School of Electrical Engineering, Computing and Mathematical Sciences

Towards Detection of Redshifted 21-cm Signal from Cosmic Dawn and Epoch of Reionisation

Jishnu Nambissan Thekkeppattu
0000-0002-3079-4046

This thesis is presented for the Degree of
Doctor of Philosophy
of
Curtin University

July 2022

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgement has been made. This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

[redacted]

Jishnu Nambissan Thekkeppattu
July 29, 2022

Statement of Contributions

The content of Chapter 3, SARAS 3 - A precision experiment to probe global 21–cm signal, is a reproduction of *SARAS 3 CD/EoR radiometer: design and performance of the receiver*, 2021, Experimental Astronomy. It is my own work, except for contributions to the development of experimental techniques and discussion from members of the SARAS team - Ravi Subrahmanyan, R. Somashekar, N. Udaya Shankar, Saurabh Singh, A. Raghunathan, B. S. Girish, K. S. Srivani and Mayuri Sathyanarayana Rao.

The content of Chapter 4, SITARA - A short spacing global 21 cm experiment, is a reproduction of *System design and calibration of SITARA - a global 21-cm short spacing interferometer prototype*, 2022, Publications of the Astronomical Society of Australia. It is my own work, except for contributions to the development of experimental techniques and discussion from Benjamin McKinley, Cathryn M. Trott, Jake Jones and Daniel C. X. Ung.

The content of Chapter 5, Singular spectrum analysis of SITARA time series, is a reproduction of a manuscript titled, *Singular spectrum analysis of time series data from low frequency radiometers, with an application to SITARA data* and is under review (on the day of thesis submission). It is my own work, except for contributions to the development of experimental techniques and discussion from Cathryn M. Trott and Benjamin McKinley.

(Signature of Candidate) (Signature of Supervisor) (Signature of Chair)

We acknowledge that Curtin University works across hundreds of traditional lands and custodial groups in Australia, and with First Nations people around the globe. We wish to pay our deepest respects to their ancestors and members of their communities, past, present, and to their emerging leaders. Our passion and commitment to work with all Australians and peoples from across the world, including our First Nations peoples are at the core of the work we do, reflective of our institutions' values and commitment to our role as leaders in the Reconciliation space in Australia.

Acknowledgements

First and foremost, I am deeply indebted to my thesis advisors Cathryn Trott and Benjamin McKinley, who offered the best advice, encouragement and support. They were always available for an interesting discussion, whether in person, email or Slack. I will definitely miss having interesting conversations with them - both the serious academic chats and light-hearted jokes !

I would like to acknowledge the support and guidance provided by my panel members - Nick Seymour, Rajan Chhetri, Amanda Wilber and Ian Morrison. They provided me important advice throughout my PhD journey and ensured that my PhD experience at CIRA is a pleasant one.

I would like to extend my sincere gratitude to my mentors Ravi Subrahmanyan (now with CSIRO) and Uday Shankar at Raman Research Institute Bengaluru (RRI). Thanks should also go to my other colleagues at RRI - Saurabh, Mayuri, Srivani, Som, Girish and Raghu. Mugundhan has always been an awesome friend and colleague.

Many thanks to the CIRA engineering team who supported the development of SITARA. Specifically, I would like to recognize the support provided by Daniel Ung, Jake Jones and Randall Wayth. David Kenney, Clinton Ward, David Emrich and Andrew McPhail assisted in various aspects of SITARA hardware development and deployment. I am also thankful to John Kennewell for offering space at Australian Space Academy, Meckering to validate the SITARA concept.

I am also grateful to my fellow PhD students at CIRA for their help and moral support. Many thanks to Steve, Mike, Kariuki, Susmita, Jai, Jaiden, Ferry, Gayatri, Himanshu, Aishwarya and Jun. Special thanks to Steve for pushing me to join a gym as well as asking the most interesting questions about radio

telescopes, and Mike and Kariuki for all the interesting discussions.

I could not have undertaken this journey without my family who supported me wholeheartedly throughout my academic career. A big thank you to my father, grandmother, uncles, aunts and cousins who provided me with support and encouragement when I needed them most. Finally, I dedicate this thesis to my late mother. She was my first teacher who taught me how to read, write and count, and someone who always gently demanded the best out of me.

Abstract

The epoch when the first stars formed and transformed the baryonic content of our Universe from being mostly neutral to almost fully ionised, appropriately called the cosmic dawn and epoch of reionisation (CD/EoR), is one of the least constrained periods in our understanding of the Universe. Observations leading to a better understanding of this epoch are therefore required. Specifically, the redshifted 21–cm signal from neutral hydrogen has the potential to provide valuable information on the baryonic evolution during this epoch. The sky-averaged or global component of the 21–cm signal is expected to be observable as a distortion to the low frequency radio spectrum. Albeit being conceptually simple to measure, precise measurements of the radio spectrum at the low frequencies are challenging due to the exquisite calibration requirements. This thesis is on experimental techniques to measure the global 21–cm signal from CD/EoR. With two different instruments, SARAS 3 and SITARA, two different approaches are explored.

SARAS 3 is a single antenna radiometer working in the CD/EoR band, with an aim to detect the global 21–cm signal. This experiment follows the design philosophy of maintaining *maximally smooth* responses for the receiver and antenna, such that the system does not introduce any spectral turnovers into the measured data, thereby making separation of the foregrounds and global 21–cm signal easier. In this thesis, the design and performance of the SARAS 3 receiver is detailed, with an emphasis on the calibration strategy, relevant equations and laboratory verification tests. With maximally smooth polynomial fits to the laboratory measurements, it is shown that there is no systematic in the SARAS 3 receiver that can mimic the expected global 21–cm signal.

SITARA follows a different experimental technique compared to other global 21–cm experiments. SITARA explores an interferometric technique, with two antennas placed in close proximity to exploit the enhancement of the mutual coherence of a uniform signal at short baselines. In this thesis, the experimental setup and initial results are detailed, including the challenges associated with calibration. SITARA is the first prototype to validate the short-spacing approach in the global 21–cm context. Analysis of SITARA data shows that there is indeed a response to the sky-averaged component, and a non-negligible coupling of receiver noise between the interferometer arms. A novel data analysis technique to understand the stability of the SITARA system is also presented in this thesis. For the first time, singular spectrum analysis (SSA) techniques are applied to time series data from low-frequency radiometers. The mathematical framework for the same is also outlined, along with some associated caveats. A few results related to the properties of anti-circulant matrices that are crucial for SSA, and the relationship between anti-circulant matrices and Fourier transforms are also presented. Simulations to validate the calculations are also provided. The techniques are subsequently applied to a single frequency channel time series from SITARA data, and the results point to diurnal variations in the system gains.

Thus, this thesis explores two different approaches to 21–cm experiments. Based on the outcomes of this thesis, I observe that the future of global 21–cm experiments will be based on a multipronged approach, with multiple experiments - each with its own unique design philosophy - working in synergy and cross-validating the observations of other experiments.

Contents

Acknowledgements **v**

Abstract **vii**

1 An introduction to cosmology **1**

1.1 The ΛCDM cosmological model 2

1.1.1 Cosmological parameters 4

1.2 Dark ages, cosmic dawn and epoch of reionisation 5

1.2.1 Observable effects of reionisation 6

1.2.1.1 Optical depth to reionisation 6

1.2.1.2 Lyman-α forest and Gunn-Peterson effect 8

1.3 The 21–cm signal 11

1.3.1 The global 21–cm signal 12

1.3.2 Low frequency foregrounds 16

1.3.3 Detection prospects and challenges 17

1.3.4 Current status of the field 19

1.3.5 A note on helium reionisation 20

1.4 Thesis outline 21

2 Low frequency experimental cosmology **23**

2.1 Measurement of radiation 23

2.1.1 Measurements with a radio telescope 25

2.2 Architecture of low frequency radio telescopes 28

2.2.1 Antennas 29

2.2.2 Analog electronics 31

2.2.3 Digitisation and Correlation 32

2.3 Low frequency experimental cosmology 33

2.3.1 Challenges 33

2.3.2 Global 21–cm experiments 35

2.3.2.1 Single antenna experiments 35

2.3.2.2 Interferometric experiments 36

3 SARAS 3 - A precision experiment to probe the global 21–cm signal 39

3.1 Abstract 39

3.2 Introduction 40

3.3 Motivation for the SARAS 3 receiver 44

3.4 System overview 46

3.4.1 Calibration considerations 53

3.5 Measurement equations 56

3.6 Implementation of the SARAS 3 receiver 60

3.7 Sensitivity of the SARAS 3 receiver 67

3.8 Laboratory tests 72

3.8.1 Absolute calibration 72

3.8.2 Termination tests 74

3.8.2.1 Modeling laboratory measurements with the measurement equation 76

3.8.2.2 Modeling laboratory measurements with maximally smooth polynomials 77

3.9 Summary of receiver tests 84

3.10 Results from SARAS 3 observations 87

4 **SITARA - A short spacing global 21 cm experiment** **89**

4.1 Abstract 89

4.2 Introduction 90

4.3 Background and motivation 91

4.3.1 Notations and conventions 95

4.4 SITARA System overview 95

4.5 Deployment and observations 99

4.6 Data calibration and analysis 102

4.6.1 A prelude on antenna radiation patterns 104

4.6.2 Calibration ignoring sky signal cross-talk 105

4.6.3 An empirical model for cross-talk 113

4.7 Results 118

4.7.1 Coupled receiver noise 119

4.7.2 Is SITARA sensitive to an all-sky signal? 122

4.8 Discussion 126

4.8.1 Future work 127

4.9 Conclusions 128

5 **Singular spectrum analysis of SITARA time series** **131**

5.1 Abstract 131

5.2 Introduction 132

5.3 Motivation 134

5.3.1 Notations and mathematical preliminaries 135

5.3.1.1 Circulant and anti-circulant matrices 136

5.3.1.2 Block matrices 137

5.3.1.3 Matrix products 137

5.4 Basics of singular spectrum analysis 137

5.5 SSA of periodic time series 140
5.5.1 SSA applied to an ideal periodic time series 140
5.5.1.1 Simulations 143
5.5.2 SSA applied to periodic time series with time-varying gains 148
5.5.3 On the use of SSA to aid calibration 152
5.5.4 SSA with diurnal gain variations 156
5.6 SSA applied to SITARA data 158
5.6.1 Data preparation 158
5.7 Discussion 160
5.7.1 Caveats and future work 162
5.7.2 A potential application of SSA for space-based 21–cm experiments 163
5.8 Conclusions 163

6 Summary and conclusions **165**
6.1 Single antenna experiments 165
6.2 Short-spacing interferometers 166
6.3 Calibration and data analysis techniques 166
6.4 The future of global 21–cm measurements 167

Appendices **169**

A Derivation of the SARAS 3 measurement equation **171**

B A physical model for cross-talk in SITARA **183**

C Eigendecomposition of anti-circulant matrices **189**

D Copyright Information **195**

Bibliography **201**

List of Figures

1.1 Quasar absorption spectra across several lines of sight (Loeb, 2008; Fan et al., 2006). As the redshifts of the sources increase, the neutral fraction and the Ly-α optical depths also increase rapidly, resulting in relatively featureless flat troughs. 10

1.2 A fiducial global 21–cm signal. The coloured regions approximately portray the redshifts where the spin temperature is determined by the indicated processes. The signal has multiple turning points as each process gives way to other competing processes in determining the 21–cm brightness temperature. In the region marked V, the Universe begins to reionise substantially, and 21–cm signal disappears. 16

1.3 Global 21–cm profiles with different f_*. The top figure has profiles when the star formation efficiency is low, leading to a slow reionisation. The bottom figure has a higher star formation efficiency, leading to faster reionisation compared to the former case. 18

2.1 A generic, high level block diagram of a low frequency radio telescope with N number of antennas. For brevity, only one signal chain is elaborated. The signal from an antenna goes through some analog processing and subsequently gets digitised and Fourier transformed (F). The Fourier transformed data from multiple antennas are then correlated (X), giving visibilities. 30

3.1 Schematic of the SARAS 3 receiver architecture. 48

3.2 SARAS 3 analog electronics at the base of the antenna. A block diagram of the front end electronics is shown in top panel. Bottom panel shows the measured gain of this unit, from the antenna terminal to the output of optical to electrical modulator housed inside remote station analog electronics. 61

3.3 SARAS 3 analog electronics at the remote station 150 m from the antenna. Top panel is a block diagram of the electronics signal chain. The bottom panel shows magnitude of measured gain of one arm of the analog electronics, from the output of the optical to electrical demodulator to sampler input. The other arm has an identical response and is not shown here. 63

3.4 The analog electronics enclosures beneath the antenna and at the remote base station. In the top panel on the left and right are shown the enclosure at the antenna base; the left panel shows its view from the top and in the right panel is shown the view from below with the bottom cover removed. The bottom panel shows the enclosure that houses the analog electronics at the remote station. 65

3.5 Expected antenna temperature for an observation with the SARAS 3 antenna, at latitude +14° and over local sidereal time (LST) from 10 to 18 hr. Also shown is the expected average sky spectrum. It may be noted that the antenna temperature is defined as per Eq.2.13, such that it is the average sky spectrum multiplied by the mismatch factor or reflection efficiency of the SARAS 3 antenna (Raghunathan et al., 2020) 70

3.6 An example of the noise component that might be present in an 8 hr mock observation. The rms noise expected across the band is also shown in orange. 70

3.7 The distribution in rms noise for smoothing to different noise equivalent widths, for the 8-hr mock observation. It may be noted that the rms noise estimate has been referred to the sky domain by correcting for the total efficiency of the radiometer; therefore, the rms noise represents 1-σ uncertainty in measurement of sky brightness temperature. 71

3.8 Fit to data acquired with termination at antenna port placed in warm and cold baths, which provides estimate of the absolute calibration scale factor T_{STEP}. 74

3.9 Modeling measured data in the 50–100 MHz and 90–180 MHz bands using the measurement equation Eq. 3.14. The sum and difference of data acquired with the antenna replaced with precision open and short terminations are modeled. 78

3.10 The result of modeling SARAS 3 systematics in the CD band 50–100 MHz. The sum of measurement data acquired with precision electrical open and short terminations at the antenna terminals is fitted using a maximally smooth polynomial form. The top panel shows the measurement data and fit together; the residuals are magnified by factor 20 for clarity. The middle panel shows the fitting residuals smoothed using kernels of increasing fwhm. The data with native resolution of 61 kHz is the lowest trace and spectra smoothed progressively to larger fwhm are shown above that with offsets of 0.25 K; traces are magnified by factors that keep the apparent rms the same on all smoothing. The legend in this middle panel lists the rms at different spectral resolutions. The continuous line in the bottom panel shows the run of variance in the residuals versus spectral resolution; the expected rate of fall in noise with smoothing is indicated by the dotted line. 80

3.11 Modelling measurement data in the 50–100 MHz CD band. Here the difference of data acquired with open and short terminations at the antenna terminals is modeled. The three panels depict analyses same as that in the previous figure. 81

3.12 Modeling laboratory measurement data in the EoR band 90–180 MHz. In this figure, the sum of calibrated spectra acquired with open and short terminations at the antenna terminals is modeled. The three panels depict analyses same as that in the previous figure; however, for the wider EoR band the data was modeled using the modified form of the maximally smooth function described in the text, which allows for one zero crossing in higher order derivatives. 82

3.13 Modeling for systematics in the EoR band 90–180 MHz. In this figure, data that is the difference of those acquired with open and short terminations at the antenna terminals is examined. Here also, as in the previous figure, the fitting function is a modified maximally smooth function. The three panels depict analyses same as that in the previous figure. 83

3.14 Modeling measurement data acquired with a 1-port antenna circuit simulator replacing the antenna. The three panels depict analyses same as that done in the previous figure. 85

4.1 A high level block diagram of SITARA; auxiliary details such as power supplies as well as attenuators used for impedance matching between various modules are not shown. The multiplication units shown in the digital receiver perform conjugate multiplication. . . 96

4.2 SITARA system as deployed at MRO. The left photograph shows SITARA antennas and fieldbox; the cables have since been tied to the ground plane. The specific dipoles used in this experiment are highlighted in blue ellipses. The inset shows the antenna orientation and relevant dimensions where the inactive dipoles have been greyed out. The right photograph shows SITARA back-end electronics inside the Telstra hut. The receiver box houses the SNAP and RPi as well as media converters for networking. Signal conditioning module (SCM) contains the amplifiers and filters to perform analog processing before digitization and correlation. . . . 100

4.3 Time-frequency (waterfall) plot of the cross-correlation data collected on May 17th-18th, 2021. Panel B is the time-frequency plot of the magnitude of the complex visibilities. Panel A is the average spectrum and panel C shows the power as a function of LST for a frequency of 70 MHz. The data are unflagged and uncalibrated. The waterfall plot shows the sky drifting through SITARA beam; the peak occurs when the Galactic plane is at the local zenith. On closer inspection, the data shown in this figure are seen to contain Solar bursts between 1-2 hours LST. 101

4.4 Variations in uncalibrated power with local sidereal time (LST) for data collected on Mar 14-15, 2021 and April 05-06, 2021. The top figure shows the power in a single frequency channel in antenna 1 auto-correlations and the bottom figure shows the magnitude of antenna 1-2 cross-correlations. The colored regions in the plots show the night time LSTs for the corresponding day. 103

4.5 Simulated antenna radiation patterns (H-plane) as a function of zenith angle for two MWA dipoles spaced 1 m apart in parallel configuration. The patterns at 90 MHz are identical to each other and are well approximated by an ideal dipole $\cos^2(\text{ZA})$ pattern while the patterns at 180 MHz have shifted peaks away from zenith. 105

4.6 Simulated SITARA auto and cross antenna patterns at two frequencies, in Mollweide projection. For comparison, patterns for an isolated MWA antenna are given in the top row. The plots are peak normalised as shown in the colour bar. The coordinate system is local altitude-azimuth with the centre of the Mollweide projection corresponding to zenith; the local directions are also shown. It can be seen that due to mutual coupling, the patterns of closely spaced SITARA antennas diverge from that of an isolated MWA dipole. 106

4.7 Receiver gains and noise temperatures as functions of frequency. The plots are semi-logarithmic to accommodate a wide dynamic range. The gains show the filtering introduced by the system at 70 MHz and 200 MHz. The gains include contributions from antennas, analog stages as well as *any scaling introduced by the digital signal processing in the correlator*, therefore the units are arbitrary. The noise temperatures are calibrated to units of kelvin. An interesting feature in the receiver noise temperatures is that the coupled receiver noise in cross-correlations is almost an order of magnitude less than receiver noise in autocorrelations. 111

4.8 Variations in calibrated and T_{Nij} subtracted data as functions of local sidereal times (LST). The top panel shows calibrated auto-correlations along with simulated auto-correlations and the bottom panel shows magnitude of the calibrated cross-correlations along with simulated cross-correlations. Only data in the shaded region are used for calibration, since those LSTs have a rapid change in the sky temperature due to Galaxy transit. The solutions derived are then used for the entire data. It may be noted that T_{Nij} subtraction also removes any 21 cm signal from the data. 112

4.9 Calibrated and T_{Nij} subtracted data for $\sim$ 174 MHz. The top panel shows calibrated auto-correlations along with simulated auto-correlations and the bottom panel shows magnitude of the calibrated cross-correlations along with simulated cross-correlations. The plot is of the same nature as Fig.4.8, however at this frequency the individual antenna radiation patterns differ. Despite this being captured by the FEE simulations, the simulated temperatures differ from calibrated data. 114

4.10 Comparison of gains estimated with and without cross-talk. The plots are semi-logarithmic to accommodate the dynamic range. As noted in the text, each "gain" in the cross-talk model is a sum of coefficients that includes cross-talk. Despite using two different formalism, it can be seen that they are in close agreement. 118

4.11 Differences between gains estimated with and without cross-talk. In this plot, the fractional differences between the gains estimated with and without cross-talk are shown as percentages. The auto-correlation gains derived with the two models have a difference less than 10% while the cross-correlation gains differ by about 20% at frequencies where the antennas patterns are dissimilar. 119

4.12 A comparison between SITARA data at 174 MHz with a model that does not consider cross-talk and one that considers cross-talk. Plots (A) and (B) are the auto-correlations and (C) is the cross-correlation magnitude. The data are forward modelled and therefore not in units of brightness temperature. Data from shaded area alone are used to compute gains and receiver noises. With the cross-talk model, the simulations match the data. 120

4.13 Comparison between SITARA data and simulations for the cross-correlation magnitude. Shown are the temperature-temperature plots between the SITARA data and simulations based on the two models. Two frequencies where the individual antenna patterns are dissimilar are chosen. We expect the simulations to follow data in a linear fashion in this plot, if the model used for simulations is accurate. While the model neglecting cross-talk fails to explain the variations in data, the cross-talk model fits the data very well. . . 121

4.14 A comparison between estimations of receiver noise with and without cross-talk considerations. The receiver noise estimates are not calibrated to units of kelvin. It is seen that when cross-talk is modelled, the estimations of coupled receiver noise are generally lower, especially below $\sim$ 150 MHz where the antenna beams are similar. 122

4.15 Ratio of estimated coupled receiver noise temperature to an auto-correlation receiver noise temperature. As expected, the cross-coupled receiver noise in data is substantially lower than auto-correlation receiver noise. The data have been smoothed with a Savitzky-Golay filter to reduce noise in the plots. 123

4.16 Comparing measured coherence (black) with simulations assuming a uniform sky (red) and a more realistic GSM foregrounds (blue). Uncalibrated data with receiver noise subtracted from auto and cross-correlations are used for this computation. 125

5.1 SITARA raw data time series as a function of Julian Date (JD) for a single frequency channel of bandwidth 61 kHz, at a frequency of 111 MHz. The data have been extracted from a concatenated SITARA dataset for the month of June 2021. Solar bursts contribute most of the RFI seen in this time series. 135

5.2 Simulated time series at a frequency of 111 MHz and the corresponding trajectory matrix formed by choosing an embedding dimension $L = 96$. 144

5.3 SVD of the trajectory matrix. Panel (A) shows singular values in a semilogarithmic scale. The same panel shows the eigenvalues of an $L \times L$ submatrix as well as the DFT spectrum of the periodic sequence. The DFT spectrum has been sorted according to descending magnitude and artificially scaled by a value of 4 to make it distinguishable from the eigenspectrum. For clarity, only the first 19 values are plotted. Panel (B) shows the first 5 left singular vectors while Panel (C) shows the first 5 right singular vectors. In both plots, the orthogonal (sine-cosine) vector pairs are plotted with the same colour but with different line-styles. 145

5.4 A comparison between the left singular vectors of the trajectory matrix and eigenvectors of a square submatrix. Only the vectors from the first orthogonal pair in each case are plotted. 146

5.5 Reconstructed series from SSA of an ideal periodic series. The 0^{th} component and the succeeding 5 grouped orthogonal pairs are given in this figure. Since the reconstructed series add up to the original time series, their x-axis is the number of days. 147

5.6 Simulated time series with gains at a frequency of 111 MHz and the corresponding trajectory matrix 150

5.7 Reconstructed series from SSA of a simulated series with time-varying gains. The 0^{th} series and the succeeding 5 components from grouping orthogonal pairs are given in this figure. 151

5.8 Injected and recovered gain templates. The plots have been mean subtracted and divided with their standard deviation for normalisation. 152

5.9 Injected and recovered gains. The recovered gains have been normalised according Eq.5.23. The recovered gains from periodic components resemble the injected gains. 154
5.10 Calibrating mean sky component using recovered gains from periodic components. For comparison, the expected mean levels with and without the 100 K receiver noise temperatures are also plotted. 155
5.11 Simulated radiometer data with gains that have a smoothly varying component and a diurnal component. 156
5.12 Injected and SSA recovered gains when the gains have a diurnal component. Appropriate normalisation has been applied. 157
5.13 Application of recovered gains from periodic components to calibrate the mean-sky term, when the injected gains have a diurnal component. 157
5.14 Reconstructed series from SITARA data SSA. A plot of the physical temperature recorded within the Murchison shire is included in the trend plot to show the anti-correlation between both. . . . 159
5.15 Recovered gains from SITARA data, June 2021. The gains differ across reconstructed series, pointing to potential diurnal gain variations. 161

A.1 Simplified noise model for the SARAS 3 radiometer, when connected to the antenna. 172
A.2 Simplified noise model for the SARAS 3 radiometer, when connected to the reference termination and calibration source. 174
A.3 Simulations of the first 5 leading terms of the receiver noise alone using Eq.A.30, and assuming a 1 m lossless cable. The top panel shows when the termination is open, and the bottom one shows when the termination is short. It may also be noted that the terminations considered here do not contribute to the system temperature and hence the resulting spectra are a result of the receiver noise and reflections of it. 181

B.1 Cross-talk factor f_c as a percentage. Please see the text for caveats associated with this calculation. 186

D.1 Co-author attribution for Chapter 3 199
D.2 Co-author attribution for Chapter 4 200

Chapter 1

An introduction to cosmology

> ***Socrates*** *: Shall we set down astronomy among the objects of study?*
> ***Glaucon*** *: I think so, to know something about the seasons, the months and the years is of use for military purposes, as well as for agriculture and for navigation.*
> ***Socrates*** *: It amuses me to see how afraid you are, lest the common herd of people should accuse you of recommending useless studies.*
>
> - Plato, *The Republic*

Cosmology is the study of our Universe at large. The term Universe encompasses everything known to us, from the largest building blocks known to us to the smallest quantum of matter. However, cosmology is concerned predominantly with the study of the Universe at large, that is the scales set by galaxies, clusters and superclusters with studies at other scales left for other fields.

It can be safely claimed that every civilisation on Earth had their own model of the Universe with mythology combined with observations, but modern cosmology as a science can be said to have origins in the revolutions initiated by the groundbreaking heliocentric models put forth by Copernicus and the subsequent observations conducted by Galileo. The tool used by Galileo to observe the heavens - the telescope - has since evolved over centuries into sophisticated instruments with highly advanced optics. Each new wavelength window opened up by advances in technology has resulted in corresponding advances in astronomy and cosmology. Today telescopes are conceived, designed and built at wavelengths well larger and smaller than the optical wavelengths (∼600 nanometers).

Modern cosmology is a well established scientific discipline with a body of theory validated by experiments and observations. Quite surprisingly, this cosmology is a recent development with the first major turning point perhaps being the serendipitous discovery of the cosmic microwave background (CMB) by Penzias and Wilson in 1965 and its interpretation (Penzias & Wilson, 1965; Dicke et al., 1965). The presence of an isotropic radiation field with a near perfect blackbody spectrum, that could be explained only by an epoch of the Universe when it was much smaller than the current epoch, led to the acceptance of what is called the "big bang" cosmological model. The Big Bang puts forth a cosmological model with an evolving Universe that expands with time. Along with the discovery of dark matter and dark energy, this has resulted in what is called the lambda cold dark matter (ΛCDM) cosmology.

1.1 The ΛCDM cosmological model

Over the past few decades, intensive theoretical modelling and precision measurements have resulted in a standard cosmological model. The emergent cosmological model - called the ΛCDM model - has relatively few parameters, however it does a remarkable job of explaining a wide variety of observations. This model has a spatially flat Universe that is homogeneous and isotropic at large scales, is composed of ordinary matter, radiation, nonbaryonic cold dark matter, and dark energy. The growth of large scale structure is seeded by tiny, nearly scale invariant primordial adiabatic Gaussian fluctuations.

The ΛCDM model, which is the concurrent cosmological model, can be summarised as follows. The beginning of the Universe - the so called big bang - remains a mystery and our current knowledge of physics is insufficient to understand it. This was followed by a period of rapid exponential expansion of the Universe - called the inflation. This epoch lasted from about 10^{-36} to 10^{-32} seconds after the origin. Inflation seeds the structure formation by magnifying the microscopic quantum fluctuations to cosmic scales. The next major milestone in the history of the Universe is nucleosynthesis. During this epoch, that lasted for several 10s of seconds after the origin, nuclei of elements (or isotopes) heavier than that of hydrogen (i.e. protons) are formed. The major nuclei synthesised during this epoch are deuterium, helium-3 and 4 and lithium-7. These are still in the ionised form, as the temperature is too high for neutral atoms to form. When the temperature falls to approximately 4000K, the nuclei combine with the electrons to form the very first neutral atoms in our Universe. This epoch is called the epoch of recombination. Along with this, the Universe became transparent to radiation, thereby “releasing” the primordial radiation, which was essentially “trapped” in the primordial plasma. This relic radiation is what we now observe as the CMB. Consequently, the CMB that we now see reveals the conditions when the Universe recombined. Combining our (redshift $z = 0$) knowledge of CMB temperature $T_\gamma = 2.725$ K and the temperature for recombination to occur as ~ 4000 K, the redshifts when the Universe recombined can be estimated to be $z \sim 1400$.

Before recombination, matter (in the form of plasma) and radiation were in a tightly coupled condition, with processes such as Compton scattering, double Compton scattering and bremsstrahlung ensuring equilibrium between them and ironing out departures from equilibrium states. Even if energy is injected, Compton scattering distributes the energy among photons. Double Compton and bremsstrahlung adjusts the photon number density, as they do not conserve photon numbers. These processes therefore result in the primordial radiation acquiring a Planckian, or black body spectrum. It should be mentioned here that equilibration resulting in a perfect black-body spectrum is an idealized view of the early Universe. Spectral distortions are expected due to energy injections when the thermalisation processes became inefficient at maintaining equilibrium (Tashiro, 2014). These spectral distortions, categorized into μ and y distortions, are yet to be discovered experimentally. Indeed, one of the cornerstones of the ΛCDM model is the presence of an isotropic radiation possessing a blackbody spectral shape with very small spectral and spatial fluctuations. Measurements of the CMB spectrum along with its spectral distortions and spatial fluctuations can validate as well as determine the parameters of the cosmological model.

COBE-FIRAS, launched in 1989, measured the spectrum of the CMB and found it to be near-perfect blackbody with a temperature of 2.72548 ± 0.00057 K (Fixsen, 2009), and set upper limits of $|\mu| < 9 \times 10^{-5}$ and $|y| < 1.5 \times 10^{-5}$ (Fixsen et al., 1996). An early verification of the black-body nature of the CMB spectrum came with rocket-borne measurements with COBRA (Gush et al., 1990). With a series of precision space-borne experiments such as COBE-DMR (Smoot et al., 1990), WMAP (Bennett et al., 2003) and Planck (Tauber et al., 2010), CMB angular characteristics and thereby the cosmological parameters have been accurately determined; some of these parameters are given in Sec.1.1.1. In general, measurements of the CMB spectrum at short wavelengths or its angular fluctuations are of limited utility in probing the epochs following recombination, as the Universe became largely transparent to CMB photons after recombination. However, post-recombination processes that absorb from or emit to the CMB can still result in spectral and spatial fluctuations, and detection of them can yield information on the post-recombination evolution of the Universe. In Sec.1.2.1.1 the utility of CMB measurements of a specific parameter (τ) to constrain the post-recombination evolution of the Universe is provided.

1.1.1 Cosmological parameters

The basic ΛCDM model requires at least six independent parameters to describe the Universe. The choice of the parameters is quite flexible; in practice parameterisations are chosen to avoid potential issues such as degeneracies that may arise when fitting experimental data with models. Consequently, parameterisation can vary from experiment to experiment. Nonetheless, once a set of six parameters have been obtained, the others can be derived. For example, the fit parameters adopted by Hinshaw et al. (2013) using WMAP with priors from other experiments (eCMB, BAO, H0) are given in Table.1.1.

Parameter	Symbol	Value
Dark matter density	$\Omega_c h^2$	0.1153 ± 0.0019
Baryon density	$\Omega_b h^2$	0.02223 ± 0.00033
Dark energy density parameter	Ω_Λ	$0.7135^{+0.0095}_{-0.0096}$
Scalar spectral index	n_s	0.9608 ± 0.0080
Optical depth to reionisation	τ	0.081 ± 0.012
Curvature fluctuation amplitude	Δ^2_R	$2.464 \pm 0.072 \times 10^{-9}$

Table 1.1: Cosmological parameters from WMAP analysis, Hinshaw et al. (2013)

1.2 Dark ages, cosmic dawn and epoch of reionisation

Recombination resulted in the baryonic matter content of the Universe becoming mostly neutral and hence, largely transparent to electromagnetic radiation including CMB. The epochs following recombination had structure formation from the primordial fluctuations, resulting in the stars and galaxies that we see in the present Universe. The period in between recombination and when the first stars turned on had no sources of radiation except the CMB, and hence is called the dark ages.

The tiny matter density fluctuations created during inflation act as the seeds for structure formation during the dark ages. The first galaxies formed when primordial gas collapsed onto the dark matter potential wells at the peaks of the matter density field. The gas radiatively cooled predominantly via Lyman-α transitions. The condensing of gas resulted in the first stars and black holes. The birth of these sources resulted in the emission of electromagnetic radiation in our Universe (apart from the CMB), and therefore this epoch is referred to as the cosmic dawn. It is important to note that there were no heavier elements (so-called metals in astrophysics) to assist in cooling during these epochs, therefore the stars formed were of a different nature to the metal-enriched population I stars or metal-poor population II we observe in our local Universe. The first stars are hence called population III stars. The emission of radiation from these first sources ionised the Universe, and the corresponding epoch is called the epoch of reionisation. Henceforth, the abbreviation of DA/CD/EoR will be used to denote the epochs of dark ages, cosmic dawn and reionisation.

To begin with, it can be assumed that the reionisation occurred at a specific redshift as is assumed in instantaneous reionisation models (Griffiths et al., 1999). Then two regions can be identified, one where the hydrogen is all in neutral form and the other where it is fully ionised. In the region with neutral hydrogen, the atoms would interact with any radiation through the levels in the hydrogen, governed by the relevant dynamics. In the region where the hydrogen is fully ionised, the ions would interact with the radiation through mechanisms such as Compton scattering, inverse Compton scattering and the non-relativistic variation of Thomson scattering. These different mechanisms can therefore be used to probe the Universe before and after reionisation.

1.2.1 Observable effects of reionisation

As mentioned previously, the cosmological phase transition caused by the reionisation leaves behind certain signatures, which are observable. While multiple probes have been proposed to study the Universe before and after reionisation, the dominant ones are the 21–cm signal emitted by neutral hydrogen, the Gunn-Peterson effect and the optical depth to reionisation constrained by observations of the CMB. The latter two will be discussed in this section, while the 21–cm signal will be discussed in detail in the next section. In an instantaneous reionisation model, the 21–cm signal and the Gunn-Peterson effect pertain to epochs when the Universe was fully neutral while the optical depth pertains to the ionised Universe. However, this division is not a straightforward one.

1.2.1.1 Optical depth to reionisation

The CMB photons released during recombination have to travel through the electrons in the reionised Universe. Consequently, these photons encounter Thomson scattering with these electrons. The scattering of the CMB photons by these electrons can result in multiple observable effects (Haiman & Knox, 1999). As scattering leads to blending of photons from different lines of sight, the primary anisotropies in the CMB are damped. However, the CMB photons can also acquire some momentum from inverse Compton scattering with hot plasma, resulting in a secondary anisotropy. Thomson scattering has a polarisation dependency, therefore this scattering leaves behind imprints in the polarisation of the CMB anisotropies. Interestingly, these can aid in a measurement of the extent of the Thomson scattering in the post-recombination epochs. The relevant measurand is the integrated optical depth to reionisation, given in Eq.1.1.

$$\tau = n_H(0) c \sigma_T \int_0^{z_{max}} x_e(z) \frac{(1+z)^2}{H(z)} \, dz \tag{1.1}$$

with $x_e(z) \equiv n_e^{reion}(z)/n_H(z)$. The number density of free electrons from reionisation is given by n_e^{reion} and the total number of hydrogen nuclei by $n_H(z)$. σ_T is the Thomson scattering cross-section and $H(z)$ is the Hubble parameter (Planck Collaboration et al., 2020). The optical depth is a dimensionless quantity (Rybicki & Lightman, 1979).

As described in Sec.1.1.1, the optical depth to reionisation is one of the key parameters that can be obtained from precision measurements of the CMB angular fluctuations as well as polarisation. Interestingly, compared to other cosmological parameters τ is somewhat distinct in the sense that the value is not determined by the primordial physics, but rather by the post-recombination evolution of the Universe.

As can be seen from Eq.1.1, the optical depth depends on the limits of the redshift. Therefore obtaining an optical depth to Thomson scattering from CMB data can constrain the EoR, using some assumed model for the evolution of $x_e(z)$. Combined with polarisation data, Planck Collaboration et al. (2020) set the following limits on the optical depth and reionisation redshift.

$$\begin{aligned} \tau &= 0.0544^{0.0070}_{0.0081} \\ z_{re} &= 7.68 \pm 0.79 \end{aligned} \tag{1.2}$$

where z_{re} is the mid-point redshift of reionisation, assuming a `tanh` reionisation model.

Prior to WMAP, there was no actual measurement of τ. With the launch of precision cosmology experiments WMAP and Planck, measurements of τ became available. Based on the first year WMAP data alone, the best-fit value of optical depth was measured in Spergel et al. (2003) to be $\tau = 0.166^{+0.076}_{-0.071}$. However the optical depth got better constrained to $\tau = 0.089 \pm 0.030$ in Spergel et al. (2007), which made use of three years of WMAP data. The difference arises from the inclusion of polarisation power spectra (EE) in Spergel et al. (2007) that has more power to constrain τ. In Hinshaw et al. (2013), a value of $\tau = 0.081 \pm 0.012$ was reported based on WMAP data, along with some other experiments (same as given in Sec. 1.1.1). However, in Planck Collaboration et al. (2020), the optical depth was reported as $\tau = 0.0544^{0.0070}_{0.0081}$ using Planck data. As can be seen, this measurement is substantially smaller than the value initially reported in Spergel et al. (2003). The differences over the years can be attributed to various factors such as tighter control of systematics, inclusion of polarisation data and larger frequency coverage. Consequently, measurements of the mid-point redshift of reionisation also got smaller, with the values reported in Planck Collaboration et al. (2020) in agreement with measurements based on quasar absorption spectra.

1.2.1.2 Lyman-α forest and Gunn-Peterson effect

The spectral lines in the ultraviolet (UV) range related to electron transitions between $n \geq 2$ to $n = 1$ energy levels in neutral hydrogen form the Lyman series. The transitions between $n = 2$ to $n = 1$ give rise to the Lyman-α line at a wavelength of 121.567 nm or 1215.67 Å. Due to their association with neutral hydrogen, Lyman lines can be used to study neutral hydrogen, specifically Ly-α is a very prominent line that can yield information on the EoR.

If there is a cloud of neutral hydrogen with a source of strong UV radiation (such as qausars) behind it, the radiation interacts with the cloud giving rise to absorption features in the spectrum of the background source in the Ly-α wavelength. If the background source is receding away from the cloud, an interesting effect becomes noticeable. The radiation from the background source with energy greater than (or wavelengths blueward of) Ly-α gets redshifted and some of it falls in the Ly-α wavelength range in the rest-frame of the cloud and gets absorbed. Therefore in an expanding Universe, the spectra of distant sources of UV radiation will show a “forest” of absorption lines in their spectra blueward of Ly-α, as the radiation passes through clouds of neutral hydrogen at various redshifts in the intergalactic medium (IGM) between us and them.

However, if there is a substantial amount of neutral hydrogen between the source and us, the forest of lines will blend into an absorption trough. This effect, first predicted by in Gunn & Peterson (1965), is therefore aptly named the Gunn-Peterson trough. Since the wavelengths blueward of Ly-α are absorbed in the trough, other Lyman lines of smaller wavelengths can be ignored.

The first observations of the Gunn-Peterson trough in a $z = 6.28$ quasar (Becker et al., 2001) provided evidence for the presence of neutral hydrogen in large quantities in the IGM at those epochs. Subsequently, the Gunn-Peterson trough has been observed in several other high-redshift ($z > 6$) quasars - see Fig.1.1 for a compilation of observations of multiple quasars from the Sloan digital sky survey (SDSS). These observations of the Gunn-Peterson trough in multiple high redshift quasars provide strong observational evidence that the Universe had a much higher fraction of neutral hydrogen at redshifts beyond $z \sim 6$. On the other hand, the lack of absorption troughs in the spectra of quasars at lower redshifts shows that the present Universe does not have sufficient neutral hydrogen in the IGM. These constraints firmly show that the Universe transitioned from being fully neutral to fully ionised at around $z \sim 6$, thereby marking the end of epoch of reionisation.

However, despite being a valuable tool in constraining the EoR, Ly-α is of limited utility in studying the evolution of the reionisation process itself. For this, let us look at the relevant Gunn-Peterson optical depth τ_{GP} given in Eq.1.3 (Becker et al., 2001).

$$\tau_{GP} = 1.8 \times 10^5 \; h^{-1} \Omega_m^{-1/2} \frac{\Omega_b h^2}{0.02} \left(\frac{1+z}{7} \right)^{\frac{3}{2}} \left(\frac{n_{HI}}{n_H} \right) \tag{1.3}$$

For a mean redshift of $z \sim 10$, and taking cosmological parameters as $\Omega_b \sim 0.05$, $\Omega_m \sim 0.3$ and $h \sim 0.7$, it can be seen that $\tau_{GP} > 1$ for $\frac{n_{HI}}{n_H} > 10^{-6}$. Therefore, even for a small neutral fraction, the IGM becomes optically thick to the Ly-α

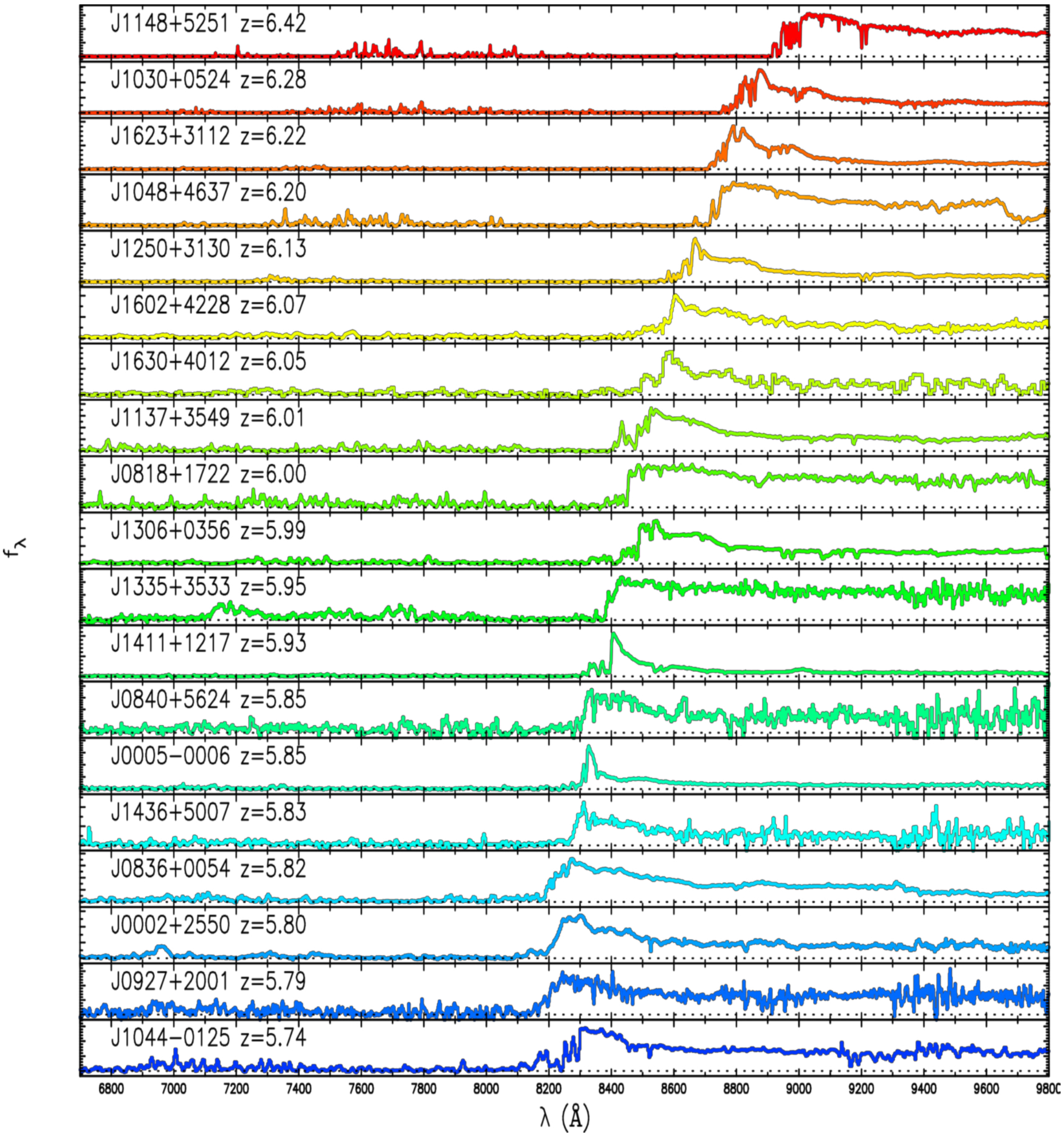


Figure 1.1: Quasar absorption spectra across several lines of sight (Loeb, 2008; Fan et al., 2006). As the redshifts of the sources increase, the neutral fraction and the Ly-α optical depths also increase rapidly, resulting in relatively featureless flat troughs.

transition, thereby diminishing the amount of information on higher redshifts that can be extracted from the spectra. This effect can be noticed in Fig.1.1 where the high redshift quasars show relatively flat G-P troughs.

Hence, what is required is a spectral line that is optically thin at all relevant epochs. Such a probe is provided by the 21–cm microwave spectral line arising from the hyperfine transition of neutral hydrogen, redshifted to metre wavelengths.

1.3 The 21–cm signal

The hyperfine splitting of the ground-state (1s) of neutral hydrogen leads to a spectral line. This line arises from the spin-flip transition of the electron in the ground state and the associated energy is quite small, resulting in the line belonging in the microwave regime. This spectral line, at 1420 MHz or a wavelength of 21–cm, is extensively used for the study of neutral hydrogen in our local Universe. It is therefore one of the most important tools used in radio astronomy. However, the utility of the 21–cm line is not limited to our local Universe; the 21–cm line from the early Universe has been recognized as one of the most important probes of the DA/CD/EoR (Scott & Rees, 1990). The line is optically thin even for a fully neutral Universe (Furlanetto et al., 2006), making it an ideal probe to study the neutral Universe before reionisation. Owing to cosmological expansion, the line would be redshifted and ought to be observable in what is the very high frequency (VHF) range of the radio spectrum (30–300 MHz).

Since the formation of the first sources of radiation that ionised the Universe occurred as a result of matter clumping in gravitational potential wells, the 21–cm signal is expected to have angular fluctuations. While it is tempting to draw similarities between these fluctuations and CMB angular fluctuations, some differences have to be noticed. CMB fluctuations offer a “snapshot” image of the Universe when it recombined, and the radiation is of black-body nature. Therefore the CMB fluctuations do not vary as a function of the observing wavelength (ignoring the distortions from other secondary effects). On the other hand, 21–cm fluctuations vary as a function of redshift as the formation of the structure progresses. As the 21–cm line is a narrow spectral line, mapping these fluctuations as a function of frequency therefore is - to a first order - equivalent to mapping the formation of the first stars and the associated ionised bubbles as a function of redshift. Creation of such maps, often called 21–cm tomography, offers an unprecedented view of the early Universe. However, obtaining such maps requires sensitivities that are difficult to achieve with the current generation of radio telescopes such as the MWA (Tingay et al., 2013) and LOFAR (van Haarlem et al., 2013); it is one of the key science goals of the Square Kilometre Array

(SKA, Dewdney et al. (2009)) that is expected to achieve the required sensitivity. Therefore, the current generation of radio telescopes attempt to measure the angular fluctuations of the 21–cm statistically, via a measurement of their spatial power spectrum, setting upper limits on plausible reionisation scenarios (Trott et al., 2020; Mertens et al., 2020). In the same vein, it is interesting to note that HERA, which is a dedicated experiment for 21–cm angular fluctuations (DeBoer et al., 2017), measures a variant of the power spectrum in the form of the delay spectrum (The HERA Collaboration et al., 2021).

1.3.1 The global 21–cm signal

The potential of the sky-averaged component of the redshifted 21–cm signal as a tool to probe CD/EoR was first pointed out in Shaver et al. (1999), though Varshalovich & Khersonskii (1977) had pointed out the utility of the 21–cm signal in probing the dark ages. As an average over the fluctuations, the global signal sets the mean level of the cosmological 21–cm transition at each frequency or redshift. Though not as detailed in complexity as the power spectrum, the global component is complementary to power spectrum measurements. For example, the global signal can unambiguously determine emission or absorption as its unit is mK, while power spectra are in units of mK^2. Besides, the global signal is expected to have considerably higher amplitude, while requiring conceptually simple radiometers to measure, as the signal being sought is a sky-averaged component. This makes dedicated experiments to probe the global component of the 21–cm signal attractive.

It is useful to detail some basics of the 21–cm signal and the expected spectral features. The brightness temperature δT_b of the 21–cm signal from a patch, relative to the CMB, is given by Eq.1.4.

$$\delta T_b = 27 x_{HI}(1+\delta)\left(\frac{\Omega_b h^2}{0.023}\right)\left(\frac{0.15}{\Omega_m h^2}\frac{1+z}{10}\right)^{1/2}\left(\frac{T_s - T_\gamma}{T_s}\right), \qquad [\mathrm{mK}] \tag{1.4}$$

where x_{HI} is the hydrogen neutral fraction, and T_γ is the CMB temperature.

The parameter δ is the fractional overdensity and is not of consequence to the global 21–cm signal. The spin temperature T_s defined in Eq.1.5, is a convenient handle to describe the relative fine-structure level populations, however it is not a physical temperature.

$$\frac{n_2}{n_1} = 3\ exp\Big(- \frac{h\ \nu_{21}}{k_B\ T_s}\Big) \tag{1.5}$$

From a simplistic perspective, most of the features in the global 21–cm signal during DA/CD/EoR can be viewed as a tug-of-war competition between coupling spin temperature to the CMB temperature and gas kinetic temperature. When coupled to the CMB the signal is invisible, but when coupled to the gas temperature, the signal is visible in absorption or emission (against the CMB) depending on the gas temperature. Three competing mechanisms determine the coupling and thereby the evolution of the hydrogen spin temperature; these three processes are :

1. Scattering of CMB photons - this causes absorption/emission of 21–cm photons from/to the CMB. This results in 21–cm spin temperature approaching CMB temperature i.e. $T_S \sim T_\gamma$. The assumption here is that the background radiation in the Universe consists of the CMB alone and no other "excess" emissions.

2. Collisions with other hydrogen atoms and electrons. This results in the 21–cm spin temperature approaching the kinetic temperature of the gas i.e. $T_S \sim T_K$.

3. Scattering of Ly-α photons that effectively couples spin temperature to gas kinetic temperature. This effect is called the Wouthuysen-Field effect.

While scattering and collisions are easy to visualise, the Wouthuysen-Field effect is a somewhat distinctive process, in that it indirectly couples the spin temperature to kinetic temperature, mediated by Ly-α radiation (Wouthuysen, 1952; Field, 1958). A simple explanation of this effect is as follows. Absorption of Ly-α radiation causes the neutral atoms to change from 1s energy state to 2p states. Emission of Ly-α causes the atoms to return to ground state. However, the atom before absorption can have been in a certain hyperfine state (say singlet), while the emission of Ly-α can result in it returning to a different state (say triplet). Thus, absorption and emission of Ly-α photons by neutral hydrogen can result in a change in the spin state. This effectively couples the spin temperature to a "colour temperature", that is approximately the kinetic temperature of the gas (Field, 1959).

The dependence of the spin temperature on the three processes is given by

$$T_s^{-1} = \frac{T_\gamma^{-1} + x_\alpha T_c^{-1} + x_c T_K^{-1}}{1 + x_c + x_\alpha}, \qquad [\mathrm{K}^{-1}] \tag{1.6}$$

where T_γ is the CMB temperature, x_c is the coupling coefficient for collisions, T_K is the gas kinetic temperature, x_α is the coupling coefficient for Wouthuysen-Field effect and T_c is the colour temperature; $T_c \sim T_K$ (Field, 1959). The approximate regions where each of these processes determine the 21–cm brightness temperature are given in Fig.1.2. The global 21–cm profile in this figure has been generated with the help of the `ares` [1] simulation tool (Mirocha, 2020, 2014). A description of the various regions in Fig.1.2 and the characteristics of the physical processes determining the signal is given below (Pritchard & Loeb, 2012).

1. At high redshifts ($1100 \geq z \geq 200$), the residual free electrons leftover from recombination maintain thermal coupling between the gas temperature and the CMB. This is also a region where the collisions lead to spin temperature getting coupled to gas temperature. Thus, during these redshifts, the spin temperature is the same as the CMB temperature, leading to no detectable 21–cm signal. This region is marked I in Fig.1.2.

2. At redshifts of $z \sim 200$, the efficacy of thermal coupling (mediated by electrons) between gas and the CMB diminishes, as the residual electrons recombine. As a result, the gas decouples from the CMB and cools adiabiatically at a rate faster than the CMB. However, collisions still determine the spin temperature. Thus, the spin temperature is coupled to cooling gas and is effectively lower than that of the CMB, leading to absorption. This region is marked II in Fig.1.2.

---

[1] https://github.com/mirochaj/ares

3. Owing to cosmological expansion, collisions become ineffective in coupling spin temperature to gas temperature. Therefore the spin temperature couples back into CMB temperature leading to a weak 21–cm signal as seen in III in Fig.1.2. This, along with the absorption in II, leads to an absorption trough prediction as seen in the same figure.

4. With sufficient matter collapsing into dark matter haloes, the first radiative sources in our Universe turn on. The soft-UV radiation from them couples the spin temperature to the gas kinetic temperature again, via the Wouthuysen-Field mechanism. However, the bulk of the IGM is still cold. The gas in the IGM is much colder than the dark ages and therefore the absorption is expected to be much larger in amplitude. With the first sources emitting radiation, the X-rays begin to heat the gas such that the IGM temperature begins to rise; however the signal is still in absorption. X-rays, being effective in penetrating the gas, heat the IGM more effectively than UV. The heating of the cold gas introduces a turning point in the spectrum, resulting in a 21–cm absorption trough. This signature can be seen in the region marked IV in Fig.1.2.

5. As X-rays continue to heat the IGM, the gas temperature rises to above of that of the CMB. However, the spin temperature is still coupled to the gas temperature via the Wouthuysen-Field mechanism, therefore the 21–cm signal appears in emission. By now, the UV radiation starts carving out reionised regions in the IGM, marking the epoch of reionisation. Since the 21–cm signal can originate only from neutral hydrogen, the emission signature diminishes as the entire IGM undergoes ionisation, and then the signal drops to undetectable levels. This region is marked V in Fig.1.2. However, a limit on the redshifts at which reionisation is complete can be obtained from Ly-α observations, thereby overcoming the limitations of the 21–cm signal at these epochs.

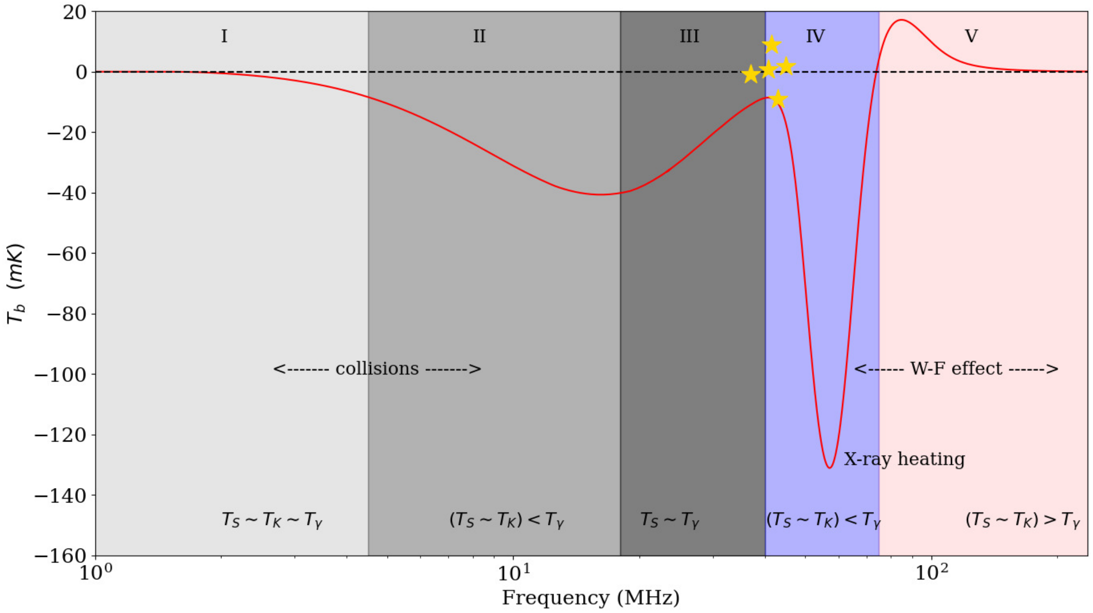


Figure 1.2: A fiducial global 21–cm signal. The coloured regions approximately portray the redshifts where the spin temperature is determined by the indicated processes. The signal has multiple turning points as each process gives way to other competing processes in determining the 21–cm brightness temperature. In the region marked V, the Universe begins to reionise substantially, and 21–cm signal disappears.

1.3.2 Low frequency foregrounds

Perhaps the most important difference between CMB observations above a few GHz and low frequency observations (at $\sim$ 100 MHz) is the strong foregrounds at radio frequencies. The low frequency radio sky is dominated by emissions from relativistic electrons and cosmic rays, spiralling through Galactic magnetic fields. At the radio frequencies, the spectrum of emissions from these relativistic charged particles follow their energy distribution. Since their energy distribution follows a power-law, the radio emission also follows a power-law. This emission has been given the name of synchrotron radiation, owing to its resemblance to the mechanism by which synchrotrons produce strong radiation (Rybicki & Lightman, 1979). These emissions can have substantially higher power relative to the 21–cm signal; while the 21–cm signal is of the order of a few 100s of mK, the foregrounds can have brightness temperatures of a few 1000s of K at frequencies below 100 MHz. The spectra measured by experiments are a sum of the foregrounds, along with the 21–cm signal. This necessitates a signal separation operation to remove foregrounds from the spectra, with a dynamic range of about 10^6. However, physically motivated simulations of the low frequency foregrounds have shown that they have smooth spectra with frequency, while the global 21–cm has multiple turning points (Sathyanarayana Rao et al., 2017b). Therefore, in principle, foregrounds can be fully modelled with smooth polynomials, including maximally smooth polynomials (Sathyanarayana Rao et al., 2015) or low-order non-smooth polynomials, while the 21–cm signal leaves behind a residual due to its multiple turning points. Ideally, the basis functions describing the foregrounds would be fully orthogonal to the signal being sought. However, this is difficult in practice as the signal and foregrounds share some similarities in their spectral structures, and some signal loss is unavoidable in most circumstances. Nonetheless, useful information can be extracted from the 21–cm signal residuals.

1.3.3 Detection prospects and challenges

It should be emphasised that Fig.1.2 and the associated description of the various epochs provide a very generic picture of the DA/CD/EoR era. In reality, these epochs have extremely poor constraints from our observations (or lack thereof). Therefore the overall shape, strength and the turning points of the actual global 21–cm signal - that can be revealed only with experiments - can be very different from theoretical modelling with several simplifications. Indeed, the nature of the global 21–cm signal depends on several astrophysical parameters, most of which are poorly understood. While an extensive investigation of the dependence of the signal on the various parameters including their mutual dependencies and degeneracies is beyond the scope of this thesis, I will demonstrate the effect of two parameters, namely the star formation efficiency f_* and the X-ray luminosity parameter f_X. The parameter f_X determines the X-ray emissions and therefore X-ray heating. For low values of f_X, the heating is slow and spans a large redshift range, while larger values of f_X induce rapid heating.

