

# Applied Radio Frequency Interference Removal in Radio Astronomy

Jonathon Kocz

A thesis submitted for the degree of

Doctor of Philosophy

of The Australian National University



THE AUSTRALIAN NATIONAL UNIVERSITY

Research School of Astronomy and Astrophysics

June 2009

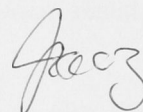


## Disclaimer

I hereby declare that the work in this thesis is that of the candidate alone, except where indicated below or in the text of the thesis.

The base code used for reading, correlating and correcting the raw data in Chapter 2 was written in Fortran by Frank Briggs, and adapted by the candidate. The correlation code for the Murchison Widefield Array data used in Chapter 5 was also written by Frank Briggs, although all subsequent processing software for this data was written by the candidate.

All other software for reading and processing data was written by the candidate in MATLAB<sup>®</sup> and Fortran.



Jonathon Kocz

11th June 2009



# Acknowledgements

There are several people I have become indebted to over the course of this thesis. First and foremost I must express my appreciation to my supervisor, Prof. Frank Briggs. It has been a pleasure to work with you over the last few years, and I look forward to doing so again in the future.

In no particular order, I would also like to thank:

- The Australian National University and Research School of Astronomy and Astrophysics for providing funding during this research.
- Mark Waterson and Andre de Gans for the sage technical advice, tips and tricks.
- Tim Ruckley and Brett Preisig at Parkes Observatory for all your assistance in building and connecting up new hardware.
- John Reynolds for continually finding me observing time in the Parkes telescope maintenance schedule.
- Haley, Josh, Nic, Grant, Arved and Mira for the late night discussions, encouragement and humour.

Finally, this thesis would have taken much longer to complete without the unstinting support of Amanda. Thank you, for more reasons that it is possible to list here.



# Abstract

This thesis addresses the practical application of Radio Frequency Interference (RFI) excision algorithms to observational radio astronomy. The research emphasises spatial filtering methods that are common to mathematical signal processing and the use of reference antennas to isolate RFI signals from the astronomical data.

RFI is a growing problem in radio astronomy. With the advent and proliferation of wireless communication systems the portions of the radio spectrum that can be observed without interference are being progressively limited.

Interference subtraction with a reference antenna and spatial filtering are two methods that attempt to recover the astronomical signal drowned out by interference, instead of blanking the contaminated portion of the spectrum and discarding the scientific information that it might contain.

This thesis explores RFI subtraction using a reference antenna through the application of closure relations to gain an accurate deterministic coupling of the interfering signal to the contaminated astronomical data so that it can be subtracted away. Subtraction using reference antennas is performed on contaminated data from the Parkes radio telescope to illustrate the effectiveness of this method.

Spatial filtering, traditionally used on telescope arrays, is performed on data from the Parkes Multibeam Receiver, demonstrating for the first time that it is possible to use spatial filtering on an array feed receiver on a single dish.

Spatial Filters rely on the relative arrival times of a signal at multiple sensors to identify and separate signals from different directions. The method is a consequence of the Karhunen-Loève Theorem and relies on the eigen decomposition of the covariance matrix formed from the multiple signal paths.

Both reference antenna and spatial filtering methods are found to be effective at removing the interference from data. This is demonstrated on several different pulsar observations.

In the case of spatial filtering, particular attention is paid to the risk of removing astronomical signal-of-interest along with the interference. The potential to implement the algorithm as a filter is also explored, both in order to reduce computational requirements, and to develop the ability to apply corrections on shorter timescales than would be possible if computing the correction for each integration time.





---

# CONTENTS

Disclaimer . . . . .	i
Acknowledgements . . . . .	iii
Abstract . . . . .	v
Nomenclature . . . . .	xvii
Acronyms . . . . .	xix
<b>1 Introduction</b>	<b>1</b>
1.1 Types of observations and the effect of RFI . . . . .	1
1.2 RFI elimination techniques . . . . .	2
1.2.1 Subtraction using a reference antenna . . . . .	5
1.2.2 Spatial filtering . . . . .	6
1.3 Principal Component Analysis . . . . .	7
1.4 Signal processing and FPGAs in radio astronomy . . . . .	7
1.5 Thesis outline . . . . .	8
<b>2 Interference subtraction</b>	<b>9</b>
2.1 RFI subtraction using reference antennas . . . . .	9
2.2 Parkes 50 cm CPSR2 data . . . . .	12
2.2.1 RFI environment . . . . .	12
2.2.2 RFI subtraction . . . . .	13
2.2.3 Pulsar detection . . . . .	18
Pulsar J0437-4715 . . . . .	18

	Pulsar J0630-2834 . . . . .	20
2.2.4	Subtraction effects on data . . . . .	21
	Effects of subtraction on 50 cm Pulsar observations . . . . .	22
2.3	Conclusions . . . . .	24
<b>3</b>	<b>Spatial filtering</b>	<b>27</b>
3.1	Phased arrays and covariance matrices . . . . .	27
3.2	Singular Value Decomposition . . . . .	29
3.3	Spatial filtering implementation options . . . . .	30
3.3.1	Expected Performance . . . . .	35
3.4	Data calibration requirements . . . . .	35
3.5	Conclusions . . . . .	39
<b>4</b>	<b>Spatial filtering using the Parkes Multibeam Receiver</b>	<b>41</b>
4.1	Experiment outline . . . . .	41
4.2	Instrumentation . . . . .	42
4.2.1	Data processing software . . . . .	47
4.3	Results . . . . .	48
4.3.1	RFI environment . . . . .	48
4.3.2	Interference removal . . . . .	48
	Reference antenna . . . . .	52
4.3.3	Projection variations and dealing with strong astronomy sources . . . . .	52
4.4	Applying the projection as a filter for pulsar detection . . . . .	64
4.4.1	Interference-free region . . . . .	66
4.4.2	Interference contaminated region . . . . .	70
4.4.3	Summary . . . . .	71
4.5	Future work . . . . .	72
4.6	Conclusions . . . . .	73
<b>5</b>	<b>MWA 32 T data</b>	<b>75</b>
5.1	Data acquisition and hardware . . . . .	76
5.2	Hydra A observations . . . . .	78
5.3	Solar observations . . . . .	83
5.4	Conclusions . . . . .	88

<b>6</b>	<b>Signal processing and Principal Component Analysis</b>	<b>89</b>
6.1	Passband calibration techniques . . . . .	90
6.2	Introduction to Principal Component Analysis . . . . .	92
6.2.1	Method applied to the Parkes 50 cm dataset . . . . .	92
6.3	Applications to feature detection demonstrated with the Parkes 50 cm data . . . . .	93
6.3.1	Simulation Results . . . . .	94
6.3.2	Semi-Simulated Results . . . . .	95
6.3.3	Real Data . . . . .	97
6.4	Conclusions . . . . .	99
<b>7</b>	<b>Conclusions and Future Work</b>	<b>101</b>
7.1	Main results . . . . .	101
7.2	Future work . . . . .	103
7.3	Finally . . . . .	104
	<b>Appendices</b>	<b>105</b>
<b>A</b>	<b>FPGAs in radio astronomy</b>	<b>107</b>
A.1	An FPGA spectrometer . . . . .	107
A.1.1	Hardware selection . . . . .	107
A.1.2	Development tool selection . . . . .	108
A.1.3	Design . . . . .	108
A.1.4	Implementation . . . . .	109
A.1.5	Results & Conclusion . . . . .	111
<b>B</b>	<b>PCB and Schematic Diagrams</b>	<b>115</b>
<b>C</b>	<b>ATIM Design for MWA</b>	<b>121</b>
<b>D</b>	<b>Observation Data</b>	<b>123</b>
	<b>Bibliography</b>	<b>126</b>



---

# LIST OF FIGURES

1.1	RFI environment at Sydney, Narrabri and Mileura . . . . .	5
2.1	Interference subtraction illustration . . . . .	10
2.2	Location and type of reference antennas relative to Parkes 64 m dish	13
2.3	RFI example for different dish orientations at Parkes . . . . .	14
2.4	Total power spectra of J0437-4715 observation pre- and post-RFI subtraction . . . . .	15
2.5	Total power spectra of J0630-2834 observation pre- and post-RFI subtraction . . . . .	15
2.6	Total power spectra of zenith observation pre- and post-RFI subtraction . . . . .	16
2.7	Total power spectra of J0437-4715, J0630-2834 and zenith position observations pre- and post-RFI subtraction, channels 1000-4000 . .	17
2.8	Reverse subtraction results for J0437-4715 observations . . . . .	18
2.9	Time and frequency detection of pulsar J0437-4715 . . . . .	19
2.10	Comparison of time and frequency detection of pulsar J0437-4715 pre- and post-correction in interference-free region of spectrum .	20
2.11	Comparison of time and frequency detection of pulsar J0437-4715 pre- and post-correction a region of spectrum contaminated by interference . . . . .	21
2.12	Detection of pulsar J0630-2834 . . . . .	22
2.13	Detection of pulsar J0630-2834 in interference-free subset of bandwidth . . . . .	23

2.14	Detection of pulsar J0630-2834 in interference contaminated subset of bandwidth . . . . .	24
3.1	Example of a simple telescope array. . . . .	28
3.2	Total power spectra for one polarisation of a reference antenna connected to the ATCA receiving system and one telescope input . . . . .	36
3.3	Gain calibration effects on eigenvalue effectiveness . . . . .	37
3.4	Pre- and post-correction spectra for a reference antenna connected to the ATCA receiving system, and one telescope input, in the presence of wideband interference . . . . .	38
4.1	Beam layout for the Parkes Multibeam Receiver . . . . .	42
4.2	Signal chain for each polarisation of each receiver feed . . . . .	43
4.3	Custom Multibeam recording system dataflow . . . . .	45
4.4	Top side of populated converter circuit board . . . . .	46
4.5	Bottom side of populated converter circuit board . . . . .	46
4.6	Converter circuit boards in Parkes chassis . . . . .	47
4.7	RFI variation demonstration in Parkes dataset . . . . .	49
4.8	Autocorrelation power spectrum example from Parkes dataset . . . . .	50
4.9	Interference free power spectrum example from Parkes dataset . . . . .	50
4.10	Eigenvalues from section of Parkes dataset . . . . .	51
4.11	Power spectrum snapshot from observations of Vela pulsar at Parkes with varying corrections applied . . . . .	55
4.12	Eigenvalues from a power spectrum snapshot from observations of Vela pulsar at Parkes . . . . .	56
4.13	A comparison of the Vela pulsar detection strength in RFI unaffected spectral channels after applying different RFI correction methods . . . . .	57
4.14	A comparison of the Vela pulsar detection strength in RFI affected spectral channels after applying different RFI correction methods . . . . .	59
4.15	First 3 eigenvalues for summed over the 300-450 spectral channel range . . . . .	60
4.16	First 3 eigenvalues for summed over the 100-250 spectral channel range . . . . .	60
4.17	Eigenvector coefficient one: $u_{11}$ and $u_{18}$ as a function of frequency channel for timesteps 231-238, for two of the Multibeam Receiver beams. One beam is observing blank sky, the other the Vela pulsar. . . . .	61

4.18	Eigenvector coefficient two: $u_{21}$ and $u_{28}$ as a function of frequency channel for timesteps 231-238, for two of the Multibeam Receiver beams. One beam is observing blank sky, the other the Vela pulsar.	62
4.19	Eigenvector coefficient three: $u_{31}$ and $u_{38}$ as a function of frequency channel for timesteps 231-238, for two of the Multibeam Receiver beams. One beam is observing blank sky, the other the Vela pulsar.	63
4.20	Example of all information being contained in a single eigenvalue for each frequency channel . . . . .	65
4.21	Filter results from the application of the noise space projection to the Vela pulsar data . . . . .	67
4.22	Pulsar detection using spectral channels 300-450 . . . . .	68
4.23	Pulsar detection using spectral channels 100-250 . . . . .	69
4.24	Comparison of the detection of Vela after using short timescale corrections and correcting using a long timescale filter in the 300-450 channel range . . . . .	70
4.25	Comparison of the detection of Vela after using short timescale corrections and correcting using a long timescale filter in the 100-250 channel range . . . . .	72
5.1	MWA 32 T prototype receiving system interior . . . . .	77
5.2	MWA Hydra A total power spectra at 261.12 MHz, pre-correction	79
5.3	MWA Hydra A observations, eigenvalue matrix . . . . .	80
5.4	MWA Hydra A total power spectra at 261.12 MHz, correction comparison . . . . .	81
5.5	MWA Hydra A total power spectra at 261.12 MHz, post-correction	82
5.6	MWA Solar observations at 261.12 MHz, pre-correction . . . . .	84
5.7	MWA Solar observations, eigenvalue matrix . . . . .	85
5.8	MWA Solar observations at 261.12 MHz, correction comparison .	86
5.9	MWA Solar observations at 261.12 MHz, post-correction . . . . .	87
6.1	Simple illustration for the purposes of explaining gain calibration	90
6.2	Detection of 1% absorption line using principal components . . .	94
6.3	Detection of 0.3% absorption line with principal components using an unsimulated bandpass . . . . .	95
6.4	Difference in feature detection using a variable number of principal components . . . . .	96
6.5	Example of principal components for a spectrum . . . . .	97

6.6	Detection of 1% absorption line using principal components in the presence of RFI . . . . .	98
6.7	Detection of 0.3% absorption line using principal components in the presence of RFI . . . . .	99
A.1	Xilinx® ML402 development board. . . . .	108
A.2	FFT Spectrometer design for FPGA . . . . .	109
A.3	System Generator example . . . . .	110
A.4	FFT implementation in Simulink . . . . .	110
A.5	Serial link VHDL black box . . . . .	111
A.6	Spectrometer output from a test sine wave with a frequency of 10 MHz . . . . .	112
A.7	Spectrometer output from a test square wave with a frequency of 1 MHz . . . . .	112
B.1	Trigger and clock input schematic for Multibeam converter board	116
B.2	Data input and conversion schematic for Multibeam converter board	117
B.3	Decoupling capacitors schematic for Multibeam converter board .	118
B.4	Circuit layout for Multibeam converter board . . . . .	119
C.1	Schematic for the ATIM . . . . .	122



---

# LIST OF TABLES

4.1	Bit mapping for 1-32 bit stream recorded by the VSIB recorders . . . . .	47
4.2	Vela single pulse SNR for different correction methods . . . . .	58
4.3	Vela single pulse SNR for different filtering methods . . . . .	67
B.1	Circuit elements and their mapping to the circuit board . . . . .	120
D.1	Observational data times and receiver details . . . . .	124
D.2	Observers and sources . . . . .	125



---

# NOMENCLATURE

$\langle \cdot \rangle$	Time average
$(\cdot)^H$	Conjugate (Hermitian) transpose
$\vec{\cdot}$	Vector
$\bar{\cdot}$	Average
$\otimes$	Correlation
<b>A</b>	Boldface capitals indicate a matrix
<b>P<sub>I</sub></b>	Interference subspace projection matrix
<b>P<sub>N</sub></b>	Noise subspace projection matrix
<b>C</b>	Covariance matrix
$j$	$\sqrt{-1}$



---

# ACRONYMS

ACMA	Australian Communications and Media Authority
ANU	The Australian National University
ASIC	Application Specific Integrated Circuits
ATCA	Australia Telescope Compact Array
ATIM	Antenna Time Interface Module
ATNF	Australia Telescope National Facility
CPLD	Complex Programmable Logic Device
FFT	Fast Fourier Transform
FPGA	Field Programmable Gate Array
INR	Interference to Noise Ratio
LVDS	Low Voltage Differential Signalling
MECL	Motorola Emitter Coupled Logic
MWA	Murchison Widefield Array
PCA	Principal Component Analysis
PCB	Printed Circuit Board
PECL	Positive Emitter Coupled Logic
PFB	Polyphase Filter Bank
RRI	Raman Research Institute
RFI	Radio Frequency Interference
SNR	Signal to Noise Ratio
SKA	Square Kilometre Array
SVD	Singular Value Decomposition
UART	Universal Asynchronous Receiver/Transmitter
VHDL	VHSIC Hardware Description Language
VHSIC	Very High Speed Integrated Circuits
VLBI	Very Long Baseline Interferometry
VSI	VLBI Standard Interface
VSIB	VLBI Standard Interface B

# CHAPTER 1

---

## INTRODUCTION

This thesis concerns the issue of Radio Frequency Interference (RFI) in Radio Astronomy.

RFI is a growing problem in radio astronomy. With the advent and proliferation of wireless communication systems the portions of the radio spectrum that can be observed without interference are being progressively limited.

RFI can occur in several different ways. These range from terrestrial signals, which will be designated as any Earth-based signal (for example, signals from a nearby mobile phone or digital television transmitter) to signals generated by satellites passing through the area of interest in the sky. As eloquently put forth by Fred Harris, “One person’s trash is another person’s treasure (one person’s interference is another person’s signal)” (Bell et al., 2000). Following this theme, for the duration of this thesis the term “RFI” will refer to any undesired signal in the frequency band of interest.

This chapter introduces RFI in the context of radio astronomy. Several different observation types, the effect of RFI on those observations, and some of the different ways the literature suggests RFI can be characterised and excised are considered. The background presented is focused on the two specific methods applied in this thesis: RFI subtraction with a reference antenna and spatial filtering.

Finally, an outline is presented for the remainder of the thesis.

### **1.1 Types of observations and the effect of RFI**

There are several different types of observations in radio astronomy, as well as several different types of RFI. Each combination of the type of observation and the type of RFI has a different effect on the data. Because of this, it is important

to note that there is no one RFI excision method that will work equally well in all cases. The effect of RFI and the method used to correct it depends on the type of observations, the instrumentation used, and often the availability of computational capacity. Different observation types include spectral line or continuum observations in the case of single dish, phased array, and Very Long Baseline Interferometry (VLBI) observations. The effects of RFI can be broad or narrowband in nature, transient and intermittent, and varying in strength. Time variable astronomical phenomena, such as pulsars, add another dimension of complexity. When doing continuum observations, for example, broadband RFI will raise the effective noise level of receivers, degrading their sensitivity. Narrowband RFI may imitate spectral lines (Fridman, 2001a).

The strength of the RFI also plays a role in any attempt to remove it. Strong RFI may saturate receivers, driving the signal into a non-linear regime and making the recovery of any astronomical signal underneath impossible. Weak RFI may be difficult to identify, and therefore remove, from the signal.

Just considering these simple differences, almost fifty different scenarios for RFI correction are created. Greater subtleties ensue when the exact nature of the observations are taken into account.

Single dish observations are perhaps the most vulnerable to RFI. Single dish observations typically measure the total power of the receiver signal, and because of this any RFI present in the input will be included in the output (Fridman, 2001a; Mitchell, 2004).

Interferometers, in the case of a stationary RFI source, are less affected (Thompson, 1982). The rapid phase winding of the RFI signal in the interferometer visibility discriminates against the RFI in time averages. This is effectively a form of spatial filtering (see §3). In VLBI observations, there is little chance of the same interfering signal being received by multiple receivers. This creates a substantial resistance to RFI, but the presence of the interferer will still limit the sensitivity of the local receiver by increasing the system noise (Thompson, 1982). Short baseline interferometers can be almost as vulnerable as a single dish to RFI.

An overview of some of the different types of RFI characterisation and removal mechanisms is given in the following section §1.2.

## 1.2 RFI elimination techniques

There are many different methods for both determining the presence of RFI in an observation (characterisation), and also for removing it (excision). These include both online techniques, where the correction is applied in real-time as the signal is received, detected and integrated, and offline techniques, with the correction occurring after the signal has been recorded.

The methods available also depend on the system being used for observations. Existing methods for single dish receivers, for example, include the use of reference antennas to record the RFI signals. These recordings can be used to assist in the creation of a real-time adaptive filter (Barnbaum & Bradley, 1998; Mitchell et al., 2002), or for offline subtraction (Briggs et al., 2000; Briggs & Kesteven, 2002; Mitchell et al., 2002).

For array observations, as in the single dish case, it is also theoretically possible to make use of a reference antenna for offline subtraction (Briggs et al., 2000).

Other methods that are especially well suited to observing with an array include a range of spatial filtering techniques (Leshem et al., 2000; Boonstra et al., 2002; Raza et al., 2002). These spatial filtering techniques can further be divided into two types, *deterministic* and *non-deterministic* nulling. Deterministic nulling involves pre-configuring an array to ignore signals from a specific location. This, however, requires the location of the RFI source to be known beforehand, and the system to be calibrated with high precision. Non-deterministic nulling is accomplished by adaptively steering the nulls of an array toward an interference source during operation. As mentioned in §1.1, the spacing of antennas in an interferometer results in the decorrelation of the interfering signals, forming a natural method of RFI suppression.

Ellingson & Cazemier (2003) explore the theory of using Fast Fourier Transform (FFT) beamforming as an efficient synthesis method when an array is capable of generating multiple beams on the sky. They also explore the use of spatial filtering in this situation.

When the interferer characteristics are known, it is also possible to make a model of the signal to subtract. This is demonstrated by Ellingson et al. (2000) using GLONASS satellite interference.

Fridman (2001b) investigates the use of statistics, analysing the probability distribution of the power spectra, and taking advantage of the fact it should be Gaussian in nature to characterise and remove the RFI.

In some situations for either arrays or single dishes, where the time-frequency variability of the RFI is characterised, the RFI can then be blanked and not included at all in any of the future signal processing steps. Jeffs et al. (2006), for example, uses statistics to help determine the presence, and predict the location of, RFI at the Greenbank telescope caused by radar reflections from aircraft and blank these regions of the spectrum.

However, this may defeat the purpose of the “excision” portion of “radio frequency characterisation and excision”, since, ideally, the goal is to remove the interference and recover the astronomical signal underneath. Also, while it may be acceptable to sacrifice some portion of the bandwidth in continuum observations, if there is RFI in the region of interest in a spectral line observation, blanking that portion of the spectrum may be an unacceptable solution.



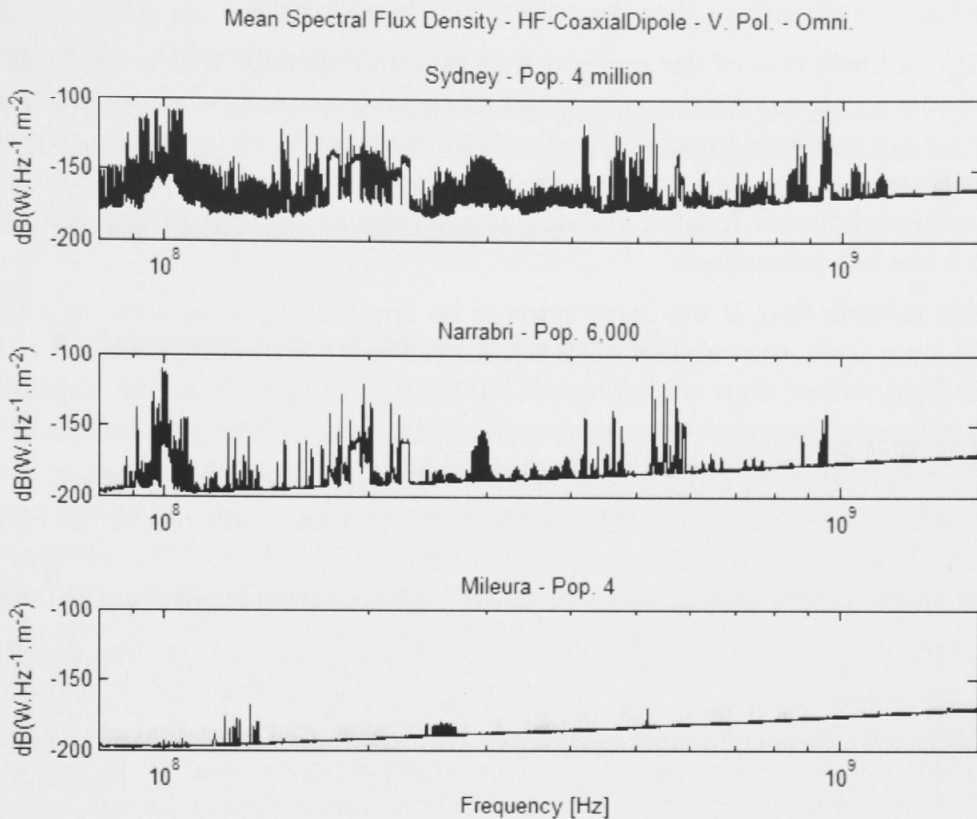
RFI characterisation and excision algorithms are in continual development. Recent publications by Athreya (2009) and Eatough et al. (2009) demonstrate this, and this thesis explores novel applications. Athreya (2009) has developed a method for identifying “quasi-constant” sources of RFI (from transmitters that are fixed to the surface of the Earth), based on the fact that the phase of the RFI does not change, since the RFI enters the interferometer at “zero fringe rate” (see for example Thompson (1982)). Separating the signal from the Earth-based source on this basis leads to the formulation of an RFI estimate, which is then subtracted from the data. Eatough et al. (2009) use the fact that broadband RFI is not dispersed by propagation through the Interstellar Medium (as it is generally “local”) as a discriminant to separate the RFI from pulsar signals, which are dispersed.

It is becoming highly desirable to implement a reliable system of real-time RFI removal. This suffers a disadvantage compared to post-processing RFI removal, where the original data remains intact and any processing can always be undone should errors in removal occur. With the new generation of telescopes under development, however, the input data rates will be such that it may not be possible to keep the raw data in all cases. It is desirable to remove the RFI as early as possible in the signal processing chain, and so a real-time removal system that can be trusted to leave the celestial data intact will be required.

One last approach, that does not require any processing capability at all, and is still one of most successful, is simply isolation of the telescope. This is demonstrated well by a plot created by Chippendale & Beresford (2005) in figure 1.1. This figure shows the RFI environment for Sydney, Narrabri (the location of the Australia Telescope Compact Array (ATCA)), and Mileura, one of the sites considered in Australia for the Square Kilometre Array (SKA), and the initial site of the Murchison Widefield Array (MWA) (Lonsdale et al., 2009; Bowman et al., 2007). This plot shows that by putting the telescope in a low interference environment, much effort on RFI removal can be saved.

A starting condition for any new radio astronomy facility will be situating it at the most radio-quiet location possible. This way, the RFI excision processing can be devoted to eliminating contamination such as satellite downlinks, that cannot be simply avoided anywhere on Earth. Of course, it is not always possible to do this, as a radio quiet location in the desired frequency range may be too impractical, nor would it be desirable to move already established observatories.

Two of the more generally applicable techniques are addressed in this thesis: subtraction with a reference antenna and spatial filtering. A more detailed description of the current state and accomplishments of these two techniques will be given in §1.2.1 and §1.2.2 respectively.



**Figure 1.1** RFI environment at Sydney, Narrabri and Mileura. A radio quiet location for the telescope makes a large difference to the amount of RFI that needs to be characterised and excised. Shown courtesy of A. P. Chippendale and R. Beresford (taken as part of the CSIRO ATNF SKA Site Monitoring Program).

### 1.2.1 Subtraction using a reference antenna

Barnbaum & Bradley (1998) introduced the concept of using a reference antenna to receive an interfering signal in parallel with telescope data so that it could be subtracted from the data. The specific variation pursued in this thesis is that introduced by Briggs et al. (2000), where closure relations (Cornwell, 1989; Thompson et al., 2001) are used to create a reference signal to be subtracted from the telescope receiver signal.

After the data is recorded by the reference antenna and telescope receivers, a covariance matrix (see §2.1 for the mathematical description) is formed. The cross power elements of this matrix that combine the reference and telescope receivers are used to form the closure relations used for subtraction. With enough computing power, or specific purpose hardware (for example a Field Programmable Gate Array (FPGA) processing system), it would be possible to do this in real-time. This closure-relation approach is a post-correlation method and in this thesis it is applied offline. In general, performing the subtraction

offline has the advantage that the subtraction is optional.

Note it is a limitation of the method that only interference that is captured by the reference antenna can be subtracted. If the reference antenna is directional, and there are multiple interferers, only those signals picked up by the reference antenna can be cancelled. However, this also adds an element of safety, as astronomical sources located outside the reception pattern of the reference antenna are left untouched.

It is also notable that, if the corrections to be applied are computed on a short enough time scale, the subtraction is equivalent to a coherent cancellation of the electric field, rather than an incoherent subtraction of powers, as the correlation function retains the information pertaining to the phases between the RFI in the reference antenna and the telescope receiver (Briggs et al., 2000).

This method has previously been successfully applied to RFI test datasets (see Bell et al. (2000)), and was also successfully used to remove interference from the dataset when a simulated galaxy was added, without harming the galaxy signal.

RFI subtraction using a reference antenna in this way to aid in pulsar detection is the focus of Chapter 2. One difference, however, is that the data used in this thesis are from two, single polarisation reference antennas that are spatially separated, instead of the single, dual polarisation horn used in Briggs et al. (2000). This has the effect of removing correlated noise from cross talk between the nominally orthogonal polarisations of the horn. This additional correlated power had formed a limitation in the earlier experiments conducted by Briggs et al. (2000).

## 1.2.2 Spatial filtering

In reference to the radio frequency signals incident on radio interferometer arrays, Thompson (1982) states that “the three components of the visibility cannot be separated”. The three components to which he refers in this statement are celestial signal, noise (for example  $T_{sys}$ ), and interference. Spatial filtering theory, used for the majority of the experiments in this thesis (Chapters 3, 4, 5), defies this statement. The specific mathematics and implementation details of spatial filtering will be discussed in these respective chapters.

Spatial filtering is not a new concept in engineering terms (see, for example, Hung & Turner (1983)). However, it has only recently been applied in detail to radio astronomy by Leshem et al. (2000). Spatial filtering is a technique that can be used with multi-feed systems, such as an array of antennas or a multibeam receiver. Using the cross-correlation information from these independent sensors, the “spatial signature” of interfering signals can be determined. It is then possible to create a vector projection that removes signals originating from that spatial signature. This is a technique known as subspace projection. In the case of a

telescope array, the spatial signature can normally be mapped to the direction of arrival of the interfering signal wavefront.

Spatial filtering in this way has been successfully applied to phased array systems (e.g. Leshem et al. (2000); Kocz (2004); Briggs & Kocz (2005); Boonstra (2005)), however to date it has not been applied to a multibeam system on a single dish. This thesis addresses that gap in Chapter 4.

### 1.3 Principal Component Analysis

While not technically Principal Component Analysis (PCA) by definition, many of the processes involved in spatial filtering rely on similar mathematics. A covariance matrix is computed, then decomposed into its eigenvector and eigenvalue components. In PCA, the first few eigenvector and eigenvalue pairs are known as the principal components. This can be compared to spatial filtering, where the eigenvalues represent signal power, and the eigenvectors the coupling of that power to the various antenna input feeds. As the interference is generally orders of magnitude stronger than the astronomical signal of interest, the first few eigenvectors and eigenvalues (the principal components of the system) represent the interference. Using PCA for bandpass calibration and feature detection is discussed in Chapter 6.

### 1.4 Signal processing and FPGAs in radio astronomy

The field of radio astronomy is closely tied to that of communications and electronics. This is evidenced in the number of papers published in engineering journals that are applicable to radio astronomy, or indeed are focused on it. These journals include: IEEE Transactions on Antennas and Propagation, IEEE Transactions of Information Theory, IEEE Transactions on Aerospace and Electronic Systems and IEEE Transactions on Signal Processing. Early publications on spatial filtering theory that is used for much of this thesis appeared in IEEE Transactions on Aerospace and Electronic Systems (see Hung & Turner (1983)).

In addition to communications and algorithm development, the electronics used for data processing in radio astronomy have also undergone development in recent times. FPGA chips have become powerful enough to allow them to start replacing Application Specific Integrated Circuits (ASIC) (for example, Stanko et al. (2005)). An overview of the situation and the potential of FPGAs is provided by Parsons et al. (2005). It is expected that FPGAs will play a large role in implementing real-time RFI excision in the future. To gain some further experience with the devices in an astronomy context, a simple FPGA spectrometer was developed as an exercise using a Xilinx® development board

and MATLAB<sup>®</sup> System Generator software. A brief discussion of FPGAs and an outline of the spectrometer development process are given in Appendix A.

## 1.5 Thesis outline

The principal goal of this thesis is to test RFI cancellation techniques in a range of radio astronomical applications. This leads to an assessment of the relative merits of the different techniques and a determination of the requirements for applicability of the different approaches.

Specifically, Chapter 2 deals with the mathematics and results of practical application of RFI subtraction using a reference antenna for the purpose of pulsar observations. Data from the Parkes 50 cm receiver is used to examine the RFI environment and perform RFI subtraction.

Chapter 3 introduces the theory of spatial filtering and Chapter 4 applies the spatial filtering theory of Chapter 3 to the Parkes Multibeam Receiver. The effects of spatial filtering on astronomical data are examined, and the ability of spatial filtering to remove interference is shown. An actual “filter” implementation is then demonstrated. Chapter 5 applies the spatial filtering theory of Chapter 3 to data taken with the MWA, showing some of the methods limitations.

Chapter 6 examines the similarity between PCA and spatial filtering, and explores the concept of PCA with respect to bandpass calibration and feature detection on both simulated and real data from the Parkes telescope.

Chapter 7 draws together the conclusions from the different chapters and expands upon the future directions of this work.

Appendix A looks at the concept of FPGAs in radio astronomy, and specifically the construction of a basic spectrometer.

Appendix B contains the schematics and Printed Circuit Board (PCB) layout of the equipment developed for the experiment in Chapter 4.

Appendix C contains the schematics for the Antenna Tile Interface Module (ATIM) subsystem of the MWA, discussed in Chapter 5.

Appendix D contains tables of all observations used in this thesis. These tables describe the dates, times, instruments and observers.

# CHAPTER 2

---

## INTERFERENCE SUBTRACTION

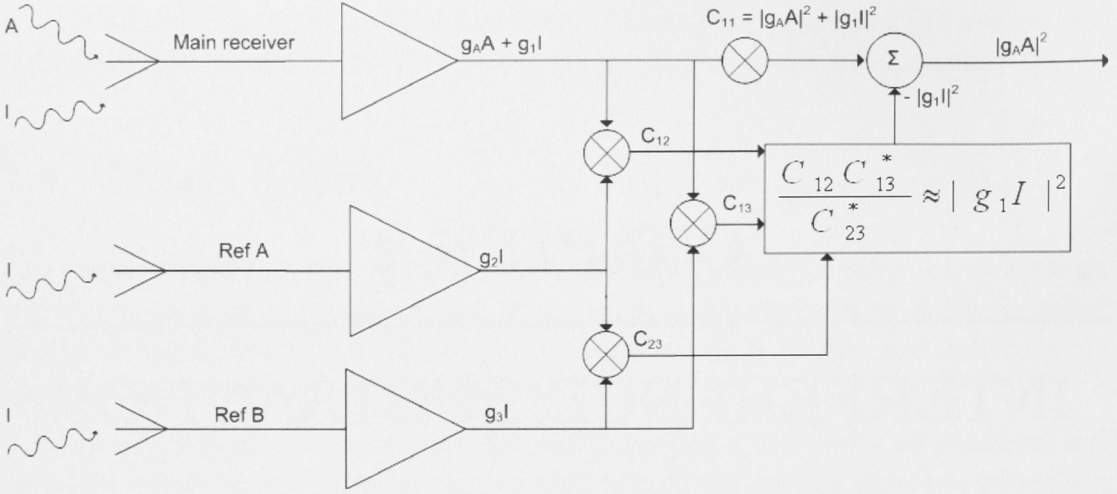
The battle against RFI contamination in astronomical data is continually evolving. As introduced in Chapter 1, one method of removing interference from data using a reference antenna was pioneered by Briggs et al. (2000). This method involves the application of closure relations to gain an accurate deterministic coupling of the interfering signal to the contaminated astronomical data so that it can be subtracted away. This chapter details this method and assesses its application to observations taken at the Parkes telescope.

The observations were taken using the dual polarisation 50 cm receiver on the 64 m dish antenna, a separate 3 m dish for one of the reference antennas, and a Yagi antenna for the other reference antenna. A 64 MHz bandwidth was used, ranging from 643-707 MHz. The CPSRII baseband recorder (van Straten et al., 2000) was used to record the voltages from the input feeds with 2-bit precision at the Nyquist rate (see tables D.1 and D.2 for further observation details).

Analysis of the data was performed offline. This was so that the data could be reprocessed under different conditions. Covariance matrices were formed with 8192 channel resolution using both FFT and Polyphase Filter Bank (PFB) techniques with varying integration times. The details of the processing required to obtain each result are outlined in the appropriate section.

### 2.1 RFI subtraction using reference antennas

The basic premise behind using one or more reference antennas is that a “clean” copy of the interfering signal is obtained separately from the astronomical data. It is hence available for subtraction from the combined astronomical and interference data received by the main dish without affecting the astronomical part of the data. This concept is illustrated in figure 2.1.



**Figure 2.1** Interference subtraction illustration.  $A$  represents the astronomical signal, and  $I$  the interference.  $g_A$  and  $g_1$  are the gains applied to the astronomical and interfering signal respectively in the telescope receiver, and  $g_2$  and  $g_3$  are the gains applied to the interferer in the reference receivers respectively. Closure relations (detailed in §2.1), are used to find the  $g_1$ ,  $g_2$  and  $g_3$  multiplying factors to get an estimate of the interference in terms of  $g_1 I$  to subtract from the contaminated astronomical data.

In the case of a telescope receiver tracking a celestial source, the input voltage to the receiver  $s_1(t)$  can be represented as

$$s_1(t) = G_{A_1} a_A(t) + \sum_k G_{1,k} i(t - \tau_{1,k}) + n_1(t) \quad (2.1)$$

where  $G_A$  is the gain applied to the celestial signal  $a_A$ , by the telescope receiver,  $G_{1,k}$  is the gain applied to the interfering signal  $i$ , and  $n_1$  is the noise.  $\tau$  is the time delay for the signal components due to multipathing.

Transforming these signals into the frequency domain gives

$$S_1(f) = g_{A_1}(f) A_A(f) + g_1(f) I(f) + N_1(f) \quad (2.2)$$

where

$$g_1(f) = \sum_k g_{1,k}(f) e^{j2\pi f L_{1,k}/c} \quad (2.3)$$

where the  $L_{1,k}/c$  term comes from the multiple signal path lengths.  $L$  is the length of the path and  $c$  the speed of light. For further details regarding the derivation of these equations see Briggs et al. (2000) and Briggs & Kocz (2005).

