

$$R_x(t_1, t_2) = \overline{x_1 x_2} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_{x_1 x_2}(x_1, x_2) dx_1 dx_2. \quad (\text{B.8})$$

It will be shown in section B.3, that the autocorrelation function of (B.8) leads to the spectral content of a random process. Before this is shown it is necessary to classify a random process

as stationary.

A random process is defined as stationary if its statistical characteristics is independent of time, and that a shift of the time origin will be impossible to detect. For the random process to be strictly stationary, the n -th order ($n \rightarrow \infty$) statistic has to be shown to be stationary. This is difficult to verify, but fortunately a suitable condition is to at least have the mean value, and the autocorrelation function independent of a shift in time. This is mathematically expressed as

$$\overline{x(t)} = K, \quad (\text{B.9})$$

$$R_x(t_1, t_2) = R_x(\tau), \quad (\text{B.10})$$

where K is a constant, and $\tau = t_2 - t_1$. A random process that fulfils both these requirements are known to be Wide-Sense Stationary (WSS).

B.3 PSD of a random process

The spectral content of a stationary (or at least WSS) random process is the Fourier transform of its autocorrelation function, $R_x(\tau) \iff S_x(\omega)$ [80]. This is known as the Wiener-Khinchin theorem,

$$S_x(\omega) = \int_{-\infty}^{\infty} R_x(\tau) e^{-j\omega\tau} d\tau. \quad (\text{B.11})$$

Only the result of the Wiener-Khinchin theorem is presented here, but a derivation of the theorem is presented in [19, 80]. It should be appreciated that the computed PSD is meaningful only when the random process is stationary (WSS). The total average power of a random process can further be derived from the inverse Fourier transform of (B.11),

$$R_x(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_x(\omega) e^{j\omega\tau} d\omega, \quad (\text{B.12})$$

and then by adding all the spectral components of $S_x(\omega)$ ($\tau = 0$ in (B.12)) results in

$$P_x = \overline{x^2} = R_x(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_x(\omega) d\omega. \quad (\text{B.13})$$

B.4 Ergodic processes

Equations (B.7) and (B.8), are the ensemble averages of the statistics of the random process $x(t)$. A large number of sample functions (waveforms) are required to compute these ensemble averages. However, to obtain a large number of sample functions is usually not practically possible. A solution it to ensure that the random process is ergodic, since the ensemble- and time-averages are then equal

$$\overline{x(t)} = \langle x(t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t) dt, \quad (\text{B.14})$$

$$R_x(\tau) = r_x(\tau) = \langle x(t)x(t + \tau) \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)x(t + \tau) dt, \quad (\text{B.15})$$

where the acute braces $\langle \cdot \rangle$ indicate the time averages of the mean and autocorrelation function. If a random process is ergodic, then only one sample function is needed to determine the ensemble averages, which is clearly advantageous. Since, the signals received from a radio source is usually assumed to be ergodic- the correlator output is a time average, which is equal to the ensemble average according to (B.14) and (B.15).

Appendix C

Hardware details of the dual-conversion superheterodyne receiver

In this appendix the details of the receiver packaging are discussed. This includes, the receiver enclosure, the mechanical layout of the components, the coaxial cable connections, and the power distribution. Furthermore, the details of the power limiting circuit are presented.

C.1 Enclosure

The receiver enclosure is constructed from soft steel, and further tin-plated to protect it from corrosion. The components are directly mounted onto the base-plate of the chassis. Furthermore, L-plates are used to connect the RF signals to the receiver chassis. A detailed mechanical drawing of the receiver chassis is depicted in figure C.1. From this figure the overall dimensions of the receiver chassis are $L \times W \times H = 700 \text{ mm} \times 500 \text{ mm} \times 50 \text{ mm}$. Additionally, the low-noise front-end units are housed in separate metal box enclosures, with dimensions $L \times W \times H = 330 \text{ mm} \times 100 \text{ mm} \times 45 \text{ mm}$. Similarly, the components are directly mounted onto the base of these enclosures. Photos of both the receiver board and the low-noise front-end units are shown in figures E.2 and E.3, respectively.

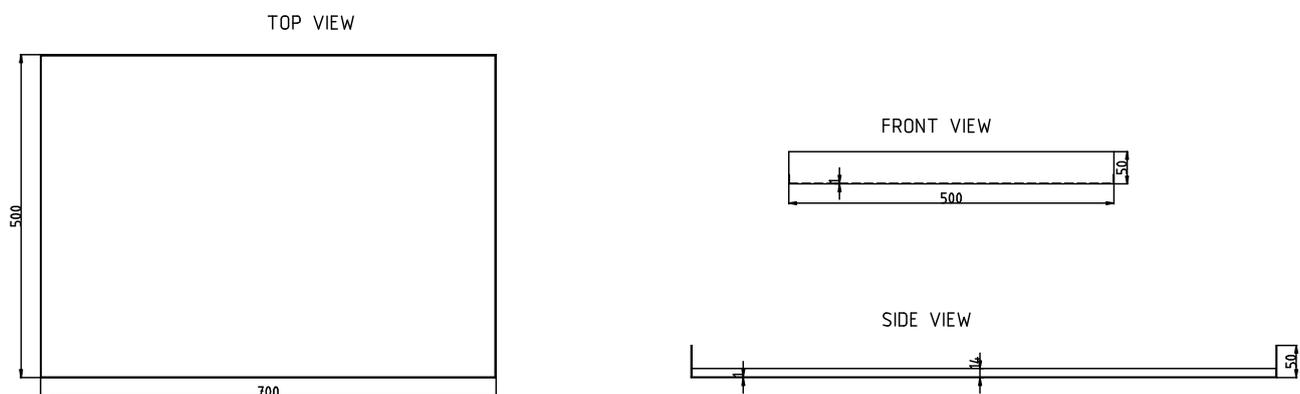


Figure C.1: Mechanical drawing of the receiver chassis.

C.2 Power requirements and distribution

The power supply requirements for the receiver are as follows:

- +5 V@390 mA (2 W),
- +12 V@1.1 A (13.2 W).

The DC operating voltages and currents for the components are summarised in table 4.2. The use of switch-mode power suppliers are prohibited, due to switching noise. Linear power supplies must therefore be used. The +5 V rail is supplied by the Callex 32005AR (5 Vdc@3 A) linear supply, and the +12 V rail is provided by the Callex 32012AR (15 Vdc@1.7 A) linear supply.

