Searching for the Global 21 cm Cosmic Dawn Absorption Signal with the Long Wavelength Array

Christopher DiLullo

University of New Mexico - Main Campus

Follow this and additional works at: https://digitalrepository.unm.edu/phyc_etds

Part of the Astrophysics and Astronomy Commons, and the Physics Commons

Recommended Citation


This Dissertation is brought to you for free and open access by the Electronic Theses and Dissertations at UNM Digital Repository. It has been accepted for inclusion in Physics & Astronomy ETDs by an authorized administrator of UNM Digital Repository. For more information, please contact disc@unm.edu.
Christopher DiLullo

Candidate

Physics and Astronomy

Department, The University of New Mexico

This dissertation is approved and is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Dr. Gregory B. Taylor, Chair

Dr. Jayce Dowell

Dr. Ylva Pihlström

Dr. Rouzbeh Allahverdi

Dr. Frank Schinzel

Defended June 24, 2021.
Searching for the Global 21 cm Cosmic Dawn Absorption Signal with the Long Wavelength Array

by

Christopher DiLullo

M.S. Physics, University of New Mexico, 2018

DISSERTATION

Submitted in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy
Physics

The University of New Mexico
Albuquerque, New Mexico

July, 2021
Dedication

To my father,

who inspired my love of learning.
Acknowledgments

This PhD dissertation would not have been possible without the incredible support group I have in the form of both family and friends.

First, I’d like to thank Greg Taylor for having been an excellent advisor and having listened to me during my first year at UNM when I told him I primarily wanted to study cosmology. When the opportunity to work on this project presented itself he offered it to me and was very honest in what a difficult undertaking it would be. He truly listens to his students and I could not be more appreciative. He has always offered wise advice along the way and has supported me by encouraging me to take on challenges, whether they be hunting for a tiny cosmological signal or surviving the sweltering heat of the United Arab Emirates in July.

I also could not have done any of this without the massive amount of help I received from Jayce Dowell. Jayce is known for his mastery at programming, which has undoubtedly made me a better programmer than any class ever could have, but what I really have learned from Jayce is how to be a better scientific thinker. His ability to analyze a difficult problem and devise a plan to proceed is unparalleled and has helped me tremendously throughout graduate school. His help was crucial in developing the methodology presented in this dissertation.

I would also like to thank all of my friends at UNM who have helped keep me sane over the last few years. Whether it was struggling together in classes, grabbing beers at the breweries, or hiking out in the wilderness, you all have a special place in my heart. A special thanks goes out to past and present members of my research group, especially Joe Malins, Savin Varghese, Karishma Bansal, Seth Bruzewski, Ken Obenberger, and Pratik Kumar, for the stimulating discussions and help.

I’d like to thank my fiancée, Adah, and my mom, Camille. Adah, thank you for the constant love and support even though we spent hours in confusion as my physics jargon clashed with your statistics jargon. Thank you for taking this life journey with me. Mom, I can’t even properly thank you for all that you have done to help get me to this point. You always encouraged me to chase my dreams, no matter how far away I had to go to achieve them. Thank you for everything.

Last, but certainly not least, I’d like to thank my undergraduate advisor, Leslie (Doc) Brown. Doc, thank you for all the support you have given me over the years. It is because of you that I even looked at New Mexico for graduate school and now I am following in your footsteps as a radio astronomer. I will always cherish my time in the physics department at Connecticut College and some of my best memories are chatting with you in the lounge. You could not have prepared me better for this journey.
Searching for the Global 21 cm Cosmic Dawn Absorption Signal with the Long Wavelength Array

by

Christopher DiLullo

M.S. Physics, University of New Mexico, 2018
Ph.D., Physics, University of New Mexico, 2022

Abstract

The redshifted 21 cm signal from neutral hydrogen offers one of the best observational probes of Cosmic Dawn and the Epoch of Reionization. This dissertation presents an effort to detect the redshifted 21 cm signal using the Long Wavelength Array station located on the Sevilleta National Wildlife Refuge in New Mexico, USA (LWA–SV). The major goal is to validate the potential detection reported by the EDGES collaboration. This measurement requires a dynamic range on the order of $10^5$ in order to disentangle the cosmological signal from the Galactic foregrounds. The beamforming capability of LWA–SV is novel to this search. The presented work introduces observational and data analysis methodology as well as an achromatic beamforming framework for LWA–SV. Residual RMS limits on the order of a few Kelvin, a factor of $10^3$ times lower than the foreground brightness, are achieved and possible strategies for improvement are presented.
# Contents

List of Figures xi

List of Tables xiv

List of Symbols xv

List of Acronyms xvi

1 Introduction 1

1.1 Cosmological Overview 1

1.1.1 Cosmic Dawn and The Epoch of Reionization 3

1.1.2 21 cm Cosmology 7

1.2 The Experiment to Detect the Global EoR Signature 19

1.3 The Long Wavelength Array 22

2 Using the Long Wavelength Array to Search for Cosmic Dawn 26

2.1 Introduction 27

2.2 The Long Wavelength Array 30

2.3 Observations 33
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.4 Results and Current Limits</td>
<td>36</td>
</tr>
<tr>
<td>2.5 Discussion</td>
<td>40</td>
</tr>
<tr>
<td>2.5.1 Calibration</td>
<td>40</td>
</tr>
<tr>
<td>2.5.2 Modelling</td>
<td>42</td>
</tr>
<tr>
<td>2.6 Custom Beamforming</td>
<td>45</td>
</tr>
<tr>
<td>2.7 Future Work</td>
<td>47</td>
</tr>
<tr>
<td>2.8 Summary</td>
<td>48</td>
</tr>
<tr>
<td>2.9 Acknowledgments</td>
<td>50</td>
</tr>
<tr>
<td>3 Improvements to the Search for Cosmic Dawn Using the Long Wavelength Array</td>
<td>51</td>
</tr>
<tr>
<td>3.1 Introduction</td>
<td>52</td>
</tr>
<tr>
<td>3.2 Recent Improvements</td>
<td>54</td>
</tr>
<tr>
<td>3.2.1 Observing Strategy and Data Analysis</td>
<td>54</td>
</tr>
<tr>
<td>3.2.2 Upgrades to LWA–SV</td>
<td>57</td>
</tr>
<tr>
<td>3.3 Achromatic Beamforming</td>
<td>58</td>
</tr>
<tr>
<td>3.3.1 Implementation on LWA–SV</td>
<td>59</td>
</tr>
<tr>
<td>3.3.2 Beam Simulator: A Python Package to Simulate Array Beam Patterns</td>
<td>62</td>
</tr>
<tr>
<td>3.4 Recent Observing Campaign and Results</td>
<td>64</td>
</tr>
<tr>
<td>3.5 Discussion and Conclusion</td>
<td>68</td>
</tr>
<tr>
<td>3.6 Acknowledgments</td>
<td>70</td>
</tr>
<tr>
<td>4 Achromatic Beamforming at LWA–SV</td>
<td>71</td>
</tr>
</tbody>
</table>
4.1 Introduction .............................................................................. 71
4.2 Theoretical Framework ............................................................. 72
  4.2.1 Choice of Weighting Function ............................................. 73
4.3 The Advanced Digital Processor ................................................ 76
4.4 Results and Discussion ............................................................. 79
4.5 Beam Simulator: Description and Tutorial ................................. 80
  4.5.1 Representing An Antenna Array .......................................... 80
  4.5.2 Simulating the Array Beam Pattern ..................................... 82
  4.5.3 Tutorial .............................................................................. 83
5 Measurements of the LWA Dipole Impedance Mismatch .................. 85
  5.1 Introduction ............................................................................ 85
  5.2 Measurement Setup ............................................................... 86
  5.3 Results and Conclusion .......................................................... 87
6 Conclusions and Future Work ...................................................... 91
  6.1 Summary and Conclusions ...................................................... 91
  6.2 Future Work ......................................................................... 93
    6.2.1 Developing a Robust RFI Detection Algorithm ................. 93
    6.2.2 Measuring the LWA Dipole Gain Pattern ......................... 94
    6.2.3 Alternative Achromatic Beamforming Frameworks ........... 95
    6.2.4 Improving Modelling Robustness with Simulations ............ 96
    6.2.5 Choice of Sky Model ....................................................... 96
## List of Figures

1.1 Timeline of cosmic history. ......................................................... 2  
1.2 A series of quasar spectra taken by the Sloan Digital Sky Survey (Fan et al., 2006). ................................................................. 5  
1.3 The evolution of the 21 cm signal with redshift (Pritchard and Loeb, 2012). ................................................................. 9  
1.4 The hyperfine structure of hydrogen (Pritchard and Loeb, 2012). .... 11  
1.5 Various models of the 21 cm global signal which assume different astrophysics (Cohen et al., 2017). ................................. 13  
1.6 Dimensionless 21 cm power spectrum for various mean ionization fractions (Furlanetto, 2016, Lidz et al., 2008). ............................. 17  
1.7 Observable Fourier modes in the $k_{\parallel}$–$k_{\perp}$ plane. ...................... 18  
1.8 Summary of the EDGES detection (Bowman et al., 2018). ............ 20  
1.9 LWA antennas deployed at LWA–SV. ........................................... 23  
1.10 The frequency response of each ARX filter mode at LWA–SV (Dowell and Taylor, 2020b). ................................................................. 25  

2.1 Median power across frequency vs ASP electronics temperature. .... 32  
2.2 Uncorrected and corrected median drifts. ...................................... 33
2.3 Log-log plot of r.m.s. noise vs integration time for a Gaussian noise source. ................................................................. 34

2.4 Raw spectra of the Science Field (top) and Virgo A (bottom) for XX and YY polarizations. .................................................. 37

2.5 Temperature scale factors derived from the observed power of Virgo A and simulated beam temperatures from the Global Sky Model. . . 38

2.6 Calibrated spectra in units of Kelvin. ........................................ 39

2.7 N=5 smooth polynomial model and residuals. ............................ 41

2.8 Power law model and residuals. .............................................. 42

2.9 Average residual r.m.s. vs. integration time for each polarization. . 43

2.10 3-d weighting plot showing the antenna locations (crosses) and their weight values (dots). .................................................. 46

2.11 Simulated XX polarization beam patterns for an unshaped and shaped beam. ................................................................. 48

2.12 Observed drift curves for an unshaped beam and a custom shaped beam. ................................................................. 49

3.1 Model dynamic spectrum of Virgo A. ........................................ 56

3.2 Interferometric basket weave results from observations of Cygnus A. 63

3.3 Calibrated datasets for the entire observing campaign at 30 second integration time. .................................................. 66

3.4 Results of the MCMC fit. ..................................................... 67

3.5 Posterior distributions for foreground model parameters for XX polarization. ................................................................. 68

3.6 Results of the maximally smooth function foreground model fit. . . 69
4.1 Gaussian and Blackman-Harris weighting functions. 75
4.2 Simulated beam patterns using both Gaussian and Blackman-Harris weighting functions. 75
4.3 Beamforming example code snippet. 84
4.4 Standard beam pattern of LWA–SV generated by Beam Simulator. 84
5.1 Testing fixtures designed by NRL. 86
5.2 Impedance measurement done at the LWA–SV outrigger. 88
5.3 Locations of the antenna stands with measured stands marked. 89
5.4 Real and imaginary components of the complex impedance measured at three different stands. 89
5.5 Impedance mismatch efficiency of the three measured stands. 90
List of Tables

2.1 Radiometer Equation Assumptions and Results for $T_{rms} = 50$ mK. . . 30
2.2 Foreground Model Best Fit Parameters . . . . . . . . . . . . . . . . . . 40
3.1 Beam-Dipole Mode Basket Weave Results . . . . . . . . . . . . . . . . 62
3.2 Residual RMS Across Frequency . . . . . . . . . . . . . . . . . . . . . . 68
List of Symbols

- $z$: redshift
- $\Omega$: density parameter
- $H_0$: Hubble constant
- $H$: Hubble parameter
- $h$: “little h”, $\equiv H_0 / 100$ km/s/Mpc
- $T_B$: brightness temperature
- $T_S$: spin temperature
- $T_K$: kinetic temperature
- $\nu$: frequency
- $t$: time
- $A_e$: effective area
- $\sigma$: standard deviation
- $k_B$: Boltzmann constant
- $S_\nu$: flux density
- $Z$: impedance
## List of Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADP</td>
<td>Advanced Digital Processor</td>
</tr>
<tr>
<td>ASP</td>
<td>Analog Signal Processor</td>
</tr>
<tr>
<td>ASSASSIN</td>
<td>All-Sky SignAl Short-Spacing INterferometer</td>
</tr>
<tr>
<td>BIGHORNS</td>
<td>Broadband Instrument for Global HydrOgen ReioNisation Signal</td>
</tr>
<tr>
<td>BFU</td>
<td>Beamformer Unit</td>
</tr>
<tr>
<td>DAPPER</td>
<td>Dark Ages Polarimeter PathfinER</td>
</tr>
<tr>
<td>DFT</td>
<td>Discrete Fourier Transform</td>
</tr>
<tr>
<td>EDGES</td>
<td>Experiment to Detect the Global EoR Signature</td>
</tr>
<tr>
<td>EoR</td>
<td>Epoch of Reionization</td>
</tr>
<tr>
<td>FEE</td>
<td>Front End Electronics</td>
</tr>
<tr>
<td>FFT</td>
<td>Fast Fourier Transform</td>
</tr>
<tr>
<td>FWHM</td>
<td>Full Width at Half Maximum</td>
</tr>
<tr>
<td>FWFM</td>
<td>Full Width at Fifth Maximum</td>
</tr>
<tr>
<td>GMRT</td>
<td>Giant Metrewave Radio Telescope</td>
</tr>
<tr>
<td>GMROSS</td>
<td>Global Model for the Radio Sky Spectrum</td>
</tr>
<tr>
<td>GSM</td>
<td>Global Sky Model</td>
</tr>
<tr>
<td>HERA</td>
<td>Hydrogen Epoch of Reionization Array</td>
</tr>
<tr>
<td>IME</td>
<td>Impedance Mismatch Efficiency</td>
</tr>
<tr>
<td>LEDA</td>
<td>Large–Aperture Experiment to Detect the Dark Ages</td>
</tr>
<tr>
<td>LFSM</td>
<td>Low Frequency Sky Model</td>
</tr>
<tr>
<td>LOFAR</td>
<td>Low Frequency Array</td>
</tr>
<tr>
<td>LSL</td>
<td>LWA Software Library</td>
</tr>
<tr>
<td>LST</td>
<td>Local Sidereal Time</td>
</tr>
</tbody>
</table>
LWA  Long Wavelength Array
LWA–SV  Long Wavelength Array Sevilleta
LWA–OVRO  Long Wavelength Array Owens Valley Radio Observatory
MCMC  Markov Chain Monte Carlo
MSF  Maximally Smooth Function
MWA  Murchison Widefield Array
PAPER  Precision Array to Probe the Epoch of Reionization
RMS  Root Mean Square
RFI  Radio Frequency Interference
ROACH  Reconfigurable Open Architecture Computing Hardware
SARAS  Shaped Antenna measurement of the background RA dio Spectrum
SCI–HI  Sonda Cosmológica de las Islas para la Detección de Hidrógeno Neutro
SEFD  System Equivalent Flux Density
SKA  Square Kilometre Array
SNR  Signal to Noise Ratio
UTC  Universal Coordinated Time
Chapter 1

Introduction

1.1 Cosmological Overview

Standard cold dark matter cosmology ($\Lambda$CDM) begins with the Big Bang and attempts to describe the evolution of the Universe throughout cosmic history until the present. Figure 1.1 shows a simplified overview of cosmic history with major events labeled. After the Big Bang, the Universe is a largely homogenous distribution of protons, electrons, photons, dark matter, and neutrinos. Small inhomogeneities which are present in the density field of the Universe eventually lead to the growth of large scale structure which we observe today. During this early period in cosmic history, the high temperature of the Universe prevents electrons from binding to protons. The mean free path of a photon is very small due to Thompson scattering in such a dense plasma and so the Universe is opaque to electromagnetic radiation. However, roughly 370,000 years after the Big Bang, the Universe had eventually expanded and cooled sufficiently that electrons can effectively bind to protons to form neutral hydrogen during the epoch known as Recombination. This marks the creation of the Cosmic Microwave Background (CMB) and an important phase transition where the Universe changes from being opaque to electromagnetic radiation to being transparent. After Recombination, the mean free path for a photon changed from nearly zero to the Hubble scale in a very short amount of time.
Chapter 1. Introduction

The conditions of the early Universe are well understood through high precision measurements of the CMB (Aghanim et al., 2020, Bennett et al., 1996, 2013) which measure anisotropies in the CMB to one part in 10,000. The CMB is the relic heat left over in the Universe after the Big Bang and studying it with extreme precision yields insights into the seeds of structure which evolve into the Universe we observe today. After the CMB formed and the Universe became transparent, the photons decoupled from the matter and gravitational interactions began to dominate the dynamics of the matter. This period in cosmic history is known as the Dark Ages since the Universe mainly consisted of neutral hydrogen and helium which did not emit light. The Dark Ages lasted until the Universe expanded and cooled enough for the gravitational collapse of the first overdensities to form the first luminous sources. The formation of the first luminous sources is known as either First Light or Cosmic Dawn. After Cosmic Dawn, the light from the first luminous sources ionized most of the surrounding gas during the Epoch of Reionization (EoR).

Cosmic Dawn and the Epoch of Reionization remain some of the final frontiers of observational cosmology. We know from observations of the CMB that the Universe
was once a homogeneous distribution of neutral gas and we know that the Universe today consists of stars and galaxies and that most of the hydrogen present in the Universe exists in an ionized state; but, understanding how the Universe transitioned between these neutral and ionized states is not straightforward. It is believed, from observations of the Gunn–Peterson trough (Gunn and Peterson, 1965), that the EoR is mostly finished by redshift $z \approx 6$, which places it tantalizingly just beyond current direct detection limits in the optical with the Hubble Space Telescope (HST) due to the declining number of bright galaxies at high redshift (Ishigaki et al., 2018, McLure et al., 2010). The major questions surrounding Cosmic Dawn and the EoR are:

- When and how did the first stars form?
- What are the first galaxies?
- When and how did the first active galactic nuclei form?

The detection of Cosmic Dawn and the EoR is a major science goal of agencies around the world and a major science driver for international projects such as the James Webb Space Telescope (JWST; Gardner et al., 2006), the Square Kilometre Array (SKA; Braun et al., 2015), and the Hydrogen Epoch of Reionization Array (HERA; DeBoer et al., 2017).

### 1.1.1 Cosmic Dawn and The Epoch of Reionization

The prevailing idea before the detection of the first quasars was that the majority of the Universe’s matter density was in the form of neutral hydrogen. However, by the early 1960’s a number of quasars had been observed with quasar 3C 9 holding the distance record at $z = 2.01$ (Schmidt, 1965). Quasar 3C 9 showed two emission lines which were deduced to be the Lyman–$\alpha$ and C IV spectral lines. Gunn and Peterson (1965) realized that the presence of Lyman–$\alpha$ in the spectrum of a quasar at $z = 2.01$ meant that the amount of neutral hydrogen present in the Universe must be far below what was expected at the time or else this emission line would have been absorbed by
neutral hydrogen in the intergalactic medium (IGM). Thus, hydrogen must mainly exist in an ionized form in the local Universe.