Fig.1.3 shows the effect of these two parameters on the global 21–cm signal. As in previous sections, the signal templates have been generated with the `ares` tool. It can easily be seen that the signal can vary considerably from one astrophysical

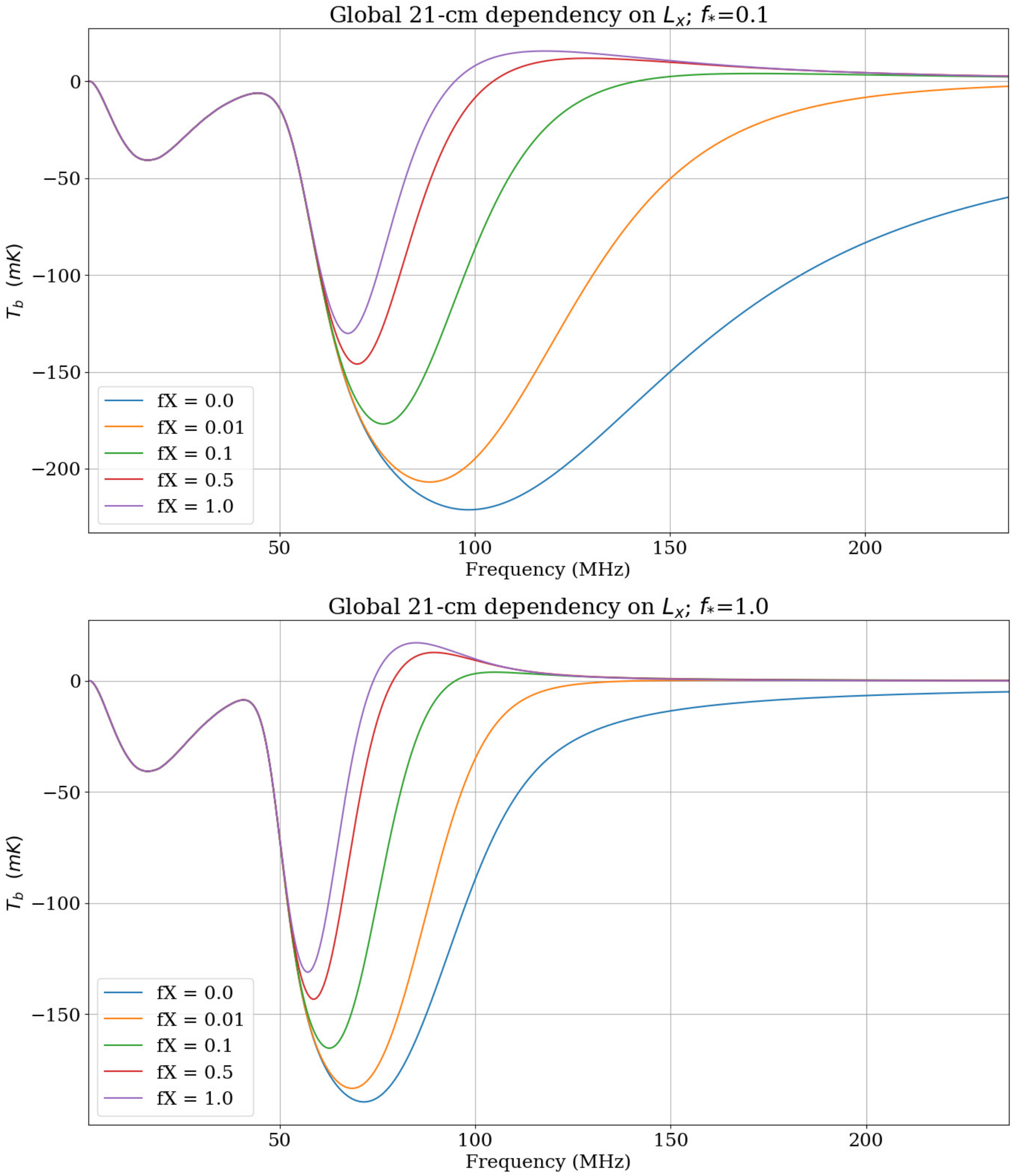


Figure 1.3: Global 21–cm profiles with different f_*. The top figure has profiles when the star formation efficiency is low, leading to a slow reionisation. The bottom figure has a higher star formation efficiency, leading to faster reionisation compared to the former case.

setting to another. Consequently, some conditions can be easier to detect or rule out with experiments while others can be hard to infer from experiments. For example, if the X-ray heating was efficient, the cosmic dawn absorption trough would be weak in amplitude, requiring larger integration times to reduce noise, and a detection would be considerably more difficult than otherwise. As described in Sec.1.3.2, the bright low-frequency radio foregrounds complicate the detection problem. If the global 21–cm signal has smooth/broad features that span a large frequency range, they tend to be correlated and the signal gets absorbed by the terms describing foregrounds, thereby making a detection difficult. Optimally, if an experiment covers the entire CD/EoR frequency range encompassing all the potential turning points, the 21–cm signal and foregrounds become less correlated and the prospects for detection increase considerably.

1.3.4 Current status of the field

The search for the global 21–cm signal has attracted considerable attention internationally in the recent past. Experiments such as EDGES (Rogers & Bowman, 2012; Monsalve et al., 2017) and SARAS (Patra et al., 2013; Singh et al., 2018b), having progressed through multiple iterations, provided constraints on the 21-cm signal models. Analysis of EDGES high-band data in Monsalve et al. (2017) rejected several reionisation scenarios. Specifically, Monsalve et al. (2017) ruled out at 2σ significance models with duration of up to $\Delta z = 1$ at $z \approx 8.5$ and higher than $\Delta z = 0.4$ for reionisation scenarios with 21–cm spin temperature higher than that of the CMB. For "cold" reionisation scenarios, with 21–cm spin temperature is perfectly coupled to the IGM temperature which is not heated by early stars or their remnants, they reject reionisation models of duration $\Delta z < 2$. These two constraints were made with `tanh` models of reionisation. Monsalve et al. (2017) also rejected several Gaussian models for the absorption trough.

Results from the SARAS-2 experiment rejected reionisation models with inefficient X-ray heating, along with rapid reionisation (Singh et al., 2018a). The models rejected by SARAS-2 have rapid reionisation with no X-ray heating or late heating due to very inefficient sources. In these models, due to lack of sufficient heating of the IGM, the gas does not have enough time to reach the temperature of CMB, thus resulting in the 21–cm signal being in absorption throughout the EoR. It may be noted that while Monsalve et al. (2017) predominantly used a `tanh` model for reionisation, Singh et al. (2018a) used simulated profiles of the 21–cm profiles by Cohen et al. (2017).

Other single antenna global 21–cm experiments include BIGHORNS (Sokolowski et al., 2015), SCI-HI(Voytek et al., 2014), PRIZM (Philip et al., 2019) and LEDA (Price et al., 2018). Upcoming single antenna global 21–cm experiments include REACH, MIST and HYPEREION (Patra et.al., in prep).

It may be noted that the global 21–cm experiments had been providing constraints and upper limits on the reionisation scenarios and not a measurement of the 21–cm signal profile itself. However, an analysis of the EDGES low-band data resulted in a tentative detection of the 21–cm absorption trough from cosmic dawn (Bowman et al., 2018). The recovered template with a flattened profile has an amplitude of about 500 mK at a centre frequency of 78 MHz, and is at least twice as large as the deepest absorption troughs from theoretical modelling. Combined with the shape of the signal not conforming to the expected Gaussian-like signals, the detection posed as a challenge requiring independent verification. Part of the work described in this thesis has led to an independent experimental investigation into the aforesaid signal, resulting in a non-detection of the same in the SARAS-3 data. Therefore, more experiments with much better controlled systematics are required, if an unambiguous detection is to be made. This thesis also proposes a novel experimental strategy using interferometers in short spacing, as opposed to isolated single antennas, for global 21–cm detection.

1.3.5 A note on helium reionisation

In this thesis, the word reionisation is used to refer to the reionisation of hydrogen in our Universe. Similar to hydrogen, the helium in the IGM also has undergone a reionisation process. Since helium has an atomic number of two, the reionisation in this context refers to the singly ionised helium (He II) getting further ionised into doubly ionised helium (He III). The first ionisation (He I → He II) is expected to have occurred along with hydrogen reionisation, as the ionisation potential of He I is only 24.58 eV which is 1.8 times that of hydrogen (Wyithe & Loeb, 2003). This epoch of helium reionisation, however does not correspond directly to structure formation or birth of the first stars. He II has an ionisation potential of 54.4 eV, while H I has a potential of only 13.6 eV. Therefore the “softer” photons from the first stars could not have ionised He II (Furlanetto & Oh, 2008). Indeed, the ionisation of He II to He III is expected to have happened at later redshifts of about $3 \leq z \leq 4$, with ionisation driven by quasars (Sokasian et al., 2002).

Similar to neutral hydrogen, singly ionised helium-3 (^{3}He II) has a hyperfine transition, at a frequency of about 8.7 GHz (Schuessler et al., 1969). Combined with the redshifts of helium reionisation being around 3, this implies that a spectral line tracing the evolution of helium reionisation ought to be observable in the GHz frequencies. While observational prospects of this line are somewhat uncertain - some reviews can be found in McQuinn & Switzer (2009), Bagla & Loeb (2009) and Khullar et al. (2020) - the GHz frequency range offers substantially lower foregrounds along with availability of better instrument calibration techniques (compared to lower frequencies). Therefore, future experiments targeting helium reionisation spectral lines can provide complementary information on the baryon evolution in our Universe.

1.4 Thesis outline

This thesis on experimental cosmology is based on two separate experiments, both targeting a detection of the global 21–cm signal. The current chapter (Chapter 1) introduces some basic cosmology to set the experimental goals. Chapter 2 presents concepts of low-frequency experimental cosmology, with some high level descriptions of various constituent components of a practical experiment. Chapter 3 describes the SARAS 3 experiment, with emphasis on the receiver design and laboratory validation that resulted in a non-detection of the cosmic dawn signal detected by EDGES. Chapter 4 introduces a novel experimental strategy for detection of the global 21–cm signal using interferometers instead of single-element designs such as SARAS. Chapter 5 describes a novel analysis technique for low frequency radiometric data using singular spectrum analysis (SSA). The mathematical foundations to apply SSA to radiometric time series data are explained, and as an application the techniques are demonstrated with SITARA data. Chapter 6 concludes the thesis.

Chapter 2

Low frequency experimental cosmology

> *An experiment is a question which science poses to Nature, and a measurement is the recording of Nature's answer.*
>
> \- Max Planck

In this chapter, I outline some concepts related to experimental cosmology. I begin with an introduction to the concepts associated with measurement of radiation. I then proceed to the principles upon which radio telescopes measure radiation. The general architecture of low frequency radio telescopes is discussed afterwards. I touch upon challenges related to global 21–cm experiments, and end the chapter with a brief description of the various techniques used for global 21–cm investigation.

2.1 Measurement of radiation

One of the fundamental measurands for a telescope is the specific intensity I_ν. Consider an area dA of the telescope that is normal to the incoming radiation. The energy dE collected by the telescope from radiation coming within a solid angle $d\Omega$, in a bandwidth $d\nu$ in time dt is then given as

$$dE = I_\nu \; dA \; dt \; d\Omega \; d\nu \tag{2.1}$$

where I_ν is the specific intensity (Rybicki & Lightman, 1979). In other words specific intensity is energy per unit time, per unit bandwidth, per solid angle, per area. Therefore specific intensity has units of $\mathrm{W\,m^{-2}\,Hz^{-1}\,sr^{-1}}$. Specific intensity is a conserved quantity; it remains the same from a source to the receptor, provided energy is not added to or subtracted from the ray via emission, absorption or scattering.

For black-body radiation, specific intensity is related to source temperature. The specific intensity of radiation from a black-body at temperature T can be written as

$$I_\nu(T) = \frac{2h\nu^3}{c^2}\frac{1}{exp(\frac{h\nu}{k_BT}) - 1}, \qquad [\mathrm{W\,m^{-2}\,Hz^{-1}\,sr^{-1}}] \tag{2.2}$$

where h is Planck's constant, k_B is the Boltzmann constant, c is the speed of light, ν is the observing frequency and T is the temperature. As can be seen, the specific intensity has a non-linear relation to the observing frequency and temperature. However, when the observing frequencies are low ($<$ a few GHz) and the source temperatures higher than a few K, the black-body radiation equation can be approximated as given in Eq.2.3.

$$I_\nu(T) = \frac{2\nu^2 k_B T}{c^2} = \frac{2k_BT}{\lambda^2} \tag{2.3}$$

This linear relation between source temperature and specific intensity, called the Rayleigh-Jeans approximation, allows the use of temperature units (K) as a proxy for the specific intensity.

A related quantity often used in astronomy is the flux density. Flux density is the specific intensity integrated over the observed solid angle, given in Eq.2.4

$$S_\nu = \int_{\Omega_s} I_\nu(\Omega) d\Omega, \qquad [\mathrm{W\,m^{-2}\,Hz^{-1}}] \tag{2.4}$$

where Ω_s is the observed solid angle. If the source has a constant specific intensity over the solid angle, Eq.2.4 simplifies to

$$S_\nu = I_\nu \Omega_s \tag{2.5}$$

To measure specific intensity the source solid angle has to be known. This condition is met when the telescope beam is smaller than the source size, and the source is said to be *resolved* and is generally considered as an *extended* source. When the celestial object under study fills a solid angle smaller than that of the telescope, the source solid angle is unknown and the source is said to be *unresolved.* In this case the flux density becomes the useful measurand. When radio telescopes observe compact, unresolved sources the preferred unit of measurements is the flux density. The unit of flux density in radio astronomy is the jansky (Jy), with 1 Jy = 10^{-26} W m^{-2} Hz^{-1}.

2.1.1 Measurements with a radio telescope

A radio telescope antenna receives celestial electromagnetic radiation and converts that into electrical signals. An isotropic antenna, by definition, picks up radiation from all 4π steradians of solid angle. A non-isotropic antenna picks up radiation from less than 4π steradians of solid angle, this leads to a definition of *directivity* for antennas. The unitless directivity D of an antenna is given by :

$$D = \frac{4\pi}{\Omega_A} \tag{2.6}$$

where Ω_A is solid angle of the antenna beam (Kraus & Marhefka, 2002). A related antenna parameter is the effective aperture A_e which can be viewed as the collecting area for an antenna. The effective aperture need not be the same as the physical area occupied by the antenna; as a matter of fact antennas such as dipoles used in low frequency radio astronomy occupy far less physical area than their effective apertures. The effective aperture is related to directivity as follows.

$$A_e = \frac{\lambda^2 D}{4\pi}, \qquad [\mathrm{m}^2] \tag{2.7}$$

For an unpolarised source of flux density S_ν, the power per unit bandwidth or the power spectral density (psd) measured by a single polarisation of a radio telescope is then given by:

$$P_t = \frac{1}{2} S_\nu A_e, \qquad [\mathrm{W\,Hz^{-1}}] \tag{2.8}$$

Using Eqs 2.5, 2.6 and 2.7 in Eq.2.8 and assuming that the source has uniform temperature, we obtain:

$$P_t = \frac{1}{2} I_\nu \lambda^2 \Big(\frac{\Omega_s}{\Omega_A}\Big) \tag{2.9}$$

When the source is resolved, the part of the source outside the antenna beam gets attenuated. Then the observed source solid angle becomes the antenna beam solid angle $\Omega_s = \Omega_A$ and the received power does not vary if the antenna beam solid angle reduces further, if the source is featureless and has no subcomponents.

$$P_{t,resolved} = \frac{1}{2} I_\nu \lambda^2 \tag{2.10}$$

Eq.2.10 shows that the power received by a radio telescope from a resolved source scales as λ^2. From Eq.2.3, the specific intensity for Rayleigh-Jeans approximation to black-body radiation has a scaling of $\frac{1}{\lambda^2}$. Therefore, the psd obtained becomes independent of the observing frequency and only a function of the source temperature.

$$\begin{aligned} P_{t,resolved} &= \frac{1}{2} \frac{2 k_B T}{\lambda^2} \lambda^2 \\ &= k_B\, T. \end{aligned} \tag{2.11}$$

It is interesting to draw parallels to a similar relation in electrical systems. The psd of the Johnson-Nyquist noise (Johnson, 1928; Nyquist, 1928b) measured at the ends of a resistor kept at a temperature T is given by Eq.2.12:

$$P_{JN} = k_B\, T, \qquad [\mathrm{W\,Hz^{-1}}] \tag{2.12}$$

where P is the power measured in a bandwidth of 1 Hz. A strong connection between Eqs.2.11 and 2.12 is obvious; in both equations the power is proportional to some temperature.

The relation between Eqs.2.11 and 2.12 is used to define *antenna temperature*. Consider an antenna measuring celestial radiation and yielding a power of P_A. We then replace the antenna with a resistor and heat/cool the resistor to a temperature T_A to obtain the exact same power as P_A. The temperature of the resistor that yields the same power as the antenna is defined as the antenna temperature. In general, antennas can have varying impedances as a function of frequency, the resulting mismatch losses need to be corrected for while calculating the antenna temperature.

The mismatch between an antenna and the transmission line connected to it is expressed by a voltage reflection coefficient Γ_a (Pozar, 2011). If the beam weighted sky temperature is given by T, the measured antenna temperature would be:

$$T_a = T(1 - |\Gamma_a|^2), \qquad \text{[K]} \tag{2.13}$$

From this, the beam weighted sky temperature can be recovered by measuring Γ_a and correcting for it, assuming the antenna to be lossless. It has to be mentioned here that the inclusion or exclusion of the mismatch into the antenna temperature definition often depends on the field of study (see for example Randa, 2008). However, in radio astronomy the term antenna temperature usually refers to the measurements that do not have a mismatch correction applied to them.

The concept of antenna temperature has important implications for observations, as the measurements from a radio telescope of extended sources can be directly calibrated into units of temperature by comparison with resistors of known physical temperature. If the nature of emission is thermal, this gives a measurement of source temperature; in other words the radio telescope acts as a remote thermometer! Indeed, using this technique the physical temperatures of resolved sources with black-body like emission such as the Sun and Moon have been measured since the dawn of radio astronomy (Dicke & Beringer, 1946). Such systems find use beyond radio astronomy and are often used as radiometers for remote sensing.

For global 21–cm observations, the signal of interest is an average over the sky and therefore can be treated as an extended source. As Eq.2.10 shows, the received power becomes independent of the antenna beam when the source fills a solid angle larger than that of the antenna. This leads to an important design choice for global 21–cm antennas - increasing the antenna collecting area does not fundamentally improve the sensitivity to the global 21–cm signal. Therefore, all else being equal, the chances of detection of the global 21–cm signal are the same for a simple dipole and a large sized dish antenna. However, the antenna beams of dish antennas have undesirable features for global 21–cm research, such as beams that vary with frequency (chromaticity). When antennas become too small, other factors such as the antenna resistive losses begin to dominate and affect detection. Therefore, specialised dipole or monopole antennas are the antennas of choice for global 21–cm experiments.

Also, as the global 21–cm signal is an extended source, the preferred unit of measurements is that of temperature (K). As this thesis is on global 21–cm measurements, units of Jy seldom appear for this reason.

2.2 Architecture of low frequency radio telescopes

The electrical signals of celestial origin are almost always noise-like; if a radio astronomer were to listen to it after amplitude demodulation, the celestial signals would sound noise-like with no intelligible information. Therefore radio telescopes are designed to extract useful information from the noise - such as the spatial, temporal or spectral nature of the noise.

Historically, radio telescopes have targeted higher and higher frequencies of observation, driven by the need for better angular resolution. Nonetheless, the emergence of low frequency science goals - especially cosmological ones - have motivated new experiments at lower frequencies in recent decades. The target frequency range for most of the modern low frequency telescopes is the metre wavelengths, frequencies of 10–300 MHz where the redshifted 21–cm signal from CD/EoR is expected.

While there are several designs for low frequency radio telescopes, each with their own design features suitable for the intended science goals, some of the common design features of their signal chains are outlined here. Importance is given to the signal flow through the analog and digital signal processing chains. The architectures of radio telescopes have evolved over decades, from adding interferometers with analog backends and stripcharts to modern systems with sophisticated supercomputers processing petabytes of data. A high level block diagram of a generic low frequency radio telescope with N number of antennas is given in Fig.2.1. The design shown is for continuum and spectral line studies where the spectral behaviour of the sky signals is of importance. Similar architectures have been implemented in the current generation of low frequency radio telescopes with 21–cm cosmology as one of the key science goals, such as MWA (Tingay et al., 2013; Wayth et al., 2018), LOFAR (van Haarlem et al., 2013), LWA (Hallinan et al., 2015), HERA (DeBoer et al., 2017) etc. In the case of uGMRT(Gupta et al., 2017), being a dish based interferometer, the architecture is different to aperture arrays. An overview of the sub-components of a radio telescope is given below.

2.2.1 Antennas

The sensing elements in radio telescopes are antennas that convert electromagnetic waves into electrical waves. The antenna could be a single antenna element such as a dipole/monopole or a dish or a phased array consisting of multiple simple antennas electronically combined into an effective single antenna using a beamformer. Modern radio telescopes observe large instantaneous bandwidths, and designing antennas to be sensitive across their entire operational range is often extremely challenging. For aperture arrays, the common strategies used to obtain sensitivities across large bandwidths are electrically small antennas and frequency independent antennas.

Electrically small antennas have dimensions that are small compared to their operating wavelengths. For example, LWA antennas can be considered electrically short at their operational wavelengths (Hicks et al., 2012). Frequency independent antennas have their dimensions specified by angles instead of lengths (Rumsey, 1957). Many frequency independent antennas therefore tend to have

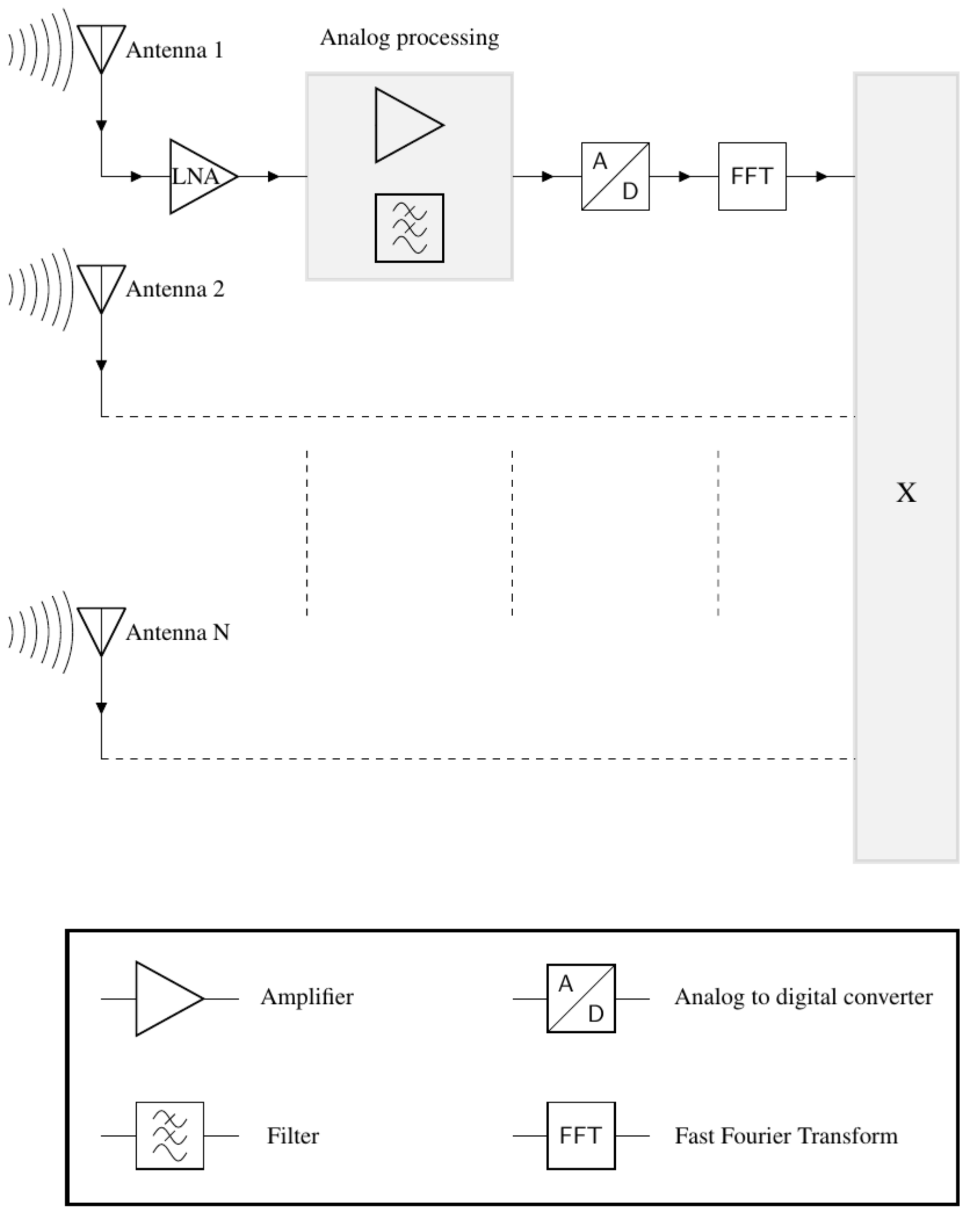

Antenna 1
Analog processing
LNA
A
D
FFT
Antenna 2
X
Antenna N
A
D
Amplifier
Analog to digital converter
Filter
FFT
Fast Fourier Transform

Figure 2.1: A generic, high level block diagram of a low frequency radio telescope with N number of antennas. For brevity, only one signal chain is elaborated. The signal from an antenna goes through some analog processing and subsequently gets digitised and Fourier transformed (F). The Fourier transformed data from multiple antennas are then correlated (X), giving visibilities.

an overall conical shape, such as conical log-spirals. A use case in astronomy is the SKA-low antenna design (de Lera Acedo et al., 2015), which is a log periodic dipole array (LPDA). It has to be noted that the distinction between electrically small antennas and frequency independent antennas is quite thin. Often, antenna designs combine aspects of both to improve performance. The bow-tie antenna designs used in MWA and LOFAR high band array (HBA) are examples where aspects of electrically small antennas and frequency independence are combined. For example, the MWA bow-tie design has a span of 74cm, which would be half wavelength at about 200 MHz. However, as the bow-tie is identical to a bi-conical structure, the antenna has a larger useful bandwidth and supports frequencies much lower than 200 MHz where it is electrically short.

2.2.2 Analog electronics

Low-frequency radio telescopes are largely *software telescopes*, i.e. most of the data processing is done in the digital domain. However analog electronics still play a major role in ensuring performance. The major analog subsystems used in radio telescopes are amplifiers, attenuators, filters and cables. It has to be mentioned that the specifics of analog electronics vary substantially from one telescope design to another.

The antenna is generally followed by an amplifier to amplify the electrical signals with very minimal excess noise added. These amplifiers are hence called low noise amplifiers (LNAs); their job is to minimise the impact of noisier components further down in the signal path on the overall system temperature. In modern radio telescopes the LNAs are typically placed close to each of the individual antenna elements without any intermediate cables, constituting the so-called active antennas. Often, the power to these active antennas is supplied though the same coaxial cable used for signal transport, employing *bias-tees*. Filters limit the signal into desired frequencies, called the passband. Filters can be of low-pass, band-pass, band-reject, or high-pass nature. Multiple filters can be cascaded to ensure adequate rejection of undesired signals or to control the levels of systematics.

The processed analog signals are transported over cables to digitisers. The typical transport medium of choice is coaxial cable, though systems using radio frequency over fibre (RFoF) are also rapidly gaining traction (Beresford et al., 2017; Perini et al., 2022). When RFoF is employed, radio signals are converted into intensity modulated optical signals and carried over optical fibres. At the receiving end, the optical signals are demodulated back into electrical signals. The main advantage of RFoF over coaxial cables is low loss over large distances.

2.2.3 Digitisation and Correlation

An analog to digital converter (ADC) performs sampling of the analog signals at a sampling rate, and converts them into a digital form suitable for processing with computers, field programmable gate arrays (FPGAs), graphical processing units (GPUs) etc. The sampling rate is typically at least twice that of the highest desired frequency in the passband, to satisfy the Nyquist-Shannon sampling criteria (Nyquist, 1928a; Shannon, 1949). For example, if the passband is 0-250 MHz the sampling rate is chosen to be 500 MHz or higher. Prior to digitisation, the frequencies higher than the highest frequency in the passband are filtered out using a bandpass filter. This filter prevents *aliasing* of the signals, and hence is called an anti-aliasing filter. The digital signals are then channelised into a set of frequency channels. This operation is typically performed with fast Fourier transforms (FFT), sometimes preceded by a polyphase filter bank (PFB) to improve channel to channel isolation and reduce spectral leakage (Price, 2021).

After channelisation, the resulting complex-valued spectral data from various antennas are correlated in a *correlator*, shown as block X in Fig.2.1. The correlator performs pairwise cross multiplication and accumulation of digitised and channelised data to yield *visibilities*. The visibilities formed by multiplication of channelised data from the same antenna constitute *autocorrelations*, while those formed between different antennas constitute *crosscorrelations*. When cross multiplying, data from one of the antennas in the pair are complex conjugated. Therefore, autocorrelations are real-valued while crosscorrelations are in general complex-valued.

The architecture presented here follows the FX design, where the channelisation (F) happens before correlation (X). A different order of operations is implemented in XF designs, where lag correlations are computed first and then Fourier transformed. However, computation of FFTs is inexpensive in modern computing environments and therefore FX is the de-facto standard for almost all radio telescopes. Apart from this, the sampled data are sometimes digitally beamformed to obtain directional beams. Such beams are then steered towards compact sources of interest such as pulsars to capture voltage data instead of visibilities for science cases such as pulsars and transient research. In any case, modern telescopes provide flexibility for such operations to be carried out in the digital domain and the rest of the signal chain remains more or less the same.

2.3 Low frequency experimental cosmology

Experimental cosmology often refers to measurements pertaining to that of the CMB, carried out at frequencies above $\sim$ 1GHz. Low frequency experimental cosmology is also associated with similar measurements, however the experiment designs are often different to that of the higher frequency ones.

For specialised science cases, such as global 21–cm research, the systems differ from the one given in Fig.2.1. The systems used for global 21–cm experiments are typically specialised radio-spectrometers that sense the incoming radiation and generate spectra. They typically use a single antenna that is custom designed along with additional electronics to perform precise calibration to remove systematic structures in the measured spectra. Sometimes, dedicated hardware calibrators such as noise sources are incorporated into the analog signal chain to obtain better calibration. Before discussing global 21–cm experiments, it is useful to look into the challenges specific to low-frequency cosmology.

2.3.1 Challenges

As mentioned in Chapter 1, foregrounds constitute a major challenge for low-frequency cosmology, especially for global 21–cm experiments. The foreground emissions are noise-like in nature, therefore they are generally the dominant noise source in the data. This necessitates long integration times to obtain sensitivities required to find any cosmological signal. Ironically, strong foregrounds simplifies instrument design to some extent. The fact that the foregrounds - which are external to the instrument - are strong implies that the noise in measured data is dominated by foregrounds. Therefore signal to noise ratios (SNR) are determined largely by the foregrounds and reducing the noise temperature of the receivers does not provide a substantial improvement in SNR. This enables the use of room temperature LNAs in the receiver front end electronics.

Foregrounds are expected to have smooth spectral features, therefore instrument designs have to ensure that the spectral nature of foregrounds is not modified by the instrument and confused with the 21–cm signal. Two common modifications to foregrounds introduced by a radiometer are multiplicative bandpass and mode-coupling. Across a wide bandwidth, the receiving electronics' bandpass can have spectral variations that multiplicatively corrupt the signals. To remove the bandpass, flat spectrum noise sources can be switched in and out - this scheme is conceptually similar to Dicke switching. The antenna beams can vary with frequency (chromaticity) and point to different sky regions at different frequencies. As the radio foregrounds have brightness temperature variations, this effectively couples structures in the sky into the frequency domain. This is often called mode-coupling and results in time-varying structures in the measured spectra that are hard to remove. A potential solution to this problem is to use achromatic antenna designs that have very low beam variations with frequency.

Apart from modifying the foregrounds, the instrument adds excess noise temperature to the measured spectra. Most of this noise can be attributed to the first amplifier in the receiver signal chain. To reduce this noise contribution, LNAs are employed as the first amplifiers. At radio frequencies, inexpensive un-cooled LNAs that achieve noise temperatures less than 100K are available; however the spectral structure of the added noise may not always be flat or smooth. Besides, the contribution of an LNA to the net system temperature depends on the antenna impedance seen at its input (a parameterisation of this can be found in Haus et al. (1960)). An antenna impedance that has non-smooth variations with frequency can result in an LNA noise temperature spectrum with non-smooth features. Therefore, the LNA and antenna designs often go hand-in-hand to ensure an overall spectrally-smooth response.

Metre wavelength radio astronomy carried out on Earth has several additional challenges, some of which are distinct from observations at higher frequencies. Perhaps the most important one is the presence of terrestrial radio frequency interference (RFI) due to high usage of the low frequency radio spectrum for licensed and unlicensed transmissions. If the RFI signals are confined in time or frequency or both, they may be flagged via algorithms designed to do so (Offringa et al., 2010). However, if the RFI signals are strong enough to drive the receiver electronics into their non-linear regimes, they can cause non-linear intermodulation products. This necessitates frequency domain filtering to suppress RFI and/or moving the receivers to a radio quiet zone to avoid interfering signals altogether.

Yet another challenge with ground-based metre wavelength cosmology is the ionosphere, which acts as a phase screen that alters the incoming wavefront. The effects of the ionosphere increase as the observing frequencies are reduced, and for frequencies below 30 MHz, the ionosphere becomes largely opaque to celestial radiation. The ionosphere also has temporal and spatial variations that complicate 21–cm power spectrum observations (Mevius et al., 2016; Jordan et al., 2017). Some of the effects of the ionosphere on global 21–cm experiments have been investigated in Vedantham et al. (2014) and Sokolowski et al. (2015).

2.3.2 Global 21–cm experiments

2.3.2.1 Single antenna experiments

As the global signal is an all-sky measurement, a single well-calibrated radio telescope may be used to measure it. Conceptually, this method is not vastly different from experiments that aim at detection of spectral distortions in the CMB at frequencies of a few 100s of GHz, such as COBE-FIRAS (Mather et al., 1994). Therefore, a single antenna connected to a well calibrated receiver that collects wideband spectra over time - a radio spectrometer - is all that is required to measure this signal. By averaging the collected spectra, radiometric noise in the measured spectra can be reduced, enabling a detection of the signal. However, in practice, several other effects such and antenna beam chromaticity, receiver noise systematics etc. have to be taken into account. Chapter 3 outlines one such single antenna experiment targeting the global 21–cm signature. Ways to mitigate some of the systematics associated with such experiments are also detailed.

2.3.2.2 Interferometric experiments

Motivated partly by the complexities involved in the single antenna method, detection of the 21–cm global signal using interferometers has been proposed as an alternative (Presley et al., 2015). A major advantage of interferometers when compared to single element instruments is that there are multiple antennas with their own LNAs and associated signal chains. Since the noise contributions from the signal chains are expected to be uncorrelated, instrumental noise can be reduced by cross-correlating voltages from the antennas. The two major techniques coming under interferometric experiments targeting the global 21–cm signal are explained below.

Using conventional interferometers

Due to the long baselines used, conventional low frequency interferometers that make use of dishes (GMRT) or phased array tiles (MWA, LOFAR) are insensitive to a monopole signal in the sky. However, a source such as the Moon can locally block the global signal. Observations of a sky patch with the source blocking a portion of the sky and without this blocking can be subtracted to yield an image which would contain the global signal within the blocked region. In other words, the Moon increases the spatial coherence of a global signal by blocking it. This method has been adopted in Vedantham et al. (2015) and McKinley et al. (2018).

Using short-spacing interferometers

At sufficiently short baselines, an interferometer has a response to an all sky signal. The origin of this response has been debated. Presley et al. (2015) attribute the response to finite beams of the antennas used. However, using the coherence equation, Singh et al. (2015) show that even isotropic antennas can give rise to a response at short baselines. It is also worthwhile noting that a spherical harmonic expansion of the coherence equation, given in Vedantham et al. (2015) provides the expected *sinc* response at short baselines. All these works, however, ignore effects of mutual coupling between antennas. A more theoretical investigation into this effect by Venumadhav et al. (2016) predicts that mutual coupling between antenna elements is crucial to this response. The same work also predicts that the receiver noise coupled across antennas has a spectral shape anti-correlated with the sky response. Despite these predictions, experimental investigations into short-spacing interferometry have been sparse. Notable exceptions are the ZEBRA experiment (Mahesh et al., 2015) using a semi-transparent resistive screen to increase the coherence as well as the ASSASSIN experiment (McKinley et al., 2020) employing EDA-2 (Wayth et al., 2022). Chapter 4 presents a dedicated broadband interferometer experiment to validate the concept of short-spacing interferometry.

Chapter 3

SARAS 3 - A precision experiment to probe the global 21–cm signal

> *Science progresses best when observations force us to alter our preconceptions.*
>
> \- Vera Rubin

This chapter is a reproduction of Jishnu Nambissan, T.; Ravi Subrahmanyan; R. Somashekar; N. Udaya Shankar; Saurabh Singh; A. Raghunathan; B. S. Girish; K. S. Srivani; Mayuri Sathyanarayana Rao, *SARAS 3 CD/EoR radiometer: design and performance of the receiver*, 2021, Experimental Astronomy, https://doi.org/10.1007/s10686-020-09697-2 . Minor alterations have been made to the original material for consistency within this thesis. The reader may encounter some repetition of material in the introductory sections.

3.1 Abstract

SARAS is an ongoing experiment aiming to detect the redshifted global 21-cm signal expected from the Cosmic Dawn and the Epoch of Reionisation (CD/EoR). Standard cosmological models predict the signal to be present in the redshift range z ∼6–35, corresponding to a frequency range 40–200 MHz, as a spectral distortion of amplitude 20–200 mK in the 3 K cosmic microwave background. Since the signal might span multiple octaves in frequency, and this frequency range is dominated by strong terrestrial Radio Frequency Interference (RFI) and astrophysical foregrounds of Galactic and Extragalactic origin that are several orders of magnitude greater in brightness temperature, design of a radiometer for measurement of this faint signal is a challenging task. It is critical that the instrumental systematics do not result in additive or multiplicative confusing spectral structures in the measured sky spectrum and thus preclude detection of the weak 21-cm signal. Here we present the system design of the SARAS 3 version of the receiver. New features in the evolved design include Dicke switching, double differencing and optical isolation for improved accuracy in calibration and rejection of additive and multiplicative systematics. We derive and present the measurement equations for the SARAS 3 receiver configuration and calibration scheme, and provide results of laboratory tests performed using various precision terminations that qualify the performance of the radiometer receiver for the science goal.

3.2 Introduction

Following cosmological recombination of the primordial hydrogen and helium, the Dark Ages is expected to have ended with the formation of the first stars in the first ultra-faint galaxies, which lit up the Universe. During this transformational epoch spanning redshifts $z \sim 6$–35, commonly referred to as the Cosmic Dawn (CD) and the Epoch of Reionisation (EoR), the baryons in the Universe transitioned from being mostly neutral during the Dark Ages to being almost completely ionised by the end of EoR (Sethi, 2005; Shaver et al., 1999). However, the nature of the sources driving this transition, the timing of events, and the physical—light-matter and gravitational hydrodynamic—processes that govern the evolving gas transition are poorly understood. A key reason for the uncertainty is the lack of observational constraints.

In this context, it has been recognised that the redshifted 21-cm signal from neutral hydrogen at those epochs could act as a direct probe to trace the evolving gas properties of the Intergalactic Medium (IGM) in the DA/CD/EoR. Specifically, the sky-averaged or global component of this 21-cm signal has been shown to be an extremely powerful tool (Cohen et al., 2017; Pritchard & Loeb, 2010); therefore, several experiments are currently underway to detect this signal. SARAS (Shaped Antenna measurement of the background RAdio Spectrum) is a spectral radiometer experiment aiming to measure the spectrum of the global 21–cm signal. Apart from SARAS, which is the subject of this chapter, other ongoing experiments include EDGES (Bowman et al., 2018), SCI-HI (Voytek et al., 2014), BIGHORNS (Sokolowski et al., 2015), PRIZM (Philip et al., 2019), LEDA (Price et al., 2018), ASSASSIN (McKinley et al., 2020), REACH (de Lera Acedo, 2019), MIST (MIST, 2020), and HIGH-z (HIGH-Z, 2020) .

The first version, SARAS 1 (Patra et al., 2013), operated in the band 87.5–175 MHz and employed a frequency-independent fat-dipole antenna above ferrite-tile absorbers as the sensor of the sky radiation. The signal received at the antenna terminal was split as the first stage of analog signal processing, propagated through independent receiver chains, and finally digitised and correlated using a 1024-channel digital cross-correlation spectrometer to measure the sky spectrum. Adoption of a correlation radiometer concept, usage of optical fibre links and phase switching were aimed at suppression of internal systematics. SARAS 1 provided an improved calibration for the 150 MHz all-sky map of Landecker & Wielebinski (Patra et al., 2015); however, RFI, non-smooth spectral behaviour of the balun, limited absorbtivity of the ferrite-tile ground and spectral confusion arising from multi-path interference within the signal path limited the sensitivity.

The second version SARAS 2 employed a 8192 channel cross correlation spectrometer along with an electrically short spherical monopole antenna (Singh et al., 2018b). Evolution from dipole to monopole antenna allowed the elimination of the balun and also substantial reduction in signal transmission length between the antenna and receiver. As a consequence, confusion from multi-path interference within the signal path was reduced and hence spectral smoothness in the instrument response was improved. Additionally, the use of a monopole antenna instead of a dipole avoided beam chromaticity that arises from reflections off the ground plane beneath the antenna, thereby reducing spectral confusion due to mode coupling of spatial structures in the sky foreground into frequency structure. Based on 63 hr of data collected at the Timbaktu Collective in Southern India, SARAS 2 placed constraints on the global 21–cm signal, disfavouring models with low X-ray heating and rapid reionisation (Singh et al., 2017, 2018a). The poor efficiency of the short monopole, which worsened at low frequencies, limited the scientifically useful band of SARAS 2 to 110–200 MHz; additionally, the sensitivity suffered due to the excess system temperature of the adopted correlation receiver configuration.

A detection of the 21-cm absorption during cosmic dawn has been claimed using the EDGES low-band system (Bowman et al., 2018). The recovered EDGES signal has an amplitude of about 500 mK at a frequency of 78 MHz, corresponding to a redshift of 17, and has a flattened profile. The absorption depth is greater than the maximum allowed in standard ΛCDM cosmology, which is about 220 mK. The unexpectedly deep absorption profile has been interpreted by Barkana (2018) as evidence for new physics: millicharge in a small fraction of Dark Matter (DM), which would then scatter off and act as an additional cooling mechanism for the baryons. In this case, during cosmic dawn, strong Wouthuysen-Field coupling (Wouthuysen, 1952) of the 21-cm spin to low baryon kinetic temperatures results in a deep absorption. However, it has been pointed out in Barkana et al. (2018) that this hypothesis of dark-matter baryon interaction may violate constraints from other astrophysics data.

The EDGES absorption may be compatible with expected baryon cooling in standard cosmology if there is an additional radiation background during cosmic dawn, apart from the CMB (Feng & Holder, 2018). Thus, EDGES has been viewed as providing support for the interpretation of the ARCADE-2 measurements (Fixsen et al., 2011) that previously suggested an unaccounted for radio background. However, the analysis of the ARCADE-2 data has itself been questioned (Subrahmanyan & Cowsik, 2013).

Together with the tension between theory and the 21-cm profile suggested by EDGES for cosmic dawn, reanalysis of the publicly available EDGES data have raised the possibility that the claimed detection has an instrumental origin. Independent modeling of the EDGES spectrum by Hills et al. (2018) demonstrated that the fitting procedure is not unique and that an uncalibrated systematic in the form of a sinusoidal spectral structure with period 12.5 MHz might be present in the data. Another independent analysis by Singh & Subrahmanyan (2019) that uses maximally smooth functions (Sathyanarayana Rao et al., 2017b) to model foregrounds indicates that allowing for such a systematic is preferred by Bayesian information criterion (Schwarz, 1978). Joint modeling that takes into account potential errors in applying beam chromaticity corrections and a more detailed noise model has shown that Bayesian evidence prefers a model for the EDGES spectrum with this added complexity, while favouring absorption profiles of depth < 209 mK that are within standard predictions (Sims & Pober, 2019) . Inhomogeneous ground beneath the EDGES antenna has been suggested as a potential instrumental artefact that might be an alternate non-astrophysical explanation for the EDGES residual (Bradley et al., 2019). All this constitutes motivation for an experiment with a different design and calibration philosophy to verify the claimed detection.

In this chapter, we present an improved version of SARAS, referred to as SARAS 3, which has an architecture different from earlier versions. First, introducing a double differencing strategy improves calibration and cancellation of unwanted additive features in the spectra. Second, in order to reduce the additive noise from the splitter of the correlation spectrometer, splitting of the signal received by the antenna is carried out post signal amplification. The receiver system is designed to operate over frequencies 40–230 MHz, and intended to be used with scaled conical monopole antennas that operate in octave bands and cover the range in staggered bands. At any instant, the receiver would be paired with a conical monopole antenna designed to operate over an octave band located within the range of the receiver, thus providing science data covering an octave band. The system would be deployed with different monocones at different times, thus sweeping across the multi-octave band. Independent analysis of data in different octave bands, as well as joint analysis of the combined data, is expected to provide constraints on the baryonic evolution during CD/EoR.

In the following sections, we describe the various considerations that have gone into the design of SARAS 3, emphasising improvements in design and performance. In Section 3.3, we discuss the motivation for the SARAS 3 version of the receiver and in Section 3.4, we present an overview of the system. In Section 3.5, we derive the measurement equation. The implementation of the analog receiver is described in Section 3.6. Section 3.7 presents the measurement sensitivity wherein we also demonstrate that SARAS 3 is capable of detection of EoR signals with amplitudes and spectral complexity predicted in standard models. Laboratory tests that qualify receiver performance are discussed in Section 3.8 to demonstrate adequate control of internal systematics.

3.3 Motivation for the SARAS 3 receiver

SARAS 2, the second version of SARAS, was a cross correlation differential spectrometer (Singh et al., 2018b). The spectra measured were differential measurements made between the antenna and an internal reference load. The antenna was a spherically shaped monopole antenna of height 33 cm over a ground disc of radius 43.5 cm. The antenna was designed to be electrically short at all frequencies below 200 MHz thus capable of observing over the entire CD/EoR band with a fairly frequency independent beam pattern; however, the downside was the substantial loss in radiative and reflection efficiencies at longer wavelengths.

The first element in the front-end electronics was a four-port cross over switch with the antenna connected to one input port and a calibration noise source connected to the second port through a series of attenuators, which also served as the internal reference load. The outputs of the cross-over switch were connected to a 180° hybrid, which produced sum and differences of the input antenna and reference signals. These sum and difference signals were then amplified and transmitted via a pair of 100 m radio frequency over fibre (RFoF) links to the back-end electronics. The front-end electronics were in a shielded enclosure underground beneath the antenna. At the back-end receiver, the pair of optical signals were demodulated to electrical signals and then processed by two identical arms of signal processing electronics consisting of amplifiers and band limiting filters. The signals were band-limited to 40–230 MHz where the CD/EoR signal is predicted to be. The processed signals were sampled, digitised and Fourier transformed using a 8192-point Fast Fourier Transform (FFT) algorithm on a Virtex-6 FPGA, then cross correlated to yield the correlation spectra. Bandpass as well as absolute calibration of the spectra was implemented by switching on and off the noise source, which introduced a step of calibrated noise temperature in the measured spectra. For details of the SARAS 2 system and its performance, we refer the reader to (Singh et al., 2018b).

The SARAS design philosophy has been to purpose design all multiplicative gains and additive systematics, which survive calibration with the noise injection, to be spectrally smooth. We define smoothness as given by the class of polynomials having no zero crossings in second and higher-order derivatives; such polynomials are called maximally smooth (MS) polynomials (Sathyanarayana Rao et al., 2017b). SARAS 2 included a 180-degree hybrid in the front-end receiver that served to split the signal from the antenna between the two arms of the correlation spectrometer. This hybrid impressed frequency-dependent loss as well as additives on to the sky signal and, therefore, the spectral smoothness of the receiver was limited to be somewhat less than an octave. Additionally, the loss in the hybrid, which was ahead of any amplifiers in the signal path, resulted in loss of sensitivity. As a consequence of the antenna and receiver design limitations, SARAS 2 was limited to EoR science in the band 115–185 MHz.

The SARAS 3 receiver, described in detail below, is designed to mitigate these limitations, taking constructive lessons from the SARAS 2 experience. In order to better cancel additives and systematics, the signal from the antenna is directly coupled to a low-noise amplifier through a Dicke switch, which alternately connects the receiver between the antenna and an internal reference load. Spectral structure arising from multi-path interference within the receiver is reduced by reducing the total electrical length in the front-end by connecting the antenna terminal directly to a miniature RF switch and high-gain modular MMIC amplifier that are followed immediately by an RFoF modulator. Phase switching to cancel additives in the analog receiver chain and digital samplers is achieved, without loss of sensitivity and without multi-path interference, using a 180-degree hybrid following the optical RFoF demodulator. Thus SARAS 3 has both Dicke switching as well as phase switching—a double differencing—to more effectively cancel the additives in the measured spectra. Additionally, the digital receiver for SARAS 3 has been improved to perform a 16384-point FFT instead of the earlier 8192-point, so as to improve detection and rejection of data corrupted by narrow-band RFI. The FFT implementation architecture was also changed in firmware to be 2 × 8k instead of 8 × 2k, so as to avoid errors at 2k boundaries in the 8k spectrum.

3.4 System overview

SARAS 3 is a double differencing radiometer, with Dicke-switching between antenna and reference load plus phase switching. The receiver is designed to be split between a front-end unit that is located immediately below the ground plane of a shaped monopole antenna, and connects via optic fiber to a remote back-end some distance away, typically about 150 m.

This antenna type has been selected so that the antenna terminals are at ground level and there is no significant length of transmission line between the antenna and receiver. The wideband antenna will almost certainly not have an excellent impedance match across the entire band of operation, which results in internal reflection of the receiver temperature at the antenna terminal. Avoiding a transmission line between the antenna terminals and receiver avoids delays in the signal path and hence avoids frequency dependent structures arising from multi-path interference. Second, monopoles are inherently unbalanced at their terminals and hence do not require a balanced-to-unbalanced (balun) transformer to connect to a receiver with a coaxial input terminal. Third, monopoles with finite conductive ground planes inherently have nulls towards the horizon, which aids in suppression of unwanted terrestrial interference that mostly arrives from low elevations. Lastly, monopoles may be designed to be electrically short and thus have spectrally smooth reflection and radiation efficiency characteristics, which preserve the spectrally smooth nature of the foregrounds as they couple to antenna temperature. Short monopoles also have achromatic beams, which prevent mode coupling of foreground spatial features into the frequency spectrum.

The signal flow in SARAS 3 is depicted in Fig. 3.1. The antenna design used in conjunction with the SARAS 3 receiver is a monopole cone-disk floating on water (Raghunathan et al., 2020). A Dicke switch selects between the antenna and a reference termination. This reference is a flat-spectrum noise source that is connected to the Dicke switch via attenuations, which ensure that the reference presents a constant impedance independent of whether the calibration noise source is on or off. Thus the Dicke switch sequentially presents, to the receiver, the antenna temperature, the noise temperature of the reference with noise source off and the noise temperature of the reference with calibration noise source on; we refer to these three switch states and corresponding noise temperatures as OBS, REF and CAL respectively. Differencing the measurement data with antenna connected with that when the reference is connected provides a differential measurement, which cancels most of the unwanted additives in the receiver chain. Differencing the measurement data with calibration noise source on with that when it is off provides a bandpass calibration.

The selected signal voltage—from either the antenna, reference termination or calibration noise source—is amplified and directly intensity modulates a semiconductor laser source, thus providing transmission of the RF signal via single-mode optical fiber from the front-end to the remote part of the receiver electronics. Conversion from electrical to optical followed by optical to electrical provides excellent galvanic and reverse isolation of the front-end electronics from that at the back end. This optical isolation is essential since the Dicke switch presents different impedances to the subsequent receiver electronics when it switches between the antenna and reference termination. Consequently, any standing waves

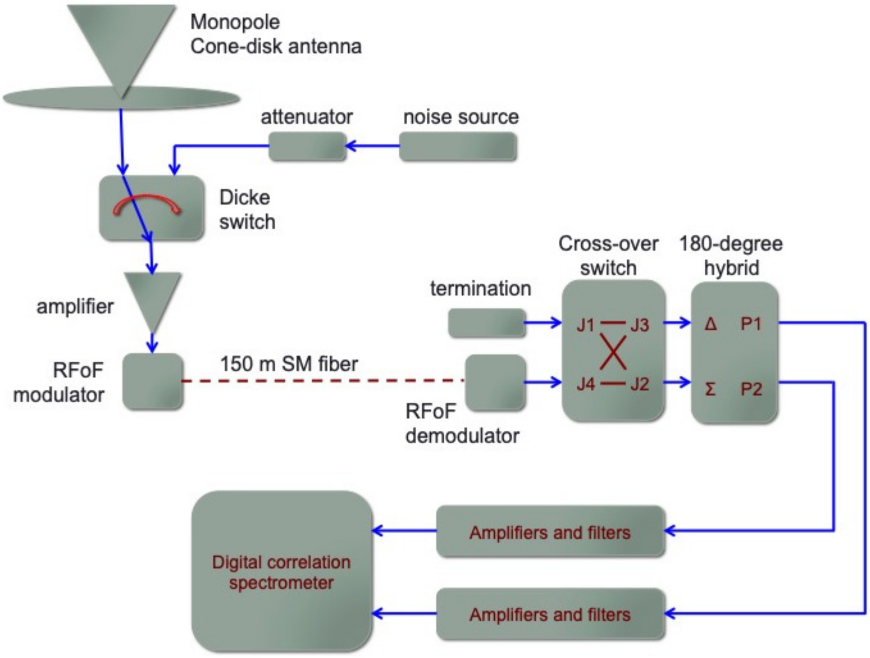

Monopole
Cone-disk antenna
attenuator
noise source
Dicke
switch
amplifier
RFoF
modulator
150 m SM fiber
termination
RFoF
demodulator
Cross-over
switch
180-degree
hybrid
J1
J3
J4
J2
Δ
P1
Σ
P2
Digital correlation
spectrometer
Amplifiers and filters
Amplifiers and filters

Figure 3.1: Schematic of the SARAS 3 receiver architecture.

in the receiver chain that reflect from the antenna/reference terminations and result in multi-path interference would change with the position of the Dicke switch, thus resulting in calibration errors. Standing waves are inevitable owing to impedance mismatches along the signal path, and long transmission lines between the front-end and remote electronics would result in complex and high order spectral structure. Amplifiers do provide isolation; however, they are limited to the difference between their forward gain and reverse isolation. Optical modulators/demodulators provide excellent isolation, prevent standing waves existing across the back end and front end electronics. Thus spectral structure in SARAS 3 is limited to the path length between the antenna terminal, which is coincident with the receiver input, and the optical modulator. All standing waves downstream of the optical isolation are accurately calibrated out as part of the bandpass calibration and the Dicke switching.

Digital control signals that operate the Dicke switch and calibration noise source, located in the front-end electronics, are transmitted from the back-end electronics to the front-end receiver over multi-mode optical fibres. The switch, calibrated noise source, low-noise amplifiers, RF optical transmitters, optical receivers for the digital control signals and the control and monitor circuits of the front-end receiver unit are all powered by a set of Li-Ion battery banks. The entire front-end receiver is housed in an environmentally protected and electromagnetically shielded enclosure, which is attached below the ground plane of the monopole antenna terminals. The electromagnetic shielding is vital to prevent coupling and feedback of the signal from the high-gain analog electronics beneath the antenna back to the antenna, which would result in calibration errors. This unit is also waterproofed, as the receiver is co-located with an antenna that floats in water.

The remote electronics consist of an analog signal-conditioning unit (ASU) and a digital correlation spectrometer. The ASU implements phase switching to cancel common-mode additives in this part of the electronics, including the samplers; the ASU and digital spectrometer together operate as a correlation receiver. To achieve this, the RFoF optical signal is first converted back to electrical in an optical demodulator, and the electrical signal goes to a cross-over switch. The other input of the cross over switch has a matched termination. It may be noted here that the unwanted constant additive from this termination is about 6 mK equivalent antenna temperature, owing to the 44 dB gain between the antenna and the hybrid; this small additive is nevertheless canceled in the Dicke switching. The cross over switch alternately connects the antenna signal to the sum (Σ) and difference (Δ) ports of a 180° hybrid. Thus the hybrid alternately provides a pair of in-phase or a pair of out-of-phase electrical voltage signals at its two output ports, depending on the position of the cross over switch. The pair of electrical signals are then processed in separate receiver chains: filtered using pairs of low pass and high pass filters to band limit the signal to a 40–230 MHz range and amplified to levels appropriate for sampling by the analog to digital converters (ADCs) of the digital spectrometer. The gains within the ASU are optimally distributed so that all the amplifiers operate in their linear regimes with intermodulation contributing less than a mK equivalent distortion. Power levels are set so that the samplers contribute less than about 1% additional noise power as a result of quantisation, while leaving sufficient head room within the

Table 3.1: SARAS 3 system states.

State	Noise source	Dicke switch	Cross-over switch
OBS00	OFF	0	0
OBS11	OFF	0	1
CAL00	OFF	1	0
CAL01	OFF	1	1
CAL10	ON	1	0
CAL11	ON	1	1

sampler full scale for RFI.

The SARAS 3 digital back-end is an FX correlator based on Virtex 6 field programmable gate array (FPGA) (Girish et al., 2020). High-speed 10-bit ADCs sample and digitise the pair of analog signals from the ASU; the ADCs operate at a sampling rate of 500 MS/s to provide an analog bandwidth of 250 MHz. The samples are first apodised using a four-term Blackman-Nuttall window and then the digital receiver processes blocks of 16k samples in pipelined FFTs. As the FFT has Hermitian symmetry and the noise equivalent bandwidth of the window is 1.98, the total number of independent channels in the computed spectrum across 250 MHz band is 4096, with a resolution bandwidth of 61 kHz. The channel data corresponding to the two signals are integrated separately to provide auto-correlation or power spectra; the channel data are also complex multiplied and integrated to provide complex cross-correlation spectra. The real auto- and complex cross-correlation spectra are transferred to a laptop acquisition computer through high speed ethernet and stored as a MIRIAD-format multi-channel visibility data set.

In the complex cross correlation spectrum, the signal amplitude flips sign when the cross over switch alternates. When the switch feeds the antenna signal to the Σ port, the sky signal appears with positive sign and when the switch feeds to the Δ port the sign of the sky signal flips to negative. However, any common mode additives and coupling between the pair of channels within the ASU electronics appear in the cross spectra with constant complex correlation; therefore, these unwanted systematics are canceled on differencing the spectra obtained in the two positions of the cross over switch.

SARAS 3 cycles through six states. The front end cycles between three states, one in which the antenna is connected to the receiver, second where the reference is connected to the receiver, and third when the calibration noise source is on and this noise power flows via the reference attenuation to the receiver. In each of these states, the cross over switch in the ASU is toggled through its two states. These six states are given in Table 3.1. In each state, 16 spectral measurements are recorded, each with integration time of 67.11 ms. Including overheads associated with delays introduced to account for switch settling times and for reading the data into the acquisition computer, the total time for completing a cycle of 6 states is 8.23 s.

Clocks for the ADCs and the FPGA are derived from a Rubidium atomic standard. This ensures low clock jitter and sufficient frequency stability so that, given the gradients in filter passbands, frequency drifts within the bandpass calibration cycle time would only result in spectral structure well below 1 mK (Girish et al., 2020).

The electrical switching signals within the digital spectrometer are potentially a major source of radio frequency interference, if they enter the signal path radiatively via the antenna. Coupling of self-generated RFI into the analog electronics chains in the ASU is expected to largely cancel in the double differencing. For this reason, the digital receiver is in a high quality shielded enclosure, with welded walls, absorber lining, and carefully designed enclosure door. Fans are provided to reject internal heats via RF filtered vents. RF, optical and DC lines are taken through a panel with filtering. DC power to the digital correlator and ASU is provided from a battery pack, with RFI blocking filters inserted in these power lines. Additionally, the antenna is located at a distance greater than 100 m from the digital correlator. We have made field measurements of the RFI leakage from the digital spectrometer and established that the shielding and separation together suppress the RFI contamination of antenna temperature to below 1 mK (Girish et al., 2020).

Since readers may find some similarities between SARAS 3 and EDGES, we provide a brief comparison of the SARAS 3 receiver system with that of EDGES, specifically EDGES-2. EDGES implements a 3-state calibration, with Dicke switching between the antenna and a reference load, which may be switched to a high noise temperature state for calibration. SARAS 3 has a similar Dicke switch between antenna and reference, with a reference that switches also to a higher noise temperature state; however, SARAS 3 implements double differencing rather than single differencing. EDGES uses a single ADC in auto-correlation mode, with out-of-band noise injection to maintain proper operation of the ADC, while SARAS 3 uses a cross-correlation scheme. In SARAS 3, apart from the Dicke switch, the signal is split using a cross-over switch and 180-degree hybrid, and the pair of signals is processed in parallel chains of a correlation spectrometer. This cross-correlation and double differencing architecture provides additional switching states to remove additive systematics that may be present in the spectra, by subtracting two cross-correlated spectra acquired with identical input powers to the ADCs but with a 180° phase shift. EDGES also includes a mechanism for in-situ measurement of the antenna reflection efficiency (using a vector network analyser or VNA) and depends on precision measurement and correction of the data for its complex antenna reflection coefficient, noise wave parameters, and transmission line. However, SARAS 3 evades in-situ calibration and correction for these by careful design that uses maximally smooth functions to evaluate the antenna and receiver responses. Also, EDGES and SARAS 3 differ in the types of antennas used; EDGES uses a dipole and balun design, while SARAS 3 receiver is designed to be used with unbalanced antennas. Another major design difference between EDGES and SARAS 3 is the use of RFoF in SARAS 3 to transport signals from the antenna base electronics to the back-end electronics. This introduces galvanic isolation between the two units that are separated by about 100 m, thereby mitigating effects due to ground loops as well as mitigating deleterious effects caused by common mode currents in co-axial cables. However, the use of RFoF is not without limitations, as they have a limited dynamic range. Besides, optical fibre systems demand much more care during deployments and can suffer from reflections if the connectors are not clean. These issues are taken care of in SARAS 3 with proper maintenance of the optical fibre and cleaning the fibres during deployments. EDGES is also designed to operate in a fixed deployment configuration with data collection occurring for long duration, while SARAS 3 is designed as a self contained system (similar to SARAS 2) that can be transported to radio quiet zones with deployment campaigns spanning a few days to a few weeks. While advantageous, in the sense of flexibility in selecting radio quiet zones, this also sets constraints on the power, mass and volume of the SARAS 3 system.