A centralised power distribution approach was followed, whereby the regulated +5 V and +12 V supplies are directly connected to the receiver chassis. The complexity of this approach is low and well suited for a small localised systems, as in this case. Because of the low current requirements, the +5 V and +12 V supply rails were simply implemented with AWG-24 conductors. The resistive loss of these wires are $29.3 \Omega/\text{kft}$, resulting in a voltage drop of $\sim 105 \text{ mV}$ and $\sim 37.4 \text{ mV}$ over the length of the +5 V and +12 V rails, respectively. The receiver chassis served as a ground plane, since all the components were grounded to their casings. Secondly, the base-plate of the receiver chassis acts as a heat sink, and dissipates the heat generated by the components. The supply rails were attached right onto the chassis in order to minimise potential ground loops. Furthermore, the respective power rails for each channel were separated, to try an improve the isolation between the two channels. An illustration of the power distribution for the receiver chassis is shown in figure C.2.

The low-noise front-end units are situated at the antenna feeds. The two ZX60-1614LN low-noise amplifiers housed in each unit must be powered. The solution chosen to distribute power to these components consisted of using bias-tees (Mini-Circuits ZFBT-4R2G-FT). The voltage drop over the 35 m LMR-600 cable was measured to be less than 0.5 volts, and this was acceptable. The power distribution within the low-noise front-end units is shown in figure C.3

C.3 Details of the power limiting circuit

The maximum allowable input power for the receiver was determined to be -34 dBm. At this input level, the maximum power rating of the downconverting mixer (ZX05-73L) is reached, and permanent damage may occur. To protect the receiver from exceeding this power threshold, a power limiting circuit was integrated into the receiver chain. Figure 4.12 shows that the power limiting circuit is placed after the low-noise front-end. The block diagram of the power limiting circuit is shown in figure C.4, and a summary of the components used appear in table 4.2.

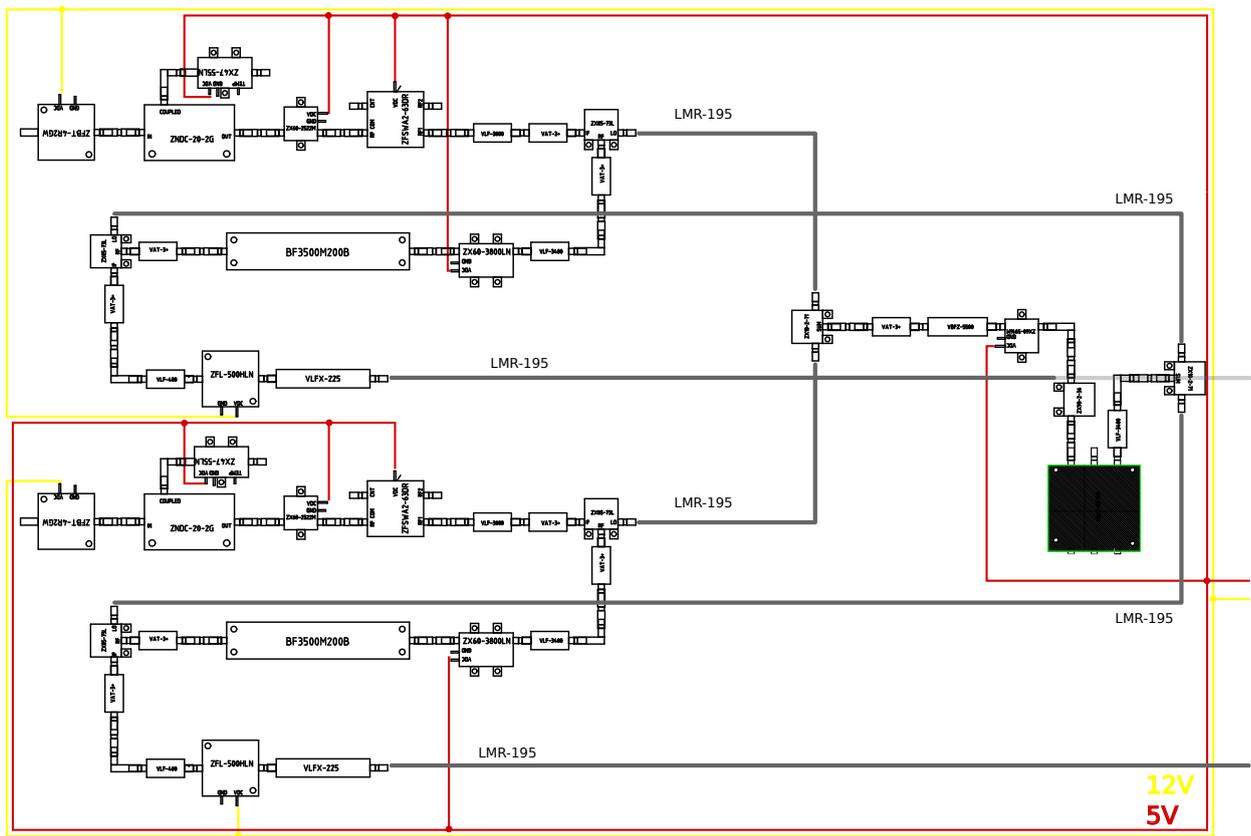


Figure C.2: An illustration of the power distribution throughout the receiver chassis. Furthermore, the coaxial cable connections within the receiver chassis is shown.

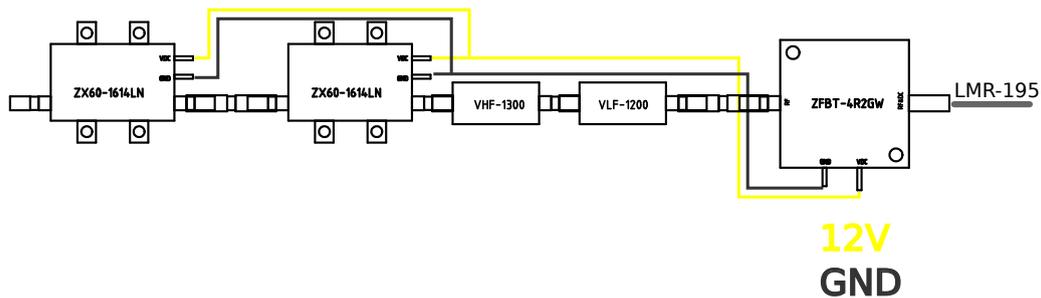


Figure C.3: An illustration of the power distribution within the low-noise front-end.

The power limiting function is implemented entirely with analogue circuitry. The basic operation of this circuit is as follows, if the input power exceeds -40 dBm then the low-noise front-end is switched out of the receiver chain. This effectively isolates the rest of the receiver. Once, the input power has dropped by at least 3 dB then only does the low-noise front-end get switched back into the receiver path. The SPDT ZFSAW2-63DR (Mini-Circuits) RF switch is controlled by a Schmitt trigger. The circuit components used in the Schmitt trigger design is listed in

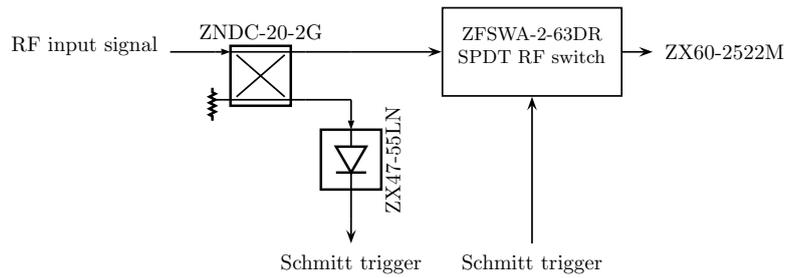


Figure C.4: Block diagram of the power limiting circuit.

