Carnegie Mellon University
MELLON COLLEGE OF SCIENCE

THESIS

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN THE FIELD OF PHYSICS

TITLE: **"**Studying the History of the Intergalactic Medium with the SCI-HI Experiment."

PRESENTED BY: Tabitha Voytek

ACCEPTED BY THE DEPARTMENT OF PHYSICS

Jeff Peterson 03/27/15

JEFF PETERSON, CHAIR PROFESSOR DATE

Stephen Garoff 03/27/15

STEPHEN GAROFF, DEPT HEAD DATE

APPROVED BY THE COLLEGE COUNCIL

Fred Gilman 03/27/15

FRED GILMAN, DEAN DATE

Studying the History of the Intergalactic Medium with the SCI-HI Experiment

by

Tabitha Christine Voytek

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
at
Carnegie Mellon University
Department of Physics
Pittsburgh, Pennsylvania

Advised by Professor Jeffrey Peterson

March 23, 2015

Abstract

The Cosmic Dawn ($z \sim 15-35$) is the period in the history of our universe when stars first began to form in small Dark Matter minihalos. Light from these first stars is too dim for telescopes to see, which means that the Cosmic Dawn has never been directly measured. However, the first stars impacted the gas, or intergalactic medium (IGM), around them. The impact of the first stars was heating and eventual ionization of the IGM. The process of heating and ionization creates a spectrum that varies over redshift, namely the spatially averaged brightness temperature spectrum of 21-cm light from the IGM. Measurement of this spectrum will give us a first glimpse of the Cosmic Dawn.

The ″Sonda Cosmologica de las Islas para la Deteccion de Hidrogeno Neutro″ (SCI-HI) experiment is a collaboration between Carnegie Mellon University (CMU) and Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE) in Mexico and was designed to make this measurement. The SCI-HI experiment is a small-scale system which travels with the team to remote locations for deployments. These remote locations are necessary to avoid radio frequency interference and other environmental impacts on the system.

This thesis describes the development and deployment of the SCI-HI experiment. It starts with the original design and covers development of the system over time. Deployment location selection is then discussed, including the results of site evaluations. In addition, the thesis outlines the data analysis process used for the system and shows results from data collected during the June 2013 deployment of the experiment. Finally, the thesis describes plans for the future of the SCI-HI experiment, including deployment to South Africa in 2015.

Acknowledgments

Any graduate thesis is a collaboration between the writer and the people around them who have supported them over the years. So I'd like to take a moment to recognize all those whose collaboration and help has made this thesis possible.

First, I would like to thank my advisor, Jeff Peterson, who has always been a great support and guide through the perils of graduate school. Thanks for helping me find the adventurous side of cosmology. I can't believe I get to travel the world and do science at the same time.

Second, I would like to thank all of the professors, post-docs and students who have worked on the SCI-HI project over the past several years. I would particularly like to acknowledge the hard work of my fellow graduate student, Jose-Miguel Jauegui-Garcia, who was my co-conspirator in developing much of the SCI-HI system. Also, Aravind Natarajan, the post-doc who worked with me on the data analysis side of the SCI-HI project. Some of the other folks who have worked hard to make the SCI-HI project a reality include, but are not limited to; Omar Lopez-Cruz, Edgar Castillo, Amy Stetten, Alma Gonzalez, Taylor Merritt, Aleksander Popstefanija, and Bruce Taylor.

Third, I would like to thank all of my collaborators on my other research projects who helped me to succeed and develop skills that I have used in my thesis. These collaborations include the GBT-IM team, the CHIME team, the Hydrogen Sky Planetarium Show team and many others.

Fourth, I would like to thank all of the students, post-docs and professors around the Physics department here at Carnegie Mellon University; especially those in the McWilliams Center for Cosmology. I never would have survived without the support, advice and feedback I have recieved from all of you. Special thanks to the members of my thesis committee, who helped keep me on track and get to this point.

Fifth, I would like to thank all of my extended support network who have helped me through this roller coaster. Thanks to all of my SWEsters, who constantly remind me that anything is possible and that women in STEM are awesome.

Finally, I would like to thank all of my family for their constant support and encouragement. Special thanks to Mom, Dad, Josh, Abe, Dave and Tim for putting up with my craziness when I needed it most. Look, look, I did it!

Contents

1 Introduction **1**

1.1 Hydrogen 21-cm Science . . . 2

1.1.1 Hydrogen Atomic States . . . 3

1.1.2 Spin Temperature (T_S) . . . 4

1.1.3 Emission and Absorption . . . 4

1.1.4 Brightness Temperature (T_b) . . . 6

1.1.5 Opacity (τ_ν) . . . 6

1.2 The Intergalactic Medium (IGM) . . . 7

1.2.1 IGM Fundamentals . . . 7

1.2.2 Wouthuysen-Field Mechanism . . . 8

1.2.3 History of the IGM . . . 10

1.3 Measuring the 21-cm Signal . . . 15

1.3.1 Non-Cosmological Foregrounds . . . 15

1.3.2 Average (Global) Frequency Spectrum ($\langle \delta T_b \rangle$) . . . 16

1.3.3 Fluctuation Power Spectrum ($\tilde{\delta}_{T_b}(\vec{k})$) . . . 17

1.4 Hydrogen 21-cm Cosmology Experiments . . . 17

1.4.1 Global (All-Sky Average) Experiments . . . 18

1.4.2 Mapping Experiments . . . 18

2 SCI-HI System Development **20**

2.1 Overview . . . 20

2.2 Antenna . . . 21

2.2.1 Antenna Design Constraints . . . 21

2.2.2 Trombone Antenna . . . 22

2.2.3 HIbiscus Antenna Design . . . 26

2.2.4 HIbiscus Antenna Construction . . . 31

2.3 RF Electronics . . . 38

2.3.1 RF Electronics Chain Overview . . . 39

2.3.2 Antenna Electronics . . . 39

2.3.3 Faraday Cage Electronics . . . 47

2.3.4 SCI-HI Faraday Cage and System Power . . . 49

2.4 Data Processing System . . . 51

2.4.1 Data Processing Pathway 51
2.4.2 System Control and User Interface 52
2.4.3 Control Computer 53
2.4.4 Power Supply and Consumption 53
2.4.5 System Noise Generation 54
2.5 Summary 55

3 Radio Frequency Interference (RFI) and Site Testing 56
3.1 Overview 56
3.1.1 Site Testing Kit 57
3.2 Existing Site Evaluations 59
3.2.1 Carnegie Mellon University Pittsburgh, PA, USA 59
3.2.2 National Radio Astronomy Observatory (NRAO) Green Bank 60
3.2.3 Dominion Radio Astrophysical Observatory (DRAO) 62
3.2.4 Algonquin Radio Observatory (ARO) 63
3.3 New Site Evaluations 64
3.3.1 La Zona del Silencio 64
3.3.2 Isla Guadalupe 67
3.4 Future Sites 73
3.4.1 New Sites in Mexico 74
3.4.2 New Sites in South Africa and Antarctica 75

4 Low Frequency Radio Astronomy and the Earth's Ionosphere 76
4.1 Overview 76
4.2 Earth's Ionosphere 76
4.2.1 Ionosphere Layers 76
4.2.2 Ionosphere Properties 78
4.3 Total Electron Content (TEC) 79
4.3.1 Measuring TEC 79
4.4 Ionospheric Impacts 81
4.4.1 Refraction 81
4.4.2 Absorption 82
4.5 Earth's Ionosphere and the SCI-HI system 83

5 SCI-HI Data Analysis 84
5.1 Preparing the Data 84
5.1.1 Integration and Sampling 84
5.1.2 Truncation 85
5.1.3 RFI Excision 86
5.1.4 Time-variable RFI 89
5.1.5 Rebinning 90
5.2 Calibrating the Data 90
5.2.1 Calibration Datasets 90

5.2.2 On-site Impedance Measurement . . . 91
5.2.3 Milky Way Galaxy (GSM) Modeling . . . 93
5.2.4 Calibration Factor Calculation . . . 95
5.3 Removing the Foregrounds . . . 101
5.3.1 Polynomial Fitting . . . 101
5.3.2 Daily Residuals . . . 104
5.3.3 Frequency Limitations . . . 104
5.4 21-cm Signal Attenuation in the SCI-HI Data . . . 104
5.5 Ionosphere Impacts to the SCI-HI Data . . . 104
5.5.1 Structure Due to Refraction . . . 105
5.5.2 Structure Due to Absorption . . . 105
5.5.3 Identifying Ionosphere Structure . . . 106

6 Looking Toward the Future 107
6.1 Deployment of the Improved SCI-HI System . . . 107
6.2 Development of Expanded Global 21-cm Signal Experiments . . . 108

A Teaching 21-cm Cosmology to the Public 109
A.1 Hydrogen Sky Planetarium Show Overview . . . 110
A.2 Storyboard and Script . . . 110
A.2.1 Final Script . . . 111
A.3 On-site Filming . . . 120
A.3.1 Green Bank, WV . . . 120
A.3.2 Penticton, BC, Canada . . . 121
A.4 21-cm Data Maps . . . 124
A.5 Working in the Planetarium Environment . . . 126
A.5.1 Image Distortion . . . 126
A.5.2 Motion . . . 126
A.5.3 Using pop-outs . . . 127
A.6 Audio Production . . . 129
A.7 Feedback and Evaluation . . . 129
A.7.1 Process Feedback . . . 129
A.7.2 Script Development . . . 130
A.7.3 Intermediate Viewings . . . 130
A.7.4 Public Viewings . . . 130

List of Figures

1.1 Time-line of cosmic history with redshift and approximate time since the Big Bang for the different cosmological eras. 1

1.2 Simple atomic state diagram representing the hyperfine energy splitting of the Hydrogen ground state. 2

1.3 Simple atomic state diagram for the Wouthuysen-Field Mechanism. The pathway in blue corresponds to a $n_1 \rightarrow n_0$ transition while the pathway in red corresponds to a $n_0 \rightarrow n_1$ transition. Solid lines are the stimulated emission pathways, while the dashed lines are the absorption pathways. 9

1.4 (a) Plot of T_S, T_γ, and T_K during the Dark Ages, as calculated from initial conditions given by the CMB and (b) δT_b during the same period with only collisional coupling present. Plots come from Furlanetto et. al. [13] 11

1.5 Plots of (a) x_K and x_α, (b) T_γ and T_K with and without X-ray heating, and (c) δT_b during the Cosmic Dawn for different models of star formation from Natarajan et. al. [21] 12

1.6 Snapshots at z =14,12,10,8 (from top left to bottom right) from a simulation run using SIMFAST from Natarajan et. al. [21]. The simulation shows the expected δT_b signal in units of Kelvin during the Epoch of Reionization. 14

1.7 Plots of radio sky measurements used to construct the Global Sky Model [10] of radio foregrounds. The number in the upper left corner of each plot is the frequency (in GHz) at which the map data was collected, while the color scale is the log of the sky temperature in Kelvin. 16

2.1 High level block diagram of the SCI-HI instrument, showing the three primary sections of the experiment. 21

2.2 Trombone antenna set-up in its smallest (highest center frequency) configuration. 22

2.3 Trombone antenna set-up in its largest (lowest center frequency) configuration. 22

2.4 Mounting for the Trombone antenna, with a Lucite mount point and fiberglass support structure. 23
2.5 Tabitha in the process of tuning the Trombone antenna between measurements. 23
2.6 SCI-HI setup with Trombone antenna on site at Green Bank in August 2011. 24
2.7 SCI-HI setup with Trombone antenna on site at the Zona del Silencio in January 2012. 24
2.8 SCI-HI setup with Trombone antenna on site at the Algonquin Radio Observatory in August 2012. 24
2.9 SCI-HI setup with Trombone antenna on site at Isla Guadalupe in October 2012. 24
2.10 Simulated Trombone antenna design using cocoaNEC. Simulation software utilizes line segments which intersect to create structures. . . . 24
2.11 Circular ground plane with radials, simulated using cocoaNEC. The center, filled section has a $5m$ radius and the radials extend an additional $5m$. 24
2.12 Trombone antenna simulation results with an idealized infinite ground plane. (a) Shows the azimuth (left) and elevation (right) antenna beam for different frequencies with the antenna at its shortest (35 cm) length. (b) Shows the same plots with the antenna at is longest (86 cm) length. (c) Shows a Smith chart of the antenna complex reflectivity ($S11$), where the 1.0 point corresponds to $Z = 50\Omega$, for the shortest (35 cm) length. (d) Shows a Smith chart of the antenna complex reflectivity for the longest (86 cm) length. 25
2.13 Trombone antenna simulation results with a spider web ground plane included in the simulation. (a) Shows the azimuth (left) and elevation (right) antenna beam for different frequencies with the antenna at its shortest (35 cm) length. (b) Shows a Smith chart of the antenna complex reflectivity, where 1.0 is set to 50Ω, for the shortest (35 cm) length. 26
2.14 Simulated HIbiscus antenna as viewed from above. 26
2.15 Simulated HIbiscus antenna as viewed from one side. 26
2.16 Simulated HIbiscus antenna beam slice in azimuth along the passive petals of the antenna (H-plane), shown for different frequencies. (Red $= 55MHz$, Green $= 66MHz$, Pink $= 78MHz$, and Blue $= 89MHz$) . 27
2.17 Simulated HIbiscus antenna beam slice in azimuth along the active petals of the antenna (E-plane), shown for different frequencies. (Red $= 55MHz$, Green $= 66MHz$, Pink $= 78MHz$, and Blue $= 89MHz$) . 27
2.18 Reflectivity of the simulated HIbiscus antenna when connected to a 50Ω transmission line. The narrow dip at $\sim 55MHz$ is an antenna resonance. $S11$ is below -10 dB for wide frequency band. 27

2.19 Smith chart of the simulated HIbiscus antenna complex $S11$ relative to 50Ω for $30-180MHz$. The loop around the horizontal axis corresponds to frequencies with small $S11$. 27
2.20 Tabitha with one scale model HIbiscus antenna in the CMU Project REAL Chamber. 29
2.21 Jose-Miguel with a second scale model HIbiscus antenna in the CMU Project REAL Chamber. 29
2.22 Scale model antenna gain measured in the PREAL chamber. Data is taken along the plane of the dipole (active) petals of the antenna, the E-plane. 29
2.23 Scale model antenna gain measured in the PREAL chamber. Data is taken perpendicular to the plane of the dipole (active) petals of the antenna, the H-plane. 29
2.24 Scale model antenna $S11$ measured in the PREAL chamber and in the lab for one model. Short period structure in the lab data comes from reflections off the lab walls, while large period structure is the actual $S11$. 30
2.25 Scale model antenna $S11$ measured in the lab for different strip line gap widths. Changing the gap width tunes the center frequency and bandwidth of the region where the magnitude of the antenna reflectivity is low. 30
2.26 Scale model antenna $S11$ Smith chart as measured in the PREAL chamber and in the lab for one of the models. Frequency range for the data is $300-700MHz$. In this model, the $S11$ data is always highly complex. This is because the strip line gap width was relatively small (0.047 inches). 30
2.27 Scale model antenna $S11$ Smith chart as measured in the lab for different strip line gap widths. Frequency range for the data is $300-700MHz$. Wider strip line spacing moves the center of the main loop closer to purely real reflectivity and higher magnitude impedance. . . 30
2.28 HIbiscus antenna scaled for $70MHz$ as it was set-up for $S11$ measurement at CMU. 31
2.29 HIbiscus antenna scaled for $70MHz$ as it was set-up for $S11$ measurement on-site at Green Bank. 31
2.30 HIbiscus antenna PCB panels being cut. 32
2.31 Soldering the strip line to the HIbiscus antenna panels. 32
2.32 Sight line down one of the strip lines for the HIbiscus antenna with spacers in place. 33
2.33 Center of the HIbiscus antenna with spacers and center mount block in place. 33
2.34 HIbiscus antenna $S11$ reflectivity as measured on-site at Isla Guadalupe during the June 2013 deployment. 33

2.35 HIbiscus antenna $S11$ Smith chart for $30-200MHz$ as measured on-site at Isla Guadalupe during the June 2013 deployment. The loop around 1.0 corresponds to the frequency range of $f \sim 50-90MHz$. 34
2.36 Measured HIbiscus reflectivity for the new antennas scaled for $70MHz$ and $100MHz$. Measurements were made without the metal ground plane while on-site in Pittsburgh. 35
2.37 Measured Smith chart for the HIbiscus antennas scaled for $70MHz$ and $100MHz$. Measurements were made without the metal ground plane while on-site in Pittsburgh. Data shown has a frequency range of $20-250MHz$. 35
2.38 New HIbiscus antenna scaled for $70MHz$ laid out in the lab. $400MHz$ Scale model is shown in the picture for comparison. 36
2.39 New HIbiscus antenna scaled for $100MHz$ set-up on the lawn at CMU while its reflectivity was measured. 37
2.40 HIbiscus petals with hinged joints. 38
2.41 HIbiscus petals folded up in preparation for packing. 38
2.42 Block diagram of the RF Electronics for the SCI-HI instrument. . . . 38
2.43 SCI-HI system with HIbiscus antenna set up on site at Green Bank in May 2013. 39
2.44 SCI-HI system with HIbiscus antenna set up on site at Isla Guadalupe in June 2013 40
2.45 Center of the Trombone antenna as set-up with the calibration switch mounted directly below the antenna. 41
2.46 RF electronics box at the base of the Trombone antenna, which contains the second stage amplifier and filters. 41
2.47 Block diagram of the RF Electronics located at the base of the antenna. 41
2.48 New Lucite box containing all the antenna RF electronics. 42
2.49 RF electronics box attached to one of the HIbiscus antenna center mounts. 42
2.50 Amplifier reflectivity ($S11$) in dB for the two amplifiers used in the SCI-HI system. 43
2.51 Smith chart of the reflectivity of the two amplifiers used in the SCI-HI system. Data is shown for $20-250MHz$. 43
2.52 Amplifier gain ($S21$) in dB for the two amplifiers used in the SCI-HI system. 44
2.53 Large scale oscillation in the short calibration data due to a cable between the first and second stage amplifiers. Blue is the data with the original 1 meter cable, green is the data with a 2 meter cable, and red is the data with the amplifiers right next to each other. 45
2.54 Wiring block diagram for the antenna RF electronics. 46
2.55 Block diagram of the RF Electronics located inside the Faraday Cage. 47
2.56 Wiring block diagram for the Faraday Cage RF electronics. 48

2.57 Circuit diagram for the parallel port switch control circuit. 49
2.58 Faraday Cage around the SCI-HI data processing system as it was set-up in October 2012. 50
2.59 One of the power and switch control filter boxes for the SCI-HI system. Capacitive high-pass filters with $\geq$ 40 dB attenuation above $30MHz$ are used in the box. 50
2.60 Faraday Cage around the data processing system as it was set-up in June 2013. 50
2.61 Faraday Cage with double shielding around data processing system, currently under development. 50
2.62 User interface for the SCI-HI data processing system. Data shown in the interface was collected when the data processing system was not connected to the antenna. 52
2.63 Raspberry Pi control computer, which connects to the data processing system via ethernet and is used for field monitoring. 52
2.64 Most recent update to the data processing system. System is assembled inside a robust Faraday Cage (shown without lid in place). 54
2.65 Copper tubing with brass mounting used as part of the data processing system and Faraday Cage's water cooling system. 54

3.1 Site testing kit laid out in the lab (portable Spectrum Analyzer in inset). 57
3.2 Tabitha collecting RFI data with the site testing equipment while at one of the testing sites. 57
3.3 Resistor measurement with the site testing kit. Blue line shows measured power from a 50Ω resistor. In the data, FM spikes are RF leakage due to the measurement being made in the lab in Pittsburgh where the FM band is extremely loud. Meanwhile, the red line is the expected signal level of a 50Ω resistor based on Equation 3.3. The difference between the expected signal and actual signal can be attributed to limitations of the Spectrum Analyzer in measuring flux density. . . . 58
3.4 Map of evaluated sites in North America, created using Google Maps. 60
3.5 RFI measurement at Carnegie Mellon University in Pittsburgh, Pennsylvania collected April 23rd, 2009. In the data, there are many RFI sources that produce signals that have a magnitude above the plotted range. These strong RFI signals produce inter-modulations that corrupt data outside of the actual frequencies of the RFI. For example, strong FM signals cause inter-modulation below $88MHz$ and above $108MHz$. 61
3.6 Extent of the U.S. National Radio Quiet Zone around the Green Bank Site. 62
3.7 Robert C. Byrd Green Bank Radio Telescope as viewed from the site observation deck. 62

3.8 RFI measurement at the NRAO Green Bank site collected on May 18th, 2010. In the plot, colored boxes indicate bands of well known RFI sources. Red is the FM band, Purple is the Orbcomm satellite band, and Cyan is the military satellite band. Additional spikes in the data come from RFI sources not in the indicated bands. Large scale sinusoidal variation (or ringing) in the spectrum is caused by cable reflections in the system and is common to all data collected with the site testing kit. 63
3.9 Some of the DRAO facilities and telescopes (CHIME pathfinder in foreground). 64
3.10 RFI measurement at the DRAO Penticton site collected on December 14th, 2009. DRAO has excessive RFI in the entire spectrum below $400MHz$ and is generally noisier than Green Bank. Because some of these signals (particularly the FM band) are above the inter-modulation cutoff, some of the smaller spikes in known clear bands are believed to be caused by inter-modulation. In the plot, colored boxes again indicate bands of well known RFI sources (Red= FM band, Purple = Orbcomm satellite band, and Cyan = military satellite band). 65
3.11 RFI measurement at the ARO Algonquin site collected on September 12th, 2012. Like the GBT site, the ARO site is quiet at high frequencies compared to DRAO or Pittsburgh. However, significant RFI is still present below $300MHz$. In the plot, colored boxes again indicate bands of well known RFI sources (Red= FM band, Purple = Orbcomm satellite band, and Cyan = military satellite band). 66
3.12 View of the Zona del Silencio from one of the nearby peaks. The ecologist's camp where we stayed while running our tests is in the center of the picture. 67
3.13 RFI measurement near the ecologist's camp in the Zona del Silencio, collected on May 5th, 2010. The Zona del Silencio site has even less RFI than the previous sites at low frequencies, although some RFI spikes are still present. In the plot, colored boxes again indicate bands of well known RFI sources (Red= FM band, Purple = Orbcomm satellite band, and Cyan = military satellite band). 68
3.14 Map of Isla Guadalupe showing the sites of interest, made using Google Maps. 69
3.15 Collecting data with the site testing equipment on Isla Guadalupe. . . 69
3.16 Airplane used for access to Isla Guadalupe. 69
3.17 View of the plateau with the fishing village from the airplane. 69
3.18 Mexican naval vessel as it arrived at Isla Guadalupe to deliver supplies and pick us up. 70
3.19 View from on board the Mexican naval vessel during the trip back to the Port of Ensenada. 70

3.20 RFI measurement from the ecology camp at the summit of Isla Guadalupe collected on November 1st, 2012. At this site, the elevation means that the line of sight is longer and RFI can be seen from further distances than at lower elevations. Despite this fact, the spectrum is mostly free from RFI above $300MHz$. In the plot, colored boxes again indicate bands of well known RFI sources (Red= FM band, Purple = Orbcomm satellite band, and Cyan = military satellite band). 71
3.21 RFI measurement from the lava flow near the fishing village at Isla Guadalupe collected on November 1st, 2012. At this site, the lower elevation and the mountain between the lava flow and the mainland act as an RFI shield. This means that even the lowest frequency bands are relatively clear of RFI. In the plot, colored boxes again indicate bands of well known RFI sources (Red= FM band, Purple = Orbcomm satellite band, and Cyan = military satellite band). 73
3.22 Potential future site testing locations in the Southern Hemisphere. Image made using Google Maps. 74

4.1 Idealized average free electron density distribution in Earth's atmosphere. Layers are shown during day and night. Each layer corresponds to an inflection point in the free electron density distribution, referred to in the literature as a maximum. 77
4.2 Global Map of GPS Network stations used for the quasi-real time maps of TEC available online at iono.jpl.nasa.gov 79
4.3 TEC map for October 13th, 2014 at 18:40 UTC. (Image found at iono.jpl.nasa.gov) 80
4.4 TEC map for February 13th, 2015 at 01:10 UTC. (Image found at iono.jpl.nasa.gov) 80
4.5 North American TEC map for June 1st, 2013 at 21:00 UTC. (Image found at www.swpc.noaa.gov) 80
4.6 North American TEC map for June 1st, 2013 at 12:15 UTC. (Image found at www.swpc.noaa.gov) 80
4.7 Ray tracing for incident light with a single ionospheric layer. Sizes are not to scale. 81
4.8 Refraction angle ($\delta\theta$) for $TEC = 10TECU$ as a function of zenith angle and frequency from Vedantham et al [36]. The predicted level of refraction due to the Earth's ionosphere is below the detection limit of the SCI-HI data collected in June 2013. 82

5.1 Single dataset (1 second integration) of the entire frequency spectrum of raw data from the SCI-HI system. Also present are single datasets from the different calibration sources. 85

5.2 Truncated spectrum of a single dataset of raw data from the SCI-HI system. Also present are single datasets from the different calibration sources. 86
5.3 RFI comparison in the FM band between Isla Guadalupe and Green Bank, West Virginia. 87
5.4 Calibrated full day of data taken on June 6th, 2013 showing RFI variance for $85 - 100MHz$. Note that the vertical lines correspond to FM stations, while horizontal lines correspond to times of meteor activity. 88
5.5 Short (~ 1 hour) subset of data taken on June 6th, 2013. Data shows RFI variance for $89 - 92MHz$ during a period of high meteor activity. 88
5.6 Short (~ 1hr) subset of data taken on June 6th, 2013. Data shows RFI variance for $89 - 92MHz$ during a period of low meteor activity. . . . 88
5.7 SCI-HI system transmission efficiency between the first stage amplifier and the different switch inputs. Transmission efficiency is calculated using $S11$ measurements. Blue (η) is the transmission efficiency for the HIbiscus antenna, green ($\eta_{50\Omega}$) is the transmission efficiency for the 50Ω terminator, and red ($\eta_{100\Omega}$) is the transmission efficiency for the 100Ω terminator. Here $\eta = 1$ means that all the power from the switch input is transmitted through the first stage amplifier. 92
5.8 Simulated HIbiscus antenna beam on the sky in RA, DEC coordinates at Isla Guadalupe's latitude. Each map corresponds to the antenna beam at $70MHz$ for a different LST. 93
5.9 Sky temperature calculated using the GSM model from de Oliveira-Costa et al [10] at $70MHz$. The temperature map is in units of log_{10}(Kelvin), while the map coordinates are in RA, DEC. 94
5.10 Single day plot of uncalibrated data collected with the antenna on June 6th, 2013 at $70MHz$. 95
5.11 High frequency resolution time average of all the calibration datasets for June 3rd, 2013. 96
5.12 Calibrated high frequency resolution time average of all the calibration datasets for June 3rd, 2013. 97
5.13 Comparison calibrated high frequency resolution time averages of all the calibration datasets for June 3rd, 2013 before/after including the current noise contribution to the calibration. 99
5.14 JNC calibrated high frequency resolution time average of all the data taken on June 1st, 2013; along with the fits for the average calibrated datasets. 100
5.15 Single day fit of the $K_{\Delta GSM}$ calibration term with data collected on June 6th, 2013 at $70MHz$. 101
5.16 Fits for $K_{\Delta GSM}$ calibration term for multiple days in June 2013 at $70MHz$. Days where a smaller percentage of the data is available have poorer fits. 102

5.17 Daily mean of data from June 4th, 2013 calibrated using $K_{\Delta GSM}$, along with the foreground fit polynomial calculated for the data. 102
5.18 Daily mean of data from June 4th, 2013 calibrated using K_{JNC}, along with the foreground fit polynomial calculated for the data. 102
5.19 Log magnitude average and variance of the daily residual signals measured using data calibrated with either the Galaxy calibration (red) or Johnson Noise calibration (blue). Residual data is compared to the log magnitude of different models of the 21-cm or δT_bsignal (black). Model data shown has the frequency mean of δT_b for $60 - 90MHz$ subtracted to mimic the effect of foreground removal. This means that in the model data, the two minima in each model line correspond to the two sides of the absorption dip in the δT_b spectrum for that model. 103
5.20 Simulated 21-cm signal with $100K$ magnitude (red) and residuals (green) after calibration and foreground removal. 105

A.1 Alex and Meng checking an image during filming. 121
A.2 Alex, Meng and Hsiu-Hsien taking multiple shots of the GBT from its base. 121
A.3 Some of the monitors in the GBT control room, captured with the rectangular lens. 121
A.4 GBT as seen from the observation deck, captured with the rectangular lens. 121
A.5 GBT as seen from the observation deck, captured with the fisheye lens. 122
A.6 GBT as seen from its base, composited from images with the fisheye lens. 122
A.7 Tabitha on top of one of the CHIME cylinders during the DRAO site visit. 122
A.8 Starling nest hollowed out of an RFI antenna on one of the CHIME cylinders. 122
A.9 CHIME pathfinder cylinders, captured with the rectangular lens. CHIME control room is inside the shipping container on the left side of the cylinders. 123
A.10 Close up view of one of the CHIME cylinders, captured with the fisheye lens. 123
A.11 CHIME cylinders at dusk, along with the neighboring 26 m dish, captured with the fisheye lens. 123
A.12 Map of the Green Bank Telescope Intensity Mapping project "deep" field after foreground subtraction. In the image, X and Y are the Right Ascension and Declination of the data (centered at $14^h 31^m 28.5^s$ RA and $2°0'$ Dec). 124

A.13 Simulated CHIME map in green with blue and yellow temperature variation placed in a shell for the planetarium show. Image center has a model of the large scale structure we've seen so far with optical telescopes, and the red structure along the edge represents the Cosmic Microwave Background. 125
A.14 Panorama image of the CHIME telescope made using footage shot with the rectangular lens. Color saturation in the image is cranked up to fit the planetarium environment. 127
A.15 Snapshot of CHIME cylinder animation inside a pop-out with an image of the CHIME telescope as the background. 128

Chapter 1

Introduction

When we open the history book of the universe, we find blank pages in the middle – periods in cosmic history where we have no direct data. One such time period is the ″Dark Ages″ ($z \sim 1100$ to $z \sim 30$), which started when light and matter decoupled during recombination and ended with the births of the first stars in the universe. Following that came the ″Cosmic Dawn″ ($z \sim 30$ to $z \sim 10$), when the first stars lived, died and impacted the universe around them (see Figure 1.1). Thus far, very little is known about these cosmic periods since no one has been able to observe them directly.

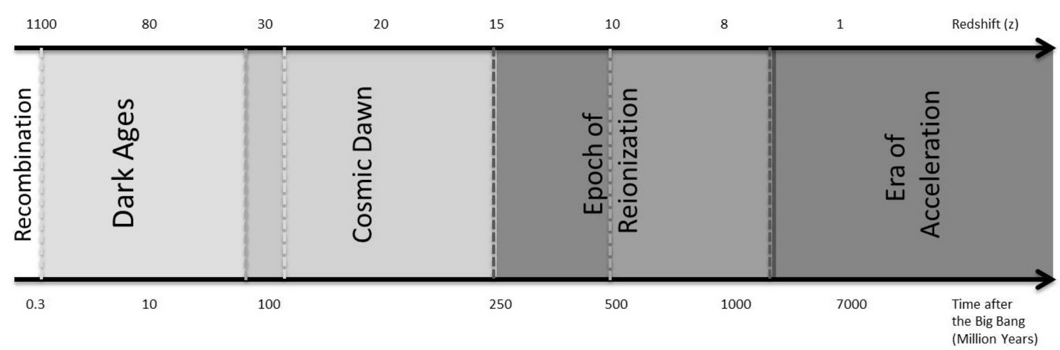


Figure 1.1: Time-line of cosmic history with redshift and approximate time since the Big Bang for the different cosmological eras.

However, the recent development of 21-cm cosmology has provided a potential tool for measuring these eras of our history and filling in some of the blanks. There are a number of experiments which seek to utilize the 21-cm line of neutral Hydrogen to study the universe, as will be discussed in Section 1.4. These experiments focus on a number of cosmological eras, including the Dark Ages, Cosmic Dawn, and the

Epoch of Reionization and Era of Acceleration that immediately follow them (see the time-line in Figure 1.1). We will discuss the cosmological eras mentioned above in detail in Section 1.2.3.

In this thesis I focus on one particular experiment in 21-cm cosmology called SCI-HI, which observes the end of the Dark Ages and the Cosmic Dawn. In Chapter 1 the science motivation for the experiment is outlined. A detailed description of the SCI-HI system is given in Chapter 2. Site selection for data collection with the SCI-HI system is discussed in Chapters 3 and 4. The data analysis procedure and results from SCI-HI data collected in June 2013 are discussed in Chapter 5. Plans for the future of SCI-HI are described in Chapter 6.

Moving beyond the SCI-HI experiment, Appendix A describes a strategy used to educate the public on 21-cm cosmology. This strategy is a planetarium show called ″The Hydrogen Sky″.

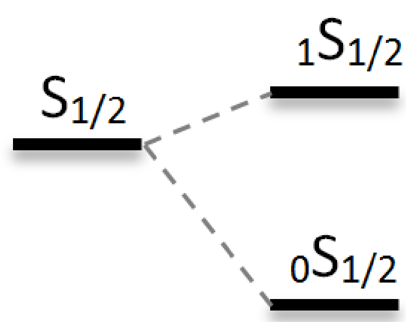


Figure 1.2: Simple atomic state diagram representing the hyperfine energy splitting of the Hydrogen ground state.

1.1 Hydrogen 21-cm Science

To set a background for 21-cm cosmology, we first need to understand the basic physics behind the signals that we are studying. There are four concepts that need to be outlined to facilitate this understanding.

1. We need to review the atomic states of Hydrogen and in particular the 21-cm hyperfine splitting of the ground state.

2. We need an overview of spin temperature including a definition of the spin temperature of a cloud of neutral Hydrogen atoms and a discussion of how emission and absorption change the spin temperature.

3. We need an overview of brightness temperature and how that temperature can change when it passes through a cloud of Hydrogen gas.

4. We need to define the relationship between the spin temperature, brightness temperature, and opacity of a cloud of Hydrogen gas.

1.1.1 Hydrogen Atomic States

The Hydrogen atom is the simplest atom, with only one proton and one electron. The proton and electron each have a spin $s_p, s_e = \pm^1/_2$. Interactions between the two spins lead to a splitting of the Hydrogen ground state into two energy levels. Using the Hamiltonian of a two spin system, we can calculate the eigenstates of the atom. These eigenstates can be written in terms of the spin state of the proton and electron ($|s_p, s_e\rangle$) as:

$$|1\rangle = |+\frac{1}{2}, +\frac{1}{2}\rangle \tag{1.1}$$

$$|2\rangle = |-\frac{1}{2}, -\frac{1}{2}\rangle \tag{1.2}$$

$$|3\rangle = \frac{1}{\sqrt{2}}(|+\frac{1}{2}, -\frac{1}{2}\rangle + |-\frac{1}{2}, +\frac{1}{2}\rangle) \tag{1.3}$$

$$|4\rangle = \frac{1}{\sqrt{2}}(|+\frac{1}{2}, -\frac{1}{2}\rangle - |-\frac{1}{2}, +\frac{1}{2}\rangle) \tag{1.4}$$

States $|1\rangle$, $|2\rangle$, and $|3\rangle$ correspond to the higher energy eigenvalue, while state $|4\rangle$ corresponds to the lower energy eigenvalue. The energy difference between the two eigenstates is called "hyperfine splitting", and has a value $E = h\nu_{10}$, where $\nu_{10} = 1420.405 MHz$ and $\lambda_{10} = c/\nu_{10} = 21\ cm$ [35].

Figure 1.2 shows a graphical representation of the energy levels of these two spin states, where atomic quantum number notation is used to represent the states. In this notation the azimuthal angular momentum (l) is represented by the letter (S means that $l = 0$). The sub-script after the letter is the total angular momentum of the electron ($j = l + s$), where s is the spin angular momentum of the electron. The sub-script before the letter is the total spin angular momentum ($F = 0, 1$) of the atom.

1.1.2 Spin Temperature (T_S)

When working with a cloud of Hydrogen atoms rather than a single atom, we need to compare the number of atoms in each of the two hyperfine states. To do this, we define a spin temperature using Boltzmann's law for a cloud in thermodynamic equilibrium. Spin temperature (T_S) is defined using the ratio of the the two states n_1/n_0, as shown in Equation 1.5. The factor of 3 in the equation comes from the ratio of eigenstates for the different spin energy levels of the ground state of Hydrogen (3:1) and we define a reference temperature: $T_* = h\nu_{10}/k$ [12].

$$\frac{n_1}{n_0} \equiv 3e^{-h\nu_{10}/kT_S} = 3e^{-T_*/T_S} \tag{1.5}$$

1.1.3 Emission and Absorption

Change of the spin temperature occurs through emission and absorption of photons whose energy matches the transition energy between the two spin states. Transitions occur when the atom either spontaneously emits a photon or is induced to emit or absorb a photon due to external forces. These transitions can be described using a differential equation (Equation 1.6) [13], and each type of transition has a different rate coefficient X_{ij}^m. The rate coefficient notation here is: (a) the type of transition, m, (b) the initial state of the atom, i, and (c) the final state of the atom, j.

$$\Big(\frac{dn_i}{dt}\Big)_m = X_{ij}^m n_i \tag{1.6}$$

If the cloud is in equilibrium, the rates of change are equal, holding $n_{total} = n_0 + n_1$ constant. So $dn_0/dt = dn_1/dt$, allowing us to calculate the spin temperature using Equation 1.7 [12].

$$n_0 \sum^m X_{01}^m = n_1 \sum^m X_{10}^m \tag{1.7}$$

Spontaneous Emission

Spontaneous emission due to a transition between the spin states is caused by quantum interactions with the electromagnetic environment. Because the spin transition is a forbidden one, the lifetime of the higher energy (triplet) state is over 10 million years. The rate coefficient for spontaneous emission is the Einstein A coefficient ($X_{10}^A = A_{10}$), defined using Equation 1.8, where β is the Bohr magneton [12].

$$A_{10} = \frac{64\pi^4\beta^2}{3h\lambda_{10}^3} = 2.85x10^{-15}sec^{-1} \tag{1.8}$$

Absorption and Stimulated Emission

Since absorption and stimulated emission are transitions which depend on external triggers, they can occur at a higher rate than spontaneous emission. In the types of Hydrogen gas clouds that we are considering, there are three primary triggers: (a) 21-cm radiation, (b) baryon collisions, and (c) optical radiation [12].

21-cm Radiation Incident radiation from an external source of 21-cm photons such as the Cosmic Microwave Background (CMB) will trigger absorption and stimulated emission. The rate coefficients of absorption and stimulated emission are set by the Einstein B coefficients ($X_{01}^R = B_{01}I_\nu$ and $X_{10}^R = B_{10}I_\nu$), where I_ν is the intensity of the external radiation. Since the CMB is a Blackbody, we know that $B_{10}I_\nu = A_{10}\lambda^2 I_\nu / h\nu_{10}$. Also, the relative number of eigenstates for the two spin states means that $B_{01} = 3B_{10}$ [12].

Baryon Collisions Collisions between the atoms and electrons in a Hydrogen gas cloud causes transitions described by coefficients $X_{01}^m = C_{01}$ and $X_{10}^m = C_{10}$. Because the system is in thermodynamic equilibrium, we can define a kinetic temperature (T_K) using the transition rate coefficients [12].

$$\frac{C_{01}}{C_{10}} \equiv 3e^{-T_*/T_K} \tag{1.9}$$

Optical Radiation Transitions to and from higher energy states of the Hydrogen atom due to optical radiation will also cause transitions between spin states. This is known as the Wouthuysen-Field Mechanism [38][12] and will be discussed in greater detail in Section 1.2.2. For now, we'll just define an optical brightness temperature (T_L). This gives us Equation 1.10, where the transition rate coefficients are $L_{01} = X_{01}^L$ and $L_{10} = X_{10}^L$ [12].

$$\frac{L_{01}}{L_{10}} = 3e^{-T_*/T_L} \tag{1.10}$$

Spin Transition Rate Equation

To calculate the spin temperature we combine the contributions from each of the sources of emission and absorption. Using the four sources described in the previous section, we get an expansion of Equation 1.7 for Hydrogen gas in equilibrium.

$$n_1(A_{10} + B_{10}I_\nu + C_{10} + L_{10}) = n_0(B_{01}I_\nu + C_{01} + L_{01}) \tag{1.11}$$

By rearranging and using Equation 1.5 to substitute for n_1/n_0, we get a direct relationship between the transition rates and the spin temperature.

$$\frac{n_1}{n_0} = 3e^{-T_*/T_S} = \frac{B_{01}I_\nu + C_{01} + L_{01}}{A_{10} + B_{10}I_\nu + C_{10} + L_{10}} \tag{1.12}$$

Since $T_* = 68.1mK$ and $T_S \gg T_*$, we can approximate the exponentials ($e^{-T_*/T} \simeq 1 - T_*/T$) and rearrange to get Equation 1.13. In this equation, the coupling coefficients are $x_K = T_*C_{10}/A_{10}T_K$ and $x_L = T_*L_{10}/A_{10}T_L$, and T_R is the incident 21-cm radiation temperature [12].

$$T_s \cong \frac{T_R + x_K T_K + x_L T_L}{1 + x_K + x_L} \tag{1.13}$$

1.1.4 Brightness Temperature (T_b)

The primary sources of incident radiation for Hydrogen gas clouds are Blackbody sources such as the CMB. These Blackbody sources have a spectral energy distribution ($B_\nu(T)$) defined by the Planck spectrum in Equation 1.14. For long wavelengths, we use the Rayleigh-Jeans approximation ($B_\nu(T) \simeq (2k\nu^2/c^2)T$) to give us a brightness temperature ($T_b(\nu)$) [8].

$$I_\nu(T) = B_\nu(T) = \frac{2h\nu^3/c^2}{e^{h\nu/kT} - 1} \tag{1.14}$$

Change in Brightness Temperature

Brightness temperature will change as the light travels through the universe. Such change can be caused by the expansion of the universe or by radiation sources and sinks such as gas clouds. Expansion of the universe simply causes the brightness temperature at each frequency to change at a rate proportional to the expansion rate (a).

In comparison, radiation sources and sinks cause a frequency dependent shift in the brightness temperature. The rate of change of the brightness temperature is represented as a cloud opacity (τ_ν). Using opacity allows us to write the brightness temperature after passing through a radiation source using Equation 1.15, where T_S is the source temperature and T_R is the original brightness temperature when it enters the source. As observers, what we actually care about and can measure is $\delta T_b(\nu) = (T_b(\nu))_{final} - T_R(\nu)$.

$$(T_b(\nu))_{final} = T_S(\nu)(1 - e^{-\tau_\nu}) + T_R e^{-\tau_\nu} \tag{1.15}$$

1.1.5 Opacity (τ_ν)

The change in brightness temperature caused by passing through a cloud of gas is determined by the cloud's opacity at a given frequency. Opacity (τ_ν) is defined as an integral along a line of sight through the cloud of the absorption rate of photons at the frequency ν. The opacity for Hydrogen gas is defined for frequencies ($\nu \approx \nu_{10}$) using Equation 1.16, where $\phi(\nu)$ is the line profile of the Hydrogen 21-cm spectral line at each position along the line of sight.

$$\tau_\nu = \frac{3c^2 A_{10}}{8\pi\nu^2}(1 - e^{-T_*/T_S}) \int ds\ \phi(\nu) n_0(s) \tag{1.16}$$

The line profile must have $\int \phi(\nu) d\nu = 1$ at each position. Meanwhile, $n_0(s)$ is the number of Hydrogen atoms at each position that are in the spin 0 state [13]. In order to determine $n_0(s)$ and $\phi(\nu)$ at each position, we need to know the details of the Hydrogen gas cloud.

1.2 The Intergalactic Medium (IGM)

Since Hydrogen is a dominant component of the Intergalactic Medium (IGM), the behavior of the IGM at 21-cm wavelengths can be modeled using a purely Hydrogen gas cloud. Using the physics discussed in the previous section, we know that a Hydrogen gas cloud changes the brightness temperature of 21-cm photons from the Cosmic Microwave Background (CMB) as they pass through the cloud. Therefore, the modified brightness temperature has a spectrum in frequency and space that depends on the history of the IGM. To make a prediction of this spectrum, we need a model of the history of the IGM during the time periods that we want to study. This model allows us to calculate a predicted T_S and τ_ν for the IGM.

1.2.1 IGM Fundamentals

Opacity (τ_ν)

To calculate the opacity of the IGM, we need to use Equation 1.16 with the appropriate distribution of Hydrogen atoms in the lower spin state (n_0). The distribution, defined as $N_{HI} = 4 \int ds\ \phi(\nu) n_0(s)$, is the column density of Hydrogen gas. The factor of 4 in the column density comes from the total number of eigenstates in the ground state of Hydrogen. The column density can also be written as the fraction of Hydrogen that is neutral (x_{HI}) times the number of Hydrogen atoms along the line of sight ($n_H(z)$) times the line of sight (s).