The picture became clearer in the early 1970’s with the detection of quasar 4C 05.34 ($z = 2.877$) whose spectrum showed a series of strong absorption lines to the short–wavelength side of the Lyman–$\alpha$ emission line (Bahcall and Goldsmith, 1971, Lynds, 1971). It was posited that the presence of multiple absorption lines on the short–wavelength side of the Lyman–$\alpha$ transition was due to presence of clouds of neutral hydrogen which exist at different redshifts within the IGM along the line of sight to the quasar. As emission from these distant sources traverses the IGM, it encounters these neutral clouds and the fraction of the emission which is at the local Lyman–$\alpha$ resonance for that relative redshift difference is absorbed. This series of absorption lines is known today as the Lyman–$\alpha$ forest.

The Lyman–$\alpha$ forest plays an important role in understanding the distribution of neutral hydrogen in the Universe. More distant sources show denser “forests” with more absorption lines since the emission has had to pass through more neutral clouds in the IGM. However, at redshifts larger than $z \approx 6$, the Lyman–$\alpha$ forest becomes so dense that the absorption lines bleed into one another until almost all emission on the short–wavelength side of the Lyman–$\alpha$ line has been completely suppressed. This is known as the Gunn–Peterson trough. Figure 1.2 shows the spectra of multiple quasars taken by the Sloan Digital Sky Survey. The presence of a dense Lyman–$\alpha$ forest can be seen in the spectra of quasars whose redshift is $z \lesssim 6$ and a more complete Gunn–Peterson trough can be seen in those spectra whose redshift is $z \gtrsim 6$ (Fan et al., 2006).

The presence of the Gunn–Peterson trough at high redshifts implies that in the past there was so much neutral hydrogen present that essentially all light with wavelengths between the Lyman–$\alpha$ line and the Lyman limit were immediately absorbed. This is strong evidence for the Universe having previously been almost entirely neutral and then having undergone some transition which left it mostly ionized by $z \approx 6$. However, as more became understood about the hydrogen ionization history, the question of “What caused this ionizing transition?” still remained. The answer to
Figure 1.2: A series of quasar spectra taken by the Sloan Digital Sky Survey (Fan et al., 2006). Each spectrum shows suppression of emission at wavelengths shorter than that of the Lyman–α wavelength. At $z \lesssim 6$, the suppression is incomplete across wavelength and is known as the Lyman–α forest. At $z \gtrsim 6$, the suppression is more complete and is known as the Gunn-Peterson trough.
this question was pinned down using measurements of the CMB spectrum. If the low–density IGM was very hot ($T_{\text{IGM}} \gtrsim 10^5$ K), like what was proposed in Gunn and Peterson (1965), then this would lead to large electron scattering which would impose large distortions on the spectral shape of the CMB (Hogan et al., 1982, Sunyaev and Zel’Dovich, 1980). Thus, measurements of the CMB were used to constrain the temperature of the low–density IGM to $T_{\text{IGM}} \approx 10^4$ K. This temperature is more consistent with a background UV flux which arises from stars and quasars which continually keeps the low–density IGM ionized (Haardt and Madau, 1995).

Understanding the nature of this ionizing UV background remains a major goal. UV photons with energies greater than or equal to the Lyman limit are capable of photoionizing neutral hydrogen, so measuring the ionization history of hydrogen yields information about the UV background history. The major sources of this UV background are starburst galaxies and quasars. At redshifts of $z \lesssim 3$, quasars dominate in their contribution to the UV background, but it appears starburst galaxies dominate at higher redshifts (Faucher-Giguere et al., 2008, Madau et al., 1999, Shapiro and Giroux, 1987). Measuring the luminosity function, which measures the number of sources for a given luminosity range, of UV emitting galaxies at high redshifts is of particular interest as this would shed light on what types of sources drive reionization. Current surveys of high redshift galaxies combined with measurements of the optical depth of reionization from the Planck telescope suggest that the UV luminosity function decreases at redshifts of $z \gtrsim 5$ (Bouwens et al., 2015a,b, Dickinson et al., 2004, Khusanova et al., 2020). JWST is expected to tightly constrain the UV luminosity function in the redshift range $4 \leq z \leq 10$ as it will probe much deeper than previous observations using the Hubble Space Telescope (HST). Recent work has focused on making model–informed predictions of what JWST will see in its survey of the high redshift Universe (Mason et al., 2015, Vogelsberger et al., 2020, Yung et al., 2019).
Although measurements of the Lyman–α forest/Gunn-Peterson trough and CMB are strong probes of the Epoch of Reionization, they are not perfect (Furlanetto, 2016). First, the Lyman–α optical depth is huge, which means that any small amount of neutral hydrogen, anything larger than 1 part in 1000, will cause a complete GP-trough. Therefore, while it is a good predictor that the Universe was once neutral, it does not offer a good probe into redshifts during and before the EoR. On the other hand, measurements of the CMB are only sensitive to the total line of sight from the current epoch to the surface of last scattering. Thus, they do not offer the ability to discern what is happening at different redshifts. An alternative method of studying the early Universe during Cosmic Dawn and the EoR is to study the spectral signature of neutral hydrogen that was present during these times.

The 21 cm hyperfine transition in neutral hydrogen offers the strongest observational probe of the redshifts during and right before the Epoch of Reionization. This hyperfine transition occurs when the electron in the 1s ground state of a HI atom changes its spin orientation in such a way that the spin of the electron goes from being parallel to that of the proton to anti-parallel. This transition is known as the spin–flip transition. These two states have a small energy difference equal to $\Delta E \approx 5.874 \, \mu\text{eV}$ which corresponds to a wavelength of $\lambda \approx 21.1 \, \text{cm}$ and a frequency of $\nu \approx 1420 \, \text{MHz}$. However, for Cosmic Dawn and the EoR the 21 cm signal is expected to have redshifted to frequencies below 200 MHz. See Furlanetto et al. (2006), Morales and Wyithe (2010), Pritchard and Loeb (2012) for a detailed description of the uses of the 21 cm signal as a probe of cosmology.

The 21 cm transition is mostly used in cosmology to probe neutral hydrogen along the line of sight to some background radio source. If the background source is taken to be the CMB, then a standard radiative transfer analysis can be carried out within the Rayleigh-Jeans limit to express the intensity of the 21 cm transition as a brightness temperature. Specifically, a differential brightness temperature between the 21 cm
transition and the background CMB can be expressed as

$$\delta T_b \approx 27x_{HI}(1 + \delta_b) \left( \frac{\Omega_b h^2}{0.023} \right) \left( \frac{0.15}{\Omega_m h^2} \frac{1 + z}{10} \right)^{1/2} \left( \frac{T_S - T_R}{T_S} \right) \left[ \frac{\partial_r v_r}{(1 + z)H(z)} \right] \text{mK}, \quad (1.1)$$

where $x_{HI}$ is the fraction of neutral hydrogen, $\delta_b$ is the fractional overdensity of baryons, $\Omega_b$ is the density of baryons, $\Omega_m$ is the density of matter, $h$ is the Hubble constant divided by 100 km/s/Mpc, $z$ is redshift, $T_R$ is the brightness temperature of the background radiation field, and $T_S$ is the spin temperature of the neutral hydrogen. The final term accounts for the velocity gradient along the line of sight, $\partial_r v_r$. When $\delta T_b$ is positive the 21 cm signal is seen as emission above that of the CMB and when it is negative the 21 cm signal is seen as absorption against the CMB.

Most terms in Equation 1.1 are set by the choice of a cosmological model. The only terms which directly relate to the nature of the hydrogen are the neutral fraction, $x_{HI}$, and the spin temperature, $T_S$. The term spin temperature is unique to the study of HI, but $T_S$ is simply the excitation temperature of the neutral hydrogen. The excitation temperature of a population is defined by the Boltzmann factor, which is the ratio of probabilities of two states. If we define an “upper” state which has more energy than a “lower” state, then the Boltzmann factor is written as

$$\frac{p_u}{p_l} = \frac{g_u}{g_l} \exp \left( -\frac{E_u - E_l}{kT_{ex}} \right), \quad (1.2)$$

where $p$ is the probability of a given state, $g$ is the statistical weight of a state, $E$ is the energy of a state, $k$ is the Boltzmann constant, and the subscripts $u$ and $l$ denote the upper and lower states, respectively. Thus, the spin temperature of HI describes the relative population difference between HI atoms in the parallel spin state to HI atoms in the anti-parallel spin state.

Understanding which astrophysical processes affect $T_S$ and their dependence on redshift is important if an accurate history of the 21 cm signal is to be made. Three processes determine the spin temperature: 1) absorption and emission of 21 cm photons, 2) collisions between two HI atoms and between a HI atom and an electron, and 3) scattering of Lyman–α photons which can induce a spin–flip transition, known
as the Wouthuysen–Field (WF) effect (Field, 1958, Wouthuysen, 1952). These three process can be quantitatively expressed via

\[ T_S^{-1} = \frac{T_\gamma^{-1} + x_c T_K^{-1} + x_\alpha T_\alpha^{-1}}{1 + x_c + x_\alpha}, \]  

(1.3)

where \( T_\gamma \) is the temperature of the background radiation field, which is mainly the CMB, \( T_K \) is the kinetic temperature of the gas, \( T_\alpha \) is the color temperature\(^1\) of the Lyman–\( \alpha \) emission, and \( x_c \) and \( x_\alpha \) are coupling coefficients which describe the strength of collisional and Lyman–\( \alpha \) coupling, respectively. Throughout Cosmic Dawn and the EoR, as the first luminous sources begin to alter the astrophysics of the Universe, each of these above processes becomes dominant at different times and guides the evolution of the 21 cm signal. A model of evolution of the 21 cm signal (Pritchard and Loeb, 2012) is shown in Figure 1.3 and a discussion of the timeline follows below.

\[ \text{Figure 1.3: The evolution of the 21 cm signal with redshift (Pritchard and Loeb, 2012). This signal is a differential brightness temperature with respect to the CMB. The top panel shows a 2-d slice of cosmic volume with the signal strength shown in color (blue is absorption and red is emission). The bottom panel shows the 1-d amplitude of the signal with frequency. This 1-d signal is known as the global 21 cm signal. Major events throughout Cosmic Dawn and the EoR are labeled.} \]

\[ \text{During the Dark Ages at redshifts } z \gtrsim 200 \text{ the Universe was dense enough that } T_S \text{ was coupled to } T_K \text{ via collisional coupling. However, free electrons were coupled to} \]

\(^1\)A color temperature represents the “temperature” of a photon distribution in some frequency range assuming it is from a blackbody. This is not a true thermodynamic temperature as it can vary greatly from the effective temperature of the emitting body.
the CMB via Compton scattering, so $T_K = T_\gamma$. Thus, at these redshifts $T_S = T_\gamma$ and there is no differential 21 cm signal. However, between redshifts $40 \lesssim z \lesssim 200$ the Universe was still dense enough that collisional coupling dominated and set $T_S = T_K$, but the temperature of the gas adiabatically cooled faster than that of the CMB since $T_K \propto (1 + z)^2$ and $T_\gamma \propto 1 + z$. This drove $T_S < T_\gamma$ and so the 21 cm signal is seen in absorption. This absorption signal is labeled in Figure 1.3 as “Dark Ages.” This signal would yield insights into the conditions of the Universe during the Dark Ages and offers one of the only observational probes of this epoch; however, the amplitude is so small that this signal is generally considered undetectable with current technology.

Once the Universe had expanded enough that the density had fallen to where collisional coupling no longer dominated $T_S$, absorption of 21 cm photons from the CMB began to dominate. This reset $T_S = T_\gamma$ and once again there is no detectable 21 cm signal. However, things began to change when the first luminous sources began to turn on during Cosmic Dawn. The first stars are expected to have had large Lyman–$\alpha$ fluxes (Schaerer, 2002, 2003) which coupled $T_S$ to $T_\alpha$ via the WF effect. The WF effect couples $T_S$ to $T_\alpha$ by randomly inducing spin-flips via the absorption and emission of Lyman–$\alpha$ photons. This is illustrated in Figure 1.4 which shows the hyperfine structure of hydrogen. When a HI atom in one of the two ground states absorbs a Lyman–$\alpha$ photon, it is excited into one of the $n = 2$ states. Depending on the excited state it ends up in, the HI atom can emit a Lyman–$\alpha$ photon and end up in the other ground state. This is not guaranteed to happen, but if it does, a spin-flip has occurred. Thus, simply absorbing and emitting Lyman–$\alpha$ photons can induce spin-flips and affect the spin temperature of HI. The color temperature of the Lyman–$\alpha$ emission, $T_\alpha$, can be assumed to be equal to the kinetic temperature of the gas, $T_K$, since the surrounding gas was optically thick with respect to Lyman–$\alpha$ emission during this period. Thus, the WF effect drove $T_S = T_K < T_\gamma$ and so the 21 cm signal is once again seen in absorption. This second absorption feature can be seen in Figure 1.3. This is known as the Cosmic Dawn absorption signal. Methods to detect it will be discussed in Section 1.1.2.

After enough stars had formed and there was sufficient Lyman–$\alpha$ emission, this
coupling eventually saturated and the coupling strength between $T_S$ and $T_K$ reached a maximum. After saturation occurred, the X-ray flux from stars and the first black holes began to heat the gas and drove $T_K$ back towards $T_\gamma$; this is labeled as “Heating begins” in Figure 1.3. Eventually, the kinetic temperature of the gas was driven above the temperature of the CMB and the 21 cm signal is seen as emission for the first time. This was the beginning of the Epoch of Reionization. The ionization fraction began to become important here as HII regions grew around the first galaxies. The signal becomes largely spatially inhomogeneous, which can be seen in the top panel of Figure 1.3. The nature of the inhomogeneous EoR signal and methods to attempt to measure it are discussed further in Section 1.1.2. After the EoR ended the Universe was left primarily ionized with only a small number of clouds of neutral hydrogen remaining. These are the clouds which are observed in the Lyman–$\alpha$ forest. The cosmological 21 cm signal vanishes once the Universe is ionized.

Now that the timeline of Cosmic Dawn and the EoR has been laid out, it is useful to break the discussion of the 21 cm signal into two “components”. These are
the global 21 cm signal, which is the average amplitude of the differential brightness temperature averaged over the sky, and the 21 cm power spectrum, which attempts to measure the statistical fluctuations of the 21 cm signal as HII regions formed and grew throughout the EoR. The expected evolution of the global 21 cm signal is shown in the bottom panel of Figure 1.3 and the spatial fluctuations which are measured using the 21 cm power spectrum can be seen toward the end of the top panel of Figure 1.3. The latter is akin to power spectrum measurements of anisotropies in the CMB. The methods used to detect each signal vary greatly, but a detection of either would yield deep insights into the astrophysical conditions which governed the state of neutral hydrogen throughout Cosmic Dawn and the EoR.

The Global 21 cm Signal

The global 21 cm signal traces the lifespan of the first stars, so its spectral shape is highly sensitive to their properties. The global signal consists of the Cosmic Dawn absorption signal, which is caused by strong Lyman–α coupling between the background UV radiation field and HI, and the subsequent Epoch of Reionization emission signal, which is driven by heating due to the X-ray background originating from the very first active galactic nuclei. It eventually fades as the Universe becomes completely ionized. Thus, a detection of this signal across frequency would lead to direct observational constraints on the three major radiation fields which govern its evolution. These radiation fields are: 1) the UV background, 2) the X-ray background, and 3) the ionizing background.

Three major points in the evolution of the global signal have been identified which make it easier to discuss how a measurement of the signal would help constrain the associated astrophysics (Cohen et al., 2017). The three main events are: 1) the advent of the first stars, 2) the moment when X-ray heating begins to sufficiently heat the IGM, and 3) when the X-ray heating hits a maximum. The features in the signal corresponding to these events are called the high-z maximum, the minimum, and the low-z maximum, respectively. They can be seen labeled in Figure 1.5.
Chapter 1. Introduction

Figure 1.5: Various models of the 21 cm global signal which assume different astrophysics (Cohen et al., 2017). Major features which correspond to important events throughout Cosmic Dawn and the EoR are labeled.

The location of the high-z maximum in frequency places tight constraints on when the first stars began to form. This, relative to the location of the minimum in frequency, would also constrain how quickly Lyman–α coupling saturates and the spin temperature can no longer effectively decrease. This could help constrain the star formation efficiency and the minimum halo mass which is required for gravitational collapse at these redshifts. The amplitude of the Cosmic Dawn absorption signal at the minimum would place constraints on the relative strength of Lyman–α cooling to X-ray heating. The location of the minimum in frequency would place constraints on when X-ray heating begins to become sufficient to drive the gas temperature up. This relies on astrophysical parameters such as the X-ray heating efficiency and the
spectral energy distribution of X-rays. Finally, the location of the low-z maximum would place constraints on when the Universe becomes appreciably ionized such that the 21 cm signal begins to decline. It’s amplitude is sensitive to how efficiently X-rays can heat the surrounding IGM before ionization becomes dominant.

Detecting the global 21 cm signal is a major challenge as the expected amplitudes of both the Cosmic Dawn absorption signal and the EoR emission signal are extremely small. Figure 1.5 shows that even the most extreme models place the absorption amplitude to be less than $\approx 250 \text{ mK}$ and the emission signal to be on the order of a few tens of mK. In fact, most experiments attempting to detect the global 21 cm signal focus on detecting the Cosmic Dawn absorption signal because of its larger amplitude relative to the EoR emission signal. The small amplitudes of these signals, combined with the extremely bright emission ($T_B \geq 10^3 - 10^4 \text{ K}$) from the Milky Way at these low frequencies, means that any measurement of the global signal will require a dynamic range of $10^5$ in order to disentangle the cosmological signal from galactic foregrounds. The foregrounds consist of extragalactic point sources and galactic synchrotron and free-free emission/absorption. The spectral shape of these foreground contributions is expected to be smooth with frequency. Most experiments attempting to detect the global 21 cm signal try to reach the required dynamic range by using a single antenna element which is extremely well calibrated and measures the average spectrum of the sky. The hope is that if the antenna is sufficiently well calibrated, its response will be well enough understood to properly remove the spectrally smooth foreground contribution leaving the 21 cm signal to be detected in the residuals.

Experiments attempting to detect the global 21 cm signal include the Sonda Cosmológica de las Islas para la Detección de Hidrógeno Neutro (SCI–HI; Voytek et al., 2014), the Broadband Instrument for Global HydrOgen ReioNisation Signal (BIGHORNS; Sokolowski et al., 2015), the Shaped Antenna measurement of the background RAdio Spectrum 2 (SARAS 2; Singh et al., 2018), the Large–Aperture Experiment to Detect the Dark Ages (LEDA; Price et al., 2018), and the Experiment to Detect the Global EoR Signature (EDGES; Bowman et al., 2018) which all use
individual antennas to measure the sky-averaged spectrum. A new radio telescope named NenuFAR which uses Long Wavelength Array antennas is also coming online and is attempting to measure both the global signal and the 21 cm power spectrum (Fialkov et al., 2014). The All-Sky SignAl Short-Spacing INterferometer (ASSASSIN; McKinley et al., 2020) is attempting a new approach of observing the sky interferometrically using an array of closely spaced dipoles and searching for the signal in the observed visibilities. Finally, there is even an idea to build a lunar orbiter radio telescope which would be far away from anthropogenic radio frequency interference (RFI) and the Earth’s ionosphere known as the Dark Ages Polarimeter PathfinER (DAPPER; Burns, 2021).