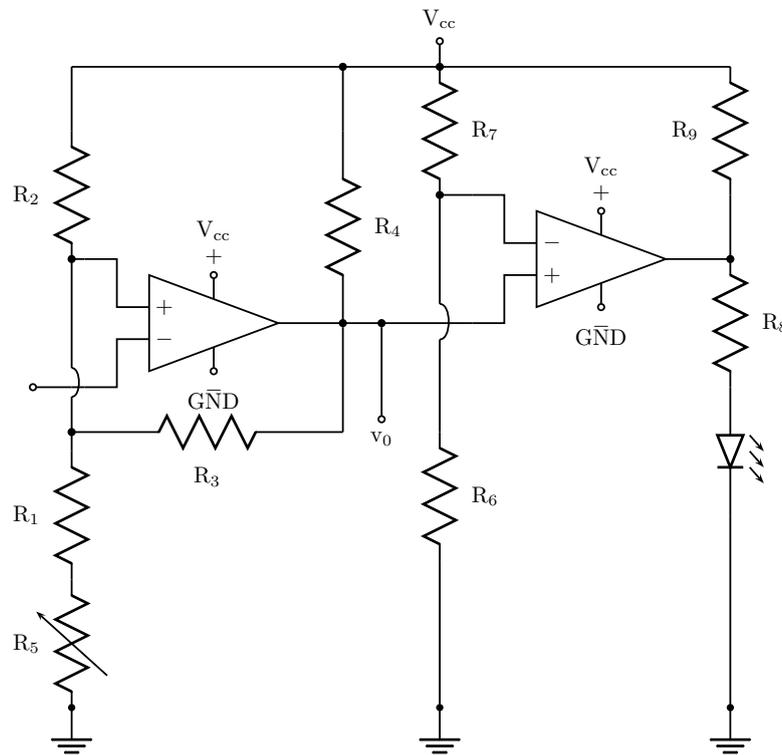
Designator	Value	Description
R_1	1 k Ω	1/4W carbon film leaded resistor
R_2	3.3 k Ω	1/4W carbon film leaded resistor
R_3	100 k Ω	1/4W carbon film leaded resistor
R_4, R_6, R_7, R_9	2.2 k Ω	1/4W carbon film leaded resistor
R_5	≤ 1 k Ω	Trimming potentiometer
R_8	500 Ω	1/4W carbon film leaded resistor
C_1	100 nF	Tantalum leaded capacitor
LM339	–	Quad comparator IC
BL-B5134	–	5.0 mm LED

Table C.1: A list of the Schmitt trigger circuit components.

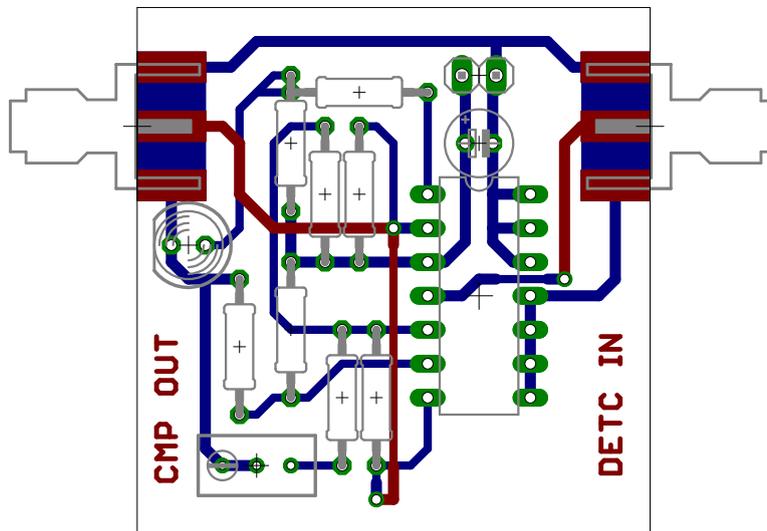
table C.1. The circuit schematic and the PCB design of the Schmitt trigger circuit is presented in figure C.5.

C.4 Coaxial cables

The various coaxial cables used in the system for carrying the signals are listed in table C.2. All the coaxial cables are of the flexible type. Since, this simplified the board layout design. The cable connections within the receiver chassis are shown in figure C.2.



(a)



(b)

Figure C.5: The (a) circuit schematic and the (b) PCB of the Schmitt trigger. The ZNDC-20-2G couples some of the input power, and the ZX47-55LN power detector outputs a corresponding DC voltage. The Schmitt trigger is calibrated to toggle the ZFSWA2-63DR RF switch if the input power exceeds -40 dBm.

Designator	length [mm]	Quantity	Atten [dB/100m]	Connectors	Description
LMR-240	1000	2	31.46	SMA M-M	Feed antenna to low-noise front-end unit
LMR-195	150	2	41.31	SMA F-M	Bias-Tee (antenna side) to N-type adapter
LMR-600	35000	2	10.55	N-type M-M	Signal transmission system
LMR-195	150	2	41.31	SMA M-M	LO1
LMR-195	700	2	41.31	SMA M-M	LO2
LMR-195	500	2	41.31	SMA M-F	Receiver output to ROACH board
LMR-195	150	1	41.31	SMA M-F	Valon reference to ROACH board
RG-316	1000	7	102.89	SMA M-M	Receiver to ROACH board

Table C.2: A list of the coaxial cables used throughout the interferometer system. All attenuation values are specified for 1420 MHz.

Appendix D

User guide for the FX correlator

This appendix serves as a user guide for setting up and operating the FX correlator. The instructions for connecting up the hardware correctly, and configuring the data-logging and supervisory control software are presented.

D.1 Hardware setup

The ROACH board setup for the FX correlator is shown in figure D.1, additionally the relevant connectors and ports are highlighted. The hardware setup procedure for the FX correlator is summarised by the following steps:

1. Connect the KatADC to the Z-DOK0 connector. This is the connector closest to the power supply.
2. Connect the two input signals to channels i and q of the KatADC (the two outer connectors). The KatADC is fitted with 50Ω SMA connectors. To safely operate the KatADC avoid input power levels of $+7.6\text{ dB}^1$. This may cause permanent damage. Furthermore, the full scale range of the ADC is $+3.3\text{ dBm}$. ideally, the power level for the FX correlator is -27 dBm , this provides enough headroom for the ADC in the presence of strong RFI and ensures that the FFT does not overflow.
3. Next connect the output of the Valon 5007 synthesiser to the clock input of the KatADC. The clock signal must have a power level of 0 dBm at 800 MHz . The Valon 5007 synthesiser can be programmed using either the Windows Configuration Manager software or the Linux python interface. Both sets of software are available at <http://www.valontechnology.com/5007/5007.htm>.
4. Connect the 100MbE port of the ROACH board to a control computer ². Any standard Ethernet twisted-pair cable may be used.