With all of these factors, we can re-write the opacity as shown in Equation 1.17, where $(1+\delta)$ is the matter density at a given position relative to the average matter density. In addition, $dv_\parallel/dr_\parallel$ is the gradient of proper velocity along the line of sight and $H(z)$ is the Hubble parameter [13].

$$\tau_\nu \approx 0.0092(1+\delta)(1+z)^{3/2} \frac{x_{HI}}{T_S} \Big[\frac{H(z)/(1+z)}{dv_\parallel/dr_\parallel}\Big] Kelvin \tag{1.17}$$

Brightness Temperature Change (δT_b)

As long as the opacity is small ($\tau_\nu \ll 1$), we can approximate ($e^{-\tau_\nu} \approx 1 - \tau_\nu$). As will be discussed in Section 1.2.3, this low opacity condition is expected to hold true during the eras of interest. In general, low opacity in the IGM occurs because the spin temperature is increasing with time during those eras and τ_ν is inversely proportional to the spin temperature (see Figures 1.4 and 1.5). Therefore, we can write the change in brightness temperature of CMB photons due to passing through the IGM using Equation 1.18, where $T_\gamma$is the CMB photon temperature, which also acts as the incident radiation temperature (T_R) in Equation 1.13 [13].

$$\delta T_b = \frac{T_S - T_\gamma}{1+z}(1-e^{-\tau_\nu}) \approx 9(1+\delta)(1+z)^{1/2}\Big(1-\frac{T_\gamma}{T_S}\Big)x_{HI}\Big[\frac{H(z)/(1+z)}{dv_\parallel/dr_\parallel}\Big]mK \quad (1.18)$$

1.2.2 Wouthuysen-Field Mechanism

Given that the spin temperature of the IGM is a significant term in the brightness temperature equation, we need to understand how it may evolve with time. As discussed in Section 1.1.3, the spin temperature of Hydrogen in the IGM is coupled to three sources during its history. The first source is the CMB (T_γ), which provides a source of external 21-cm radiation. The second source is the kinetic temperature of the gas (T_K), which characterizes the thermal motion of the atoms in the gas. The third and final source is external optical photons corresponding to transition energies between the ground and excited states of the Hydrogen atom. The most important transition for our purposes is the Lyman-α transition.

The mechanism for coupling between the Hydrogen spin temperature and Lyman-α photons is called the Wouthuysen-Field Mechanism [38][12]. To understand this mechanism, we must go back to the atomic states of the Hydrogen atom. Just as we've discussed for the 21-cm transition, external photons with an energy corresponding to the Lyman-α wavelength ($\lambda_{Ly-\alpha} = 121.6$ nm) can be absorbed by the atom. This induces a transition from the ground state of Hydrogen to its first excited state. The excited state is not stable, so the Hydrogen atom will eventually decay back to its ground state and emit a photon. However, conservation of angular momentum limits these stimulated emission transitions.

Figure 1.3 shows the possible transition chains for the Lyman-α absorption and stimulated emission process. One transition chain occurs when the original $S_{1/2}$ state is the $F = 0$ state, and can be written as ${}_0S_{1/2} \rightarrow {}_1P_{1/2} \rightarrow {}_1S_{1/2}$. The other transition chain occurs when the original $S_{1/2}$ state is the $F = 1$ state, and can be written as ${}_1S_{1/2} \rightarrow {}_1P_{3/2} \rightarrow {}_0S_{1/2}$.

So the absorption and stimulated emission of Lyman-α photons redistributes the the neutral Hydrogen atoms in the ground state such that $n_{tot} = n_0 + n_1$ is unchanged but the spin temperature increases or decreases. The exact magnitude of the change

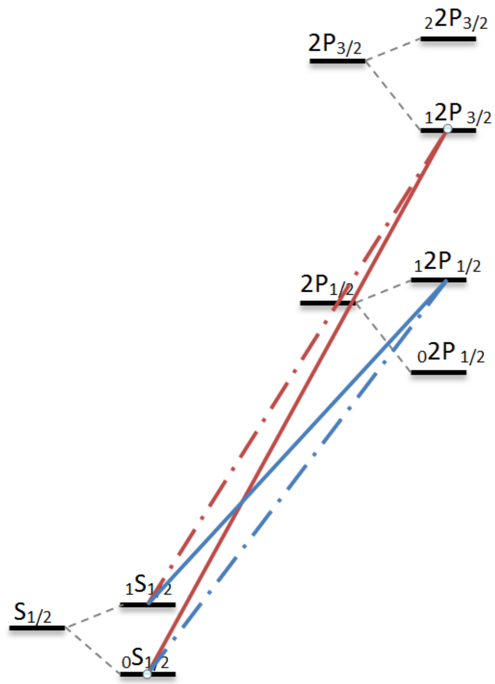


Figure 1.3: Simple atomic state diagram for the Wouthuysen-Field Mechanism. The pathway in blue corresponds to a $n_1 \rightarrow n_0$ transition while the pathway in red corresponds to a $n_0 \rightarrow n_1$ transition. Solid lines are the stimulated emission pathways, while the dashed lines are the absorption pathways.

in spin temperature will depend on the distribution of photons with wavelengths around the Lyman-α wavelength. This is because the absorbed photon frequency for $n_0 \rightarrow n_1$ is not identical to the absorbed photon frequency for $n_1 \rightarrow n_0$ due to hyperfine splitting.

Determining the rate of redistribution by the Wouthuysen-Field Mechanism requires the spectral distribution of Lyman-α photons in the IGM and the Lyman-α coupling coefficient. The spectral distribution of Lyman-α photons is set by collisions within the IGM, and approximates a Blackbody curve with temperature $T_{Ly-\alpha} = T_K$. Meanwhile, the Lyman-α coupling term ($x_L = x_\alpha$) is proportional to the intensity of Lyman-α photons produced by external sources of light. The exact Lyman-α coupling coefficient is related to the rate of scattering of Lyman-α photons by Equation 1.19, where P_α is the scattering rate [13].

$$x_\alpha = \frac{4 P_\alpha T_*}{27 A_{10} T_\gamma} \tag{1.19}$$

Calculation of the scattering rate requires knowledge of the absorption cross section of photons (σ_ν). It also requires knowledge of the intensity of incident radiation from external optical radiation sources (J_ν). This is shown in Equation 1.20 [13]. Optical radiation sources present during different cosmological eras will be discussed in Section 1.2.3.

$$P_\alpha = 4\pi \int d\nu J_\nu(\nu)\sigma_\nu(\nu) \tag{1.20}$$

1.2.3 History of the IGM

Initial Conditions

In the early universe before galaxies came into existence, the entire universe was filled with a medium that was the progenitor of the IGM. From measurements of the Cosmic Microwave Background, we know that at $z \sim 1100$ the IGM was a nearly homogeneous gas of protons, free electrons, and atoms; most of which were neutral Hydrogen (HI) atoms. Within the gas were small inhomogeneities in the matter density that would eventually grow into galaxies, but the average gas density was high enough that collisions between baryons were common throughout the medium [13].

Dark Ages

During the Dark Ages, adiabatic gas cooling and collisional coupling created structure in the spin temperature relative to the CMB temperature. Right after recombination ($z \sim 1100$), the IGM was dense enough for Compton scattering between CMB photons and free electrons in the IGM to set T_K=T_γ=T_S. But as the universe expanded, the gas adiabatically cooled faster than the CMB. This led to thermal decoupling at $z \sim 200$, after which T_K decreased below T_γ[13].

When this thermal decoupling occurred, collisions between baryons in the IGM were still frequent enough to keep x_K large. Therefore, the spin temperature began to drop below the CMB temperature thanks to coupling with the kinetic temperature, as described in Equation 1.13. However, x_K coupling gradually decreased as the rate of collisions between baryons decreased. In addition, $x_L \approx 0$ during the Dark Ages as there were not yet any stars to produce optical photons. Going back to Equation 1.13, we see that with $x_K \approx 0$ and $x_L \approx 0$, T_S= T_γ.

The process of cooling and collisional decoupling created a dip in T_S, which most models predict should be centered around $z \sim 80$ [13]. Figure 1.4 shows one model prediction for the relevant temperatures including T_γ, T_K, T_S and δT_b before star formation.

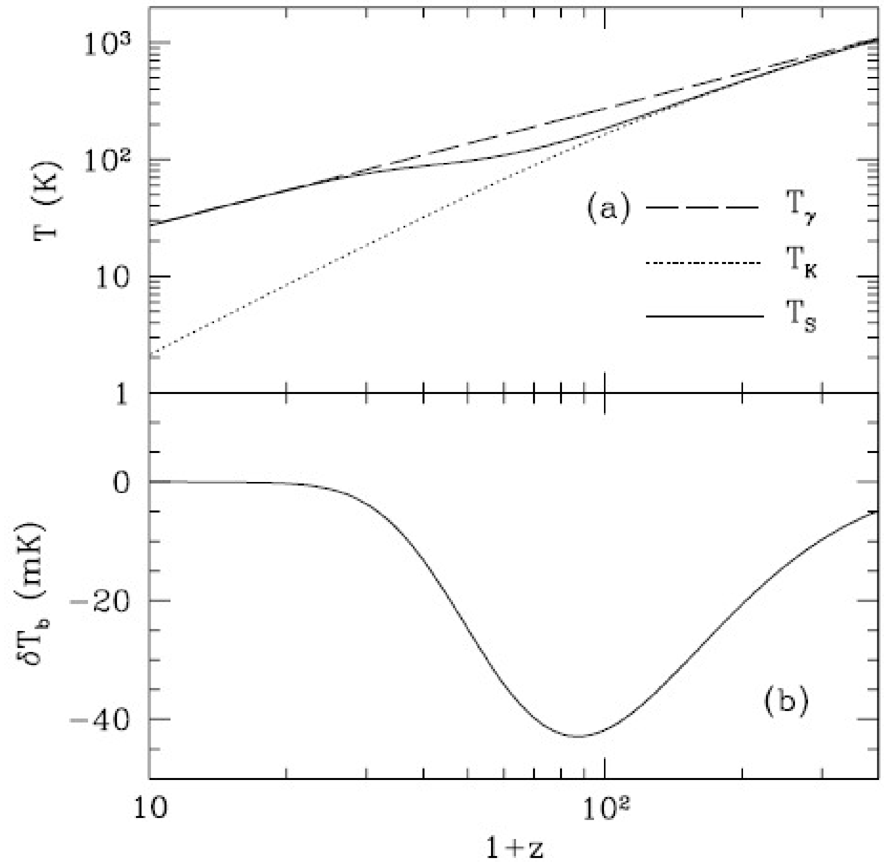


Figure 1.4: (a) Plot of T_S, T_γ, and T_K during the Dark Ages, as calculated from initial conditions given by the CMB and (b) δT_b during the same period with only collisional coupling present. Plots come from Furlanetto et. al. [13]

Cosmic Dawn

The beginning of the "Cosmic Dawn" is defined as the time when the first stars in the universe began to form. During the Dark Ages, hierarchical structure formation led to more and more massive dark matter minihalos. The first (PopIII.1) stars are believed to have formed in Dark Matter minihalos with mass $M_h \sim 10^6 - 10^8 M_\odot$. The minihalo mass range is set by the minimum mass needed for a minihalo to gravitationally collapse to a state of Virial equilibrium. Virial equilibrium occurs when $GM_h/R_{vir} \sim v_{vir}^2$, where R_{vir} is the radius of the minihalo after gravitational collapse and v_{vir} is the velocity of Dark Matter particles at the Virial radius.

Gas in the minihalos then cools by forming molecular Hydrogen (H_2) and eventually forms into PopIII.1 stars. Using ΛCDM simulations, the time required for small

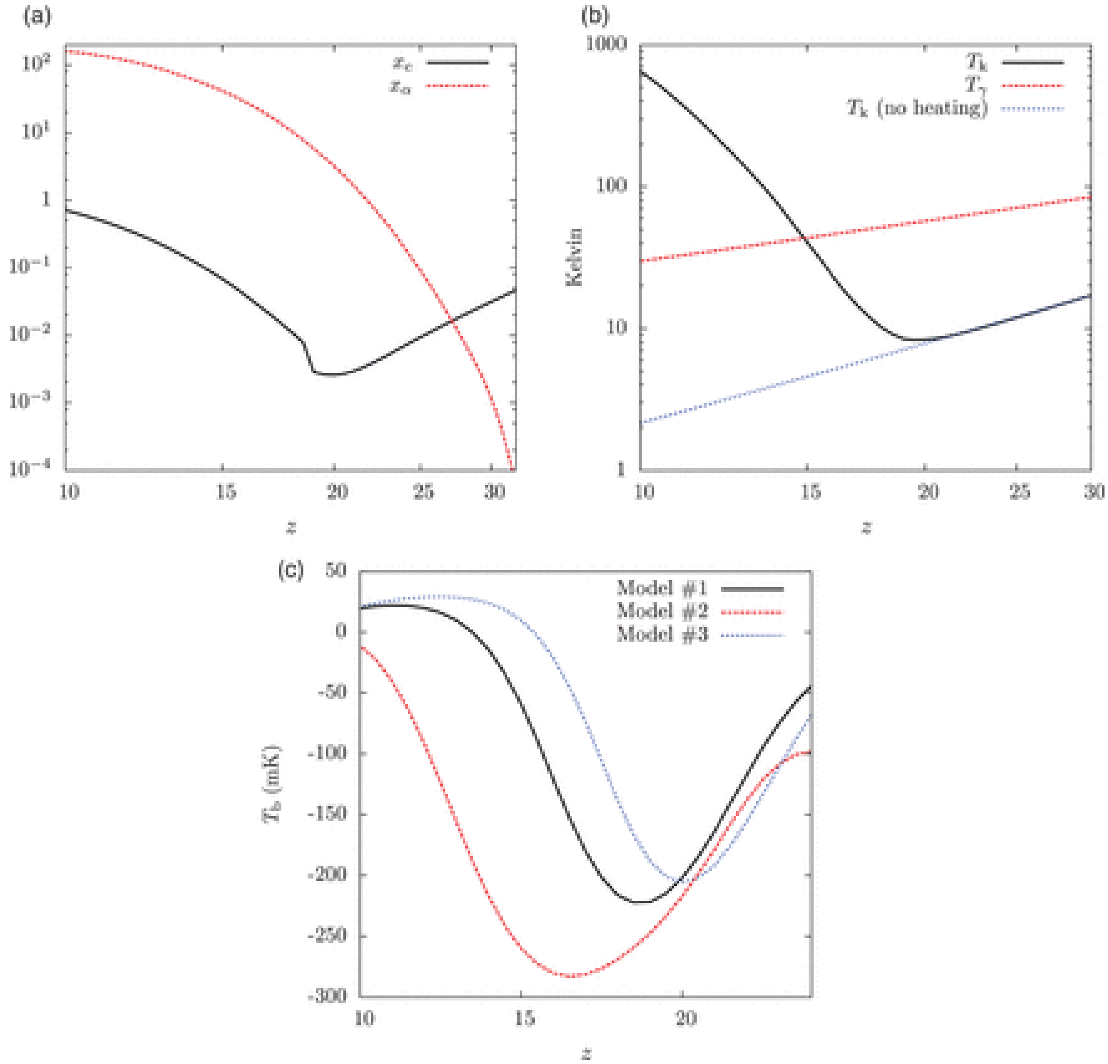


Figure 1.5: Plots of (a) x_K and x_α, (b) T_γ and T_K with and without X-ray heating, and (c) δT_b during the Cosmic Dawn for different models of star formation from Natarajan et. al. [21]

perturbations in the matter density to grow to the size needed for gravitational collapse can be calculated. This time corresponds to a collapse redshift of $z \approx 20 - 30$ [5]. PopIII.1 stars had three major impacts on the IGM: (a) PopIII.1 stars produced UV photons which ionized Hydrogen in the IGM, (b) UV photons from PopIII.1 stars also increased Wouthuysen-Field coupling and collisional coupling to the spin temperature, and (c) evolution of the first PopIII.1 stars also led to the first X-ray sources, which heated the IGM and increased its kinetic temperature.

Ionization by UV photons created bubbles of ionized Hydrogen (HII) around the stars. These bubbles gradually expanded and merged until the IGM was fully ionized. This process of ionization leads to the Epoch of Reionization, which will be discussed in the next section.

Meanwhile, UV photons included Lyman-α photons, which caused increasing coupling between T_S and T_K due to the Wouthuysen-Field Mechanism. The total amount of coupling from the first stars (x_α), shown in Equations 1.19 and 1.20, is related to the Lyman-α photon intensity J_α. This photon intensity comes from the emissivity of Lyman-α photons by the stars. Additional photon intensity comes from the emissivity of photons in other Lyman transitions which cascade down to affect the Lyman-α transition rate. Emissivity is heavily dependent on the rate of first star formation and the initial mass function (IMF) of PopIII.1 stars. Some models suggest that PopIII.1 stars may have a heavier IMF than is observed in the modern universe. If this is true, PopIII.1 stars would have produced more Lyman-α photons per star, changing the evolution of x_α [21].

During this era of first star formation, X-ray photons were produced by two types of sources among the first PopIII.1 stars. One source of X-ray photons was electrons accelerated by supernovae to relativistic speeds. When these electrons undergo inverse Compton scattering, X-ray photons are produced. The other source of X-ray photons was high-mass X-ray binary stars, which occur when massive stars on the main sequence lose material through accretion onto a compact neighbor such as a neutron star, black hole, or white dwarf star. These high energy photons heated the IGM through a combination of photoionization of Hydrogen or Helium atoms and collisions with IGM components. In most models of cosmic history, X-ray sources appear later than the first PopIII.1 stars [13].

Depending on the time delay between the appearance of the first PopIII.1 stars and the first X-ray sources, as well as the relative rates of X-ray heating and Lyman-α coupling, the predicted variance of the average spin temperature over time is modified. The spin temperature is predicted to dip during the period where $x_K > 0$ and T_K is less than T_γ. However, by varying the model of first star formation that sets these two rates for PopIII.1 stars, the exact shape of the dip in the spin temperature will be modified.

Most models predict a dip centered around $z \sim 20 - 30$, when the first PopIII.1 stars began to form. Models predict a dip width in the range of $0 \leq z \leq 10$ and a depth in the range of $0 \leq \Delta T \leq 300mK$. Figure 1.5 shows a few models of global δT_b during the Cosmic Dawn, as well as some of the parameters that feed into the brightness temperature. These models were run by Aravind Natarajan using SIMFAST code [26][20][21].

Epoch of Reionization (EoR)

As the kinetic temperature of the IGM continued to rise, it reached the threshold where Equation 1.18 breaks down because the kinetic and spin temperatures of the

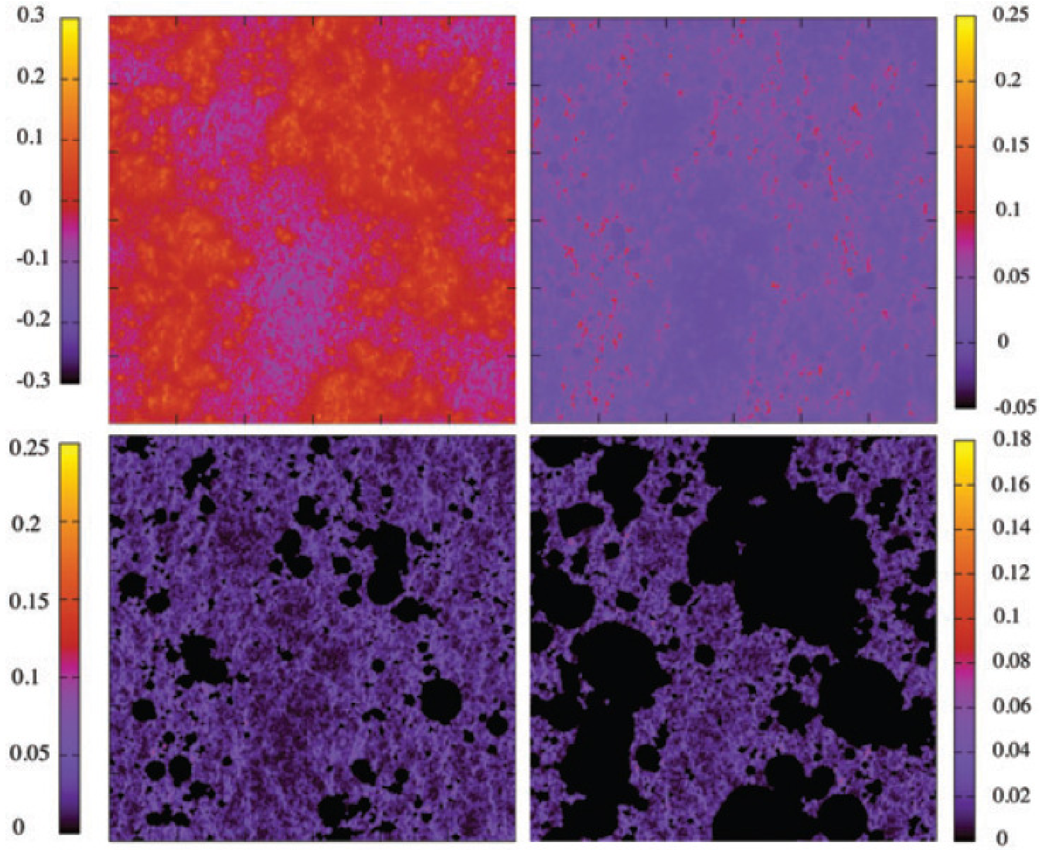


Figure 1.6: Snapshots at z =14,12,10,8 (from top left to bottom right) from a simulation run using SIMFAST from Natarajan et. al. [21]. The simulation shows the expected δT_b signal in units of Kelvin during the Epoch of Reionization.

IGM are large enough that the approximation used for τ_ν is no longer applicable. At this point δT_b reaches a maximum because of saturation of the signal.

Propagation of X-ray and UV photons through the IGM also began to ionize the Hydrogen atoms, starting in regions near the galaxies and spreading gradually throughout the IGM (see Figure 1.6). This period of time is also known as the Epoch of Reionization (EoR) because the IGM is resuming the ionized state that it had in the early universe. Since the brightness temperature of the 21-cm signal is also proportional to the amount of neutral Hydrogen (x_{HI}) present in the IGM, as shown in Equation 1.18, the ionization process caused a gradual decrease in the average 21-cm signal.

However, the gradual nature of the ionization means that there are interesting structures in the spatial distribution of the 21-cm signal during the EoR. The exact spatial distribution over time provides insight into the distribution of the PopIII.1

stars, since larger stars produce more ionizing photons compared to the number of Lyman-α photons [13].

Era of Acceleration

In the post-reionization universe, neutral Hydrogen gas in the IGM is limited to the regions around galaxies with a large column density. The minimum column density ($N_{HI} > 10^{21} cm^{-2}$) is set by the minimum density required to self-shield the gas from ionizing UV photons. This gas can still be mapped using the 21-cm signal, but the magnitude of the signal is much smaller than in previous eras.

In this regime, mapping of the spectral structure can be done through intensity mapping. Intensity mapping is a process by which the signal from neutral Hydrogen can be measured in aggregate around many galaxies. This is done with data that is collected in large voxels. Each voxel contains many galaxies and the 21-cm signal that is measured is the sum of the 21-cm signal from all the galaxies in that voxel [19]. This period of cosmic history is called the Era of Acceleration, because it was during this time ($z \sim 1$) that the expansion of the universe began to accelerate.

1.3 Measuring the 21-cm Signal

After passing through the IGM at a given time, the 21-cm photons travel toward the Earth. As they travel, the photons' frequencies are redshifted ($\nu_{meas} = \nu_{10}/(1 + z)$) from their original frequencies. This shift allows us to measure a spectrum $\delta T_b(\nu)$, where ν is the frequency at which the signal is measured, corresponding to the redshift where the signal was produced.

Measurements of the 21-cm spectrum are classified in two types. The first type are measurements of the sky-averaged spectrum ($\langle \delta T_b \rangle$). The second type are measurements of the spatial fluctuations in the 21-cm spectrum ($\delta_{T_b}(\theta, \phi, \nu) = \delta T_b(\theta, \phi, \nu)/T_b(\nu)$) using the power spectrum of the fluctuations ($\tilde{\delta}_{T_b}(\vec{k})$) [21]. This thesis focuses on a measurement of the first type, a sky-averaged spectrum.

1.3.1 Non-Cosmological Foregrounds

One of the greatest challenges of making any 21-cm measurement is the presence of foreground sources in the sky. Foreground sources include astrophysical objects which emit in the low frequency radio band ($\nu \leq \nu_{10}$), interference from Earth's ionosphere, and radio frequency interference (RFI) from man-made sources. I will discuss RFI in Chapter 3 and ionospheric impacts in Chapter 4, so in this section I will focus on astrophysical foregrounds.

For $\nu \leq \nu_{10}$ the dominant astrophysical foreground is synchrotron radiation from the Milky Way Galaxy. This radiation comes from free electrons in the Milky Way Galaxy and has a strongly downward sloping spectrum near the Galactic Poles [13].

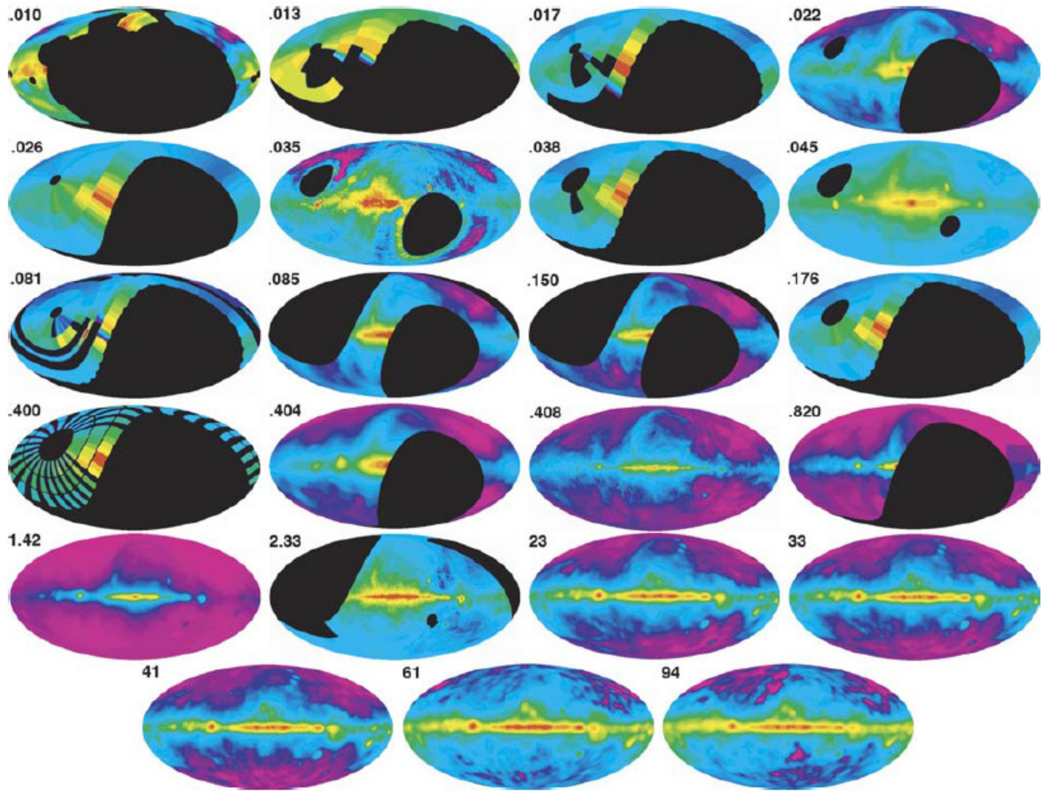


Figure 1.7: Plots of radio sky measurements used to construct the Global Sky Model [10] of radio foregrounds. The number in the upper left corner of each plot is the frequency (in GHz) at which the map data was collected, while the color scale is the log of the sky temperature in Kelvin.

$$T_{sky} \approx 180\Big(\frac{\nu}{180MHz}\Big)^{-2.6}K \tag{1.21}$$

Maps of astrophysical foregrounds have been made at a number of frequencies, including the ″Haslam″ map at $408MHz$ and the WMAP maps (after removal of the CMB signal) at 23,33,41,61 and $94GHz$. These maps are shown in Figure 1.7. A model of the foregrounds from $10MHz - 100GHz$ was constructed using the maps, as outlined in de Oliveira-Costa et al [10]. This sky model is commonly used for foreground assessment and even calibration by 21-cm experiments. Section 5.2.3 details how the GSM model is used in the SCI-HI experiment.

1.3.2 Average (Global) Frequency Spectrum ($\langle\delta T_b\rangle$)

To make a measurement of $\langle\delta T_b\rangle$, experiments measure the total sky brightness temperature spectrum (T_{Sky}) over a wide range of frequencies. The combination of signals collected by an antenna is given in Equation 1.22, where T_{fg} is the astrophysical foreground signal, T_{RFI} is the man-made signal, T_{sys} is the system temperature, and T_b is the brightness temperature from the CMB.

$$T_{Ant} = T_{fg} + T_b + T_{sys} + T_{RFI} = T_{Sky} + T_{sys} \tag{1.22}$$

The system temperature (T_{sys}) will be a combination of thermal noise and instrument noise. Thermal noise is set by the radiometer equation ($T_{thermal} = T_{Ant}/\sqrt{BW\tau}$), where BW is the system bandwidth and τ is the integration time [8]. Instrument noise comes from the electronics in the system (such as the amplifiers).

Extracting the 21-cm brightness temperature spectrum (T_b) requires removing all of the other terms in the data. Different strategies have been proposed for removing these signals, but all of them rely on the fact that the other signals have a different type of spectral structure than the 21-cm signal. I will discuss one such strategy for foreground removal with the SCI-HI experiment in Section 5.3.

1.3.3 Fluctuation Power Spectrum ($\tilde{\delta}_{T_b}(\vec{k})$)

Like $\langle \delta T_b \rangle$ experiments, power spectrum measurements are also dominated by foregrounds. Unlike $\langle \delta T_b \rangle$ experiments, these foregrounds are primarily specific galactic and extra-galactic point sources. These sources have a brightness temperature spectrum that can be fitted with a power law; although each point source has a different power law coefficient.

There are a number of different specific foreground removal strategies that have been developed, including principal component analysis [19][32]. All of these strategies rely on the removal of the dominant foregrounds through their smooth spectra. The 21-cm spectrum is left behind when the foregrounds are removed because it will have complex structure in frequency. This is because each frequency corresponds to the signal from a different set of galaxies in the intensity map.

1.4 Hydrogen 21-cm Cosmology Experiments

There are a number of experiments whose goal is measurement of the 21-cm global spectrum or 21-cm power spectrum. Some of these experiments are currently gathering data, while others are still under development. Preliminary constraints on 21-cm spectra have been placed for several cosmological eras using current experiments. Future measurements should continue to tighten constraints on the 21-cm global spectrum and 21-cm power spectrum during the Dark Ages, Cosmic Dawn, Epoch of Reionization and Era of Acceleration.

1.4.1 Global (All-Sky Average) Experiments

Experiments that are trying to measure $\langle \delta T_b \rangle$ focus on redshifts where $\langle \delta T_b \rangle$ is large and/or has a large first derivative. This is predicted to occur during the Dark Ages, Cosmic Dawn, and the Epoch of Reionization. EDGES [3] is a single antenna experiment focused on frequencies from $100 - 200 MHz$. It has placed limits on the duration of reionization by setting a minimum width for the 21-cm global signal emission bump during the Epoch of Reionization (EoR). Similar experiments, such as BIGHORNS[4][29], are also under development and are targeting the EoR spectrum.

Meanwhile, LEDA [17][2] and DARE [6] are currently under development, and are targeting the Dark Ages and Cosmic Dawn. LEDA is designed to operate on the ground using a combination of an interferometer and a single antenna, while DARE intends to launch a satellite in orbit around the moon with a single antenna.

Both of these experiments have large budgets of over 1 million dollars and require significant resources and technology development. In contrast, the EDGES experiment is on a much smaller scale and can be deployed in the field with minimal infrastructure.

SCI-HI Experiment

We were inspired by the EDGES experiment to develop a similar system that focuses on a different part of the 21-cm spectrum ($40 \leq \nu \leq 130 MHz$, or $34 \leq z \leq 9$), going after the global spectrum during the Cosmic Dawn. This experiment is the SCI-HI experiment, and is the main focus of the rest of this thesis.

1.4.2 Mapping Experiments

Experiments which seek to measure the 21-cm power spectrum are typically many-element interferometers, each experiment with its own antenna design and configuration. Early projects, including the GMRT-EoR [23] project, made use of existing telescopes to make measurements. Meanwhile, a number of projects were designed and constructed such as the Precision Array for Probing the Epoch of Reionization (PAPER) [24][15], the Murchison Widefield Array (MWA) [1][34], and the Low Frequency Array for Radio Astronomy (LOFAR) [16][18]. These projects have put first constraints on the power spectrum during the EoR. Future projects targeting the EoR signal include the Hydrogen Epoch of Reionization (HERA) [14][2] array and the Square Kilometer Array (SKA) [30].

Beyond the EoR projects, there are a number of power spectrum projects focusing on lower redshifts during the Era of Acceleration. The 21-cm signal targeted by these projects is much smaller than the EoR signal, but the foregrounds at these frequencies are also much smaller. First constraints on the power spectrum at ($z \sim 0.8$) were placed by the GBT-IM project [19][32], which used the Robert C. Byrd Green Bank

Telescope to make a map of a few degree patch of sky and construct a power spectrum from that map.

As with the EoR experiments, full sky mapping at later times also requires a dedicated instrument. One such instrument is the Canadian Hydrogen Intensity Mapping Experiment (CHIME) [28][7], currently under development. CHIME uses a unique cylindrical interferometer design to map the sky. Several other experiments are in earlier stages of planning and development, including the TIANLAI project in China and the HIRAX project in South Africa. Appendix A includes some discussion of the Green Bank Telescope and CHIME, and how they were used in an outreach project educating the public on 21-cm cosmology.

Chapter 2

SCI-HI System Development

2.1 Overview

″Sonda Cosmologica de las Islas para la Deteccion de Hidrogeno Neutro″ (SCI-HI) is an experiment which seeks to measure the 21-cm global spectrum during the end of the Dark Ages and the Cosmic Dawn before reionization. SCI-HI is a collaboration between Carnegie Mellon University (CMU) and Instituto Nacional de Astrofísica, Óptica y Electrónica (INAOE) in Mexico. The SCI-HI instrument was designed to meet several distinct constraints:

1. The instrument must be low cost (under $10,000).

2. The instrument must be low power (under 200 Watts) so that it can be run with an independent power source separate from external supplies.

3. The instrument must be simple and easy to deploy to remote locations.

4. The instrument must be broadband to cover the wide bandwidth of interest ($40 - 130MHz$).

5. The instrument must include a system response that varies smoothly with frequency.

In order to meet these constraints, the instrument design was broken down into three sections based upon their purpose in the overall design (see Figure 2.1). These sections are: (a) an antenna for signal collection, (b) radio frequency (RF) electronics for calibration and signal transmission, and (c) a computer to sample the signal, then process and store its spectrum.

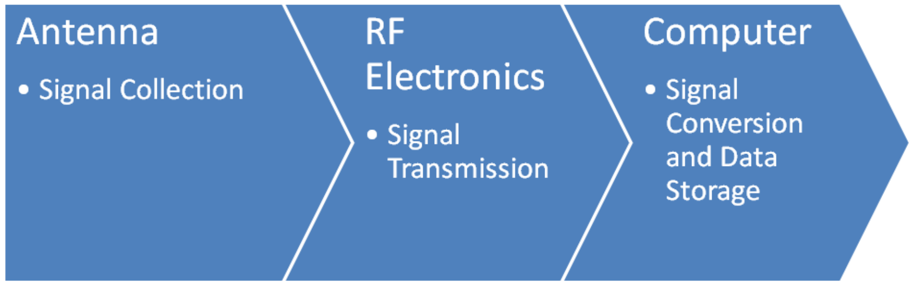


Figure 2.1: High level block diagram of the SCI-HI instrument, showing the three primary sections of the experiment.

2.2 Antenna

2.2.1 Antenna Design Constraints

The SCI-HI experiment needs a single dipole antenna that meets four design constraints.

1. The antenna radiation (beam) pattern must be stable over a wide bandwidth.
2. The antenna beam must have a main lobe centered at zenith and small sidelobes.
3. The antenna impedance must be stable and have a small reactance over its bandwidth.
4. The antenna must be small enough for easy transport and set up in remote locations.

To identify an antenna that meets the first three constraints, we utilized two design strategies. First, we simulated the antenna design using commercially available software to identify a basic design that met our criteria. Second, we used an antenna chamber to measure the antenna's beam and impedance.

Antenna Beam Shape Measurement

Measurement of an antenna's beam shape requires a special facility called an antenna range. The antenna range has a transmitting antenna with well known properties and an anechoic chamber large enough for the transmitting antenna and antenna-under-test to be in the far-field regime. Measurement is made by rotating the antenna-under-test with respect to the transmitting antenna. Meanwhile, the distance between the two antennas is held constant. The magnitude of the signal received by the antenna-under-test will change as a function of rotation angle. Mapping this magnitude over many angles gives the antenna's beam pattern.

Antenna Impedance Measurement

Antenna impedance is measured at an antenna range by connecting a Vector Network Analyzer to the antenna's terminal. A signal is sent down a cable from the analyzer into the antenna-under-test. The fraction of the original signal that is reflected back into the analyzer is measured. This reflection is called the reflectivity, or $S11$, and can be converted into impedance by comparing it to the reflection of a reference with known impedance. The conversion to impedance will be discussed in Section 2.2.4 [31].

Figure 2.2: Trombone antenna set-up in its smallest (highest center frequency) configuration.

Figure 2.3: Trombone antenna set-up in its largest (lowest center frequency) configuration.

2.2.2 Trombone Antenna

For the first version of the SCI-HI antenna, we started with a simple ″Trombone″ antenna. This design is based on the U.S. Naval Research Laboratory's Low-frequency Test Array (NLTA) dipole [11]. The design is a dipole with fat, angled elements over a ground plane and is shown in Figures 2.2 and 2.3. Changing the frequency range of the antenna (aka tuning) requires shifting the length of the dipole elements and their height above a ground plane.

We originally selected the Trombone antenna for its tuning capabilities. Tuning of the antenna causes features in the measured spectrum from the antenna design to shift in frequency. On the other hand, features in the measured spectrum which are from the sky stay at the same frequencies when the antenna is tuned. Lack of shifting during tuning allows us to distinguish real structure on the sky such as the 21-cm signal from structure due to the antenna.

Figure 2.4: Mounting for the Trombone antenna, with a Lucite mount point and fiberglass support structure.

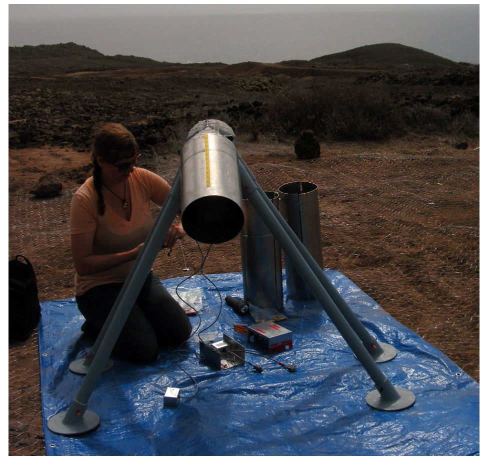


Figure 2.5: Tabitha in the process of tuning the Trombone antenna between measurements.

Antenna Construction

The Trombone antenna used in the SCI-HI experiment was constructed by Edgar Castillo, a graduate student visiting CMU from INAOE in 2011. The trombone shapes were constructed out of welded aluminum with a Lucite mounting block at the connection point of the two cones (see Figure 2.4), The dipoles and mounting block were supported with a structure constructed out of fiberglass and PVC pipes. The entire system was placed above a ground plane composed of two parts: (a) metal mesh (aka chicken wire) with a small $\sim 9m^2$ area and (b) long radial extensions ($\sim 10m$) extending out from the center like a spider web. Tuning the Trombone antenna length was done by adding an external aluminum tube that slid over the main dipole elements (see Figure 2.5). Meanwhile, tuning the height of the Trombone antenna was done using extendable PVC pipe support legs.

Antenna Deployment

The Trombone antenna was used throughout the initial stages of the SCI-HI project. This included deployments at Green Bank in West Virginia (Figure 2.6), the Zona del Silencio in Mexico (Figure 2.7), Algonquin Radio Observatory in Canada (Figure 2.8), and Isla Guadalupe in Mexico (Figure 2.9). For more discussion of these sites and why they were selected, see Chapter 3.

Antenna Simulation

To better understand and assess the behavior of the Trombone antenna, simulations were run by Amy Stetten, a graduate student at CMU. Figure 2.10 shows the sim-

Figure 2.6: SCI-HI setup with Trombone antenna on site at Green Bank in August 2011.

Figure 2.7: SCI-HI setup with Trombone antenna on site at the Zona del Silencio in January 2012.

Figure 2.8: SCI-HI setup with Trombone antenna on site at the Algonquin Radio Observatory in August 2012.

Figure 2.9: SCI-HI setup with Trombone antenna on site at Isla Guadalupe in October 2012.

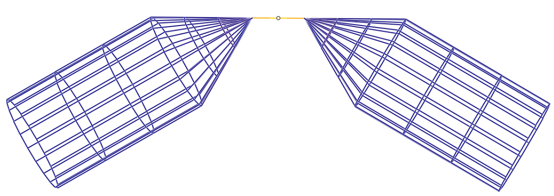


Figure 2.10: Simulated Trombone antenna design using cocoaNEC. Simulation software utilizes line segments which intersect to create structures.

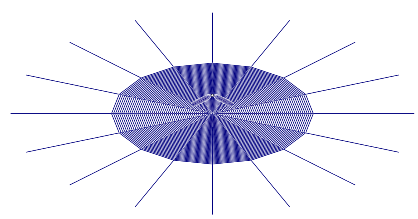


Figure 2.11: Circular ground plane with radials, simulated using cocoaNEC. The center, filled section has a $5m$ radius and the radials extend an additional $5m$.

ulated Trombone design. Simulations were run using cocoaNEC,[1] a free antenna modeling code designed for use in a Macintosh OS environment.

[1] http://www.w7ay.net/site/Applications/cocoaNEC/

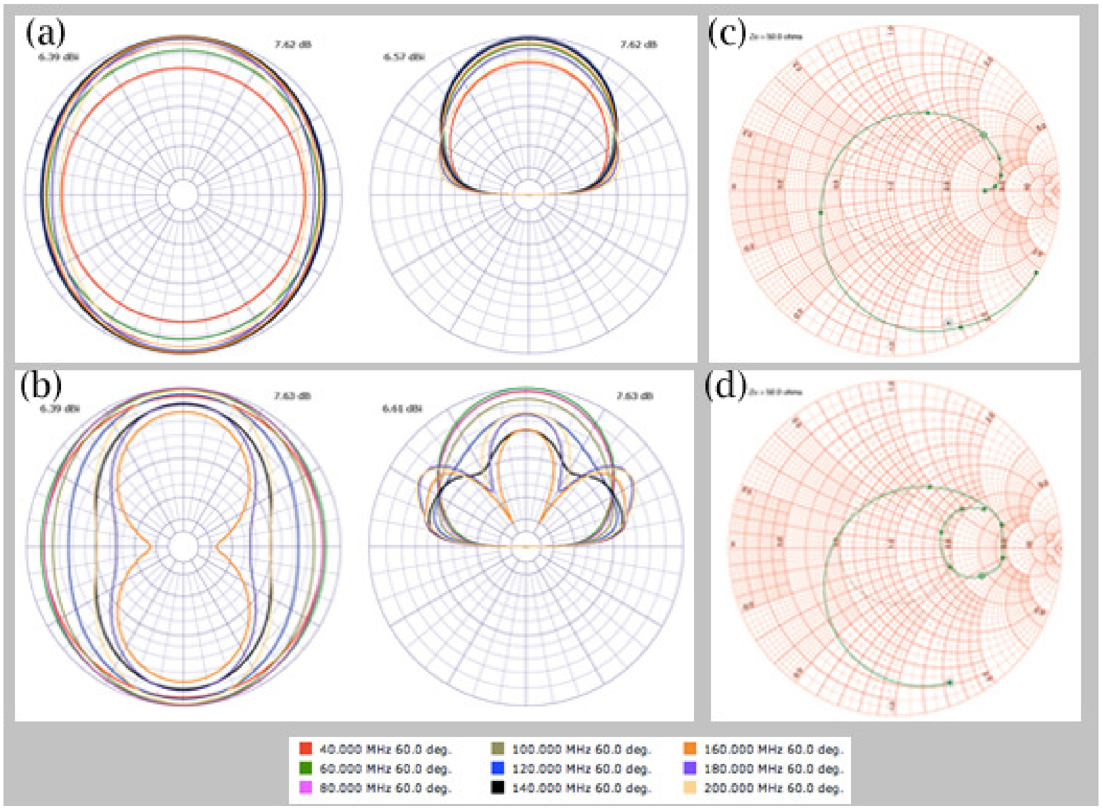


Figure 2.12: Trombone antenna simulation results with an idealized infinite ground plane. (a) Shows the azimuth (left) and elevation (right) antenna beam for different frequencies with the antenna at its shortest (35 cm) length. (b) Shows the same plots with the antenna at is longest (86 cm) length. (c) Shows a Smith chart of the antenna complex reflectivity ($S11$), where the 1.0 point corresponds to $Z = 50\Omega$, for the shortest (35 cm) length. (d) Shows a Smith chart of the antenna complex reflectivity for the longest (86 cm) length.