The 21 cm Power Spectrum

The global 21 cm signal can be thought of being a monopole approximation to the true 21 cm signal since it is averaged over large angular scales. The true nature of the 21 cm signal is 3-dimensional and inhomogeneous as it traces the angular growth of HII regions and fluctuations in the Lyman–α and X-ray fields throughout redshift. A tomographic measurement of the full 3-dimensional 21 cm signal would probe the evolving astrophysics which are present throughout the EoR. The power spectrum is a strong statistical tool which is used in attempts to measure the full 21 cm signal because of its ability to extract information from low signal–to–noise maps. This is partially motivated by previous successes in using the power spectrum to study the CMB and galaxy surveys.

If the fractional perturbation to the brightness temperature of the 21 cm signal is written as

$$\delta_{21}(x) \equiv \frac{T_b(x) - \bar{T}_b}{\bar{T}_b},$$

(1.4)

where $T_b(x)$ is the brightness temperature along some direction $x$ and $\bar{T}_b$ is the average brightness temperature, then its power spectrum can be written as

$$\left\langle \tilde{\delta}_{21}(k_1) \tilde{\delta}_{21}(k_2) \right\rangle \equiv (2\pi)^3 \delta_D(k_1 - k_2) P_{21}(k_1),$$

(1.5)
where \( \tilde{\delta}_{21}(k_1) \) denotes the Fourier transform of Equation 1.4, \( \delta_D(k_1 - k_2) \) is the Dirac delta function, the brackets denote an ensemble average, and \( P_{21}(k_1) \) denotes the 21 cm power spectrum for Fourier mode \( k_1 \). It can be shown that the dimensionless power spectrum, defined as

\[
\Delta^2_{21}(k) = \frac{k^3 P_{21}(k)}{(2\pi)^3} ,
\tag{1.6}
\]

can be expressed as a sum of terms which describe perturbations in each relevant physical parameter which dominate during the EoR (Furlanetto, 2016). This can be written as

\[
\Delta^2_{21}(k) = T_0^2 \langle x_{HI} \rangle^2 \left[ \Delta^2_{\delta\delta}(k) + 2\Delta^2_{\delta x}(k) + \Delta^2_{xx}(k) \right] ,
\tag{1.7}
\]

where \( x_{HI} \) is the neutral fraction, \( \Delta^2_{\delta\delta}(k) \) and \( \Delta^2_{xx}(k) \) are the dimensionless power spectra of the density field and ionized fraction, respectively, and \( \Delta^2_{\delta x}(k) \) is the cross-power spectrum of the two quantities. The parameter \( T_0 \) is the mean brightness temperature in a fully neutral medium, given by \( T_0^2 \approx [28^2(1 + z)/10] \) mK\(^2\). The changing astrophysics in the Universe after Cosmic Dawn and through the EoR will change each of these terms differently and determine the shape of the 21 cm power spectrum. An example of the 21 cm power spectrum for various mean ionization fractions is shown in Figure 1.6.

Just like the global 21 cm signal, detection of the 21 cm power spectrum is obscured by Galactic foregrounds. Removing these foregrounds from the power spectrum has been extensively studied (Morales et al., 2006) with removal strategies focusing on using the spectral smoothness of foreground emission. While the galactic foregrounds have very smooth spectra, the cosmological signal varies rapidly as any line of sight passes through fluctuations in the density, temperature, or neutral fraction. Thus, a smooth function should be able to be fit to each line of sight and be removed so that the non-smooth cosmological signal remains in the residuals. However, this limits the signal in Fourier space by limiting small line of sight modes which get confused with variations in the foregrounds. There are more practical challenges that limit the observable Fourier modes. They are shown and summarized in Figure 1.7.
Experiments attempting to measure the 21 cm power spectrum include LEDA (Garsden et al., 2021), the Giant Metrewave Radio Telescope (GMRT; Paciga et al., 2013) in India, the Low Frequency Array (LOFAR; Harker et al., 2010) in Europe, the Murchison Widefield Array (MWA; Yoshiura et al., 2021) in Australia, and the Hydrogen Epoch of Reionization Array (HERA; DeBoer et al., 2017) which stands on the shoulders of the older Precision Array to Probe the Epoch of Reionization (PAPER; Parsons et al., 2010). The upcoming Square Kilometre Array (SKA; Braun et al., 2015) is being specifically designed to measure the 21 cm power spectrum.
Figure 1.7: Observable Fourier modes in the $k_{\parallel}-k_{\perp}$ plane. $k_{\parallel}$ represents modes along the line of sight and $k_{\perp}$ represents angular modes perpendicular to the line of sight. $k_{\perp}$ is limited on the low and high ends by the instrument’s finite field of view and longest baseline, respectively. $k_{\parallel}$ is limited on the low and high ends by confusion due to foreground variance and the instruments finite spectral resolution, respectively. At larger values of $k_{\perp}$, foregrounds leak out into higher $k_{\parallel}$ modes in a feature known as the foreground wedge (Furlanetto, 2016, Liu et al., 2014).
1.2 The Experiment to Detect the Global EoR Signature

The Experiment to Detect the GLobal EoR Signature (EDGES) collaboration has reported a potential detection of the global 21 cm Cosmic Dawn absorption signal (Bowman et al., 2018). The reported detection is summarized in Figure 1.8. The best fit absorption signal is a flattened Gaussian centered at 78.1 MHz with a full width at half maximum (FWHM) of 9.8 MHz and an amplitude of 0.53 K. This absorption feature is unexpected as it significantly differs in both shape and amplitude from theoretical predictions, something which is evident via comparing Figures 1.5 and 1.8 (d). The amplitude of 0.53 K makes it more than 2 times larger in amplitude than the most extreme models in Figure 1.5. The shape is also a flattened Gaussian which is very different than the normal Gaussian shape predicted by models. This flattened shape would imply that Lyman–α coupling saturates much earlier than X-ray heating can heat the IGM. This implies that the star formation efficiency is very high during Cosmic Dawn.

The abnormally large amplitude of the absorption feature implies, from Equation 1.1, that either the spin temperature, \( T_S \), of the HI is much lower than models predict or the temperature of the background radiation field, \( T_\gamma \), is much higher. Both options have been explored in the scientific literature. Since the spin temperature is coupled to the kinetic temperature of the surrounding gas, an explanation which involves more efficient gas cooling is required to explain a lower spin temperature. Attempts have been made to explain this excess cooling via an interaction between dark matter and baryons (Barkana, 2018, Muñoz and Loeb, 2018). If some fraction of dark matter particles were mini-charged, i.e., they carried a small electric charge (\( 10^{-6} \) times that of the electron), then the scattering cross-section between dark matter and baryons becomes similar to that of Rutherford scattering; i.e., growing as the relative velocity between dark matter and baryons becomes slower. This would cause a cooling effect on the gas, thus decreasing the spin temperature of the HI. However, Berlin et al. (2018) constrained the parameter space in which models like this can exist and found
that only $\sim 0.3\%-2\%$ of dark matter particles with mass in the range $\sim 10-80$ MeV could have a mini-charge without violating cosmological constraints placed by CMB measurements, light element abundances, supernovae measurements, and laboratory based experiments. On top of these constraints, non-standard depletion mechanisms must be invoked in order to avoid overabundance of these mini-charged dark matter particles.

The second option of increasing the temperature of the background radiation field at these redshifts does not need to invoke any new physics. This simply requires the presence of an extra radio component to the total radiation background (Feng and
Dowell and Taylor (2018) looked at data taken by the first station of the Long Wavelength Array (LWA1) to search for evidence of a low frequency radio background that is not accounted for. They use all-sky maps taken by the LWA1 Low Frequency Sky Survey (LFSS; Dowell et al., 2017) to search for a diffuse excess radio background which was suggested by previous results from ARCADE 2 (Fixsen et al., 2011). They found evidence for a strong diffuse radio background which is well modelled by a power law and has a temperature of 603 mK at the rest frame 21 cm frequency.

The unexpected nature of the EDGES signal and the profound impact it would have on models of the early Universe has opened it up to much scrutiny over the validity of the detection. Many attempts have been made to either reanalyze the EDGES data or provide a physical, non-cosmological, explanation for the observed signal. Hills et al. (2018) reanalyzed the EDGES data and have raised concerns over the modelling. They agree that the EDGES data contains non-smooth spectral structure which is not consistent with smooth foreground emission; however, they find that the best fit parameters for the physically motivated foreground polynomial used in Bowman et al. (2018) lead to non-physical properties for the foregrounds. They also found that constraining the priors for the foreground model parameters to realistic values leads to a poor fit. This could imply that the reported signal is not physical and only a remnant of the modelling. They were also able to recover a similar residual root-mean-square (RMS) after subtracting a model which does not contain any absorption signature.

Another reanalysis of the EDGES data was carried by Singh et al. (2018) who used maximally smooth functions (Rao et al., 2015, 2017a) to represent the foregrounds. They found that maximally smooth function foregrounds led to an even more extreme absorption profile with an amplitude of 921 mK. However, if they assume a small sinusoidal systematic in the spectrum has gone unnoticed in the EDGES data and remove it, then the resulting fit leads to an absorption profile with much more standard cosmological parameters. If the sinusoidal systematic has a wavelength of 12 MHz and an amplitude of 60 mK, then the best fit absorption profile has an amplitude
of 133 mK, a center frequency of 72.5 MHz, and a FWHM of 9 MHz.

Finally, Bradley et al. (2019) propose that a small ground screen artifact could create absorption features which creep into the observed data. They simulate the beam response of a simple patch antenna which is formed via a gap between the EDGES ground screen and the ground. A simple 2nd-order foreground polynomial combined with the 3 resonances which fall within the frequency range of interest lead to a fit which was just as good as the flattened Gaussian reported by EDGES. This could imply that the EDGES signal is not cosmological in nature and is an artifact of the antenna.

The unexpected nature of the EDGES absorption signal and its wide ranging implications for astrophysics in the early Universe mean that a validation study which aims to search for the signal using a different instrument is imperative. While there are other experiments similar to EDGES which can probe these low frequencies, an instrument which can offer complimentary information utilizing a different methodology would be most beneficial. The Long Wavelength Array (LWA; Taylor et al., 2012) in New Mexico, USA is such an instrument. Not only can it observe the low frequencies required to detect the EDGES signal, but it is an antenna array which offers new methodology which provides benefits over single element experiments.

1.3 The Long Wavelength Array

The LWA was originally designed to be an interferometer whose constituent elements were antenna arrays themselves. The individual antenna arrays are known as “stations” and consist of 256 antennas which consist of two sets of orthogonally polarized dipoles. The V-shaped blades of the antenna are designed to smooth the spectral response and improve sensitivity over a wider frequency range compared to a single resonance frequency for a classical dipole. They are also angled downward to increase sensitivity at low elevations above the horizon. The antenna design can be seen in Figure 1.9. There are currently two LWA stations in use: LWA1 which is the origi-
inal station that is colocated with the Karl J. Jansky Very Large Array in Socorro, New Mexico, USA and the LWA station located on the Sevilleta National Wildlife Refuge in New Mexico, USA (LWA–SV; Cranmer et al., 2017, Dowell and Taylor, 2020b, Price, 2017). The placement of the antennas within the array is pseudorandom to avoid beam artifacts such as grating lobes and to minimize sidelobes (Craig, 2009). The work presented here makes use of LWA–SV and so an overview of how that specific LWA station works will be given. LWA–SV shares much of its hardware design with LWA1, however there are some significant differences. The signal path of LWA–SV can be broken into the analog signal path and the digital signal path. The digital signal path will be described in more detail in Chapter 4 and so will be skipped in this overview.

The analog signal processor (ASP) consists of the front end electronics (FEE) and
the analog receiver (ARX). The FEE boards are located on the individual antennas. They are responsible for applying an initial gain of about 36 dB via an active balun. The amplified signal is then received in the electronics shelter by the ARX boards which apply another gain stage to the signal and a notch filtering operation which filters out FM radio signals. The ARX gain stage is tunable, so the amount of gain can be set by the user with the maximum being 60 dB. The current ARX board design allows for a switchable signal path which allows for different observing filter modes. The different filter modes allow varying sensitivity within the observable frequency range of 3–88 MHz. They are:

- Split Bandwidth
- Full Bandwidth
- Reduced Bandwidth
- Split Bandwidth @ 3 MHz
- Full Bandwidth @ 3 MHz

where the identifier “@ 3 MHz” is used to identify modes which allow observations down to the lower frequency cutoff of 3 MHz. The response of each filter mode is shown in Figure 1.10. After the ARX boards have filtered the RF signal, it is sent to the digital backend which is known as the Advanced Digital Processor (ADP). It is here where the data is digitized and the final data products such as beamformed and correlated data are produced. The ADP is covered in Chapter 4 which discusses modifications that were made in order to implement an achromatic beamforming framework at LWA–SV.

This dissertation details work which has been carried out to search for the global 21 cm Cosmic Dawn absorption profile using LWA–SV. The following chapters detail observational methodology and data analysis strategies which have been developed in order to minimize the residual RMS. It presents novel techniques which could potentially help in the search for this signal. This work has also led to a deeper
Figure 1.10: The frequency response of each ARX filter mode at LWA–SV (Dowell and Taylor, 2020b).

understanding of the challenges which LWA–SV faces in making such a sensitive and high dynamic range measurement.
Chapter 2

Using the Long Wavelength Array to Search for Cosmic Dawn

The contents of this chapter were originally published as part of DiLullo et al. 2020, Journal of Astronomical Instrumentation, 9, 2050008

The search for the spectral signature of hydrogen from the formation of the first stars, known as Cosmic Dawn or First Light, is an ongoing effort around the world. The signature should present itself as a decrease in the temperature of the 21-cm transition relative to that of the Cosmic Microwave Background and is believed to reside somewhere below 100 MHz. A potential detection was published by the Experiment to Detect the Global EoR Signal (EDGES) collaboration with a profile centered around 78 MHz of both unexpected depth and width (Bowman et al., 2018). If validated, this detection will have profound impacts on the current paradigm of structure formation within $\Lambda$CDM cosmology. We present an attempt to detect the spectral signature reported by the EDGES collaboration with the Long Wavelength Array station located on the Sevilleta National Wildlife Refuge in New Mexico, USA (LWA–SV). LWA–SV differs from other instruments in that it is a 256 element antenna array and offers beamforming capabilities that should help with calibration and detection. We report first limits from LWA–SV and look toward future plans to improve these limits.
2.1 Introduction

The spin-flip transition in neutral hydrogen provides the opportunity to study the ionization history of the Universe via its corresponding 21-cm emission. This transition occurs when an electron in the 1S ground state changes from having its spin parallel to that of the proton to antiparallel. The small energy difference between these configurations corresponds to a photon with a wavelength of 21.1 cm, or a frequency of 1420 MHz. This provides a useful probe for observing the early Universe where neutral hydrogen is abundant. See Furlanetto et al. (2006), Morales and Wyithe (2010), Pritchard and Loeb (2012) for an in-depth review of cosmology using the 21-cm line.

The formation of the first stars, known as Cosmic Dawn, marks the beginning of a phase transition for the Universe where it changes from being predominantly neutral beforehand to predominantly ionized after the Epoch of Reionization (EoR). Tracing the relative brightness of the 21-cm emission with respect to the Cosmic Microwave Background (CMB) as a function of redshift gives insight into which processes dominate the excitation state of the hydrogen during times when the first structures are forming in the Universe. The differential brightness temperature of the 21-cm signal relative to the CMB, $\delta T_B$, can be shown to follow:

$$\delta T_B \approx 27 \cdot x_{HI} \sqrt{\frac{1+z}{10} \left( \frac{T_S - T_{CMB}}{T_S} \right)} \text{mK},$$

where $T_S$ is the spin temperature of the hydrogen, $T_{CMB}$ is the temperature of the CMB, and $x_{HI}$ is the fraction of neutral hydrogen (Pritchard and Loeb, 2012). These quantities are all highly dependent on redshift, $z$. We have ignored terms dealing with spatial density fluctuations.

The driving factor that determines the detectability of the 21-cm transition, either as absorption or emission relative to the CMB at a given redshift, is the spin temperature of the hydrogen. The spin temperature is affected by three main processes: absorption and emission of 21-cm photons, collisions with other hydrogen atoms and with free electrons, and resonant scattering of Ly$\alpha$ photons that can induce a spin-flip,
known as the Wouthuysen-Field effect (Field, 1958, Wouthuysen, 1952).

It is expected that the ultraviolet radiation from the first stars couples to the surrounding neutral hydrogen via the Wouthuysen-Field effect. This decouples the HI spin temperature from the temperature of the CMB and instead couples it to the colder kinetic temperature of the gas. The kinetic temperature of the gas, \( T_K \), is colder than that of the CMB, \( T_\gamma \), since \( T_K = T_0(1 + z) \) for a non-relativistic gas, but \( T_\gamma = T_0(1 + z)^2 \) for photons, where \( T_0 \) is the respective temperature of each measured in the current epoch. This would drive the spin temperature to be less than \( T_{CMB} \) and so, from Equation 2.1, the 21-cm signal is expected to show in absorption against the CMB for these redshifts. Astrophysical properties of the early Universe correlated with the features of the 21-cm signal suggest that the absorption feature should be present at frequencies less than 100 MHz (Cohen et al., 2017). While most experiments have historically focused on detecting the later EoR emission signal, experiments around the world are beginning to look for this absorption feature as well (e.g. Price et al., 2018, Singh et al., 2018, Sokolowski et al., 2015). These experiments rely on using a small number of elements to observe the entire sky and return the sky-averaged spectrum.

After the first galaxies begin to form, their UV flux ionizes regions of the surrounding medium. This heating drives the 21-cm signal into emission and the signal becomes dominated by the filling fraction of HII regions (Furlanetto et al., 2004). Experiments attempting to measure the three-dimensional k-space power spectrum of the 21-cm line at these redshifts are searching for this emission signal (e.g. DeBoer et al., 2017, Harker et al., 2010, Parsons et al., 2010).

The Experiment to Detect the Global EoR Signal (EDGES) collaboration has published a potential detection of the absorption signal from Cosmic Dawn (Bowman et al., 2018). They have reported a flattened Gaussian profile centered at 78.1 MHz with a width of 18.7 MHz and an amplitude of 0.53 K. Both the shape and amplitude of this profile is unexpected from conventional models (Furlanetto, 2006) and, if validated, could imply interactions between baryons and dark matter (Barkana, 2018, Berlin et al., 2018, Muñoz and Loeb, 2018) or the presence of an unaccounted
component in the radio background at these redshifts (Dowell and Taylor, 2018, Feng and Holder, 2018, Mirocha and Furlanetto, 2018). There has also been much debate about the validity of this profile (Bradley et al., 2019, Hills et al., 2018, Sims and Pober, 2019, Singh and Subrahmanyan, 2019). This debate, coupled with the wide ranging implications of this potential detection, warrants a follow up with a different instrument in order to validate the reported profile parameters. The Long Wavelength Array (LWA) offers an opportunity for follow up as the profile center of 78.1 MHz falls within its observable band of 10-88 MHz. The full LWA currently consists of 3 stations: LWA1, which is colocated with the Karl J. Jansky Very Large Array in New Mexico, USA; LWA–SV, which is located on the Sevilleta National Wildlife Refuge in New Mexico, USA; and LWA–OVRO, which is located at the Owens Valley Radio Observatory in the Owens Valley, California, USA. The constituent stations can operate independently and offer a beamforming mode which is described in this paper.