¹All power level ratings presented in this guide assumes an attenuator setting of 20 dB for the KatADC.

²The ROACH board with serial number *020217*, has been added to the Stellenbosch University network. The ROACH board can therefore be connected to any available network port, that is part of the University's network. And, any computer on the network can then access the ROACH board.

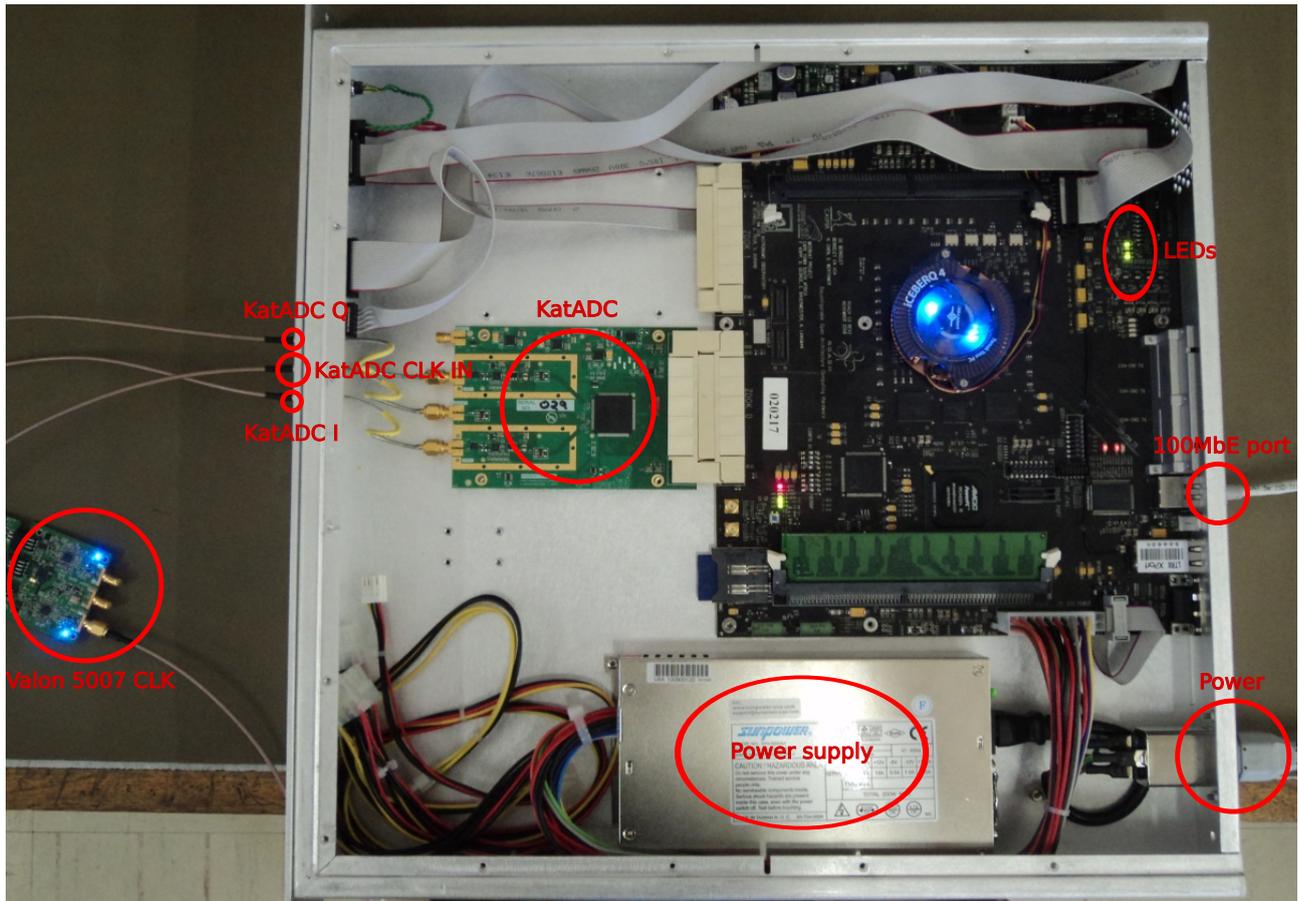


Figure D.1: ROACH board with the labels.

5. Supply power to the ROACH board using a standard kettle cord. The ROACH board is then powered on by briefly pressing the power button.

Additionally, the FX correlator has 4 user LEDs that are configurable from the FPGA, see figure D.1. The descriptions for the LEDs that were programmed are as follows:

- LED0 (left most) is wired to the sync pulse. This LED can be used to verify that the FPGA is clocking.
- LED1 is connected to the accumulation sync pulse. This obviously gives a visual indication of when a new accumulation starts.
- LED2 lights up when the ADC starts to clip.

Lastly, note that when powering down the ROACH board, the power button should be pressed once for only a short period. This initiates the ROACH board's own shutdown routine. Holding the power button in forces a sort of "hard" shutdown, this can corrupt the Linux filesystem.

D.2 Software configuration

The series of steps to follow in order to successfully configure the software of the FX correlator is discussed below.

1. The FX correlator design (boffile) must firstly be transferred to the ROACH board. The name of the boffile is *fx_corr_400mhz_1024.bof*. This step is not necessary if the ROACH board with serial number *020217* is used. Since, the correct boffile already appears on this ROACH board. The DNS name of this ROACH board is **gauss.cemagg.sun.ac.za**³ (*IP address: 146.232.221.192*). Also note, that this guide does not give instructions on how to setup the ROACH board. If the ROACH board must be configured, please refer to the CASPER wiki here: https://casper.berkeley.edu/wiki/Getting_Started_with_ROACH.

a) To transfer the boffile, open a terminal and enter the following command: `scp ./fx_corr_400mhz_1024.bof root@gauss.cemagg.sun.ac.za:/boffiles/fx_corr_400mhz_1024.bof`. For this command to work the user must be the superuser, if this is not the case then add `sudo` before the command. Furthermore, this command assumes that the current directory contains the specified boffile.

b) The boffile must be marked as executable before the FPGA can be programmed. This is done by entering the following command⁴:

```
chmod a+x ./fx_corr_400mhz_1024.bof
```

If this was successful, entering `ls -al` should print a list of available files with their permissions. There should appear *x*'s in the left most column next to the *fx_corr_400mhz_1024.bof* file name, similarly to

```
-rw-r--r--  1  root root  4470282 Jun 30 1961  fx_corr_1024nq_shr_2012_Aug_28_1821.bof
-rwxr-xr-x  1  root root  4469913 Jul 1 1961    fx_corr_400mhz_1024.bof
-rw-r--r--  1  root root  4469778 Jun 29 1961  fx_corr_400mhz_1024_2012_Jun_25_2150.bof
```

2. Before the data-logging and supervisory control script can be executed, ensure that the following dependencies are installed. The easiest way to install these dependencies is with the *easy_install* tool.