Simulations showed that the Trombone antenna beam was wide, but well behaved, for the shorter length as shown in Figure 2.12(a). In contrast, the longer length antenna beam had quite a bit of structure, as shown in Figure 2.12(b). In addition, the impedance of the antenna was not well matched to 50Ω and varied significantly across the band for all lengths (Figures 2.12(c) and 2.12(d)).

Beyond the Trombone dipole antenna design, it was important to examine the effect of a non-ideal ground plane. Preliminary work with a square ground plane led to significant additional structure in the antenna beam. To counteract this structure, the ground plane was modified to the ″spider-web″ design shown in Figure 2.11, with a center filled region and longer radials.

Adding this style of ground plane to the system did not have a big impact on

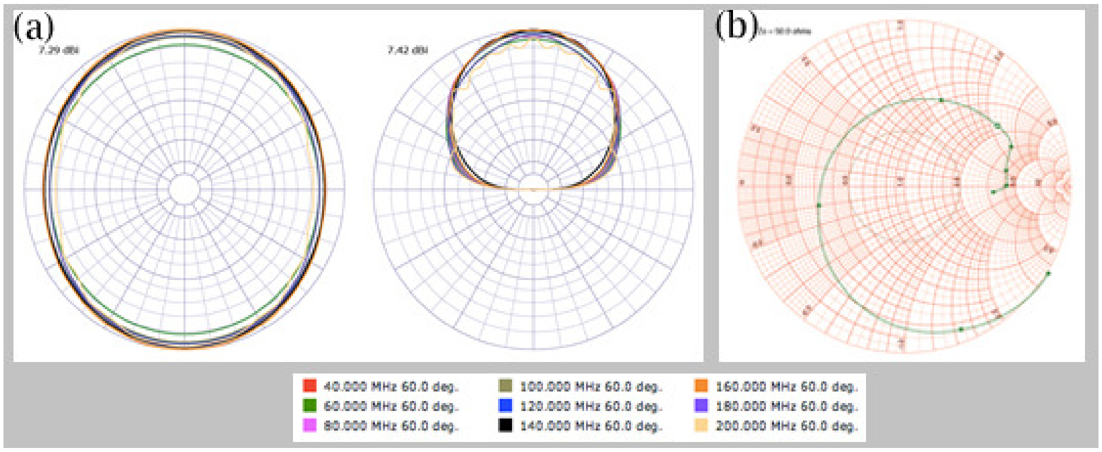


Figure 2.13: Trombone antenna simulation results with a spider web ground plane included in the simulation. (a) Shows the azimuth (left) and elevation (right) antenna beam for different frequencies with the antenna at its shortest (35 *cm*) length. (b) Shows a Smith chart of the antenna complex reflectivity, where 1.0 is set to 50Ω, for the shortest (35 *cm*) length.

the simulation results (see Figure 2.13) at most frequencies. There was additional structure in the beam at higher frequencies, but these frequencies are above the target frequency range of $40 - 130 MHz$.

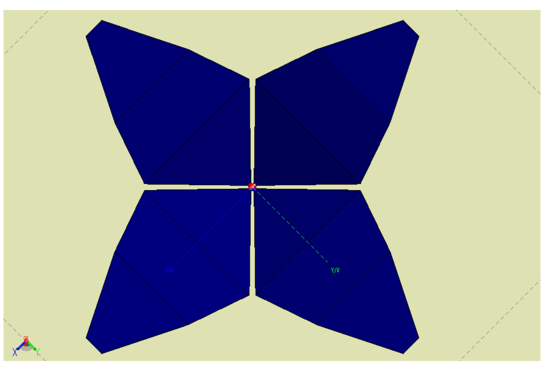


Figure 2.14: Simulated HIbiscus antenna as viewed from above.

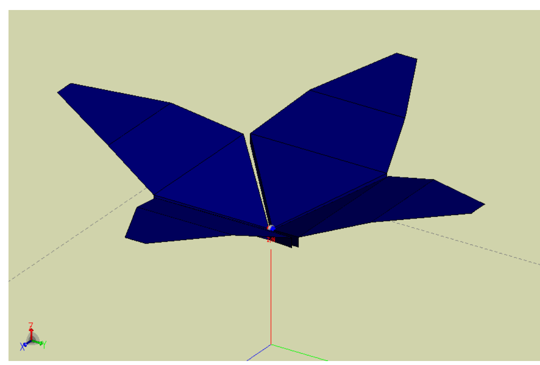


Figure 2.15: Simulated HIbiscus antenna as viewed from one side.

2.2.3 HIbiscus Antenna Design

Simulation and testing demonstrated that the Trombone antenna has a wide beam and is not well behaved at higher frequencies. In addition, the impedance of the antenna is complex and strongly varying over our frequency band.

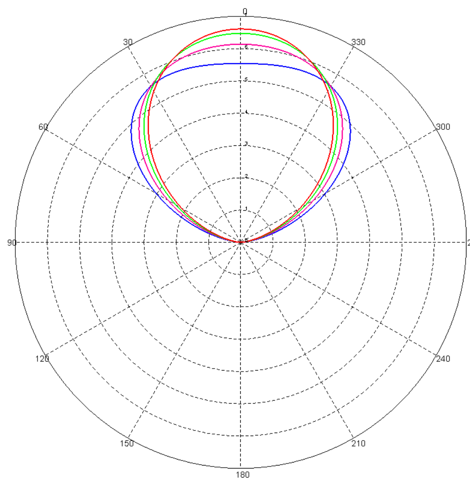


Figure 2.16: Simulated HIbiscus antenna beam slice in azimuth along the passive petals of the antenna (H-plane), shown for different frequencies. (Red = $55MHz$, Green = $66MHz$, Pink = $78MHz$, and Blue = $89MHz$)

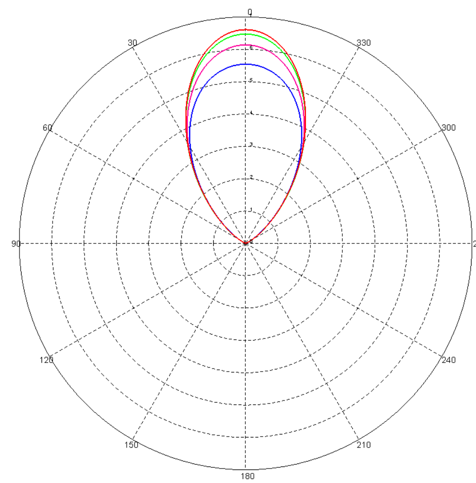


Figure 2.17: Simulated HIbiscus antenna beam slice in azimuth along the active petals of the antenna (E-plane), shown for different frequencies. (Red = $55MHz$, Green = $66MHz$, Pink = $78MHz$, and Blue = $89MHz$)

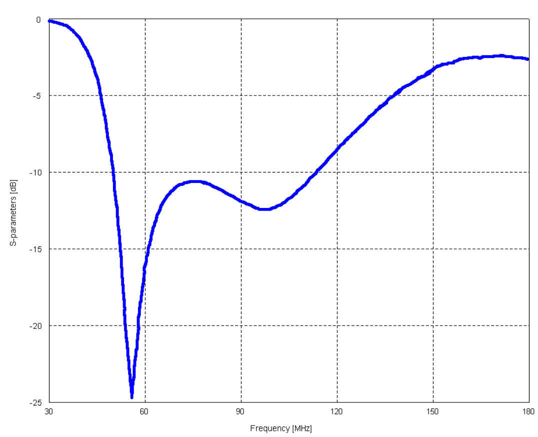


Figure 2.18: Reflectivity of the simulated HIbiscus antenna when connected to a 50Ω transmission line. The narrow dip at $\sim 55MHz$ is an antenna resonance. $S11$ is below -10 dB for wide frequency band.

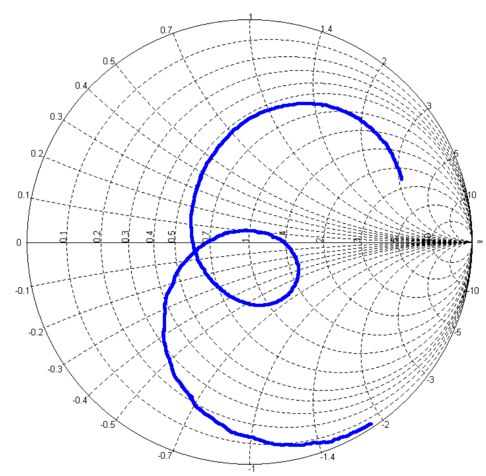


Figure 2.19: Smith chart of the simulated HIbiscus antenna complex $S11$ relative to 50Ω for $30 - 180MHz$. The loop around the horizontal axis corresponds to frequencies with small $S11$.

Because of these problems, we decided to take our antenna design in a new direction. For this new direction, we started with a modified four-square antenna design used by the EDGES experiment [3][25], and modified it to fit our needs. Simulations were run by Jose-Miguel Juaregui-Garcia, a graduate student visiting CMU from INAOE. He ran the simulations using FEKO, a finite element computational electromagnetics software package.[2]

Based on the FEKO simulation results, we settled on a design that used the EDGES design as a starting point, but changed the exact shapes of the four squares. We took each square and turned it into an inclined petal composed of three trapezoidal shapes. Each shape was connected to its neighbors on the petal, but had a different angle with respect to the ground. Additional panels were added to each side of the petal to create a strip line with a fixed gap between the petals. The entire antenna was then placed above a flat ground plane. We call this design a HIbiscus antenna, as shown in Figures 2.14 and 2.15.

This HIbiscus design has a $\sim 55°$ full width half-maximum (FWHM) and no sidelobes over a wide frequency band, as shown in Figures 2.16 and 2.17. The beam is narrower along the active elements (E-plane) of the antenna. The reflectivity of the HIbiscus antenna is complex, but is centered around $Z = 50\Omega$ such that it has less than $-10dB$ of reflectivity over a $60MHz$ band. The complex structure of the simulated reflectivity is shown in Figure 2.19, and the magnitude of the simulated antenna reflectivity is shown in Figure 2.18.

Scale Model Testing

In order to test our simulation to see if it matched the real antenna performance, Jose-Miguel built a set of scaled HIbiscus designs tuned for higher frequencies around $400 - 500MHz$. This allowed us to use an antenna range to measure the antenna beam shape and impedance. Jose-Miguel and I used the PREAL[3] (Project Remote Educational Antenna Laboratory) facility at Carnegie Mellon University to measure the antenna response of different scaled HIbiscus models. Figures 2.20 and 2.21 show two of the scale models being measured in the chamber.

Measuring scale models gave us the antenna gain for two azimuthal planes of the antenna. One, the E-plane, is along the active petals of the antenna. The other, the H-plane, is perpendicular to the E-plane and lies along the passive petals of the antenna. Figures 2.22 and 2.23 show the measured antenna gain for the scale model shown in Figure 2.20. Reflectivity of the scale model antenna was also measured using the PREAL facility, as shown in Figures 2.24 and 2.26.

Scale model testing indicated that antenna performance was particularly sensitive to two antenna shape parameters. These shape parameters are the strip line gap width and the height of the antenna with respect to the ground plane. Using measurements made after incremental adjustments of each parameter on one of the scale models, we optimized the antenna performance for our needs. Figures 2.25 and 2.27 show

[2] `https://www.feko.info/`

[3] `http://www.preal.ece.cmu.edu/`

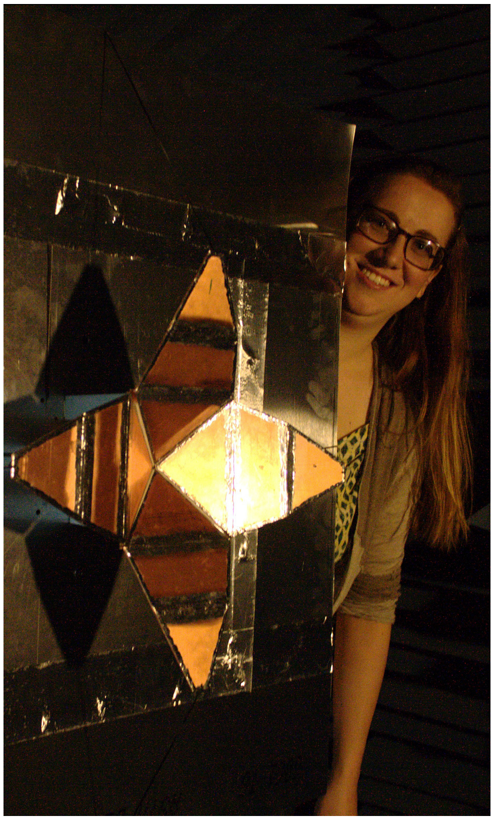


Figure 2.20: Tabitha with one scale model HIbiscus antenna in the CMU Project REAL Chamber.

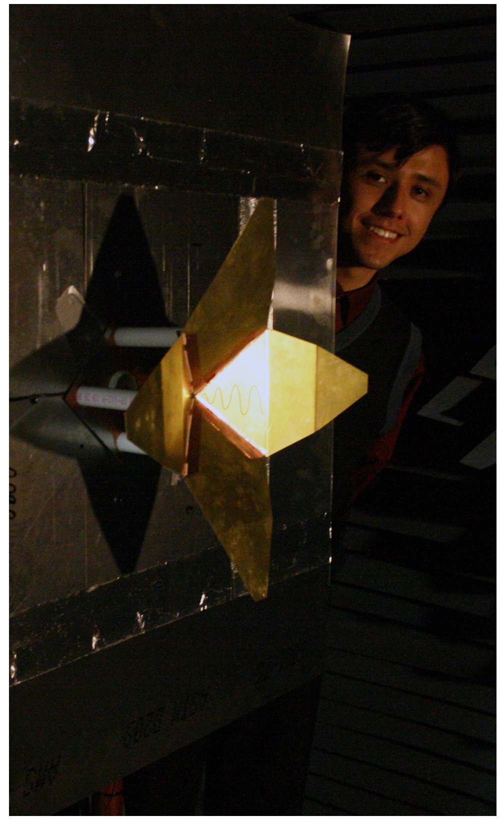


Figure 2.21: Jose-Miguel with a second scale model HIbiscus antenna in the CMU Project REAL Chamber.

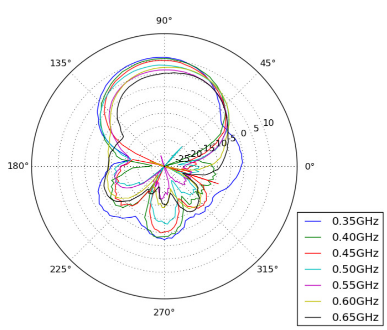


Figure 2.22: Scale model antenna gain measured in the PREAL chamber. Data is taken along the plane of the dipole (active) petals of the antenna, the E-plane.

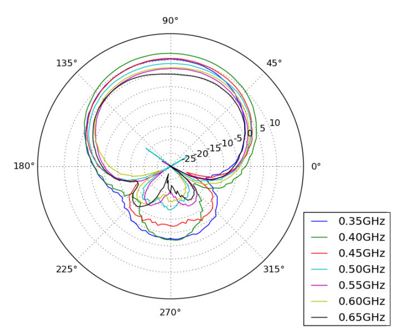


Figure 2.23: Scale model antenna gain measured in the PREAL chamber. Data is taken perpendicular to the plane of the dipole (active) petals of the antenna, the H-plane.

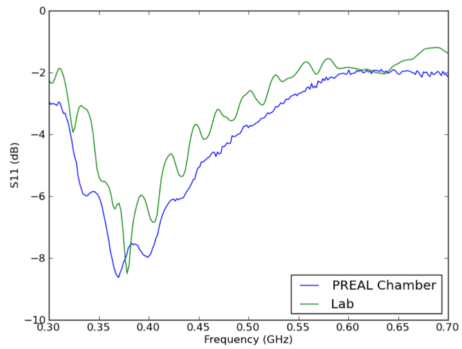


Figure 2.24: Scale model antenna $S11$ measured in the PREAL chamber and in the lab for one model. Short period structure in the lab data comes from reflections off the lab walls, while large period structure is the actual $S11$.

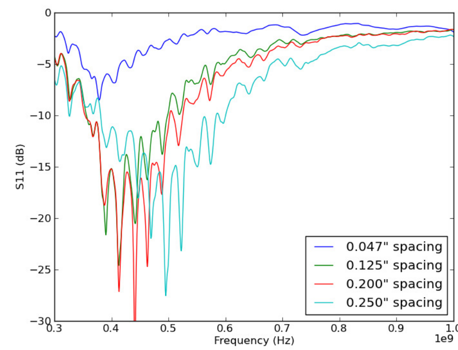


Figure 2.25: Scale model antenna $S11$ measured in the lab for different strip line gap widths. Changing the gap width tunes the center frequency and bandwidth of the region where the magnitude of the antenna reflectivity is low.

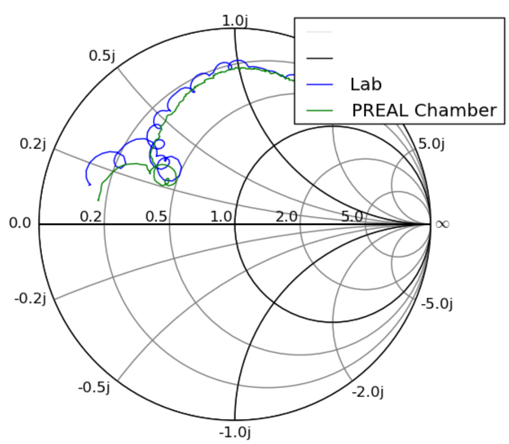


Figure 2.26: Scale model antenna $S11$ Smith chart as measured in the PREAL chamber and in the lab for one of the models. Frequency range for the data is $300-700MHz$. In this model, the $S11$ data is always highly complex. This is because the strip line gap width was relatively small (0.047 inches).

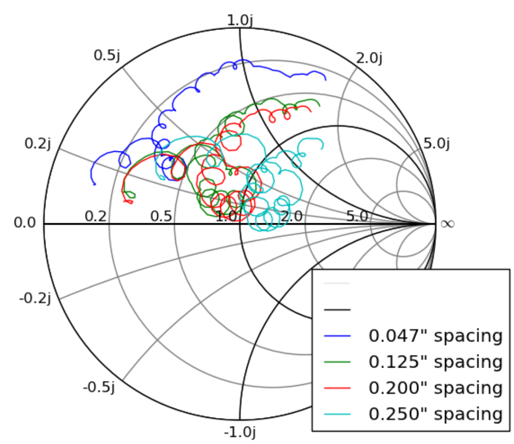


Figure 2.27: Scale model antenna $S11$ Smith chart as measured in the lab for different strip line gap widths. Frequency range for the data is $300-700MHz$. Wider strip line spacing moves the center of the main loop closer to purely real reflectivity and higher magnitude impedance.

one example of measurements made for incremental adjustments of the strip line gap width.

One challenge of making $S11$ measurements is the presence of short period oscillations in the $S11$ data. These oscillations come from reflections off the walls of the lab where the $S11$ data was measured. Removing these oscillations requires a very large room with minimal reflection or outdoor measurements. Figures 2.24 and 2.26 show a decrease in the magnitude of the oscillations when measurements were made using the PREAL chamber. This is because the chamber has lower reflection due to the anechoic foam on the chamber walls.

Beyond the small scale oscillations, the large scale structure of the $S11$ measurement can be seen to shift as we change the strip line gap width. In the Smith chart (Figure 2.27) we see that increasing the gap width moves the $S11$ data closer to the real axis and toward larger impedance around 1.0 or 50Ω. Using the scale model data, Jose-Miguel was also able to edit the antenna simulation to better match the actual antenna response. This gave us an excellent antenna model to use in our data analysis, as discussed in Chapter 5.

Figure 2.28: HIbiscus antenna scaled for $70MHz$ as it was set-up for $S11$ measurement at CMU.

Figure 2.29: HIbiscus antenna scaled for $70MHz$ as it was set-up for $S11$ measurement on-site at Green Bank.

2.2.4 HIbiscus Antenna Construction

Initial Assembly

Once we had a design finalized, Jose-Miguel and I set out to construct a HIbiscus antenna with its center frequency tuned to $70MHz$ (in the middle of the SCI-HI observation band). The entire antenna construction cost was $\leq$\$500.

Antenna petals were constructed in sections. To construct the antenna petals, large sheets of printed circuit board (PCB) material were used (see Figure 2.30). This material is double sided FR4-4 fiberglass sandwiched between two thin sheets of copper. Trapezoidal shapes were cut out of the large sheets, then brass flanges were added to connect the shapes. Underneath the center line of each petal we placed a rib made out of aluminum angle with the appropriate bends to match the angles between trapezoids. Finally, the strip line panels were added using brass sheets soldered to some of the trapezoids (see Figure 2.31). The entire system is bolted together using screws, and can be separated into sections for travel.

Figure 2.30: HIbiscus antenna PCB panels being cut.

Figure 2.31: Soldering the strip line to the HIbiscus antenna panels.

The next step after petal construction was mounting of the petals. First, a small Lucite mount was built for the connection points at the center of the antenna (see Figure 2.33). Next, Lucite spacers were placed between each of the petals to keep the petal separation even and correct (see Figure 2.32). Finally, the entire system was raised above the ground using a set of five PVC pipe legs; a central column and one leg for each petal (see Figure 2.28).

Beyond the antenna structure, a ground plane was also required to avoid variation in the antenna pattern and impedance due to weather conditions. Like with the Trombone antenna, we used a ground plane made out of metal mesh (chicken wire). Using simulations we found that the HIbiscus antenna was not as sensitive to the shape of the ground plane. Simulations showed that a $\sim 16m^2$ square ground plane gave good results and and was easy to lay out.

Weather-proofing was also important for the design. The entire system (antenna plus ground plane) was anchored to the ground to minimize the wind impact. Each of the four support legs had an anchoring rope (see Figure 2.29), and the chicken wire was anchored with many small stakes. This had the added advantage of keeping the ground plane flat to the ground.

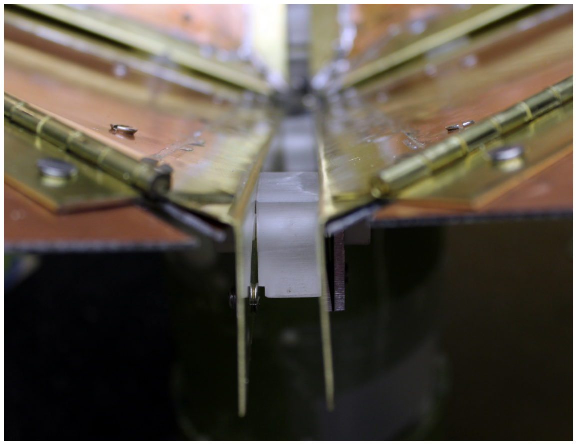


Figure 2.32: Sight line down one of the strip lines for the HIbiscus antenna with spacers in place.

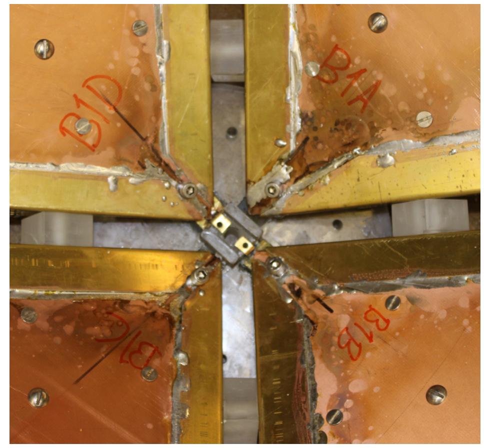


Figure 2.33: Center of the HIbiscus antenna with spacers and center mount block in place.

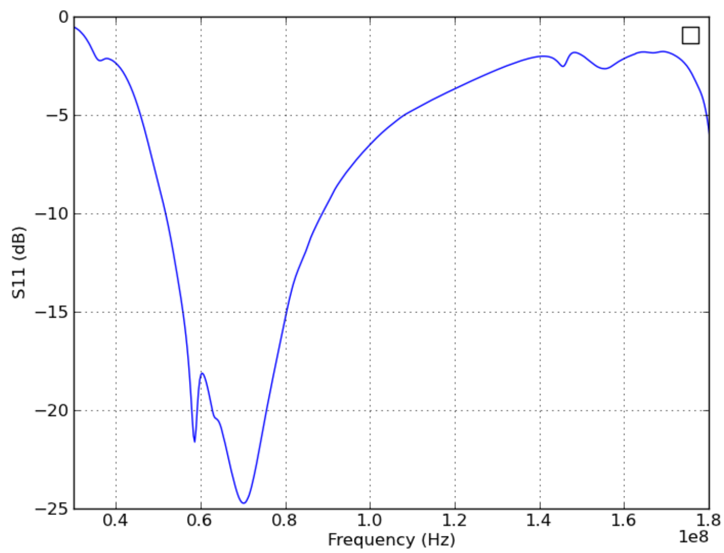


Figure 2.34: HIbiscus antenna $S11$ reflectivity as measured on-site at Isla Guadalupe during the June 2013 deployment.

S-parameters and Impedance

Following antenna construction, reflectivity was measured for the full size antenna. This allowed us to compare the full size results to the scale model measurements. I set up the antenna outdoors and made measurements with a portable Vector Network Analyzer (VNA). Measurements had to be made outdoors to avoid having signal reflections from the walls of a room interfere with the results. We saw this previously

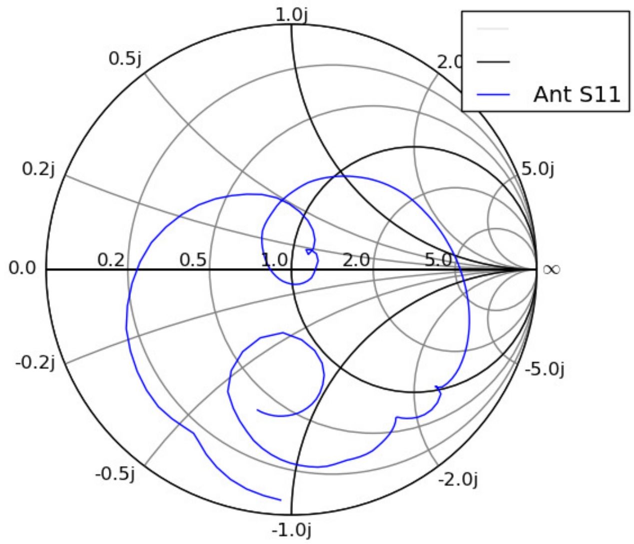


Figure 2.35: HIbiscus antenna $S11$ Smith chart for $30-200MHz$ as measured on-site at Isla Guadalupe during the June 2013 deployment. The loop around 1.0 corresponds to the frequency range of $f \sim 50-90MHz$.

with the scale model testing, and the longer wavelengths ($\sim 1-10m$) of the SCI-HI band make reflections an even larger problem. Figures 2.28 and 2.29 show examples of how I made this measurement at CMU and Green Bank with the help of Jose-Miguel and others.

The full size HIbiscus $S11$ measurement is shown in Figures 2.34 and 2.35. Data for the figures was collected on-site at Isla Guadalupe with the antenna ground plane in place. We can compare this measurement to the simulated reflectivity shown in Figure 2.18. The simulation successfully identified the antenna resonance between $50-60MHz$ and the overall shape of the antenna reflectivity matches the simulation. Like in the scale model data shown in Figures 2.24 and 2.25, the minimum reflectivity magnitude is much lower than the simulation predicts.

When the antenna was upgraded, as will be described later in this section, we again measured the $S11$ data for the antennas scaled to $70MHz$ and $100MHz$. These measurements assessed the new antennas and checked that they match expectations. $S11$ measurements for the new antennas are shown in Figures 2.36 and 2.37.

From the magnitude of the reflectivity in Figure 2.36 we see that measurement without the ground plane in the system removes the antenna resonance around $55MHz$. Meanwhile, the overall shape of the reflectivity matches the earlier antenna. Also as expected, the $100MHz$ antenna reflectivity looks like a scaled version of the $70MHz$ antenna reflectivity. Complex structure of the antenna reflectivity, shown in the Smith chart in Figure 2.37, also matches past measurements and simulation. The loop around 50Ω corresponds to frequencies where the reflectivity is at a

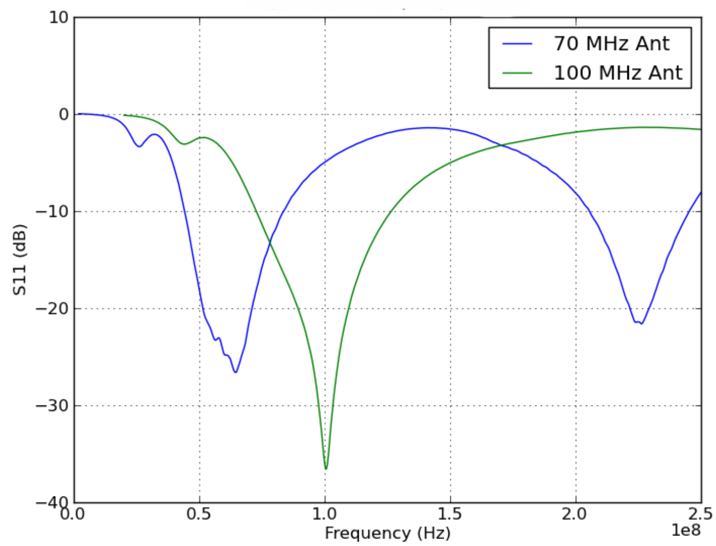


Figure 2.36: Measured HIbiscus reflectivity for the new antennas scaled for $70MHz$ and $100MHz$. Measurements were made without the metal ground plane while on-site in Pittsburgh.

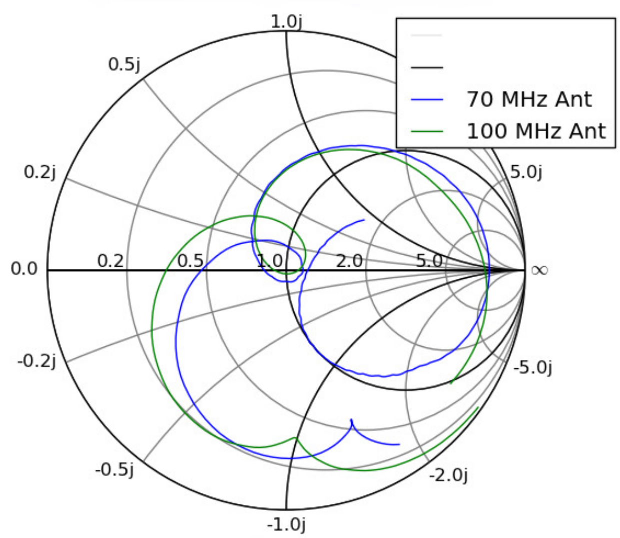


Figure 2.37: Measured Smith chart for the HIbiscus antennas scaled for $70MHz$ and $100MHz$. Measurements were made without the metal ground plane while on-site in Pittsburgh. Data shown has a frequency range of $20 - 250MHz$.

minimum.

Traveling with the Antenna

In order to use the HIbiscus antenna on-site at remote locations, it needs to be transportable to those locations and easily assembled upon arrival. Design of the mechanical structure of the antenna incorporated techniques that allowed us to meet this requirement.

First, the entire antenna structure is very lightweight. This is because we built the antenna petals out of PCB material, which has the stiffness of sheet metal without the weight.

Second, we laid out each of the antenna petals as a series of flat panels that can be easily assembled into a full petal or packed away into small bags. Each petal separates out into three trapezoids that screwed together for assembly, but could be packed as mostly flat sheets. The Lucite spacers between the petals are mounted using nylon screws, while the central mount uses metal screws to attach the petals to a short piece of coaxial cable. This allows the petals to be assembled or disassembled and packed rapidly. The PVC pipe supports are also mounted using screws and can also be rapidly assembled or disassembled and packed.

Third, the entire antenna structure is designed to disassemble into small, light pieces that allow it to be packed up into three suitcases plus an extra bag for the ground plane chicken wire. These suitcases meet the limitations for regular checked baggage on commercial airlines, typically 50 pounds and ∼62 linear inches.

Figure 2.38: New HIbiscus antenna scaled for $70MHz$ laid out in the lab. $400MHz$ Scale model is shown in the picture for comparison.

Figure 2.39: New HIbiscus antenna scaled for $100MHz$ set-up on the lawn at CMU while its reflectivity was measured.

Antenna Upgrades

Deployment of the SCI-HI system in June 2013 demonstrated that the HIbiscus antenna could use some improvements prior to future deployments. The first and most significant change was the addition of a second antenna scaled to have a center frequency of $100MHz$. By having both scales ($70MHz$ and $100MHz$), data can now be collected over a wider range of frequencies. In addition, the overlap region where both antennas work well ($\sim 85-95MHz$) provides a cross check for the signal. This cross check will help us to make sure that any spectral structure measured is not coming from the antenna.

We built a new $70MHz$ HIbiscus antenna (see Figure 2.38) as well as the $100MHz$ HIbiscus antenna (see Figure 2.39) based on the same simulated antenna design as the original HIbiscus antenna. However, our deployment helped us to identify areas of improvement to the mechanical design of the antenna. These changes are intended to make assembly and travel easier.

First, we replaced the brass flanges that had to be screwed together to assemble each petal with hinged joints made out of piano hinge that do not have to be disassembled and reassembled, as shown in Figure 2.40. When prepared for travel, the petals are folded and stacked as shown in Figure 2.41, and placed inside a bag for travel.

Second, we modified the support PVC pipes so that we simply add an extension

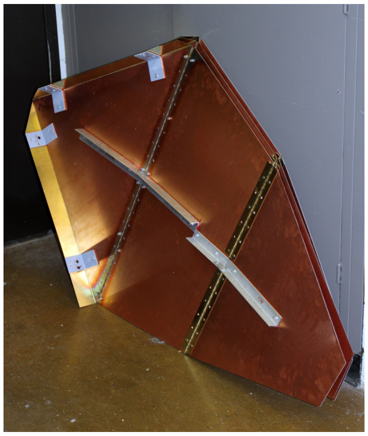


Figure 2.40: HIbiscus petals with hinged joints.

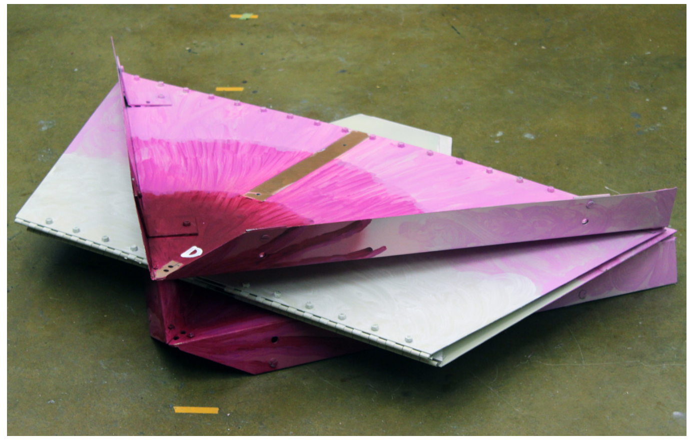


Figure 2.41: HIbiscus petals folded up in preparation for packing.

to the middle of the pipe to adjust the height for the $70 MHz$ HIbiscus antenna compared to the $100 MHz$ antenna. Therefore, adding the $100 MHz$ antenna to the experiment only requires one additional bag beyond the three required to transport the $70 MHz$ antenna.

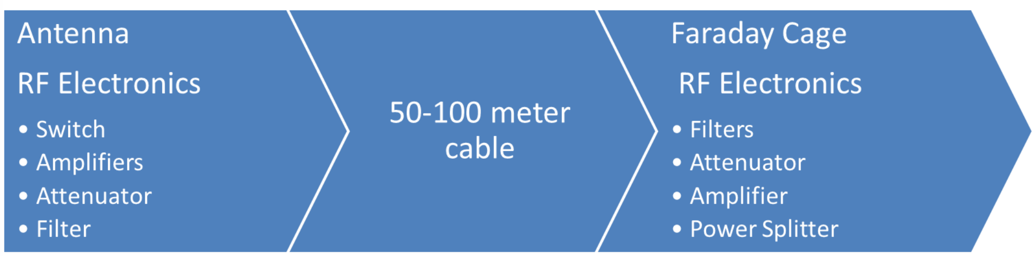


Figure 2.42: Block diagram of the RF Electronics for the SCI-HI instrument.

2.3 RF Electronics

Once the signal from the sky has been collected by the antenna, it has to be transmitted to the data processing system in a form that the system is capable of processing. This transmission happens by passing the antenna signal through a chain of radio frequency (RF) electronics between the antenna and the data processing system.

Figure 2.43: SCI-HI system with HIbiscus antenna set up on site at Green Bank in May 2013.

2.3.1 RF Electronics Chain Overview

For the SCI-HI system, the RF electronics chain components are a switch, amplifiers, high and low pass filters, attenuators and a power splitter (along with the attendant cabling). Each of these components serves a specific purpose in the system.

We divide the RF electronics chain into two sections based upon physical location, see Figure 2.42. The first section is the antenna electronics, which sit as close to the antenna terminal as physically possible. The second section is the Faraday Cage electronics, which sit next to the data processing system within a Faraday Cage. Between the two sections is a long signal cable ($\sim 50 - 100m$), allowing the antenna and data processing system to be placed far enough apart that the transmission of self-generated noise from the data processor is significantly decreased (see Figures 2.43 and 2.44).

2.3.2 Antenna Electronics

The antenna electronics are a calibration switch with known signal sources on some of the terminals, and two stages of amplification. In some versions of the electronics system, at least one of the RF filters was also placed at the antenna end of the system. The layout of the antenna electronics is shown in Figure 2.47.

Calibration Switch

Because the SCI-HI system has a single antenna with a relatively large beam, calibrating the data from the antenna can be difficult. One way to do this calibration is to use a noise source of known temperature whose signal is sent through the same RF chain as the antenna data (see Chapter 5).

Figure 2.44: SCI-HI system with HIbiscus antenna set up on site at Isla Guadalupe in June 2013

In order to facilitate this calibration, a switch was placed at the terminal of the antenna. By making the antenna one of the inputs of the switch and connecting one or more known temperature sources to the other switch inputs, data can be collected for all inputs with the same RF chain.

We used a 6 input electromechanical switch purchased on eBay (Narda SEM066 SP6T[4]) for the system. One input to the switch was connected directly to the antenna (see Figures 2.45 and 2.49). The other inputs were connected to a short terminator, a 50Ω load terminator, a 100Ω load terminator, and an artificial noise source of known power. One input was left open.

Switching between inputs is controlled through a cable from the Faraday Cage that carries both the power for the antenna electronics and the control signal. In addition, the artificial noise source is turned off when that input is not enabled (to

[4] `http://www.nardamicrowave.com/east/index.php`

Figure 2.45: Center of the Trombone antenna as set-up with the calibration switch mounted directly below the antenna.

Figure 2.46: RF electronics box at the base of the Trombone antenna, which contains the second stage amplifier and filters.

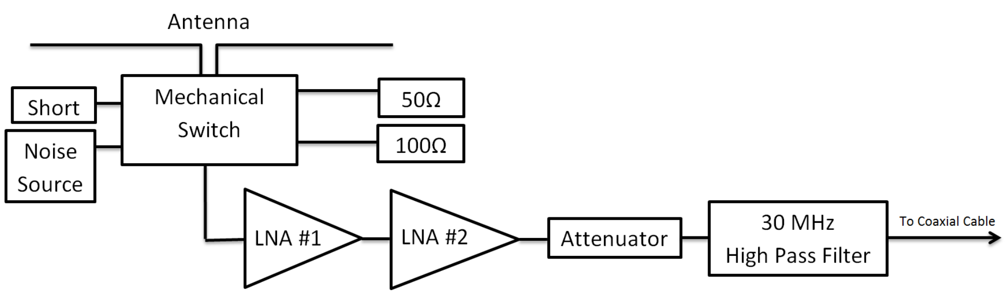


Figure 2.47: Block diagram of the RF Electronics located at the base of the antenna.

avoid noise bleeding into the antenna). We also have a manual switch control box that can be hooked up to the switch for testing and runs off a small battery pack.

Amplifiers

Immediately following the switch in the RF chain is the first stage low noise amplifier (LNA). This amplifier should have a low noise figure, as amplifier noise will be added to the incoming RF signal before amplification. In addition, the LNA should have an impedance that is well matched to the antenna. Matching is necessary for signals from the antenna to be transmitted by the LNA instead of being reflected back out the antenna.

Impedance Impedance (Z) is the complex ratio ($Z = V/I$) of a circuit element. It is measured using the circuit element reflectivity ($S11$) relative to a reference

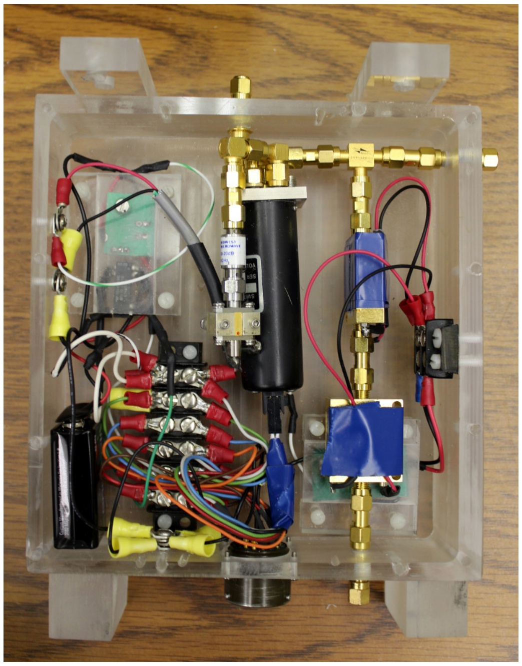


Figure 2.48: New Lucite box containing all the antenna RF electronics.

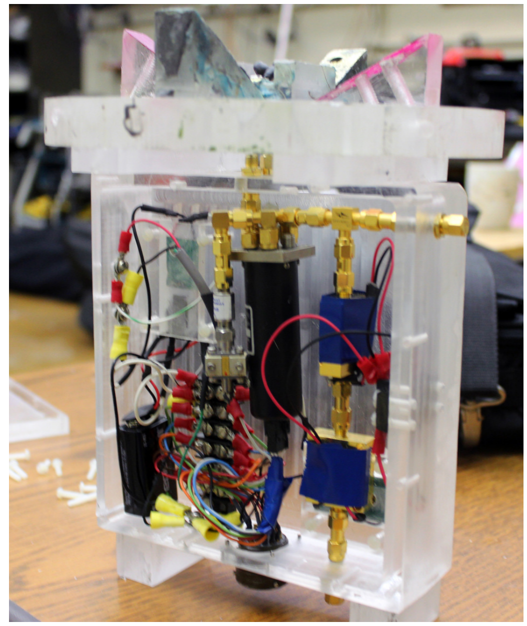


Figure 2.49: RF electronics box attached to one of the HIbiscus antenna center mounts.

impedance, which is usually 50Ω [31].

$$Z(\nu) = R(\nu) + iX(\nu) \tag{2.1}$$

$$R(\nu) = \frac{50(1 - |S11(\nu)|^2)}{(1 - \Re(S11(\nu)))^2 + \Im(S11(\nu))} \tag{2.2}$$

$$X(\nu) = \frac{100\Im(S11(\nu))}{(1 - \Re(S11(\nu)))^2 + \Im(S11(\nu))} \tag{2.3}$$

Impedance Matching The presence of reflectivity means that not all of the signal sent down a circuit will reach its final destination. Some of the signal will instead be reflected back toward its source. The exact magnitude of this reflectivity at each interface between circuit elements is set by the match between the impedances of the two elements.

Transmission Efficiency (η) If $Z_{Source} = Z^*_{Load}$, then all of the signal is transmitted. Otherwise, some of the signal power is reflected. The fraction of signal power that is transmitted is called the transmission efficiency (η). Equation 2.4 gives the relationship between the impedances of two circuit elements and the percentage of signal which is transmitted.

$$\eta(\nu) = \sqrt{\frac{4|R_S * R_L|}{(R_S + R_L)^2 + (X_S + X_L)^2}} \tag{2.4}$$

Cable Delay In our case the interface of interest is the one between the antenna and the first stage amplifier, $Z_S = Z_{ant}$ and $Z_L = Z_{amp}$. However, the antenna impedance as seen by the LNA will be modified by the cable between the antenna and LNA.

This modification will not change the magnitude of Z, but will affect its complex phase as $Z_{mod} = |Z_{orig}|e^{2\pi i\tau\nu}$. The magnitude of τ is set by the length of the cable and its propagation constant (C_{cable}), $\tau = l_{cable}C_{cable}$. The propagation constant C_{cable} is based upon the cable's index of refraction n and has a value of $C_{cable} \approx 0.125$ ns/inch for coaxial cable.

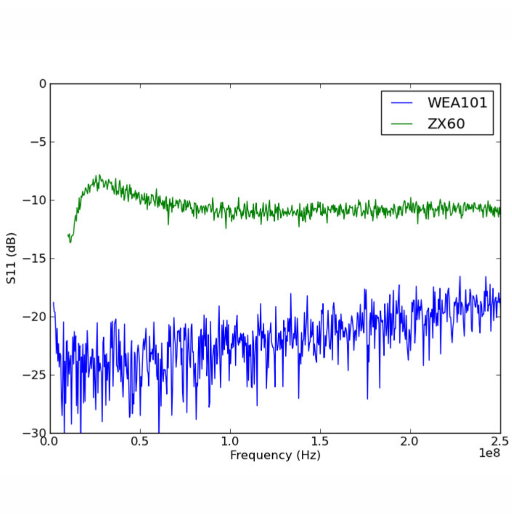


Figure 2.50: Amplifier reflectivity ($S11$) in dB for the two amplifiers used in the SCI-HI system.

