21 cm Cosmology with Optimized Instrumentation and Algorithms

by

Haoxuan Zheng

B.S., University of Richmond (2011)

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2016

Author . Signature redacted ....
Department of Physics
April 29, 2016

Signature redacted
Certified by . ...
Max Tegmark
Professor of Physics
Thesis Supervisor

Accepted by . . . Signature redacted ....
Nergis Mavalvala
Professor of Physics
Associate Department Head for Education

21 cm Cosmology
with Optimized Instrumentation and Algorithms

by

Haoxuan Zheng

Submitted to the Department of Physics
on April 29, 2016, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

Precision cosmology has made tremendous progress in the past two decades thanks to a large amount of high quality data from the Cosmic Microwave Background (CMB), galaxy surveys and other cosmological probes. However, most of our universe's volume, corresponding to the period between the CMB and when the first stars formed, remains unexplored. Since there were no luminous objects during that period, it is called the cosmic "dark ages". 21 cm cosmology is the study of the high redshift universe using the hyperfine transition of neutral hydrogen, and it has the potential to probe that unchartered volume of our universe and the ensuing cosmic dawn, placing unprecedented constraints on our cosmic history as well as on fundamental physics.

My Ph.D. thesis work tackles the most pressing observational challenges we face in the field of 21 cm cosmology: precision calibration and foreground characterization. I lead the design, deployment and data analysis of the MIT Epoch of Reionization (MITEoR) radio telescope, an interferometric array of 64-dual polarization antennas whose goal was to test technology and algorithms for incorporation into the Hydrogen Epoch of Reionization Array (HERA). In four papers, I develop, test and improve many algorithms in low frequency radio interferometry that are optimized for 21 cm cosmology. These include a set of calibration algorithms forming redundant calibration pipeline which I created and demonstrated to be the most precise and robust calibration method currently available. By applying this redundant calibration to high quality data collected by the Precision Array for Probing the Epoch of Reionization (PAPER), we have produced the tightest upper bound of the redshifted 21 cm signals to date. I have also created new imaging algorithms specifically tailored to the latest generation of radio interferometers, allowing them to make Galactic foreground maps that are not accessible through traditional radio interferometry. Lastly, I have improved on the algorithm that synthesizes foreground maps into the Global Sky Model (GSM), and used it to create an improved model of diffuse sky emission from 10 MHz through 5 THz.

Thesis Supervisor: Max Tegmark
Title: Professor of Physics

Many are stubborn in pursuit of the path they have chosen, few in pursuit of the goal.

FRIEDRICH NIETZSCHE

Contents

List of Figures

List of Tables

Acknowledgments

I am truly grateful to my advisor, Prof. Max Tegmark, without whom I could not have enjoyed such an exciting and fulfilling five years as a graduate student. The many lessons I learned from him have not only been invaluable assets to my Ph.D. career, but will remain so to all my years ahead. I would also like to thank my thesis committee members, Prof. Jacqueline Hewitt and Prof. Paolo Zuccon, for their time and guidance that helped shape this thesis.

I would like to thank my lovely wife, Sherry, for her unending love and support. She has made many sacrifices for me so that I can pursue my passion and curiosity. Having her in my life is the best thing that has ever happened to me.

I would like to thank my mom and dad: nature or nurture, they are the ones who helped me become who I am today, and I'm truly grateful for having such wonderful parents.

Last but not least, I would like to thank my academic brothers, Adrian Liu and Josh Dillon, who have spent countless hours helping me with my projects and papers, and from whom much of this thesis inherit from. I also want to thank my friend Abraham Neben, with whom I have shared so many enjoyable conversations in and outside physics.

Chapter 1

Introduction

1.1 Precision Cosmology

Precision cosmology has made tremendous progress in the past two decades, thanks to a large amount of high quality data from the Cosmic Microwave Background (CMB) as well as the study of nearby stars and galaxies. As my advisor Max likes to say, when he was a graduate student 25 years ago, people were arguing over whether our universe is 10 billion years old or 20 billion years old, whereas nowadays, people debate over whether it is 13.7 or 13.8 billion years old. This drastic improvement in our knowledge of our universe's age is only the tip of the iceberg that is the triumph of precision cosmology. As another example, we now know that our universe consists of only about 5% ordinary matter that constitutes everything around us here on Earth, and the other 95% is split between about 26% dark matter and 69% dark energy [134]. The physical natures of dark matter and dark energy remain mysteries to us, but perhaps another two decades of precision cosmology could unveil their secrets. In the next two sections, I will briefly review two of the most important topics in precision cosmology: the Cosmic Microwave Background (CMB) and the 3D mapping of galaxies through the Sloan Digital Sky Survey.

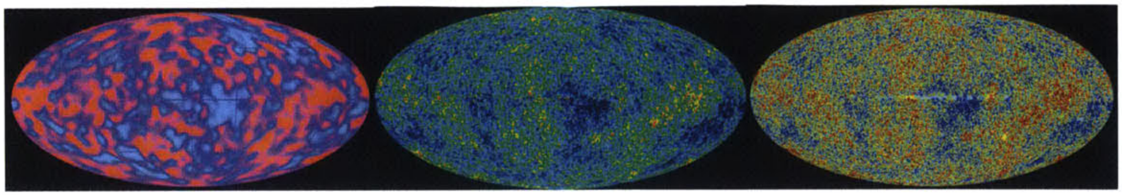

Figure 1-1: The evolution of CMB anisotropy measurements made by the last three generations of CMB experiments. *Image credits: the COBE, WMAP, and Planck Collaborations, respectively.*

1.1.1 The Far End: the Cosmic Microwave Background

The CMB is a blackbody radiation emitted by cooling plasma in the early universe when it cooled to about 3,000 K, less than half a million years after our Big Bang. As the CMB radiation traveled to us, it cooled down by a factor of about 1100, so it appears to us as 2.725 K blackbody radiation. Because the 2.725 K blackbody spectrum peaks in the microwave band, it is called the Cosmic Microwave Background. This radiation was discovered back in the 1960s. Most of the recent progress in cosmology has come from studying very slight variations in CMB temperature throughout the sky, typically no more than 0.1 mK. Fig. 1-1 clearly shows the improvement in quality, specifically angular resolution and sensitivity, of the last three generations of CMB experiments, which are all microwave detectors mounted on satellites. People study the CMB anisotropy by investigating the statistical property of the anisotropies, captured by the CMB power spectrum, as shown in Fig. 1-2. The strength of the power spectrum represents the amount of fluctuation concentrated at various angular scales, and the amplitudes and locations of the peaks in the power spectrum are determined by what happened both before the CMB was emitted, as well as when it was on its way to us. Thus, careful study of the power spectrum can and has revealed a great deal about the history of our universe.

1.1.2 In the Neighborhood: the Galaxy Surveys

Observing stars in our Galaxy and other distant galaxies have been the study of astronomy for centuries. Edwin Hubble's observation in 1929 that the majority of distant galaxies are moving away from us is perhaps one of the most impactful dis-

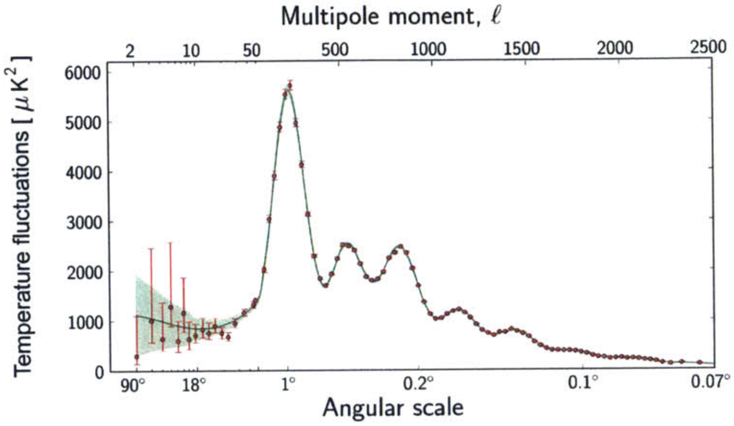

Figure 1-2: The CMB power spectrum measured by the Planck satellite. Our current cosmological model (green curve) can fit it very well with just six parameters. *Image credit: the Planck Collaborations.*

coveries in the history of cosmology. The constant development of more and more advanced telescopes and spectroscopes has allowed us to observe more galaxies that are farther away. One of the most ambitious efforts in systematically studying galaxies is the Sloan Digital Sky Survey [190] (SDSS). Using the 2.5 m telescope at Apache Point Observatory, the SDSS created the largest volume three-dimensional map of galaxies to date (see Fig. 1-3).

1.1.3 Is There Anything in between?

The 13.8 billion years old CMB is at the far end of our universe accessible to us through light. This is because before the time when CMB was emitted, the hot plasma that filled the universe did not allow light to travel freely. On the other hand, most of the galaxies surveyed by SDSS are less than 2 billion years old[1], which is very close to us compared to the CMB. Thus, one must be tempted to ask: is there anything we can observe in between the CMB and the oldest galaxies? If observations of the farthest slice (CMB) and the nearby volume (SDSS) have propelled the last

---

[1] SDSS has been carrying out larger and larger surveys, and other instruments have observed galaxies more than 10 billion years old. However, volume in cosmology is not proportional to the age cubed, so the gap between the CMB and the oldest galaxy is much larger than their age gaps suggest. See Fig. 1-4.

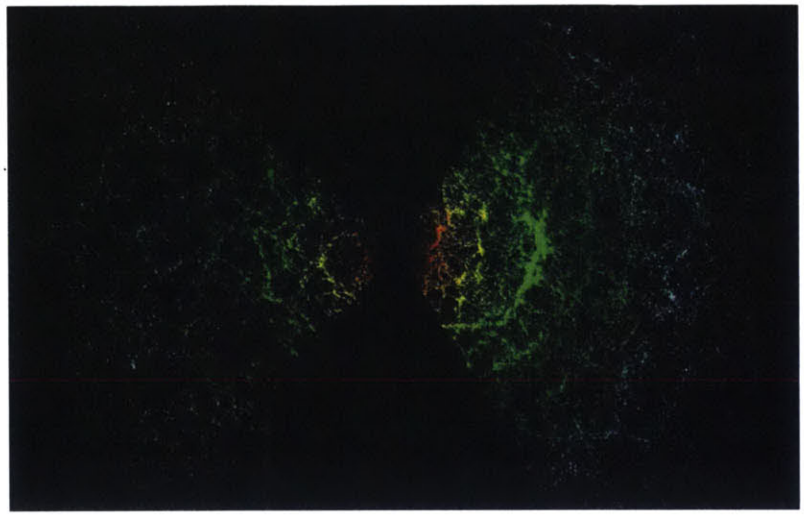

Figure 1-3: The 3D galaxy map made by the SDSS Legacy Survey. Each point in this "cosmic web" is a galaxy, and the farthest galaxies are about 2 billion years old. *Image credit: the Sloan Digital Sky Survey.*

two decades of cosmological discoveries, observing the volume in between certainly has the potential to revolutionize cosmology again.

As shown in Fig. 1-4, the missing gap actually makes up more than 80% of our universe's comoving volume. This gap period has been called the cosmic dark ages, because there was very little light being generated then, compared to the later times when our universe is populated by shiny stars and galaxies. Fortunately, the universe was not a complete void, but was filled by neutral hydrogen gas, which later gravitated to form the stars and galaxies. What is more, the neutral hydrogen gas did emit radiation that we can observe today. Unlike the CMB which obeys a continuous spectrum characteristic of the blackbody radiation, the neutral hydrogen atoms have a hyperfine transition that emit light at a very specific wavelength: 21 cm. As the 21 cm radio waves traveled towards us, their wavelength got stretched due to the universe's expansion, and the earlier they were emitted, the more they were stretched. This, combined with the fact that the neutral hydrogen is optically thin, allows us to determine the age of any emission from neutral hydrogen by simply looking at how much longer its wavelength is compared to 21 cm, so in principle we can make a 3D map of the entire gap shown in Fig. 1-4.

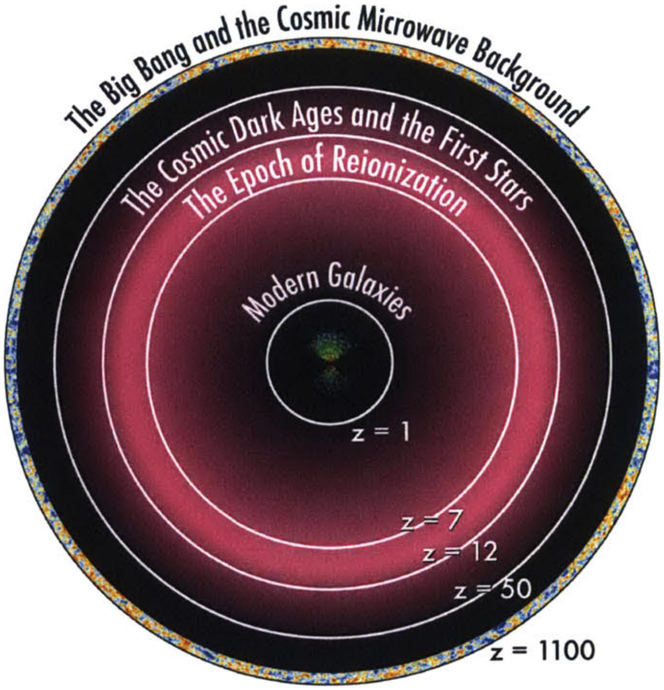

Figure 1-4: The CMB anisotropy maps and the galaxy surveys only probe a small fraction of the volume of the observable universe. 21 cm cosmology, the probe that this thesis focuses on developing, may one day make the entire pink region accessible to direct observation. *Image credits: Josh Dillon and Tegmark and Zaldarriaga [166].*

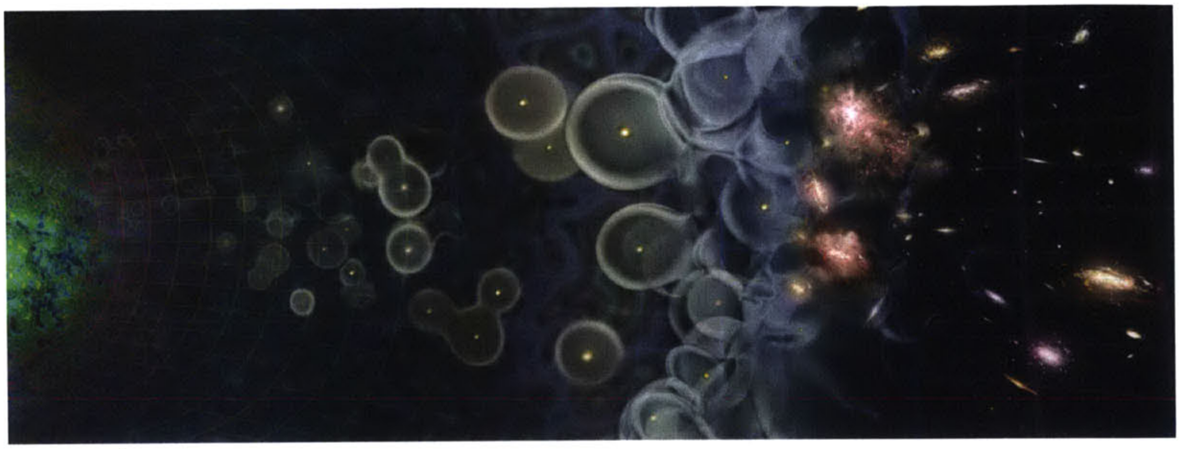

Figure 1-5: The Epoch of Reionization is a period in the early history of the universe between the cosmic microwave background (on the left) and modern stars and galaxies (on the right). During this time, the first stars and galaxies form and ionize the neutral hydrogen gas around them, which creates merging bubbles of ionized hydrogen. *Image credit: Abraham Loeb and Scientific American.*

Being able to map the neutral hydrogen throughout the history of our universe can shed much light on both cosmology and astrophysics. The neutral hydrogen gas underwent a lot of changes during the dark ages, especially towards the end when the first stars and galaxies just started forming (which we call the "cosmic dawn"). The young stars and galaxies emitted strong radiation that ionized the neutral hydrogen around them (in a process called reionization), and the hydrogen gas stopped emitting 21 cm radiation once ionized. Thus, if we can make a 3D hydrogen map, we can see that more and more empty "bubbles" appeared and merged during the cosmic dawn, as shown in Fig. 1-5. Currently we have a rough idea of how it might have happened, but we don't even know exactly when this "bubble period" was, how long it lasted, or any further details. Thus, direct measurements of this transition period will teach us about both the structure of our universe, as well as the formation of the first stars and galaxies.

1.2 Probing the Red-shifted Neutral Hydrogen Gas and the Potential of 21 cm Cosmology

In this section I briefly describe the more technical aspects of 21 cm cosmology. For a much more detailed review, see the introduction of Josh Dillon's thesis [35] and reviews in the literature [54, 144],

1.2.1 Temperature Evolution of the Neutral Hydrogen Gas

The aforementioned 21 cm radiation has been observed and used to trace neutral hydrogen in our Galaxy since its first detection in 1951 by Ewen and Purcell [44]. However, it is different to probe the neutral hydrogen before and during the epoch of reionization (EoR) with red-shifted 21 cm signals. For periods before and during the EoR, the hydrogen gas is observed in the form of either emission or absorption relative to the CMB. What we directly measure is I_ν, the specific intensity of emission at the frequency ν. Since the red-shifted 21 cm signals have frequencies on the order of 10^8 Hz, much lower than the peak of the CMB at around 10^{11} Hz, we can use the Rayleigh-Jeans limit of the blackbody spectrum to represent observed specific intensities as brightness temperatures T_b, where

$$I_\nu \equiv \frac{2k_B T_b \nu^2}{c^2}. \tag{1.1}$$

The observed brightness temperature is a combination of the CMB temperature T_γ and the spin temperature of the hydrogen gas T_S, which is defined in terms of the Boltzmann factor for the spin-singlet and spin-triplet hyperfine levels of the ground state of hydrogen,

$$\frac{n_{\text{triplet}}}{n_{\text{singlet}}} = 3e^{-h\nu_0/k_B T_S}. \tag{1.2}$$

Following Furlanetto et al. [54], the equation of radiative transfer through a cloud of hydrogen backlit by the CMB is

$$T_b(z,\nu) = T_S(z)\left(1 - e^{-\tau_\nu}\right) + T_\gamma(z) e^{-\tau_\nu}, \tag{1.3}$$

where τ_ν is the optical depth of the cloud due to the 21 cm transition.

Since the neutral hydrogen gas is considered optically thin ($\tau_\nu \ll 1$) [54], contrast in the 21 cm signal observed today relative to the CMB is then given by

$$\begin{aligned}\delta T_b^{\text{obs}}(z) &= \frac{T_b(z)}{1+z} - T_\gamma(z=0) \\ &= \frac{(T_S(z) - T_\gamma(z))\left(1 - e^{-\tau_{\nu_0}}\right)}{1+z} \\ &\approx \frac{T_S(z) - T_\gamma(z)}{1+z}\tau_{\nu_0},\end{aligned} \tag{1.4}$$

where τ_{ν_0} is the integrated optical depth over frequency. Finally, we skip some technical calculations shown in Furlanetto et al. [54] and Pritchard and Loeb [144], and arrive at the final equation

$$\delta T_b^{\text{obs}}(z) \approx (27\,\text{mK})x_{\text{HI}}(z)\left(\frac{T_S(z) - T_\gamma(z)}{T_S(z)}\right)(1+\delta_b(z))\sqrt{\frac{1+z}{10}}\left[\frac{(1+z)H(z)}{\partial v_\parallel(z)/\partial r_\parallel}\right], \tag{1.5}$$

where x_{HI} is the neutral fraction of the hydrogen gas (with 1 meaning fully neutral), δ_b is the baryon over-density, $H(z)$ is the Hubble parameter, and $\partial v_\parallel/\partial r_\parallel$ is the gradient of the proper velocity along the line of sight.

Equation 1.5 shows that the observed brightness temperature depends on the relative amplitude of the spin temperature and the CMB temperature. As discussed in Furlanetto et al. [54], Pritchard and Loeb [144], the spin temperature is driven by three major processes: interaction with CMB photons, collisions between neutral hydrogen atoms and other particles, and absorption and remission of Lyman-alpha photons known as the *Wouthuysen-Field effect* [187, 48]. In equilibrium, the spin temperature is given by

$$T_S^{-1} = \frac{T_\gamma^{-1} + x_c T_K^{-1} + x_\alpha T_\alpha^{-1}}{1 + x_c + x_\alpha}, \tag{1.6}$$

where x_c and x_α are the collisional and Lyman-alpha coupling coefficients, T_K is the kinetic temperature related to particle collisions, and T_α is the Lyman-alpha temperature related to the Wouthuysen-Field effect. Since many of these quanti-

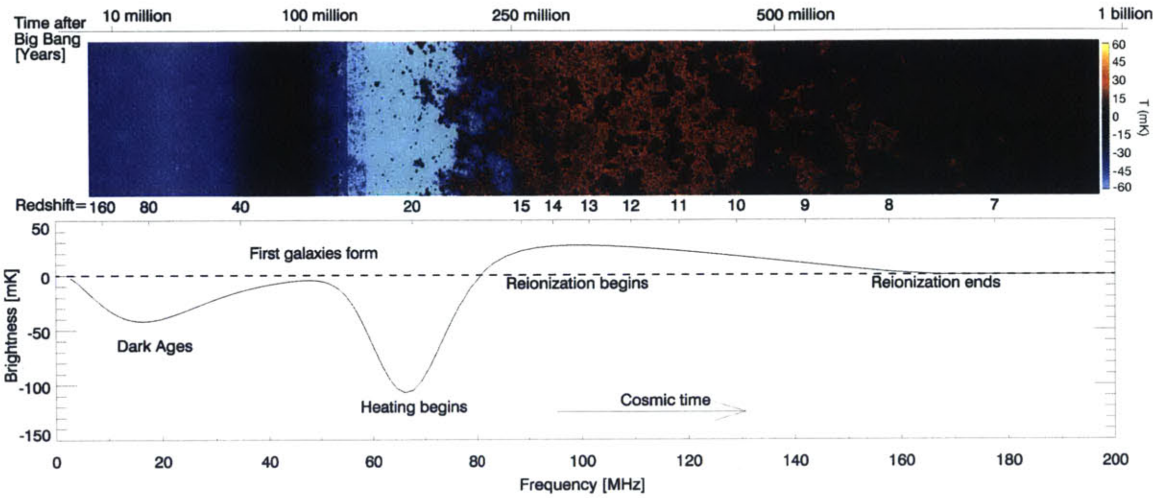

Figure 1-6: Top: simulated evolution of the brightness temperature, driven by the formation of the first stars and galaxies. Bottom: the sky-averaged global 21 cm signal calculated from the simulated result in the top panel. *Reproduced from Pritchard and Loeb [144].*

ties are inhomogeneous and change across cosmological time, so does the brightness temperature we observe. The top panel of Fig. 1-6 shows simulated evolution of the brightness temperature. However, the precise physical processes that drive this evolution are poorly understood, and precise measurements on the evolution of brightness temperature during EoR will shed light on these topics [54, 144].

1.2.2 The 21 cm Power Spectrum

While making high quality 3D maps of the brightness temperature will be possible in the future, current efforts are focusing on reduced data products due to low signal-to-noise ratio in today's instruments. The simplest data product that probes the evolution of the brightness temperature is the global 21 cm signal, which is the brightness temperature averaged over the whole sky, as shown in the bottom panel of Fig. 1-6. Another useful reduced data product is the power spectrum of the brightness temperature. This power spectrum $P(\mathbf{k})$ reflects the amount of correlation on various length scales $\mathbf{k}$ in the 21 cm brightness temperature, and it is defined by

$$\langle \widetilde{\delta T}_b^*(\mathbf{k}) \widetilde{\delta T}_b(\mathbf{k}') \rangle \equiv (2\pi)^3 \delta(\mathbf{k}-\mathbf{k}') P(\mathbf{k}), \tag{1.7}$$

where $\widetilde{\delta T}_b(\mathbf{k})$ is the spatial Fourier transform of $\delta T_b(\mathbf{r})$. If we approximate the spatial distribution of $\delta T_b(\mathbf{r})$ as isotropic, we turn $P(\mathbf{k})$ into $P(k)$, and obtain the dimensionless power spectrum

$$\Delta_{21}^2(k) \equiv \frac{k^3}{2\pi^2} P(k). \tag{1.8}$$

As reionization progresses, both the number and the size of ionized bubbles increase, so we expect to see more and more power at lower k, and less and less power overall. Fig. 1-7 shows simulated evolution of the power spectrum from 10% ionized to 98% ionized. Measuring the evolution of the 21 cm power spectrum is the focus of many current instruments, including the ones presented in this thesis.

1.2.3 The Potential of 21 cm Cosmology

Precise measurements of the red-shifted 21 cm signals will not only shed light on the astrophysical processes driving reionization, but also teach us a lot more about our universe. Pober et al. [139] has shown that a suitably designed instrument with a tenth of a square kilometer of collecting area will allow tight constraints on the timing and duration of reionization and the astrophysical processes that drove it. As shown by Mao et al. [93], a future radio array with a square kilometer of collecting area, maximal sky coverage, and good foreground maps could improve the sensitivity to cosmological parameters, such as spatial curvature and neutrino masses, by up to two orders of magnitude. It also has the potential to shed new light on the early universe by measuring the running of the spectral index related to the theory of inflation. In the shorter term, Liu et al. [88] has shown that with the instruments currently under construction, it is possible to constrain some of the cosmological parameters much better than the CMB experiments can.

1.3 Low-frequency Radio Interferometry

The first generation of instruments designed to measure the 21 cm power spectrum are all radio interferometers [75, 172, 125, 188, 115, 60], and they have inherited

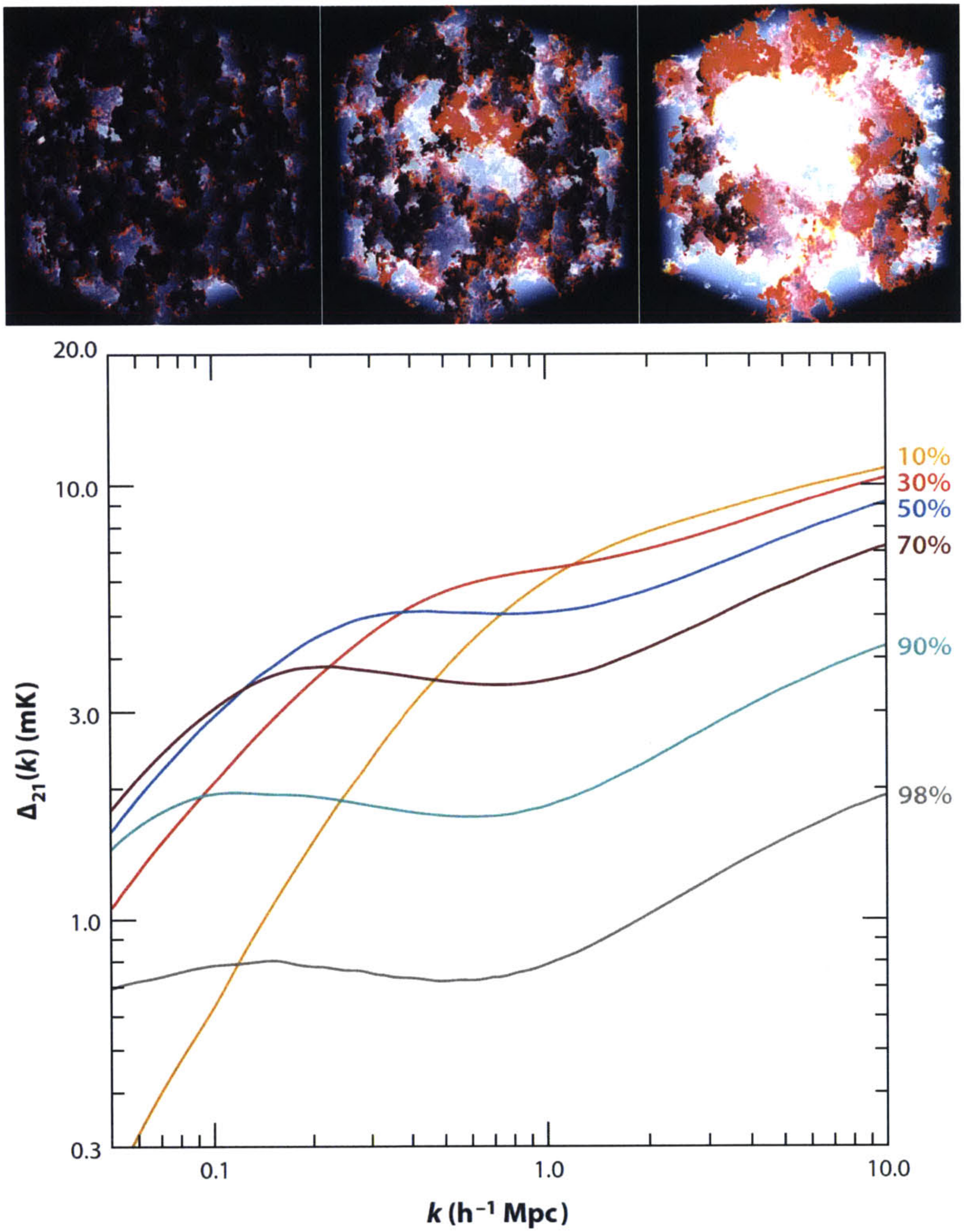

Figure 1-7: Simulated evolution of a hydrogen cube undergoing reionization (top), and simulated evolution of the power spectrum (bottom). As the number and the size of ionized bubbles increase, there is more and more power at lower k, and less and less power overall. *Image credit: Marcelo Alvarez, Ralf Kaehler, and Tom Abel (top); reproduced from Morales and Wyithe [108] (bottom).*

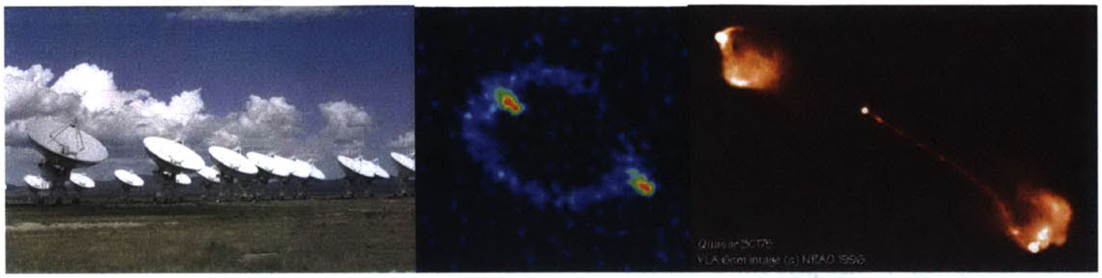

Figure 1-8: The Very Large Array (VLA, on the left) is a radio interferometer that consists of 27 radio antennas located at the NRAO site in Socorro, New Mexico. Each antenna in the array measures 25 meters in diameter and weighs about 230 tons. The first direct observation of an Einstein ring (middle figure) was made using the VLA by a team led by Prof. Jacqueline Hewitt, whom I have the honor to have on my thesis committee. This 1986 observation of the quasar 3C175 (on the right) has a 0.35 arcsecond resolution. *Image credit: National Radio Astronomy Observatory and the Very Large Array.*

many ideas from traditional radio interferometry developed in the past century. One example of an exquisite radio interferometer and its amazing achievements are shown in Fig. 1-8. Before we discuss more about new challenges in 21 cm instrumentation and how I optimized algorithms and instrumentations for detecting 21 cm signals, I would like to give a brief introduction to fundamental radio interferometry in this section, which will help us understand the need for new algorithms in 21 cm interferometry.