The work detailed in this paper uses LWA–SV, which is the second station of the larger LWA. The beamforming mode of LWA–SV offers a new method for detecting the potential Cosmic Dawn signal. As mentioned above, other experiments that are searching for this signal observe the entire sky and generate sky-averaged spectra. This can create challenges as sources of contamination, namely extragalactic point sources and galactic synchrotron and free-free emission, can obscure the signal. The beamforming capability of LWA–SV allows for spatial selection on the sky in order to avoid sources of contamination, such as bright sources and the Galactic plane. This will not fully remove contamination, since the beam side lobes will pick up some unwanted signal, but it should improve overall performance. Beamforming also allows us to try a different method for calibrating the array, namely simultaneously observing a bright calibrator source with a second beam. This allows for in situ astronomical calibration of the entire system instead of relying on laboratory measurements like other experiments.
Table 2.1: Radiometer Equation Assumptions and Results for $T_{\text{rms}} = 50 \text{ mK}$.

<table>
<thead>
<tr>
<th>$\nu$</th>
<th>$\Delta \nu$</th>
<th>$A_e$</th>
<th>$T_{\text{sys}}$</th>
<th>$\Delta t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>55 MHz</td>
<td>9.57 kHz</td>
<td>1900 m$^2$</td>
<td>3840 K</td>
<td>73.4 s</td>
</tr>
<tr>
<td>65 MHz</td>
<td>9.57 kHz</td>
<td>1360 m$^2$</td>
<td>2500 K</td>
<td>32.0 s</td>
</tr>
<tr>
<td>75 MHz</td>
<td>9.57 kHz</td>
<td>1020 m$^2$</td>
<td>1740 K</td>
<td>15.5 s</td>
</tr>
</tbody>
</table>

2.2 The Long Wavelength Array

LWA–SV (Cranmer et al., 2017, Taylor et al., 2012) is an antenna array consisting of 256 dual-polarization antennas which are arranged in a pseudo-random layout and observe within the frequency range 3-88 MHz. It can support two simultaneous beam pointings with each beam containing two tunings with 9.8 MHz of bandwidth\(^1\). The array is roughly 110 m $\times$ 100 m in the N/S and E/W directions, respectively.

The system must be stable in time in order to integrate the data long enough to achieve significant signal to noise. If the system is not stable, fluctuations in the data can obscure the signal as the data is averaged over times longer than the fluctuation timescale. Single element experiments suffer from the need for integration times on the order of hours. However, LWA–SV benefits from its large effective area and should be able to reach a residual r.m.s. of 50 mK, one tenth of the signal amplitude reported by the EDGES collaboration, within an integration time of 10’s of seconds and so the electronics are not required to be stable over large fractions of a day or over multiple days. We estimate this using the radiometer equation within the Rayleigh-Jeans limit to relate integration time to an observed brightness temperature via:

$$\Delta t = \frac{1}{2\Delta \nu} \left[ \frac{c^2 T_{\text{sys}}}{\nu^2 T_{\text{rms}} A_e} \right]^2,$$

where $\Delta \nu$ is the channel bandwidth, $\nu$ is the tuning frequency, $A_e$ is the effective area of the array, and $T_{\text{sys}}$ and $T_{\text{rms}}$ are the system and r.m.s. brightness temperatures, respectively. Assumed values\(^2\) for a few frequencies and results for the integration time are summarized in Table 2.1.

\(^1\)LWA–SV has been upgraded since this work was completed. It now supports three simultaneous beams with each beam containing two tunings with 19.6 MHz of bandwidth.

\(^2\)http://lwa.phys.unm.edu/obsstatus/obsstatus006.html#toc14
The signal chain of LWA–SV is straightforward, but there are many places that can introduce instabilities. The entire chain can be broken into the front end electronics and the back end electronics. The FEEs are dual polarization receiver boards located on each of the dipoles. They provide initial amplification and low-pass filtering of the signal before it is passed to the back end electronics. The back end of LWA–SV consists of two components, the analog signal processor (ASP) and the advanced digital processor (ADP). The ASP electronics apply a second gain stage to the dual polarization analogue signal from each dipole and apply a bandpass to the signal. These signals are then passed along to ROACH-2 boards within ADP which digitize them and compute the Fast Fourier Transform (FFT) in order to output a complex spectrum. These data are then passed into the beamformer, after which the full beamformed data is recorded onto disk.

To ensure that the data remains unobscured by any fluctuations in the electronics chain of LWA–SV, we verified that an unstructured noise signal injected into the back end electronics integrated as expected. A Noisecom, Inc. model 3201K noise source was connected to one input of ASP and all other inputs were zero weighted. The choice of which ASP input channel the noise source is injected into should not affect the results. The ASP electronics are designed to have channel independent outputs and this has been verified through lab measurements. If the station is generally stable over a given length of time, the data output by the system for a noise-like input should integrate with an r.m.s. which goes like:

$$\sigma \propto t^{-1/2},$$  

(2.3)

since this is true for Gaussian data. It is important to note that this only tests the stability of the back end electronics and does not account for potential instabilities in the FEE.

Temperature variations in the electronics shelter, due to the air conditioning cycle, cause the amplifier responses in the electronics to vary creating power variations in

---

nc3200-coaxial-noise-sources

4. See LWA Memo #201
the data. We correct for these variations by fitting a linear relation between the observed median drift and the median temperature of the ASP electronics which is used to detrend the observed drift. The detrended drifts are then multiplied by the mean power of the original drift. The linear fit and drifts, before and after correction, are shown in Figures 2.1 and 2.2, respectively.

Figure 2.1: Median power across frequency vs ASP electronics temperature. The linear fit is used to detrend air conditioning cycle effects.

After correcting for power variations induced by the air conditioning cycle, we iteratively compute the r.m.s. noise of the data at different integration times to test overall system performance. In order to test how the data as a whole integrates, we use a bootstrapping method to randomly select 75% of frequency channels at each iteration. We then compute the r.m.s. in each of these channels and report the mean r.m.s. noise for that integration time. The results are shown in Figure 2.3 for integration times up to 24 minutes. We fit a line using the data for integration times of \( t \leq 120 \) s and report the best fit slopes. It is apparent that while we do not integrate down with the ideal slope of \(-0.5\) (see Equation 2.3), we have good performance with slopes that are \( \sim 0.05 \) away from ideal. It is unclear why we begin to plateau after integration times larger than a few minutes. The larger variations in the curves at
Figure 2.2: Uncorrected and corrected median drifts. The uncorrected drifts have large temporal structure in them as a result of the air conditioning cycle within the electronics shelter of LWA–SV. This causes variations in the response of the ASP electronics which induce power variations. The corrected drifts have been detrended and are centered about the mean of the uncorrected drifts.

The largest integration times are likely statistical in origin. The uncertainty in the r.m.s. at these integrations is dominated by statistical uncertainty arising from small sampling size. A 4 hour observation will only have 10 data points with which to compute the r.m.s. for an integration time of 24 minutes. We conclude that the system is stable at least on the order of a few minutes. This should be enough to detect the Cosmic Dawn signal (see Table 2.1).

2.3 Observations

The LWA1 Low Frequency Sky Survey (LFSS; Dowell et al., 2017) covered the entire sky visible to the LWA and was used to identify the coldest region on the sky. The reasoning for this was to minimize foreground contamination. The coldest region on the sky should offer the best chance to disentangle the Cosmic Dawn signal from
galactic foregrounds. The cold region that was identified and observed has J2000 coordinates of RA of 9$^h$ 38$^m$ 40.56$^s$ and DEC of +30$^\circ$ 49$'$ 1.4$''$, hereafter referred to as the “Science Field”. In order to calibrate the observations, we simultaneously point a second beam at Virgo A, located at J2000 coordinates of a RA of 12$^h$ 30$^m$ 49.42$^s$ and a DEC of +12$^\circ$ 23$'$ 28$''$. This allows for simultaneous in situ calibration using a source which is $\sim$ 43.7$^\circ$ away on the sky. This helps calibrate out time dependent systematics. The angular separation between Virgo A and the Science Field is large which means that this calibration will not accurately account for ionospheric effects since the two beams will each suffer unique perturbations as the signal paths through the ionosphere are different. This can be addressed with an ionospheric model, but that will be considered in the future. However, Virgo A is the closest object which is also significantly bright enough to use for calibration.

Observations were taken on 2019, October 11$^{th}$ for 3 hours beginning at 15:58:00 UTC. This time range captures both Virgo A and the Science Field at high elevation, with their elevations being similar during the midpoint of the observation. This
minimizes any elevation dependent effects during the middle of the observation. Both beams contain two tunings centered at 67 and 75 MHz, each with a bandwidth of 9.8 MHz. Accounting for rolloff at the edges of the band, which limits us to the inner 80%, this yields almost continuous coverage between 63 and 79 MHz. There is a small gap in coverage with a width of about 0.5 MHz around 71 MHz.

We chose LWA–SV’s spectrometer mode\(^5\), which channelizes the data and averages individual integrations to output time averaged spectra, for data acquisition. Setting the number of frequency channels to 1024 and the number of integrations per spectrum to 768 provided us with spectra with 9.56 kHz frequency resolution and 80 ms time resolution. We obtained spectra for the two linear polarizations XX and YY, which are elongated in the E/W and N/S directions, respectively. All data reduction was done using modules in the LWA Software Library\(^6\) (LSL; Dowell et al., 2012).

We flag RFI using a pseudo-spectral kurtosis flagging criterion (Nita and Gary, 2010a,b) that flags data with a spectral kurtosis outside of 3-σ from the mean. The spectral kurtosis is “pseudo” because the spectra that is output by the spectrometer mode is an average over our chosen number of 768 integrations per spectrum. We also use a smooth-bandpass model, created by smoothing the data along the frequency axis, to flag any frequency channels with an average power greater than 15-σ from the mean. This helps flag narrow-band RFI arising from digital television pilot tones that fall within the observed band that are not captured by the pseudo-spectral kurtosis criterion. One RFI source was identified as originating from a noisy electrical pole located roughly a mile from the station, but it is unclear what effect this has on the overall quality of the data.

The spectra output by LWA–SV are initially in units of power on an arbitrary scale. It is therefore necessary to calibrate the data in order to yield a temperature spectrum. The calibration using Virgo A works by simulating LWA–SV’s beam pattern at a given frequency on Virgo A at the midpoint of the observation. This beam pattern is then used in conjunction with the Global Sky Model (GSM; de Oliveira-

\(^5\)See LWA Memo #177.
\(^6\)http://fornax.phys.unm.edu/lwa/trac/wiki
Costa et al., 2008) to obtain a simulated measurement of the temperature of Virgo A at that frequency. This temperature is then divided by the arbitrary power reported by LWA–SV to generate a scaling coefficient in units of [K / Power] for that frequency. This is carried out on a per frequency channel basis, in order to get a set of scaling coefficients that can be applied to the spectrum of the Science Field. These scaling coefficients are expected to be independent of the spectral structure of Virgo A and should maintain the spectral shape of the Science Field.

The accuracy of our beam models for each pointing on the sky and the accuracy of the GSM temperatures for Virgo A can obscure the calibration and induce spectral structure in the calibrated spectra. The beam model used incorporates an electromagnetic simulation of the beam pattern of a single dipole and models of the responses of the electronics along the signal path. However, it is important to note that the accuracy of the beam model is extremely difficult to verify as measuring the pattern of the entire station is a major challenge. The accuracy of the dipole simulation is also difficult to measure in the field. Possible methods of measuring the station beam pattern are discussed in Section 2.6. The accuracy of the GSM is dependent on frequency and position on the sky. However, at these low frequencies and Galactic latitudes, the authors report the accuracy being $\sim 10\%$ or better with respect to the input maps. Dowell et al. (2017) also found a similar level of agreement between a 74 MHz map made with the first station of the LWA (LWA1) and the GSM in the region of the sky near Virgo A.

### 2.4 Results and Current Limits

The observed uncalibrated spectra of the Science Field and Virgo A are presented in Figure 2.4. These are the median spectra of 2 minutes of data selected from the middle of the observation described in Section 2.3. It is apparent that the low brightness Science Field suffers from more spectral structure than the brighter Virgo A field. The scale factors derived from our calibration scheme, described in Section 2.3, are presented in Figure 2.5. These factors should help calibrate structures, whether
Figure 2.4: Raw spectra of the Science Field (top) and Virgo A (bottom) for XX and YY polarizations. These spectra consist of 2 minutes of integrated data.

celestial, ionospheric, or system-based in nature, and should smooth the spectra of the Science Field.

It is important to note that the scale factors do have spectral structure that we believe should not be present. The scaling should be spectrally smooth in nature, but it is clear from Figure 2.5 that this is not the case as many bumps are present. The source of this structure is believed to be mostly related to the chromaticity of the beam of LWA–SV. Potential sources of spectral structure which are not captured in the current temperature calibration scheme, such as beam chromaticity, will be discussed in Section 2.5.

The temperature calibrated Science Field and Virgo A spectra are presented in Figure 2.6. The spectra are generally smooth in nature with some small structure present in the Science Field spectra. These unsmooth features will not be captured by a smooth polynomial fit and will remain present in the residuals.

We investigate the performance of two foreground models of differing complexity
in order to determine the extent of model dependence in the r.m.s. of the residuals. The first model we investigate is a simple power law of the form:

$$T(\nu) = k \left( \frac{\nu}{\nu_0} \right)^\alpha,$$

(2.4)

where $k$ is a proportionality constant, $\nu_0$ is a reference frequency, and $\alpha$ is the spectral index. The second model is a smooth 5-term polynomial of the form:

$$T(\nu) = \sum_{n=0}^{4} a_n \left( \frac{\nu}{\nu_0} \right)^{n-2.5},$$

(2.5)

where $\nu_0$ is a reference frequency. We chose $\nu_0$ to be the center frequency of the band for both models. These functions should capture the overall smooth shape of the spectrum, which is expected in this frequency regime. The best fit parameters for both models are presented in Table 2.2. The best fit models, with residuals, are presented in Figures 2.7 and 2.8.

We check the performance of the system by iteratively computing the average residual r.m.s. of time-averaged spectra for different integration times, ranging from the native 80 ms resolution of the observation to 2 minutes. The data corresponding to the central 20 minutes of the observation are selected and time-averaged accordingly.

Figure 2.5: Temperature scale factors derived from the observed power of Virgo A and simulated beam temperatures from the Global Sky Model.
at each iteration to produce a number of time-averaged spectra. The foreground model is then subtracted from these time-averaged spectra and a random sample of 75% of the residual spectra are selected. We then compute the residual r.m.s. across frequencies for these residuals and compute the average residual r.m.s. across the chosen sample. This allows us to estimate the behavior of the entire distribution of residuals as a function of integration time. We look for the r.m.s. to follow a similar trend to that seen in Figure 2.3. This would mean that we have correctly accounted for all forms of spectral contamination both from the system and from contaminating sources on the sky. The average residual r.m.s. vs. integration time for both models is shown in Figure 2.9. It is apparent that the data does not integrate like unstructured data, but rather we begin to plateau after $\sim 10$ seconds of integration.
Table 2.2: Foreground Model Best Fit Parameters

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter</th>
<th>XX Polarization</th>
<th>YY Polarization</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( a_0 )</td>
<td>( 7.49 \times 10^4 \pm 1.43 \times 10^3 )</td>
<td>( 2.29 \times 10^4 \pm 2.51 \times 10^4 )</td>
</tr>
<tr>
<td></td>
<td>( a_1 )</td>
<td>( -2.69 \times 10^5 \pm 5.78 \times 10^4 )</td>
<td>( -5.96 \times 10^4 \pm 1.01 \times 10^5 )</td>
</tr>
<tr>
<td></td>
<td>( a_2 )</td>
<td>( 3.66 \times 10^5 \pm 8.74 \times 10^4 )</td>
<td>( 5.21 \times 10^4 \pm 1.53 \times 10^5 )</td>
</tr>
<tr>
<td></td>
<td>( a_3 )</td>
<td>( -2.18 \times 10^5 \pm 5.86 \times 10^4 )</td>
<td>( -1.10 \times 10^4 \pm 1.02 \times 10^5 )</td>
</tr>
<tr>
<td></td>
<td>( a_4 )</td>
<td>( 4.81 \times 10^4 \pm 1.47 \times 10^4 )</td>
<td>( -2.72 \times 10^3 \pm 2.57 \times 10^4 )</td>
</tr>
</tbody>
</table>

N=5 Smooth Polynomial

| Power-Law | \( \alpha \) | \( -2.26 \pm 1.89 \times 10^{-3} \) | \( -2.14 \pm 3.83 \times 10^{-3} \) |
| Power-Law | \( k \)     | \( 3.26 \pm 5.36 \times 10^{-5} \) | \( 3.23 \pm 1.09 \times 10^{-4} \) |

2.5 Discussion

The r.m.s. of the residuals in Figures 2.7 and 2.8 are much higher than the required levels to verify the potential detection published by EDGES. In order to truly verify the claim, we require the residual r.m.s. to be at most on the order of 50 mK. This means that our reported r.m.s. level of \( \approx 10 \) K for the XX polarization is two orders of magnitude above what we require. This high r.m.s. is likely an effect of our inability to improve signal to noise (SNR) via integration times longer than a few seconds. It is useful to break the discussion into general categories that help narrow focus onto specific issues that are contributing to the limits reported above. The issues generally fall into two categories: Calibration and Modelling.

2.5.1 Calibration

Figure 2.9 highlights that something in the system is currently prohibiting us from integrating data for long periods of time and this severely limits our SNR. The results of our initial investigation that injected a noise source into the back end of LWA–SV, summarized in Figure 2.3, seem to suggest that the issue is not the stability of LWA–SV back end electronics. This isolates the problem to be somewhere in either the front end electronics or the chromaticity of the station beam.

The front end of LWA–SV can suffer from issues such as the impedance mismatch between the dipoles and the FEE, standing waves between the dipoles and their
ground screens, and mutual coupling between individual dipoles within the array. There have been impedance measurements made on LWA dipoles (Hicks et al., 2012), but these measurements are difficult to make for a single dipole and near impossible for the entire constructed array. Nevertheless, these models are used to take the impedance mismatch into account with the assumption that all dipoles behave the same. Standing waves between the dipoles and their ground screen would manifest as waves within the observed spectra. These should stand out in the residuals and are not seen in our spectra. Mutual coupling between elements has been shown to cause significant differences in response for each antenna, but should average out across the entire array for beamformed observations (Ellingson, 2011).

The frequency dependence of the dipole gain pattern⁷ and beam size contributes a major challenge for these types of experiments. This chromaticity leads to varying responses, both from the main lobe and the side lobes, with frequency which manifest as spectral features that obscure the signal. The full width at half maximum (FWHM)

---

⁷See LWA Memos #175 and #178.
of the main lobe response can vary by \( \approx 0.5^\circ \) across our band of \( \approx 63 - 79 \) MHz. The total beam pattern will also change as a function of pointing on the sky which can cause time-dependent variations in the response as the beam tracks a source over the duration of an observation. Work aiming to solve the chromaticity and directionality of the beam is presented in Section 2.6.