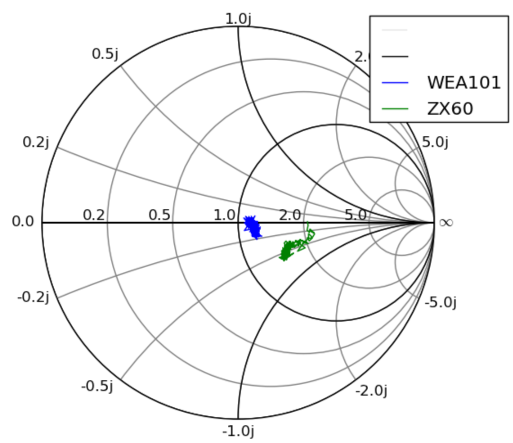


Figure 2.51: Smith chart of the reflectivity of the two amplifiers used in the SCI-HI system. Data is shown for $20 - 250MHz$.

Amplifier Selection In order to minimize loss due to impedance mismatch, we want an amplifier whose impedance is the complex conjugate of the antenna impedance. I tested a number of LNAs to find one whose impedance was matched to our antenna over the entire frequency band of interest for SCI-HI.

I started with a Minicircuits[5] packaged amplifier ($ZX60-33LN-S+$) with a low noise figure (~ 1 dB) and wide bandwidth. Unfortunately this amplifier's impedance

[5] http://www.minicircuits.com

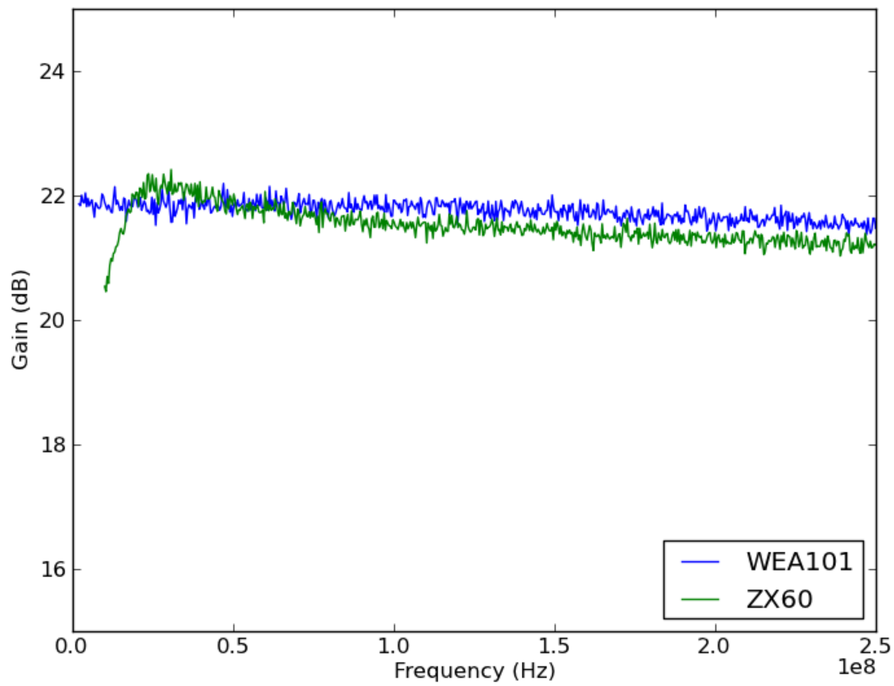


Figure 2.52: Amplifier gain ($S21$) in dB for the two amplifiers used in the SCI-HI system.

was a poor match for the antenna, but we were able to make use of the amplifier later in the RF chain (see Section 2.3.3).

Next, I tried to use the Avago Technologies[6] amplifier chip ($MGA-633P8$), which we had previously used for higher frequency applications. The benefit of using a chip versus a pre-packaged amplifier was that I could modify the circuit around the chip to tune the impedance to match the antenna. However, I was unable to tune the chip circuit to simultaneously get the desired impedance and maintain the amplifier gain in our frequency band. This was primarily due to the fact that the chip was designed for higher frequencies around $700MHz$.

Therefore, I went hunting again for a packaged amplifier that fit my specific impedance needs. Richardson RFPD[7] has a packaged low noise amplifier ($WEA-101$), whose impedance was a better match to the antenna than the Minicircuits amplifier. Figures 2.52, 2.50, and 2.51 show the $S11$ (reflectivity) and $S21$ (gain) of the two amplifiers for the SCI-HI frequency band.

[6] `http://www.avagotech.com/`

[7] `http://www.richardsonrfpd.com/`

Unlike the $ZX60$ amplifier, the $WEA101$ amplifier has an almost entirely real impedance, which is nearly constant over our frequency band. It also has a very low reflectivity. This makes the $WEA101$ amplifier an excellent choice as a first stage amplifier.

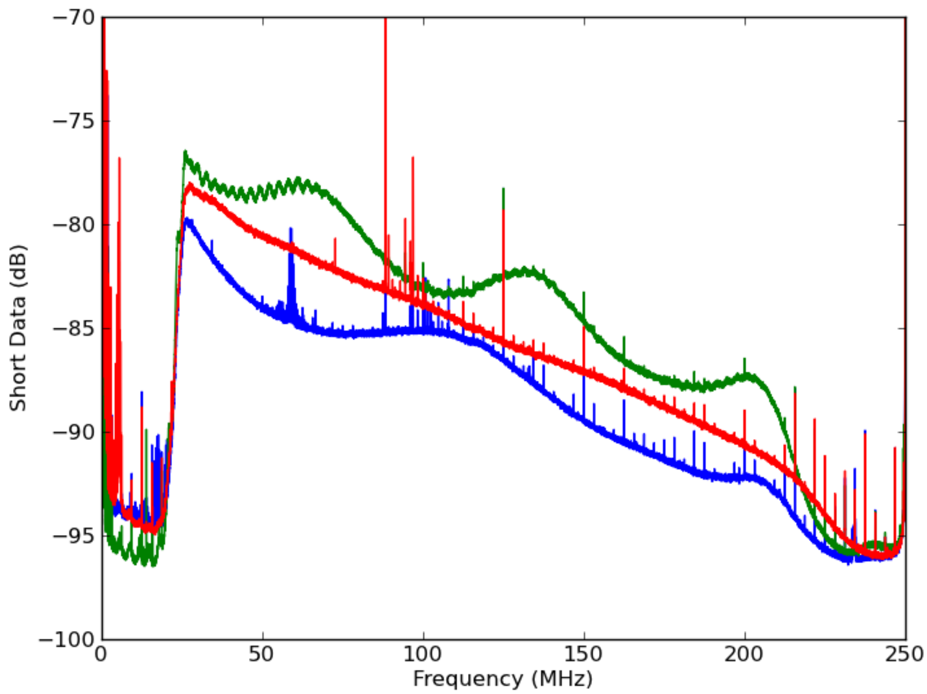


Figure 2.53: Large scale oscillation in the short calibration data due to a cable between the first and second stage amplifiers. Blue is the data with the original 1 meter cable, green is the data with a 2 meter cable, and red is the data with the amplifiers right next to each other.

Multi-Stage Amplification Because the $\sim 50 - 100m$ signal cable will attenuate signals traveling down its length, a second stage amplifier was added to the system to increase the gain of the signal. Initially the second stage amplifier was placed at the base of antenna (see Figure 2.46). However, this placement resulted in reflections in the signal whose period corresponded to the cable length between the two amplifiers. This is shown in Figure 2.53. We confirmed that the cable between the amplifiers was the source of the oscillation by replacing the cable with one twice as long. When we did this the period of the oscillation was cut in half, as is also shown in Figure 2.53.

Therefore, we moved the second stage amplifier to a position immediately following the first stage amplifier (see Figure 2.48). This removed the cable reflections. The total gain of the amplifiers was $\sim$ 50 dB over the entire frequency band. Both amplifiers use DC power supplied by the same power cable as the switch. As different levels of DC voltage are needed for the individual electronic components, voltage regulation circuits are also included in the antenna electronics housing.

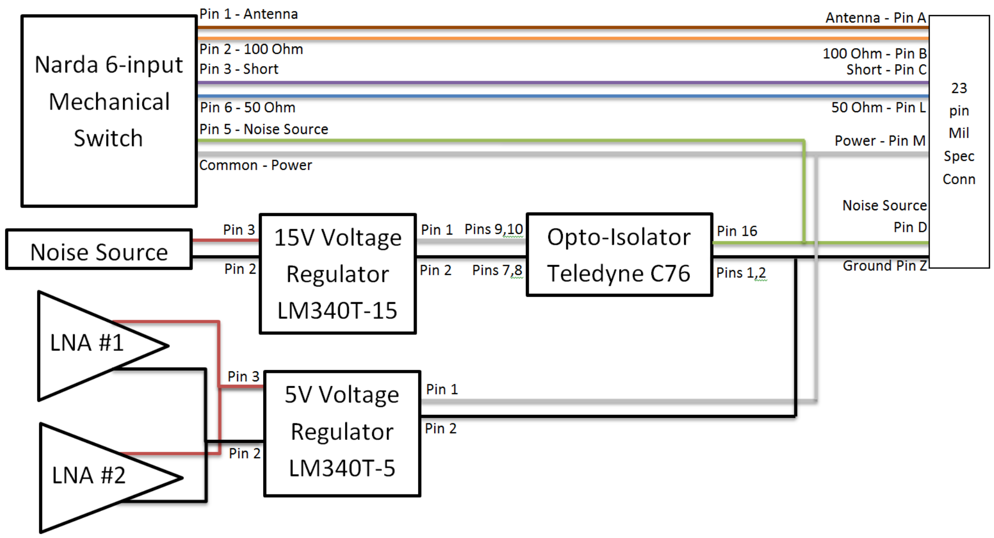


Figure 2.54: Wiring block diagram for the antenna RF electronics.

Antenna Power and Switch Control

Power and switch control signals are sent along a cable that runs parallel to the signal cable from the Faraday Cage to the base of the antenna. This cable has an Amphenol[8] 23 pin circular mil spec connectors on each end, of which 7 pins are used. One pin is ground, one is DC power, and the other five are control signals for each of the switch inputs.

Power to each of the antenna RF components is supplied following the diagram in Figure 2.54. The two amplifiers use 5 V regulated power, while the noise source has 15 V regulated power which is only supplied when the noise source is being measured through the switch. Control of the noise source power input is handled by a Teledyne[9] optically isolated relay module.

In earlier versions of the SCI-HI system, the power from the Faraday Cage was $V <$ 15V, so a 9V battery was added to the noise source circuit to supply the additional voltage. This battery was only in use when the noise source was switched on. For the current version of the system, the switch needs $V \geq 18$V. This negates the need

[8] http://www.amphenol.com/
[9] http://www.teledynerelays.com/

for a supplemental battery, as the power sent to the antenna electronics will always be $V \geq 15$V.

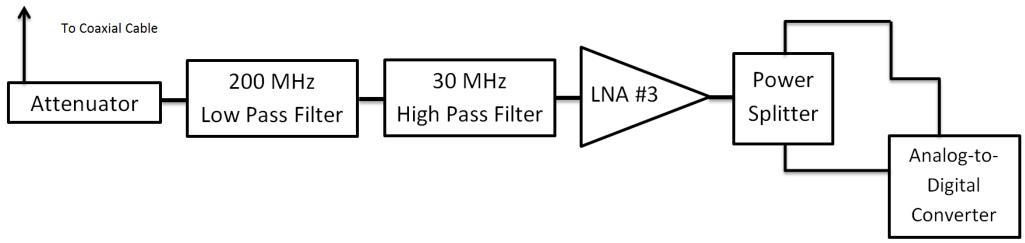


Figure 2.55: Block diagram of the RF Electronics located inside the Faraday Cage.

2.3.3 Faraday Cage Electronics

Once the signal has traveled down the long cable from the antenna, there are a few electronic stages that it must pass through prior to entering the data processing system. A third stage amplifier is needed to compensate for the attenuation caused by the long cable, filters are needed to keep signals from outside the frequency band of interest from overloading the data processing system or leaking into the band, and a signal power splitter used to achieve the necessary data sampling rate. The layout of these electronics is shown in Figure 2.55. All of the RF electronics, as well as the data processing system, were placed inside a Faraday Cage designed specifically to accommodate them.

RF Filters and Amplification

Once the signal from the antenna entered the Faraday Cage, the first part of the RF chain was a set of filters. These filters were a low pass filter at $200MHz$ (Minicircuits $SLP-200+$) and a high pass filter at $30MHz$ (Minicircuits $SHP-25+$). The low pass filter was placed to prevent signals from higher frequencies ($\geq 200MHz$) from aliasing into the frequency range where we were observing. The high pass filter was placed to prevent signals from the shortwave band and below overloading the sampling system, which could affect the overall system gain or create spurious signals in the observing band.

After the signal has passed through the long cable and the filters, it is necessary to add another level of amplification to the system. This final stage of amplification is to bring the signal levels to a point where they can be clearly distinguished from system noise by the data processing system and adds ~ 25 dB of gain. The amplifier used here is a Minicircuits amplifier ($ZX60-33LN-S+$), which is supplied with 5V regulated power from the DC source.

RF Signal Splitting

The Analog-to-Digital Conversion (ADC) card used in the data processing system is limited to a 250 MSamples-per-sec sampling rate, but the desired bandwidth of the system is over $125MHz$. In order to expand the operating range of the ADC, a hardware ″trick″ is used.

The signal from the antenna is split into two identical signals with a Minicircuits power splitter ($ZFSC-2-1$). Both signals are then sent into the ADC through two of the input ports. However, the lengths of the cables placed between the two outputs of the splitter and the two input ports of the ADC are slightly different. The difference ($\sim 50\ cm$) is exactly calculated so that a sample taken from one input port is has been delayed by 2 ns compared to the other input port. If the signals from both ports are combined, the data has now been sampled at 500 MSamples-per-sec. This interleaved sampling is common in high bandwidth ADC systems.

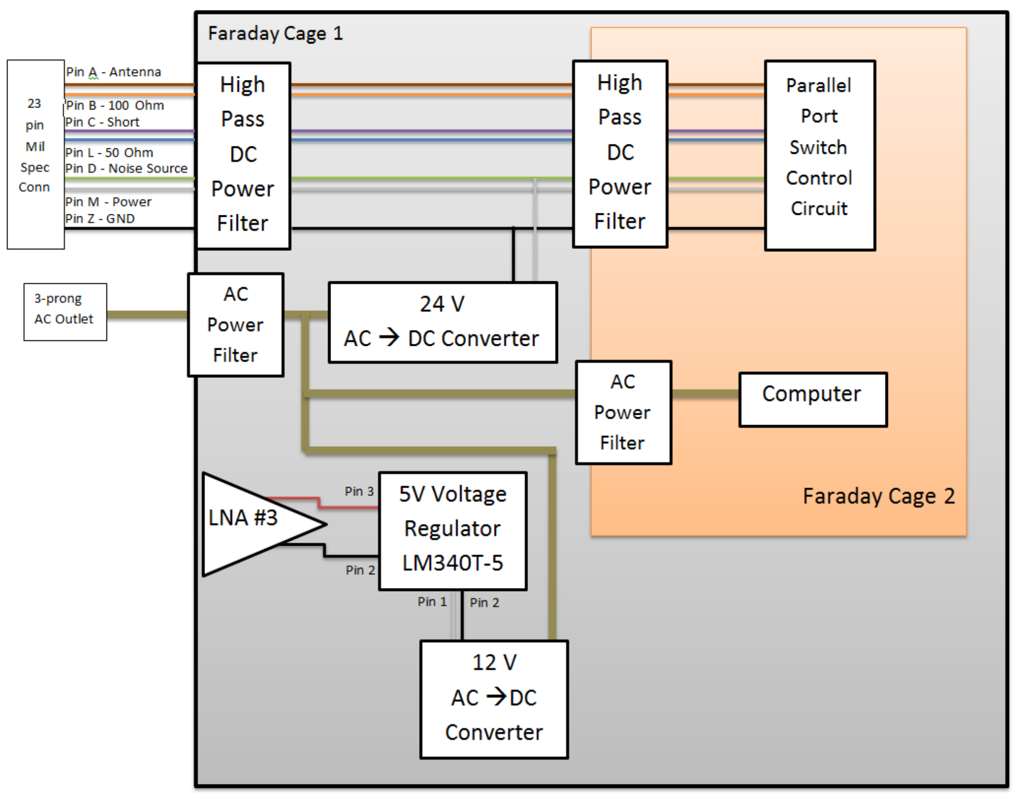


Figure 2.56: Wiring block diagram for the Faraday Cage RF electronics.

Faraday Cage RF Electronics Power Supply

Power is supplied to the RF electronics inside the Faraday Cage and at the base of the antenna through a common source (DC battery or AC generator, see Section 2.3.4). Assuming that the Faraday Cage input power is AC, the power is distributed as shown in Figure 2.56. Some of the power travels to the antenna RF electronics via the system power/control cable that runs parallel to signal cable. This power is discussed in full detail in Section 2.3.2.

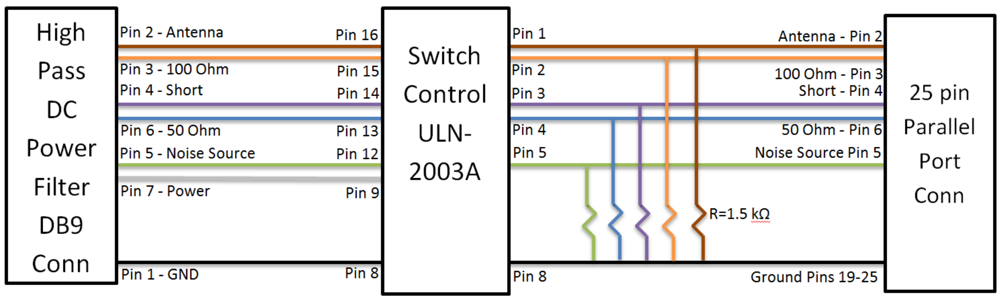


Figure 2.57: Circuit diagram for the parallel port switch control circuit.

Switch Control Circuit

In order to get the control signal for the switch from the data processing system into the RF chain, a small circuit is necessary. This circuit, shown in Figure 2.57, uses a transistor array chip ($ULN2003A$) to change the switch control voltages to the necessary levels based on the parallel port output signals. The control circuit lives inside the Faraday Cage with the computer (see Figure 2.56).

2.3.4 SCI-HI Faraday Cage and System Power

Faraday Cage Design

The design of the Faraday Cage for the SCI-HI system was inspired by the dual requirements of space and portability. Because the wavelengths being studied with the system are $\lambda \geq 1$ meters, metal mesh is used in place of solid metal for the sides of the cage. This mesh is considerably lighter than solid metal, but does not hold its structure. So I came up with a design that uses a metal mesh bag with a PVC pipe framework placed inside to hold its shape (see Figure 2.58). The mesh bag folds up like a paper bag and the PVC framework can be disassembled and packed into a suitcase.

Figure 2.58: Faraday Cage around the SCI-HI data processing system as it was set-up in October 2012.

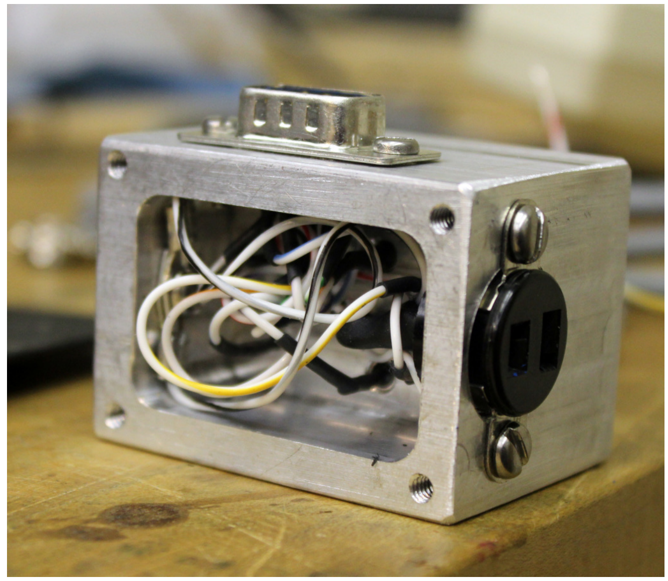


Figure 2.59: One of the power and switch control filter boxes for the SCI-HI system. Capacitive high-pass filters with $\geq$ 40 dB attenuation above $30 MHz$ are used in the box.

Access to the Faraday Cage is supplied through the front panel, which attaches to the rest of the bag through a set of metal snaps spaced less than 20 cm apart. This front panel includes access ports for the signal cable, as well as the power and switch control cable. Multiple versions of the Faraday Cage were constructed, all of which used the mesh bag concept (see Figures 2.58, 2.60, and 2.61).

Figure 2.60: Faraday Cage around the data processing system as it was set-up in June 2013.

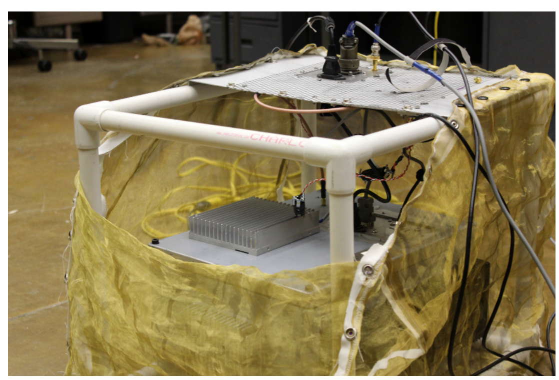


Figure 2.61: Faraday Cage with double shielding around data processing system, currently under development.

Faraday Cage Power Filters

Because the power and switch control cable carries DC power out to the antenna electronics, it needs to be filtered to keep RF signals from using the cable to escape the Faraday Cage. To accomplish this, a set of filter boxes are placed in the DC power and switch control signal lines. These filter boxes (see Figure 2.59) are low pass capacitive filters with over 40 dB of attenuation above $40MHz$. Two stages of filtering are included to ensure that there is no signal transmission through the DC power lines.

Initially the power supply battery lived inside the Faraday Cage (see Figure 2.58). Eventually, it was moved outside the cage to simplify battery swapping (see Section 2.3.4). At this point filters also had to be used between the battery and the electronics inside the cage. The same filters used for the DC power and switch control cable were also used for the battery cable. This ensures that RF signals don't use the battery cable to transmit outside the Faraday Cage.

System Power Selection

The entire system was powered using deep-cycle automotive batteries. These batteries supplied 12 V unregulated power, which was used by both the data processing system and the RF electronics. A fully charged battery could power the entire system for $\sim 12-15$ hours.

Charging batteries was a big part of the support process for the SCI-HI system. Keeping the system running continually required at least 3 batteries with regular access to a battery charger and transport between the SCI-HI site and the battery charging location 2-3 times a day. In addition, recent upgrades to the SCI-HI system include components that need DC voltages larger than 12 V. Therefore, alternative power sources have been investigated.

One alternative power sources is a small portable gasoline or diesel generator. Using this generator would cut down on the number of site visits required, as the generator would run constantly with only occasional stops to top off the fuel. The generator would also allow us to supply larger DC voltages as needed.

However, using a generator leads to a new source of RF noise (particularly a gasoline generator with spark plugs). Placing the generator inside its own Faraday Cage should attenuate this noise, but it is a factor that must be accounted for in selecting a power source.

2.4 Data Processing System

2.4.1 Data Processing Pathway

As discussed in Section 2.3.3, signals from the electronics enter the data processing system via two ports of an Analog-to-Digital Conversion (ADC) card. For the SCI-

HI system, we used a GE PCIe digitizer board ($ICS1650$). Data is collected for one second of integration, with the interleaving strategy producing 500 MSamples-per-sec of data.

Storing the entire dataset for each second would require a great deal of space. Instead, a Fast Fourier Transform (FFT) is performed on the sampled time streams, then the power signals are averaged. The averaged power spectrum produced by the FFT is much smaller than the raw data. This spectrum is stored by the system, after which the process is repeated for another second of data.

Because of the limited bandwidth from the ADC to the rest of the system, sampling of a new data stream requires the FFT of the previous data stream to be complete. This means that the system is less than 100% efficient, it takes longer than 1 second to record the power spectrum of 1 second of data. Initially the FFT calculation was quite time consuming ($\sim 30-60$ seconds), but improvements to the software and hardware by Jose-Miguel have decreased this calculation time to ~ 2 seconds. This means that the system has a duty cycle of $\sim 30\%$.

All of the frequency spectrum data is stored locally on the hard drive of the processing system. The relatively small data volume (~ 2 GB per day) means that data transfer can be done using small USB drives. During deployment, data was removed from the hard drive 2-3 times a day and stored locally on multiple laptop computers and external hard drives before being uploaded to a server upon return to "civilization" (aka the lab).

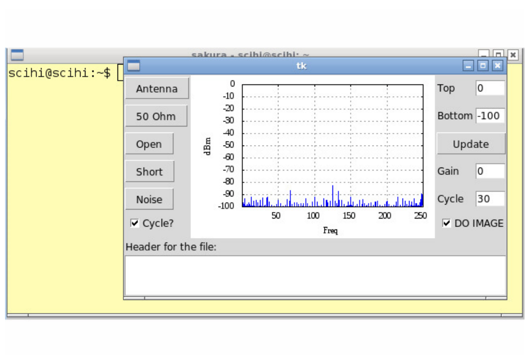


Figure 2.62: User interface for the SCI-HI data processing system. Data shown in the interface was collected when the data processing system was not connected to the antenna.

Figure 2.63: Raspberry Pi control computer, which connects to the data processing system via ethernet and is used for field monitoring.

2.4.2 System Control and User Interface

In addition to the data processing software, a graphical user interface was designed by Jose-Miguel for diagnostic and control purposes. This interface displays the most recent frequency spectrum collected by the system and lets the user set the current data collection mode (including switch control). The user can choose to either take a single data set with any of the switch positions (Antenna, 50 Ω, Short, Noise Source, 100 Ω), or take continual data with a set number of antenna datasets followed by a single dataset from each of the calibration sources. Figure 2.62 shows an example screenshot of the user interface. In this screenshot, the data is ~ 0 because the electronics and antenna are not connected to the system.

The system default upon start-up is to run in continual data collection mode with the number of antenna iterations per cycle set by the previous data collection run. To lower the power consumption and space requirements of the system, the software can run without a monitor, mouse and keyboard. Instead, the system can be controlled by an external system through an ethernet port. This port is only used during system monitoring, and can be sealed to prevent RF leakage at all other times.

2.4.3 Control Computer

The data processing system can be monitored using a personal computer (such as a laptop) with an ethernet port. However, an additional diagnostic computer was also developed by Jose-Miguel for the SCI-HI system. This computer is a Raspberry Pi[10] with a small monitor, mouse and ethernet port, built by Jose-Miguel and shown in Figure 2.63. Using this small computer, we can quickly check that the system is working properly and run simple diagnostics.

2.4.4 Power Supply and Consumption

In order to power the data processing system, a computer power supply had to be selected that matched the system power source (DC battery). Initially, we started with a typical AC computer power supply and a DC to AC inverter that converted the incoming DC power from batteries into AC power that the computer could handle.

While using an inverter is the simplest solution, it can be very energy inefficient. For example, the 800 W inverter that we have previously used in the field is $\sim$ 70% efficient. Additionally, any failure of the inverter can crash the entire system. Therefore, we decided to switch to an entirely DC system by replacing the AC power supply for the computer with a M4-ATX DC power supply. This power supply takes input DC power of 8-30 V and converts it to the DC power that the motherboard and other systems need.

Utilizing this power supply lowered our power consumption. However, we found that the power supply we selected was not reliable on long time scales. Since most of the DC power supplies on the market are designed for automotive applications, they are designed for reliability when run for short durations on higher battery (voltage) levels. The DC power supplies were particularly unhappy when running on a nearly

[10] www.raspberrypi.org

drained battery, which had a voltage level close to the minimum required voltage for the power supply. We are currently exploring alternatives for the system, as discussed in Section 2.3.4.

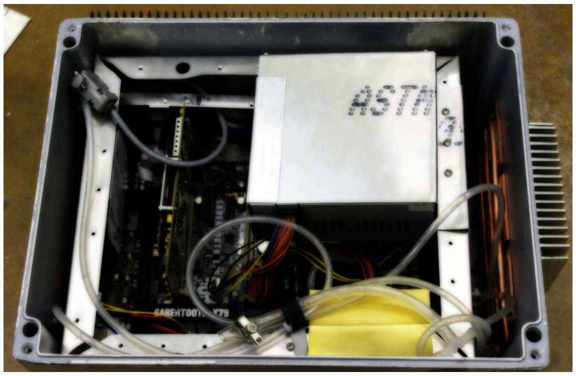


Figure 2.64: Most recent update to the data processing system. System is assembled inside a robust Faraday Cage (shown without lid in place).

Figure 2.65: Copper tubing with brass mounting used as part of the data processing system and Faraday Cage's water cooling system.

2.4.5 System Noise Generation

Deployment to Isla Guadalupe in June 2013 indicated that our Faraday Cage design had insufficient attenuation of self-generated RFI from the data processing system. This problem was not identified until deployment at the final site due to masking from external RFI at our testing sites.

Self-generated RFI was particularly noticeable above $90MHz$, as can be seen in Figure 5.1. It is also believed to be present in the data at lower frequencies as well. Our data indicated that the strength of the self-generated RFI was correlated with the voltage (or charge level) of the DC batteries.

In order to address this problem, I designed a more robust Faraday Cage for the SCI-HI system. This Faraday Cage is a double layer cage (see Figure 2.64) with the data processing system placed inside a solid aluminum sealed box and a second Faraday cage with the rest of the electronics placed around the inner box.

Using a set of two walkie-talkies, the aluminum box was tested to show attenuation $\geq$70 dB before modification. This was measured by placing one walkie talkie inside the box in receive mode and transmitting a signal with the other walkie talkie. If the attenuation was smaller than the transmitted signal, then the receiving signal would make a noise. Even transmitting from less than 1 meter from the box, the signal was not picked up by the receiver.

However, the problem with using a solid metal Faraday Cage is that it becomes very difficult to cool the system. we designed a water cooling system using copper tubing mounted to the inside of the box (see Figure 2.65) and heat fins mounted to the outside of the box. A fan with its own Faraday cage was also mounted to the outside of the box to aid heat transfer. This system has been found sufficient for cooling in lab tests, but has not yet been tested in the field.

2.5 Summary

The SCI-HI system has been developed to meet the constraints set in Section 2.1 and features the optimized HIbiscus antenna, a compact RF electronic chain, and a robust, low-power data processing system.

Chapter 3

Radio Frequency Interference (RFI) and Site Testing

3.1 Overview

One of the challenges of radio astronomy is locating observing sites that meet the environmental, atmospheric, and ionospheric requirements of a particular set of observations. Potential sites must be assessed for their viability prior to significant experiment development and observation at those sites. Site assessment must include four elements, each of which must be addressed to evaluate the site quality.

First, does the site have any nearby man-made sources of time independent or continual radio frequency interference (RFI) in the frequency band of the observations? Initial evaluation of this requirement can be done using a simple broadband antenna and spectrum analyzer on site at multiple locations. Deployment of an identical system at different sites allows a much more reliable comparison than can be made with independent systems.

Second, is the site logistically accessible for the type of equipment needed for a set of observations? Some important considerations include the availability of power, roads or other transport into/out of the location, housing and other observer requirements and site access permissions (such as permits). Assessment of these considerations often requires an in-person visit and an ongoing relationship with the agencies controlling the site.

Third, are there atmospheric effects that must be considered in assessing the viability of a site (e.g. contributions of meteor scatter, thickness of ionosphere, inclement weather)? Tracking data may be available from external sources such weather surveys, but it is often limited to broad trends instead of local details.

Fourth, are there sources of time variable RFI visible from the site? This can be more difficult to evaluate as it requires a long period of data collection. In cases where time-variable RFI is expected to play a significant role in the data collection, semi-permanent systems may need to be installed to track the RFI environment over time.

In the following sections, I will evaluate both existing telescope sites and new sites based upon the requirements listed above. Using data collected in Pittsburgh, Pennsylvania, which as a major metropolitan area is not expected to be a suitable radio astronomy site, I will demonstrate the site testing system. I will then look at several existing radio telescope sites and examine their strengths and weaknesses. Finally, I will report on several new radio sites, comparing them to the existing sites to demonstrate viability.

Figure 3.1: Site testing kit laid out in the lab (portable Spectrum Analyzer in inset).

Figure 3.2: Tabitha collecting RFI data with the site testing equipment while at one of the testing sites.

3.1.1 Site Testing Kit

One major element of site evaluations was a set of single time sweep measurements of the RFI over a wide frequency band at each site. To do this, a site testing kit was assembled. The kit includes a broadband antenna, amplifiers, and a portable Spectrum Analyzer for data collection.

RFI signals are received by the antenna and passed along the RF electronics into the Spectrum Analyzer. RFI is first received by the antenna, a Workman $T-601$ discone antenna with a vertical polarization and a 3 m PVC pipe mast (see Figure

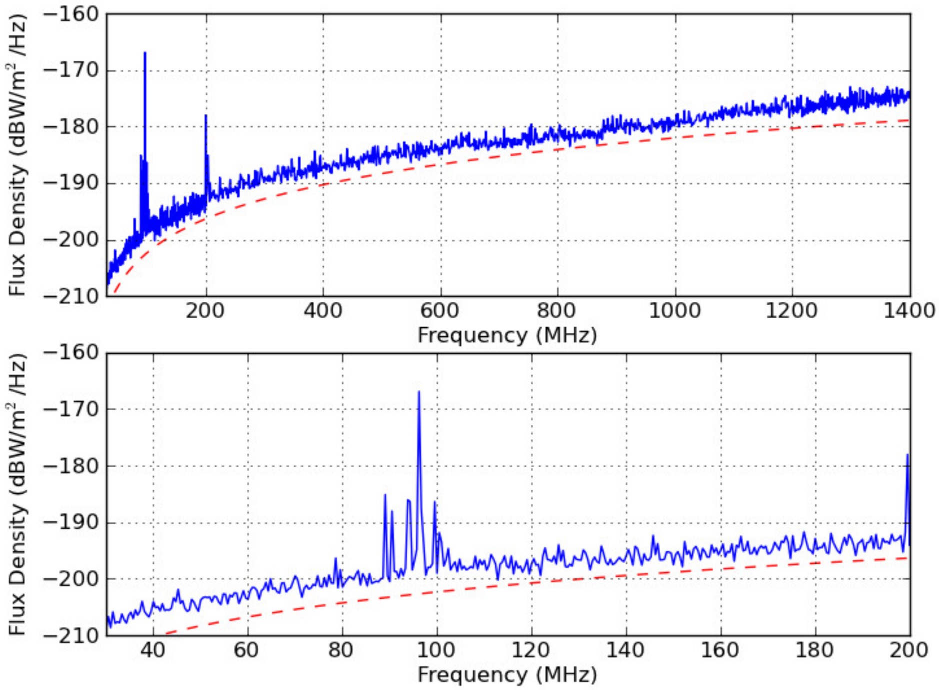


Figure 3.3: Resistor measurement with the site testing kit. Blue line shows measured power from a 50Ω resistor. In the data, FM spikes are RF leakage due to the measurement being made in the lab in Pittsburgh where the FM band is extremely loud. Meanwhile, the red line is the expected signal level of a 50Ω resistor based on Equation 3.3. The difference between the expected signal and actual signal can be attributed to limitations of the Spectrum Analyzer in measuring flux density.

3.2). From the antenna, the received signal is then sent through a 50 cm cable to a set of amplifiers powered by a DC battery pack. The amplifiers are Minicircuits $ZX60-33LN$ followed by Minicircuits $ZX60-4016E$. The signal is then sent down a $\sim 7\ m$ cable to the Spectrum Analyzer. The Spectrum Analyzer (Anritsu $MS2711A$) was enclosed in a brass mesh bag (Faraday Cage) to minimize self-generated RF contamination.

Data is collected by sweeping individual $200MHz$ bands from $\sim 1MHz$ to $1600MHz$, with a video and resolution bandwidth of $100kHz$. Only 400 data points are saved by the Spectrum Analyzer, requiring the data to be rebinned down to $500kHz$ resolution. The highest peak of the ~ 50 data points within each $500kHz$ band is stored, which can be accounted for by changing the resolution bandwidth to an effective bandwidth of $30kHz$.

In order to convert the data to flux density (S in $dB\ W/m^2Hz$), Equation 3.1

was applied to the data. In the equation, G is the amplifier gain in dB, BW is the effective spectrum analyzer bandwidth in Hz and G_{ant} is the antenna gain in dB.

$$S\left(\frac{dB\ W}{m^2 Hz}\right) = P_{meas}(dBm) - 30\left(\frac{dB\ W}{dB\ m}\right) - G - 10log_{10}[BW] - G_{ant} \tag{3.1}$$

Amplifier gain (G) was measured in the lab using a noise figure meter and antenna gain (G_{ant}) can be calculated using Equation 3.2, where ν is the frequency in Hz, and A_g is the isotropic antenna gain in m^2. We used a value for A_g corresponding to the typical value for our antenna, which is $1dB\ i$ or $10^{1dB\ i/10}m^2$.

$$G_{ant} = 10log_{10}\left[\left(\frac{c}{\nu}\right)^2\left(\frac{A_g}{4\pi}\right)\right] \tag{3.2}$$

To test the conversion equation, we replaced the antenna with a 50Ω resistor. That resistor should have a flux density which can be calculated using Equation 3.3, where k is the Boltzmann Constant, T is the room temperature in Kelvin and NF is the amplifier noise as measured using a noise figure meter.

$$S\left(\frac{dB\ W}{m^2 Hz}\right) = 10log_{10}[kT] + NF \tag{3.3}$$

The results of this measurement are shown in Figure 3.3. The measured signal was slightly higher than expected signal. Some of this difference was tied to the effective Spectrum Analyzer bandwidth ($30kHz$), which is larger than the $10kHz$ bandwidth quoted in the analyzer specifications from the manufacturer. The remainder of the difference may come from additional noise in the system or uncertainty in the noise figure meter values, but it is roughly flat in frequency.

Because we use the same kit at all the sites, any systematic noise contribution such as the uncertainty in the 50Ω data found in Figure 3.3 is constant across all datasets.

3.2 Existing Site Evaluations

Using the site testing kit discussed in Section 3.1.1, data was taken at each of the sites shown in Figure 3.4. The portability of the kit (it packs up into the small suitcase and poster tube shown in Figure 3.1) allows it to be easily transported to each of the sites.

3.2.1 Carnegie Mellon University Pittsburgh, PA, USA

Carnegie Mellon University is located in the city of Pittsburgh (40°26′30″ N, 80°00′00″ W), home to several universities and possessing a population of over 300,000. As

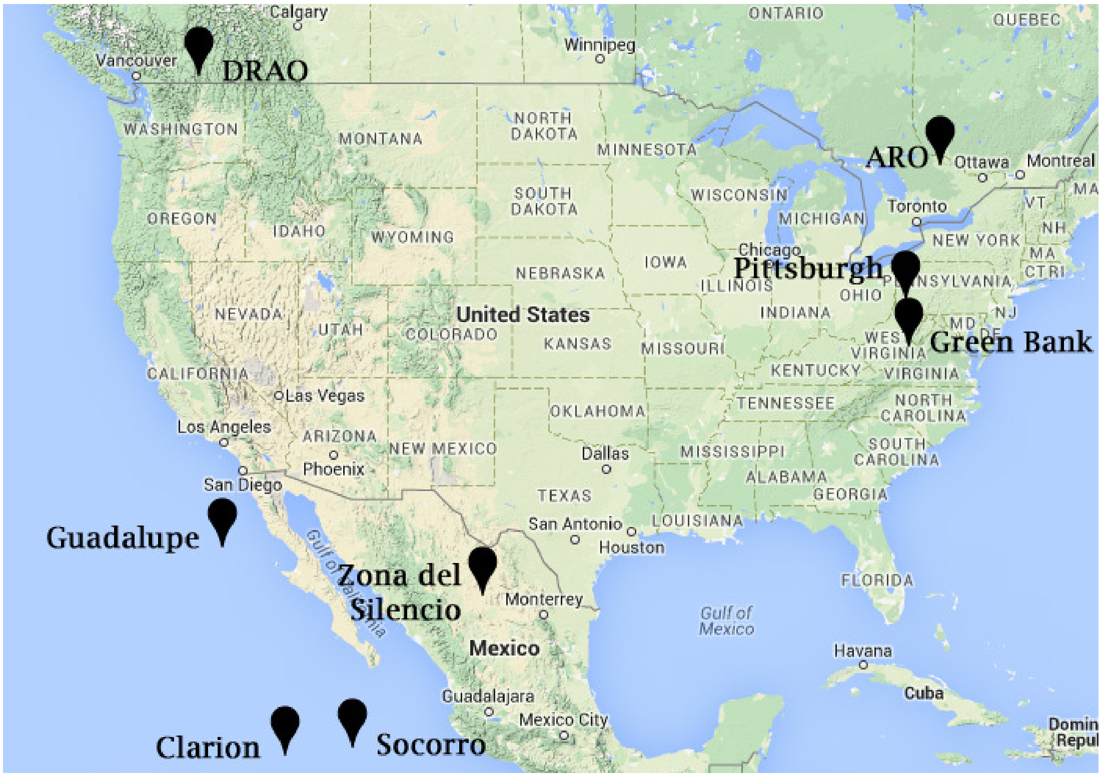


Figure 3.4: Map of evaluated sites in North America, created using Google Maps.

should be expected, the radio environment in Pittsburgh is full of RFI with signals of such magnitude that they overload test equipment. Figure 3.5 shows the RFI environment in Pittsburgh as measured with the site testing kit. The RFI signals are so loud in Pittsburgh that they overload the Spectrum Analyzer.

To try and lower the RFI levels to measurable levels, we removed one stage of amplification from the system for the low frequency data. This is why the noise floor in Figure 3.5 is lower at low frequencies than it is in the other datasets. Amplification removal was designed to minimize the inter-modulation which begins to occur when there are RFI signals larger than $-170\frac{dB\,W}{m^2 Hz}$. However, even with the amplifier removed from the system some of the RFI signals were still above the inter-modulation cutoff. For example, strong FM signals cause inter-modulation below $88MHz$ and above $108MHz$ in the Pittsburgh data.

3.2.2 National Radio Astronomy Observatory (NRAO) Green Bank

The National Radio Astronomy Observatory (NRAO) is a research center funded by the U.S. National Science Foundation (NSF). It maintains several telescopes in radio

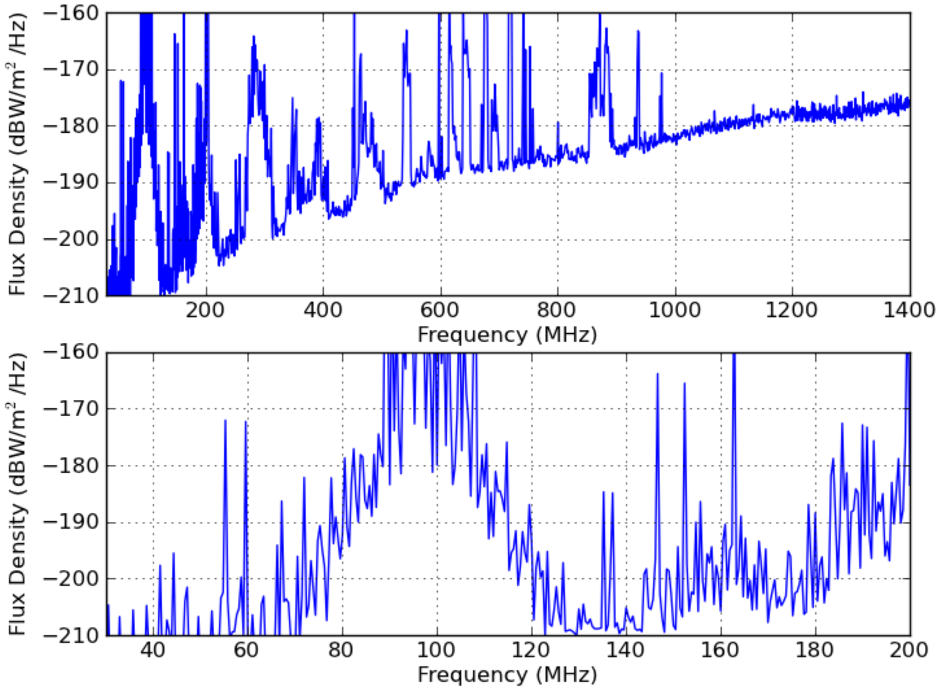


Figure 3.5: RFI measurement at Carnegie Mellon University in Pittsburgh, Pennsylvania collected April 23rd, 2009. In the data, there are many RFI sources that produce signals that have a magnitude above the plotted range. These strong RFI signals produce inter-modulations that corrupt data outside of the actual frequencies of the RFI. For example, strong FM signals cause inter-modulation below $88MHz$ and above $108MHz$.

quiet locations. One of these locations is in Green Bank, West Virginia inside the U.S. National Radio Quiet Zone in Virginia and West Virginia (see Figure 3.6). Within this zone, radio broadcasts at all frequencies are minimized by law. This provides a relatively quiet RFI environment.