1.3.1 Basic Instrumentation

Very roughly speaking, the signal chain in radio interferometers can be divided into four stages: detection, amplification, digitization, and cross-correlation. Radio frequency emissions are usually detected using antennas, which themselves are typically similar in size to the wavelength they are designed to detect, ranging from millimeters to meters. The large dishes, such as those in Fig. 1-8, are reflectors designed to increase sensitivity by reflecting much more radiation onto the antennas, which are located at the focal points of the dishes. The signals picked up by the antennas are first amplified by low noise amplifiers, before they are digitized by analog-to-digital converters (ADCs). At this stage, the output of an ADC connected to the ith antenna can be described using incident plane wave, summed over all directions in the sky:

$$E_i(t) = \int_\Omega s(\hat{\mathbf{k}}, t) b(\hat{\mathbf{k}}) e^{i(\mathbf{k}\cdot\mathbf{r}_i - \omega t)} d\Omega, \tag{1.9}$$

where $\mathbf{k}$ is the wave vector of radiation coming from sky direction $\hat{\mathbf{k}}$, $s(\hat{\mathbf{k}}, t)$ is the strength of signal coming from that direction at the moment t in time, $b(\hat{\mathbf{k}})$ is the antenna's sensitivity in that direction decided by the shape of the antenna and its reflector, and $\mathbf{r}_i$ is the antenna's position on the ground. For simplicity I have omitted the noise term here.

Since the radio frequency is in the MHz-GHz range, it is not convenient to record millions or billions of samples of $E_i(t)$ for every antenna onto a hard drive directly. However, since the signals oscillate rapidly, any averaging over time would severely reduce their amplitude. To solve this problem, people cross-correlate signals to form so-called *visibilities* from pairs of antennas:

$$v_{ij}(t) = \langle E_i^*(t) E_j(t) \rangle = \int_\Omega S(\hat{\mathbf{k}}, t) B(\hat{\mathbf{k}}) e^{i\mathbf{k}(\mathbf{r}_j - \mathbf{r}_i)} d\Omega, \tag{1.10}$$

where * means complex conjugate, $S = |s|^2$, and $B = |b|^2$. The $e^{i\mathbf{k}(\mathbf{r}_j - \mathbf{r}_i)}$ term is the interference pattern between the two antennas; thus the name interferometer. In this form, the fast varying ωt-terms are canceled, and the only remaining t-dependence is in $S(\hat{\mathbf{k}}, t)$, which varies on the time scale of Earth's rotation. Thus, we can safely integrate the visibilities over tens of milliseconds. Even longer integration is possible, depending on the angular resolution of the instrument. For high resolution instruments like the VLA, integration over timescales of minutes can be done through techniques such as delay tracking, which compensates for the movement of the source in the sky.

1.3.2 Calibration

As Niels Bohr said: in theory, theory and practice are the same, but in practice, they are not. In practice, the measured visibilities are rather different from what is written in Eq. (1.10). There are many effects at play, including but not limited to sky noise, amplifier noise, cable delay, reflections between antennas, cross-talk between signal channels, radio frequency interference (RFI) from airplanes or satellites, and so on. Just as in many other experiments in physics, noise can be averaged down by repeated measurements over many nights, assuming that all the other systematic effects are accounted for. Among these effects, the strongest are modeled as a set of complex gain parameters, $g_i(t)$'s, where each antenna has a different one, and they vary over time. Conceptually, the amplitude of $g_i(t)$ corresponds to fluctuating amplifier gains, whereas the phase of $g_i(t)$ is a consequence of fluctuating cable delays. These gain parameters change our measured visibilities from Eq. (1.10) to

$$v_{ij}(t) = g_i^*(t) g_j(t) \int_\Omega S(\hat{\mathbf{k}}, t) B(\hat{\mathbf{k}}) e^{i\mathbf{k}(\mathbf{r}_j - \mathbf{r}_i)} d\Omega. \tag{1.11}$$

Calibrating the instrument is then finding the complex gains for all of the antennas. A lot of work has been put into this topic and many sophisticated algorithms have been developed [157, 185, 145, 23, 114]. The simplest form of calibration is to point the instrument at a known bright source that has been carefully measured, and then use the knowledge of S, B, and v_{ij}'s to solve for the g_i's. Assuming those g_i's do not change very much over time, the instrument is then pointed to the object of interest. A more sophisticated form is called self-cal, where one first uses the uncalibrated data to form an image of the sky, and using the approximation that there are only point-like objects in the sky, one can correct the image and use that as the model. One then iterates this operation until both the image and the gain solutions g_i's converge. This brings us to our next topic: how do we form an image using the measured visibilities?

1.3.3 Imaging

Radio interferometry typically takes advantage of Eq. (1.10) by first performing a coordinate transformation from $\boldsymbol{k}$ on the celestial sphere to its projection on the

xy-plane, the horizontal plane in the observer's local coordinate system:

$$v_{ij} = \int \frac{S(\boldsymbol{q})B(\boldsymbol{q})}{\sqrt{1-|\boldsymbol{q}|^2}} e^{i2\pi \boldsymbol{q}\cdot\boldsymbol{u}_{ij}} d\boldsymbol{q}, \tag{1.12}$$

where $\boldsymbol{q} = \frac{\lambda}{2\pi}(k_x, k_y)$, $\boldsymbol{u}_{ij} = \frac{r_j - r_i}{\lambda}$, and λ is the wavelength of the radiation. We see that in this form, the visibilities v_{ij} (or written in a more explicit form, $v(\boldsymbol{u}_{ij})$) and the sky-beam image, $\frac{S(\boldsymbol{q})B(\boldsymbol{q})}{\sqrt{1-|\boldsymbol{q}|^2}}$, form a Fourier pair. Performing 2D Fourier transforms on measured v_{ij}'s is the core step in making images. The Fourier approach, however, comes with one important limitation. Generally speaking, without knowing the specific form of B or what is in the sky, the sky-beam image is band-limited to the unit circle $|\boldsymbol{q}| \leq 1$, so by Nyquist theorem, one has to have the shortest baseline shorter than half a wavelength to avoid aliasing in the image (see Section 2.A for more details and illustrations). In reality, it is difficult to have any baselines shorter than half a wavelength due to the physical size of the antenna dishes. What is more, the size of $B(\boldsymbol{q})$, which determines the band-limitedness of the sky-beam image, is roughly the inverse of the antenna size[2]. Since the shortest baseline has to be longer than the diameter of the antenna, the largest angular scale available is always smaller than the primary beam width (or the band-limitedness of the sky-beam image), making aliasing inevitable.

Fortunately for instruments like the VLA, aliasing is not much of a problem, and that is due to the nature of the objects these instruments are trying to observe. As seen in Fig. 1-8, the objects of interest are typically compact objects that are much brighter than the background. For the rightmost image in Fig. 1-8 of 3C175, the field of view is about 1 arcminute, and the primary beam width is around 10 arcminutes. First of all, there are no objects comparable in brightness to 3C175 within 10 arcminutes. Secondly, before Fourier transforming them to obtain the image, we can apply an anti-aliasing filter [163] to the visibilities to artificially suppress any signal outside the field of view, since we are only interested in studying the structure of 3C175. This way, many stunning high resolution images have been captured using

---

[2] If there's a reflector, then it is the reflector size.

radio interferometers.

1.3.4 Challenges in 21 cm Cosmology Instruments

Unfortunately, the cosmological 21 cm signal is so faint that none of the current experiments around the world (LOFAR [75], MWA [172], PAPER [125], 21CMA [188], GMRT [115]) have detected it yet, although increasingly stringent upper limits have recently been placed [116, 38, 127, 4]. The major challenge comes from our own Galaxy. In the frequency range where the redshifted 21 cm signals are thought to be the strongest, our Galaxy is emitting synchrotron radiation that is thought be four orders of magnitude larger than the cosmological hydrogen signal [34, 4].

To put this in perspective, recall our discussion on how aliasing does not affect traditional radio interferometric imaging. As mentioned, in the case of 3C175 shown in Fig. 1-8, the background that makes up the majority of space in the beam is much lower than the bright structures of 3C175, and both imaging and calibration rely on that fact. For 21 cm cosmology, the neutral hydrogen signals we want to map are ten thousand times weaker than the "negligible background" in the 3C175 image, and these signals are of comparable strength across most of the sky. In this case, the traditional imaging algorithm cannot work very well due to effects such as aliasing, and because of that, calibration will not work very well either, as no maps or collections of point sources can be used as the model in the first place.

In addition to calibration and foreground removal, correlator cost will become a bottleneck when experiments scale up in the future. Since steerable single-dish radio telescopes become prohibitively expensive beyond a certain size, the aforementioned experiments have all opted for interferometry, combining N (generally a large number) independent antenna elements which are (except for GMRT) individually more affordable. The problem with scaling interferometers to high N is that all of these experiments use standard hardware cross-correlators whose cost grows quadratically with N, since they need to correlate all $N(N-1)/2 \sim N^2/2$ pairs of antenna elements. This cost is reasonable for the current scale $N \sim 10^2$, but will completely dominate the cost for $N \gtrsim 10^3$, making traditionally designed precision cosmology arrays with $N \sim 10^6$ as discussed in Mao et al. [93] infeasible in the near future.

1.4 Thesis Outline

The work that constitutes this thesis was originally written as four different papers. The papers appear here as Chapters 2 through 5 and are reproduced verbatim with the permission of their primary co-authors. I played a significant role in the development and writing of all these papers and served as the first author on three of them—in this thesis, Chapters 2, 4, and 5. Two of them have already been published in peer-reviewed journals; Chapter 4 and 5 will be submitted in May 2016.

This thesis is organized into three thematic parts. In Part I, *Building MITEoR and OMNICAL*, I describe my work leading the MIT effort of building a new radio interferometer: the MIT Epoch of Reionization (MITEoR) experiment. I demonstrate many new instrument design ideas and algorithms through this experiment, and one of the most notable is the redundant calibration algorithms. I show that the redundant calibration algorithms are able to perform calibration without the need for a sky model, and can achieve optimal precision. MITEoR is also a precursor to what we call "omniscopes", a type of radio interferometer that will not be limited by the quadratic correlator cost in the future.

In Part II, *Latest Epoch of Reionization Science Results*, I apply the redundant calibration algorithms to the latest data collected by the Precision Array for Probing the Epoch of Reionization (PAPER). The redundant calibration algorithms dramatically improve the quality of the PAPER data, which were used by the PAPER team to place the most stringent constraints to date on 21 cm power spectra.

Finally, in Part III, *Novel Imaging and The New Global Sky Model*, I present two new algorithms optimized for 21 cm cosmology. The first is a new imaging method that does not use the Fourier approach. Rather, the new method takes on Eq. (1.10) as a linear system of equations, and uses precise mathematical tools developed by the CMB community to make high precision, large field of view images of the radio sky. The second algorithm builds on what is called the Global Sky Model (GSM), which combines sky maps from many different frequencies to model the radio sky. The new GSM algorithm improves both the precision and the flexibility of the original, allowing us to combine many more datasets in a more accurate fashion.

Part I

Building MITEoR and OMNICAL

Chapter 2

MITEoR: A Scalable Interferometer for Precision 21 cm Cosmology

The content of this chapter was submitted to the Monthly Notices of the Royal Astronomical Society *on June 12, 2014 and published [194] as* MITEoR: a scalable interferometer for precision 21 cm cosmology *on October 8, 2014. The authors are: Haoxuan Zheng, M. Tegmark, V. Buza, J. S. Dillon, H. Gharibyan, J. Hickish, E. Kunz, A. Liu, J. Losh, A. Lutomirski, S. Morrison, S. Narayanan, A. Perko, D. Rosner, N. Sanchez, K. Schutz, S. M. Tribiano, M. Valdez, H. Yang, K. Zarb Adami, I. Zelko, K. Zheng, R. P. Armstrong, R. F. Bradley, M. R. Dexter, A. Ewall-Wice, A. Magro, M. Matejek, E. Morgan, A. R. Neben, Q. Pan, R. F. Penna, C. M. Peterson, M. Su, J. Villasenor, C. L. Williams, and Y. Zhu.*

2.1 Introduction

Mapping neutral hydrogen throughout our universe via its redshifted 21 cm line offers a unique opportunity to probe the cosmic "dark ages," the formation of the first luminous objects, and the epoch of reionization (EoR). A suitably designed instrument with a tenth of a square kilometer of collecting area will allow tight constraints on the timing and duration of reionization and the astrophysical processes that drove it [139]. Moreover, because it can map a much larger comoving volume of our universe, it has the potential to overtake the Cosmic Microwave Background (CMB) as our most sensitive cosmological probe of inflation, dark matter, dark energy, and neutrino masses. For example [93], a radio array with a square kilometer of collecting area, maximal sky coverage, and good foreground maps could improve the sensitivity to spatial curvature and neutrino masses by up to two orders of magnitude, to $\Delta\Omega_k \approx 0.0002$ and $\Delta m_\nu \approx 0.007$ eV, and shed new light on the early universe by a 4σ detection of the spectral index running predicted by the simplest inflation models favored by the BICEP2 experiment [2].

Unfortunately, the cosmological 21 cm signal is so faint that none of the current experiments around the world (LOFAR [75], MWA [172], PAPER [125], 21CMA [188], GMRT [115]) have detected it yet, although increasingly stringent upper limits have recently been placed [116, 38, 127]. A second challenge is that foreground contamination from our galaxy and extragalactic sources is perhaps four orders of magnitude larger than the cosmological hydrogen signal [34]. Any attempt to accurately clean it out from the data requires even greater sensitivity as well as more accurate calibration and beam modeling than the current state-of-the-art in radio astronomy (see Furlanetto et al. [54], Morales and Wyithe [108] for reviews).

Large sensitivity requires large collecting area. Since steerable single dish radio telescopes become prohibitively expensive beyond a certain size, the aforementioned experiments have all opted for interferometry, combining N (generally a large number) independent antenna elements which are (except for GMRT) individually more affordable. The LOFAR, MWA, PAPER, 21CMA and GMRT experiments currently have comparable N. The problem with scaling interferometers to high N is that all of these experiments use standard hardware cross-correlators whose cost grows quadratically with N, since they need to correlate all $N(N-1)/2 \sim N^2/2$ pairs of antenna elements. This cost is reasonable for the current scale $N \sim 10^2$, but will completely dominate the cost for $N \gtrsim 10^3$, making precision cosmology arrays with $N \sim 10^6$ as discussed in Mao et al. [93] infeasible in the near future, which has motivated novel correlator approaches such as Morales [107].

For the particular application of 21 cm cosmology, however, designs with better cost scaling are possible, as described in Tegmark and Zaldarriaga [166, 167]: by arranging the antennas in a hierarchical rectangular or hexagonal grid and performing the correlations using Fast Fourier Transforms (FFTs), thereby cutting the cost scaling to $N \log N$. This is particularly attractive for science applications requiring exquisite sensitivity at vastly different angular scales, such as 21 cm cosmology (where short baselines are needed to probe the cosmological signal[1] and long baselines are needed for point source removal). Such hierarchical grids thus combine the angular resolution advantage of traditional array layouts with the cost advantage of a rectangular Fast Fourier Transform Telescope. If the antennas have a broad spectral response as well and their signals are digitized with high bandwidth, the cosmological neutral hydrogen gets simultaneously imaged in a vast 3D volume covering both much of the sky and also a vast range of distances (corresponding to different redshifts, *i.e.*, different observed frequencies.) Such low-cost arrays have been called *omniscopes* [166, 167] for their wide field of view and broad spectral range.

Of course, producing such scientifically rich maps with any interferometer depends crucially on our ability to precisely calibrate the instrument, so that we can truly understand how our measurements relate to the sky. Traditional radio telescopes rely on a well-sampled Fourier plane to perform self-calibration using the positions and fluxes of a number of bright point sources. At first blush, one might think that any highly-redundant array would be at a disadvantage in its attempt to calibrate the gains and phases of individual antennas. However, we can use the fact that redundant baselines should measure the same Fourier component of the sky to improve the

---

[1] It has been shown that the 21 cm signal-to-noise ratio (S/N) per resolution element in the *uv*-plane (Fourier plane) is $\ll 1$ for all current 21 cm cosmology experiments, and that their cosmological sensitivity therefore improves by moving their antennas closer together to focus on the center of the *uv*-plane and bringing its S/N closer to unity [106, 20, 98, 93, 82]. Error bars on the cosmological power spectrum have contributions from both noise and sample variance, and it is well-known that the total error bars on a given physical scale (for a fixed experimental cost) are minimized when both contributions are comparable, which happens when the $S/N \sim 1$ on that scale. This is why more compact 21 cm experiments have been advocated. This is also why early suborbital CMB experiments focused on small patches of sky to get $S/N \sim 1$ per pixel, and why galaxy redshift surveys target objects like luminous red galaxies that give $S/N \sim 1$ per 3D voxel.

calibration of the array dramatically and quantifiably. In fact, we find that the ease and precision of redundant baseline calibration is a strong rationale for building a highly-redundant array, in addition to the improvements in sensitivity and correlator speed.

Redundant calibration is useful both for current generation redundant arrays like MITEoR and PAPER and for future large arrays that will need redundancy to cut down correlator cost. Omniscopes must be calibrated in real time, because they do not compute and store the visibilities measured by each pair of antennas, but effectively gain their speed advantage by averaging redundant baselines in real time. Individual antennas therefore cannot be calibrated in post-processing. No calibration scheme used on existing low frequency radio interferometers has been demonstrated to meet the speed and precision requirements of omniscopes. Thus, the main goal of the MIT Epoch of Reionization experiment (MITEoR) and this paper is to demonstrate a successful redundant calibration pipeline that can overcome the calibration challenges faced by current and future generation instruments by performing automatic precision calibration in real time.

Building on past redundant baseline calibration methods by Wieringa [185] and others, some of us recently developed an algorithm which is both automatic and statistically unbiased, able to produce precision phase and gain calibration for all antennas in a hierarchical grid (up to a handful of degeneracies) without making any assumptions about the sky signal [87]. Once obtained, precision calibration solutions can in turn produce more accurate modeling of the synthesized and primary beams[2] [137], which has been shown to improve the quality of the foreground modeling and removal which is so crucial to 21 cm cosmology. It is therefore timely to develop a pathfinder instrument that tests how well the latest calibration ideas works in practice.

MITEoR is such a pathfinder instrument, designed to test redundant baseline calibration. We developed and successfully applied a real-time redundant calibration

---

[2] For tile-based interferometers like the MWA and 21CMA, gain and phase errors in individual antennas (as opposed to tiles) do not typically get calibrated in the field, adding a fundamental uncertainty to the tile sky response.

pipeline to data we took with our 64 dual-polarization antenna array during the summer of 2013 in The Forks, Maine. The goal of this paper is to describe the design of the MITEoR instrument, demonstrate the effectiveness of our redundant baseline calibration and absolute calibration pipelines, and use the calibration results to obtain an optimal scheme for estimating calibration parameters as a function of time and frequency.

This paper is organized as follows. We first describe in Section 2.2 the instrument, including the custom developed analog components, the 8 bit 128 antenna-polarization correlator, the deployment, and the observation history. In Section 2.3, we focus on precision calibration. We explain and quantitatively evaluate relative redundant calibration, and address the question of how often calibration coefficients should be updated. We also examine the absolute calibration, including breaking the degeneracies in relative calibration, mapping the primary beam, and measuring the array orientation. In Section 2.4, we summarize this work and discuss implications for future redundant arrays such as HERA [139].

2.2 The MITEoR Experiment

In theory, a very large omniscope can be built following the generalized architecture in Figure 2-1. On the other hand, it is crucial to demonstrate that automatic and precise calibration is possible in real-time using redundant baselines, since the calibration coefficients for each antenna must be updated frequently to allow the FFTs to combine the signals from the different antennas without introducing errors. In this section, we will present our partial implementation of this general design, including both the analog and the digital systems. Because the digital hardware is powerful enough to allow it, the MITEoR prototype correlates all 128 input channels with one another, rather than just a small sample as mentioned in the caption of Figure 2-1. This provides additional cross-checks that greatly aid technological development, where instrumentation may be particularly prone to systematics. This also allows us to explore the question of exactly how often and how finely in frequency we must measure

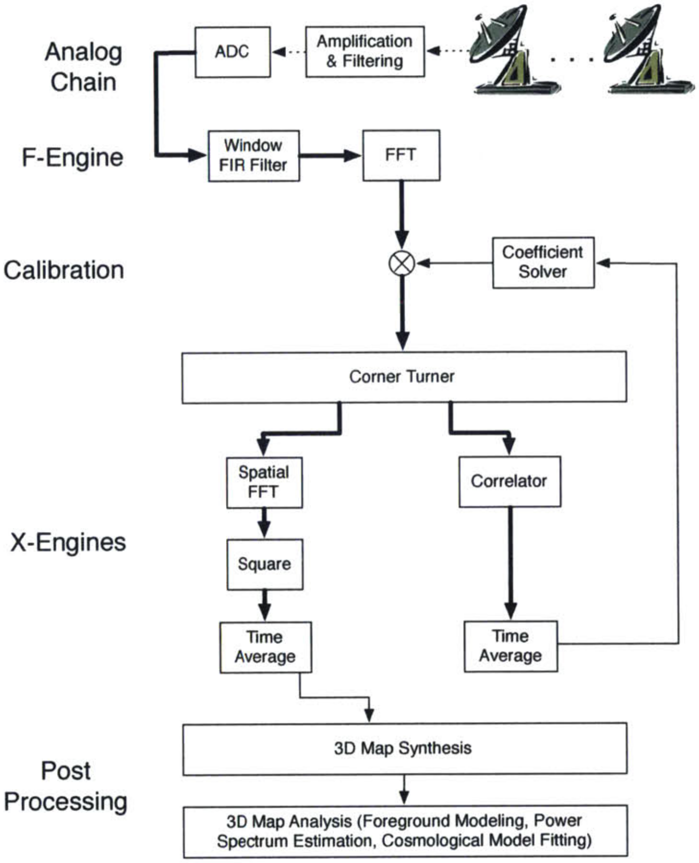

Figure 2-1: Data pipeline for a large omniscope that implements FFT correlator and redundant baseline calibration. First, a hierarchical grid of dual-polarization antennas converts the sky signal into volts, which get amplified and filtered by the analog chain, transported to a central location, and digitized every few nanoseconds. These high-volume digital signals (thick lines) get processed by field-programmable gate arrays (FPGAs) which perform a temporal Fourier transform. The FPGAs (or GPUs) then multiply by complex-valued calibration coefficients that depend on antenna, polarization and frequency, then spatially Fourier transform, square and accumulate the results, recording integrated sky snapshots every few seconds and thus reducing the data rate by a factor $\sim 10^9$. They also cross-correlate a small fraction of all antenna pairs, allowing the redundant baseline calibration software [87, 112] to update the calibration coefficients in real time and automatically monitor the quality of calibration solutions for instrumental malfunctions. Finally, software running on regular computers combine all snapshots of sufficient quality into a 3D sky ball or "data cube" representing the sky brightness as a function of angle and frequency in Stokes (I,Q,U,V) [167], and subsequent software accounts for foregrounds and measures power spectra and other cosmological observables.

visibilities to solve for calibration coefficients, a question we return to in Section 2.3. Since we chose to implement a full correlator, an additional FFT correlator would bring no extra information (simply computing the same redundant-baseline-averaged visibilities faster), so we leave the digital implementation of an FFT correlator to future work. In general, our mission is to empirically explore any challenges that are unique to a massively redundant interferometer array. Once these are known, one can reconfigure the cross-correlation hardware to perform spatial FFTs, thereby obtaining an omniscope with $N \log N$ correlator scaling.

2.2.1 The Analog System

MITEoR contains 64 dual-polarization antennas, giving 128 signal channels in total. The signal picked up by the antennas is first amplified by two orders of magnitude in power by the low noise amplifiers (LNAs) built-in to the antennas. It is then phase switched in the swapper system, which greatly reduces cross-talk downstream. The signal is then amplified again by about five orders of magnitude in the line-drivers before being sent over 50 meter RG6 cables to the receivers. The receivers perform IQ demodulation on a desired 50 MHz band selected between 100 MHz and 200 MHz, producing two channels with adjacent 25 MHz bands, and sends the resulting signals into the digitization boards containing 256 analog-to-digital converters (ADCs) sampling at 50 MHz. The swappers, line-drivers and receivers we designed are shown in Figure 2-2.

When designing the components of this system, we chose to use commercially-available integrated circuits and filters whenever possible, to allow us to focus on system design and construction. In some cases (such as with the amplifiers) the cost of the IC is less than the cost of enough discrete transistors to implement even a rough approximation of the same functionality. Less expensive filters could be made from discrete components, but the characteristics of purchased modules tend to be better due to custom inductors and shielding. When we needed to produce our own boards as described below, our approach was to design, populate and test them in our

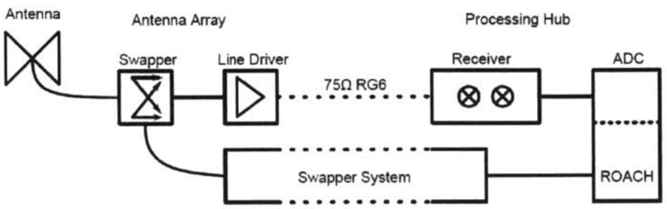

Figure 2-2: System diagram of the analog system. The signal received with an MWA "bow-tie" antenna is first amplified by the built-in low noise amplifier, then Walsh-modulated in the swapper module controlled by the swapper system. The signal is amplified again in the line driver and sent to the processing rack through 50 m long coaxial cables. In the processing rack, the signal first goes into the receiver, where it undergoes further amplification, frequency down-mixing and I/Q modulation from the 120-180 MHz range to the 0-25 MHz range. The analog chain ends with digitization on ADC connected to ROACH boards.

laboratory, then have them affordably mass-produced for us by Burns Industries[3].

2.2.1.1 Antennas

The dual-polarization antennas used in MITEoR were originally developed for the Murchison Widefield Array [91, 172], and consist of two "bow-tie"-shaped arms as can be seen in Figure 2-8. They are inexpensive, easy to assemble, and sensitive to the entire band of our interest. The MWA antennas were designed for the frequency range 80-300 MHz, and have a built-in low noise amplifier with 20 dB of gain. The noise figure of the amplifier is 0.2 dB, and the 20 dB of gain means that subsequent gain stages do not contribute significantly to the noise figure[4].

2.2.1.2 Swappers (Phase Switches)

As with many other interferometers, crosstalk within the receivers, ADCs, and cabling significantly affects signal quality. We observe the cross-talk to depend strongly on the physical proximity of channel pairs, reaching as high as about −30 dB between nearest neighbor receiver channels. Our swapper system is designed to cancel out crosstalk during the correlator's time averaging by selectively inverting analog signals

---

[3] `http://www.burnsindustriesinc.com`

[4] In a multi-stage amplifier, the contribution of each stage's noise figure is suppressed by a factor that is equal to the total gain of previous stages.