### 2.5.2 Modelling

Modelling plays a crucial role in the work presented above both in our temperature calibration and in our subtraction of galactic foregrounds. Our chosen temperature calibration scheme makes use of the GSM (de Oliveira-Costa et al., 2008) at 1° resolution; however, this is not the only sky model that we can use. The LFSS has been used to create the Low Frequency Sky Model (LFSM). This would seem to be the most logical sky model to use in order to maintain consistency of instrumentation, but the LFSM has an angular resolution of 5.1° which is currently larger than the size
of the beam main lobe across our entire band. If the beam size is sub-resolution of the chosen sky model, then the temperatures derived for Virgo A will not accurately match the observed power and the scaling will be in error. The LFSM should be investigated more in the future as we further develop custom achromatic beams.

Concerns have also been raised about the GSM using a cubic-spline interpolation between input surveys which does not take physical constraints into account (Rao et al., 2017b). The Global Model for the Radio Sky Spectrum (GMOSS; Rao et al., 2016) has been proposed as an alternative which takes physical constraints into account when interpolating between input surveys. However, GMOSS is also only available at 5° resolution and so we have not investigated its potential as our beam size is sub-resolution. This would lead to inaccurate temperature scaling, as stated above. Therefore, we leave the investigation into GMOSS’s performance for future work.

There is also debate surrounding how to properly model Galactic foregrounds that obscure the Cosmic Dawn signal. Bowman et al. (2018) use a physically motivated
model which accounts for galactic synchrotron emission and Earth’s ionosphere. A derivation of this physically motivated model can be found in Hills et al. (2018). However, Singh and Subrahmanyan (2019) express concerns that radiometers that observe the entire sky can produce spectra which will not be fit well by physically motivated polynomials. This is due to output spectra of the radiometer being some combination of spectra of various objects across the sky that have different spectral shapes. Thus, the output will not be well described by a single physically motivated power law. Instead, they propose the usage of maximally smooth functions (Rao et al., 2015, 2017b). These should remove the smooth structure in the sky-averaged spectrum and reveal the Cosmic Dawn signal in the residuals; however, it is unclear what degree polynomial is needed to properly model systematic effects without over-modelling and washing out the Cosmic Dawn signal. It is for this reason we have looked at the performance of two different models of varying complexity in this work.

The simple power law foreground model seems to model the observed spectrum almost as well as the 5th order smooth polynomial model. The similarities between the residuals presented in Figures 2.7 and 2.8 show that most of the residual structure is on relatively small spectral scales. However, both residual r.m.s. limits presented here are overly optimistic since we have not considered the effects of signal loss due to over-modelling (Cheng et al., 2018). This is especially true for the 5th order smooth polynomial model as it will be highly correlated with smaller spectral scales. Bernardi et al. (2015) investigated the effects of instrumental response over a large bandwidth and found the assumption that low order foreground models will appropriately account for instrumental effects is too optimistic. They claim the signal should be detectable and will not be over-modelled by a foreground polynomial of order \( \approx 5 \). We expect even more on-sky structure with a beamforming approach, so a higher order polynomial model may be necessary. However, they jointly model the foreground emission and the 21-cm signal, which will account for correlations between the model parameters, and consider a much larger bandwidth than observed here. A 21-cm signal with spectral structure on the order of the bandwidth observed here can lead to large correlations between the 5th order smooth polynomial foreground model and
the 21-cm signal, leading to possibly significant signal loss. These effects will be less significant with a more simple foreground model like the power law shown in Figure 2.8. We leave a more in depth analysis of signal loss to future work.

### 2.6 Custom Beamforming

The chromaticity and directionality of the beam are two major sources of spectral structure for our current results. The development of a framework for custom beamforming is necessary if we wish to push our limits further into the regime were the Cosmic Dawn signal is detectable. LWA–SV offers an advantage as an antenna array since elements can be weighted in such a way to shape the beam into a custom configuration. This should allow for not only custom shaped beams for some given pointing, but also achromatic beams at that pointing across the full band. We have begun work on simulating and forming custom beams with LWA–SV. The simplest first case is to make a beam with a circular main lobe with some defined FWHM for a given frequency at any pointing on the sky. This can be achieved by weighting the array in such a way that compensates for the projection of its elliptical shape onto the sky. We begin by defining a new coordinate system, \((x', y')\), cocentered on the array with the \(x'\)- and \(y'\)-axes corresponding to directions along and perpendicular to the line of sight for a given azimuth, respectively. We then compute the physical diameter that would correspond to a defined beam width via:

\[
D_1 = \frac{c}{\nu \theta},
\]

where \(c\) is the speed of light, \(\nu\) is frequency, and \(\theta\) is the user-defined beam width in radians. A second diameter is then computed which accounts for projection effects along the line of sight, given by:

\[
D_2 = \frac{D_1}{\sin(e)},
\]

where \(e\) is the elevation of the beam pointing center. The weighting is set by defining an ellipse that is aligned with the \(x'\)- and \(y'\)-axes and has a major axis of length \(D_2\)
Figure 2.10: 3-d weighting plot showing the antenna locations (crosses) and their weight values (dots). The weighting function is a Gaussian concentric with the array. The area of the array within the full width at fifth maximum (FWFM) is shaded blue. This area has a major axis of 53.1m in the N/S direction and a minor axis of 51.3m in the E/W direction. This was generated for a pointing at 180° azimuth and 75° elevation to obtain a 5° beam at 67 MHz.

and minor axis of length $D_1$. We apply a Gaussian taper from the center of the array with full widths at fifth maximum (FWFM) of $D_2$ and $D_1$ along the $x'$- and $y'$- axes, respectively. An example of this weighting scheme can be seen in Figure 2.10.

The weighting scheme shown in Figure 2.10 has been computed for a pointing center located at 180° azimuth and 75° elevation for a tuning at 67 MHz. The simulated beam pattern for this weighting can be seen at the bottom of Figure 2.11. It is apparent from the simulations that adjusting the weights to shape the main lobe is feasible, but the cost is generally stronger side lobes, albeit the side lobes are spatially smoother. This can be an issue if a bright source enters the side lobes as a low-level Cosmic Dawn signal can be washed out. However, the main lobe becomes more isolated as the nearby side lobes seem to be pushed outward. This might be beneficial if we can make sure the response from the high side lobes is minimized by observing during times when no bright sources are in these regions of the sky.
We checked the validity of these simulations by attempting to actually generate a custom beam with FWHM of 5° at 67 MHz pointed at 0° azimuth and 83.5° elevation. This pointing allows for Cygnus A to drift through the center of the beam at its transit. We pointed the custom beam and collected data for 3 hours centered around the transit of Cygnus A to generate a drift curve. We can use the shape of the drift curve to test the shape of the beam. The results of this are shown in Figure 2.12. We compare the shapes of observed drift curves generated with simultaneous unshaped and shaped beams to a simulated drift curve generated using a simulation of the shaped beam pattern convolved with the LFSM. All drift curves have been normalized with respect to the peak value since the observed data is in arbitrary power units and the simulated drift curve has units of Kelvin. The LFSM has a resolution of 5.1° and so we look for similarity between the shaped beam drift curve and the simulated one. We indeed find that the shaped beam drift curve matches the shape predicted by the LFSM and therefore conclude that the beam is 5° in size.

2.7 Future Work

The work done thus far has highlighted potential barriers that limit our ability to integrate to the SNR required for a significant verification of the claims made in Bowman et al. (2018). We believe the largest contribution to this comes from the beam shape of the array and its chromaticity. The simulations and observations presented in Section 2.6 have shown that the beam is shapable, but this does not solve the chromaticity problem. In theory, this should be straightforward as we can just compute how the dipoles must be weighted for all frequencies across the band for any given pointing. However, this will be more challenging in practice since the system will be required to compute and store all the weighting coefficients and apply them on a per channel basis. Implementing this into LWA–SV is in progress. Shaping the beam size will also allow us to investigate the performance of different sky models. If we choose a beam size of > 5°, we will no longer be sub-resolution for the LWA1 LFSS and GMOSS and can properly test their performance for this application.
Chapter 2. Using the Long Wavelength Array to Search for Cosmic Dawn

Figure 2.11: Simulated XX polarization beam patterns for an unshaped (top) and shaped (bottom) beam. Both patterns are for beam pointings at 180° azimuth and 75° elevation at 67 MHz on sinusoidal projection of the sky where $X = (a - a_{pointing})\cos(e)$ and $Y = e$ where $a$ is azimuth, $a_{pointing}$ is the beam pointing azimuth, and $e$ is elevation. The bottom plot is a circular beam with a FWHM of 5° and is a result of the weighting shown in Figure 2.10.

We also need to accurately measure the beam and dipole responses and chromaticity. This is essential if we want to verify that our beam is behaving in a way consistent with the simulations. Consistency is imperative since the simulations are necessary for our temperature calibration scheme. We have begun to experiment with measuring the dipole radiation pattern using a test drone (Chang et al., 2015, Jacobs et al., 2016). This work is underway, but this is still a new method and will need to be validated in other ways.

2.8 Summary

We have presented first limits from LWA–SV which employs a new method for attempting to detect the global 21-cm absorption profile associated with Cosmic Dawn.
We have detailed the stability of the system and highlighted the capabilities of a beamforming array as opposed to a single element radiometer which observes the entire sky. Simultaneous beams allow for \textit{in situ} calibration of the system which should allow for temperature calibration which is robust against systematics such as instrumental response and ionospheric effects.

We have selected a large cold region on the sky as our Science Field and use Virgo A as a temperature calibrator. We use LWA–SV’s spectrometer mode to obtain data products for the XX and YY polarizations. We are able to calibrate the observed spectrum of the Science Field and find residual r.m.s. limits of $\approx 10$ K for XX polarization and $\approx 20$ K for YY polarization after fitting a simple power law to model foreground structure.
Chapter 2. Using the Long Wavelength Array to Search for Cosmic Dawn

The largest challenge that we face is the chromaticity and directionality of the dipole response and the beam. We have presented theoretical simulations and early work showing that a direction independent beam can be made and have outlined what should be needed to move forward to an achromatic beam.

2.9 Acknowledgments

Construction of the LWA has been supported by the Office of Naval Research under Contract N00014-07-C-0147 and by the AFOSR. Support for operations and continuing development of the LWA is provided by the Air Force Research Laboratory and the National Science Foundation under grants AST-1835400 and AGS-1708855.
Chapter 3

Improvements to the Search for Cosmic Dawn Using the Long Wavelength Array

The contents of this chapter are part of DiLullo et al. 2021, submitted to the Journal of Astronomical Instrumentation.

We present recent improvements to the search for the global Cosmic Dawn signature using the Long Wavelength Array station located on the Sevilleta National Wildlife Refuge in New Mexico, USA (LWA–SV). These improvements are both in the methodology of the experiment and the hardware of the station. An improved observing strategy along with more sophisticated temperature calibration and foreground modelling schemes have led to improved residual RMS limits. A large improvement over previous work is the use of a novel achromatic beamforming technique which has been developed for LWA–SV. We present results from an observing campaign which contains 29 days of observations between March 10th, 2021 and April 10th, 2021. The reported residual RMS limits are approaching the required levels needed in order to validate the potential detection reported by the Experiment to Detect the Global EoR Signature (EDGES) collaboration.
3.1 Introduction

Detecting the formation of the first stars, known as Cosmic Dawn, and probing the subsequent Epoch of Reionization (EoR) remain some of the biggest goals of observational cosmology. The efforts to detect signals from these times in cosmic history focus either on direct imaging of the first luminous sources in the infrared, a major science goal of the upcoming James Webb Space Telescope (Gardner et al., 2006), or detecting the redshifted 21 cm signal from neutral hydrogen present during these periods (Furlanetto et al., 2006, Morales and Wyithe, 2010, Pritchard and Loeb, 2012). The redshifted 21 cm signal offers the only ground-based approach to studying Cosmic Dawn and the EoR since Earth’s atmosphere is opaque in the infrared.

The 21 cm signal provides an especially strong probe which traces the evolution of neutral hydrogen throughout the Epoch of Reionization. Measuring the signal across frequency would give insights into how astrophysical conditions evolve as the first luminous sources begin to dominate the Universe (Cohen et al., 2017). The 21 cm signal is generally broken into two observable signals: a global signal which is detectable over large angles on the sky and a spatially inhomogenous signal which corresponds to the growth of HII regions during the EoR. The global Cosmic Dawn signal is expected to manifest as a small absorption feature in the average spectrum of the sky below 100 MHz. Experiments to detect the 21 cm signal focus on detecting either the global signal by measuring the average spectrum of the sky and accurately modelling and removing foreground signals (McKinley et al., 2020, Price et al., 2018, Singh et al., 2018, Sokolowski et al., 2015) or the inhomogeneous signal by measuring the three dimensional k-space power spectrum whose structure arises from the formation of ionized regions during the Epoch of Reionization (DeBoer et al., 2017, Harker et al., 2010, Parsons et al., 2010). Novel interferometric techniques have also been developed for future use in detecting the global signal (McKinley et al., 2020).

The potential detection of a global absorption signal reported by the Experiment to Detect the Global EoR Signature (EDGES) collaboration (Bowman et al., 2018) has sparked much interest and debate due to its unexpected shape and amplitude
Chapter 3. Improvements to the Search for Cosmic Dawn Using the Long Wavelength Array

(Bradley et al., 2019, Hills et al., 2018, Sims and Pober, 2019, Singh and Subrahamanyan, 2019). If validated, the EDGES absorption signal would imply that current cosmological models fail to accurately predict the astrophysical processes which govern the 21 cm physics of neutral hydrogen throughout Cosmic Dawn and the Epoch of Reionization. This hints to either new physics in the form of interactions between dark matter and baryons (Barkana, 2018, Berlin et al., 2018, Muñoz and Loeb, 2018) or the presence of a previously unaccounted for radio background (Dowell and Taylor, 2018, Feng and Holder, 2018, Mirocha and Furlanetto, 2018).

The implications of the unexpected nature of the absorption signal reported by EDGES warrant independent validation using a different instrument. The Long Wavelength Array (LWA; Cranmer et al., 2017, Taylor et al., 2012) located in New Mexico, USA is one of the few radio telescopes in the world whose frequency coverage allows it to possibly validate the EDGES signal. Initial work using the LWA station located on the Sevilleta National Wildlife Refuge, NM, USA (LWA–SV) has been previously presented (DiLullo et al., 2020). In that work, a novel beamforming approach was presented which allows for increased sensitivity, improved foreground minimization, and in situ astronomical calibration. It was concluded that the dependence of beam shape with frequency is a major challenge and that a custom beamforming framework would have to be developed for LWA–SV which would allow for achromatic beamforming.

The work detailed here describes recent improvements to the methods described in DiLullo et al. (2020) which have lowered the residual RMS limits by an average factor of 3. These include hardware improvements to LWA–SV, changes to the observational strategy, improvements to the data analysis, and the successful adoption of achromatic beamforming. The paper is structured as follows: Section 3.2 details the recent changes to LWA–SV and to the observational strategy; Section 3.3 details the achromatic beamforming framework, discusses tests to compare the sensitivity of a standard LWA–SV beam to that of an achromatic beam, and presents a software package which can be used to simulate the beam pattern, both standard and custom, for a general antenna array; Section 4.4 will present results from a recent observing
campaign which employs all the recent improvements; and Section ?? discusses these results and looks towards future work.

### 3.2 Recent Improvements

The improvements to the methods described in DiLullo et al. (2020) fall into two categories: methodological improvements to the observing strategy and data analysis and improvements to the hardware and software of LWA–SV.

#### 3.2.1 Observing Strategy and Data Analysis

The observing strategy laid out in DiLullo et al. (2020) focuses on the use of simultaneous observations of a cold region on the sky (RA $9^h 38^m 40.56^s$, Dec $+30^\circ 49' 1.4''$; J2000), referred to as the Science Field, and Virgo A (RA $12^h 30^m 49.42^s$, Dec $+12^\circ 23' 28''$; J2000). Simultaneous observations of Virgo A allowed for astronomical calibration which would help account for any time dependent variation in the system. However, this did not account for the differences between the shapes of the science and the calibration beams as they were pointed towards different locations on the sky. In addition to the instantaneous differences in beam shapes, the different declinations of the Science Field and Virgo A meant that the science and calibration beams would trace different paths across the sky. These two effects caused inaccuracies in the calibration of the science beam.

A new Science Field center pointing (RA: $11^h 00^m 49.42^s$, Dec: $+12^\circ 23' 28''$; J2000) has been identified. The choice of coordinates still places the Science Field in a large cold region on the sky, but guarantees that the science and calibration beams trace out the same arc on the sky. The observations of the Science Field and Virgo A are also no longer done simultaneously, but rather during times when the targets align in terms of position on the sky. This ensures that the derived calibration better accounts for the changing beam shape throughout a single observation.
The data analysis methodology has been improved through the adoption of fully
time dependent temperature calibration and Bayesian foreground modelling. Time
dependent temperature calibration is achieved by modelling the beam pattern of
LWA–SV for every 5th pointing throughout the run at every 5th observed frequency
in the band and convolving the models with the Global Sky Model (GSM; de Oliveira-
Costa et al., 2008). This creates a set of model spectra at 4 minute intervals between
the relevant Local Sidereal Time (LST) ranges which can be linearly interpolated in
both time and frequency at the same temporal and spectral resolution of an observation to yield a model dynamic spectrum which is used to calibrate the raw data. It should be noted that the choice to use every 5th pointing and frequency is purely
arbitrary and was chosen to reduce required compute time. These models can be
generated at higher temporal and spectral resolution before interpolation, but this
becomes quite computationally expensive and is not expected to significantly change
the results since the sky changes smoothly over time and frequency and this smooth
nature is accurately captured through interpolation. An example of the model dy-
namic spectrum of Virgo A, interpolated at the time and frequency resolution of a
single observation, is shown in Figure 3.1.

Bayesian foreground modelling allows for the posterior distributions of the model
parameters to be explored to search for correlations between foreground model pa-
rameters and 21 cm model parameters. Correlations between the foreground and 21
cm models would be non-physical and would imply issues with our foreground mod-
els. We use the Python package emcee (Foreman-Mackey et al., 2013) which uses the
Affine Invariant Markov chain Monte Carlo (MCMC) Ensemble Sampler (Goodman
and Weare, 2010) to maximize the likelihood function, which essentially measures
the probability of a dataset given a certain set of model parameters. Bayes’ theorem
relates the posterior probability distribution, $P(\Theta|D, H)$, of a set of parameters $\Theta$
given the data $D$ and some model $H$ to the likelihood $L(D|\Theta, H)$ via

$$P(\Theta|D, H) = \frac{L(D|\Theta, H)\Pi(\Theta|H)}{Z(D|H)},$$

(3.1)

where $\Pi(\Theta|H)$ is the prior distribution that encodes information about the model
parameters and $Z(D|H)$ is the evidence, which is the integral of the likelihood func-
Figure 3.1: Model dynamic spectrum of Virgo A. These are generated by simulating the beam pattern at every 5\textsuperscript{th} pointing for every 5\textsuperscript{th} frequency and convolving these simulations with the Global Sky Model at 7.5 s resolution to yield temperature spectra. These spectra are then interpolated at the frequency and time resolution of an observation. The median spectrum and drift are shown at the top and right of each plot, respectively.