3.4.1 Calibration considerations

The basic calibration of a spectrometer involves bandpass calibration and absolute calibration, which correct for the variation in receiver gain over frequency and provide an absolute temperature scale for the measured spectra. The spectral power available at the antenna terminal is multiplied by the spectral gain of the entire receiver chain of the spectrometer, impressing a multiplicative band shape on the antenna temperature. Bandpass calibration removes this instrumental bandpass response. The spectral data recorded by a spectrometer is usually in arbitrary units that is determined by the net electronics gain of the receiver chain and also the arithmetic in the digital signal processing. The process of absolute calibration sets a scale to the counts so that the measurement is converted to be in units of Kelvins of antenna temperature. Both the bandpass and also absolute gain of the receiver may be time varying, usually because the physical temperature of the amplifiers and other components may change over time; the SARAS 3 receiver is not actively temperature stabilised.

In SARAS 3, calibration data are derived using the internal calibration noise source, which is connected to the Dicke switch via fixed attenuators. In the CAL00 and CAL01 states the noise source is off and the switch is connected to the 50 Ω reference termination that is at ambient temperature, which we refer to as T_{REF}. The recorded spectral powers, in arbitrary units, are termed P_{CAL00} for the state when the cross over switch is in position 0, and P_{CAL01} when the cross over switch is in position 1. Switching on the noise source injects noise of equivalent temperature T_{CAL} into the receiver and the corresponding spectral powers recorded are denoted by P_{CAL10} for the case when the cross over switch is in position 0 and P_{CAL11} for when the cross over switch is in position 1. We denote the spectral powers recorded when the Dicke switch connects the antenna to the receiver by P_{OBS00} and P_{OBS11} respectively for the cases where the cross over switch is in position 0 and 1. We may write

$$\begin{aligned}
P_{OBS00} &= -|G|^2(T_A + T_N) + P_{cor}, & (3.1)\\
P_{OBS11} &= \quad |G|^2(T_A + T_N) + P_{cor}, & (3.2)\\
P_{CAL00} &= -|G|^2(T_{REF} + T_N) + P_{cor}, & (3.3)\\
P_{CAL01} &= \quad |G|^2(T_{REF} + T_N) + P_{cor}, & (3.4)\\
P_{CAL10} &= -|G|^2(T_{CAL} + T_N) + P_{cor},\ and & (3.5)\\
P_{CAL11} &= \quad |G|^2(T_{CAL} + T_N) + P_{cor}, & (3.6)
\end{aligned}$$

where T_A is the antenna temperature, T_N is the receiver noise added by the analog electronics, $|G|^2$ is the power gain of the receiver and P_{cor} is the unwanted spectral power added by the samplers of the digital receiver.

P_{cor} is not expected to change either in power or spectral characteristics between system states; therefore, differencing the powers recorded in the two crossover switch positions would cancel it:

$$\begin{aligned}
P_{OBS} &= P_{OBS11} - P_{OBS00},\\
&= 2|G|^2(T_A + T_N). & (3.7)\\
P_{REF} &= P_{CAL01} - P_{CAL00},\\
&= 2|G|^2(T_{REF} + T_N). & (3.8)\\
P_{CAL} &= P_{CAL11} - P_{CAL10},\\
&= 2|G|^2(T_{CAL} + T_N). & (3.9)
\end{aligned}$$

A second differencing of the two spectra obtained in Eqs. 3.9 and 3.8 yields a measure of the receiver bandpass:

$$\begin{aligned} P_{CAL} - P_{REF} &= 2|G|^2(T_{CAL} + T_N) - 2|G|^2(T_{REF} + T_N), \\ &= 2|G|^2(T_{CAL} - T_{REF}). \end{aligned} \tag{3.10}$$

The factor $(T_{CAL} - T_{REF})$ is the excess noise injected by the calibration noise source when on and we refer to this excess spectral power as T_{STEP}.

Similarly, a second differencing of the two spectra obtained in Eqs. 3.7 and 3.8 yields a measure of the difference between antenna temperature and that of the internal reference:

$$\begin{aligned} P_{OBS} - P_{REF} &= 2|G|^2(T_A + T_N) - 2|G|^2(T_{REF} + T_N), \\ &= 2|G|^2(T_A - T_{REF}). \end{aligned} \tag{3.11}$$

Calibrating $(P_{OBS} - P_{REF})$ with $(P_{CAL} - P_{REF})$ yields the measured temperature

$$\begin{aligned} T_{meas} &= \frac{P_{OBS} - P_{REF}}{P_{CAL} - P_{REF}} T_{STEP}, \\ &= \frac{T_A - T_{REF}}{T_{CAL} - T_{REF}} T_{STEP}. \end{aligned} \tag{3.12}$$

The systematic additives have been canceled by the double differencing. The calibration factor T_{STEP}, which determines the absolute scale for the measured spectrum, is determined by a procedure involving a “calibration of the internal calibrator” by referencing it to noise power from an external termination whose physical temperature is controlled; this is described in Section 3.8.1.

Eq. 3.12 is the spectrum that SARAS 3 measures. While double differencing removes additive systematics, the resulting calibrated spectrum is offset by T_{REF}. In principle T_{REF} can be added back to obtain T_A, however for this the value of T_{REF} must be known by other means. In practice, this can be accomplished by measuring the *physical* temperature of the reference with an accurate thermometer such as a platinum resistance thermometer, assuming that value to be T_{REF} and adding it to T_{meas} to obtain T_A.

3.5 Measurement equations

The calibration of the data recorded in the different switch states is described above in Section 3.4.1 and that leads to Eq. 3.12, which defines the measured spectrum of the antenna temperature. In an ideal system design, all the terms except P_{OBS} in the aforementioned equation are frequency independent and hence the calibrated spectrum has a frequency dependence arising purely from the antenna temperature. However, owing to internal reflections of receiver noise caused by impedance mismatches in the signal path within the receiver chain, the measured spectrum will have frequency dependent terms that may have complex spectral structure. An understanding of these subtle effects is critical in system design as well as selection of suitable components for its realization. In this section, we expand on the description of the measurement equations, taking into account reflections at interfaces. For brevity, only final expressions are given here, details of the derivation may be found in Appendix A.

In the SARAS 3 analog receiver, reflections that occur before the first stage of amplification—in the interconnect and signal path between the antenna and first low-noise amplifier—result in the dominant complex features in the measured spectrum. The important factors to be considered are:

1. The back and forth reflections of signal power corresponding to the antenna temperature, between the antenna and LNA.

2. The backward propagation of receiver noise towards the antenna and its back and forth reflections between the LNA and antenna.

The impedance matching of a wideband antenna to a transmission line is relatively more difficult technically compared to matching the transmission line to a wideband low-noise amplifier. Taking into account first-order reflections of the receiver noise at the antenna, and assuming that the receiver is perfectly matched to the cable connecting it to the antenna, the equation for the calibrated temperature T_{meas} may be written as

$$\begin{aligned} T_{meas} = & T_{STEP}\Big[\frac{P_A - P_{REF}}{P_{CAL} - P_{REF}}\Big] + \\ & T_{STEP}\Big[\frac{P_N}{P_{CAL} - P_{REF}} \times \Big\{2|f||\Gamma_A|cos(\phi_f + \phi_A + \phi) + |f|^2|\Gamma_A|^2\Big\}\Big], \end{aligned} \tag{3.13}$$

where P_A and P_{REF} are the noise powers corresponding respectively to the antenna temperature T_A and the reference termination temperature T_{REF}. P_N is the receiver noise power, corresponding to the receiver noise temperature T_N, that is added to the signal power P_A in the low-noise amplifier. $f = |f|e^{i\phi_f}$ is the fraction of this receiver noise voltage that emerges from the input of the amplifier and back propagates towards the antenna. $\Gamma_A = |\Gamma_A|e^{i\phi_A}$ is the complex reflection coefficient of the antenna: the scattering matrix element S11. ϕ is the phase change arising from 2-way signal propagation in the transmission line connecting the amplifier and the antenna: $\phi = (4\pi\nu l)/(v_f c)$ where l is the physical length of the line, c is the speed of light in vacuum and v_f is the velocity factor of the transmission line. All the terms may have a frequency dependence.

The last two terms in Eq. 3.13 give the spectral additives due to first order reflections of receiver noise at the antenna terminals. If the antenna were perfectly matched to the transmission line connecting to the receiver at all frequencies, $|\Gamma_A| = 0$, the contributions from the last two terms would vanish, and this equation reduces to the case in Eq. 3.12. However, this condition cannot be satisfied at all frequencies for an antenna, particularly for a wideband antenna, and hence contributions from the reflection terms would inevitably appear in the final spectrum. The cosine term arises from the interference between the forward and reflected noise waves of the receiver, these have a relative phase $(\phi_f + \phi_A + \phi)$. Although ϕ_f and ϕ_A would have a variation over the CD/EoR band, it is the length of the transmission line that usually dominates the total change in $(\phi_f + \phi_A + \phi)$ across the band. Therefore, it is critical to limit the length l and a design goal is to keep the total phase change across the band to within a fraction of 2π, to maintain the spectral smoothness in T_{meas}. With this aim, in SARAS 3 the length l has been reduced so that the contributions of the phase terms add up to less than $\pi/2$ so that after calibration, the additive contribution of the receiver noise to the measured spectrum is maximally smooth. Additionally, the antenna is designed to have a maximally smooth reflection coefficient, and the first stage amplifiers are selected to be ultra wideband, so that the characteristics—S11 and also the complex factor f—within the CD/EoR band are maximally smooth. Separately, reduction in the magnitude of the reflected receiver noise is achieved by using a Low Noise Amplifier (LNA), with low noise temperature, as the first amplifier in the receiver.

We next relax the assumption that the LNA is matched to the transmission line. This would introduce additional components in the measured spectrum, as both the signal from the antenna and the backward-propagating receiver noise from the low-noise amplifier undergo multiple reflections between the antenna and LNA input. We may view the transmission line connecting the antenna to the amplifier as a leaky resonant cavity that supports various resonant modes, with the power coupled to the LNA and antenna as leakages. Taking into account the series of higher order reflections, the expression for the measured spectrum may be written in the form

$$\begin{aligned} T_{meas} = T_{STEP} \Big\{ & \frac{P_A[\sum_{k=0}^{+\infty} |\Gamma_N|^k |\Gamma_A|^k \sum_{l=0}^{k} cos\{(2l-k)(\phi_N+\phi_A+\phi)\}] - P_{REF}}{P_{CAL}-P_{REF}} \\ & + \frac{P_N}{P_{CAL}-P_{REF}} \times \Big[\sum_{k=0}^{+\infty} (2|f||\Gamma_A|^{(k+1)}|\Gamma_N|^k cos\{\phi_f + (k+1)(\phi_A+\phi) + k\phi_N\}) \\ & + |f|^2|\Gamma_A|^2 \sum_{k=0}^{+\infty} |\Gamma_N|^k |\Gamma_A|^k \sum_{l=0}^{k} cos\{(2l-k)(\phi_N+\phi_A+\phi)\} \Big] \Big\}. \end{aligned} \tag{3.14}$$

The above Eq. 3.14 is a detailed expression for the measured spectrum. First, it is clear that better impedance matching between the antenna, the transmission line that follows, and the low-noise amplifier at the end of the transmission line, are critical. Improving either or both of these will substantially reduce the amplitude of successive reflections and hence reduce the amplitude and complexity of unwanted structures in the measured spectrum.

In principle, it is possible for all the quantities in Eq. 3.14 to be measured—

both via laboratory calibrations and field measurements—and the measured sky spectrum may be then corrected for the multiple unwanted terms to derive an estimate of the antenna temperature. To examine the validity of the derived model, we provide below in Sec. 3.8.2.1 the results of fitting Eq. 3.14 to laboratory measurements made with the antenna replaced with precision electrical open and short terminations. To model the measurement equation to the accuracy necessary for detection of CD/EoR signals, such an approach requires precise laboratory and in-situ measurements of several quantities, necessitating complexity in system design to allow for the calibrations.

As mentioned above, the SARAS approach has been to design the receiver to avoid precise modelling of the terms in the detailed measurement equation. The SARAS 3 design strategy has taken an alternate path recognising that the complexity of the unwanted spectral structures may be substantially reduced by reducing the path length over which the receiver temperature components suffer multiple internal reflections. As the total differential path increases with multiple reflections, the interference between the signals with larger differential paths results in higher order structure in the measured spectrum.

It may be noted here that the only relevant paths are those which involve the antenna terminals at one end, since it is only those terminals that are Dicke switched over to the reference port for calibration. Internal reflections between any pair of impedance mismatches downstream of the low-noise amplifier would result in bandpass structure, which would be calibrated out. In a radiometer receiver chain where the antenna is followed by successive stages of amplification, each amplifier provides an isolation for multi-path propagation that is limited to the difference between the reverse loss and forward gain. Thus the length l over which uncalibrated multi-path reflections may occur is not limited to just the length between the antenna and first low-noise amplifier, but may be the effective length between the antenna and many stages of amplification, depending on the effective isolation provided by successive components. Therefore, the SARAS 3 design has introduced optical isolation immediately following the first amplification module, by introducing an optical modulator, fiber optic transmission line and a demodulator. Thus for SARAS 3 the only relevant path is that between the antenna and the optical isolator; this path has been made much smaller than the wavelengths of operation.

Finally, we qualify our receiver by evaluating the receiver performance by replacing the antenna with precision loads as well as a load with frequency-dependent reflection coefficient similar to that of the CD/EoR antenna and examining for the level of confusion between spectral structures arising from uncalibrated internal systematics and plausible CD/EoR signals. Details of these tests are given below in Section 3.8.

3.6 Implementation of the SARAS 3 receiver

The SARAS 3 receiver is implemented in two sections: an antenna base electronics and a remote station electronics unit that is located about 150 m away. The radio frequency (RF) signal from the antenna base electronics travels as RF over fiber (RFoF), intensity modulated on a laser carrier and in a single-mode fiber, to the remote station analog electronics units for further signal processing. The entire analog receiver chain is designed to operate in the band 40 to 230 MHz that is defined by high and low pass filters; however, when operated in sites where FM might be present with intolerable strength, the low pass filters may be replaced with units that cut the band above 87.5 MHz for CD signal detection, or the high pass filters may be replaced with units that cut the band below 110 MHz for EoR signal detection. Control signals that time the switching of states in the analog sections are generated in a control and acquisition computer that is part of the digital receiver, and sent to both analog units via multi-mode optical fiber. The digital receiver is co-located with the remote analog electronics unit about 150 m from the antenna and the electronics at its base.

The antenna base electronics unit has the components in Fig. 3.2 in the signal path. An electro-mechanical surface-mount RF switch with an insertion loss less than 0.1 dB and isolation greater than 60 dB selects between the antenna temperature and that from a reference ambient temperature termination. This reference

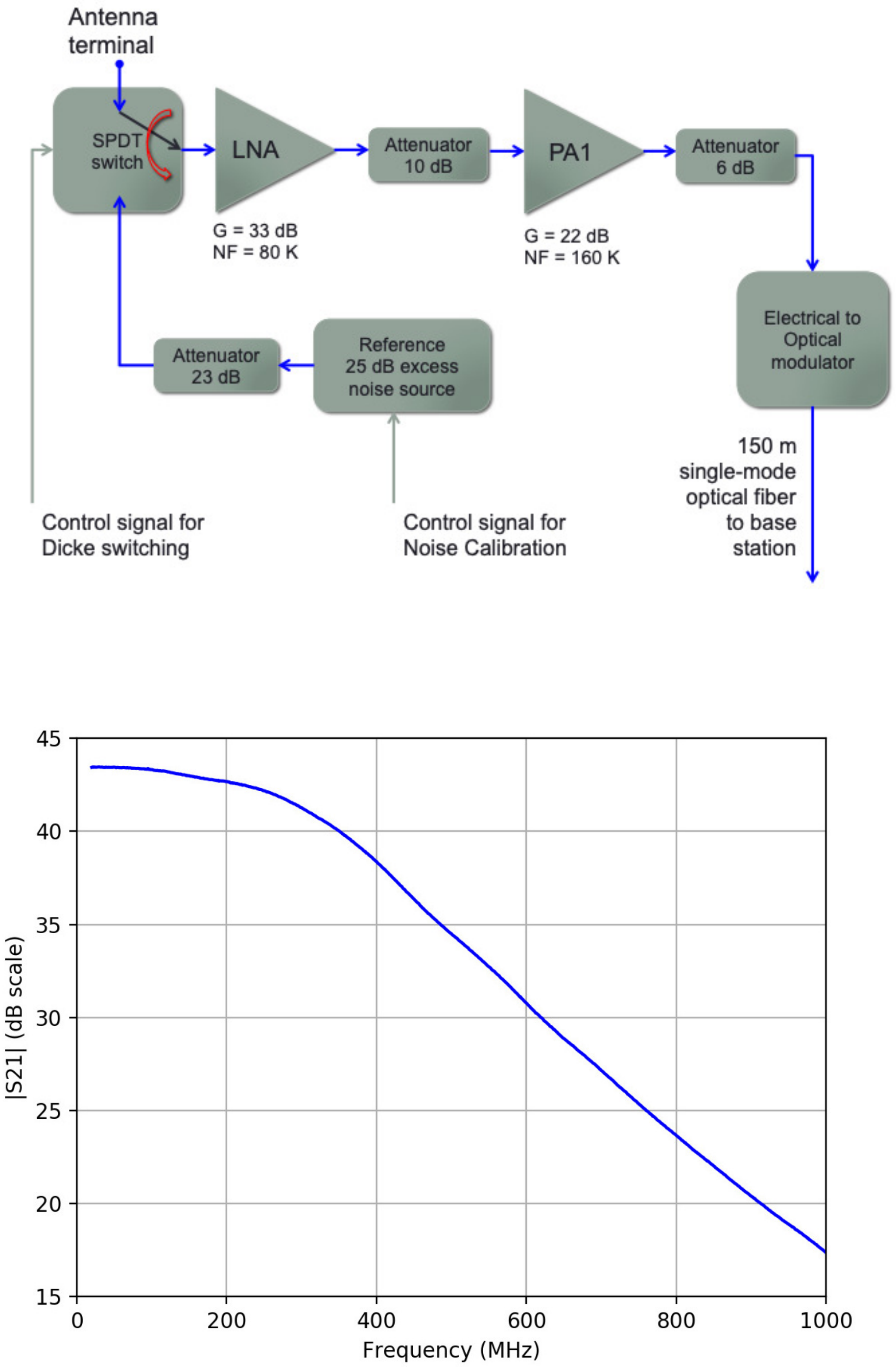

Antenna terminal
SPDT switch
LNA
G = 33 dB
NF = 80 K
Attenuator 10 dB
PA1
G = 22 dB
NF = 160 K
Attenuator 6 dB
Electrical to Optical modulator
150 m single-mode optical fiber to base station
Reference 25 dB excess noise source
Attenuator 23 dB
Control signal for Dicke switching
Control signal for Noise Calibration
|S21| (dB scale)
45
40
35
30
25
20
15
0
200
400
600
800
1000
Frequency (MHz)

Figure 3.2: SARAS 3 analog electronics at the base of the antenna. A block diagram of the front end electronics is shown in top panel. Bottom panel shows the measured gain of this unit, from the antenna terminal to the output of optical to electrical modulator housed inside remote station analog electronics.

is implemented as a matched 50-Ω 23 dB attenuator connected to a flat-spectrum noise source of 25 dB excess noise. Calibration signal is input to the radiometers when this noise source is on; the radiometer sees a reference termination when off. Interconnects between the switch and terminals of the antenna and reference use adaptor bullets, without cables, to minimize path lengths. When the control to the switch is off, the analog receiver chain is disconnected from the antenna thereby providing protection from static. Protection during observing is provided with Schottky diodes and PIN diodes to prevent environmental electrostatic discharge and high power RFI damaging the sensitive low-noise amplifier. These devices also provide protection from human body static when the antenna is manually mounted on the receiver.

The first active device is a MMIC based low-noise amplifier with gain of 33 dB, followed by a second amplifier providing an additional gain of 22 dB. A chip attenuation of value 10 dB is placed in the path between the two for improved impedance matching and isolation. The first amplifier has noise temperature of about 80 K and the second 160 K. The switch, protection devices, amplifiers, attenuations, along with a final attenuation of 6 dB following the second stage of amplification, are all accommodated in a single custom-made printed circuit board (PCB) designed in-house and mounted in a compact aluminum enclosure. A modular electrical to optical RFoF modulator follows that is based on a distributed feedback (DFB) laser and provides intensity modulated 1310 nm laser light on a single-mode fiber.

In order to keep the signal path between the antenna and optical transmitter as small as possible, no filtering is done in the antenna base electronics. The signal bandwidth at this stage is therefore determined by the antenna bandwidth and that of the pair of amplifiers. While the reduced path length helps in maintaining smoothness, the lack of filtering increases the susceptibility of the receiver to strong out-of-band RFI. As a consequence, deployments of radiometers with this SARAS 3 receiver are constrained to be at remote sites with appropriately low levels of RFI, both within the band and also out of band.

The remote station analog electronics unit has the components shown in Fig. 3.3. The key block in this unit is a cross-over switch, implemented as a high-reliability electro-mechanical coaxial switch, that is connected to the sum

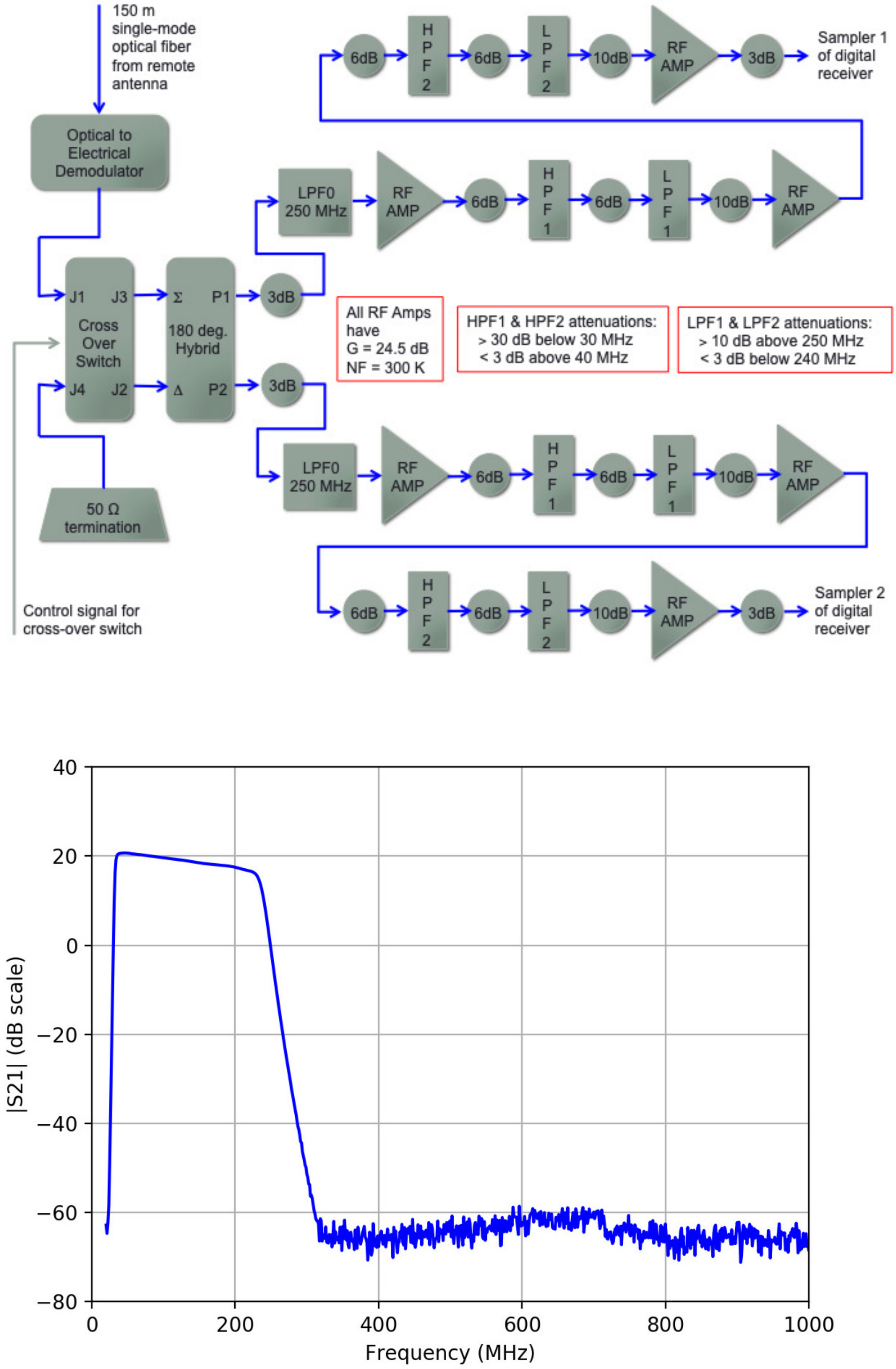

150 m single-mode optical fiber from remote antenna
Optical to Electrical Demodulator
J1
J3
Cross Over Switch
J4
J2
50 Ω termination
Control signal for cross-over switch
Σ
P1
180 deg. Hybrid
Δ
P2
3dB
3dB
LPF0 250 MHz
RF AMP
6dB
HPF1
6dB
LPF1
10dB
RF AMP
6dB
HPF2
6dB
LPF2
10dB
RF AMP
3dB
Sampler 1 of digital receiver
Sampler 2 of digital receiver
All RF Amps have G = 24.5 dB NF = 300 K
HPF1 & HPF2 attenuations: > 30 dB below 30 MHz < 3 dB above 40 MHz
LPF1 & LPF2 attenuations: > 10 dB above 250 MHz < 3 dB below 240 MHz
|S21| (dB scale)
40
20
0
−20
−40
−60
−80
0
200
400
600
800
1000
Frequency (MHz)

Figure 3.3: SARAS 3 analog electronics at the remote station 150 m from the antenna. Top panel is a block diagram of the electronics signal chain. The bottom panel shows magnitude of measured gain of one arm of the analog electronics, from the output of the optical to electrical demodulator to sampler input. The other arm has an identical response and is not shown here.

and difference ports of a 180° hybrid. The RFoF signal is first converted back to RF in an optical to electrical demodulator and this signal forms one input to the block. The ambient-temperature noise power of a 50-Ω matched termination forms the second input. The cross-over switch has an isolation better than 80 dB in the receiver operating band. The two outputs P1 and P2 are the sum and difference of these two inputs, and toggling the cross-over switch causes these outputs to swap. The block implements phase switching and provides a pair of signal outputs for a correlation receiver.

The pair of signals from the block go through identical receiver chains. The signals are first low-pass filtered to about 250 MHz with coaxial Butterworth filters. Then follows three stages of amplification with sharper low and high pass filters in between, to limit the 6-dB bandwidth to 40–240 MHz and have 60 dB attenuation below 30 MHz and above 260 MHz. These filters are designed and developed in-house using discrete inductors and chip capacitors realizing 9th-order Elliptic filter approximations. The amplifiers in the remote station receiver chains, where power levels are higher compared to the antenna base electronics chain, are built using modules with 24 dB gain and 1 dB compression point of +24 dBm power level, which leaves a headroom exceeding 40 dB for RFI within the band. The band-limited signals are available for the digital receiver.

All of the antenna base electronics are powered by a pair of Li-Ion battery packs of 18 V/20 AH rating; only one battery pack is used at any time and the power source may be switched remotely. Together they are capable of operating the antenna base electronics backend for eight nights of observing before it needs to be accessed for charging the battery packs. Linear regulators with low dropout ratings are used to supply the different voltages required; the voltages are supplied to the different modules via shielded coaxial cables to obtain about 40 dB of isolation to radiated couplings within the receiver enclosure. The power supply lines are provided with good filtering at both the source and destination ends to ensure that these lines are not a source of feedbacks for RF power.

The antenna base electronics is housed in a square aluminum welded box. The top face of the box is also part of the ground plane of the antenna and has a UHF to SMA adaptor at its center for the monopole terminal of the antenna. This top face has mounting holes along flanges at its edges where extensions to the ground plane may be fastened to extend the ground plane to any desired area around the monopole. Styrofoam blocks are glued on to guide the placement of the monopole to the connector and prevent lateral movement. This top view of the antenna base electronics box is shown in the top left panel in Fig. 3.4.

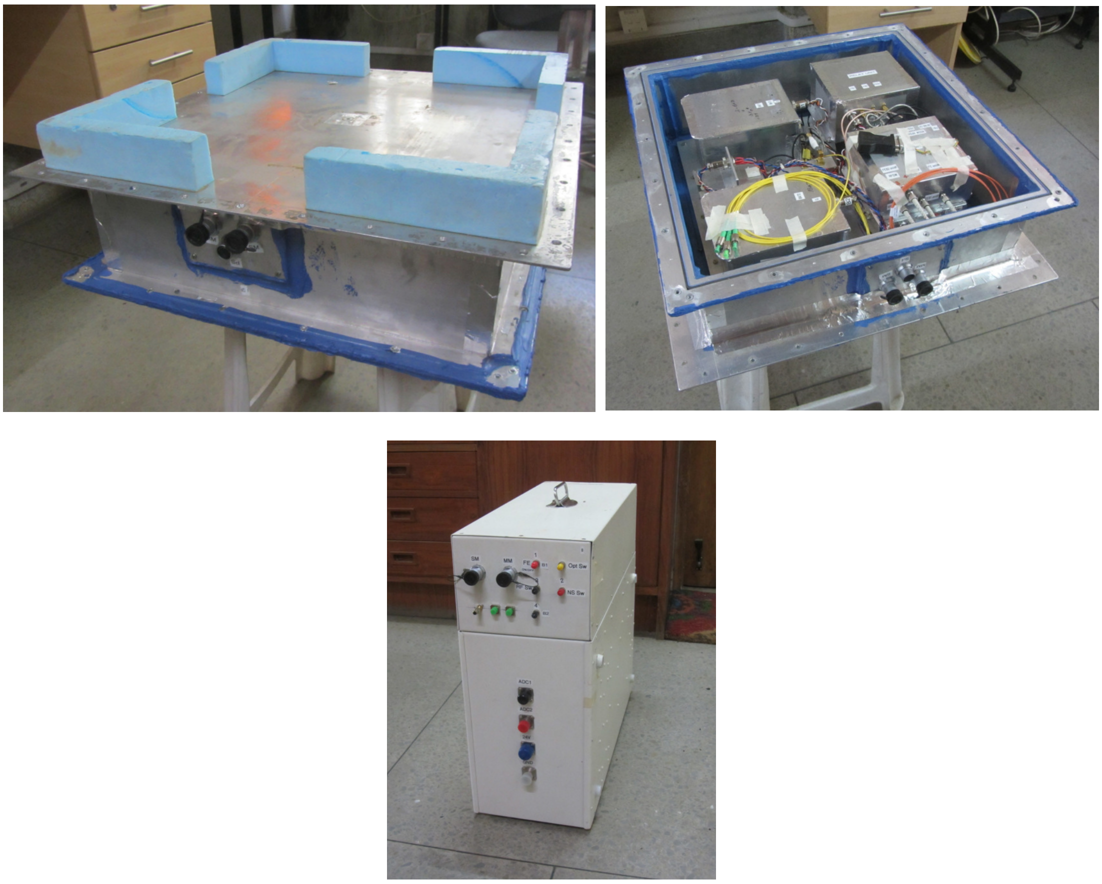


Figure 3.4: The analog electronics enclosures beneath the antenna and at the remote base station. In the top panel on the left and right are shown the enclosure at the antenna base; the left panel shows its view from the top and in the right panel is shown the view from below with the bottom cover removed. The bottom panel shows the enclosure that houses the analog electronics at the remote station.

A photograph of the antenna base electronics enclosure, flipped over and with the cover removed, is in the top right panel of Fig. 3.4. The electronics modules are mounted on a plate fixed to the top panel of the enclosure, and a flat cover is bolted on below. All components are mounted in separate aluminium chassis even within the enclosure and these are sealed with screw-on lids and covered along mated edges with aluminium tape. Electromagnetic sealing and water proofing of this enclosure is provided by having 'O'-ring grooves machined on the flange, in which gaskets are placed that provide both these protections. Additionally, silicone waterproofing sealant is applied all along edges where the lid mates with the box. Connectors for the single-mode and multi-mode fiber cables, plus an additional connector that provides access to the batteries within the enclosure for charging, are in a plate mounted on one side of the enclosure.

Within the antenna base enclosure is a temperature logger that records the physical temperature of the reference termination. This provides the noise temperature of the load that is Dicke switched in place of the antenna. The temperature of the top ground plate of the receiver enclosure is also recorded. The logger is accessed only after an observing campaign is completed and the sealed enclosure is opened.

The remote station analog electronics is in an enclosure shown in the bottom panel in Fig. 3.4. The optical components are in a separate segment at the top, and the electrical modules are mounted on the inner side walls of the enclosure. The walls are hinged and may be opened on either side to access the receiver chains that form the two arms of the correlation receiver. This enclosure receives 24 V d.c. power via a pair of coaxial cables from a battery box, and control signals via four multi-mode fibers from the digital receiver. Single-mode and multi-mode fiber cable connectors are provided for the 150-m outdoor fiber-optic cables that run from this remote station to the antenna base electronics enclosure. A pair of panel mounted coaxial connectors are provided for cabling the processed RF power to the digital correlation spectrometer.

The RFoF link is provided by a rugged waterproof fiber cable with IP65 rating and single-mode APC connectors. The RFoF link has an overall gain of 5 dB and a noise figure of 19 dB, or 23,000 K. Since the RFoF link is preceded by a gain of about 39 dB, its input referred noise temperature goes down by the same amount, contributing about 3 K to the system temperature. It is the RFoF link that limits the headroom for RFI to about 30 dB above sky noise.

3.7 Sensitivity of the SARAS 3 receiver

The sensitivity of a receiver depends on the thermal noise that the calibrated spectra have, as well as any residual systematic errors. By careful design, SARAS 3 receiver and antenna have maximally smooth responses and hence any residual systematic is expected to be less than a mK; this is demonstrated below in Section 3.8. Therefore, the sensitivity of the system depends predominantly on the random measurement noise in the spectra recorded in each of the six states through which the system cycles, and the mechanics of calibration that combines these recordings to compute calibrated spectra for the antenna temperature.

In the SARAS 3 receiver, spectral powers measured in each of the six switched states would have different associated noise. For measurement data recorded in each of these states, associated uncertainties are computed and stored as meta-data in the pre-processing stages of the software pipeline. The sensitivity of the system is then estimated by propagating these uncertainties through the calibration equation given in Eq. 3.12, which is reproduced here for reference:

$$T_{meas} = \frac{P_{OBS} - P_{REF}}{P_{CAL} - P_{REF}} T_{STEP}. \tag{3.15}$$

All terms in the above equation are functions of frequency and the calibration is computed separately for each spectral channel.

P_{OBS}, P_{REF} and P_{CAL} individually represent differences between the spectral powers measured in a pair of states, where the phase of the antenna signals is switched relative to internal additives in the arms of the correlation receiver. We denote the rms noise associated with each of them as $\Delta P_{OBS}, \Delta P_{REF}$ and ΔP_{CAL} respectively. For small perturbations, the noise in T_{meas} can then be approximately expressed as a combination of the rms noise in the three power measurements:

$$\Delta T_{meas} = T_{STEP} \sqrt{\left(\frac{\partial T_{meas}}{\partial P_{OBS}} \Delta P_{OBS}\right)^2 + \left(\frac{\partial T_{meas}}{\partial P_{REF}} \Delta P_{REF}\right)^2 + \left(\frac{\partial T_{meas}}{\partial P_{CAL}} \Delta P_{CAL}\right)^2}. \tag{3.16}$$

Evaluating the partial derivatives yields

$$\Delta T_{meas} = T_{STEP}\Big[\Big\{\frac{\Delta P_{OBS}}{P_{CAL}-P_{REF}}\Big\}^2 + \Big\{\frac{(P_{CAL}-P_{OBS})\Delta P_{REF}}{(P_{CAL}-P_{REF})^2}\Big\}^2 + \Big\{\frac{(P_{OBS}-P_{REF})\Delta P_{CAL}}{(P_{CAL}-P_{REF})^2}\Big\}^2\Big]^{1/2}. \quad (3.17)$$

In the above equation, each of the powers P_{OBS}, P_{REF} and P_{CAL} may be expressed in terms of the system temperatures $(T_A + T_N)$, $(T_{REF} + T_N)$ and $(T_{CAL} + T_N)$ in the respective states using the general form $P = 2|G|^2T$, where G is the system voltage gain. The rms noise in these measured spectral powers is related to the corresponding system temperatures by the radiometer equation:

$$\Delta P = \frac{2|G|^2T}{\sqrt{B\tau}}, \quad (3.18)$$

where τ is the integration time and B is the noise-equivalent bandwidth of the spectral channels. It may be noted here that the gain term G cancels when the substitutions are made and therefore precise information regarding the receiver gain as a function of frequency is not required for estimating the sensitivity.

We may now estimate the rms noise in calibrated spectra. As stated above, the receiver is designed to operate over frequencies 40–230 MHz, and intended to be used with scaled conical monopole antennas that operate in octave bands and cover the range in staggered bands. We compute here the sensitivity to Cosmic Dawn and Reionisation signals when fitted with the floating cone-disc antenna (Raghunathan et al., 2020) covering an octave band from 43.75 to 87.5 MHz. In this band the sky brightness is a maximum and hence sensitivity is the lowest; therefore, bands at higher frequencies will have greater sensitivities and lower rms noise for same integration times.

We first estimate the noise at 70 MHz. Away from the Galactic plane, the sky temperature at this frequency is about 2000 K. The receiver noise is about 80 K and we assume that the reflection efficiency of the antenna is 70% and radiation efficiency is 50%. The system temperature when the switch is in OBS state and the receiver is connected to the antenna would be about 930 K, with 150 K contribution coming from the resistive loss in the environment of the antenna.

The calibration step T_{STEP} of the noise injection corresponds to a temperature of 630 K when referred to the antenna terminals. In the reference (REF) and calibration (CAL) states, the system temperatures would be about 380 K and 1010 K respectively. The SARAS correlator provides spectra with a spectral resolution of 61 kHz, corresponding to 4096 spectral channels over 250 MHz (Girish et al., 2020). For spectra with this resolution, and integration time of 2.7 seconds in each of OBS, REF and CAL states, using Eqs. 3.17 and 3.18, the sensitivity is 3.2 K per channel, per spectra. This is the rms noise in a total observing time of $2.7 \times 3 = 8.2$ s. For an observation session spanning eight hours, there would be about 3400 such spectra, giving a sensitivity of about 54 mK per channel. If this averaged spectrum were smoothed to a noise-equivalent bandwidth of 4 MHz, the rms measurement noise would be 6 mK. Referred to the sky, by accounting for the total efficiency, the rms measurement noise in the estimate of sky brightness temperature would be 17 mK.

The above computation may be extended to estimate the expected distribution of rms noise across the octave band, using measured antenna efficiencies and sky models. We use measured reflection and radiation efficiencies given in Raghunathan et al. (2020) for the SARAS 3 antenna, and use GMOSS (Sathyanarayana Rao et al., 2017a) model foreground. We assume that the observing is at latitude $+14^\circ$ and over local sidereal time (LST) from 10 to 18 hr that includes a transit of the Galactic plane across the antenna beam. The expected average sky spectrum along with the expected calibrated spectrum of the antenna temperature, over an octave band from 43.75 to 87.5 MHz, is given in Fig. 3.5.

The OBS data acquired with receiver connected to sky are then calibrated using Eq. 3.15 with CAL and REF spectra acquired with receiver switched to reference and with calibration noise on, respectively. The rms noise distribution in these spectra are computed using Eqs. 3.17 and 3.18 and shown in Fig. 3.6. It may be noted that, as expected, the rms noise reduces across the band towards higher frequencies since the system temperature is sky dominated and the sky temperature reduces with frequency.

The native spectral resolution of the SARAS digital spectrometer has been

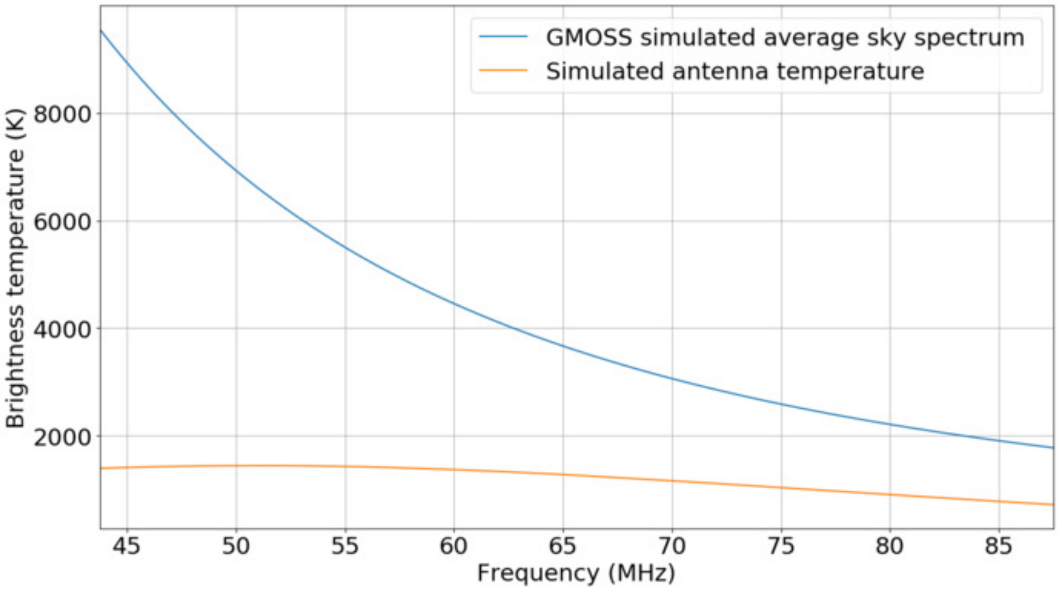


Figure 3.5: Expected antenna temperature for an observation with the SARAS 3 antenna, at latitude +14° and over local sidereal time (LST) from 10 to 18 hr. Also shown is the expected average sky spectrum. It may be noted that the antenna temperature is defined as per Eq.2.13, such that it is the average sky spectrum multiplied by the mismatch factor or reflection efficiency of the SARAS 3 antenna (Raghunathan et al., 2020)

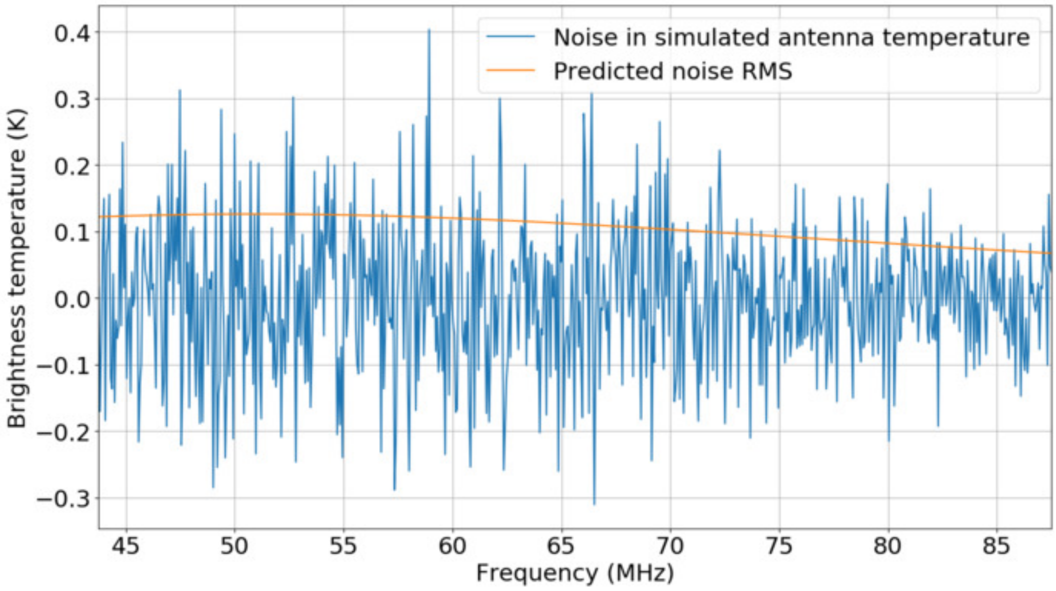


Figure 3.6: An example of the noise component that might be present in an 8 hr mock observation. The rms noise expected across the band is also shown in orange.

designed to be 61 kHz, much finer than the characteristic scale of expected global 21-cm signal, so that any man-made narrow band RFI may be identified and rejected. Since the global 21-cm signal is expected to have broad spectral shapes, we may smooth the spectra—following rejection of data corrupted by RFI—to a resolution much poorer than the native spectral resolution of 61 kHz of the digital spectrometer, to increase sensitivity without significant loss of signal. The effect of smoothing on the rms of the measurement noise is shown in Fig. 3.7, where the distribution of the rms noise across the band is plotted for smoothing to a range of noise equivalent widths. If the measurement noise is Gaussian random, the incremental gain in sensitivity is roughly proportional to the square root of number of independent points across which smoothing is performed.

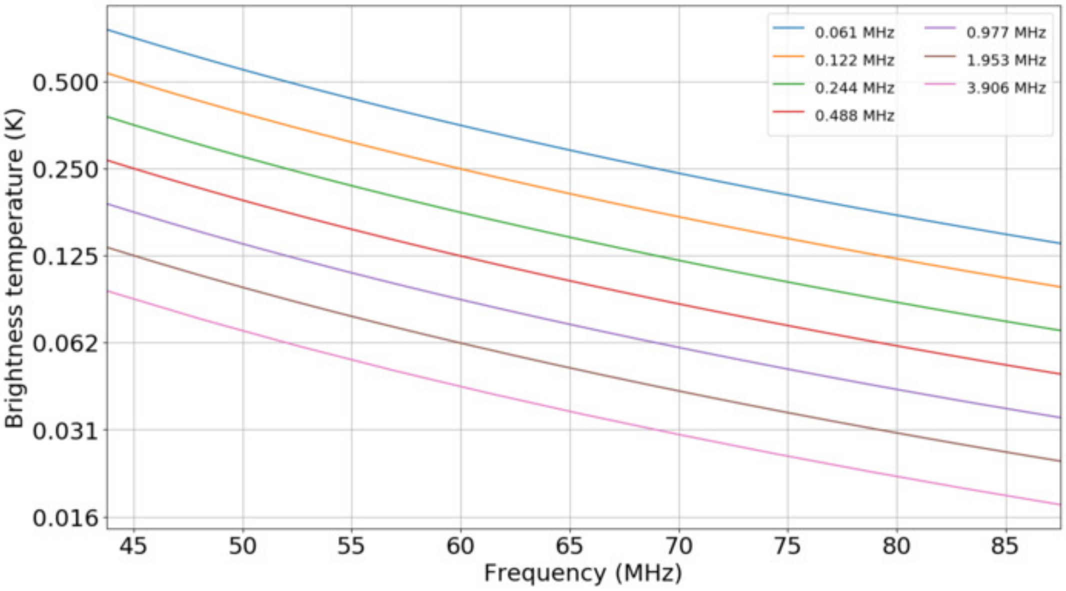


Figure 3.7: The distribution in rms noise for smoothing to different noise equivalent widths, for the 8-hr mock observation. It may be noted that the rms noise estimate has been referred to the sky domain by correcting for the total efficiency of the radiometer; therefore, the rms noise represents 1-σ uncertainty in measurement of sky brightness temperature.

3.8 Laboratory tests

3.8.1 Absolute calibration

Absolute calibration of the measured data is the process by which the acquired data in some arbitrary counts are converted to units of kelvin using a scaling factor. In the calibration equation given in Eq. 3.12, T_{STEP} is a scaling factor that sets the overall temperature scale by virtue of its multiplication with the dimensionless ratio of powers. It may be noted here that this calibration of the measurement data with T_{STEP} sets the data to be antenna temperature in kelvin scale at a reference point that is the input to the switch, which is same as the terminals of the antenna. The value of the scaling factor T_{STEP} is determined by the calibrator assembly formed by the noise source and the attenuators that follow it; the receiver is switched between the antenna and this calibrator assembly. Though it is possible to compute T_{STEP} using the published excess noise ratio (ENR) of the noise source and the values of the attenuators between the noise source and switch; for improved accuracy, a laboratory measurement of T_{STEP} is required.

Laboratory measurement of T_{STEP} is carried out by replacing the antenna with a source of RF noise whose spectral power is known. For this measurement, we replaced the antenna with a precision 50 Ω termination, and assume that the noise power per unit bandwidth from this matched impedance is given by $k_B T$, where k_B is the Boltzmann constant and T is the physical temperature of the termination. Accurate temperature probes are firmly attached to this termination and to the internal reference formed by the attenuators; thermal insulation is provided so that the thermal resistance between the probes and termination/reference attenuator is significantly lower compared to that between the probes and environment.

Raw (uncalibrated) spectral data are recorded with the receiver cycled sequentially between the termination, reference with noise source off, and reference with noise source on. Physical temperatures of both the termination and the reference are logged. For this experimental setup, when the recorded data are calibrated using Eq. 3.12, the calibrated spectrum is ideally expected to be the difference between the physical temperatures of the termination and the reference ports, if the value of T_{STEP} is accurate. The expectation is that $T_{meas} = T_{50} - T_{REF}$, where T_{50} and T_{REF} are the physical temperatures of the 50 Ω termination and reference port respectively.

Despite the experimental setup in which thermal resistance is added between the environment and termination/reference, and care is taken to bond the temperature probes to the termination/reference attenuation, in practice a finite and significant offset was unpreventable between the temperatures logged by the probes and true noise temperatures of the termination/attenuation. The offset errors may be written in the form: $T_{50} = T_{50,m} + T_{os_1}$ and $T_{REF} = T_{REF,m} + T_{os_2}$, where $T_{50,m}$ and $T_{REF,m}$ are the temperatures of the 50 Ω termination and reference as measured by the probes, and T_{os_1} and T_{os_2} denote the offsets in temperatures, which may be positive or negative. Substituting these relations including error terms into Eq. 3.12, we obtain:

$$T_{50,m} - T_{REF,m} = T_{STEP} \frac{P_{OBS} - P_{REF}}{P_{CAL} - P_{REF}} + T_{os}, \tag{3.19}$$

where $T_{os} = T_{os_2} - T_{os_1}$.

The equation to be solved is a linear equation requiring at least two measurements to solve for the unknowns. Traditionally, the termination is placed in baths of two different temperatures, and the measured data used to solve for the two unknowns T_{STEP} and T_{os}. Instead, for the SARAS 3 receiver, we made two dynamic measurements. The termination is first immersed in a hot water bath in a well insulated dewar and allowed to cool slowly over time. Separately, the termination is immersed in ice-cold water in the dewar and allowed to slowly warm over time. The recorded data constitute an overdetermined system with two unknowns, and a solution can be found with a least squares fitting of a straight line. Plot of the differential of the probe temperatures: $(T_{50,m} - T_{REF,m})$ versus ratio of the differential powers recorded $(P_{OBS} - P_{REF})/(P_{CAL} - P_{REF})$ is shown in Fig. 3.8. The fit of a straight line yields the slope and intercept, which are the two unknowns T_{STEP} and T_{os}.

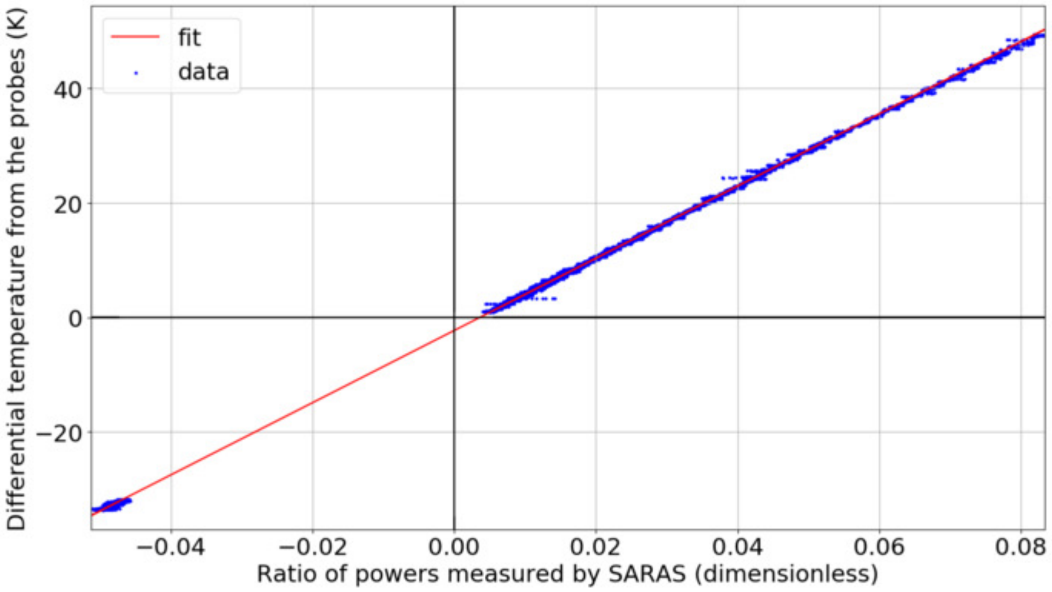


Figure 3.8: Fit to data acquired with termination at antenna port placed in warm and cold baths, which provides estimate of the absolute calibration scale factor T_{STEP}.

The fit gives the value of T_{STEP} to be 630 K and the y intercept gives T_{os} to be about -2.3 K. The goodness of fit is a confirmation of the model for the experimental setup.

3.8.2 Termination tests

Eq.3.14 gives a detailed description of the measured spectrum, including the expected systematic structures that it may contain owing to non-ideal component behavior within the receiver chain. The measured spectrum may be viewed on the whole as consisting of three components. The first is the signal from the antenna modified by the transfer function of the system. Departure in this transfer function from an ideal flat response is a multiplicative error. The measurement differences the antenna signal with the reference power. The reference is the second component and we assume in the analysis herein that it is ideal and of flat spectrum to the accuracy required for CD/EoR detection. The third component is additive noise from along the receiver path: this is dominated by the receiver amplifier noise and the internal additives appear in the measured spectrum with their band shapes multiplied by corresponding transfer functions.

In this section, we present results of laboratory measurements done as qualification tests of the SARAS 3 receiver, with the antenna replaced successively by precision terminations and a load that is a circuit simulator of the antenna characteristics. The load resembling the antenna characteristics—hereinafter called an 'antenna simulator'—is a resistance-inductance-capacitance network purpose built to have a reflection coefficient $S11$ amplitude that matches the reflection coefficient measured for the antenna. The aim of the qualification tests were to examine the limitations of the design effort in the SARAS 3 receiver, which was aimed at realizing a spectral radiometer whose unavoidable systematics were relatively smooth and hence separable from CD/EoR spectral profiles predicted in cosmological models.

In order to investigate the internal systematics of the radiometer, ideally a measurement that does not contain any sky signal is desired. This can be obtained by replacing the antenna with a perfectly mismatched termination, which may be an electrical open or short. Indeed, such terminations can provide an estimate of the maximum levels of additive systematics that any spectral measurement with the radiometer might contain. Data for such a test were acquired by replacing the antenna with precision open and short terminations and acquiring spectra with the full SARAS 3 receiver chain for about 16 hr for each termination, with the receiver cycling through the switch states exactly as designed for celestial observations. The data were then calibrated in the standard process and reduced to provide a single average spectrum for each termination. The terms in the measurement equation that depend on the reflection coefficient Γ_A at the antenna terminal will flip sign when changing from an electrical short to open at the antenna terminals. Therefore, we examine the measurement data linearly combined to yield (open-short)/2 and (open+short)/2, which would separate terms that depend on Γ_A and those that do not, respectively.

3.8.2.1 Modeling laboratory measurements with the measurement equation

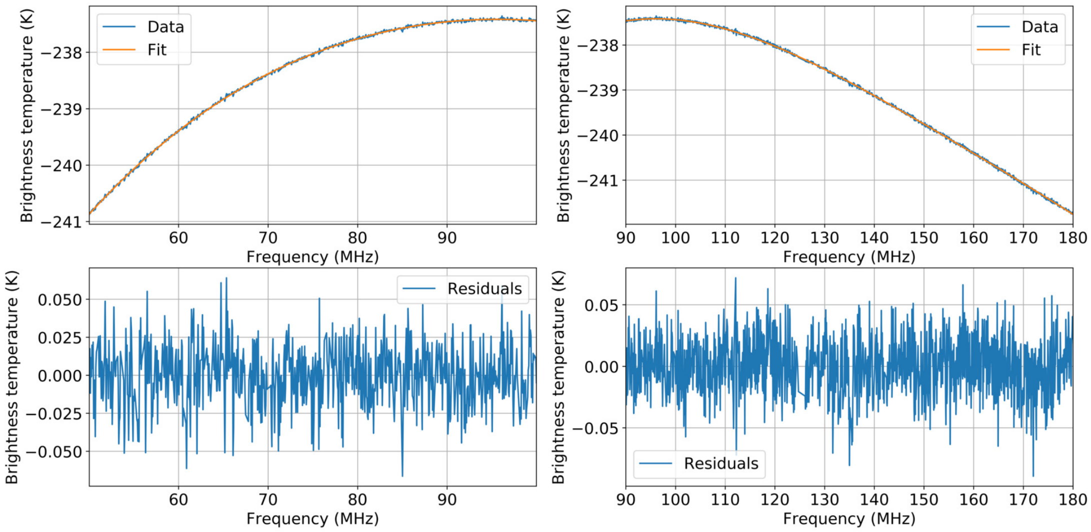
The measurement equation, given by Eq. 3.14, was fitted to the measured data in two frequency bands 50–100 MHz and 90–180 MHz, which roughly correspond respectively to the bands in which CD and EoR related 21-cm signals are expected. The reflection coefficients of the precision open and short terminations are assumed to be ideal, since at the frequencies of interest the effects of their fringing capacitance and inductance are negligible. The fraction f of the receiver noise voltage that emerges from the input of the amplifier and back propagates towards the antenna is modeled as a complex variable that is a constant, independent of frequency, throughout the band of interest. Similarly, the complex reflection coefficient at the input terminals of the receiver, Γ_N, is also modeled as a complex variable that is a constant throughout the band of interest. The path length l is a free parameter in the modeling. Spectra recorded with the switch connecting the receiver to the REF port were calibrated using the difference CAL$-$REF; this measurement is expected to represent the sum of receiver noise and ambient temperature of the matched load at the REF port. We subtract the recorded ambient temperature of the REF termination from this calibrated REF spectrum and derive an estimate of the receiver noise temperature T_N by fitting a third-order polynomial to this residual calibrated REF port spectrum. Thus we effectively adopt a five-parameter model for the system—two complex variables and one real variable—and fit this to the termination measurement data. Suitable boundary conditions are imposed to prevent the fits from returning unphysical parameters.

In Fig. 3.9, the results of fitting Eq. 3.14 to (open+short)/2 and (open$-$short)/2 are shown. It is seen that the measurement equation is indeed successful in modeling the measured spectra and its various components to within the measurement noise. However, in order to accurately model the system response to mK levels the assumptions made above regarding various model parameters have to be relaxed. The various model parameters have their own frequency dependence and if we were to attempt an accurate estimation of that as well from the measured data, the optimization problem would have a large number of free parameters. Allowing for larger numbers of free parameters in the measurement equation would result in a model that would fit the data with reduced residuals. However, the model describing the systematics would also fit out any CD/EoR signal when used to model sky spectra, and hence substantially reduce the sensitivity of the radiometer. Alternately, Γ_N and f may be measured in the laboratory or in-situ using, for example, an accurate impedance tuner at the receiver input to obtain data for various impedance states (this is sometimes referred to by the name source pulling). However, such measurements are difficult to make with the desired accuracy because the receiver would have to be switched to a different measurement apparatus for this, and the receiver parameters might change over time and between laboratory conditions and the field. Therefore, we model the laboratory and field data with maximally smooth polynomials (Sathyanarayana Rao et al., 2017b), as discussed below.