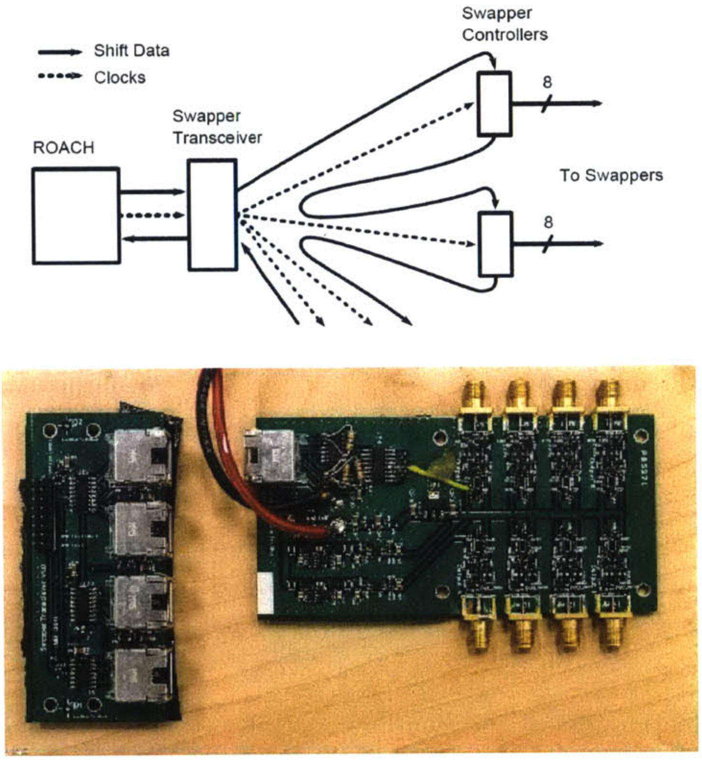

Figure 2-3: System diagram of our swapper signal system and physical components of the swapper transceiver (lower left) and swapper controller (lower right). The swapper is designed to reduce crosstalk between neighboring channels.

using Walsh modulation [154]. The signal from each antenna-polarization is inverted 50% of the time according to its own Walsh function, by an analog ZMAS-1 phase switch from Mini-Circuits located before the second amplification stage (line-driver), then appropriately re-inverted after digitization[5]. We perform the inversion once every millisecond, which is much longer than the ADC's 20 ns sample time, and much shorter than the averaging time of a few seconds[6]. This eliminates all crosstalk to first order [154]. If crosstalk reduction were the only concern, the ideal position for the swapper would be immediately after the antenna, in order to cancel as much crosstalk as possible. In practice, the swapper introduces a loss of about 3 dB, so we perform the modulation after the LNA to avoid adding noise (raising the system temperature).

---

[5] Since the undesirable crosstalk signal is demodulated with a different Walsh function than it is modulated with, it will be averaged out due to orthogonality of Walsh functions.

[6] The inversion cannot be too frequent, because we need to discard data during the analog inversion process which takes a few microseconds. At the same time, the inversion needs to be frequent enough to average out the cross-talk.

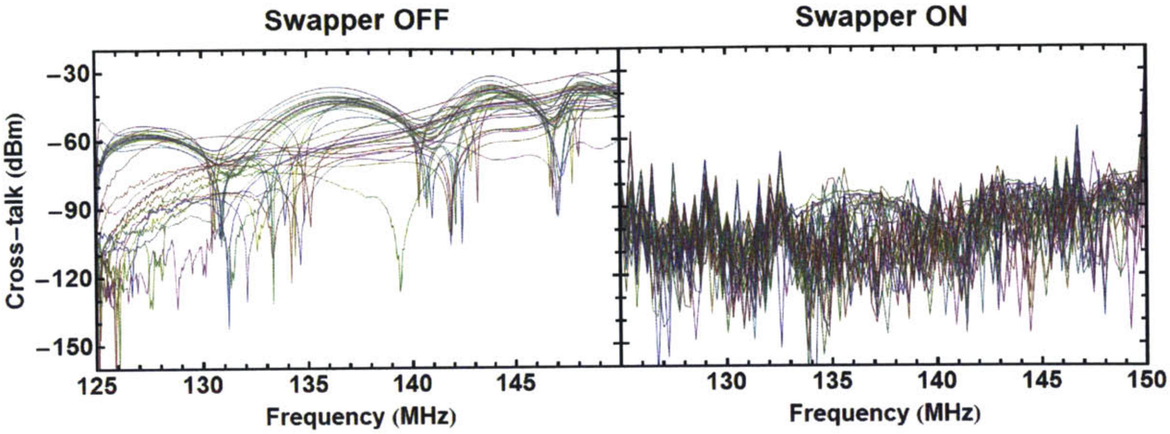

Figure 2-4: Plots of cross-talk power measured in the laboratory. The swapper suppresses crosstalk between channels by as much as 50 dB. To measure these curves we fed a 0 dBm sinusoidal signal into input channel 0 of the receivers and left the other 31 input channels open. We then measured the correlations between channel 0 and all 31 empty channels, due to crosstalk from channel 0. We repeated the procedure with input frequencies from 125-150 MHz and obtained the results shown above.

To evaluate the effectiveness of the swapper modules, we sent a monotone signal into one single channel of the receivers while leaving other channels open, and measured the correlation between the signal channel and each empty channels with the swapper turned on and off. We then repeated this while varying the signal frequency over the full range of interest. As seen in Figure 2-4, the swapper system attenuates crosstalk in the receiver and ADC by as much as 50 dB over the frequency band of interest, typically reducing it to being of order −80 dB for strongly afflicted signal pairs.

2.2.1.3 Line-Driver

A line-driver (Figure 2-5) amplifies a single antenna's signal from one of its two polarization channels while also powering its LNA. Line-drivers only handle a single channel to reduce potential crosstalk from sharing a printed circuit board. They are placed within a few meters of the antennas in order to reduce resistive losses from powering the antenna at low voltage. Additional gain that they provide early in the analog chain helps the signal overpower any noise picked up along the way to the processing hub, and maintains the low noise figure set up by the LNA. To further reduce potential radio-frequency interference (RFI), we chose to power the line-drivers

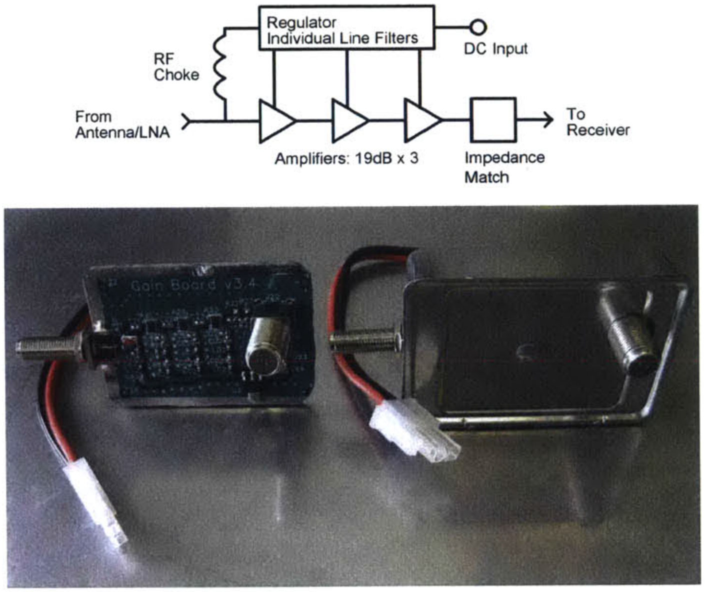

Figure 2-5: System diagram and physical components of the line drivers. The line driver we designed takes the signal in the 50Ω coaxial cable from the antenna LNA and amplifies it by 51 dB, in order to overpower noise picked up in the subsequent 75Ω coaxial cable and further processing steps up to 50 meters away. It operates on 5V DC and also provides DC bias power to the antenna's LNA through the 50 Ohm cable.

with 58Ah 6V sealed lead acid rechargeable batteries during the final 64-antenna deployment, rather than 120 VAC to 6 VDC adapters (whose unwanted RF-emission may have caused occasional saturation problems during our earlier expeditions).

2.2.1.4 Receiver

Our receivers (Figure 2-6) take input from the line-drivers, bandpass filter the incoming signals, amplify their power level by 23 dB, and IQ-demodulate them. The resulting signals go directly to an ADC for digitization. Receivers are placed near the ADCs to which they are connected to reduce cabling for local oscillator (LO) distribution and ADC connections. IQ demodulation is used, which doubles received bandwidth for a given ADC frequency at the cost of using two ADC channels, and has the advantage of requiring only a single LO and low speed ADCs. The result is

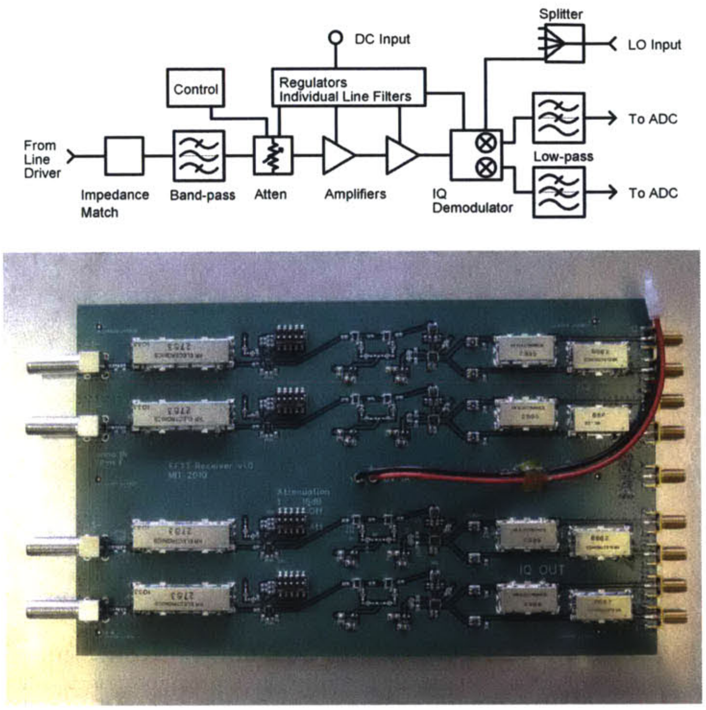

Figure 2-6: System diagram and physical component of the receiver boards. The boards take the signals arriving from four line drivers, band-pass filter and amplify them, then use a local oscillator to frequency shift them from the band of interest to a DC-centered signal suitable for input to the ADC.

40 MHz of usable bandwidth[7] anywhere in the range 110-190 MHz, with a 2-3 MHz gap centered around the LO frequency due to bandpass filters. The receiver boards have five pins allowing their signals to be attenuated by any amount between 0 dB and 31 dB (in steps of 1 dB) before the second amplification stage, to avoid saturation and non-linearities from RFI and to attain signal levels optimal for digitization.

2.2.2 The Digital System

We designed MITEoR's digital system (Figure 2-7) to be highly compact and portable. The entire system occupies 2 shock-mounted equipment racks on wheels, each measuring about 1 m on all sides. It takes in data from 256 ADC channels (64 antennas with I and Q signals for polarizations), Fourier transforms each channel, reconstructs IQ demodulated channels back to 128 corresponding antenna channels, computes the

---

[7] Due to limitations in our FPGAs' computing power, only 12.5 MHz of digitized data are correlated and stored at any instant.

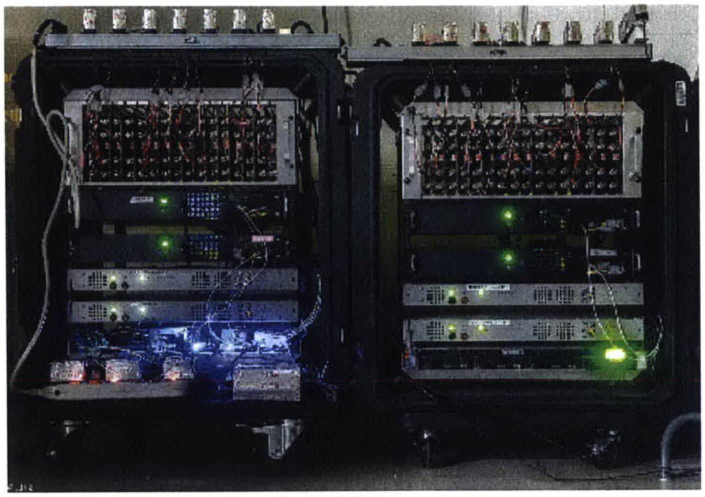

Figure 2-7: The entirety of our 128 antenna-polarization digital correlator system, packaged in two portable shock mounted racks. The two black chassis and two silver chassis in the middle of each rack are F-engines (ROACH) and X-engines (ROACH2), respectively. Above the ROACHes are 32 receiver boards that input the signals from 128 line drivers via F-cables. The blue lit area below the ROACHes contains various clocking devices responsible for synchronization, whereas the chassis below the ROACHes on the right hand side is the 8TB data acquisition server.

cross-correlations of all pairs of the 128 antenna channels with 8 bit precision, and then time-averages these cross-correlations. Although standard 4 bit correlators suffice for most astronomical observation tasks, the better dynamic range of our 8 bit correlator allows us to observe faint astronomical signals at the same time as 10^3 times brighter ORBCOMM satellites, whose enormous signal-to-noise has proved invaluable in characterizing various aspects of the system (see Sections 2.3.2.2, 2.3.2.3, and 2.3.3). The digital hardware is capable of processing an instantaneous bandwidth of 12.5 MHz with 49 kHz frequency bins. It averages those correlations and then writes them to disk every few seconds (usually either 2.6 or 5.3 seconds).

While one of the advantages of a massively redundant interferometer array is the ability to reduce costs by performing a spatial FFT rather than a full cross-correlation, we have not implemented FFT correlation in the current MITEoR prototype as the hardware is powerful enough to correlate all antenna pairs in real time (the feasibility of implementing FFT correlation on the ROACH platform has been demonstrated by Foster et al. [53]). Rather, the goal of MITEoR is to quantify the accuracy that automatic redundant baseline calibration can attain, thereby experimentally characterizing all of the unknowns in the system, such as unexpected analog chain systematics and other barriers to finding good calibration solutions.

We adopted the widely-used F-X scheme in MITEoR's digital system. We have 4 synchronized F-engines that take in data from 4 synchronized 64-channel ADC boards, which run at 12 bits and 50Ms/s. The F-engines perform the FFT and IQ reconstruction, and distribute the data onto 4 X-engines through 16 10GbE links. The 4 asynchronous X-engines each perform full correlation on 4 different frequency bands on all 128 antenna polarizations, and send the time averaged results to a computer for data storage.

To implement the computational steps of the MITEoR design, we used Field Programmable Gate Arrays (FPGAs). These devices can be programmed to function as dedicated pieces of computational hardware. Each F-engine and X-engine is implemented by one Xilinx FPGA (Virtex-5 for F-engines and Virtex-6 for X-engines). These FPGAs are seated on custom hardware boards developed by the CASPER collaboration[8] [121]. We also use the software tool flow developed by CASPER to design the digital system. The CASPER collaboration is dedicated to building open-source programmable hardware specifically for applications in astronomy. We currently use two of their newer devices, the ROACH[9] (Reconfigurable Open Architecture Computing Hardware) for the F-engines, and the ROACH 2[10] for the X-engines. The main benefit of using CASPER hardware is that it facilitates the time-consuming process of designing and building custom radio interferometry hardware. The CASPER collaboration also offers a large open-source library of FPGA blocks for commonly used signal processing structures such as polyphase filter banks, FIR filters and fast Fourier transform blocks [120]. However, due to MITEoR's ambitious architecture, involving both extreme compactness, an 8-bit correlator, and tight inter-ROACH synchronization constraints, we custom-designed most of the digital FPGA blocks. The specifications of our latest correlator are listed in Table 2.1.

---

[8] `https://casper.berkeley.edu/`

[9] `https://casper.berkeley.edu/wiki/ROACH`

[10] `https://casper.berkeley.edu/wiki/ROACH-2_Revision_2`

Antenna	MWA dual-pol bow-tie
Antenna count	64 × 2 polarizations
Array configuration	8 × 8 grid
ADC	4 × 64-channel 50 Msps
F-engine	4 ROACHes with Virtex-5
X-engine	4 ROACH2s with Virtex-6
Correlator precision	8 bits
Frequency range	110-190 MHz
Instantaneous bandwidth	12.5 MHz (50 MHz digitized)
Frequency resolution	49 kHz
Time resolution	≥ 2.68 s

Table 2.1: List of MITEoR specifications. We observed with two different 8 by 8 array configurations, one with 3 m separation and one with 1.5 m separation. We observed ORBCOMM band with 2.68 s resolution, and we chose a resolution of 5.37 s for other bands.

2.2.3 MITEoR Deployment and Data Collection

We deployed MITEoR in The Forks, Maine, which our online research suggested might be one of the most radio quiet region in the United States at the frequencies of interest[11][18]. We deployed the first prototype in September 2010, and performed a successful suite of test observations with an 8-antenna interferometer. In May 2012, we completed and deployed a major upgrade of the digital system to fully correlate $N = 16$ dual-polarization antennas. With the experience of this successful deployment, we further upgraded the digital system to accommodate $N = 64$ dual-polarization antennas, which led to our latest deployment in July 2013 and the results we describe in this paper.

The MITEoR experiment was designed to be portable and easy to assemble. The entire experiment was loaded into a 17 foot U-Haul truck and driven to The Forks. It took a crew of 15 people less than 2 days to assemble the instrument and bring it to full capacity. A skeleton crew of 3 members stayed on site for monitoring and maintenance for the following two weeks, during which we collected more than 300

---

[11] The Forks has also been successfully used to test the EDGES experiment [21], and we found the RFI spectrum to be significantly cleaner than at the National Radio Astronomy Observatory in Green Bank, West Virginia at the very low (100-200 MHz) frequency range that is our focus: the entire spectrum at The Forks is below -100 dBm except for one -89.5 dBm spike at 150MHz.

Figure 2-8: Part of the MITEoR array during the most recent deployment in the summer of 2013. 64 dual-polarization antennas were laid on a 21 m by 21 m regular grid with 3 m separation. The digital system was housed in the back of a shielded U-Haul truck (not shown).

hours of data. Subsequently, a demolition crew of 5 members disassembled and packed up MITEoR in 6 hours and concluded the successful deployment.

During the deployment, we scanned through the frequency range 123.5-179.5 MHz, with at least 24 consecutive hours at each frequency. We used two different array layouts for most of the frequencies we covered. The observation began with the antennas arranged in a regular 8 by 8 grid, with 3 meter spacing[12] between neighboring antennas, which we later reconfigured to an 8 by 8 regular grid with 1.5 meter spacing for a more compact layout (which provides better signal-to-noise ratio on the 21 cm signal). The total volume of binary data collected was 3.9TB, and in the rest of this paper, we demonstrate the results of our various calibration techniques using this data set.

2.3 Calibration Results

As we have emphasized above, the precision calibration of an interferometer is essential to its ability to detect the faint cosmological imprint upon the 21 cm signal, and the key focus of MITEoR is to determine how well real-time redundant calibration can be made to work in practice. In this section we describe the calibration scheme

---

[12] We aligned the antenna positions using a laser-ranging total station, and measured their positions with millimeter level precision. The median deviation from a perfect grid is 2 mm in the N-S direction, 3 mm E-W, and 28 mm vertical, primarily caused by the fact that the deployment site had not been leveled.

that we have designed and implemented and quantify its performance. We first constrain the *relative* calibration between antennas, utilizing both per-baseline algorithms and redundant-baseline calibration algorithms [87]. We then build on these relative calibration results to constrain the ***absolute*** calibration of the instrument, including breaking the few degeneracies inherent to redundant calibration.

2.3.1 Relative Calibration

2.3.1.1 Overview

The goal of relative calibration is to calibrate out differences among antenna elements caused by non-identical analog components, such as variations in amplifier gains and cable lengths, which may be functions of time and frequency. We parametrize the calibration solution as a time- and frequency-dependent multiplicative complex gain g_i for each of the 128 antenna-polarizations. Calibrating the interferometer amounts to solving for the coefficients g_i and undoing their effects on the data. Our calibration scheme revolves around calibration methods that heavily utilize the redundancy of our array, whose efficacy we aim to demonstrate with MITEoR. The current redundant calibration pipeline involves three steps, as illustrated in Figure 2-9:

1. Rough calibration computes approximate calibration phases using knowledge of the sky.

2. Logarithmic calibration ("logcal") decomposes roughly calibrated data into amplitudes and phases and computes least square fits for amplitude and phase separately.

3. Linear calibration ("lincal") takes the relatively precise but biased results from logcal and computes unbiased calibration parameters with even higher precision.

Although logcal and lincal have been previously proposed [185, 87, 94], they both fail in their original form if the phases of g_i are not close to 0.[13] In practice, the phases

---

[13] Logcal requires phase calibrations close to 0 to avoid phase wrapping issues, whereas lincal requires phase calibrations close to 0 to converge.

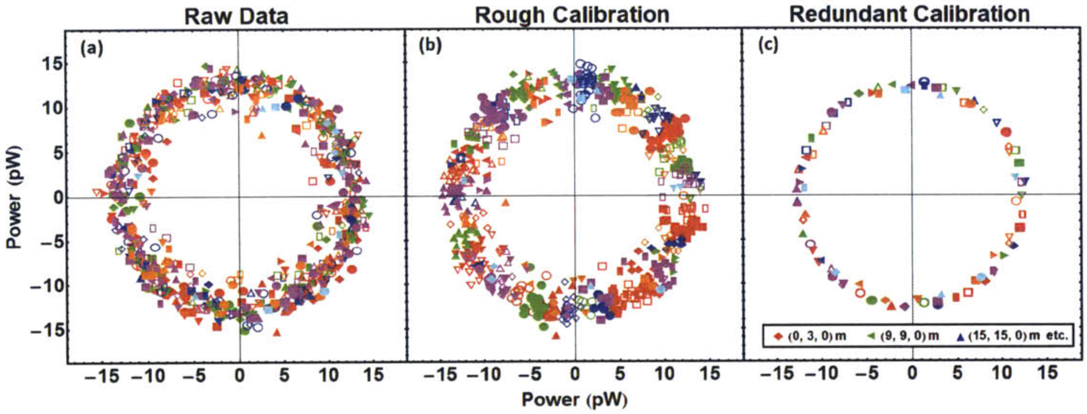

Figure 2-9: Illustration of three stages in the redundant baseline calibration pipeline. Each panel is a complex plane, and each point is a complex visibility for a specific pair of antennas at 137.1 MHz during the passage of an ORBCOMM satellite. Each unique combination of color and shape stands for one set of redundant baselines. In an ideal world, all identical symbols, such as all upright red triangles, should have the same value thus overlap exactly. Due to noise, they should cluster together around the same complex value. In panel (a) showing raw data, the redundant baselines have almost no clustering visible—for example, red filled circles can be found throughout the plot. After crude calibration in panel (b), we see most points falling into clustered segments—though the clustering is still far from exact. Finally in panel (c), after performing log calibration, we see that all points corresponding to each redundant baseline are almost exactly overlapping, with no visible deviation due to the high signal-to-noise. While the difference is not visible here, linear calibration can further improve log calibration results, as shown in Figure 2-11.

of g_i can be anywhere in the interval $[0, 2\pi)$. To overcome these practical challenges, we introduced various improvements to these algorithms. In the following sections, we describe our improvements to calibration algorithms in detail, and demonstrate the effectiveness of our calibration by obtaining $\chi^2/\mathrm{DoF} \approx 1$ for the majority of our data. We then analyze these calibration parameters to construct a Wiener filter to optimally average them over time and frequency, which also tells us how frequently we need to calibrate in time and frequency.

2.3.1.2 Rough Calibration

The goal of rough calibration is to obtain reliable initial phase estimates for the calibration parameters to enable the subsequent more sophisticated algorithms. This step does not have to involve redundancy, thus it can be done with any standard calibration techniques, for example self-calibration [145, 23]. The rough calibration algorithm that we describe below is computationally cheap and can robustly improve upon raw data even when a few antennas have failed.

At a given time and frequency, we have both the measured visibilities, v_{ij}, and v_{ij}^{model}, a rough model of the true sky signal[14], where indices i and j represent antenna number. We first compute the phase difference between each measured visibility and its prediction. We then pick one reference antenna and subtract the phases of its visibilities from the phases of other visibilities to obtain a list of estimated phase calibration for each antenna. Finally, we take the median of these calibration phases to obtain a robust phase calibration estimator for each antenna. More concretely, we use the following procedure:

1. Construct a matrix $\mathbf{M}$ of phase differences where $M_{ij} = -M_{ji} = \arg(v_{ij}/v_{ij}^{\mathrm{model}})$.
2. Define the first antenna as the reference by subtracting the first column of $\mathbf{M}$ from all columns to obtain $M'_{jk} = M_{jk} - M_{j0}$.
3. Obtain rough phase calibration parameters $\phi_k \equiv \arg(g_k)$ by computing the

---

[14] Since we are trying to obtain an initial estimate, the model does not have to be very accurate.

median angle of column k in M', defined as

$$\begin{aligned}\phi_k &\equiv \arg\left[\text{median}_j\{\exp(iM'_{jk})\}\right] \\ &= \arg\left[\text{median}_j\{\cos(M'_{jk})\}\right. \\ &\quad \left. + i\,\text{median}_j\{\sin(M'_{jk})\}\right]. \end{aligned} \tag{2.1}$$

For stable instruments, the true calibration parameters have very small variation over days, so we can use one set of rough calibration parameters from a single snapshot in time for data from all other times. Thus we pick a snapshot at noon when each v_{ij}^{model} can be easily computed from position of the Sun, and use the resulting raw calibration parameters as the starting point for logcal at all other times.

2.3.1.3 Log Calibration and Linear Calibration

To explain our redundant calibration procedure, we first need to briefly reintroduce the formalism developed in Liu et al. [87]. Suppose the i^{th} antenna measures a signal s_i at a given instant. This signal can be expressed in terms of a complex gain factor g_i, the antenna's instrumental noise contribution n_i, and the true sky signal x_i that would be measured in the limit of perfect gain and no noise:

$$s_i = g_i x_i + n_i. \tag{2.2}$$

Under the standard assumption that the noise is uncorrelated with the signal, each baseline's measured visibility is the correlation between the two signals from the two antennas:

$$\begin{aligned} v_{ij} &\equiv \langle s_i^* s_j \rangle \\ &= g_i^* g_j \langle x_i^* x_j \rangle + g_i^* \langle x_i^* n_j \rangle + g_j \langle n_i^* x_j \rangle + \langle n_i^* n_j \rangle \\ &= g_i^* g_j y_{i-j} + n_{ij}^{\text{res}}, \end{aligned} \tag{2.3}$$

where we have denoted the true correlation $\langle x_i^* x_j \rangle$ by y_{i-j},[15] the noise from each antenna by n_i, the noise for each baseline by n_{ij}^{res}, and expectation values (effectively time averages) by angled brackets $\langle \dots \rangle$. In a maximally redundant array such as MITEoR, the number of unique baselines is much smaller than the total number of baselines. Therefore, we can treat all the g_is and the y_{i-j}s as unknowns while keeping the system of equations (2.3) overdetermined, enabling fits for both despite the presence of instrumental noise.

In Liu et al. [87], some of us proposed logcal and lincal, and we have implemented both for calibrating MITEoR data. In log calibration, we take the logarithm of both sides of Equation 2.3 and obtain a linearized equation in logarithmic space. We then perform a least squares fit for the system of equations

$$\log v_{ij} = \log g_i^* + \log g_j + \log y_{i-j}, \tag{2.4}$$

where we solve for $\log g_i$ and $\log y_{i-j}$. Because the least squares fit takes place in log space whereas the noise is additive in linear space, the best fit results are biased. Linear calibration, on the other hand, is unbiased [87]. The lincal method performs a Taylor expansion of Equation 2.3 around initial estimates g_i^0 and y_{i-j}^0 and obtains a system of linearized equations

$$v_{ij} = g_i^{0*} g_j^0 y_{i-j}^0 + g_i^{1*} g_j^0 y_{i-j}^0 + g_i^{0*} g_j^1 y_{i-j}^0 + g_i^{0*} g_j^0 y_{i-j}^1, \tag{2.5}$$

where we solve for g_i^1 and y_{i-j}^1. For a detailed description of the logcal and lincal algorithms and their noise properties, we direct the reader to Liu et al. [87], Marthi and Chengalur [94]. We now describe some essential improvements to these algorithms.

Logcal was first thought to be unable to calibrate the phase component due to phase wrapping, since logcal has no way to recognize that 0° and 360° are the same quantity. Consider, for example a pair of redundant baselines that measure phases

---

[15] Following Liu et al. [87], we use y_{i-j} instead of y_{ij} to emphasize that in a redundant array, the number of unique baseline visibilities can be much smaller than number of measured visibilities. The complete expression should be $y_{u(i,j)}$, where $u(i,j)$ means that baseline ij corresponds to the uth unique baseline.

of 0.1° and 359.9° respectively. We can infer that they each only need a very small phase correction (±0.1°) to agree perfectly. However, since logcal treats the difference between them as 359.8° rather than 0.2°, it will calibrate by averaging 0.1° and 359.9° to 180°, which is completely wrong.

We made two improvements to the logcal method to guard against this. The first is to perform rough calibration beforehand, as described in Section 2.3.1.2. The second is to re-wrap the phases of v_{ij}. While rough calibration can make the phase errors relatively small[16], that improvement alone is not sufficient, since 0° and 360° are still treated as different quantities. Thus we need to intelligently wrap the phases of the input vector before feeding it into logcal. This is done in two simple steps. For a snapshot of rough calibrated visibilities at given time and frequency, v_{ij}, we first estimate the true phases of each group of redundant baselines, $\arg(y_{i-j})$, by computing median angles of measured phases using Eq. 2.1. Then for each measured phase, we add or subtract 2π until it is within $\pm\pi$ of $\arg(y_{i-j})$. This eliminates the phase wrapping problem.

Unlike logcal, lincal is an unbiased algorithm, but it relies on a set of initial estimates for the correct calibration solutions to start with. The output of lincal can be fed back into the algorithm and it can iteratively improve upon its own solution. However, the algorithm converges to the right answer only if the initial estimates are good. In practice, we find that three iterations of lincal typically produces excellent convergence, because the outputs of logcal are already decent estimates of the calibration solutions. Thus, by improving logcal, we also greatly improve lincal's effectiveness.