- Python 2.5 or later (Successfully tested with Python 2.6)
- numpy 1.6.2 (available at PyPI: <http://pypi.python.org/pypi/numpy>)
- katcp 0.3.4 (available at PyPI: <http://pypi.python.org/pypi/katcp/0.3.4>)
- construct 2.06 (available at PyPI: <http://pypi.python.org/pypi/construct/2.06>)

³The DNS name **gauss.cemagg.sun.ac.za** is used in this guide whenever a hostname is required. If the hostname of the ROACH board used is different, remember then to use that hostname instead.

⁴The user must remotely connect to the ROACH board using `ssh root@gauss.cemagg.sun.ac.za` and then change the current directory to `cd /boffiles/`

- iniparse 0.4 (available at PyPI: <http://pypi.python.org/pypi/iniparse/0.4>)
- spead 0.4.0 (available at PyPI: <http://pypi.python.org/pypi/spead/0.4.0>)
- corr 0.6.9 (available at PyPI: <http://pypi.python.org/pypi/corr>)
- bitstring 3.0.2 (available at PyPI: <http://pypi.python.org/pypi/bitstring/3.0.2>)
- pyephem 3.7.5.1 (available at PyPI: <http://pypi.python.org/pypi/pyephem>)
- matplotlib 1.1.1 (available at PyPI: <http://pypi.python.org/pypi/matplotlib/1.1.1>)

3. Open a terminal and browse to the directory containing the *dsc.py* script. At the prompt enter the following command: `python ./dsc.py --help`

This generates a description of how to execute the script and it furthermore prints a list of all the options that are configurable. The usage syntax for the script is shown to be `Usage: dsc.py < roach_hostname > [options]`. Replacing `< roach_hostname >` with the hostname of the ROACH board, which in this guide is **gauss.cemagg.sun.ac.za**. The acute brackets `< · >` indicates that this value must be specified. The `[options]` are optional, and if a parameter is not specified then the default value is used. A simple example is given of how to run the *dsc.py* script with options set for programming the FPGA with a specified boffile, the start- and stop-times, attenuator levels, accumulation length and log interval.

```
python ./dsc.py gauss.cemagg.sun.ac.za -p -b fx_corr_400mhz_1024.bof
--start_time=12:00:00 --stop_time=15:25:00 -a 15.5 -l 1 -i 5
```

This command reads as follows-

The ROACH board is programmed with the *fx_corr_400mhz_1024.bof* boffile, and the FX correlator is configured such that the attenuator levels of the ADC are set to 15.5dB, the accumulation length is 1 second. The observation will start at 12:00:00 local time and the spectral data is logged every 5 seconds, at 15:25:00 the observation will end and the *dsc.py* script will terminate. Note, that it is always a good idea to specify a start and stop time, and not to use the defaults. Furthermore, it is also useful to change the default file name to something more meaningful, using the `-d` option. The data file is written into the same directory containing the *dsc.py* script, this can not be changed. There are some other options available, but due to brevity these are not discussed here, but the `--help` option can be used to get an idea of the available options.

4. To terminate the *dsc.py* script safely during an observation, the user must press the *q* key. This will ensure that the data file is properly closed and that any hardware references are closed properly.

5. To view the measured data, the python script *data_analysis.py* must be used. The data file is in the *HDF5* format. Again, open a terminal and at the command prompt enter:

```
python ./data_analysis.py --help
```

This will again generate a list of options available and the syntax used to run the script. The usage syntax for the script is shown to be `Usage: data_analysis.py < file_name > [options]`. Replace `< file_name >` with the correct data file name, this value must be

specified. The options available for this script are `-i` and `-g`. The option `-i` specifies the integration time for the data being processed. The option `-g`, takes the following values: `autoi`, `autoq` and `crossiq`. The `autoi` and `autoq` parameters are used to display the PSD of input signals i and q , respectively. The parameter `crossiq`, displays the cross-correlation data between the two input channels against time.

Appendix E

Photos



Figure E.1: Photo of the helical beam antenna.