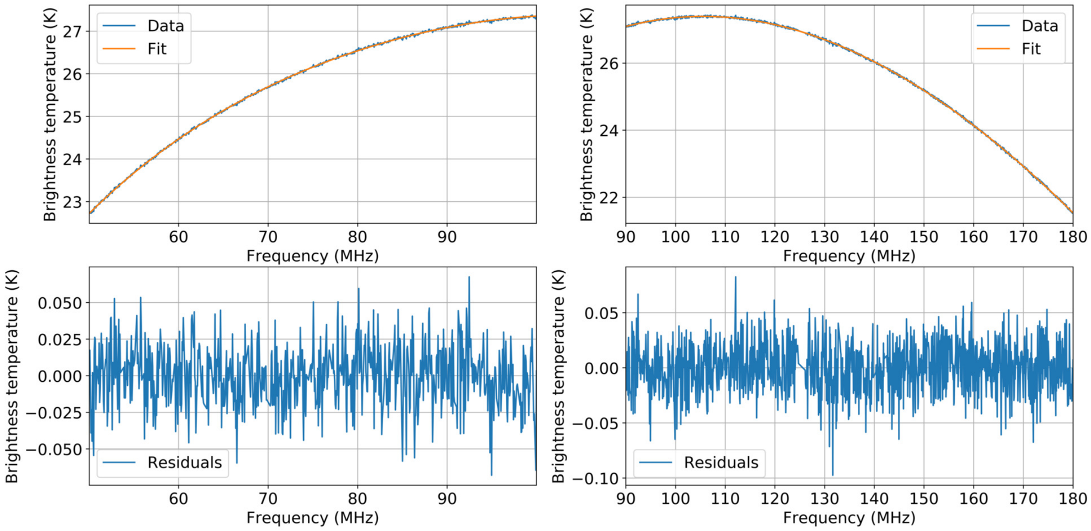
3.8.2.2 Modeling laboratory measurements with maximally smooth polynomials

The measurement data with different terminations, and in bands appropriate for detection of CD and EoR features, were fitted using maximally smooth (MS) polynomials. In its smoothest formulation, maximally smooth polynomials allow a single maximum or minimum within the band. A relaxation is to allow for a single zero crossing within the band in the second and higher order derivatives, resulting in a modified maximally smooth polynomial. These polynomials may be of arbitrarily large order; nevertheless, they would not completely fit out CD/EoR signals that might be present in measurements of sky spectra. The maximally smooth polynomials would fit out part of the signals being searched for, and hence reduce the sensitivity of the experiment. Modified maximally smooth polynomials would fit out a greater fraction. The goal of the receiver qualifying tests has been to evaluate whether the internal systematics may be modeled with maximally smooth functions, or in their modified form.

In Fig. 3.10, we fit the (open+short)/2 measurement data versus frequency



Measurement equation fit to (open+short)/2 data



Measurement equation fit to (open-short)/2 data

Figure 3.9: Modeling measured data in the 50–100 MHz and 90–180 MHz bands using the measurement equation Eq. 3.14. The sum and difference of data acquired with the antenna replaced with precision open and short terminations are modeled.

with a *maximally smooth* polynomial in the CD band 50–100 MHz. This linear combination would cancel terms in Eq. 3.14 that depend on odd powers of Γ_A; therefore, the last term in the third line of the equation survives whereas part of the series in the second line drops out. The panel on the top shows the data with MS fit overlaid, demonstrating the goodness of fit. The measurement data is with negative temperatures in the y-axis because the measurement represents difference between powers from the antenna terminal and that from the reference. For this measurement, an electrical short/open is at the antenna terminals and this has lower noise power compared to that from the reference termination, which is a matched ambient temperature load. The middle panel shows the fitting residuals. The residuals are displayed with their native resolution of 61 kHz and also shown smoothed using kernels of increasing bandwidths. Using Eq. 3.17 and adopting realistic values for the noise temperatures in the different switching states, as discussed in Sec. 3.7, we expect that the calibrated (open+short)/2 spectrum would have an rms noise of 20 mK at native resolution, which is indeed what is measured for the data. The bottom panel of the figure shows the variation in the variance of the residuals as a function of the full width at half maximum (fwhm) of the smoothing kernel. If the spectra were Gaussian random noise, this variation is expected to be a straight line with slope -1 in log-log domain; for comparison, we also show in the panel this expected rate of fall.

Fig. 3.11 shows the result of fitting an MS polynomial to the (open$-$short)/2 spectrum in the 50–100 MHz band. This linear combination is expected to cancel the entire term in the third line of Eq. 3.14, which wholly depends on even powers of Γ_A, and also part of the series in the second line where the terms depend on even powers of Γ_A. As in the case of the analysis of the (open+short)/2 spectra, here too we have displayed the fits, fit residuals smoothed to lower resolutions, and compared the run of variance with smoothing fwhm.

Figs. 3.12 and 3.13 show the results of modeling (open+short)/2 and (open$-$short)/2 measurement data respectively in the EoR band 90–180 MHz. The increased bandwidth in these cases necessitated the use of the modified MS polynomial discussed above.

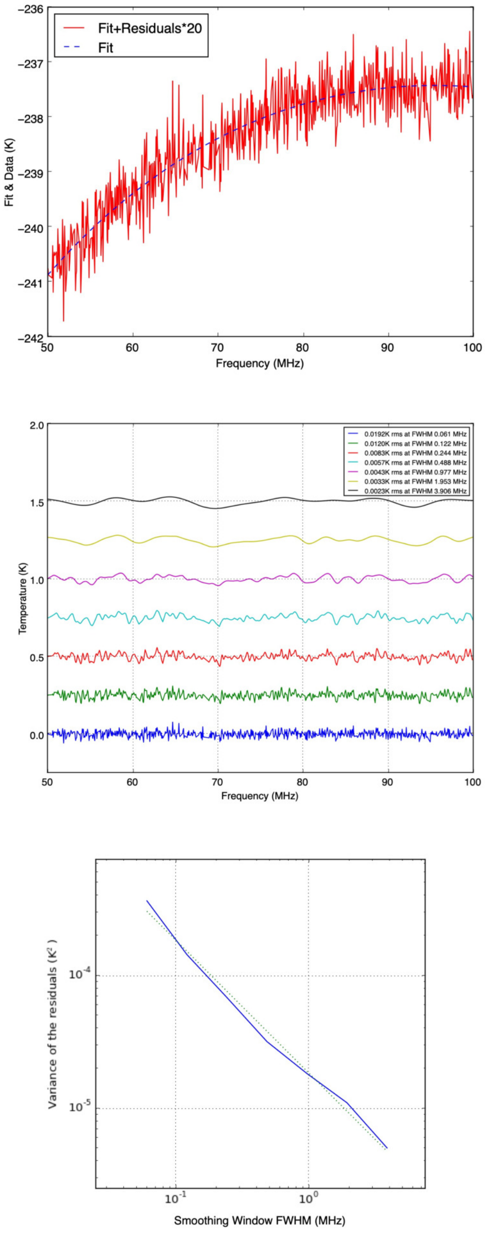


Figure 3.10: The result of modeling SARAS 3 systematics in the CD band 50–100 MHz. The sum of measurement data acquired with precision electrical open and short terminations at the antenna terminals is fitted using a maximally smooth polynomial form. The top panel shows the measurement data and fit together; the residuals are magnified by factor 20 for clarity. The middle panel shows the fitting residuals smoothed using kernels of increasing fwhm. The data with native resolution of 61 kHz is the lowest trace and spectra smoothed progressively to larger fwhm are shown above that with offsets of 0.25 K; traces are magnified by factors that keep the apparent rms the same on all smoothing. The legend in this middle panel lists the rms at different spectral resolutions. The continuous line in the bottom panel shows the run of variance in the residuals versus spectral resolution; the expected rate of fall in noise with smoothing is indicated by the dotted line.

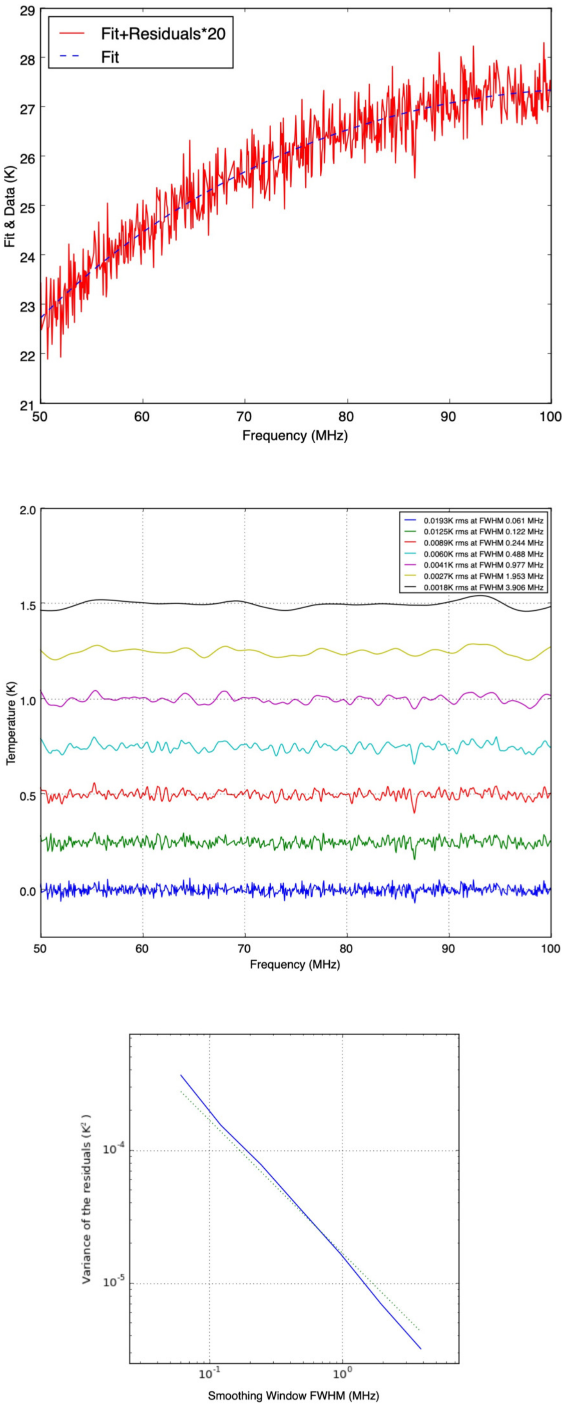


Figure 3.11: Modelling measurement data in the 50–100 MHz CD band. Here the difference of data acquired with open and short terminations at the antenna terminals is modeled. The three panels depict analyses same as that in the previous figure.

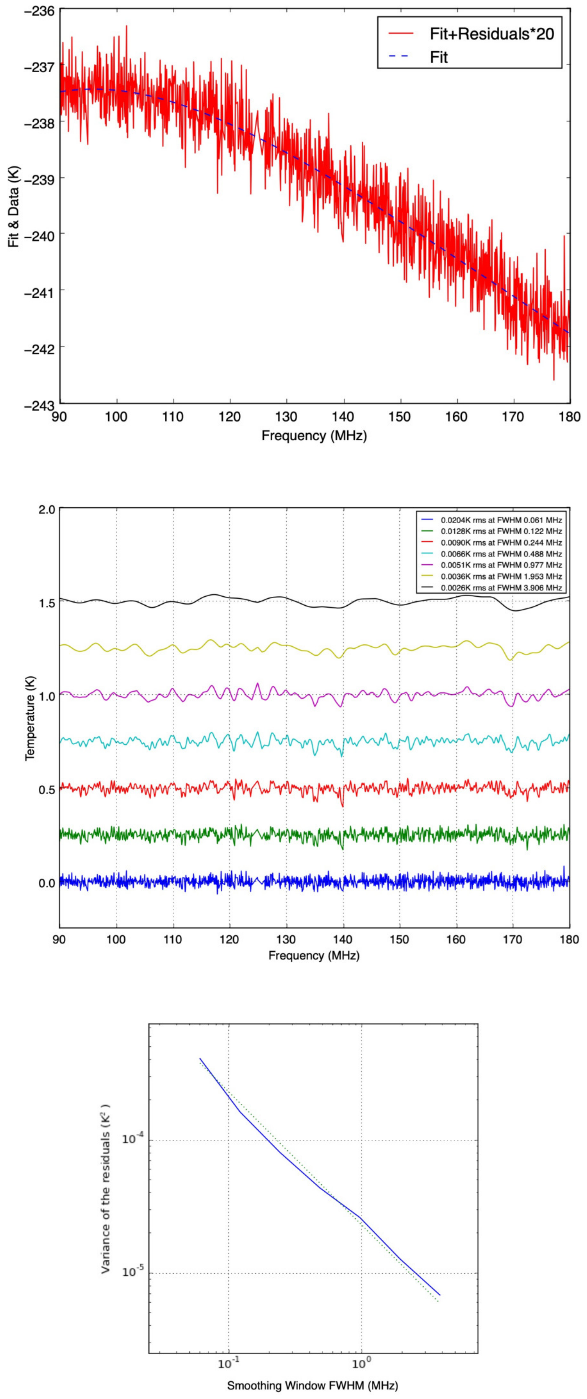


Figure 3.12: Modeling laboratory measurement data in the EoR band 90–180 MHz. In this figure, the sum of calibrated spectra acquired with open and short terminations at the antenna terminals is modeled. The three panels depict analyses same as that in the previous figure; however, for the wider EoR band the data was modeled using the modified form of the maximally smooth function described in the text, which allows for one zero crossing in higher order derivatives.

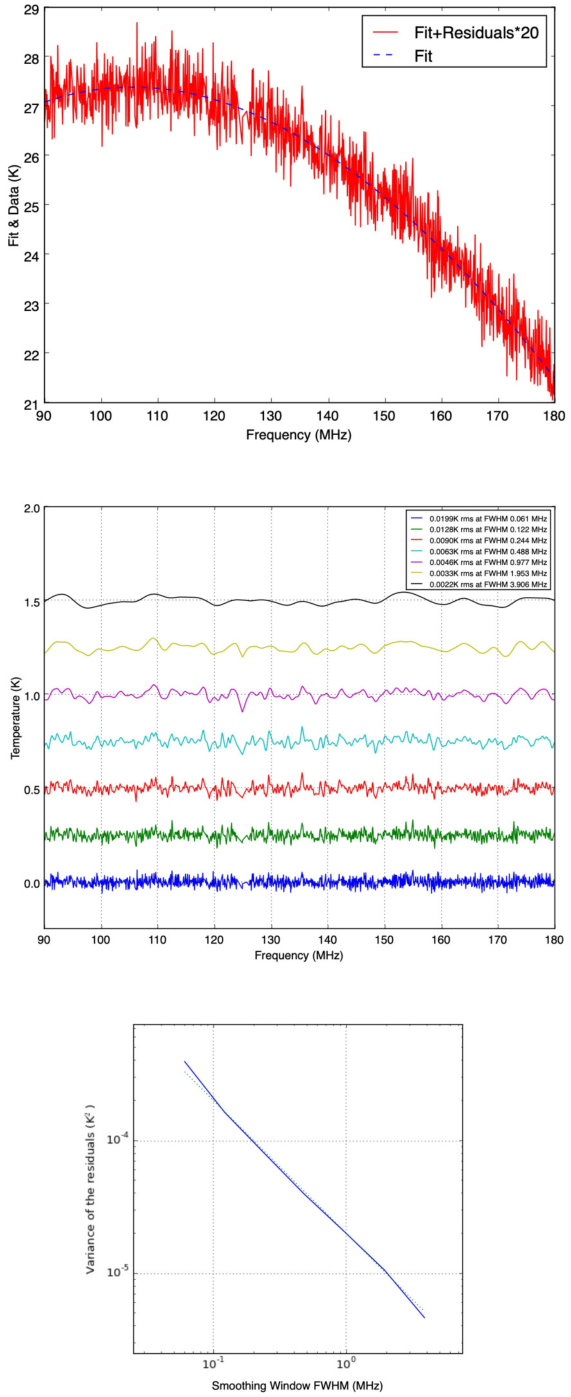


Figure 3.13: Modeling for systematics in the EoR band 90–180 MHz. In this figure, data that is the difference of those acquired with open and short terminations at the antenna terminals is examined. Here also, as in the previous figure, the fitting function is a modified maximally smooth function. The three panels depict analyses same as that in the previous figure.

In order to evaluate the internal systematics in a setup in which their characteristics would be similar to that while observing with the SARAS 3 antenna (Raghunathan et al., 2020), we examine the measurement data in a third case where the antenna is replaced with a circuit simulator: the antenna simulator discussed above. The resistive component of the antenna simulator is at ambient temperature, hence the effective antenna temperature in this case would be the product of the ambient temperature and a "reflection efficiency" for the simulator that is related to the reflection coefficient $S11$ of the 1-port network. Additionally, the measurement data would be expected to reveal any systematics that result from system temperature components suffering internal reflections at the antenna terminals. Results that follow from a maximally smooth fit to the measurement data acquired in this configuration are shown in Fig. 3.14. Once again, due to the subtraction of the reference temperature from the antenna temperature in the computation of the difference measurement, the values in the y-axis are negative.

In all the cases considered above, the fits of maximally smooth functions to the measurement data are good. There are no obvious systematics in the residuals, whose rms decreases with smoothing to within a few mK. This rms decreases with smoothing without any indication that systematics would limit the sensitivity to above a mK. Indeed the maximum deviation of the rms from that expected for Gaussian noise suggests that any residual systematic, that cannot be modeled with MS functions, has rms less than 1.2 mK. The laboratory test measurements demonstrate the receiver fidelity and qualifies the receiver for experiments aiming to detect CD/EoR structures that are distinct from maximally smooth functions and with signal amplitudes above a mK.

3.9 Summary of receiver tests

In this chapter, we have presented the design philosophy, design, and performance of the SARAS 3 radiometer receiver and evaluated the capability of the receiver to detect the global 21–cm cosmological signal from cosmic dawn and reionisation.

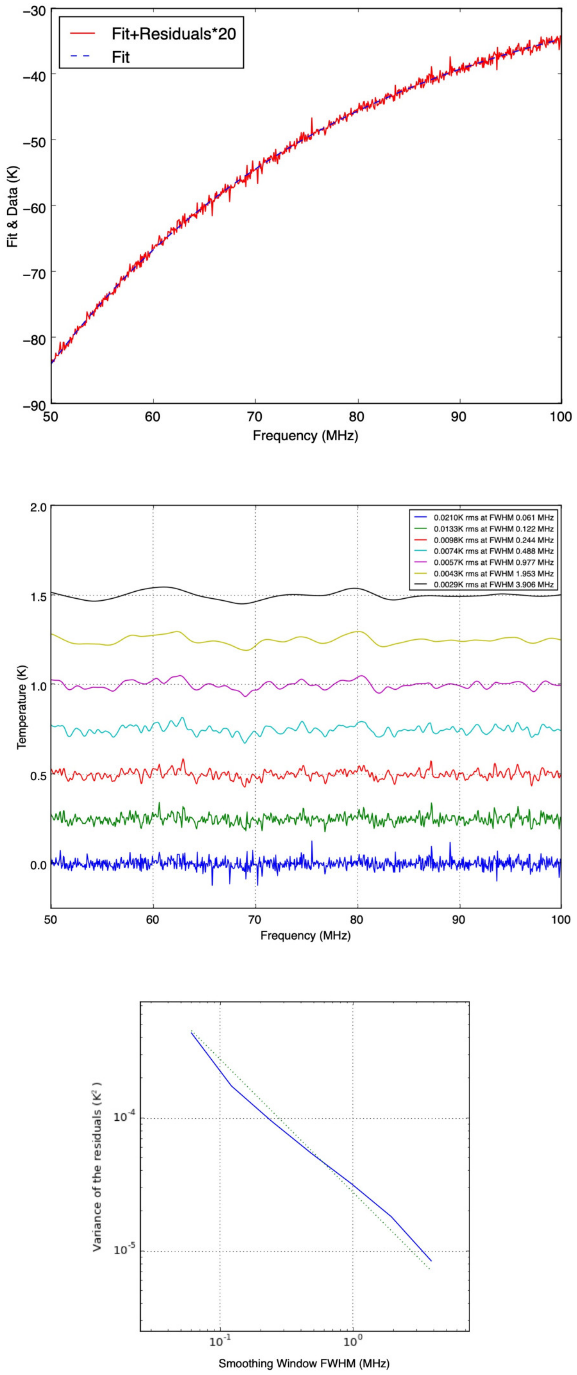


Figure 3.14: Modeling measurement data acquired with a 1-port antenna circuit simulator replacing the antenna. The three panels depict analyses same as that done in the previous figure.

We have discussed the new features in this receiver system that represent an improvement from earlier versions of the experiment. The receiver has been designed in such a manner so that inevitable systematic features—additive and multiplicative—that survive standard calibration would be maximally smooth and hence separable from CD/EoR signals. Double differencing is employed, switching the receiver between the antenna and a reference and phase switching to cancel additive spurious systematics entering the signal path. Optical isolation is provided and a digital correlation spectrometer is used. The receiver is built to be compact and located at the terminals of the antenna.

The signal path in the receiver has been analyzed in detail, leading to derivation of a measurement equation that includes spectral contributions from multi-order reflections occurring between the receiver and the antenna, within the signal path. We have also presented estimates of the sensitivity of the system taking into account the different system temperature components and the measurement equation, arising from the double differencing and the proposed switching scheme for calibration.

Finally, we report the qualification tests carried out in the laboratory to experimentally confirm that any systematics present in the data acquired with this receiver will not hinder a detection of 21–cm signals from CD/EoR. Using data acquired, with the antenna successively replaced with precision electrical terminations and an antenna simulator network, we have demonstrated that the receiver system has no residual additive systematics above 1 mK. Thus cosmological signals received by the SARAS 3 antenna and processed by the receiver would appear in measurement data without suffering confusion above 1 mK.

Within standard cosmology, star and galaxy formation models predict a range of plausible 21-cm profiles whose detection is the goal of CD/EoR receivers. We have taken the atlas of theoretically motivated signals (Cohen et al., 2017), added realistic models for the foreground (Sathyanarayana Rao et al., 2017a), and fitted out maximally smooth polynomials. The RMS residuals well exceed 1 mK, which is the sensitivity level to which the SARAS 3 receiver has been qualified in the laboratory tests. Therefore, the SARAS 3 receiver design and implementation is adequate for the detection of 21-cm signals predicted in standard models for cosmic dawn and reionisation.

3.10 Results from SARAS 3 observations

This section summarises the results from analysis of SARAS 3 data obtained from field deployments. Apart from the receiver, the SARAS 3 system consists of a few others subsystems. As shown in Fig.3.1, the antenna is a monopole cone-disk floating on water (Raghunathan et al., 2020). The front-end of the receiver unit is placed below the disk of the antenna. The back-end consists of the rest of the receiver electronics, along with the digital systems to collect data. The digital correlator and data storage has been described in Girish et al. (2020). Altogether, the SARAS 3 system is a portable radio spectrometer that can be transported and deployed at remote radio quiet locations.

During 2020 January 27-30, SARAS 3 system was deployed on Dandiganahalli Lake (latitude 13.50896667° N, longitude 77.65981667° E) and during 2020 March 10-20, on Sharavati backwaters (latitude 13.992163° N, longitude 74.876027° E), both in Southern India. These locations were chosen based on water salinity and water depth, and for relatively low RFI conditions. The collected spectral data were excised for RFI, and based on the analysis (details in Singh et al. (2022)), the best-fit EDGES absorption profile is rejected with 95.3% confidence. This rejection rules out the absorption profile having an astrophysical origin.

Chapter 4

SITARA - A short spacing global 21 cm experiment

> *Simplicity is the final achievement. After one has played a vast quantity of notes and more notes, it is simplicity that emerges as the crowning reward of art.*
>
> - Frédéric Chopin

This chapter is a reproduction of Jishnu N. Thekkeppattu, Benjamin McKinley, Cathryn M. Trott, Jake Jones, Daniel C. X. Ung, *System design and calibration of SITARA - a global 21-cm short spacing interferometer prototype*, 2022, Publications of the Astronomical Society of Australia, https://doi.org/10.1017/pasa.2022.13. Minor alterations have been made to the original material for consistency within this thesis. The reader may encounter some repetition of material in the introductory sections.

4.1 Abstract

Global 21–cm experiments require exquisitely precise calibration of the measurement systems in order to separate the weak 21–cm signal from Galactic and extragalactic foregrounds as well as instrumental systematics. Hitherto, experiments aiming to make this measurement have concentrated on measuring this

signal using the single element approach. However, an alternative approach based on interferometers with short baselines is expected to alleviate some of the difficulties associated with a single element approach such as precision modelling of the receiver noise spectrum. Short spacing Interferometer Telescope probing cosmic dAwn and epoch of ReionisAtion (SITARA) is a short spacing interferometer deployed at the Murchison Radio-astronomy Observatory (MRO). It is intended to be a prototype or a test-bed to gain a better understanding of interferometry at short baselines, and develop tools to perform observations and data calibration. In this chapter, we provide a description of the SITARA system and its deployment at the MRO, and discuss strategies developed to calibrate SITARA. We touch upon certain systematics seen in SITARA data and their modelling. We find that SITARA has sensitivity to all sky signals as well as non-negligible noise coupling between the antennas. It is seen that the coupled receiver noise has a spectral shape that broadly matches the theoretical calculations reported in prior works. We also find that when appropriately modified antenna radiation patterns taking into account the effects of mutual coupling are used, the measured data are well modelled by the standard visibility equation.

4.2 Introduction

The period in cosmological history when the first stars ionised the Universe remains one of the least constrained epochs in the concurrent cosmological models. This period known as the cosmic dawn and epoch of reionisation (CD/EoR), despite being a critical epoch in our cosmological models, lacks observational constraints. It has been recognized that the redshifted signal from the 21 cm hyperfine transition of neutral hydrogen can be an effective tracer of baryonic evolution during this period (Varshalovich & Khersonskii, 1977; Pritchard & Loeb, 2010). The coupling of this transition's spin temperature to radiation temperatures (CMB as well as any excess background radiation) via scattering, and matter kinetic temperature via collisions as well as the Wouthuysen-Field effect (Wouthuysen, 1952; Field, 1958) can give rise to absorption and emission features in the mean background spectrum. Owing to cosmological expansion, the rest-frame frequency of 1420 MHz of this transition gets redshifted to 40–230 MHz. There is considerable effort being put in to measure the spatial power spectrum of this signal, with several radio telescopes such as MWA (Tingay et al., 2013; Trott et al., 2020), LOFAR (van Haarlem et al., 2013; Mertens et al., 2020), GMRT (Swarup et al., 1991; Paciga et al., 2011), HERA (DeBoer et al., 2017; The HERA Collaboration et al., 2021), 21 centimeter Array (Peterson et al., 2004), OVRO-LWA (Hallinan et al., 2015; Garsden et al., 2021) currently operating with precision measurements of 21 cm power spectrum as one of the key science goals. The sky-averaged or global component has also been recognized as a powerful probe of this epoch (Shaver et al., 1999). Since this uniform component is an average of the angular variations, a single antenna of low angular resolution is sufficient to detect the signal. Given that the sky-averaged component has a strength of $\sim 10 - 100$ mK against Galactic and extragalactic foregrounds with $10^2 - 10^4$ K brightness temperatures, an unambiguous detection of this signal requires well calibrated instruments. Most of the experiments aiming at a measurement of the global signal, such as EDGES (Bowman et al., 2018), SARAS (Singh et al., 2018a; Nambissan T. et al., 2021), BIGHORNS (Sokolowski et al., 2015), PRIZM (Philip et al., 2019), LEDA (Price et al., 2018), REACH, MIST, HYPEREION (Patra et.al., in prep) use single well-calibrated antennas as the electromagnetic sensor. However, these experiments require precision calibration of the systems to mitigate the effects of the antenna transfer function, antenna radiation pattern variations with frequency (beam chromaticity) as well as the receiver bandpass and spectrum of the receiver noise.

4.3 Background and motivation

As an alternative to single antenna based measurement of the 21 cm signal, interferometers with closely spaced antennas have been proposed. The motivation for interferometers stems from the fact that individual receiver noise contributions, being uncorrelated, average to zero upon cross-correlation. Conventional wisdom based on a Fourier perspective is that an interferometer does not respond to a uniform sky signal. However, this argument fails at the limit when the antennas are brought to close proximity.

A radio interferometer measures the spatial coherence function. For wavelength λ corresponding to a frequency ν and for a baseline vector $\vec{b}$, the coherence is given by Eq.4.1 (see for e.g. Clark (1999)),

$$V = \frac{1}{4\pi} \int_{4\pi} T_{sky} A_a e^{-2\pi i (\frac{\vec{b}.\vec{r}}{\lambda})} d\Omega \tag{4.1}$$

where c is the speed of light, T_{sky} is the sky brightness temperature as a function of spatial coordinates, A_a is the antenna radiation pattern, assumed to be identical for both antennas. From this, we can compute the expected auto-correlation powers for the individual antennas as well as their cross power by appropriately setting $\vec{b}$. Setting $|b| = 0$, the auto-correlation powers may be recovered. Our interest is when $|b| \sim \lambda$, where Eq.4.1 yields a non-negligible non-zero value.

Indeed it is shown in Vedantham et al. (2015) using a spherical harmonic expansion (instead of a Fourier expansion) that the interferometer response to a global signal has a characteristic *sinc* shape as a function of baseline length. There also appears to be some controversy regarding the nature of this response. While Presley et al. (2015) argue that the response is due to the primary radiation pattern of the antennas, Singh et al. (2015) demonstrate using simulations with isotropic antennas that the response is an inherent property of the wavefield as opposed to being purely an instrumental response. Also in Singh et al. (2015), simulations of the coherence function as a function of baseline length $|\vec{b}|$ for various types of antennas and orientations are shown. However, these studies ignore effects such as antenna mutual coupling, noise coupling between antennas, ground, and foregrounds.

While Eq.4.1 provides a convenient starting point for short-spacing interferometry, it assumes identical radiation patterns for the antennas - a condition that is not necessarily satisfied due to mutual coupling when antennas are closely spaced. An interesting theoretical discussion on the effects of mutual coupling on the response of a short-spacing interferometer - from the perspective of the incomplete nature of Eq.4.1 - is given in Venumadhav et al. (2016). Specifically, the “shadowing” of antennas when closely spaced and mutually coupled is not considered by Eq.4.1. Therefore, Venumadhav et al. (2016) show that effects such as scattering and shadowing have to be included. It is argued in Venumadhav et al. (2016) that cross-talk between the antennas forming a short-spacing interferometer is crucial to having a response to the sky monopole, as shadowing effects obstruct the view of antennas to regions of sky that dominate the nonzero response. In the same work, it is shown that the sensitivity of closely packed antenna arrays to a sky monopole maximises in the regime where antennas couple by non-radiative fields. However, cross-talk can also result in noise coupling between the antennas, thereby invalidating the assumption of negligible noise bias in cross-correlations.

Though there have been theoretical and simulation studies on the short baseline response of an interferometer to an all sky component of the sky, only a few experiments have attempted a measurement. ZEBRA (Raghunathan et al., 2011) used a resistive spatial beamsplitter made out of discrete resistors to enhance the short spacing response (Mahesh et al., 2015). While a beamsplitter enhances the coherence of a uniform sky signal at short baselines, modelling emission from the splitter, which appears as an additive term in the spectrum, presents a formidable challenge. In McKinley et al. (2020), an alternative approach based on the Engineering Development Array (EDA-2) deployed at the Murchison Radio-astronomy Observatory (MRO) has been employed to evaluate the potential of this idea. However, the presence of a large number of antennas in close proximity introduces complicated mutual coupling responses between the antennas, the effects of which are in general hard to characterise in-situ. To the best of the authors’ knowledge no such study has been undertaken in literature wherein the nature of short baseline interferometer response to an all sky signal has been experimentally investigated with a dedicated experiment.

In this context, it was recognized that a dedicated broadband interferometer to study the effects of mutual coupling, noise coupling and foregrounds, and their effects in an interferometer for probing the global 21 cm signal is required. Short spacing Interferometer Telescope probing cosmic dAwn and epoch of ReionisAtion - SITARA is the first in a series of experiments aiming at measurement and validation of the short spacing interferometer response with an ultimate aim of having a dedicated interferometric array with multiple short baselines, named All-Sky SignAl Short-Spacing INterferometer (ASSASSIN). As a first step in this direction we built and deployed a prototype two-element broadband interferometer at the MRO, to measure the response of an interferometer to the radio sky at short baselines ($\sim \lambda$). This version of the instrument is envisaged to be a test-bed to develop techniques for system design, calibration and data analysis at short baselines and to understand potential systematics. Experience gained from this version will feed into more advanced experiments. In this chapter we outline the SITARA system concept, deployment and data calibration strategies with a particular emphasis on the calibration of short-baseline interferometric data.

In closely spaced interferometers such as SITARA, cross-talk between the antennas becomes non-negligible. The term cross-talk can imply a wide range of phenomena; however in this chapter we use cross-talk as a blanket term for any coupling of signals from one arm of the interferometer to the other. Cross-talk can occur at multiple points in the signal chain. However for short-spacing interferometers, the dominant form of cross-talk is expected to be due to the mutual impedance between the antennas. There are two major effects expected due to such cross-talk.

1. The receiver noise from an antenna and associated electronics leaks into the other antenna. This results in a non-zero cross-correlation between receiver noises. This appears as a constant excess receiver noise in cross-correlations.

2. Similarly, sky signals will also get coupled between the antennas. This results in the autocorrelations and cross-correlations deviating from the idealised simulations using the visibility equation, even if accurate antenna radiation patterns, such as embedded element patterns (EEPs), are used.

We find that both effects are seen in SITARA data. We also find that ignoring the effects of cross-talk leads to poor modelling of data and therefore a model for cross-talk is presented that captures the complexity of the data.

4.3.1 Notations and conventions

In this chapter, we use boldface letters to denote matrices. Vectors are denoted with an arrow over the symbol, such as $\vec{b}$. The forward Fourier transform carries a negative sign. The imaginary number is denoted by i. We use T to denote temperature quantities that are expressed in kelvins (K), and powers that have arbitrary units due to scaling are denoted as P. Visibilities are represented by V and electric fields with $\vec{E}$. Frequency is denoted by ν and voltage by e; it may be noted the difference between the use of e as a voltage and as the exponential factor will be self-evident from the context. We use the term noise temperature to refer to the mean value of noise power, calibrated into units of kelvin.

4.4 SITARA System overview

Broadly, the SITARA system consists of two antennas kept in close proximity, a “fieldbox” performing initial analog signal conditioning and a back-end performing further analog signal conditioning, digitisation and correlation. No form of hardware calibration such as noise diodes is employed. The first prototype is kept simple so as to study systematics that have to be considered for more advanced designs. A block-diagram of the SITARA system is given in Fig.4.1

In order to enable rapid development, prototyping and deployment, it was decided to use system components with good pedigree, especially in the harsh field conditions of MRO. The two antennas used are standard MWA active dipoles kept over a metallic groundplane, each one consisting of two bow-ties forming a crossed dipole and the associated low noise amplifiers (LNA). The ground plane has a diameter of 35 m with 5 cm square grids and was previously used for the Engineering Development Array (EDA) - 1 (Wayth et al., 2017). Each bow-tie dipole has an end to end length of 74 cm, and a height of 40 cm. Each crossed dipole antenna, formed by two orthogonal dipoles, is held above the ground plane with four 10 cm dielectric stand-offs; one on each arm of the antenna. However, only one polarisation of each antenna is utilised for this experiment. Further details regarding the mechanical structure of the antenna can be found in Reeve

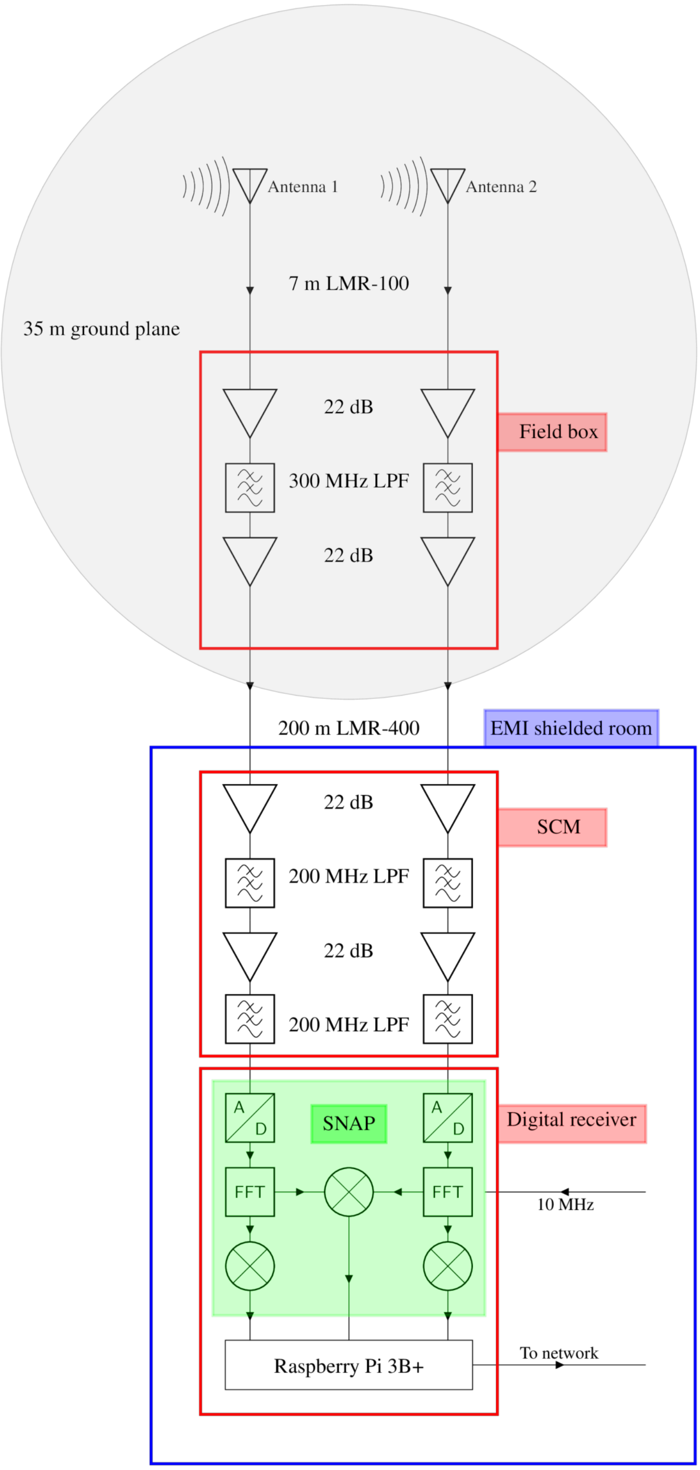

Antenna 1
Antenna 2
7 m LMR-100
35 m ground plane
22 dB
Field box
300 MHz LPF
22 dB
200 m LMR-400
EMI shielded room
22 dB
SCM
200 MHz LPF
22 dB
200 MHz LPF
A
D
SNAP
A
D
Digital receiver
FFT
FFT
10 MHz
Raspberry Pi 3B+
To network

Figure 4.1: A high level block diagram of SITARA; auxiliary details such as power supplies as well as attenuators used for impedance matching between various modules are not shown. The multiplication units shown in the digital receiver perform conjugate multiplication.

(2017). The central hub of each antenna houses a dual LNA assembly based on Broadcom ATF-54143 pHEMT, with a gain of about 20 dB per polarisation. The LNA circuit also performs differential to single ended conversion, such that the balanced antenna (dipole) can be connected to an unbalanced transmission line (coaxial cable). The signals from the antennas are transported over 7 m of KSR-100 coaxial cables (specification conforming to LMR-100, impedance 50 Ω) to a fieldbox that contains modular amplifiers and filters. Through the same coaxial cables the DC power for the LNAs is supplied by the fieldbox via bias-tees.

The fieldbox amplifies the signals further to reduce effects of a long cable on the net system temperature. To reduce effects of out-of-band radio frequency interference (RFI) on signal chain linearity (as the amplifiers used are broadband compared to the required 250 MHz bandwidth) a relatively broad 300 MHz low-pass filtering is performed. The amplified and filtered signals are transported over 200 m of coaxial cables (specification conforming to LMR-400, impedance 50 Ω) to the back-end electronics housed inside a shielded room, colloquially called the Telstra hut. Power to the fieldbox is delivered over a pair of dedicated power lines running 12V DC. This power is derived from a power supply housed in the Telstra hut and passed through dedicated filters to reduce electromagnetic interference (EMI) and meet the radiated EMI specification requirements of the MRO.

The signals arriving at the Telstra hut end are further amplified to ensure a sufficiently high signal to noise ratio to overcome the quantisation noise of the analog to digital converters (ADC) in the correlator. Two stages of 200 MHz filtering are utilised as an anti-aliasing filter to limit the bandpass to 250 MHz. Altogether, the analog section has a net gain of about 70 dB inclusive of the cables and active antenna LNA. A SNAP board (Hickish et al., 2016) sampling at 500 MSPS is used as the digitiser and correlator. Though SNAP has 12 inputs, only 6 inputs can be utilised at the sampling rate of 500 MSPS. However this is not a constraint, as only two of the six available inputs are used for the current experiment. To reduce the amount of correlated board noise, two physically different ADCs out of the three on board are used to digitise the data.

The 10 MHz clock to SNAP is provided by a White Rabbit[1] unit referenced to a master hydrogen maser. The data are channelised into 8192 channels and correlated in the SNAP to form auto and cross-correlation spectral products. As the sampling is real, only 4096 channels are useful. Thus, each of the resulting correlated spectra has 4096 channels spanning a frequency range of 0-250 MHz, with a spectral resolution of $\approx$ 61 kHz. The amplifiers in the signal chain have high pass filtering at 50 MHz and the LNAs in the active antennas have high pass filtering at 70 MHz. Therefore, frequencies between 50 and 70 MHz have reduced sensitivity. The anti-aliasing filters in the analog signal chain reduce sensitivity above 200 MHz. Thus, owing to the filtering introduced by the active antennas and the analog signal processing, only frequencies between 50 and 200 MHz have sensitivity to sky signals.

When operated at 500 MSPS, the SNAP ADCs perform interleaved sampling. Small offset, gain and phase (OGP) mismatches between the ADC cores introduce spurious tones in the spectra at sub-harmonics of the clock signal. While these tones themselves are not deleterious, as they can be flagged during data analysis, the interleaving process has been found to cause intermodulation products in the measured spectra. In principle it is possible to measure and correct for the mismatches using bin-centred tones at each frequency, however any such correction would have to be performed in the signal processing within FPGA. As the complexity required to introduce such tones without causing conducted EMI outweighs any advantage obtained, we do not perform it.

A Raspberry Pi 3B+ (henceforth RPi) with 32 GB microSD storage, connected over GPIO controls the SNAP. The same connection is utilised to transfer correlated data to the RPi as well as provide DC power to it. An acquisition code in the RPi acquires auto and cross data from SNAP and writes them out in `miriad` format (Sault et al., 1995) into the RPi microSD card, at a time cadence of about 3 seconds. These data are also appropriately time (UTC) and local sidereal time (LST) stamped for subsequent analysis. To reduce the EMI generated by the SNAP digital clocks from getting radiated and conducted via power lines, a dedicated switching mode power supply (SMPS) along with input EMI filters is enclosed within the correlator rack chassis. To further reduce EMI from the correlator, an off-the-shelf media converter is enclosed that converts the electrical ethernet connection from the RPi into an optical fibre connection. Thus, this unit forms a low-EMI networked correlator that can be accessed over internet.

---

[1] https://ohwr.org/projects/white-rabbit

Efforts have been made to keep the analog signal chains symmetric in their amplitude and phase responses, nonetheless there could be an excess delay between the arms due to component tolerances. The effect of an excess delay is to decorrelate the signals, however even for an excess path length of the order of a few metres, decorrelation is expected to be minimal as the signals are fine channelised to a resolution of $\sim$ 61 kHz and correlated.

4.5 Deployment and observations

SITARA was deployed at the MRO in March 2021 with first light achieved on March 10th. The antenna spacing has been chosen as 1 m between the dipole centres, with the dipoles oriented parallel to each other - the so-called parallel configuration in Singh et al. (2015) - along local East-West. In this configuration, the baseline is oriented along the local E-W, while the specific dipoles used have their nulls oriented along N-S. However, after one night of observations an amplifier in one of the signal chains failed on March 11th and had to be removed. To preserve symmetry, the corresponding amplifier in the other signal chain was also removed and brought back for further investigation of the failure. This resulted in a gain reduction of about 22 dB and the receiver temperature increased by a factor of about 2. Though the data collected before amplifier failure on March 10th have low levels of RFI, they do not have sufficient time coverage so as to enable calibration and hence are not used in this chapter. Thus, in this chapter a 24 hour span of data collected after the amplifier removal are presented. A few photographs of SITARA as deployed at the MRO are shown in Fig.4.2.

Fig.4.3 is a time-frequency plot of the cross-correlation data collected on May 17th-18th, 2021. Also shown are the averages of the powers along time and fre-

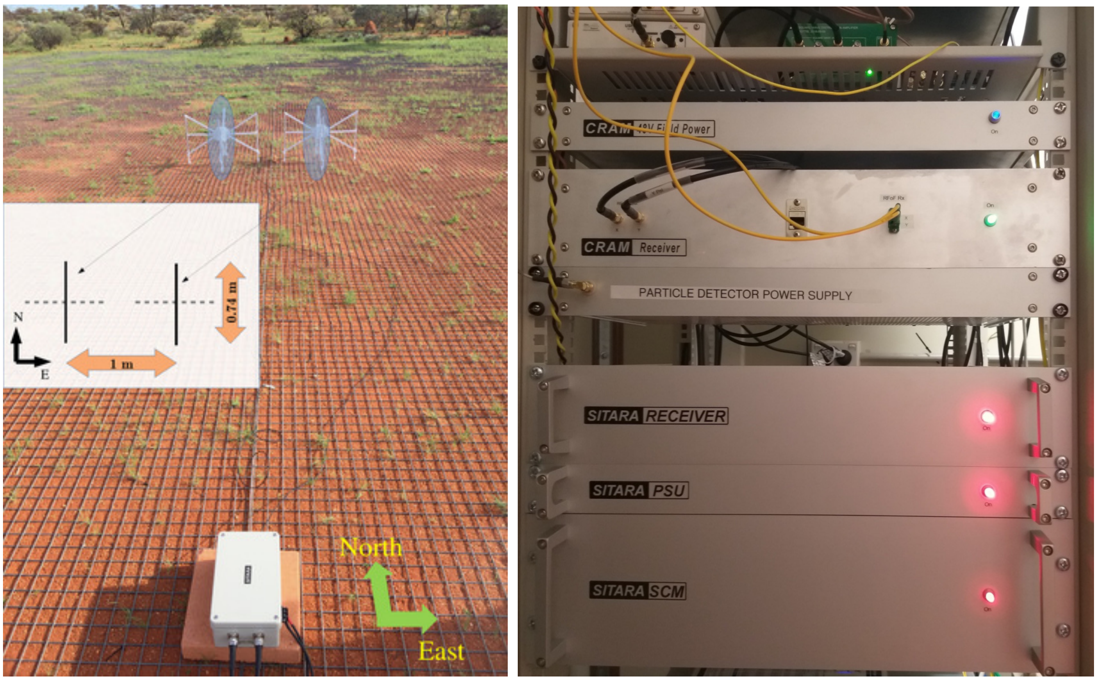


Figure 4.2: SITARA system as deployed at MRO. The left photograph shows SITARA antennas and fieldbox; the cables have since been tied to the ground plane. The specific dipoles used in this experiment are highlighted in blue ellipses. The inset shows the antenna orientation and relevant dimensions where the inactive dipoles have been greyed out. The right photograph shows SITARA back-end electronics inside the Telstra hut. The receiver box houses the SNAP and RPi as well as media converters for networking. Signal conditioning module (SCM) contains the amplifiers and filters to perform analog processing before digitization and correlation.

quency axes. The LNAs in the active antennas have a lower cutoff of 70 MHz, and the analog anti-aliasing filters in the analog signal chains low-pass filter the data above 200 MHz. The effects of both filters are visible in the data. Also visible are tones from ADC clocking at 62.5 MHz, 125 MHz and 187.5 MHz. These tones are of narrow-band nature and are easily flagged, however owing to the reduced analog gain due to amplifier failure there could be intermodulation products due to these tones mixing with the analog signal at frequencies close to these tones. The predominant sources of RFI at the MRO are satellites with downlink frequencies around 137 MHz such as Orbcomm, NOAA-APT and METEOR-LRPT weather transmissions. Amateur radio satellite downlinks around 145 MHz and aircraft communications below 130 MHz are also seen in the data. Signals from FM transmitters appear sporadically, perhaps reflected by overhead flights, meteor trails or through some VHF propagation modes such as tropospheric ducting or sporadic E-layer propagation (Hitney et al., 1985; Jessop, 1983). Details about the RFI conditions at the MRO can be found in Offringa et al. (2015) and Sokolowski et al. (2015). A recent study of RFI at MRO in the broadcast FM bands is reported in Tingay et al. (2020).

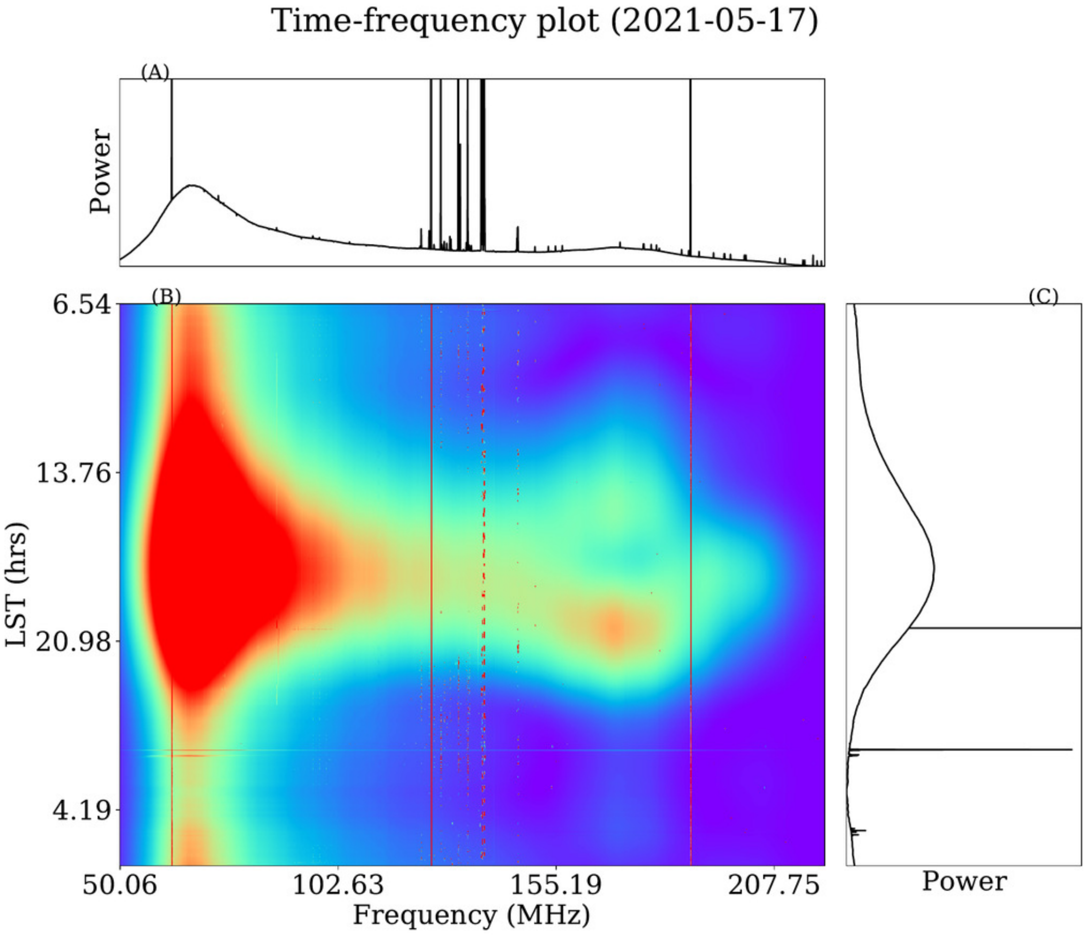


Figure 4.3: Time-frequency (waterfall) plot of the cross-correlation data collected on May 17th-18th, 2021. Panel B is the time-frequency plot of the magnitude of the complex visibilities. Panel A is the average spectrum and panel C shows the power as a function of LST for a frequency of 70 MHz. The data are unflagged and uncalibrated. The waterfall plot shows the sky drifting through SITARA beam; the peak occurs when the Galactic plane is at the local zenith. On closer inspection, the data shown in this figure are seen to contain Solar bursts between 1-2 hours LST.

Data are continuously collected and the timestamped data are accessed for analysis at regular intervals. With more than 2500 hours of operation and data collection, no major glitch has been noticed. In Fig.4.4, the measured uncalibrated powers as a function of LST in a single 61 kHz frequency channel centred at 111.05 MHz are shown (for auto-correlation for antenna 1, as well the cross-correlation between the two antennas). The data used are after the amplifier failure. As expected, data collected with the system over a span of few weeks show variations in the power levels with time, however we do not find any significant drift with time in the system performance. During the ongoing observational run, we had a few serendipitous high signal to noise ratio detections of solar bursts. Analysis of those bursts are beyond the scope of this chapter and will be reported elsewhere.

4.6 Data calibration and analysis

In this section, we describe the procedures adopted to calibrate and analyse data. The observations in `miriad` format are flagged with `pgflag` using the SumThreshold algorithm (Offringa et al., 2010). Further calibration and analysis of data are carried out in custom python codes with data read using the `aipy`[2] package.

This section is organised as follows. Before inspecting the data, we visit the antenna radiation patterns from FEE simulations in Sec.4.6.1, where we find that the individual antenna patterns cannot be treated as frequency invariant. In Sec.4.6.2, a simple model for measurements that considers cross-talk for (internal) receiver noise but not (external) sky signals is presented. We find that while this simplistic model is able to represent the variations in data, certain shortcomings are evident. The differences seen between the mock data and SITARA data are attributed to the cross-talk of sky signals between the antennas and an empirical model for it is introduced in Sec.4.6.3. This model brings the coupling of receiver noises and sky signals between the antennas under the same formalism. Interestingly, effects of sky signal cross-talk become evident in the data only at frequencies where the antenna patterns differ, as the individual auto-correlations are identical when the antenna patterns are identical. As a by-product of the model, we obtain the coupled receiver noise at all frequencies. Comparing the models with and without cross-talk, we find that the empirical cross-talk model

---

[2] https://github.com/HERA-Team/aipy

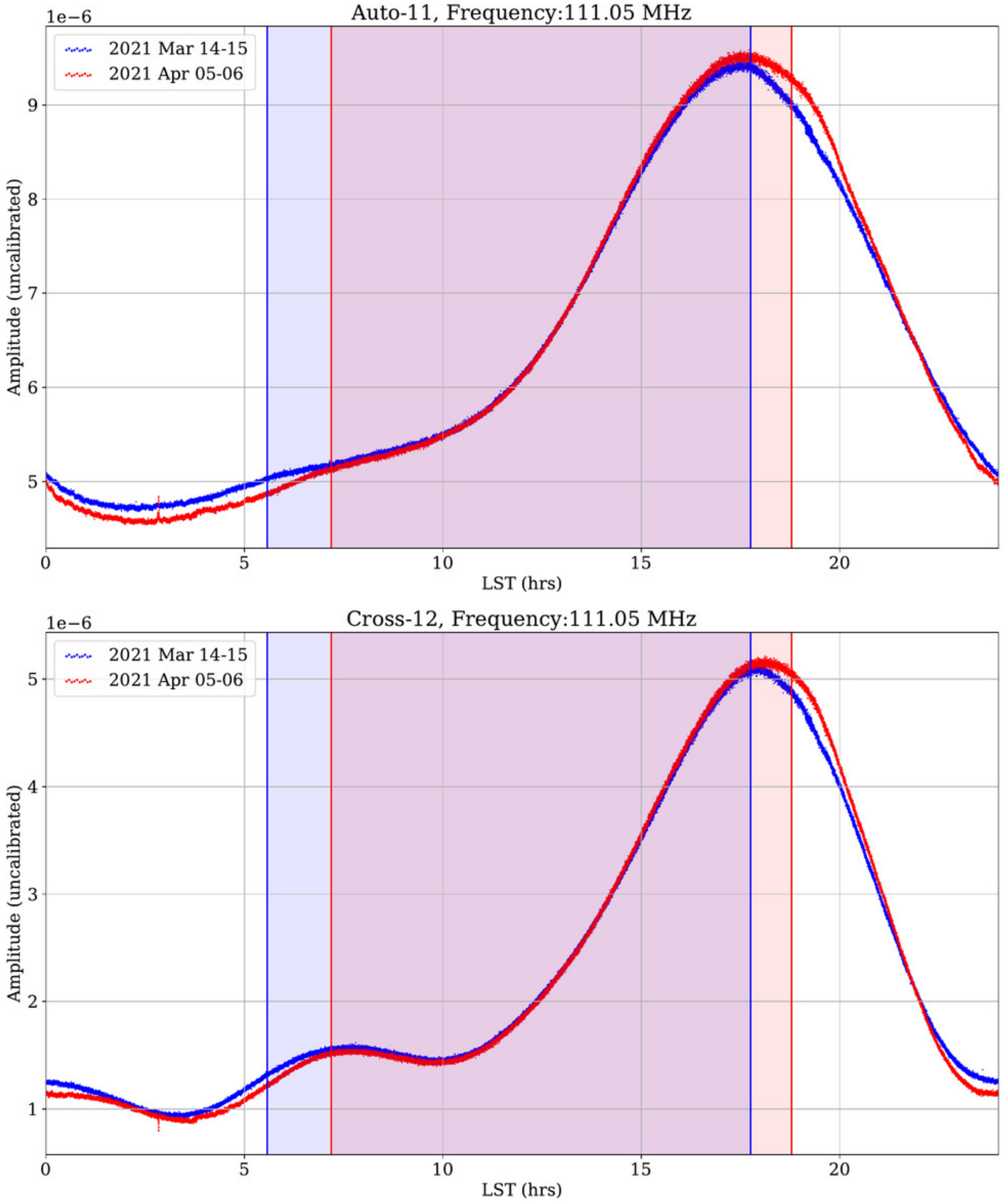


Figure 4.4: Variations in uncalibrated power with local sidereal time (LST) for data collected on Mar 14-15, 2021 and April 05-06, 2021. The top figure shows the power in a single frequency channel in antenna 1 auto-correlations and the bottom figure shows the magnitude of antenna 1-2 cross-correlations. The colored regions in the plots show the night time LSTs for the corresponding day.

captures the variations in the data as a function of LST accurately.

4.6.1 A prelude on antenna radiation patterns

The radiation patterns (also called beams) of isolated MWA dipoles over a large ground plane have a peak at local zenith. With closely spaced antennas, the effects of mutual coupling between elements cause the patterns to deviate from those of isolated dipoles. Even if the beams of the isolated antennas are achromatic, mutual coupling would induce some chromaticity when they are brought together for a short spacing experiment. Hence, if radiation patterns of isolated MWA dipoles (that are largely achromatic) are used in Eq.4.1 to simulate visibilities for SITARA, the measured visibilities would deviate significantly from the simulations. For frequencies where the baseline $|\vec{b}| > \lambda/2$, the antenna patterns are seen to vary rapidly with frequency with the peak shifting away from the zenith. Moreover, owing to the intrinsic symmetry, the peak shifts in opposite directions for each antenna and thus their overlapping beam solid angles vary as a function of frequency. Nonetheless, electromagnetic simulations model these and hence it is possible to use the simulated antenna patterns to compute visibilities as given in 4.6.2. For this work, we use antenna patterns simulated with FEKO [3]. The simulation is for the full structure of the SITARA antenna system, which consists of two MWA dipoles at a separation of 1 m. The dipoles are assumed to be placed over an infinite ground plane. Similar to the procedure adopted in Sokolowski et al. (2017), the antenna ports are loaded with lumped circuit models of the LNA. The simulations are generated at a frequency resolution of 1 MHz. This ensures that the simulation yields the patterns of each antenna in the presence of the other, including the effects of mutual coupling, and the resulting patterns are therefore embedded element patterns (EEP). In Fig.4.5, FEKO-simulated SITARA antenna radiation patterns at two representative frequencies of 90 MHz and 180 MHz are shown as cross-sectional plots. The patterns at 90 MHz are analogous to each other and are well approximated by an ideal dipole ($\cos^2$(ZA)) radiation pattern, while the patterns at 180 MHz are not identical to each other. A more insightful representation of the antenna patterns is given in Fig.4.6, which shows intensity maps in Mollweide projection

---

[3] https://www.altair.com/feko/

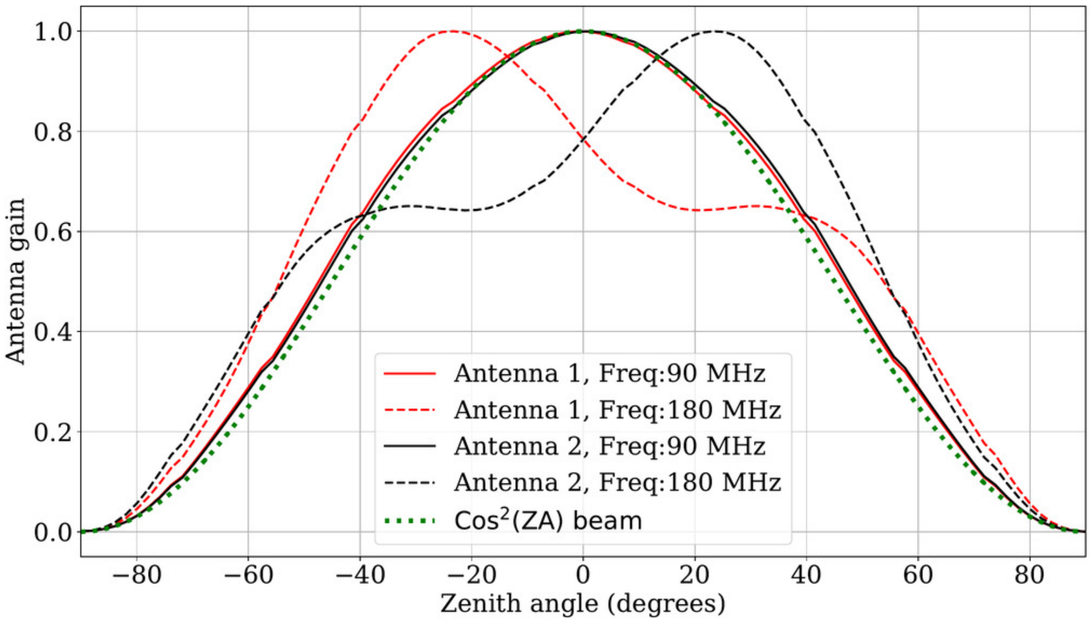


Figure 4.5: Simulated antenna radiation patterns (H-plane) as a function of zenith angle for two MWA dipoles spaced 1 m apart in parallel configuration. The patterns at 90 MHz are identical to each other and are well approximated by an ideal dipole $\cos^2$(ZA) pattern while the patterns at 180 MHz have shifted peaks away from zenith.

at 90 MHz and 180 MHz of the power patterns given by

$$|E_{j,\theta}(\theta,\phi)E^*_{k,\theta}(\theta,\phi) + E_{j,\phi}(\theta,\phi)E^*_{k,\phi}(\theta,\phi)| \tag{4.2}$$

where $E_{j,\theta}$ and $E_{j,\phi}$ are the two orthogonal components of the E-field patterns of antenna j; similarly $E_{k,\theta}$ and $E_{k,\phi}$ are the components of antenna k. In Eq.4.2, when $j = k$, the patterns are of individual antennas, while $j \neq k$ gives the cross-correlated beam. At 90 MHz, the individual patterns are similar to that of an isolated MWA antenna, while at 180 MHz, they deviate substantially from the pattern of an MWA dipole. Moreover, the antenna patterns have a mirror symmetry owing to the inherent symmetry of a two antenna system. For subsequent analysis, we use these EEPs to simulate the expected sky response.