Our current calibration pipeline performs all steps of redundant calibration in less than 1 millisecond on a single processor core for a data slice at one time and one frequency channel, which is an order of magnitude faster than the rate data is saved onto disk. It is carried out by our open source `Omnical` package, coded in C++/Python.[17] Thus there should be no computational challenge in performing the

---

[16] In our experience, they need to be less than about 20 degrees to ensure that the subsequent calibration steps converge reliably.

[17] The package supports the miriad file format and is easily adapted to work with other file formats.

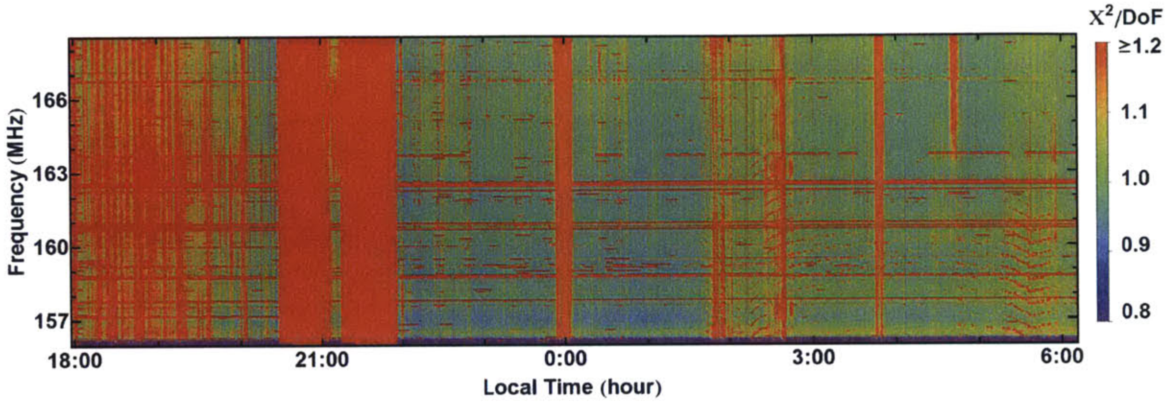

Figure 2-10: Waterfall plot of χ^2/DoF for a day's data. This demonstrates the stability of our instrument as well as the effectiveness of using χ^2/DoF as a indicator of data quality. We evaluate χ^2/DoF every 5.3 seconds and every 49 kHz. For the majority of the night time data, χ^2/DoF is close to 1. We flag all data with χ^2 larger than 1.2, which are marked red in this plot and account for 20% of this data set. The amount of detailed structure in the flagged area (around 18:00 for example) shows the χ^2 flagging technique's sensitivity to rapidly changing data quality.

above described calibration procedure in real-time for any array with less than 10^3 elements. For a future omniscope that has as many as 10^6 elements, there are two ways to reduce the computational cost. The first is to calibrate less frequently in time and frequency, and we will discuss in detail the minimal sampling frequency in Section 2.3.1.5. The other is to adapt a hierarchical redundant calibration scheme, where instead of calibrating all visibilities at the same time, one can calibrate the array in a hierarchical fashion whose computational cost scales only linearly with the number of elements. We discuss more details regarding hierarchical redundant calibration in Appendix 2.B.

2.3.1.4 χ^2 and Quality of Calibration

One of the many advantages of redundant calibration is it allows for the calculation of a χ^2 for every snapshot to quantify how accurate the estimated visibilities are for each unique baseline, even without any knowledge of the sky. For a set of visibilities at a given time and frequency, v_{ij}, with calibration results g_i and y_{i-j}, we define χ^2

To obtain a copy, please contact jeff_z@mit.edu.

as

$$\chi^2 = \sum_{ij} \frac{|v_{ij} - y_{i-j} g_i^* g_j|^2}{\sigma_{ij}^2}, \tag{2.6}$$

where σ_{ij}^2 is the noise contribution to the variance of the visibility v_{ij}. The effective number of degrees of freedom (DoF) is

$$\begin{aligned} \text{DoF} &= N_{\text{measurements}} - N_{\text{parameters}} \\ &= N_{\text{baselines}} - (N_{\text{antennas}} + N_{\text{unique baselines}}). \end{aligned} \tag{2.7}$$

The numerator in Equation 2.6 represents the deviation of measured data, v_{ij}, from the best fit redundant model, $y_{i-j} g_i^* g_j$. Thus χ^2/DoF can be interpreted as the non-redundancy in measured data divided by the expected non-redundancy from pure noise. If the data agrees perfectly with the redundant model (with noise) and is free from systematics, then χ^2/DoF is drawn from a χ^2 distribution with mean 1 and variance $2/\text{DoF}$ [1].

With a smooth model for σ_{ij} which we describe below, we compute χ^2/DoF for the results of rough calibration, log calibration, and linear calibration using all of our data. The χ^2 distributions of our calibrations for one day's data are shown in Figures 2-10 and 2-11. Each calibration algorithm significantly reduces the χ^2/DoF, and lincal's produces a distribution of χ^2/DoF consistently centered around 1. We automatically flag any data with χ^2/DoF larger than 1.2, which accounts for about 20% of the data. Among the data that is not flagged, 85% is accounted for by the theoretical χ^2 distribution. The 15% in the right tail is mostly attributable to a slightly optimistic noise model designed to avoid underestimating χ^2. This close agreement between predicted and observed χ^2-distributions for the lincal results suggests that except during periods that get automatically flagged, our instrument and analysis pipeline is free from significant systematic errors. The fully automatic nature of our calibration pipeline and data quality assessment is encouraging for future instruments with data volume too large for direct human intervention.

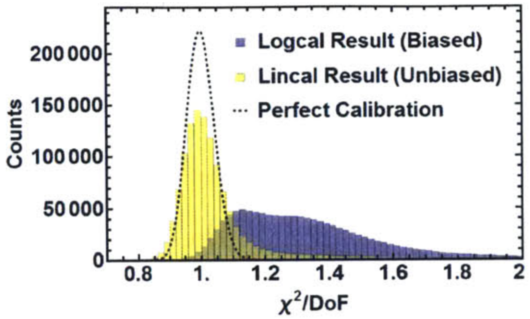

Figure 2-11: Histograms of the distributions of χ^2/DoF of logcal results (mean 1.31) and lincal results (mean 1.05, median 1.01), together with the theoretical distribution of χ^2/DoF (mean 1). They contain one night of data in a 12.5MHz frequency band (21:00-5:00 in Figure 2-10). We evaluate χ^2/DoF for every 5.3 seconds and every 49 kHz. We set the flagging threshold to $\chi^2 = 1.2$, and 80% of the lincal result is below the threshold (majority of the 20% flagged data have χ^2 much larger than 2, thus not shown in this figure). Among the data that is not flagged, 85% is accounted for by the theoretical χ^2 distribution. The right tail in lincal's distribution is due to the noise model sometimes underestimating the noise in order to minimize false negatives in the flagging process. The fact that χ^2/DoF for lincal is so close to the theoretical distribution means that both the instrument and the calibration algorithms are working exactly as we expect.

Calculating χ^2/DoF for flagging and data quality assessment requires an accurate model of noise in the measured visibilities. To compute the noise σ_{ij}, we approximate σ_{ij}^2 by $\langle\sigma^2\rangle$, where the average is over all baselines. This assumption that all antennas have the same noise properties drastically deceases the computational cost of calculating χ^2/DoF. Because we have 10^3 baselines, and the variation of σ_{ij} between baselines is less than 20% (due to slightly different amplifier gains), this approximation should cause only about a 1% error in the final χ^2/DoF.

To compute $\langle\sigma^2\rangle$, we perform linear regression on each visibility v_{ij} over one minute to obtain its estimated variance σ_{ij}^2, and then average all σ_{ij}^2 to obtain σ^2. Thus we have $\langle\sigma^2\rangle$ at all frequencies every minute. Before we plug $\langle\sigma^2\rangle$ into Equation 2.6, we model it as a smooth and separable function: $\langle\sigma^2\rangle(f,t) = F(f)T(t)$, where $F(f)$ and $T(t)$ are polynomials. The smooth model has three advantages. The first is that it is physically motivated to model thermal fluctuation as a smooth and separable function. Secondly, a smooth noise model makes the χ^2/DoF a much more sensitive flagging device. Theoretically, χ^2/DoF should not rise above 1 when unwanted radio events such as radio frequency interference (RFI) occur, because they are far field signals that do not violate any redundancy. However, since RFI events make both the signal and noise stronger, by demanding a smooth noise model, the $\langle\sigma^2\rangle$ we use will underestimate the noise during RFI events and give abnormally high χ^2/DoF, which can then be successfully flagged with the $\chi^2/\mathrm{DoF} < 1.2$ threshold. Thirdly, seasonal changes aside, the noise model is expected to largely repeat itself from day to day, so for future experiments that will operate for years, it suffices to use the same model repeatedly without recomputing σ_{ij} *in situ* for all the data. Thus, by using a smooth noise model, one can drastically reduce the occurrence of false negatives (since it is better to flag good data than it is to fail to flag bad data) as well as the computational cost of calculating χ^2/DoF.

2.3.1.5 Optimal Filtering of Calibration Parameters

While the above-mentioned estimates of the calibration parameters that we obtain from redundant baseline calibration vary over time and frequency, much of that variation is due to the noise in raw data. To minimize the effect of instrumental noise on the calibration parameters, we would like to optimally average information from nearby times and frequencies to estimate the calibration parameters for any particular measurement.

As we will show below, the optimal method for performing this averaging is Wiener filtering. In the rest of this section, we first measure the power spectrum of the calibration parameters over time and frequency, and make a determination of how to decompose this into contributions from signal (true calibration changes) and noise. We then weight the Fourier components in a way that is informed by their signal-to-noise ratio, and quantify how this Wiener filtering procedure improves upon more naive averaging over time and/or frequency. Finally, we discuss the implications for how regularly (in time and frequency) we should calibrate. It is important to note that while these methods are applied only to MITEoR below, they are applicable to any current or future experiment.

We model the measured calibration parameter $g_i(f, t)$ for the i^{th} antenna as the sum of a true calibration parameter $s_i(f,t)$ (the "signal") and uncorrelated noise $n_i(f,t)$:

$$g_i(f,t) = s_i(f,t) + n_i(f,t). \tag{2.8}$$

We choose our estimator $\hat{g}_i(f,t)$ of the true calibration parameter $s_i(f,t)$ to be a linear combination of the observed calibration parameters g_i at different times and frequencies:

$$\hat{g}_i(f,t) \equiv \int\int W(f,t,f',t')g_i(f',t')df'dt' \tag{2.9}$$

for some weight function W. We optimize the estimator $\hat{g}_i$ by choosing the weight function W that minimizes the mean-squared estimation error $\langle|\hat{g}_i(f,t) - s_i(f,t)|^2\rangle$. Assuming that the statistical properties of the signal and noise fluctuations are stationary over time[18], all correlation functions become diagonal in Fourier space:

$$\begin{aligned}
\langle\tilde{s}_i(\tau,\nu)^*\tilde{s}_i(\tau',\nu')\rangle &= (2\pi)^2\delta(\tau'-\tau)\delta(\nu'-\nu)S(\tau,\nu),\\
\langle\tilde{n}_i(\tau,\nu)^*\tilde{n}_i(\tau',\nu')\rangle &= (2\pi)^2\delta(\tau'-\tau)\delta(\nu'-\nu)N(\tau,\nu),\\
\langle\tilde{s}_i(\tau,\nu)^*\tilde{n}_i(\tau',\nu')\rangle &= 0,
\end{aligned} \tag{2.10}$$

where tildes denote Fourier transforms and S and N are the power spectra of signal and noise, respectively. This means that the optimal filter becomes a simple multiplication $\hat{\tilde{g}} = \tilde{W}\tilde{g}$ in Fourier space, corresponding to the weighting function $\tilde{W}(\tau,\nu)$ that minimizes the mean-squared error

$$\langle|\tilde{W}(\tau,\nu)\tilde{g}_i(\tau,\nu) - \tilde{s}_i(\tau,\nu)|^2\rangle. \tag{2.11}$$

Requiring the derivative of this with respect to $\tilde{W}$ to vanish gives the Wiener filter [184]

$$\tilde{W}(\tau,\nu) = \frac{S(\tau,\nu)}{S(\tau,\nu)+N(\tau,\nu)}. \tag{2.12}$$

Since S and N are to a reasonable approximation independent of the antenna number

---

[18] We perform this analysis on data over 12 MHz and two hours, where the signal and noise power are empirically found to be approximately time-independent.

i, we have dropped all subscripts i for simplicity. Back in real space, the optimal estimator $\hat{g}_i$ for the i^{th} calibration parameter is thus g_i convolved with the 2D inverse Fourier transform of $\tilde{W}$.

To demonstrate this technique, we show the above process carried out in the time dimension in Figure 2-12. In practice we perform the analysis on time and frequency dimensions simultaneously through a 2D FFT. The noise power spectrum $N(\nu)$ is seen to be constant to an excellent approximation, corresponding to white noise (uncorrelated noise in each sample). The signal power spectrum $S(\nu)$ is seen to be well fit by a combination of two power laws: $S(\nu) \approx (\nu/2.9 \times 10^{-5}\,\text{Hz})^{-2.7} + (\nu/4.8 \times 10^{-17}\,\text{Hz})^{-0.46}$. The optimal convolution kernel W is seen to perform a weighted average of the data on the timescale of roughly 200 s and frequency scale of 0.15 MHz, giving the greatest weight to nearby times and frequencies, resulting in an order-of-magnitude noise reduction.

To quantify the effectiveness of the obtained filter compared to naive "boxcar" averages, we use the 2D power spectrum and noise floor of the calibration parameters obtained from real data to simulate many realizations of calibration parameters $g(f,t) = s(f,t) + n(f,t)$, apply various averaging/convolution schemes $W(f,t)$ on the simulated data, and compare their effectiveness by computing the RMS error $\langle |(W \star g)(f,t) - s(f,t)|^2 \rangle$ normalized by $\langle |n|^2 \rangle$. Due to our limited frequency bandwidth as well as frequent RFI contamination, power spectrum modeling in the frequency dimension is very challenging, so the frequency Wiener filter appears to be less effective than the time filter. In Table 2.2 we list the normalized noise powers using frequency Wiener filter, time Wiener filter, 2D Wiener filter, as well as traditional boxcar averaging, and the 2D Wiener filter produces results three times less noisy than that of the traditional boxcar averaging.

We have described how to optimally average calibration parameters when we calibrate very regularly in time and frequency. For a future instrument such as an omniscope with 10^6 antennas, calibration will pose a serious computational challenge, so it is important to know what the minimal frequency one needs to calibrate the instrument. The above analysis conveniently provides an answer to this question. As

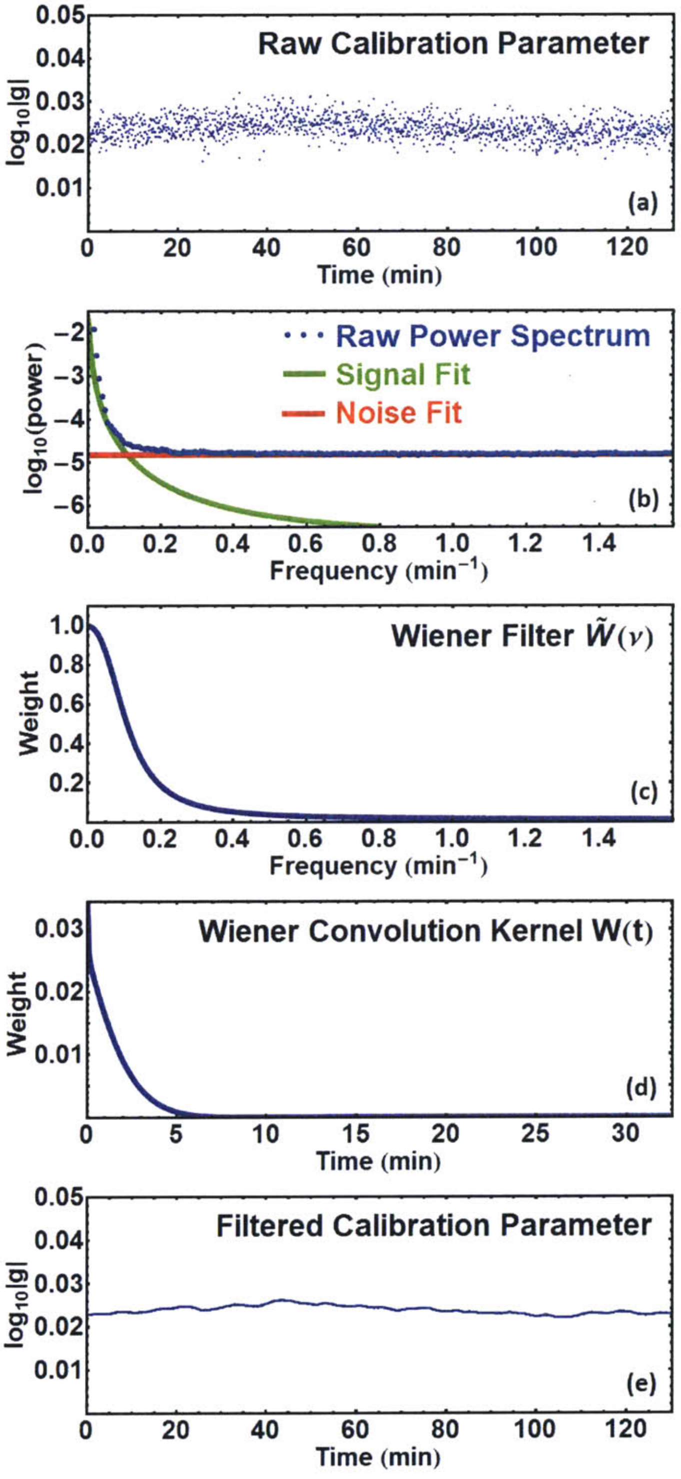

Figure 2-12: Illustration of 1D Wiener filtering of calibration parameters at different times. Panel (a) shows the amplitude of calibration parameters measured for one antenna over two hours. Panel (b) shows that the average power spectrum across all antennas (blue dots) is well fit by a white noise floor (red horizontal line) plus a sum of two power laws (green curve). Panel (c) shows the Wiener filter in frequency domain computed using Eq. 2.3.1.5 and the power spectra from panel (b). Panel (d) shows the Wiener convolution kernel in the time domain, the Fourier transform of the filter in Panel (c). Panel (e) shows the best estimates of the true calibration amplitude. The effectiveness of this filter is compared with that of other filters in Table 2.2.

Averaging method	Relative noise power
No average	1
Frequency Wiener filter	0.33
Time Wiener filter	0.12
Time and frequency Wiener filter	0.09
Time and frequency boxcar average	0.32

Table 2.2: Wiener filtering reduces the noise contribution to the calibration parameters by an order of magnitude. This table lists residual noise power (normalized by original noise power) after applying various filters to average the amplitude calibration parameters in time and/or frequency. The optimal two-dimensional Wiener filter indeed performed the best, lowering the noise power by an order of magnitude. In comparison, the naive boxcar average, using the characteristic scales of the optimal Wiener filter (200 s and 0.15 MHz), has more than 3 times residual noise power than the Wiener filtered result.

shown in the second panel of Figure 2-12, the signal[19] is band limited. By the Nyquist theorem, one needs to sample with at least double the frequency of signal bandwidth, so in our case we could measure the calibration parameters without aliasing problems as long as we calibrate once per minute. Calibrating more frequently than this simply helps average down the noise. Although this one-minute timescale depends on the temporal stability characteristics of the amplifiers and other components used in our particular experiment, it provides a useful lower bound on what to expect from future experiments whose analog chains are even more stable.

2.3.2 Absolute Calibration

The absolute calibration of the array involves two separate tasks. One is to find the overall gain and to break phase degeneracies that redundant baseline calibration is unable to resolve, and the other to calibrate fixed properties of the instrument such as the orientation of the array and shape of the primary beam. The former is done by comparing the data to a sky model comprised of the global sky model (GSM) [34] and published astronomical catalogs (see Jacobs et al. 73 for example). The latter is done using bright point sources with known positions. While we can take advantage of the

---

[19] We only show results for amplitude calibration parameters for brevity, as the phase calibration results have nearly identical power spectrum.

extremely high signal-to-noise data in the ORBCOMM channels (around 137 MHz), thanks to the dynamic range provided by our 8 bit correlator, it is important to note that all the algorithms described here are applicable to astronomical point sources as well.

This section is divided into three parts. The first part describes how we use prior knowledge of the sky to break the degeneracies in redundant calibration results, a vital step to obtain usable data products. The second and third part each describe one aspect of absolute calibration using satellite data: primary beam measurement and array orientation.

2.3.2.1 Breaking Degeneracies in Redundant Calibration

Redundant calibration alone cannot produce directly usable data products, due to the degeneracies intrinsic to the algorithms. There is one degeneracy in the amplitude of the calibration coefficients g_i, since scaling the amplitude of everything up by a common factor does not violate any redundancy (the degeneracies discussed here are per frequency and per time, as are the calibration solutions). There are three degeneracies in phase, corresponding to three degrees of freedom in a two dimensional linear field (see Appendix 2.A for a detailed discussion). In general, breaking these degeneracies requires prior knowledge of the sky. In this section, we briefly describe our algorithm that uses the global sky model (GSM) of de Oliveira-Costa et al. [34] to remove these degeneracies. Doing so requires efficiently simulating the response of the instrument to the GSM; we summarize a fast algorithm for doing so in Appendix 2.C. We defer detailed comparison of our data and the GSM to a future publication.

Our degeneracy removal procedure is an iterative loop that repeats two steps. The first step is to fit for the amplitude degeneracy factor. The knowledge of the GSM and bright point sources give us a set of model visibilities, m_{ij}^a, where index a denotes different modeled components such as the GSM or Cygnus A. A linear combination of these models should be able to fit our measurements[20]. Thus we fit for the weights

---

[20] We allow each model to have a separate weight to guard against potential calibration offsets between existing models.

w_a of the models by minimizing

$$\left| v_{ij} - \sum_a w_a m_{ij}^a \right|^2 . \tag{2.13}$$

The second step is to break the degeneracy in redundant phase calibration by fitting for the degeneracy vector $\boldsymbol{\Phi}$ and the constant ψ defined in Appendix 2.A. We assume that the error in the first step's fitting is mostly due to the phase degeneracies, so we take the w_a from step one and fit for $\boldsymbol{\Phi}$ and ψ by minimizing

$$\left| \arg(v_{ij}) - \arg\left(\sum_a w_a m_{ij}^a \right) - \boldsymbol{d}_{i-j} \cdot \boldsymbol{\Phi} - \psi \right|^2 , \tag{2.14}$$

where $\boldsymbol{d}_{i-j}$ is the position vector for baseline $i - j$.

Note that the two fitting processes described above are not independent of one another, so we repeat these steps until convergence is reached. We find that in practice, the errors converge within two iterations. Our preliminary result is illustrated in Figure 2-13, which shows that the data agree very well with current models.

2.3.2.2 Beam Measurement Using ORBCOMM Satellites

In general, *in situ* measurements of antenna primary beams over large fields of view pose a challenge to 21 cm cosmology, as primary beam uncertainties are intimately related to calibration, imaging, and catalog flux uncertainties [72]. Motivated by these difficulties, Pober et al. [137] present a solution that uses celestial point sources and assumes reflection symmetry of the beam, whereas Neben, Bradley, and Hewitt (in preparation) demonstrate high dynamic range beam measurement using the constellation of ORBCOMM satellites. Here, we present *in situ* primary beam measurements of the MWA bow-tie antennas using the ORBCOMM constellation. We take advantage of both the high signal-to-noise ratio of ORBCOMM signals, and of our full cross-correlation measurements (rather than auto-correlations alone) to determine the beam.

In order to measure their primary beam profile $B_{\mathrm{mwa}}(\hat{\mathbf{r}})$, we compare measure-

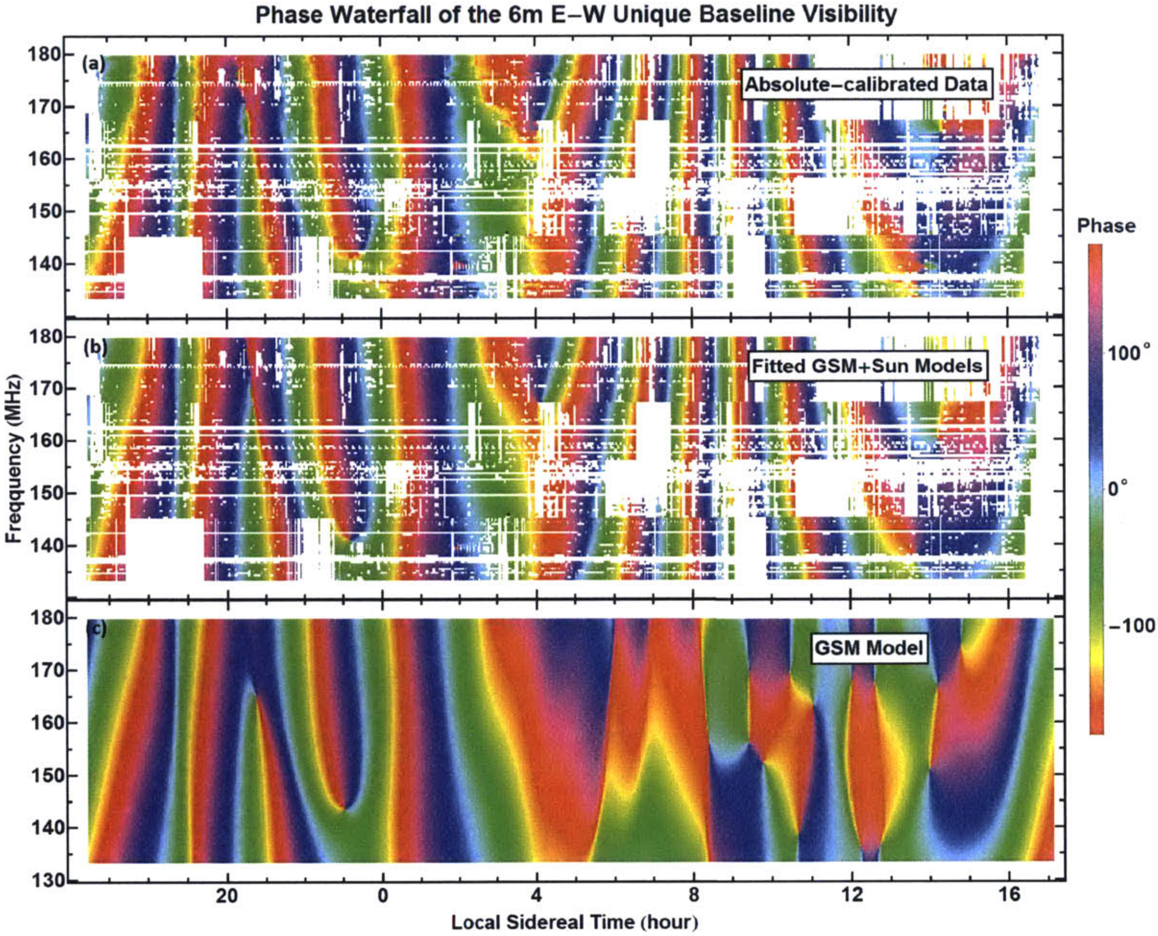

Figure 2-13: Waterfall plots of phases on the 6 m E-W baseline. These show that our absolute calibration successfully matches the data (panel (a)) with a linear combination of global sky model and known point sources, including the Sun (panel (b)). Panel (c) shows the global sky model alone. The white areas are flagged out using χ^2 criterion described in Figure 2-10. Each plot is stitched together using four independently measured and calibrated frequency bands, aligning local sidereal time. Thus the discontinuities between hours 4 and 12 are due to the Sun rising at different local sidereal times on different days of our observing expedition.

ments with MWA antennas to simultaneous measurements with simple center-fed dipoles, whose beam pattern B_{dipole} is known analytically. When there is a single extremely bright point source in the sky, such as an ORBCOMM satellite, we can compute the ratio of the visibilities of select baselines to obtain the ratio of the MWA antenna beam to the analytically known dipole antenna beam — thus determining the MWA antenna beam itself. To perform this analysis, two dipole antennas, one orientated along the x-polarization axis of the array and the other along the y-polarization axis, are added to the array and cross-correlated with all other MWA antennas.

The rationale behind this technique is as follows. At an angular frequency ω, the electric field from a sky signal at the position of a receiving antenna can be encoded in the Jones vector $\boldsymbol{S}(\hat{\boldsymbol{k}})$, where $\hat{\boldsymbol{k}}$ is the position vector of the source [31]. With a primary beam matrix $\mathbf{B}_j(\hat{\boldsymbol{k}})$, the signal measured by the j^{th} antenna at position $\boldsymbol{r}_j$ is

$$s_j = \int e^{-i[\boldsymbol{k}\cdot\boldsymbol{r}_j+\omega t]}\mathbf{B}_j(\hat{\boldsymbol{k}})\boldsymbol{S}(\hat{\boldsymbol{k}})\ \mathrm{d}\Omega. \tag{2.15}$$

When a single ORBCOMM satellite is above the horizon[21], its signal strength is so dominant at its transmit frequency that $\boldsymbol{S}(\hat{\boldsymbol{k}})$ becomes well-approximated by a point source at the satellite's location. The measured signal can then be written as:

$$s_j \approx e^{-i[\boldsymbol{k}_s\cdot\boldsymbol{r}_j+\omega t]}\mathbf{B}_j(\hat{\boldsymbol{k}}_s)\boldsymbol{S}_s, \tag{2.16}$$

where $\boldsymbol{k}_s$ is the wave vector of the satellite signal, and $\boldsymbol{S}_s$ is the Jones vector encoding the satellite signal strength.