In much the same way, the signal recorded by a reference antenna targeted at an interfering source can be represented by

$$S_{ref}(f) = g_{ref}(f) I(f) + N_{ref}(f) \quad (2.4)$$

provided the reference antenna does not admit any signal from the celestial source. Again, the  $g_{ref}(f)$  term includes the signal multipaths in the same way that  $g_1(f)$  does in equation 2.3. Thus for a system with two reference antennas,  $S_{ref1}$  and  $S_{ref2}$ , where  $S_{ref1}$  and  $S_{ref2}$  are represented by equation 2.4, the covariance matrix,  $\mathbf{C}$ , formed would be as follows:

$$\mathbf{C} = \begin{pmatrix} \langle S_{ref1} S_{ref1}^* \rangle & \langle S_{ref1} S_{ref2}^* \rangle & \langle S_{ref1} S_1^* \rangle & \langle S_{ref1} S_2^* \rangle \\ \langle S_{ref2} S_{ref1}^* \rangle & \langle S_{ref2} S_{ref2}^* \rangle & \langle S_{ref2} S_1^* \rangle & \langle S_{ref2} S_2^* \rangle \\ \langle S_1 S_{ref1}^* \rangle & \langle S_1 S_{ref2}^* \rangle & \langle S_1 S_1^* \rangle & \langle S_1 S_2^* \rangle \\ \langle S_2 S_{ref1}^* \rangle & \langle S_2 S_{ref2}^* \rangle & \langle S_2 S_1^* \rangle & \langle S_2 S_2^* \rangle \end{pmatrix} \quad (2.5)$$

with the first two rows and first two columns being the reference correlations, and  $S_1$  and  $S_2$  the two polarisations of the telescope feed.

Expanding out element  $(i, i)$  of the matrix gives a time-averaged power

$$\langle S_i S_i^* \rangle = |g_{A_i}|^2 \langle |A_A| \rangle + |g_i|^2 \langle |I|^2 \rangle + \langle |N_i|^2 \rangle \quad (2.6)$$

where  $|g_i|^2 \langle |I|^2 \rangle$  is the interference term to be eliminated by the RFI cancellation algorithm. This term can be expressed in the form

$$|g_i|^2 \langle |I|^2 \rangle = \frac{g_i g_j^* g_i^* g_k}{g_j^* g_k} \langle |I|^2 \rangle \quad (2.7)$$

where  $j$  and  $k$  represent reference channels. When the interference to noise power ratio is high ( $|g_i|^2 \langle |I|^2 \rangle \gg \langle |N_i|^2 \rangle$ ) this term can be approximated as

$$|g_i|^2 \langle |I|^2 \rangle \approx \frac{C_{ij} C_{ik}^*}{C_{jk}^*} \quad (2.8)$$

where  $C_{ij}$  represents element  $(i, j)$  in the correlation matrix, with  $i$  denotes the row, and  $j$  the column.

To be most effective, the cross power spectra come from a finite integration time, limited such that the multipath scattering of the interfering signals (and hence  $g_i$  and  $g_{j,k}$ ) will be constant over the integration time. As such, the noise terms in the cross spectra form the ultimate limitations on the accuracy of the cancellation. This is because there is generally a noise floor in the covariance spectra caused by low-level coupling in the receivers and stray radiation (Briggs et al. (2000)). The assumption made here is that only the interfering signal is strongly correlated in the different feeds. In practice, care must be taken when the Interference to Noise Ratio (INR) is low, as the  $C_{jk}$  term in the denominator could be small. This would mean that the correction would diverge to infinity, as the numerator term is now being divided by a term approaching zero. In practice,  $|C_{jk}|^2$  is a positive, non-zero noise power, but to ensure a non-zero denominator, the adjustable, positive parameter  $\psi$  is added. Equation 2.8 is then extended to become



$$|g_i|^2 \langle |I|^2 \rangle \approx \frac{C_{jk} C_{ij} C_{ik}^*}{\psi + C_{jk} C_{jk}^*} \quad (2.9)$$

where  $\psi$  is approximately equal to the square of the power spectrum of the noise in  $C_{jk}$ . This factor can be approximated by a constant  $\psi_0$  and is typically on the order of 0.001.

This estimate (equation 2.9) of the RFI contribution can then be subtracted away from equation 2.6. Note that the reference antenna can only ever be effective at cancelling RFI that is included in the gain reception patterns at a high Signal to Noise Ratio (SNR). One advantage of this is that the celestial source being observed by the dish cannot be harmed, since its signal is excluded from the reference antenna.

## 2.2 Parkes 50 cm CPSR2 data

In an effort to quantify the effect RFI subtraction using a reference antenna has on data, and at the same time examine the RFI environment in a 64 MHz wide band at the 50 cm wavelength at Parkes, three short observations (of approximately 1 minute each) were taken. The observations were recorded by John Reynolds using CPSRII on June 30, 2005. The first two observations were of pulsars J0437-4715 and J0630-2834. The final observation was taken with the telescope positioned at zenith (not observing an astronomical source). The positioning of the reference antennas relative to the Parkes dish is shown in figure 2.2.

### 2.2.1 RFI environment

At 50 cm, the RFI environment at Parkes is unfriendly. Blocks of interference from digital television transmitters which take out several megahertz of bandwidth, coupled with a high spectral occupancy of interference spikes, results in an environment in which it can be hard to take larger bandwidth observations. Interference comes from many sources, two of the main ones show up prominently in spectrum in figure 2.3, which shows an average of the RFI present for three different orientations of the dish on a 10 ms timescale. These strong, broadband interferers are digital television transmitters situated on Mt Ulandra, approximately 200 km away. At 666-673 MHz there is TV channel 48, a 200 kW transmission falling roughly in channels 2950–3850 in the 8192 channel spectrum. At 652-659 MHz there is TV channel 46, a 600 kW transmission approximately corresponding to the 1150 – 2050 channel range in the spectrum.

In addition to these transmitters on Mt Ulandra, there are other transmitters at these same frequencies of varying power and distance. In the 666-673 MHz range for example there is also a transmitter on Mt Sugarloaf. While further away

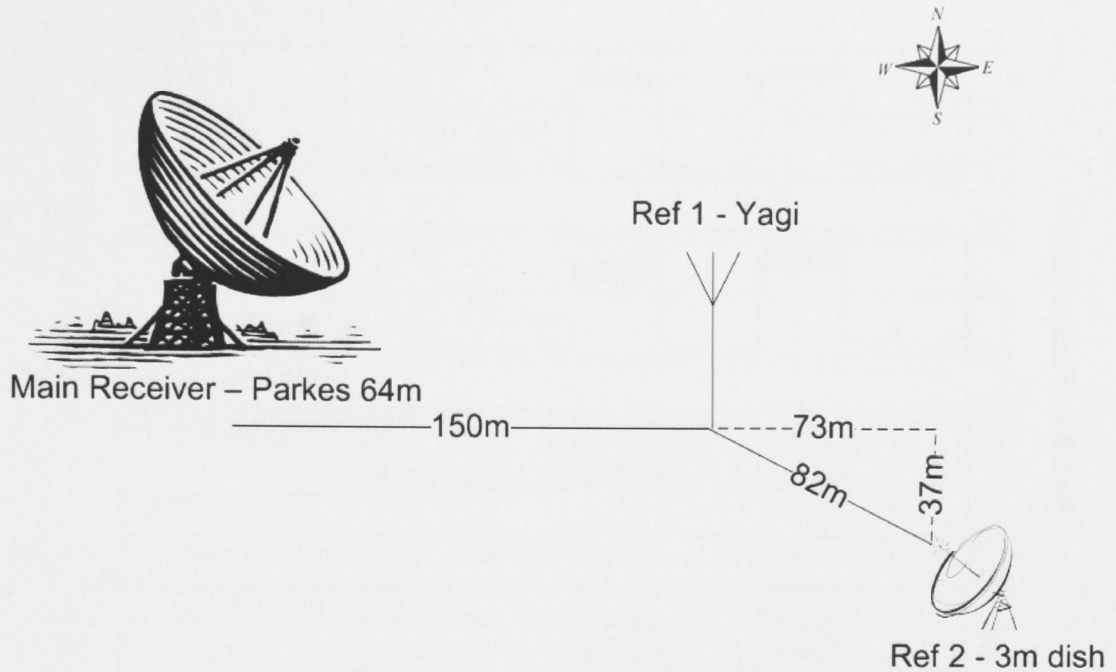


Figure 2.2 Location and type of reference antennas relative to Parkes 64 m dish.

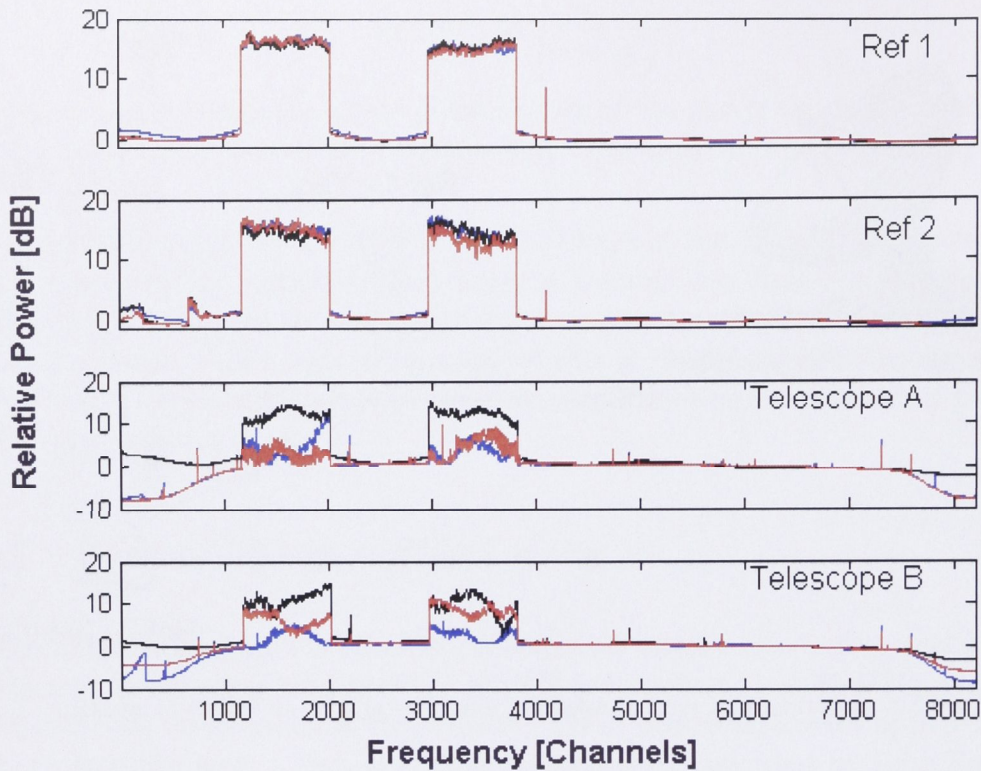
(by around 100 km), the transmitted power is 1200 kW in strength, and can also interfere.<sup>1</sup>

The spectra were created simply using an FFT spectrometer implemented in software, with an integration time of 10 ms. The black line in figure 2.3 shows the spectrum when the dish was pointed at zenith. This position allows RFI to be received from all directions. As the dish tilts to observe each of the other sources, the back side of the dish can act as a shield to the RFI coming from some directions, reducing the overall RFI in the spectrum, and also reducing the magnitude of some of the RFI received (notably shown in the 1000-2000 and 3000-4000 channel ranges contaminated by digital television transmissions).

### 2.2.2 RFI subtraction

In this section the method outlined in §2.1 of using reference signals for RFI cancellation is applied to the pulsar data described in §2.2. Initially a simple analysis of the power spectrum is undertaken, followed in §2.2.3 by an assessment of the effectiveness of the subtraction with respect to pulsar detection.

<sup>1</sup>The transmitter information listed here was correct at January 2006, and taken from "Radio and Television Broadcasting Stations (Internet Edition, January 2006)" Published by the Australian Communications and Media Authority (ACMA). The current listing can be downloaded from the ACMA website. <http://www.acma.gov.au>

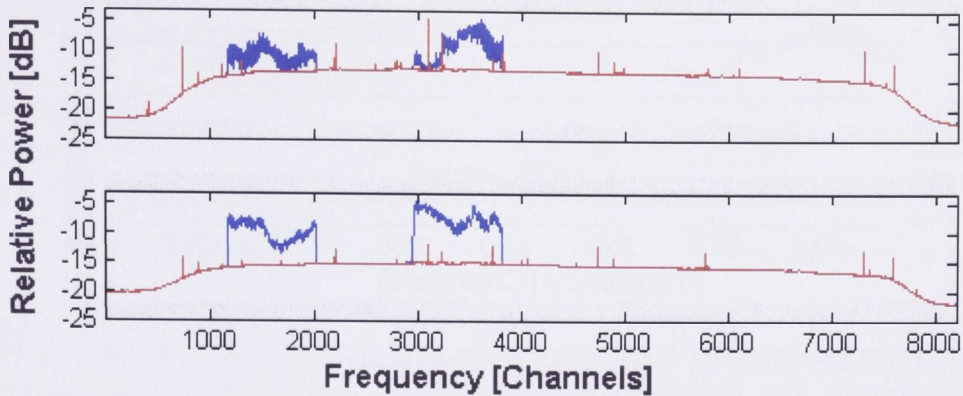


**Figure 2.3** Observed total power spectra for 643 - 707 MHz for a 10 ms integration time. Black line: Observations with the antenna pointed at zenith. Red line: Observations of J0437-4715. Blue line: Observations of J0630-2834. A normalising factor has been applied to make the noise level consistent in each of the orientations. Top panel: Reference antenna one. Second panel: Reference antenna two. Third panel: Telescope receiver, polarisation A. Bottom Panel: Telescope receiver, polarisation B.

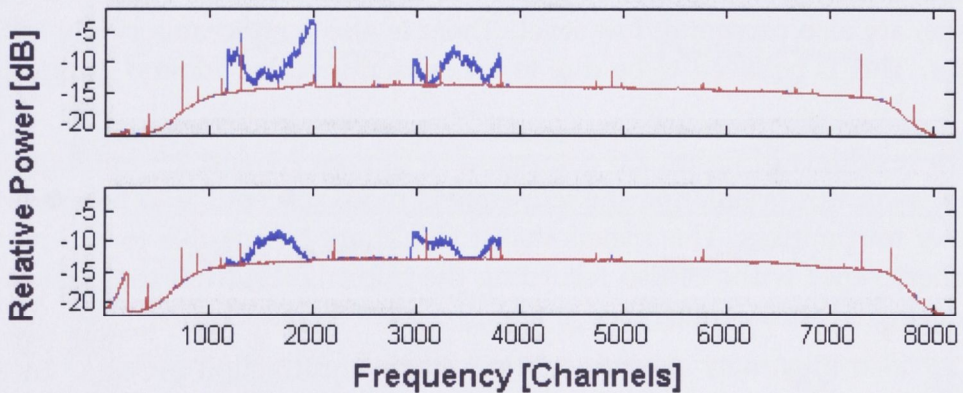
Figures 2.4, 2.5 and 2.6 show the total power spectra, pre- and post-correction for each of the three orientations of the telescope. The total power spectra are taken from a covariance matrix constructed using a PFB spectrometer, implemented in software.

In the case of J0437-4715 (figure 2.4) the majority of the interference was subtracted well, with minor residuals found from TV channel 48, at 666-675 MHz. These residuals were faint enough to only be visible after averaging (also shown in figure 2.7). The spectra in figure 2.4 and 2.7 were created by correcting on 10 ms timescales, then averaging 1024 spectra. As each of the observations themselves lasted only for one minute, the orientation of the dish would not have drastically changed. This means that the continuous interference, caused by the television stations, should remain roughly constant within each observation.

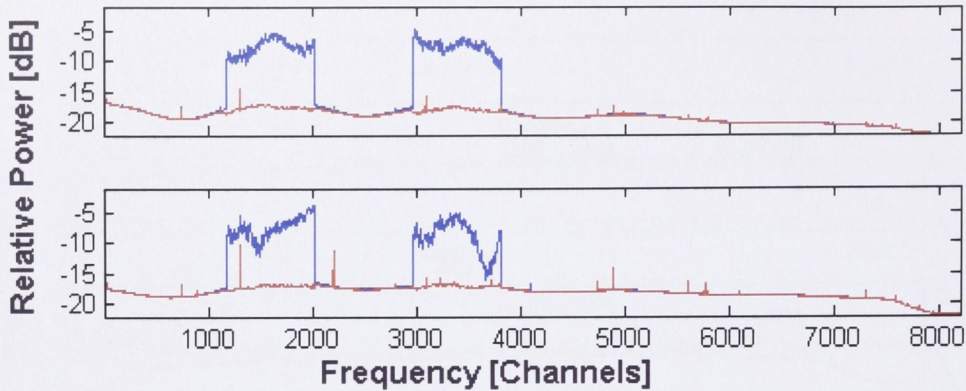
When the orientation of the telescope is changed to observe J0630-2834, interference residuals remain from TV channel 46 at 652-659 MHz (figure 2.5). A closer



**Figure 2.4** Total power spectra of J0437-4715 observation pre- and post-RFI subtraction on a 10 ms timescale. Blue: Original spectra. Red: Corrected spectra. Top panel: Telescope receiver, polarisation A. Bottom Panel: Telescope receiver, polarisation B. The subtraction is very effective, with only very minor residuals showing up after averaging in the channel range 3400-3600



**Figure 2.5** Total power spectra of J0630-2834 observation pre- and post-RFI subtraction on a 10 ms timescale. Blue: Original spectra. Red: Corrected spectra. Top panel: Telescope receiver, polarisation A. Bottom Panel: Telescope receiver, polarisation B. While also generally effective, the subtraction here leaves slightly more residue than in figure 2.4, in this case in polarisation A of the telescope receiver, in the approximate channel range 1900-2000. This interference is believed to be from an alternate source to those observed with the reference antennas, and hence is not possible to be removed in the subtraction process. A magnified view of this portion of the spectrum is shown in figure 2.7.



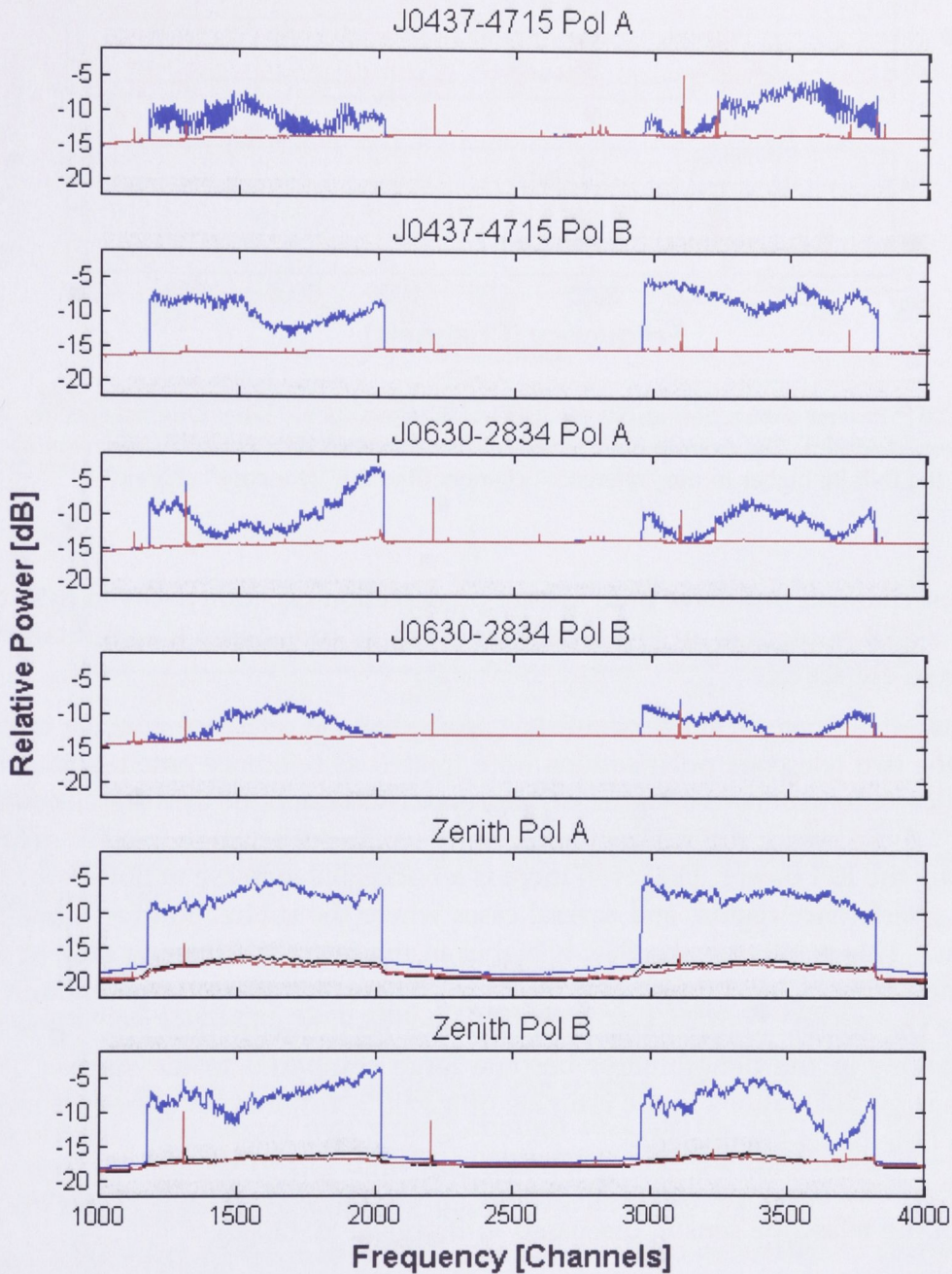
**Figure 2.6** Total power spectra of the zenith observation pre- and post-RFI subtraction. Blue: Original spectra. Red: Corrected spectra. Top panel: Telescope receiver, polarisation A. Bottom Panel: Telescope receiver, polarisation B. The RFI subtraction with the telescope at zenith was less effective than when observing J0437-4715 or J0630-2834. This interference is believed to be from an alternate source to those observed with the reference antennas, and hence is not possible to be removed in the subtraction process.

look at this area of the spectra is shown in figure 2.7. The zenith observations (figure 2.6) show significant residuals in both of the digital TV transmitter channel ranges. This supports the explanation that at certain orientations of the receiver, not only are the nearby TV stations at Mt Ulandra received as interference, but that stations further away, such as those transmitting from Mt Sugarloaf near Newcastle, are also present at low level. There is also a noticeable ripple in the bandpass. This is believed to be due to an unidentified broadband interfering source.

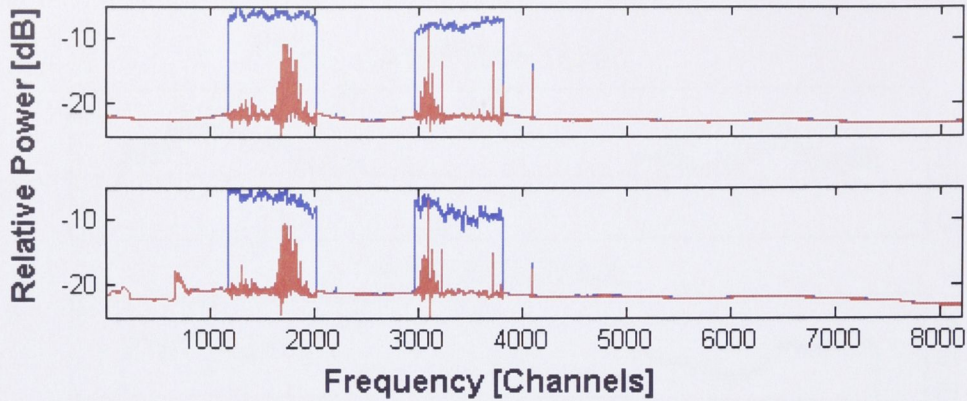
The extra residuals of figure 2.6 compared to figures 2.4 and 2.5 are logically expected, as in this orientation the telescope is more susceptible to interference from other transmitters. This means that it would not be possible to remove all of the interference without also recording the information from these faraway sources using additional reference antennas.

Figure 2.7 also illustrates another subtlety in the subtraction process. In the bottom two panels of this figure, the RFI subtraction was computed using different spectra. One correction was calculated after using a simple  $n$ -point FFT to construct each complex voltage spectrum (black). The other correction was calculated using a PFB in place of the FFT (red). The PFB by nature gives better isolation between the spectral channels. This resulted in the improved application of closure relations, giving a better subtraction.

It should be noted however, that despite the PFB appearing to give an overall better correction, the use of an FFT or PFB in constructing the correlation matrix made little difference to the actual pulsar detection (performed in §2.2.3). This makes sense as the pulsar is a continuum source, so signal bled through into



**Figure 2.7** Total power spectra of J0437-4715, J0630-2834 and zenith position observations pre- and post-RFI subtraction, channels 1000-4000, with an integration time of 10 ms. Blue: Original spectra. Red: Corrected spectra. Top panel: J0437-4715, polarisation A. Second panel: J0437-4715, polarisation B. Third panel: J0630-2834, polarisation A. Fourth panel: J0630-2834, polarisation B. Fifth panel: Zenith position, polarisation A, Black: FFT spectrometer correction. Red: PFB spectrometer correction. Bottom Panel: Zenith position, polarisation B, Black: FFT spectrometer correction. Red: PFB spectrometer correction.



**Figure 2.8** Reverse subtraction results for J0437-4715 observations. Blue: Original spectra. Red: Corrected spectra. The damage done to the reference antenna data illustrates the requirement that the INR be higher in the “reference” channel than the “telescope” channel.

adjacent channels will have little impact on detection capability. Given this, the use of a PFB in place of an FFT should be carefully considered if computational resources are scarce.

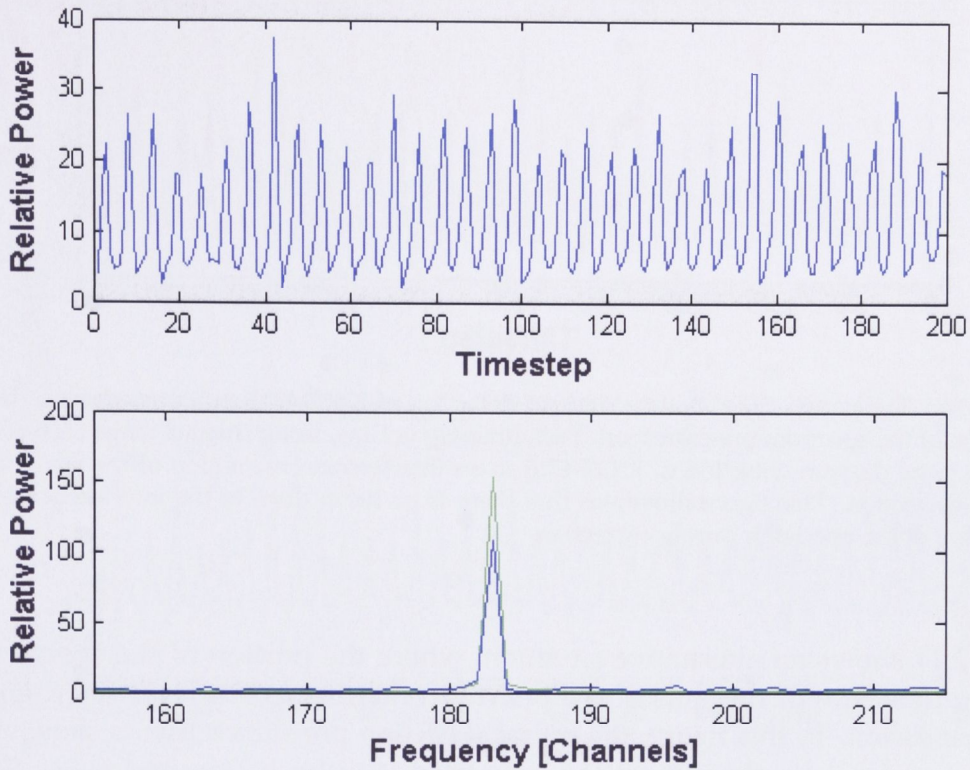
For interest, a “reverse subtraction” was applied to the reference antenna inputs. Here the two telescope polarisations were treated as reference antenna, and the reference antennas as the signal of interest. The results of this are shown in figure 2.8. In places, the subtraction actually works surprisingly well at greatly reducing the RFI power. However, there is a noticeable increase in noise over the whole interference region, and several cases where the subtraction has damaged the data. This is because the INR is higher in the reference antenna than in the telescope sensors, and thus noise terms in the correction are large enough to impact the signal. Mathematically this can be seen in equation 2.9: the noise fluctuations in the denominator become large compared to the signal. This causes large fluctuations in the estimate of  $|g_i|^2 \langle |I|^2 \rangle$ , resulting in either too much or too little being subtracted.

This illustrates the requirement for the INR to be higher in the reference antenna than in the telescope sensor, discussed in Briggs et al. (2000).

### 2.2.3 Pulsar detection

#### Pulsar J0437-4715

The first pulsar observed, J0437-4715 (RA: 04:37:15 and DEC: -47:15:08), was a strong source (304 mJy at 685 MHz) with a low dispersion measure ( $2.6 \text{ cm}^{-3} \text{ pc}$ ) and a period of 5.757 ms. This made the method of detection straightforward. Initially the total power spectrum for each of the inputs was computed. Each



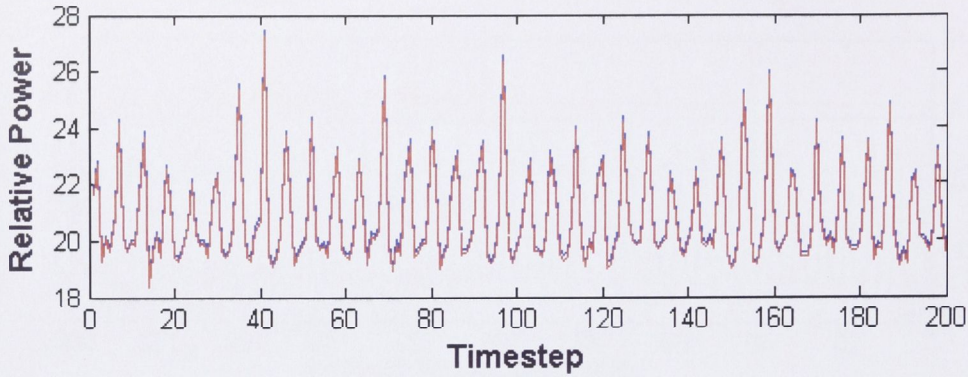
**Figure 2.9** Top panel: A time domain detection of J0437-4715. Each timestep is 1 ms, channels 4000 - 7000 used. Bottom panel: Blue: Frequency domain representation of the time domain detection. Green: Simulated sine wave of the same frequency as the pulsar. Each frequency channel is 1 Hz.

spectrum was averaged such that it represented a time of about 1 ms. For initial detection and verification of the data, a portion of the spectrum that was interference-free was averaged for each of the spectra. This time series is plotted in figure 2.9. The signal was strong enough that for the purposes of simple detection, no folding of the data was necessary, and the dispersion was low enough that it did not significantly impact the detection process.

A sine wave of same frequency as the pulsar was also generated to verify that the signal detected was at the appropriate location in the frequency domain (the Fourier transformed sine wave and pulsar data can be seen in figure 2.9).

In figure 2.10 a time series plot of J40437-4715 in an interference-free region of the spectrum is shown. Overplotted is a detection in the same time range after the subtraction algorithm has been applied. This figure illustrates that the subtraction has negligible effect on the data that is free of interference, and therefore that subtraction can be implemented with the confidence that the process does not harm the astronomical signal of interest.





**Figure 2.10** Top panel: Blue: A time domain detection of J0437-4715 in an interference-free region of the spectrum pre-correction. Each timestep is 1 ms, using channel range 5500-6000. Red: Time domain detection of J0437-4715 in an interference-free region of the spectrum post-correction. This figure illustrates that there is no harm done to the interference-free portion of the spectrum during correction.

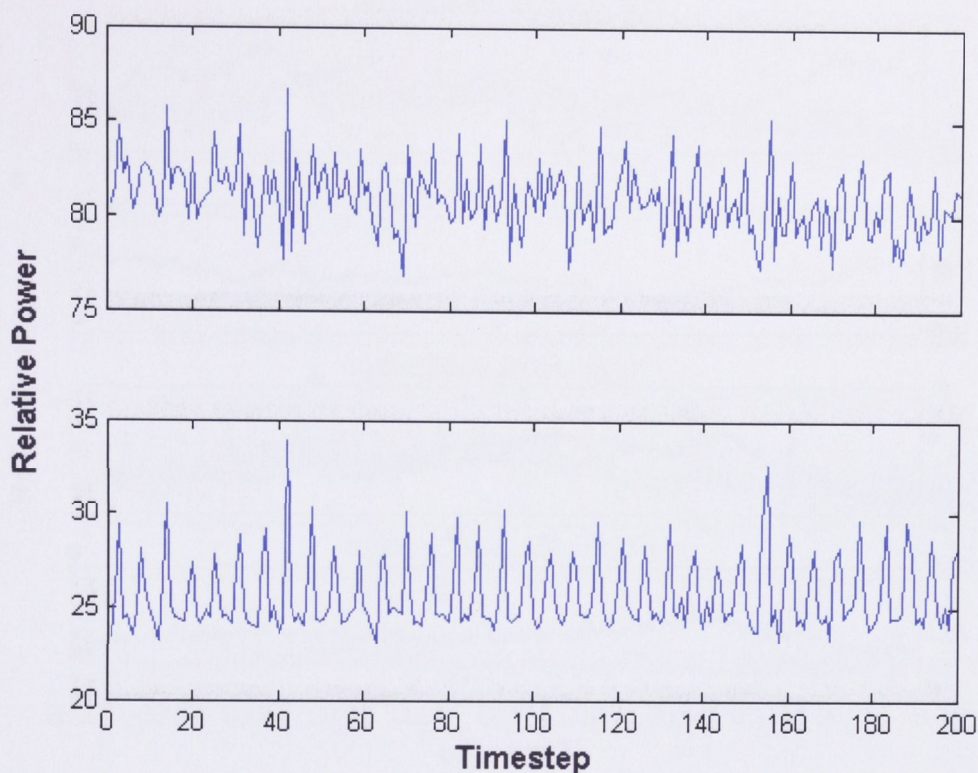
Figure 2.11 shows an alternative situation, where the portion of the spectrum used for detection of the pulsar was heavily contaminated by a nearby digital TV transmission. In this figure, the pulsar detection pre-subtraction is shown in the top panel. Given the integration time used, a pulse is expected every five samples. This signal is not readily apparent in the figure. The bottom panel shows the spectrum after subtraction, where the pulsar is now clearly visible.

### Pulsar J0630-2834

Pulsar J0630-2834 (RA: 06:30:49 and DEC: -28:34:43) is a much weaker source (64.6 mJy at 675 MHz) with a dispersion measure of  $24.4 \text{ cm}^{-3}\text{pc}$  and a period of 1.24 s. Despite the lower flux and larger dispersion, it was still possible to find this pulsar using the simple detection method mentioned in §2.2.3. The only change was that due to the longer period, spectra were computed on a 10 ms timescale.

The four panels in figure 2.12 show the detected power, averaged across the entire 64 MHz band as a function of time. It is suspected that the data here suffered from an additional interfering source, resulting in the instability of the noise power level. The working hypothesis is that this second RFI source was not targeted with a reference antenna, and hence could not be subtracted. The positive effect of the subtraction on the pulsar detection can still clearly be seen, however, when the principal interferers are removed.

Some difficulty was encountered when looking for pulsar J0630-2834 using only small portions of the total bandwidth. The overall goal was to detect the pulsar using a channel range that had been previously affected by RFI, and compare it

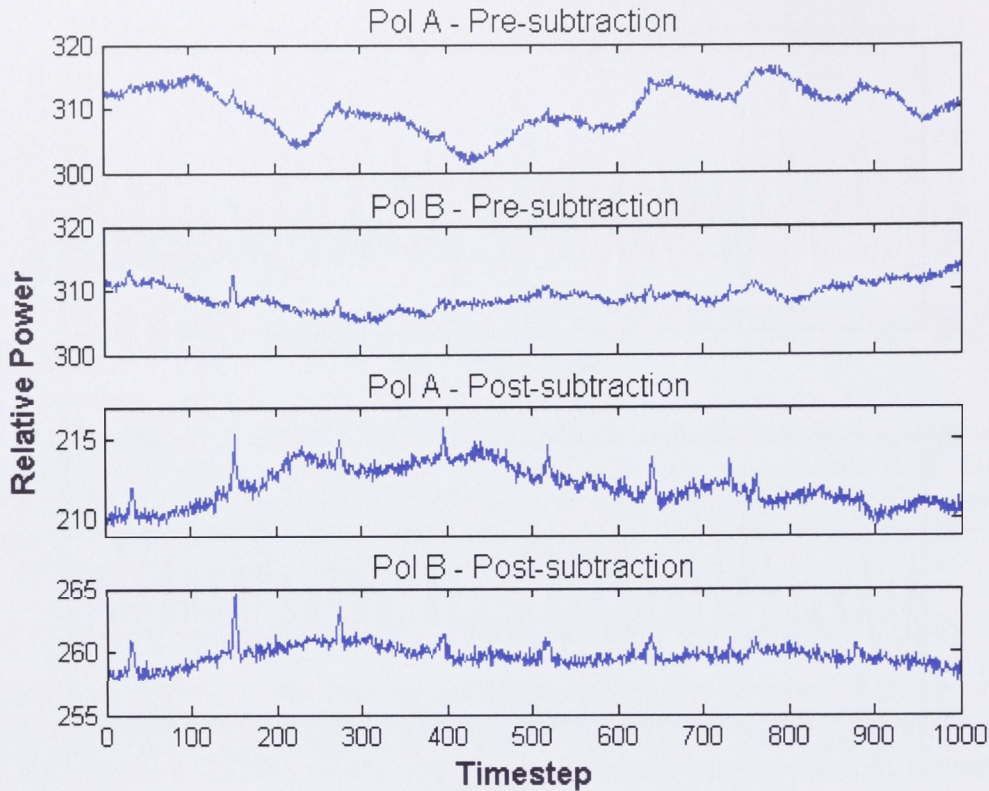


**Figure 2.11** Top panel: Relative power as a function of time for Pulsar J0437-4715 in a region of the spectrum contaminated by interference. The bandwidth taken is the same size as that of figure 2.10, using channel range 3200-3700. Each timestep is 1 ms. Bottom panel: A time domain detection of J0437-4715 in the same region of the spectrum following subtraction of the interference. This figure demonstrates the effectiveness of interference subtraction in removing the interfering signal, whilst leaving the pulsar signal intact.

to one of similar bandwidth that was free of interference. Figures 2.13 and 2.14 illustrate the results of this attempt. In figure 2.13, where an interference-free region was used, it can be clearly seen that the pulsar is more strongly detected in polarisation A. In contrast to this, post-subtraction in figure 2.14, a previously contaminated region, the pulsar is more strongly present in polarisation B, and is not easily detectable in polarisation A at all. The reason for this is still unknown, however the variation in the noise power over time (resulting in the significant ripple seen in figures 2.12 and 2.14) suggests that there may have been another interfering source.