The Green Bank site (38°25′59″ N, 79°50′23″ W) hosts a number of radio telescopes including the 100 m Robert C. Byrd Green Bank Radio Telescope (see Figure 3.7). As an NRAO facility, the site has a full staff and facilities including housing, power, and other amenities. Located off a local highway, the site is a 4-5 hour drive from Pittsburgh.

Looking at the data from the site testing kit, we find that the size of the radio quiet zone (about 34,000 km^2) is sufficient for higher frequencies but is too small at lower frequencies. As you can see in Figure 3.8, there are specific bands of frequencies below $600MHz$ where there is still a great deal of RFI. Some of these bands are indicated

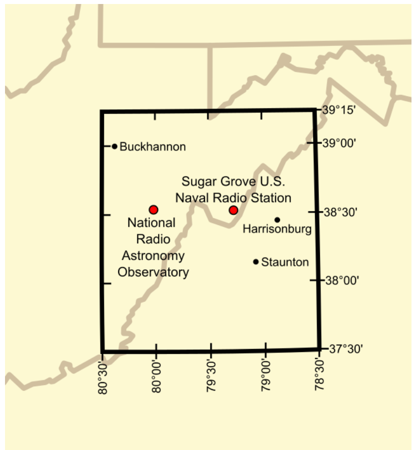


Figure 3.6: Extent of the U.S. National Radio Quiet Zone around the Green Bank Site.

Figure 3.7: Robert C. Byrd Green Bank Radio Telescope as viewed from the site observation deck.

with colored boxes (red is the FM band, magenta is an Orbcomm satellite band and cyan is a military satellite band).

Beyond the RFI, one of the other features of the site test data is a large scale sinusoidal variation (or ringing) in frequency. This variation is caused by reflections in the 50 cm cable between the antenna and first stage amplifier. One of the useful things about this ringing is that it provides a check that the system is working, because if one or more of the amplifiers has blown there is no ringing present in the data. Such a cross check is particularly useful in radio quiet areas, where there may not be much of a difference in the spectrum if the amplifier is malfunctioning.

3.2.3 Dominion Radio Astrophysical Observatory (DRAO)

The Dominion Radio Astrophysical Observatory (DRAO) is a Canadian radio astronomy site (49°19′15.6″ N, 119°37′26.4″ W). Located near Penticton, British Columbia in the south-central part of the province, DRAO has a number of radio telescopes on site and the Canadian Hydrogen Intensity Mapping Experiment (CHIME) system is currently being built there. Figure 3.9 shows some of the facilities at DRAO including part of the CHIME pathfinder in the foreground.

The site can be easily accessed by car, with on-site facilities available for researchers. It is also a ∼30 minute drive from Penticton, a town with a population of over 30,000.

Like the NRAO Green Bank site, the DRAO site has insufficient isolation from civilization to provide a radio quiet environment at the lower frequencies. There is significant RFI contamination for most frequencies below 500MHz at this site (see Figure 3.10). Given DRAO's location close to Penticton, some of the signals in the FM

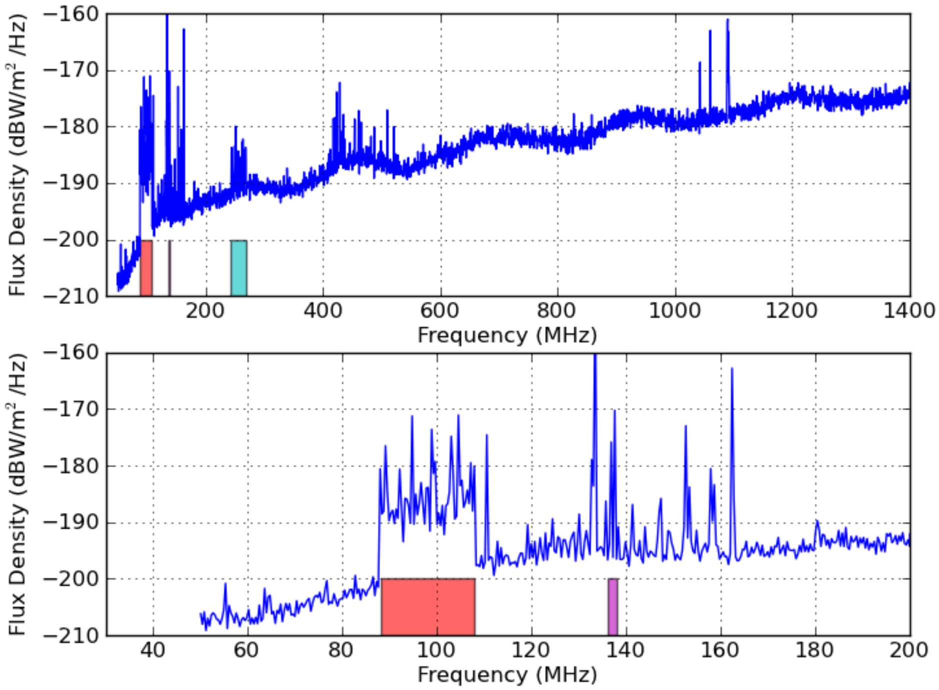


Figure 3.8: RFI measurement at the NRAO Green Bank site collected on May 18th, 2010. In the plot, colored boxes indicate bands of well known RFI sources. Red is the FM band, Purple is the Orbcomm satellite band, and Cyan is the military satellite band. Additional spikes in the data come from RFI sources not in the indicated bands. Large scale sinusoidal variation (or ringing) in the spectrum is caused by cable reflections in the system and is common to all data collected with the site testing kit.

band are loud enough to cause intermodulation of signals into some of the frequencies that are actually clean. This can be seen in the band between $108 - 136 MHz$, the dedicated aerospace band, which should be clear of RFI everywhere in the world.

3.2.4 Algonquin Radio Observatory (ARO)

The Algonquin Radio Observatory (ARO) is a single instrument Canadian radio astronomy site (45°57′19.8″ N, 78°4′23″ W). Located in the center of Algonquin Provincial Park in Ontario, Canada; the site is only accessible by logging roads, which are gravel but can be driven by most vehicles during the summer. ARO is also accessible in an emergency via helicopter, but this mode of transportation is not commonly used. ARO does have power and full amenities on site including housing for researchers.

Figure 3.9: Some of the DRAO facilities and telescopes (CHIME pathfinder in foreground).

Although the site has excellent radio quiet properties at higher frequencies, it is still relatively close (about $200 - 250$ km) to the major Canadian metropolitan areas of Toronto and Ottawa. When we set up our site test at ARO, we found that although the rest of the spectrum is quite clean there is still significant RFI below $300MHz$ (see Figure 3.11), including in the FM radio band.

3.3 New Site Evaluations

Evaluation of these existing radio quiet sites demonstrated a clear need for sites whose RFI environments are cleaner at frequencies below $500MHz$. However, locating such sites can be difficult as the distances required begin to grow large. Here, I report on a couple of potential sites in Mexico that have significant improvement in their low frequency RFI strength compared to the existing sites.

3.3.1 La Zona del Silencio

″Zona del Silencio″ ($26°41'10.3''$ N, $103°44'50.9''$ W) is a radio quiet region in the Mapimi part of the Chihuahuan desert in Northern Mexico. It has a reputation as a mysterious radio quiet zone due to historic events similar to the Bermuda Triangle,

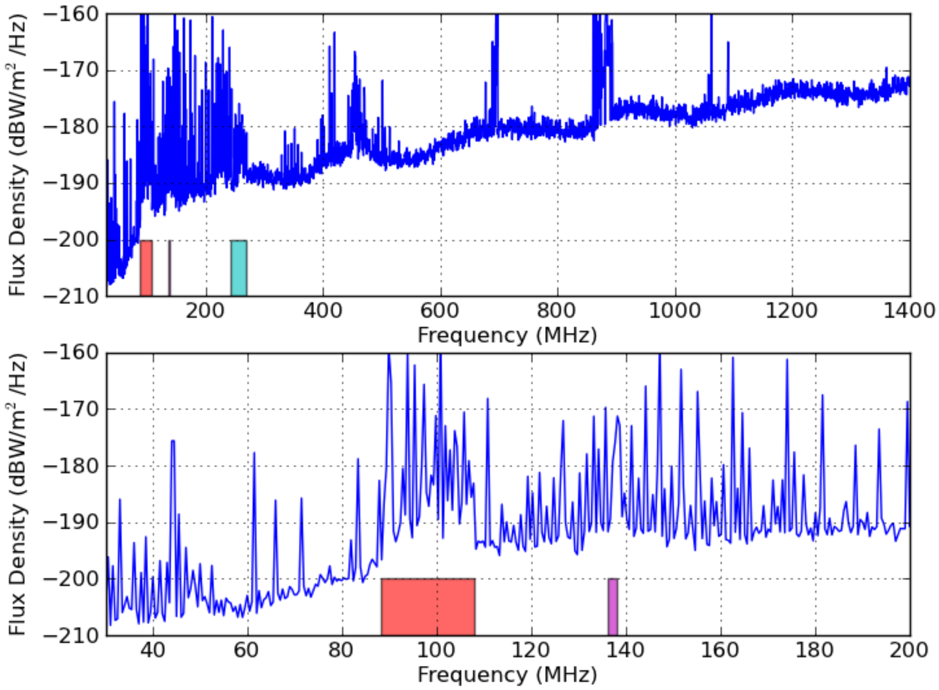


Figure 3.10: RFI measurement at the DRAO Penticton site collected on December 14th, 2009. DRAO has excessive RFI in the entire spectrum below $400MHz$ and is generally noisier than Green Bank. Because some of these signals (particularly the FM band) are above the inter-modulation cutoff, some of the smaller spikes in known clear bands are believed to be caused by inter-modulation. In the plot, colored boxes again indicate bands of well known RFI sources (Red= FM band, Purple = Orbcomm satellite band, and Cyan = military satellite band).

but in a desert setting. Some of its radio quiet status may be due to the local geography as the desert is a flat plateau, but there are mountains between it and any significant civilization. The closest major metropolitan area is Torreón, over $150km$ south of the zone center. The zone is also a protected biosphere reserve maintained by Mexico's ″Comisión Nacional de Áreas Naturales Protegidas″ (CONANP).

Logistics and Current Infrastructure

While major highways can be found along the outside of the region, the only roads in and out of the site are poorly maintained dirt roads that require 4-wheel drive.

Permanent settlements are not allowed within the biosphere reserve. However, at the center of the site is a camp for ecologists studying the biosphere found in the

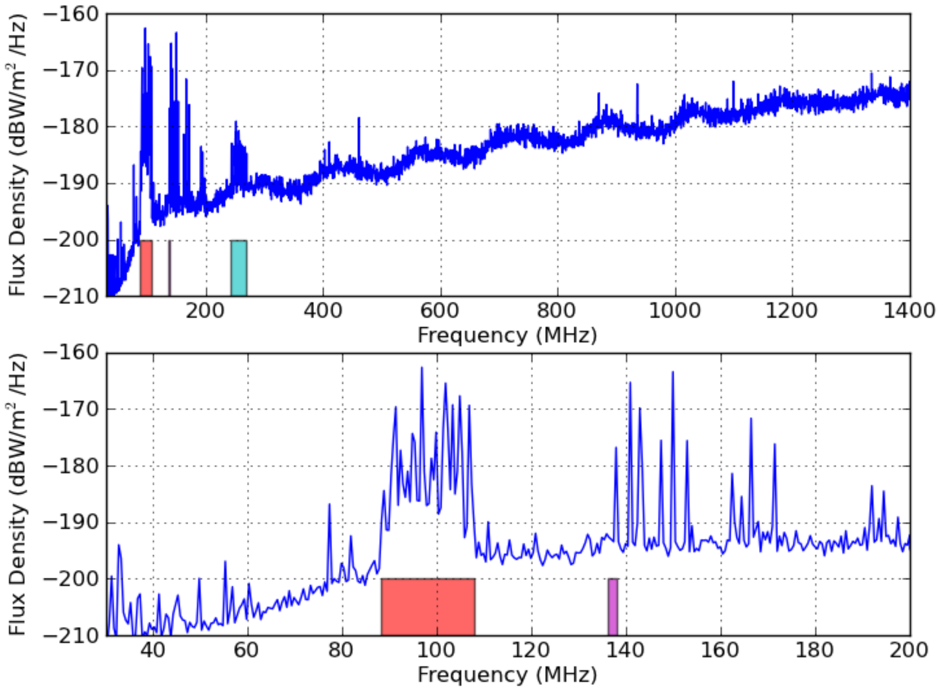


Figure 3.11: RFI measurement at the ARO Algonquin site collected on September 12th, 2012. Like the GBT site, the ARO site is quiet at high frequencies compared to DRAO or Pittsburgh. However, significant RFI is still present below $300 MHz$. In the plot, colored boxes again indicate bands of well known RFI sources (Red= FM band, Purple = Orbcomm satellite band, and Cyan = military satellite band).

reserve. The site has minimal housing with solar cells charging batteries for power but all water on site has to be brought in from outside.

For our site test, we got special permission to stay at the ecologist's camp called ″Laboratorio del Desierto″, shown in Figure 3.12.

Environmental Impacts

During the site testing, we were able to observe the general climate of the site as a consideration for future deployments. One of the challenges we observed was the prevalence of dust due to the arid climate. All our equipment had to be well protected from dust, especially during transport to and from the site. If not properly shielded, dust can cause electronic equipment to malfunction.

Another potential challenge is the temperature variation associated with the climate. Even within a single day we saw a strong difference between night and day

Figure 3.12: View of the Zona del Silencio from one of the nearby peaks. The ecologist's camp where we stayed while running our tests is in the center of the picture.

temperatures. Temperature variation can create variance in collected data from a telescope, making the desired signals more difficult to detect.

Measurements

On this first site test, we chose to measure the site quality at a specific location near the center of the zone where RFI was expected to be at a minimum. Setting up near the ecology camp but far enough away to prevent contamination by the local electronics, we measured extremely low RFI levels as is shown in Figure 3.13.

There is still some noise at lower frequencies, but the FM band is considerably quieter than at any of the existing radio quiet sites that we had tested. Further testing at this site should include a survey of a wide range of locations within the ″Zona del Silencio″ to map out the region's radio quiet properties.

3.3.2 Isla Guadalupe

Isla Guadalupe ($29°1'51''$ N, $118°16'48''$ W) is a small volcanic island located about 250 km west of Baja California in Mexico. The island has an area of $\sim$250 km^2, with two significant peaks along the north-south axis of the island. A biosphere reserve, access to Guadalupe is limited to a few groups. Namely, the Mexican government and Navy (″Secretaría de Gobernación″ and ″Secretaría de Marina″), ecologists studying

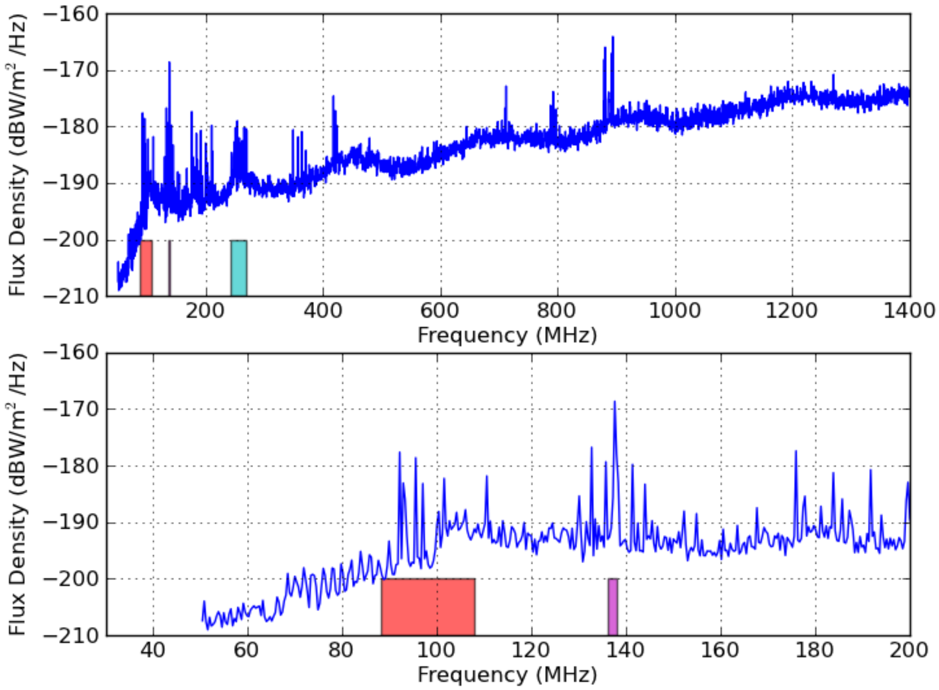


Figure 3.13: RFI measurement near the ecologist's camp in the Zona del Silencio, collected on May 5th, 2010. The Zona del Silencio site has even less RFI than the previous sites at low frequencies, although some RFI spikes are still present. In the plot, colored boxes again indicate bands of well known RFI sources (Red= FM band, Purple = Orbcomm satellite band, and Cyan = military satellite band).

the land and marine life such as the ″Grupo de Ecología y Conservación de Islas A.C.″ (GECI) and CONANP, and the local fishing cooperative (″Sociedad Cooperativa d Producción Pesquera de Participación Estatal Abuloneros y Langosteros, S.C.L.″). We were able to travel to Isla Guadalupe with support from these organizations.

Logistics and Current Infrastructure

Access to Isla Guadalupe requires one of two transport methods. First, small planes such as the one shown in Figure 3.16 can fly from the city of Ensenada in Baja California to the island, where there is a small landing strip. This flight takes 1-2 hours and can only be made during good weather. On several of our visits to Isla Guadalupe, this was our method of transport. However, each flight costs about $2000 and can only carry ∼600 kg including both people and supplies.

A much cheaper alternative is transport with the supply ship that the Mexican

Figure 3.14: Map of Isla Guadalupe showing the sites of interest, made using Google Maps.

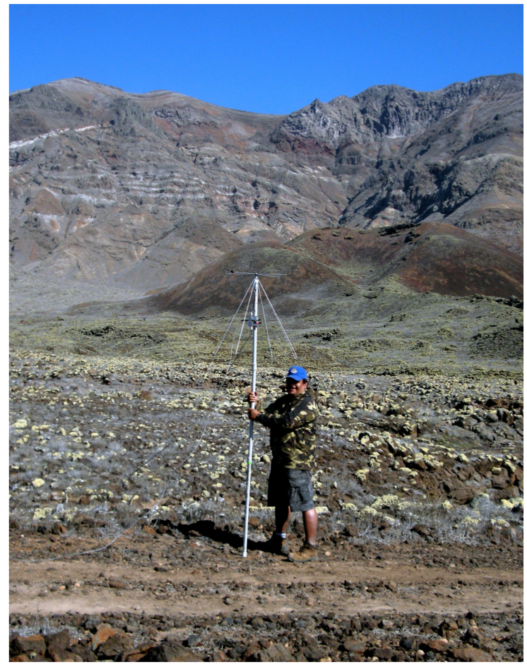


Figure 3.15: Collecting data with the site testing equipment on Isla Guadalupe.

Figure 3.16: Airplane used for access to Isla Guadalupe.

Figure 3.17: View of the plateau with the fishing village from the airplane.

Navy uses to support its base on Guadalupe. This supply ship, shown in Figures 3.18 and 3.19, deploys once a month from the port of Ensenada; stopping first at Isla Guadalupe, then Isla Cedros, then returning to Ensenada with a total travel time of about three days. Travel via this route requires passengers to ″camp out″ by sleeping on the deck of the ship during transit. In addition, since passengers are hitching a ride with the ship they have no control over changes in ship deployment (eg delays or

Figure 3.18: Mexican naval vessel as it arrived at Isla Guadalupe to deliver supplies and pick us up.

Figure 3.19: View from on board the Mexican naval vessel during the trip back to the Port of Ensenada.

re-routing) that may change the departure or arrival schedule. This places strictures on any deployments to Isla Guadalupe that must be accounted for in planning the trip. In comparison to the flight costs, passage on the supply ship is only $50 per person (for food) with minimal weight limits as some of the other passengers have been known to even transport cars via the supply ship.

While on Guadalupe we had the option of staying with the ecologists, the fishing village or the navy base. Each provided some level of logistical support, but all had limited resources.

The ecology camp is a small camp with about 5-15 researchers in residence at any given time. Housing, including running water, is available at the site for the researchers and their visitors but there is little to no plumbing, so the bathroom is a dry toilet. Since power is supplied by solar panels and batteries, it is pretty limited, especially at night. However, during the day there is regular internet access via satellite. Since it is a small camp, food is served communally, with everyone taking turns for cooking and cleaning responsibilities.

In contrast, the fishing village has a semi-permanent population of about 100 people and can be seen in Figure 3.17. The fishing village is a cooperative, which means that resources and profits are shared within the community. Houses are assigned to individual families who are currently living on site, as opposed to their main homes on the mainland. Furnishings in the houses are haphazard since all of the furnishings had to be brought in on boats. There is a sewage plant for the village, but no running water (so flushing the toilet means dumping sea water into the bowl). Instead, water is supplied by a desalination plant and each family has barrels of both clean water and sea water at their homes.

Power in the village is supplied by a large generator that runs throughout the day, except for a few hour siesta in the mid-afternoon and in the middle of the night. Food supplies are purchased by the cooperative and shipped in via the supply boat.

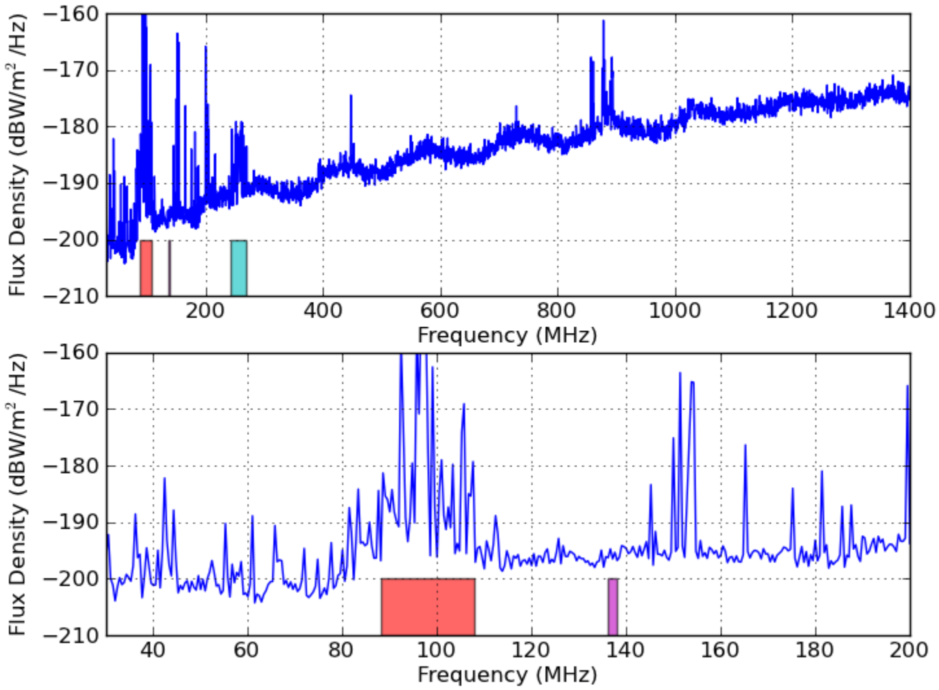


Figure 3.20: RFI measurement from the ecology camp at the summit of Isla Guadalupe collected on November 1st, 2012. At this site, the elevation means that the line of sight is longer and RFI can be seen from further distances than at lower elevations. Despite this fact, the spectrum is mostly free from RFI above $300MHz$. In the plot, colored boxes again indicate bands of well known RFI sources (Red= FM band, Purple = Orbcomm satellite band, and Cyan = military satellite band).

Bulk supplies are stored at the community store where each family can ″purchase″ food by signing out the items they need and debiting the cost to their share of the cooperative's profits. When we stayed in the fishing village, we were given use of one of the houses that was currently unoccupied. Meanwhile, we paid one of the fishermen's wives to cook for us.

Like the ecology camp, the military (naval) base is extremely minimal in scope. There is only a small contingent of personnel (∼10) at any time, so there are only a few buildings with a small generator and everyone takes turns cooking, etc. Water is also limited at the base, since the only natural water source on the island is located at the peak near the ecology camp.

Environmental Impacts

Located in the Pacific Ocean, in the midst of the California current, Isla Guadalupe is quite temperate for its latitude. The high elevation of its peaks (nearly 1300 m) means that there are two distinct micro-climates (one near sea level and one at high elevations). One reason for this contrast is that the island's peaks sit above the low cloud layer, making the higher altitudes warmer and generally clearer. Additional impacts include flash flooding in the lower areas of the island during the wet season and wind and dust interfering with system at any time.

Guadalupe's location puts it north of the main Pacific hurricane impact zone, but during the hurricane season storms pass over the island. In addition, the naval supply vessel often has its schedule changed during this season due rough seas from the storms. As an example, we had to change our deployment strategy from boat transit to plane in October 2012 due to Hurricane Paul, which hit the island as a weak tropical depression. This means that the optimal time to visit Isla Guadalupe is during the hurricane off-season (November to June).

Measurements

Upon arrival on Guadalupe, several sites were studied for potential deployment. Site 1 was near the ecology camp at the summit of the northern peak of the island, site 2 was near the fishing village on the western side of the island and site 3 was near the military base on the southern tip of the island. The exact positions of these sites are shown in Figure 3.14. I am only showing results from sites 1 and 2 as they have the most dramatic differences in RFI quality.

Figure 3.20 shows the RFI signals from site 1 at the northern summit. Just as at existing radio quiet sites the spectrum is quite clean at high frequencies. However, as we move to low frequencies there is still some significant RFI, particularly in the FM radio band. Much of this noise is coming from the radio stations in San Diego and Ensenada, including some channels which can actually be heard with a hand-held radio. In this case, the elevation is actually a detriment as the height extends the line of sight for the RFI testing antenna.

In contrast, Figure 3.21 shows the RFI signals from site 2 near the fishing village. Here the combination of low elevation, distance from the mainland and the peaks of Guadalupe act as an excellent shield to minimize the RFI in the FM band to nearly undetectable levels. As seen from the plane in Figure 3.17, this plateau has significant elevations to the north, south, and east, effectively shielding it from mainland Mexico and Baja California.

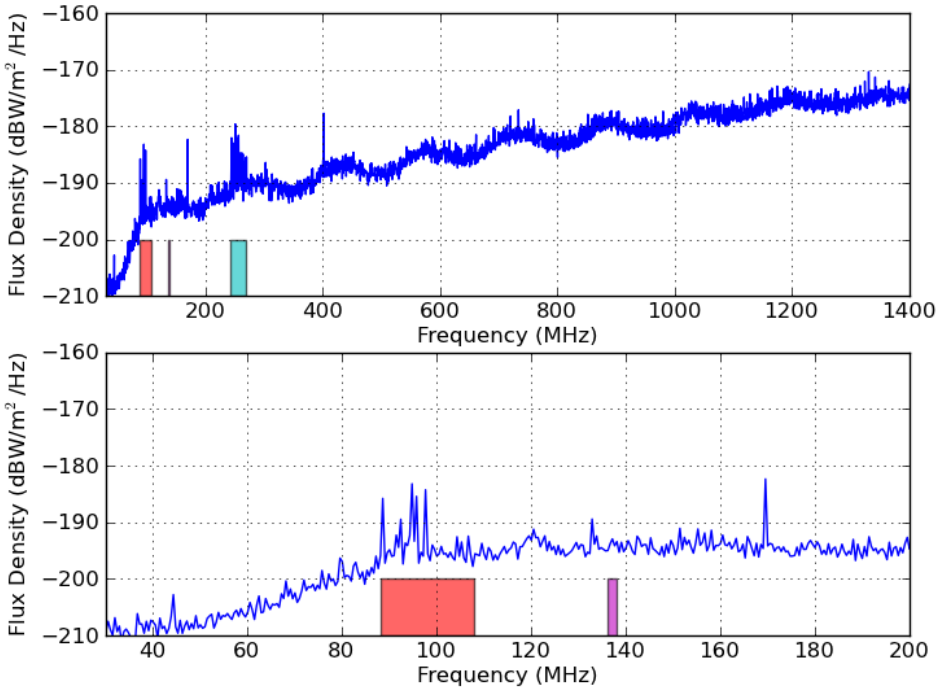


Figure 3.21: RFI measurement from the lava flow near the fishing village at Isla Guadalupe collected on November 1st, 2012. At this site, the lower elevation and the mountain between the lava flow and the mainland act as an RFI shield. This means that even the lowest frequency bands are relatively clear of RFI. In the plot, colored boxes again indicate bands of well known RFI sources (Red= FM band, Purple = Orbcomm satellite band, and Cyan = military satellite band).

3.4 Future Sites

Deployment in June 2013 to Isla Guadalupe with the SCI-HI experiment demonstrated that while the island has very low RFI in general, there is still some residual RFI in the FM band ($88MHz \leq f \leq 108MHz$). This RFI makes the band un-usable for the SCI-HI experiment. In order to continue the SCI-HI experiment, several potential sites have been identified.

1. Isla Socorro and Isla Clarión, off the west coast of Mexico, (see Figure 3.4) are under investigation as further remote sites in the northern hemisphere.

2. Marion Island, off the coast of South Africa, (see Figure 3.22) has been identified as an excellent potential site in the southern hemisphere.

Figure 3.22: Potential future site testing locations in the Southern Hemisphere. Image made using Google Maps.

3. Additional sites in the southern hemisphere, such as the Antarctic bases and Gough Island, have longer term potential as future sites for testing.

3.4.1 New Sites in Mexico

Like Isla Guadalupe, Socorro and Clarión are small volcanic islands in the Pacific Ocean off the coast of Mexico. Isla Socorro (18°47′4″ N, 110°58′30″ W) has an area of ∼130 km^2 and is about 600 km off the western coast of Mexico. Isla Clarión (18°22′ N, 114°44′ W) has an area of ∼20 km^2 and is over 700 km from the mainland.

Possessions of Mexico, both islands are ecological reserves with no permanent population. Both islands also have naval installations, although the base on Socorro is significantly larger than the one on Clarión. Access to these islands requires permits from the Mexican government, and can be achieved through passage with the Mexican Navy. The Socorro and Clarión bases are supported by twice monthly supply boats, and passage can be arranged with the Navy using these boats.

Weather plays a significant role in limiting visits to Socorro and Clarión because most of the Pacific hurricanes impact the islands each year. Therefore, access for research is limited to the off-season (December to May). Site testing on Socorro and Clarión is planned for the future.

3.4.2 New Sites in South Africa and Antarctica

Meanwhile, Marion Island ($46°52'34''$ S, $37°51'32''$ E) is a small volcanic island in the sub-Antarctic Indian Ocean owned by South Africa (see Figure 3.22). It has an area of $\sim 225\ km^2$, with a single peak $\sim$1 km in height. Located >2000 km from the South African coast, this island is expected to have excellent isolation from RFI in the FM band.

Access to the island is provided by the South African National Antarctic Programme (SANAP), which oversees research done on Marion and other islands. Most of the research on the island is focused on the native species and climate. Travel to the island happens once a year, with most of the research occurring over the ~1 month deployment. A small crew remains on the island over the rest of the year (~11 months).

Professor Jonathan Sievers at the University of Kwa-Zulu Natal in South Africa was recently awarded a grant through SANAP to investigate and use Marion Island as a site for radio astronomy. As a part of this grant, I will be traveling to Marion Island on the 2016 SANAP trip, which takes place from April to May 2016. During this trip, I will make measurements around the island. These measurements will be used to evaluate the overall RFI environment and identify the best location to place an experiment. I will also deploy the SCI-HI experiment at the location I identify during my evaluation.

Since this trip is over a year away, we will deploy the SCI-HI instrument to the Karoo desert in South Africa in April 2015. The Karoo desert site is the future location for the Square Kilometer Array [30] in South Africa, and as such is a protected radio quiet site. The Karoo site is not expected to be as quiet in the FM band as Marion island, but is much more readily accessible.

Chapter 4

Low Frequency Radio Astronomy and the Earth's Ionosphere

4.1 Overview

Beyond the effects of RFI, the Earth's atmosphere has additional impacts on low frequency radio signals from the universe. These impacts come from interactions between the incoming radio signals and free electrons in the ionosphere. Impacts include refraction and absorption at radio frequencies.

Ionospheric impacts can mask the 21-cm signal by introducing additional frequency dependent structure into the sky signal. This frequency structure is proportional to ν^{-2}, as explained below, and may be larger than the 21-cm structure for the SCI-HI frequency band. In addition, ionospheric structure is time and latitude dependent. Therefore, it is necessary to quantify and address ionospheric impacts to any 21-cm measurement.

4.2 Earth's Ionosphere

To understand ionospheric impacts on the 21-cm signal we need to understand the ionosphere. Earth's ionosphere is located $60 - 1000+$ km above the surface of the Earth and is made up of free electrons and ionized atoms. Neutral atoms in Earth's atmosphere are photoionized by solar radiation, which leads to an altitude dependent distribution of free electrons and ions in the atmosphere.

4.2.1 Ionosphere Layers

We can separate the ionosphere into altitude-dependent layers based on the distribution of free electrons. Each layer is defined by an altitude where the free electron distribution has an inflection point, commonly described in the literature as a local

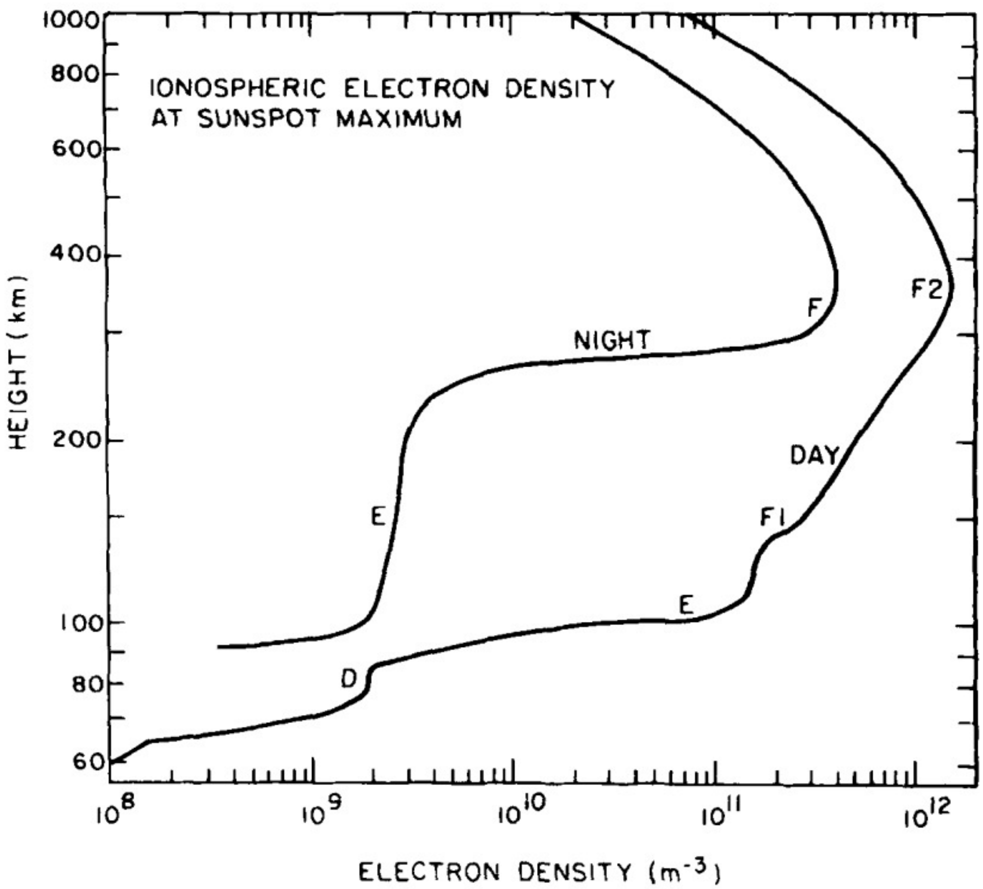


Figure 4.1: Idealized average free electron density distribution in Earth's atmosphere. Layers are shown during day and night. Each layer corresponds to an inflection point in the free electron density distribution, referred to in the literature as a maximum.

maximum. The altitude corresponding to the inflection point is the center altitude of the layer. Depending on the time of day and geographic location, the number of layers present in the ionosphere will vary as shown in Figure 4.1. There are four layers defined using inflection points (D, E, F1 and F2). A fifth topside layer is defined as the altitude above the F2 layer where the dominant ion in the ionosphere becomes O^+. The absolute maximum free electron density (n_e) is reached in the F2 layer and is $n_e \cong 10^5 cm^{-3}$ [27].

Earth's ionosphere comes from solar radiation photoionizing neutral atoms, which is balanced by recombination of ions and free electrons. This ionization and recombination balance leads to fluctuations in the free electron density and ionospheric layers that depend on the amount of solar radiation currently impacting a given latitude and longitude. There are three major sources of variability in solar radiation levels that affect the ionosphere [27].

First, solar radiation is maximized during daylight. During those times, the free electron density is large, and more ionospheric layers are present (see Figure 4.1). In contrast, there is little to no solar radiation at night, so fewer ionospheric layers are present [27].

Second, solar radiation levels are also always much lower at polar latitudes than they are at mid-latitudes. This is due to the lower sun angle at polar latitudes. During the polar winter, solar radiation levels and the number of free electrons in the atmosphere reach a minimum [27].

Third, solar radiation levels have a periodic cycle of $\sim$ 11 years as tracked by sunspot frequency. During a period of high sunspot activity, the corresponding levels of solar radiation are also higher, leading to larger free electron density [27].

4.2.2 Ionosphere Properties

Plasma Frequency

Free electrons in the ionosphere can be treated as a plasma which modifies electromagnetic wave propagation. Change in propagation due to a plasma is quantified using the plasma frequency (ν_p), which has a value given in Equation 4.1, where n_e is in m^{-3} [33].

$$\nu_p = \frac{e}{2\pi}\sqrt{\frac{n_e}{\varepsilon_0 m}} \cong 9\sqrt{n_e} Hz \tag{4.1}$$

This plasma frequency changes the index of refraction for the ionosphere compared to free space ($n = \sqrt{1 - \nu_p^2/\nu^2}$), where ν is the frequency of the radiation that is passing through the atmosphere. The plasma frequency is typically $\nu_p \leq 12MHz$, below the SCI-HI frequency band. Above the plasma frequency, electromagnetic waves are refracted by the ionosphere. Meanwhile, below the plasma frequency the ionosphere reflects electromagnetic waves, preventing transmission of signals from the sky [33].

Cyclotron Frequency

In addition to the plasma frequency due to free electrons, the Earth's magnetic field leads to cyclotron motion of free electrons and further modification of radiation propagation. Free electron motion is quantified using a cyclotron frequency, also called a gyrofrequency. The magnitude of the gyrofrequency is ($\nu_B = eB/2\pi m$), and has a typical magnitude of $\nu_B \cong 1.4MHz$, lower than the plasma frequency. [33].

This is the source of Faraday rotation of signals due to Earth's ionosphere [33]. However, since the impact is proportional to ν_B/ν and $\nu \gg \nu_B$ for the SCI-HI frequency band we will neglect the impact from the Earth's magnetic field in the remainder of the thesis.

4.3 Total Electron Content (TEC)

Measurement of the number of free electrons versus altitude (n_e) can be quite difficult. Instead, ionosphere experts typically summarize the state of the ionosphere using the total electron content (TEC). TEC is the projected column density of free electrons in the atmosphere. It is defined using Equation 4.2, where h is the altitude [33].

$$TEC = \int_0^\infty n_e(h)dh \tag{4.2}$$

TEC is typically quoted in $TECU$, where $1TECU$ corresponds to an average $n_e = 10^{16}$ electrons per m^2 [36]. Total electron content is continually monitored on Earth using measurements from GPS stations around the world. This data is collected by a number of agencies that convert the data into maps.

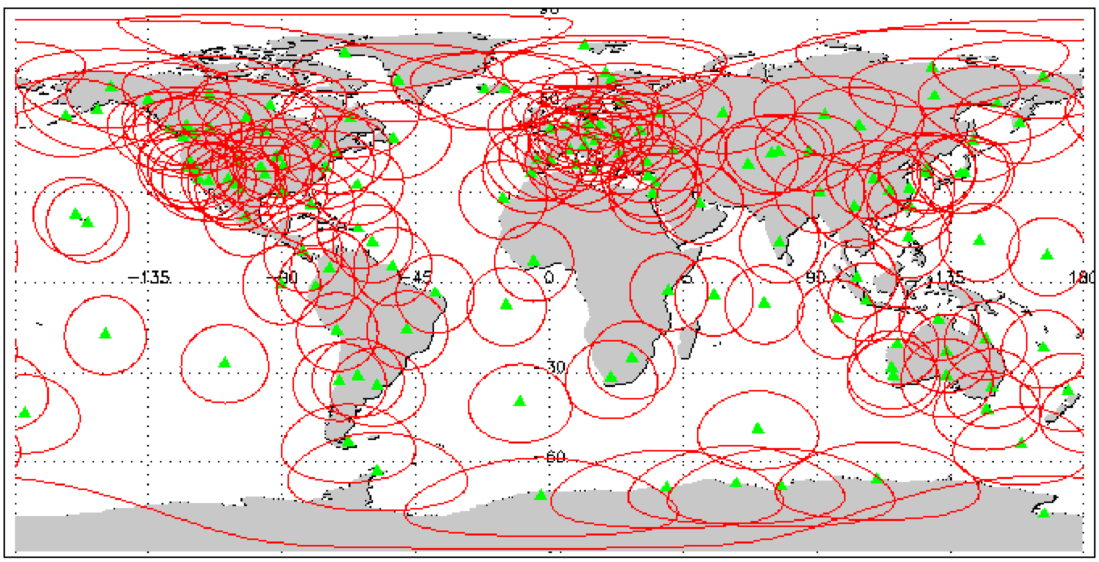


Figure 4.2: Global Map of GPS Network stations used for the quasi-real time maps of TEC available online at iono.jpl.nasa.gov

4.3.1 Measuring TEC

Real time maps of TEC over the entire world are made by NASA JPL using the GPS stations shown in Figure 4.2 and are available online[1]. The plots are based on a five minute average, and the image is updated online every five minutes.

Example TEC maps for two different times of day (and year) are shown in Figures 4.3 and 4.4. The bright regions with high TEC correspond to the part of the globe

[1] `http://iono.jpl.nasa.gov/latest_rti_global.html`

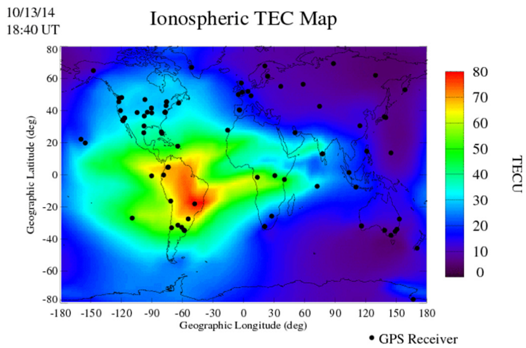


Figure 4.3: TEC map for October 13th, 2014 at 18:40 UTC. (Image found at iono.jpl.nasa.gov)

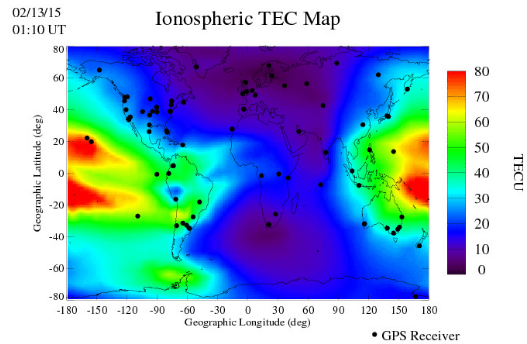


Figure 4.4: TEC map for February 13th, 2015 at 01:10 UTC. (Image found at iono.jpl.nasa.gov)

currently receiving sunlight at the time the data was collected. Areas close to the equator have much higher TEC than areas near the poles in both the day and night regions.

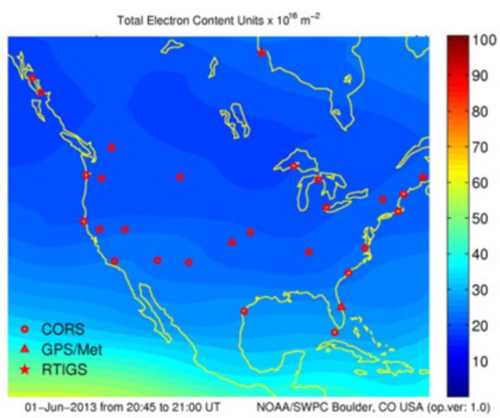


Figure 4.5: North American TEC map for June 1st, 2013 at 21:00 UTC. (Image found at www.swpc.noaa.gov)

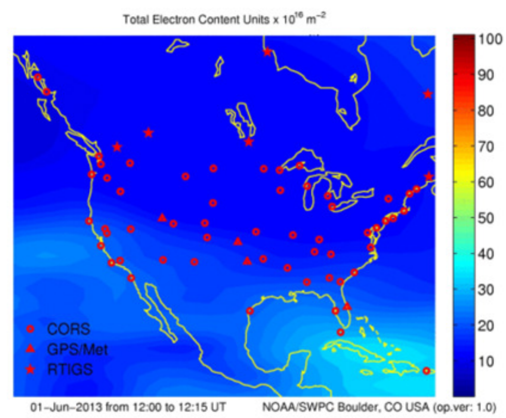


Figure 4.6: North American TEC map for June 1st, 2013 at 12:15 UTC. (Image found at www.swpc.noaa.gov)

North American TEC maps are also generated by the Space Weather Prediction Center of the National Oceanic and Atmospheric Administration (NOAA)[2]. Beyond the regularly updated images, NOAA makes the most recent data ($\sim 24 - 72hrs$) available for easy download. Plots of North American TEC for the Isla Guadalupe deployment in June 2013 are shown in Figures 4.5 and 4.6.