We follow Harker et al. (2012) and Bernardi et al. (2016) by assuming Gaussian measurement noise in a single frequency channel which allows us to write the likelihood \( L_i \) of observing the sky temperature \( T_{\text{sky}}(\nu_i) \) at frequency \( \nu_i \) as

\[
L_i(T_{\text{sky}}(\nu_i)|\Theta) = \frac{1}{\sqrt{2\pi\sigma^2(\nu_i)}} e^{-\frac{(T_{\text{sky}}(\nu_i)-T_m(\nu_i,\Theta))^2}{2\sigma^2(\nu_i)}},
\]

where \( T_m \) is the sky model and \( \sigma \) is the standard deviation of the instrumental noise is a given frequency channel given by

\[
\sigma(\nu_i) = \frac{T_{\text{sky}}(\nu_i)}{\sqrt{\Delta\nu \Delta t}},
\]

where \( \Delta\nu \) and \( \Delta t \) are the channel bandwidth and integration time, respectively.

Given Equation 3.2, the log-likelihood of the entire temperature spectrum can be
represented as

\[ \ln \mathcal{L}(T_{\text{sky}}|\Theta) = \sum_{i=1}^{N} \ln \mathcal{L}_i(T_{\text{sky}}(\nu_i)|\Theta), \quad (3.4) \]

where we have switched to the log-likelihood due to computational efficiency. The sky model is a superposition of the foreground contribution and the 21 cm signal, written as

\[ T_m(\nu_i) = T_f(\nu_i) + T_{21\text{cm}}(\nu_i), \quad (3.5) \]

but for the work presented in this paper we do not attempt to fit the 21 cm signal as our residual root mean square (RMS) is still too large to meaningfully fit a 21 cm signal (See Section 4.4). We use a slightly modified foreground polynomial than that used in Bernardi et al. (2016); however, it is still a \(N^{th}\) order log-polynomial given by

\[ T_f(\nu_i) = \sum_{n=0}^{N} a_n \left[ \log_{10} \left( \frac{\nu_i}{\nu_0} \right) \right]^n, \quad (3.6) \]

where \(\nu_0\) is a reference frequency that is set to the center of the band. We choose an uninformative prior for each model parameter \(a_n\) which only guarantees that each parameter is real.

### 3.2.2 Upgrades to LWA–SV

There have been a few upgrades to the hardware of LWA–SV which should help in the search for Cosmic Dawn. The observational bandwidth of LWA–SV has been improved to 20 MHz per tuning per beam. This new bandwidth allows for uninterrupted frequency coverage in the range 50 – 85 MHz with center tuning choices of 60 and 75 MHz. Accounting for bandpass rolloff, we select the inner 80% of the band which yields a usable frequency range of 52 – 83 MHz for the output spectra. This frequency range should allow for the capture of the lower edge of the EDGES signal through the center frequency of 78 MHz. This is a large improvement over the spectra presented in DiLullo et al. (2020), which had incomplete frequency coverage in the smaller range of \(\approx 63 – 79\) MHz.
Along with the upgraded bandwidth capabilities of LWA–SV, a new weather station was also deployed at the station which records the outside temperature. We have previously seen that temperature variations inside the electronics shelter can affect performance of the system, namely the analog receiver (ARX) boards, and so we also expect outside temperature variations to affect the performance of the front end electronics on each dipole. Outside temperature data from the new weather station can be combined with the temperature data from the ARX boards to allow for a 2-dimensional fit which better captures gain fluctuations than the previous simple 1-dimensional fit with ARX temperature. However, we have found this effect to be so small in magnitude that it is difficult to accurately remove this trend from the data which is dominated by the changing sky. The temperature change, both outside and in the electronics shelter, over a single 1.5 hour observation is less than 1% compared to the change in sky brightness due to the rising Galactic plane. Perhaps as we push our residuals below 1 K, this correction will become necessary, but for now this approach does not seem to improve data quality. Another improvement to note is the installation of second temperature sensor for the HVAC controller within the electronics shelter. This has led to much more stable temperatures within the shelter. We now see typical temperature variations on the order of $\approx 2^\circ$ F, as opposed to older temperature variations which were on the order of $\approx 10^\circ$ F.

### 3.3 Achromatic Beamforming

A major challenge that all experiments attempting to detect Cosmic Dawn and the EoR face is the frequency dependence of the antenna response. This “chromaticity” of the receiving element produces spectral structure which can obscure any potential cosmological signal. The array nature of LWA–SV helps as beamforming allows for better rejection of foreground contributions by focusing the main lobe in a region of the sky away from the Galactic plane; however, this introduces a second chromaticity factor which can obscure the signal. The beam pattern of an antenna array is not only a function of the antenna gain pattern, which is chromatic, but also has intrinsic
chromaticity. This second chromaticity factor can be combated through a custom
beamforming framework which modifies the required complex beamforming coeffi-
cients in order to tailor the array response to be more constant across the observed
frequency band. However, this still does not completely account for the chromaticity
of the gain pattern of the individual antennas within the array so that will need to
be combated some other way.

In DiLullo et al. (2020), a framework for custom beamforming at a single frequency
was presented. The beam pattern of LWA–SV can be shaped by modifying the
amplitude of the complex gains of each dipole in the array depending on the frequency
of interest and the desired pointing center on the sky. Model custom beam patterns
were presented and driftscan observations of Cygnus A (RA 19h 59m 428.36s, Dec
+40° 44′ 2.1″; J2000) at transit confirmed that the beam full width at half maximum
(FWHM) could be predictably shaped using the framework. However, the beam
was only properly tuned for a single frequency. It was concluded that in order to
achieve near-achromatic beamforming, modifications would need to be made to the
beamformer pipelines of LWA–SV.

3.3.1 Implementation on LWA–SV

The beamformer pipelines are part of the digital backend of LWA–SV, known as the
Advanced Digital Processor (ADP; Cranmer et al., 2017, Dowell and Taylor, 2020b,
Price, 2017). The ADP beamformer pipelines have been modified in order to allow for
pre-computed custom complex gains to be read in and overwrite the complex gains
which would normally be computed by the Monitor and Control System at LWA–SV.
Therefore, for an observation which consists of some number of pointings which track
a source on the sky, the necessary custom complex gains can be computed for every
observed frequency at each pointing. This yields a set of custom complex gains which
keep the beam pattern mostly constant in both frequency and time as the beam
tracks the source across the sky. However, the beam pattern can still not be said
to be completely constant as the response pattern of the individual dipoles has both
The major trade-off for the ability to shape the beam pattern of an antenna array is a loss in sensitivity. The sensitivity of an array is given by its System Equivalent Flux Density (SEFD)

\[
SEFD = \frac{k_B T_{sys}}{A_e} 10^{26} \text{ Jy},
\]

where \(k_B\) is the Boltzmann constant, \(T_{sys}\) is the system temperature, and \(A_e\) is the effective area of the array. The SEFD is an estimate of the flux density of a source which would double the system temperature, which means that a lower SEFD corresponds to a higher sensitivity. \(A_e\) is difficult to accurately measure for an antenna array, but the simple estimate of

\[
A_e \approx N_{dip} \times A_{ant},
\]

where \(N_{dip}\) is the number of antennas in the array and \(A_{ant}\) is the effective area of the antenna, is an idealized estimate. However, effects such as mutual coupling of individual antennas can drastically change the effective area of the array as a whole (Ellingson, 2007).

Compared to standard beamforming at LWA–SV where every antenna is equally weighted, adjusting the amplitudes of the complex gains to shape the beam effectively reduces the effective area of the entire array. Therefore, it is expected for the SEFD of the array to increase as more dipoles are down weighted to maintain beam shape with frequency. The SEFD of the array can be estimated by observing a bright source, such as Cygnus A, and comparing the observed power when the source is in and out of the beam. The SEFD can be estimated with

\[
SEFD = \frac{S_\nu}{P_{on}} \frac{P_{off}}{P_{off}} - 1,
\]

where \(P_{on}\) is the observed power when the beam is pointed on source, \(P_{off}\) is the observed power when the beam is pointed off source, and \(S_\nu\) is the flux density of the source at the measured frequency, \(\nu\). This is typically done by pointing a beam at the transit position of a bright source and letting the source drift into and out of the
Chapter 3. Improvements to the Search for Cosmic Dawn Using the Long Wavelength Array

beam. Driftscans can also be used to measure the FWHM of the beam main lobe. The major limitation of driftscan observations is that only a single dimension of the beam can be probed throughout the observation, so measuring the beam shape in both dimensions on the sky requires either multiple beams or multiple days of observation where the beam pointing center is below, on, and above the source transit position so the source can probe different parts of the beam main lobe.

A method to avoid the need for multiple days of observations is to continuously repoint the beam in a basket weave pattern. This approach is based on methods used by higher frequency arrays such as the Very Large Array. The idea is to make two cuts along the source in different directions on the sky. This allows for the 2-dimensional beam shape to be estimated at various points in a single observation. However, at low frequencies ionospheric scintillation on $\sim 10$ s timescales can make things difficult since the observed power at any position is a combination of the beam pattern and the local ionospheric conditions. Thus, we modify the above approach to interleave the off-source pointings with on-source pointings. These on-source pointings serve as an ionospheric references which bracket each off-source pointing and can be interpolated between to determine the amount of ionospheric scintillation.

The modified basket weave approach is utilized on a weekly basis at LWA–SV in order to monitor any changes in the station’s SEFD. These weekly observations utilize standard beamforming on Cygnus A, so Cygnus A was a natural choice to use when testing the sensitivity of the station with an achromatic beam. A circular beam with FWHM of $4^\circ$ was chosen for a basket weave observation of Cygnus A; however, estimating the SEFD of LWA–SV turned out to be challenging as the larger beam size meant that more power from the Galactic plane was being picked up by the beam. This generally smoothed out the contribution of Cygnus A to the total observed power and resulted in unrealistic SEFD estimates. To avoid this, we used the beam-dipole observing mode which LWA–SV offers to observe Cygnus A interferometrically using the modified basket weave to determine the phase centers. Beam-dipole mode correlates the beamformed output with the signal from a single antenna in the array. We chose to correlate the beamformed output with the outrigger antenna which is
Chapter 3. Improvements to the Search for Cosmic Dawn Using the Long Wavelength Array

Table 3.1: Beam-Dipole Mode Basket Weave Results

<table>
<thead>
<tr>
<th>Frequency</th>
<th>Standard Beamforming</th>
<th>Achromatic Beamforming</th>
</tr>
</thead>
<tbody>
<tr>
<td>60 MHz</td>
<td>SEFD: 4.0 kJy</td>
<td>SEFD: 4.8 kJy</td>
</tr>
<tr>
<td></td>
<td>FWHM: 3°22'10.6″</td>
<td>FWHM: 5°06'09.3″</td>
</tr>
<tr>
<td>75 MHz</td>
<td>SEFD: 4.3 kJy</td>
<td>SEFD: 6.7 kJy</td>
</tr>
<tr>
<td></td>
<td>FWHM: 2°42'16.1″</td>
<td>FWHM: 4°44'22.7″</td>
</tr>
</tbody>
</table>

located ~ 300 m to the West of the station center. This two-element interferometer produces a fringe pattern on the sky which is a function of the observed frequency, the baseline geometry, and the geometric mean of the beam patterns of the station beam and the outrigger dipole. These interferometric basket weave observations resolve out the diffuse structure from the Galactic plane and help isolate the contribution from Cygnus A. Results from observations using standard and achromatic beamforming can be seen in Figure 3.2. The final SEFD estimates are averages between the two values derived from the RA and Dec slices, respectively. The SEFD and FWHM estimates for beam-dipole mode observations using standard and achromatic beamforming are presented in Table 3.1. It is clear that the observations using achromatic beams suffer a loss in sensitivity; however, the loss in sensitivity is less than a factor of 2.

3.3.2 Beam Simulator: A Python Package to Simulate Array Beam Patterns

The achromatic beamforming work presented above might be of interest to a broader community that is interested in either simply simulating the beam pattern of a general antenna array or developing similar custom frameworks for an antenna array. We have developed a Python package, aptly named Beam Simulator, to help simulate the beam pattern of a general antenna array. It can represent the elements of an antenna array down from a single cable up to the entire array and is also compatible with output files from the Numerical Electromagentics Code (NEC) which is used to numerically model the gain pattern of antennas. This allows for the simulated
Source: CygA at LWA-SV  
Az: 0.0 degrees; El: 83.6 degrees

Figure 3.2: Interferometric basket weave results from observations of Cygnus A. The cuts along declination and right ascension are shown in each figure in the left and right columns, respectively. The lower and upper tunings are shown in each figure in the top and bottom rows, respectively. The achromatic beam shows less variation in FWHM across frequency in both directions on the sky.
beam pattern to encapsulate the gain pattern of the individual array elements and not just account for the geometry of the array. 

*Beam Simulator* also allows for custom beams to be simulated using the same framework that has been developed for LWA–SV. There are convenience functions built in to easily represent a LWA station, so this should be of particular interest to any institutions that plan on hosting a future LWA Swarm station (*Taylor et al., 2019*). However, it should also be of interest to the more general community interested in representing the beam pattern of any phased array feed. The package can be found and downloaded on GitHub at [https://github.com/cdilullo/beam_simulator](https://github.com/cdilullo/beam_simulator).

### 3.4 Recent Observing Campaign and Results

An observing campaign was carried out between March 10th, 2021 and April 10th, 2021 in order to build a sufficiently large data set in time. A single observation consists of an achromatic beam with a main lobe FWHM of 4° which tracks the Science Field (see Section 3.2.1) between the local sidereal time (LST) range 10.28 – 11.78 h and a second achromatic beam with a similar main lobe FWHM which tracks Virgo A between the LST range 11.78 – 13.28 h. We use the spectrometer observing mode offered by LWA–SV which yields dynamic spectra with 40 ms time resolution and 1024 frequency channels each having 19.1 kHz of bandwidth. The individual observations from each day are flagged for RFI in both time and frequency by applying a median smoothing window to the median drift and spectrum, respectively, to create smooth models of each and then flagging times or frequencies which deviate from these models. This allows for the RFI flagging to be controlled through 4 simple parameters: the smoothing window sizes and the threshold values in both time and frequency. It was found that the parameters which led to good RFI detection were a time smoothing window of 30 s, a frequency smoothing window of 250 kHz, and time and frequency thresholds of 3-σ and 5-σ, respectively, where σ is the standard deviation of the data after the smoothed model has been divided out. These parameters led to relatively aggressive flagging for a single observation, ≈ 20% of data for each day, but this was
intentional to try and capture low level RFI.

A common source of RFI in the relevant frequency band is ionospheric reflections from terrestrial digital television (DTV) signals which manifest as flares with a characteristic 6 MHz bandwidth. DTV channels 2–6 all lie within our observed band and can be seen quite frequently. Due to the short timescales of the bursts and the relatively small bandwidth that is affected, these DTV flares are not captured in the above RFI flagging method which is better suited at capturing long duration narrow-band RFI or short duration wide-band RFI. To better capture these events, we search along the pilot tone frequencies for each channel for times where the pilot tone is $2\sigma$ above the median value and flag the entire 6 MHz sub-band corresponding to the DTV channel at these times.

After each day was flagged for RFI, we performed a moving median window smoothing operation on the data in time in order to smooth out variations due to ionospheric scintillation. Ionospheric scintillation causes random variations in the brightness of our targets with time, which obscure the temperature calibration. We chose a window size of 5 minutes in order to properly smooth out scintillation. The individually flagged and smoothed datasets were then temperature calibrated by multiplying a model temperature dynamic spectrum of Virgo A (see Section 3.2.1) with the ratio between the raw Science Field and Virgo A data. The calibrated Science Field data were then integrated in time into 30 second bins before being combined into a single dataset which contains the entire observing campaign. An extra RFI flagging step was applied which flags any channel that has been flagged for more than 75% of the entire campaign. The fully flagged, calibrated, and integrated dataset for the entire observing campaign is shown in Figure 3.3.

We generated a measured spectrum by employing a bootstrapping algorithm to generate samples from the full calibrated dataset. This bootstrapping algorithm ensured that the output spectrum was the most “typical” and was not influenced by random features, such as unflagged RFI, which were present in only some of the integrated spectra. In order to generate the samples, we first fit a smooth bandpass model to each integrated spectrum which was then divided out to create a copy of
Chapter 3. Improvements to the Search for Cosmic Dawn Using the Long Wavelength Array

Figure 3.3: Calibrated datasets for the entire observing campaign at 30 second integration time. The apparent ripples correspond to each of the individual 1.5 hour observations. The full dataset consists of 29 days of data taken between March 10th, 2021 and April 10th, 2021.

The data which highlighted the unsmooth features in each integrated spectrum. We then computed the standard deviation across frequency for each of these residual spectra and computed the first and third quartiles, $Q_1$ and $Q_3$, for the distribution of standard deviations. We flagged outlier spectra whose standard deviation was above the standard outlier threshold given by

$$\sigma \geq 1.5 \times (Q_3 - Q_1).$$

A subset of the non-outlier integrated spectra were then chosen at random and an average spectrum was computed. This was repeated 10,000 times in order to generate a set of bootstrap sampled spectra for each tuning and polarization. The average of these 10,000 spectra was taken to be the “typical” spectrum. The two average spectra for a given polarization were then combined to yield a single average spectrum which spans the entire bandwidth of 52–83 MHz. A foreground model was then fit using the Bayesian modelling framework described in Section 3.2.1. We chose to start 100 walkers, which are independent stochastic processes which explore the parameter
space, around the zero vector and run a chain which was 25,000 steps long. It was observed that the model struggled to properly fit the data across the entire frequency band, with frequencies below \( \approx 57.5 \) MHz having the poorest fit, so the lower edge of the band below 57.5 MHz was discarded before the fitting. The average bootstrapped spectrum after trimming, along with the fit foreground model and residuals, for each polarization is shown in Figure 3.4. Figure 3.5 shows a typical plot of the posterior distributions of the model parameters for the XX polarization. A similar plot is generated for the YY polarization.

We also investigate the performance of maximally smooth functions (MSFs; Rao et al., 2015, 2017a) as foreground models. The Python package `maxsmooth` (Bevins et al., 2021) is used to fit the same 5th order log-polynomial foreground model (See Equation 3.6). We find that the MSF model performs well, but not as well as the MCMC approach. The observed spectrum along with the MSF model and residuals are shown in Figure 3.6. The residual RMS across frequency for each polarization using both modelling methods is summarized in Table 3.2.

![Average Observed Spectra and Residuals](image)

Figure 3.4: Results of the MCMC fit.
Chapter 3. Improvements to the Search for Cosmic Dawn Using the Long Wavelength Array

Figure 3.5: Posterior distributions for foreground model parameters for XX polarization. The 1-dimensional posteriors show the median value along with the 16th and 84th percentile bounds. This plot was made with the Python package *corner* (Foreman-Mackey, 2016).

<table>
<thead>
<tr>
<th>Foreground Model Method</th>
<th>XX</th>
<th>YY</th>
</tr>
</thead>
<tbody>
<tr>
<td>MCMC</td>
<td>2.47 K</td>
<td>3.81 K</td>
</tr>
<tr>
<td>MSF</td>
<td>3.29 K</td>
<td>5.26 K</td>
</tr>
</tbody>
</table>

### 3.5 Discussion and Conclusion

The residual RMS limits reported in Table 3.2 are encouraging as they are beginning to approach the sub-Kelvin level that is required for a proper validation of the reported EDGES signal. The desired goal is to attain a residual RMS of \( \approx 50 \text{mK} \) so that the EDGES signal could be validated at the 10-\( \sigma \) level. However, pushing the residual RMS to the sub-Kelvin level is not expected to be simple.