4.6.2 Calibration ignoring sky signal cross-talk

We first attempt to model the measurements with a simple model that does not take into consideration the cross-talk of sky signals between the antennas. How-

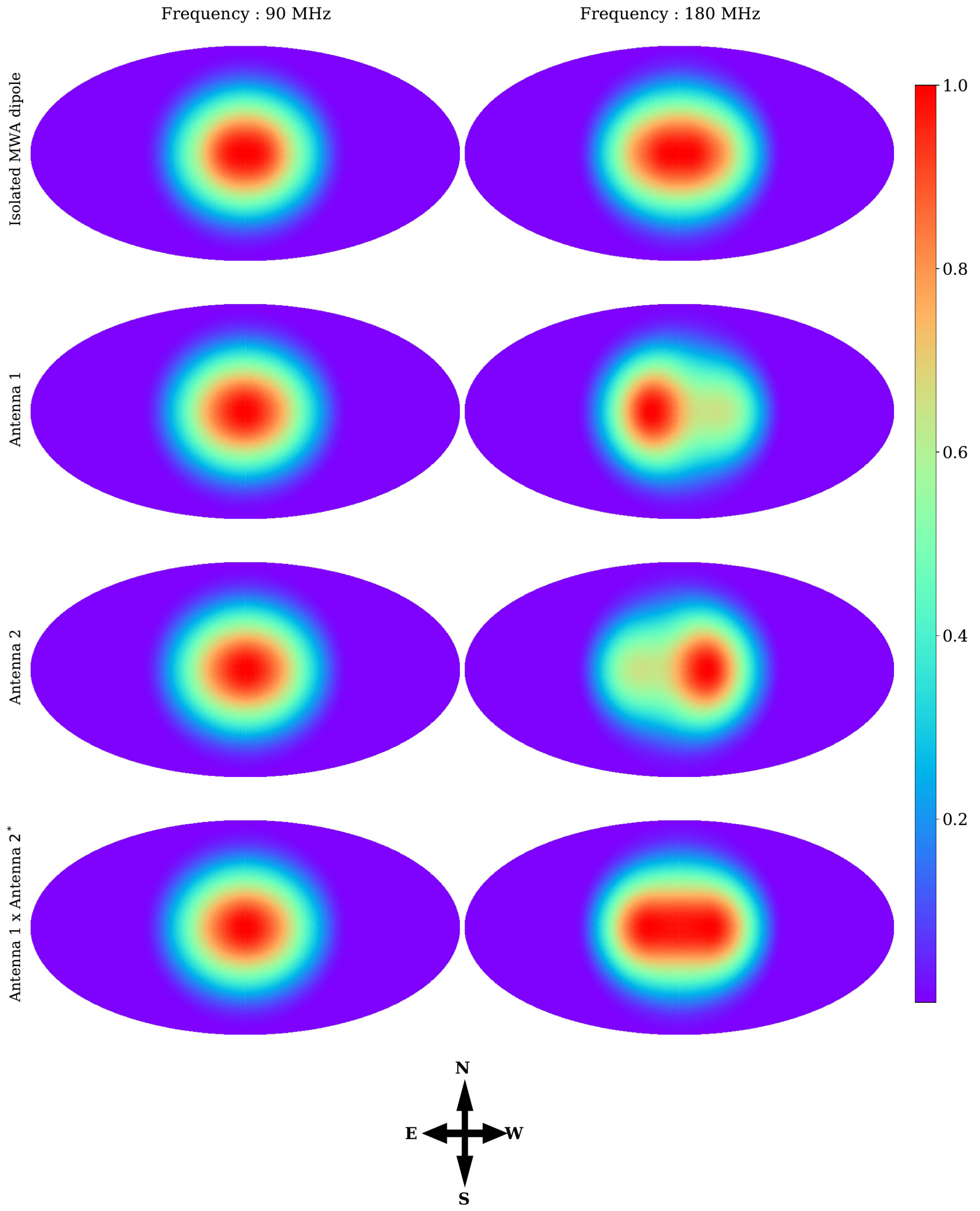


Figure 4.6: Simulated SITARA auto and cross antenna patterns at two frequencies, in Mollweide projection. For comparison, patterns for an isolated MWA antenna are given in the top row. The plots are peak normalised as shown in the colour bar. The coordinate system is local altitude-azimuth with the centre of the Mollweide projection corresponding to zenith; the local directions are also shown. It can be seen that due to mutual coupling, the patterns of closely spaced SITARA antennas diverge from that of an isolated MWA dipole.

ever, excess noise temperature in cross-correlations due to cross-talk of receiver noise is considered, since neglecting it is seen to yield poor results. We model the power measured in auto-correlations and cross correlations at each frequency as affine equations as given in Eqs.4.3.

The set of equations in Eqs.4.3 is an adaptation of a commonly used system model in single element global 21 cm experiments where the measured data are modelled as an ideal sky signal along with a multiplicative gain and an additive constant. In this model, the gains include all the multiplicative factors in the system such as the antenna efficiencies, analog gains, and any scaling introduced by the correlation and digital signal processing. The constant additive comprises of forward and reflected receiver noise and losses in the system. Similar models have been widely adopted for calibration of single element global 21 cm experiments such as EDGES (Rogers & Bowman, 2012) and SARAS (Nambissan T. et al., 2021).

$$\begin{aligned} P_{11} &= (T_{A11} + T_{N11})|G_1|^2 \\ P_{22} &= (T_{A22} + T_{N22})|G_2|^2 \\ P_{12} &= (T_{A12} + T_{N12})|G_1||G_2|e^{i(\phi_1 - \phi_2)} \end{aligned} \tag{4.3}$$

where T_{An} are the respective beam-weighted sky brightness temperatures, $G_1 = |G_1|e^{i\phi_1}$ and $G_2 = |G_2|e^{i\phi_2}$ are the complex gains of the signal chains, T_{N11} and T_{N22} are the excess noise powers in the individual auto-correlations in temperature units, with the dominant contribution from the active antenna LNA. Interested readers are referred to Sec.3.4.1 where a similar model was adopted for the SARAS 3 system. We have dropped the frequency terms for brevity. We consider the coupled receiver noise due to cross-talk in T_{N12}, however we ignore the cross-talk of sky signals. All temperatures are referred to the sky plane (i.e. the beam weighted sky temperature at the antenna terminals) and hence are in units of brightness temperature.

The absence of an in-situ absolute calibration, coupled with the wide radiation patterns of the antennas, motivates a calibration with diffuse sky models. Calibration of radio telescopes using models of the diffuse sky has been successfully utilised in low frequency astronomy (Rogers et al., 2004) as well as in a single antenna global 21 cm context (Singh et al., 2017). We adopt a similar procedure to obtain the gains and use them to scale the measured data to units of kelvin (K).

To calibrate auto-correlations, expected sky spectra are computed as follows. `pyGDSM`[4], a python implementation of the global sky model (GSM; de Oliveira-Costa et al., 2008; Zheng et al., 2017), is employed as the sky model. In this work we use the de Oliveira-Costa et al. (2008) version of GSM. For antenna radiation patterns, FEKO simulated EEPs of two MWA dipoles over a ground plane as discussed in Sec.4.6.1 are used. As mentioned in Sec.4.2, while the visibility equation given by Eq.4.1 is a convenient starting point, it is insufficient to compute expected sky spectra when the antennas are mutually coupled. Specifically, the assumption of identical antenna patterns for both of the antennas fails, as shown in Fig.4.6. Therefore an appropriately modified visibility equation, given by Eq.4.4, has to be employed.

$$V_{jk}(\nu) = \int J_j(\nu,\hat{n})C(\nu,\hat{n})J_k^H(\nu,\hat{n})e^{(\frac{-2\pi i\nu\vec{b}.\hat{n}}{c})}d\hat{n} \tag{4.4}$$

where $J_j(\nu,\hat{n})$ and $J_k(\nu,\hat{n})$ are the Jones matrices for the two antennas and $C(\nu,\hat{n})$ is the coherency matrix (Hamaker et al., 1996; Smirnov, 2011). Since only a single linear polarisation is utilised in our experiment, Eq.4.4 can be reduced into

$$\begin{split} V_{jk}(\nu) = \int T(\nu,\hat{n})[E_{j,\theta}(\nu,\hat{n})E^*_{k,\theta}(\nu,\hat{n}) + E_{j,\phi}(\nu,\hat{n})E^*_{k,\phi}(\nu,\hat{n})] \\ e^{(\frac{-2\pi i\nu\vec{b}.\hat{n}}{c})}d\hat{n} \end{split} \tag{4.5}$$

where we also simplify the coherency matrix to consist of unpolarised radiation of brightness temperature $T(\nu,\hat{n})$. Comparing Eqs.4.1 and 4.5 we can readily see that the antenna pattern A_a in Eq.4.1, assumed to be identical for both the antennas, can be replaced by the quantity $A_{j,k}(\nu,\hat{n})$ given by Eq.4.6.

$$A_{jk}(\nu,\hat{n}) = E_{j,\theta}(\nu,\hat{n})E^*_{k,\theta}(\nu,\hat{n}) + E_{j,\phi}(\nu,\hat{n})E^*_{k,\phi}(\nu,\hat{n}). \tag{4.6}$$

---

[4] https://github.com/telegraphic/pygdsm

It has to be noted that the visibilities given as per Eq.4.5 are not normalised. All sky interferometry as performed by SITARA is greatly benefited by the use of the HEALPix (Górski et al., 2005) framework; therefore the sky maps as well as antenna patterns are manipulated in HEALPix format. In this case, the vector $\hat{n}$ pointing to a celestial coordinate can be mapped to a pixel in a HEALPix map. With EEP simulations of antennas performed with a common origin - which is the mid point of the two antennas - the complex E-fields contain the geometrical phase referred to that common origin. Therefore the exponential factors corresponding to geometrical phase in Eq.4.5 can be removed, leading us to Eq.4.7.

$$V_{jk}(\nu, LST) = \frac{\sum_{n=1}^{N_{pix}} T(n, \nu, LST) A_{jk}(n, \nu)}{\sum_{n=1}^{N_{pix}} |A_{jk}(n, \nu)|} \tag{4.7}$$

Eq.4.7 is the discretized form of Eq.4.5 where the visibilities computed are also normalised. The normalisation adopted is such that when presented with a uniform sky temperature, the autocorrelation visibilities computed as per Eq.4.7 have the same uniform temperature. $T(\nu, LST)$ is the $N_{pix} \times 1$ sky map at frequency ν. The sky maps are rotated to bring the right ascension corresponding to the LST and the declination corresponding to the site latitude, to the zenith.

The computation in Eq.4.7 is repeated at a time cadence of 6 minutes and interpolated to all timestamps for which the data are collected to yield a simulated dataset similar to the measured data in time-frequency resolution, but devoid of instrumental noise and systematics. We assume that the multiplicative receiver gains and spectrum of the additive receiver noise remain constant throughout the observations and that the antenna radiation patterns are well known. Under these assumptions, the observed data along with the computed sky temperature form an overdetermined set of linear equations that may be solved to yield the system gain as well as additives with associated errors. In practice, a simple polynomial fit to simulation *vs* SITARA data is performed at each frequency. It may be noted that this technique bears a resemblance to the hot-cold or Y-factor measurement commonly employed in RF noise figure measurements, however in our case, there are multiple temperature states available by virtue of sky drift. The analog electronics in the field do not have any temperature regulation, therefore temperature fluctuations can induce gain variations. Daytime data are observed to be of poorer quality due to temperature rise in the electronics, as well as ionospheric fluctuations. This, along with the fact that during this specific observation carried out in May 2021 the highest sky temperature change occurs during local night time with the Galaxy transit, prompts us to perform calibration using night time data alone.

The gains derived with this technique can be assumed to consist of two multiplicative components - the overall system bandpass and a multiplicative gain, though in this analysis we do not attempt to separate the two components. The overall bandpass solution can be expected to the temporally stable. However the gains have a dependence on the ambient temperature, which has subsequently been shown with the singular spectrum analysis in Chapter 5. Therefore application of the gain solutions to data outside the calibration interval results in an overall bandpass calibration, but the resulting temperature scale would be incorrect.

A similar procedure is adopted to calibrate cross-correlations; however for cross-correlations the simulated and measured visibilities are complex valued. Therefore, we perform linear regression for the magnitude and phase of the visibilities separately to derive the complex gain. In addition to estimating gains, this calibration also yields an estimate of the cross-coupled receiver noise temperature. The results from calibration, viz the signal chain gains $|G_1|^2$, $|G_2|^2$ and $|G_1 G_2^*|$, as well as receiver noise temperatures are shown in Fig.4.7. It has to be emphasised here that the gains will have an arbitrary scaling depending on the normalisation in the FPGA firmware. Therefore the gains *do not* represent the analog system gain; rather they are merely the calibration coefficients to convert observed data into units of kelvin. Also, we place less confidence on the receiver noise temperature estimates, as the model used is deemed to be incomplete since it does not take into account cross-talk between antennas. Estimated absolute gains $|G_1|^2$ and $|G_2|^2$ are used to calibrate the auto-correlations as shown in

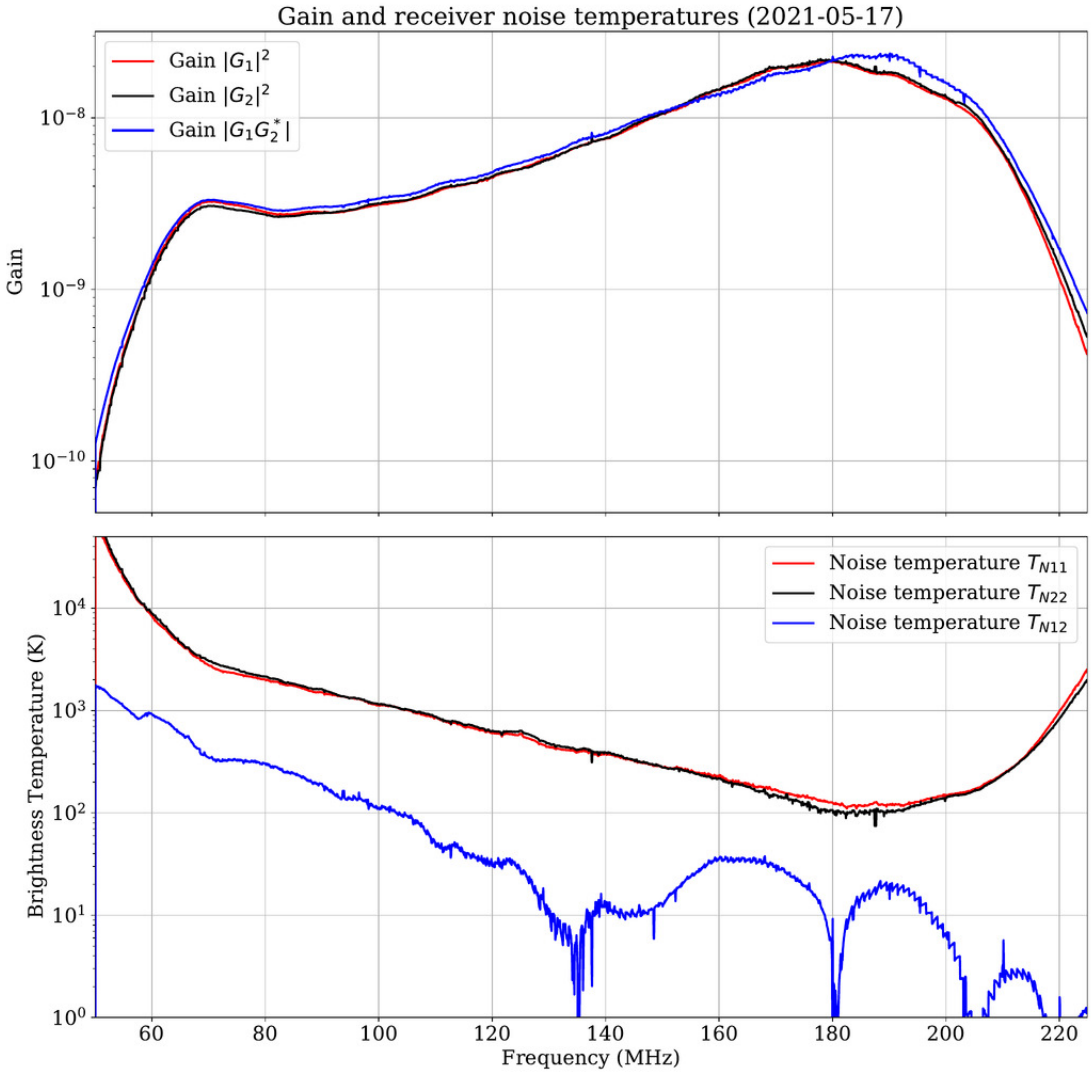


Figure 4.7: Receiver gains and noise temperatures as functions of frequency. The plots are semi-logarithmic to accommodate a wide dynamic range. The gains show the filtering introduced by the system at 70 MHz and 200 MHz. The gains include contributions from antennas, analog stages as well as *any scaling introduced by the digital signal processing in the correlator*, therefore the units are arbitrary. The noise temperatures are calibrated to units of kelvin. An interesting feature in the receiver noise temperatures is that the coupled receiver noise in cross-correlations is almost an order of magnitude less than receiver noise in autocorrelations.

Eq.4.8.

$$
\begin{aligned}
T_{11,meas} &= \frac{P_{11}}{|G_1|^2} \\
T_{22,meas} &= \frac{P_{22}}{|G_2|^2}
\end{aligned}
\tag{4.8}
$$

Similarly cross correlation visibilities are scaled with $G_1 G_2^*$.

$$T_{12,meas} = \frac{P_{12}}{G_1 G_2^*} \tag{4.9}$$

Receiver noise temperatures T_{Nij} may be subtracted out from $T_{ij,meas}$ for sake of comparison with simulations. Fig.4.8 shows the results of the calibration based on Eqs.4.8 and 4.9 at a frequency of about 111 MHz. In Fig.4.8 we have also

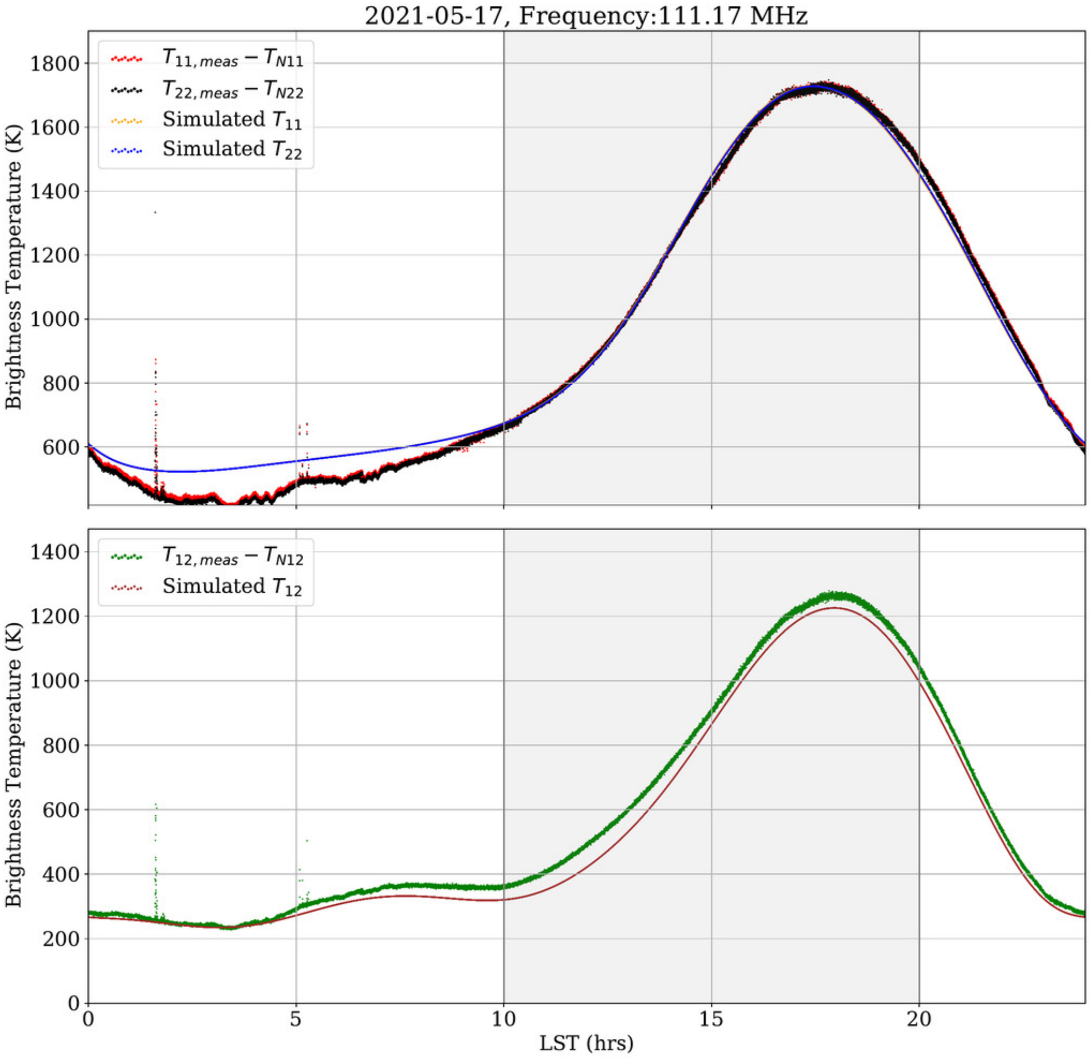


Figure 4.8: Variations in calibrated and T_{Nij} subtracted data as functions of local sidereal times (LST). The top panel shows calibrated auto-correlations along with simulated auto-correlations and the bottom panel shows magnitude of the calibrated cross-correlations along with simulated cross-correlations. Only data in the shaded region are used for calibration, since those LSTs have a rapid change in the sky temperature due to Galaxy transit. The solutions derived are then used for the entire data. It may be noted that T_{Nij} subtraction also removes any 21 cm signal from the data.

subtracted out the individual receiver noise temperatures T_{Nij}. We find that the simple model that we have adopted is able to capture the variations in SITARA data, at frequencies where the individual antenna patterns are somewhat similar.

We now inspect the result of calibration at frequencies where the antenna radiation patterns differ substantially. Consider the plots in Fig.4.9 which are identical to those in Fig.4.8 except that the frequency is now about 174 MHz. Despite scaling the temperatures as well as subtracting excess receiver noise temperatures, the shape of temperature *vs* LST does not exactly follow the simulations, unlike the plots for 111 MHz. In inteferometers with closely spaced antennas, cross-talk between the antennas becomes non-negligible and has to be taken into account. Since SITARA has antennas spaced at 1 m, we attribute the differences between SITARA data and mock data to cross-talk between the antennas and attempt to model its effects. With a model including cross-talk, we expect to obtain better estimates of the receiver noise temperatures.

4.6.3 An empirical model for cross-talk

In this section, we present an empirical model for the cross-talk in our system. We choose to model the data empirically such that the model can extended for future versions of SITARA with multiple antennas. The aim of this modelling is to enable *forward modelling* of global 21 cm templates into the instrument plane and to search for them in the data after foreground subtraction etc.

If there were no cross-talk, individual antenna voltages would consist only of the signals induced on the specific antennas. In the presence of cross-talk, cross-correlations would have non-negligible amounts of auto-correlations and vice-versa. We may therefore model the resulting measurements as a combination of "ideal" auto-correlations and cross-correlations. If we ignore reflections, the problem can be linearised. Thus, the equations for auto and cross correlations in

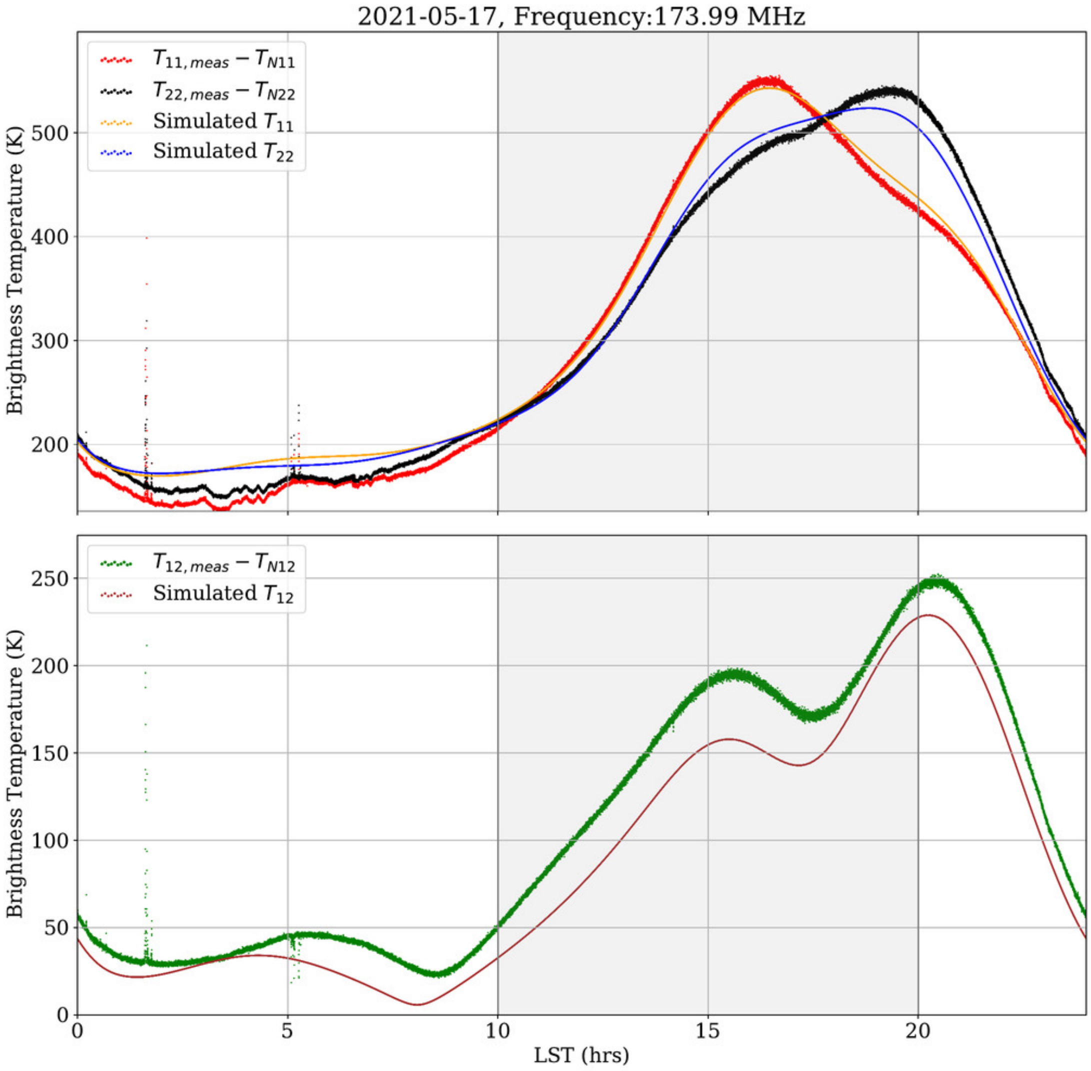


Figure 4.9: Calibrated and T_{Nij} subtracted data for ~ 174 MHz. The top panel shows calibrated auto-correlations along with simulated auto-correlations and the bottom panel shows magnitude of the calibrated cross-correlations along with simulated cross-correlations. The plot is of the same nature as Fig.4.8, however at this frequency the individual antenna radiation patterns differ. Despite this being captured by the FEE simulations, the simulated temperatures differ from calibrated data.

the presence of cross-talk at each frequency and LST can be written as

$$T_{11} = a_{11}V_{11} + a_{12}V_{12} + a_{21}V_{21} + a_{22}V_{22} + T_{n11} \tag{4.10}$$

$$T_{12} = b_{11}V_{11} + b_{12}V_{12} + b_{21}V_{21} + b_{22}V_{22} + T_{n12}$$

$$T_{21} = c_{11}V_{11} + c_{12}V_{12} + c_{21}V_{21} + c_{22}V_{22} + T_{n21}$$

$$T_{22} = d_{11}V_{11} + d_{12}V_{12} + d_{21}V_{21} + d_{22}V_{22} + T_{n22}$$

where V are the expected (simulated) visibilities in the absence of cross-talk, T are the visibilities in the presence of cross-talk. Though $T_{21} = T^*_{12}$ and $V_{2,1} = V^*_{1,2}$, we have included them for the sake of completeness. For a drift scan instrument such as SITARA, T and V vary as a function of LST, while their coefficients are expected to be constant. We may thus write them as matrices as follows

$$\boldsymbol{T} = \boldsymbol{V}\boldsymbol{B}, \tag{4.11}$$

with $\boldsymbol{T}$, $\boldsymbol{V}$ and $\boldsymbol{B}$ given by Eqs.4.13, 4.14 and 4.12 respectively.

$$\boldsymbol{T} = \begin{bmatrix} T_{11}(t_i) & T_{12}(t_i) & T_{21}(t_i) & T_{22}(t_i) \end{bmatrix} ; i = 1 \ to \ n \tag{4.12}$$

$$\boldsymbol{V} = \begin{bmatrix} V_{11}(t_i) & V_{12}(t_i) & V_{21}(t_i) & V_{22}(t_i) & 1 \end{bmatrix} ; i = 1 \ to \ n \tag{4.13}$$

$$\boldsymbol{B} = \begin{bmatrix} a_{11} & b_{11} & c_{11} & d_{11} \\ a_{12} & b_{12} & c_{12} & d_{12} \\ a_{21} & b_{21} & c_{21} & d_{21} \\ a_{22} & b_{22} & c_{22} & d_{22} \\ T_{n11} & T_{n12} & T_{n21} & T_{n22} \end{bmatrix} \tag{4.14}$$

where $\boldsymbol{T}$ is a $n \times 4$ matrix of measurements, V is a $n \times 5$ matrix of simulated visibilities and $\boldsymbol{B}$ is a 5×4 matrix of model coefficients. In this model, $\boldsymbol{B}$ is a matrix of coefficients that form the linear model for cross-talk. It may be noted that $\boldsymbol{B}$ has no direct physical interpretation. However, in Appendix. B, we present a plausible physical model for the terms in the empirical model that provides a plausible interpretation for the terms in $\boldsymbol{B}$. If the measurements are not calibrated to brightness temperature scale i.e. if the measured and expected visibilities are not in the same units, we may write them as $\boldsymbol{P} = \boldsymbol{T}\boldsymbol{G}$ where $\boldsymbol{P}$ is a matrix with the measured, uncalibrated data. $\boldsymbol{G}$ is a $4x4$ complex diagonal gain matrix that has the signal chain gains as diagonal elements as given in Eq.4.15.

Being a diagonal matrix, the effect of $\boldsymbol{G}$ is just a scaling of the data.

$$G = \begin{bmatrix} |G_1|^2 & 0 & 0 & 0 \\ 0 & G_1 G_2^* & 0 & 0 \\ 0 & 0 & G_2 G_1^* & 0 \\ 0 & 0 & 0 & |G_2|^2 \end{bmatrix} \tag{4.15}$$

where G_1 and G_2 are the gains of the individual signal chains. This gives us the equation describing SITARA data as :

$$\boldsymbol{P} = \boldsymbol{TG} = \boldsymbol{VBG} \tag{4.16}$$

If $\boldsymbol{G}$ is accurately known, it can be inverted to calibrate our measurements $\boldsymbol{P}$ to units of kelvin, as $\boldsymbol{G}$ is generally non-singular and invertible. Though gains obtained in Sec.4.6.2 can be used to construct $\boldsymbol{G}$, it is also shown that the model used to obtain those gains is incomplete. Therefore, we will use raw measurements $\boldsymbol{P}$ in our subsequent analysis.

If we have $n > 5$ independent measurements, it is possible to find a least-squares solution to Eq.4.16 to obtain the matrix of coefficients $\boldsymbol{BG}$. SITARA observations have a cadence of about 3 seconds and each observation spans several hours with good LST coverage and therefore, the $n > 5$ condition is easily satisfied for $\boldsymbol{P}$. Also, visibilities simulated for each observational timestamp in the same fashion as in Sec.4.6.2 form $\boldsymbol{V}$, and for the same considerations given there, we use LSTs between 10 and 20 hours for $\boldsymbol{T}$ and $\boldsymbol{V}$. Eq.4.16 is then solved at each frequency using a least-squares algorithm (such as `numpy.linalg.lstsq`). It has to be noted that since we have not corrected data for $\boldsymbol{G}$, the solution that we obtain is $\boldsymbol{BG}$ which is a product of coefficient matrix and gain matrix. In this work, we are interested in finding whether simple cross-talk considerations can better model the data.

Before proceeding to inspect the results of the least-squares fit, it is instructive to compare the above formalism with the procedure given in Sec.4.6.2. It is easy to see that if cross-talk is neglected, the principal diagonal elements of $\boldsymbol{B}$ will have a value of unity, the last row will have values of the receiver noise temperatures and all other terms will be zero. Then, a least-squares solution provides an estimate of the gain matrix $\boldsymbol{G}$ as well as receiver noise temperatures that, as we have already noticed, are also less accurate. Therefore, it can be inferred that the procedure given in Sec.4.6.2 is a simplified form of the more generalised procedure given here.

We now inspect the resultant $\boldsymbol{BG}$ matrix. We have to note that due to the lack of accurate estimations of $\boldsymbol{G}$, "gain" becomes a concept which is not well defined in the cross-talk model. Besides, there could be linear dependencies in the matrix $\boldsymbol{V}$ that introduce degeneracies in the fitted parameters. For example, as the antenna patterns are nearly identical at frequencies less than 150 MHz, we expect $V_{11} \approx V_{22} \approx |V_{12}|$ and therefore their corresponding coefficients in matrix $\boldsymbol{BG}$ will be degenerate. Nonetheless, we expect a sum of the coefficients to mitigate such degeneracies. Therefore, to enable comparisons with the gains derived in Sec.4.6.2, we use the sum of the coefficients in each column of $\boldsymbol{BG}$ (except the receiver noise temperatures) as a proxy for gains. Doing so also enables us to look at the differences between the two gain models. Fig.4.10 compares the gains derived with and without cross-talk considerations and Fig.4.11 shows the fractional difference as a percentage. We find that the differences are less than 10% for auto-correlation gains and less than 20% for cross-correlations, thus implying that the impact of cross-talk is less than 20%.

It is interesting to know whether the cross-talk model does a better job at accurately representing the measurements. While it is tempting to compute a pseudoinverse of $\boldsymbol{BG}$ and "calibrate" SITARA measurements $\boldsymbol{P}$, such an operation is erroneous. Therefore we choose to perform forward modelling to avoid issues of matrix inversion from affecting our results. For this, simulated $\boldsymbol{V}$ and fitted $\boldsymbol{BG}$ are multiplied to generate mock SITARA data and compared with SITARA measurements $\boldsymbol{P}$. We also compare them with $\boldsymbol{V}$ modified by gains and receiver noise temperatures from the no cross-talk model (Sec.4.6.2).

Fig.4.12 shows the results of this forward modelling from which it is evident

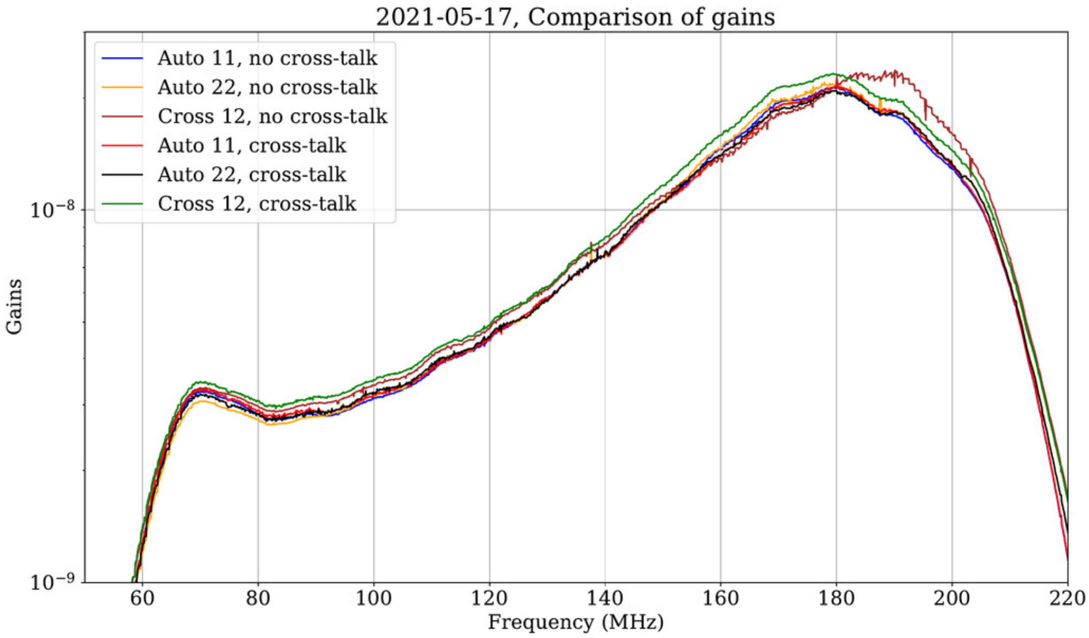


Figure 4.10: Comparison of gains estimated with and without cross-talk. The plots are semi-logarithmic to accommodate the dynamic range. As noted in the text, each "gain" in the cross-talk model is a sum of coefficients that includes cross-talk. Despite using two different formalism, it can be seen that they are in close agreement.

that the data are better modelled with a cross-talk model. A more informative way to represent the same data is to plot measurements against simulations employing different models. In Fig.4.13, the magnitude of the cross-correlations of the sky drift data over 24 hours is scatter plotted against simulated data spanning the same duration. The two colours in the plots correspond to the two different models tested - the red curve is for a model that ignores cross-talk while the black curve is for an empirical cross-talk model. If the measured data are well represented by a specific model, variations in the measured data due to sky drift will closely follow those of the simulations, and corresponding curve in a temperature-temperature plot would be a straight line. Inspecting Fig.4.13, we find that the cross-talk model represents the measured data better.

4.7 Results

In this section, we present some of the results obtained from the data analysis given in previous sections. Specifically, we present measurements of coupled re-

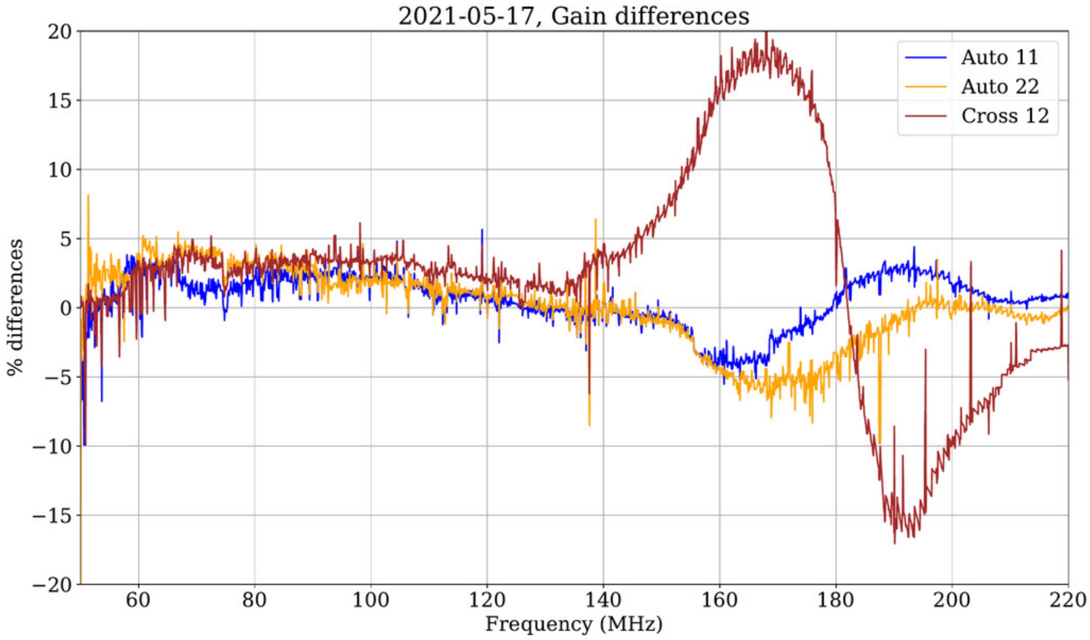


Figure 4.11: Differences between gains estimated with and without cross-talk. In this plot, the fractional differences between the gains estimated with and without cross-talk are shown as percentages. The auto-correlation gains derived with the two models have a difference less than 10% while the cross-correlation gains differ by about 20% at frequencies where the antennas patterns are dissimilar.

ceiver noise in SITARA as well as attempt to answer the question of whether SITARA is sensitive to all sky signals.

4.7.1 Coupled receiver noise

As a by-product of the modelling performed in Sec.4.6.3 that includes cross-talk, we obtain estimates of the receiver noise temperatures T_{n11}, T_{n22} and T_{n12}. However, these receiver noises are not calibrated to units of kelvin, and therefore have an arbitrary scaling introduced by the instrumental gains. Thus, we refer to them as P_{n11}, P_{n22} and P_{n12}. In a similar vein, we forward model the receiver noise temperatures obtained in Sec.4.6.2 where cross-talk has been ignored, to enable a comparison between the two methods, shown in Fig.4.14.

We find that below ~ 150 MHz where the antenna beam patterns are similar, the coupled receiver noise estimations with a model containing cross-talk are lower than the ones estimated without cross-talk. However, Fig.4.14 shows data that are not in units of temperature as the gains are not accurately known, thereby

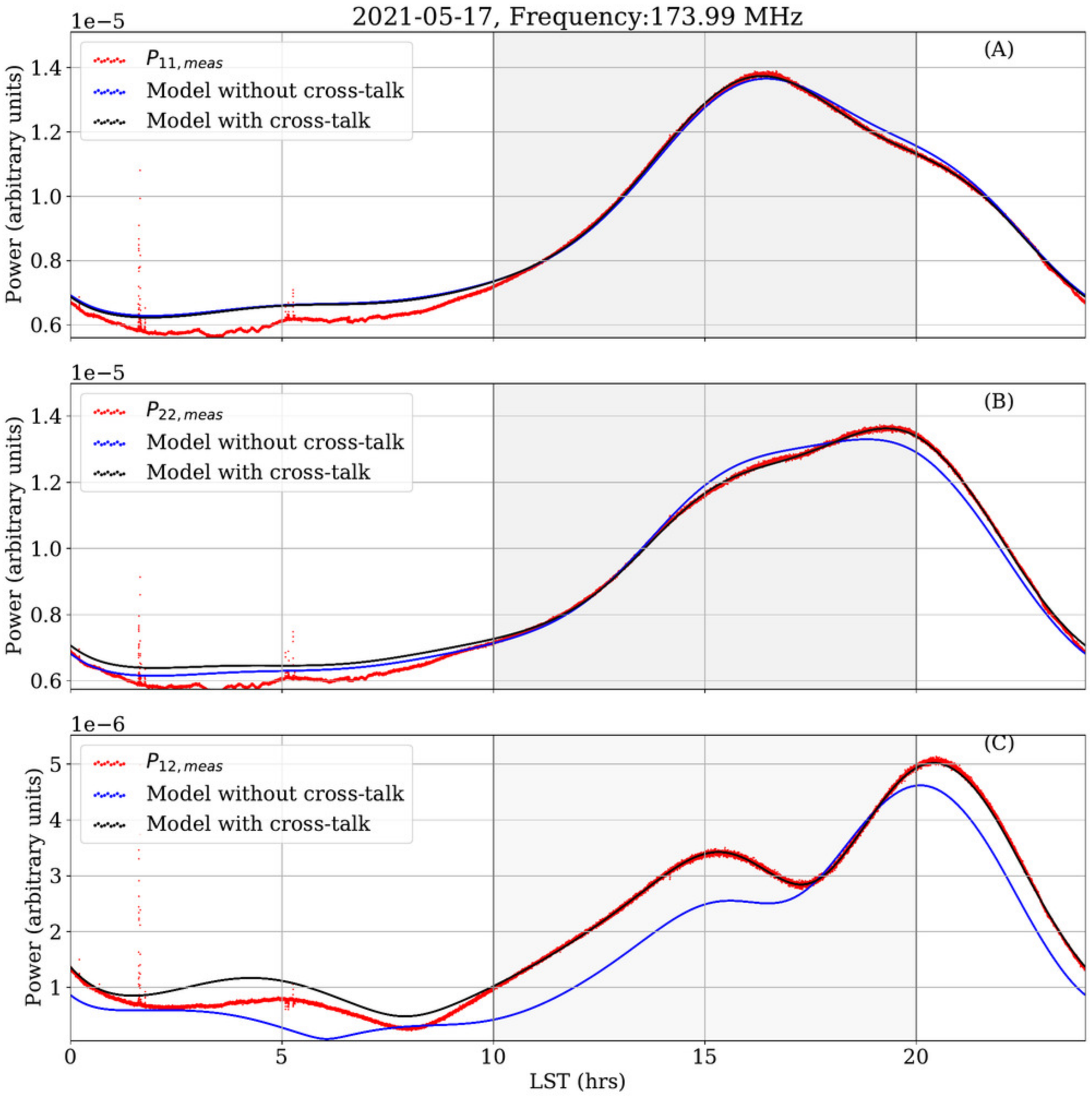


Figure 4.12: A comparison between SITARA data at 174 MHz with a model that does not consider cross-talk and one that considers cross-talk. Plots (A) and (B) are the auto-correlations and (C) is the cross-correlation magnitude. The data are forward modelled and therefore not in units of brightness temperature. Data from shaded area alone are used to compute gains and receiver noises. With the cross-talk model, the simulations match the data.

limiting its utility. We therefore use a ratio of the coupled and self noises as shown in Fig.4.15 to mitigate the effects of gain. It is seen that the cross-coupled receiver noise is approximately 10% of the receiver noise in auto-correlations; which has some interesting consequences. As discussed in Sec.4.7.2, interferometers have sensitivity to a uniform component that is a function of frequency. Frequencies less than 150 MHz where SITARA has more than 10% sensitivity to uniform components also happen to be where the ratio of coupled to self noise is less than

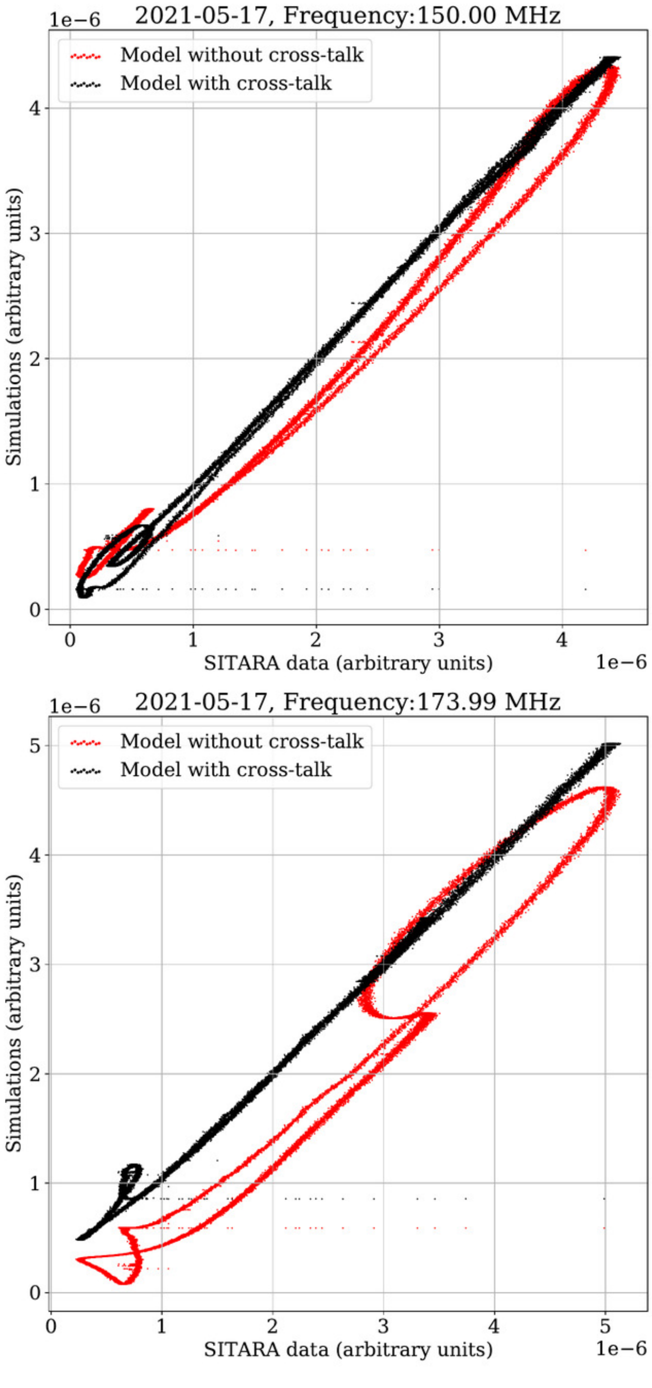


Figure 4.13: Comparison between SITARA data and simulations for the cross-correlation magnitude. Shown are the temperature-temperature plots between the SITARA data and simulations based on the two models. Two frequencies where the individual antenna patterns are dissimilar are chosen. We expect the simulations to follow data in a linear fashion in this plot, if the model used for simulations is accurate. While the model neglecting cross-talk fails to explain the variations in data, the cross-talk model fits the data very well.

5%. A subtraction of coupled receiver noise requires detailed modelling of the

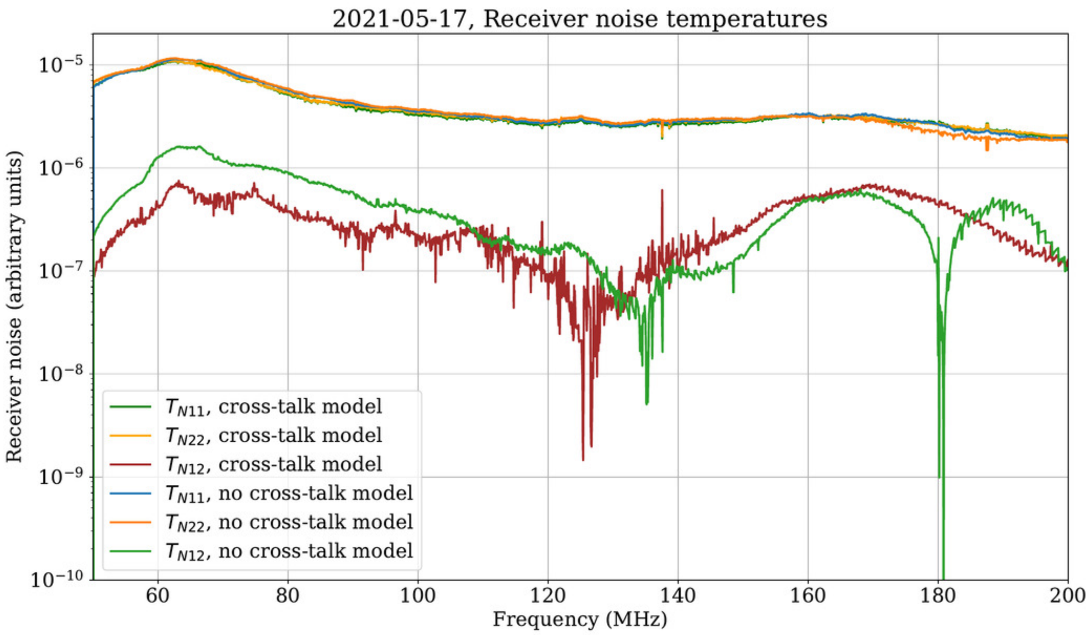


Figure 4.14: A comparison between estimations of receiver noise with and without cross-talk considerations. The receiver noise estimates are not calibrated to units of kelvin. It is seen that when cross-talk is modelled, the estimations of coupled receiver noise are generally lower, especially below ~ 150 MHz where the antenna beams are similar.

receiver noise coupling between interferometer arms (see for e.g. Sutinjo et al. (2020)). Nonetheless, their reduced levels by an order of magnitude compared to the receiver temperatures in auto-correlations may reduce their deleterious impact in detecting 21 cm signals, provided the coupled receiver noise does not contain high frequency ripples arising from multiple reflections. With better calibration, the coupled receiver noise temperature may be estimated with higher accuracy, and a delay spectrum of it can reveal such high frequency structures.

4.7.2 Is SITARA sensitive to an all-sky signal?

A global component of the sky is that average of the sky temperature over the entire solid angle visible to an observer. While previous works have demonstrated that an interferometer does respond to a uniform component by numerically integrating Eq.4.1 for various antenna types and orientations, no experiment has conclusively demonstrated these simulations. In this section, we attempt to answer the question of whether an interferometer becomes sensitive to an all sky

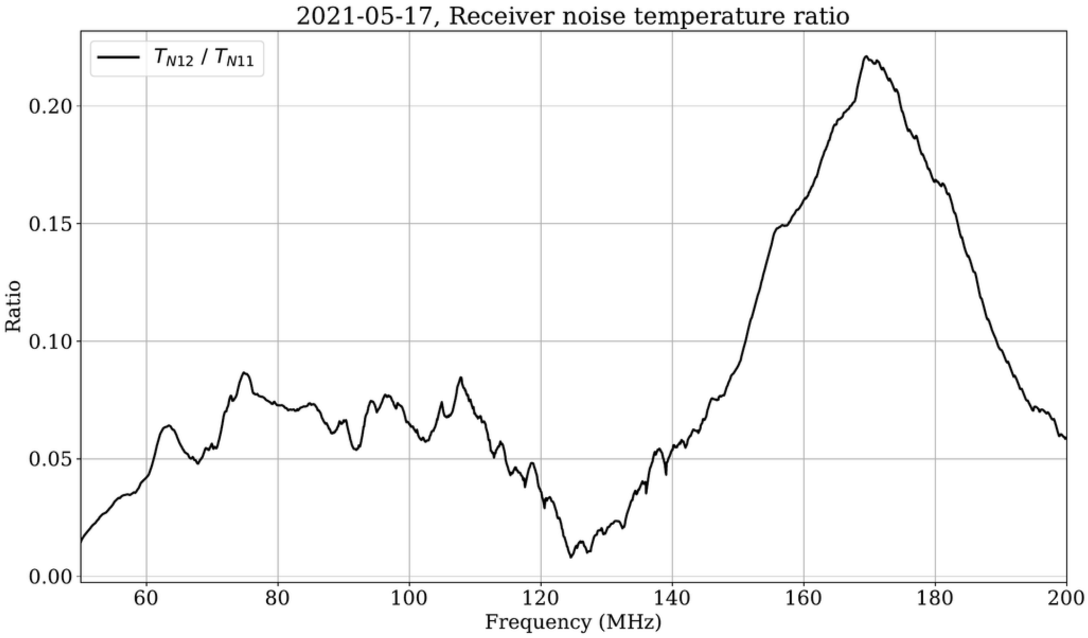


Figure 4.15: Ratio of estimated coupled receiver noise temperature to an auto-correlation receiver noise temperature. As expected, the cross-coupled receiver noise in data is substantially lower than auto-correlation receiver noise. The data have been smoothed with a Savitzky-Golay filter to reduce noise in the plots.

signal at short baselines using broadband data collected with SITARA.

In order to enable comparison between measured data and simulations such as in Singh et al. (2015), we define a quantity that is called coherence in Eq.4.17. The ratio of the measured auto-correlations and cross correlations provides an estimate of the coherence that an interferometer would measure at short baselines. In Eq.4.17, the coherence provided is devoid of the individual receiver gains. Owing to a formal resemblance between this computed quantity and the "complex degree of coherence" in optics (Born & Wolf, 1959), we refer to this quantity as the degree of coherence. We use the letter C to denote the degree of coherence instead of Γ used in optics, as Γ stands for reflection coefficient in electromagnetics.

$$\begin{aligned} C(\nu) &= \frac{T_{12}(\nu)}{\sqrt{T_{11}(\nu)T_{22}(\nu)}} \\ &\approx \frac{P_{12}(\nu)}{\sqrt{P_{11}(\nu)P_{22}(\nu)}} \end{aligned} \tag{4.17}$$

where $T_{12}(\nu)$ is the measured cross-correlations, $T_{11}(\nu)$ and $T_{22}(\nu)$ are corresponding auto-correlations, all with receiver noise subtracted. Since coherence is a ratio of temperatures, we expect it to be independent of instrument gains, and a ratio of raw powers can be used instead (with receiver noise subtracted). For a uniform sky of unit temperature $T_{A11} = T_{A22} = 1$, coherence is simply the visibility in cross-correlations. As shown in Singh et al. (2015), the visibility as a function of antenna spacing (or equivalently as a function of frequency for a fixed physical baseline) is expected to have a characteristic *sinc* shape for a uniform sky temperature. Viewed in this light, the simulated cross-correlations for a uniform sky can be interpreted as the degrees of coherence for the same.

The simplified picture given above is complicated by the presence of foregrounds having spatial structures, antennas with anisotropic radiation patterns, antenna pattern variations as a function of frequency (the so called "chromaticity" which couples spatial structures into the measured visibilities), noise coupling and cross-talk between the antennas and signal reflections within the receiver chains. Of these, foregrounds have the largest impact on our measurements. The low frequency radio sky above $\sim$ 10 MHz is dominated by Galactic synchrotron emission largely following a power law spectrum that also has spatial variations in the spectral index (Mozdzen et al., 2018; Sathyanarayana Rao et al., 2017a). If the foregrounds were spatially uniform, it is easy to see that their effect on coherence would be similar to that of a uniform sky. However, foregrounds have substantial structure, especially on the Galactic plane, and the measured visibilities would be the antenna beam weighted sum of these structures.

Thus, it is imperative to use sky regions with minimal structures - that can be considered close to a uniform sky - to carry out coherence computations. We expect imaginary components of visibilities to be minimal when the sky has no substantial off-zenith structures. If a minimum in imaginary is caused by bright compact sky regions at zenith - which is the nominal phase centre of SITARA - such a minimum would be short lived as the compact regions drift away from the zenith. Therefore, we choose a subset of data with the lowest imaginary component across all frequencies in cross-correlations, which also have the least variation in imaginary components as a function of time. We find LSTs between 4 to 5 hours to satisfy these conditions.