If we limit our attention to either x-polarization or y-polarization and approximate the off diagonal terms of $\mathbf{B}(\hat{\boldsymbol{k}})$ as zero, the visibility for two antennas can be written as

$$v_{jk} \approx S^2 B_j(\hat{\boldsymbol{k}}_s)^* B_k(\hat{\boldsymbol{k}}_s)e^{-i\boldsymbol{k}_s\cdot(\boldsymbol{r}_k-\boldsymbol{r}_j)}. \tag{2.17}$$

If we take one visibility v_{ij} formed by correlating a simple center-fed dipole with an MWA bow-tie antenna and another visibility v_{kl} for the same baseline vector formed

---

[21] There is typically more than one ORBCOMM satellite above the horizon at any one time, but they are coordinated so that they do not transmit in the same frequency band simultaneously.

by correlating two MWA antennas, then their ratio is simply

$$\frac{|v_{ij}|}{|v_{kl}|} \approx \frac{|B_{\mathrm{mwa}}(\hat{\boldsymbol{k}}_s)|}{|B_{\mathrm{dipole}}(\hat{\boldsymbol{k}}_s)|}, \tag{2.18}$$

because the satellite intensity S and one MWA beam factor B_{mwa} all cancel out, and the phase factor $e^{-i\boldsymbol{k}_s\cdot(\boldsymbol{r}_j-\boldsymbol{r}_i)}$ is removed due to taking absolute values of the visibilities. This means that when a single point source dominates the sky, the ratio of visibility amplitudes is simply the ratio of the antenna beams at the direction of the point source. Since we already know the beam B_{dipole} of a center-fed dipole over a ground screen, we can directly infer the magnitude of MWA primary beam $|B_{\mathrm{mwa}}(\hat{\boldsymbol{k}}_s)|$.

In order to fully map out the MWA primary beam, we need to take data during many satellite passes until we have direction vectors that densely cover the entire sky. Satellite signals from 27 ORBCOMM satellites at 5 frequencies in the range of 137.2-137.8 MHz were identified. Their orbital elements are publicly available[22], so we can calculate $\hat{\boldsymbol{k}}_s(t)$ straightforwardly. With 40 hours of data taken at the frequencies of interest, we were able to obtain 248 satellite passes, shown in Figure 2-14.

We compared our measurements of the MWA primary beam using Equation 2.18 to numerical calculations using the FEKO electromagnetic modeling software package. Fixing an azimuth angle ϕ, we can plot and compare the simulated and measured beam at different polar angles θ (the angle between the direction vector and zenith). Figure 2-15 shows how the beam changes with θ for four different ϕ-values, where $\phi = 0$ correspond to North and increases clockwise. Our measurements of the MWA beams are seen to agree well with the numerical predictions for both polarizations. The small differences between the predicted and measured beams are larger than the statistical noise, implying that the main limitation is not noise but one or more of the above-mentioned approximations, or approximations in the electromagnetic antenna modeling.

---

[22] We obtained the TLE files from CelesTrak, a company that archives TLEs of many civil satellites.

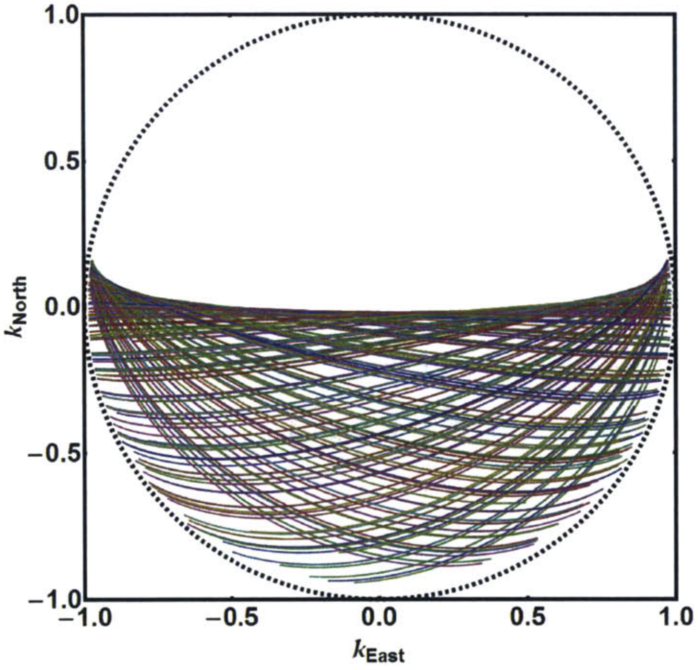

Figure 2-14: Projected trajectories of 248 passes of ORBCOMM satellites over 40 hours. With these passes we obtain sufficiently dense sampling of the MWA antennas primary beam that we can robustly map its response, especially at high elevations where the response is strongest. With a map of the southern half of the primary beam, we can use the reflection symmetry of the antennas to infer the entire beam at the ORBCOMM transmission frequencies. Each curve is a satellite pass projected onto the x-y plane, and the different colors specify sets of data taken at different times.

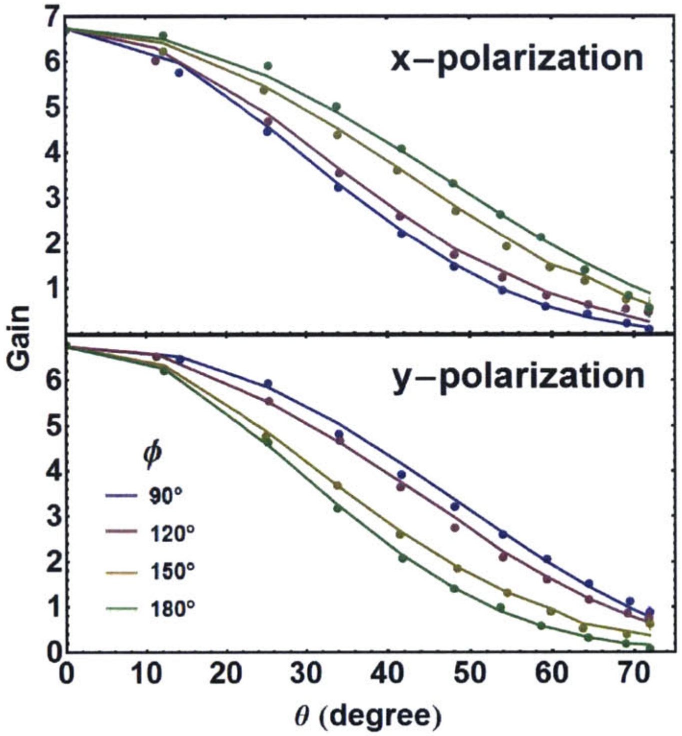

Figure 2-15: Measured MWA primary beam patterns compared to those obtained from numerical modeling. The two panels show the predictions (curves) and measurements (points) of the primary beam for the x-polarized and y-polarized MWA antennas. Each curve shows how the primary beam changes with the polar angle θ for a fixed azimuth angle ϕ. To reduce noise, the measurements have been averaged in 10 square degree bins.

2.3.2.3 Calibrating Array Orientation Using ORBCOMM Satellites and the Sun

The orientation of the array is very important, because the degeneracy removal process relies on the predicted measurement for each unique baseline, which in turn relies on precise knowledge of the baselines' orientations. Although we measured the relative position of each antenna to millimeter level precision with a laser-ranging total station, we did not measure the absolute orientation of the array to better than the $\sim 1^\circ$ accuracy obtainable with a handheld compass. To improve upon this crude measurement, we make use of both the known positions of both ORBCOMM satellites and the Sun. As we show in Figure 2-16, the exceptional signal-to-noise in the ORBCOMM data allows us to fit for a small array rotation as a first order correction to a model based on our crude measurement. Our method for finding the true orientation of the array is as follows. For a given baseline during the peak few minutes of an ORBCOMM satellite pass at frequency ν, we measure a phase $\phi(t)$. We also know the satellite's position vector $\boldsymbol{k}(t)$. However, we only have crude knowledge of baseline vector $\boldsymbol{d}_0$ in units of wavelength, where vectors are in horizontal coordinates with x, y, z that correspond to south, east and up. We can therefore only compute a crude prediction of the phase measurement

$$\phi_0(t) = 2\pi \boldsymbol{k}(t) \cdot \boldsymbol{d}_0. \tag{2.19}$$

We assume that the difference between the measurement $\phi(t)$ and our crude prediction $\phi_0(t)$ is due to a small angle rotation of the baseline vector $\boldsymbol{d}_0$ around the axis $\boldsymbol{\theta} = (\theta_x, \theta_y, \theta_z)$ by an angle $\theta = |\boldsymbol{\theta}|$, ignoring a constant cable length delay.[23] In the small

---

[23] Here it is important to use data before redundant baseline calibration to avoid phase degeneracy. We remove the phase delay from cables by allowing a constant offset that matches $\phi(t_M)$ with the crude prediction at time t_M when the satellite has the strongest signal during the pass.

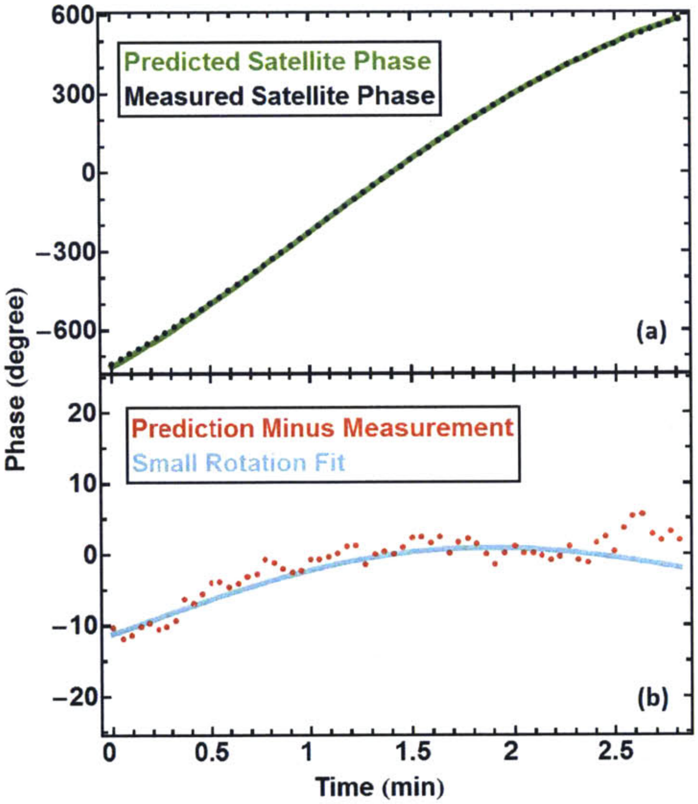

Figure 2-16: Illustration of using the high signal-to-noise ORBCOMM data to calculate any small rotation in the array relative to the field-measured orientation. Panel (a) shows the rapidly wrapping phase of the raw data (black) from one baseline at the ORBCOMM frequency during the peak three minutes of a single satellite pass. In green, we see the predicted values computed with the field-measured array orientation and publicly available satellite positions. The residual between the model and the data is plotted in red points in panel (b). Finally, the cyan curve shows the best fit using small angle rotations of the array. In practice we use hundreds of satellite passes and all the baselines to obtain a single accurate fit for the true orientation of the array.

θ regime, we have that

$$\begin{aligned}\phi(t) - \phi_0(t) &= 2\pi \boldsymbol{k}(t) \cdot (\mathbf{R}(\boldsymbol{\theta}) \cdot \boldsymbol{d}_0 - \boldsymbol{d}_0) \\ &\approx 2\pi \boldsymbol{k}(t) \cdot (\boldsymbol{\theta} \times \boldsymbol{d}_0) \\ &= 2\pi (\boldsymbol{d}_0 \times \boldsymbol{k}(t)) \cdot \boldsymbol{\theta}, \end{aligned} \tag{2.20}$$

where $\mathbf{R}(\boldsymbol{\theta})$ is the rotation matrix. Since we have a set of equations each representing a different time, the problem of finding $\boldsymbol{\theta}$ can be reduced to that of finding a least squares fit. With 117 satellite passes, we obtained the following best fit for the array rotation around the vertical axis:

$$\theta_z^{\text{sat}} = 0.66^\circ \pm 0.0005^\circ_{\text{stat}} \pm 0.07^\circ_{\text{sys}}.$$

While this method is very precise for solving the main problem we were worried about—the direction of North (θ_z) which we approximated in the field with a hand-held compass—it is less useful for measuring rotations of the array in the other two directions. Our instrument's absolute timing precision is only ~ 0.5 seconds, which makes it hard to distinguish rotations about the North-South axis from timing errors, as most ORBCOMM passes are East-West. This issue can of course be easily addressed in future experiments; for our experiment, we solve it using a more slowly moving bright point source: the Sun.

By using one day of solar data at 139.3 MHz, we obtained

$$\begin{aligned}(\theta_x, \theta_y, \theta_z)^\odot = &(-0.08^\circ, -0.12^\circ, 0.672^\circ) \\ &\pm(0.01^\circ, 0.03^\circ, 0.004^\circ)_{\text{stat}} \\ &\pm(0.04^\circ, 0.003^\circ, 0.005^\circ)_{\text{sys}}.\end{aligned}$$

Although solar data is noisier, in part because the Sun is not as bright as the ORBCOMM satellites in a given channel, timing errors are no longer important. These results agree with and complement the satellite-based results and allow us to confidently pin down the orientation of the array and thus improve the quality of the calibration of all of our data. The excellent agreement between the independent measurements $\theta_z^{\mathrm{sat}} \approx 0.66°$ and $\theta_z^{\odot} \approx 0.67°$ provides encouraging validation of both the satellite and solar calibration techniques.

2.3.3 Systematics

As we discussed in section 2.3.1.4, most of our data are well-calibrated with $\chi^2/\mathrm{DoF} < 1.2$, which means that any systematic effects should lie well below the level of the thermal noise. In this section we aim to identify all the systematic effects present in the system, and describe our efforts to quantify and, whenever possible, remove them. The systematics can be categorized into two groups:

1. Signal-dependent systematics that grow as the signal becomes stronger, such as cross-talk, antenna position errors and antenna orientation errors.
2. Signal-independent systematics, such as radio frequency interference (RFI) from outside or inside the instrument.

Below we find a strict upper bound of 0.15% for the signal-dependent component, as well as a signal-independent component which is easy to remove.

To quantify signal-dependent systematics, we again use ORBCOMM satellite data. Because the ORBCOMM signals are 10^3 times brighter than astronomical signals, and we know that any signal-independent systematics must be weaker than the astronomical signals (otherwise they would have been blatantly apparent in the data), any signal-independent systematics must be negligible compared to the ORBCOMM signal. We therefore investigate how the discrepancies between calibrated visibilities and the models for each unique baseline depend on ORBCOMM signal strength. We define the average fitting error per baseline at a given time and frequency to be

$$\epsilon = \langle |v_{ij} - y_{i-j} g_i^* g_j| \rangle, \tag{2.21}$$

which is a combination of antenna noise and systematic errors. If we compute ϵ at different times with different signal strength and compute its signal dependence, we can derive an upper bound on the signal dependence of systematic errors. To do this, we take all data at the ORBCOMM satellite frequency over a day and compute ϵ after performing redundant calibration. We then bin the ϵ-values according to the average signal strength $s = \langle |y_{i-j}| \rangle$, and obtain the results shown in Figure 2-17[24]. The result is seen to be well fit by a constant noise floor plus a straight line $\epsilon \approx 0.0015s$. This slope implies that the combined effect of all signal-dependent systematic effects is at most 0.15%. This is merely an upper bound on the systematics, since it is possible that the increase in ϵ is mainly due not to systematics but to an increase in instrumental noise caused by an increase in the system temperature during the ORBCOMM passes.

There is one signal-dependent systematic that is not included in the above analysis: deviation from redundancy caused by imperfect positioning of antenna elements. This is because the data we used to derive the upper bound is always dominated by a single point source, the ORBCOMM satellite, and redundant calibration cannot detect any deviation of antenna position when the sky is dominated by a single point source[25]. We have two ways to quantify the error in our data due to antenna position errors. Firstly, the laser-ranging measurements of antenna positions in the field indicate an average of 0.037 m deviation from perfect redundancy, which translates to about 2% average error in phase on each visibility. Since the deviations are in random directions, the variance of phase error in the unique baseline fits should be brought down by a factor equal to the number of redundant baselines, resulting in phase errors much less than 1% for most of the unique baselines. Secondly, although satellite calibration cannot detect position error in a given snapshot, over time the position errors would create very rapidly changing calibration parameters, which we do not observe in our

---

[24] Another way of describing these data points is that, if we look at the third panel in Figure 2-9, we are plotting the average small spread in each unique baseline group versus the radius of the circle, and as the satellites pass over, both the circle size and the amount of average spread change over time, forming the data set in question.

[25] This is because for any arbitrary position deviation $\Delta \boldsymbol{r}_i$ for antenna i, one can add a phase equal to $\boldsymbol{k} \cdot \Delta \boldsymbol{r}_i$ to the calibration parameter g_i to perfectly "mask" this deviation. Note that this "mask phase" depends on $\boldsymbol{k}$ and thus changes rapidly over time when the ORBCOMM satellite moves across the sky.

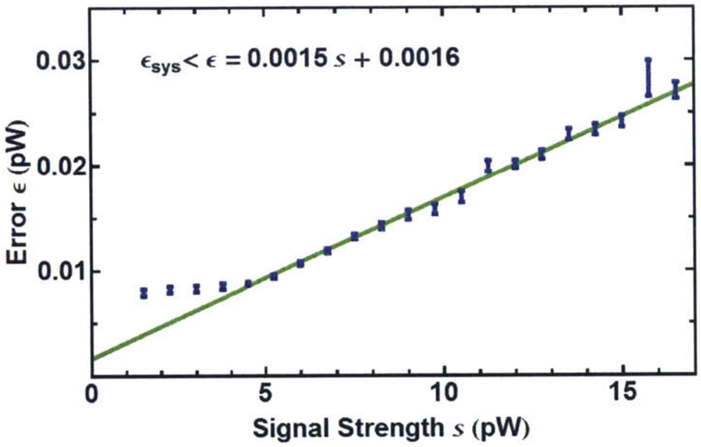

Figure 2-17: Signal-dependent systematic error and its linear fit. By comparing the modeled and calibrated visibilities during ORBCOMM satellite passes, we conclude that signal-dependent systematic effects account for no more than 0.15% of our measurement. We calculate the average fitting error per baseline $\epsilon = \langle|v_{ij} - y_{i-j}g_i^* g_j|\rangle$ and the average signal strength $s = \langle|y_{i-j}|\rangle$ binned over one day's data (blue points). The green line fits the data points above the noise floor. While many systematic errors, such as cross-talk, can contribute to the fitting error in addition to thermal noise, the best-fit slope of 0.0015 puts an upper bound on the sum of all signal-dependent errors. Since the ORBCOMM signal is so strong, any signal-independent systematic errors are negligible in this analysis. The high noise floor of ~ 0.01 pW is due to our digital tuning in the ORBCOMM frequency channels to maximize dynamic range.

data. Lastly, a formalism exists [87] to treat errors in antenna placement as small perturbations when redundantly calibrating, although though we did not need to take advantage of this technique for the present paper.

We first identified a signal-independent systematic when we obtained consistent $\chi^2/\mathrm{DoF} \sim 4$ for much of our data[26], which means that the fitting error was on average twice as large as the thermal noise in each visibility. This implies a systematic (or a combination of systematics) at the level of $10^{-6}\,\mathrm{pW/kHz}$, about 10% of the total astronomical signal. Given the above analysis, we can exclude the possibility of any signal-dependent explanations such as cross-talk between channels or antennas. While we are unable to offer any conclusive explanation of this systematic, it appears consistent with persistent near-field RFI, perhaps originating from our electronics. Fortunately, we found this additive signal to vary only very slowly over time, typically remaining roughly constant over 5-minute periods, which made it easy to remove. After calibrating the data with logcal, we average the fitting errors $\epsilon_{ij} = \langle v_{ij} - y_{i-j} g_i^* g_j \rangle_t$ over time and subtract them from the data before we run logcal again. We perform the averaging over 5 minute segments, corresponding to 112 independent time samples, and iterate the calibration-subtraction process three times. This corresponds to less than a 1% increase in the number of effective calibration parameters we fit for during logcal. Because many baselines probe the same unique baseline, the procedure described above exploits the redundancy of the array to robustly remove this slowly varying, signal-independent systematic, leaving us with $\chi^2/\mathrm{DoF} \sim 1$.

2.4 Summary and Outlook

We have described the MITEoR experiment, a pathfinder "omniscope" radio interferometer with 64 dual-polarization antennas in a highly redundant configuration. We have demonstrated a real-time precision calibration pipeline with automatic data quality monitoring that uses χ^2/DoF as a data quality metric to ensure that redundant

---

[26] This was before we obtained a consistent $\chi^2/\mathrm{DoF} \sim 1$ in Section 2.3.1.4, which occurred after we were able to remove the systematic described here.

baselines are truly seeing the same sky. We have also implemented various instrumental calibration techniques that utilize the ORBCOMM constellation of satellites to measure the primary beam and precise orientation of the array. Our success bodes well for future attempts to perform such calibration in real-time instead of in post-processing, and thus clears the way for FFT correlation that will make interferometers with $\gtrsim 10^3$ antennas cost-efficient by reducing the computational cost of correlating N antennas from an N^2 scaling to an $N \log N$ scaling. It also suggests that the extreme calibration precision required to reap the full potential of 21 cm cosmology is within reach.

The various calibration techniques that MITEoR successfully demonstrates are now being incorporated into the much more ambitious HERA project[27] [139], a broad-based collaboration among US radio astronomers from the PAPER, MWA, and MITEoR experiments. Our results are also pertinent to the design of the SKA low-frequency aperture array[28]. HERA plans to deploy around 331 14-meter dishes in a close-packed hexagonal array in South Africa, giving a collecting area of more than 0.05 km^2, virtually guaranteeing not only a solid detection of the elusive cosmological 21 signal but also interesting new clues about our cosmos.

2.A Appendix: Phase Degeneracy in Redundant Calibration

Both of our redundant baseline calibration algorithms, logcal and lincal (see Section 2.3.1.3), have the same set of phase degeneracies that require additional absolute calibration that must incorporate knowledge of the sky. When calibrating a given unique baseline, the quantity that logcal minimizes is

$$\sum_{jk} |(\theta_{j-k} - \phi_j + \phi_k) - \arg(v_{jk})|^2, \tag{2.22}$$

---

[27] `http://reionization.org/`
[28] `http://skatelescope.org/`

where we define $\theta_{j-k} \equiv \arg(y_{j-k}), \phi_j \equiv \arg(g_j)$. Similarly, lincal minimizes

$$\begin{aligned}
&\sum_{jk} |(y_{j-k}g_j^* g_k) - v_{jk}|^2 \\
=&\sum_{jk} \left||y_{j-k}g_j^* g_k| \exp\left[i(\theta_{j-k} - \phi_j + \phi_k)\right] - v_{jk}\right|^2. \qquad (2.23)
\end{aligned}$$

Unfortunately, for all values of θ_{j-k} and ϕ_k, one can add any linear field $\boldsymbol{\Phi}\cdot\boldsymbol{r}_j+\psi$ to the ϕ_j across the entire array while subtracting $\boldsymbol{\Phi}\cdot\boldsymbol{d}_j$ from the θ_{j-k} without changing the minimized quantities:

$$\begin{aligned}
\theta'_{j-k} - \phi'_j + \phi'_k \equiv&(\theta_{j-k} - \boldsymbol{\Phi}\cdot\boldsymbol{d}_{j-k}) - (\phi_j + \boldsymbol{\Phi}\cdot\boldsymbol{r}_j + \psi) \\
&+ (\phi_k + \boldsymbol{\Phi}\cdot\boldsymbol{r}_k + \psi) \\
=&\theta_{j-k} - \phi_j + \phi_k. \qquad (2.24)
\end{aligned}$$

Here $\boldsymbol{r}_j$ is the position vector of antenna j and $\boldsymbol{d}_{j-k} \equiv \boldsymbol{r}_k - \boldsymbol{r}_j$ is the baseline vector for the unique baseline with best-fit visibility y_{j-k}. Thus, the quantities in expressions 2.22 and 2.23 that the calibrations minimize are degenerate with changes to the linear phase field $\boldsymbol{\Phi}$ and the scalar ψ. This means that there are, in general, 4 degenerate phase parameters that need absolute calibration: one overall phase ψ and three related to the three degrees of freedom of the linear function $\boldsymbol{\Phi}$ (which reduces to two for a planar array).

In an ideal instrument, the measured visibilities for a given unique baseline would be

$$y_{i-j} = \int_{k_x,k_y} e^{i\boldsymbol{k}\cdot\boldsymbol{d}_{i-j}} S_B(k_x,k_y)dk_x dk_y, \qquad (2.25)$$

where $\boldsymbol{k} = (k_x, k_y, k_z)$ is the wave vector of incoming radiation and $S_B(k_x, k_y)$ is the product of the incoming signal intensity and the primary beam in the direction $\hat{\boldsymbol{k}}$ normalized by kk_z (which comes from the Jacobian of the coordinate transformation from spherical coordinates; see Tegmark and Zaldarriaga 166). When the array is coplanar, we can take an inverse Fourier transform of y_{i-j} and obtain an image of

$S_B(k_x, k_y)$. Above we saw that the best fit y_{i-j} computed by logcal and lincal is multiplied by an unknown linearly varying phase $\boldsymbol{\Phi} \cdot \boldsymbol{d}_{i-j}$. Since multiplication in *uv* space is a convolution in image space, this means that the image generated using these y_{i-j} is the true image convolved with a Dirac delta function centered at $\boldsymbol{\Phi}$, which corresponds to a simple shift by the unknown vector $\boldsymbol{\Phi}$ in the $S_B(k_x, k_y)$ image space.

To calibrate these last few overall phase factors, one can either make sure that bright radio sources line up properly in the image, or match phases between measured visibilities and predicted visibilities, as we described in Section 2.3.2.1. However, there may be another complementary way to remove this phase degeneracy without any reference to the sky. We know that physically the true image $S_B(k_x, k_y)$ is only non-zero within the disk $|k_x^2 + k_y^2|^{1/2} \leq k$ centered around the origin, and a shift caused by $\boldsymbol{\Phi}$ would move this circle off center. This suggests that we should be able to reverse engineer $\boldsymbol{\Phi}$ by looking at how much the image circle has been shifted, without knowing what $S_B(k_x, k_y)$ is. Figure 2-18 demonstrates how the image is shifted by $\boldsymbol{\Phi}$ using simulated data.

Unfortunately, this simple approach to identifying and removing the effect of $\boldsymbol{\Phi}$ suffers from a few complications. By far the most important one is the requirement of very short baselines. In the example in Figure 2-18, the shortest separation between antennas is 0.21λ, and it is easy to show that the sky disk is only clearly demarcated when the shortest separation is less than 0.5λ[29]. This sets a limit on the physical size of each element, which makes achieving a given collecting area proportionately more difficult. As Figure 2-19 shows, the deployed configuration of MITEoR cannot be used to reverse engineer the degeneracy vector $\boldsymbol{\Phi}$ without knowledge of the true sky.

---

[29] This is the 2D imaging counterpart of the well-known fact that, in signal processing, one must sample with a time interval shorter than $0.5\nu^{-1}$ to avoid aliasing in the spectrum of maximum frequency ν.

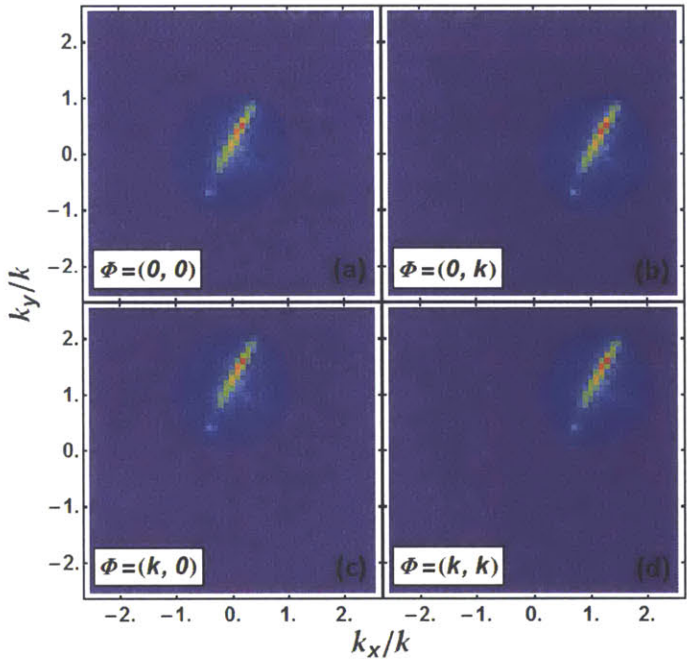

Figure 2-18: Illustration of phase degeneracies shifting the sky image where the sky disk is demarcated. The linear phase degeneracy, which takes the form $\boldsymbol{\Phi} \cdot \boldsymbol{r}_i$ in each antenna for any $\boldsymbol{\Phi}$, corresponds to a shift of the reconstructed image. These simulated images demonstrate shifts of fiducial sky image at 160 MHz caused by four different $\boldsymbol{\Phi}$, where the fiducial array's shortest baseline is 0.2 m. Panel (a) shows the image obtained from visibilities with no $\boldsymbol{\Phi}$, and the sky image is centered at 0. In the other three panels, the sky image is shifted by amount $\boldsymbol{\Phi}$. Even if one has no knowledge of what the true sky is, it is still possible to determine $\boldsymbol{\Phi}$ from where the sky image is centered.