#### 2.2.4 Subtraction effects on data

One of the reasons for conducting these experiments was to determine the effect, if any, that the interference removal would have on the received astronomical



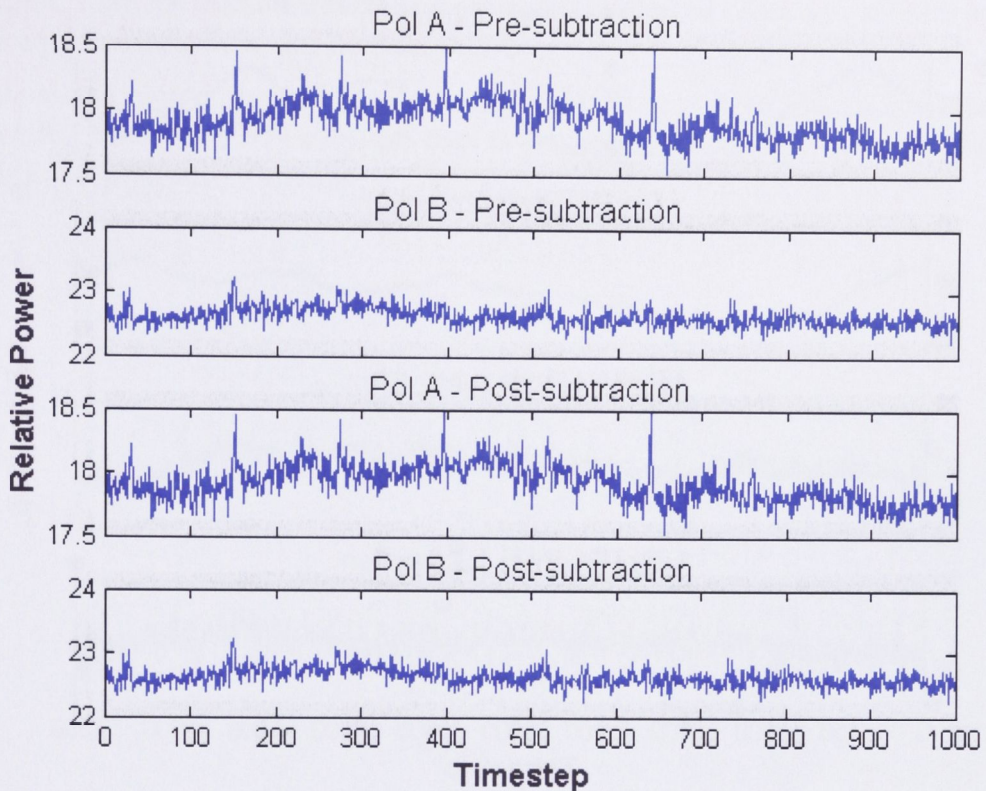
**Figure 2.12** J0630-2834: Relative power as a function of time for two polarisations of the telescope input for the pre- and post-subtraction cases. The power average was taken across the majority of the 64 MHz band, channels 1500-7000. One timestep is 10 ms. Top panel: Polarisation A, pre-subtraction. Second panel: Polarisation B, pre-subtraction. Third panel: Polarisation A, post-subtraction. Bottom panel: Polarisation B, post-subtraction.

data of interest. For example, in a highly contaminated region, if the RFI was removed, would the astronomical data in that region still be valid, or would the process contaminate any results that might have been obtained?

### Effects of subtraction on 50 cm Pulsar observations

As far as could be determined with the data available, the subtraction was not harmful to the data. However, the pulsar J0437-4715 was a strong source. As was noted in §2.2.3 there were problems in detecting the pulsar J0630-2834 in one polarisation. This is still believed to be independent of the cancellation method used however, particularly since pulsar signal was clearly detected in the other polarisation.

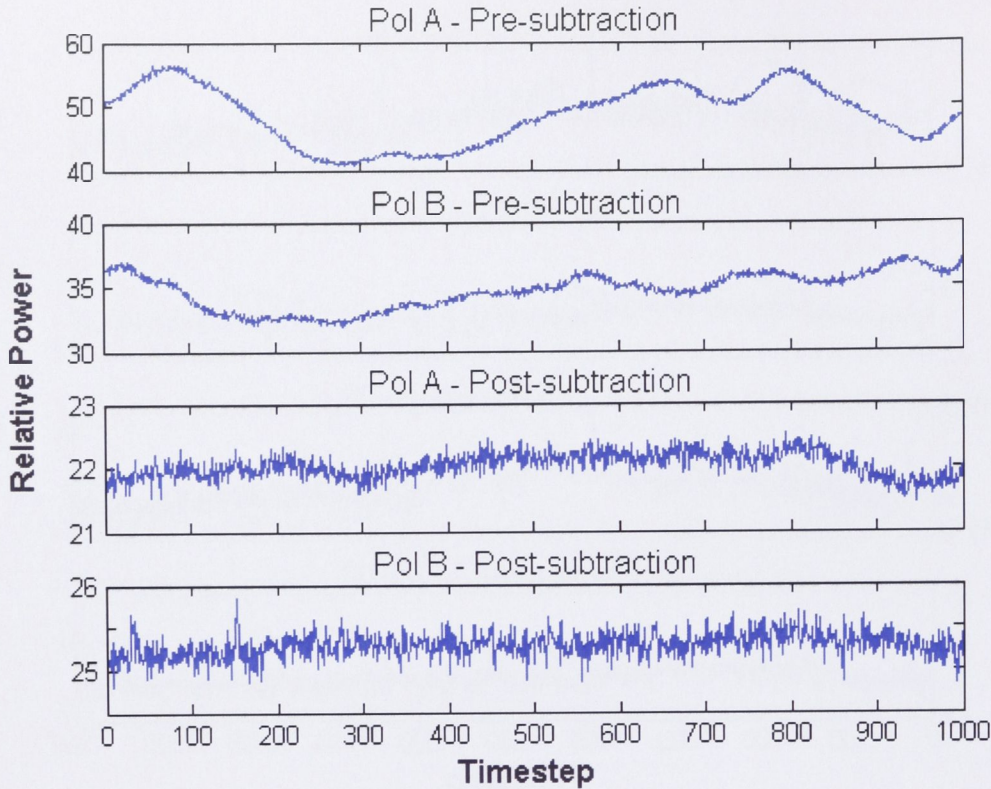
One particular caveat to consider when performing the subtraction however, is the integration time of the correlation spectra. If the integration time is too short the calculated subtraction will be noisy and become inaccurate. If the integration



**Figure 2.13** J0630-2834: Relative power as a function of time in the interference-free 5500-6000 spectral channel region of the spectrum with a timestep of 10 ms. The plot demonstrates the weak pulsar detection: particularly that it stands out in one polarisation more than the other. Top panel: Polarisation A, pre-subtraction. Second panel: Polarisation B pre-subtraction. Third Panel: Polarisation A post-subtraction. Bottom panel: Polarisation B, post-subtraction. Furthermore, the applied correction algorithm has no effect on the interference-free spectral channels.

time is too long, the assumption that the interfering signal and any scattering along the propagation path from the transmitter of this signal are stationary over the integration time breaks down. This problem has been encountered and described previously in Briggs et al. (2000).

One way in which to avoid this problem, if it was absolutely necessary to process the data in short timescales, would be to create a long term filter (based on say one seconds worth of data). This would then be applied to the short integrations, rather than recomputing the subtraction for each individual spectrum. While not tested here, this would make an interesting future experiment. A similar method was implemented for spatial filtering techniques in Chapter §3.



**Figure 2.14** Relative power as a function of time for J0630-2834. In this figure all of the pulsar detection is done in the contaminated 3200-3700 spectral channel region of the spectrum. One timestep is 10 ms. This area of the spectrum is heavily contaminated by a digital TV station. Pre-subtraction detection of the pulsar is difficult. Note that the pulsar is still difficult to detect in the corrected polarisation A (third panel), whereas this channel had the pulsar more strongly in the 5500-6000 channel region (figure 2.13). Top panel: Polarisation A, pre-subtraction. Second panel: Polarisation B pre-subtraction. Third Panel: Polarisation A post-subtraction. Bottom panel: Polarisation B, post-subtraction.

## 2.3 Conclusions

This chapter assessed the effectiveness of using a reference antenna and closure relations to enhance the detection of pulsars J0437-4715 and J0630-2834 in data taken with the Parkes telescope 50 cm receiver.

After performing the subtraction on the observations at three telescope orientations (§2.2.3) it was found that there seemed to be some small residuals left over in the spectra. The amount of residue differed for the three orientations of the dish.

Visually, the interference seemed to be completely subtracted in the case of J0437-4715, an almost entirely removed in the J0630-2834 observations, with a suspected weak second source remaining. This demonstrated a limitation in the

method, namely that an interfering source outside the reception pattern of the reference antenna will not be subtracted.

In the next chapter the technique of spatial filtering will be discussed. This is a powerful method of RFI excision that is accomplished without the need for an additional reference antenna.



# CHAPTER 3

---

## SPATIAL FILTERING

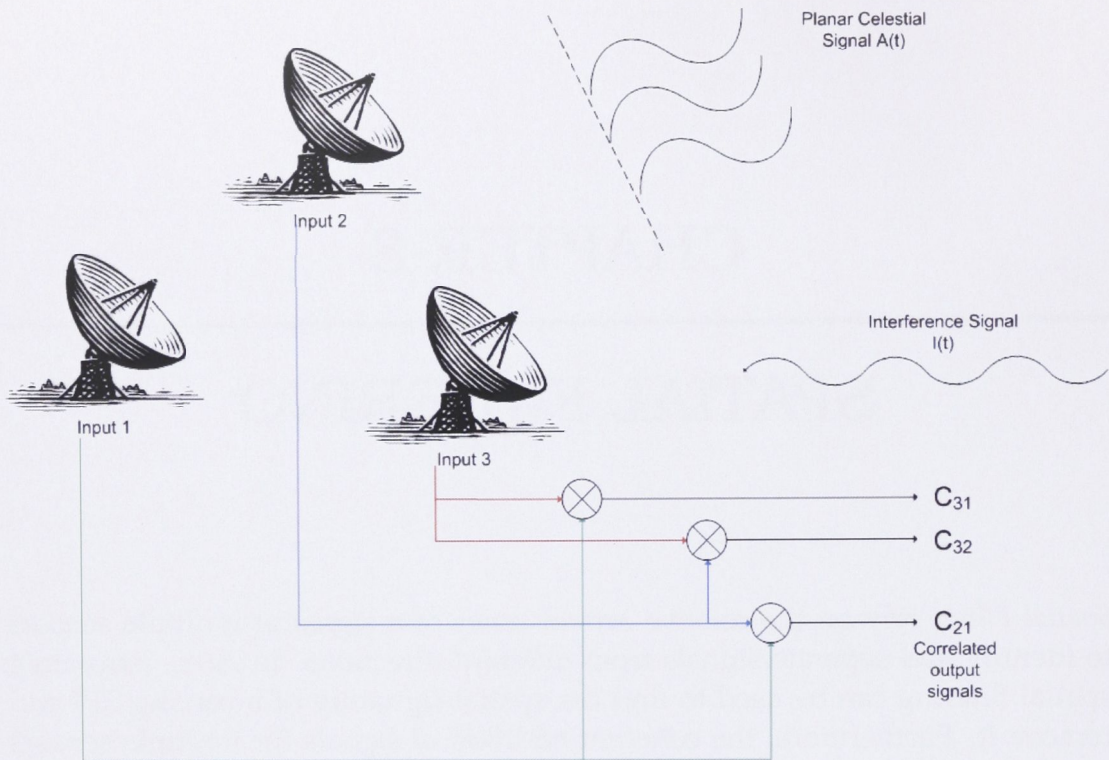
*Spatial Filters* rely on the relative arrival times of a signal at multiple sensors to identify and separate signals from different directions. In radio astronomy, spatial filtering can be used to find the spatial signature of incoming RFI and remove it. Furthermore, the coherent addition of signals for multiple sensors that is inherent in this method greatly enhances the recognition of weak signals. RFI mitigation via spatial filtering can be done both preemptively if the location of the interferer is known and the system can be well calibrated (deterministic nulling), or adaptively, without prior knowledge of interference or detailed system behaviour (non-deterministic nulling).

In this chapter the theory of spatial filtering is reviewed and some of the intricacies for its practical implementation are discussed. The different ways in which spatial filtering can be implemented are explored, and its effect on data is examined. Much of the theory presented in this chapter was developed in the series of papers by Leshem, van der Veen and Boonstra (see Boonstra (2005) for details and references). As mentioned in Chapter 1 there is extensive literature on this topic that is steadily evolving.

### 3.1 Phased arrays and covariance matrices

The method explored in this chapter relies on the information carried by correlation coefficients between multiple receivers, many or all of which have receptivity to interfering signals. Correlation coefficients are statistical estimates of the commonality of two signals formed from the time average of the covariance of the signals in the presence of noise. In the present application these are organised into a matrix of complex numbers that measure the strength of signals in common to the receivers and information on the relative arrival times, which is contained in the phases of the complex coefficients. An example of a simple





**Figure 3.1** Example of a simple telescope array. The interfering signal is assumed to approach from a different direction to the astronomical source. The  $\otimes$  symbol represents cross-correlation.

array is shown in figure 3.1. The interfering signal is assumed to approach from a different direction to the astronomical source. The data from each antenna are correlated, creating a covariance matrix.

The covariance matrix captures information about relative power of the received signal as a snapshot over the time interval of the integration for the complex coefficients. There are efficient techniques for extracting information from the correlation matrix such as Singular Value Decomposition (SVD) (introduced in the next section). SVD is effective at adaptively recovering the interfering signal strength and the complex gains by which it is coupled to each sensor.

Applications in radio astronomy that rely on the treatment of correlation coefficients as “visibilities” for use in synthesis imaging are also a form of spatial filtering. In this case, prior knowledge of the calibration of the array of receivers allows the integration time to be arbitrarily extended in specific directions in the form of a map or synthesis image of the sky. The imaging application relies on accurate knowledge of the antenna locations for mapping arrival time information to directions in the sky. Spatial filtering methods also have a more general application where sensor locations need not be known, provided the signal is strong enough to stand out from the noise in a coherent integration

time.

## 3.2 Singular Value Decomposition

Extraction of information from the covariance matrix requires “solving” or “decomposing” the matrix. In mathematical terms, the solution comprises a set of eigenvalues and eigenvectors. The mathematics provides a further framework for applications that enhance or null the signals that appear in the decomposition.

SVD is one method of determining the eigenvectors and eigenvalues of a matrix. The strength of SVD is that it produces useful solutions from matrices that are singular and from systems of equations that are poorly constrained. Once this decomposition has been accomplished, the extracted eigenvectors (corresponding to complex gain factors) and eigenvalues (signal powers) of the system can be analysed and manipulated.

In the present application, SVD is performed on a covariance matrix. In a similar manner to §2.1, the covariance matrix,  $\mathbf{C}$ , is constructed for each spectral frequency channel such that

$$\mathbf{C} = \begin{pmatrix} \langle S_1 S_1^* \rangle & \langle S_1 S_2^* \rangle & \cdots & \langle S_1 S_p^* \rangle \\ \langle S_2 S_1^* \rangle & \langle S_2 S_2^* \rangle & \cdots & \langle S_2 S_p^* \rangle \\ \vdots & \vdots & \cdots & \vdots \\ \langle S_p S_1^* \rangle & \langle S_p S_2^* \rangle & \cdots & \langle S_p S_p^* \rangle \end{pmatrix} \quad (3.1)$$

where there are  $p$  receiver inputs, and  $S_i(f)$  is the Fourier transform of the input voltage signal  $s(t)$ .  $S_i(f)$  is defined as

$$S_i(f) = g_{Ai} A_A(f) + g_{Ii} I(f) + N_i(f) \quad (3.2)$$

where  $g_{Ai}$  is the gain applied by channel  $i$  to the celestial signal  $A_A$ , by the telescope receiver,  $g_{Ii}$  is the gain applied to the interfering signal  $I$ , and  $N$  is the noise (assuming one celestial source, one interferer, and independent noise in each sensor). The computation  $\langle S_i(f) S_i(f)^* \rangle$  is equal to  $|g_{Ai}|^2 \langle |A_A|^2 \rangle + |g_{Ii}|^2 \langle |I|^2 \rangle + \langle |N_i|^2 \rangle$ , and forms a power spectrum.

The cross terms (e.g.  $g_{Aj} g_{Ai}^* A_A I^*$ ) have been neglected in the calculation of  $\langle S_1(f) S_1(f)^* \rangle$  as they average toward zero (as shown in Briggs et al. (2000)). Note also that  $\langle N_i N_i^* \rangle$  goes to noise power  $\sigma_n^2$ , but in a cross-correlation,  $\langle N_i N_j^* \rangle$ , averages to zero for  $i \neq j$ .

Performing SVD on a covariance matrix decomposes it into the form

$$\mathbf{C} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^H \quad (3.3)$$

where  $\mathbf{C}$  is the covariance matrix, and  $\mathbf{U}$  and  $\mathbf{U}^H$  are the  $p \times p$  eigenvector matrix and its Hermitian conjugate respectively. This matrix  $\mathbf{U}$  can be expanded into the eigenvectors

$$\mathbf{U} = (\vec{U}_1 \vec{U}_2 \dots \vec{U}_N). \quad (3.4)$$

The elements of  $\mathbf{U}$  are themselves vectors, of dimension  $p$  (that is,  $\vec{U}_1 = (U_{11}, U_{12}, \dots, U_{1p})$ ). The eigenvectors form an orthonormal set of vectors having the property  $\vec{U}_i \bullet \vec{U}_i = 1$  and  $\vec{U}_i \bullet \vec{U}_j = 0$  for  $i \neq j$ . The eigenvalue matrix,  $\mathbf{\Lambda}$ , is of the form

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_p \end{pmatrix} \quad (3.5)$$

where  $\lambda_1 > \lambda_2 > \dots > \lambda_p$ .

In the present application, the eigenvalues can be viewed as signal powers. SVD orders the eigenvalues so that the first eigenvalue contains the power due to the strongest signal. In the case of RFI affected radio astronomy observations, this is commonly the interfering signal, as it is generally orders of magnitude greater than the celestial signal of interest.

These eigenvector and eigenvalue matrices are computed for each spectral channel from its covariance matrix,  $\mathbf{C}(f)$ . Since the decomposition responds dynamically to changes in the relative strengths of the correlated signals and noise powers, the eigenvectors reorient as a result. Furthermore, normalisation of the eigenvectors to unit length means that these eigenvector coefficients that might be conveniently mapped into gain factors are coupled by the decomposition, making an intuitive understanding difficult.

### 3.3 Spatial filtering implementation options

The signals in the covariance matrix are made up of several components: receiver gains, astronomy and interfering signals, and noise (see for example equation 3.2). This can be stated as

$$\mathbf{C} = \mathbf{C}_A + \mathbf{C}_I + \mathbf{C}_N \quad (3.6)$$

where  $\mathbf{C}_A$  represents the astronomy signal,  $\mathbf{C}_I$ , the interference, and  $\mathbf{C}_N$  the system noise (as described in §3.2). For simplicity, consider the case where the integration time of the covariance matrix is short enough that the signal  $\mathbf{C}_A$  is negligible with respect to the interference and noise (as discussed in section 4.3.2 this is not always a practical assumption). The covariance matrix can then be approximated as

$$\mathbf{C} \approx \mathbf{C}_I + \mathbf{C}_N. \quad (3.7)$$

Assuming independent noise signals of equal power,  $\langle \sigma_{n_i}^2 \rangle = \sigma_n^2$  in each of the  $p$  signal paths. The eigenvalue and eigenvector representation of the covariance matrix can then be partitioned to give

$$\mathbf{C} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H \quad (3.8)$$

$$= \left( \mathbf{U}_I \mid \mathbf{U}_n \right) \left( \begin{array}{c|c} \mathbf{\Lambda}_I + \sigma_n^2 \mathbf{I}_q & 0 \\ \hline 0 & \sigma_n^2 \mathbf{I}_{p-q} \end{array} \right) \begin{pmatrix} \mathbf{U}_I^H \\ \mathbf{U}_n^H \end{pmatrix} \quad (3.9)$$

where there are  $q$  interferers,  $p$  inputs, and  $\mathbf{I}_p$  denotes a  $p \times p$  identity matrix. Here  $\mathbf{U}_I$  represents the “signal” (or interferer) subspace eigenvector matrix, and is of dimension  $p \times q$ .  $\mathbf{U}_n$  the noise (or null) subspace eigenvector matrix of dimension  $p \times (p - q)$ .  $\mathbf{\Lambda}_I$  is the interference power eigenvalue matrix, and is a diagonal matrix of dimension  $q \times q$ , and  $\sigma_n^2$  is the noise power.

In an ideal, noise free system,  $\sigma_n^2 = 0$ . However, equation 3.9 shows that when  $\sigma_n^2 \neq 0$  the interference eigenvalues are biased to be too large. The effects of this bias, and methods of compensating for it are outlined later in equations 3.18 through 3.26.

Mathematically there are several ways in which spatial filtering can be accomplished. While producing identical outputs, the methods vary in computational requirements. These methods include:

- Zeroing the appropriate eigenvalue(s) (those that specify the RFI power), then reconstructing the covariance matrix.
- Projecting out the interference subspace.
- Subtracting the interference subspace from the signal subspace.

Conceptually, the simplest way to remove the interference is to null the primary eigenvalue and the reconstruct the correlation matrix. For example, take a  $2 \times 2$  covariance matrix that can be represented by some eigenvalues and eigenvectors such that

$$\mathbf{C} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix} \begin{pmatrix} a_1^* & a_2^* \\ b_1^* & b_2^* \end{pmatrix} \quad (3.10)$$

Reconstructing this matrix gives

$$\mathbf{C} = \begin{pmatrix} Aa_1a_1^* + Bb_1b_1^* & Aa_1a_2^* + Bb_1b_2^* \\ Aa_2a_1^* + Bb_2b_1^* & Aa_2a_2^* + Bb_2b_2^* \end{pmatrix} \quad (3.11)$$

Assuming that the interference power was represented in the first eigenvalue, setting this value,  $A$ , to zero would remove the interference from the reconstructed covariance matrix,  $\tilde{\mathbf{C}}$ , such that

$$\tilde{\mathbf{C}} = \begin{pmatrix} Bb_1b_1^* & Bb_1b_2^* \\ Bb_2b_1^* & Bb_2b_2^* \end{pmatrix} \quad (3.12)$$

A similar effect is achieved by creating a noise subspace projection, although in a less intuitive manner. A projection is achieved by first determining the spatial signature of the signals to be removed (the interference subspace). A vector projection is then defined that is orthogonal to this spatial signature, so that when applied to the interference space, the information for the interference subspace is zeroed. An in depth discussion on the formation of projections can be found in Strang (1998).

Consider the  $2 \times 2$  case again of equation 3.11. Because the eigenvectors are orthogonal, the noise and interference subspaces form a natural basis for projections. The projection operator  $\mathbf{P}_N$  (itself a  $2 \times 2$  matrix) to extract the noise subspace from the covariance matrix is defined as

$$\mathbf{P}_N = \vec{U}_n \vec{U}_n^H \quad (3.13)$$

and therefore

$$\tilde{\mathbf{C}} = \mathbf{P}_N \mathbf{C} \mathbf{P}_N \quad (3.14)$$

The projection operation produces an identical result for  $\tilde{\mathbf{C}}$  as described in equation 3.12.

Writing the projection matrix formed in terms of the vector elements of equation 3.10 gives

$$\mathbf{P}_N = \vec{U}_n \vec{U}_n^H = \begin{pmatrix} b_1 b_1^* & b_1 b_2^* \\ b_2 b_1^* & b_2 b_2^* \end{pmatrix} \quad (3.15)$$

Labelling the elements of the covariance matrix

$$\mathbf{C} = \begin{pmatrix} W & X \\ Y & Z \end{pmatrix} \quad (3.16)$$

and then multiplying through the projection gives a new first element in the covariance matrix of

$$\tilde{C}_{11} = b_1 b_1^* W b_1 b_1^* + b_1 b_2^* Y b_1 b_1^* + b_1 b_1^* X b_2 b_1^* + b_1 b_2^* Z b_2 b_1^* \quad (3.17)$$

where it is shown that there are now portions of each of the elements in the corrected covariance matrix located in the first value  $C_{11}$ . This will occur in all of the elements of  $\tilde{\mathbf{C}}$ .

It is difficult to see the equivalence of equations 3.12 and 3.17 intuitively, so to demonstrate the equivalence of the projection and nulling, the following specific covariance matrix is taken from Mitchell (2004). In this example, the covariance matrix is defined as

$$\mathbf{C} = \begin{pmatrix} \sigma_n^2 + \sigma_I^2 & \sigma_I^2 \\ \sigma_I^2 & \sigma_n^2 + \sigma_I^2 \end{pmatrix} \quad (3.18)$$

where  $\sigma_n^2$  and  $\sigma_I^2$  are the receiver noise and interference power respectively.

Decomposing this matrix, assuming only one interferer and assigning the values to the general case equations 3.10 and 3.14 gives, for 3.10:

$$(a_1, a_2)^T \text{ is } \vec{U}_1 = \vec{U}_I = \frac{1}{\sqrt{2}} (1, 1)^T.$$

$$(b_1, b_2)^T \text{ is } \vec{U}_2 = \vec{U}_n = \frac{1}{\sqrt{2}} (1, -1)^T.$$

The eigenvalue  $A = \Lambda_I = \lambda_1 = \sigma_n^2 + 2\sigma_I^2$  and  $B = \lambda_2 = \sigma_n^2$ .

Setting  $A = 0$ , and recombining the decomposed matrix gives

$$\tilde{\mathbf{C}} = \frac{1}{2} \begin{pmatrix} +\sigma_n^2 & -\sigma_n^2 \\ -\sigma_n^2 & +\sigma_n^2 \end{pmatrix} \quad (3.19)$$

In the case of 3.14, creating the projection matrix  $\mathbf{P}_N = \vec{U}_n \vec{U}_n^H$  gives

$$\mathbf{P}_N = \frac{1}{2} \begin{pmatrix} +1 & -1 \\ -1 & +1 \end{pmatrix} \quad (3.20)$$

and applying this projection to the covariance matrix

$$\mathbf{P}_N \mathbf{C} \mathbf{P}_N = \frac{1}{2} \begin{pmatrix} +\sigma_n^2 & -\sigma_n^2 \\ -\sigma_n^2 & +\sigma_n^2 \end{pmatrix}. \quad (3.21)$$

which is equivalent to the result obtained in 3.19. Comparing both these results to the original covariance matrix (equation 3.18) shows that too much information has been removed from all of the elements of the matrix. The desired case was to have

$$\tilde{\mathbf{C}} = \begin{pmatrix} \sigma_n^2 & 0 \\ 0 & \sigma_n^2 \end{pmatrix}. \quad (3.22)$$

This discrepancy was foreshadowed in the discussion surrounding equation 3.9. Simply zeroing the eigenvalue  $\lambda_1$ , or projecting into the noise subspace, removes both the RFI and also whatever noise and celestial signal that happens to have a component in the direction of signal space delivered from  $\lambda_1$ . For example, as was just shown, if the first eigenvalue,  $A$ , of equation 3.10 is removed in its entirety, the interference is taken away, along with any system noise contribution (and potentially, in the case where the assumption that the astronomy signal is negligible breaks down, any weak celestial signal contributing to the noise power). While ideally there would be no astronomical or noise signal in  $\lambda_1$ , there is normally some bleed-through between the signals. Therefore, some effort must go into compensating for the overcorrection, when the noise is subtracted along with the interference.

This can be achieved by modifying the nulling or projection scheme slightly.

Instead of removing the first eigenvalue, it can be replaced with an estimate of the noise, calculated via an average of the noise power represented by the noise space eigenvalues. That is

$$\lambda_{p_i} = \frac{\sum^j \lambda_{q_j}}{q} \quad (3.23)$$

where  $\lambda_{p_i}$  is the new value,  $\sum^j \lambda_{q_j}$  is a sum of all the noise space eigenvalues and  $q$  is the number of noise space eigenvalues. In the above example, substitution of  $A = \sigma_n^2$  instead of  $A = 0$  indeed gives the desired result.

It is difficult to correct the noise projection in an equivalent way. However, instead of projecting onto the noise subspace, the interference free covariance matrix can also be obtained by subtracting away the interference subspace. That is

$$\tilde{\mathbf{C}} = \mathbf{C} - \mathbf{P}_I \mathbf{C} \mathbf{P}_I^H \quad (3.24)$$

Ideally, this subtraction of the interference subspace from the covariance matrix is equivalent to

$$\tilde{\mathbf{C}} = \mathbf{C} - \vec{\mathbf{U}}_I \Lambda_I \vec{\mathbf{U}}_I^H \quad (3.25)$$

In this case, one way ensure the correct subtraction is made is to remove the bias from  $\Lambda_I$  due to the excess noise, i.e.  $\Lambda_I$  is replaced by  $\Lambda_I - \sigma_n^2$  (see Leshem & van der Veen (1999) and Mitchell (2004) for further details). The corrected covariance matrix is then

$$\tilde{\mathbf{C}} = \mathbf{C} - \vec{\mathbf{U}}_I (\Lambda_I - \sigma_n^2) \vec{\mathbf{U}}_I^H \quad (3.26)$$

Comparing equation 3.24 to the cases given in equation 3.10 and 3.17 it is now trivial to show that a substitution of a common covariance matrix into any of these equations gives an equivalent result, and that setting  $A = \sigma_n^2$  in equation 3.10 gives the same result as scaling  $\Lambda_I$  in equation 3.26. A good comparison of the two projection cases is given in Mitchell (2004).

Leshem et al. (2000) and Boonstra (2005) also derive a method of correcting the noise space projection case by storing all the projections as they are made. The record of projections can be used to compensate for distortion in the covariance matrix imposed by the nulling the eigenvalues.

Chapters 4 and 5 explore different applications of these concepts in terms of bandwidth, adaptive nulling and a more slowly adapting filter.

### 3.3.1 Expected Performance

In the case where the spatial signature of the interfering signal is perfectly characterised, the interference removal should be complete. In practice, however, this signature will always be an estimate whose precision is limited by noise. The residual interference after filtering is a function of the integration time and the number of receivers, and it can be quantified by equation 3.27 (Boonstra, 2005) describing the relationship of the interference to noise ratio in the corrected dataset  $INR_{after}$  to the interference to noise ratio prior to cancellation  $INR_{before}$ .

$$INR_{after} = \frac{1}{N} \left( 1 + \frac{1}{pINR_{before}} \right) \quad (3.27)$$

where  $N$  is the number of spectra in the average, and  $p$  is the number of receivers. The residuals after cancellation of the interference will be spatially white (Boonstra, 2005).

This metric will be revisited in Chapter 4 for each of the different filtering variations attempted.

## 3.4 Data calibration requirements

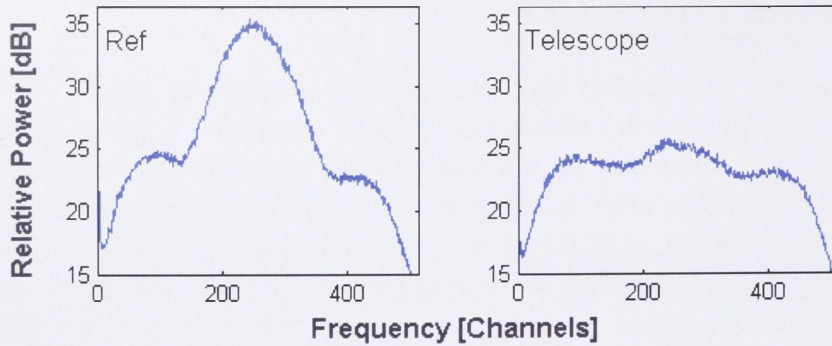
An important factor to note is that the successful separation of the various input signals into eigenvalues requires the data be calibrated (or normalised) as accurately as possible. In particular, the gains between the various inputs need to be calibrated. This process is also referred to in Leshem & van der Veen (1999) as “noise whitening” and is discussed in detail therein.

A practical example showing the effect of gain calibration on the effectiveness of the SVD is given here. The data used was dataset SRTCA01 taken by Bell et al. (2001) for the express purpose of testing RFI mitigation algorithms. A covariance analysis, decomposition and RFI elimination was performed and the results published in Kocz (2004). The results shown here have been reprocessed to a different time resolution, but essentially reproduce the same figures.

The data were taken on the ATCA at Narrabri NSW. The ATCA dataset consists of 12 simultaneously recorded signals: Five receivers with dual polarisation observed a common point in the sky, and one reference antenna with dual polarisation was directed at an interfering terrestrial transmitter. A 4 MHz bandwidth centered on frequency 1503 MHz was sampled with 4-bit precision at the Nyquist rate (see tables D.1 and D.2 for further observation details).

The input signals were Fourier transformed to obtain voltage spectra with 8 kHz resolution. The data for each frequency channel were cross-multiplied to form covariance matrices, with an integration time of 100 ms. An example of the total





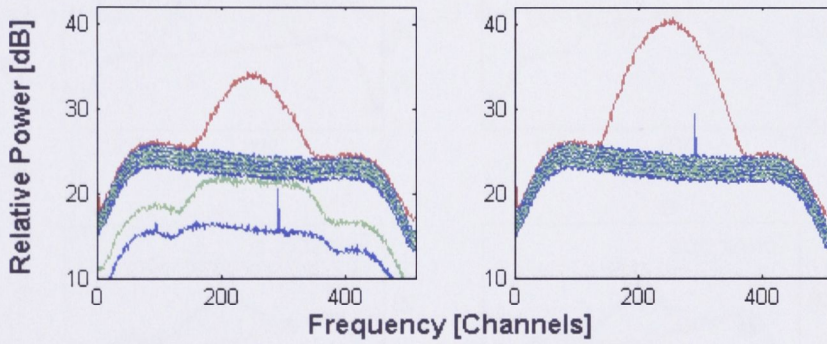
**Figure 3.2** Total power spectra for one polarisation of the reference antenna (left) and one ATCA telescope input (right). In the reference channel, the interference is mainly present in spectral channels 150-380. In the telescope input, mainly in spectral channels 180-380. There are 512 spectral channels across 4 MHz total bandwidth, centred at 1503 MHz.

power spectra of one of the reference antenna polarisations and one of the sky observations are shown in figure 3.2.

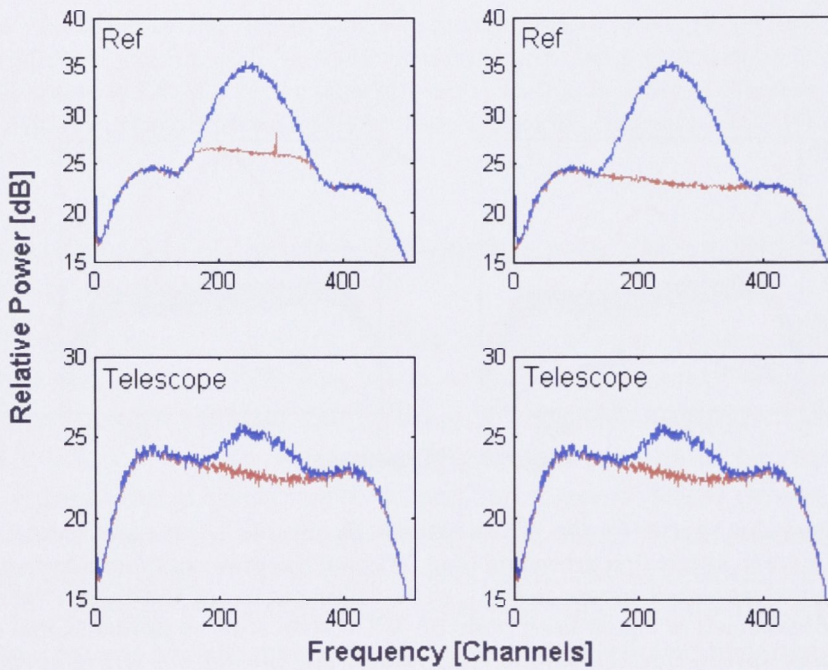
Two separate versions of the covariance matrices were decomposed using SVD. In the first version, the SVD was taken without any attempt at gain calibration. The eigenvalues for each of the frequency channels are shown on the left in figure 3.3. Ideally if there was a single interferer, it would be present only in the first eigenvalue. If there was a second interferer of lesser power, it would be represented in the second eigenvalue and so on. Here the interference appears to be spread throughout the first eigenvalue, with related residuals in eigenvalues 11 and 12. Clearly the process has failed to fully isolate the RFI making it harder to remove the interference. It is also problematic because if the lower, less significant, eigenvalues are ignored, there is then residual interference that is not removed.

A second decomposition of the covariance matrices undertaken after a rudimentary gain calibration was performed resulted in the eigenvalue spectra on the right-hand side of figure 3.3. In this case the calibration consisted of simple scaling of the spectra by an appropriate factor. Comparing this figure to the one on the left, the interference is now located principally in a single eigenvalue, the first one, with a second, narrowband interferer now located in the second eigenvalue.

The “corrected” spectra (using the method outlined in equation 3.10 and incorporating the modification of equation 3.23), for both the reference antenna and one telescope receiver are shown in figure 3.4. The differences between the calibrated and uncalibrated results are more noticeable in the reference antenna case, as this was where the majority of gain calibration had to take place.



**Figure 3.3** Eigenvalue spectra for the ATCA dataset. Eigenvalue  $\lambda_1$  for each channel is in red, then  $\lambda_2$  to  $\lambda_{12}$  are in alternating green and blue. Without the presence of an interfering signal, all of the eigenvalues would appear as  $\lambda_3$ - $\lambda_{12}$  in the righthand figure, that is, as a flat bandpass shape. The breakdown of eigenvalues from the ATCA data without normalisation (left) has the interference spread through several eigenvalues. The breakdown where normalisation has been performed (right), has all of the interference in the first two eigenvalues.  $\lambda_1$  shows a broadband interfering signal, stretching from channels 150-380.  $\lambda_2$  shows a narrow interferer at around channel 300. There are 512 spectral channels across 4 MHz total bandwidth, centred at 1503 MHz.



**Figure 3.4** Pre- and post-correction spectra for the reference antenna connected to the ATCA receiving system and one input from the telescope. The bandwidth is 4 MHz and the centre frequency 1503 MHz. Blue: Original spectra. Red: Corrected spectra. Top Left: Correction for the reference antenna formed without normalising the covariance matrix. Top Right: Correction for the reference antenna formed using a normalised covariance matrix. Bottom Left: Correction for the telescope feed formed using a non-normalised covariance matrix. Bottom Right: Correction for the telescope feed formed using a normalised covariance matrix.

## 3.5 Conclusions

In this chapter the RFI excision method of spatial filtering was introduced. SVD was used to extract and sort the eigenvalues and eigenvectors from the covariance matrix. The theoretical and practical application of spatial filtering to astronomical data was discussed, and the importance of gain calibration shown. The methods of noise and interference subspace projection were shown to be equivalent to nulling the primary eigenvalue(s), and the danger of removing noise and astronomical signal along with the interference was highlighted. A method of compensating for this by replacing the interference subspace eigenvalues with noise estimates was then reviewed.

Chapters 4 and 5 deal with the application of spatial filtering to data from the Parkes radio telescope and the MWA and the results obtained.



# CHAPTER 4

---

## SPATIAL FILTERING USING THE PARKES MULTIBEAM RECEIVER

Chapter 3 introduced the application of spatial filtering to multi-receiver systems and how it is applied to RFI suppression. The literature on this topic to date has focused on antenna arrays where both the RFI and celestial signals appear at all receivers. This chapter explains the first application of this technique to an array feed system, where each feed has independent celestial signal and system noise, but only RFI in common among the receivers. The study tests the effectiveness of RFI removal centred around the Multibeam Receiver on the Parkes 64 m dish.