[2] `http://www.swpc.noaa.gov/products/us-total-electron-content`

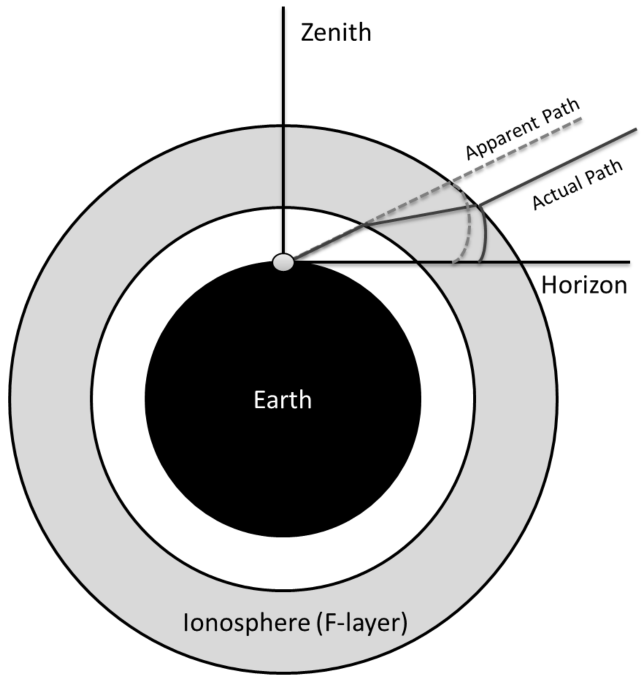


Figure 4.7: Ray tracing for incident light with a single ionospheric layer. Sizes are not to scale.

4.4 Ionospheric Impacts

4.4.1 Refraction

From basic optics, Snell's law tells us that radiation passing through an interface between materials having different indices of refraction will be bent. The magnitude of this bending depends only on the difference between the indices of the two materials and the incidence angle of the radiation ($n_0 sin\theta_0 = n_1 sin\theta_1$). If radiation passes through a flat plane of material with index n_1, having air ($n = 1$) on either side, the angle of the radiation coming out the other side of the material is unchanged. However, the ionosphere is a set of spherical layers rather than a flat plane, leading to changes in the angle of the radiation [33].

Therefore, when radiation enters the ionosphere at some angle $\theta_0 \neq 0$, the radiation reaches the ground with an apparent incidence angle that differs from its arriving incidence angle. This is shown in Figure 4.7, where I've used a single ionospheric layer for simplicity. We can define the deflection angle $\delta\theta = \theta_{apparent} - \theta_{actual}$ [33].

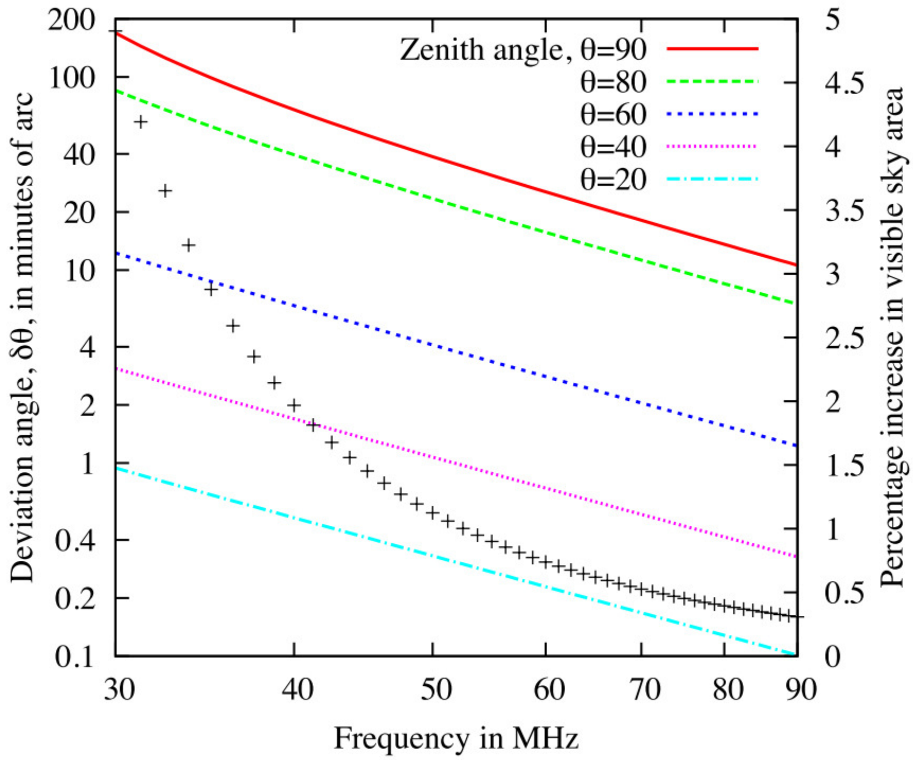


Figure 4.8: Refraction angle ($\delta\theta$) for $TEC = 10TECU$ as a function of zenith angle and frequency from Vedantham et al [36]. The predicted level of refraction due to the Earth's ionosphere is below the detection limit of the SCI-HI data collected in June 2013.

To compute $\delta\theta$ requires a model for the ionosphere. The simplest approximation is a single ionospheric layer with constant n_e. Using this approximation, $\delta\theta \propto \nu^{-2}cos(\theta)(sin^2\theta + 2h_F/R_e)^{-1.5}$, where h_F is the mean height of the ionosphere and R_e is the Earth's radius. Vedantham et al [36] used this approximation, assuming that the total electron content was $10TECU$ and $h_F = 300km$, to calculate the deviation angle $\delta\theta$ and percentage increase in visible sky area (beyond the normal horizon), as shown in Figure 4.8.

4.4.2 Absorption

Radiation passing through Earth's ionosphere can also be absorbed due to collisions between electrons and other particles in the atmosphere. Most of this absorption will occur at the lowest level of the ionosphere (D-layer) where the number of neutral particles is largest. Absorption is small for $\nu \gg \nu_p$, on the order of 1-2% [36]. In addition, absorption is frequency dependent and $\propto \nu^{\alpha-2}$, where α is the spectral index of the radiation from the sky [36].

4.5 Earth's Ionosphere and the SCI-HI system

Both refraction and absorption due to the ionosphere have the potential to affect the SCI-HI experiment. Refraction will appear as a change in antenna beam pattern compared to the ionosphere-free pattern, while absorption will appear as a multiplicative loss factor that adds additional frequency structure to the data.

Ionospheric impacts can be assessed and mitigated through two strategies. The first strategy is collection of supplemental data to quantify the expected ionospheric impact. The strategy uses a time dependent estimate of TEC at the SCI-HI deployment site and a good model for absorption and refraction given TEC. This allows calculation of $\delta\theta$ and an absorption attenuation factor. TEC estimates can be sourced from publicly available data such as the NASA JPL and NOAA libraries.

The second strategy is to decrease the ionospheric impact on the data to below 21-cm signal levels. This decrease requires data that is collected during low TEC times and in low TEC locations. Limiting analysis to data collected during the night will greatly decrease average TEC and ionospheric impacts in the data. In addition, usage of a polar latitude site such as Marion Island, see Figure 3.22, will help to further decrease average TEC.

Chapter 5

SCI-HI Data Analysis

5.1 Preparing the Data

Before the data can be analyzed for potential signals, it first has to go through an analysis pipeline to remove man-made RFI and instrument noise and make the datasets easier to handle.

5.1.1 Integration and Sampling

The frequency spectrum collected for each second of data is stored in a single file. The file has a header containing basic information: (a) time-stamp, (b) DC voltage level (only collected when the system is powered by a DC Power Supply), (c) system temperature, and (d) switch position (antenna vs calibration sources). The spectrum covers the frequency range ($0 \leq f \leq 250MHz$), with a frequency resolution ($\Delta f \sim 7.63kHz$ or 32769 samples). Data is stored in units of power (dB). An example of the data collected by the system is shown in Figure 5.1. Also included is a single sample of data from each of the calibration sources (Short, 50 Ω, 100 Ω, and Noise Source).

Examining the raw data, we can identify several frequency dependent structures that are connected to the SCI-HI system design. The sharp signal cutoffs at $30MHz$ and $200MHz$ are caused by the high and low pass filters. The comb of narrow spikes present in both the antenna and calibration data is due to the interleaved sampling strategy described in Section 2.3.3. Long period frequency structures from $150-200MHz$ are caused by the HIbiscus antenna beam and impedance. The ″bump″ in the antenna data around $95MHz$ is time-dependent and has been tied to the system power, so it is believed to be caused by RF leakage from the SCI-HI system.

Additional spikes present only in the antenna data are man-made RFI picked up by the antenna. These RFI spikes can be matched to RFI measured during site testing such as the FM band ($88 - 108MHz$) and the Orbcomm satellite band ($136 - 139MHz$). Looking at the structure of the data over time, we can also see that the antenna data varies depending on what part of the sky it sees. This is shown

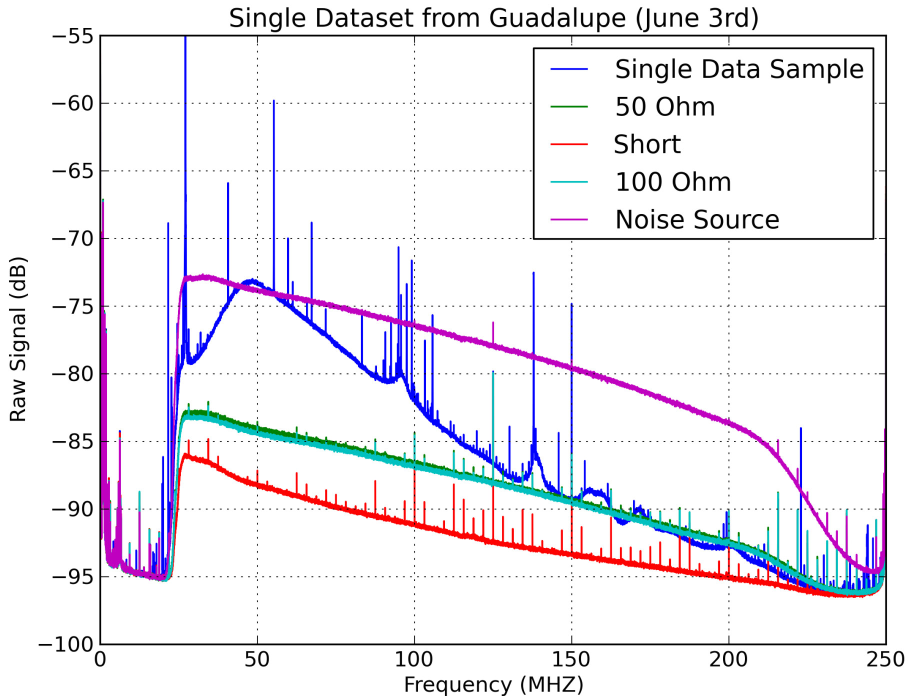


Figure 5.1: Single dataset (1 second integration) of the entire frequency spectrum of raw data from the SCI-HI system. Also present are single datasets from the different calibration sources.

in Figure 5.10, where the daily variance of a single frequency is plotted.

5.1.2 Truncation

Because the data we are interested in is a subset of the overall spectrum, one of the first things that we do is truncate the frequency range of the data to make the dataset smaller and easier to handle. We can throw away data for frequencies where the HIbiscus antenna doesn't have a stable beam or the data is dominated by signals other than those from the sky.

In the 2013 data, the reliable band corresponds to $(50 - 120MHz)$, as is shown in Figure 5.2. In this part of the spectrum, the linear slope of the antenna dataset compared to the calibration datasets stands out. This linear slope is caused by the spectrum of synchrotron radiation in the Milky Way Galaxy, as will be discussed in later sections of this chapter.

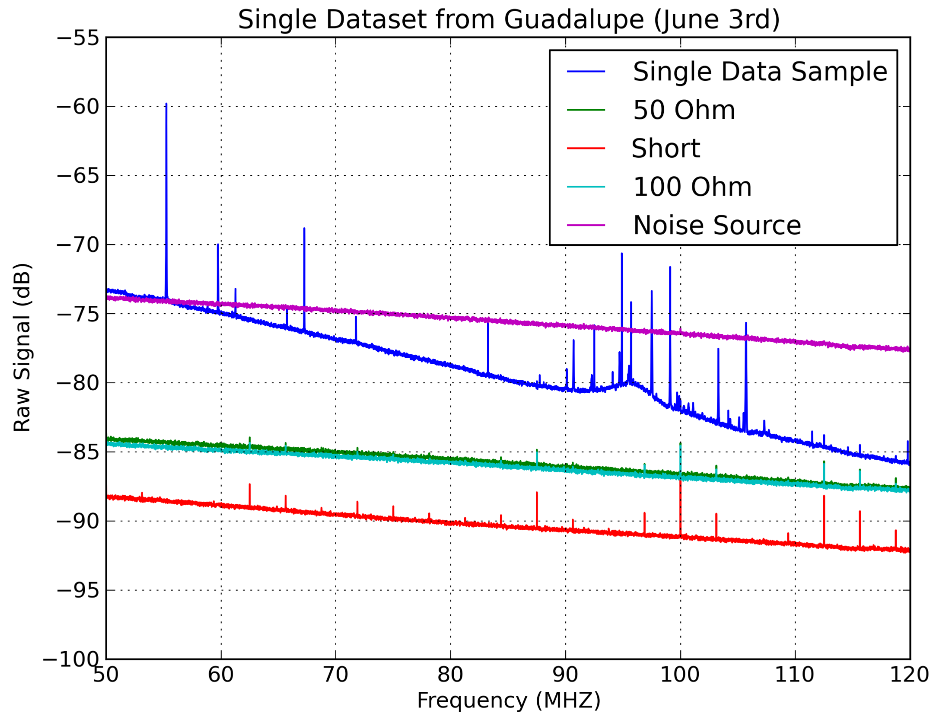


Figure 5.2: Truncated spectrum of a single dataset of raw data from the SCI-HI system. Also present are single datasets from the different calibration sources.

5.1.3 RFI Excision

Another early stage of the pipeline is RFI excision (both in frequency and time) of narrow features or spikes in the data.

RFI Frequency Flagger

RFI excision in frequency is done by flagging each dataset according to a threshold. The threshold is calculated for each frequency (f_c) in the spectrum by fitting a small subset of the spectrum ($f_c - n \leq f \leq f_c + n$), where n is the number of nearest neighbors used for the fit. This fit is a two term polynomial. Dividing out the fit gives a flattened dataset whose mean is one and whose variance can be calculated. The variance is used to set a threshold (eg 3σ), and if the flattened data at f_c is outside the threshold then f_c is flagged.

Flagging identifies a list of frequencies between $50 - 120MHz$, which have data that should be ignored. The data at these frequencies has a spike above the threshold in a given dataset. The exact threshold and number of neighbors used to set the mean can be tweaked, but the typical values I used were a 3σ threshold and 32 neighbors on either side. If a data point is flagged, it is masked out of the data and will not be included in further analysis.

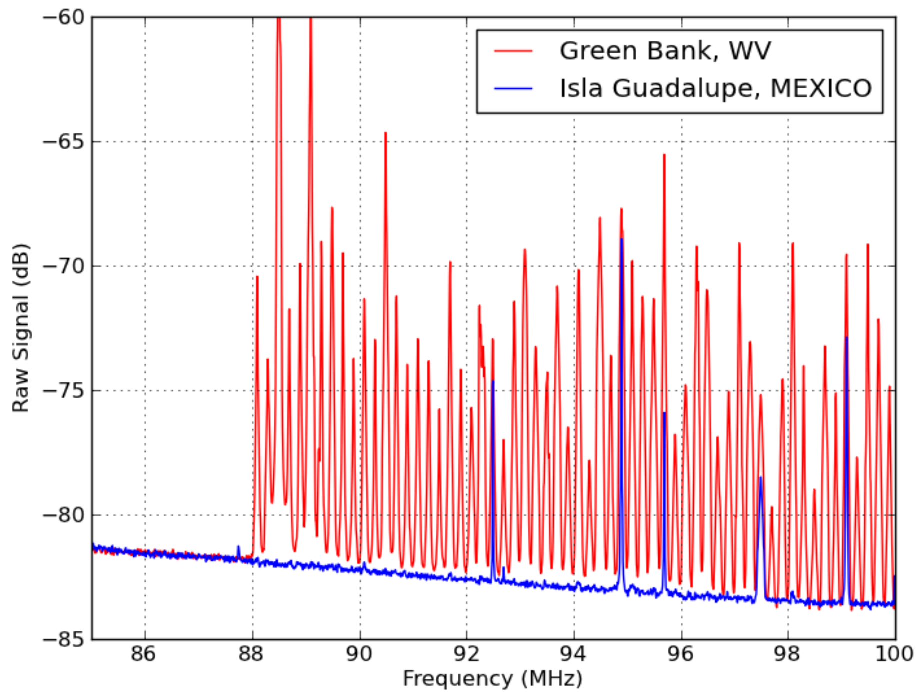


Figure 5.3: RFI comparison in the FM band between Isla Guadalupe and Green Bank, West Virginia.

RFI in the FM Band

One of the frequency bands where RFI excision is particularly challenging is the FM band ($88 - 108MHz$). In this band, the variance of the signal can be quite large, making the RFI excision procedure less reliable. For many of our testing sites like Green Bank, WV the entire FM band is occupied with large RFI signals (as discussed in Chapter 3). This can be seen in Figure 5.3, where the RFI signals are on average $\sim$10 dB above the floor.

On the other hand, the FM band RFI signals from Isla Guadalupe are mostly $\leq$1 dB. These RFI signals are small enough to be missed by the RFI flagger and may be run through the rest of the pipeline stages. The few large spikes $\geq$5 dB above the floor will be removed by the RFI excision code.

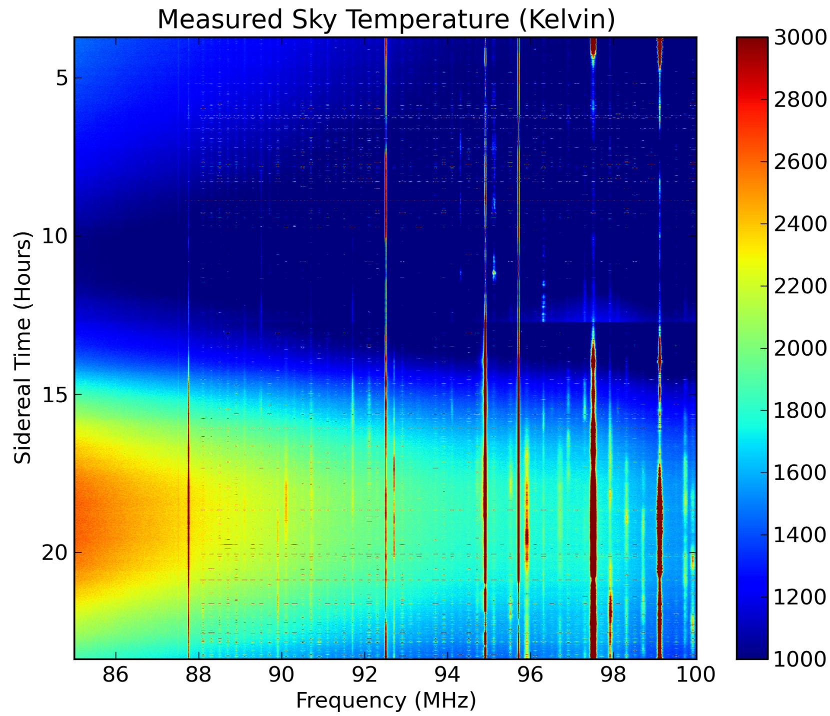


Figure 5.4: Calibrated full day of data taken on June 6th, 2013 showing RFI variance for $85 - 100MHz$. Note that the vertical lines correspond to FM stations, while horizontal lines correspond to times of meteor activity.

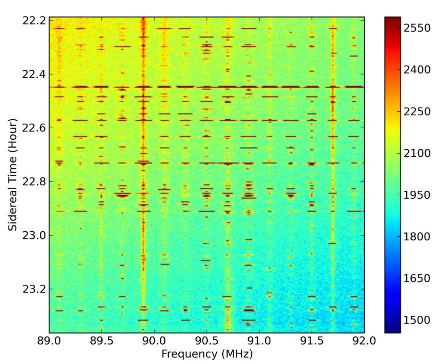


Figure 5.5: Short ($\sim$ 1 hour) subset of data taken on June 6th, 2013. Data shows RFI variance for $89-92MHz$ during a period of high meteor activity.

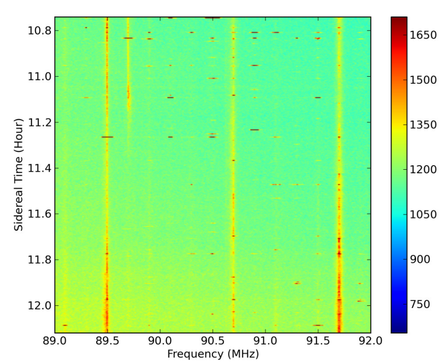


Figure 5.6: Short ($\sim$ 1hr) subset of data taken on June 6th, 2013. Data shows RFI variance for $89 - 92MHz$ during a period of low meteor activity.

5.1.4 Time-variable RFI

While most of the RFI in the data is time independent and can be flagged out in each dataset independently, some of the RFI has a strong time dependence. Time dependence comes from two major sources: (a) meteor scatter and (b) local sources.

Time-variability of RFI due to Meteor Scatter

Meteoroids are always passing through Earth's atmosphere. These meteoroids have a wide spectrum of sizes starting from $\leq$10 nm. As these meteoroids pass through the Earth's atmosphere, some of their atoms are vaporized. Inelastic collisions between these atoms and air molecules in the atmosphere can lead to ionization [9].

Ionization due to meteoroids leads reflection of FM signals off the meteor trail. This occurs because ionization causes an increase in the column density of free electrons in the Earth's ionosphere along the meteor trail. The increase in ionization temporarily increases the plasma frequency of the ionosphere along that trail, such that the ionosphere is opaque at higher frequencies along the trail. Opacity allows reflection of signals such as FM radio off the ionosphere, extending the range of visibility for the FM band beyond what would normally be accessible. Observation of meteor scatter in the FM band has been previously observed with low frequency radio telescopes [22].

In the SCI-HI data, meteor scatter leads to a time-variable increase in the FM band RFI. For short time periods after a meteor passes overhead the strength of RFI signals from FM radio towers becomes larger. The duration of these signals is relatively small, typically lasting only a few seconds and diminishing as free electrons diffuse away from the original meteor trail. The magnitude of the increase in the FM signals is also variable, and different RFI signals are amplified depending on the location of the meteor trail. The rate at which meteor events occur is highly time variable, as meteors don't fall evenly, but rather travel in groups (aka meteor showers). Also, smaller meteoroids produce smaller trails, which allow for less transmission of FM signals.

Meteor scatter effects can be seen in Figure 5.4, which shows half of the FM band over one day of data collection. In this waterfall plot, strong FM stations are seen as vertical lines, while meteor scatter events show up as horizontal lines. When we zoom in, as shown in Figures 5.5 and 5.6, we can see differences in meteor scatter rate and meteor trail size at different times.

Time-variability of RFI from Local Sources

The other source of time-variable RFI is local noise from our surroundings. Our site at Isla Guadalupe was a few km from the local fishing village, which meant that RFI was occasionally picked up from the village. This mainly occurred when the diesel generator was on. Additionally, the site was a few hundred meters from the road that led from the village to the rest of the island. When vehicles drove past the site (a few times each day), RFI from the vehicles such as broadband RF from spark plugs could be seen in the data.

Time-variable RFI Flagging

In order to remove such time-variable RFI, I wrote a second RFI flagger identical to the frequency flagger, except that it works along the time axis. I typically used the same threshold and number of neighbors as the frequency flagger.

5.1.5 Rebinning

Because the data is collected at a much higher resolution than is needed for the analysis, it needs to be rebinned to lower resolution. This can be done either in frequency or time. Compression is done after RFI flagging in order to keep a single ″bad″ channel from affecting the overall signal. The mean of the unflagged data for a given bin scale becomes the new data point. In addition, a new data point is left as a flagged point if over half of the data used to make it were flagged.

5.2 Calibrating the Data

After RFI removal and compression, the data is ready for calibration. Calibration is used both to put the data into correct units and also remove instrument sourced structure in the frequency spectra.

5.2.1 Calibration Datasets

Part of the SCI-HI system, as laid out in Chapter 2, is an electromechanical-mechanical switch that allows data collection from both the antenna and known temperature sources placed on the different switch input terminals. Figures 5.1 and 5.2 show the different datasets measured with the system, which are also known as ″Johnson Noise″ datasets. In these datasets, the temperature of each source can be separated into multiple terms.

Short Terminator For the ″Short″ signal, a shorting terminator is placed on the switch input terminal. This means that all of the signal seen in the data processor is coming from the SCI-HI system. This includes both instrument noise and thermal noise. Therefore, we'll call the instrument noise power P_{short}.

50Ω **Terminator** For the ″50Ω″ signal, a load is placed on the switch input with an impedance of 50Ω. The load provides an input of the current ambient temperature, and instrument and thermal noise adds additional signal. There is another noise term

(P_Z), which depends on the match between the impedance of the first stage amplifier and the load. We write the 50Ω power as:

$$P_{50\Omega} = P_{amb} + P_{Z50} + P_{short} + P_{thermal} \tag{5.1}$$

100Ω Terminator For the ″100Ω″ signal, a load is placed on the switch input with an impedance of 100Ω. This load provides a nearly identical power as the 50Ω load, with the exception of the impedance noise term. We write the 100Ω power as:

$$P_{100\Omega} = P_{amb} + P_{Z100} + P_{short} + P_{thermal} \tag{5.2}$$

Artificial Noise Source For the ″Noise Source″ signal, an artificial noise source with 50Ω impedance is placed on the switch input. This source provides both an ambient temperature signal and an additional artificial noise signal (P_{Noise}) which can be independently measured. The impedance noise term matches the 50Ω impedance term. We write the noise source power as:

$$P_{NS} = P_{Noise} + P_{amb} + P_{Z50} + P_{short} + P_{thermal} \tag{5.3}$$

Antenna For the signal from the HIbiscus antenna, we write a similar equation for the power. In this case, we have thermal and instrument noise and an impedance term, as well as signals from the sky.

$$P_{Ant} = P_{Sky} + P_{Zant} + P_{short} + P_{thermal} \tag{5.4}$$

Impedance (current) Noise The impedance noise (P_Z) is also called the current noise of the system. To understand this definition, we go back to the basic circuit diagram of an amplifier. If the amplifier input is hooked up to a short, the measured noise in units of power is $P \propto V^2$, where V is the voltage sent into the amplifier.

On the other hand, when the 50Ω (or 100Ω) terminator is connected the additional noise present in the system beyond the short noise also has a voltage. The magnitude of this voltage is $V \propto IZ$, where I is the current through the system and Z is the impedance of the terminator [31]. Given Figure 5.1, which shows that $P_{50\Omega} \cong P_{100\Omega}$ while $Z_{100\Omega} = 2Z_{50\Omega}$, the current noise term must be small for our system.

5.2.2 On-site Impedance Measurement

Impedance can be measured using a Vector Network Analyzer (VNA) looking at the reflectivity ($S11$) data for the input sources. The calibration sources have an impedance that is entirely real with zero phase and constant in frequency (50Ω or 100Ω).

In comparison, the HIbiscus antenna has a complex impedance that varies with frequency and has a phase component (see Section 2.2.4 and Figures 2.34 and 2.35). This impedance must be measured in-situ, since it depends on the exact layout of the system and the shape of the terrain around the antenna.

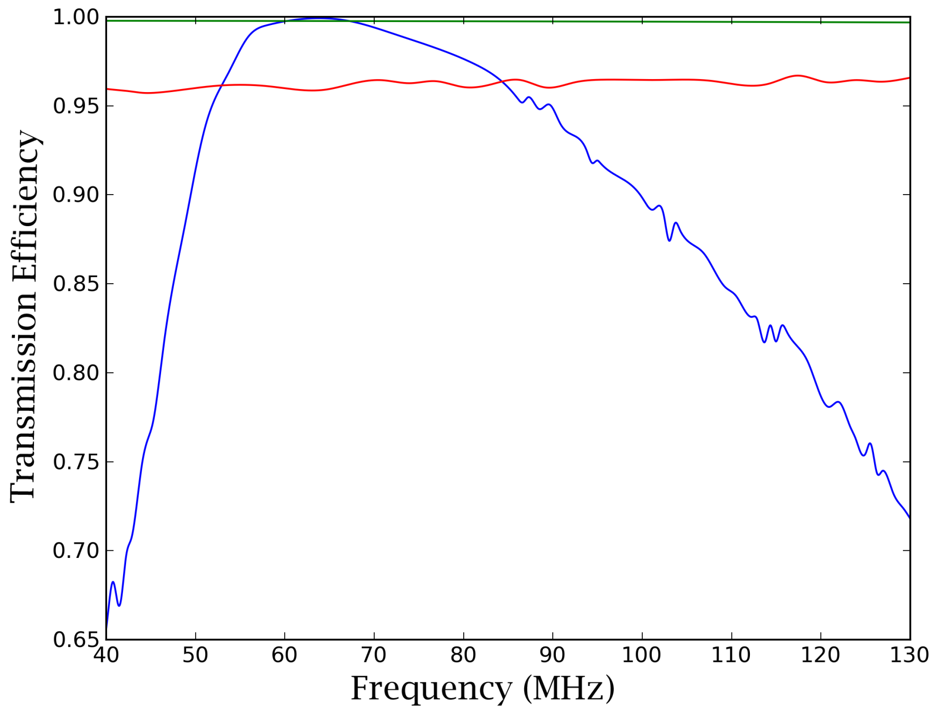


Figure 5.7: SCI-HI system transmission efficiency between the first stage amplifier and the different switch inputs. Transmission efficiency is calculated using $S11$ measurements. Blue (η) is the transmission efficiency for the HIbiscus antenna, green ($\eta_{50\Omega}$) is the transmission efficiency for the 50Ω terminator, and red ($\eta_{100\Omega}$) is the transmission efficiency for the 100Ω terminator. Here $\eta = 1$ means that all the power from the switch input is transmitted through the first stage amplifier.

Measured Transmission Efficiency (η)

Because the antenna impedance is not the complex conjugate of the first stage amplifier impedance, the sky signal measured by the antenna does not all make it into the first stage amplifier. What the amplifier actually sees is ($P_{Sky}\eta$), where the transmission efficiency (η) was defined in Section 2.3.2 and is $0 \leq \eta \leq 1$. For the June 2013 deployment data, the efficiency is shown in Figure 5.7. With the efficiency, our antenna power equation (5.4) can be re-written using Equation 5.5.

$$P_{Ant} = P_{Sky}\eta + P_{short} + P_{thermal} \tag{5.5}$$

Systematic errors in the efficiency calculation can occur in three ways. First, errors in the measured reflectivity of the antenna and first stage amplifier will propagate into the transmission efficiency. Second, errors can also arise from the fact that the antenna reflectivity must be measured when the switch and first stage amplifier are not present in the system. Third, errors can also arise from inaccurate estimation of the cable delay through the switch between the antenna and first stage amplifier.

5.2.3 Milky Way Galaxy (GSM) Modeling

Calibration using the switch and known temperature sources assumes that the antenna is perfectly efficient at collecting radiation from the sky. This efficiency, also known as antenna gain, is separate from the transmission efficiency (η) and is $0 \leq e \leq 1$. If we include this term, the antenna power equation (Equation 5.5) becomes Equation 5.6. To account for $e < 1$, it is necessary to use a known temperature source that is seen through the antenna. For the SCI-HI experiment, the only astronomical source accessible given our $\sim 55°$ antenna beam is radiation from the Milky Way Galaxy.

$$P_{Ant} = eP_{Sky}\eta + P_{short} + P_{thermal} \tag{5.6}$$

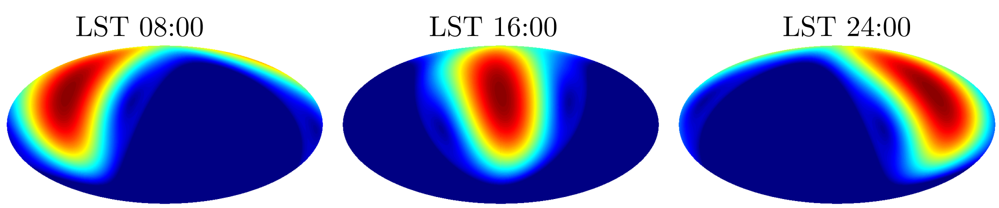


Figure 5.8: Simulated HIbiscus antenna beam on the sky in RA, DEC coordinates at Isla Guadalupe's latitude. Each map corresponds to the antenna beam at $70MHz$ for a different LST.

HIbiscus Beam Coverage

Once the SCI-HI system has been set-up on-site at a particular location, the Earth rotates as the antenna constantly points towards the zenith. Due to this rotation, the beam of the Hibiscus antenna looks at different parts of the sky at different times of day.

The exact beam coverage can be calculated using the simulated HIbiscus beam and the site latitude. Antenna beam coverage on the sky for Isla Guadalupe at a few different times is shown in Figure 5.8. It is important to note that here we use Local Sidereal Time (LST), not Coordinated Universal Time (UTC), because we care about the motion of the Milky Way Galaxy rather than the sun.

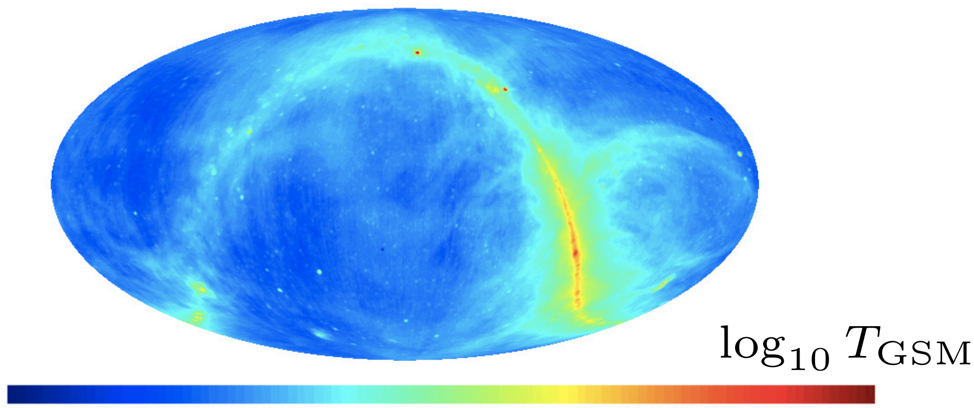


Figure 5.9: Sky temperature calculated using the GSM model from de Oliveira-Costa et al [10] at $70MHz$. The temperature map is in units of log_{10}(Kelvin), while the map coordinates are in RA, DEC.

GSM Model

The main signal from the sky at $40 - 130MHz$ is the Milky Way Galaxy, which has temperatures ≥ 1000 Kelvin for most of the SCI-HI frequency band. However, most of the Milky Way Galaxy signal comes from the plane of the galaxy, which moves from horizon to zenith to horizon due to Earth's rotation. This motion causes the galactic plane to drift in and out of the HIbiscus beam over time.

The Galactic Global Sky Model (GSM) is the current best model for the sky, including the Milky Way Galaxy, at these frequencies. It is a model created by interpolating data from many different publicly available radio surveys with a frequency range of $10MHz - 100GHz$, as is discussed in Section 1.3.1. The model and software [10], can be used to make maps in our frequency band. One such map is shown in Figure 5.9, displayed using equatorial coordinates (right ascension, RA, and declination, DEC) to match the beam maps.

Expected Sky Signal

Using the combination of the GSM model and the simulated HIbiscus antenna beam, the beam-averaged sky temperature can be calculated using Equation 5.7, where $GSM(\theta, \phi, \nu)$ is the model data and $\mathcal{B}(\theta - \theta_0(t), \phi - \phi_0(t), \nu)$ is the simulated beam, with $(\theta_0(t), \phi_0(t))$ being the beam center.

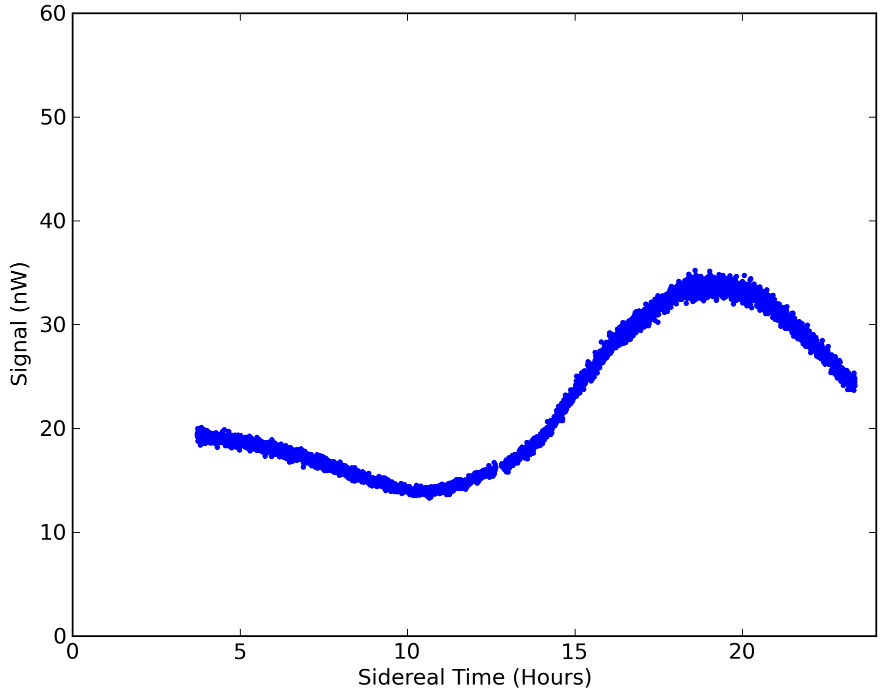


Figure 5.10: Single day plot of uncalibrated data collected with the antenna on June 6th, 2013 at $70MHz$.

$$T_{GSM}(\nu, t) = \frac{\int d\Omega GSM(\theta, \phi, \nu)\mathcal{B}(\theta - \theta_0(t), \phi - \phi_0(t), \nu)}{\int d\Omega \mathcal{B}(\theta - \theta_0(t), \phi - \phi_0(t), \nu)} \tag{5.7}$$

This signal is expected to reach a maximum when the Galactic plane is at zenith, and be at a minimum when the Galactic plane is along the horizon. Figure 5.10 shows that the antenna signal matches this behavior using a single day of data from June 2013, plotted at $70MHz$.

5.2.4 Calibration Factor Calculation

All of our datasets have units of power (dB), as is discussed in Section 5.1.1, but we want to know the sky signal in units of temperature (Kelvin). Therefore, we need to convert our signals from power to temperature. This conversion can be written with a simple calibration factor $K(\nu)$ such that $T(\nu, t) = K(\nu) * P(\nu, t)$. This means that the sky temperature (T_{Sky}) can be written using Equation 5.8. Calculating the calibration factor can be done with either known source datasets or the Milky Way Galaxy model.

$$T_{sky}(\nu,t) = K(\nu) * P_{sky}(\nu,t) = K(\nu)\Big[\frac{P_{Ant}(\nu,t) - P_{Short}(\nu) - P_{thermal}(\nu)}{\eta(\nu)\ e(\nu)}\Big] \tag{5.8}$$

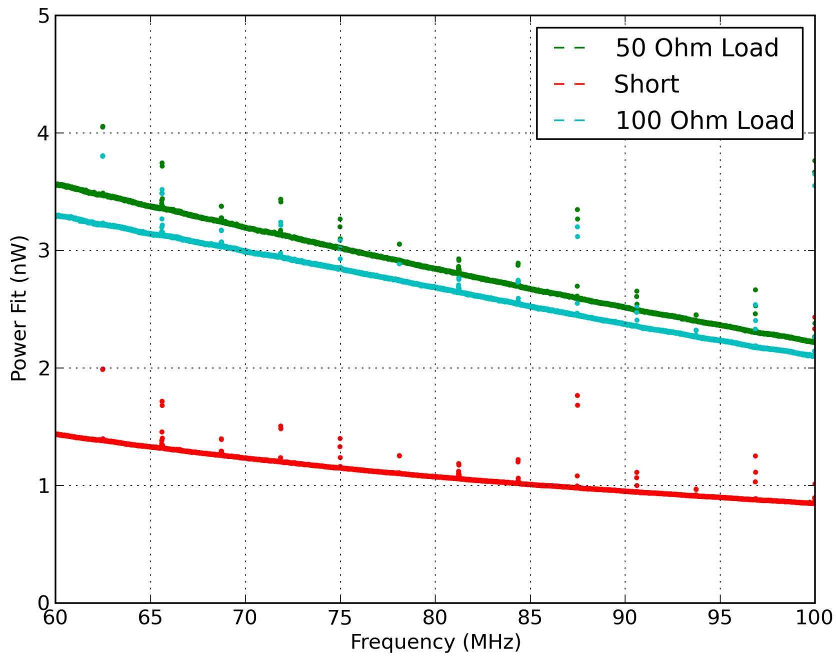


Figure 5.11: High frequency resolution time average of all the calibration datasets for June 3rd, 2013.

Johnson Noise Calibration (K_{JNC})

Using the Johnson Noise datasets, we determine a calibration factor $K_{JNC}(\nu)$. To do this, we start with the following four datasets:

$$T_{short} = K_{JNC} * (P_{short} + P_{thermal}) \tag{5.9}$$

$$T_{50\Omega} = K_{JNC} * P_{50\Omega} = K_{JNC} * (P_{amb} + P_{Z50} + P_{short} + P_{thermal}) \tag{5.10}$$

$$T_{100\Omega} = K_{JNC} * P_{100\Omega} = K_{JNC} * (P_{amb} + P_{Z100} + P_{short} + P_{thermal}) \tag{5.11}$$

$$T_{NS} = K_{JNC} * P_{NS} = K_{JNC} * (P_{Noise} + P_{amb} + P_{Z50} + P_{short} + P_{thermal}) \tag{5.12}$$

By re-arranging the datasets, we get three equations for calculating the calibration factor.

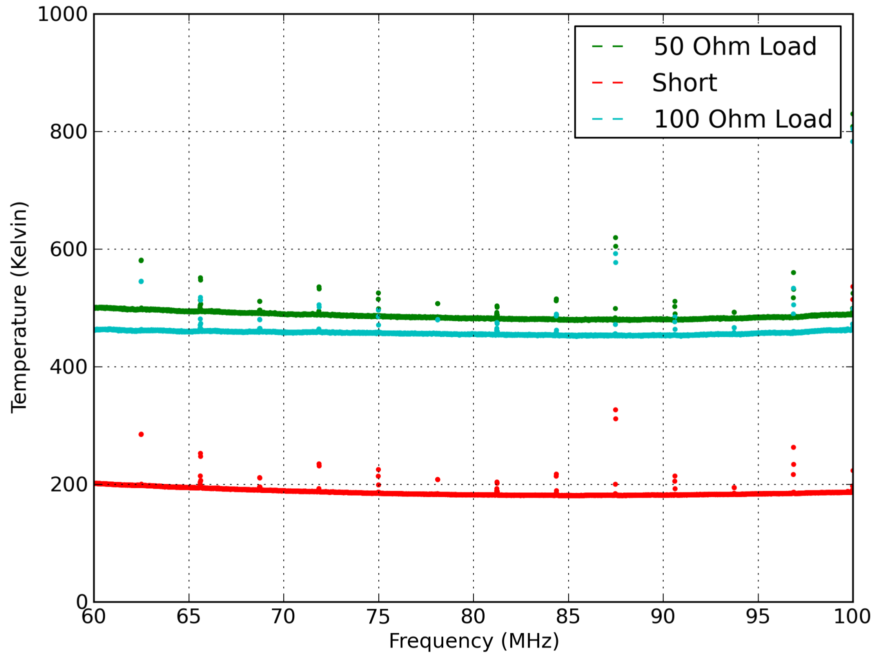


Figure 5.12: Calibrated high frequency resolution time average of all the calibration datasets for June 3rd, 2013.

Calibration Equation 1:

$$K_{JNC} = \frac{T_{amb}}{P_{50\Omega} - P_{short} - P_{Z50}} \tag{5.13}$$

Calibration Equation 2:

$$K_{JNC} = \frac{T_{amb}}{P_{100\Omega} - P_{short} - P_{Z100}} \tag{5.14}$$

Calibration Equation 3:

$$K_{JNC} = \frac{T_{Noise}}{P_{NS} - P_{50\Omega}} \tag{5.15}$$

We want to use one (or more) of these equations for our calibration. The noise temperature $T_{Noise} = K_{JNC} * P_{Noise}$ can be measured in the lab prior to installing the noise source, which would allow us to calculate K_{JNC} using Calibration Equation 3. However, if we can't get a high enough quality measurement of P_{Noise} then we will need to use either Calibration Equation 1 or 2.