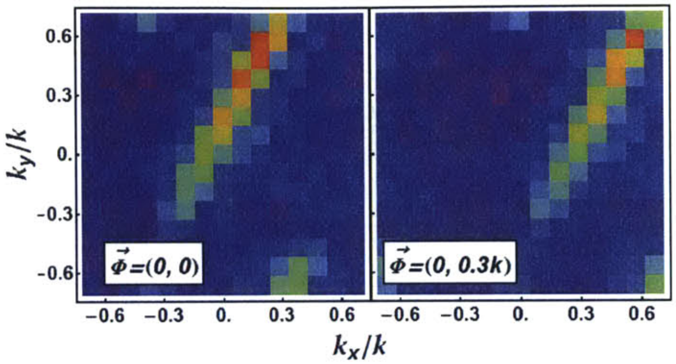

Figure 2-19: Illustration of phase degeneracies shifting the sky image where the sky disk is not demarcated. With any practical array configuration, including that of MITEoR, distinguishing image shifts caused by the $\boldsymbol{\Phi}$-degeneracy becomes significantly more difficult. These images demonstrate shifts of fiducial sky image at 160 MHz just as in Figure 2-18, but with MITEoR's compact configuration where the shortest baseline is 1.5 m. In the left panel, the image is obtained from visibilities with $\boldsymbol{\Phi} = (0, 0)$, and in the right panel the sky image is shifted by and amount $\boldsymbol{\Phi} = (0, 0.3k)$. Because the shortest baseline is too long (0.8λ), the Fourier transform of the visibilities only cover up to about 0.7 in k_x and k_y, so in contrast with Figure 2-18, it is impossible to determine $\boldsymbol{\Phi}$ by merely looking at where the sky image is centered without prior knowledge of the sky.

2.B Appendix: A Hierarchical Redundant Calibration Scheme with $\mathcal{O}(N)$ Scaling

One of the major advantages of an omniscope is its $N \log N$ cost scaling where N is the number of antennas. However, existing calibration techniques, including the ones presented in this paper, require all of the visibilities to compute the calibration parameters. Since the cost for computing the visibilities alone scales as N^2, this is a lower bound to the computational cost of existing calibration schemes regardless of the actual algorithm. While current instruments with less than 10^3 elements can afford full N^2 cross-correlation, such computation will be extremely demanding for a future omniscope with 10^4 or more elements. Thus, to take advantage of the $N \log N$ scaling of an omniscope with large N, it is necessary to have a calibration method whose cost scaling is less than $N \log N$. In this section, we describe a such a method using a hierarchical approach, and show that its computational cost scales only linearly with the number of antennas.

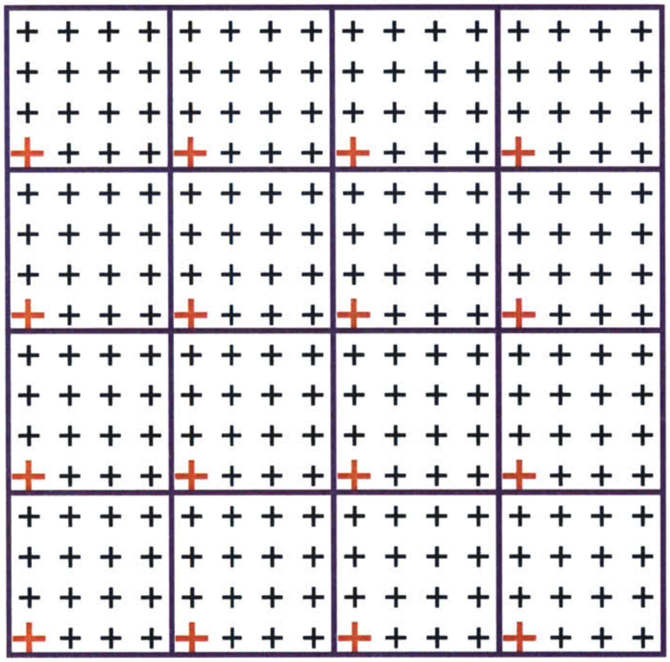

Figure 2-20: Example of the hierarchical calibration method for 256 antennas (marked by +-symbols) viewed as a 2-level hierarchy of 4 × 4 arrays ($m = 16, n = 2$). Our method first calibrates each sub-array independently with both relative and absolute calibrations. This produces calibration parameters for every antenna, up to one phase degeneracy ψ per sub-array. We can remove these 16 phase degeneracies among sub-arrays by choosing one antenna out of each sub-array (marked red and thick) and performing calibration on these 16 antennas. Thus we have calibrated the whole 256 antenna array by performing 16-antenna calibration 16+1=17 times. This can be generalized to a hierarchy with more levels by placing 16 such 256-antenna arrays in a 4 × 4 grid to get a 4096-antenna array, and then repeating to obtain arrays of exponentially increasing size. As shown in the text, the computational cost for this calibration method scales only linearly with the number of antennas.

Figure 2-20 illustrates the hierarchical calibration method for an example with a 256 antennas in a 16×16 regular grid, viewed as a 2-level hierarchy of 4×4 grids. More generally, consider an n-level hierarchy with m sub-arrays at each level, containing a total of $N = m^n$ antennas; the example in Figure 2-20 corresponds to $m = 16$, $n = 2$ and $N = 256$.[30] Let B_m denote the computational cost of calibrating a basic m-antenna array[31] Let C_n denote the computational .cost of calibrating the entire n-level hierarchy containing all N antennas. We have $C_1 = B_m$ by definition and

$$C_{n+1} = mC_n + B_m \tag{2.26}$$

since, as explained in the caption of Figure 2-20, we can calibrate the m sub-arrays at cost C_n each and then calibrate the m reference antennas (one from each sub-array) at cost B_m. Solving this recursion relation gives

$$\begin{aligned} C_n &= B_m(1+m(1+m(1+m(1+\cdots)))) \\ &= B_m\sum_{k=0}^{n-1} m^k = \frac{m^n-1}{m-1}B_m \\ &= \frac{N-1}{m-1}B_m = \mathcal{O}(N). \end{aligned} \tag{2.27}$$

Eq. 2.27 implies that for a fixed m, the computational cost for calibrating a 10^5 antenna array will be 10 times that of a 10^4 antenna array. Intuitively, the cost reduction comes from not computing cross-correlations among most pairs of antennas. In the simple case in Figure 2-20, only one visibility is computed between each pair of sub-arrays, rather than 256 visibilities in a full correlation scheme. Because of the reduced number of cross-correlations computed, we expect the quality of calibration parameters to be worse than that in the full correlation case. Since both the precision of calibration parameters and the computational cost depend on m, one can tune the

---

[30] It is easy to see that for a regular grid of N antennas, N need not an exact power of m to obtain the scaling that we will derive.

[31] B_m includes the cost to compute cross-correlations between the m antennas, as well as both relative and absolute calibrations. The cost B_m is unimportant for the scaling as long as it is independent of n.

value of m to achieve an optimal balance between precision and computational cost.

2.C Appendix: Fast Algorithm to Simulate Visibilities Using Global Sky Model

For both traditional self-calibration and the absolute calibration described in this paper, it is crucial to have accurate predictions for the visibilities. This requires simulation of both the contributions of bright point sources and diffuse emission, which can be added together due to the linearity of visibilities. While it is computationally easy to compute the contributions of point sources of known flux, it is much harder to compute visibilities from diffuse emission such as that modeled by the global sky model (GSM, de Oliveira-Costa et al. 34). Dominated by Galactic synchrotron radiation, this diffuse emission is especially important for the low frequencies and angular resolutions typical of current 21 cm experiments.

In general, we want to compute visibilities

$$y_u(f,t) = \int s(\hat{\boldsymbol{k}}, f, t) B(\hat{\boldsymbol{k}}, f) e^{i\boldsymbol{k}\cdot\boldsymbol{d}_u} d\Omega_k, \tag{2.28}$$

where $s(\hat{\boldsymbol{k}}, f, t)$ is the magnitude squared of the global sky model at time t in horizontal coordinates, and $B(\hat{\boldsymbol{k}}, f)$ the magnitude squared of the primary beam at a given frequency. Performing the integral by summing over all n_{pix} pixels in the GSM takes $\mathcal{O}(n_{\text{pix}} n_b n_f n_t)$ computational steps, where n_b is the number of unique baselines one simulates, n_f is the number of frequency bins, and n_t is the number of visibilities one simulates for one sidereal day.

The faster algorithm we describe here takes only $\mathcal{O}(n_{\text{pix}} n_b n_f)$ steps, by taking advantage of the smoothness of the primary beam as well as the diurnal periodicity in Earth's rotation. It applies only to drift-scanning instruments, so $B(\boldsymbol{k}, f, t) = B(\boldsymbol{k}, f)$ in horizontal coordinates, and is similar in spirit to the ideas proposed by Shaw et al. [158].

The key idea is to decompose Equation 2.28 as follows:

$$y_u(f,t) = \sum_{\ell,m} a^f_{\ell m} \mathcal{B}^{uf}_{\ell m} e^{im\phi(t)}, \tag{2.29}$$

where each $a^f_{\ell m}$ is a spherical harmonic component of the GSM at a given frequency, and each $\mathcal{B}^{uf}_{\ell m}$ is a spherical harmonic component of $B(\hat{\boldsymbol{k}}, f)e^{i\boldsymbol{k}\cdot\boldsymbol{d}_u}$, both in equatorial coordinates. In this appendix, we describe precisely how to perform this decomposition and why it decreases the computational cost of calculating visibilities from the GSM.

2.C.1 Spherical Harmonic Transform of the GSM

The GSM of de Oliveira-Costa et al. [34] is composed of three HEALPIX maps of size n_{side} describing different frequency-independent sky principal components $s^c(\hat{\boldsymbol{k}})$ and the relative weights of each component $w^c(f)$ that encode the frequency dependence. We can decompose the spatial dependence into spherical harmonics,

$$a^c_{\ell m} = \int Y^*_{\ell m}(\hat{\boldsymbol{k}}) s^c(\hat{\boldsymbol{k}}) d\Omega_k \tag{2.30}$$

in $\mathcal{O}(n_{\text{pix}}^{\frac{3}{2}})$ steps, due to the advantage of HEALPIX format [59]. The frequency dependence of the spherical harmonic coefficients of the sky is given by

$$a^f_{\ell m} = \sum_c a^c_{\ell m} w^c(f), \tag{2.31}$$

and the total complexity of computing the coefficients $a^f_{\ell m}$ is $\mathcal{O}(n_{\text{pix}}^{\frac{3}{2}}) + \mathcal{O}(n_f)$.

2.C.2 Spherical Harmonic Transform of the Beam and Phase Factors

Next, we would like to compute the spherical harmonics components of $B(\hat{\boldsymbol{k}}, f)e^{i\boldsymbol{k}\cdot\boldsymbol{d}_u}$:

$$\mathcal{B}^{uf}_{\ell m} = \int Y^*_{\ell m}(\hat{\boldsymbol{k}}) B(\hat{\boldsymbol{k}}, f) e^{i\boldsymbol{k}\cdot\boldsymbol{d}_u} d\Omega_k. \tag{2.32}$$

Substituting the spherical harmonic decompositions of $B(\hat{\boldsymbol{k}}, f)$ and $e^{i\boldsymbol{k}\cdot\boldsymbol{d}_u}$ gives

$$\begin{aligned}
\mathcal{B}^{uf}_{\ell m} = & \int Y^*_{\ell m}(\hat{\boldsymbol{k}}) \sum_{\ell' m'} B^f_{\ell', m'} Y_{\ell' m'}(\hat{\boldsymbol{k}}) \\
& \times \sum_{\ell'' m''} 4\pi i^{\ell''} j_{\ell''}\left(\frac{2\pi f}{c} d_u\right) Y^*_{\ell'' m''}(\hat{\boldsymbol{d}}_u) Y_{\ell'' m''}(\hat{\boldsymbol{k}}) d\Omega_k \\
= & \sum_{\ell' m'} \sum_{\ell'' m''} 4\pi i^{\ell''} j_{\ell''}\left(\frac{2\pi f}{c} d_u\right) B^f_{\ell' m'} Y^*_{\ell'' m''}(\hat{\boldsymbol{d}}_u) \\
& \times \int Y^*_{\ell m}(\hat{\boldsymbol{k}}) Y_{\ell' m'}(\hat{\boldsymbol{k}}) Y_{\ell'' m''}(\hat{\boldsymbol{k}}) d\Omega_k \\
= & \sum_{\ell' m'} \sum_{\ell'' m''} 4\pi i^{\ell''} j_{\ell''}\left(\frac{2\pi f}{c} d_u\right) B^f_{\ell' m'} Y^*_{\ell'' m''}(\hat{\boldsymbol{d}}_u) \\
& \times \sqrt{\frac{(2\ell+1)(2\ell'+1)(2\ell''+1)}{4\pi}} \\
& \times (-1)^m \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell'' \\ -m & m' & m'' \end{pmatrix},
\end{aligned} \tag{2.33}$$

where $j_\ell(x)$ is the spherical Bessel function, $\ell' m'$ represent quantum numbers when expanding the primary beam, $\ell'' m''$ represent quantum numbers when expanding $e^{i\boldsymbol{k}\cdot\boldsymbol{d}_u}$, and the Wigner-3j symbols are results of integrating the product of three spherical harmonics. Because the Wigner-3j symbols vanish unless $\ell - \ell' \leq \ell'' \leq \ell + \ell'$

and $-m+m'+m''=0$, the above sum simplifies to

$$\begin{aligned}\mathcal{B}^{uf}_{\ell m} &= \sum_{\ell' m'} \sum_{\ell''=\ell-\ell'}^{\ell+\ell'} 4\pi i^{\ell''} j_{\ell''}\left(\frac{2\pi f}{c} d_u\right) B^{f}_{\ell' m'} Y^{*}_{\ell'' m''}(\hat{\boldsymbol{d}}_u) \\ &\times \sqrt{\frac{(2\ell+1)(2\ell'+1)(2\ell''+1)}{4\pi}} \\ &\times (-1)^m \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell'' \\ -m & m' & m'' \end{pmatrix}, \end{aligned} \tag{2.34}$$

where $m''=m-m'$. Note that ℓ', m' and ℓ'' in this sum are all limited to the range of ℓ-values where the spherical harmonics components for the primary beam are non-zero, so the complexity for this triple sum is $n_{\mathcal{B}\text{pix}}^{\frac{3}{2}}$, where $n_{\mathcal{B}\text{pix}}$ is the number of non-zero spherical harmonics components for the primary beam. Since the cost for each $\mathcal{B}^{uf}_{\ell m}$ is $n_{\mathcal{B}\text{pix}}^{\frac{3}{2}}$, and there are $n_b n_f n_{\text{pix}}$ of them, the computational complexity of calculating all $\mathcal{B}^{uf}_{\ell m}$-coefficients scales like $\mathcal{O}(n_b n_f n_{\text{pix}} n_{\mathcal{B}\text{pix}}^{\frac{3}{2}})$.

2.C.3 Computing Visibilities

By performing a coordinate transformation on Equation 2.28 from horizontal coordinates (corresponding to the local Horizon at the observing site) to equatorial coordinates, the time dependence of $s(\hat{\boldsymbol{k}})$ is transferred to $B(\hat{\boldsymbol{k}})$ and $\boldsymbol{d}_u$. We can now calculate $y_u(f,t)$ by computing

$$\begin{aligned} y_u(f,t) &= \int s(\hat{\boldsymbol{k}}) B^{ft}(\hat{\boldsymbol{k}}) e^{i\boldsymbol{k}\cdot\boldsymbol{d}_u(t)} d\Omega_k \\ &= \sum_{\ell m} a^{*}_{\ell m} \mathcal{B}^{uft}_{\ell m}. \end{aligned} \tag{2.35}$$

Since the time dependence of $\mathcal{B}^{uft}_{\ell m}$ is a constant rotation along the azimuthal direction, we can write the above as

$$y_u(f,t) = \sum_{\ell m} a^{*}_{\ell m} \mathcal{B}^{uf}_{\ell m} e^{im\phi(t)} = \sum_{m} c^{uf}_{m} e^{im\phi(t)}, \tag{2.36}$$

where we have defined

$$c_m^{uf} \equiv \sum_\ell a_{\ell m}^* \mathcal{B}_{\ell m}^{uf}, \tag{2.37}$$

which can be computed in $\mathcal{O}(n_b n_f n_{\text{pix}})$ steps. Given c_m^{uf}, it is clear that we can evaluate Equation 2.36 using a fast Fourier Transform (FFT), whose cost is

$$\mathcal{O}(n_b n_f n_t \log(n_t)). \tag{2.38}$$

Note that this FFT in Equation 2.36 has no n_{pix} dependence, because we always need to zero-pad c_m to length n_t before the FFT. In summary, the total complexity of all of the above steps is

$$\begin{aligned}&\mathcal{O}\left(n_{\text{pix}}^{\frac{3}{2}}\right) + \mathcal{O}(n_f) + \mathcal{O}\left(n_b n_f n_{\text{pix}} n_{B\text{pix}}^{\frac{3}{2}}\right) \\ &+ \mathcal{O}\left(n_b n_f n_{\text{pix}}\right) + \mathcal{O}(n_b n_f n_t \lg(n_t)) \\ \approx &\mathcal{O}\left(n_b n_f n_{\text{pix}} n_{B\text{pix}}^{\frac{3}{2}}\right).\end{aligned} \tag{2.39}$$

This does not scale with n_t, unlike the naive integration's $\mathcal{O}(n_b n_f n_{\text{pix}} n_t)$. Thus with a spatially smooth beam whose $n_{B\text{pix}} \ll n_t^{\frac{2}{3}}$, the algorithm described here is much faster than the naive numerical integration approach described at the beginning of this Appendix.

Part II

Latest Epoch of Reionization Science Results

Chapter 3

PAPER-64 Constraints on Reionization: the 21 cm Power Spectrum at $z = 8.4$

The content of this chapter was submitted the The Astrophysical Journal *in February 2015 and published [4] as* PAPER-64 Constraints on Reionization: the 21 cm Power Spectrum at $z = 8.4$ *in August 2015. The authors are: Zaki S. Ali, Aaron R. Parsons, Haoxuan Zheng, Jonathan C. Pober, Adrian Liu, James E. Aguirre, Richard F. Bradley, Gianni Bernardi, Chris L. Carilli, Carina Cheng, David R. DeBoer, Matthew R. Dexter, Jasper Grobbelaar, Jasper Horrell, Daniel C. Jacobs, Pat Klima, David H. E. MacMahon, Matthys Maree, David F. Moore, Nima Razavi, Irina I. Stefan, William P. Walbrugh, and Andre Walker.*

3.1 Introduction

The *cosmic dawn* of the universe, which begins with the birth of the first stars and ends approximately one billion years later with the full reionization of the intergalactic medium (IGM), represents one of the last unexplored phases in cosmic history. Studying the formation of the first galaxies and their influence on the primordial IGM during this period is among the highest priorities in modern astronomy. During our cosmic dawn, IGM characteristics depend on the matter density field, the mass and clustering of the first galaxies [81], their ultraviolet luminosities [99], the abundance of X-ray sources and other sources of heating [143, 102], and higher-order cosmological effects like the relative velocities of baryons and dark matter [97, 180].

Recent measurements have pinned down the bright end of the galaxy luminosity function at $z \lesssim 8$ [16, 156] and have detected a few sources at even greater distances [43, 113]. In parallel, a number of indirect techniques have constrained the evolution of the neutral fraction with redshift. Examples include integral constraints on reionization from the optical depth of Thomson scattering to the CMB [132, 134], large-scale CMB polarization anisotropies [117], and secondary temperature fluctuations generated by the kinetic Sunyaev-Zel'dovich effect [101, 192, 8, 118, 55]. Other probes of the tail end of reionization include observations of resonant scattering of Lyα by the neutral IGM toward distant quasars (the 'Gunn-Peterson' effect) [46], the demographics of Lyα emitting galaxies [156, 173, 45], and the Lyα absorption profile toward very distant quasars [14, 15]. As it stands, the known population of galaxies falls well short of the requirements for reionizing the universe at redshifts compatible with CMB optical depth measurements [150, 151], driving us to deeper observations with, e.g., JWST and ALMA, to reveal the fainter end of the luminosity function.

Complementing these probes of our cosmic dawn are experiments targeting the 21 cm "spin-flip" transition of neutral hydrogen at high redshifts. This signal has been recognized as a potentially powerful probe of the cosmic dawn [54, 108, 144] that can reveal large-scale fluctuations in the ionization state and temperature of the IGM, opening a unique window into the complex astrophysical interplay between the first luminous structures and their surroundings. Cosmological redshifting maps each observed frequency with a particular emission time (or distance), enabling 21 cm experiments to eventually reconstruct three-dimensional pictures of the time-evolution of large scale structure in the universe. While such maps can potentially probe nearly the entire observable universe [93], in the near term, 21 cm cosmology experiments are focusing on statistical measures of the signal.

There are two complementary experimental approaches to accessing 21 cm emission from our cosmic dawn. So-called "global" experiments such as EDGES [19], the LWA [42], LEDA [60, 13], DARE [25], SciHi [181], BigHorns [160], and SARAS [129] seek to measure the mean brightness temperature of 21 cm relative to the CMB background. These experiments typically rely on auto-correlations from a small number of dipole elements to access the sky-averaged 21 cm signal, although recent work is showing that interferometric cross-correlations may also be used to access the signal [177, 142]. In contrast, experiments targeting statistical power-spectral measurements of the 21 cm signal employ larger interferometers. Examples of such interferometers targeting the reionization signal include the GMRT [116], LOFAR [175], the MWA [172], the 21CMA [130, 188], and the Donald C. Backer Precision Array for Probe the Epoch of Reionization (PAPER; [125]).

PAPER is unique for being a dedicated instrument with the flexibility to explore non-traditional experimental approaches, and is converging on a self-consistent approach to achieving both the level of foreground removal and the sensitivity that are required to detect the 21cm reionization signal. This approach focuses on spectral smoothness as the primary discriminant between foreground emission and the 21cm reionization signal and applies an understanding of interferometric responses in the delay domain to identify bounds on instrumental chromaticity ([126], hereafter P12b). This type of "delay-spectrum" analysis permits data from each interferometric baseline to be analyzed separately without requiring synthesis imaging for foreground removal. As a result, PAPER has been able to adopt new antenna configurations that are densely packed and highly redundant. These configurations are poorly suited for synthesis imaging but deliver a substantial sensitivity boost for power-spectral measurements that are not yet limited by cosmic variance ([123], hereafter P12a). Moreover, they are particularly suited for redundancy-based calibration [185, 87, 194], on which PAPER now relies to solve for the majority of the internal instrumental degrees of freedom (dof). The efficacy of this approach was demonstrated with data from a 32-antenna deployment of PAPER, which achieved an upper limit on the 21 cm power spectrum $\Delta^2(k) \leq (41\,\text{mK})^2$ at $k = 0.27h$ Mpc^{-1} ([127], hereafter P14). That upper limit improved over previous limits by orders of magnitude, showing that the early universe was heated from adiabatic cooling, presumably by emission from high-mass X-ray binaries or mini-quasars.

In this paper, we improve on this previous result using a larger 64-element deployment of PAPER and a longer observing period, along with better redundant calibration, an improved fringe-rate filtering technique, and an updated power-spectrum estimation pipeline. The result is an upper limit on $\Delta^2(k)$ of $(22.4\,\text{mK})^2$ in the range $0.15 < k < 0.5h$ Mpc^{-1} at $z = 8.4$. This result places constraints on the spin temperature of the IGM, and as is shown in a forthcoming paper, [140], this supports and extends previous evidence against extremely cold reionization scenarios. In Section 3.2 we describe the observations used in this analysis. In Sections 3.3 and 3.4, we discuss the calibration and the stability of the PAPER instrument. We then move on to a discussion of our power-spectrum analysis pipeline in Section 3.5. We present our results in Section 3.6 along with new constraints on the 21cm power spectrum. We discuss these results in Section 3.7 and conclude in Section 3.8.

3.2 Observations

We base our analysis on drift-scan observations with 64 dual-polarization PAPER antennas (hereafter, "PAPER-64") deployed at the Square Kilometre Array South Africa (SKA-SA) reserve in the Karoo desert in South Africa (30:43:17.5° S, 21:25:41.8° E). Each PAPER element features a crossed-dipole design measuring two linear (X,Y) polarizations. The design of the PAPER element, which features spectrally and spatially smooth responses down to the horizon with a FWHM of 60°, is summarized in [125] and [137]. For this analysis, we use only the XX and YY polarization cross-products.

As shown in Figure 3-1, PAPER-64 employs a highly redundant antenna layout where multiple baselines measure the same Fourier mode on the sky (P12a; P14). We rely on all 2016 baselines for calibration, but only use a subset of the baselines for the power spectrum analysis. This subset consists of three types of baselines: the 30-m strictly east-west baselines between adjacent columns (e.g. 49-41, black

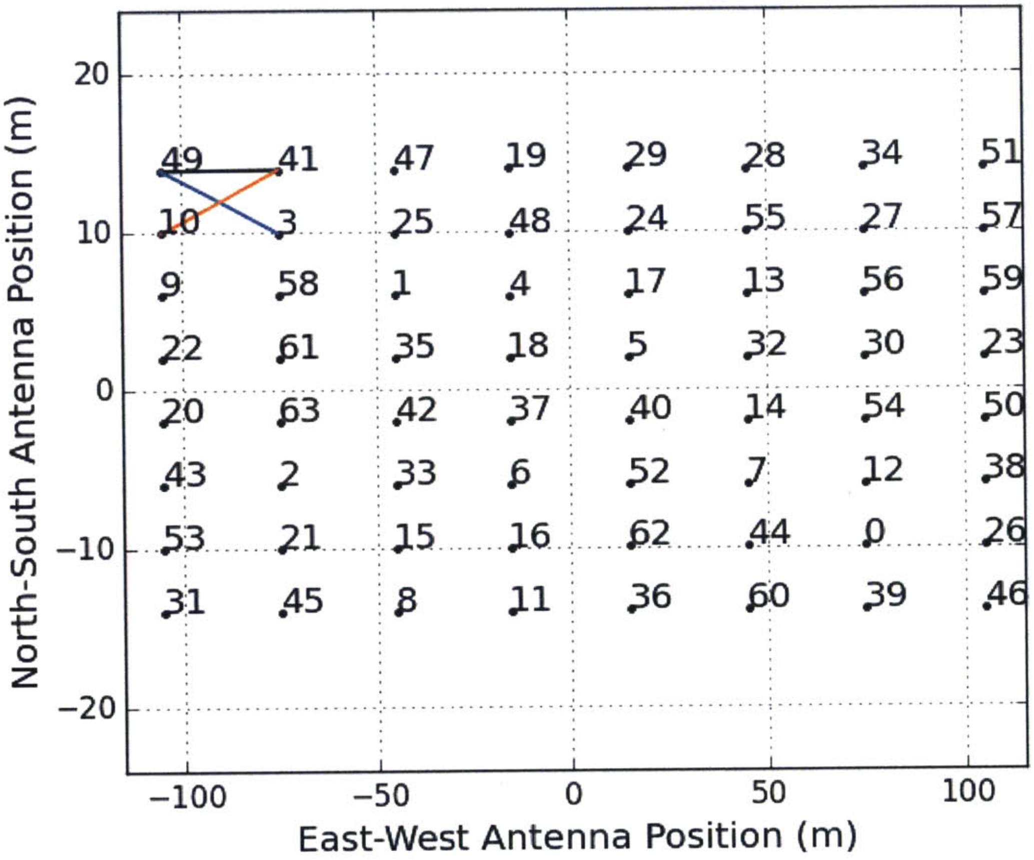

Figure 3-1: Antenna position within the PAPER-64 array. This analysis only makes use of east-west baselines between adjacent columns that have row separations of zero (black; e.g. 49-41, 41-47, 10-3, . . .) one in the northward direction (orange; e.g. 10-41, 3-47, 9-3, . . .) or one in the southward direction (blue; e.g. 49-3, 41-25, 10-58, . . .). Because of their high levels of redundancy, these baselines constitute the bulk of the array's sensitivity for power spectrum analysis.

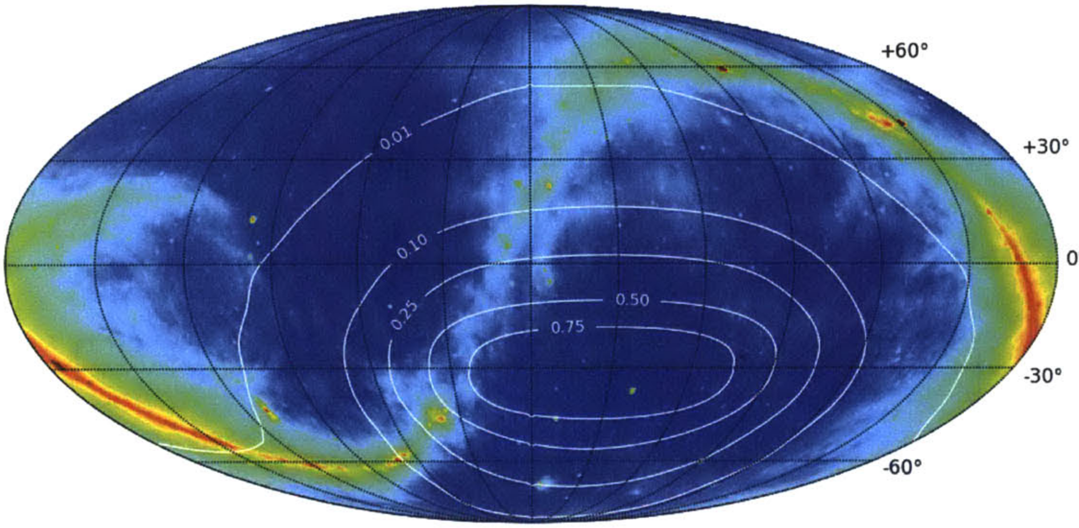

Figure 3-2: The Global Sky Model [34], illustrating foregrounds to the 21cm cosmological signal, with contours indicating beam-weighted observing time (relative to peak) for the PAPER observations described in Section 3.2. The map is centered at 6:00 hours in RA.

in Figure 3-1; hereafter referred to as *fiducial baselines*), 30-m east-west baselines whose eastern element is staggered one row up (e.g. 10-41, orange in Figure 3-1), and those whose eastern element is one row down (e.g. 49-3, blue in Figure 3-1). These baseline groups consist of 56, 49, and 49 baselines, respectively. We define a redundant group of baselines as being the set of baselines that have the same grid spacing; baselines in each of the three redundant groups described above are instantaneously redundant and therefore measure the same Fourier modes on the sky. Thus, within a redundant group, measurements from baselines may be coherently added to build power-spectrum sensitivity as N rather than $\sqrt{N}$, where N is the number of baselines added.