We believe that the largest systematic that is limiting our residual RMS is the discrepancy between the beam model used in the temperature calibration and the true beam pattern of LWA–SV on the sky. The model beam pattern predicts a
very smooth response which is unlikely to be realistic. We tested these differences by taking an observation from a single day of the observing campaign and fitting a smooth model to the data. Many spectral and temporal structures remained after the smooth model was removed from the data which implies that there are underlying low level non-smooth spectral structures present in the data. These structures are most likely related to the sidelobe pattern of the array which is difficult to measure and quantify.

There are also discrepancies between the model used for the gain pattern of a single LWA dipole and the true gain pattern. The model is directly used in the simulation of the beamformed pattern of the station as a whole and so is also present in the temperature calibration. The model is based on Numerical Electromagnetics Code (NEC) simulations of the LWA dipole, but directly measuring it is extremely difficult at these low frequencies. There is an ongoing collaboration with the External Calibrator for Hydrogen Observatories (ECHO) team at Arizona State University whose goal is to measure the gain pattern of the LWA dipole through use of a drone...
which transmits a tone of known power at various altitudes and azimuths (Chang et al., 2015, Jacobs et al., 2016). Preliminary results show discrepancies between the measured pattern and the model pattern, but there is still much more to understand before a definitive conclusion can be made.

Another concern is the presence of low level RFI which is difficult to detect and flag via the median spectrum and drift method described in Section 4.4. Any unflagged RFI will significantly increase the residual RMS, especially as we approach the sub-Kelvin level. The bootstrapping algorithm described in Section 4.4 should help reduce the contribution of random low level RFI which is only present on some days of the observing campaign, but more systematic low level RFI which is present throughout the entire observing campaign at the same frequencies will still be problematic. As we decrease our residual RMS limits, we will need to better understand the RFI environment of LWA–SV, whether it is external or internal to the station itself.

The work presented above shows encouraging improvements to the previously reported limits of LWA–SV. Future work will focus on better quantifying the LWA dipole gain pattern and understanding discrepancies between models of the station beam and the true beam pattern. This is crucial if the residual RMS is to be reduced to the level where the 21 cm signal can be jointly fit with the foreground model. A more robust RFI detection algorithm may have to be developed in order to capture any low level RFI that is present throughout the entire observing campaign.

3.6 Acknowledgments

Construction of the LWA has been supported by the Office of Naval Research under Contract N00014-07-C-0147 and by the AFOSR. Support for operations and continuing development of the LWA is provided by the Air Force Research Laboratory and the National Science Foundation under grants AST-1835400 and AGS-1708855.
Chapter 4

Achromatic Beamforming at LWA–SV

4.1 Introduction

The beam pattern of an antenna array such as the LWA has complicated structure which changes with both pointing direction and frequency. These dependencies imply that the sensitivity of the beam can change across the observed frequency band and throughout the duration of an observation which tracks a source on the sky. This effect is most likely negligible for observations of bright sources at high elevations, however it can create spectral structure large enough to obscure the detection of very weak signals, such as the global Cosmic Dawn absorption signature. The chromaticity of the elements being used in single antenna experiments to detect Cosmic Dawn has been shown to be a major challenge which obscures the signal. We expect the chromaticity of the beam pattern of LWA–SV to similarly be a major factor which can obscure the detection of Cosmic Dawn.

Luckily, this effect can be minimized by manipulating the complex gains of the antennas in the array. The complex gains required for a given pointing on the sky can be manipulated such that the resulting beam pattern will have the same main
lobe size as a function of frequency. The sidelobes of the beam pattern are more
difficult to control and so the entire beam pattern is not completely constant across
frequency, however the resulting beam pattern is more consistent across frequency
than the standard beam pattern.

In this chapter, a custom beamforming framework which has been developed at
LWA–SV in order to make achromatic beams is presented. The theoretical frame-
work is first discussed and then changes made to the digital backend of LWA–SV,
the Advanced Digital Processor (ADP). Finally, a description and tutorial are given
for Beam Simulator, a Python package which has been developed to help simulate
antenna array beam patterns.

4.2 Theoretical Framework

The standard beamforming equation for a general antenna array of $N$ elements is
given by

$$Y(\nu, \theta, \phi) = R(\nu, \theta, \phi) \times \sum_{j=1}^{N} w_j e^{-2\pi i \vec{r}_j \cdot \vec{k}},$$ (4.1)

where $Y$ is the beam pattern of the array, $R$ is the antenna gain pattern, $w_j$ is the
weight assigned to the $j^{th}$ antenna, $\vec{r}_j$ is the position vector of the $j^{th}$ antenna, $\vec{k}$ is
the wave vector pointing in the direction of the source, $\nu$ is frequency, and $\theta$ and $\phi$ are
angular coordinates representing a position on the sky. This equation has implicitly
invoked pattern multiplication which is only valid if the gain patterns of the individual
antennas can be assumed to be identical. This allows for the individual antenna gain
patterns to be represented as a single term which is pulled out of the summation.

The only term in Equation 4.1 that is not set by either the design of the antenna
or the geometry of the array itself is the individual antenna weights, $\{w_j : j \in [1, N]\}$. Standard beamforming at LWA–SV sets these all to unity so that every antenna has
equal weighting, however these can be changed to modify the shape of the beam. A
framework for this shaping has been developed and is described in detail in DiLullo
et al. (2020), but a broad overview is presented here.

Given a pointing center on the sky, a new coordinate system \((x', y')\) can be defined on the array such that the \(x'\)-axis lies perpendicular to the azimuthal line of sight and the \(y'\)-axis lies parallel to the azimuthal line of sight. The required dimensions for a desired beam full width at half maximum (FWHM) are then given by

\[
D_{x'} = \frac{c}{\nu \theta},
\]

and

\[
D_{y'} = \frac{D_{x'}}{\sin(e)},
\]

where \(\theta\) is the desired FWHM of the beam main lobe, \(\nu\) is the frequency, \(e\) is the elevation of the center pointing, and \(c\) is the speed of light. The weights of the antennas can then be tapered in such a way so that the effective size of the array matches those of Equations 4.2 and 4.3.

### 4.2.1 Choice of Weighting Function

Equation 4.1 is essentially the Discrete Fourier Transform of the antenna weighting function, so educated choices can be made about the weighting function in order to achieve desired results in the beam pattern that is generated. A straightforward choice for the weighting function is a Gaussian taper whose maximum is located at the center of the array. This has a functional form given by

\[
w(x', y') = e^{-\left(\frac{x'^2}{2\sigma_{x'}^2} + \frac{y'^2}{2\sigma_{y'}^2}\right)},
\]

where \(\sigma_{x'}\) and \(\sigma_{y'}\) are the standard deviation of the Gaussian in each dimension. These are set such that the full width at fifth maximum (FWFM) of the Gaussian is equal to \(D_{x'}\) and \(D_{y'}\); i.e.

\[
\sigma_{x'(y')} = \frac{D_{x'(y')}}{2\sqrt{2 \ln 5}}.
\]

However, more exotic weighting functions exist such as the Blackman-Harris windowing function which is designed to minimize the sidelobe response of its Fourier
Transform. The 1-dimensional Blackman-Harris window function for the $x'$- and equivalently $y'$- axis is given by

$$f(x') = a_0 - a_1 \cos \left( \frac{2\pi(x' + D_{x'}/2)}{D_{x'}} \right) + a_2 \cos \left( \frac{4\pi(x' + D_{x'}/2)}{D_{x'}} \right) - a_3 \cos \left( \frac{6\pi(x' + D_{x'}/2)}{D_{x'}} \right),$$

(4.6)

where $a_0 = 0.35875$, $a_1 = 0.48829$, $a_2 = 0.14128$, and $a_3 = 0.01168$. The full 2-dimensional function is a product of two 1-dimensional functions for both the $x'$ and $y'$ dimensions; i.e.

$$w(x', y') = f(x') \cdot f(y').$$

(4.7)

Examples of the different weightings can be seen in Figure 4.1.

The differences between simulated beam patterns using Gaussian weighting versus Blackman-Harris weighting were investigated. The simulated beam patterns corresponding to the weightings shown in Figure 4.1 are shown in Figure 4.2. It is apparent that the Blackman-Harris beam has a more diffuse main lobe response with generally larger and stronger sidelobes. This may suggest that the weighting function, when applied at discrete points corresponding to the antenna locations within the array, does not yield the desired effect of a true Blackman-Harris weighting. Further investigation is needed to see if a possibly scaled version of the Blackman-Harris function can lead to a truly minimized sidelobe response with a main lobe that better matches the desired FWHM. The Gaussian weighting generally results in a more isolated main lobe with generally weaker sidelobes relative to those generated by the Blackman-Harris weighting. It is therefore in our best interest to choose the Gaussian function especially as it tends to preserve more antennas in the array due to its slower taper.
Chapter 4. Achromatic Beamforming at LWA–SV

(a) Gaussian Weighting

(b) Blackman-Harris Weighting

Figure 4.1: Gaussian and Blackman-Harris weighting functions for a center pointing at 180° azimuth, 68° elevation at 60 MHz. The blue shaded region represents the effective geometric area of the array.

Figure 4.2: Simulated beam patterns using both Gaussian and Blackman-Harris weighting functions. It is clear that the Blackman-Harris weighting leads to stronger sidelobes and a less isolated and more diffuse main lobe.
Chapter 4. Achromatic Beamforming at LWA–SV

4.3 The Advanced Digital Processor

The Advanced Digital Processor (ADP) is the digital portion of the backend electronics which handles the quantization of the analog signal and outputs data from either the DRX or TBN pipelines. See Cranmer et al. (2017), Price (2017), and Dowell and Taylor (2020a) for a comprehensive overview of the ADP, but a brief overview of how the ADP works is as follows:

- Output from the Analog Receiver (ARX) is quantized by ROACH2\(^1\) boards which also convert the data to the frequency domain by performing a Fast Fourier Transform (FFT). Afterwards, the data are sent to the seven computing nodes of the ADP system for processing.

- The DRX pipelines run on six of the computing servers and are responsible for multiple data products: beamformed data, the wideband correlator data, and transient buffer – Frequency Domain (TBF) data. Two DRX pipelines run simultaneously on each server to handle each independent tuning. A single tuning is defined by a center frequency and consists of 19.6 MHz of data.

- The Transient Buffer – Narrow (TBN) pipeline runs on the same six servers as the DRX pipelines and outputs voltage time-series data from all of the antennas with 100 kHz bandwidth.

- The T-Engine runs on the ADP head computing node and aggregates the intermediate beamformed product output by the beamforming units, computes an inverse FFT to return the data to the time domain, performs sub-channel tuning, filters the data to set the bandpass, and then packetizes the data.

Since the beamformer units (BFUs) exist within the DRX pipelines, creating an achromatic beamforming framework for LWA–SV only required modifications to the existing DRX pipelines.

---

\(^1\)https://casper.ssl.berkeley.edu/wiki/ROACH2
The DRX pipelines run on six of the ADP computing nodes where each node receives one sixth of the total bandwidth from the ROACH2 boards for each tuning. This means that on any given node there are two DRX pipelines running simultaneously for each of the two tunings. The BFUs in each DRX pipeline perform a matrix multiplication between a matrix containing the complex beamforming coefficients with the data from the ROACH2 boards. The station Monitor and Control System (MCS) precomputes the delays and gains required for each pointing of an observation and stores them in separate files which are read by the DRX pipelines and used to compute the complex beamforming coefficients. Every time the center pointing is updated throughout an observation, a “BAM” command is issued by MCS which updates which gains and delays the DRX pipelines use to compute new complex beamforming coefficients. It is here where the DRX pipelines are modified to receive custom complex coefficients which have been precomputed for each observed frequency at each pointing in an observation for a desired beam FWHM as described in Section 4.2.

The custom complex coefficients required for an achromatic beam pointing of a given FWHM are first computed for all observed frequencies in the band on the ADP head computing node. The coefficients are then split across six files. The six computing nodes which run the DRX pipelines read the corresponding file which contains the custom coefficients for the frequencies which fall within the subband that that computing node processes from the ROACH2 boards. These files each contain an array of the frequency dependent custom coefficients required for each pointing defined within a given observation. This allows the computing nodes running the DRX pipelines to simply iterate through the array as the beam pointing is updated. The main issue with directing the DRX pipelines to a set of custom coefficients is the possibility of another observation being ruined if a different observer wishes to use the standard beamforming coefficients generated by MCS. This can happen if two beamformed observations are simultaneously scheduled. A dynamic method of setting a custom achromatic beam which can flexibly update the coefficients with each new beam pointing was therefore required.
In order to create a dynamic achromatic beamforming framework which only reads the precomputed custom coefficients when desired, the DRX pipelines were further modified to search for a “trigger” which would tell the pipelines that an observation was to use achromatic beamforming and to read in the custom coefficients. This trigger was tied into individual BAM commands, the commands which tell the system to update the beam pointing, so that the pipelines could easily keep track of the current pointing in an observation and the proper coefficients from the files could be used. The trigger used is a specific gain setting for a single step in STEPPED mode observation which is also sufficiently uncommon so that achromatic beamforming is not accidentally triggered. The STEPPED observing mode of LWA–SV allows an observer to define specific “beam steps” which contain the center pointing information, beam dwell time, and custom delays and real valued gains which are used to compute the complex coefficients for beamforming. To avoid confusion, it is worth noting that the custom delays and gains which can be defined in a STEPPED observation are applied to the whole frequency band so this mode does not offer a straightforward path to achromatic beamforming without modifications to the DRX pipelines like we have done. An individual beam step corresponds to a single BAM command which passes the custom delays and gains defined for that step to the DRX pipelines. The custom gains for a beam step are stored as 256 $2 \times 2$ matrices, one matrix for each of the antennas in the array, and one delay vector of length 512 which stores the delays for each polarization input (2 polarizations per antenna). Triggering the achromatic beamforming pipelines therefore requires the first custom gains matrix, corresponding to the gains which are used to compute the $[XX \ XY \ YX \ YY]$ data products, to be the matrix 

\[
\begin{bmatrix}
8191 & 16383 \\
32767 & 65535
\end{bmatrix}
\]

These values are 16-bit big endian integers which are read by MCS and represent an extremely non-standard gain setup that is extremely unlikely to be used for any reason except to trigger achromatic beamforming. The second custom gain matrix for a given step stores the pointing number so the achromatic beamforming pipelines can use this to properly index the file with the precomputed custom complex coefficients.

Relating the reading of the precomputed custom complex beam coefficients re-
quired for achromatic beamforming to an individual BAM command as described above makes achromatic beamforming a dynamic framework which exists all the time in a dormant manner at LWA–SV. It coexists with standard beamforming and allows for one beam to be achromatic while the other remains standard. While there is no high level observing mode which can be easily accessed by any user, all the necessary lower level tools exist and only require an observer who is interested in utilizing achromatic beamforming to precompute the necessary gains and set up a custom STEPPED mode observation whose beam steps have the “secret code” to trigger the modified pipelines.

4.4 Results and Discussion

The changes to ADP described in Section 4.3 have been successfully implemented and are currently running on LWA–SV. We have carried out tests which use interferometric basket weave observations to quantify the beam shape and loss of sensitivity due to the decreased effective area of the array. See Section 3.3.1 for a detailed description of the interferometric basket weave observations. The results are presented in Figure 3.2.

It is difficult to quantify how precisely the main lobe of the beam can be shaped. The fringe pattern cast by the station in beam-dipole mode is a function of the geometry of the baseline and the geometric mean of the station beam pattern and the gain pattern of a single LWA dipole. The method described above only accounts for the geometry of the array and not the intrinsic chromaticity of the dipole gain pattern. Thus, the beam is not completely achromatic across the observed frequency band, but it is apparent that its shape is more uniform.

The LWA dipole has intrinsic chromaticity which is not accounted for in the above beamforming framework. This chromaticity is still present in the custom array beam pattern, so a truly achromatic beam will need to account for this. Accounting for this is not straightforward as this chromaticity is related to the physical antenna design.
The best way to account for this would be to have a very accurate model of the gain pattern of the dipole with frequency which can be used in the computation of the custom gains.

The discrepancies between the model and true LWA dipole gain pattern, mentioned previously in Section 3.5, are also an issue. If accounting for the dipole chromaticity requires an accurate model of the dipole gain pattern, accurate measurements must be made. At these low frequencies, anechoic chambers are too small to be of use and more clever solutions are required. The Murchison Widefield Array (MWA) has made use of the 137–138 MHz ORBCOMM satellite constellation to measure the power pattern of an MWA antenna tile by comparing the observed power of satellites between an antenna tile and a reference antenna (Neben et al., 2015). Another option, mentioned in Section 3.5, is the use of drones which carry a transmitter of known power which can be used to measure the gain pattern of a single dipole.

4.5 Beam Simulator: Description and Tutorial

*Beam Simulator*, mentioned in Section 3.3.2, is a Python package which has been developed to help simulate the beam pattern of an antenna array. It can be downloaded at [https://github.com/cdilullo/beam_simulator](https://github.com/cdilullo/beam_simulator). It can hierarchically represent an entire array using only a few objects and simulate the beam pattern of said array with only a few simple function calls. The hierarchical structure of the objects and a tutorial on how to simulate a beam pattern is presented below.

4.5.1 Representing An Antenna Array

The classes which represent the various levels of an antenna array are located in the *station.py* module. The relevant objects which represent an array are *Station*, *Antenna*, and *Cable*. They are structured hierarchically so that a *Station* object is made up of a list of *Antenna* objects which each have a corresponding *Cable* object.
This structure allows for a simple representation of potentially complex arrays.

A Station object simply stores the name of the array and a list which contains the individual Antenna objects. There is also a convenience function called plot_antennas() which plots the locations of the constituent Antenna objects. The Antenna objects store more specific information about the antennas themselves. They store the physical location of each antenna in an \((x, y, z)\)-coordinate system, which corresponds to the antenna’s position Eastward, Northward, and Elevation above some defined reference point, respectively; and other information such as antenna stand number, polarization (1 or 2), and status (1 for “good” or 0 for “bad”). An Antenna object also stores the Cable object which is associated with that antenna. The Cable object represents a coaxial cable and therefore requires a number of parameters which define a coaxial cable such as its length, velocity factor, dielectric constant, and the sizes and conductivities of its inner and outer conductors. A Cable object has two functions attenuation() and delay() which compute the attenuation and delay factors induced by a coaxial cable on a signal. These functions make use of the coaxial cable model presented in (Liu and Ellinson, 2012).

These three objects are all that is needed to fully represent an antenna array from a beamforming standpoint. However, fully populating a Station object with all the antennas and cables can be a bit tedious, so some convenience functions have been added which make fully populating a Station object simple and fast. The first two are the LMR200 and LMR400 subclasses. These create Cable objects for two very common cable types used in radio arrays, LMR200 and LMR400 cables, respectively. The other two convenience functions, named load_station() and load_LWA(), focus on building a Station object. The function load_station() reads in a text file template which describes the station and its constituent antennas and cables. An example of the template is included in Beam Simulator. The final convenience function, load_LWA(), reads in a LWA Station Static MIB Initialization File\(^2\) (SSMIF) and generates the Station object which represents a LWA station.