Since LSTs of 4-5 hours for the data used in this work corresponded to local day time when we expect the system gains to be different to the ones from night time, we estimate $\boldsymbol{BG}$ matrix with all 24 hours of data. From $\boldsymbol{BG}$ we obtain estimates of receiver noise temperatures P_{n11}, P_{n22} and P_{n12} and subtract them from the respective averaged auto and cross correlations. However, we do not remove other cross-talk components from the calibrated data to keep the computations simple. The coherence, as a function of frequency as given in Eq.4.17, is then computed, and the results are given in Fig.4.16. The same figure also shows simulated coherence employing the method given in Sec. 4.6.2, with a uniform sky model as well as a more realistic GSM foreground model.

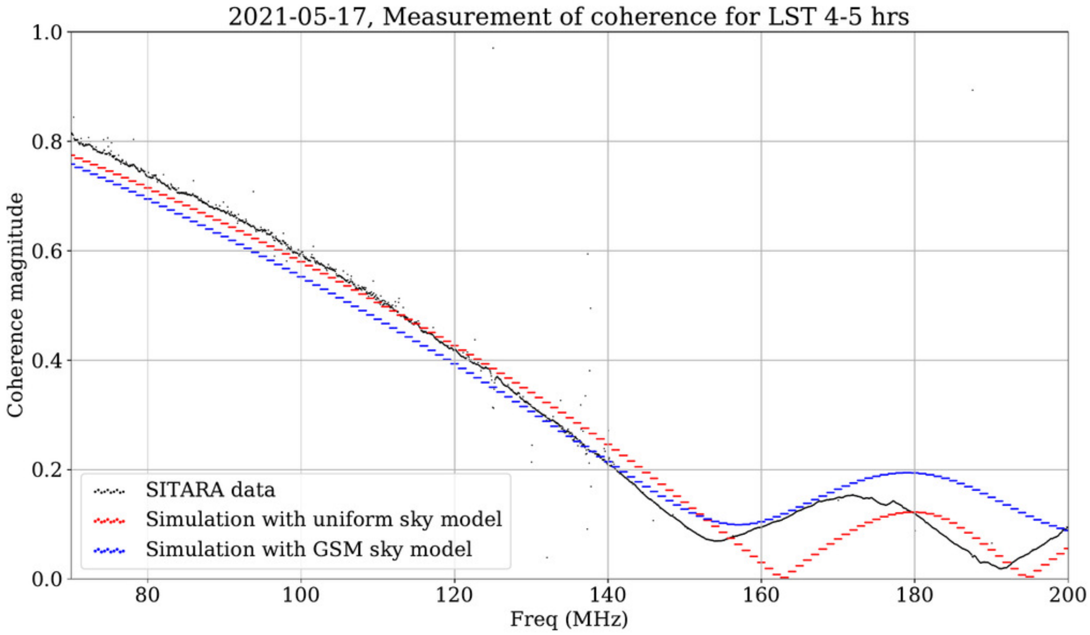


Figure 4.16: Comparing measured coherence (black) with simulations assuming a uniform sky (red) and a more realistic GSM foregrounds (blue). Uncalibrated data with receiver noise subtracted from auto and cross-correlations are used for this computation.

The close similarity between the measured coherence and the simulated ones, despite neglecting the cross-talk in the measurements, shows that SITARA is indeed sensitive to an all-sky component.

The sensitivity of an experiment to detect the global 21–cm signal depends on the foreground subtraction technique used and the global 21–cm template being searched for. In general, the sensitivity also depends on the observing bandwidth, with larger bandwidths providing better sensitivities. However as Fig.4.16 shows, coherence of a short spacing interferometer has an inflection point that depends on the baseline length. This inflection point in the coherence limits the usable bandwidths for conventional (polynomial based) foreground subtraction techniques, which rely on smoothness of the spectrum to separate foregrounds from the global 21–cm signal. Therefore, a correction for the coherence may be required to search for global 21–cm profiles, or the data may have to be band-limited, or novel analysis techniques need to be developed to account for the inflection point.

4.8 Discussion

Based on our modelling and analysis of SITARA data, we draw the following inferences.

1. Interferometers with short baselines are sensitive to all sky signals such as uniform components, with a response that closely matches simulations following Presley et al. (2015) and Singh et al. (2015).

2. Such interferometers also have non-negligible cross-talk as well as internal noise coupling between the antennas that has to be considered when they are employed for precision cosmology.

It is interesting to compare Fig.4.15 and Fig.4.16 with theoretical predictions of Venumadhav et al. (2016), specifically their Fig.5 that predicts that the spectral shapes of coherence and noise have an anti-correlation. Despite the differences between the SITARA setup and the case studied in Venumadhav et al. (2016), we find that the trend followed by the spectral shapes reported in this work are identical to the ones shown in Venumadhav et al. (2016). It is seen that at frequencies larger than 150 MHz where the coherence is low, the coupled receiver noise peaks. However, this observation has to be treated with some caution as Venumadhav et al. (2016) consider electrically short dipoles while SITARA antennas are not electrically short at all frequencies. Besides, the specifics of the curve in Fig.4.15 depend on the characteristics of the LNAs used, and so will be different to the idealised simulations of Venumadhav et al. (2016).

The main utility of this work is in short spacing interferometry in a 21 cm context. Nonetheless, the systematic effects seen in SITARA are expected to be seen by interferometers with closely packed antennas such as EDA-2 and SKA-low (Turner, 2016; van Es et al., 2020) stations. Extensions of the work given in this chapter may also aid in the analysis of systematics in those instruments as well as aid in their calibration. In this context, we also would like to point out that some of the cm-wave CMB instruments with closely packed elements observed excess spurious contributions in the data that were never fully explained (Watson et al., 2003; Padin et al., 2001). While the optics and electronics of the cm-wave instruments differ from low-frequency instruments such as SITARA, analysis based on adaptions of the empirical cross-talk model that we have outlined may aid in understanding the systematics in such instruments better.

4.8.1 Future work

We have not provided error estimates in this work. While it is possible to provide *formal* fitting errors based on the covariances of the fits performed, we have found them to be less reliable and frequently underestimating the errors. In future work, we expect to present an extensive error model devoid of such issues. A caveat with the approach given in this work is that the calibration and temperature scales are tied to diffuse sky models provided by the GSM. Many of the maps used as inputs to the GSM themselves have errors that are poorly understood, and re-calibrations of these maps are required when compared with precision radiometric data (see for e.g. Patra et al. (2015)). Noise-source based bandpass calibration of signal chains can be employed, along with in-situ measurements of complex antenna scattering parameters, to enable better modelling of various effects such as cross-talk. Use of compact, integrated circuitry deployed in active antennas such as LEDA (Price et al., 2018) can provide stable (with careful designs, traceable) calibration states to perform high time cadence calibration. However, since SITARA is an interferometer, techniques to provide *correlated* noise to the antennas have to be explored.

A potential source of error in low frequency radiometric calibration is the antenna pattern model. Since SITARA relies on simulated radiation patterns and sky models to calibrate the instrument, errors in either lead to mis-calibration. In-situ measurements of antenna patterns with unmanned aerial vehicles (UAVs) equipped with radio frequency instrumentation (Ninni et al., 2020) or satellites (Chokshi et al., 2020) can aid in this aspect.

Alternative calibration strategies include using multiple antennas and closure relations. The techniques developed in this chapter are currently being attempted with sky drift data acquired with EDA-2 at select frequencies (McKinley et.al. in prep). We also plan to carry out observations with different antenna spacings and orientations to quantify the response of interferometers at various spacings.

4.9 Conclusions

In this chapter we have detailed the system design, development and deployment of a short spacing interferometer - SITARA. We have also described the calibration strategies and some initial results employing those strategies. We find that interferometers with short baselines do have a response to all-sky signals. We also find that they have non-negligible cross-talk as well as noise coupling, with noise coupling less than 20% of the individual receiver noises in the current configuration of SITARA. We plan to modify the SITARA system and carry out observations as well as evolve the techniques so as to make them useful for similar closely packed interferometers such as EDA-2 and SKA-low.

Acknowledgements

This work is funded through CT's ARC Future Fellowship, FT180100321 - "Unveiling the first billion years: enabling Epoch of Reionisation science". The International Centre for Radio Astronomy Research (ICRAR) is a Joint Venture of Curtin University and The University of Western Australia, funded by the Western Australian State government. This scientific work makes use of the Murchison

Radio-astronomy Observatory, operated by CSIRO. We acknowledge the Wajarri Yamatji people as the traditional owners of the Observatory site. SITARA is an external instrument at the MRO, and is supported by MWA infrastructure at the MRO. Support for the operation of the MWA is provided by the Australian Government (NCRIS), under a contract to Curtin University administered by Astronomy Australia Limited. We acknowledge the Pawsey Supercomputing Centre which is supported by the Western Australian and Australian Governments.

We would like to acknowledge the expertise of the following people who contributed to the development of SITARA. David Kenney and Clinton Ward contributed to the hardware development and EMI chamber measurements of SITARA. A test deployment of SITARA to validate its end-to-end performance was carried out at John Kennewell's Australian Space Academy[5], Meckering, WA, with the help of John Kennewell and Randall Wayth. The MWA operations team at ICRAR/Curtin assisted in the deployment of SITARA; specifically Andrew McPhail and David Emrich played a crucial role in SITARA deployment at the MRO and troubleshooting.

This work uses the following python packages and we would like to thank the authors and maintainers of these packages :- `numpy` (Harris et al., 2020), `scipy` (Virtanen et al., 2020), `healpy` (Zonca et al., 2019), `astropy` (Astropy Collaboration et al., 2013, 2018), `aipy` (Parsons, 2016), `matplotlib` (Hunter, 2007), `ephem` (Rhodes, 2011) and `pygdsm` (Price, 2016).

---

[5] https://spaceacademy.net.au

Chapter 5

Singular spectrum analysis of SITARA time series

> *The task is not to see what has never been seen before, but to think what has never been thought before about what you see everyday.*
>
> - Arthur Schopenhauer, *Parerga und Paralipomena*

The work in this chapter is being developed for a publication entitled *Singular spectrum analysis of time series data from low frequency radiometers, with an application to SITARA data*, with authors Jishnu N. Thekkeppattu, Cathryn M. Trott, and Benjamin McKinley.

5.1 Abstract

Understanding the temporal characteristics of data from low frequency radio telescopes is of importance in devising suitable calibration strategies. Application of time series analysis techniques to data from radio telescopes can reveal a wealth of information that can aid in calibration. In this chapter, we investigate singular spectrum analysis (SSA) as an analysis tool for radio data. We show the intimate connection between SSA and Fourier techniques. We develop the relevant mathematics starting with an idealised periodic dataset and proceeding to include various non-ideal behaviours. We propose a novel technique to obtain long-term gain changes in data, leveraging the periodicity arising from sky drift through the antenna beams. We also simulate several plausible scenarios and apply the techniques to a 30-day time series data collected during June 2021 from SITARA - a short-spacing two element interferometer for global 21–cm detection. Applying the techniques to real data, we find that the first reconstructed component - the trend - has a strong anti-correlation with the local temperature suggesting temperature fluctuations as the most likely origin for the observed variations in the data. We also study the limitations of the calibration in the presence of diurnal gain variations and find that such variations are the likely impediment to calibrating SITARA data with SSA.

5.2 Introduction

There is a renewed interest in low-frequency (< 300 MHz) radio astronomy, due to the multitude of science cases that benefit from low-frequency observations. Several low-frequency radio telescopes such as LOFAR (van Haarlem et al., 2013), MWA (Tingay et al., 2013), HERA (DeBoer et al., 2017) and LWA(Hallinan et al., 2015) have been constructed and are observing, with the low frequency Square Kilometre Array, SKA-low (Dewdney et al., 2009) in its construction phase. Some of the key science goals for these metre wavelengths radio telescopes are cosmic dawn and epoch of reionisation (CD/EoR)(Trott, 2017), solar and heliospheric science (Nindos et al., 2019), cosmic magnetism (Gaensler et al., 2004) etc. Modern low frequency radio telescopes differ from their higher frequency counterparts in that they consist of aperture arrays constructed from large numbers of antennas, that are often beamformed in the analog or digital domain and correlated against each other to observe the radio sky. The calibration requirements, calibration models and the complexity are different to higher frequency (cm wavelengths and above) radio telescopes.

For dish-based interferometers with a small number of antennas, single dish telescopes and specialised low frequency radiometers, calibration techniques such as Dicke switching or noise injection can be employed. Specifically, global 21–

cm experiments constitute a group of low-frequency radiometers requiring precise and accurate calibration of the systems to limit the systematics to less than one part in a million. Most of the global 21–cm experiments use single antennas as sensors and employ Dicke-switching ambient temperature thermal loads and noise diodes (see Rogers & Bowman (2012); Nambissan T. et al. (2021)) or noise injection (Singh et al., 2018b) for bandpass calibration. The same calibrators are used to compensate for receiver gain drifts over extended periods of time. However, in order to maintain a stable excess noise ratio (ENR), the calibrators have to be maintained in temperature controlled environments or specialised ENR compensating designs used (see for e.g. Oh et al. (2017)), and the noise diodes themselves require periodic re-calibration with laboratory standards to mitigate drift and ageing, especially if they are to be used in applications demanding high accuracy such as cosmology. Besides, with future telescopes such as the SKA-low potentially employing thousands of low cost active antennas, calibration using dedicated noise diodes at each antenna becomes impossible. The aperture array nature of these instruments, i.e. large numbers of stationary antennas of simple construction, often with integrated low-cost low noise amplifiers (LNAs), necessitates calibration based on sky models. Therefore, development of novel mathematical tools to explore the long-term stability of these low-frequency telescopes and radiometers, and to determine the limits of calibration and data integration, become essential.

In this chapter, we explore the potential of one such tool in analysing time series data from low frequency radiometers, and an application of it to real valued time series data. The instrument under study is SITARA, a broadband two element interferometer targeting global 21–cm detection employing short spacing interferometry. SITARA consists of two MWA style active antennas kept 1 m apart over a large ground-plane (35 m) and a correlator that records auto-correlations and cross-correlations with spectral resolution of about 61 kHz each, across 0–250 MHz. The usable band is limited to 70–200 MHz, with reduced sensitivity in the 50–70 MHz band. SITARA is deployed within the radio quiet zone of the Murchison Radioastronomy Observatory (MRO) in Western Australia, which is also the site for future SKA-low. The data are collected round the clock and timestamped data are written out in `miriad` format (Sault et al., 1995). SITARA has been conceived as the first prototype to evolve tools and techniques for short-spacing interferometry; further details about SITARA can be found in Thekkeppattu et al. (2022). Single frequency data from SITARA auto-correlations can essentially be treated as radiometric data, with SITARA behaving as an uncalibrated total power radiometer.

5.3 Motivation

As an example to motivate this study, we consider real valued auto-correlations from SITARA data. A plot of single frequency channel time series data from SITARA auto-correlations, at a frequency of 111 MHz for the month of June 2021, is shown in Fig.5.1. The frequency of 111 MHz has relatively low radio frequency interference (RFI) and is also a frequency where mutual coupling between antennas does not cause large beam shifts (see Thekkeppattu et al. (2022) for details). The data have a periodic nature arising from the drifting of various regions of the radio sky through the antenna beams as the Earth rotates. If the radio telescope was perfectly calibrated, and there was no RFI, Fig.5.1 would have shown a perfectly repeating pattern.

As can be seen, this is not the case. The data have some multiplicative gain variations as well as some additive RFI. Therefore, our aim is to decompose this time series into some components that can help us understand it better. A plausible decomposition is the Fourier transform, i.e. into sines and cosines. However, radiometric data can have aperiodic patterns such as trends that do not become evident in Fourier analysis. Besides, sines and cosines are an artificial basis in some sense. Therefore a method of analysis that is independent of such assumptions is required.

Singular spectrum analysis (SSA) is a set of data driven tools that can decompose a time series into elementary patterns such as trend and oscillatory components. The raw data from a low frequency radio telescope with a fixed

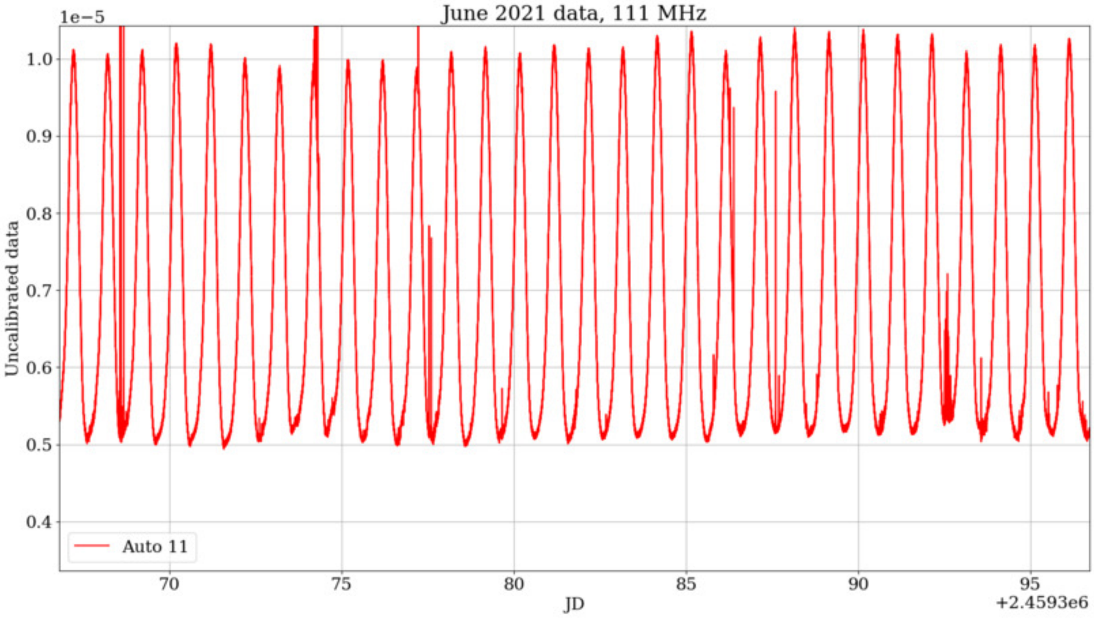


Figure 5.1: SITARA raw data time series as a function of Julian Date (JD) for a single frequency channel of bandwidth 61 kHz, at a frequency of 111 MHz. The data have been extracted from a concatenated SITARA dataset for the month of June 2021. Solar bursts contribute most of the RFI seen in this time series.

pointing are expected to have oscillatory components with a period corresponding to a sidereal day, while owing to environmental changes trend-like patterns (drift) are also expected. This makes SSA an ideal tool for analysis of such data.

Singular spectrum analysis (SSA) techniques appear in the analysis of dynamical systems (Broomhead & King, 1986; Vautard & Ghil, 1989) and SSA has been a popular tool for time series analysis in a variety of fields such as meteorology and climate science (Ghil et al., 2002) and geophysics (Dokht et al., 2016). However the application of SSA in radio astronomy has been limited (Donskikh et al., 2016; Gürel et al., 2018). Indeed, to the best of our knowledge, SSA has not been applied for time series analysis of radiometric time series data. This chapter aims to detail the necessary mathematical tools for SSA of time series data from a low frequency radiometer and demonstrates them with SITARA data.

5.3.1 Notations and mathematical preliminaries

In this section, we describe the notation employed and define certain matrices that are useful for the subsequent analysis. Many of these definitions can be found in Davis (1979) as well as Olson et al. (2014). We use bold capital Roman letters such as $\boldsymbol{X}$ to denote matrices, and small Roman letters such as n are used for indexing. We use zero-based numbering such that the indices start at 0. Small Roman letters with an arrow such as $\vec{u}$ denote vectors. Integers are denoted by capital Roman letters such as L. We now proceed to define some basic matrices and associated linear algebra.

5.3.1.1 Circulant and anti-circulant matrices

An $N \times N$ square matrix $\boldsymbol{C}$ is circulant if each row of the matrix is a *right* shifted version of the previous row as shown in Eq.5.1.

$$\boldsymbol{C} = \begin{bmatrix} c[0] & c[1] & ...... & c[N-1] \\ c[N-1] & c[0] & ...... & c[N-2] \\ c[N-2] & c[N-1] & ...... & c[N-3] \\ .. & .. & ...... & .. \\ c[1] & c[2] & ...... & c[0] \end{bmatrix} \tag{5.1}$$

An $N \times N$ square matrix $\boldsymbol{C}$ is anti-circulant if each row of the matrix is a *left* shifted version of the previous row as shown in Eq.5.2.

$$\boldsymbol{C_a} = \begin{bmatrix} c[0] & c[1] & ...... & c[N-1] \\ c[1] & c[2] & ...... & c[0] \\ c[2] & c[3] & ...... & c[1] \\ .. & .. & ...... & .. \\ c[N-1] & c[0] & ...... & c[N-2] \end{bmatrix} \tag{5.2}$$

Both the circulant and anti-circulant matrices can be obtained from a sequence $c[n]; n = 0, 1, ...N-1$ and are completely specified by that sequence. While useful in a wide variety of analyses, circulant matrices are not of much utility for our analysis. We will be dealing with anti-circulant matrices instead.

5.3.1.2 Block matrices

A matrix can be interpreted to have been broken into blocks or submatrices which are themselves matrices.

5.3.1.3 Matrix products

The notations employed in this chapter for the various matrix products are listed below.

1. Regular matrix-matrix multiplication is denoted with no specific operator.

2. $\otimes_O$ denotes the outer product of two vectors that results in a matrix.

3. $\odot$ represents an element-by-element multiplication of two matrices, known as the Hadamard product.

5.4 Basics of singular spectrum analysis

Though the theory of SSA is covered in detail in references such as Golyandina et al. (2001), we provide a basic description of the steps involved for completeness. In this chapter we follow the SSA approach known as the Broomhead-King (BK) version, with the alternate being the Vautard-Ghil (VG) version. The VG version is only suitable for the analysis of a stationary time series and therefore is not discussed here. Following Golyandina & Korobeynikov (2014), the four major steps in BK SSA are given below.

1. Convert the 1-D time series into a 2-D matrix called a *trajectory matrix*. This step is called *embedding* in time series analysis.

2. Decompose the trajectory matrix with singular value decomposition (SVD). The result consists of a set of left and right singular vectors and associated singular values.

3. Reconstruct the constituent components of the trajectory matrix with selected singular vectors and singular values.

4. Reconstruct the time series with these components by performing diagonal averaging.

In the embedding step, the trajectory matrix is constructed with columns consisting of elements from sliding a window of length L across the original time series. The criteria for selection of this window length is given in Sec.5.5.1.1. For each sliding, the elements inside the window are made into one column of the trajectory matrix, yielding a matrix with L rows and $K = N - L + 1$ columns. Consider a time series $x[n], n = 0, 1, ...N - 1$. The embedding step converts this time series of length N into an $L \times K$ matrix that has the elements of the time series as given in Eq.5.3.

$$\boldsymbol{X} = \begin{bmatrix} x[0] & x[1] & ...... & x[K-1] \\ x[1] & x[2] & ...... & x[K] \\ x[2] & x[3] & ...... & x[K+1] \\ .. & .. & ...... & .. \\ x[L-1] & x[L] & ...... & x[N-1] \end{bmatrix} \tag{5.3}$$

The trajectory matrix is like a Hankel matrix, though it is not square in general. It may be noted that some implementations pad the original time series with zeros to obtain $K = N$, though we do not employ this.

In the second step, the trajectory matrix is decomposed via SVD to yield left and right singular vectors as well as the corresponding singular values. This can be written as:

$$\boldsymbol{X} = \boldsymbol{U\Sigma V}^T \tag{5.4}$$

The SVD operation decomposes the $L \times K$ matrix $\boldsymbol{X}$ into three matrices $\boldsymbol{U}, \boldsymbol{\Sigma}$ and $\boldsymbol{V}$; where $\boldsymbol{U}$ is an $L \times L$ unitary matrix, $\boldsymbol{V}$ is a $K \times K$ unitary matrix and $\boldsymbol{\Sigma}$ is an $L \times K$ diagonal matrix consisting of the singular values. Since we consider only real valued matrices $\boldsymbol{X}$ in this work, the matrices $\boldsymbol{U}$ and $\boldsymbol{V}$ are real orthogonal matrices. Each singular value σ_i and the corresponding singular vectors $\vec{u_i}$ and $\vec{v_i}$ form an eigentriple $(\sigma_i, \vec{u_i}, \vec{v_i})$. The decomposition step can also be written as:

$$\boldsymbol{X} = \sum_{i} \boldsymbol{X_i};\ \boldsymbol{X_i} = \sigma_i(\vec{u_i} \otimes_O \vec{v_i}) \tag{5.5}$$

where $\otimes_O$ is the vector outer product of two vectors yielding a matrix.

These eigentriples can be grouped and used to reconstruct the various components of the time series. An inspection of the singular values can give an insight into the complexity of the data and the number of eigentriples to consider, and this is also related to the rank of the trajectory matrix.

The next step is to reconstruct the trajectory matrix with the selected eigentriples. Once this is accomplished, a reconstructed time series is obtained by performing an anti-diagonal averaging over the reconstructed matrix - this step is called "diagonal averaging" (Golyandina & Korobeynikov, 2014) [1] The diagonal averaging step applied to a trajectory matrix $\boldsymbol{y}$ yields a series y_s as shown in Eq.5.6.

$$y_s[n] = \frac{\sum_{(l,k)\in A_s} y[l][k]}{|\mathbb{A}_n|} \tag{5.6}$$

where $\mathbb{A}_n = \{(l,k); l+k = n, 0 \leq l \leq L-1, 0 \leq k \leq K-1\}$ and $|\mathbb{A}_n|$ is the number of elements in the set $\mathbb{A}_n$. This step can be treated as the reverse of the embedding step, conversion of an anti-diagonal, Hankel-like matrix back into a time series. When applied to each $\boldsymbol{X_i}$ in Eq.5.5, this results in *reconstructed series.* However, one can also choose to group the eigentriples to obtain the reconstructed series, and indeed the grouping approach is what we employ. Therefore, the result is the decomposition of the original time series into a sum of reconstructed series as shown in Eq.5.7

$$x[n] = \sum_{i=0}^{h} \widetilde{x}_i[n], n = 0, 1, ...N-1 \tag{5.7}$$

where h depends on the grouping of eigentriples. For elementary grouping, $h = L - 1$ as there can only be a maximum of L singular values for a matrix of dimensions $L \times K$. Depending on the selection of eigentriples, the reconstructed time series reveals the corresponding aspect of the data. For example, the first eigentriple contains information on the trend, while the subsequent ones are associated with oscillatory patterns. Higher eigentriples are typically associated with noise.

---

[1] We suggest to name it anti-diagonal averaging to avoid confusion, however we use the term "diagonal averaging" to be consistent with SSA literature.

It is interesting to compare some of the techniques used in SSA to a related technique, principal component analysis (PCA). With PCA, a set of new orthonormal basis vectors (called principal components) is found by diagonalising the data covariance matrix. The first few principal components with the largest associated variance can then be used to represent the data. In practice, PCA is computed with eigendecomposition of the covariance matrix formed from the data, or with an SVD of the data themselves. Seen in this light, SSA can be viewed as a PCA of the trajectory matrix formed from a time series, with an associated diagonal averaging of a reconstructed trajectory formed out of principal components. However, it should be noted that SSA is different to a straightforward application of PCA to a one-dimensional time series; with the key differences being embedding and diagonal averaging which convert a time-series into a matrix and back.

5.5 SSA of periodic time series

In this section, we obtain the mathematical form of SSA when applied to a time series containing periodic data. We first consider the case where we have an ideal periodic time series. We find that in this case, the singular vectors obtained are sinusoidal in nature. The case where a periodic series is corrupted by multiplicative element is then considered. We also perform simulations to validate the algebra.

5.5.1 SSA applied to an ideal periodic time series

Consider a time series $x[n], n = 0, 1, ...N-1$ of length N that is strictly periodic with a period given by L.

$$x[k + pL] = x[k] \tag{5.8}$$

We also assume that the length N of the time series obeys the relation $N = zL - 1$, where z is an integer. Forming a trajectory matrix out of this sequence gives the following matrix.

$$\boldsymbol{X} = \begin{bmatrix} x[0] & x[1] & \ldots\ldots & x[L-1] \\ x[1] & x[2] & \ldots\ldots & x[0] \\ .. & .. & \ldots\ldots & .. \\ x[L-1] & x[0] & \ldots\ldots & x[L-2] \end{bmatrix}_{L\times K} \tag{5.9}$$

As K is an integer multiple of L, the trajectory matrix can be interpreted as a block matrix, and partitioned into $M = \frac{K}{L}$ square submatrices as given in Eq.5.10, where $\boldsymbol{X_{ac}}$ is an anti-circulant matrix of dimensions $L \times L$.

$$\boldsymbol{X} = \begin{bmatrix} \boldsymbol{X_{ac}} & \boldsymbol{X_{ac}}.. \end{bmatrix}_{L\times M} \tag{5.10}$$

$$\boldsymbol{X_{ac}} = \begin{bmatrix} x[0] & x[1] & \ldots\ldots & x[L-1] \\ x[1] & x[2] & \ldots\ldots & x[0] \\ x[2] & x[3] & \ldots\ldots & x[1] \\ .. & .. & \ldots\ldots & .. \\ x[L-1] & x[0] & \ldots\ldots & x[L-2] \end{bmatrix} \tag{5.11}$$

$$\tag{5.12}$$

Before proceeding further, some observations can be made.

1. We find that $rank(\boldsymbol{X}) \leq L$, as there can only be a maximum of L linearly independent columns in $\boldsymbol{X}$. Consequently there can be a maximum of L non-zero singular values for $\boldsymbol{X}$.

2. The anti-circulant matrix form given in Eq.5.10 is real symmetric. Therefore, it has an eigendecomposition given as $\boldsymbol{X_{ac}} = \boldsymbol{Q\Lambda Q^T}$, with $\boldsymbol{Q}$ being orthonormal.

This enables us to write the trajectory matrix as $\boldsymbol{X} = \begin{bmatrix} \boldsymbol{Q\Lambda Q^T} & \boldsymbol{Q\Lambda Q^T}.. \end{bmatrix}$. Without going through the pedagogical details, we state that the block matrix $\boldsymbol{X}$ can be decomposed into three matrices as given below.

$$\boldsymbol{X} = \boldsymbol{Q}\left[\sqrt{M}\boldsymbol{\Lambda} \quad \boldsymbol{0} \quad \boldsymbol{0} \quad ...\right] \frac{1}{\sqrt{M}} \begin{bmatrix} \boldsymbol{Q^T} & \boldsymbol{Q^T} & \boldsymbol{Q^T} & .. \\ \boldsymbol{Q^T} & \boldsymbol{Q^T} & \boldsymbol{Q^T} & .. \\ \boldsymbol{Q^T} & \boldsymbol{Q^T} & \boldsymbol{Q^T} & .. \\ .. & .. & .. & .. \end{bmatrix} \tag{5.13}$$

$$= \boldsymbol{U\Sigma V^T} \tag{5.14}$$

where, in the last step we recall the SVD of $\boldsymbol{X}$ to facilitate a direct comparison.

It can be seen the decomposition given may be treated as an SVD of the trajectory matrix, provided some caution is exercised. Since the singular vector matrices have to be unitary, the right singular matrix gets divided by a scaling factor of $\sqrt{\frac{K}{L}}$ while the singular values get multiplied the same factor. Also, the left singular matrix in SVD is unitary, while the eigenvector form given in Appendix.C has unit amplitude. The corresponding scaling applies to the singular value, however it is inconsequential when the eigendecomposition and SVD are computed with numerical packages. Moreover, eigenvalues can be positive and negative, and as given in Appendix.C, they occur in positive and negative pairs in this context. However, singular values are always non-negative and therefore the signs of the eigenvalues get moved into the singular vectors when equated with the SVD given in Eq.5.13. Also, we assume that both the singular values and eigenvalues (and corresponding vectors) have been ordered in the same fashion - typically in descending magnitude. Nonetheless, these subtleties are inconsequential to an important observation described below.

The equivalence of decompositions in Eq.5.13 reveals an interesting aspect of the SSA of a periodic time series. As given in Appendix.C, the orthonormal matrix $\boldsymbol{Q}$ consists of sinusoidal eigenvectors. Therefore, the left singular vectors in the matrix $\boldsymbol{Q}$ are all sinusoidal. Besides, the vectors in $\boldsymbol{Q}$ are periodic in L and thus when arranged as blocks in the right singular vector matrix, they form continuous sinusoids. Therefore, we conclude that *when SSA is applied to a strictly periodic sequence, the singular vectors obtained are sinusoidal.*

5.5.1.1 Simulations

To validate the above calculations, we analyse a simulated periodic time series. Since our aim is to gain a better understanding of SSA as applied to radiometric time series data, the simulations are beam multiplied sky temperature as would be seen by a radiometer. The simulation methodology used to obtain the time series is described in Thekkeppattu et al. (2022), which we have extended to yield a time series spanning 30 days. For this, we make an important assumption that the radio sky is static and therefore, the true sky temperature is exactly the same across all days for each local sidereal time. The contributions from the Sun are ignored, the position of which relative to the Galaxy changes with time. The simulations are for auto-correlations at a frequency of 111 MHz with a time cadence of 15 minutes of sidereal time. Since the sky has a periodicity of one sidereal day, it is important that the cadence is chosen in sidereal time units.

The simulation yields $N = 2879$ data points which are then converted into a trajectory matrix. Key to embedding data is selection of an appropriate window length. Typically, a window length is chosen such that it is divisible by the fundamental of the known periodicity. As we have prior information that data have a periodicity corresponding to a sidereal day, a choice of the window length is the number of samples that correspond to one sidereal day. Therefore a suitable embedding dimension is $L = 96$; as it corresponds to one sidereal day. This gives a matrix of dimensions 96×2784 which can be partitioned into $M = \frac{K}{L} = 29$ block matrices. The simulated time series and a representation of the trajectory matrix are shown in Fig.5.2.

The trajectory matrix is then decomposed with SVD. The results from the SVD are shown in Fig.5.3. Independently, an $L \times L$ submatrix of the trajectory matrix (with $L = 96$) is eigendecomposed. The eigenspectrum of this anti-circulant matrix is also given in Fig.5.3. Moreover, the DFT of the underlying periodic sequence of length $L = 96$ is also computed and plotted in the same figure with the DFT values sorted in descending order. As the sequence is real, the DFT spectrum is Hermitian and the sorted values appear twice in the spectrum.

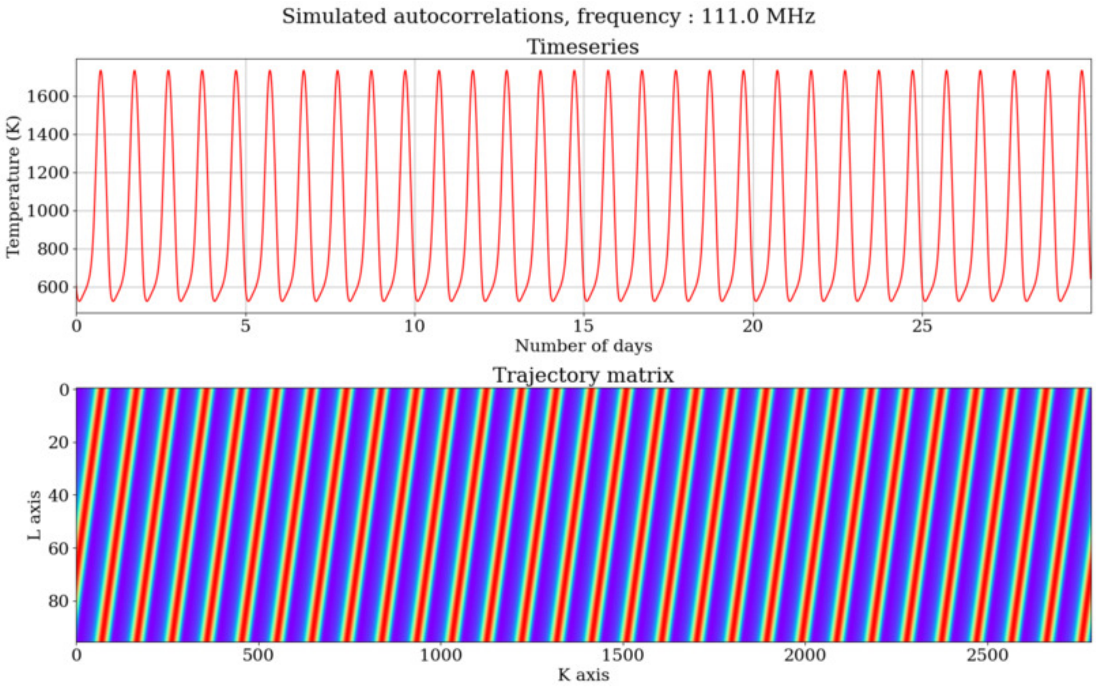


Figure 5.2: Simulated time series at a frequency of 111 MHz and the corresponding trajectory matrix formed by choosing an embedding dimension $L = 96$.

Several observations can be made from Fig.5.3. The spectrum of singular values matchs the spectrum of eigenvalue magnitudes exactly, except for a scaling (which is $\sqrt{29}$). The spectrum of eigenvalue magnitudes matches exactly the DFT magnitude spectrum, thereby validating the results from Appendix.C. The same scaling is also evident in the amplitudes of the right singular vectors. We also find that the singular vectors, except the $n = 0$ component, are purely sinusoidal and occur in sine-cosine pairs as the calculations showed. We have also verified that the vector pairs are indeed orthogonal by calculating the inner product between such pairs. The singular vectors are also periodic in L while the $n = 0$ component is essentially the DC component of the data, similar to the zeroth component in Fourier transforms.

In Fig.5.4, we compare the first pair of orthogonal left singular vectors with the corresponding pair of the eigenvectors. We find that the eigenvectors are exactly the same as the singular vectors, except for a sign reversal in one of the vectors in the pair. It has been verified that the corresponding eigenvalue carries a negative sign, thus it is inconsequential to the overall analysis, and we can safely consider the equivalence between SVD and eigendecomposition in Eq.5.13 to be

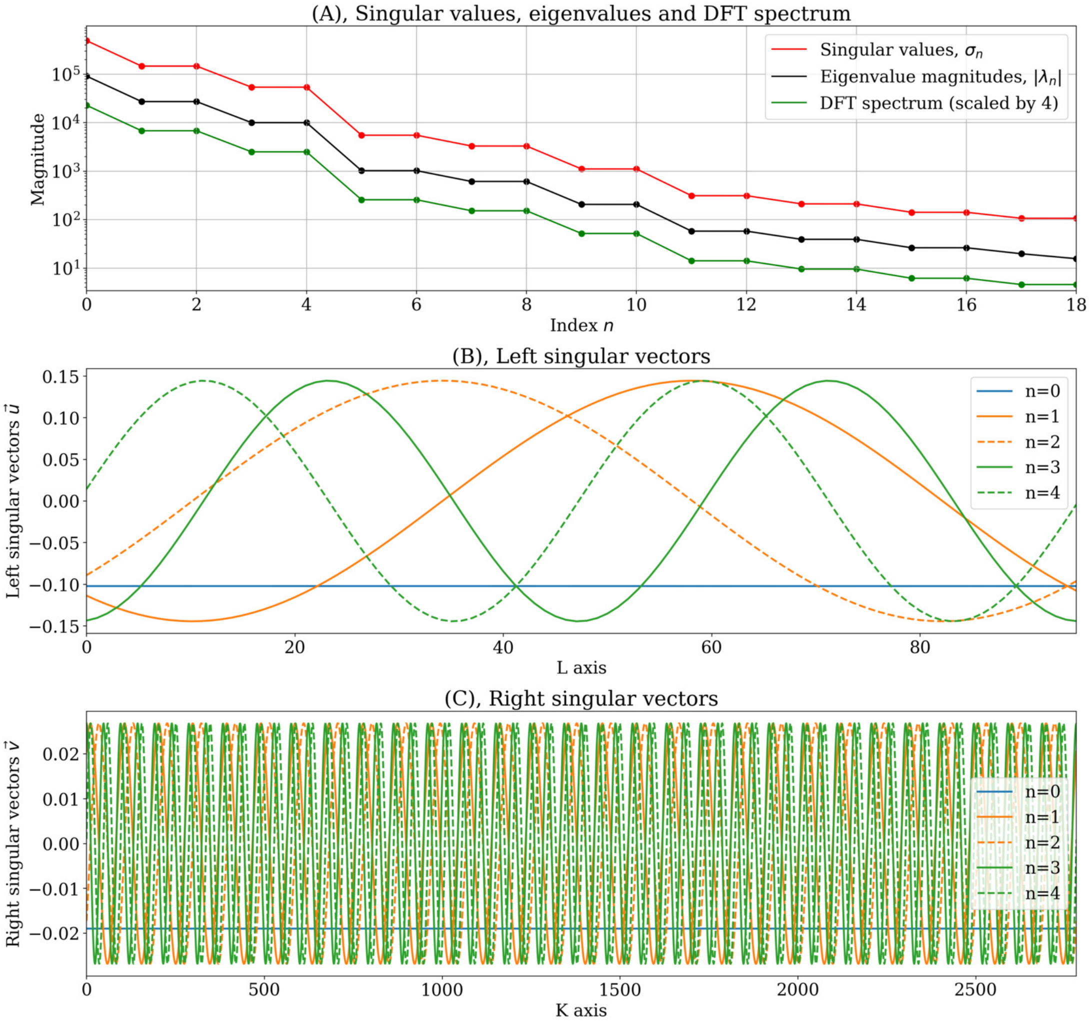


Figure 5.3: SVD of the trajectory matrix. Panel (A) shows singular values in a semilogarithmic scale. The same panel shows the eigenvalues of an $L \times L$ submatrix as well as the DFT spectrum of the periodic sequence. The DFT spectrum has been sorted according to descending magnitude and artificially scaled by a value of 4 to make it distinguishable from the eigenspectrum. For clarity, only the first 19 values are plotted. Panel (B) shows the first 5 left singular vectors while Panel (C) shows the first 5 right singular vectors. In both plots, the orthogonal (sine-cosine) vector pairs are plotted with the same colour but with different line-styles.

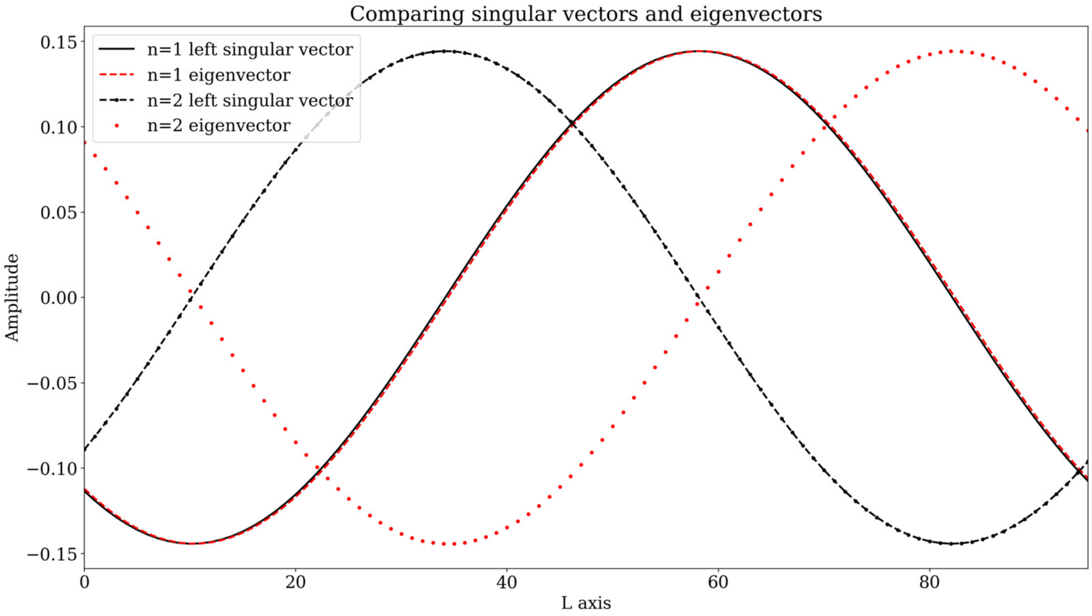


Figure 5.4: A comparison between the left singular vectors of the trajectory matrix and eigenvectors of a square submatrix. Only the vectors from the first orthogonal pair in each case are plotted.

valid.

Thus, we find that the singular vectors obtained from SSA of a periodic sequence are purely sinusoidal with them occurring in orthogonal pairs. It may be also noted that such orthogonal features have been noticed in SSA literature (see for e.g. Broomhead & King (1986); Golyandina et al. (2001)). We now inspect the reconstructed series given in Fig.5.5. As we know that the singular vectors and corresponding values occur in pairs, we group the eigentriples into pairs (except the 0^{th} component) and apply the diagonal averaging. As can be seen, the reconstructed series are also purely sinusoidal since the vectors are sinusoidal. Another important observation with Fig.5.5 is the 0^{th} component, which is equivalent to the DC term in Fourier transforms. This term, often called ”trend” in SSA literature, accounts for the sky-averaged component of the radio data and is the relevant observable for global 21–cm research.

In the above exercise, it is essential that the embedding dimension (window length) corresponds to exactly one sidereal day. If not, the decomposition will not lead to sinusoidal patterns of appropriate periodicity, and the various relations that we obtained between SVD, eigendecomposition and Fourier analysis

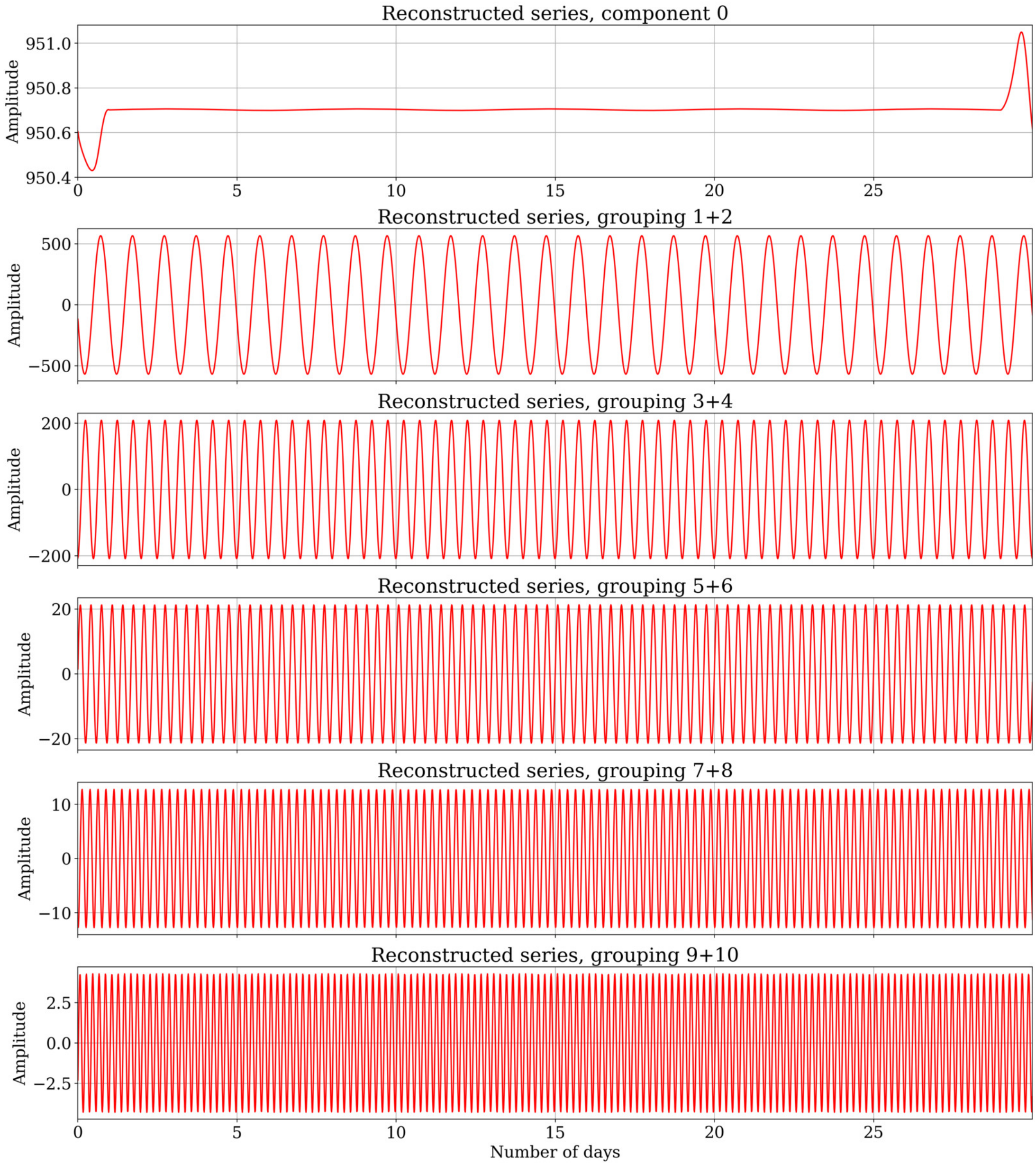


Figure 5.5: Reconstructed series from SSA of an ideal periodic series. The 0^{th} component and the succeeding 5 grouped orthogonal pairs are given in this figure. Since the reconstructed series add up to the original time series, their x-axis is the number of days.

become invalid. With this important result, we now proceed to introduce some nonidealities into the time series.

5.5.2 SSA applied to periodic time series with time-varying gains

The calculations and simulations so far assume a case where the radiometer system has been fully calibrated to yield data calibrated to a reference plane outside the Earth's ionosphere. Observational radiometric data have multiple nonidealties, the dominant ones are additive components such as receiver noise temperatures and mutliplicative factors that are often called gains. Both of them can be time-varying and therefore are required to be known to correct the data. We introduce a formalism to incorporate such non-idealities into our data. We first consider smoothly varying gains that do not have any periodic components; in other words, they are smooth on time scales of days. Gains with periodic or diurnal components, and their effects on data are discussed in later sections.

Assuming a simple system model for SITARA data (see Sec. 5.2 of Thekkeppattu et al. (2022)), the trajectory matrix of the measured data can be expressed as a Hadamard product, given in Eq.5.15.

$$\boldsymbol{X'} = (\boldsymbol{X_s} + \boldsymbol{R}) \odot \boldsymbol{G} \tag{5.15}$$

where $\boldsymbol{X_s}$ is the sky temperature trajectory matrix, $\boldsymbol{R}$ is the receiver noise temperature trajectory matrix and $\boldsymbol{G}$ is the trajectory matrix of time-varying system gains. Since the low frequency radio sky is bright, the dominant time-varying component in the data is due to sky temperature multiplied with time-varying system gains. Therefore, we can apply a simplifying assumption that the receiver noise temperature is constant with time and absorb it into the "ideal" sky matrix, and the sum can be written as $\boldsymbol{X}$. Thus we write:

$$\boldsymbol{X'} = \boldsymbol{X} \odot \boldsymbol{G} \tag{5.16}$$

While Eq.5.15 and 5.16 may be written without resorting to a matrix formulation, expressing them with trajectory matrices enables SSA.

Based on the summation of $\boldsymbol{X}$ from Eq.5.5 and Eq.5.16, the following relations can be derived.

$$\begin{aligned} \boldsymbol{X'} &= \sum_i \boldsymbol{X'_i}, \text{and} \\ \boldsymbol{X'} &= \Big(\sum_i \boldsymbol{X_i}\Big) \odot \boldsymbol{G} \\ &= \sum_i (\boldsymbol{X_i} \odot \boldsymbol{G}) \end{aligned} \tag{5.17}$$

where we have used the distributive property of Hadamard products. Thus we have:

$$\sum_i \boldsymbol{X'_i} = \sum_i (\boldsymbol{X_i} \odot \boldsymbol{G}) \tag{5.18}$$

Though Eq.5.18 appears trivial, it reveals a powerful aspect of the decomposition. Applying diagonal averaging to each $\boldsymbol{X'_i}$ leads to gain multiplied reconstructed series, as the Hadamard product is an element-by-element multiplication. How-

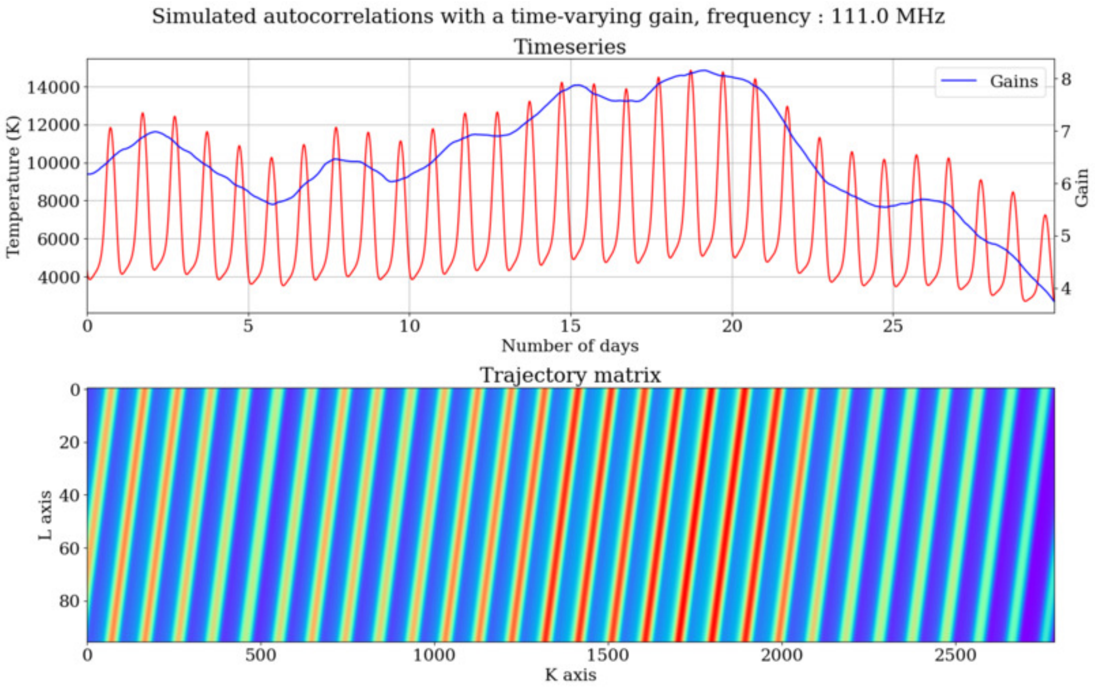


Figure 5.6: Simulated time series with gains at a frequency of 111 MHz and the corresponding trajectory matrix

ever, we already know the "true" reconstructed series to be purely sinusoidal from Sec.5.5. Therefore the gain multiplication of sinusoids resembles the process of double side-band amplitude modulation (DSBAM) in communication theory.

Before proceeding further, it is instructive to verify the above calculations with simulations. For this, we begin with the basic time series from Sec.5.5. To this series, we add a constant of 100 K as the receiver noise temperature and simulated radiometric noise (integration time of 15 minutes and bandwidth of 61 kHz) for the overall system temperature. We generate a smooth gain variation as an integral of artificially generated Gaussian white noise, which is subsequently smoothed with box-car averaging. The gains have been normalised to avoid negative excursions. The time series is then multiplied with the gains and the resulting series and the associated trajectory matrix are shown in Fig.5.6. It may be noted that the dimensions of the matrices and time series have all been kept the same. The trajectory matrix is then decomposed with SVD, and the eigentriples grouped in the same manner as in Sec.5.5 and diagonally averaged. It is easy to notice that the resulting reconstructed series shown in Fig.5.7 are indeed amplitude modulated sinusoids.

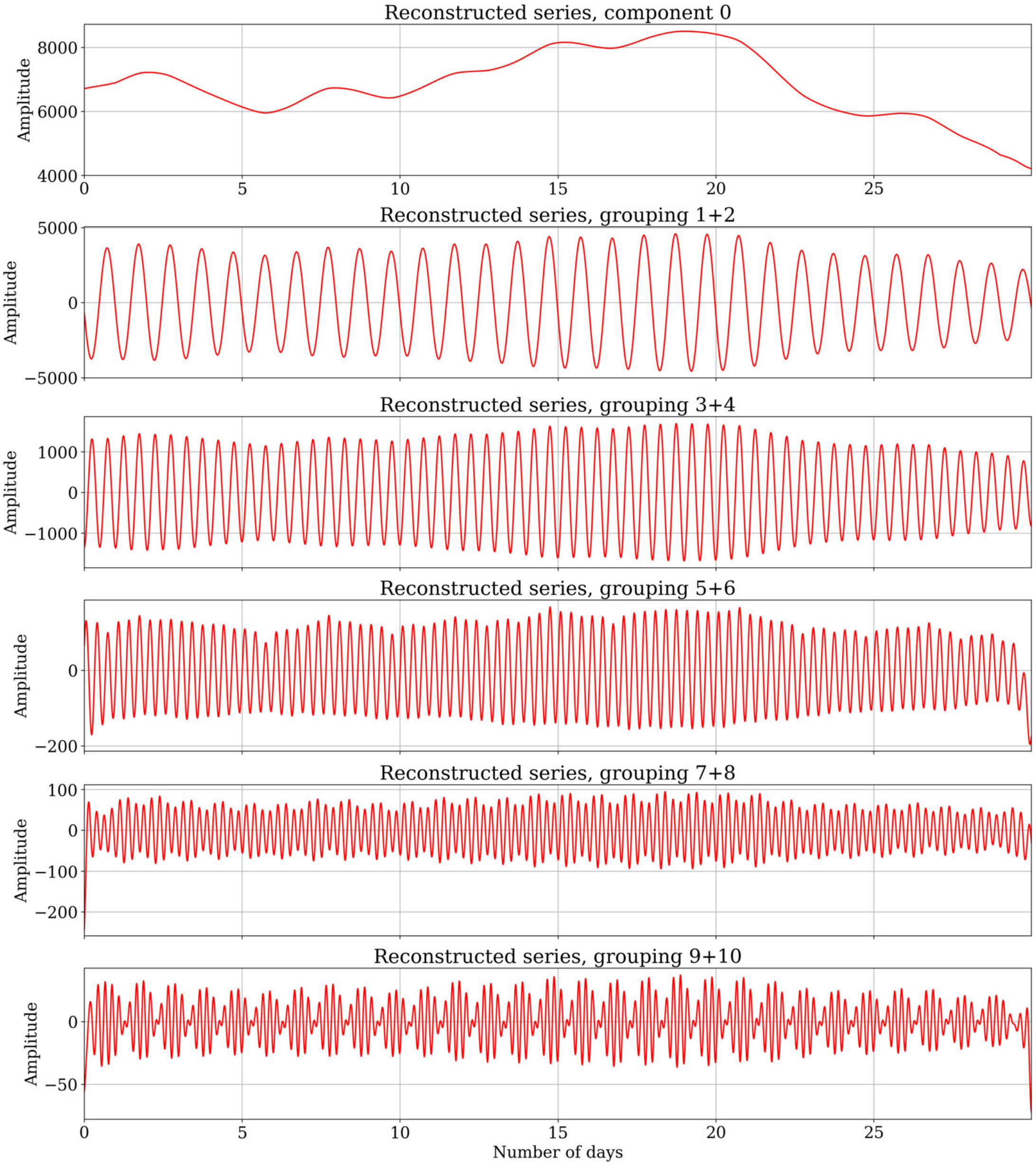


Figure 5.7: Reconstructed series from SSA of a simulated series with time-varying gains. The 0^{th} series and the succeeding 5 components from grouping orthogonal pairs are given in this figure.

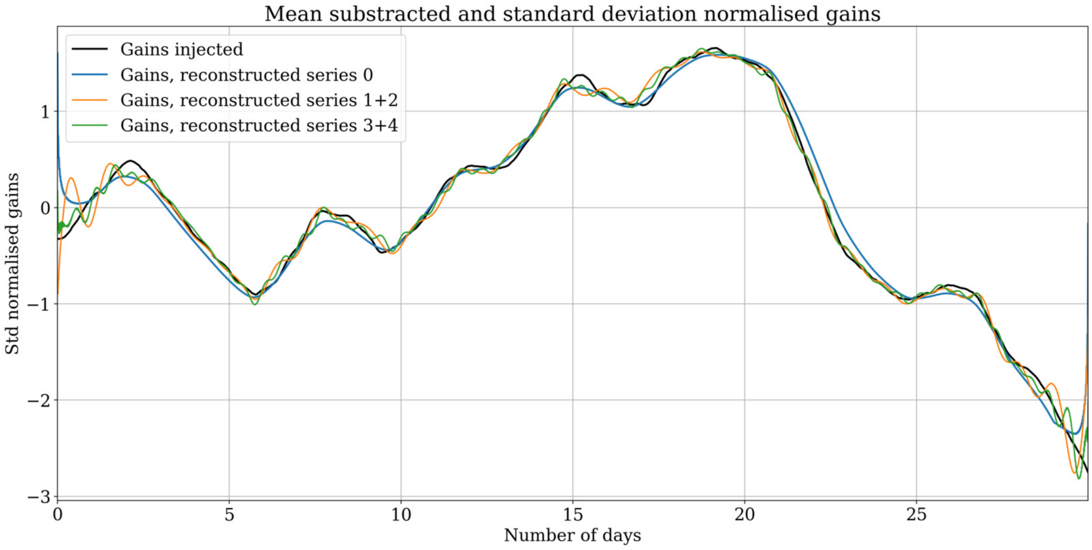


Figure 5.8: Injected and recovered gain templates. The plots have been mean subtracted and divided with their standard deviation for normalisation.