In order to use Calibration Equations 1 or 2, we need two things. First, we need to know the ambient temperature $T_{amb} = K_{JNC} * P_{amb}$. Second, we need to know the current noise term P_{Z50} or P_{Z100}. The ambient temperature can be measured by installing a temperature sensor near the system, or approximated using a constant ambient temperature. Since Figure 5.1 shows $P_{50\Omega} \approx P_{100\Omega}$, we know that the current noise term is small. This allows us to run the calibration using Calibration Equation 1, with $P_{Z50} \equiv 0$ and $T_{amb} \equiv 300$K.

Thermal Noise Contribution

For all of these Calibration Equations I've dropped the thermal noise term ($P_{thermal}$). I can do this because thermal noise is random Gaussian fluctuations in the signal whose magnitude is defined by the radiometer equation ($P_{thermal} = P_{total}/\sqrt{\Delta\nu\Delta t}$), where $\Delta\nu$ is the frequency resolution of the data in Hz and Δt is the integration time in seconds [31].

Therefore, thermal noise can be lowered by averaging many calibration datasets (increasing Δt). Figure 5.11 shows the averaged data for one full day. Once we've taken the averages, we can further lower the thermal noise contribution by fitting the averages to a simple polynomial in frequency over a large frequency band, increasing $\Delta\nu$. We used $P = b_0 + b_1\nu + b_2\nu^2$ for our fit for the frequency band of $50 \leq \nu \leq 100MHz$. Figure 5.12 shows the average of the calibration datasets after Calibration Equation 1 is used to calculate K_{JNC}. Note that our assumption that $P_{Z50} = 0$ is approximately, but not exactly, correct since $P_{50\Omega} \neq P_{100\Omega}$.

Using K_{JNC} to Calibrate

To apply K_{JNC}, it is necessary to assume that $e = 1$ and use the measured η for the system in Equation 5.8. We'll neglect $P_{thermal}$ because we'll remove it later by taking time averages of the antenna data. We will also use the assumption that $P_{Z_{ant}} = 0$, just as we used $P_{Z50} = P_{Z100} = 0$. Figure 5.14 shows the calibrated average of one day of antenna data and the reference spectra used for calibration under these assumptions.

Current Noise Contribution

In order to include a non-zero P_Z component in the calibration, we can use the difference between $P_{50\Omega}$ and $P_{100\Omega}$. $P \propto |Z|^2\eta_Z$, where η_Z is the transmission efficiency if the amplifier is attached to a 50Ω or 100Ω terminator. By taking the difference $P_{50\Omega} - P_{100\Omega}$, we can calculate a proportionality constant (A) using Equation 5.16.

$$A = \Big[\frac{P_{50\Omega} - P_{100\Omega}}{\eta_{50\Omega} - 4\eta_{100\Omega}}\Big] \tag{5.16}$$

Using the proportionality constant (A) we can calculate $P_Z = A|Z|^2\eta_Z$, which can be plugged into Calibration Equations 1 and 2. We can then calculate the contribution to the system noise from this correction. This is done by comparing T_{short} for the

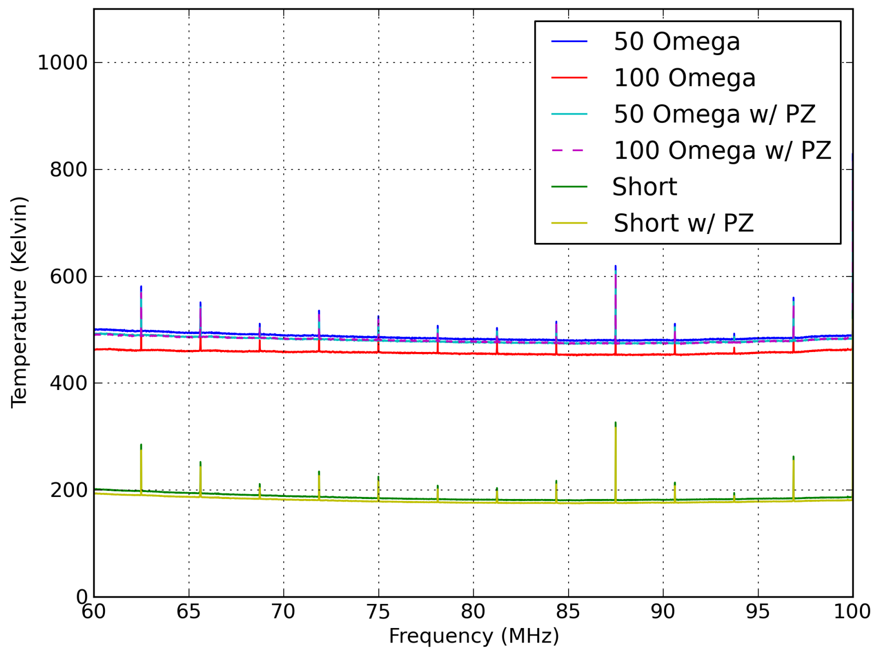


Figure 5.13: Comparison calibrated high frequency resolution time averages of all the calibration datasets for June 3rd, 2013 before/after including the current noise contribution to the calibration.

original and modified calibration. The modified calibration gives a lower T_{short} across the entire frequency band of $60 \leq \nu \leq 100MHz$. This value is lower for all days of data and the difference has an average magnitude of $\Delta T_{short} = 8.359 \pm 1.020$ Kelvin. Figure 5.13 matches Figure 5.11 with the addition of data calibrated with a calculated P_Z.

Daily Variance with GSM Modeling ($K_{\Delta GSM}$)

Johnson Noise calibration (K_{JNC}) assumes that $e = 1$ and η is measured accurately. Calibration using an astronomical source removes the necessity of this assumption. From our GSM model, we have a predicted sky temperature (T_{GSM}). Meanwhile, we have defined P_{Sky} in Equation 5.6. This allows us to calculate $K_{\Delta GSM}$, where we have dropped the $P_{Z_{ant}}$ and $P_{thermal}$ terms and allowed e to be incorporated into $K_{\Delta GSM}$.

$$K_{\Delta GSM} = \frac{T_{GSM}}{P_{Sky}} = \frac{T_{GSM}\,\eta}{P_{Ant} - P_{Short}} \tag{5.17}$$

In order to maximize the accuracy of T_{GSM}, it is better to use the combination of data from a full sidereal day rather than a single time step and remove the time

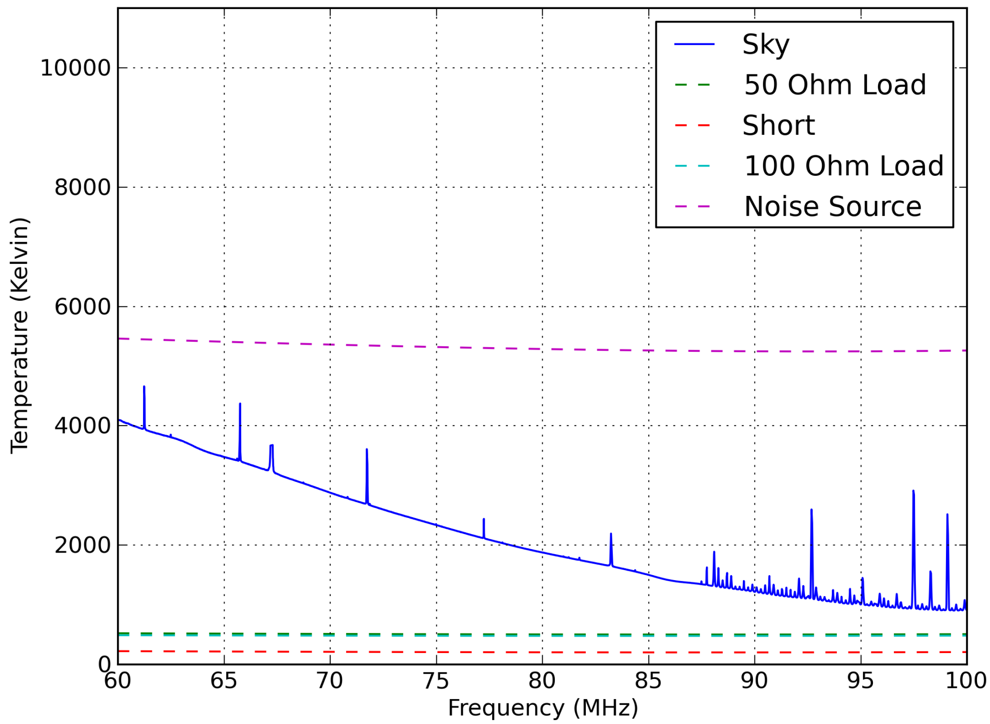


Figure 5.14: JNC calibrated high frequency resolution time average of all the data taken on June 1st, 2013; along with the fits for the average calibrated datasets.

independent component of the data. A χ^2 fitting can then be done for each frequency independently to get a $K_{\Delta GSM}$ value for that frequency. The χ^2 fit equation is Equation 5.18 where $\Delta T_{Sky}(\nu, t) = T_{Sky}(\nu, t) - \langle T_{Sky} \rangle_{DAY}(\nu)$ and $\Delta T_{GSM}(\nu, t) = T_{GSM}(\nu, t) - \langle T_{GSM} \rangle_{DAY}(\nu)$.

$$\chi^2(\nu) = \sum_t \left[\Delta T_{Sky}(\nu, t) - \Delta T_{GSM}(\nu, t)\right]^2 \tag{5.18}$$

This calibration strategy can be considered analogous to the traditional radio calibration strategy where the telescope points on and off a well known source. Once the fit has been calculated, it can be applied to the data (as shown in Figure 5.15). Optimal use of the calibration strategy requires a full day of data. When the calibration strategy is applied to a dataset where less than a full day of data is available, inaccuracies in the simulated beam or GSM model have a larger impact on the calibration. This can be seen in Figure 5.16, where the magnitude of $K_{\Delta GSM}$ at a particular frequency is clearly different when less of the day's data is available.

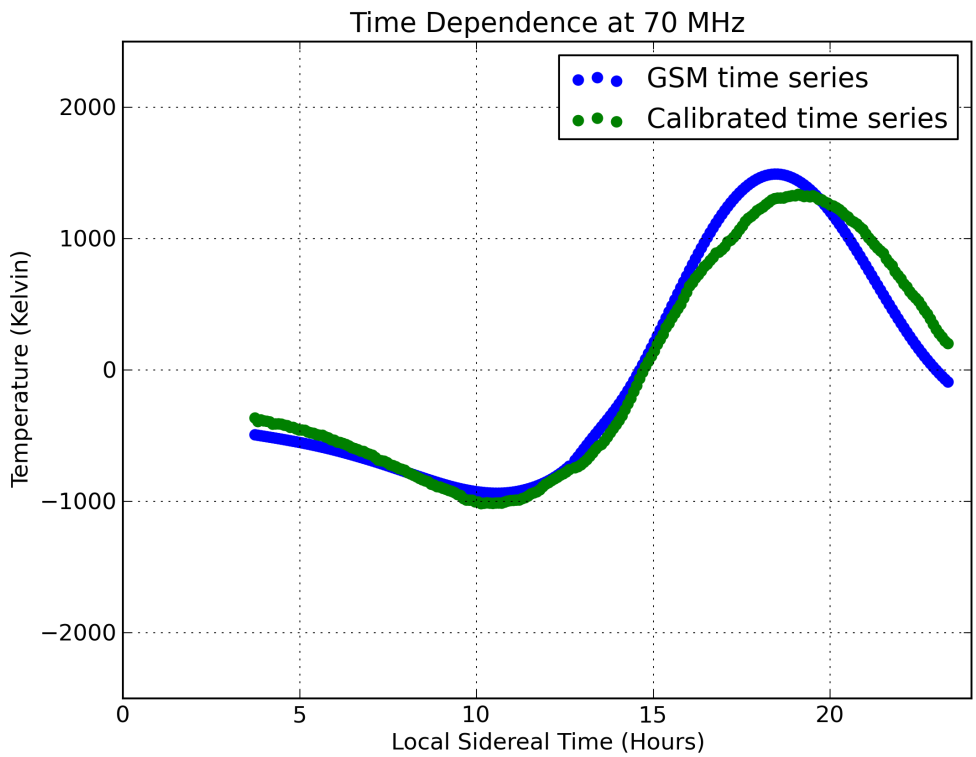


Figure 5.15: Single day fit of the $K_{\Delta GSM}$ calibration term with data collected on June 6th, 2013 at $70MHz$.

5.3 Removing the Foregrounds

5.3.1 Polynomial Fitting

Once the data has been calibrated, the Milky Way Galaxy and other foregrounds have to be removed to reveal the 21-cm structure. This removal is possible thanks to the structural simplicity of the foreground temperature ($T_{GM}(\nu)$), namely $T_{GM}(\nu) = \langle T_{Sky}(\nu, t)\rangle_{DAY} - \delta T_b(\nu) - T_{resid}(\nu)$. We can model $T_{GM}(\nu)$ as a simple polynomial, as shown in Equation 5.19.

$$log_{10}T_{GM}(\nu) = \sum_{k=0}^{n} a_k \Big[log_{10}\Big(\frac{\nu}{70MHz}\Big)\Big]^k \tag{5.19}$$

A $n = 2$ polynomial captures the expected foreground brightness temperature at the log-center of the frequency band (a_0), a power law spectral shape (a_1), and a synchrotron self-absorption correction term (a_2). When we use the GSM calibration, the foreground values for the June 2013 data are $a_0 = 3.3826 \pm 0.1825$, $a_1 = -2.3747 \pm$

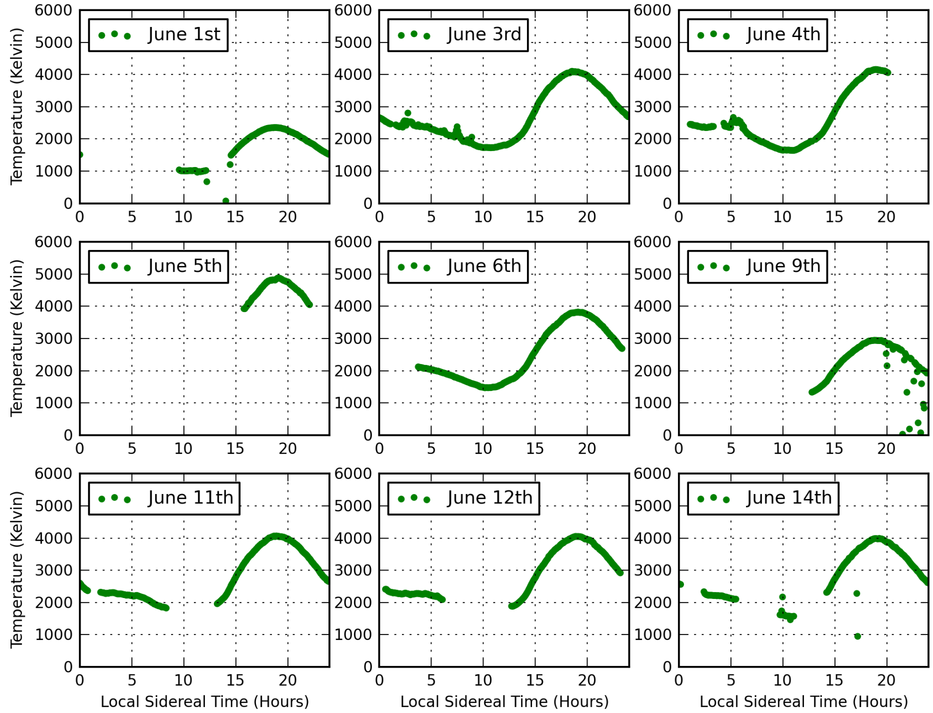


Figure 5.16: Fits for $K_{\Delta GSM}$ calibration term for multiple days in June 2013 at $70MHz$. Days where a smaller percentage of the data is available have poorer fits.

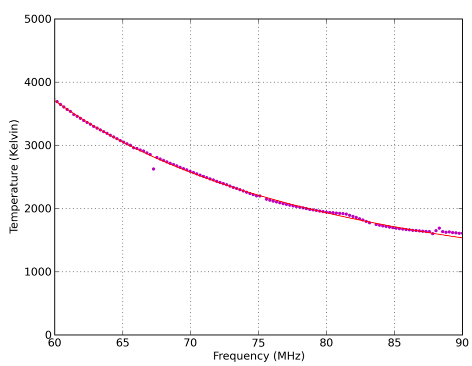


Figure 5.17: Daily mean of data from June 4th, 2013 calibrated using $K_{\Delta GSM}$, along with the foreground fit polynomial calculated for the data.

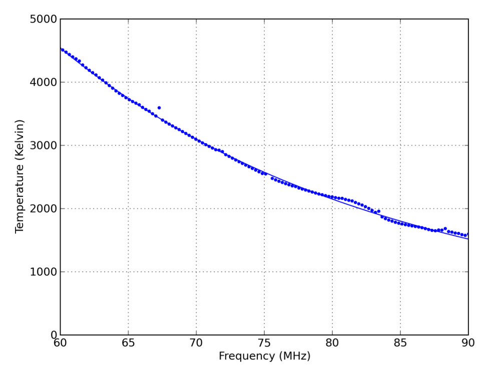


Figure 5.18: Daily mean of data from June 4th, 2013 calibrated using K_{JNC}, along with the foreground fit polynomial calculated for the data.

0.3178, and $a_2 = 0.3903 \pm 1.9351$. When we use the JNC calibration, the foreground values for the June 2013 data are $a_0 = 3.5186 \pm 0.0432$, $a_1 = -2.6205 \pm 0.0533$, and $a_2 = -2.5400 \pm 1.0962$.

Some of the variance in the foreground fit parameters comes from the fact that the daily averages are calculated for different fractions of a full 24 period depending on the day selected. Additional variance is seen in the GSM calibration parameters due to the weakness of the calibration strategy when less than a full day of data is observed. This was previously discussed in Section 5.2.4.Once we subtract this polynomial, we are left with residuals $\Delta T(\nu) = \delta T_b(\nu) + T_{resid}(\nu)$. Figures 5.17 and 5.18 show the daily mean of one day of data from June 2013 for each calibration strategy, along with its corresponding polynomial fit.

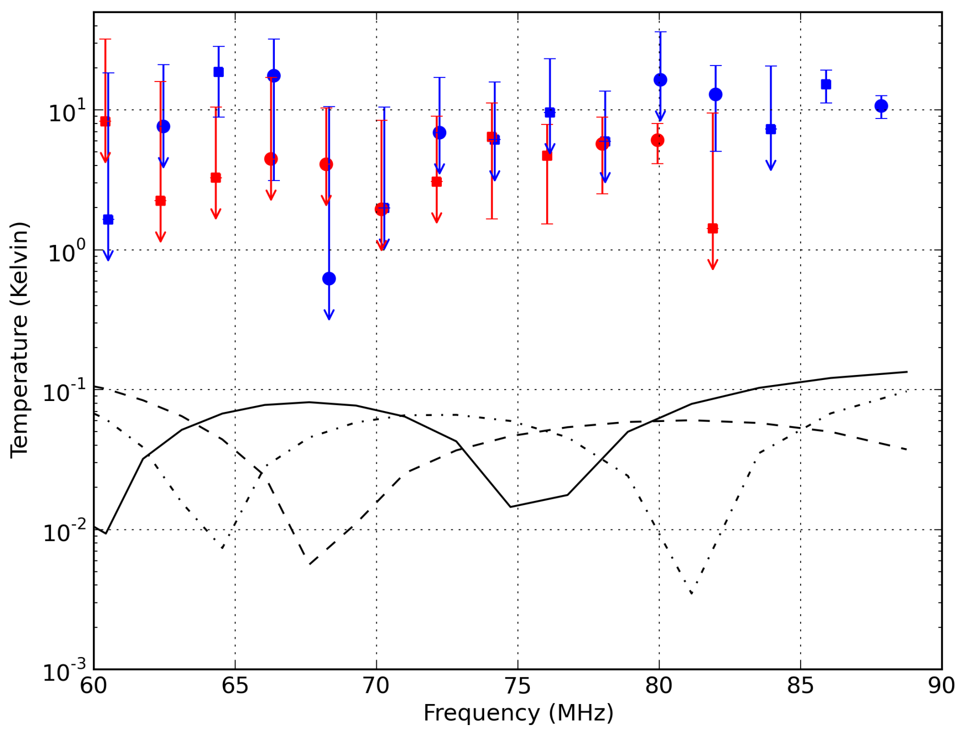


Figure 5.19: Log magnitude average and variance of the daily residual signals measured using data calibrated with either the Galaxy calibration (red) or Johnson Noise calibration (blue). Residual data is compared to the log magnitude of different models of the 21-cm or δT_bsignal (black). Model data shown has the frequency mean of δT_b for $60 - 90 MHz$ subtracted to mimic the effect of foreground removal. This means that in the model data, the two minima in each model line correspond to the two sides of the absorption dip in the δT_b spectrum for that model.

5.3.2 Daily Residuals

We calculate and apply the foreground fit separately for each daily mean. This gives us a distribution of residual data at each frequency for the entire dataset. Measurement of the δT_b spectrum requires residuals whose dominant component is the 21-cm structure. Since theoretical models of δT_b have a maximum structure $\sim 100mK$, residuals must reach this limit for δT_b models to be constrained. The distributions of daily residuals from the June 2013 data are too large to constrain the theoretical models, as is shown in Figure 5.19.

Current residuals are $T_{resid} \sim 1 - 10K$, which is still much smaller than the foreground signal ($T_{Sky} \sim 1000K$) at these frequencies. The residuals are the lowest published results in this frequency band and demonstrate the potential of the SCI-HI system for future measurements [37].

5.3.3 Frequency Limitations

In order to achieve reasonable residuals, it was necessary to limit the frequency range of the analysis to avoid frequencies where T_{resid} is large due to RFI from the SCI-HI system or other external sources. In the data collected in June 2013, the usable frequencies were limited to $f \sim 60 - 85MHz$. Below $60MHz$, there was additional large scale structure in the spectra, which may be due to either ionospheric impacts on the beam shape or inaccuracies in the simulated HIbiscus beam compared to the actual beam. Above $85MHz$ there was contamination from both external FM signals and self-generated noise from the SCI-HI system.

5.4 21-cm Signal Attenuation in the SCI-HI Data

An important cross-check for any 21-cm measurement experiment is a simulation calibration measurement of signal attenuation [23]. This check is done by adding a simulated 21-cm signal to the raw data and then running the combination data through the signal pipeline. Figure 5.20 shows what happens when a simulated 21-cm signal, magnified to $100K$ to make it easy to identify, is run through the analysis pipeline.

Because the 21-cm signal is constant with time, with complex structure in frequency, there is minimal signal attenuation caused by the calibration strategy and foreground removal process.

5.5 Ionosphere Impacts to the SCI-HI Data

All of the data analysis above was done assuming no ionospheric impacts on the data. However, Figures 4.5 and 4.6 indicate that there should be some ionospheric impact on the data, particularly during the daytime hours.

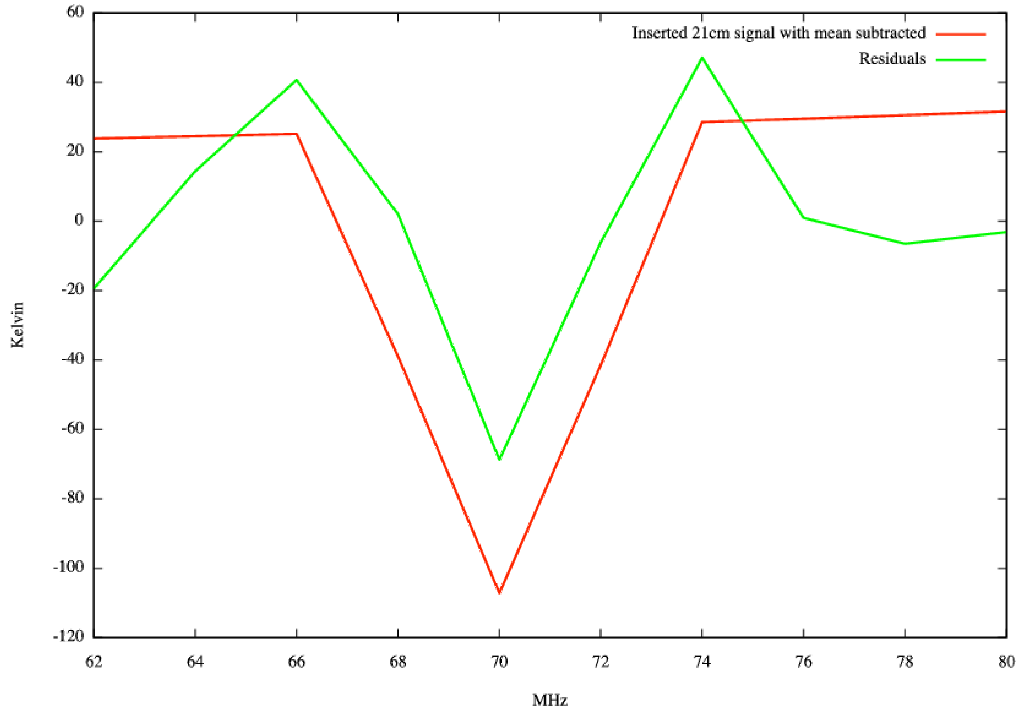


Figure 5.20: Simulated 21-cm signal with $100K$ magnitude (red) and residuals (green) after calibration and foreground removal.

5.5.1 Structure Due to Refraction

Since ionospheric refraction affects the direction of sky signals, refraction will lead to inaccuracies in T_{GSM}, particularly at large θ. These inaccuracies will propagate into the $K_{\Delta GSM}$ calibration factor, leading to inaccuracies in T_{Sky}.

Because we don't use a sky signal for calibration with the Johnson Noise datasets, the impact of refraction on K_{JNC} is zero. However, refraction may introduce frequency independent structure in e, which the Johnson Noise calibration strategy is incapable of measuring. This structure would then lead to inaccuracies in T_{Sky}, just as with $K_{\Delta GSM}$ calibration.

5.5.2 Structure Due to Absorption

Frequency dependent absorption can also introduce additional structure to the mean signal ($\langle T_{Sky}(\nu, t)\rangle_{DAY}$). If TEC is constant for the data, then the structure will simply add an additional term in T_{GM}. But if TEC is variable, then the absorption structure can be quite complex. This structure will also not be removed by either calibration strategy.

5.5.3 Identifying Ionosphere Structure

The residual levels shown in Figure 5.19 set a maximum level for the total ionospheric impact in the frequency band used ($f = 60 - 85MHz$). Ionospheric impact may be larger if the a_2 term defined in the T_{GM} polynomial is caused by the ionosphere rather than Galactic structure. It is possible to quantify the impact using the time dependence of a_2. If a_2 is from the ionosphere, its magnitude should increase with higher TEC. On the other hand, if a_2 is from Galactic structure, then a_2 should depend on the part of the Milky Way Galaxy currently in the beam.

However, self-generated RFI and time gaps in the SCI-HI data from June 2013 add time dependent structure to a_2 with a magnitude larger than any potential ionospheric effect. Therefore, ionospheric effects cannot be detected in the June 2013 data. However, data collected with an improved system should have significantly less self-generated RFI. This may make it possible to detect ionospheric structure in a_2 and/or T_{resid}.

Chapter 6

Looking Toward the Future

6.1 Deployment of the Improved SCI-HI System

Current limits with the SCI-HI system from the data collected in June 2013, as shown in Chapter 5, are above the threshold needed to measure the 21-cm signal during the Cosmic Dawn. In addition, the results only cover a relatively narrow frequency band. Limits are currently dominated by contamination from self-generated RF and RFI in the FM band.

Self-generated RF comes from instrumental defects including insufficient Faraday Cage shielding, limited antenna bandwidth, and power sources (batteries) that last less than 24 hours. These issues can can be addressed by building improvements to the SCI-HI system. Improvements include the new HIbiscus antennas discussed in Section 2.2.4 to expand the antenna bandwidth. Also, additional Faraday Cage shielding, discussed in Section 2.4.5, and a new power source have been added to the system. This new power source is a gasoline fueled generator, which is more reliable than batteries, as discussed in Section 2.3.4. In addition, the data processing system computer has been upgraded since the June 2013 deployment. The new system has a duty cycle of 30%, compared to the previous duty cycle of 5-10%.

We plan to deploy this updated system to the SKA site in the Karoo desert in April 2015, as discussed in Section 3.4.2. During the Karoo deployment, data will be collected for full 24 hour cycles using each of the two HIbiscus antennas. This will allow us to optimize the Milky Way Galaxy calibration strategy and utilize a larger range of frequencies. Given the improvements to the SCI-HI system, we expect that self-generated RF contributions to the sky temperature will be decreased to levels below the 21-cm signal. In addition, the higher duty cycle will allow us to collect enough data to bring $P_{thermal}$ below the 21-cm signal levels with only a few days of data.

However, the Karoo desert site is not expected to be sufficiently remote to allow use of the FM band. Therefore, we plan to deploy the SCI-HI system to Marion Island in April 2016, as discussed in Section 3.4.2. The RFI levels in the FM band at this site are expected to be below the current residual limits measured by the SCI-HI system.

Deployment at the Karoo desert should yield data with self-generated RF for $f < 88MHz$ to below 21-cm signal levels. Once this component in the data is removed, the residuals will be dominated by one of three components. These three components are: (a) ionospheric refraction and absorption, (b) variance in the spectral index of the Milky Way Galaxy, and (c) the 21-cm signal.

Data collected in the Karoo will allow us to identify which contribution (a, b or c) is dominant in the residuals between $(40 \leq f \leq 88MHz)$. If the 21-cm signal is the dominant residual in the Karoo data at those frequencies, we will make a first detection of the 21-cm signal. Otherwise, we will be able to place tighter constraints on the 21-cm signal. The level of these constraints will be set by the level of the dominant component in the residuals.

6.2 Development of Expanded Global 21-cm Signal Experiments

Funding for the Karoo and Marion deployments is part of a three-year grant from the South African National Antarctic Program (SANAP). The grant is also intended to support further development of the SCI-HI experiment, which will be called ″SCI-HI in the Sub-Antarctic″ (SHISA).

SHISA plans to continue improving the SCI-HI system design. Some potential areas of improvement include a multi-element version of the system for improved spatial resolution, full polarization data collection, and a more sophisticated data processing system. These developments will allow us to address contributions to the residuals from the ionosphere and foregrounds at levels beyond the capability of the current SCI-HI system.

One of the advantages to the Marion Island site is the potential for diminished ionospheric impacts, particularly during the Antarctic winter. As part of the SHISA project, While we are on the island in April 2016, we plan to deploy a system for evaluating the ionosphere at Marion Island's latitude during the entire year. This system will focus on frequencies $\sim 5 - 50MHz$, where the ionospheric impact is large. Using the frequency spectrum of the signal we will be able to identify the lowest frequency where we can still see the sky through the ionosphere as a function of time. This will allow us to better quantify the ionospheric impacts on global 21-cm experiments and other low frequency radio astronomy projects deployed to the island.

Appendix A

Teaching 21-cm Cosmology to the Public

Cosmology using the 21-cm line of Hydrogen is a relatively new concept in the world of astronomy. It relies on radio telescopes, which have only been in general use for $\sim 50-100$ years. Measurements of the 21-cm signal from our own Milky Way Galaxy and other nearby galaxies have been used extensively to study the neutral Hydrogen gas distribution around local galaxies.

On cosmological scales, where the 21-cm signals are $10^4 - 10^6$ times smaller than the foregrounds, the field is still in its infancy. There are many telescopes seeking to measure the 21-cm signal at redshifts ($z > 1$), but currently the results are at preliminary levels placing loose constraints on the signal. In the future, scientists hope to build full maps of the Hydrogen sky at many cosmological redshifts as has been discussed in Chapter 1.

However, to make such maps requires a significant amount of effort and funding and public approval and interest is key. Improving visibility and interest in 21-cm cosmology can be done through public outreach activities including lectures, open houses, and demonstrations. We decided to focus our efforts on public outreach through a planetarium show.

Planetariums are one of the most common outreach tools in astronomy. They bring the night sky and the universe to the public in a controlled environment where the spectacular beauty of the sky can be showcased. Historically, planetariums projected the night sky onto a circular dome using a single central projector such as a Zeiss projector. These projectors could produce accurate pictures of the night sky at different locations and times of the year.

Today, planetariums have evolved to use a set of video projectors placed around the base of the dome. These video projectors display digital stills and videos on the dome, greatly increasing the potential of the planetarium environment. Instead of simply showing the night sky, digital videos allow audiences to leave the earth and travel out into the universe in an immersive environment.

A.1 Hydrogen Sky Planetarium Show Overview

As part of a grant from the National Science Foundation(AST-1009615), we received a \$20,000 budget to create a $5-10$ minute planetarium show entitled ″The Hydrogen Sky″. This show would introduce audiences to astronomy using the 21-cm Hydrogen line and educate the public on the research supported by the grant.

To facilitate this project, I have been working with a team of animators and artists as show producer. I recruited two animators from Carnegie Mellon University's Entertainment Technology Center (ETC) masters program[1], Alexander Moser and Meng Zhang. I also worked with Tom Casey and Warren Casey from Home Run Pictures[2], a local Pittsburgh animation company which specializes in full-dome planetarium productions. Additional support was provided by the Carnegie Science Center's Buhl Planetarium[3] and its staff, specifically Frank Mancuso and Dr. Brendan Mullan.

A.2 Storyboard and Script

In order to complete the show, the first thing that we had to develop was a storyboard with visual queues to help the animators start creating content. From this first attempt at a storyboard, we identified a few key components for the show:

- Zoom out from the night sky to emphasize the connection to what we see every day.
- Show the large scale structure, then show gaps where we don't have data. This is meant to demonstrate the need for 21-cm maps.
- Have a sequence showing the interactions between Hydrogen atoms and 21-cm photons in a cloud of gas. Include both stimulated emission and absorption, as well as a discussion of atomic spin.
- Show photons traveling from the Hydrogen gas cloud to earth, being redshifted along the way.
- Have a sequence of photons interacting with the Green Bank Telescope and being collected in the computers to make maps.
- Make maps of the data from the Green Bank Telescope Intensity Mapping project and show the difference between the foregrounds and 21-cm maps.
- Show that the Green Bank maps are a very small fraction of the overall sky.

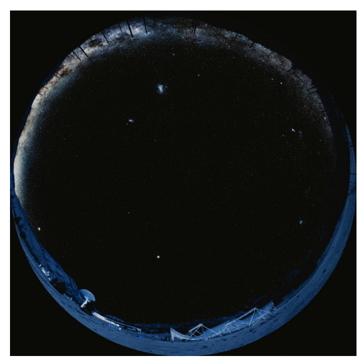
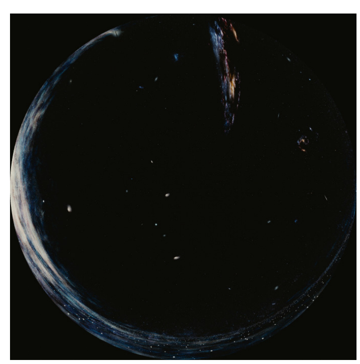
[1] `http://www.etc.cmu.edu/`
[2] `http://www.hrpictures.com/`
[3] `http://www.carnegiesciencecenter.org/planetarium/`

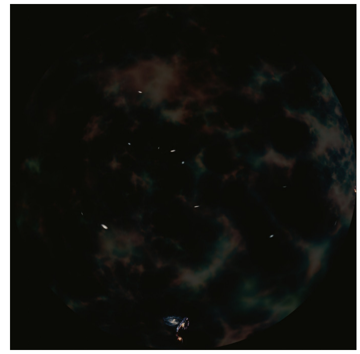
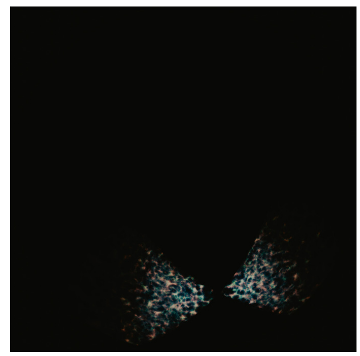
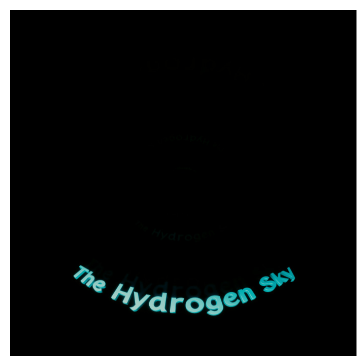
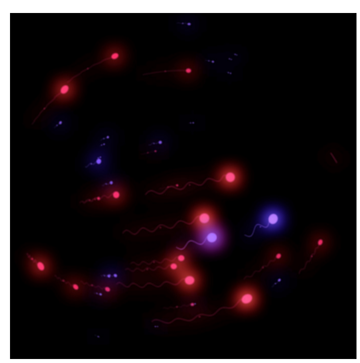
- Have some video of the Canadian Hydrogen Intensity Mapping Experiment (CHIME).
- Show how CHIME collects much more of the sky in a single day (aka drift scanning).
- Figure out a good conclusion to capture all of the science that can be done with 21-cm maps. Emphasize that CHIME is still under development.

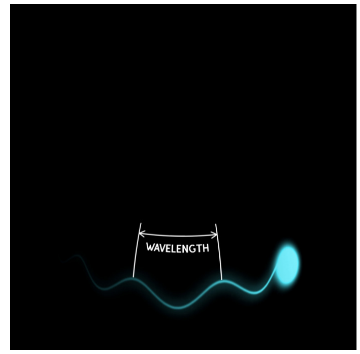
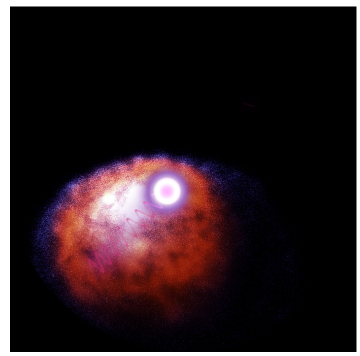
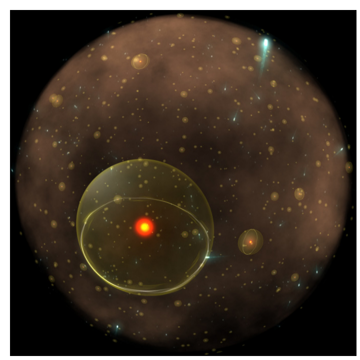
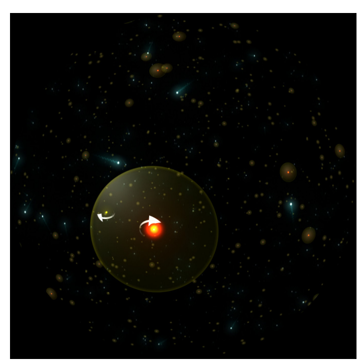
Once we had a general storyboard, we began refining it and writing a script that matched the visual components and told a clear story for a general audience. We went through a number of iterations on the script, trying to find a balance in communicating the important information without making the story too complicated.

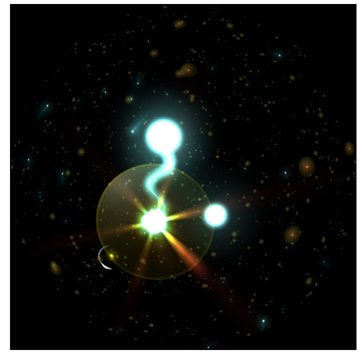
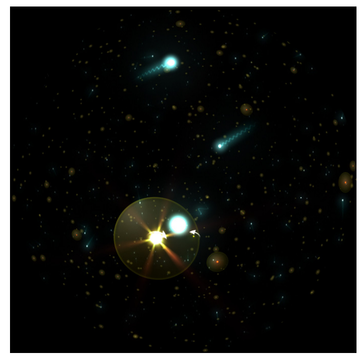
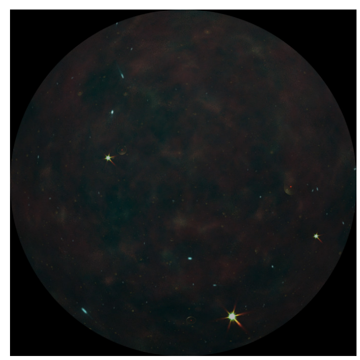
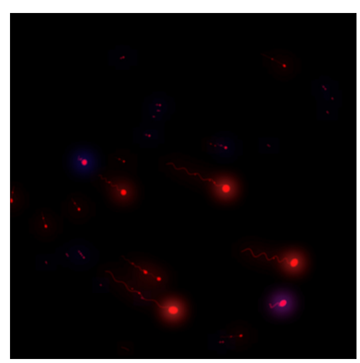
A.2.1 Final Script

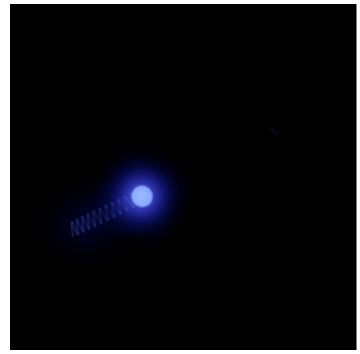
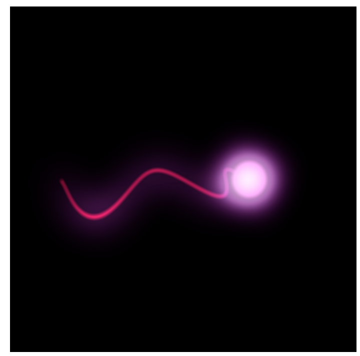
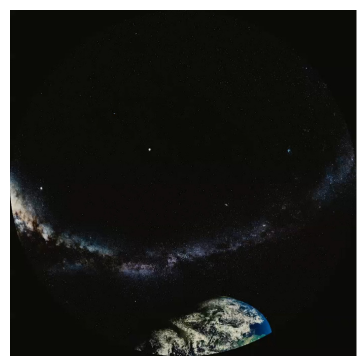
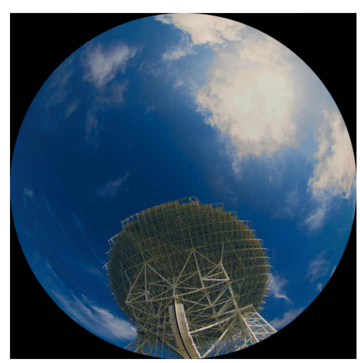
The final version of the script is laid out below. Along with each paragraph of the script is a single image snapshot from the portion of the video corresponding to that paragraph.

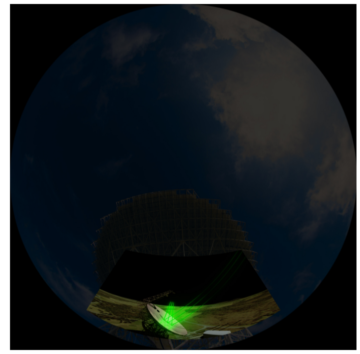
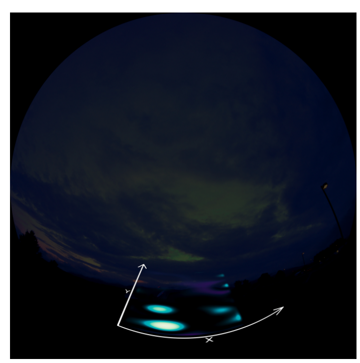
Script	Image
When we look out into the night sky we see the stars. But the stars we see are only a very small fraction of the stars in the universe.	
Beyond the hundreds of billions of stars that make up our own galaxy, the Milky Way; we see many other galaxies, each with hundreds of billions of stars of their own.	

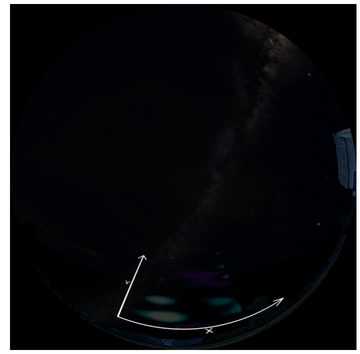
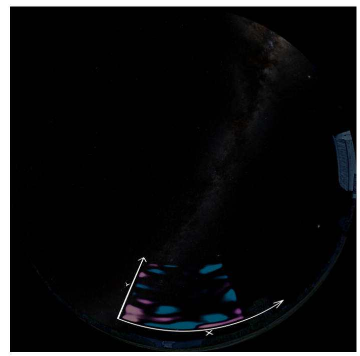
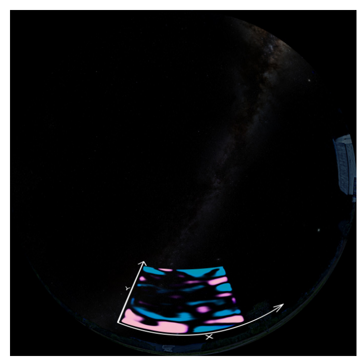
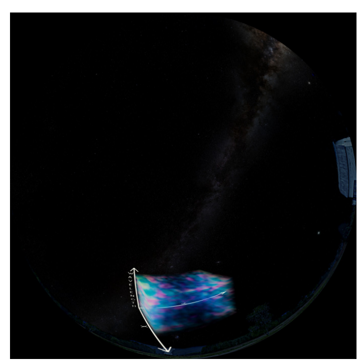
Script	Image
Because light has a set speed, the light from these more distant galaxies takes longer to reach Earth than the light from our own galaxy. Therefore, we are seeing these galaxies as they were long ago.	
If we keep looking further and further away, and therefore further back in time, eventually we stop seeing anything. This is because the very distant galaxies are not bright enough for us to see, or they haven't even formed stars yet.	
To study these very distant, and very young galaxies, scientists like me need a new way to see the sky, we need the Hydrogen Sky.	
From early times, photons of light were able to freely travel through the universe.	