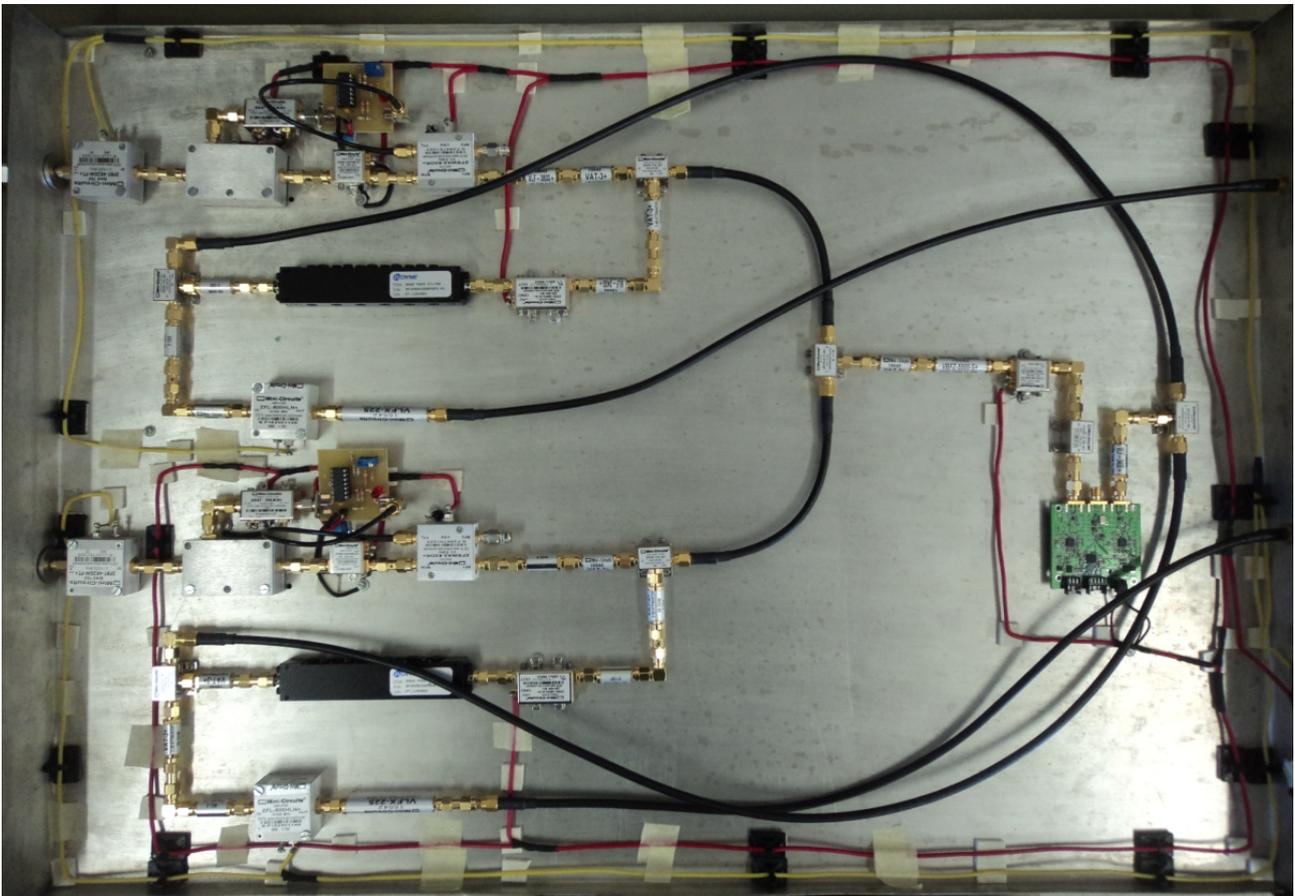


Figure E.2: Photo of the receiver board.



Figure E.3: Photo of the low-noise front-end units.

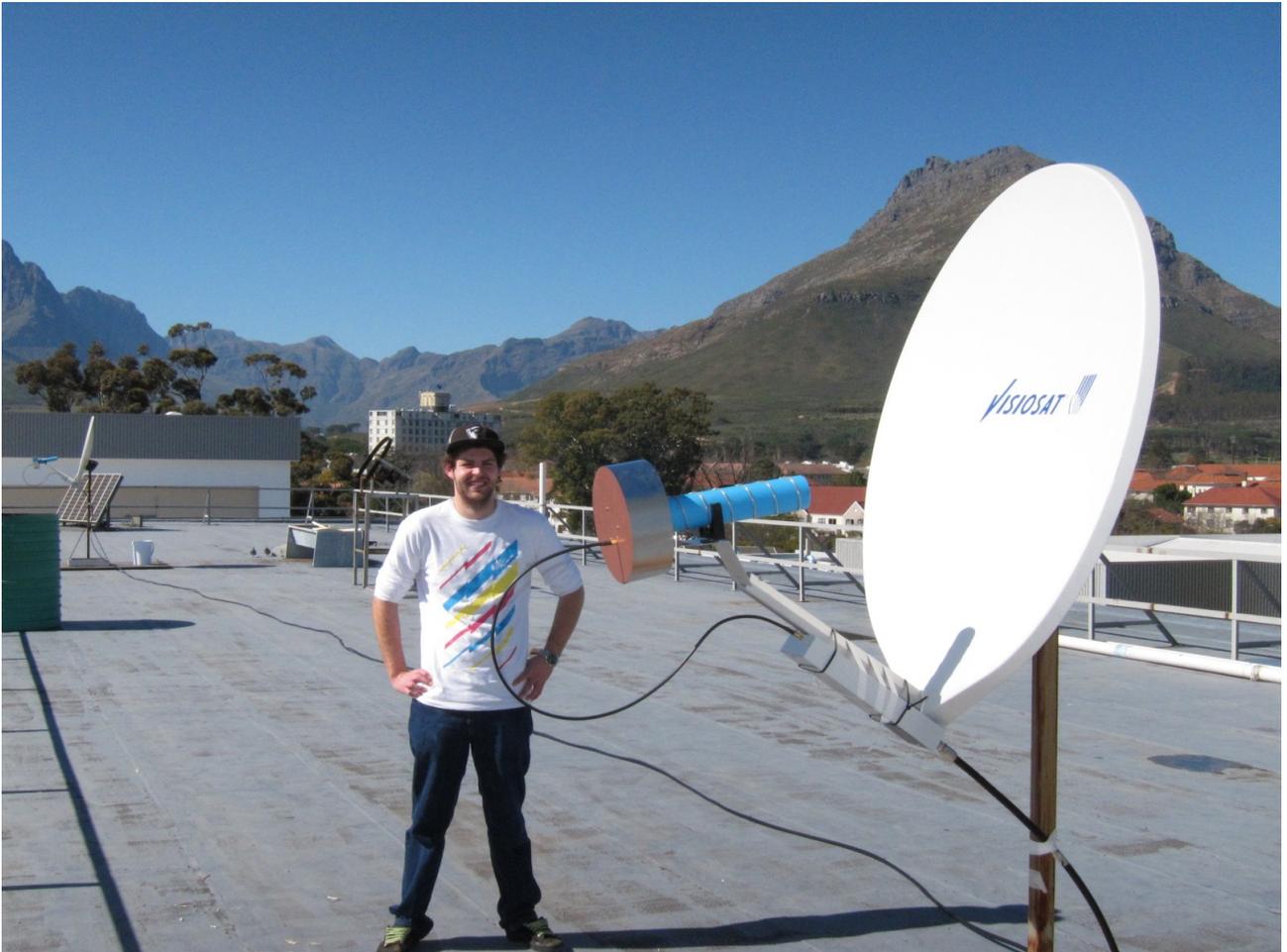


Figure E.4: Photo of the interferometer.

List of References

- [1] J. W. M. Baars, L. R. D’Addario, and A. R. Thompson, “Radio astronomy in the early twenty-first century,” *Proc. IEEE*, vol. 97, no. 8, pp. 1377–1381, Aug. 2009.
- [2] T. Wilson, “Techniques of radio astronomy,” *Astro-ph.IM*, Nov. 2011. [Online]. Available: <http://arxiv.org/abs/1111.1183>
- [3] J. C. Webber and M. W. Pospieszalski, “Microwave instrumentation for radio astronomy,” *IEEE Trans. Microw. Theory Tech.*, vol. 50, no. 3, pp. 986–995, Mar. 2002.
- [4] J. N. Chengalur, Y. Gupta, and K. S. Dwarakanath. (2003, May) Low frequency radio astronomy. [Online]. Available: http://gmrt.ncra.tifr.res.in/gmrt_hpage/Users/doc/WEBLF/LFRA/
- [5] P. E. Dewdney, P. J. Hall, R. T. Schilizzi *et al.*, “The square kilometre array,” *Proc. IEEE*, vol. 97, no. 8, Aug. 2009. [Online]. Available: [\[Online\]. Available: http://www.skatelescope.org](http://www.skatelescope.org)
- [6] P. J. Hall, *The Square Kilometre Array: An Engineering Perspective*. Dordrecht, Netherlands: Springer, 2005.
- [7] A. R. Taylor, “The square kilometre array,” *Proceedings IAU Symposium*, no. 248, 2007. [Online]. Available: [\[Online\]. Available: http://www.skatelescope.org](http://www.skatelescope.org)
- [8] J. Lazio. The square kilometer array. [Online]. Available: <http://www.skatelescope.org/>
- [9] J. L. Jonas, “MeerKAT - the south african array with composite dishes and wide-band single pixel feeds,” *Proc. IEEE*, vol. 97, no. 8, pp. 1522–1530, Aug. 2009.
- [10] D. R. DeBoer, R. G. Gough, J. D. Bunton, T. J. Cornwell *et al.*, “Australian SKA pathfinder: A high-dynamic range wide-field of view survey telescope,” *Proc. IEEE*, vol. 97, no. 8, pp. 1507–1521, Aug. 2009.
- [11] A. Gallerani and G. Minarelli, “A multiplying interferometer for 1.3GHz using two Yagi antennas,” *Measurement Science and Technology*, vol. 18, no. 2, May 2007.
- [12] K. Tapping, “A correlation interferometer for 408MHz,” *Journal of the Royal Astronomical Society of Canada*, Nov. 2001. [Online]. Available: <http://www.ukaranet.org.uk/projects/index.htm>