Several algorithms exist to demodulate such AM signals to obtain their envelopes. Here we use an algorithm that is mathematically simple to interpret. The first step in this algorithm is converting a real valued signal into an analytic signal, which can be accomplished with Hilbert transforms. Taking the magnitude of this analytic signal yields the modulation envelope. We apply this procedure to the series given in Fig.5.7 and the resulting gain profiles are plotted in Fig.5.8. The gain plots have been normalised with subtraction of means and division with standard deviations.

Fig.5.8 validates the interpretation of Eq.5.18 that the reconstructed series have similar gain templates that can be recovered with amplitude demodulation. Indeed, a major outcome of this exercise is that the gain template for the DC component is similar to those of the periodic components. Therefore the time-variations in the DC term may be corrected for by using the gain templates from periodic components. This leads us to a novel calibration strategy for global 21–cm radiometers that can correct long-term drift and related systematics.

5.5.3 On the use of SSA to aid calibration

The gain templates in Fig.5.8 lack normalisation constants. To make them useful for calibration, certain normalisation constants have to be incorporated. We begin with Eq.5.18 and incorporate Eq.5.5 as shown in Eq.5.19, where we also include the eigentriple grouping. The indices have also been modified for convenience.

$$\begin{aligned} \boldsymbol{X_i'} &= \boldsymbol{G} \odot \sum_{j=0}^{1} \sigma_{i+j}(\vec{u_{i+j}} \otimes_O \vec{v_{i+j}}), i = 1, 3... \\ \boldsymbol{X_i'} &= \sigma_i \boldsymbol{G} \odot \sum_{j=0}^{1} (\vec{u_{i+j}} \otimes_O \vec{v_{i+j}}), i = 1, 3... \end{aligned} \tag{5.19}$$

where we used the property that singular values are identical for the eigentriples in a grouped pair. The vectors $\vec{u}$ and $\vec{v}$ are from unitary matrices and therefore their individual inner products equal to unity. However, for subsequent analysis, these vectors need to be normalised to have unity *amplitude* when diagonally averaged after grouping. The normalisation can be achieved with the following operations, where the sinusoidal terms in the brackets have been normalised to have unity amplitude. As the lengths of a vector $\vec{u}$ is L, the scaling required to have unity amplitude (as against unity inner product) is $\sqrt{L}$. Similarly, the scaling for $\vec{v}$ is $\sqrt{K}$ as it has length K.

$$\begin{aligned} \vec{u_i} &= \frac{1}{\sqrt{L/2}} \left(\vec{u_i} \sqrt{L/2} \right) \\ \vec{v_i} &= \frac{1}{\sqrt{K/2}} \left(\vec{v_i} \sqrt{K/2} \right) \end{aligned} \tag{5.20}$$

The grouped and diagonally averaged reconstructed series is given as:

$$\widetilde{x_i}[n] = \frac{\sigma_i}{\sqrt{L/2}\sqrt{K/2}} \, g[n] c_i[n] \tag{5.21}$$

where $g[n]$ is the gain series and $c_i[n]$ is the "carrier" sinusoid. The normalisation in Eq.5.20 ensures unity amplitude for $c_i[n]$, and therefore the gain template $g_i'[n]$ from demodulation of a specific series $\widetilde{x_i}[n]$ is given as:

$$g_{i,t}[n] = \frac{\sigma_i}{\sqrt{L/2}\sqrt{K/2}} \, g[n] \tag{5.22}$$

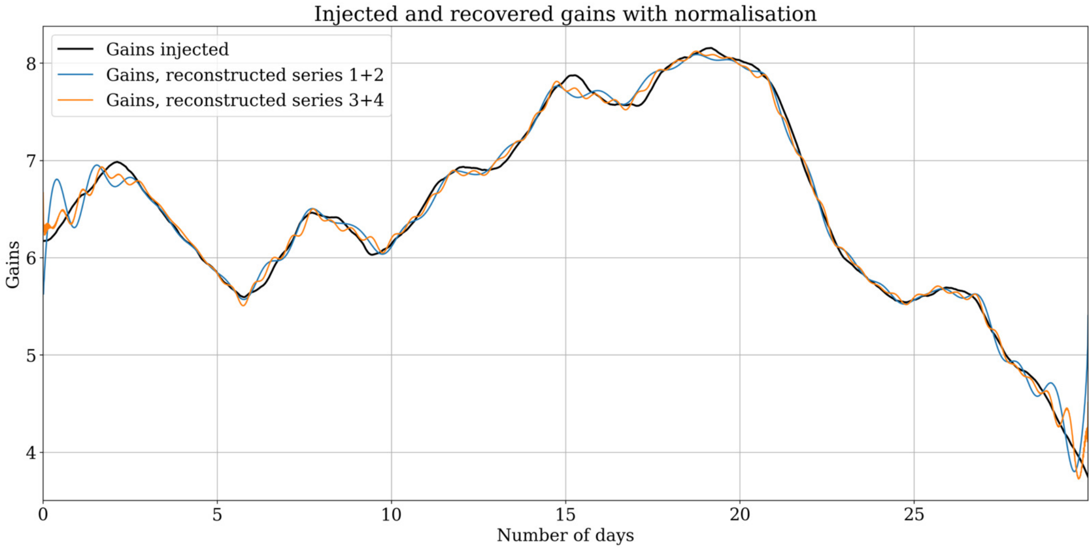


Figure 5.9: Injected and recovered gains. The recovered gains have been normalised according Eq.5.23. The recovered gains from periodic components resemble the injected gains.

Thus, the recovered gains can be written:

$$g'_i[n] = g_{i,t}[n] \frac{\sqrt{L/2}\sqrt{K/2}}{\sigma_i} \tag{5.23}$$

For Eq.5.23 to be useful for calibration, the only auxiliary information required is the singular value σ_i. This is the singular value of the component with the same periodicity that we expect from the true sky, and can be estimated from the SSA of the simulated ideal sky time series from Sec.5.5. Using those singular values, we obtain the recovered and appropriately normalised gains that are plotted in Fig.5.9. We find that the recovered gains closely resemble the injected gains. Thus, we are now in a position to apply those gains to the 0^{th} reconstructed series to achieve calibration and compare the calibration with the expected levels. The results are given in Fig.5.10. We find that the mean-sky component has been calibrated solely using the periodic component(s) of the sky with the aid of SSA. It has to be noted that the prescription provided here to calibrate the mean-sky relies only on the singular values of the simulated periodic components. Even without using the singular values, and only using the data, it is possible to achieve relative gain calibration, while applying gains with appropriate singular values to the mean-sky component establishes its proper brightness temperature scale.

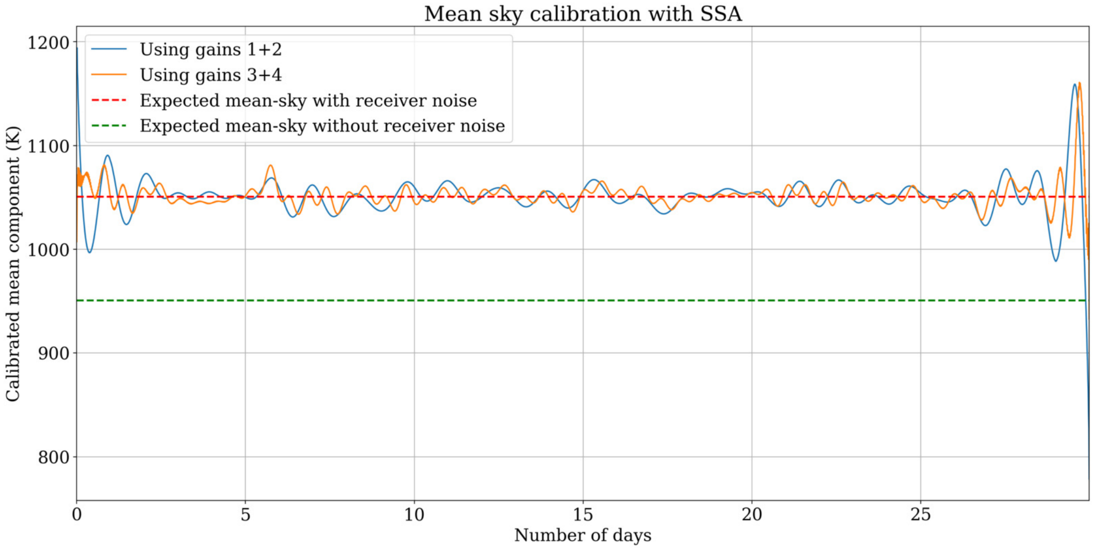


Figure 5.10: Calibrating mean sky component using recovered gains from periodic components. For comparison, the expected mean levels with and without the 100 K receiver noise temperatures are also plotted.

Let us now try to obtain an intuitive appreciation of this calibration. The trend component obtained from SSA is essentially a box-car averaging of time series data. If the window used for averaging encompasses exactly one sidereal day, the periodic variations caused by the Galaxy transiting through the beam are averaged out. If the system were perfect the result would be a constant temperature, which is the sum of mean-sky and receiver noise temperatures. Thus the trend that we observe in the non-ideal system may be modelled by a constant value multiplied with the time-varying gains.

The gain variations that we have studied so far have smooth evolution over time with no periodicity. However, for real radio telescope systems that do not have temperature regulation, gain variations with local temperature are expected - details given in Sec.5.6. Indeed, the rising and setting of the Sun can heat and cool the components in the analog signal chain and induce diurnal variations in the gain that may be correlated with the sky-drift. Yet another potential origin for diurnal patterns in gains is the ionosphere. Therefore it is imperative to study such a scenario, which we perform next.

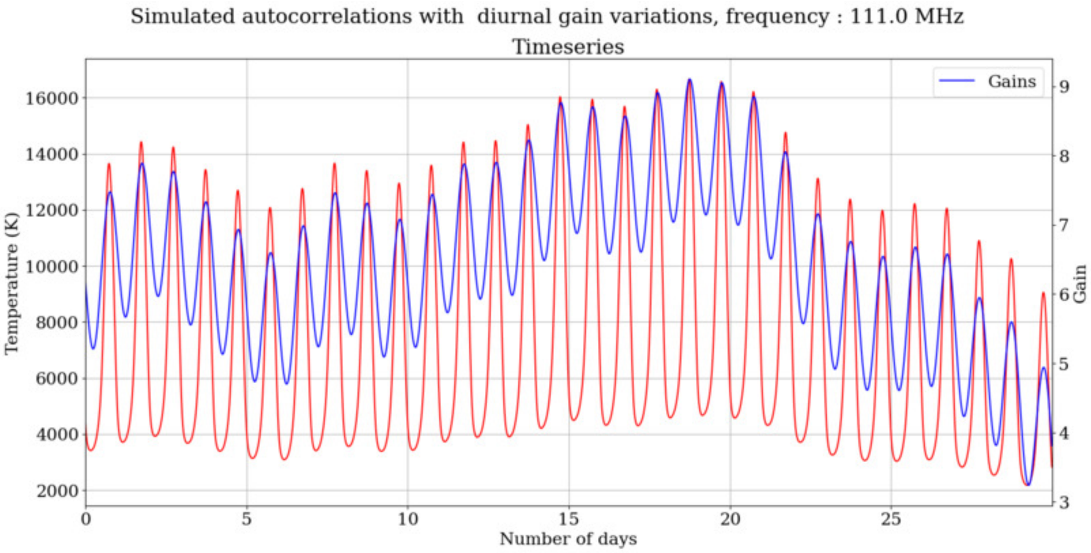


Figure 5.11: Simulated radiometer data with gains that have a smoothly varying component and a diurnal component.

5.5.4 SSA with diurnal gain variations

To the gains simulated in the previous section, we add a sinusoidal diurnal gain component. For a time series that spans a month, the diurnal gain variations are expected to show a strong correlation with Galaxy transit. For longer time series, the difference between sidereal time and civil time reduces this correlation. The modified gains and the resulting time series are shown in Fig.5.11. We perform SSA in the exact same manner as in the previous case, and recover normalised gains, as shown in Fig.5.12. It is interesting to note that the correlation between gains and sky patterns reduces the separability of the two, and the recovered gains from the reconstructed series differ from one another. We attempt a calibration of the mean component with these gains and the results are given in Fig.5.13. The differences in the recovered gains in Fig.5.12 as well as the calibrated temperatures in Fig.5.13 demonstrate a fundamental limitation of the SSA technique in calibrating data when the gain variations are correlated with sky patterns. Lacking auxiliary information on such gain variations, the calibration technique fails and yields erroneous results, as demonstrated. On the other hand, a disagreement between the recovered gains can be used as a practical indicator of gain variations that have frequencies coinciding with the periodicity induced by sky drift, thus pointing to a necessity for subsidiary information. Having demon-

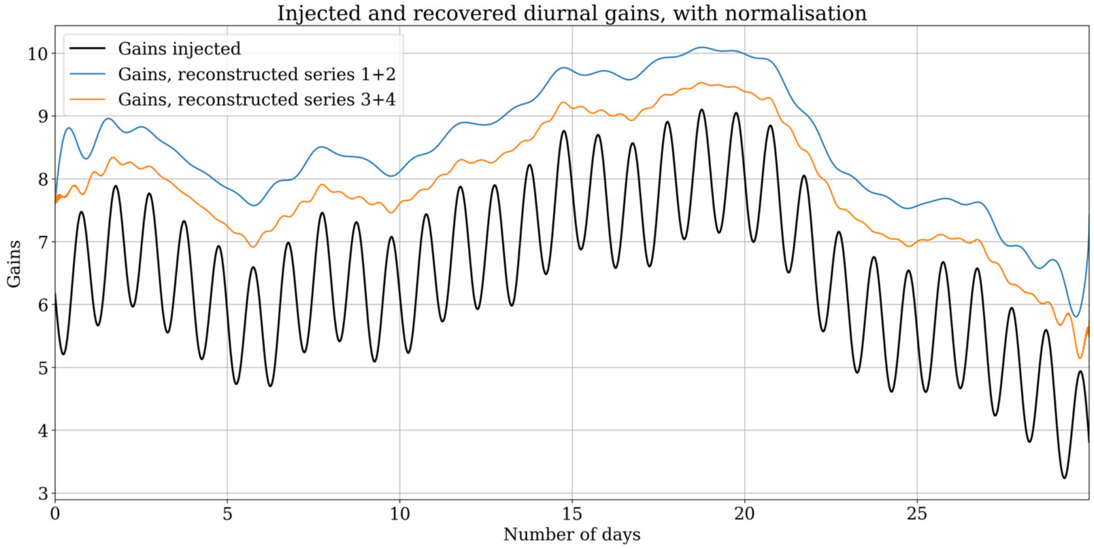


Figure 5.12: Injected and SSA recovered gains when the gains have a diurnal component. Appropriate normalisation has been applied.

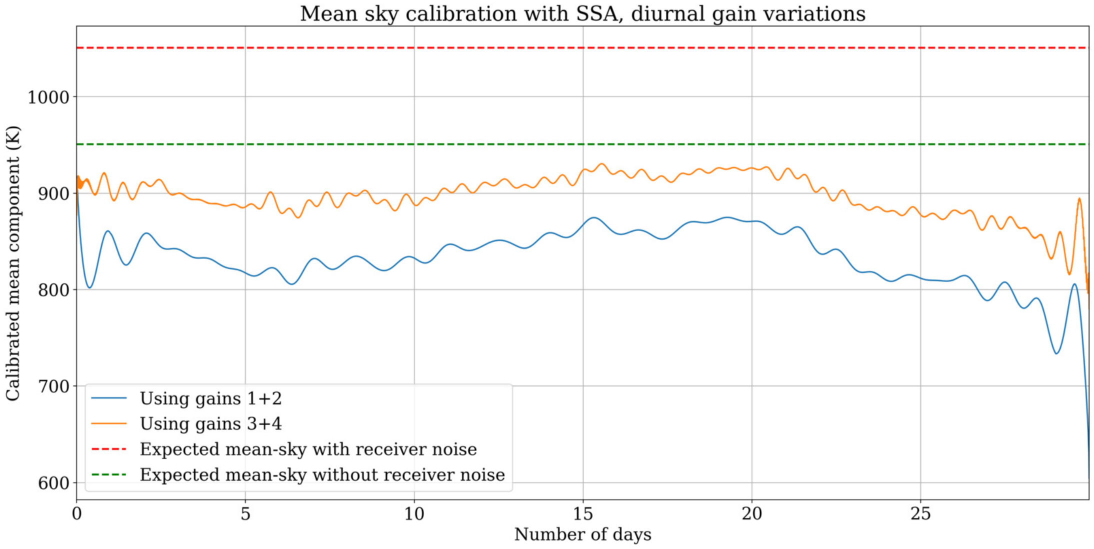


Figure 5.13: Application of recovered gains from periodic components to calibrate the mean-sky term, when the injected gains have a diurnal component.

strated the various cases of SSA, we now apply these techniques to data from SITARA observations.

5.6 SSA applied to SITARA data

5.6.1 Data preparation

We use a time series of SITARA spectral data from the concatenated 30 days of June 2021 data. The individual data files in `miriad` format are converted to `hdf5` format and concatenated using custom tools written in `python`. From the concatenated dataset, auto-correlation data for a frequency of 111 MHz are extracted, as it is a frequency with relatively low RFI occupancy. The major source of RFI observed during this period was Solar radio bursts that are transient in nature. This raw data time series as a function of Julian Date (JD) is shown in Fig.5.1.

The raw SITARA data spanning a month have about 749,000 points, which makes calculations difficult. Therefore, the data are binned into the same number of bins (2879) as done in the simulations, giving one sample per 15 minutes or 96 samples per sidereal day. Once again, in performing binning it is important to use sidereal time - and not Julian time - to ensure an equal number of samples for each cycle of sky-drift. The data are then embedded into a trajectory matrix and SSA is applied in the same manner as before. The eigentriples are grouped and reconstructed series formed.

Fig.5.14 shows the first 4 groups of the reconstructed series. The trend plot also has a low time cadence record of the temperature recorded within the Murchison shire, which is available from NCEI-NOAA [2]. The temperature scale has been inverted, as the trend shows an anti-correlation with temperature. Further, the radio data and temperature data fall into two different families with different sampling, and therefore the temperature data have been interpolated to the same Julian day bins as the radio data. The Pearson correlation coefficient computed between the trend and the temperature data is -0.83, which denotes a strong

---

[2] https://www.ncei.noaa.gov/access/search/index

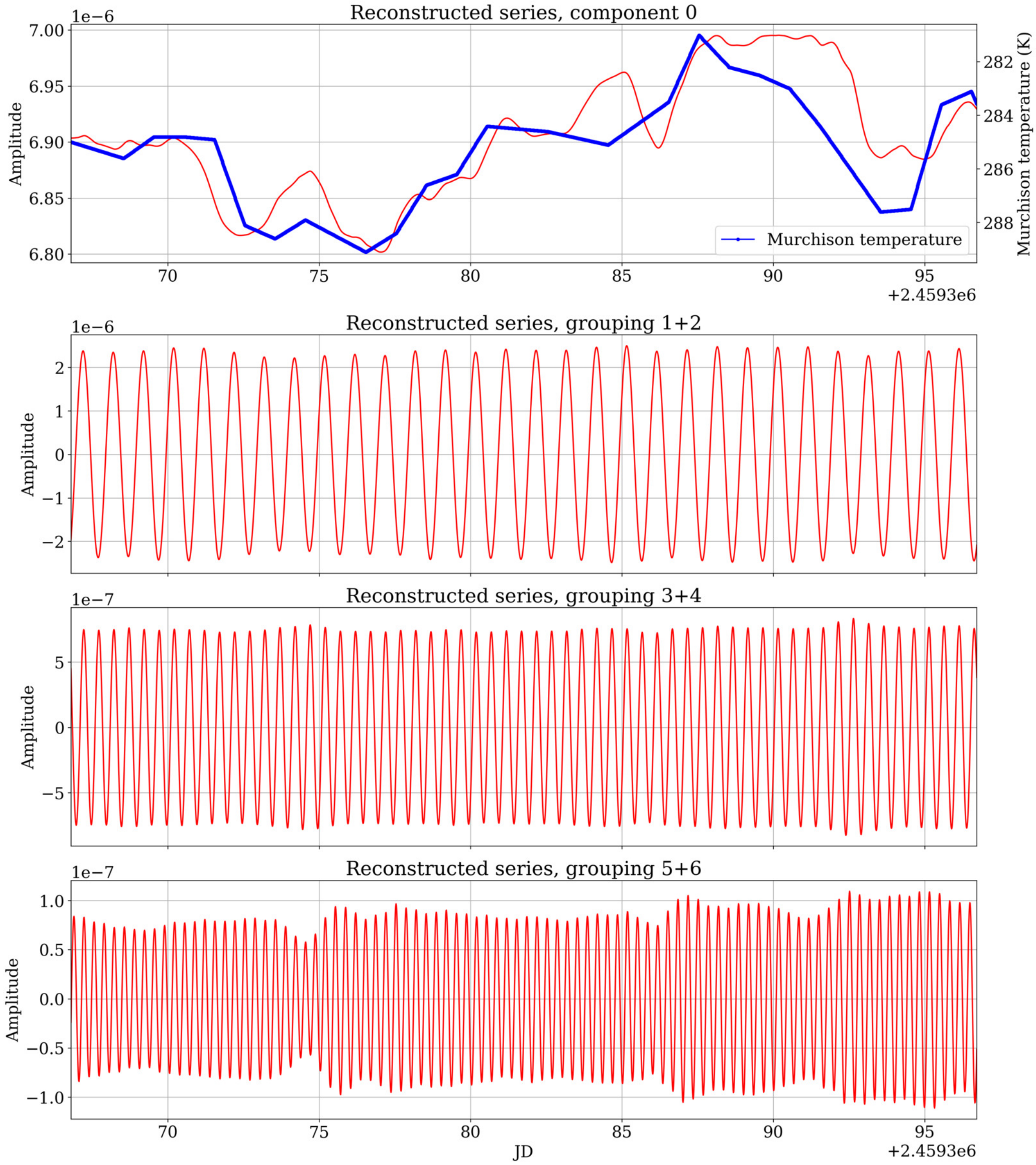


Figure 5.14: Reconstructed series from SITARA data SSA. A plot of the physical temperature recorded within the Murchison shire is included in the trend plot to show the anti-correlation between both.

anti-correlation between the trend and temperature.

Apart from component ageing effects, there are two major reasons for such temperature-induced variations in radiometric data.

1. The noise temperatures of active devices used in amplifiers increase with physical temperature[3]. The same holds true for passive components such as attenuators. If this is the cause of drift, the pattern so obtained is expected to be *correlated* with physical temperature.

2. The gain of amplifiers reduces as temperature is increased. In this case, the trend pattern and physical temperature would be *anti-correlated.*

It is therefore evident from Fig.5.14 that the major contributor to the trend is temperature-induced gain variations. The receiver noise would inevitably vary as a function of physical temperature, however when the overall system temperature is sky-dominated the impact of this would be secondary to gain variations. This informs our choice of calibration model given in Sec.5.5.3, where we assume a constant receiver noise temperature and a time-varying gain.

Subsequently, the gain patterns are recovered from the reconstructed series with the demodulation technique. These recovered gains for the first two periodic components are shown in Fig.5.15. As can be seen, the normalised gains are different between the series, thus pointing to potential diurnal gain variations. Therefore, we refrain from applying these gains to the mean component to establish its brightness temperature scale.

5.7 Discussion

In this chapter we developed the mathematical framework for SSA of a radiometric time series and demonstrated its application in analysing radiometric time series data from a radio telescope at a frequency with relatively low RFI. A major outcome of this work is a novel method to calibrate the mean-sky component of radio data using the periodic component of sky-drift patterns. For this,

---

[3] It may be noted that this is one of the reasons for cryogenic cooling of radio telescope front-ends to achieve a low overall system temperature.

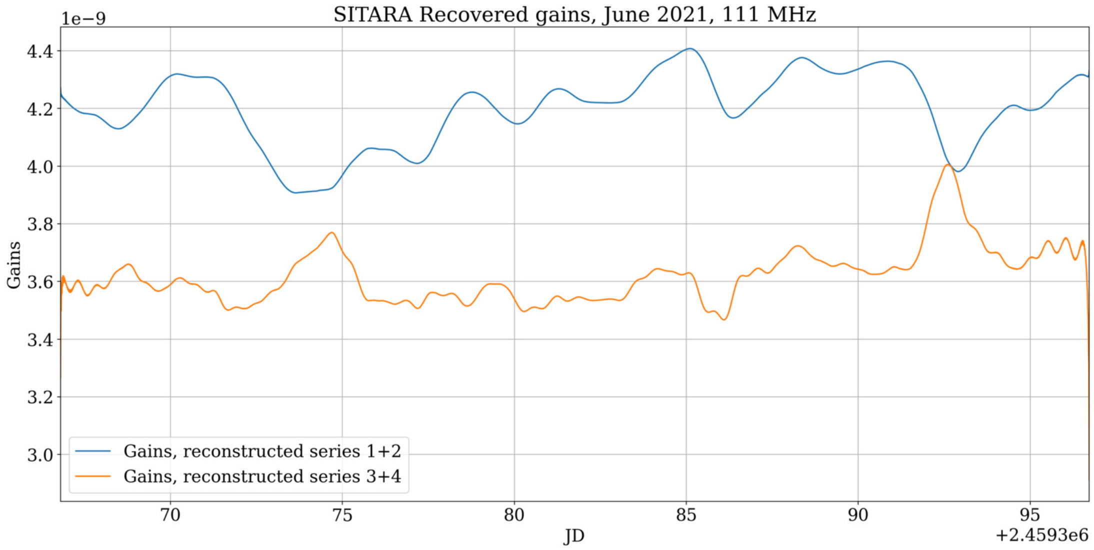


Figure 5.15: Recovered gains from SITARA data, June 2021. The gains differ across reconstructed series, pointing to potential diurnal gain variations.

let us examine the implications of the gain recovery in Fig.5.9. The gain patterns that we recovered are purely from the periodic components of the sky and the singular values of simulated periodic components. The procedure does not need simulations that include the mean-sky or the *zero-point* of the sky maps. Indeed, zero-point errors have been noticed in sky maps, since these maps are often products of combination of sky surveys conducted with different telescopes. Disagreements exist between estimates of the zero-points of sky maps across experiments, for example, Monsalve et al. (2021) report a different zero-point offset for the 150 MHz sky map (Landecker & Wielebinski, 1970) compared to the measurements of Patra et al. (2015). Therefore, the methods outlined here can be applied to calibrate experiments where zero-point levels have to be accurately known to verify that the gains do not have variations that can confuse sky signals. However, in this case the goodness of calibration depends on the zero-levels being consistent throughout the sky. If the zero-levels in the sky maps used for calibration vary across the maps, the result would be a mis-calibration of the observed data.

5.7.1 Caveats and future work

Since this is the first work exploring SSA for radiometric data analysis and calibration, we have kept the analysis simple. Specifically, low frequency radio data can have RFI and we ignored the flagging that has been performed. Flagging considerations lead to gaps in a time series dataset and may necessitate data in-painting or interpolation. Such explorations will be taken up in future work.

We have developed the mathematical background and demonstrated an application of SSA for auto-correlations at a single frequency, while SITARA has nearly two thousand usable frequency channels in auto-correlations and cross-correlations each. For application of SSA to global 21 cm research, it is essential to have capabilities to perform broadband analysis and calibration of time series data. However, to use all the frequency channels and/or consider complex cross-correlations, the mathematical framework developed in this chapter has to be expanded. Specifically, when multi-frequency data are considered, the trajectory matrix will become three dimensional; such a multidimensional structure can be thought of as a ”tensor” in a signal processing sense Cichocki et al. (2015). For a proper treatment of multiple frequency channels, multivariate SSA techniques will be explored in future work.

While we demonstrated via simulations a calibration technique to remove gain variations that evolve smoothly with time, the application of it to SITARA data is hampered by diurnal gain variations. This shows the necessity to maintain temperature regulation or inclusion of thermometers at points in the signal chain to track such gain variations - this will be incorporated in future revisions of the SITARA system. Additionally, the data in frequency channels below 50 MHz, where receiver temperature dominates over sky temperature, show diurnal periodicity arising from such gain variations which may be exploited as a template to improve calibration.

5.7.2 A potential application of SSA for space-based 21–cm experiments

As we demonstrated, a major limitation with the SSA technique for calibration is the confusion between diurnal gain variations and sky drift, as both have approximately the same periodicity. If the periodicities can be made to differ, the calibration can be improved significantly. Specifically, if the sky-drift can be made faster than the gain variations, substantially better calibration can be expected. While it is difficult for ground-based experiments to introduce such a separation, a radiometer payload on a spin-stabilized satellite can have beams that rapidly scan the sky, thereby increasing the sky-drift rate. For example, a broadband dipole antenna could be placed on a satellite spinning such that the nulls sweep the Galactic plane at a rate much faster than any gain variations, including the flicker noise of the electronic systems. Spinning a satellite at a rate faster than than the $1/f$ knee frequency of the radiometric system, to reduce the deleterious impact of gain fluctuations on the images, has been employed in CMB missions such as Planck (Bersanelli et al., 2010). As there are a few projects proposed, planned or launched targeting the global 21–cm signal from a satellite platform such as DARE (Burns et al., 2012), Longjiang/Chang'e-4 (Jia et al., 2018), DSL (Chen et al., 2021), PRATUSH etc., we opine that SSA would be an ideal tool for analysis and calibration of time series data from such radiometers in space. For ground based observations, by placing an antenna on a rotating pedestal with a highly stable periodicity such that the beam sweeps various parts of the sky, a periodicity can be imparted to the measured data that is different to diurnal variations.

5.8 Conclusions

In this chapter, we introduced singular spectrum analysis (SSA) as a powerful tool to study radiometric data. We showed the deep connections between the SSA techniques and Fourier transforms and leverage that to obtain long-term gain evolution from radio data. A novel technique to calibrate 21–cm experiments using periodicity in the sky drift patterns has been proposed and simulated. The limitations of that technique in the presence of diurnal variations that can confuse with sky-drift is studied. Upon application of SSA to SITARA data, we find that the obtained decomposition has the features as expected, with the trend showing strong anti-correlation with temperature. However, the gains obtained point to diurnal variations and pose a limiting factor in using SSA for gain calibration of SITARA data.

Acknowledgements

This work is funded through CT's ARC Future Fellowship, FT180100321 - "Unveiling the first billion years: enabling Epoch of Reionisation science". This work makes use of the Murchison Radio-astronomy Observatory, operated by CSIRO. We acknowledge the Wajarri Yamatji people as the traditional owners of the Observatory site. JNT acknowledges Dr. Adrian T. Sutinjo for his Linear Algebra over Lunch lectures. This work uses the following python packages and we would like to thank the authors and maintainers of these packages :- `numpy` (Harris et al., 2020), `scipy` (Virtanen et al., 2020), `healpy` (Zonca et al., 2019), `astropy` (Astropy Collaboration et al., 2013, 2018), `aipy` (Parsons, 2016), `matplotlib` (Hunter, 2007), `ephem` (Rhodes, 2011), `pygdsm` (Price, 2016) and `h5py` (https://www.h5py.org/).

Chapter 6

Summary and conclusions

Don't think about why you question, simply don't stop questioning.

- Albert Einstein

In this chapter, I summarise outcomes from this thesis. Short descriptions of results from the two different experiments, as well as the novel data analysis technique form this chapter. I end with a note on what the future holds for global 21–cm experiments.

6.1 Single antenna experiments

Single antenna experiments constitute the most common method to constrain the global 21–cm signal. This thesis shows that the receivers intended for the same can be made to have maximally smooth responses for their transfer functions and any residual systematics can be suppressed to mK levels. Key to ensuring a maximally smooth response is the adoption of a simple set of rules; keeping the path lengths in the front-end analog electronics as short as possible, adopting RFoF technology for galvanic isolation between the analog front-end electronics and "noisier" digital back-end electronics and a hierarchical double differencing scheme to cancel out spectral structures. A detailed measurement equation for the response of a Dicke-switched radiometer, taking into account multiple reflections, has been derived. It is also shown with laboratory tests that when combined with antennas that have smooth responses, the SARAS 3 system can constrain the global 21–cm signal. The SARAS 3 system described in this thesis has resulted in a non-detection of the 21–cm absorption signal reported by EDGES.

6.2 Short-spacing interferometers

In this thesis, it is shown that interferometers with closely spaced antennas are indeed sensitive to the averaged sky. The coherence of such interferometers can be well approximated by the visibility equation, taking into account the effect of mutual coupling on the beams. Such a simple model can approximate the measured data, including most of the variations seen as functions of LST as well as frequency. However, despite accurate beam simulations, differences are seen between the measured data and simulations, that are attributed to excess cross-talk effects not captured by the simulations. Nonetheless, the cross-talk can be modelled in a linear-algebraic sense. Such a model can be fit with data having antenna temperature variations induced by the transit of hot and cold regions of the radio sky. Regardless of the model used - with or without crosstalk - modelling the cross-correlations requires inclusion of coupled receiver noise. Though this shows that the basic assumption in standard interferometry (at long baselines) of lack of receiver noise in cross-correlations fails with short-spacing interferometry, the cross-coupled noise is significantly lower than the receiver noise in auto-correlations. Therefore, it is still advantageous to employ short-spacing interferometry for global 21–cm experiments.

6.3 Calibration and data analysis techniques

Data analysis techniques for global 21–cm always concentrated on *averaging* the data. For the same reason, leveraging the highly predictable periodicity of the Galaxy transit to improve calibration, or to understand systematics better, has not been given attention. However, real data have variations induced by other effects too. In this thesis, SSA has been explored as a powerful tool to separate periodic data from average data, and to use the obtained periodic patterns to understand the system better. While theoretically long term drift can be corrected using this approach, diurnal patterns that are correlated with sky drift on short timescales are seen to be a roadblock with this approach.

6.4 The future of global 21–cm measurements

The detection of an absorption profile by the EDGES experiment has resulted in a flurry of activity in global 21–cm research. The profile, albeit being anomalous in shape and strength, has led to novel ideas about the early Universe and theoretical efforts have been channelled into hitherto unexplored physical processes in the early Universe. However, despite the individual systems having exquisite calibration, the EDGES absorption profile detection could not be replicated by SARAS 3, a system with a different design philosophy. This non-detection of the EDGES profile poses some serious questions to an experimenter, regarding the origin of the profile. Is the signal of instrumental origin, or is it external to the system? If external, what are the roles of ionosphere, polarised radio foregrounds, low-lying RFI etc.? If internal, why does such a systematic elude calibration? Future global 21–cm experiments need to consider these questions, and in this context, it is worthwhile outlining some directions for future global 21–cm experiments.

When it comes to single antenna experiments, if a receiver system is designed with maximally smooth transfer functions (like SARAS 3), and with subsystems for in-situ calibration (like EDGES), the ambiguities associated with a detection can be substantially reduced. Perhaps the ultimate frontier of single antenna global 21–cm research is such a system in a Lunar orbit, far from anthropogenic RFI and the ionosphere. The question then remains as to what if one reaches the limits of single antenna experiments, specifically whether system calibration can be made near perfect. Rather than measuring the receiver noise temperature - a process that requires inclusion of precision hardware calibrators into the signal chain - cross-correlation between antennas to reduce receiver noise is a viable alternative. Despite the effects of mutual coupling hampering data analysis, the reduction in some systematics can help in exploring parameter spaces unreachable by single antenna experiments. In this case, if the signal chains include in-situ calibrators such as correlated noise sources, the calibration of the system may be comparable or better than that of single antenna experiments. Also, the use of multiple antennas and closure relations is a niche area for the interferometric approach to global 21–cm research. However, the complexity of such a system may offset the gains.

Based on the work given in this thesis, I opine that the future of global 21–cm research relies on synergy between multiple instruments, each one working with a different philosophy.

Appendices

Appendix A

Derivation of the SARAS 3 measurement equation

The Dicke switch in the SARAS 3 receiver alternately connects the radiometer receiver to the antenna and to a reference load. The reference load is a noise source followed by an attenuation, so that the reference noise temperature may be switched between ambient and high temperature states depending on whether the noise source is on or off, while maintaining the impedance of the reference constant.

We first consider the case in which the Dicke switch is connected to the antenna and define the following terms to describe the noise model:

- Z_N is the input impedance of the low noise amplifier (LNA) and Γ_N is the reflection coefficient of the LNA as referred to a $Z_0 = 50\ \Omega$ measuring system impedance.

- Z_A is the input impedance of the antenna and Γ_A is the reflection coefficient of the antenna, again assuming a $Z_0 = 50\ \Omega$ impedance for the measuring instrument.

- G is the power gain of the front-end amplifier in the radiometer.

- V_A is the voltage from the antenna terminals that is coupled into the transmission line, which is assumed to be of $Z_0 = 50\ \Omega$ impedance. V_A is a voltage waveform in the transmission line at the LNA input resulting from the coupling of antenna temperature to the line.

- V_N is the voltage of the noise wave generated in the LNA, referred to the amplifier input.
- f is the fraction of that noise wave voltage that gets coupled in the reverse direction into the $Z_0 = 50\ \Omega$ transmission line connecting the antenna to the amplifier, which is of length l.

The amplifiers connected to the antenna in SARAS 3 are all in a compact module that is followed immediately by an optical modulator and hence is optically isolated from all electronics that follows. Therefore, the total amplification—that of the first low-noise amplifier, a second amplification stage that follows, and the amplifier associated with the optical modulator—may be treated as lumped, referred to as the front-end amplifier, and represented by a single noise wave V_N. The analysis parameters are depicted in Fig. A.1.

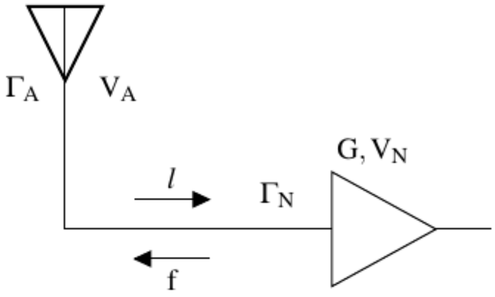


Figure A.1: Simplified noise model for the SARAS 3 radiometer, when connected to the antenna.

Taking into account the first order reflections of front-end amplifier noise from the antenna, the voltage V_S at the input of the amplifier can be written as:

$$\begin{aligned} V_S &= V_A(1+\Gamma_N) + V_N(1+\Gamma_N) + fV_N\Gamma_A e^{i\phi}(1+\Gamma_N) \\ &= V_A(1+\Gamma_N) + V_N(1+\Gamma_N)[1+f\Gamma_A e^{i\phi}], \end{aligned} \tag{A.1}$$

where ϕ is the phase difference between the forward propagating wave and the reflected wave due to the finite length l of the transmission line connecting the amplifier and the antenna. ϕ and l are related as $\phi = (4\pi\nu l)/(v_f c)$ where c is the speed of light in vacuum and v_f is the velocity factor of the transmission line.

Taking into account the power gain of the amplifier, the time-averaged power flow out of the front-end amplifier is:

$$\begin{aligned} P_S &= \left\langle G\ Re\Big(\frac{V_S V_S^*}{Z_N}\Big)\right\rangle \\ &= G\ Re\{P_A[1 - 2iIm(\Gamma_N) - |\Gamma_N|^2] \\ &\quad +P_N[1 - 2iIm(\Gamma_N) - |\Gamma_N|^2][1 + f\Gamma_A \mathrm{e}^{i\phi}][1 + f^*\Gamma_A^* \mathrm{e}^{-i\phi}]\} \\ &= G\ (1 - |\Gamma_N|^2)\{P_A + P_N[1 + 2Re(f\Gamma_A \mathrm{e}^{i\phi}) + |f|^2|\Gamma_A|^2]\}, \end{aligned} \tag{A.2}$$

where the following definitions are used

$$P_A = \left\langle \frac{V_A V_A^*}{Z_0} \right\rangle \tag{A.3}$$

and

$$P_N = \left\langle \frac{V_N V_N^*}{Z_0} \right\rangle. \tag{A.4}$$

The relation $Z_N(1-\Gamma_N) = Z_0(1+\Gamma_N)$ is used in the above derivation to express Z_N in terms of Z_0.Additionally, it may be noted here that P_A represents the available power from the antenna: the power corresponding to the antenna temperature that couples into the transmission line, with characteristic impedance Z_0,connecting to the receiver. P_N corresponds to the receiver noise, referred to the input of the LNA.

The correlation receiver response contains unwanted additives as a result of coupling of any common mode self-generated RFI or noise into the two arms of the correlation receiver. For example, the samplers on the digital receiver board that digitise the analog signals of the two arms would inevitably have common mode noise of the digital board, which results in an unwanted additive component in the response. This additive is expected to be constant in time and we denote the net unwanted common mode response as P_{corr}. With this power included, the measurements in each of two switch states, P_{OBS00} and P_{OBS11}, are:

$$P_{OBS00} = -G\ (1 - |\Gamma_N|^2)\{P_A + P_N[1 + 2Re(f\Gamma_A \mathrm{e}^{i\phi}) + |f|^2|\Gamma_A|^2]\} + P_{corr}$$

and

$$P_{OBS11} = G\ (1 - |\Gamma_N|^2)\{P_A + P_N[1 + 2Re(f\Gamma_A \mathrm{e}^{i\phi}) + |f|^2|\Gamma_A|^2]\} + P_{corr}.$$

Their difference P_{OBS} is:

$$\begin{aligned} P_{OBS} &= P_{OBS11} - P_{OBS00} \\ &= 2G\ (1 - |\Gamma_N|^2)\{P_A + P_N[1 + 2Re(f\Gamma_A \mathrm{e}^{i\phi}) + |f|^2|\Gamma_A|^2]\}. \end{aligned} \tag{A.5}$$

Consider the case in which, instead of an antenna, an impedance matched $Z_0 = 50\ \Omega$ calibration noise source is connected. This also serves as an ambient temperature reference termination when the noise source is off. As there is no mismatch between the transmission line and noise source or reference termination, the noise wave from the amplifier that is coupled into the transmission line is absorbed at the calibration noise/reference termination. The analysis parameters in this case are depicted in Fig. A.2. For this case, we may write the time-averaged power flowing out of the front-end amplifier as:

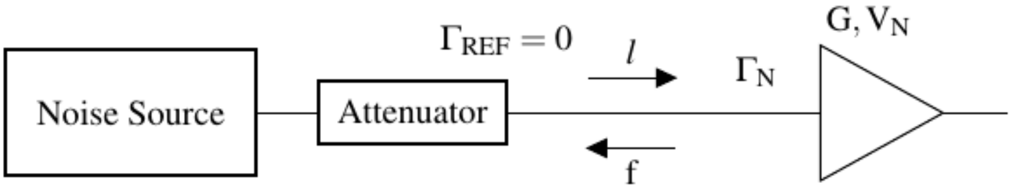


Figure A.2: Simplified noise model for the SARAS 3 radiometer, when connected to the reference termination and calibration source.

$$P'_S = G\ (P_{REF} + P_N)(1 - |\Gamma_N|^2) \tag{A.6}$$

when the noise source is off and

$$P_S'' = G\ (P_{CAL} + P_N)(1 - |\Gamma_N|^2) \tag{A.7}$$

when the noise source is on. In deriving this, we have used Eq. A.2 and set $\Gamma_A = 0$ and replaced P_A with P_{REF} or P_{CAL}. P_{REF} represents the noise power from the reference termination that couples into the transmission line, and P_{CAL} that from the termination when the calibration source is on. We may write:

$$P_{REF} = \left\langle \frac{V_{REF} V_{REF}^*}{Z_0} \right\rangle, \tag{A.8}$$

where V_{REF} is the noise voltage from the reference port, and

$$P_{CAL} = \left\langle \frac{V_{CAL} V_{CAL}^*}{Z_0} \right\rangle, \tag{A.9}$$

where V_{CAL} is the noise voltage from the reference port when the calibration noise source is on.

Taking into account the unwanted common-mode noise from the digital boards, which are inevitably added, the measurement data provided by the correlation receiver in each of the states CAL00, CAL01, CAL10 and CAL11 may be written as:

$$P_{CAL00} = -G(P_{REF} + P_N)(1 - |\Gamma_N|^2) + P_{corr}, \tag{A.10}$$

$$P_{CAL01} = G(P_{REF} + P_N)(1 - |\Gamma_N|^2) + P_{corr}, \tag{A.11}$$

$$P_{CAL10} = -G(P_{CAL} + P_N)(1 - |\Gamma_N|^2) + P_{corr} \tag{A.12}$$

and

$$P_{CAL11} = G(P_{CAL} + P_N)(1 - |\Gamma_N|^2) + P_{corr}. \tag{A.13}$$

Differencing the measurements recorded in the two switch positions gives:

$$\begin{aligned} P_{CAL0} &= P_{CAL01} - P_{CAL00} \\ &= 2G(P_{REF} + P_N)(1 - |\Gamma_N|^2) \end{aligned} \tag{A.14}$$

and

$$\begin{aligned} P_{CAL1} &= P_{CAL11} - P_{CAL10} \\ &= 2G(P_{CAL} + P_N)(1 - |\Gamma_N|^2). \end{aligned} \tag{A.15}$$

The correlation spectrometer thus provides three differenced measurements: P_{OBS} corresponding to when the antenna is connected to the receiver, P_{CAL0} when the reference is connected, and P_{CAL1} when the calibration noise is on. Together with T_{STEP}, these yield a calibrated measurement of the antenna temperature:

$$\begin{aligned} T_{meas} &= \frac{P_{OBS} - P_{CAL0}}{P_{CAL1} - P_{CAL0}} T_{STEP} \\ &= \frac{T_{STEP}\Big[P_A - P_{REF} + P_N[2Re(f\Gamma_A e^{i\phi}) + |f|^2|\Gamma_A|^2]\Big]}{\Big[(P_{CAL} - P_{REF})\Big]}. \end{aligned} \tag{A.16}$$

This Eq. A.16 may be written in the form

$$\begin{aligned} T_{meas} &= T_{STEP}\Big[\frac{P_A - P_{REF}}{P_{CAL} - P_{REF}}\Big] + \\ & T_{STEP}\Big[\frac{P_N}{P_{CAL} - P_{REF}} \times \Big\{2|f||\Gamma_A|cos(\phi_f + \phi_A + \phi) + |f|^2|\Gamma_A|^2\Big\}\Big], \end{aligned} \tag{A.17}$$

where ϕ_f is the phase associated with the complex f and ϕ_A is the phase associated with the scattering parameter S11 of the antenna.

So far, we have considered only first order reflection of the LNA noise from the antenna, which introduces sinusoidal standing waves with a single period within the transmission line and, consequently, sinusoidal modulation of the measured spectrum with a single period. However, reflections of the LNA noise as well as the antenna signal that occurs at the input of the LNAs leads to higher order reflections and standing waves in the transmission line. We now proceed to quantify these reflections and associated spectral structure.

We begin with Eq.A.1 and introduce higher order reflection terms. For clarity, we split the voltage at the input of the LNA into two parts, a part originating in the antenna and a second part corresponding to the LNA noise, and superpose the responses to get the resultant. Since the antenna signal and noise from the LNA are uncorrelated, this separation can be extended to the powers as well.

The voltage due to the antenna, denoted as V_{SA}, may be written as:

$$\begin{aligned}
V_{SA} = {} & V_A(1+\Gamma_N) + V_A(\Gamma_N\Gamma_A \mathrm{e}^{i\phi})(1+\Gamma_N) + V_A(\Gamma_N^2\Gamma_A^2 \mathrm{e}^{i2\phi})(1+\Gamma_N).... \qquad \text{(A.18)} \\
& + V_A(\Gamma_N^n\Gamma_A^n \mathrm{e}^{in\phi})(1+\Gamma_N) + ... \\
= {} & V_A \sum_{n=0}^{+\infty} (\Gamma_N\Gamma_A \mathrm{e}^{i\phi})^n (1+\Gamma_N). \qquad \text{(A.19)}
\end{aligned}$$

The time averaged power flow out of the system, due to the signal from the antenna, can be written as:

$$\begin{aligned}
P_{SA} &= \Big\langle G\ Re\Big(\frac{V_{SA}V_{SA}^*}{Z_N}\Big)\Big\rangle \\
&= \Big\langle G\ Re\Big[\frac{V_A V_A^*(1+\Gamma_N)(1+\Gamma_N^*)\{\sum_{m=0}^{+\infty}(\Gamma_N\Gamma_A \mathrm{e}^{i\phi})^m\}\{\sum_{n=0}^{+\infty}(\Gamma_N^*\Gamma_A^* \mathrm{e}^{-i\phi})^n\}}{Z_N}\Big]\Big\rangle .
\end{aligned} \qquad \text{(A.20)}$$

Using Cauchy product to evaluate the product of the two infinite series, the above expression may be simplified to:

$$\begin{aligned}
P_{SA} = {} & G\ Re\Big[P_A\{1 - 2iIm(\Gamma_N) - |\Gamma_N|^2\}\{\sum_{k=0}^{+\infty} |\Gamma_N|^k|\Gamma_A|^k\}\{\sum_{l=0}^{k} e^{i(2l-k)(\phi_N+\phi_A+\phi)}\}\Big] \\
= {} & G\ P_A \sum_{k=0}^{+\infty} |\Gamma_N|^k|\Gamma_A|^k \sum_{l=0}^{k} \Big[cos\{(2l-k)(\phi_N+\phi_A+\phi)\}(1-|\Gamma_N|^2) \\
& + 2Im(\Gamma_N)sin\{(2l-k)(\phi_N+\phi_A+\phi)\}\Big].
\end{aligned} \qquad \text{(A.21)}$$

Since the last term containing the sine function is anti-symmetric, the summation of all of the sine terms is zero. Therefore, the expression for the power corresponding to the antenna becomes:

$$P_{SA} =G\ P_A(1-|\Gamma_N|^2)\sum_{k=0}^{+\infty}\ |\Gamma_N|^k|\Gamma_A|^k\sum_{l=0}^{k}cos\{(2l-k)(\phi_N+\phi_A+\phi)\}. \tag{A.22}$$

In a similar fashion, we may derive the voltage and power for the additive noise from the front-end amplifier. The voltage originating in the LNA is:

$$\begin{aligned} V_{SN} &= V_N(1+\Gamma_N)+fV_N\Gamma_A\mathrm{e}^{i\phi}(1+\Gamma_N)+fV_N\Gamma_A^2\Gamma_N\mathrm{e}^{i2\phi}(1+\Gamma_N).... \qquad \text{(A.23)}\\ &\quad +fV_N\Gamma_N^{n-1}\Gamma_A^{n}\mathrm{e}^{in\phi}(1+\Gamma_N)+....\\ &= V_N(1+\Gamma_N)\big\{1+f\sum_{n=0}^{+\infty}(\Gamma_A^{n+1}\Gamma_N^{n}e^{i(n+1)\phi})\big\}. \qquad \text{(A.24)} \end{aligned}$$

The power due to this voltage is given as:

$$\begin{aligned} P_{SN} &= \Big\langle G\ Re\Big(\frac{V_{SN}V_{SN}^*}{Z_N}\Big)\Big\rangle\\ &= G\ Re\Big[\frac{V_NV_N^*}{Z_N}(1+\Gamma_N)(1+\Gamma_N^*)\big\{1+\sum_{m=0}^{+\infty}(f\ \Gamma_A^{(m+1)}\Gamma_N^{m}e^{i(m+1)\phi})\big\} \qquad \text{(A.25)}\\ &\qquad \times\big\{1+\sum_{n=0}^{+\infty}(f^*\ \Gamma_A^{*(n+1)}\Gamma_N^{*n}e^{-i(n+1)\phi})\big\}\Big]\\ &= G\ Re\Big[P_N\big\{1-2iIm(\Gamma_N)-|\Gamma_N|^2\big\}\big\{1+\sum_{m=0}^{+\infty}(f\ \Gamma_A^{(m+1)}\Gamma_N^{m}e^{i(m+1)\phi})\big\} \qquad \text{(A.26)}\\ &\qquad \times\big\{1+\sum_{n=0}^{+\infty}(f^*\ \Gamma_A^{*(n+1)}\Gamma_N^{*n}e^{-i(n+1)\phi})\big\}\Big]. \end{aligned}$$

The above equation simplifies to:

$$\begin{aligned} P_{SN} = G\ P_N\ Re\Big[\big\{1-2iIm(\Gamma_N)-|\Gamma_N|^2\big\}\big\{1+\sum_{m=0}^{+\infty}2Re(f\ \Gamma_A^{m+1}\Gamma_N^{m}e^{i(m+1)\phi})+ \qquad \text{(A.27)}\\ |f|^2|\Gamma_A|^2\sum_{k=0}^{+\infty}|\Gamma_N|^k|\Gamma_A|^k\sum_{l=0}^{k}e^{i(2l-k)(\phi_N+\phi_A+\phi)}\big\}\Big]. \end{aligned}$$

Expanding terms, identifying $m = k$, and using arguments similar to that used in derivations above for the case of a single reflection at the antenna, we obtain:

$$P_{SN} = G\ P_N(1-|\Gamma_N|^2)\Big[1+\sum_{k=0}^{+\infty}(2|f||\Gamma_A|^{(k+1)}|\Gamma_N|^k cos\{\phi_f+(k+1)(\phi_A+\phi)+k\phi_N\}) \tag{A.28}$$
$$+|f|^2|\Gamma_A|^2\sum_{k=0}^{+\infty}\ |\Gamma_N|^k|\Gamma_A|^k\sum_{l=0}^{k}cos\{(2l-k)(\phi_N+\phi_A+\phi)\}\Big].$$

The total power flow out of the system can then be expressed as:

$$\begin{aligned}P_S &= P_{SA}+P_{SN} \qquad\qquad (\text{A.29})\\ &= G(1-|\Gamma_N|^2)\Big[P_N\big[1+\sum_{k=0}^{+\infty}(2|f||\Gamma_A|^{(k+1)}|\Gamma_N|^k cos\{\phi_f+(k+1)(\phi_A+\phi)+k\phi_N\})\\ &\qquad +|f|^2|\Gamma_A|^2\sum_{k=0}^{+\infty}\ |\Gamma_N|^k|\Gamma_A|^k\sum_{l=0}^{k}cos\{(2l-k)(\phi_N+\phi_A+\phi)\}\big]\\ &\qquad +P_A\sum_{k=0}^{+\infty}\ |\Gamma_N|^k|\Gamma_A|^k\sum_{l=0}^{k}cos\{(2l-k)(\phi_N+\phi_A+\phi)\}\Big]\end{aligned}$$

It may be noted here that equations A.14 and A.15 for the calibration states remain unchanged since it is assumed here that the reference port is impedance matched to the transmission line and both have impedances Z_0; there are no reflections of voltage waveforms at the reference port.

Omitting the pedagogical steps, the calibrated spectrum may thus be written as:

$$T_{meas} = T_{STEP}\Big\{\frac{P_A[\sum_{k=0}^{+\infty}\ |\Gamma_N|^k|\Gamma_A|^k\sum_{l=0}^{k}cos\{(2l-k)(\phi_N+\phi_A+\phi)\}]-P_{REF}}{P_{CAL}-P_{REF}} \tag{A.30}$$
$$+\frac{P_N}{P_{CAL}-P_{REF}}\times\Big[\sum_{k=0}^{+\infty}(2|f||\Gamma_A|^{(k+1)}|\Gamma_N|^k cos\{\phi_f+(k+1)(\phi_A+\phi)+k\phi_N\})$$
$$+|f|^2|\Gamma_A|^2\sum_{k=0}^{+\infty}\ |\Gamma_N|^k|\Gamma_A|^k\sum_{l=0}^{k}cos\{(2l-k)(\phi_N+\phi_A+\phi)\}\Big]\Big\}.$$

If we set $k = 0$ in the above equation, we recover Equation A.17 that represents the measured temperature assuming single reflection at the antenna and neglecting higher order terms.

Fig.A.3 shows simulations carried out with Eq.A.30 of the first 5 leading terms of the receiver noise alone. To demonstrate the nature of each term, Fig.A.3 shows the individual components inside the summation in Eq.A.30; the final spectrum would be a sum of all of them, ideally up to $k = \infty$. Further, the simulations have been simplified to have $P_N = 1$, $f = 0.1$, $\phi_f = 0$, $\Gamma_N = 0.4$, $\phi_N = 0$. This is a scenario with a receiver having a flat receiver noise spectrum, and spectrally flat characteristics for rest of the parameters, with receiver input having a reasonable match to the system impedance and presenting only a real impedance. The simulations in the top panel assume that the receiver is terminated with a *open*, at the end of a 1 m lossless cable with velocity factor equal to unity, thus $\Gamma_A = 1.0$ and $\phi_A = 0$. The bottom panel assumes a similar setup with a *short* termination, with $\Gamma_A = 1.0$ and $\phi_A = \pi$. As can be seen, the resulting terms have sinusoidal nature, with the phase depending on the termination. It can also be noticed that they are not perfect sinusoids.

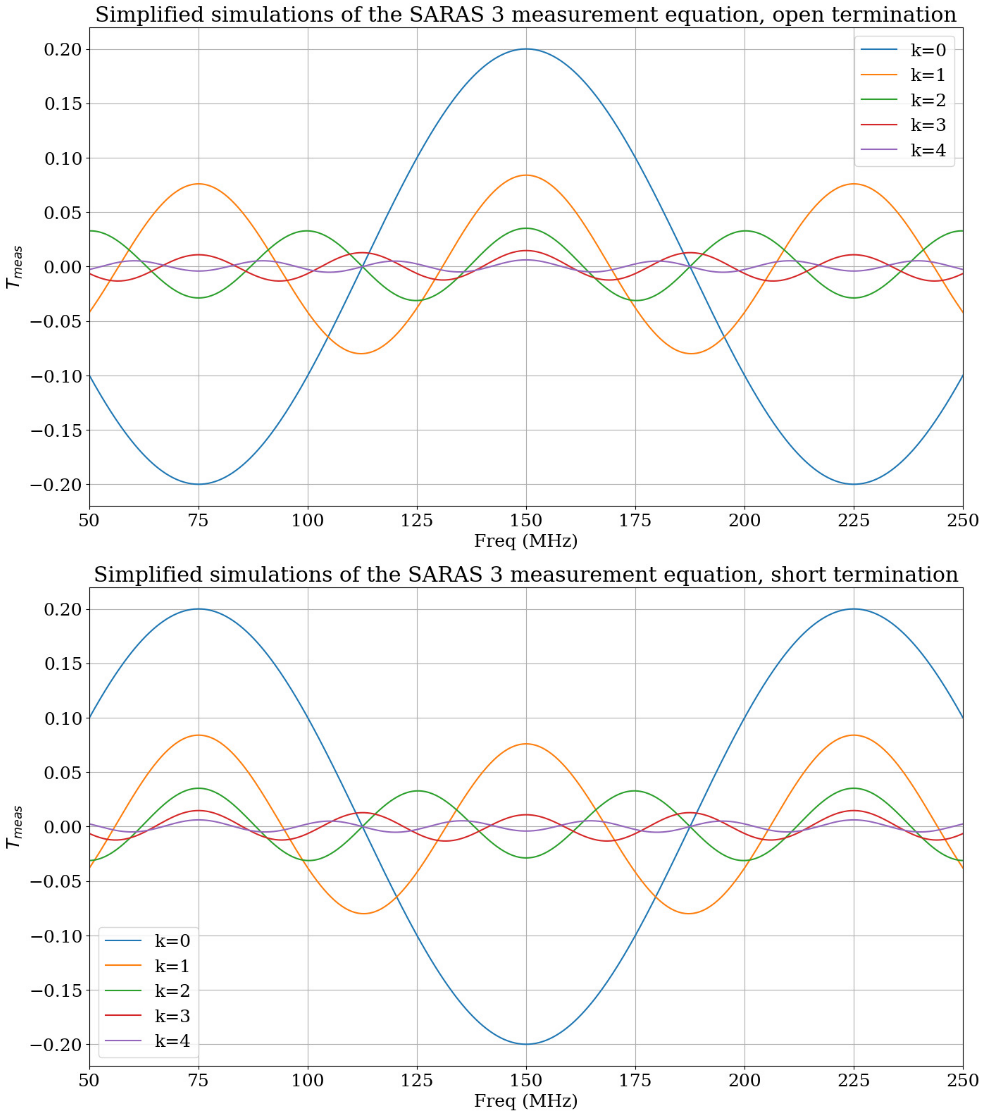


Figure A.3: Simulations of the first 5 leading terms of the receiver noise alone using Eq.A.30, and assuming a 1 m lossless cable. The top panel shows when the termination is open, and the bottom one shows when the termination is short. It may also be noted that the terminations considered here do not contribute to the system temperature and hence the resulting spectra are a result of the receiver noise and reflections of it.

Appendix B

A physical model for cross-talk in SITARA

In Sec.4.6.3, an empirical model for cross-talk in SITARA is provided. While the empirical model does describe the data, a model that is physically motivated would help in system design and analysis. Here we provide a plausible physical model, though we do not fit this model to our data. We would like to emphasise that the model presented here is a very simplistic one; in reality it is *not* possible to parameterise the cross-talk into a single factor as the cross-talk depends on the LNA input impedance and noise parameters. A more detailed model will be explored in future work.