Since the signal processing capacity of the Parkes Observatory's standard spectrometer for the Multibeam Receiver is inadequate to compute the covariances for all signal paths, a customised baseband recording system was constructed to take data directly from the samplers (analog to digital converters) in the 13 feed dual polarisation Multibeam Receiver. Covariances were computed by playing back the recordings into offline correlation and analysis software. A reference antenna was also used to record data simultaneously from one of the interfering transmitters nearby, while excluding celestial sources. The effectiveness of eigen analysis was then tested for various integration times and with varying levels of RFI. The results demonstrate that spatial filtering techniques provide powerful tools for RFI removal for single dish telescopes when using an array feed receiver.

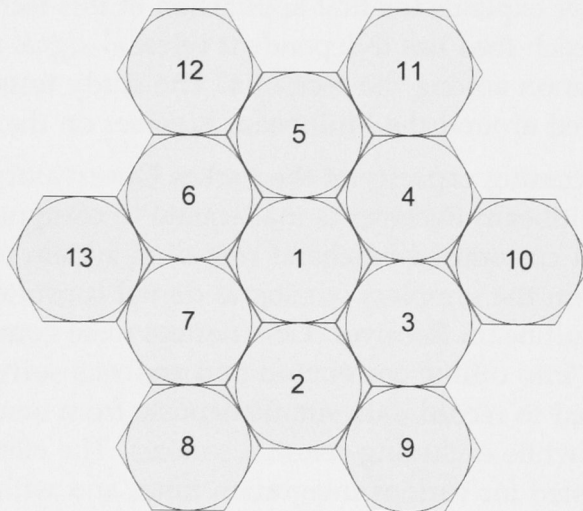
### 4.1 Experiment outline

There are two components to implementing spatial filtering techniques on the Parkes Multibeam Receiver system. The first concerns the required hardware modifications to the receiving and recording system, and the second involves the new applications of eigen decomposition for RFI suppression on a single

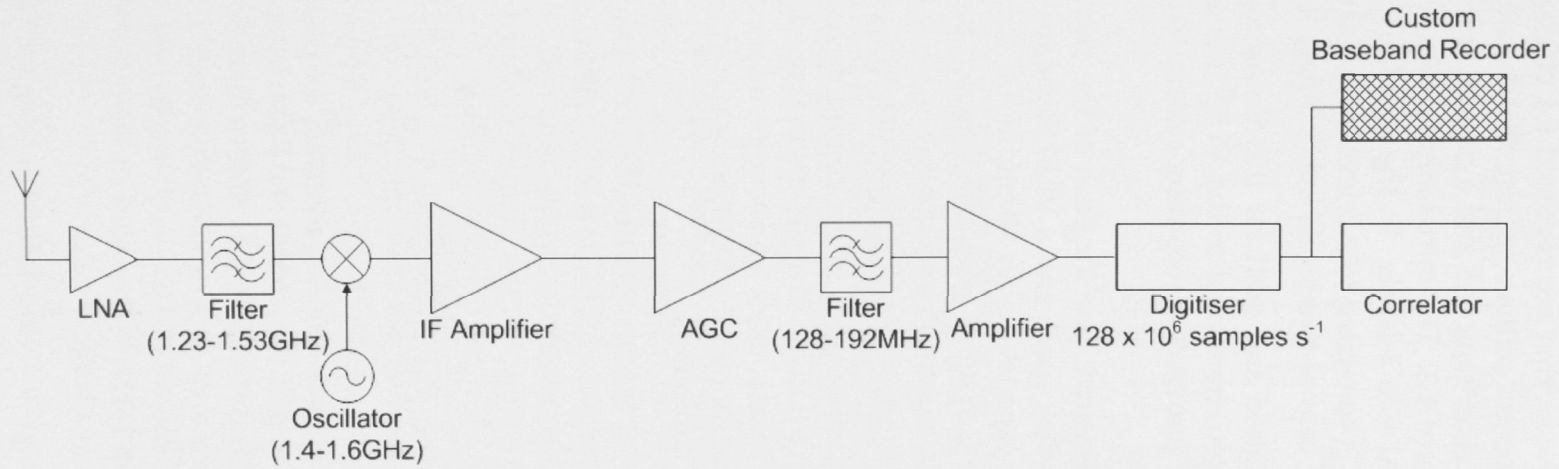
dish receiver. The instrumentation requirements are discussed in §4.2, and the results of the spatial filtering application in §4.3.

## 4.2 Instrumentation

The Multibeam Receiver consists of 13, dual polarisation beams co-located with hexagonal symmetry in the focal plane of the Parkes telescope (figure 4.1). The receivers cover a frequency range from 1.23-1.53 GHz, with baseband filters of 8, 16, 32 and 64 MHz (shown in figure 4.2). In this experiment, each of the 26 signals was bandlimited to 8 MHz and sampled at the Nyquist rate with 2 bits per sample. The bandwidth was limited by the recording capability of the two computers that supported this exercise. The net data rate was 1024 Mbits/sec.



**Figure 4.1** Beam layout for the Parkes Multibeam Receiver. The receiver consists of 13, dual polarisation beams, co-located with hexagonal symmetry in the focal plane of the Parkes telescope.



**Figure 4.2** Signal chain for each polarisation of each receiver feed. In this experiment the data was concurrently sent into a computer recording system after sampling, as well as passing on to the conventional Parkes Multibeam Correlator.



The Parkes Multibeam correlator can compute power spectra for all 26 data channels for bandwidths as high as 64 MHz, but does not have the capability to form the cross power spectra that are required for eigen analysis. In order to form the cross power spectra offline, an interface board was built to run between the Multibeam Receiver samplers and a custom recording system. The signal path is shown in figure 4.2. The initial signal voltages are captured by the receivers, and quantized into 4-levels. In the standard signal path of the Parkes correlator the signal is converted to a 3-level quantization mode prior to correlation by grouping the central two levels into a single level.

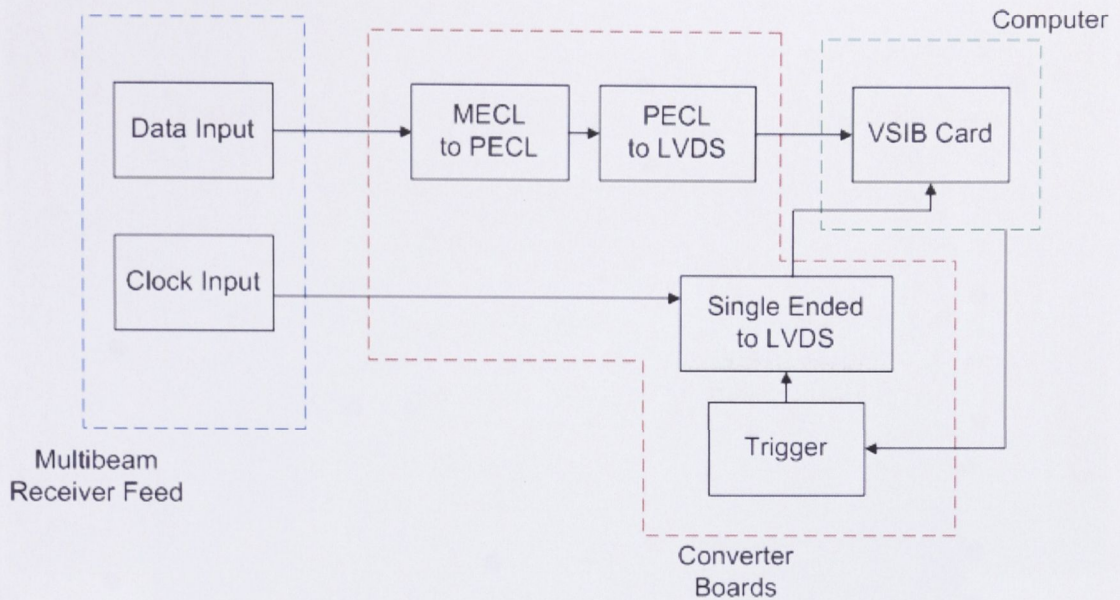
In the custom system, designed and built for this thesis research, the 4-level (2-bit) samples are intercepted, translated to the VLBI Standard Interface (VSI) standard used in VLBI, and rerouted to recording systems similar to those now used in e-VLBI observations.

Two PC computers were required, one for each of the two Multibeam racks. Each PC contained 1 TB of disk space and a VLBI Standard Interface B (VSIB) (Ritakari & Mujunen, 2002) card taking an 80-pin input. Capable of recording 512 Mbits per second, both PC systems were able to record the data clocked into the VSIB interface as 32 parallel bits at 16 MHz.

The digital output from the samplers was in Motorola Emitter Coupled Logic (MECL), while the VSIB recording cards in the computer system required Low Voltage Differential Signalling (LVDS). The signals that needed to be translated included the digital data streams, clock and trigger information. The trigger signal initialises the recording through the VSIB cards and precisely synchronises data acquisition on the two computers.

The output from the sampling system was in MECL format, but unfortunately off-the-shelf components to translate MECL into LVDS were not readily available. This meant that the MECL signals needed to be converted into an intermediate stage prior to the LVDS conversion. Positive Emitter Coupled Logic (PECL) was chosen for this task. The dataflow for this conversion system is represented by the block diagram in figure 4.3. Both magnitude and sign bits were required to be translated for each of the receiver polarisations. Separately from these signals, the clock input needed to be converted from a  $0.2V_{pp}$  single-ended signal into LVDS. To ensure that the clock was properly received, a comparator was used to steady the clock input. The use of this comparator restricted the input clock to 32 MHz, which was in excess of the 16 MHz requirement for this experiment. It was later verified in laboratory experiments that with a sufficiently stable clock signal, the comparator could be bypassed, allowing the board to drive the signals at much higher ( $> 128$  MHz) frequencies in the event that increased data recording capacity becomes available.

The final element of the design was the ability to “trigger” the VSIB card and start data recording. As there were two computer systems each with its own VSIB card, they needed to be triggered simultaneously in order to keep the data



**Figure 4.3** Required dataflow and input signals for the VSIB recorders. The data signals are taken from the Multibeam samplers. The clock is fed through a separate cable. There are two conversion stages, from the native MECL exiting the samplers, to PECL, and then into LVDS format for the VSIB card readers. The VSIB card also needs to be told when to start recording, which is done by a trigger signal: a pulse is generated via a computer serial port, which switches a transistor, allowing a voltage signal to pass to the card.

streams from both Multibeam Receiver racks synchronised. This was done by using a pulse generated via the serial port on one of the computer systems that was connected to both of the VSIB cards. The pulse from the serial port effectively turned on a transistor, allowing an appropriate input voltage to be translated into LVDS and triggering the VSIB cards.

Photos of the populated board are shown in figures 4.4 and 4.5. Figure 4.6 shows the boards in the chassis for the Multibeam Receiver rack (this chassis was built by Tim Ruckley at Parkes). The schematics and PCB layout of the entire design can be found in Appendix B.

The recording system implemented had the ability to take 32 input data streams. As only 26 inputs were required, one signal was duplicated by recording it on both of the computers to check for consistency and phase lock. A separate, single polarisation reference antenna feed was also added. This left a set of 28 connected inputs and four blank. With each PC recording at 512 Mbits per second, the computers were able to record 8 MHz of the 64 MHz bandwidth available (32, 2-bit inputs, 16 MSample/sec).

The Multibeam Receiver data path includes variable attenuators to set the power levels into the samplers so that each of the 2-bit sampler outputs can be mapped to a relative voltage level as shown in table 4.1.

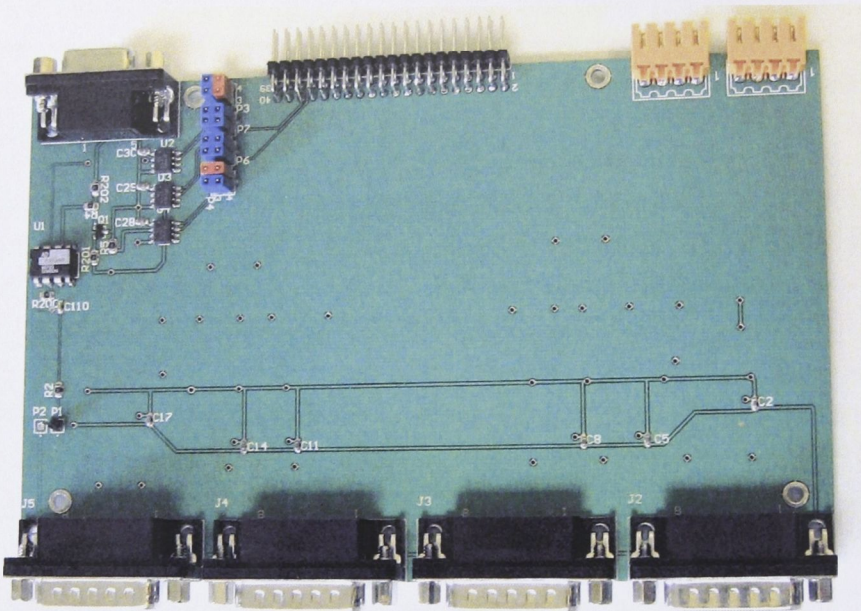


Figure 4.4 Top side of populated converter circuit board.

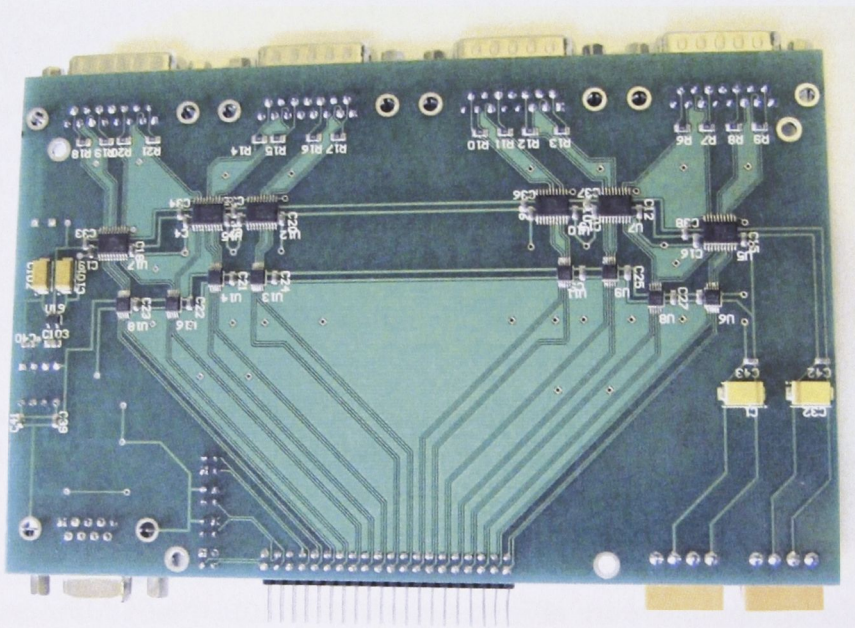


Figure 4.5 Bottom side of populated converter circuit board.

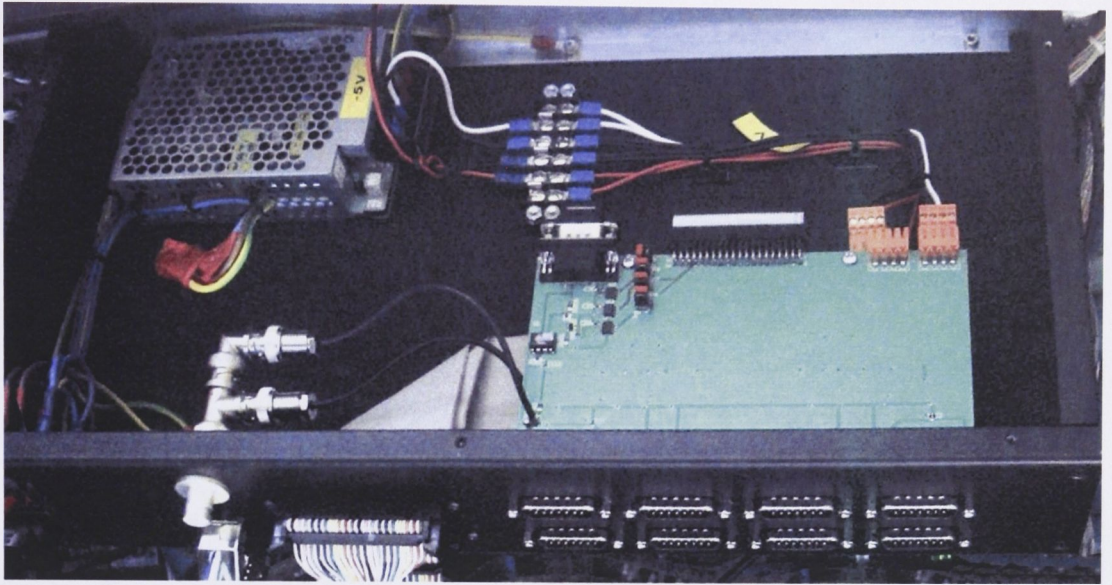


Figure 4.6 Converter circuit boards in Parkes chassis.

Bits 1-16	Mapping	Bits 17-32	Mapping
00	4	00	-1
01	-4	01	1
10	1	10	-4
11	-1	11	4

Table 4.1 Bit mapping for 1-32 bit stream recorded by the VSIB recorders.

### 4.2.1 Data processing software

The first step of the offline software process was to decode the bit streams from each receiver. Then, the 28 data lines were Fourier transformed into the frequency domain and auto- and cross-correlation spectra were computed. These spectra typically had a channel resolution of 15.625 kHz, providing 512 spectral channels for the 8 MHz of bandwidth.

At this point the data was normalised and underwent the required decomposition and interference detection. These two steps are outlined in greater detail in section §4.3.2.

The initial software to accomplish all of these steps was written in MATLAB®. Later, the mapping and correlation functions were translated into Fortran for more efficient processing.

## 4.3 Results

### 4.3.1 RFI environment

The RFI environment in the frequency bands for the multibeam receiver (1.23-1.53 GHz) is generally clean at Parkes. There are some areas of the spectrum, however, containing strong interfering sources. Two of these interferers, both transmitting at 1438.5 MHz, are located around 100 km away, one to the North East (Wellington), the other to the South (Grenfell).

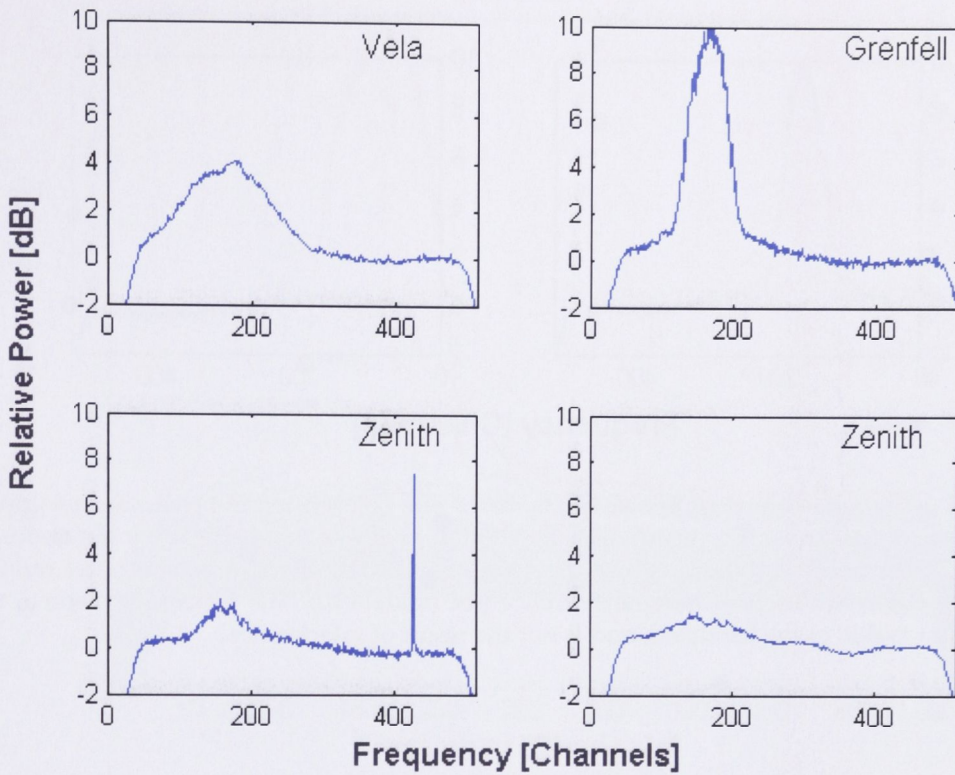
The Grenfell transmitter creates a strong Gaussian peak approximately 1.25 MHz wide. The interference signal from Wellington is broader, about 3 MHz in width, and varies in strength. The effect of these signals on the total power spectrum is demonstrated for three different orientations of the dish in figure 4.7. Details of the observations taken with the Multibeam Receiver can be found in entries 5-8 in tables D.1 and D.2.

### 4.3.2 Interference removal

Various methods for detecting the presence of interference in a frequency band are discussed in Leshem et al. (2000). For the purposes of this Parkes experiment, automatic detection of interference was not implemented. Once the data were recorded both total and cross power spectra were calculated. A sample total power spectrum is shown in the left panel in figure 4.8. An interference free spectrum, is shown for reference purposes in figure 4.9.

The singular value decomposition of the resulting matrices was then taken. The eigenvalues for the total power spectrum shown in figure 4.8 are given in figure 4.10. Once the spectra are successfully decomposed, the RFI signals can be identified and nulled. The reconstructed spectrum after nulling is shown in the right panel of figure 4.8. In this case the first two eigenvalues shown in figure 4.10 represent the interference from the two 1438.5 MHz transmitters at Grenfell and Wellington. Reconstruction making use of the reference antenna showed that the stronger peak originated from Grenfell, and the weaker from Wellington in this orientation of the dish.

The Grenfell RFI is located approximately in channels 120-200 of the first eigenvalue spectrum. It is then represented in the second eigenvalue spectrum in the 100-110 and the 200-210 channel ranges. The Wellington source is located in the second eigenvalue spectrum in the 120-200 channel range, and the first eigenvalue spectrum in channels 90-110 and 200-250. This is because it is a broader source than the Grenfell transmission, and in these channel ranges is the strongest coherent signal in the covariance matrix. This demonstrates that a signal does not always stay in the same eigenvalue, but can switch between

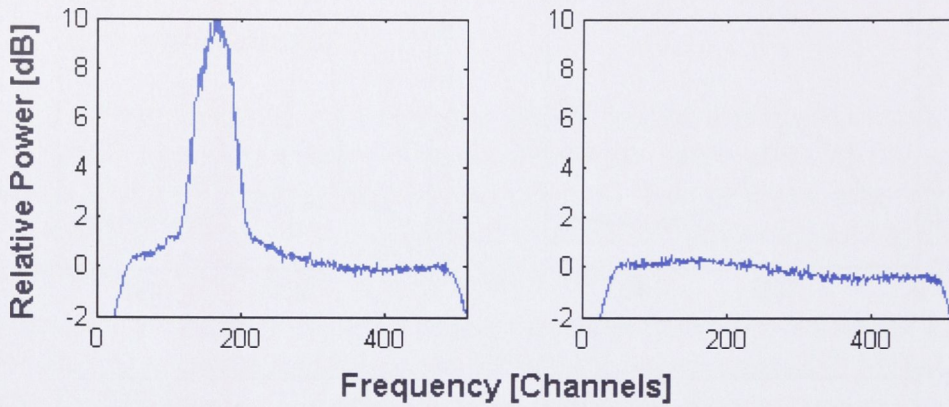


**Figure 4.7** Total power spectra demonstrating the RFI variation in Parkes dataset. Each figure shows a total power spectrum from one polarisation of the multibeam receiver (beam one) at different orientations and different times, with a centre frequency of 1440 MHz. Top Left: Interference present while observing Vela pulsar. (Taken 1027 AEST, 2nd Jul, 2008, RA: 08:35:20, DEC: -45:10:34,  $T_{int} = 1$  s.) Top Right: Dish directed toward the interferer at Grenfell (Taken 1237 AEST, 28th Feb, 2008, AZ: 340, ZEN: 20,  $T_{int} = 0.1$  s.). Bottom Left: Dish at zenith (Taken 1344 AEST, 28th Feb, 2008,  $T_{int} = 0.1$  s.). Bottom Right: Dish at zenith (Taken 1311 AEST, 1st Jul, 2008,  $T_{int} = 1$  s.). The two bottom plots demonstrate particularly the variable strength of the RFI sources.

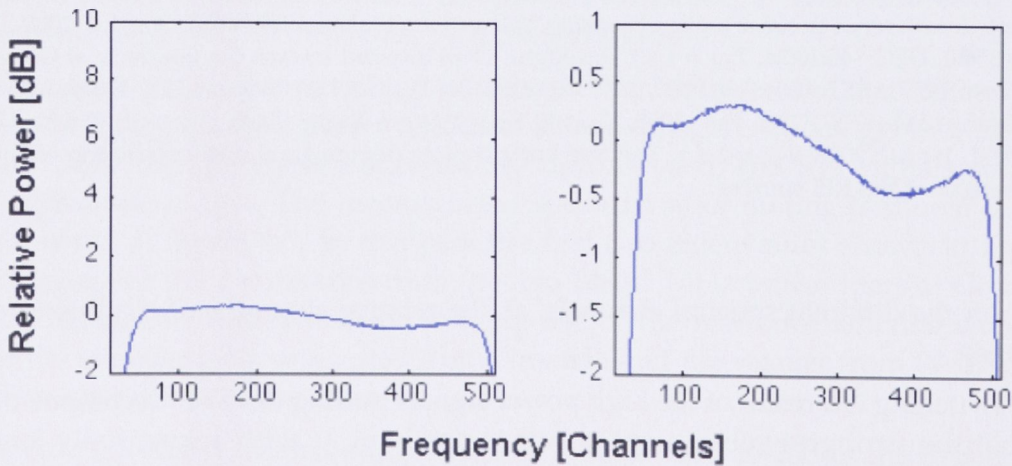
them in the different spectral channels, as the relative strengths of the interferers changes.

This switching is a result of the high power signals pulling the SVD decomposition so that the strongest subsequent power is partly in  $\lambda_1$  with sequentially lower powers present in  $\lambda_2, \lambda_3, \dots, \lambda_{26}$ . These eigenvalues are calculated subject to the constraint that their corresponding eigenvectors are orthogonal to those belonging to the higher power eigenvalues.

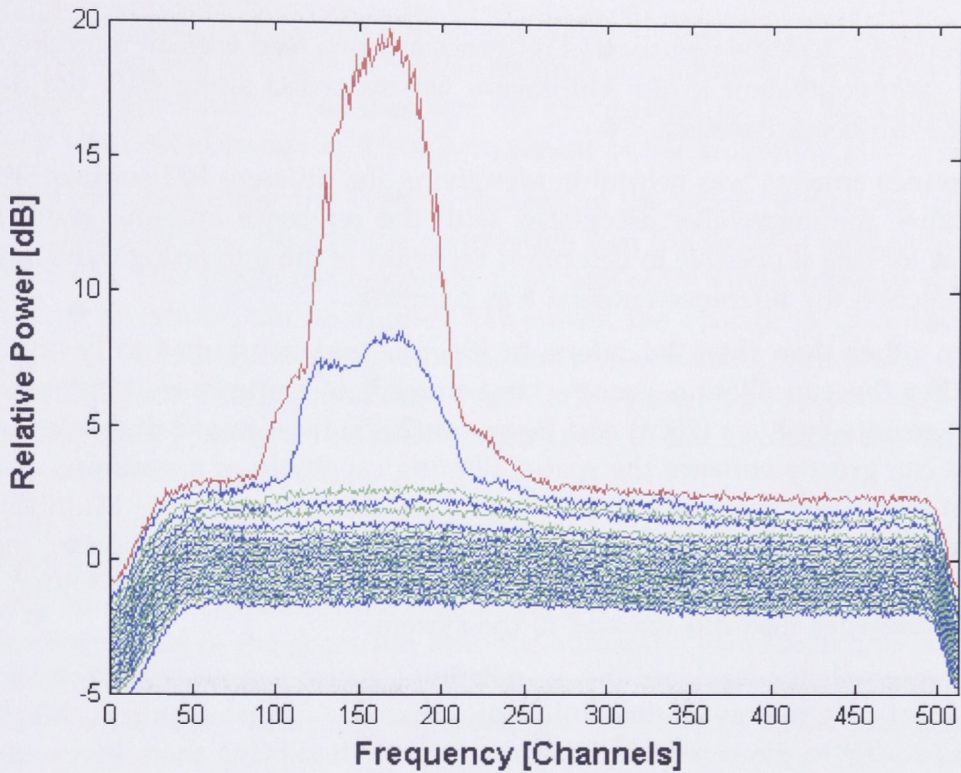
In the RFI contaminated part of the spectrum, the RFI occupies  $\lambda_1$  and  $\lambda_2$ , and their orthogonal eigenvectors are aligned accordingly. The SVD algorithm then moves on to the next strongest power, which is likely to be close to that of the peak  $\lambda_1$  in the RFI-free part of the spectrum. Some deviation is expected, since



**Figure 4.8** The autocorrelation power spectrum for one polarisation of beam one, integrated for about 0.1 seconds. The bandwidth for the figure is 8 MHz. Left: The interference at 1438.5 MHz stands out clearly above the noise. Right: The power spectrum after nulling the two eigenvalues representing interference (see figure 4.10). The remaining shape in the spectrum is due to the bandpass, and is not the result of interference.



**Figure 4.9** Interference free power spectrum example from the Parkes dataset. The bandwidth for the figure is 8 MHz. Left: The power spectrum at a scale comparable to figure 4.8. Right: The same power spectrum with improved scale.



**Figure 4.10** Eigenvalue spectra for the full 13 beam dual polarisation inputs. The top (red) line is the first eigenvalue, plotted as a function of frequency. As expected it contains the majority of the interference. Significant power is also contained in the second eigenvalue, which is a result of the second transmitter at the same frequency creating a different RFI footprint. Subsequent observations indicate that where the “second” transmitter extends in frequency beyond the first, the wings of the profile are contained in the first eigenvalue spectra. The lines below indicate the 3-26th eigenvalues.

the orthogonality of each additional eigenvector provides an addition constraint. This means that eigenvalues  $\lambda_3, \lambda_4, \dots, \lambda_{26}$  in the contaminated spectral channel range (around 90-250) must capture the high noise powers that were captured by  $\lambda_1$  and  $\lambda_2$  in the RFI free spectral channels.

One consequence of this method of decomposition is that further attention must be given to the transition between the region where  $\lambda_1$  contains RFI and where  $\lambda_1$  represents noise and celestial signals. As will be addressed in §4.3.3, nulling  $\lambda_1$  across the entire spectrum has the potential to throw away valid celestial signal.



## **Reference antenna**

As stated in §4.2, a single polarisation reference antenna feed with an interference to noise ratio equivalent to the Multibeam was recorded along with the data from the Multibeam Receiver.

This reference antenna was helpful in identifying the different RFI sources. This was because the eigenvalue associated with the reference antenna could be identified, making it possible to determine the effect of the interfering transmitter in the direction the reference antenna was pointing.

However, other than this, the reference antenna was not found to be useful in assisting the cancellation process, and indeed, in some cases hindered it. Previous results by Kocz (2004) and Boonstra (2005) have found that reference antennas can greatly enhance the spatial filtering capability of a synthesis array, aiding in separating the interference eigenvalues. In the case of the Multibeam Receiver, because the beams are aimed at different directions on the sky, they are all effectively reference antennas containing the same RFI, so the addition of the extra reference feed did not add to the system.

The reference antenna hindered the cancellation because it recorded interference that was not seen by any of the Multibeam receivers. This additional interference was added to the eigenvalue spectra, further disrupting them by causing additional switching of eigenvalues between noise and interference.

### **4.3.3 Projection variations and dealing with strong astronomy sources**

As stated in previously in §3.3, problems occur when the assumption that the astronomy signals are negligible for the duration of the covariance matrix integration breaks down. An example of this is demonstrated with observations of the Vela pulsar (J0835-4510). These observations also illustrate limitations, advantages and disadvantages of some of the different ways the interference projection can be applied.

The Vela dataset was taken with the Multibeam Receiver, with the central beam observing the source. Approximately five minutes of data were taken. The analyses presented here are based on subsets of the data, dealing with it in one second intervals, as this is a convenient timescale for illustration purposes. On long timescales reorientation of the dish and variation in atmospheric propagation alter the effective gain for the interferers so that separate decompositions are required to track the changes, effectively “refreshing” the eigenvector space.

The dataset was taken by John Reynolds and Jonathon Kocz, 1027 AEST, 2nd July 2008. Unrecognised at the time, an automatic gain correction was being applied to the input. The purpose of this automatic gain correction is to take

maximum advantage of the number of quantisation bits available. The effect of the automatic gain correction is noticeable in the higher-frequency pulsar detection figures as a dip in power immediately following the pulse (see for example figure 4.22). This dip, and the corresponding rise back to the noise power level corresponds to the time constant of the automatic gain correction.

Other than assaulting the eyes by disturbing the pulse profile, the automatic gain correction also violates the assumption that the receiver gains are constant with time. This limits the degree of precision with which some of the spatial filtering variations can be applied. However, the change in gain affects both  $g_A$  and  $g_I$ , which allows a demonstration of the vulnerabilities and strengths of the different filtering methods. Specifically, some applications of the spatial filtering theory show a remarkable immunity to the effects of the unplanned gain fluctuations (see §4.4).

For lower-frequency pulsar detection (figure 4.23), the automatic gain correction causes a dip in power both before and after the pulse. This dip before the pulse is due to the dispersion of the pulsar. As the pulsar signal reaches the higher frequency areas of the spectrum first, the automatic gain correction activates to keep the integral power in the band constant. This causes a reduction in the power in the interference region, creating a dip relative to the previous power level. While unintended, this automatic correction has the affect of making the constant RFI in the band behave like an RFI source with varying power. Consequently, in the correction the effects of the automatic correction are removed along with the interference, further illustrating the robustness of this type of cancellation.

Application of the method from §3.3 requires that the algorithms are applied to each frequency channel separately using the full covariance matrix containing the 26 inputs from the Multibeam Receiver. The covariance matrix needs to be recalculated for each timestep, and the duration of each timestep is a function of the type of science being extracted from the observation and the stability of the eigenvectors (gains) for the RFI signal paths.

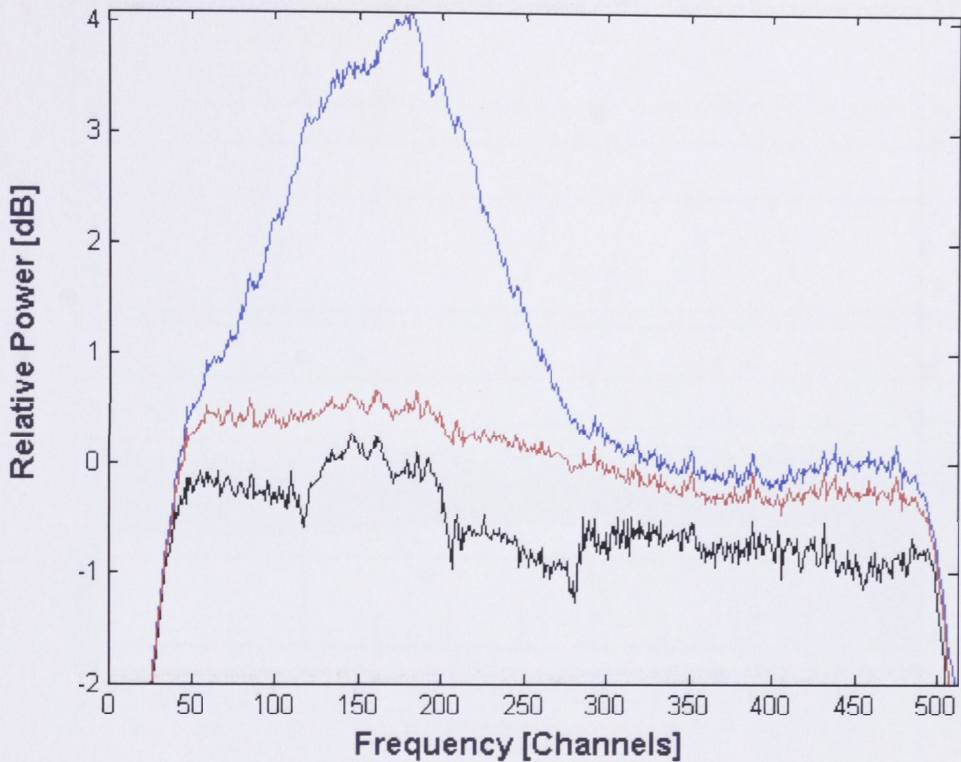
Furthermore, in practice, there are a range of possible strategies for applying excision algorithms. These choices include, but are not limited to:

- Application of the algorithm blindly across the entire frequency spectrum versus identifying and targeting only contaminated channels.
- Accounting for the noise power and astronomical signal that is included with the RFI when it is nulled and restoring it in a statistical sense, versus ignoring the lost signal.
- Recomputing the decomposition on short time-scales (which may be noisy), versus “creating an average correction” (a filter) from a decomposition over a long time-scale, that can be applied on arbitrarily short time steps.

The primary aim of course, is to leave the portions of spectra without interference untouched, and remove the interference from those areas that are contaminated. The different applications mentioned accomplish this with varying degrees of success. Care must be taken, however, to ensure that the improvements in a spectrum are not just cosmetic. By removing more eigenvalues than necessary, and scaling by the appropriate noise variance, it is possible to get a spectrum that appears completely interference free (for example a flat passband for a continuum source). In this case it is probable that astronomical signal of interest would also be removed. The amount of signal removed depends on the strength of the source, and the number of eigenvalues taken. The stronger the source the more likely it is that it will be removed, as it will be present in a higher eigenvalue. To minimise this it is essential that only the eigenvalues representing the interference are removed, although even these are likely to take a portion of the astronomy signal away (see §3.3). The Multibeam is more robust to this than a normal array, as the beams look at different portions of the sky, and the therefore will not have correlated power from the astronomical source in the covariance matrix elements.

The rest of this section examines in detail the different effects of spatial filtering on the portions of the spectrum that contain interference and those that are interference free. The covariance matrix decomposition is performed on 1 ms timescales (the pulse period for Vela is approximately 0.0909 seconds). Also shown is the effect of accounting for noise covariance and astronomy bleed-through into the eigenvalues that are nulled. The strength of the Vela source (mean flux density 1100 mJy at 1400MHz) and the integration time of 1 ms means that in the ideal interference-free case, the pulse would be detected with an average SNR of around 30, if a selection of 150 channels (2.34 MHz) was taken from the total 8 MHz bandwidth.