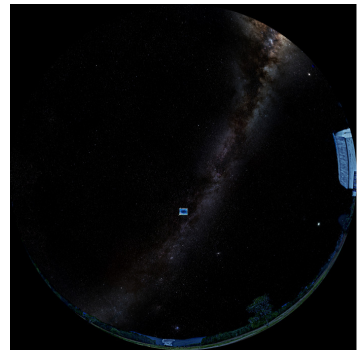
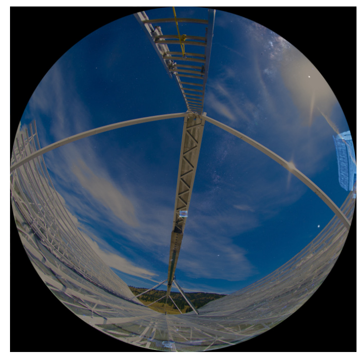
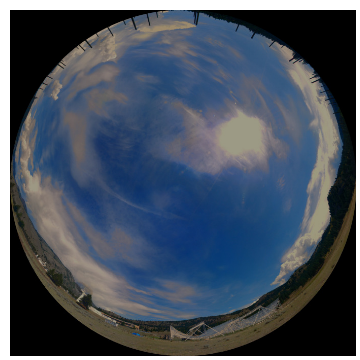
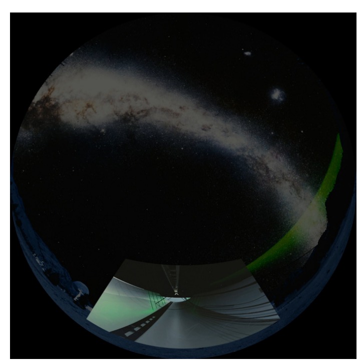
Script	Image
These photons each had a specific wavelength, or color. For some, that wavelength was 21 centimeters, or a little over 8 inches.	
As they traveled, photons passed through galaxies and the clouds of Hydrogen gas that surrounded the galaxies.	
Inside these clouds are many trillions of individual Hydrogen atoms. Each atom has a proton and an electron.	
Both the proton and the electron have a special property called spin. This spin has two directions, clockwise and counter-clockwise.	

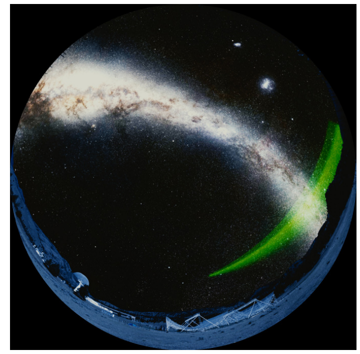
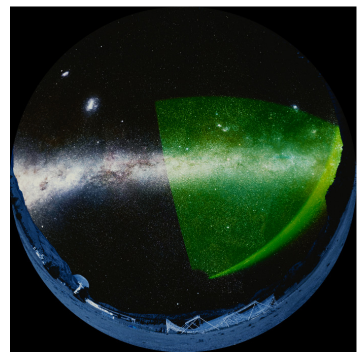
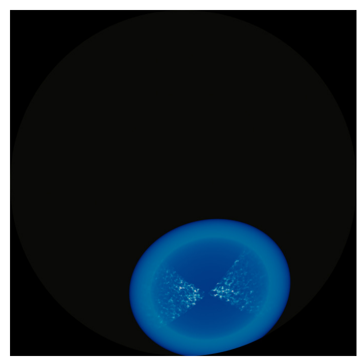
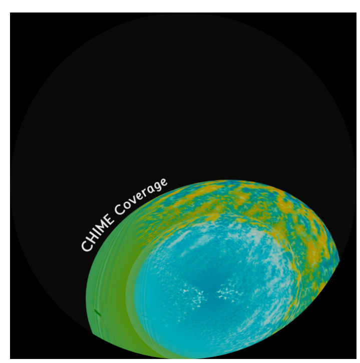
Script	Image
Photons sometimes collide with these Hydrogen atoms. When they do, one of two things can happen. If the proton and electron spin in the same direction, a spin flips, and two photons fly away.	
But, if the electron and proton spin in opposite directions, one of the spins flips and the photon is absorbed by the atom.	
Because of this, when light passes through a Hydrogen gas cloud the number of 21 centimeter photons changes.	
After passing through the Hydrogen gas cloud, the photons that remain continue to travel through the universe. As they travel, the universe expands around them.	

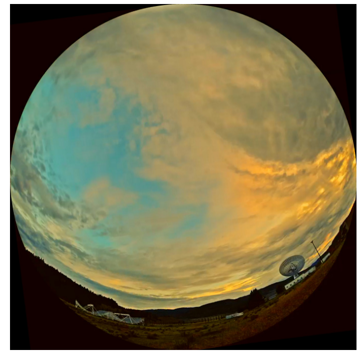
Script	Image
As the universe expands, or stretches, the wavelength of the photon also expands and grows longer.	
We call this expansion redshifting, because red light has longer wavelengths than blue light.	
Eventually the photons reach the Milky Way Galaxy and approach the Earth, where they are measured by radio telescopes.	
One such radio telescope is the Robert C. Byrd Green Bank Radio Telescope in southern West Virginia.	

Script	Image
At the Green Bank Telescope, light is captured by a receiver located above the huge dish of the telescope.	
From there it is sent down along cables into nearby computers. Once inside the computers, the light is chopped up in time, and then further separated into groups by wavelength. In this way, we can measure a spectrum of light at each second.	
The telescope looks at a different part of the sky in each second. So we build an image of the sky, one point at a time.	
The brightest objects in the sky are nearby things, like the stars in the Milky Way Galaxy. If we want to see dimmer things, like the photons from Hydrogen clouds around distant galaxies; we have to remove the bright stuff.	

Script	Image
We can do this thanks to redshifting, which changes the wavelength of the photons from the Hydrogen clouds.	
Unlike the bright nearby stars that produce light at all wavelengths; photons from the Hydrogen clouds have a specific wavelength.	
More distant clouds have their photons wavelengths redshifted to longer wavelengths than closer clouds. In this way, we can separate the light from the Hydrogen sky from the light from nearby bright sources.	
This allows us to build a three-dimensional map of the Hydrogen sky. In this map the longer, and more redshifted wavelengths show us the Hydrogen sky around younger and more distant galaxies.	

Script	Image
We spent over one hundred hours with the Green Bank Telescope collecting the light for this map. But the map only covers a very small fraction of the overall sky. To make maps of the whole sky we need to find faster ways to measure this light.	
One way to do this is to build new radio telescopes which can look at multiple parts of the sky at the same time. The Canadian Hydrogen Intensity Mapping Experiment, or CHIME for short, is one such new telescope.	
Located at the Dominion Radio Astrophysical Observatory, near Penticton in British Columbia, Canada; CHIME has cylinders instead of a single, large dish.	
With many receivers along the center line of each cylinder, CHIME collects light from a stripe of the sky instead of a single point.	

Script	Image
Meanwhile, the rotation of the Earth causes the cylinder to see a slightly different stripe of sky each second.	
During the course of a day, photons of light are collected for the entire sky as seen from that location; which is about half the overall sky. So maps of the Hydrogen sky can now be made much more quickly than with the Green Bank Telescope.	
In the future, we will use the maps of the Hydrogen sky in combination with observations from other telescopes to fill in gaps in our picture of the universe.	
This will allow us to study many things including the basic building blocks of the universe, young galaxies, and even the first stars in the universe; which lived and died before the Earth existed.	

Script	Image
Because there are so many interesting scientific applications for the Hydrogen sky, new radio telescopes such as CHIME are currently under development around the world. With these new telescopes the future is bright for the Hydrogen sky.	

A.3 On-site Filming

One of the key components of this show was on-site filming at a couple of real radio telescopes, the Robert C. Byrd Green Bank Radio Telescope (GBT)[4] and the Canadian Hydrogen Intensity Mapping Experiment (CHIME)[5]. For filming, we focused on still images and time lapses rather than video footage. We found that video footage did not translate well into the dome environment, as any irregular motion in the videos was magnified by the shape of the dome.

We used two Nikon DSLR cameras (an old D80 and a new D800), with a rectangular lens on the D80 and a circular fisheye lens on the D800. The D80 was used primarily to capture images for creating panoramas of the locations we visited, while the fisheye was used for capturing single images and doing time-lapse videos.

A.3.1 Green Bank, WV

I visited the National Radio Astronomy Observatory (NRAO) Green Bank site twice to gather footage in June and August 2014. For the June visit, Alex and Meng accompanied me to the site. We were able to stay on-site at the telescope and get footage both from the general site and from up close to the main 100m Green Bank Telescope. Footage had to be collected during the telescope maintenance time, as the cameras are a source of radio frequency interference (RFI) for the telescope when it is actually running.

Mike Holstine, the business manager at the Green Bank site, was our primary contact for the trips. He was able to get us a tour of the Robert C. Byrd Green Bank Radio Telescope, including taking two elevators up to the receiver on top of the

[4]`https://science.nrao.edu/facilities/gbt/`
[5]`http://chime.phas.ubc.ca/`

Figure A.1: Alex and Meng checking an image during filming.

Figure A.2: Alex, Meng and Hsiu-Hsien taking multiple shots of the GBT from its base.

Figure A.3: Some of the monitors in the GBT control room, captured with the rectangular lens.

Figure A.4: GBT as seen from the observation deck, captured with the rectangular lens.

telescope. Figures A.1 and A.2 show us filming at the GBT, while Figures A.3, A.4, A.5, and A.6 show some of the images we were able to capture at the site with both the rectangular and fisheye lenses.

A.3.2 Penticton, BC, Canada

In June 2014, Alex and I visited the Dominion Radio Astrophysical Observatory (DRAO) site near Penticton in British Columbia, Canada where the CHIME telescope

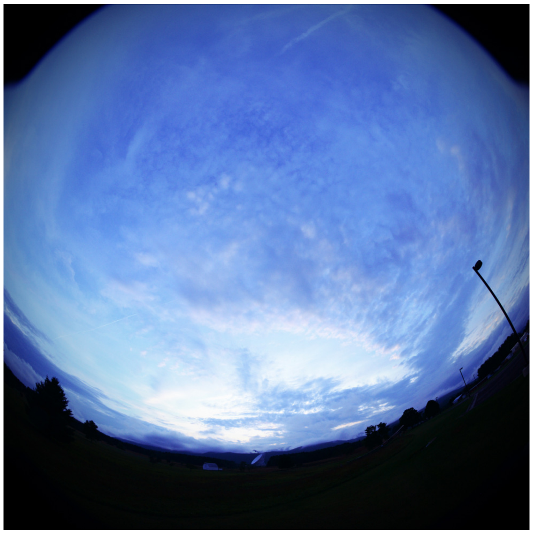


Figure A.5: GBT as seen from the observation deck, captured with the fisheye lens.

Figure A.6: GBT as seen from its base, composited from images with the fisheye lens.

Figure A.7: Tabitha on top of one of the CHIME cylinders during the DRAO site visit.

Figure A.8: Starling nest hollowed out of an RFI antenna on one of the CHIME cylinders.

pathfinder is located and the full CHIME telescope will be built. We stayed locally in the city of Penticton and were able to capture footage despite battling with rain for part of our visit. We were able to capture footage while the CHIME team was working on the computer system so that we didn't interfere with observations.

Permission to film at the site was provided by Tom Landecker, the primary CHIME contact at the DRAO, as well as the rest of the CHIME team. David Hanna from McGill University was able to show us around the site and get us access during off

Figure A.9: CHIME pathfinder cylinders, captured with the rectangular lens. CHIME control room is inside the shipping container on the left side of the cylinders.

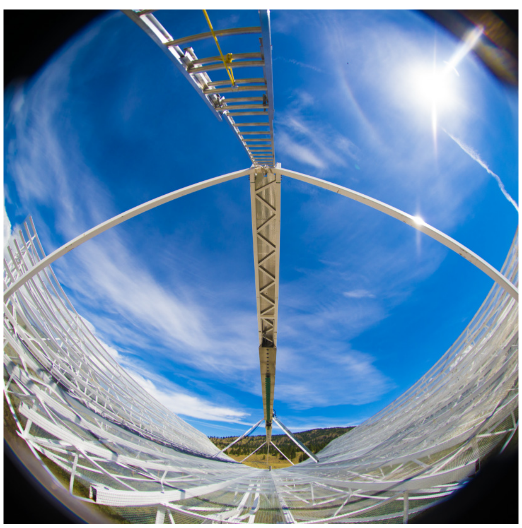


Figure A.10: Close up view of one of the CHIME cylinders, captured with the fisheye lens.

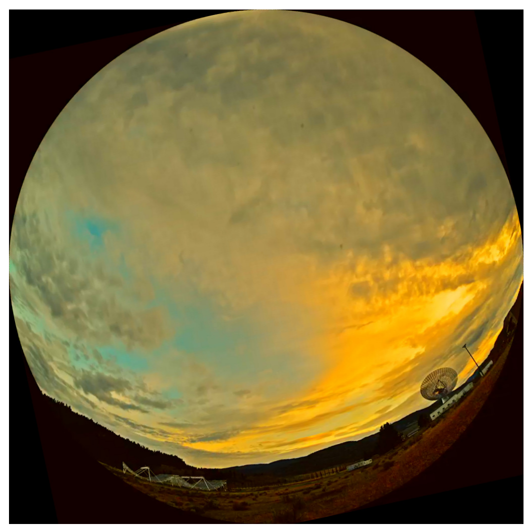


Figure A.11: CHIME cylinders at dusk, along with the neighboring 26 m dish, captured with the fisheye lens.

hours so that we could film around dusk. Figure A.7 shows telescope during filming, while Figures A.9, A.10, and A.11 show some of the images we were able to capture at the site with both the rectangular and fisheye lenses.

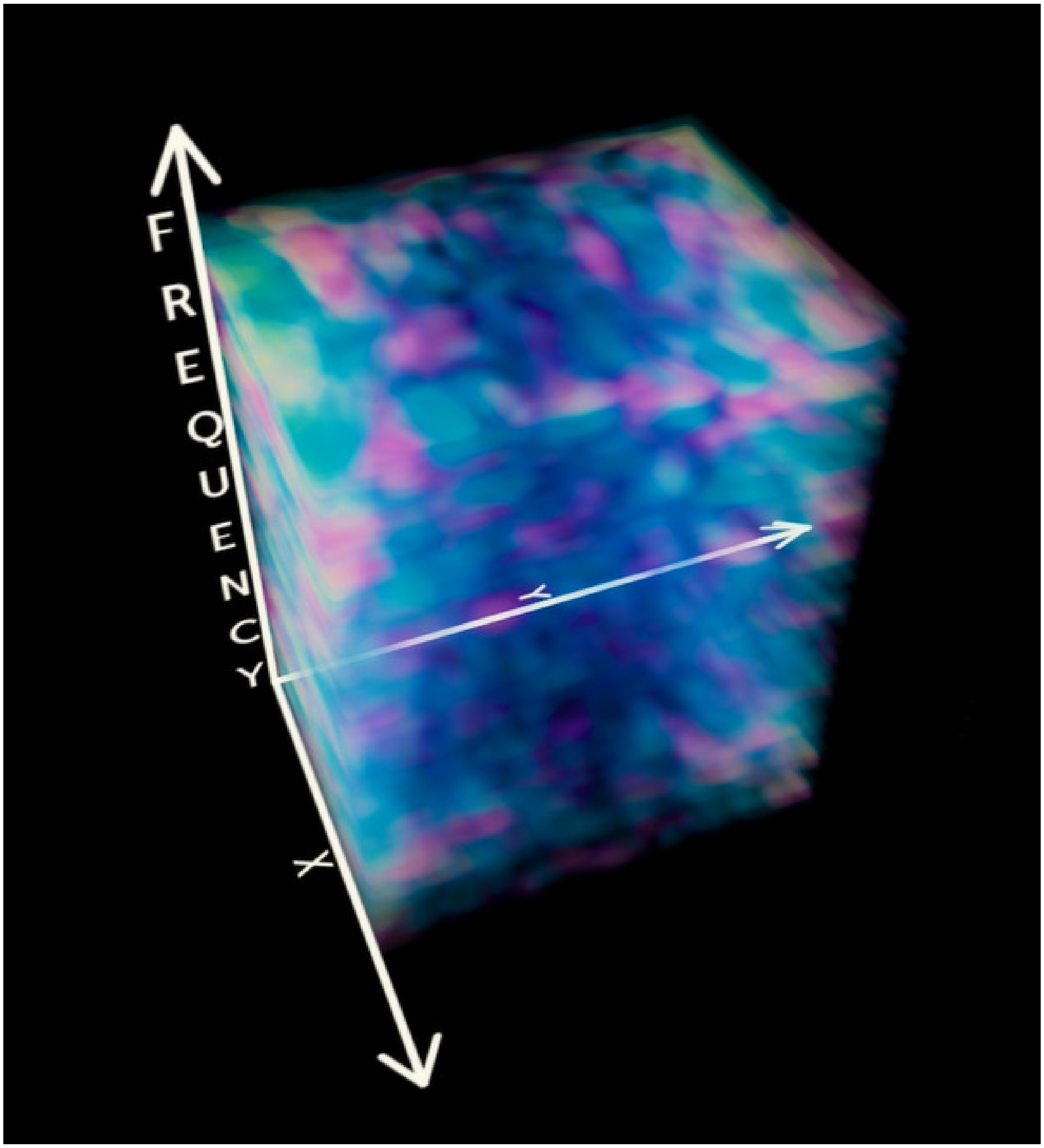


Figure A.12: Map of the Green Bank Telescope Intensity Mapping project ″deep″ field after foreground subtraction. In the image, X and Y are the Right Ascension and Declination of the data (centered at $14^h31^m28.5^s$ RA and $2°0'$ Dec).

A.4 21-cm Data Maps

Because the show is all about the Hydrogen sky, we wanted to be able to show images of that sky. However, there isn't currently a large map of the 21-cm sky at cosmological redshifts. Instead, we decided to use a combination of data and simulation to show what the Hydrogen sky should look like.

We started with some of the data from the Green Bank Telescope Intensity Mapping project [19][32], which looked at a small part of the sky and made maps of the 21-cm signal with foregrounds. The project used principal component analysis to remove most of the foregrounds, leaving residuals that contain the 21-cm signal. Figure A.12 shows a 3D representation of one of these foreground subtracted maps (the 15hr ″deep″ field), which contains the Hydrogen sky for an area of a few square degrees.

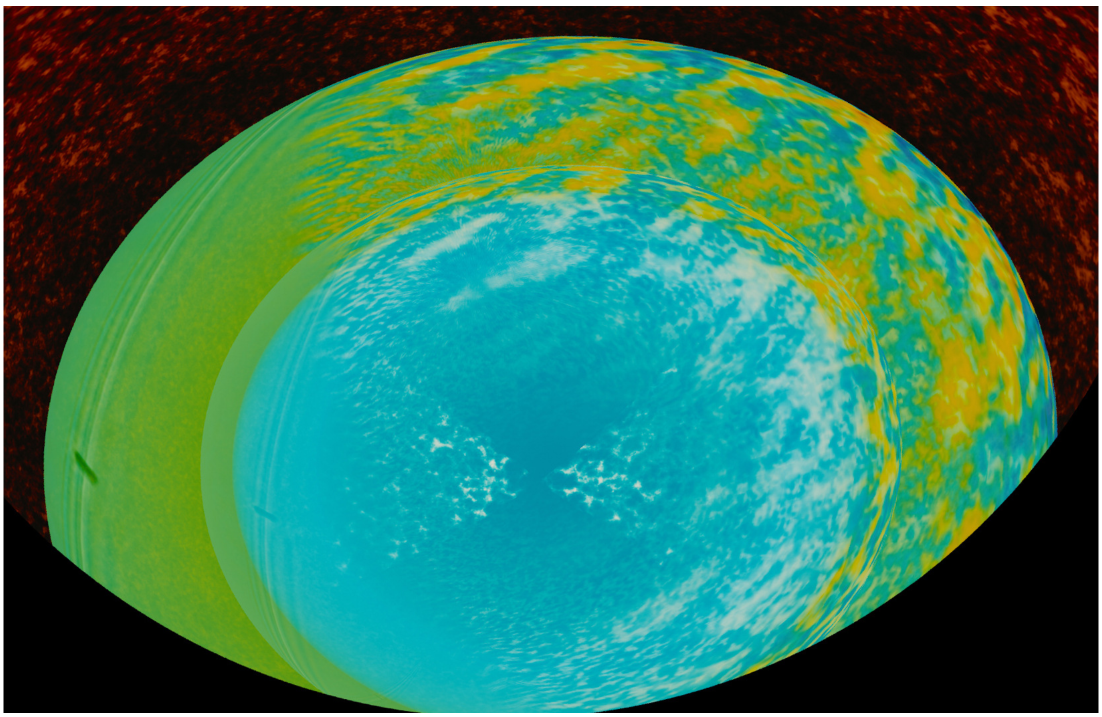


Figure A.13: Simulated CHIME map in green with blue and yellow temperature variation placed in a shell for the planetarium show. Image center has a model of the large scale structure we've seen so far with optical telescopes, and the red structure along the edge represents the Cosmic Microwave Background.

For the CHIME segment of the show, we decided to use simulated data to show what we expect the Hydrogen sky data to look like. We were given simulated CHIME 21-cm maps made by Richard Shaw [28] to use to create images. These simulations are of the 21-cm residuals after foreground removal, in the same way that the GBT maps are data residuals after foreground removal. Figure A.13 show the simulated CHIME maps as they are used in the planetarium show, placed in a shell corresponding to the redshifts that the maps cover.

A.5 Working in the Planetarium Environment

One of the big challenges of this project was learning how to work in a planetarium environment, and how it differs from normal video production. Typical planetariums have a ″2k″ or ″4k″ resolution, with an area of roughly 4-8 million square pixels. This means that any images have to be captured at very high resolution, and any animation takes a long time to render (can take days for a few seconds of data).

A.5.1 Image Distortion

Beyond sheer size, the planetarium environment also distorts images projected on it because of the shape of the dome. The dome shape means that viewers focus is primarily on the bottom third of the image, while the top third of the image is barely seen since it is behind the viewers. Anything placed along the edge of the image is expanded, while anything in the center of the image is shrunk compared to how it looks on a flat screen. You can see this in Figures A.5 and A.11, where the telescopes sit in the bottom of the image.

Image hue and brightness is also different from a regular monitor when viewed on the planetarium screen. If the image is too white or bright, the brightness bleeds into the parts of the image that is supposed to be dark and the image looks washed out. Colors generally tend to look washed out on the dome, so images prepared for the dome should have their saturation cranked up high. In other words, the images should look like they were created with crayons. This can be seen in Figure A.14, where the blue of the sky is much more intense than in the original images.

A.5.2 Motion

Motion reads differently in a dome environment than it does on a regular screen. Because the dome surrounds so much of a person's vision, if anything is moving too fast it tends to make the viewers nauseous. This means that motion should be at least half as fast as you would use in a regular video. When the whole image moves it needs to be smooth (no hand held camera footage), and motion is best when kept in one direction. We learned this the hard way when we tried to film at the Green Bank Telescope with the fisheye camera. We didn't have a steady cam, but were carrying the camera around while filming. In the dome, the motion caused by the act of walking made the footage unusable; even causing slight nausea during viewing.

Motion can be used quite effectively if you are moving only small parts of the image. Because the rest of the image is still, small objects can move quite fast without causing disorientation. In this case, it is useful to take advantage of the shape of the dome; allowing objects to move in and out of the observer's field of view without leaving the image. In the show, we play with this by having moving Hydrogen atoms with orbiting electrons and photons that travel through the sky.

Figure A.14: Panorama image of the CHIME telescope made using footage shot with the rectangular lens. Color saturation in the image is cranked up to fit the planetarium environment.

A.5.3 Using pop-outs

Because creating three-dimensional animation can be quite time consuming and tricky, one common trick that planetarium shows use is pop-out windows. This involves projecting a background image on the dome, then opening a polygon in the bottom center of the image. A two-dimensional animation sequence is run inside the polygon while the background is held constant.

Figure A.15 shows an example of how we used this for the show, where the an-

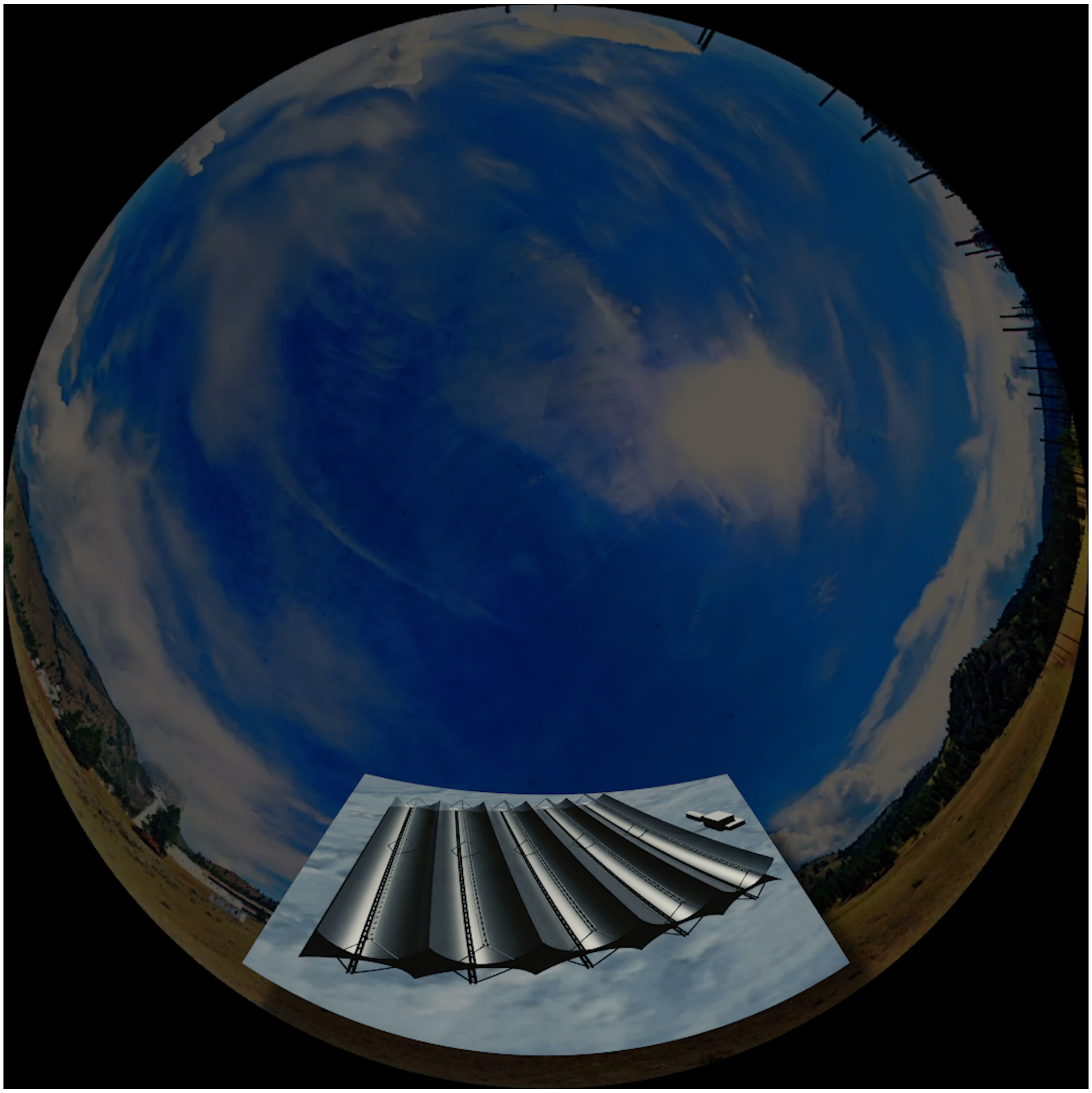


Figure A.15: Snapshot of CHIME cylinder animation inside a pop-out with an image of the CHIME telescope as the background.

imation here is a demonstration of how light interacts with the CHIME cylinders. The snapshot shows a moment when the animation is zoomed out to the full system. We used this technique for our animations that show how light is collected by the different telescopes, CHIME and the Green Bank Telescope.

A.6 Audio Production

Just as with any other video, sound is an important part of a planetarium show. The first part of the sound is the narration. Narrators must speak slowly and clearly and enunciate so that none of the words become garbled. Breath control is also important, in order to avoid audible breaths in the recording. It is also important to speak naturally, and convey enthusiasm for the material. As is indicated by the script, I acted as the show narrator to provide a personal element to the story.

Usage of sound booth can provide a way to control the environment and avoid environmental noise. It is also important to have a good quality microphone and control the sound levels to avoid clipping. For this show, we were able to use the sound booth at the CMU Entertainment Technology Center to record the narration.

Beyond the narration, the audio needs depth and complexity. The first step is to add a musical soundtrack behind the narration. This music should be set low enough not to drown out the narrator, but can become dominant during pauses in the narration. Narrative pauses are important, as they give the audience a chance to process what they've just heard. These pauses can also give the audience a chance to focus on the visuals.

Sound effects can be added to the soundtrack to match the action on the screen and further emphasize the visual story. Such auditory queues help the audience understand what they are seeing. For all of these components, it is particularly important that the sound and audio match. If the narration doesn't match what is happening on the screen, the audience can become lost or confused. With the level of technical detail present in our show, we had to be particularly careful with our synchronization as even a few seconds makes a big difference in comprehension.

A.7 Feedback and Evaluation

One of the most important parts of any show is getting feedback from various audiences. This is particularly important in a planetarium show, where the goal is to reach and educate the public. Feedback can take many forms and should occur throughout the development process.

A.7.1 Process Feedback

Concept Review

The first stage of feedback was getting approval for our concept. After we had created a rough storyboard, we took the story to Jeff Peterson and Peter Timbie, the principal investigators for the NSF grant funding the project. We identified our primary goals for the project: to educate the public about the Hydrogen sky as a new way to see the universe and how we are able to use radio telescopes to see the Hydrogen sky.

Planetarium Testing

Throughout the project, we needed to test our footage in the planetarium environment. This was important because it allowed us to see what things looked like in the planetarium compared to the flat screen and make adjustments. In addition, this gave us the opportunity to meet regularly with Tom Casey and Frank Mancuso, who acted as technical advisers for the project.

A.7.2 Script Development

In writing the script, one of the big challenges was describing technical and scientific details in language that would be accessible to the general public. The rule of thumb for planetarium shows is to write toward a middle school audience comprehension level. When writing the script, it was important to capture the ideas with analogies that would be accessible, and to match the visual story being told.

We went through several iterations of the script, including multiple recordings of the narration. One of the things that we found was that sometimes a line that made sense on paper didn't translate right when spoken. We also got feedback on the script from both scientists and educators, including Dr. Brendan Mullen at the Carnegie Science Center.

Because of the length and goals of the show, we ended up with a more technical show that is suitable for high-school and college age students. While this wasn't our original intent, the focus on the why and how behind 21-cm cosmology ended up being better suited for an older audience.

A.7.3 Intermediate Viewings

Once we had a show that was mostly complete, we went through a number of intermediate viewings. In these viewings, we invited some of our advisers to come and see the show and give recommendations of what we should change or keep. We tried to get advice from different perspectives, including technical details, creative and storytelling feedback, and scientific accuracy. These viewings helped us to revise the show and create a coherent story that worked together.

A.7.4 Public Viewings

A preview of the completed show can be viewed online (http://youtu.be/4E5luX1dCtA). At this time it is not yet available for public viewing, but we plan to have an open showing at the Carnegie Science Center in 2015. Following the viewing, we plan to ask viewers to give feedback via a survey. The exact form of this survey is still under development.

Getting public feedback will help us to assess the success of our show in meeting the stated goal of educating the public on the Hydrogen Sky. The show may also be used as a pilot for a full size show ($\sim 30-60$ minutes) as 21-cm science efforts continue to expand and larger 21-cm map datasets become available.

Bibliography

[1] G. Bernardi, L. J. Greenhill, D. A. Mitchell, S. M. Ord, B. J. Hazelton, B. M. Gaensler, A. de Oliveira-Costa, M. F. Morales, N. Udaya Shankar, R. Subrahmanyan, R. B. Wayth, E. Lenc, C. L. Williams, W. Arcus, B. S. Arora, D. G. Barnes, J. D. Bowman, F. H. Briggs, J. D. Bunton, R. J. Cappallo, B. E. Corey, A. Deshpande, L. deSouza, D. Emrich, R. Goeke, D. Herne, J. N. Hewitt, M. Johnston-Hollitt, D. Kaplan, J. C. Kasper, B. B. Kincaid, R. Koenig, E. Kratzenberg, C. J. Lonsdale, M. J. Lynch, S. R. McWhirter, E. Morgan, D. Oberoi, J. Pathikulangara, T. Prabu, R. A. Remillard, A. E. E. Rogers, A. Roshi, J. E. Salah, R. J. Sault, K. S. Srivani, J. Stevens, S. J. Tingay, M. Waterson, R. L. Webster, A. R. Whitney, A. Williams, and J. S. B. Wyithe. A 189 MHz, 2400 deg^2 Polarization Survey with the Murchison Widefield Array 32-element Prototype. *The Astrophysical Journal*, 771:105, July 2013.

[2] G. Bernardi, M. McQuinn, and L. J. Greenhill. Foreground Model and Antenna Calibration Errors in the Measurement of the Sky-Averaged \lambda 21 cm Signal at z˜20. *ArXiv e-prints*, Apr. 2014.

[3] J. D. Bowman, A. E. E. Rogers, and J. N. Hewitt. Toward Empirical Constraints on the Global Redshifted 21 cm Brightness Temperature During the Epoch of Reionization. *The Astrophysical Journal*, 676:1–9, Mar. 2008.

[4] Broadband Instrument for Global HydrOgen ReioNisation Signal (BIGHORNS). http://www.caastro.org/research/evolving/bighorns.

[5] V. Bromm. Formation of the first stars. *Reports on Progress in Physics*, 76(11):112901, Nov. 2013.

[6] J. O. Burns, J. Lazio, S. Bale, J. Bowman, R. Bradley, C. Carilli, S. Furlanetto, G. Harker, A. Loeb, and J. Pritchard. Probing the first stars and black holes in the early Universe with the Dark Ages Radio Explorer (DARE). *Advances in Space Research*, 49:433–450, Feb. 2012.

[7] Canadian Hydrogen Intensity Mapping Experiment(CHIME). http://chime.phas.ubc.ca/.

[8] B. Carroll and D. Ostlie. *An Introduction to Modern Astrophysics*. Pearson Addison-Wesley, 2007.

[9] Z. Ceplecha, J. Borovička, W. G. Elford, D. O. Revelle, R. L. Hawkes, V. Porubčan, and M. Šimek. Meteor Phenomena and Bodies. *Space Science Reviews*, 84:327–471, Sept. 1998.

[10] A. de Oliveira-Costa, M. Tegmark, B. M. Gaensler, J. Jonas, T. L. Landecker, and P. Reich. A model of diffuse Galactic radio emission from 10 MHz to 100 GHz. *Monthly Notices of the Royal Astronomical Society*, 388:247–260, July 2008.

[11] S. W. Ellingson. Antennas for the Next Generation of Low-Frequency Radio Telescopes. *IEEE Transactions on Antennas and Propagation*, 53:2480–2489, Aug. 2005.

[12] G. B. Field. Excitation of the Hydrogen 21-CM Line. *Proceedings of the IRE*, 46:240–250, Jan. 1958.

[13] S. R. Furlanetto, S. P. Oh, and F. H. Briggs. Cosmology at low frequencies: The 21 cm transition and the high-redshift Universe. *Physics Reports*, 433:181–301, Oct. 2006.

[14] Hydrogen Epoch of Reionization Array (HERA). http://reionization.org/.

[15] D. C. Jacobs, J. C. Pober, A. R. Parsons, J. E. Aguirre, Z. Ali, J. Bowman, R. F. Bradley, C. L. Carilli, D. R. DeBoer, M. R. Dexter, N. E. Gugliucci, P. Klima, A. Liu, D. H. E. MacMahon, J. R. Manley, D. F. Moore, I. I. Stefan, and W. P. Walbrugh. Multi-redshift limits on the 21cm power spectrum from PAPER. *ArXiv e-prints*, Aug. 2014.

[16] V. Jelić, A. G. de Bruyn, M. Mevius, F. B. Abdalla, K. M. B. Asad, G. Bernardi, M. A. Brentjens, S. Bus, E. Chapman, B. Ciardi, S. Daiboo, E. R. Fernandez, A. Ghosh, G. Harker, H. Jensen, S. Kazemi, L. V. E. Koopmans, P. Labropoulos, O. Martinez-Rubi, G. Mellema, A. R. Offringa, V. N. Pandey, A. H. Patil, R. M. Thomas, H. K. Vedantham, V. Veligatla, S. Yatawatta, S. Zaroubi, A. Alexov, J. Anderson, I. M. Avruch, R. Beck, M. E. Bell, M. J. Bentum, P. Best, A. Bonafede, J. Bregman, F. Breitling, J. Broderick, W. N. Brouw, M. Brüggen, H. R. Butcher, J. E. Conway, F. de Gasperin, E. de Geus, A. Deller, R.-J. Dettmar, S. Duscha, J. Eislöffel, D. Engels, H. Falcke, R. A. Fallows, R. Fender, C. Ferrari, W. Frieswijk, M. A. Garrett, J. Grießmeier, A. W. Gunst, J. P. Hamaker, T. E. Hassall, M. Haverkorn, G. Heald, J. W. T. Hessels, M. Hoeft, J. Hörandel, A. Horneffer, A. van der Horst, M. Iacobelli, E. Juette, A. Karastergiou, V. I. Kondratiev, M. Kramer, M. Kuniyoshi, G. Kuper, J. van Leeuwen, P. Maat, G. Mann, D. McKay-Bukowski, J. P. McKean, H. Munk, A. Nelles,

M. J. Norden, H. Paas, M. Pandey-Pommier, G. Pietka, R. Pizzo, A. G. Polatidis, W. Reich, H. Röttgering, A. Rowlinson, A. M. M. Scaife, D. Schwarz, M. Serylak, O. Smirnov, M. Steinmetz, A. Stewart, M. Tagger, Y. Tang, C. Tasse, S. ter Veen, S. Thoudam, C. Toribio, R. Vermeulen, C. Vocks, R. J. van Weeren, R. A. M. J. Wijers, S. J. Wijnholds, O. Wucknitz, and P. Zarka. Initial LOFAR observations of epoch of reionization windows. II. Diffuse polarized emission in the ELAIS-N1 field. *Astronomy & Astrophysics*, 568:A101, Aug. 2014.

[17] Large Aperture Experiment to detect the Dark Ages (LEDA). http://www.cfa.harvard.edu/leda/science.html.

[18] Low Frequency Array for Radio Astronomy(LOFAR). http://www.lofar.org/.

[19] K. W. Masui, E. R. Switzer, N. Banavar, K. Bandura, C. Blake, L.-M. Calin, T.-C. Chang, X. Chen, Y.-C. Li, Y.-W. Liao, A. Natarajan, U.-L. Pen, J. B. Peterson, J. R. Shaw, and T. C. Voytek. Measurement of 21 cm Brightness Fluctuations at z ˜ 0.8 in Cross-correlation. *The Astrophysical Journal Letters*, 763:L20, Jan. 2013.

[20] A. Mesinger, S. Furlanetto, and R. Cen. 21CMFAST: a fast, seminumerical simulation of the high-redshift 21-cm signal. *Monthly Notices of the Royal Astronomical Society*, 411:955–972, Feb. 2011.

[21] A. Natarajan and N. Yoshida. The Dark Ages of the Universe and hydrogen reionization. *Progress of Theoretical and Experimental Physics*, 2014(6):060000, June 2014.

[22] S. Nedeljkovic and U. of Toronto (Canada). *TREX: A Small Antenna RF Spectrometer*. Canadian theses. University of Toronto (Canada), 2007.

[23] G. Paciga, J. G. Albert, K. Bandura, T.-C. Chang, Y. Gupta, C. Hirata, J. Odegova, U.-L. Pen, J. B. Peterson, J. Roy, J. R. Shaw, K. Sigurdson, and T. Voytek. A simulation-calibrated limit on the H I power spectrum from the GMRT Epoch of Reionization experiment. *Monthly Notices of the Royal Astronomical Society*, 433:639–647, July 2013.

[24] J. C. Pober, A. R. Parsons, J. E. Aguirre, Z. Ali, R. F. Bradley, C. L. Carilli, D. DeBoer, M. Dexter, N. E. Gugliucci, D. C. Jacobs, P. J. Klima, D. MacMahon, J. Manley, D. F. Moore, I. I. Stefan, and W. P. Walbrugh. Opening the 21 cm Epoch of Reionization Window: Measurements of Foreground Isolation with PAPER. *The Astrophysical Journal Letters*, 768:L36, May 2013.

[25] A. E. E. Rogers and J. D. Bowman. Spectral Index of the Diffuse Radio Background Measured from 100 to 200 MHz. *Astronomical Journal*, 136:641–648, Aug. 2008.

[26] M. G. Santos, L. Ferramacho, M. B. Silva, A. Amblard, and A. Cooray. Fast large volume simulations of the 21-cm signal from the reionization and pre-reionization epochs. *Monthly Notices of the Royal Astronomical Society*, 406:2421–2432, Aug. 2010.

[27] R. W. Schunk and A. F. Nagy. *Ionospheres*. Cambridge University Press, 2000. Cambridge Books Online.

[28] J. R. Shaw, K. Sigurdson, M. Sitwell, A. Stebbins, and U.-L. Pen. Coaxing Cosmic 21cm Fluctuations from the Polarized Sky using m-mode Analysis. *ArXiv e-prints*, Jan. 2014.

[29] M. Sokolowski, S. E. Tremblay, R. B. Wayth, S. J. Tingay, N. Clarke, P. Roberts, M. Waterson, R. D. Ekers, P. Hall, M. Lewis, M. Mossammaparast, S. Padhi, F. Schlagenhaufer, A. Sutinjo, and J. Tickner. BIGHORNS - Broadband Instrument for Global HydrOgen ReioNisation Signal. *ArXiv e-prints*, Jan. 2015.

[30] Square Kilometer Array (SKA). http://www.skatelescope.org/.

[31] W. Stutzman and G. Thiele. *Antenna theory and design*. Wiley, 1981.

[32] E. R. Switzer, K. W. Masui, K. Bandura, L.-M. Calin, T.-C. Chang, X.-L. Chen, Y.-C. Li, Y.-W. Liao, A. Natarajan, U.-L. Pen, J. B. Peterson, J. R. Shaw, and T. C. Voytek. Determination of z ~ 0.8 neutral hydrogen fluctuations using the 21 cm intensity mapping autocorrelation. *Monthly Notices of the Royal Astronomical Society*, 434:L46–L50, July 2013.

[33] A. Thompson, J. Moran, and G. Swenson. *Interferometry and Synthesis in Radio Astronomy*. A Wiley-Interscience publication. Wiley, 2001.

[34] S. Tingay, R. Goeke, J. N. Hewitt, E. Morgan, R. A. Remillard, C. L. Williams, J. D. Bowman, E. Emrich, S. M. Ord, T. Booler, B. Crosse, D. Pallot, W. Arcus, T. Colegate, P. J. Hall, D. Herne, M. J. Lynch, F. Schlagenhaufer, S. Tremblay, R. B. Wayth, M. Waterson, D. A. Mitchell, R. J. Sault, R. L. Webster, J. S. B. Wyithe, M. F. Morales, B. J. Hazelton, A. Wicenec, A. Williams, D. Barnes, G. Bernardi, L. J. Greenhill, J. C. Kasper, F. Briggs, B. McKinley, J. D. Bunton, L. deSouza, R. Koenig, J. Pathikulangara, J. Stevens, R. J. Cappallo, B. E. Corey, B. B. Kincaid, E. Kratzenberg, C. J. Lonsdale, S. R. McWhirter, A. E. E. Rogers, J. E. Salah, A. R. Whitney, A. Deshpande, T. Prabu, A. Roshi, N. Udaya-Shankar, K. S. Srivani, R. Subrahmanyan, B. M. Gaensler, M. Johnston-Hollitt, D. L. Kaplan, and D. Oberoi. Realisation of a low frequency SKA Precursor: The Murchison Widefield Array. In *Resolving The Sky - Radio Interferometry: Past, Present and Future*, page 36, 2012.

[35] J. Townsend. *A Modern Approach to Quantum Mechanics*. University Science Books, 2000.

[36] H. K. Vedantham, L. V. E. Koopmans, A. G. de Bruyn, S. J. Wijnholds, B. Ciardi, and M. A. Brentjens. Chromatic effects in the 21 cm global signal from the cosmic dawn. *Monthly Notices of the Royal Astronomical Society*, 437:1056–1069, Jan. 2014.

[37] T. C. Voytek, A. Natarajan, J. M. Jáuregui García, J. B. Peterson, and O. López-Cruz. Probing the Dark Ages at z ˜ 20: The SCI-HI 21 cm All-sky Spectrum Experiment. *The Astrophysical Journal Letters*, 782:L9, Feb. 2014.

[38] S. A. Wouthuysen. On the excitation mechanism of the 21-cm (radio-frequency) interstellar hydrogen emission line. *Astronomical Journal*, 57:31–32, 1952.