- [13] D. M. Lucero, “N²I²: The new mexico tech and nrao instructional interferometer,” *Physics Dept., New Mexico Institute of Mining and Technology, Socorro*, May 2003.
- [14] J. Kraus, *Radio Astronomy*, 2nd ed. Powell, Ohio: Cygnus-Quasar, 1986.
- [15] A. R. Thompson, J. M. Moran, and G. W. Swenson, Jr., *Interferometry and Synthesis in Radio Astronomy*, 2nd ed. Wiley-VCH, 2004.
- [16] H. Nyquist, “Thermal agitation of electric charge in conductors,” *Phys. Rev.*, vol. 32, pp. 110–113, 1928.
- [17] *Handbook of Frequency Allocations and Spectrum Protection for Scientific Uses*. Washington, DC: National Academies Press, 2007. [Online]. Available: <http://www.nap.edu/catalog/11719.html>
- [18] D. Miller, *Basics of Radio Astronomy for the Goldstone-Apple Valley Radio Telescope*. JPL, Jet Propulsion Laboratory, Apr. 1998.
- [19] T. L. Wilson, K. Rohlf, and S. Huttemeister, *Tools of Radio Astronomy*, 5th ed. Springer-Verlag, 2009.
- [20] R. Conway, K. Kellerman, and R. Long, “The radio frequency spectra of discrete radio sources,” *Mon. Not. Royal Astronomical Society*, vol. 125, pp. 261–284, 1963.
- [21] B. F. Burke and F. Graham-Smith, *An Introduction to Radio Astronomy*, 3rd ed. Cambridge, 2010.
- [22] J. Bolton, F. Gardner, and M. Mackey, “The parkes catalogue of radio sources declination zone -20° to -60° ,” *Aust. J. Phys.*, vol. 17, 1964.
- [23] G. Day, A. Shimmins, R. Ekers, and D. Cole, “The parkes catalogue of radio sources declination zone 0° to $+20^\circ$,” *Aust. J. Phys.*, vol. 19, 1965.
- [24] B. Mills, O. Slee, and E. Hill, “A catalogue of radio sources between declinations -20° and -50° ,” *Aust. J. Phys.*, vol. 11, no. 70, 1960.
- [25] J. Baars, P. Mezger, and H. Wendker, “The flux density values of standard sources used for antenna calibration,” *NRAO Electronics Division Internal Report*, no. 35, Aug. 1964.
- [26] J. Pawsey, “A catalogue of reliably known discrete sources of cosmic radio waves,” *The Astrophysical Journal*, vol. 121, no. 1, Jan. 1955.
- [27] H. Zirin, B. Baumert, and G. Hurford, “The microwave brightness temperature spectrum of the quiet sun,” *The Astrophysical Journal*, vol. 370, Apr. 1991.
- [28] J. Fielding, *Amateur Radio Astronomy*. Radio Society of Great Britain, 2008.
- [29] ICASA, “South african table of frequency allocations,” Government gazette, Tech. Rep. 31264, Jul. 2008.

- [30] G. W. Swenson., Jr., and N. C. Mathur, “The interferometer in radio astronomy,” *Proc. IEEE*, vol. 56, no. 12, pp. 2114–2130, Dec. 1968.
- [31] M. Ryle, “A new radio interferometer and its application to the observation of weak radio stars,” *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, vol. 211, no. 1106, pp. 351–375, Mar. 1952.
- [32] S. Ananthakrishnan, “Antennas for radio astronomy,” *Proc. IEEE*, 2009.
- [33] P. J. Napier, A. R. Thompson, and R. D. Ekers, “The Very Large Array: Design and performance of a modern synthesis radio telescope,” *Proc. IEEE*, vol. 71, no. 11, pp. 1295–1320, Nov. 1983.
- [34] M. A. Holdaway and T. T. Helfer, “Interferometric array design,” *ASP Conference Series*, vol. 180, pp. 537–563, 1999.
- [35] G. B. Taylor, C. L. Carilli, and R. A. Perley, *Synthesis Imaging in Radio Astronomy II*. ASP Conference Series, 1999, vol. 180.
- [36] B. Rowson, “High resolution observations with a tracking radio interferometer,” *Royal Astronomical Society*, vol. 125, no. 2, pp. 180–188, Sep. 1963.
- [37] R. H. Dicke, “The measurement of thermal radiation at microwave frequencies,” *The Review of Scientific Instruments*, vol. 17, no. 7, pp. 268–275, Jul. 1946.
- [38] M. Tiuri, “Radio astronomy receivers,” *IEEE Trans. Mil. Electron.*, pp. 264–272, Oct. 1964.
- [39] S. J. Goldstein, “A comparison of two radiometer circuits,” *Proceedings of the IRE*, pp. 1663–1666, Nov. 1955.
- [40] A. Papoulis, *Probability, Random Variables and Stochastic Processes*, 3rd ed. McGraw-Hill, Inc., 1991.
- [41] H. T. Friis, “Noise figures of radio receivers,” *Proceedings of the IRE*, vol. 32, no. 7, pp. 419–422, Jul. 1944.
- [42] D. M. Pozar, *Microwave Engineering*, 3rd ed. John Wiley & Sons, Inc., 2005.
- [43] M. Trier, “Radiometer receiver linearity characterization by the intermodulation distortion method,” *Instrumentation and Measurement Technology Conference, IMTC, IEEE*, pp. 99–102, 1993.
- [44] W. Egan, *Practical RF System Design*. John Wiley & Sons, Inc., 2003.
- [45] G. Bianchi, F. Perini, C. Bortolotti *et al.*, “ADC bit number and input power needed, in new radio-astronomical applications,” *13th Workshop on ADC Modelling and Testing*, Sep. 2008.