The original and corrected spectra in figure 4.11 illustrate that the spatial filtering algorithm is able to remove interference from the spectrum. Two things are readily apparent in this figure. The first is the disparity between the result when the subtracted noise is ignored or not. It is obvious that simply removing the first two eigenvalues, results in too much signal being removed from the interference area. It can also be seen that in the area outside the interference, the spectrum is also lacking in power, regardless of the method. In the case of the black spectrum, this is due to the straight removal of noise along with the interference. In the case of the red spectrum, while an attempt was made to scale the interference removal by the appropriate  $\Lambda_I - \sigma_n^2$  (as per equation 3.26) the estimate of  $\sigma_n$  was not made accurately enough. The estimate was slightly low biased as the eigenvalues containing the strongest contributions to the noise power were the same as those removed to cancel the interference. This can be compensated for, but as discussed below is unnecessary in the pulsar observation. Figure 4.12 shows the eigenvalues from the decomposition of the covariance matrix from which figure 4.11 is taken. There are two distinct interfering sources,

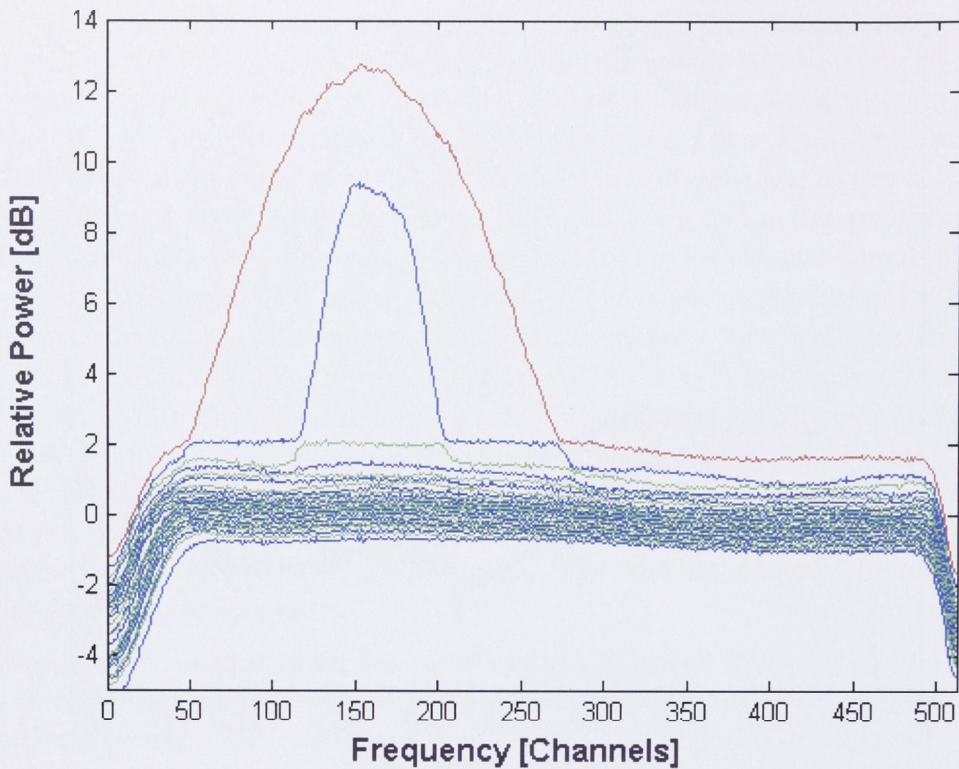


**Figure 4.11** Power spectrum snapshot from observations of Vela pulsar at Parkes. The observations were made using beam one of the Multibeam, with a central frequency of 1440 MHz, and a bandwidth of 8 MHz. The integration time was 1 s. Blue: The original spectrum. Red: The corrected spectrum where the noise power estimate has replaced the first two eigenvalues. Black: The corrected spectra where the first two eigenvalues have been nulled.

located in the first two eigenvalues.

After removing the interference it is necessary that it is still possible to detect the pulsar. Due to the strength of the Vela pulsar the simple method of determining the total power in a frequency range over time was used for detection. Figure 4.13 illustrates the results of detecting the Vela pulsar in an interference free section of the spectrum (channels 300-450) after various spatial filtering methods have been applied. The purpose of this figure is to illustrate any risks of applying the correction to an interference-free region. There are four panels to this figure, where different correction strategies have been used. The first panel shows the pulsar detection without applying any correction. The second panel shows the effect of completely removing the two primary eigenvalues. In the third panel, the two primary eigenvalues have been replaced with a noise variance estimate. Lastly, the correction was applied only to the frequency channels with visible interference (rather than blindly across the spectrum).

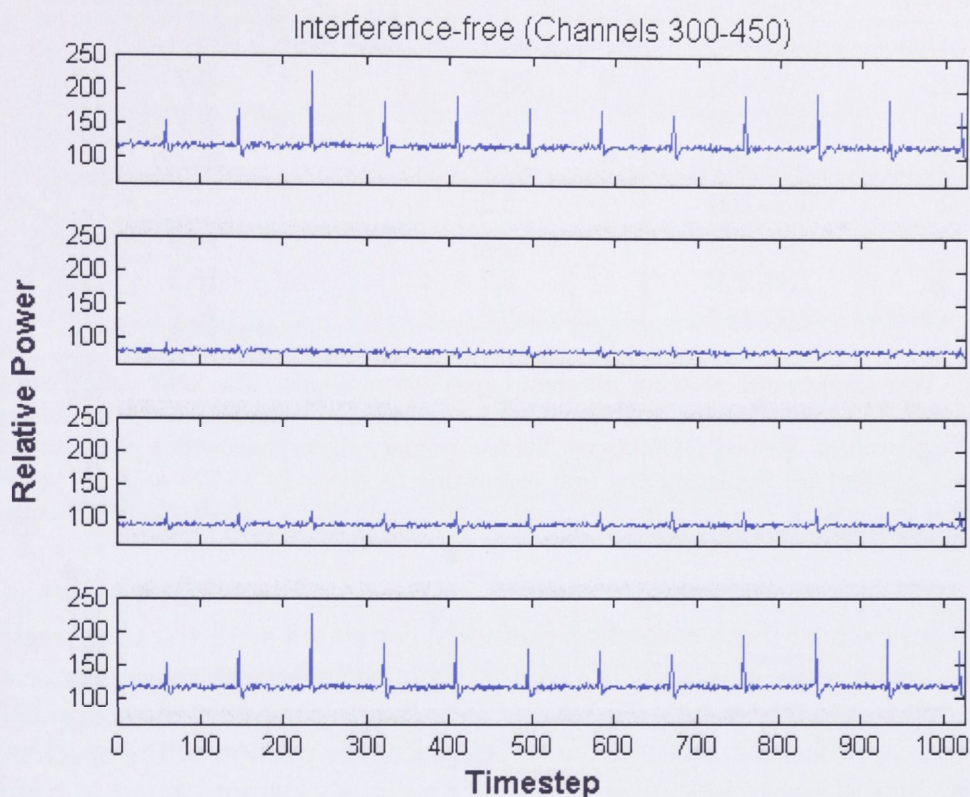
It was expected that the top and bottom plots in this figure would be identical,



**Figure 4.12** Eigenvalues from a power spectrum snapshot from observations of Vela pulsar at Parkes. The observation was taken at a centre frequency of 1440 MHz, with an 8 MHz bandwidth. The 26 eigenvalues are from a decomposition of a covariance matrix constructed from the 13 dual polarisation inputs from the Multibeam Receiver, and integrated for 1 s. In particular it should be noted that there are two interfering sources represented by the first and second eigenvalue (red and first blue line). It can also be seen that there is power above the noise represented in the first eigenvalue in the 300-500 channel range, which is represented by the second and third eigenvalues respectively in the areas of the spectrum contaminated by one, and then two, interferers.

as they are, as no correction was applied to this portion of the spectrum. In the second panel, the pulsar is almost entirely removed. This is because in the interference-free region, the two primary eigenvalues represent the pulsar signal, rather than any interference, and they have been completely removed. The third panel still shows some signs of the pulsar. This is because the two primary eigenvalues have not been entirely removed, but rather replaced with a noise variance estimate. As this estimate is based on the “interference-free” eigenvalues (3-26), the power is reduced. The pulsar is clearly visible in both the top and bottom panels. This figure shows that blind application of the correction to interference-free regions of the spectrum can destroy information that it is desirable to keep.

Figure 4.14 shows attempts to detect the Vela pulsar in a portion of the spectrum



**Figure 4.13** A comparison of the Vela pulsar detection strength in RFI non-affected spectral channels (channels 300-450) after applying different RFI correction methods. Top: No correction applied. Second: Nulling the two primary eigenvalues. Third: Replacing the two primary eigenvalues with a noise variance estimate. Bottom: Replacing the first eigenvalue in channels 46-476 with a noise estimate computed for each frequency channel, and replacing the second eigenvalue in channels 114-205 with a noise estimate computed for each frequency channel. The average SNR for each method is given in table 4.2.

that was contaminated by interference (channels 100-250). Again, there are four panels to the figure, showing the effect of different correction strategies. The detection of the pulsar in the top panel, where no correction has been applied looks almost like an inverse pulsar detection due to the automatic gain calibration. The second panel, in contrast to the case where an interference-free channel region was used (figure 4.13), shows the pulsar clearly. This makes sense as the two primary eigenvalues in this case represent the interference. The removal of these eigenvalues does not greatly disturb the pulsar detection, lowering the average power level, but not affecting the signal to noise ratio. The third panel has an equivalent signal to noise ratio as the second, but in this case the estimate of the noise power was returned to the spectrum, using eigenvalues 3-26 to form the estimate. The bottom panel has a slightly improved detection over the other attempts. This is because a more accurate estimate of the noise power was made by taking into account the different widths of the two interferers.

Method	Channel range	SNR pre-correction	SNR post-correction
1	300-450	30.77	-
2	300-450	30.77	3.7
3	300-450	30.77	9.7
4	300-450	30.77	30.77
1	100-250	5.2	-
2	100-250	5.2	14.0
3	100-250	5.2	16.2
4	100-250	5.2	16.5

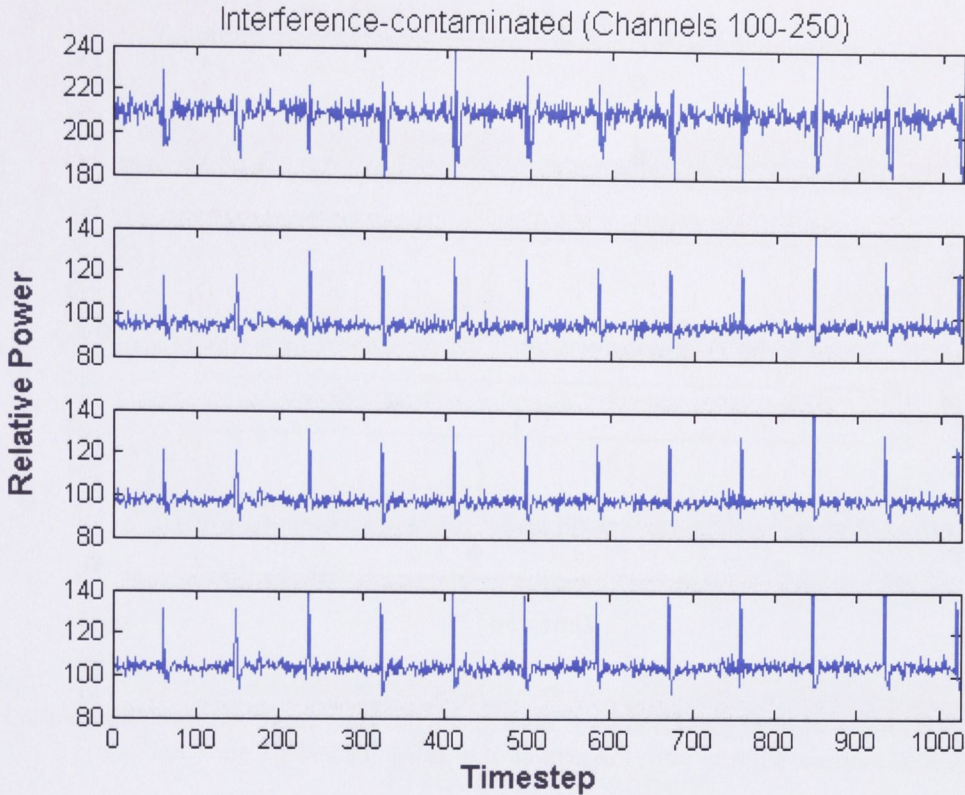
**Table 4.2** Vela single pulse SNR for different correction methods. The SNR values are an average, taken over five pulses. Method (1): No correction. Method (2): Nulling the two primary eigenvalues. Method (3) Replacing the two primary eigenvalues with a noise variance estimate. Method (4) Replacing the first eigenvalue in channels 46-276 and the second eigenvalue in channels 114-205 with a noise estimate computed for each frequency channel.

The SNR for each of these correction methods, compared with the uncorrected SNR and interference free SNR is given in table 4.2. The values given are an average, taken over five pulses. The main factor keeping the post-correction SNR values from approaching those of the interference-free portion of the spectrum, is the low bias of the noise power estimate, previously demonstrated in figure 4.11. The residual interference, from equation 3.27, will be suppressed to at least -12dB (a much higher suppression can be achieved by using longer averages to estimate the interference, as will be demonstrated in §4.4). Any increase in the system noise due to these residuals will not noticeably effect the SNR, which is dominated by the noise power estimate.

Figures 4.15 and 4.16 show the behaviour of the eigenvalues for timesteps 0-400 in the interference-free and contaminated regions of the spectrum respectively. In each of these figures the power in the eigenvalue spectra, over channels 300-450 for figure 4.15, and channels 100-250 for figure 4.16, are added in the same manner as when detecting the pulsar.

Four pulses from the Vela pulsar arrive during this time. As the pulse arrives, the eigenvalues change. There is also a difference in the eigenvalue response between the contaminated and interference-free regions. Taking the interference-free region first (figure 4.15), there is a noticeable increase in the eigenvalue power across the spectrum when the pulse is present. In the region contaminated by interference (figure 4.16), there is a decrease in power in the first eigenvalue during the pulse, and an increase in the third eigenvalue. This decrease is due to the automatic gain calibration maintaining the constant power under the band, causing the interference (dominating the first eigenvalue) to be slightly reduced.

To gain insight into the relationship between the eigenvectors and eigenvalues and the effect of the pulse arrival on them, figures 4.17 to 4.19 show the behaviour



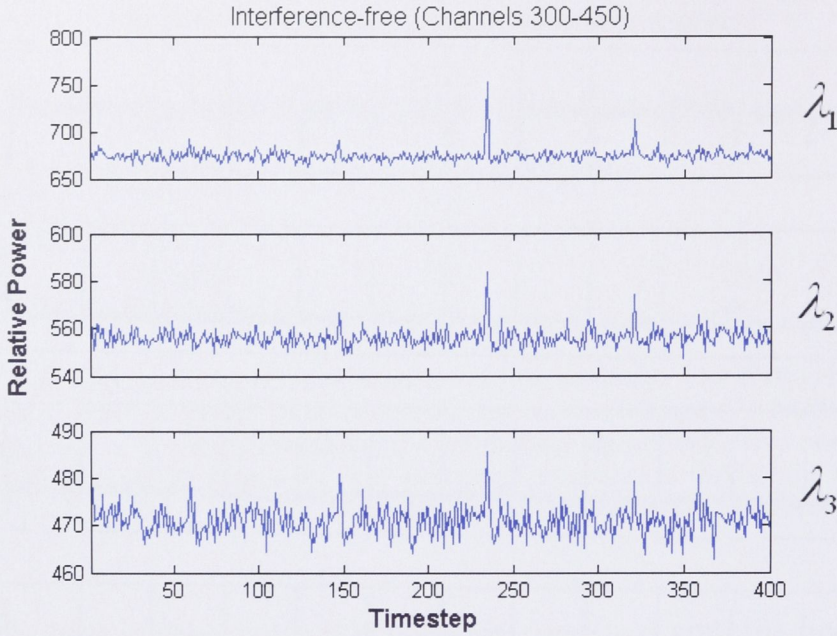
**Figure 4.14** A comparison of the Vela pulsar detection strength in RFI affected spectral channels (channels 100-250) after applying different RFI correction methods. Top: No correction applied. Second: Nulling the two primary eigenvalues. Third: Replacing the two primary eigenvalues with a noise variance estimate. Bottom: Replacing the first eigenvalue in channels 46-276 with a noise estimate computed for each frequency channel, and replacing the second eigenvalue in channels 114-205 with a noise estimate computed for each frequency channel. The average SNR for each method is given in table 4.2.

of the eigenvector coefficients for the timesteps immediately before, during, and after a pulse (timesteps 231-238). This relationship becomes particularly important for the filter implementation to be introduced in §4.4.

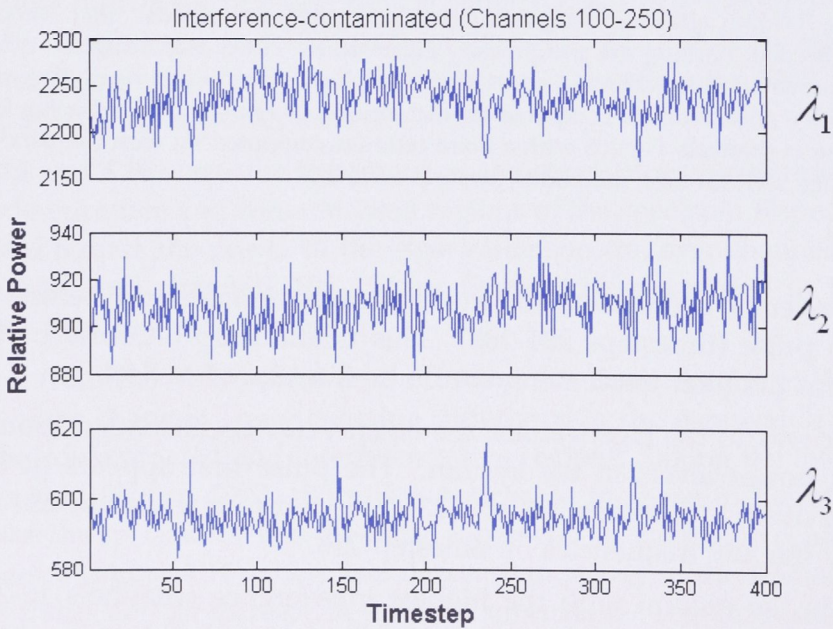
In a similar way to the eigenvalues, the eigenvector coefficients respond to the changes in power levels in the system. The pulse first appears in the high frequencies during timesteps 233 and 234, the mid-frequencies in 234 and 235, then finally the low frequencies by timestep 236.

Note that in the section with the highest interference (channels 10-250), the majority of the increase in magnitude due to the pulse arrival is in the second and third eigenvectors (figure 4.18 and figure 4.19), rather than the first eigenvector (figure 4.17). This corresponds to the pulse power being in the second and third eigenvalues in the interference contaminated region (figure 4.16), whereas in the interference-free region it is in the the primary eigenvalue (figure 4.15).

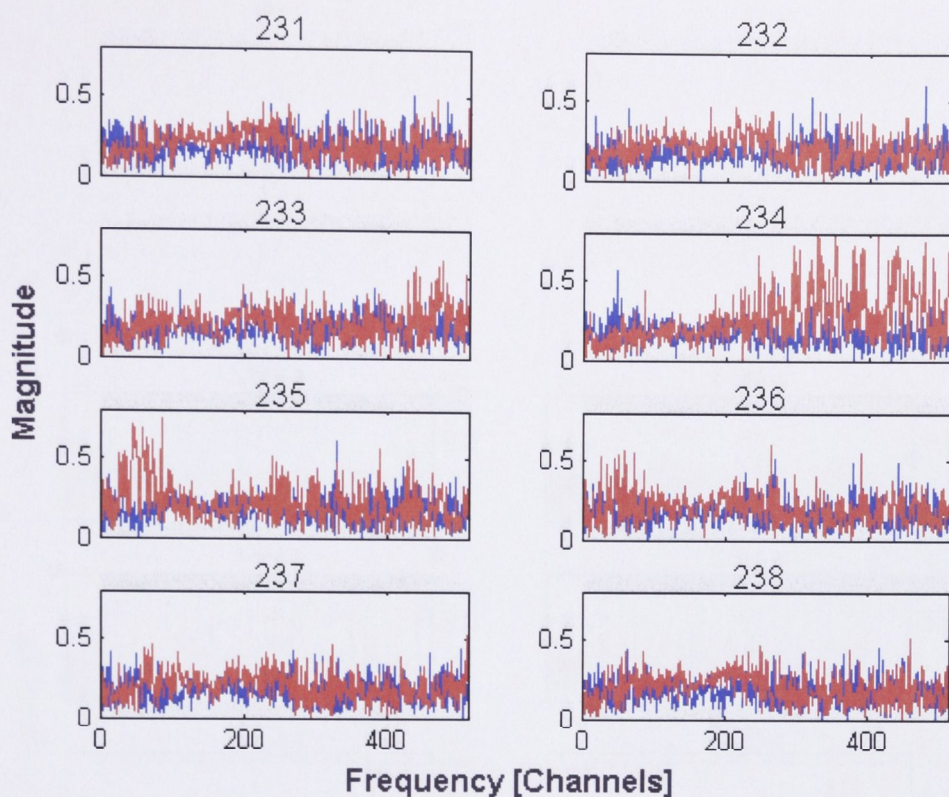




**Figure 4.15** First three eigenvalues for summed over the 300-450 spectral channel range. Top: Eigenvalue one. Middle: Eigenvalue two. Bottom: Eigenvalue three. All three eigenvalues are shown to increase slightly with the arrival of a pulse. Timesteps are 1 ms.

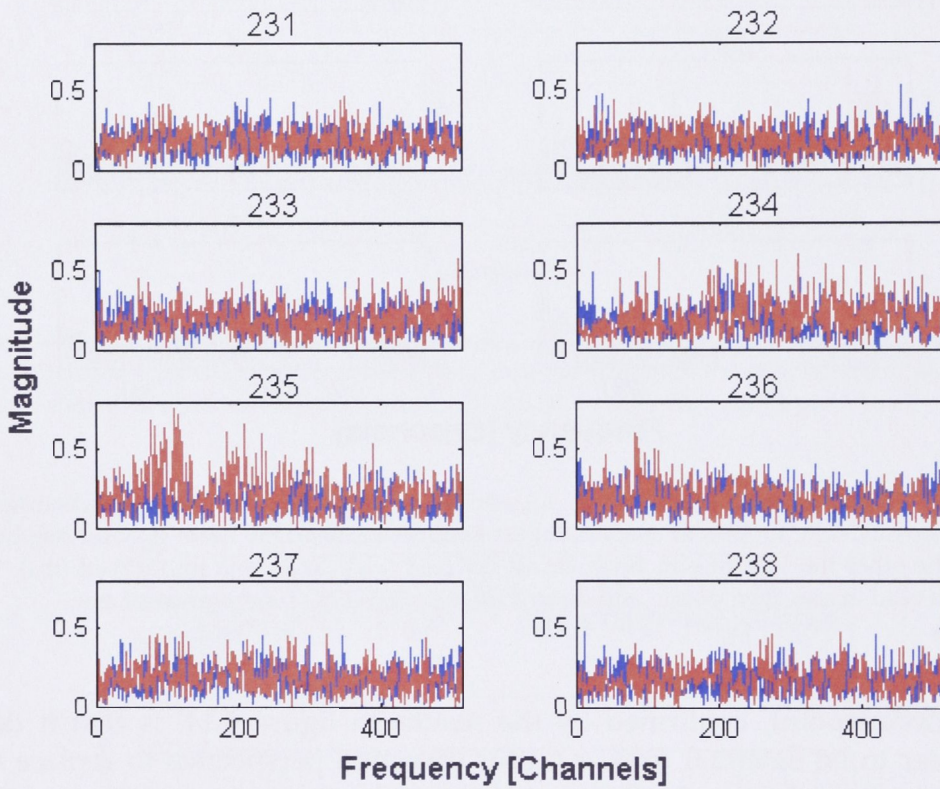


**Figure 4.16** First three eigenvalues summed over the 100-250 spectral channel range. Top: Eigenvalue one. Middle: Eigenvalue two. Bottom: Eigenvalue three. Only the third eigenvalue shows an increase in power with the arrival of a pulse. Timesteps are 1 ms.

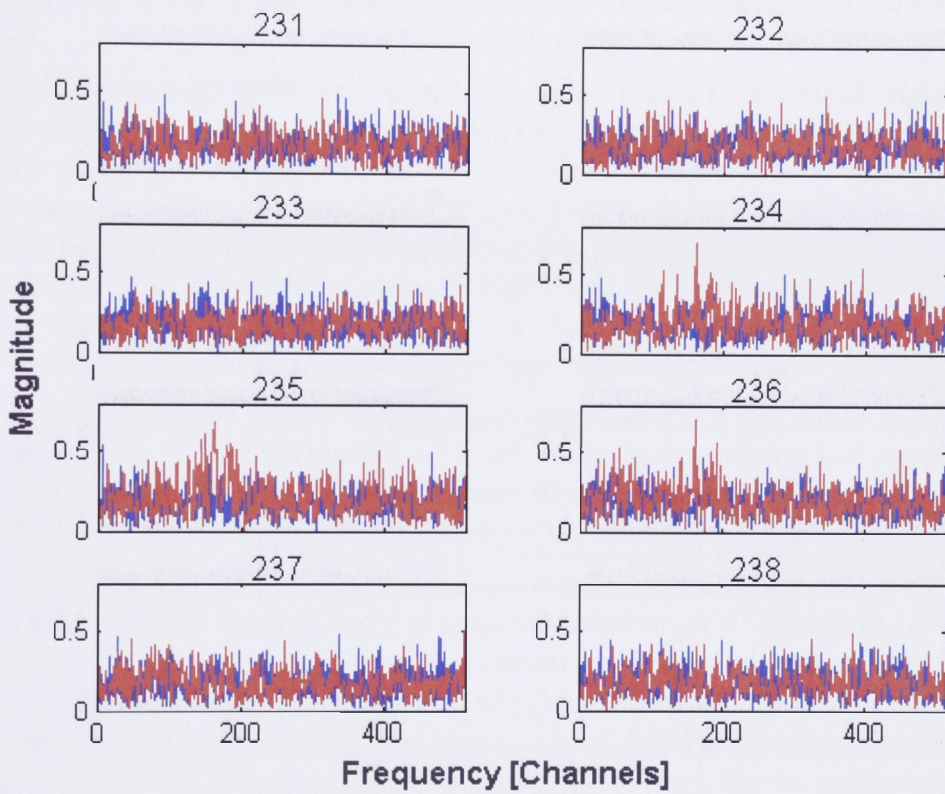


**Figure 4.17** Eigenvector coefficient one:  $u_{11}$  and  $u_{18}$  as a function of frequency channel for timesteps 231-238, for two of the Multibeam Receiver beams. One beam is observing blank sky, the other the Vela pulsar. Blue: Blank sky feed ( $u_{11}$ ). Red: Vela pulsar feed ( $u_{18}$ ). The panels read across then down, and cover timesteps 231-238. Timesteps are 1 ms.

An important point, confirmed by the results in figure 4.14, is that it does not appear to be essential from a pulsar detection perspective to replace the subtracted noise when working in the region contaminated by interference. This assessment becomes important later in terms of filter implementations. As foreshadowed in §3.3, figures 4.17 through 4.19 showed that the decomposition responds to changing power levels, affecting eigenvectors as well as eigenvalues.



**Figure 4.18** Eigenvector coefficient two:  $u_{21}$  and  $u_{28}$  as a function of frequency channel for timesteps 231-238, for two of the Multibeam Receiver beams. One beam is observing blank sky, the other the Vela pulsar. Blue: Blank sky feed ( $u_{21}$ ). Red: Vela pulsar feed ( $u_{28}$ ). The panels read across then down, and cover timesteps 231-238. Timesteps are 1 ms.



**Figure 4.19** Eigenvector coefficient three:  $u_{31}$  and  $u_{38}$  as a function of frequency channel for timesteps 231-238, for two of the Multibeam Receiver beams. One beam is observing blank sky, the other the Vela pulsar. Blue: Blank sky feed ( $u_{31}$ ). Red: Vela pulsar feed ( $u_{38}$ ). The panels read across then down, and cover timesteps 231-238. Timesteps are 1 ms.

## 4.4 Applying the projection as a filter for pulsar detection

One of the difficulties with the practical application of the method in §4.3.2 is the computational requirements for creating the full set of covariance spectra, taking the SVD, computing the appropriate correction projection and applying it in real time. A second difficulty is that pulsar studies rely on finely sampled time sequences of pulsed power, and these samples appear on such short time scales that adaptive filters cannot be derived without high levels of uncertainty, due to the noise level in short integrations.

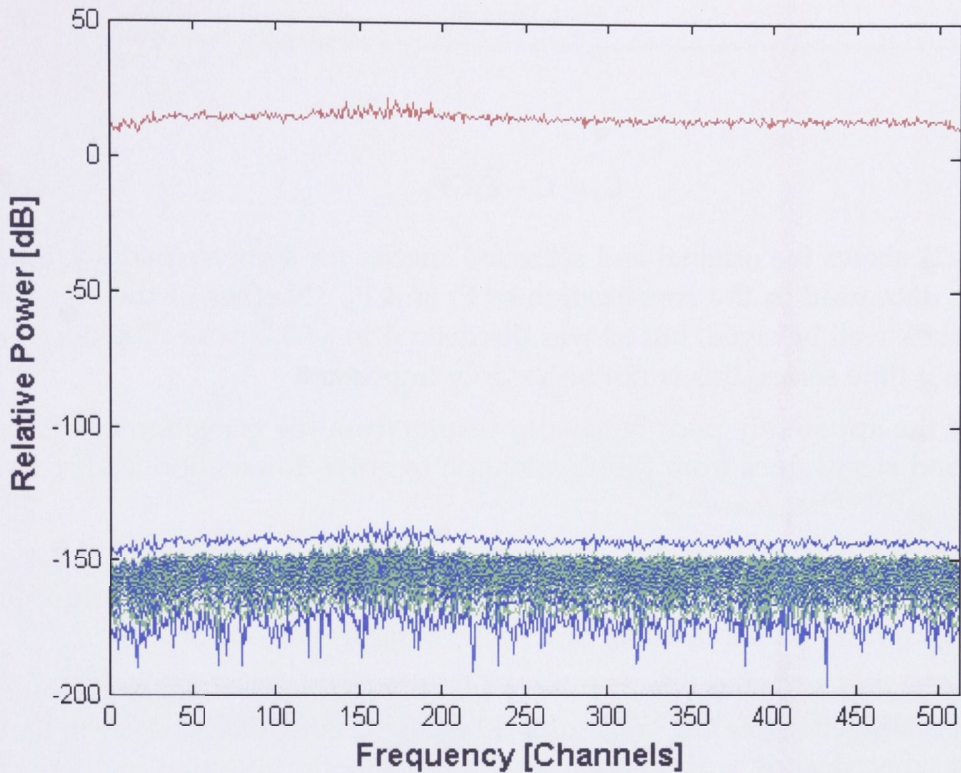
It is possible, however, to create a high SNR projection filter that is valid over a longer time and which can be applied to the data stream on short timescales. This approach relies on the stability of gain factors in equation 3.2 that apply to the interferers and that are in effect contained in the eigenvectors that describe the propagation of interference signal along the  $N$  propagation paths. Experiments in varying the integration time imply that these gains are adequately stable for time spans approaching one second. In a system with higher dynamic range, the cancellation would be more susceptible to changing gain with time.

This method takes the SVD and filter construction out of the high speed processing loop, but requires that either there be a large buffer for storage of the fine time samples so that they can be circulated through the filter after it has been constructed or that the filter be applied to future data as a prediction.

There is an additional mathematical consideration when computing the SVD of the covariance matrix. It is necessary to have at least as many independent time steps contributing to the covariance matrix as there are input signals if the matrix is to be non-singular. For example, if the entire covariance matrix,  $C(f)$ , were derived from spectra computed from a single integration time, the SVD solution at each frequency would find a covariance matrix of rank one, and only a single eigenvalue and eigenvector pair for each frequency channel would be required to contain all of the information. The remaining eigen components would be indeterminate. This is illustrated in figure 4.20. If there was any interference in the data, it would not be possible to create a projection that would eliminate it using the single value, as the interference and signal would not be separated. When looking for an event that occurs on a short timescale this could mean it would not be possible to correct the spectra at the desired resolution.

A filter is now introduced that is stable on a longer term average, but which can still be applied on a shorter timescale, allowing for the data to be processed at high time resolution.

The filter for the Vela pulsar observation is constructed by integrating the covariance matrix for one second of data to obtain  $\bar{C}$ , then constructing the projection matrices, which are applied to correct the data stream on 1 ms



**Figure 4.20** Example of all information being contained in a single eigenvalue for each frequency channel. All eigenvalues after the first are effectively zero at the precision of the digital algorithm. This occurs because there were not enough independent spectra in the covariance matrix to constrain a solution for a greater number of eigenvalues.

integrations. It makes no discernible difference to the outcome if the time periods where a pulse is present in one of the inputs are excluded from this long term average. This is because the pulse is present for such a small fraction of the total spectra that it is not strong enough to create an influence on the overall average. The average projections created will be denoted  $\bar{\mathbf{P}}_N$  for the noise space projection and  $\bar{\mathbf{P}}_I$  for the interference space projection.

The covariance matrix is made up of the noise and interference subspace

$$\mathbf{C} = \mathbf{P}_N \mathbf{C} \mathbf{P}_N + \mathbf{P}_I \mathbf{C} \mathbf{P}_I. \quad (4.1)$$

In a similar manner

$$\bar{\mathbf{C}} = \bar{\mathbf{P}}_N \bar{\mathbf{C}} \bar{\mathbf{P}}_N + \bar{\mathbf{P}}_I \bar{\mathbf{C}} \bar{\mathbf{P}}_I. \quad (4.2)$$

The filter implementations that apply  $\bar{\mathbf{P}}_N$  and  $\bar{\mathbf{P}}_I$  to  $\mathbf{C}$  on short time scales result in two approaches to deriving corrected, short-integration covariance matrices:

$$\tilde{\mathbf{C}}_N = \bar{\mathbf{P}}_N \mathbf{C} \bar{\mathbf{P}}_N \quad (4.3)$$

and

$$\tilde{\mathbf{C}}_I = \mathbf{C} - \bar{\mathbf{P}}_I \mathbf{C} \bar{\mathbf{P}}_I. \quad (4.4)$$

Figure 4.21 shows the original and corrected spectra for a one second average of the Vela data used in the construction of  $\bar{\mathbf{P}}_I$  and  $\bar{\mathbf{P}}_N$ . Neither of the corrected spectra look well behaved, but as was discovered in §4.3.2, when looking for a pulsar in a time series, this is not necessarily important.

Much of the apparently poor behaviour results from the transition of the first and second eigenvalues from RFI domination to noise domination as discussed in §3.2.

The results from the filter implementations given in equations 4.3 and 4.4 are given in figures 4.22 and 4.23, and are compared with the results from performing the full correction on short, 1 ms, timescales in figures 4.24 and 4.25.

Figures 4.22 and 4.23 show the results of plotting the time series for the pulsar signal in interference-free and contaminated regions respectively. Notice in figure 4.22 that when dealing with interference-free data both projection methods are successful, in that the data is left unharmed for pulsar detection. There is a reduction in the general noise level, but the signal to noise ratio of the pulsar detection remains the same.

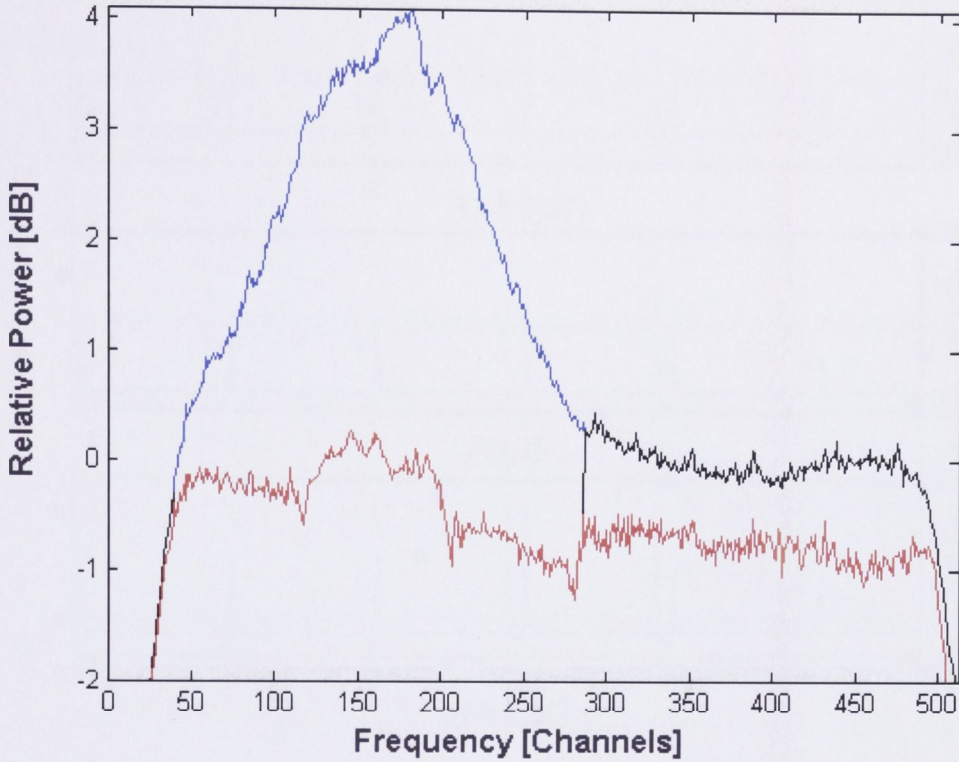
In figure 4.23 it is readily apparent that in the presence of strong RFI the projection method of equation 4.3 is much more successful than that of equation 4.4 as a filter implementation. The noise projection filter is highly successful at eliminating the RFI power and substantially enhancing the SNR of the pulses (to be on par with the SNR for the RFI free channels of figure 4.22). The interference projection, however, provides no net improvement in pulse detectability.

This is a non-intuitive result, especially given equations 4.3 and 4.4. The discrepancy is due to the way in which the pulsar power is represented during a pulse in both the eigenvectors and eigenvalues, and the way in which the eigenvectors, which create the projection, are a fixed average in the filter implementation.

The cases of the interference and noise projection in the interference-free region and contaminated region will be considered separately in the following sections. A comparison of the SNR for the different filters is given in table 4.3.

#### 4.4.1 Interference-free region

In principle, as there is no interference in the channel range (300-450), it would be possible to simply not apply the projection to this portion of the spectrum.