PAPER-64 conducted nighttime observations over a 135 day period from 2012 November 8 (JD 2456240) to 2013 March 23 (JD 2456375). Since solar time drifts with respect to local sidereal time (LST), this observing campaign yielded more samples of certain LSTs (and hence, sky positions) than others. For the power spectrum analysis, we use observations between 0:00 and 8:30 hours LST. This range corresponds to a "cold patch" of sky away from the galactic center where galactic synchrotron power is minimal, but also accounts for the weighting of coverage in LST. Figure 3-2 shows our observing field with the contours labeling the beam weighted observing time relative to the peak, directly over head the array.

The PAPER-64 correlator processes a 100–200 MHz bandwidth, first channelizing the band into 1024 channels of width 97.6 kHz, and then cross multiplying every antenna and polarization with one another for a total of 8256 cross products, including auto correlations. Following the architecture in [122], this correlator is based on CASPER[1] open-source hardware and signal processing libraries [121]. Sixteen ROACH boards each hosting eight 8-bit analog-to-digital converters digitize and channelize antenna inputs. New to this correlator relative to previous PAPER correlators [125], the cross multiplication engine is implemented on eight servers each receiving channelized data over two 10 Gb Ethernet links. Each server hosts two NVIDIA GeForce 580 GPUs running the open-source cross-correlation code developed by [30]. Visibilities are integrated for 10.7 s on the GPUs before being written to disk. All polarization cross-products are saved, although the work presented here only made use of the XX and YY polarization products.

3.3 Calibration

Foreground contamination and signal sensitivity represent the two major concerns for 21 cm experiments targeting power spectrum measurements. Sources of foregrounds include galactic synchrotron radiation, supernova remnants, and extragalactic radio sources. In the low-frequency radio band (50–200 MHz) where 21 cm reionization experiments operate, emission from these foregrounds is brighter than the predicted reionization signal by several orders of magnitude [155, 3, 34, 75, 11, 12, 56]. However, the brightest foregrounds are spectrally smooth, and this provides an important hook for their isolation and removal [85, 131, 84]. Unfortunately, interferometers, which are inherently chromatic instruments, interact with spectrally smooth foregrounds to produce unsmooth features that imitate line of sight Fourier modes over cosmolog-

---

[1] `http://casper.berkeley.edu`

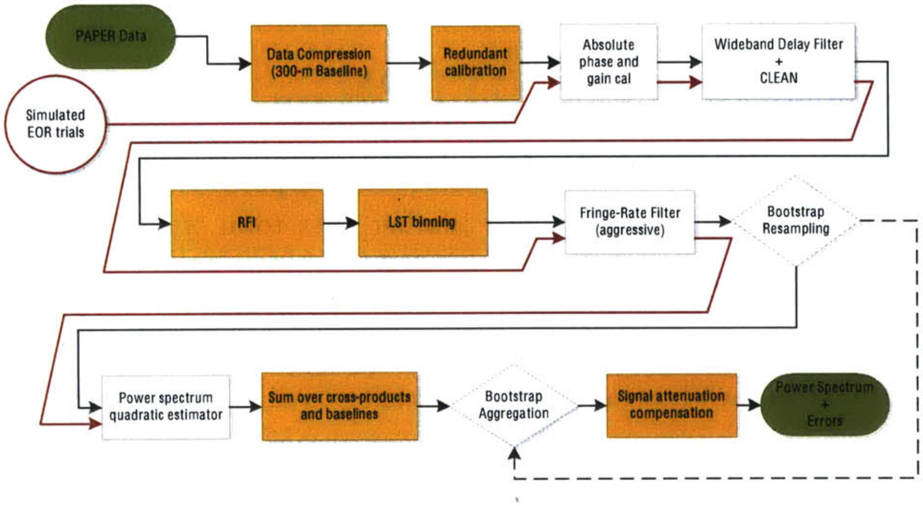

Figure 3-3: The stages of power-spectrum analysis. Black lines indicate data flow; red lines indicate Monte Carlo simulations used to measure signal loss. Yellow boxes indicate stages that by construction have negligible signal loss. Signal loss in other stages is tabluted in Table 3.1.

ical volumes (P12b; [109, 22]). One approach to solving this problem involves an ambitious calibration and modeling approach to spatially localize and remove foreground contaminants [86, 22, 61, 161, 29]. Perhaps the most impressive example of this approach is being undertaken by LOFAR, where dynamic ranges of 4.7 orders of magnitude have been achieved in synthesis images [189], although it is expected that additional suppression of smooth-spectrum foreground emission will be necessary [29].

The analysis for this paper employs a contrasting approach based on the fact that the chromaticity of an interferometer is fundamentally related to the length of an interferometric baseline. This relationship, known colloquially as "the wedge," was derived analytically (P12b; [176, 170, 89, 90]), and has been confirmed in simulations [32, 64] and observationally [138, 38]. As described in P12b, the wedge is the result of the delay between when a wavefront originating from foreground emission arrives at the two antennas in a baseline. The fact that this delay is bounded by the light-crossing time between two antennas (which we call the "horizon limit" since such a wavefront would have to originate from the horizon) places a fundamental bound on the chromaticity of an interferometric baseline. So far, PAPER has had the most success in exploiting this bound (P14; [74]). In this analysis, we continue to use the properties of the wedge in order to isolate and remove smooth spectrum foregrounds.

As illustrated in Figure 3-3, our analysis pipeline begins by running a compression algorithm to reduce the volume of our raw data by a factor of 70. As described in Appendix A of P14, this is achieved by first performing statistical flagging to remove radio frequency interference (RFI) at the 6σ level, applying low-pass delay and fringe-rate filters that limit signal variation to delay scales of $|\tau| \lesssim 1\mu$s and fringe-rate scales of $f \lesssim 23$ mHz, and then decimating to critical Nyquist sampling rates of 493 kHz along the frequency axis and 42.9 s along the time axis. We remind the reader that while information is lost in this compression, these sampling scales preserve emission between $-0.5 \leq k_{\parallel} \leq 0.5h\,\mathrm{Mpc}^{-1}$ that rotates with the sky, making this an essentially lossless compression for measurements of the 21 cm reionization signal in these ranges.

After compression, we calibrate in two stages, as described in more detail below. The first stage (Section 3.3.1) uses instantaneous redundancy to solve for the majority of the per-antenna internal dof in the array. In the second stage (Section 3.3.2), standard self-calibration is used to solve for a smaller number of absolute phase and gain parameters that cannot be solved by redundancy alone. After suppressing foregrounds with a wide-band delay filter (Section 3.3.3) and additional RFI flagging and crosstalk removal, we average the data in LST (Section 3.3.4) and apply a fringe-rate filter (Section 3.3.5) to combine time-domain data. Finally, we use an OQE (Section 3.5) to make our estimate of the 21 cm power spectrum.

3.3.1 Relative Calibration

Redundant calibration has gained attention recently as a particularly powerful way to solve for internal dof in radio interferometric measurements without simultaneously having to solve for the distribution of sky brightness ([185, 87, 112, 94, 194]; P14). The grid-based configuration of PAPER antennas allows a large number of antenna calibration parameters to be solved for on the basis of redundancy (P14; P12a; [194]). Multiple baselines of the same length and orientation measure the same sky signal.

Differences between redundant baselines result from differences in the signal chain, including amplitude and phase effects attributable to antennas, cables, and receivers. Redundant calibration only constrains the relative complex gains between antennas and is independent of the sky. Since redundant calibration preserves signals common to all redundant baselines, this type of calibration does not result in signal loss.

In practice, redundant calibration often takes on two flavors: log calibration (LOGCAL) and linear calibration (LINCAL) [87, 194]. LOGCAL uses logarithms applied to visibilities,

$$v_{ij} = g_i^* g_j y_{i-j} + n_{ij}^{res}, \tag{3.1}$$

where g denotes the complex gain of antennas indexed by i and j, and y represents the "true" visibility measured by the baseline, to give a linearized system of equations

$$\log v_{ij} = \log g_i^* + \log g_j + \log y_{i-j}, \tag{3.2}$$

In solving for per-antenna gain parameters with a number of measurements that scales quadratically with antenna number, redundancy gives an over-constrained system of equations that can be solved using traditional linear algebra techniques. While LOGCAL is useful for arriving at a coarse solution from initial estimates that are far from the true value, LOGCAL has the shortcoming of being a biased by the asymmetric behavior of additive noise in the logarithm [87].

LINCAL, on the other hand, uses a Taylor expansion of the visibility around initial estimates of the gains and visibilities,

$$v_{ij} = g_i^{0*} g_j^0 y_{i-j}^0 + g_i^{1*} g_j^0 y_{i-j}^0 + g_i^{0*} g_j^1 y_{i-j}^0 + g_i^{0*} g_j^0 y_{i-j}^1, \tag{3.3}$$

where 0 denotes initial estimates and 1 represents the perturbation to the original estimate and is the solutions we fit for. Using initial estimates taken from LOGCAL, LINCAL constructs an unbiased estimator.

Redundant calibration was performed using OMNICAL[2] — an open-source redun-

---

[2] https://github.com/jeffzhen/omnical

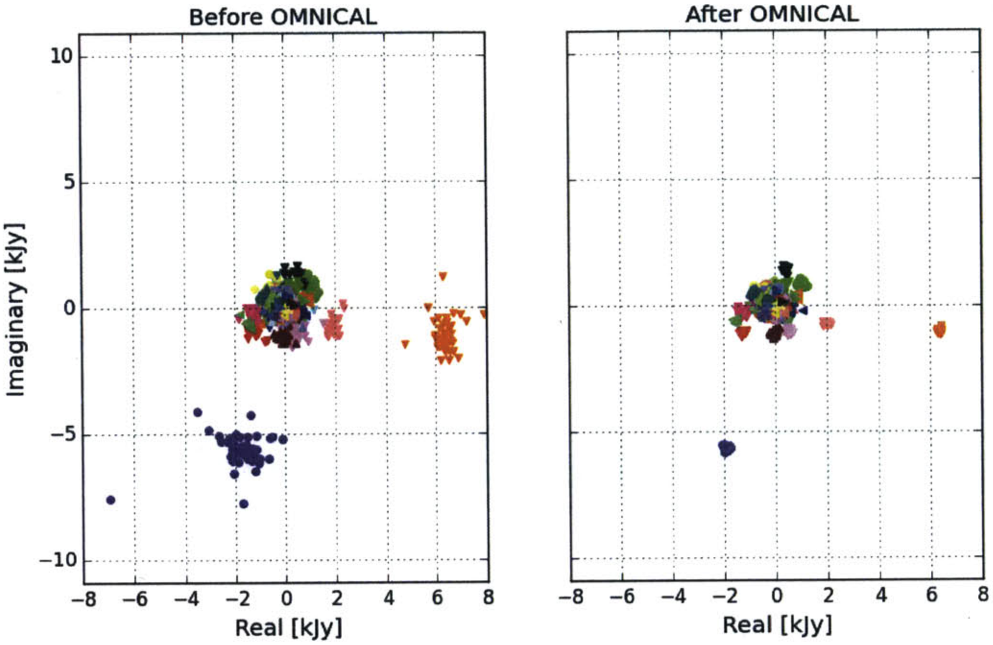

Figure 3-4: PAPER visibilities plotted in the complex plane before (left) and after (right) the application of the improved redundancy-based calibration with OMNICAL [194]. All baselines in the array measured at 159 MHz for a single time integration are plotted. Instantaneously redundant baselines are assigned the same symbol/color. The tighter clustering of redundant measurements with OMNICAL indicates improved calibration.

dant calibration package that is relatively instrument agnostic [194]. This package implements both LOGCAL and LINCAL, solving for a complex gain solution per antenna, frequency, and integration. The solutions are then applied to visibilities and the results are shown in Figure 3-4.

In addition to solving for gain solutions, OMNICAL also characterizes the quality of the calibration parameters by calculating the χ^2 for every integration. As defined in [194],

$$\chi^2 = \sum_{ij} \frac{|v_{ij} - y_{i-j} g_i^* g_j|^2}{\sigma_{ij}^2}, \tag{3.4}$$

where σ^2 is the noise in the visibilities. The χ^2 measures sum of the deviation of measured visibilities to that of the best fit model derived from the LINCAL relative to a noise model, and gives us a tool to use in order to check the quality of our data.

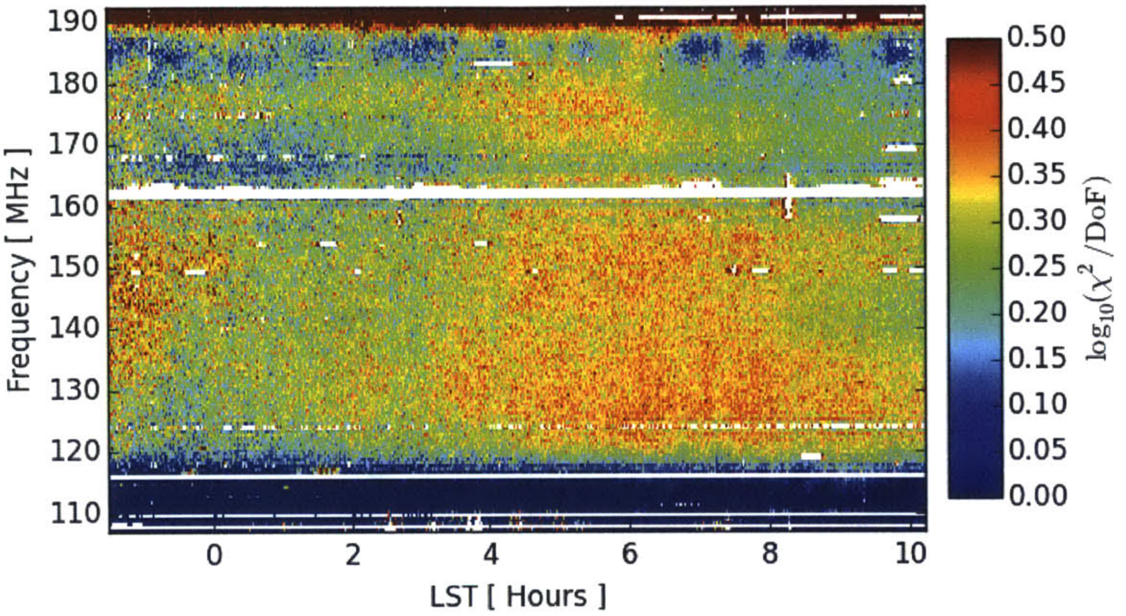

Figure 3-5: Log of χ^2 per degree of freedom of all baseline residuals after the application of OMNICAL. The plot comprises a observations over one day, with a frequency resolution of 493 kHz and a time resolution of 42.9 s.

The number of dof, as defined in [194], is given by

$$\begin{aligned} \mathrm{dof} &= N_{\mathrm{measurements}} - N_{\mathrm{parameters}} \\ &= 2N_{\mathrm{baselines}} - 2(N_{\mathrm{antennas}} + N_{\mathrm{unique\ baselines}}), \end{aligned} \tag{3.5}$$

and is effectively the number of visibilities for which χ^2 is calculated. If the data are noise-dominated, χ^2/dof is drawn from a χ^2 distribution with $\mu = 1$ and $\sigma^2 = 2/\mathrm{dof}$. The calculated χ^2/dof for every frequency and integration of a fiducial day of observation in this season and for the fiducial power spectrum baselines is shown in Figure 3-5, demonstrating the stability of the PAPER instrument.

We measure a mean χ^2/dof of 1.9. This indicates that the redundant calibration solutions, while a substantial improvement over the previous PAPER-32 calibration (P14), do not quite result in residuals that are thermal noise dominated. Possible sources of this excess include instrumental crosstalk and poorly performing signal chains. While the latter will be down-weighted by the inverse of the estimated signal covariance described in Section 3.5, crosstalk is a defect in the data that must be addressed. Crosstalk caused by the cross-coupling of signals between antennas reveals itself as a static complex bias to a visibility that varies on timescales much longer than typical fringe rates. This effect skews the distribution of the χ^2 of the residuals away from 1. To minimize crosstalk, we first use OMNICAL to solve for antenna-dependent gains, and then average the residual deviations from redundancy over 10 minute windows before subtracting the average from the original visibilities. This crosstalk removal preserves signals common to redundant baseline groups (such as the 21 cm signal). Unfortunately, it also preserves a term that is the average of the crosstalk of all baselines in the redundant group. This residual crosstalk is removed by a fringe-rate filter later in the analysis.

3.3.2 Absolute Calibration

After solving for the relative complex gains of the antennas using redundant calibration, an overall phase and gain calibration remains unknown. We use the standard self calibration method for radio interferometers to solve for the absolute phase calibration. We used Pictor A, Fornax A, and the Crab Nebula to fit for the overall phase solutions. Figure 3-6 shows an image of the field with Pictor A (5:19:49.70, -45:46:45.0) and Fornax A (3:22:41.70,-37:12:30.0).

We then set our over all flux scale by using Pictor A as our calibrator source with source spectra derived in [73],

$$S_\nu = S_{150} \times \left(\frac{\nu}{150MHz}\right)^{\alpha}, \tag{3.6}$$

where $S_{150} = 381.88$ Jy $\pm\, 5.36$ and $\alpha = -0.76 \pm 0.01$, with 1σ error bars.

To derive the source spectrum from our measurements, we use data that have been LST-averaged prior to the wide-band delay filter described in Section 3.3.3, for the hour before and after the transit of Pictor A. We image a 30° × 30° field of view for every frequency channel for each 10 minute snapshot and apply uniform weights to the gridded visibilities. We account for the required three-dimensional Fourier transform in wide field imaging by using the w-stacking algorithm implemented in WSclean [114] although we note that the standard w-projection algorithm implemented in CASA[3] gives similar performances as the PAPER array is essentially instantaneously coplanar. A source spectrum is derived for each snapshot by fitting a two-dimensional Gaussian to Pictor A by using the PyBDSM[4] source extractor. Spectra are optimally averaged together by weighting them with the primary beam model evaluated in the direction of Pictor A. To fit our bandpass, we divide the model spectrum by the measured one and fit a 9th order polynomial over the 120-170 MHz frequency range. Figure 3-7 shows the calibrated Pictor A spectrum and the model spectrum from [73]. Also plotted are the 1σ error bars derived from the PyBDSM source extractor and averaged over the multiple snapshots used after being weighted by the beam-squared.

Fitting a polynomial to the bandpass has the potential for signal loss which would include suppressing modes that may contain the cosmological signal. In order to quantify the signal loss associated with fitting a ninth degree polynomial to the bandpass, we run a Monte Carlo simulation of the effect the bandpass has on a model 21-cm reionization signal. We construct a model baseline visibility as a Gaussian random signal multiplied by the derived bandpass for every independent mode measured. We calculate the total number of independent modes by counting the number of independent uv-modes sampled for the different baseline types over the two hour time interval used to measure the bandpass. We average each mode together and fit a 9th degree polynomial. Using this as our measured bandpass for this simulated signal, we finally compare the power spectrum from the output of the simulated signal to the input power spectrum as a function fo k-mode. We find that between $-0.06 < k < 0.06$, the width of our wideband delay filter described below, the signal loss is less than 3% and at the mode right outside the above limit is $2 \times 10^{-7}\%$. We apply the latter correction factor for all modes outside the width of the delay filter to the final power spectrum.

---

[3] http://casa.nrao.edu

[4] http://www.lofar.org/wiki/doku.php?id=public:user_software:pybdsm

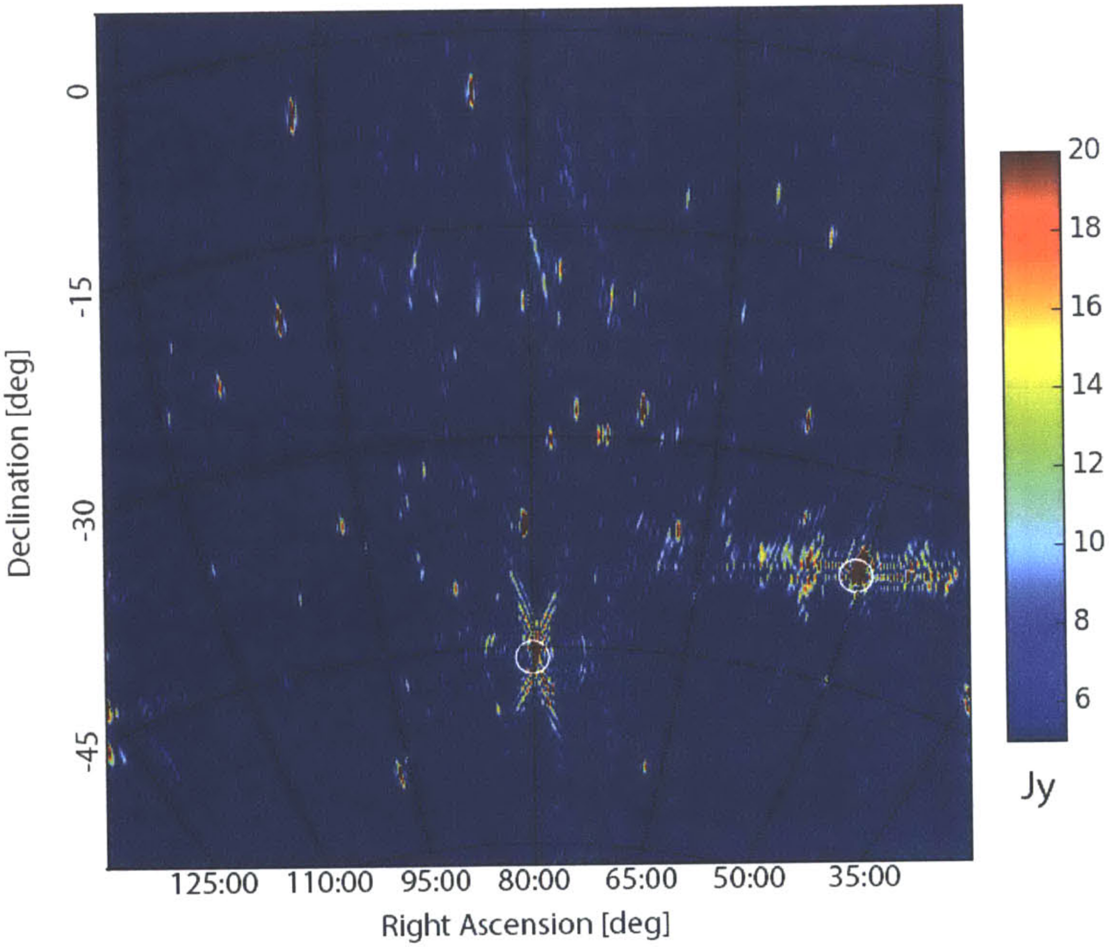

Figure 3-6: PAPER-64 image of a field including Pictor A and Fornax A, with white circles indicating catalog positions [71]. Image was synthesized with two hours of visibilities while Pictor A was in transit and 53 MHz of instantaneous bandwidth from 120 to 173 MHz. Image quality is limited by the redundant configuration of the array (e.g. grating lobes as a result of periodic antenna spacing, elongated lobes arising from poor uv-coverage in the north-south direction). Nonetheless, this image demonstrates accurate phase calibration over a wide field of view.

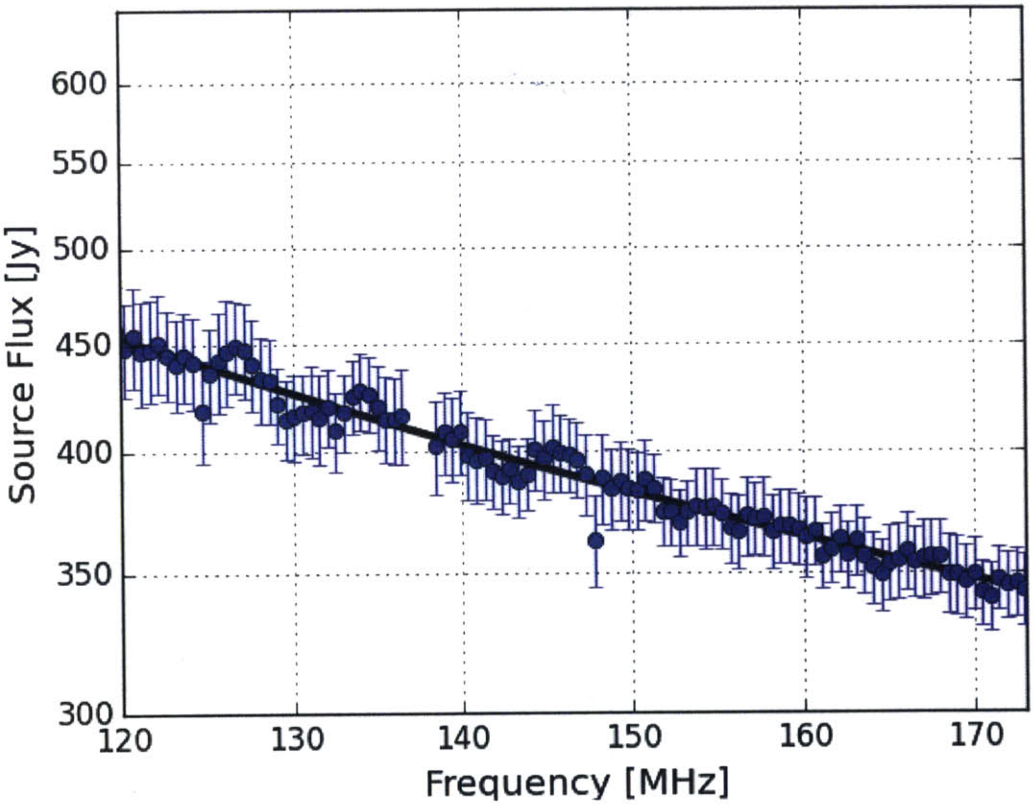

Figure 3-7: Measured spectrum of Pictor A in Stokes I (blue) relative to its catalog value (black; [73]). Flux density measurements are extracted from images of Pictor A, made independently for each frequency channel in 10 minutes snapshots as Pictor transits between hour angles of -1:49 and 1:10. Each measurement is then divided by the PAPER beam model and averaged to obtain the measured spectrum, which serves to characterize the flux scale of the PAPER-64 observations. Error bars indicate 68% confidence intervals, derived from the Gaussian fits in the source extractor used to measure the flux density in PyBDSM, combined from all snapshots.

3.3.3 Wideband Delay Filtering

Before implementing our foreground removal techniques, we combine the two linear polarizations for an estimate of Stokes I as per [105]. Namely, Stokes I can be estimated as

$$V_{\mathrm{I}} = \frac{1}{2}(V_{\mathrm{XX}} + V_{\mathrm{YY}}), \tag{3.7}$$

where V_{XX} and V_{YY} are the visibilities of the two linear polarizations measured by the interferometer. There are some important caveats to the estimate of Stokes I provided by Equation (3.7). One important caveat is that it neglects the beam asymmetry between the two linear polarization states. This mismatch can cause polarization leakage from Stokes Q into Stokes I, thus contaminating our measurement of the power spectrum with any polarized emission from the sky. This effect for PAPER, as shown in [105], leaks 4% of Q in to I in amplitude (2.2×10^{-3} in the respective power spectra). We take the conservative approach and do not correct for this effect, noting that the leakage of Q in to I will result in positive power, increasing our limits.

Foreground removal techniques discussed in the literature include spectral polynomial fitting [182, 22, 85], principal component analysis [115, 83, 116, 95], non-parametric subtractions [61, 29], and inverse covariance weighting [83, 37, 38, 89, 90], Fourier-mode filtering [131], and per-baseline delay filtering described in P12b. This delay-spectrum filtering technique is well-suited to the maximum redundancy PAPER configuration which is not optimized for the other approaches where high fidelity imaging is a prerequisite. The delay-spectrum foreground filtering method is described in detail by P14; its application is unchanged here. In summary; we Fourier transform each baseline spectrum into the delay domain

$$\begin{aligned}\tilde{V}_\tau &= \int W_\nu A_\nu I_\nu e^{-2\pi i \tau_g} \cdot e^{2\pi i \tau \nu} d\nu \\ &= \tilde{W}_\tau * \tilde{A}_\tau * \tilde{I}_\tau * \delta(\tau_g - \tau),\end{aligned} \tag{3.8}$$

where A_ν is the frequency dependent antenna response, W_ν is a sampling function that includes RFI flagging and a Blackman-Harris tapering function that minimizes delay-domain scattering from RFI flagging, and I_ν is the source spectrum. In the delay domain, a point source appears as a δ-function at delay τ_g, convolved by the Fourier transforms of the source spectrum, the antenna response, and the sampling function. We note that the antenna response effectively determines a finite bandpass, which imposes a lower bound of $1/B \approx 10\,\mathrm{ns}$ on the width of any delay-domain convolving kernel. As per [124] and P14, we deconvolve the kernel resulting from $W(\tau)$ using an iterative CLEAN-like procedure [68] restricting CLEAN components to fall within the horizon plus a 15-ns buffer that includes the bulk of the kernels convolving the δ-function in Equation (3.8). To remove the smooth spectrum foreground emission we subtract the CLEAN components from the original visibility.