- [46] “National semiconductor ADC08D1520 low power, 8-bit, dual 1.5 GSPS or single 3.0 GSPS A/D converter,” Nov. 2009.
- [47] K. McClaning and T. Vito, *Radio Receiver Design*. Atlanta, GA: Noble Pub. Corp., 2000.
- [48] W. Stutzman and G. Thiele, *Antenna Theory and Design*, 2nd ed. John Wiley & Sons, Inc., 1998.
- [49] R. Parthasarathy, C. Frank, R. Treffers *et al.*, “A rooftop radio observatory: An undergraduate telescope system at the university of california at berkeley,” *Am. J. Phys*, vol. 66, no. 9, pp. 768–771, Sep. 1998.
- [50] J. F. Aurand, “A new design procedure for optimum gain pyramidal horns,” *IEEE Trans. Antennas Propag.*, vol. 46, no. 5, pp. 828–831, 1988.
- [51] C. A. Balanis, *Antenna Theory Analysis and Design*. John Wiley & Sons, Inc., 2005.
- [52] D. E. Isbell, “Log periodic dipole arrays,” *IRE Transactions on Antennas and Propagation*, vol. AP-8, pp. 260–267, May 1960.
- [53] R. L. Carrel, “The design of log-periodic dipole antennas,” *IRE Int. Conv. Rec.*, vol. 1, pp. 61–75, 1961.
- [54] R. H. DuHamel and D. G. Berry, “Logarithmically periodic antenna arrays,” *IRE Wescon Convention Record*, pp. 161–174, 1958.
- [55] R. B. Dybdal, “Defocusing loss for a log periodic-fed reflector,” *IEEE Trans. Antennas Propag.*, vol. AP-33, no. 7, pp. 809–812, Jul. 1985.
- [56] J. D. Kraus, “The helical antenna,” *Proceedings of the IRE*, pp. 263–272, Mar. 1948.
- [57] H. E. King and J. L. Wong, “Characteristics of 1 to 8 wavelength uniform helical antennas,” *IEEE Trans. Antennas Propag.*, vol. AP-28, no. 2, pp. 291–296, Mar. 1980.
- [58] A. R. Djordjević, A. G. Zajić, and M. M. Ilić, “Enhancing the gain of helical antennas by shaping the ground conductor,” *IEEE Antennas Wireless Propag. Lett.*, vol. 5, pp. 138–140, Jan. 2006.
- [59] A. W. Rudge and N. A. Adatia, “Offset-parabolic-reflector antennas: A review,” *Proc. IEEE*, vol. 66, no. 12, pp. 1592–1618, Dec. 1978.
- [60] V. Jamnejad-Dailami and Y. Rahmat-Sammi, “Some important geometrical features of conic-section-generated offset reflector antennas,” *IEEE Trans. Antennas Propag.*, vol. AP-28, no. 6, pp. 952–957, Nov. 1980.
- [61] M. Uhm, A. Shishlov, and K. Park, “Offset-paraboloid geometry: Relations for practical use,” *IEEE Antennas Propag. Mag.*, vol. 38, no. 3, pp. 77–79, Jun. 1996.
- [62] V. Wongpaibool, “Experiment on triangular copper-strip as impedance matching for 2.4-GHz helical antenna,” *Assumption University Journal of Technology*, vol. 11, no. 4, pp. 209–214, Apr. 2008.

APPENDIX E - PHOTOS

- [63] —, “Improved axial-mode-helical-antenna impedance matching utilizing triangular copper strip for 2.4-GHz WLAN,” *IEEE Trans. Antennas Propag.*, pp. 869–873, 2008.
- [64] W. L. Stutzman and M. Terada, “Design of offset-parabolic-reflector antennas for low cross-pol and low sidelobes,” *IEEE Antennas Propag. Mag.*, vol. 35, no. 6, pp. 46–49, Dec. 1993.
- [65] S. Sander and D. K. Cheng, “Phase center of helical beam antennas,” *IRE National Convention Record*, no. 1, pp. 152–157, 1958.
- [66] J. D. Kraus, “A 50-ohm input impedance for helical beam antennas,” *IEEE Trans. Antennas Propag.*, vol. AP-25, no. 6, p. 913, Nov. 1977.
- [67] S. Weinreb, “A digital spectral analysis technique and its application to radio astronomy,” Ph.D. dissertation, Department of Electrical Engineering, Massachusetts Institute of Technology, 1963.
- [68] L. R. D’Addario, “Cross correlators,” *ASP Conference Series*, pp. 59–82, 1989.
- [69] Y. Chikada, “Correlators for interferometry - today and tomorrow,” *ASP Conference Series*, vol. 19, pp. 26–32, 1991.
- [70] J. D. Bunton, “SKA correlator advances,” in *The Square Kilometre Array: An Engineering Perspective*, P. J. Hall, Ed. Springer, 2004.
- [71] —, “New generation correlators,” *Proceedings of the XXVIIth General Assembly of International Union Radio Science (URSI)*, Oct. 2005.
- [72] A. Parsons, D. Werthimer, D. Backer *et al.*, “Digital instrumentation for the radio astronomy community,” *Astro2010 Decal Survey Committee, TEC: Technology Development for RMS: Radio, Millimeter and Submillimeter from the Ground Discipline Program Panel*, Apr. 2009.
- [73] D. Werthimer, “The casper collaboration for high-performance open source digital radio astronomy instrumentation,” *IEEE Trans. Microw. Theory Tech.*, 2011.
- [74] D. Backer, “Quantization with four bits,” *EoR Experiment - Memo*, Mar. 2007. [Online]. Available: <https://casper.berkeley.edu/wiki/Memos>
- [75] A. Parsons, D. Backer, A. Siemion *et al.*, “A scalable correlator architecture based on modular FPGA hardware, reusable gateway, and data packetization,” *Astro-ph*, Mar. 2009.
- [76] CASPER. [Online]. Available: https://casper.berkeley.edu/wiki/Main_Page
- [77] J. Chennamangalam, “The polyphase filter bank technique,” *CASPER Memo 41*, Aug. 2011.
- [78] H. Chen, P. McMahon, and A. Parsons. (2008, Aug.) Sync pulse usage in CASPER DSP blocks. [Online]. Available: <https://casper.berkeley.edu/wiki/Memos>

- [79] D. Laney. (2009, November) Approximate position of the sun (altitude and azimuth) from any location at any time (for low accuracy calculation). [Online]. Available: <http://www.sao.ac.za/public-info/sun-moon-stars/sun-index/how-to-calculate-altaz/>
- [80] B. P. Lathi, *Modern Digital and Analog Communication Systems*, 3rd ed. Oxford University Press, 1998.