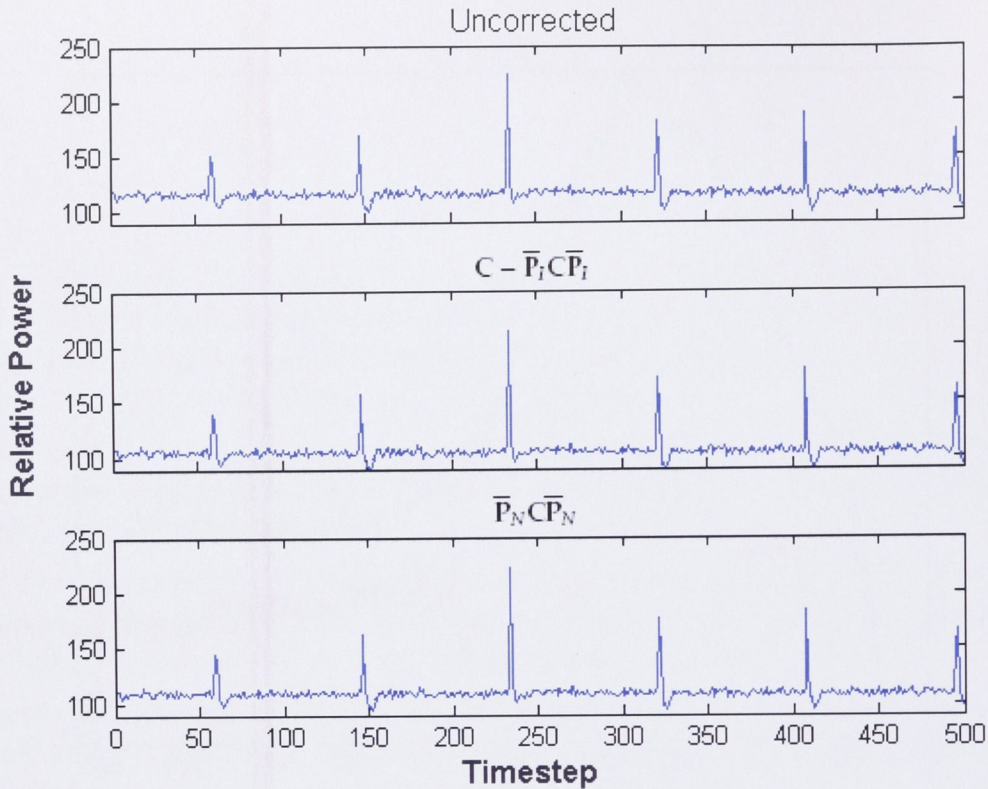


**Figure 4.21** Filter results from the application of the noise space projection to Vela pulsar data with an integration time of one second. As expected from earlier discussion (see §4.3.2), if the projection is applied across the whole band (red), there is a significant difference in power in both the interference and interference-free regions. If applied to only the interference contaminated region (black), there is still an overall reduction in the noise power, but only in the interference contaminated region. The reduction in noise power over what was desired occurs because too much power is been taken away in the projection that is not replaced or accounted for. However, this does not affect the pulsar detection, as it is focused on the change in the spectrum over time.

Method	Channel range	SNR pre-correction	SNR post-correction
1	300-450	30.77	-
2	300-450	30.77	26.44
3	300-450	30.77	30.36
1	100-250	5.2	-
2	100-250	5.2	6.2
3	100-250	5.2	23.23

**Table 4.3** Vela single pulse SNR for different filtering methods. The SNR values are an average, taken over five pulses. Method (1): No correction. Method (2): Filter based on Interference space projection. Method (3): Filter based on Noise space projection.

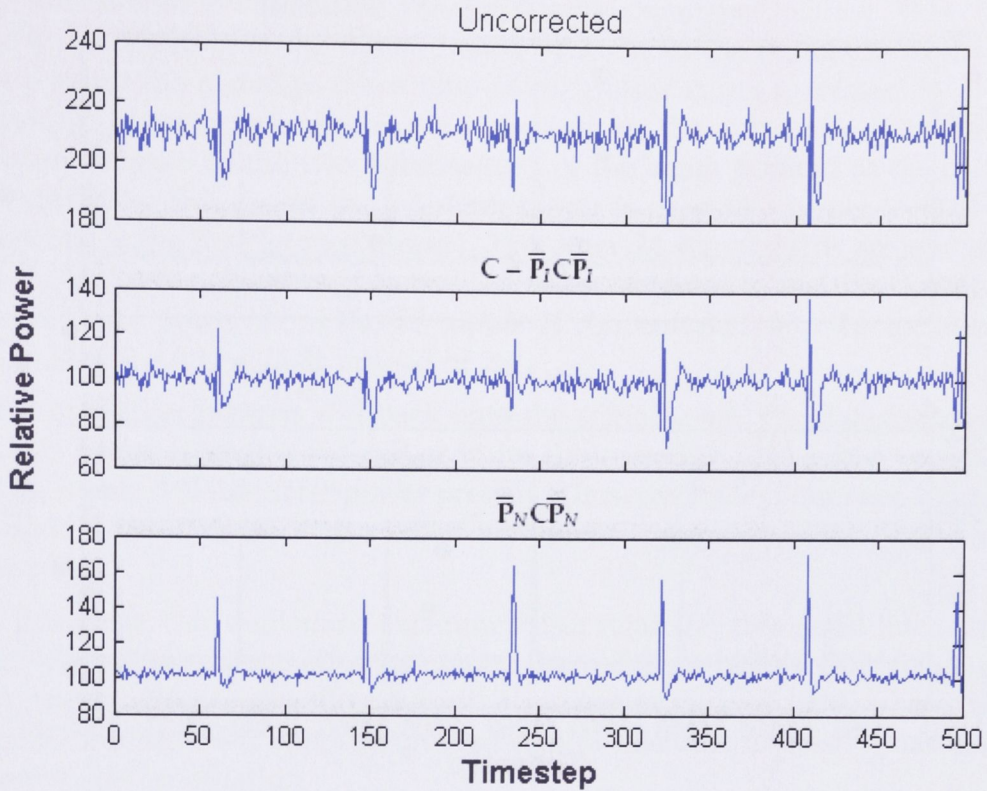




**Figure 4.22** Interference-free region: Detected pulsar power as function of time for an average of spectral channels 300-450. In contrast to the case where the projection was recomputed for each timestep (figure 4.13), the pulsar is left untouched. The spectra were integrated for 1 ms, and the filter calculated over one second. Top: The pulsar was detected from a series of spectra with no correction applied. Middle: Detection after calculating the interference by projecting onto the interference subspace then subtracting away the estimate (equation 4.4). Bottom: Detection after projecting directly onto the noise space (equation 4.3).

However, this requires a detection method for determining the areas of the spectrum that are contaminated. While a simple version of such a method is easy to envision (for example comparing the power in the spectrum to a pre-determined threshold), this proves unnecessary, as demonstrated in figure 4.22, which shows that neither the  $\bar{P}_N$  or  $\bar{P}_I$  filters harm the interference-free portion of the data.

In this section of the spectrum, a projection  $\bar{P}_I C \bar{P}_I$  can be created, which consists of the interference subspace. As there is no actual interference in this band, the first two noise eigenvalues and eigenvectors are effectively taken as the interference. This projection projects out the remaining noise space. When the pulsar pulse arrives, there is now additional information in what was previously dubbed the interference subspace. However, the projection is not updated, as it is a “filter” created from a previous time average. This means that the projection

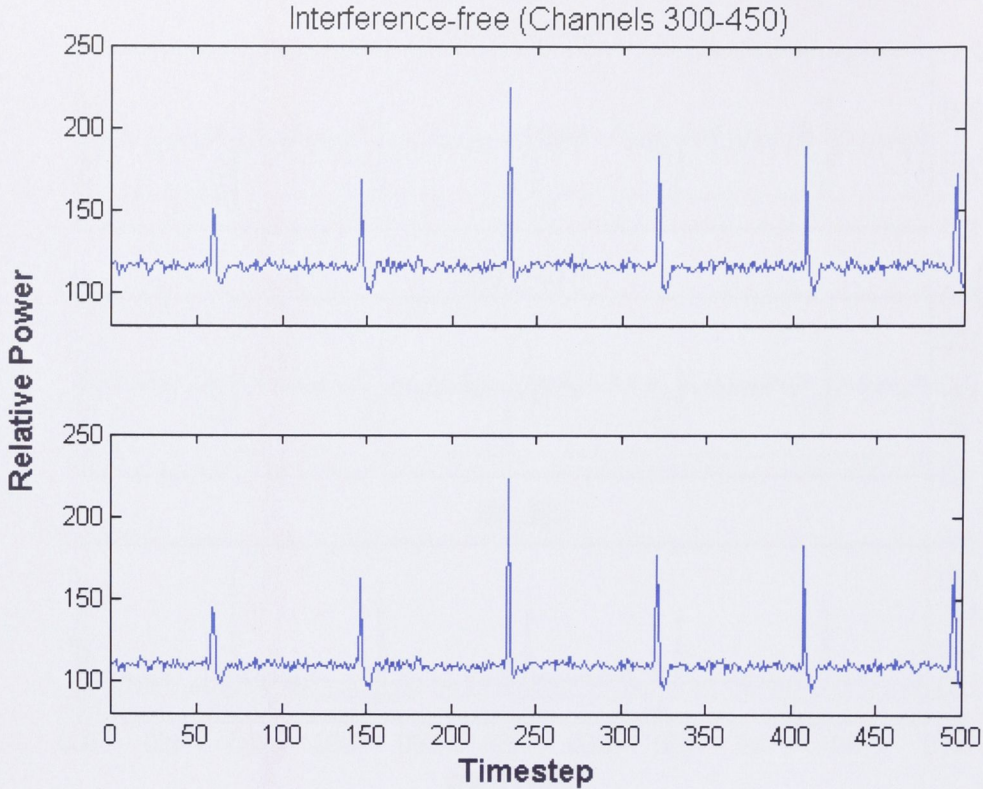


**Figure 4.23** Interference contaminated region: Detected pulsar power as function of time for an average of spectral channels 100-250. The spectra were integrated for 1 ms, and the filter calculated over one second. Top: The pulsar was detected from a series of spectra with no correction applied. Middle: Detection after calculating the interference by projecting onto the interference subspace then subtracting away the estimate (equation 4.4). Bottom: Detection after projecting directly onto the noise space (equation 4.3).

still only isolates the “interference”. This means that only a small amount of noise power with a spatial signature matching that of the RFI carries through the filter to be subtracted from  $\mathbf{C}$  ( $\mathbf{C}_{corrected} = \mathbf{C}_{original} - \bar{\mathbf{P}}_I \mathbf{C}_{original} \bar{\mathbf{P}}_I$ ), leaving the pulsar largely untouched.

In the case of the noise projection,  $\bar{\mathbf{P}}_N \mathbf{C} \bar{\mathbf{P}}_N$ , the projection is created that removes the interference by virtue of projection onto the noise subspace. This results in only the noise remaining. When the pulsar arrives, the filter is not updated, so the pulsar power remains dispersed through the eigenvalues, most of which are passed through the filter and then used to reconstruct the time series. The pulsar signal is not significantly affected.

Figure 4.24 gives a direct comparison between computing and applying the correction on short (1 ms) timescales and computing a long term (1 s) filter and applying it on short (1 ms) timescales for data in the interference-free region. There is no discernible difference, indicating that the astronomical signal has



**Figure 4.24** Interference-free region: Detected pulsar power as function of time for an average of spectral channels 300-450. Top: Detection after the first eigenvalue was replaced with a noise estimate for each frequency channel in channels 46-276, the second eigenvalue replaced with a noise estimate for each frequency channel in channels 114-205. The correction was recomputed and applied on 1 ms timescales. Bottom: Detection after projecting directly onto the noise space (equation 4.3). The spectra were integrated for 1 ms, and the filter calculated over one second.

passed through both filters equally well.

#### 4.4.2 Interference contaminated region

In the interference contaminated region, it is not possible to use the filters in the same way due to the effect of the automatic gain correction. Consider first the interference subspace projection  $\bar{\mathbf{P}}_I \mathbf{C} \bar{\mathbf{P}}_I$ . As before, the projection filter is computed to pass the interference subspace to be subtracted. When the pulse from the pulsar is present, the filter does not update and in theory only the interference should be removed. In the interference-free region, discussed in the previous section, this was not a problem, as only small components of the noise and pulsar signal that passed the RFI filter were subtracted. In the interference contaminated region, however, the automatic gain correction causes

the interference to be reduced, but only in this channel of interest. This confuses the filter, resulting in the signal passed by the filter including too much power. As a result the signal to noise ratio of the pulsar is not increased by a useful amount. More specifically, as the pulse arrives, the automatic gain correction lowers the gain in the two polarisations of the beam pointed at the pulsar by around 10%. This causes the interferer power to drop to a power level of  $0.9 \times I$  (where  $I$  is the interference power). The other 24 signal paths are not affected, as the pulse is present only in the one beam. Internal to the filter process, the interference power from the 26 inputs is concentrated into  $\lambda_1$ , meaning that  $\lambda_1 = 26I - 2 \times 0.1I = 25.8I$  instead of  $26I$ .

When the filter projects this back onto the signal space, the datastream for the pulsar beam estimates the power to be subtracted is  $(25.8 \div 26)I = 0.992I$ , but there is only  $0.9I$  interferer power present. Thus, the  $\overline{\mathbf{P}}_I \overline{\mathbf{C}} \overline{\mathbf{P}}_I$  filter over-subtracts by around 9.8%. This leads to a broad temporal suppression of the region around the pulse.

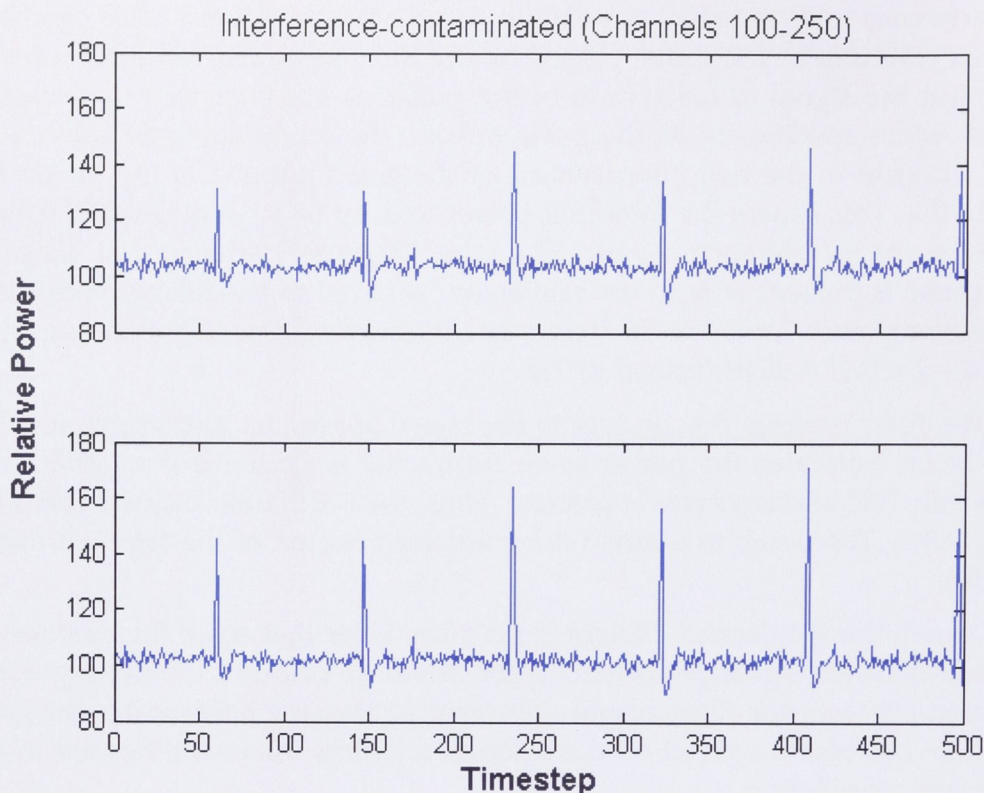
In this sense, this unplanned experiment is a reminder that small fluctuations in gain on individual signal paths among the reference beams will have a greatly attenuated effect on the filter output. However, in the case here, where the gain fluctuation has been imposed on the beam of scientific interest, it has led to an inadequate cancellation.

In the  $\overline{\mathbf{P}}_N \overline{\mathbf{C}} \overline{\mathbf{P}}_N$  case, when the pulsar signal is not present, the interference is projected out. When the pulse arrives, the interference is still removed, but in contrast to the  $\overline{\mathbf{P}}_I \overline{\mathbf{C}} \overline{\mathbf{P}}_I$  filter, the noise projection is robust to the effects of the automatic gain correction. This is because the projection onto the noise space gives the desired output immediately, rather than subtracting the estimate as in the  $\overline{\mathbf{P}}_I \overline{\mathbf{C}} \overline{\mathbf{P}}_I$  case.

Figure 4.25 gives a best case comparison between computing and applying the correction on short (1 ms) timescales and computing a long term (1 s) filter and applying it on short (1 ms) timescales in the interference contaminated region. The long term filter has slightly better performance, as there was no need to make a noise estimate to replace the eigenvalues as there was in the short term corrections. The longer integration also allows for a more accurate estimate of the interference, giving a theoretical suppression of more than 40 dB.

### 4.4.3 Summary

In summary, the  $\overline{\mathbf{P}}_I \overline{\mathbf{C}} \overline{\mathbf{P}}_I$  projection only ever makes a calculation of interference based on the original interference spatial signature and thus only ever takes the interference away. The  $\overline{\mathbf{P}}_N \overline{\mathbf{C}} \overline{\mathbf{P}}_N$  projection takes a calculation of everything except the original interference and therefore always keeps everything except for the interference.



**Figure 4.25** Interference contaminated region: Detected pulsar power as function of time for an average of spectral channels 100-250. Top: Detection after the first eigenvalue was replaced with a noise estimate for each frequency channel in channels 46-276, the second eigenvalue replaced with a noise estimate for each frequency channel in channels 114-205. The correction was recomputed and applied on 1 ms timescales. Bottom: Detection after projecting directly onto the noise space (equation 4.3). The spectra were integrated for 1 ms, and the filter calculated over one second.

Because the projections are based on the spatial signature of the signals rather than a fixed profile, the  $\bar{\mathbf{P}}_N \mathbf{C} \bar{\mathbf{P}}_N$  and  $\bar{\mathbf{P}}_I \mathbf{C} \bar{\mathbf{P}}_I$  methods of filtering should also be robust to a variable power interference source.

## 4.5 Future work

New instruments coming online at the Parkes telescope (for example the APSR and BPSR systems developed by Swinburne University of Technology) will provide the ability for greater experimentation and in ascertaining the limitations of the filtering.

Future experiments might include investigation into ways in which the filtering could be improved by taking advantage of this specific system. It would also be

interesting to look at the ways in which small or weak signal detection could be enhanced using eigen decomposition. Finally, development of a filter that would have applications other than in pulsar detection would be useful from a computation standpoint.

## 4.6 Conclusions

The application of spatial filtering to the Parkes Multibeam Receiver was found to be highly successful in the removal of RFI. This chapter discussed the hardware development required to record raw voltages from the telescope, and then several different variations of spatial filtering were tested. These included the ways in which noise contributions to the interference subspace eigenvalues were accounted for, and the effects of applying the filter across an entire bandwidth, as opposed to just the portions of the bandwidth containing interference. Care must be taken when calculating and applying the filter on short timescales. If eigenvalues in interference-free portions of the spectrum are removed, there is a high risk of removing the astronomical signal.

A long term filter implementation was also tested, where the spatial filter was computed on longer (hundreds of milliseconds) timescales and applied to data on millisecond timescales. This was also found to be successful in removing the interference and enhancing the detection of the Vela pulsar. However, it was found that the influence of the automatic gain correction caused the interference projection filter to malfunction. It is expected that this filter would work satisfactorily were the automatic gain correction not present.

Finally, it should be noted that this method did not allow for a correction factor to be applied for any noise removed along with interference during the filtering process, and this may limit the application of the filter to studies such as pulsar observations.



# CHAPTER 5

---

## MWA 32 T DATA

The MWA is a telescope currently under construction in remote Western Australia. It is a low frequency telescope designed to operate in the 80 – 300 MHz range. Lonsdale et al. (2009) and Bowman et al. (2007) provide an overview of the system. <sup>1</sup>

At the time of writing, the MWA has a 32 antenna prototype array in operation. These dual polarisation antennas (each consisting of 16 dipoles) provide 64 signal streams from the output of the receiving system.

There are a number of spectrum users in the MWA band, providing an opportunity for RFI cancellation experiments. This chapter demonstrates subspace projection techniques with a dataset in a 1.28 MHz band centred on 261.12 MHz which finds heavy use by satellite transmissions.

This analysis is included here as significant work was put into the development of MWA hardware as part of this thesis research. This included an evaluation of, and gaining familiarity with, FPGA hardware (see FPGA spectrometer development in Appendix A), PFB evaluation work, early deployment assistance, and code development for the Complex Programmable Logic Device (CPLD) in one of the receiver subsystems (see Appendix C for schematic). This subsystem is primarily responsible for configuring the beamformers (see figure 5.1).

The goals of this chapter are to: 1) illustrate the use of SVD in characterising a complex RFI environment, and 2) explore some of the difficulties that can be encountered in the use of sub-space projection when the source strength is comparable to the RFI, such as in the case of Solar observations.

---

<sup>1</sup>More information can be found at [www.mwatelescope.org](http://www.mwatelescope.org)



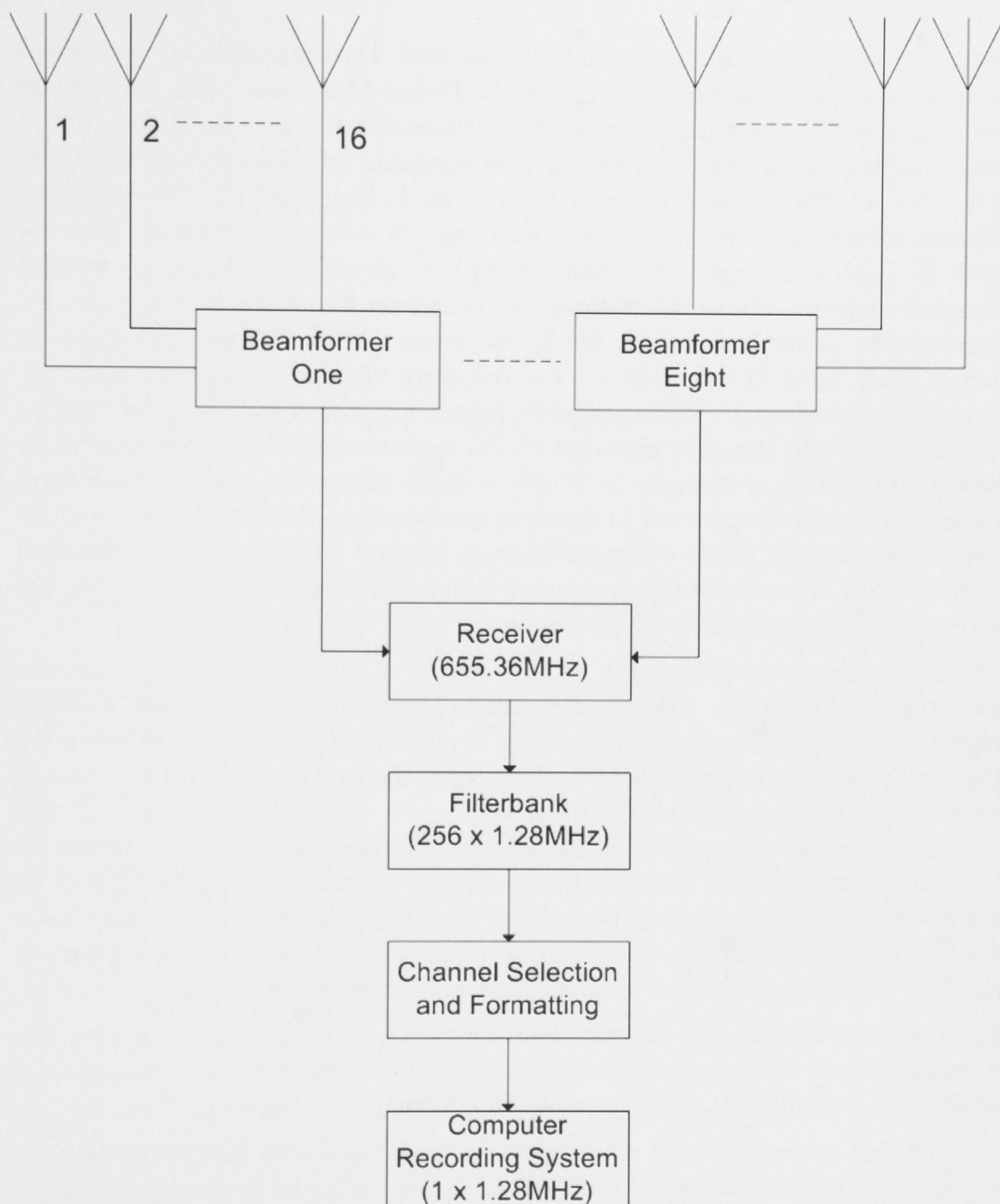
## 5.1 Data acquisition and hardware

On 19 November 2008 some sample data of the Sun (00:19:24 UTC) and Hydra A (03:48:47 UTC) were taken with the MWA in the 261.12 MHz band, which is saturated in satellite communication signals (see tables D.1 and D.2 for further observation details).

The datasets were taken by members of the MWA consortium, using a temporary receiver chain to enable the recording of data in the “32 T” prototype array. The acquisition instrumentation used in the experiment is shown in figure 5.1.

Each beamformer takes the output from 16 dipole antennas which occupy a  $4 \times 4$  critically sampled grid on the ground plane. Each receiver unit takes output from eight beamformers. Four of the acquisition systems shown in figure 5.1 are required to serve the 32 antenna “tiles” in the prototype array. The receiver subsystem was developed at The Australian National University (ANU) in collaboration with the Raman Research Institute (RRI). The Very High Speed Integrated Circuits Hardware Description Language (VHDL) for one component of this subsystem, the ATIM, was developed as part of the graduate research program described in this thesis. The MWA receiving chain relies heavily on FPGA hardware. An introduction to FPGAs and a description of the ATIM are given in Appendices A and C.

The 261.12 MHz data taken was limited in bandwidth to 1.28 MHz with 128 spectral channels, and was riddled with interference. Solar observations and observations of Hydra A were taken. The data has only been subjected to preliminary analysis, but illustrates both some of the potential, and limitations, of the spatial filtering method. This data is slightly different to that taken with the Parkes Multibeam, in that each of the beams of the Multibeam observes a different direction in the sky. In the case of the MWA, a synthesis array, all of the beams are pointed to the same direction.



**Figure 5.1** MWA 32 T prototype receiving system interior. Each beamformer takes the output from 16 dipole antennas, and each receiver takes output from eight beamformers. The data is then processed using a PFB, with 256 channels of 1.28 MHz resolution. This digital output for all 8 dual-polarised antenna tiles for a single 1.28 MHz channel is then recorded through a VSIB (Ritakari & Mujunen, 2002) interface to a large disk array attached to a PC computer system.

## 5.2 Hydra A observations

Hydra A observations are analysed in this section. The input data was processed in much the same manner as that from the Parkes Multibeam. The system was being commissioned during the period of these observations, so there were 51 data channels (of 64 total) performing with adequate reliability to be included in this analysis. The voltages from the various feeds were read into the computer, Fourier transformed and correlated, forming 128 covariance matrices (one for each frequency channel) of dimension  $51 \times 51$  averaged to have one second integration time. These dimensions are based on 51 usable inputs (out of a theoretically possible 64) from the 32 tile system, with 128 spectral channels across the 1.28 MHz bandwidth. The resulting total power spectra from the covariance matrices are represented in figure 5.2. Apart from a short portion of the bandwidth (roughly channels 65-75) the entire band is contaminated by heavy RFI having a strength of 10 dB or more above the system noise level. Given the system temperature at this frequency (around 300 K), the approximate aperture efficiency of the entire 32T system (around 0.1 K/Jy) and the strength of the Hydra A source at this frequency (120 Jy), the Hydra A detection would be expected to stand about 10 dB above the noise floor as well.

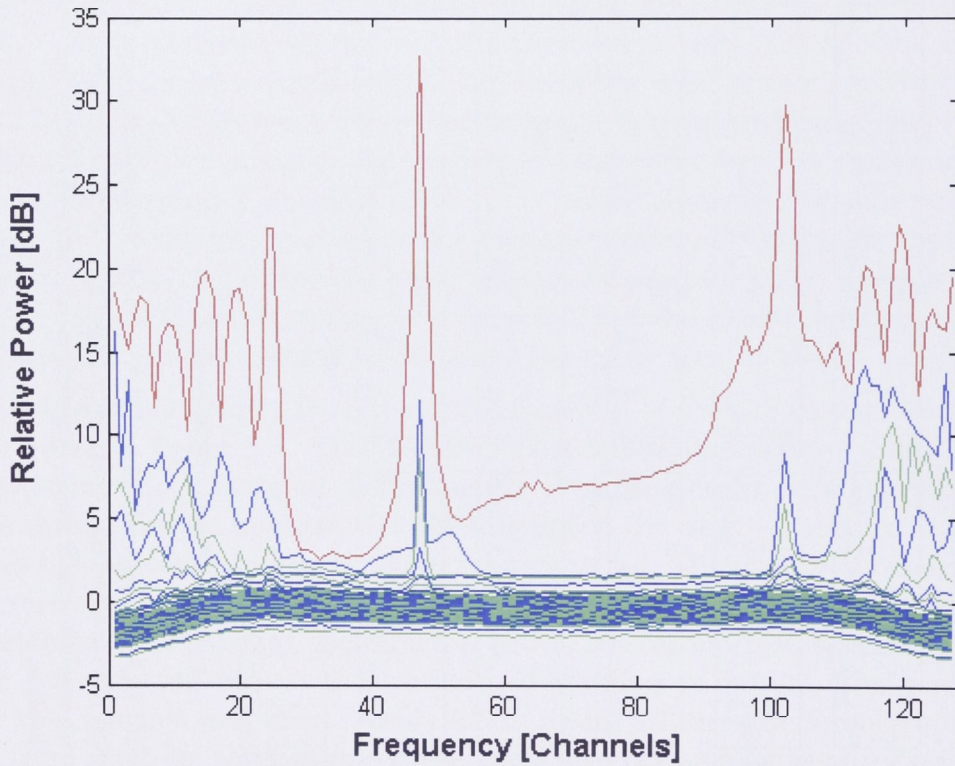
The eigenvalue spectra from the decomposition of the 128 covariance matrices are given in figure 5.3. The interference is spread through as many as eight eigenvalues depending on the frequency. Projecting out these eigenvalues gives the results shown in figure 5.4. In this figure the original spectra are shown, with the corrected spectra on the same axis in red. The majority of the channels correct very well. A re-scaled version of corrected spectra is shown in figure 5.5. Certain spectra (such as those on the end of the second row, spectra 15 and 16) are noisy and still appear to have residual interference present. The exact reason for this remains unknown. As shown in figure 5.3, the interference appears to be from multiple sources, and has a great deal of complex frequency structure.

In an effort to unravel this complexity, future work could be done analysing this data to identify the specific eigenvalues with directions in the sky through use of the eigen coefficients and the known locations of the telescope tiles.

The strength of the Hydra A source also makes it highly likely that the primary eigenvalues contain some information about the celestial source, particularly in the nominally interference-free 65 – 75 channel region, adding additional complications to the correction. The extent to which the Hydra A signal is reduced or otherwise damaged will need to be determined in future work. Subsequent to the acquisition of this data it has also been recognised that the 256 channel filterbank that isolates the 1.28 MHz bandwidth that was processed here was not delivering the theoretical level of channel-to-channel rejection. Thus some of the “interference” in the eigenvalue spectra may be due to aliased power and leakage from other nearby channels.



**Figure 5.2** MWA Hydra A total power spectra at 261.12 MHz, pre-correction. Relative power [dB] is plotted as a function of frequency channel 1-128. The channel spacing is 10 kHz and the integration time one second. Apart from a slight lull in the interference in the 65-75 spectral channel range, the entire spectrum is heavily contaminated with satellite interference. For comparison, the red plot in the bottom right shows a representative passband for a 1.28 MHz band centred at 231.68 MHz.



**Figure 5.3** MWA Hydra A observations, eigenvalue matrix. Interference is clearly seen in at least the first eight eigenvalues, and seems to be present across the whole band. Even the “lull” in channels 65-75 shows the presence of strong power in eigenvalue one that stands out by more than 5 dB.



**Figure 5.4** MWA Hydra A total power spectra at 261.12 MHz, correction comparison after replacing the first 8 eigenvalues with a noise estimate. Channel spacing is 10 kHz and the integration time one second. Blue: The original spectra. Red: The corrected spectra.



**Figure 5.5** MWA Hydra A total power spectra at 261.12 MHz, post-correction (a closer look at figure 5.4). Channel spacing is 10 kHz and the integration time one second. While appearing to remove most of the interference, several channels (most notably 15 and 16), appear to still contain some residuals, although these might be related to poor performance by both polarisations of a this specific antenna tile.

## 5.3 Solar observations

Despite being at the same frequency (261.12 MHz) as the Hydra A observations of §5.2, the cancellation of the interference in the solar data was much less successful. The total power spectra for these observations, computed in an identical manner to those of the Hydra observations, with a one second integration time are shown in figure 5.6.

Once again the covariance matrices were decomposed into eigen components. The eigenvalue spectra are shown in figure 5.7. The first five eigenvalues were replaced with a noise estimate and the covariance matrices reconstructed. The results of this are shown in comparison to the original spectra in figure 5.8 and in more detail in figure 5.9.

The high source strength and size of the Sun created complications in the cancellation process. At this frequency, the 32 T array resolves the Sun into approximately 10 independent regions of emission. This spreads the power though 10 eigenvalues. The high source strength means that these eigenvalues also have comparable power to the RFI, making it difficult to separate the RFI and source power in the multiple eigenvalues, despite attempts to pre-whiten the noise.

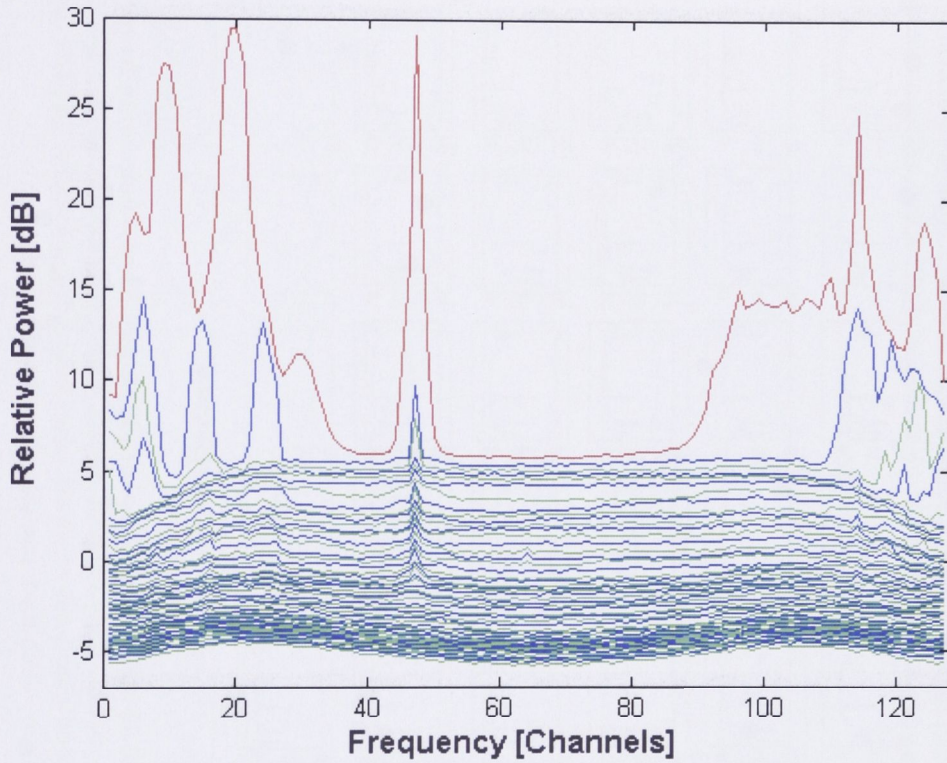
One way in which this separation may be accomplished is to use the eigenvectors (and a traditional aperture synthesis approach) to map the sky. It is expected that the  $\lambda_i(f)$  frequency ranges that correspond to the Sun could then be identified and kept, removing the  $\lambda_i(f)$  frequency ranges that map to other coordinates in the sky, where the satellite transmitter is located.

Apart from any damage caused to the solar signal, which will need to be tested, these results raise other interesting questions. Particularly, why some input signals can be corrected better than others. For example, input signals one and nine appear to have similar interference, however signal nine is far more successfully corrected than signal one (see figure 5.8). As with the Hydra A data, channels 15 and 16 are also left particularly contaminated. This raises the question as to whether the position of the tiles in the array, and hence the baseline length, is having an effect on the cancellation success. Other parameters, such as integration time, need to be explored further, but initial experiments suggest this is not the dominant factor. An investigation into the link between the location of the antenna tiles and residual interference would also provide useful results.





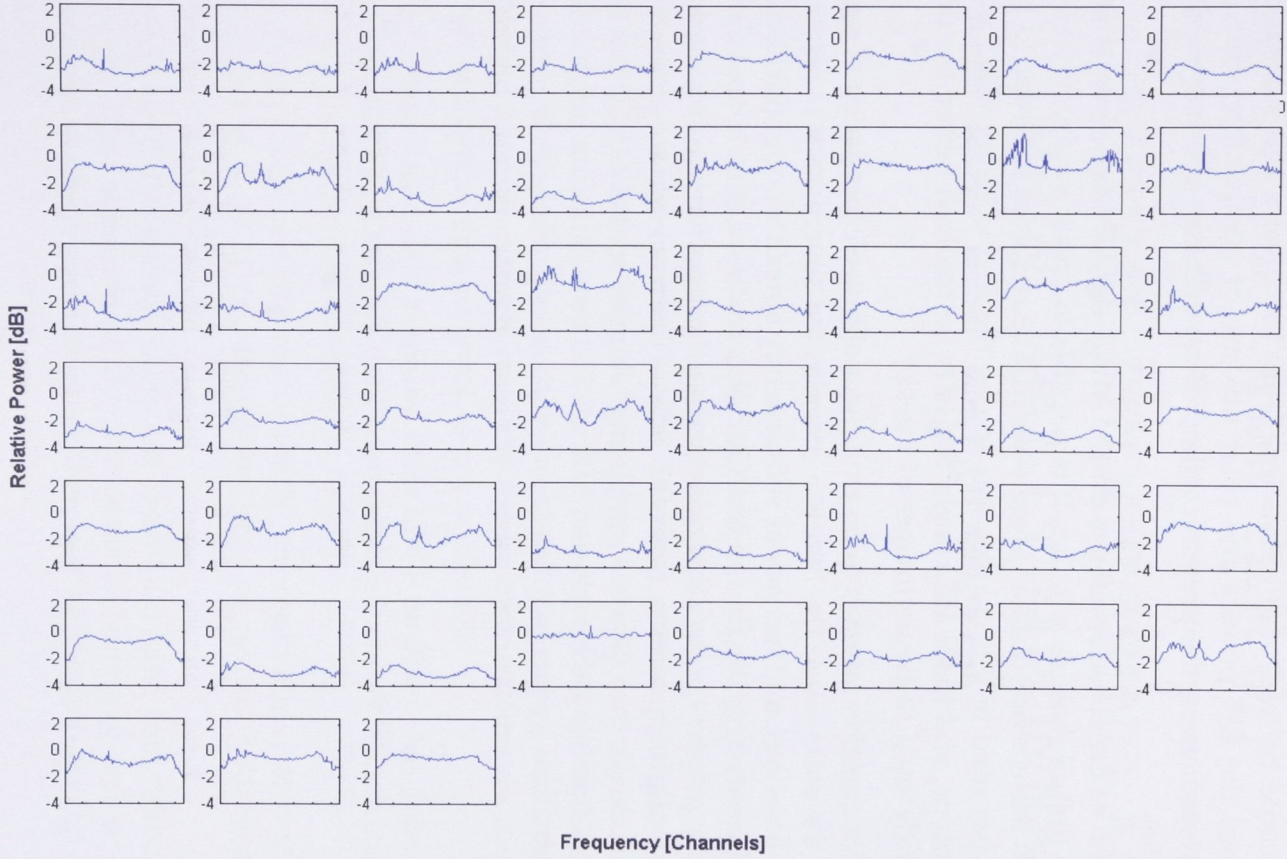
**Figure 5.6** MWA Solar total power spectra at 261.12 MHz, pre-correction. Relative power [dB] is plotted as a function of frequency channel 1-128. The channel spacing is 10 kHz and the integration time one second. As with the Hydra observations in the same band (§5.2), the only portion of the spectrum that appears interference free is the narrow band in the 65-75 spectral channel region.