We parameterise the cross-talk by a factor f_c which is the fraction of voltage that gets coupled from one antenna (or any where along its signal chain) to the other. The parameter f_c will, in general, be complex valued with a frequency dependency. We assume the two-antenna system to be reciprocal and hence f_c is same for both antenna 1-2 and antenna 2-1 paths. Under these assumptions, we may write the voltages at the antenna terminals as:

$$\begin{aligned} e_1(\nu, LST) &= e_{1,sky}(\nu, LST) + e_{1,RX}(\nu) + f_c(\nu)e_2(\nu, LST) \\ e_2(\nu, LST) &= e_{2,sky}(\nu, LST) + e_{2,RX}(\nu) + f_c(\nu)e_1(\nu, LST) \end{aligned} \tag{B.1}$$

where $e_{n,sky}$ are the voltages induced by external radiation (sky) on the individual antennas and $e_{n,RX}$ are the internal receiver noise voltages. However, Eqs.B.1 are coupled to each other and pose challenges in their application. Therefore we ignore cross-talk that arises from multiple couplings back and forth between the antennas, leading to Eqs.B.2.

$$\begin{aligned} e_1 &= e_{1,sky} + e_{1,RX} + f_c[e_{2,sky} + e_{2,RX}] \\ e_2 &= e_{2,sky} + e_{2,RX} + f_c[e_{1,sky} + e_{1,RX}] \end{aligned} \tag{B.2}$$

We can now form auto-correlations and cross-correlations from these voltages as $T_{ij} = e_i e_j^*$.

$$\begin{aligned} T_{11} = |e_{1,sky}|^2 + e_{1,sky}e^*_{2,sky}f_c^* + e^*_{1,sky}e_{2,sky}f_c + |f_c|^2|e_{2,sky}|^2 \\ +|e_{1,RX}|^2 + |f_c|^2|e_{2,RX}|^2 \\ \\ T_{22} = |e_{2,sky}|^2 + e_{2,sky}e^*_{1,sky}f_c^* + e^*_{2,sky}e_{1,sky}f_c + |f_c|^2|e_{1,sky}|^2 \\ +|e_{2,RX}|^2 + |f_c|^2|e_{1,RX}|^2 \\ \\ T_{12} = f_c^*|e_{1,sky}|^2 + e_{1,sky}e^*_{2,sky} + |f_c|^2 e_{2,sky}e^*_{1,sky} + f_c|e_{2,sky}|^2 \\ +f_c^*|e_{1,RX}|^2 + f_c|e_{2,RX}|^2 \end{aligned} \tag{B.3}$$

Eqs.B.3 may be rewritten into a matrix form as given in Eq.B.4, identifying $e_{i,sky}e^*_{j,sky} = V_{i,j}$ where $V_{i,j}$ are expected visibilities in the absence of cross-talk, as computed using Eq.4.7. Also, we use $T_{n11} = |e_{1,RX}|^2 + |f_c|^2|e_{2,RX}|^2$, $T_{n22} = |e_{2,RX}|^2 + |f_c|^2|e_{1,RX}|^2$ and $T_{n12} = f_c^*|e_{1,RX}|^2 + f_c|e_{2,RX}|^2$ to denote the noise temperatures of instrumental origin. For a drift instrument such as SITARA, expected visibilities as well as measured correlations change as a function of LST, owing to the movement of various sky regions through the antenna beams. This has also been incorporated into Eq.B.4 as $1....n$ rows in the expected visibilities as well as data matrices.

$$
\begin{aligned}
&\boldsymbol{T} = \boldsymbol{V}\boldsymbol{F}, \text{where} \\
&\qquad \boldsymbol{T} = \begin{bmatrix} T_{11}(t_i) & T_{12}(t_i) & T_{21}(t_i) & T_{22}(t_i) \end{bmatrix}; i = 1 \ to \ n \\
&\qquad \boldsymbol{V} = \begin{bmatrix} V_{11}(t_i) & V_{12}(t_i) & V_{21}(t_i) & V_{22}(t_i) & 1 \end{bmatrix}; i = 1 \ to \ n \\
&\qquad \boldsymbol{F} = \begin{bmatrix} 1 & f_c^* & f_c & |f_c|^2 \\ f_c^* & 1 & |f_c|^2 & f_c \\ f_c & |f_c|^2 & 1 & f_c^* \\ |f_c|^2 & f_c & f_c^* & 1 \\ T_{n11} & T_{n12} & T_{n21} & T_{n22} \end{bmatrix}
\end{aligned}
\tag{B.4}
$$

$\boldsymbol{T}$ is the matrix of measured auto and cross correlations, $\boldsymbol{V}$ is the matrix of expected visibilities in the absence of any cross-talk and $\boldsymbol{F}$ is the matrix with the coefficients. Given a set of simulated visibilities $\boldsymbol{V}$ and a set of measurements $\boldsymbol{T}$ in the same units as visibilities, i.e. kelvins, Eq.B.4 may be solved to obtain the matrix of coefficients $\boldsymbol{F}$. If the data are not calibrated to units of kelvins, the measurements will be treated as raw powers $\boldsymbol{P} = \boldsymbol{T}\boldsymbol{G}$ where $\boldsymbol{G}$ is the gain matrix given in B.5.

$$
G = \begin{bmatrix} |G_1|^2 & 0 & 0 & 0 \\ 0 & G_1 G_2^* & 0 & 0 \\ 0 & 0 & G_1^* G_2 & 0 \\ 0 & 0 & 0 & |G_2|^2 \end{bmatrix}
\tag{B.5}
$$

Comparing Eq.4.16 with the above formalism, it can be seen that $\boldsymbol{B}$ assumes the role of matrix $\boldsymbol{F}$ in empirical model, with the cross-talk coefficient f_c represented by the coefficients such as $a_{21}, b_{22}, c_{11}, d_{12}$ etc. However, the unconstrained least-squares fitting used in the empirical model (Sec.4.6.3) does not preserve the relations between the coefficients. This, coupled to the fact that the gain matrix

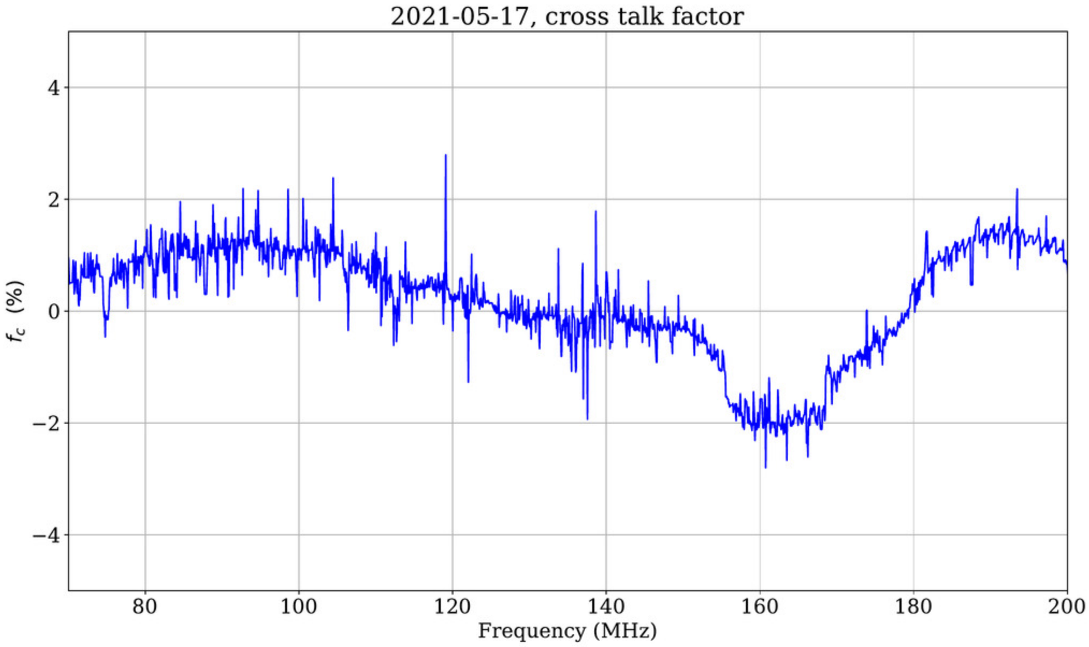


Figure B.1: Cross-talk factor f_c as a percentage. Please see the text for caveats associated with this calculation.

$\boldsymbol{G}$ is not accurately determined, makes establishing a relation between $\boldsymbol{B}$ and $\boldsymbol{F}$ a difficult exercise. However, if we assume that the auto-correlation based gains obtained in Sec.4.6.2 are reliable, it is possible to obtain some insights into the cross-talk in SITARA. For this, we sum the individual columns (except the receiver noise row) in $\boldsymbol{FG}$ into a factor y_c. For antenna-1 autocorrelations, this gives us

$$\begin{aligned} y_{c1} &= |G_1|^2\big(1 + f_c + f_c^* + |f_c|^2\big) \\ &= |G_1|^2\big(1 + f_c\big)\big(1 + f_c\big)^* \end{aligned} \tag{B.6}$$

We further assume that the cross-talk factor f_c is real valued, and write

$$y_{c1} = |G_1|^2\big(1 + 2f_c + f_c^2\big) \tag{B.7}$$

Dividing y_{c1} with the gain $|G_1|^2$ and taking the roots of the resulting quadratic equation, we obtain an estimate of f_c, which is shown in Fig.B.1 as a percentage. It has to be noted that there are caveats associated with this calculation. The assumptions that we made to obtain this estimate are not fully justified; calculation of $|G_1|^2$ is shown to be inaccurate and f_c cannot be real valued at all frequencies due to the finite path lengths for cross-talk. Therefore, the calculated f_c is given only to demonstrate an application of the physical cross-talk model.

Appendix C

Eigendecomposition of anti-circulant matrices

The relation between circulant matrices and circular convolution as well as the relation between the eigendecomposition of such matrices and Discrete Fourier Transforms (DFT) are well known - see for e.g. Gray (2006) for an approachable introduction to this relation - and put to use in several signal processing applications. However, the matrices of relevance for this paper are anti-circulant and therefore we provide some useful relations between those matrices and DFT.

Consider a real valued sequence $x[n], n = 0, 1...N-1$. This sequence can be used to construct an $N \times N$ real valued anti-circulant matrix as given in Eq.C.1.

$$\boldsymbol{X_{ac}} = \begin{bmatrix} x[0] & x[1] & ...... & x[N-1] \\ x[1] & x[2] & ...... & x[0] \\ x[2] & x[3] & ...... & x[1] \\ .. & .. & ...... & .. \\ x[N-1] & x[0] & ...... & x[N-2] \end{bmatrix} \tag{C.1}$$

The eigendecomposition of $\boldsymbol{X_{ac}}$ can be written as:

$$\boldsymbol{X_{ac}} = \boldsymbol{Q \Lambda Q^T} \tag{C.2}$$

where

$$\boldsymbol{X_{ac}}\, y_n = \lambda_n y_n \tag{C.3}$$

where λ_n are the eigenvalues and y_n corresponding eigenvectors with n being the index. The following set of equations are obtained from eigendecomposition.

$$\begin{gathered} x[0]y_n[0] + x[1]y_n[1]...... + x[N-1]y_n[N-1] = \lambda_n y_n[0] \\ x[1]y_n[0] + x[2]y_n[1]...... + x[0]y_n[N-1] = \lambda_n y_n[1] \\ .... \\ x[N-1]y_n[0] + x_n[0]y_n[1]...... + x[N-2]y_n[N-1] = \lambda_n y_n[N-1] \end{gathered} \tag{C.4}$$

which can be rearranged into:

$$\begin{gathered} x[0]y_n[0] + x[1]y_n[1]...... + x[N-1]y_n[N-1] = \lambda_n y_n[0] \\ x[0]y_n[N-1] + x[1]y_n[0]...... + x[N-1]y_n[N-2] = \lambda_n y_n[1] \\ .... \\ x[0]y_n[1] + x[1]y_n[2]...... + x[N-1]y_n[0] = \lambda_n y_n[N-1] \end{gathered} \tag{C.5}$$

The above set of equations can be recast into a summation equation as given in Eq.C.6.

$$\sum_{k=0}^{m-1} x[k]y_n[k-m+N] + \sum_{k=m}^{N-1} x[k]y_n[k-m] = \lambda_n y_n[m] \tag{C.6}$$

where $m = 0, 1....N-1$. The matrix is real symmetric and therefore the eigenvalues are real valued. The eigenvectors are fully real and we assume them to be N periodic as given in Eq.C.7.

$$y_n[k-m+N] = y_n[k-m] \tag{C.7}$$

This simplifies Eq.C.6 into:

$$\sum_{k=0}^{N-1} x[k]y_n[k-m] = \lambda_n y_n[m]; m = 0, 1....N-1. \tag{C.8}$$

It may be noted that the LHS of the above equation is a case of circular correlation between x and y_n. Therefore we can write:

$$(y_n \circledast x) = \lambda_n y_n \tag{C.9}$$

where $\circledast$ denotes correlation. Applying Fourier transforms, we get:

$$\mathcal{F}(y_n \circledast x) = \mathcal{F}(\lambda_n y_n) \tag{C.10}$$

Fourier transform of correlation can be written as a product, giving us:

$$\overline{\mathcal{F}(y_n)}\mathcal{F}(x) = \lambda_n \mathcal{F}(y_n) \tag{C.11}$$

We know that y_n and λ_n are real. Since y_n is periodic, a suitable choice is sinusoidal functions.

C.0.1 Cosine case

We first assume a cosine form for y_n such that $y_n[m] = cos\big(\omega_n m - \phi_i\big)$, where $\omega_n = \frac{2\pi}{N}n$. The Fourier transform of such a cosine function is a pair of delta functions as given in Eq.C.12.

$$\mathcal{F}\{cos\big(\omega_n m - \phi_i\big)\} = e^{-j\phi_i}\delta(\omega - \omega_n) + e^{j\phi_i}\delta(\omega + \omega_n) \tag{C.12}$$

The real valued normalisation factors have been ignored for the Fourier transform as they cancel out in the subsequent steps. The correlation equation can now be written as:

$$\begin{aligned}\big[e^{j\phi_i}\delta(\omega - \omega_n) + e^{-j\phi_i}\delta(\omega + \omega_n)\big]\mathcal{F}(x) = \lambda_n\big(e^{-j\phi_i}\delta(\omega - \omega_n)\\ + e^{j\phi_i}\delta(\omega + \omega_n)\big)\end{aligned} \tag{C.13}$$

Note that the above equation becomes non-zero only when $\omega = \omega_n$ or $\omega = -\omega_n$. Let us take $\omega = \omega_n$ case. We have:

$$e^{j\phi_i}\mathcal{F}(x)[\omega_n] = \lambda_n e^{-j\phi_i} \tag{C.14}$$

where the ω_n component of $\mathcal{F}(x)$ is

$$\mathcal{F}(x)[\omega_n] = |\mathcal{F}(x)[\omega_n]|e^{j\angle\mathcal{F}(x)[\omega_n]} \tag{C.15}$$

Then

$$e^{j\phi_i}|\mathcal{F}(x)[\omega_n]|e^{j\angle\mathcal{F}(x)[\omega_n]} = \lambda_n e^{-j\phi_i} \tag{C.16}$$

Leading to

$$|\mathcal{F}(x)[\omega_n]|e^{j\left(\angle\mathcal{F}(x)[\omega_n]+2\phi_i\right)} = \lambda_n \tag{C.17}$$

Since λ_n is real, we have $\angle\mathcal{F}(x)[\omega_n] + 2\phi_i = 0$ or $\phi_i = -\frac{\angle\mathcal{F}(x)[\omega_n]}{2}$. As x is real, the Fourier spectrum is Hermitian and therefore it can be shown that the above relation holds for $\omega = -\omega_n$. Thus, the cosine eigenvector and the corresponding eigenvalue for a given n are:

$$\begin{aligned}\lambda_n &= |\mathcal{F}(x)[\omega_n]| \\ y_n[m] &= cos\Big(\omega_n m + \frac{\angle\mathcal{F}(x)[\omega_n]}{2}\Big)\end{aligned} \tag{C.18}$$

C.0.2 Sine case

Another plausible eigenvector is $y_n[m] = sin(\omega_n m - \phi_i)$, giving:

$$\mathcal{F}\{sin\big(\omega_n m - \phi_i\big)\} = -j\left(e^{-j\phi_i}\delta(\omega-\omega_n) - e^{j\phi_i}\delta(\omega+\omega_n)\right) \tag{C.19}$$

Then we have:

$$\begin{aligned}j\left[e^{j\phi_i}\delta(\omega-\omega_n) - e^{-j\phi_i}\delta(\omega+\omega_n)\right]\mathcal{F}(x) = -j\lambda_n\big[&e^{-j\phi_i}\delta(\omega-\omega_n) \\ &- e^{j\phi_i}\delta(\omega+\omega_n)\big]\end{aligned} \tag{C.20}$$

Once again, let us consider $\omega = \omega_n$ case. We have:

$$(e^{j\phi_i})\mathcal{F}(x)[\omega_n] = -\lambda_n(e^{-j\phi_i}) \tag{C.21}$$

Leading to:

$$|\mathcal{F}(x)[\omega_n]|e^{j\left(\angle\mathcal{F}(x)[\omega_n]+2\phi_i\right)} = -\lambda_n \tag{C.22}$$

Once again, since λ_n is real we have $\angle\mathcal{F}(x)[\omega_n] + 2\phi_i = 0$. Thus, the sine eigenvector and corresponding eigenvalue for a given n are:

$$\begin{aligned} \lambda_n &= -|\mathcal{F}(x)[\omega_n]| \qquad \text{(C.23)} \\ y_n[m] &= sin\Big(\omega_n m + \frac{\angle\mathcal{F}(x)[\omega_n]}{2}\Big) \end{aligned}$$

Summarising, the spectrum of an anti-circulant matrix consists of the *absolute* value of the DFT spectrum of the underlying periodic pattern, with the values occurring in positive-negative pairs. The eigenvectors are sines and cosines with phases determined by the phase of the DFT.

Appendix D

Copyright Information

Chapter 3 and the associated Appendix A of this thesis are a reproduction of a journal article published in *Springer Experimental Astronomy*. Chapter 4 and the associated Appendix B are a reproduction of a journal article published in *Publications of the Astronomical Society of Australia (PASA)*. The licensing information for the two articles are attached in the following pages.

Licence to Publish

SPRINGER NATURE

Journal Name:	Experimental Astronomy	(the 'Journal')
Manuscript Number:	EXPA-D-20-00031R1	
Proposed Title of Article:	SARAS 3 CD/EoR radiometer: design and performance of the receiver	
Author(s) [Please list all named Authors]:	Jishnu Nambissan Thekkeppattu, Ravi Subrahmanyan, R. Somashekar, N. Udaya Shankar, Saurabh Singh, A. Raghunathan, B. S. Girish, K. S. Srivani, Mayuri Sathyanarayana Rao	(the 'Author')
Corresponding Author Name:	Jishnu Nambissan Thekkeppattu	

1 Publication

Springer Nature B.V. (the 'Licensee') will consider publishing this article, including any supplementary information and graphic elements therein (e.g. illustrations, charts, moving images) (the 'Article').
Headings are for convenience only.

2 Grant of Rights

In consideration of the Licensee evaluating the Article for publication, the Author grants the Licensee the exclusive (except as set out in clauses 3, 4 and 5a) iv) and sub-licensable right, unlimited in time and territory, to copy-edit, reproduce, publish, distribute, transmit, make available and store the Article, including abstracts thereof, in all forms of media of expression now known or developed in the future, including pre- and reprints, translations, photographic reproductions and extensions.

Furthermore, to enable additional publishing services, such as promotion of the Article, the Author grants the Licensee the right to use the Article (including the use of any graphic elements on a stand-alone basis) in whole or in part in electronic form, such as for display in databases or data networks (e.g. the Internet), or for print or download to stationary or portable devices. This includes interactive and multimedia use as well as posting the Article in full or in part or its abstract on social media, and the right to alter the Article to the extent necessary for such use. The Licensee may also let third parties share the Article in full or in part or its abstract on social media and may in this context sub-license the Article and its abstract to social media users. Author grants to Licensee the right to re-license Article metadata without restriction (including but not limited to author name, title, abstract, citation, references, keywords and any additional information as determined by Licensee).

If the Article is rejected by the Licensee and not published, all rights under this agreement shall revert to the Author.

3 Self Archiving

Author is permitted to self-archive a preprint and the accepted manuscript version of their Article.

a) A preprint is the version of the Article before peer-review has taken place ("Preprint"). Prior to acceptance for publication, Author retains the right to make a Preprint of their Article available on any of the following: their own personal, self-maintained website; a legally compliant Preprint server such as but not limited to arXiv and bioRxiv. Once the Article has been published, the Author should update the acknowledgement and provide a link to the definitive version on the publisher's website: "This is a preprint of an article published in [insert journal title]. The final authenticated version is available online at: https://doi.org/[insert DOI]"

b) The accepted manuscript version, by industry standard called the "Author's Accepted Manuscript" ("AAM") is the version accepted for publication in a journal following peer review but prior to copyediting and typesetting that can be made available under the following conditions:
(i) Author retains the right to make an AAM of the Article available on their own personal, self-maintained website immediately on acceptance, (ii) Author retains the right to make an AAM of the Article available for public release on any of the following 12 months after first publication ("Embargo Period"): their employer's internal website; their institutional and/or funder repositories; AAMs may also be deposited in such repositories immediately on acceptance, provided that they are not made publicly available until after the Embargo Period.

An acknowledgement in the following form should be included, together with a link to the published version on the publisher's website: "This is a post-peer-review, pre-copyedit version of an article published in [insert journal title]. The final authenticated version is available online at: http://dx.doi.org/[insert DOI]".

4 Reuse Rights

Author retains the following non-exclusive rights for the published version provided that, when reproducing the Article or extracts from it, the Author acknowledges and references first publication in the Journal according to current citation standards. In any event the acknowledgement should contain as a minimum, "First published in [Journal name, volume, page number, year] by Springer Nature".

a) to reuse graphic elements created by the Author and contained in the Article, in presentations and other works created by them;

RETURN FORM: A **hand-signed** copy of this LTP must be returned by email to:
pasa@cambridge.org

08_ExLTP_Proprietor_Standard (Version 2.0)

CAMBRIDGE
UNIVERSITY PRESS

EXCLUSIVE LICENCE TO PUBLISH ("LTP")

This LTP records the terms under which the article specified below will be published in **Publications of the Astronomical Society of Australia (PAS)** (the "**Journal**"). The Journal is exclusively published by the Chancellor, Masters, and Scholars of the University of Cambridge acting through its department **Cambridge University Press** of University Printing House, Shaftesbury Road, Cambridge CB2 8BS, UK (the "**Publisher**"). The Journal is owned by Astronomical Society of Australia, c/o Professor J. W. O'Byrne, School of Physics, The University of Sydney, NSW 2006 Australia (the "**Proprietor**").

THE ARTICLE	*Please insert the full title of the article below.*
Article Title*:	System design and calibration of SITARA - a global 21 cm short spacing interferometer prototype (the "**Contribution**")

This LTP can be used where a Contribution has one or more authors. The lead author must complete the box below and sign this LTP on behalf of themself (and all other authors, if any).

LEAD AUTHOR'S DETAILS AND SIGNATURE			
Full Legal Name*:	Jishnu Nambissan Thekkeppattu (the "**Lead Author**")		
Pen or preferred name:			
Email address*:	j.thekkeppattu@postgrad.curtin.edu.au		
Affiliation*:	ICRAR/Curtin and RRI Bengaluru		
Country of residence*:	Australia		
Authority to sign:	By signing this LTP, I confirm and agree that: i. All information that I have entered into this LTP is correct at the time of signature. ii. **EITHER**, I am the sole author and owner of the copyright in the Contribution and I agree to the terms and conditions in this LTP. iii. **OR**, the copyright in the Contribution is jointly owned by me and the Author(s) *listed below* and I agree to (and am authorized by each Author to agree to) the terms of this LTP on behalf of all Authors; iv. **AND**, no other person nor entity has any copyright interest in the Contribution.		
Signature*:	[redacted]	**Date*:**	21/01/2022

OTHER AUTHORS' DETAILS					
*If the Contribution is written by **two or more** authors and the copyright in the Contribution is **jointly** owned by them – please enter the details of all other individuals who contributed to the authoring of the Contribution in this box.* *If necessary, please add any more authors at the end of this LTP.*	**Full Legal Name***	**Pen or preferred name**	**Email address***	**Affiliation***	**Country of residence***
	Benjamin McKinley		ben.mckinley @curtin.edu.au	ICRAR and ASTRO3D	Australia
	Cathryn M. Trott		cathryn.trott @curtin.edu.au	ICRAR and ASTRO3D	Australia
	Jake Jones		jake.jones @curtin.edu.au	ICRAR	Australia
	Daniel C. X. Ung		daniel.ung @curtin.edu.au	ICRAR	Australia

(the Lead Author and each individual listed here and at the end of this LTP is, individually and collectively, the "**Author**")

CAMBRIDGE EMPLOYEE *You **must** check this box and enter details, if applicable.*	☐ One or more Authors are **employed by** Cambridge University Press or are **related to** a Cambridge University Press employee. Please provide names and describe the relationship(s):

SUPPLEMENTARY MATERIALS	*Identify any additional materials to be published in association with the Contribution*
If the Author intends to submit or upload any additional materials for online publication in association with the Contribution, please indicate by checking the applicable boxes in this section.	☑ **NO**, Supplementary Materials will **not** be submitted or uploaded by the Author for publication/uploading in connection with the Contribution. ☐ **YES**, Supplementary Materials which have been entirely created by the Author ("Original SM") will be submitted to the Publisher for publication/uploading in connection with the Contribution. ☐ **YES,** Supplementary Materials which contain third-party materials ("Third-party SM") will be submitted to the Publisher for publication/uploading in connection with the Contribution and the Author shall include a prominent notice stating the licence terms under which those additional materials can be made available. (the "**Supplementary Material**")

PLEASE NOTE: Amended/alternative versions of this LTP will not be accepted and may delay or prevent publication

Tables of attribution

D.0.1 Chapter 3 of this thesis

Fig.D.0.1 gives the table of co-author attribution for the journal article Jishnu Nambissan, T.; Ravi Subrahmanyan; R. Somashekar; N. Udaya Shankar; Saurabh Singh; A. Raghunathan; B. S. Girish; K. S. Srivani; Mayuri Sathyanarayana Rao, *SARAS 3 CD/EoR radiometer: design and performance of the receiver*, 2021, Experimental Astronomy, https://doi.org/10.1007/s10686-020-09697-2

D.0.2 Chapter 4 of this thesis

Fig.D.0.2 gives the table of co-author attribution for the journal article Jishnu N. Thekkeppattu, Benjamin McKinley, Cathryn M. Trott, Jake Jones, Daniel C. X. Ung, *System design and calibration of SITARA - a global 21-cm short spacing interferometer prototype*, 2022, Publications of the Astronomical Society of Australia, https://doi.org/10.1017/pasa.2022.13

	Conception and Design	Acquisition of Data and Method	Data Conditioning and Manipulation	Analysis and Statistical Method	Interpretation and Discussion
Ravi Subrahmanyan	✓	✓	✓	✓	✓

Ravi Subrahmanyan Acknowledgment:
I acknowledge that these represent my contribution to the above research output and I have approved the final version.

Signed: [redacted]

	Conception and Design	Acquisition of Data and Method	Data Conditioning and Manipulation	Analysis and Statistical Method	Interpretation and Discussion
R. Somashekar	✓	✓	✓		✓

R. Somashekar Acknowledgment:
I acknowledge that these represent my contribution to the above research output and I have approved the final version.

Signed: [redacted]

	Conception and Design	Acquisition of Data and Method	Data Conditioning and Manipulation	Analysis and Statistical Method	Interpretation and Discussion
Udaya Shankar N	✓		✓	✓	✓

Udaya Shankar N Acknowledgment:
I acknowledge that these represent my contribution to the above research output and I have approved the final version.

Signed: [redacted]

	Conception and Design	Acquisition of Data and Method	Data Conditioning and Manipulation	Analysis and Statistical Method	Interpretation and Discussion
Saurabh Singh					✓

Saurabh Singh Acknowledgment:
I acknowledge that these represent my contribution to the above research output and I have approved the final version.

Signed: [redacted]

	Conception and Design	Acquisition of Data and Method	Data Conditioning and Manipulation	Analysis and Statistical Method	Interpretation and Discussion
A.Raghunathan	☑	☐			☑

A.Raghunathan Acknowledgment:
I acknowledge that these represent my contribution to the above research output and I have approved the final version.

Signed: [redacted]

	Conception and Design	Acquisition of Data and Method	Data Conditioning and Manipulation	Analysis and Statistical Method	Interpretation and Discussion
GIRISH B. S.					✓

GIRISH B. S. Acknowledgment:
I acknowledge that these represent my contribution to the above research output and I have approved the final version.

Signed: [redacted]

	Conception and Design	Acquisition of Data and Method	Data Conditioning and Manipulation	Analysis and Statistical Method	Interpretation and Discussion
Srivani K.S.	☐	☐	☐	☐	☒

Srivani K.S. Acknowledgment:
I acknowledge that these represent my contribution to the above research output and I have approved the final version.

Signed: [redacted]

	Conception and Design	Acquisition of Data and Method	Data Conditioning and Manipulation	Analysis and Statistical Method	Interpretation and Discussion
Mayuri Sathyanarayana Rao	☐	☐	☐	☐	☒

Mayuri Sathyanarayana Rao Acknowledgment:
I acknowledge that these represent my contribution to the above research output and I have approved the final version.

Signed: [redacted]

Figure D.1: Co-author attribution for Chapter 3

<table>
<tr><th></th><th>Conception and Design</th><th>Acquisition of Data and Method</th><th>Data Conditioning and Manipulation</th><th>Analysis and Statistical Method</th><th>Interpretation and Discussion</th></tr>
<tr><td>B. McKinley</td><td>☑</td><td>☑</td><td>☐</td><td>☐</td><td>☑</td></tr>
<tr><td colspan="6">B. McKinley Acknowledgment:
I acknowledge that these represent my contribution to the above research output and I have approved the final version.
Signed: [redacted]</td></tr>
<tr><td>C. M. Trott</td><td>☑</td><td>☐</td><td>☐</td><td>☑</td><td>☑</td></tr>
<tr><td colspan="6">C. M. Trott Acknowledgment:
I acknowledge that these represent my contribution to the above research output and I have approved the final version.
Signed: [redacted]</td></tr>
<tr><td>J. Jones</td><td>☐</td><td>☑</td><td>☐</td><td>☐</td><td>☐</td></tr>
<tr><td colspan="6">J. Jones Acknowledgment:
I acknowledge that these represent my contribution to the above research output and I have approved the final version.
Signed: [redacted]</td></tr>
<tr><td>D. C. X. Ung</td><td>☐</td><td>☑</td><td>☐</td><td>☐</td><td>☑</td></tr>
<tr><td colspan="6">D. C. X. Ung Acknowledgment:
I acknowledge that these represent my contribution to the above research output and I have approved the final version.
Signed: [redacted]</td></tr>
</table>

Figure D.2: Co-author attribution for Chapter 4

Bibliography

Astropy Collaboration, et al. (2018). ‘The Astropy Project: Building an Open-science Project and Status of the v2.0 Core Package’. *The Astronomical Journal* **156**(3):123.

Astropy Collaboration, et al. (2013). ‘Astropy: A community Python package for astronomy’. *Astronomy and Astrophysics* **558**:A33.

J. S. Bagla & A. Loeb (2009). ‘The hyperfine transition of 3He+ as a probe of the intergalactic medium’. *arXiv e-prints* p. arXiv:0905.1698.

R. Barkana (2018). ‘Possible interaction between baryons and dark-matter particles revealed by the first stars’. *Nature* **555**(7694):71–74.

R. Barkana, et al. (2018). ‘Strong constraints on light dark matter interpretation of the EDGES signal’. *Physical Review D* **98**:103005.

R. H. Becker, et al. (2001). ‘Evidence for Reionization at z~6: Detection of a Gunn-Peterson Trough in a z=6.28 Quasar’. *The Astronomical Journal* **122**(6):2850–2857.

C. L. Bennett, et al. (2003). ‘The Microwave Anisotropy Probe Mission’. *The Astrophysical Journal* **583**(1):1–23.

R. Beresford, et al. (2017). ‘Radio astronomy L-band phased array feed RFoF implementation overview’. In *2017 XXXIInd General Assembly and Scientific Symposium of the International Union of Radio Science (URSI GASS)*, pp. 1–4.

M. Bersanelli, et al. (2010). 'Planck pre-launch status: Design and description of the Low Frequency Instrument'. *Astronomy and Astrophysics* **520**:A4.

M. Born & E. Wolf (1959). 'Principles of Optics Electromagnetic Theory of Propagation, Interference and Diffraction of Light'.

J. D. Bowman, et al. (2018). 'An absorption profile centred at 78 megahertz in the sky-averaged spectrum'. *Nature* **555**:67 EP –.

R. F. Bradley, et al. (2019). 'A Ground Plane Artifact that Induces an Absorption Profile in Averaged Spectra from Global 21 cm Measurements, with Possible Application to EDGES'. *The Astrophysical Journal* **874**(2):153.

D. Broomhead & G. P. King (1986). 'Extracting qualitative dynamics from experimental data'. *Physica D: Nonlinear Phenomena* **20**(2):217–236.

J. O. Burns, et al. (2012). 'Probing the first stars and black holes in the early Universe with the Dark Ages Radio Explorer (DARE)'. *Advances in Space Research* **49**(3):433–450.

X. Chen, et al. (2021). 'Discovering the sky at the longest wavelengths with a lunar orbit array'. *Philosophical Transactions of the Royal Society of London Series A* **379**(2188):20190566.

A. Chokshi, et al. (2020). 'EMBERS: Experimental Measurement of BEam Responses with Satellites'. *Journal of Open Source Software* **5**(55):2629.

A. Cichocki, et al. (2015). 'Tensor Decompositions for Signal Processing Applications: From two-way to multiway component analysis'. *IEEE Signal Processing Magazine* **32**(2):145–163.

B. G. Clark (1999). 'Coherence in Radio Astronomy'. In G. B. Taylor, C. L. Carilli, & R. A. Perley (eds.), *Synthesis Imaging in Radio Astronomy II*, vol. 180 of *Astronomical Society of the Pacific Conference Series*, p. 1.

A. Cohen, et al. (2017). 'Charting the parameter space of the global 21-cm signal'. *Monthly Notices of the Royal Astronomical Society* **472**(2):1915–1931.

P. J. Davis (1979). ‘Circulant Matrices’.

E. de Lera Acedo (2019). ‘REACH: Radio Experiment for the Analysis of Cosmic Hydrogen’. In *2019 International Conference on Electromagnetics in Advanced Applications (ICEAA)*, pp. 0626–0629.

E. de Lera Acedo, et al. (2015). ‘SKALA, a log-periodic array antenna for the SKA-low instrument: design, simulations, tests and system considerations’. *Experimental Astronomy* **39**(3):567–594.

A. de Oliveira-Costa, et al. (2008). ‘A model of diffuse Galactic radio emission from 10 MHz to 100 GHz’. *Monthly Notices of the Royal Astronomical Society* **388**(1):247–260.

D. R. DeBoer, et al. (2017). ‘Hydrogen Epoch of Reionization Array (HERA)’. *Publications of the Astronomical Society of the Pacific* **129**(974):045001.

P. E. Dewdney, et al. (2009). ‘The Square Kilometre Array’. *IEEE Proceedings* **97**(8):1482–1496.

R. H. Dicke & R. Beringer (1946). ‘Microwave Radiation from the Sun and Moon.’. *The Astrophysical Journal* **103**:375.

R. H. Dicke, et al. (1965). ‘Cosmic Black-Body Radiation.’. *The Astrophysical Journal* **142**:414–419.

R. M. Dokht, et al. (2016). ‘Singular spectrum analysis and its applications in mapping mantle seismic structure’. *Geophysical Journal International* **208**(3):1430–1442.

G. I. Donskikh, et al. (2016). ‘Singular-Spectrum Analysis and Wavelet Analysis of the Variability of the Extragalactic Radio Sources 3C 120 and CTA 102’. *Astrophysics* **59**(2):199–212.

X. Fan, et al. (2006). ‘Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with z~6 Quasars. II. A Sample of 19 Quasars’. *The Astronomical Journal* **132**(1):117–136.

C. Feng & G. Holder (2018). ‘Enhanced Global Signal of Neutral Hydrogen Due to Excess Radiation at Cosmic Dawn’. *The Astrophysical Journal* **858**(2):L17.

G. B. Field (1958). ‘Excitation of the Hydrogen 21-CM Line’. *Proceedings of the IRE* **46**:240–250.

G. B. Field (1959). ‘The Spin Temperature of Intergalactic Neutral Hydrogen.’. *The Astrophysical Journal* **129**:536.

D. J. Fixsen (2009). ‘The Temperature of the Cosmic Microwave Background’. *The Astrophysical Journal* **707**(2):916–920.

D. J. Fixsen, et al. (1996). ‘The Cosmic Microwave Background Spectrum from the Full COBE FIRAS Data Set’. *The Astrophysical Journal* **473**:576.

D. J. Fixsen, et al. (2011). ‘ARCADE 2 Measurement of the Absolute Sky Brightness at 3-90 GHz’. *The Astrophysical Journal* **734**(1):5.

S. R. Furlanetto & S. P. Oh (2008). ‘The History and Morphology of Helium Reionization’. *The Astrophysical Journal* **681**(1):1–17.

S. R. Furlanetto, et al. (2006). ‘Cosmology at low frequencies: The 21 cm transition and the high-redshift Universe’. *Physics Reports* **433**(4-6):181–301.

B. Gaensler, et al. (2004). ‘The origin and evolution of cosmic magnetism’. *New Astronomy Reviews* **48**(11):1003–1012. Science with the Square Kilometre Array.

H. Garsden, et al. (2021). ‘A 21-cm power spectrum at 48 MHz, using the owens valley long wavelength array’. *Monthly Notices of the Royal Astronomical Society* stab1671.

M. Ghil, et al. (2002). ‘Advanced spectral methods for climatic time series’. *Rev. Geophys.* **40**(1):1 – 41. Cited by: 94.

B. S. Girish, et al. (2020). ‘SARAS CD/EoR Radiometer: Design and performance of the Digital Correlation Spectrometer’. *Journal of Astronomical Instrumentation, accepted* .

N. Golyandina & A. Korobeynikov (2014). ‘Basic Singular Spectrum Analysis and forecasting with R’. *Computational Statistics & Data Analysis* **71**:934–954.

N. Golyandina, et al. (2001). ‘Analysis of time series structure: SSA and related techniques’.

K. M. Górski, et al. (2005). ‘HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere’. *The Astrophysical Journal* **622**(2):759–771.

R. M. Gray (2006). ‘Toeplitz and Circulant Matrices: A Review’. *Foundations and Trends in Communications and Information Theory* **2**(3):155–239.

L. M. Griffiths, et al. (1999). ‘Cosmic microwave background constraints on the epoch of reionization’. *Monthly Notices of the Royal Astronomical Society* **308**(3):854–862.

J. E. Gunn & B. A. Peterson (1965). ‘On the Density of Neutral Hydrogen in Intergalactic Space.’. *The Astrophysical Journal* **142**:1633–1636.

Y. Gupta, et al. (2017). ‘The upgraded GMRT: opening new windows on the radio Universe’. *Current Science* **113**(4):707–714.

H. P. Gush, et al. (1990). ‘Rocket measurement of the cosmic-background-radiation mm-wave spectrum’. *Physical Review Letters* **65**(5):537–540.

N. M. Gürel, et al. (2018). ‘Towards More Accurate Radio Telescope Images’. In *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pp. 1983–19832.

Z. Haiman & L. Knox (1999). ‘Reionization of the Intergalactic Medium and its Effect on the CMB’. In A. de Oliveira-Costa & M. Tegmark (eds.), *Microwave Foregrounds*, vol. 181 of *Astronomical Society of the Pacific Conference Series*, p. 227.

G. Hallinan, et al. (2015). ‘Monitoring All the Sky All the Time with the Owens Valley Long Wavelength Array’. In *American Astronomical Society Meeting*

Abstracts #225, vol. 225 of *American Astronomical Society Meeting Abstracts*, p. 328.01.

J. P. Hamaker, et al. (1996). 'Understanding radio polarimetry. I. Mathematical foundations.'. *Astronomy and Astrophysics Supplement Series* **117**:137–147.

C. R. Harris, et al. (2020). 'Array programming with NumPy'. *Nature* **585**(7825):357–362.

H. A. Haus, et al. (1960). 'Representation of Noise in Linear Twoports'. *Proceedings of the IRE* **48**(1):69–74.

J. Hickish, et al. (2016). 'A Decade of Developing Radio-Astronomy Instrumentation using CASPER Open-Source Technology'. *Journal of Astronomical Instrumentation* **05**(04):1641001.

B. C. Hicks, et al. (2012). 'A Wide-Band, Active Antenna System for Long Wavelength Radio Astronomy'. *Publications of the Astronomical Society of the Pacific* **124**(920):1090.

HIGH-Z (2020). 'The HIGH-Z 21-cm Global Spectrum Experiment'. `https://www.usnc-ursi-archive.org/nrsm/2019/papers/J4-8.pdf`. Accessed: 2020-11-02.

R. Hills, et al. (2018). 'Concerns about modelling of the EDGES data'. *Nature* **564**(7736):E32–E34.

G. Hinshaw, et al. (2013). 'Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Parameter Results'. *The Astrophysical Journal Supplemental* **208**(2):19.

H. Hitney, et al. (1985). 'Tropospheric radio propagation assessment'. *Proceedings of the IEEE* **73**(2):265–283.

J. D. Hunter (2007). 'Matplotlib: A 2D graphics environment'. *Computing in Science & Engineering* **9**(3):90–95.

G. R. Jessop (1983). 'VHF/UHF Manual'.

Y. Jia, et al. (2018). 'The scientific objectives and payloads of Chang'E-4 mission'. *Planetary and Space Science* **162**:207–215.

J. B. Johnson (1928). 'Thermal Agitation of Electricity in Conductors'. *Physical Review* **32**(1):97–109.

C. H. Jordan, et al. (2017). 'Characterization of the ionosphere above the Murchison Radio Observatory using the Murchison Widefield Array'. *Monthly Notices of the Royal Astronomical Society* **471**(4):3974–3987.

S. Khullar, et al. (2020). 'Probing the high-z IGM with the hyperfine transition of ^{3}He$^+$'. *Monthly Notices of the Royal Astronomical Society* **497**(1):572–580.

J. D. Kraus & R. J. Marhefka (2002). 'Antennas for all applications'.

T. L. Landecker & R. Wielebinski (1970). 'The Galactic Metre Wave Radiation: A two-frequency survey between declinations +25^o and -25^o and the preparation of a map of the whole sky'. *Australian Journal of Physics Astrophysical Supplement* **16**:1.

A. Loeb (2008). 'First Light'. In A. Loeb (ed.), *First Light in the Universe: Saas-Fee Advanced Course 36. Swiss Society for Astrophysics and Astronomy*, pp. 1–159. Springer Berlin Heidelberg, Berlin, Heidelberg.

N. Mahesh, et al. (2015). 'A Wideband Resistive Beam-Splitter Screen'. *IEEE Transactions on Antennas and Propagation* **63**(11):4835–4847.

J. C. Mather, et al. (1994). 'Measurement of the Cosmic Microwave Background Spectrum by the COBE FIRAS Instrument'. *The Astrophysical Journal* **420**:439.

B. McKinley, et al. (2018). 'Measuring the global 21-cm signal with the MWA-I: improved measurements of the Galactic synchrotron background using lunar occultation'. *Monthly Notices of the Royal Astronomical Society* **481**(4):5034–5045.

B. McKinley, et al. (2020). 'The All-Sky SignAl Short-Spacing INterferometer (ASSASSIN) – I. Global-sky measurements with the Engineering Development Array-2'. *Monthly Notices of the Royal Astronomical Society* **499**(1):52–67.

M. McQuinn & E. R. Switzer (2009). 'Redshifted intergalactic He+3 8.7 GHz hyperfine absorption'. *Physical Review D* **80**(6):063010.

F. G. Mertens, et al. (2020). 'Improved upper limits on the 21cm signal power spectrum of neutral hydrogen at $z \sim 9.1$ from LOFAR'. *Monthly Notices of the Royal Astronomical Society* **493**(2):1662–1685.

M. Mevius, et al. (2016). 'Probing ionospheric structures using the LOFAR radio telescope'. *Radio Science* **51**(7):927–941.

J. Mirocha (2014). 'Decoding the X-ray properties of pre-reionization era sources'. *Monthly Notices of the Royal Astronomical Society* **443**:1211–1223.

J. Mirocha (2020). 'ARES: Accelerated Reionization Era Simulations'. Astrophysics Source Code Library, record ascl:2011.010.

MIST (2020). 'MIST - Mapper of the IGM Spin Temperature'. `http://www.physics.mcgill.ca/mist/`. Accessed: 2020-11-02.

R. A. Monsalve, et al. (2021). 'Absolute Calibration of Diffuse Radio Surveys at 45 and 150 MHz'. *The Astrophysical Journal* **908**(2):145.

R. A. Monsalve, et al. (2017). 'Results from EDGES High-band. I. Constraints on Phenomenological Models for the Global 21 cm Signal'. *The Astrophysical Journal* **847**(1):64.

T. J. Mozdzen, et al. (2018). 'Spectral index of the diffuse radio background between 50 and 100 MHz'. *Monthly Notices of the Royal Astronomical Society* **483**(4):4411–4423.

J. Nambissan T., et al. (2021). 'SARAS 3 CD/EoR radiometer: design and performance of the receiver'. *Experimental Astronomy* .

A. Nindos, et al. (2019). ‘Solar physics with the Square Kilometre Array’. *Advances in Space Research* **63**(4):1404–1424. Advances in Solar Physics.

P. D. Ninni, et al. (2020). ‘Comparison between Measured and Simulated Antenna Patterns for a LOFAR LBA array’. In *2020 14th European Conference on Antennas and Propagation (EuCAP)*, pp. 1–5.

H. Nyquist (1928a). ‘Certain Topics in Telegraph Transmission Theory’. *Transactions of the American Institute of Electrical Engineers* **47**(2):617–624.

H. Nyquist (1928b). ‘Thermal Agitation of Electric Charge in Conductors’. *Physical Review* **32**(1):110–113.

A. R. Offringa, et al. (2010). ‘Post-correlation radio frequency interference classification methods’. *Monthly Notices of the Royal Astronomical Society* **405**(1):155–167.

A. R. Offringa, et al. (2015). ‘The Low-Frequency Environment of the Murchison Widefield Array: Radio-Frequency Interference Analysis and Mitigation’. *Publications of the Astronomical Society of Australia* **32**:e008.

M. Oh, et al. (2017). ‘ENR Compensation of a Noise Source Using Calibration Coefficients With Variation of the Ambient Temperature’. *IEEE Transactions on Microwave Theory and Techniques* **65**(6):2201–2210.

B. J. Olson, et al. (2014). ‘Circulant Matrices and Their Application to Vibration Analysis’. *Applied Mechanics Reviews* **66**(4). 040803.

G. Paciga, et al. (2011). ‘The GMRT Epoch of Reionization experiment: a new upper limit on the neutral hydrogen power spectrum at $z \sim 8.6$’. *Monthly Notices of the Royal Astronomical Society* **413**(2):1174–1183.

S. Padin, et al. (2001). ‘First Intrinsic Anisotropy Observations with the Cosmic Background Imager’. *The Astrophysical Journal* **549**(1):L1–L5.

A. Parsons (2016). ‘AIPY: Astronomical Interferometry in PYthon’.

N. Patra, et al. (2013). ‘SARAS: a precision system for measurement of the cosmic radio background and signatures from the epoch of reionization’. *Experimental Astronomy* **36**:319–370.

N. Patra, et al. (2015). ‘Saras Measurement of the Radio Background At Long Wavelengths’. *The Astrophysical Journal* **801**(2):138.

A. A. Penzias & R. W. Wilson (1965). ‘A Measurement of Excess Antenna Temperature at 4080 Mc/s.’. *The Astrophysical Journal* **142**:419–421.

F. Perini, et al. (2022). ‘Radio frequency over fiber technology for SKA-low receiver’. *Journal of Astronomical Telescopes, Instruments, and Systems* **8**:011016.

J. B. Peterson, et al. (2004). ‘The Primeval Structure Telescope’. *Modern Physics Letters A* **19**(13-16):1001–1008.

L. Philip, et al. (2019). ‘Probing Radio Intensity at High-Z from Marion: 2017 Instrument’. *Journal of Astronomical Instrumentation* **8**(2):1950004.

Planck Collaboration, et al. (2020). ‘Planck 2018 results. VI. Cosmological parameters’. *Astronomy and Astrophysics* **641**:A6.

D. M. Pozar (2011). ‘Microwave Engineering’.

M. E. Presley, et al. (2015). ‘Measuring the Cosmological 21 cm Monopole with an Interferometer’. *The Astrophysical Journal* **809**(1):18.

D. C. Price (2016). ‘PyGSM: Python interface to the Global Sky Model’.

D. C. Price (2021). ‘Spectrometers and Polyphase Filterbanks in Radio Astronomy’. In A. Wolszczan (ed.), *The WSPC Handbook of Astronomical Instrumentation*, pp. 159–179. World Scientific.

D. C. Price, et al. (2018). ‘Design and characterization of the Large-aperture Experiment to Detect the Dark Age (LEDA) radiometer systems’. *Monthly Notices of the Royal Astronomical Society* p. sty1244.

J. R. Pritchard & A. Loeb (2010). ‘Constraining the unexplored period between the dark ages and reionization with observations of the global 21 cm signal’. *Phys. Rev. D* **82**:023006.

J. R. Pritchard & A. Loeb (2012). ‘21 cm cosmology in the 21st century’. *Reports on Progress in Physics* **75**(8):086901.

A. Raghunathan, et al. (2011). ‘Zero-spacing interferometer for measurement of continuum spectrum of the low radio frequency background’. In *2011 XXXth URSI General Assembly and Scientific Symposium*, pp. 1–1.

A. Raghunathan, et al. (2020). ‘A Floating Octave Bandwidth Cone-Disc Antenna for Detection of Cosmic Dawn’. *IEEE Transactions on Antennas and Propagation, submitted* .

J. Randa (2008). ‘Recommended terminology for microwave radiometry’. Tech. rep., National Bureau of Standards, Gaithersburg, MD.

W. D. Reeve (2017). ‘MWA Antenna Description as Supplied by Reeve’. Tech. rep., Reeve Observatory.

B. C. Rhodes (2011). ‘PyEphem: Astronomical Ephemeris for Python’.

A. E. E. Rogers & J. D. Bowman (2012). ‘Absolute calibration of a wideband antenna and spectrometer for accurate sky noise temperature measurements’. *Radio Science* **47**:RS0K06.

A. E. E. Rogers & J. D. Bowman (2012). ‘Absolute calibration of a wideband antenna and spectrometer for accurate sky noise temperature measurements’. *Radio Science* **47**(6).

A. E. E. Rogers, et al. (2004). ‘Calibration of active antenna arrays using a sky brightness model’. *Radio Science* **39**(2).

V. Rumsey (1957). ‘Frequency independent antennas’. In *1958 IRE International Convention Record*, vol. 5, pp. 114–118.

G. B. Rybicki & A. P. Lightman (1979). ‘Radiative processes in astrophysics’.

M. Sathyanarayana Rao, et al. (2015). 'On the Detection of Spectral Ripples from the Recombination Epoch'. *The Astrophysical Journal* **810**(1):3.

M. Sathyanarayana Rao, et al. (2017a). 'GMOSS: All-sky Model of Spectral Radio Brightness Based on Physical Components and Associated Radiative Processes'. *The Astronomical Journal* **153**(1):26.

M. Sathyanarayana Rao, et al. (2017b). 'Modeling the Radio Foreground for Detection of CMB Spectral Distortions from the Cosmic Dawn and the Epoch of Reionization'. *The Astrophysical Journal* **840**:33.

R. J. Sault, et al. (1995). 'A Retrospective View of MIRIAD'. In R. A. Shaw, H. E. Payne, & J. J. E. Hayes (eds.), *Astronomical Data Analysis Software and Systems IV*, vol. 77 of *Astronomical Society of the Pacific Conference Series*, p. 433.

H. A. Schuessler, et al. (1969). 'Hyperfine Structure of the Ground State of ^{3}He$^+$ by the Ion-Storage Exchange-Collision Technique'. *Physical Review* **187**(1):5–38.

G. Schwarz (1978). 'Estimating the Dimension of a Model'. *Annals of Statistics* **6**:461–464.

D. Scott & M. J. Rees (1990). 'The 21-cm line at high redshift: a diagnostic for the origin of large scale structure'. *Monthly Notices of the Royal Astronomical Society* **247**:510.

S. K. Sethi (2005). 'HI signal from re-ionization epoch'. *Monthly Notices of the Royal Astronomical Society* **363**(3):818–830.

C. Shannon (1949). 'Communication in the Presence of Noise'. *Proceedings of the IRE* **37**(1):10–21.

P. A. Shaver, et al. (1999). 'Can the reionization epoch be detected as a global signature in the cosmic background?'. *Astronomy and Astrophysics* **345**:380–390.

P. H. Sims & J. C. Pober (2019). 'Testing for calibration systematics in the EDGES low-band data using Bayesian model selection'. *Monthly Notices of the Royal Astronomical Society* **492**(1):22–38.

S. Singh, et al. (2022). 'On the detection of a cosmic dawn signal in the radio background'. *Nature Astronomy* **6**:607–617.

S. Singh & R. Subrahmanyan (2019). 'The Redshifted 21 cm Signal in the EDGES Low-band Spectrum'. *The Astrophysical Journal* **880**(1):26.

S. Singh, et al. (2017). 'First Results on the Epoch of Reionization from First Light with SARAS 2'. *The Astrophysical Journal* **845**(2):L12.

S. Singh, et al. (2018a). 'SARAS 2 Constraints on Global 21 cm Signals from the Epoch of Reionization'. *The Astrophysical Journal* **858**(1):54.

S. Singh, et al. (2018b). 'SARAS 2: a spectral radiometer for probing cosmic dawn and the epoch of reionization through detection of the global 21-cm signal'. *Experimental Astronomy* **45**(2):269–314.

S. Singh, et al. (2015). 'On the Detection of Global 21-cm Signal from Reionization Using Interferometers'. *The Astrophysical Journal* **815**(2):88.

O. M. Smirnov (2011). 'Revisiting the radio interferometer measurement equation. I. A full-sky Jones formalism'. *Astronomy and Astrophysics* **527**:A106.

G. Smoot, et al. (1990). 'COBE Differential Microwave Radiometers: Instrument Design and Implementation'. *The Astrophysical Journal* **360**:685.

A. Sokasian, et al. (2002). 'The epoch of helium reionization'. *Monthly Notices of the Royal Astronomical Society* **332**(3):601–616.

M. Sokolowski, et al. (2017). 'Calibration and Stokes Imaging with Full Embedded Element Primary Beam Model for the Murchison Widefield Array'. *Publications of the Astronomical Society of Australia* **34**:e062.

M. Sokolowski, et al. (2015). ‘BIGHORNS - Broadband Instrument for Global HydrOgen ReioNisation Signal’. *Publications of the Astron. Soc. of Australia* **32**:4.

M. Sokolowski, et al. (2015). ‘The statistics of low frequency radio interference at the Murchison Radio-astronomy Observatory’. In *2015 IEEE Global Electromagnetic Compatibility Conference (GEMCCON)*, pp. 1–6.

M. Sokolowski, et al. (2015). ‘The Impact of the Ionosphere on Ground-based Detection of the Global Epoch of Reionization Signal’. *The Astrophysical Journal* **813**(1):18.

D. N. Spergel, et al. (2007). ‘Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology’. *The Astrophysical Journal Supplemental* **170**(2):377–408.

D. N. Spergel, et al. (2003). ‘First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters’. *The Astrophysical Journal Supplemental* **148**(1):175–194.

R. Subrahmanyan & R. Cowsik (2013). ‘Is there an Unaccounted for Excess in the Extragalactic Cosmic Radio Background?’. *The Astrophysical Journal* **776**(1):42.

A. T. Sutinjo, et al. (2020). ‘Design Equations for Closely Spaced Two-Element Interferometer for Radio Cosmology’. *URSI Radio Science Letters* **2**.

G. Swarup, et al. (1991). ‘The Giant Metre-Wave Radio Telescope’. *Current Science* **60**:95.

H. Tashiro (2014). ‘CMB spectral distortions and energy release in the early universe’. *Progress of Theoretical and Experimental Physics* **2014**(6):06B107.

J. A. Tauber, et al. (2010). ‘Planck pre-launch status: The Planck mission’. *Astronomy and Astrophysics* **520**:A1.

The HERA Collaboration, et al. (2021). 'First Results from HERA Phase I: Upper Limits on the Epoch of Reionization 21 cm Power Spectrum'. *arXiv e-prints* p. arXiv:2108.02263.

J. N. Thekkeppattu, et al. (2022). 'System design and calibration of SITARA—a global 21 cm short spacing interferometer prototype'. *Publications of the Astronomical Society of Australia* **39**:e018.

S. J. Tingay, et al. (2013). 'The Murchison Widefield Array: The Square Kilometre Array Precursor at Low Radio Frequencies'. *Publications of the Astronomical Society of Australia* **30**:e007.

S. J. Tingay, et al. (2020). 'A survey of spatially and temporally resolved radio frequency interference in the FM band at the Murchison Radio-astronomy Observatory'. *Publications of the Astronomical Society of Australia* **37**:e039.

C. M. Trott (2017). 'The Square Kilometre Array Epoch of Reionisation and Cosmic Dawn Experiment'. *Proceedings of the International Astronomical Union* **12**(S333):92–97.

C. M. Trott, et al. (2020). 'Deep multiredshift limits on Epoch of Reionization 21 cm power spectra from four seasons of Murchison Widefield Array observations'. *Monthly Notices of the Royal Astronomical Society* **493**(4):4711–4727.

W. Turner (2016). 'SKA-TEL-SKO-0000008, SKA Phase 1 System Requirements Specification'. Tech. rep., SKA Organisation.

A. J. J. van Es, et al. (2020). 'A prototype model for evaluating SKA-LOW station calibration'. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, vol. 11445 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, p. 1144589.

M. P. van Haarlem, et al. (2013). 'LOFAR: The LOw-Frequency ARray'. *Astronomy and Astrophysics* **556**:A2.

D. A. Varshalovich & V. K. Khersonskii (1977). 'Distortion of the primordial radiation spectrum by the 21-cm hydrogen line at epochs z = 150-15'. *Soviet Astronomy Letters* **3**:155.

R. Vautard & M. Ghil (1989). 'Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series'. *Physica D: Nonlinear Phenomena* **35**(3):395–424.

H. K. Vedantham, et al. (2015). 'Lunar occultation of the diffuse radio sky: LOFAR measurements between 35 and 80 MHz'. *Monthly Notices of the Royal Astronomical Society* **450**(3):2291–2305.

H. K. Vedantham, et al. (2014). 'Chromatic effects in the 21 cm global signal from the cosmic dawn'. *Monthly Notices of the Royal Astronomical Society* **437**(2):1056–1069.

T. Venumadhav, et al. (2016). 'A Practical Theorem on Using Interferometry to Measure the Global 21-cm Signal'. *The Astrophysical Journal* **826**(2):116.

P. Virtanen, et al. (2020). 'SciPy 1.0: fundamental algorithms for scientific computing in Python'. *Nature Methods* **17**:261–272.

T. C. Voytek, et al. (2014). 'Probing the Dark Ages at z ˜ 20: The SCI-HI 21 cm All-sky Spectrum Experiment'. *The Astrophysical Journal Letters* **782**:L9.

R. A. Watson, et al. (2003). 'First results from the Very Small Array — I. Observational methods'. *Monthly Notices of the Royal Astronomical Society* **341**(4):1057–1165.

R. Wayth, et al. (2017). 'The Engineering Development Array: A Low Frequency Radio Telescope Utilising SKA Precursor Technology'. *Publications of the Astronomical Society of Australia* **34**:e034.

R. Wayth, et al. (2022). 'Engineering Development Array 2: design, performance, and lessons from an SKA-Low prototype station'. *Journal of Astronomical Telescopes, Instruments, and Systems* **8**:011010.

R. B. Wayth, et al. (2018). ‘The Phase II Murchison Widefield Array: Design overview’. *Publications of the Astronomical Society of Australia* **35**:e033.

S. A. Wouthuysen (1952). ‘On the excitation mechanism of the 21-cm (radio-frequency) interstellar hydrogen emission line.’. *The Astronomical Journal* **57**:31–32.

J. S. B. Wyithe & A. Loeb (2003). ‘Reionization of Hydrogen and Helium by Early Stars and Quasars’. *The Astrophysical Journal* **586**(2):693–708.

H. Zheng, et al. (2017). ‘An improved model of diffuse galactic radio emission from 10 MHz to 5 THz’. *Monthly Notices of the Royal Astronomical Society* **464**(3):3486–3497.

A. Zonca, et al. (2019). ‘healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in Python’. *The Journal of Open Source Software* **4**(35):1298.

Every reasonable effort has been made to acknowledge the owners of copyright material. I would be pleased to hear from any copyright owner who has been omitted or incorrectly acknowledged.