\(^2\)See LWA Engineering Memo MCS0031.
4.5.2 Simulating the Array Beam Pattern

The first step to simulating the beam pattern of an array is to represent the gain pattern of the individual antennas. This step is optional in the sense that Beam Simulator does not require an antenna gain pattern to simulate a beam pattern for an array; however, it is necessary if the simulated pattern is to be accurate. The nec.py module contains functions and classes which help read an antenna gain pattern model output by NEC simulations. The AntennaPattern class reads a NEC output file\(^3\) which contains an EXCITATION section for a single frequency and stores it. There is an associated convenience function named plot_pattern() which plots the stored gain pattern on a simple rectangular projection. The function fit_spherical_harmonics() takes in multiple AntennaPattern objects at the same frequency and represents the antenna gain pattern at that frequency as spherical harmonic decomposition. A number of these spherical harmonic decompositions can then be combined using combine_harmonic_fits() which represents the antenna gain pattern at each point on the sky as a spherical harmonic decomposition as a function of frequency and outputs a NumPy NPZ file which can be read in by the beamformer.py module when simulating the array beam pattern.

The beamformer.py module handles the actually beamforming computation. There is a suite of functions which take in a Station object and generate a weighting vector, compute the geometric delays for a given pointing on the sky at a given frequency, and compute the beamformed signal. The two functions which generate weighting vectors are generate_uniform_weights() and generate_gaussian_weights() which generate either uniform (all antennas weights are equally weighted) weights or Gaussian weights to create a custom beam (See Section 4.2). However, any array whose size is equal to number of antennas in the array can be used as a weighting vector. Once a Station object is created and a weighting vector has been created for it, a beam pattern can be easily simulated using the beamform() function which takes in the NPZ file output by

\(^3\)See the NEC User’s Guide for a description of the output file. Available for download from the National Technical Reports Library.
combine_harmonic_fits(), described above, as an optional argument to compute the antenna gain pattern at the chosen frequency.

### 4.5.3 Tutorial

A short tutorial code snippet is shown in Figure 4.3. This tutorial shows how to generate the standard beam pattern for LWA–SV at $1^\circ$ resolution. The SSMIF for LWA–SV is first read in on line 3 which creates the `Station` object named `lwasv`. The LWA dipole gain pattern is then modelled by reading four NEC output files (.out) for each modelled frequency in the `freqs` list. These files correspond to the parallel and transverse components of each polarization for the dipole. Once the models for each frequency have been decomposed into spherical harmonics on line 16 using `nec.fit_spherical_harmonics()`, they are combined using the call to `nec.combine_harmonic_fits()` on line 19. This outputs a file named `beam_coefficients.npz` which is used in the call to `beamformer.beamform()` on line 29. Once a `Station` object and the antenna gain pattern have been generated, the weighting vector is easily made with a call to `beamformer.generate_uniform_weights()` on line 26 and the simulated beam pattern for the entire array, including the antenna gain pattern via pattern multiplication, is generated via `beamformer.beamform()` on line 29. The final result is the `pwr` array which has shape (2, 361, 91) where the first axis corresponds to polarization and the second and third axes are azimuth and elevation at $1^\circ$ resolution, respectively. It is plotted on a sinusoidal projection in Figure 4.4.
Chapter 4. Achromatic Beamforming at LWA–SV

Figure 4.3: Beamforming example code snippet.

```python
from beam_simulator import station, nec, beamformer

lwsv = station.load_LWA("lwsv-smnif.txt")

# We need to read in all NEC4 output files for a series of frequencies.
freqs = [10, 20, 30, 40, 50, 60, 70, 80, 90]  # MHz

# Generate the spherical harmonic fits for each frequency separately.
# Each iteration will generate a file named "SphericalHarmonicsFit_{freq}.1fMHz.npz".
for freq in freqs:
    p1 = 'lwsv_lcp_{freq}.out'
    t1 = 'lwsv_lcp_{freq}.out'
    p2 = 'lwsv_rbp_{freq}.out'
    t2 = 'lwsv_rbp_{freq}.out'

    nec.fit_spherical_harmonics(freq, p1, p2, t1, t2, verbose=False)

    # Now the .npz files containing the spherical harmonic fits can be combined.
    nec.combine_harmonic_fits('SphericalHarmonicsFit_10 MHz.npz', 'SphericalHarmonicsFit_20 MHz.npz',
                                 'SphericalHarmonicsFit_30 MHz.npz', 'SphericalHarmonicsFit_40 MHz.npz',
                                 'SphericalHarmonicsFit_50 MHz.npz', 'SphericalHarmonicsFit_60 MHz.npz',
                                 'SphericalHarmonicsFit_70 MHz.npz', 'SphericalHarmonicsFit_80 MHz.npz',
                                 'SphericalHarmonicsFit_90 MHz.npz')

# Generate the weighting vector for the station.
    w = beamformer.generate_uniform_weights(lwsv)  # All antennas have the same weighting of 1.0

# Simulate the beam for a pointing center of az = 180 deg, el = 75 deg at 74 MHz with 1 degree resolution.
pwr = beamformer.beamform(lwsv, w, freq=74 MHz, az=180.0, el=75.0, resolution=1.0, smtgainFiles='beam_coefficients.npz', dB=True)
```

Figure 4.4: Standard beam pattern of LWA–SV for a center pointing at 180° azimuth, 75° elevation at 74 MHz. These simulations are at 1° resolution.
Chapter 5

Measurements of the LWA Dipole Impedance Mismatch

5.1 Introduction

The impedance mismatch between any antenna and the associated electronics is a crucial parameter to measure and understand. In a receiving system such as the LWA, this mismatch will lead to a reflection of the incoming power which decreases the overall sensitivity of the system. Accurate measurements of the impedance mismatch of the LWA antenna are required to improve the Low Frequency Sky Model (LFSM; Dowell et al., 2017). The LFSM is unique in that it is a model of the sky which is derived from many observations at low frequencies whereas other sky models such as the Global Sky Model (GSM; de Oliveira-Costa et al., 2008) rely on interpolation anchored on a few measurements of the low frequency sky. This, combined with the fact that it was conducted using a station of the LWA with similar hardware, makes it especially alluring as a choice of sky model to use for the temperature calibration described in Chapter 3 and so improving its accuracy is crucial.

Impedance measurements are made using a vector network analyzer (VNA) which injects a signal into the system and then measures the amount of reflected power.
Estimates of the impedance mismatch have been made in the past for the LWA antenna (Hicks et al., 2012), but these were done using older versions of the LWA antenna and were based on simulations, not *in situ* measurements on an antenna in the completed array. In order to improve upon this previous work, new measurements were taken at LWA–SV on May 18th, 2021 using custom test fixtures (See Figure 5.1) designed at the Naval Research Laboratory (NRL) in Washington, D.C. and an Anritsu VNA Master model MS2034B\(^1\).

![Figure 5.1: Testing fixtures designed by NRL.](image)

### 5.2 Measurement Setup

The custom testing fixtures designed by NRL mimic the shape of the usual LWA front end electronics, but consist only of one signal path, which connects to the E/W

\(^{1}\)anritsu.com/en-us/test-measurement/products/ms2034b
Chapter 5. Measurements of the LWA Dipole Impedance Mismatch

polarization of the antenna, and the transformer. This design allows for the fixture to be mounted to the antenna and connected to the VNA with only a single cable with SMA and N-type connections. Calibration of the VNA to the “50Ω” and “Short” fixtures was done in the lab at the University of New Mexico. The LMR200 cable which was built measured 30 ft in length. This length was chosen so the VNA could be significantly far from the antenna after running the cable under the ground screen and up the antenna mast, but not too far away as to introduce appreciable attenuation of the signal. The typical distance between the VNA and the antenna can be seen in Figure 5.2.

The VNA was used to measure the complex impedance for four antennas in the array. The antennas chosen were located in different sections of the array in order to determine if local effects such as mutual coupling changed the value of the impedance mismatch. Stands 137, 162, and 256 were chosen to capture differences between the array center, Northern edge, and outrigger, respectively. The locations of these stands can be seen in Figure 5.3.

5.3 Results and Conclusion

The measured real and imaginary components of the complex impedances and the computed impedance mismatch efficiency (IME) are shown in Figures 5.4 and 5.5. We find no significant variation in the measured impedances across stands with drastically different locations in the array. This implies that the impedance of the antennas is not strongly affected by local mutual coupling effects. The IME is computed via:

\[
IME = 1 - \left| \frac{(Z - Z_0)}{(Z + Z_0)} \right|^2
\]

where \(Z\) is the impedance of the antenna and \(Z_0\) is the impedance of the hybrid on the front end electronics, taken to be 100 Ω. There are discrepancies between the newly measured values and those reported in Hicks et al. (2012). The biggest difference is that the newly measured values are monotonically increasing with frequency as opposed to the simulated values which show a local maximum around \(\nu \approx 35\) MHz.
This is not surprising as the values reported in Hicks et al. (2012) are based on simulations. The newly measured values have been used to update models of the LWA antenna response with frequency; however, while the results appear to be better, there still appears to be room for improvement. This is most likely due to the custom test fixtures not accurately capturing the effects of the entire signal chain on a real front end electronics board. Currently, they only account for the presence of the hybrid coupler, but do not account for other components such as the amplifiers.
Chapter 5. Measurements of the LWA Dipole Impedance Mismatch

Figure 5.3: Locations of the antenna stands with measured stands marked.

Figure 5.4: Real and imaginary components of the complex impedance measured at three different stands. The red curve represents the model stored in the LWA Software Library derived from Hicks et al. (2012).
Figure 5.5: Impedance mismatch efficiency (See Equation 5.1) of the three measured stands. The newly measured values differ from those reported in Hicks et al. (2012). The ratio between the measured values and the model values is shown in the bottom panel.
Chapter 6

Conclusions and Future Work

6.1 Summary and Conclusions

The work presented in this dissertation has focused on efforts to detect the global 21 cm Cosmic Dawn absorption signal in the radio background of the sky using LWA–SV, the Long Wavelength Array station located on the Sevilleta National Wildlife Refuge in New Mexico, USA. The purpose of these efforts is to attempt an independent validation of the potential detection reported by the EDGES collaboration. If the EDGES signal is validated, then its unexpected amplitude and shape would imply deficits in models of the astrophysics governed by the formation of the first stars.

Chapter 2 details the first attempt at determining the current noise limits of LWA–SV. It presents work whose goal was to determine an observing strategy which would allow for in situ astronomical temperature calibration. It employs two simultaneous beams which point at an empty field on the sky away from the Galactic plane, known as the Science Field, and at the radio galaxy Virgo A which serves as the temperature calibrator. After a smooth foreground model was removed from the calibrated data the residual RMS was \( \sim 10 \) K, which is over two orders of magnitude above the level required to properly validate the EDGES result. It was concluded that chromaticity in the beam response of LWA–SV was a major limiting factor and that future work aimed
Chapter 6. Conclusions and Future Work

at developing an achromatic beamforming framework for LWA–SV was required.

Chapter 3 details recent improvements to the work presented in Chapter 2. These improvements include updated hardware at LWA–SV which has doubled the observational bandwidth, a new observing strategy which better accounts for differences in the beam shape, an improved temperature calibration scheme which accounts for the changing beam shape and changes in sky brightness as the Galactic plane rises, and an improved Bayesian modelling framework for the foreground contribution. It also details the implementation of an achromatic beamforming framework which was developed and deployed at LWA–SV. These improvements have led to improved residual RMS limits of $\sim 3$ K. This residual RMS is 3 orders of magnitude below the foreground signal, but still 2 orders of magnitude above the EDGES signal.

Chapter 4 expands upon the work focused on developing and deploying the achromatic beamforming framework at LWA–SV presented in Chapter 3. It describes how the digital backend electronics, the Adavanced Digital Processor, operates and which changes were made to it in order to allow for achromatic beamforming. This was achieved by amending the DRX pipelines to read in precomputed custom complex beamforming coefficients for each pointing in an observation. The work which went into simulating the beam patterns of LWA–SV led to the development of a Python package, Beam Simulator, which can simulate the beam pattern for any general antenna array. While this package could be of interest to anyone interested in simulating the beam pattern of an array, including focal plane arrays, this package has been developed specifically for the LWA and so there are a list of convenience functions which make it quick and easy to simulate the LWA beam response. A description of the package and a tutorial showing how to simulate the standard beam pattern of LWA–SV is provided in Section 4.5.

Chapter 5 describes work which is still ongoing to measure the impedance mismatch between the front end electronics (FEE) and the LWA antenna. This mismatch creates a reflection which reduces the amount of received power, thus reducing the overall sensitivity of the array. Understanding this mismatch as a function of frequency is crucial if true spectral response of LWA–SV is to be properly understood.
The newly measured values differ from previous models of the front end electronics, however these measurements only account for the hybrid coupler on the FEE. A better calibration of the vector network analyzer using calibration fixtures which account for more of the signal chain is most likely needed to properly measure the true spectral response of the impedance mismatch.

6.2 Future Work

The limits reported in Chapter 3 are encouraging as the residual RMS is approaching the sub-Kelvin level. However, pushing below the sub-Kelvin level will require very detailed knowledge of the limiting systematics. These include, but are most likely not limited to, low-level RFI which is not captured by traditional flagging methods, unaccounted variations in the electronic response of LWA–SV, the chromaticity of the LWA dipole gain pattern, and potential cross coupling of signals within the analog backend electronics. Future challenges are addressed below.

6.2.1 Developing a Robust RFI Detection Algorithm

Anthropogenic radio frequency interference is a major limiting factor for experiments attempting to detect the redshifted 21 cm signal from Cosmic Dawn and the EoR. Even very weak RFI which induces a small power variation in a subband of the observed frequency channels can raise the residual RMS by an amount which can obscure the tiny 21 cm signal. RFI can enter from either the main lobe or the sidelobes of the beam, which means that even RFI near the horizon can cause problems. Modern day digital transmissions are broadband and appear more noise-like making them difficult to detect.

Low-level RFI is extremely difficult to detect due to the nature of traditional flagging algorithms. In Chapter 3, it is explained that RFI flagging is carried out by finding outlier channels after a smooth model is fit to the data and removed. However,
low-level RFI does not clearly stand out from the data and so is not usually flagged using this method. There has been recent work on developing new RFI detection algorithms using deep learning (Itschner and Li, 2019, Kerrigan et al., 2019). These approaches should be adaptable to LWA data and could help in the detection of low-level RFI; however, developing a proper training set might be difficult.

There have also been efforts to clean up the RFI environment around LWA–SV. The majority of RFI in the area appears to come from power lines which can radiate when internal components on power poles fail. Novel RFI detection hardware has been developed (Taylor et al., 2020) and is currently being used in the area around LWA–SV; however, this work is ongoing.

A possible way of rejecting much of the observed RFI is to use both stations, LWA1 and LWA–SV, as a two element interferometer. Correlating the signals from both stations would suppress RFI since RFI present at only one station would decorrelate. However, this would require a new analysis method to search for the Cosmic Dawn signal within the measured visibilities.

### 6.2.2 Measuring the LWA Dipole Gain Pattern

The gain pattern of the LWA dipole has intrinsic chromaticity which is not accounted for in the current achromatic beamforming framework. Thus, the custom beam pattern of LWA–SV is not truly achromatic as this introduces a frequency dependent response. This intrinsic chromaticity might be able to be accounted for in how the dipoles are weighted, something that should be investigated in the future.

Besides the chromaticity of the LWA dipole, the simulations of the beam pattern of LWA–SV use a NEC model of the antenna gain pattern for a given frequency. The true gain pattern of the dipole is expected to differ from these models due to mutual coupling effects with other dipoles embedded in the array. Discrepancies between the model and true dipole gain patterns can introduce inaccuracies in the temperature calibration which will obscure the cosmological 21 cm signal. However, measuring
the gain pattern for an antenna at such low frequencies is difficult as traditional anechoic chambers are not large enough to simulate the far field radiation pattern since the wavelengths are huge ($\lambda \geq 3.4$ m). Furthermore, this would not address effects due to mutual coupling between antennas within the array. The External Calibrator for Hydrogen Observaties (ECHO) team and their work attempting to use drones to measure the gain patterns of low frequency dipoles has been mentioned in Chapters 2 and 3. Ongoing collaborative work with them will hopefully yield accurate measurements of the LWA dipole gain pattern which will update the current models used in the simulations of the station beam pattern. Initial data collection originally was carried out in October, 2019 and plans for a second observing campaign in 2020 were delayed due to the COVID-19 pandemic.

6.2.3 Alternative Achromatic Beamforming Frameworks

The achromatic beamforming framework presented in Chapters 3 and 4 is capable of shaping the beam main lobe in such a way that it is generally frequency and pointing independent, but the taper which is applied to the amplitudes of the complex beamforming coefficients introduces a decline in sensitivity by effectively decreasing the array effective area at higher frequencies. There may exist a better way of shaping the beam which does not affect the sensitivity as drastically. This might be achievable if the delay terms are changed instead of the amplitudes. This could potentially defocus the beam in an achromatic fashion, but doing this in predictable way may not be straightforward. Another possibility of an alternative framework would be the ability to steer nulls on strong extragalactic sources and RFI sources which creep into the data via the sidelobes. Techniques similar to those which have been developed for dynamically reconfigurable surfaces (Ellingson and Sengupta, 2021) might also allow for the controlling of the array sidelobe response. Dynamically reconfigurable surfaces have been proposed as a way to control the sidelobe pattern of a single reflector antenna, but it is plausible that an analogue of these techniques can be developed for antenna arrays in order to determine how to change the relative phases of the elements towards the edges of the array so that the sidelobe response is minimized.
6.2.4 Improving Modelling Robustness with Simulations

The frequency coverage of LWA–SV has some shortcomings in that it does not cover the entirety of the reported EDGES signal and only covers the lower edge through the center frequency. Modelling the absorption profile without data to constrain the upper edge of the signal is expected to be difficult. Simulations can be useful here as a mock signal which has incomplete frequency coverage can be generated and different modelling techniques can be evaluated. This is currently ongoing work. The EDGES signal is added to a very simple unchanging sky model, whose spectrum is a simple power law with spectral index of -2.5, and Gaussian noise. Preliminary results have shown difficulty in accurately recovering the injected EDGES signal.

6.2.5 Choice of Sky Model

Section 2.5 addressed the existence of other sky models which could be used for temperature calibration. The major limitation of the currently used GSM is the lack of maps, mainly at lower frequencies, with overlapping sky coverage. This means there is very little data of the sky at relevant frequencies for Cosmic Dawn and EoR science. Thus, for many points on the sky which do not have low frequency data available, an extrapolation is used to simulate the temperature. The principal component analysis (PCA), which is used to fit the frequency dependence of a given location on the sky, also only uses the best 11 maps, so the PCA components may suffer from frequency dependent biases. An “improved” GSM which uses an enhanced fitting method which builds off the previous PCA algorithm has been published (Zheng et al., 2017). The quality of calibration using this improved GSM, along with other sky models such as GMOSS, has not yet been investigated, but will be in the future.

The ideal map to use would be the Low Frequency Sky Model (LFSM) due to the fact that it is built from full sky maps at the relevant low frequencies which were made using LWA1; however, its minimum resolution of 5.1° means our beam FWHM would be less than a pixel in size. This might be remedied with the new
custom beamforming framework as the beam FWHM could be shaped to match the LFSM resolution, but simply matching the beam FWHM to the model resolution may not lead to the best results and reaching a beam FWHM of $5.1^\circ$ requires down weighting so many dipoles that the associated loss in sensitivity becomes a concern. Investigating the performance of the LFSM for the temperature calibration is left for the future as improved custom beamforming may open more possibilities.
References


G Bernardi, M McQuinn, and LJ Greenhill. Foreground model and antenna calibra-


References


References