**Figure 5.7** MWA Solar observations, eigenvalue matrix. Channel spacing is 10 kHz and the integration time one second. Interference is present in the top five eigenvalues. The spacing of the eigenvalues also clearly indicates that there is correlated noise power, presumably solar signal, in first few eigenvalues. This makes it problematic to remove the interference without damaging the solar signal or using additional constraints (such as those that would be obtained from combining the decomposition with interferer matrix synthesis imaging).



**Figure 5.8** MWA Solar total power spectra at 261.12 MHz. Relative power [dB] is plotted as a function of frequency channel 1-128. The channel spacing is 10 kHz and the integration time one second. Blue: Original spectra. Red: Corrected spectra (first five eigenvalues replaced with a noise estimate). The effectiveness of the cancellation on only some channels raises interesting questions. In particular note the similar nature of the interference in feeds one and nine, and the way in which feed nine cancels more effectively.



**Figure 5.9** MWA Solar total power spectra at 261.12 MHz post-correction (a closer look at figure 5.8). Relative power [dB] is plotted as a function of frequency channel 1-128. The channel spacing is 10 kHz and the integration time one second. The first five eigenvalues have been replaced with a noise estimate.

## 5.4 Conclusions

In this chapter data from the MWA, specifically observations of Hydra A and the Sun, were analysed using the spatial filtering methods outlined in Chapter 3. The Hydra A data appeared to have the interference completely removed, and the solar data could be improved cosmetically, but there were still substantial ambiguities since the RFI powers and solar flux are of comparable strength. Under these circumstances, the present method of cancellation does not resolve these ambiguities.

The examination of cancellation effects on the MWA data created several new questions. The first of these is whether it is possible to create a proper map of the source with the corrected data. More data will be required in order to test this, as the dataset used in this chapter was among the first to be recorded with the MWA prototype, and there are known issues to be resolved before the data can be confidently used to create images of the sky.

The second major question raised is to do with the ineffective cancellation of some of the feeds in the solar data. The Sun is a complicated source, as it is resolved by some of the baselines, and has power subsequently spread throughout several of the eigenvalues during the decomposition. This makes a simple correction by removing the primary eigenvalues containing the interference impractical, as they also contain information from the Sun. Further work is required to determine a successful, then practical, implementation of spatial filtering in this regime, probably involving the inclusion of the additional direction information that would be obtained by merging the decomposition with methods in synthesis imaging.

## CHAPTER 6

---

# SIGNAL PROCESSING AND PRINCIPAL COMPONENT ANALYSIS

This chapter outlines the use of SVD and subsequent eigenvector manipulation of spectra in order to perform passband calibration, and gives a specific example of how this might be used for the detection of absorption lines. It was found that using a subset of the principal components to reconstruct the passband of the spectra could theoretically be used as a method for highlighting any fine structure. It was also found that the presence of strong and prolific RFI in the band had a detrimental effect on the significance of detection, and needed to be removed in some manner prior to any analysis.

This chapter used data taken by Brett Little, Stephen Ord, John Reynolds and Jonathon Kocz in 2004 with the Parkes radio telescope's 50 cm receiver (see tables D.1 and D.2 for further observation details).

The project explored the use of the 50 cm receiving system in searching for redshifted ( $z \approx 1$ ) 21 cm absorption lines of neutral hydrogen against high redshift radio sources.

The recorded dataset was subjected to an automatic gain calibration as it was received. This calibration set the gain at the beginning of each scan by making a measurement of the noise statistics then locking the gain for the duration of the scan.

This dataset was used to illustrate the potential for using SVD to aid in feature detection. The spectra used have a 64 MHz bandwidth, with a centre frequency of 685 MHz and a spectral channel resolution of 7.8125 kHz with two bit resolution. Most of the spectra used have an integration time of 1200 seconds.

For computational reasons, only portions of the 64 MHz bandwidth, either 4 MHz or 6.25 MHz, were used when creating simulations and determining the principal components.

## 6.1 Passband calibration techniques

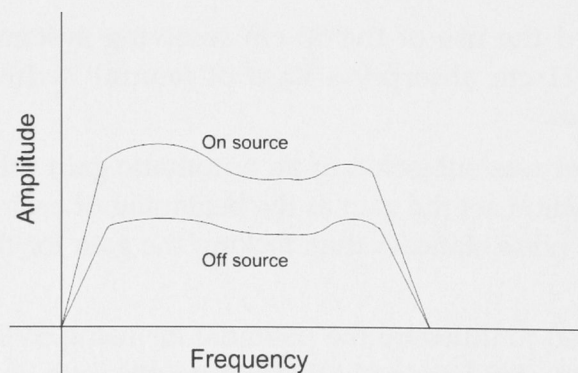
Passband calibration is the process of measuring the spectral response (gain as a function of frequency) of the radio telescope and receiving system. The spectral response is then corrected for the frequency dependent gain so that spectral features in the celestial radio sources are clearly separated from artifacts in the receiving system.

A traditional and well tested method of passband calibration involves observing portions of blank sky close to the targeted coordinates of astronomical interest, in order to obtain gain information and calibrate the spectral passband shape.

The gain calibration is applied by the simple equation  $[T_{on}(f) - T_{off}(f)]/T_{off}(f)$ , where  $T_{on}(f)$  and  $T_{off}(f)$  are power spectra measurements when the telescope is pointed at the source and the sky respectively. Figure 6.1 is a sketch of the spectral passbands for on-source and off-source scans.

The “ON-OFF” spectral calibration has served especially well in situations where the receiver noise is the dominant contribution to total power and where there is little difference in power level between the on- and off-source pointings. With modern low-noise receivers, these constraints are not always satisfied, necessitating a closer look at the gain functions that are applied to the different sources of power that reach the digital spectrometer.

To illustrate these issues, the power contributions are decomposed into two principal parts: the receiver noise power  $T_R(f)$  and the sky power  $T_a(f)$ . The sky power is expected to vary as different amounts of sky brightness and discrete



**Figure 6.1** Simple illustration for the purposes of explaining gain calibration, showing an exaggerated affect of having a continuum source in the telescope beam.

radio sources enter the telescope beam. These power levels, following the radio astronomical convention, are referenced to the point in the receiving system chain between the output of the antenna and the input to the low noise amplifiers to the receiver. The  $T_R$  and  $T_a$  are commonly called receiver temperature and antenna temperature and specified in degrees Kelvin.

The output from the spectrometer can be idealised as

$$T_{out}(f) = g_B(f) [T_R(f) + g_A(f)T_A(f)]. \quad (6.1)$$

Equation 6.1 includes frequency dependent gains  $g_B$  and  $g_A$ , where  $g_B$  is the frequency dependent portion of the receiving system that follows the input to the low noise amplifier, and  $g_A$  includes the portion ahead of the amplifiers that influences the sky signal. The  $g_B(f)$  factor is dominated by the band-limiting filter at the input to the analog-to-digital converters, but it will include any frequency dependencies of the intermediate frequency (IF) conversion system and cabling. The  $g_A(f)$  factor includes frequency structure in the passband of the feed, as well as gain modulation caused by reflections and multipaths in the antenna structure that give rise to frequency dependent ripples and the “standing wave” phenomenon.

It can be easily seen that when  $T_R$  is the dominant power, or when the  $T_A$  power level is the same in the ON and OFF scan, that the conventional  $[T_{on}(f) - T_{off}(f)]/T_{off}(f)$  approach will divide out the gain dependence appropriately.

For the purpose of the calibration process, the expression in equation 6.1 can be rewritten as

$$T_{out}(f) = T'_R(f) + g'_A(f)T_A(f) \quad (6.2)$$

where the  $g_B(f)$  factor has been absorbed into the other parameters. Passband calibration is performed by viewing either blank fields or continuum sources, which provide uniform spectral power into the antenna, so that  $T_A(f) = T_C$ , which is a constant as a function of frequency. In this circumstance

$$T_{out}(f) = T'_R(f) + g'_A(f)T_C. \quad (6.3)$$

This expression provides the motivation to search for methods that will extract the two passband shapes that appear in the data as the continuum power level  $T_C$  is varied.

Principal Component Analysis (PCA) is an efficient method for doing this.



## 6.2 Introduction to Principal Component Analysis

PCA is a common technique for finding patterns that can be used to simplify the presentation of multidimensional data sets. The general process for PCA involves calculating the appropriate covariance matrix, then solving for the eigenvectors and eigenvalues of this matrix. The eigenvectors with the highest eigenvalues are the most significant “principal components” of the data set.

PCA has previously been applied in an astronomical context, notably in the analysis of QSO spectra (see for example Francis et al. (1992)). Mittaz et al. (1990) describes a method similar to that applied in §6.2.1. A matrix  $\mathbf{D}$ , which consists of a series of QSO spectra, which are binned by wavelength and formed into a matrix (the wavelength makes the columns, and each spectrum is a row) is decomposed using SVD into its eigenvalue and eigenvector components. That is

$$\mathbf{D} = \mathbf{U}\mathbf{\Lambda}\mathbf{V}^H \quad (6.4)$$

where  $\mathbf{U}$  and  $\mathbf{V}$  are the left and right eigenvector matrices, and  $\mathbf{\Lambda}$  is the eigenvalue matrix. These eigenvectors can be multiplied by some weighting and added in a linear combination to reconstruct each individual spectrum. In the case of QSOs, this provided a simple way to determine common spectral characteristics among QSOs (Francis et al., 1992).

In principle, the 50 cm observations are suited to this method, since the database contains the recording of many spectra, of which the vast majority should have no astronomical spectral features, as redshifted 21 cm lines are very rare. Only one or two detections are expected, but the observations included more than 150 radio continuum sources with a range of  $T_C$ .

### 6.2.1 Method applied to the Parkes 50 cm dataset

For the 50 cm observations, many total power spectra were taken over time.

The total power spectra were smoothed with a median filter wide enough to remove any fine structure (including RFI) from the passband. In this application, the decomposition is implemented in an alternate, but equivalent, fashion to Mittaz et al. (1990) by constructing the covariance matrix,  $\mathbf{C}$ , of dimension  $N_{chan} \times N_{chan}$  such that

$$\mathbf{C}_{ij} = \sum_{k=1,n} C_{ij}(k) \quad (6.5)$$

where  $C_{ij}(k) = s_i(k)s_j(k)$ ,  $s_i$  and  $s_j$  are discrete spectra points, with  $i, j = 1, N_{chan}$  and  $k$  is the index of the current spectra, summing to  $n = 386$ , the number of spectra that are included in the PCA database.

Once this matrix is constructed, it is decomposed into eigen components via SVD. It is important to note that this covariance matrix is different from that computed in previous chapters. Previously the covariance matrix has been built one spectral channel at a time to examine the covariance between the different input feeds. Here it is based on the covariance between the different total power spectral passband shapes.

The process of SVD allows the matrix  $\mathbf{C}$  to be represented in the form

$$\mathbf{C} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^H \quad (6.6)$$

where  $\mathbf{C}$  is the correlation matrix,  $\mathbf{U}$  eigenvector matrix and  $\mathbf{\Lambda}$  is the eigenvalue matrix.

Each column of  $\mathbf{U}$  represents an eigenvector. The first column is the most significant, and has the most impact on the structure of the spectra. The second has less impact, and so on.

Each spectrum  $S$  is made up of a linear combination of these eigenvectors, or principal components, such that

$$S(\vec{k}) = s_1\vec{U}_1 + s_2\vec{U}_2 + \dots + s_i\vec{U}_i \quad (6.7)$$

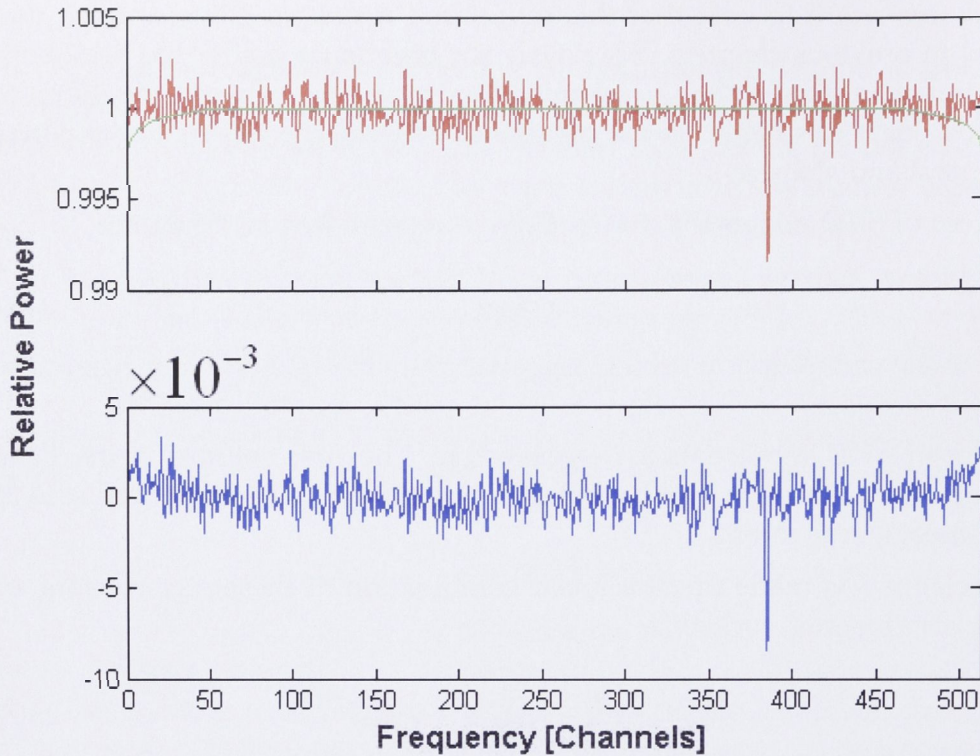
where  $s_i$  are the coefficients that can be obtained from the dot product of the spectrum with the each of the eigenvectors, for example

$$s_1 = S(\vec{k}) \bullet \vec{U}_1 \quad (6.8)$$

Once the principal components have been obtained, the passband can be reconstructed using a subset of these coefficients, for example, one through three. In this way the general passband shape is reproduced without introducing the fine structure, such as absorption lines, into the spectrum. The differences between the original spectra and that which has been reconstructed from smooth eigenvectors can then be taken. The spectral features specific to each observation should remain after the passband shapes common to the database have been removed.

## 6.3 Applications to feature detection demonstrated with the Parkes 50 cm data

In this section results of full simulations, “hybridised” simulations, and results for the Parkes 50 cm dataset are presented in the context of feature detection. The results suggest that this method could theoretically be used to implement feature detection, but the high RFI content in the test data made this difficult to illustrate.

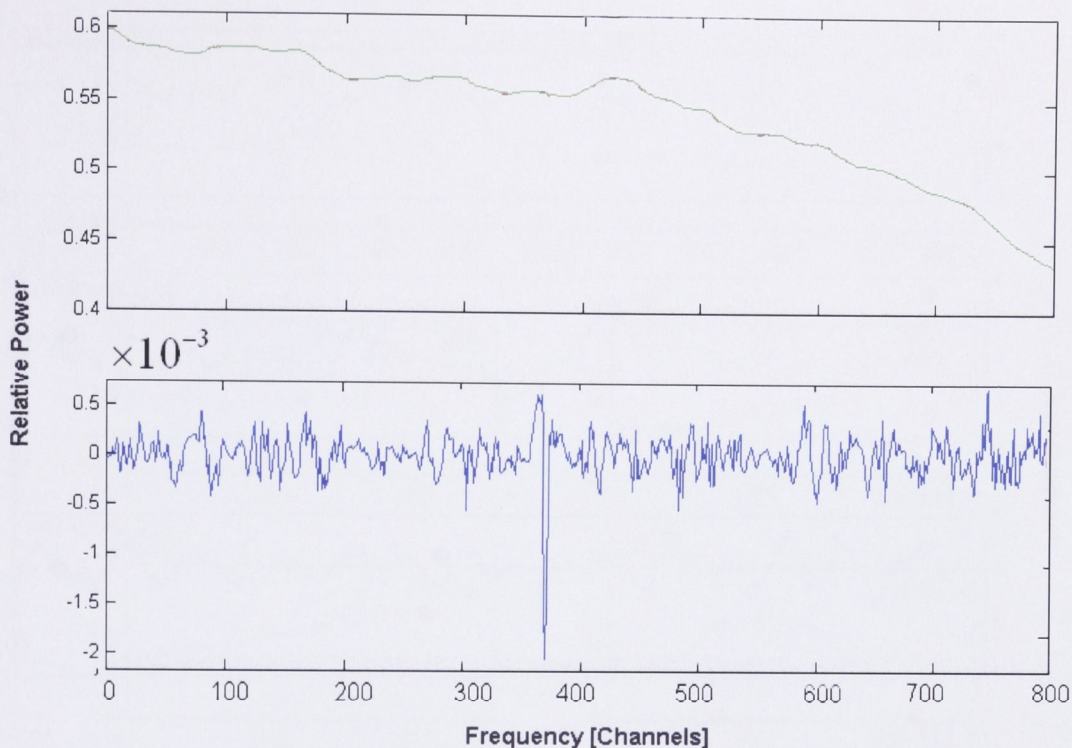


**Figure 6.2** Top: Simulated flat bandpass with Gaussian noise with 1% absorption line about 25 channels wide (red). Bandpass reconstructed using the only the first principal component (green). Bottom: Difference between the two spectra. The absorption line is easily detected.

### 6.3.1 Simulation Results

In order to test the potential of the principal component analysis in the detection of spectral features, several different simulations were run. These simulations started simply, by generating 386 spectra (the same number as available in the actual data) with a flat bandpass with 0.2% RMS Gaussian noise. An artificial absorption line 1% of the bandpass power and about  $90 \text{ km s}^{-1}$  in width was applied to one of the spectra. The value of 1% was selected in order to give a strong detection. The Parkes 50 cm observations aspired to detect absorption lines with depths of about 0.2% of the continuum.

This simulation showed for an RFI-free set of spectra, with the only deviation from the flat bandpass being the absorption line in two of the 386 spectra, using the eigen coefficients to reconstruct the bandpass was a successful method for detecting absorption lines (see figure 6.2).



**Figure 6.3** Top: Smoothed spectra with simulated 0.3% absorption line about 25 channels wide (red). Bandpass reconstructed using the first 50 principal components (green). Bottom: Difference between the two spectra. The absorption line is clearly visible.

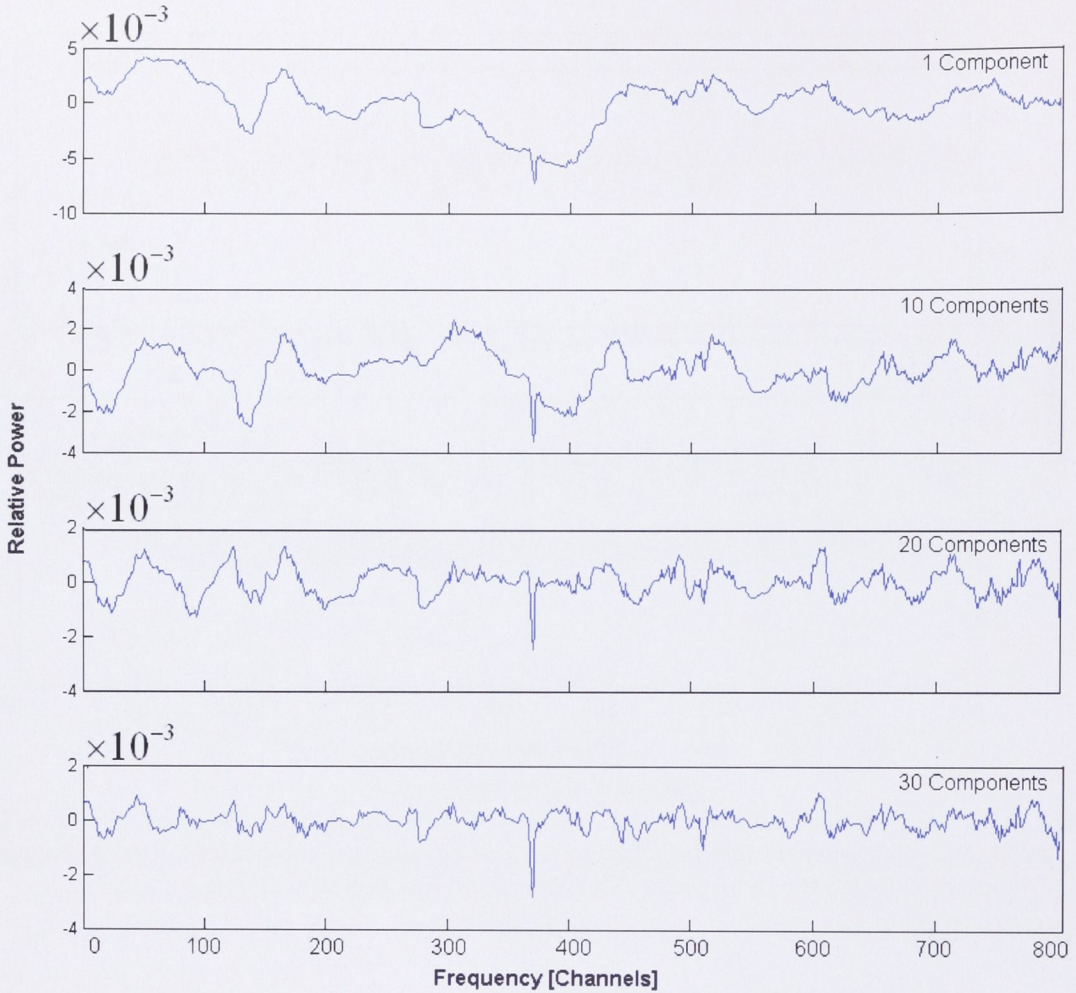
### 6.3.2 Semi-Simulated Results

Once the fully simulated results were successfully completed, the bandpass was changed to be made up of real data. Figure 6.3 is made using the actual bandpass of the data taken at Parkes. However, due to the severe nature of RFI in the band, the spectra were median smoothed so that there was no fine structure.

The purpose of this simulation was to show that if an absorption line was to be present in an otherwise uncorrupted bandpass, it would be detectable.

The main difference between the detection here, and that of the fully simulated signal in §6.3.1, was in the number of eigen coefficients needed to generate a substantial detection.

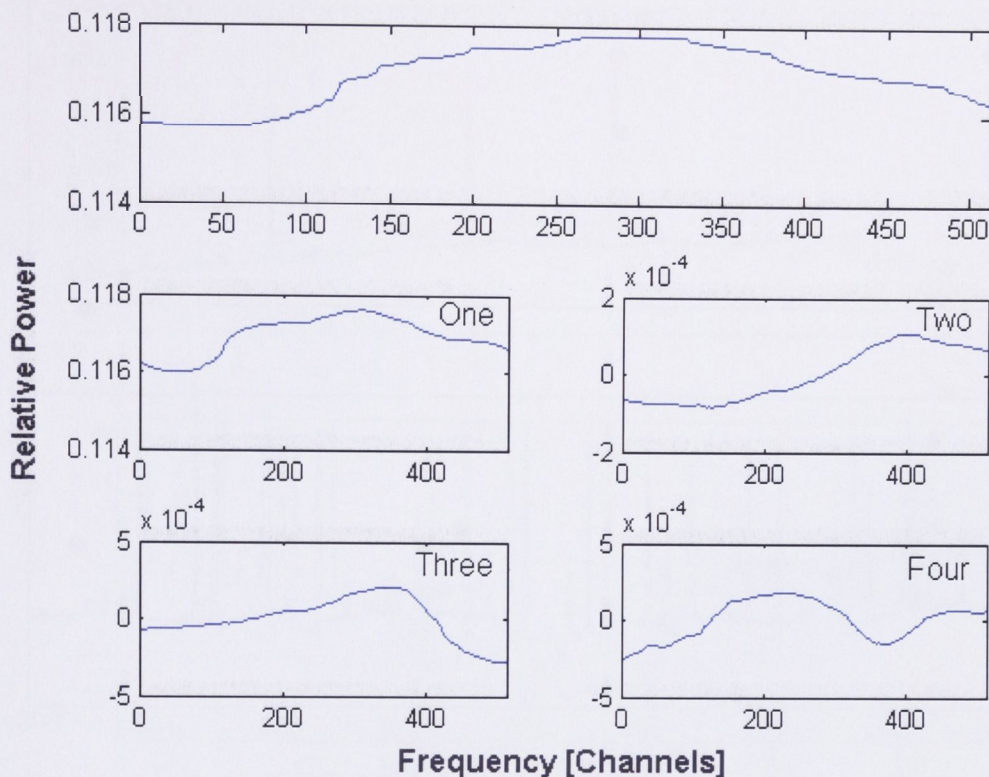
In the pure simulation, only the first principal component was required to show the absorption at a significant level relative to the bandpass. In this case, while subtracting only the first principal component showed the absorption line, considerable residuals in the bandpass were still present, which would prevent statistically significant detections. In order to reconstruct this bandpass well enough that only the absorption line was different between the original and reconstructed spectra, it was necessary to include many more of the principal



**Figure 6.4** Top to bottom: Difference between original spectrum and the passband reconstructed using 1, 10, 20 and 30 principal components. It can be seen that there is a balance to be struck in the number of principal components used. If too few are used, (as in the top 3 panels) the detection of the absorption line is suspect. If too many are used, then there is a risk of any fine structure being included in the bandpass reconstruction and thus removed with the general passband shape.

components. Figure 6.4 shows the differences between the two spectra for cases where the first one, ten, twenty and thirty components were used. If too few are used, (as in the top 3 panels) the detection of the absorption line is suspect. If too many are used, then there is a risk of any fine structure being included in the bandpass reconstruction and thus removed with the general passband shape.

An example of the way the principal components reconstruct the bandpass is given in figure 6.5, which shows a portion of a spectrum and its corresponding principal components. This figure demonstrates how the first component is typically much larger than the others.



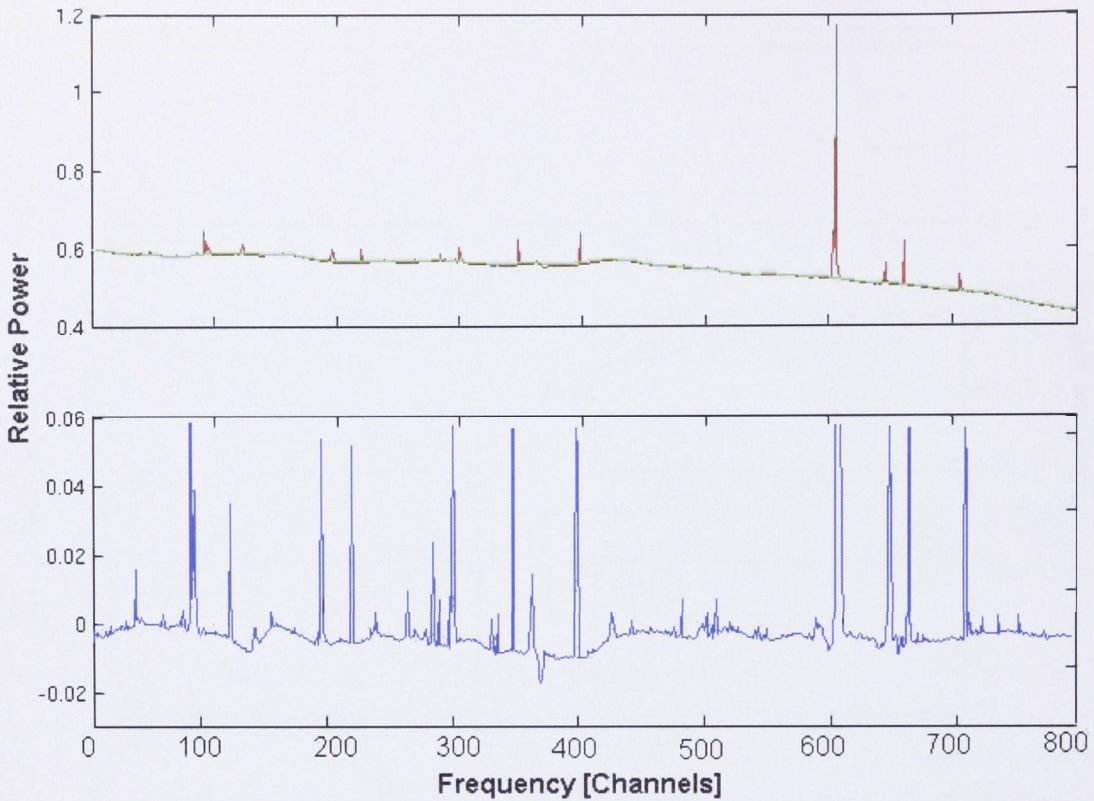
**Figure 6.5** Example of principal components for a spectrum. The top panel shows the complete spectrum, the four lower panels reading left to right show contributions made by the first four principal components to the spectrum in the top panel. The first component is typically much larger than the others.

### 6.3.3 Real Data

In this section the difficulties that were encountered when applying the eigen coefficient process to data without removing the RFI in the band are addressed. During this analysis the original data was left unchanged, except to add a 1% and then 0.3% absorption line to one of the spectra in order to test the likelihood of a correct and false detection with the presence of strong RFI. For these testing purposes, only a portion of the entire 64 MHz bandwidth was used.

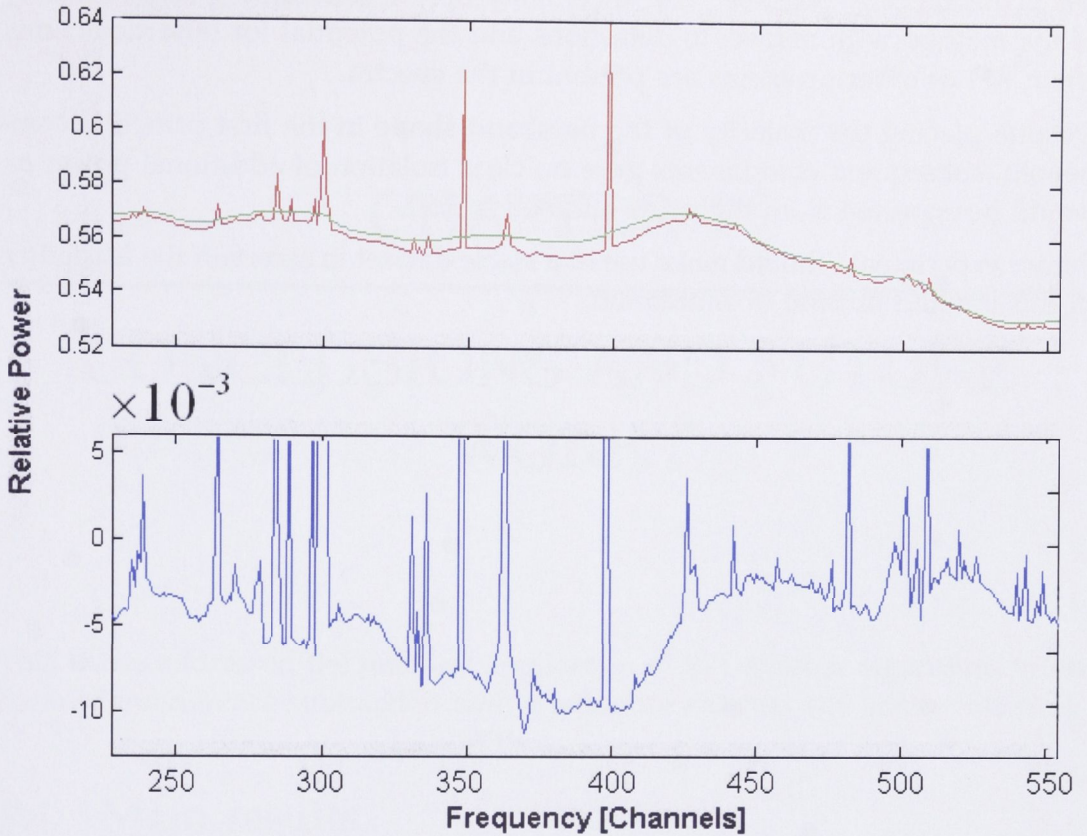
In figure 6.6 it is shown that a simulated absorption line of 1% of the bandpass power is strong enough to give a definitive ( $> 5\sigma$ ) detection. What is also noticeable is that the RFI is starting to cause false absorption lines as well, as it is included in the principal components used in reconstructing the passband.

As depicted in figure 6.7, once the strength of the absorption line was reduced to that expected from the data, (around 0.3%) the false absorption lines(s) caused by the RFI made proving a real detection difficult.



**Figure 6.6** Top: Bandpass with simulated 1% absorption line about 25 channels wide (red) around channel 370. Bandpass reconstructed using one principal component (green). Bottom: Difference between the two spectra. False absorption lines start to appear, for example, around channel 600. The vertical scale on the bottom panel has been adjusted to show the fine structure.

Various experiments, in particular attempting to automatically flag the areas of the spectrum contaminated by RFI, were also undertaken. The result of flagging was that whilst removing false detections in the RFI region, equally any chance of making a real detection in the RFI region was also removed.



**Figure 6.7** Top: Real bandpass with simulated 0.3% absorption line about 25 channels wide (red) around channel 370. Bandpass reconstructed using one principal component (green). Bottom: Difference between the two spectra. The absorption line detection is now weak enough that it could not be definitively separated from a false line cause by RFI. The vertical scale on the bottom panel has been adjusted to show the fine structure.

## 6.4 Conclusions

In this chapter an introduction to passband calibration and feature detection using SVD and PCA techniques was given.

Simulations to detect absorption lines using generated data were successful, with only a single principal component needed to distinguish the absorption line by eye. When a more realistic system was introduced involving data from observations, however, more principal components were required, due in part to the unstable nature of the bandpass. When there was a high RFI spectral occupancy, the absorption line was still found if it did not fall within an RFI feature, but was not distinguishable from false absorption lines caused by the RFI.

The initial simple simulations developed show that feature detection in this man-



ner is theoretically possible. The uncertainty lies in determining the sensitivity of the method with respect to detections and the potential for false detections when RFI or other problems are present in the spectra.

Despite placing the majority of the passband shape in the first principal component, subsequent components gave no clear isolation of additional power as would be expected from the gains outlined in §6.1.

Future experiments should make use of a stable dataset to ascertain the feasibility of this method outside of simulation.

# CHAPTER 7

---

## CONCLUSIONS AND FUTURE WORK

This thesis addressed the practical application of RFI excision algorithms in the form of interference subtraction with a reference antenna and spatial filtering.

### 7.1 Main results

Chapter 1 introduced the study and reviewed the related literature around the problem of RFI in radio astronomy, particularly focusing on the methods of interference subtraction with a reference antenna and spatial filtering which were used throughout the remainder of the thesis.

Chapter 2 assessed the effectiveness of using a reference antenna and closure relations to enhance the detection of pulsars J0437-4715 and J0630-2834 in data taken with the Parkes telescope 50 cm receiver.

Visually, the interference seemed to be completely subtracted in the case of J0437-4715, and the subtraction had a substantial positive effect on detection. The interference was also almost entirely removed in the J0630-2834 observations, with a suspected weak second source of interference remaining. This demonstrated a limitation in the method, namely that an interfering source outside the reception pattern of the reference antenna will not be subtracted. This limitation can also be an advantage: there is no danger of harming the astronomical signal of interest. This makes interference subtraction using a reference antenna a particularly useful method, as the subtraction process does not require close supervision.

In Chapter 3 the method of spatial filtering was introduced. SVD was used to extract and sort the eigenvalues and eigenvectors from the covariance matrix.

The theoretical and practical application of spatial filtering to astronomical data was discussed, and the importance of gain calibration shown. The methods of noise and interference subspace projection were shown to be equivalent to nulling the primary eigenvalue(s), and the dangers of removing noise along with the interference was demonstrated. A method of compensating for this by replacing the interference subspace eigenvalues with noise estimates was then reviewed.

Chapters 4 and 5 dealt with the application of spatial filtering to data from the Parkes radio telescope and the MWA.

The application of spatial filtering to the Parkes Multibeam shown in Chapter 4 was found to be highly successful. Two different filtering implementations were tested. These include recomputing the filter on short (millisecond) timescales, and creating a longer term (one second) filter that could be applied on short timescales. Within each of the short and long term filters there were two possible implementations: determining the interference subspace for subtraction, and projecting directly onto the noise subspace. On short timescales, both implementations were found to be successful, although care must be taken that, when applying the filter to interference-free portions of the spectrum, astronomical signal is not removed accidentally. On long timescales, the direct projection onto the noise space was successful, however it was found that when analysing pulsar data the interference projection did not properly allow for the presence of the pulsar in the interference contaminated regions. This discrepancy is believed to be due primarily to the presence of an automatic gain correction in the receiving chain. Future observations should be taken with this gain correction disabled in order to verify this.

The examination of cancellation effects on the MWA data created several new questions. The cancellation was very successful with the Hydra A data, but the first question raised that is still to be addressed is the ability to create a proper map of the source with the corrected data. A proper map would verify that the data was not harmed during the cancellation process. More data will be required in order to test this, as the dataset used in Chapter 5 was among the first to be recorded with the MWA prototype, and there are known issues to be resolved before the data can be confidently used to create images of the sky.

The second major question occurred when observing the Sun. The RFI cancellation was particularly ineffective in some of the signal feeds. The Sun is a very bright and complicated source, as it is resolved by some of the baselines, and has power subsequently spread throughout several of the eigenvalues during the decomposition. This makes a simple correction by removing the primary eigenvalues containing the interference impractical, as they also contain information from the Sun. The use of additional constraints (such as those obtained by combining the decomposition with interferer matrix synthesis imaging) to better isolate the interferer components should be investigated.

Chapter 6 introduced feature detection in the presence of RFI using SVD and PCA techniques. Simulations to detect absorption lines in generated data were successful, with only one principal component needed to detect the absorption line. When a more realistic system was used involving data from observations, however, more principal components were required. When there was a high RFI spectral occupancy, the absorption line was still present, but was not distinguishable from false absorption lines caused by the RFI.

The experiments performed using simulated spectra and absorption lines suggest that using the principal components to reconstruct the passband shape could theoretically be used as a method for feature detection. However, future experiments should make use of a stable dataset to ascertain the feasibility of this method outside of simulation.

## 7.2 Future work

Both specific near term future work, and long term goals exist for the work developed in this thesis.

With regard to interference subtraction using a reference antenna, a filtering implementation should be explored. The implementation used in this thesis has a lower limit on the integration time, meaning the correction is not reliable on short (less than a millisecond) timescales. A filter based on longer term averages that could be applied on these timescales would be useful. It would also reduce the computational costs of implementing this method of excision.

For spatial filtering, in the near term it would be desirable to obtain new data from the Parkes Multibeam Receiver with the automatic gain correction disabled. This would allow for easier analysis of the data, and a more robust determination of the improvement in SNR given by application of filtering.

Future experiments regarding spatial filtering should include investigations into ways in which the filtering could be improved by taking advantage of information known about specific receiving systems. It would also be interesting to look at the ways in which small or weak signal detection could be enhanced using eigen decomposition. The development of a filter that would have applications other than pulsar detection would be useful from a computational standpoint for other types of observations. New instruments coming online at the Parkes telescope (for example the APSR and BPSR systems developed by Swinburne University of Technology) will also provide the ability for greater experimentation and in ascertaining the limitations of spatial filtering.

It is also desirable to test the ability to create maps from RFI subtracted MWA datasets once the degree of confidence in the MWA prototype performance increases.

## 7.3 Finally

Both interference subtraction using a reference antenna and spatial filtering were applied to data from various telescopes, both proving to be effective methods of RFI excision. Interference subtraction with a reference antenna should be used when the location and number of interferes in the environment are known. Spatial filtering is more adaptable, but it is also more computationally intensive. With the continual increase in computing power, however, this will become less of a problem. The current state and development rate of FPGA chips also makes them a viable alternative to computer clusters to perform this sort of analysis in real time.

The importance of this work cannot be overstated. Wireless communication continues to expand, and with it the amount RFI in the spectrum. The proliferation of satellites now means that even the remotest locations will suffer from some interference. For radio astronomers, in the words of Jerzy Usowicz, "it is our nessun dorma"!

# Appendices



# APPENDIX A

---

## FPGAS IN RADIO ASTRONOMY

This appendix shows some of the potential uses for FPGAs in radio astronomy instrumentation. §A.1 deals with an introduction to development tools in the form of Xilinx<sup>®</sup> System Generator software and the creation of an FPGA spectrometer.

### A.1 An FPGA spectrometer

This project was initially undertaken with a two-fold purpose in mind. The first was to gain familiarity with the Xilinx<sup>®</sup> software available for programming FPGAs, and the second was to use the completed spectrometer on the MWA prototype to gain a wider band of signals than had previously been recorded. This second application was never realised, as it became unnecessary for the further development of the MWA.

#### A.1.1 Hardware selection

The first task in the development of the spectrometer was the selection of which FPGA to use. After an examination of two primary manufacturers, Xilinx<sup>®</sup> and Altera<sup>®</sup>, it was decided to use Xilinx<sup>®</sup> FPGA chips, as they could be obtained for no cost under the the Xilinx<sup>®</sup> University Donation program.<sup>1</sup>

Following this decision, the type of Xilinx<sup>®</sup> chip to use was then reviewed. At this stage the Virtex 4 chip range had just been released. This range consisted of three types, LX, SX and FX chips, which were optimised for different purposes. In summary, the LX chips provide general logic ability, the SX are geared toward signal processing and the FX chips contain both signal processing functionality and also a fast data transport ability.<sup>2</sup>

---

<sup>1</sup><http://www.xilinx.com/univ/>

<sup>2</sup><http://www.xilinx.com/products/virtex4/index.htm>



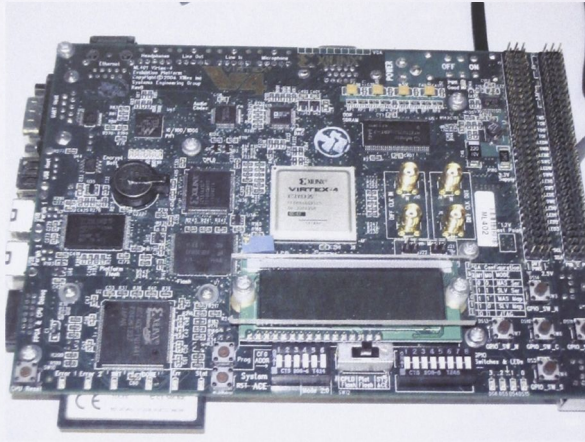


Figure A.1 Xilinx<sup>®</sup> ML402 development board.

The SX chips were chosen both due to their applicability to the task at hand, but also as they were the most likely to be used in future PFB applications and it was desired to gain a familiarity with the abilities of the chip.

The final product chosen was a Xilinx<sup>®</sup> ML402 development board.<sup>3</sup> This board was ideal for the task as it contained both a version of the SX chip to be evaluated but also various pre-routed inputs and outputs for the chip. A picture of the board is shown in figure A.1

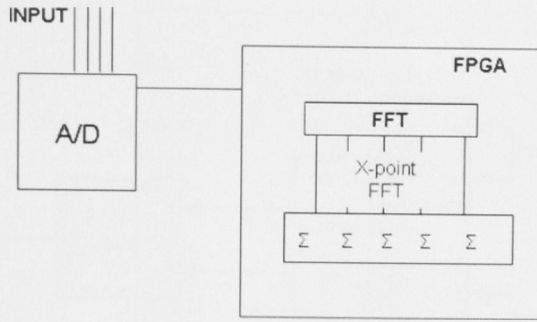
### A.1.2 Development tool selection

Once the hardware had been chosen, it was then necessary to select the software to be used in the programming of the chips. There were several options available, however the final selection involved a combination of the default Xilinx<sup>®</sup> software (ISE Foundation) along with MATLAB<sup>®</sup> Simulink combined with a Xilinx<sup>®</sup> library "System Generator" which allowed for synthesisable designs from the MATLAB<sup>®</sup> Simulink schematics.

### A.1.3 Design

The design was primarily completed using the System Generator libraries within Simulink. However there were some aspects that required the use of the ISE software. These were mainly concerned with blocks of VHDL code that needed to be inserted to facilitate various operations in the spectrometer, particularly with regard to data transmission.

<sup>3</sup><http://www.xilinx.com/products/devkits/HW-V4-ML402-UNI-G.htm>



**Figure A.2** The design of the FFT spectrometer implemented on the FPGA. A single data stream arrives from the ADC via the point to point IO pins on the ML402 board. This data stream is 10 bits wide.

The design of the spectrometer was kept as simple as possible (see figure A.2). Data is received from an analog to digital converter via the single ended pins on the development board and then passed through an FFT programmed into the FPGA. The magnitude of the coefficients from the output of the FFT are then found, squared, and the power spectrum accumulated. The data is then transmitted back to the computer via a serial link.

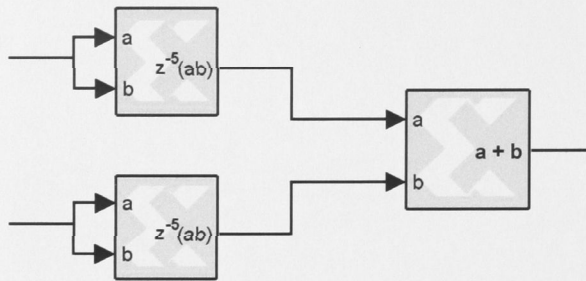
#### A.1.4 Implementation

All of the advantages offered by working with the System Generator library in the Simulink environment were exploited during the implementation of the spectrometer.

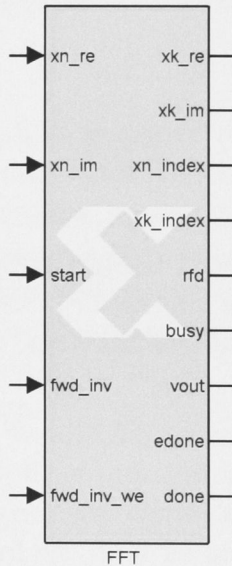
Peripheral logic was all completed using the System Generator blockset, for example, taking the magnitude of the FFT, setting the number of times to accumulate spectra or setting completion flags. An example of this is given in figure A.3

The FFT itself was implemented using the Xilinx<sup>®</sup> FFT core v3.2. Once generated, this core was imported into Simulink and is shown in figure A.4. One thing to note is that the user does not need to know anything about the internal processing of the blocks, and only needs to deal with the input and output signals.

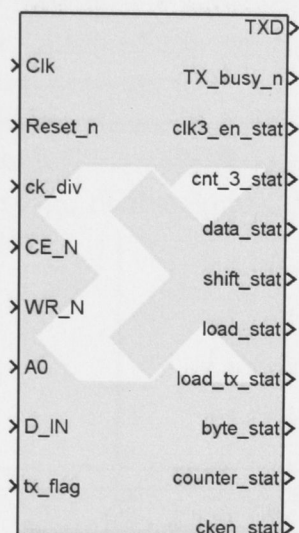
The VHDL for the Universal Asynchronous Receiver/Transmitter (UART) transmitter was implemented in a similar manner to the FFT. However, in place of a Xilinx<sup>®</sup> core, an open source UART transmitter was rewritten and simplified to suit the application. This VHDL code was then directly imported into the Simulink design (see figure A.5). The implementation of the complete design consumed approximate 25% of the FPGA resources.



**Figure A.3** This figure shows some System Generator blocks in action. The "real" and "imaginary" lines connect to the output of the FFT. The multipliers effectively square the signals, which are then combined to give the power. The "z" in the multiplier block indicated the number of clock cycles the process will take (in this case 5).



**Figure A.4** The FFT core as seen after importation into Simulink. The data input and output ports take the real and imaginary parts of the signal as required. The remaining input ports allow for the user to determine the time when FFT is performed ("start") and the type of transform (FFT or IFFT). The output ports give indications as to the status of the transform, for example, when the FFT is complete, or how far through the transform it has progressed.



**Figure A.5** The serial link VHDL black box. In much the same way as the FFT core, the user simply provides required input signals. The rest of the processing is abstracted. For this design, the data is transmitted out of the block as 8N1 UART, and was set to transmit at 115200 bps.

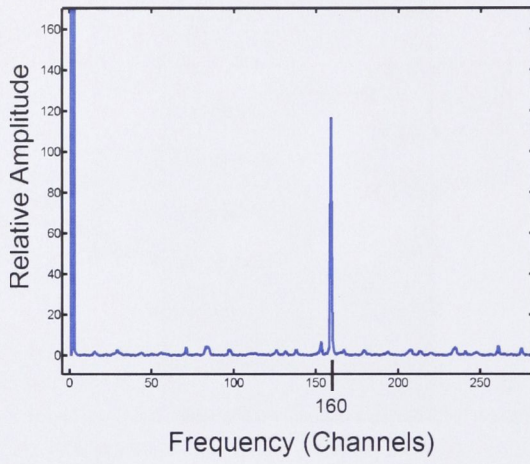
### A.1.5 Results & Conclusion

The design was tested both with input generated on the chip itself, and input from a function generator passed through an ADC clocked at 64 MHz. In both cases the data was accumulated for approximately 0.8 seconds. This accumulation rate was chosen in order to accommodate the slow transmission speed of the serial link.

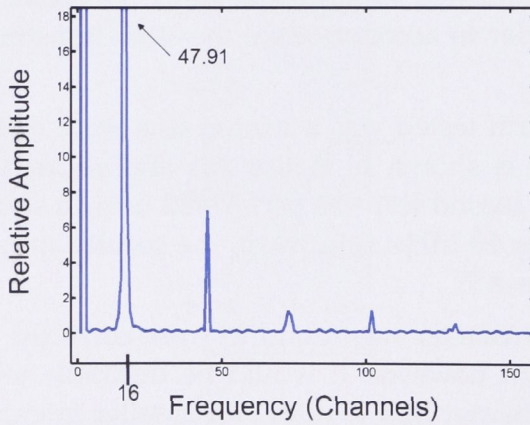
The first input waveform tested was a simple sine wave of frequency 10 MHz. The output from this is shown in figure A.6 and showed the system to be working correctly. The second test was performed using a square wave, this time at 1 MHz. As with the 10 MHz sine wave, the results appeared to be correct. This is shown in figure A.7.

In conclusion the spectrometer was found to work correctly. Were it to be used in a serious application however, it would be desirable to replace the serial connection with an ethernet link to allow for smaller integration times. Clock synthesis was also an issue. In an effort to keep things as simple as possible a separate clock (the same one as running the ADC) was used to clock the FPGA. Ideally the onboard 100 MHz clock would have been passed through a synthesizer to produce the desired frequency. The spectrometer was, however, tested up to a 32 MHz bandwidth and performed without error.

The System Generator library was also found to have great potential for the



**Figure A.6** The output from a test sine wave with a frequency of 10 MHz. The peaks shown in this diagram are due to the DC offset of the signal and the sine wave respectively.



**Figure A.7** The output from a 1 MHz square wave. The first peak in the figure is due to the DC offset. The subsequent peaks are due to the harmonic content of the square wave.

development, in particular, of prototyping software. The high level of abstraction provided by the block library allowed for common tasks to be implemented quickly. Care needs to be taken however to ensure that the design is created in such a way that it can be implemented efficiently on the FPGA chip. The other main benefit in using the System Generator library in the Simulink environment was the ability to take advantage of the Simulink and MATLAB® simulation abilities.



## APPENDIX B

---

### PCB AND SCHEMATIC DIAGRAMS

The schematics and circuit layout for the Parkes Multibeam Receiver interface board are given in figures B.1 to B.4. Table B.1 provides a description of the circuit elements used to populate the board. This board was designed in order to take the data and do the experiments of Chapter 4.



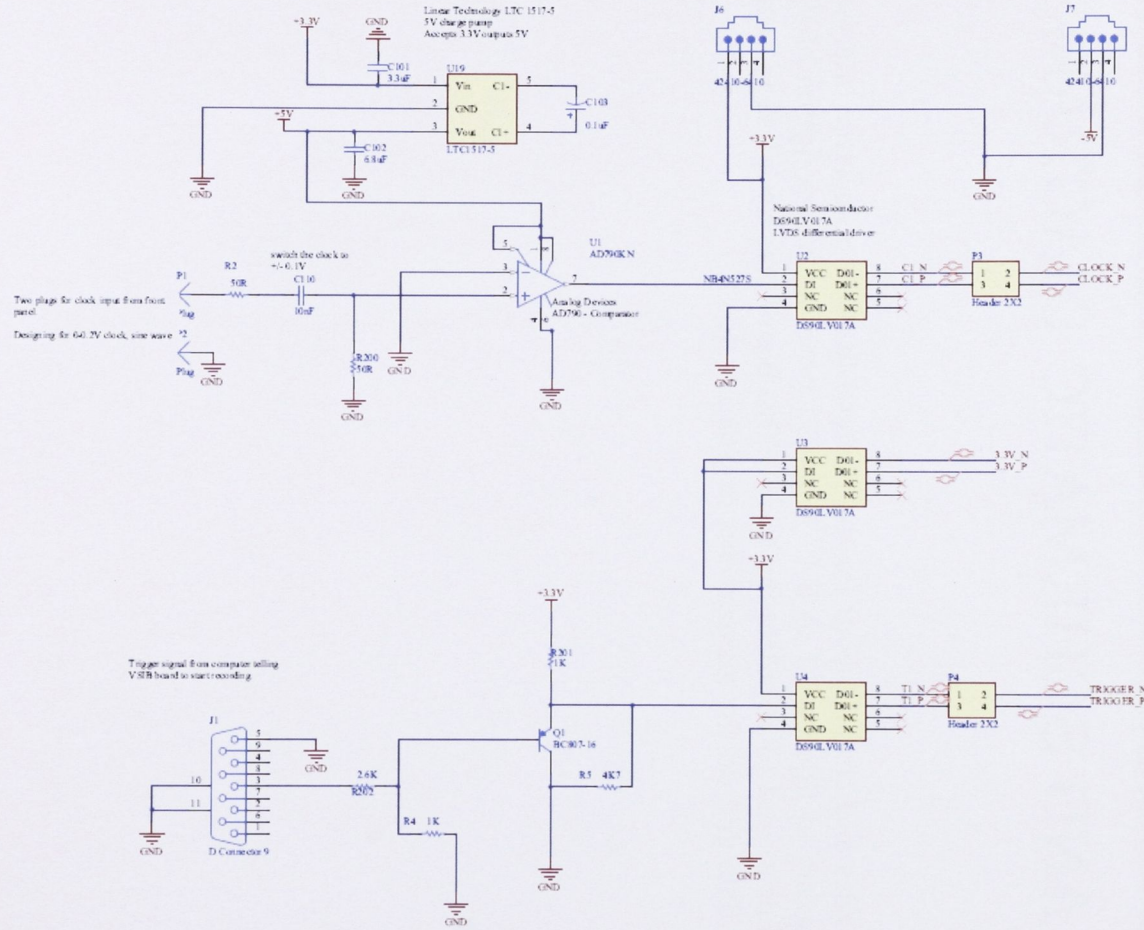


Figure B.1 Trigger and clock input schematic for Multibeam converter board.

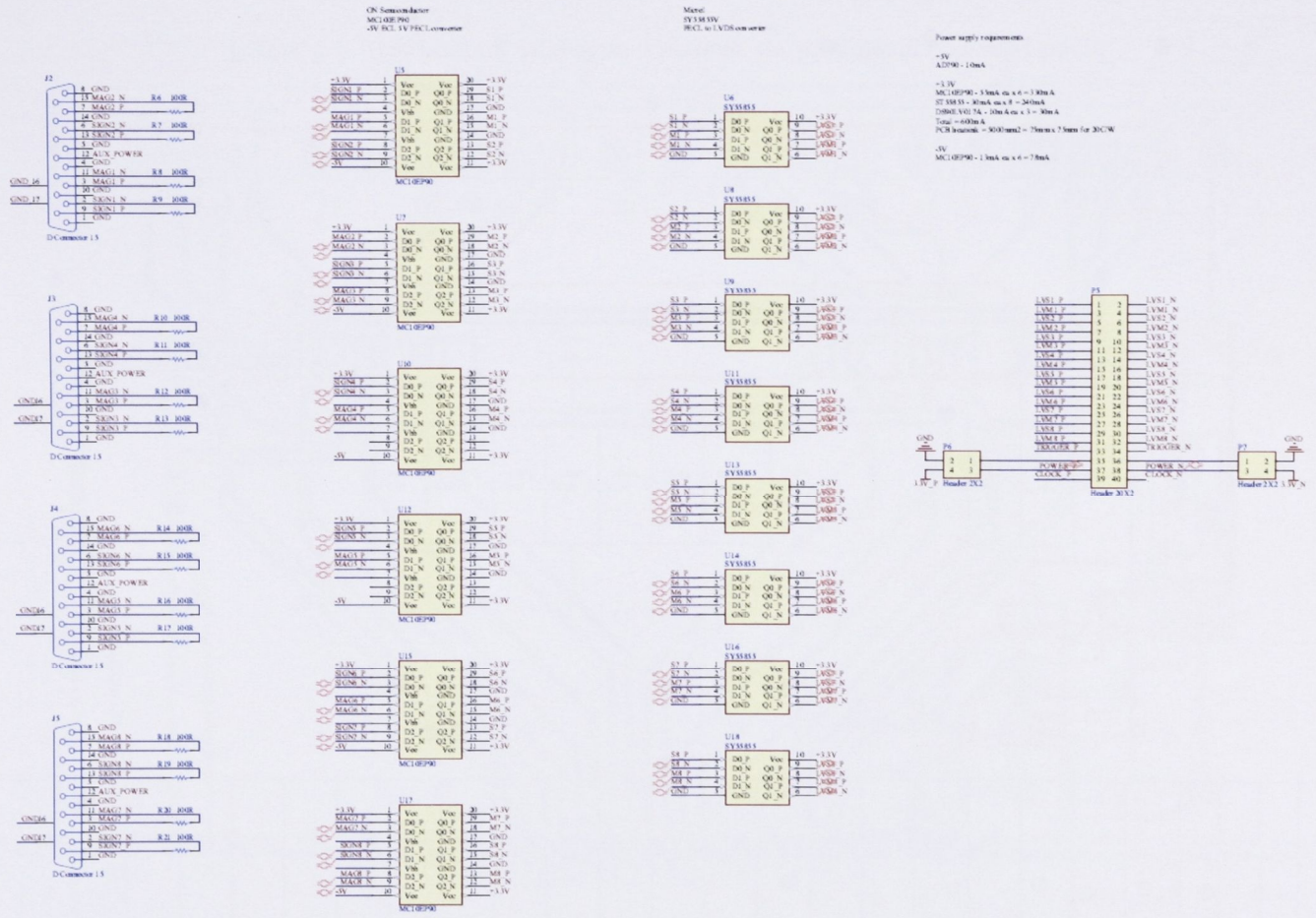


Figure B.2 Data input and conversion schematic for Multibeam converter board.

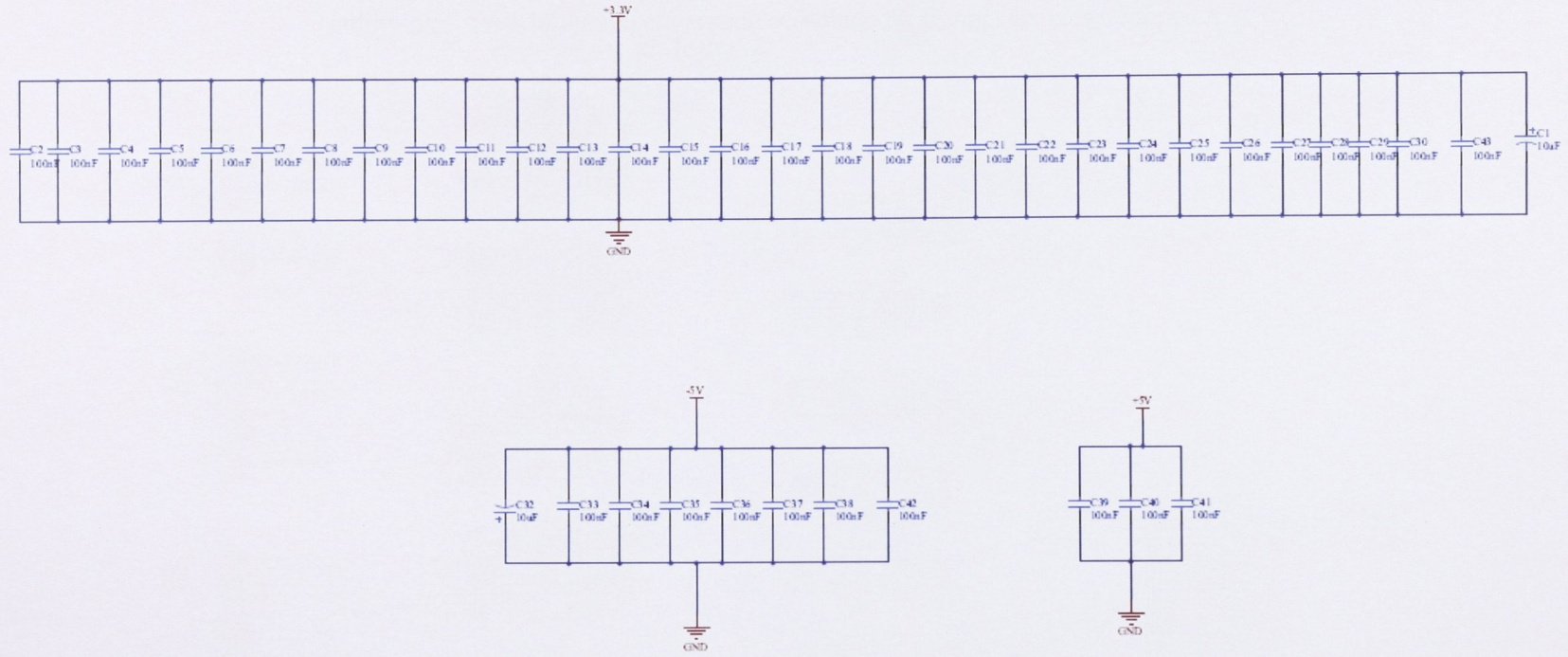


Figure B.3 Decoupling capacitors schematic for Multibeam converter board.

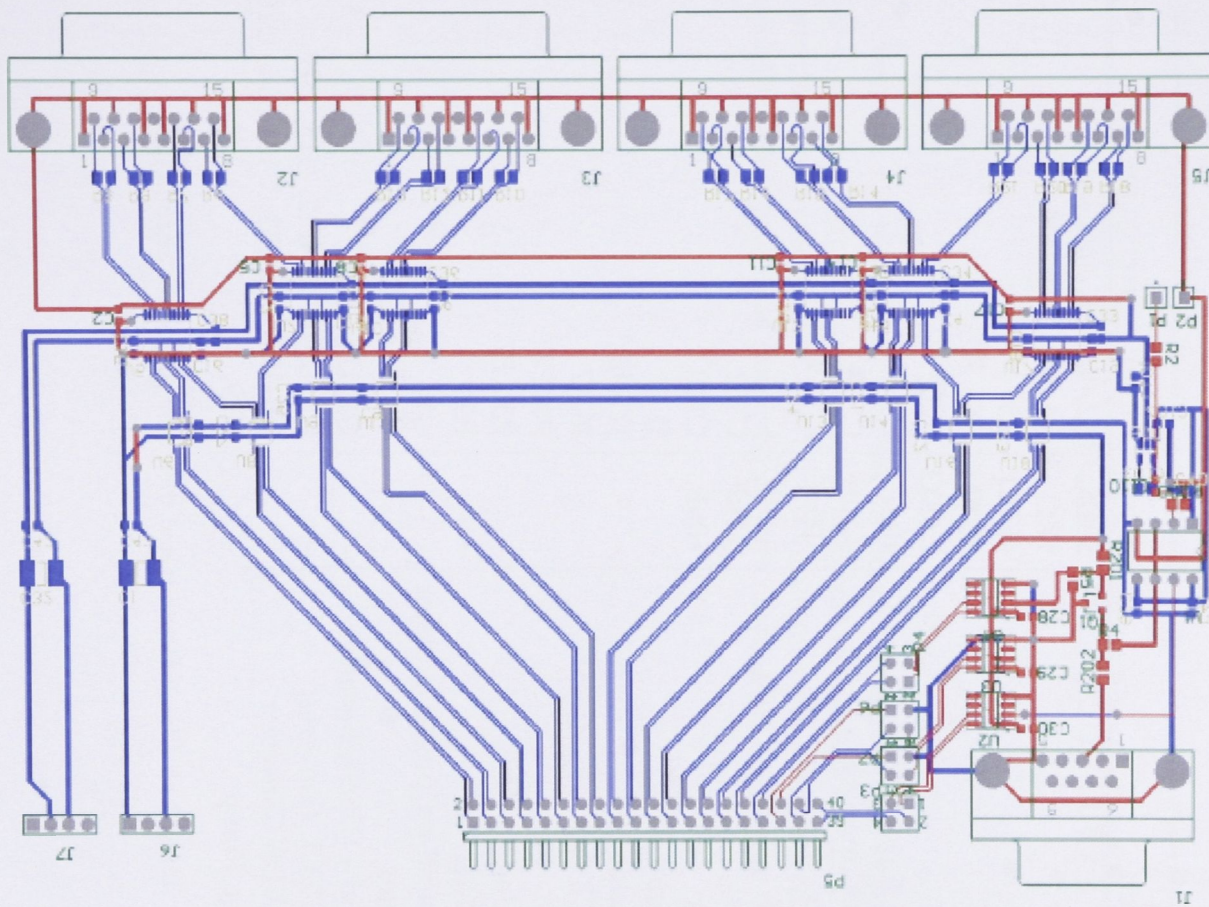


Figure B.4 Circuit layout for Multibeam converter board. Blue: Bottom of board, Red: Top of board.

Description	Value	Circuit reference number
Resistor	50Ω	R2, R200
Resistor	1KΩ	R4, R201
Resistor	4K7Ω	R5
Resistor	100Ω	R6 - R21
Resistor	2.6KΩ	R202
Polarised Capacitor	10uF	C1,C32
Polarised Capacitor	0.1uF	C103
Capacitor	100nF	C2 - C31
Capacitor	3.3uF	C101
Capacitor	6.8uF	C102
Capacitor	10nF	C110
D Connector 9	-	J1
D Connector 15	-	J2 - J5
0.125" PCB Jack	-	J6, J7
Plug	-	P1, P2
2X2 Header	-	P3 - P7
20X2 Header	-	P5
PNP General-purpose Transistor	-	Q1
Comparator	-	U1
Single End - LVDS Converter	-	U2, U3, U4
MECL - PECL Converter	-	U5, U7, U10, U12, U15, U17
PECL - LVDS Converter	-	U6, U8, U9, U11, U13, U14, U16, U18
5V Charge Pump	-	U19

**Table B.1** Circuit elements and their mapping to the circuit board.

## APPENDIX C

---

### ATIM DESIGN FOR MWA

The schematic for the ATIM showing in detail the inputs and outputs to the CPLD is given in figure C.1. This schematic was developed by Errol Kowald at Mt Stromlo Observatory, ANU, but the controlling code was written as part of this study. This circuit was designed as part of the MWA receiving system.



# APPENDIX D

---

## OBSERVATION DATA



Observation	Date	Time	Telescope	Receiver	Frequency	Bandwidth	Precision
1	30-06-2005	22:43:37 UTC	Parkes	CPSRII	675 MHz	64 MHz	2-bit
2	30-06-2005	22:48:35 UTC	Parkes	CPSRII	675 MHz	64 MHz	2-bit
3	30-06-2005	23:01:17 UTC	Parkes	CPSRII	675 MHz	64 MHz	2-bit
4	04-06-1999	06:41:00 UTC	ATCA		1503 MHz	4 MHz	4-bit
5	28-02-2008	12:37:00 AEST	Parkes	Multibeam	1440 MHz	8 MHz	2-bit
6	28-02-2008	13:11:00 AEST	Parkes	Multibeam	1440 MHz	8 MHz	2-bit
7	01-07-2008	13:11:00 AEST	Parkes	Multibeam	1440 MHz	8 MHz	2-bit
8	02-07-2008	10:27:00 AEST	Parkes	Multibeam	1440 MHz	8 MHz	2-bit
9	19-11-2008	03:48:47 UTC	MWA	32T System	261 MHz	1.28 MHz	10-bit complex
10	19-11-2008	00:19:24 UTC	MWA	32T System	261 MHz	1.28 MHz	10-bit complex
11	23-03 to 03-04-2004	-	Parkes	50 cm	685 MHz	64 MHz	2-bit

**Table D.1** Observational data times and receiver details.

Observation	Taken By	Source	Reference
1	John Reynolds	J0437-4715	Yagi & 3 m dish
2	John Reynolds	J0630-2834	Yagi & 3 m dish
3	John Reynolds	Zenith	Yagi & 3 m dish
4	Bell et al. (2000)	PKS 1934-638	1 ATCA Receiver
5	John Reynolds & Jonathon Kocz	Grenfell	Horn
6	John Reynolds & Jonathon Kocz	Zenith	Horn
7	John Reynolds & Jonathon Kocz	Zenith	Horn
8	John Reynolds & Jonathon Kocz	Vela (J0835-4510)	Horn
9	MWA Consortium members	Hydra A	none
10	MWA Consortium members	Sun	none
11	B. Little, S. Ord, J. Reynolds & J. Kocz	Various	Yagi

**Table D.2** Observers and sources.



---

## BIBLIOGRAPHY

- Athreya, R. 2009, A new approach to mitigation of radio frequency interference in interferometric data, Accepted ApJ
- Barnbaum, C., & Bradley, R. F. 1998, A New Approach to Interference Excision in Radio Astronomy: Real-Time Adaptive Cancellation, ApJ, 116, 2598
- Bell, J. F., Ekers, R. D., & Bunton, J. D. 2000, E&F White conference on Radio Frequency Interference Mitigation Strategies, arXiv:astro-ph/0005235
- Bell, J. F., Hall, P. J., Wilson, W. E., Sault, R. J., Smegal, R. J., Smith, M. R., van Straten, W., Kesteven, M. J., Ferris, R. H., Briggs, F. H., Carrad, G. J., Sinclair, M. W., Gough, R. G., Sarkissian, J. M., Bunton, J. D., & Bailes, M. 2001, Base Band Data for Testing Interference Mitigation Algorithms, Publications of the Astronomical Society of Australia, 18, 105
- Boonstra, A. J. 2005, Radio Frequency Interference Mitigation in Radio Astronomy, PhD thesis, Delft University of Technology
- Boonstra, A. J., van der Veen, A. J., & Raza, J. 2002, in IEEE International Conference on Acoustics, Speech, and Signal Processing, 2002. Proceedings. (ICASSP '02)., Vol. 3, IEEE, 2933–2936
- Bowman, J. D., Barnes, D. G., Briggs, F. H., Corey, B. E., Lynch, M. J., Bhat, N. D. R., Cappallo, R. J., Doleman, S. S., Fanous, B. J., Herne, D., Hewitt, J. N., Johnston, C., Kasper, J. C., Kocz, J., Kratzenberg, E., Lonsdale, C. J., Morales, M. F., Oberoi, D., Salah, J. E., Stansby, B., Stevens, J., Torr, G., Wayth, R., Webster, R. L., & Wyithe, J. S. B. 2007, Field Deployment of Prototype Antenna Tiles for the Mileura Widefield Array Low Frequency Demonstrator, AJ, 133, 1505

- Briggs, F. H., Bell, J. F., & Kesteven, M. J. 2000, Removing Radio Interference from Contaminated Astronomical Spectra Using an Independent Reference Signal and Closure Relations, *ApJ*, 120, 3351
- Briggs, F. H., & Kesteven, M. J. 2002, RFI subtraction with a reference horn: Application to pulsars and VLBI, 27th General Assembly of the International Union of Radio Science, Maastricht, The Netherlands
- Briggs, F. H., & Kocz, J. 2005, Overview of technical approaches to radio frequency interference mitigation, *Radio Science*, 40
- Chippendale, A., & Beresford, R. 2005, CSIRO ATNF SKA Site Monitoring Program
- Cornwell, T. J. 1989, The applications of closure phase to astronomical imaging, *Science*, 245, 263
- Eatough, R. P., Keane, E. F., & Lyne, A. G. 2009, An interference removal technique for radio pulsar searches, *MNRAS*, 395, 410
- Ellingson, S. W., Bunton, J. D., & Bell, J. F. 2000, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, Vol. 4015, *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, ed. H. R. Butcher, 400–407
- Ellingson, S. W., & Cazemier, W. 2003, Efficient multibeam synthesis with interference nulling for large arrays, *IEEE Transactions on Antennas and Propagation*, 51, 503
- Francis, P. J., Hewett, P. C., Foltz, C. B., & Chaffee, F. H. 1992, An objective classification scheme for QSO spectra, *ApJ*, 398, 476
- Fridman, P. A. 2001a, RFI excision using a higher order statistics analysis of the power spectrum, *A&A*, 368, 369
- . 2001b, RFI excision using a higher order statistics analysis of the power spectrum, *A&A*, 368, 369
- Hung, E. K., & Turner, R. M. 1983, A fast beamforming algorithm for large arrays, *IEEE Transactions on Aerospace and Electronic Systems*, AES:19, 598
- Jeffs, B. D., Lazarte, W., & Fisher, J. R. 2006, Bayesian detection of radar interference in radio astronomy, *Radio Science*, 41, 3006
- Kocz, J. 2004, *Radio Frequency Characterisation and Excision in Radio Astronomy*, Honours Thesis, FEIT, The Australian National University

- Leshem, A., & van der Veen, A. 1999, in *Perspectives on Radio Astronomy: Technologies for Large Antenna Arrays*, ed. A. Smolders & M. Haarlem (ASTRON, Netherlands Foundation for Research in Astronomy (NFRA)), 201–223
- Leshem, A., van der Veen, A., & Boonstra, A. 2000, *Multichannel Interference Mitigation Techniques in Radio Astronomy*, *ApJS*, 131, 355
- Lonsdale, C. J., Cappallo, R. J., Morales, M. F., Briggs, F. H., Benkevitch, L., Bowman, J. D., Bunton, J. D., Burns, S., Corey, B. E., deSouza, L., Doeleman, S. S., Derome, M., Deshpande, A., Gopalakrishna, M. R., Greenhill, L. J., Herne, D., Hewitt, J. N., Kamini, P. A., Kasper, J. C., Kincaid, B. B., Kocz, J., Kowald, E., Kratzenberg, E., Kumar, D., Lynch, M. J., Madhavi, S., Matejek, M., Mitchell, D., Morgan, E., Oberoi, D., Ord, S., Pathikulangara, J., Prabu, T., Rogers, A. E. E., Roshi, A., Salah, J. E., Sault, R. J., Udaya Shankar, N., Srivani, K. S., Stevens, J., Tingay, S., Vaccarella, A., Waterson, M., Wayth, R. B., Webster, R. L., Whitney, A. R., Williams, A., & Williams, C. 2009, *The Murchison Widefield Array: Design Overview*, Accepted, *Proceedings of the IEEE*
- Mitchell, Sault, B., & Kesteven, M. J. 2002, *Post correlation versus real-time adaptive RFI cancellation*, 27th General Assembly of the International Union of Radio Science, Maastricht, The Netherlands
- Mitchell, D. A. 2004, *Interference Mitigation in Radio Astronomy*, PhD thesis, University of Sydney
- Mittaz, J. P. D., Penston, M. V., & Sijnders, M. A. J. 1990, *Ultraviolet variability of NGC 4151 - A study using principal component analysis*, *MNRAS*, 242, 370
- Parsons, A., Backer, D., Chang, C., Chapman, D., Chen, H., Droz, P., de Jesus, C., MacMahon, D., Siemion, A., Werthimer, D., & Wright, M. 2005, *General Assembly of the International Union of Radio Science*
- Raza, J., Boonstra, A., & van der Veen, A. 2002, *Spatial filtering of RF interference in radio astronomy*, *IEEE Signal Processing Letters*, 9(2), 64
- Ritakari, J., & Mujunen, A. 2002, *The Metsahovi solution for gigabit VLBI*, *IVS CRL TDS News*, 20, 2
- Stanko, S., Klein, B., & Kerp, J. 2005, *A field programmable gate array spectrometer for radio astronomy. First light at the Effelsberg 100-m telescope*, *A&A*, 436, 391
- Strang, G. 1998, *Introduction to Linear Algebra*, 2nd Edition (Wellesley-Cambridge Press)
- Thompson, A. R. 1982, *The response of a radio-astronomy synthesis array to interfering signals*, *IEEE Transactions on Antennas and Propagation*, 30, 450

- Thompson, A. R., Moran, J. M., & Swenson, G. W. 2001, *Interferometry and Synthesis in Radio Astronomy*, 2nd edn. (John Wiley & Sons, Inc)
- van Straten, W., Britton, M. C., Bailes, M., Anderson, S. B., & Kulkarni, S. 2000, in *Astronomical Society of the Pacific Conference Series, Vol. 202, IAU Colloq. 177: Pulsar Astronomy - 2000 and Beyond*, ed. M. Kramer, N. Wex, & R. Wielebinski