Applying the delay filter to fiducial baselines used in the power spectrum analysis, foregrounds are suppressed by ~4 orders of magnitude in power, or −40 dB of foreground suppression, as seen in Figure 3-8. As discussed in P14, there is a small amount of signal loss associated with this filter. For the baselines and filter parameters used, the loss was found to be 4.8% for the first mode outside of the horizon, 1.3% for the next mode out, and less than 0.0015% for the higher modes.

3.3.4 Binning in LST

After the wideband delay filter, we remove a second layer of RFI which was overshadowed by the foreground signal. RFI are excised with a filter which flags values 3σ above the median using a variance calculated in a localized time and frequency window.

We then average the entire season in LST with 43-s bin widths, matching the cadence of the compressed data. The full season was 135 days long; of these, 124 days were included in the average. We make two separate LST-binned data sets, averaging every other Julian day together to obtain an "even" and "odd" dataset. The use of these two data sets allows us to construct an unbiased power spectrum estimate.

Sporadic RFI events result in measurements that, in any individual LST bin, de-

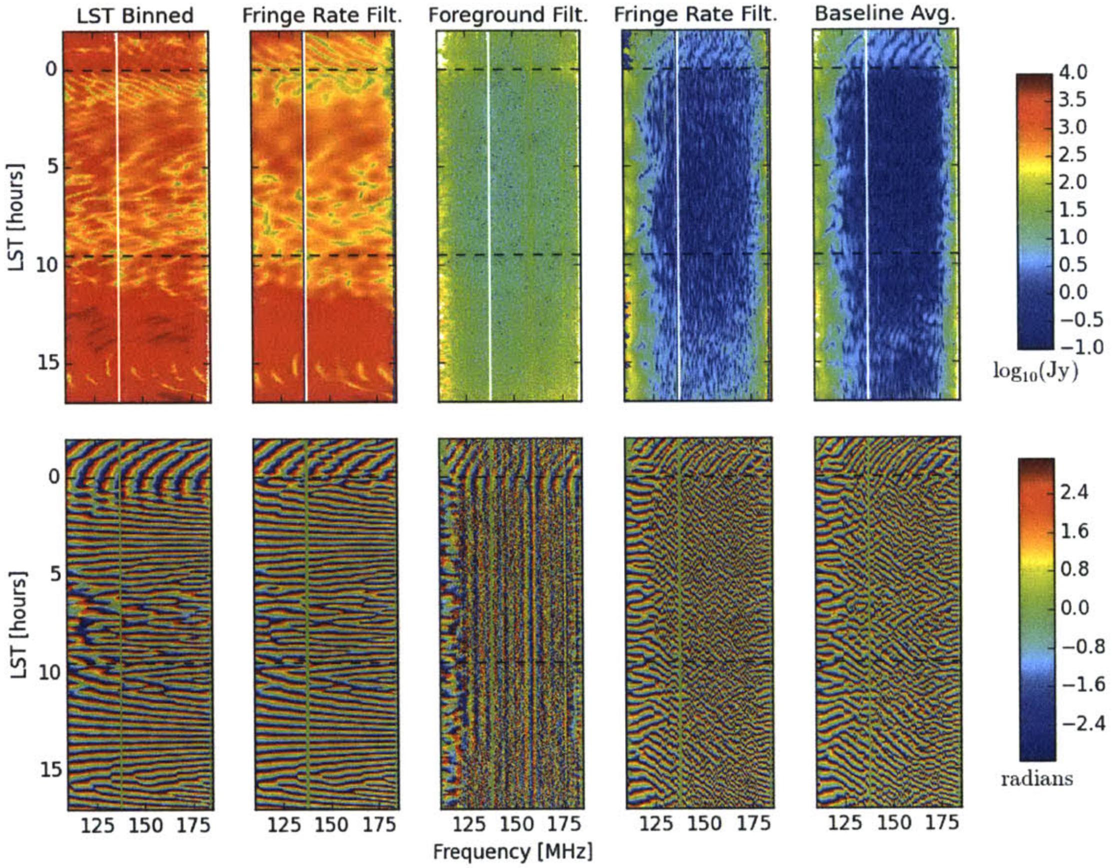

Figure 3-8: Visibilities measured by a fiducial baseline in the PAPER-64 array, averaged over 135 days of observation. From left to right, columns represent data that: (1) contain foregrounds prior to the application of a wideband delay filter or fringe-rate filtering, (2) are fringe-rate filtered but not delay filtered, (3) are delay filtered at 15 ns beyond the horizon limit but are not fringe-rate filtered, (4) are both delay and fringe-rate filtered, and (5) are delay and fringe-rate filtered and have been averaged over all redundant measurements of this visibility. The top row shows signal amplitude on a logarithmic scale; the bottom row illustrates signal phase. Dashed lines indicate the 0:00–8:30 range in LST used for power spectrum analysis. The putative crosstalk is evident in the center panel as constant phase features which do not fringe as the sky. The two right panels show some residual signal in the phase structure which is present at low delay. Away from the edges of the observing band, over four orders of magnitude of foreground suppression is evident.

viate from the Gaussian distribution characteristic of thermal noise. To catch these events, we compute the median of a LST bin for each frequency and flag values 3σ above the median, before averaging. Since we are narrowing the distribution of visibilities about the median, the measured thermal noise variance is not preserved under this filter. However, since the central value is preserved, the expectation value of the measured visibility in each LST bin is unchanged, and there is no associated signal loss for power spectrum measurements. Moreover, because errors are estimated empirically through bootstrapping (see Section 3.5.4), the slight increase in measurement error associated with truncating the tails of the Gaussian distribution are naturally accounted for.

3.3.5 Fringe-rate Filtering

By averaging visibilities in time, we aim to maximize sensitivity by coherently combining repeated measurements of k-modes before squaring these measurements and averaging over independent k-modes to estimate the power spectrum amplitude. This is mathematically similar to the more traditional process of gridding in the uv plane, but applied to a single baseline. However, rather than applying a traditional boxcar average, we can apply a kernel — a so-called "fringe-rate" filter — that weights different temporal rates by the antenna beam corresponding to the parts of the sky moving at the same rate.

For a given baseline and frequency, different parts of the sky exhibit different fringe-rates. Maximum fringe rates are found along the equatorial plane, where the rotation rate of the sky is highest, and zero fringe rates are found at the poles, where the sky does not rotate and hence sources do not move through the fringes of a baseline [124]. Fringe rates are not constant as a function of latitude. Bins of constant fringe rate correspond to rings in R.A. and decl., where the east–west projection of a baseline projected toward a patch of the sky is constant. We use this fact in conjunction with the root-mean-squared beam response for each contour of constant fringe rate to construct a time average kernel or "fringe-rate filter."

As examined in [128], it is possible to tailor fringe-rate filters to optimally combine time-ordered data for power-spectrum analysis. Fringe-rate filters can be chosen that up-weight points of the sky where our instrument is more sensitive and down-weight those points farther down in the primary beam, which are less sensitive. For white noise, all fringe-rate bins will contain the same amount of noise, but the amount of signal in each bin is determined by the primary beam response on the sky. By weighting fringe-rate bins by the rms of the beam response, we can get a net increase in sensitivity.

Applying this filter effectively weights the data by another factor of the beam area, changing the effective primary beam response[5], $A(l, m)$ [128]. By utilizing prior knowledge about the beam area, we are selectively down-weighting areas on the sky contributing little signal. This will result in a net improvement in sensitivity depending on the shape of the beam and the decl. of the array. For PAPER, this filter roughly doubles the sensitivity of our measurements.

Generally, a fringe-rate filter integrates visibilities in time. For a fringe-rate filter, $f_{\rm fr}$, the effective integration time can be calculated by comparing the variance statistic before and after filtering:

$$t_{\rm int,after} = t_{\rm int,before} \frac{\int \sigma_f^2 df}{\int \sigma_f^2 f_{\rm fr}^2 df}, \tag{3.9}$$

where $t_{\rm int,before}$ is the integration time before filtering, σ_f denotes the noise variance in fringe rate space and the integral is taken over all possible fringe rates for a given baseline and frequency. As discussed in [128], the signal re-weighting associated with this fringe-rate filter can be interpreted as a modification to the shape of the primary beam.

For the fiducial baseline at 151 MHz, the integration time, as given in equation (3.9), associated with an optimal fringe rate filter is 3430 s. The number of statistically independent samples on the sky decreases from 83 to 1 sample per hour. As discussed in section 3.5.3, empirically estimating a covariance matrix with a small number of independent samples can lead to signal loss in the OQE. In order to counteract the

---

[5] The angular area in Equation (3.24) will reflect the new angular area corresponding to the change in beam area.

signal loss, we degrade the optimal fringe-rate filter, as shown in Figure 3-9, to have an effective integration time of 1886 s, increasing the number of independent modes to 2 per hour. The fringe rate filter is now sub-optimal, but is still an improvement on the boxcar weighting as used in P14. As documented in Table 3.1, the correction factor for the associated signal loss of the filter we have chosen is 1.39.

We implement the modified filter on a per baseline basis by weighting the fringe-rate bins on the sky by the RMS of the beam at that same location. In order to obtain a smooth filter in the fringe-rate domain, we fit a Gaussian with a hyperbolic tangent tail to this filter. In addition, we multiply this response with another hyperbolic tangent function that effectively zeros out fringe rates below 0.2 mHz. This removes the slowly varying signals that we model as crosstalk. We convolve the time-domain visibilities with the Fourier transform of the resulting fringe-rate filter, shown in Figure 3-9, to produce an averaged visibility. The effect on the data can be seen in Figure 3-8.

3.4 Instrumental Performance

3.4.1 Instrument Stability

In order to build sensitivity to the 21 cm reionization signal, it is critical that PAPER be able to integrate coherently measurements made with different baselines on different days. Figure 3-10 shows the visibility repeatability between baselines and nights as a function of LST. Specifically, we histogram the real part of the visibilities for all redundant fiducial baselines in a given LST bin for foreground contained data. We see that for a given LST bin, the spread in values over all the baselines is ~50 Jy which corresponds with our observed T_{sys} ~500K. We get more samples per LST bin in the range of 2–10 hr due to our observing season, therefore the density of points in this LST region is greater, as shown by the color scale. This density plot shows that redundant baselines agree very well with one another; OMNICAL has leveled the antenna gains to within the noise.

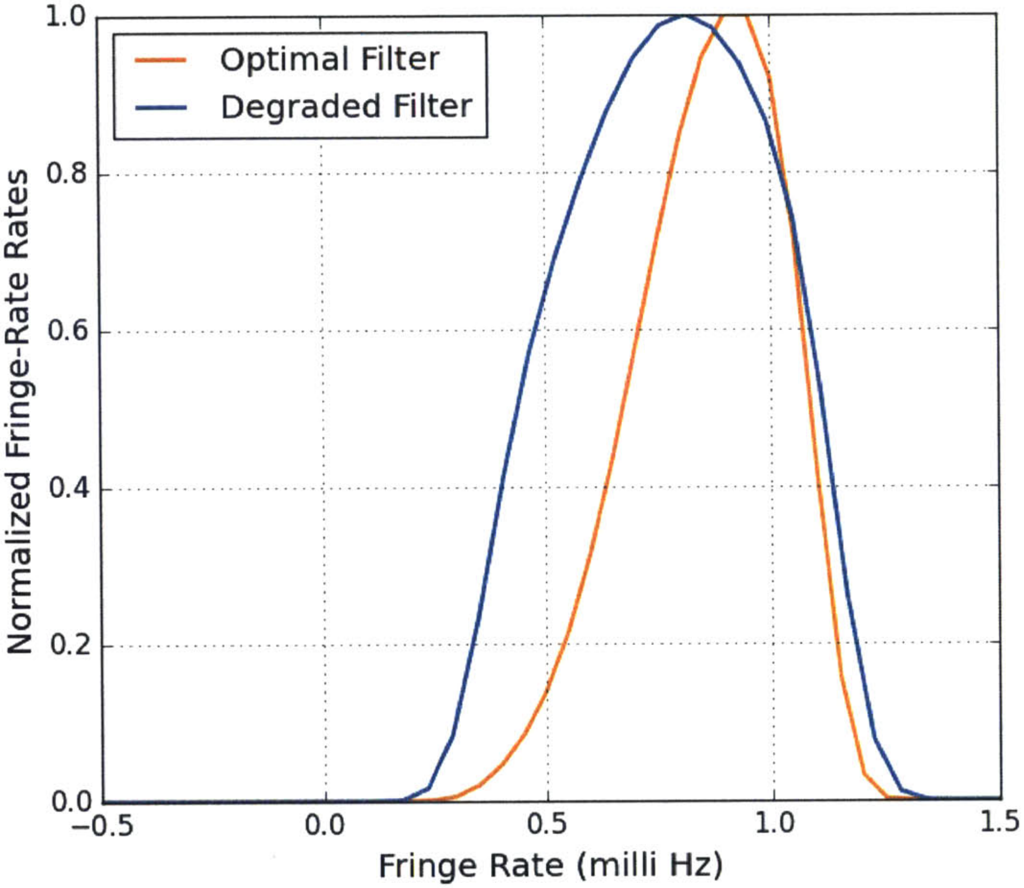

Figure 3-9: The optimal fringe-rate filter (orange) and the degraded fringe-rate filter (blue) actually used in the analysis at 151 MHz, normalized to peak at unity.

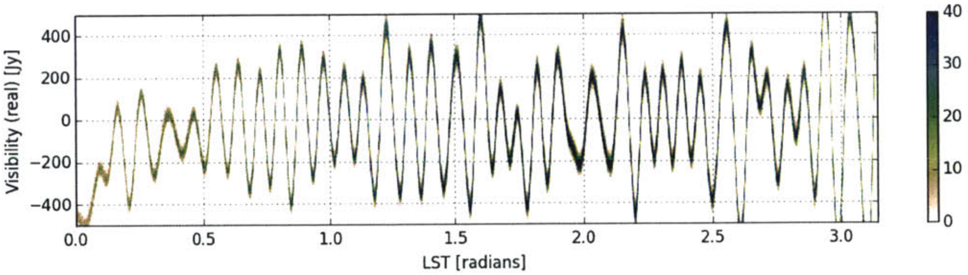

Figure 3-10: Histogram of the real component of all calibrated visibilities measured over 135 days with every redundant instance of the fiducial baseline at 150 MHz. Color scale indicates the number of samples falling in an LST/flux-density bin. This plot serves to illustrate the stability of the PAPER instrument and the precision of calibration. The temporal stability of a single LST bin over multiple days is shown in Figure 3-11.

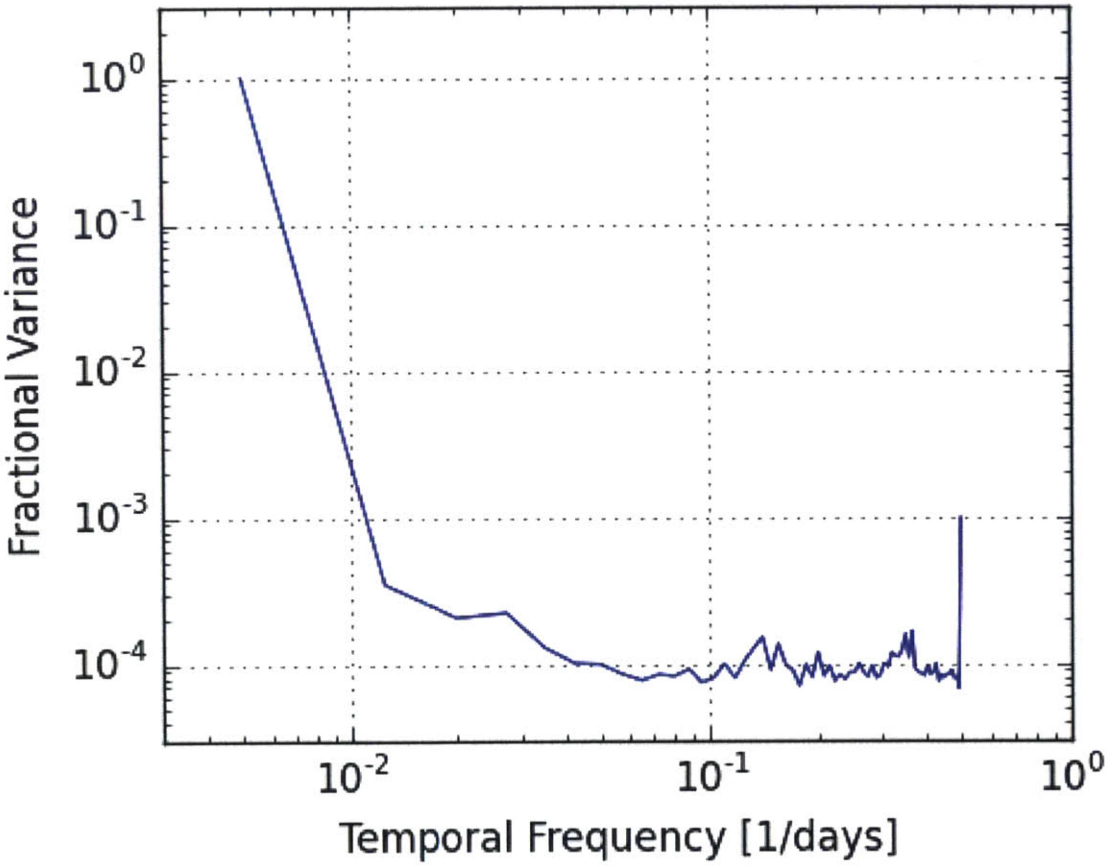

Figure 3-11: Power spectrum of 135 days of time-series data contributing to a single LST bin, illustrating the stability of measurements over the observing campaign. Relative to the average value, variation in the measured value across days (quantified by variance as a function of time period) is orders of magnitude lower. The excess at two-day timescales is a beat frequency associated with the changing alignment of integration windows in the correlator with respect to sidereal time.

Delving in a little deeper, we also examine the stability in time for measurements in a particular LST bin. In order to quantify the stability in time we extract one channel for a given baseline for every observation day and LST bin. We then Fourier transform along the time direction for every LST bin and compute the power spectrum. As shown in Figure 3-11, for time scales greater than one day, we see that signal variance drops by almost four orders of magnitude, with the exception of an excess on two-day timescales caused by the changing alignment of the 42.9 s integration timescale relative to a sidereal day. The implication of this measurement is that, after calibration, PAPER measurements are sufficiently stable to be integrated coherently over the entire length of a 135 day observation. This implies day-to-day stability of better than 1%, contributing negligibly to the uncertainties in the data.

3.4.2 System Temperature

During the LST binning step, the variance of the visibilities that are averaged together for a given frequency and LST bin are recorded. Using these variances, we calculate the system temperature as a function of LST, averaging over each LST hour.

$$T_{\rm rms} = T_{\rm sys}/\sqrt{2\Delta\nu t}, \tag{3.10}$$

where $\Delta\nu$ is the bandwidth, t is the integration time, and $T_{\rm rms}$ is the RMS temperature, or the variance statistic described above. Figure 3-12 shows the results of this calculation. In this observing season, the system temperature drops just below previous estimates as in P14 and [74] of $T_{\rm sys} = 560\,{\rm K}$, at $T_{\rm sys} = 500\,{\rm K}$ at 160 MHz. However, this estimate is more consistent with the results derived in [104], where $T_{\rm sys} = 505\,{\rm K}$ at 164 MHz. The change in the system temperature can be attributed to the reduced range of LST used in the calculation. We note that at 7:00 LST, there is an increase in the system temperature due to the rising of the galactic plane as seen in Figure 3-2.

When calculating the system temperature using the variance in the visibilities for a given LST and frequency, we take into account the fact that we flag 3σ outliers from

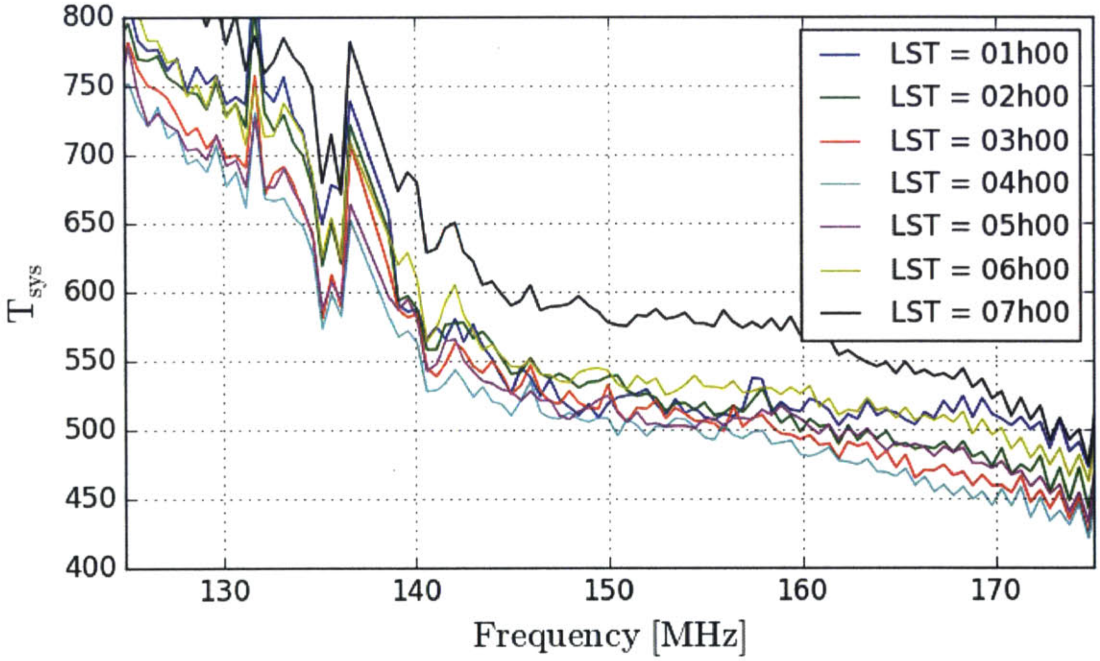

Figure 3-12: System temperature, inferred from the variance of samples falling in an LST bin, averaged over one-hour intervals in LST. The measured value in the 150–160 MHz range is consistent with previous determinations of system temperature ([74]; P14).

the median. To calculate an effective correction factor to account for the filtering, we assume the visibilities follow a Gaussian distribution which would require a correction factor of 1.34 for the removal of data points that are 3σ above the median. In other words, we are accounting for the wings of the Gaussian that would contribute to the variance in the visibility.

Previous estimates of the system temperature (P14; [74]) relied on differencing and averaging baselines, time samples, and/or frequency channels. The relative agreement between these various methods of estimating the system temperature provides a robust measure of the system temperature of the PAPER instrument. Agreement between the instantaneous measurements of the system temperature, the LST repetition variance, and the predicted power spectrum noise level (see below) indicates a robustly stable system with no significant long term instability contributing appreciable noise.

3.5 Power Spectrum Analysis

In this section we first review the OQE formalism, followed by a walk-through of our particular applications of the OQE method to our data. Finally, we discuss the effects of using an empirically estimated covariance matrix in our analysis.

3.5.1 Review of OQEs

We use the OQE method to estimate our power spectrum as done in [83], [37], [89], [90], and [174]. Here we briefly review the OQE formalism with an emphasis on our application to data, which draws strongly from the aforementioned works, but also relies on empirical techniques similar to those used in P14. The end goal of this analysis is to estimate the 21 cm power spectrum, $P_{21}(\mathbf{k})$, defined such that

$$\langle \widetilde{T}_b(\mathbf{k})\widetilde{T}_b^*(\mathbf{k}')\rangle = (2\pi)^3\delta^D(\mathbf{k}-\mathbf{k}')P_{21}(\mathbf{k}), \tag{3.11}$$

where $\widetilde{T}_b(\mathbf{k})$ is the spatial Fourier transform of the brightness temperature distribution on the sky, $\langle\rangle$ denotes an ensemble average, and δ^{D} is the Dirac delta function.

In order to make an estimate of the power spectrum in the OQE formalism, one begins with a data vector $\mathbf{x}$. This vector could, for example, consist of a list of brightness temperatures on the sky for an imaging-based data analysis, or (in our case) a list of measured visibilities. We form the intermediate quantity,

$$\hat{q}_\alpha = \frac{1}{2}\mathbf{x}^\dagger \mathbf{C}^{-1}\mathbf{Q}_\alpha \mathbf{C}^{-1}\mathbf{x} - b_\alpha, \tag{3.12}$$

which will be needed to form the OQE of our power spectrum. Here, $\mathbf{C} \equiv \langle \mathbf{x}\mathbf{x}^\dagger \rangle$ is the true covariance matrix of the data vector $\mathbf{x}$, $\mathbf{Q}_\alpha$ is the operator that takes visibilities into power spectrum k-space and bins into the αth bin, and b_α is the bias to the estimate that needs to be subtracted off. In general, $\mathbf{Q}_\alpha$ represents a family of matrices, one for each k bin indexed by α. Each matrix is defined as $\mathbf{Q}_\alpha \equiv \frac{\partial \mathbf{C}}{\partial p_\alpha}$, i.e., the derivative of the covariance matrix with respect to the band power p_α. The bandpower p_α can be intuitively thought of as the value of the power spectrum in the αth k bin. Therefore, $\mathbf{Q}_\alpha$ encodes the response of the data covariance matrix to the αth bin of the power spectrum.

The bias term b_α in Equation (3.12) will include contributions from both instrumental noise and residual foregrounds. Their presence in the data is simply due to the fact that both contributions have positive *power*. One approach to dealing with these biases is to model them and to subtract them off, as is suggested by Equation (3.12). An alternate approach is to compute a cross-power spectrum between two data sets that are known to have the same sky signal but independent instrumental noise realizations. Labeling these two data sets as $\mathbf{x}_1$ and $\mathbf{x}_2$ and computing

$$\hat{q}_\alpha = \frac{1}{2}\mathbf{x}_1^\dagger \mathbf{C}^{-1}\mathbf{Q}_\alpha \mathbf{C}^{-1}\mathbf{x}_2, \tag{3.13}$$

one arrives at a cross-power spectrum that by construction has no noise bias. There is thus no need to explicitly model and subtract any noise bias, although any residual foreground bias will remain, since it is a contribution that is sourced by signals on the sky, and therefore must exist in all our data sets.

The set of $\hat{q}_\alpha$s do not yet constitute a properly normalized estimate of the power spectrum (as evidenced, for example, by the extra factors of $\mathbf{C}^{-1}$). To normalize our results, we group the unnormalized bandpowers into a vector $\hat{\mathbf{q}}$ and apply a matrix $\mathbf{M}$ (whose exact form we specify later), so that

$$\hat{\mathbf{p}} = \mathbf{M}\hat{\mathbf{q}} \tag{3.14}$$

is a normalized estimate $\hat{\mathbf{p}}$ of the true power spectrum $\mathbf{p}$. We emphasize that the vector space that contains $\hat{\mathbf{q}}$ and $\hat{\mathbf{p}}$ is an "output" vector space over different k-bins, which is separate from the "input" vector space of the measurements, in which $\mathbf{x}$ and $\mathbf{C}$ reside.

To select an $\mathbf{M}$ matrix that properly normalizes the power spectrum, we must compute the window function matrix $\mathbf{W}$ for our estimator. The window matrix is defined such that the true bandpowers $\mathbf{p}$ and our estimates $\hat{\mathbf{p}}$ of them are related by

$$\hat{\mathbf{p}} = \mathbf{W}\mathbf{p}, \tag{3.15}$$

so that each row gives the linear combination of the true power that is probed by our estimate. With a little algebra, one can show that

$$\mathbf{W} = \mathbf{M}\mathbf{F}, \tag{3.16}$$

where

$$\mathbf{F}_{\alpha\beta} = \frac{1}{2}\mathrm{tr}(\mathbf{C}^{-1}\mathbf{Q}_\alpha\mathbf{C}^{-1}\mathbf{Q}_\beta), \tag{3.17}$$

which we have suggestively denoted with the symbol $\mathbf{F}$ to highlight the fact that this turns out to be the Fisher information matrix of the bandpowers. In order to interpret each bandpower as the weighted average of the true bandpowers, we require each row of the window function matrix to sum to unity. As long as $\mathbf{M}$ is chosen in such a way that $\mathbf{W}$ satisfies this criterion, the resulting bandpower estimates $\hat{\mathbf{p}}$ will be properly normalized.

Beyond the normalization criterion, a data analyst has some freedom over the precise form of $\mathbf{M}$, which effectively also re-bins the bandpower estimates. One popular choice is $\mathbf{M} = \mathbf{F}^{-1}$, which implies that $\mathbf{W} = \mathbf{I}$. Each window function is then a delta function, such that bandpowers do not contain leakage from other bins, and contain power from only that bin. However, the disadvantage of this becomes apparent if one also computes the error bars on the bandpower estimates. The error bars are obtained by taking the square root of the diagonal of the covariance matrix, which is defined as

$$\mathbf{\Sigma} = \mathrm{Cov}(\hat{\mathbf{p}}) = \langle \hat{\mathbf{p}}\hat{\mathbf{p}}^\dagger \rangle - \langle \hat{\mathbf{p}} \rangle \langle \hat{\mathbf{p}} \rangle^\dagger. \tag{3.18}$$

Since $\hat{\mathbf{p}} = \mathbf{M}\hat{\mathbf{q}}$, it is easily shown that

$$\mathbf{\Sigma} = \mathbf{M}\mathbf{F}\mathbf{M}^\dagger. \tag{3.19}$$

The choice of $\mathbf{M} = \mathbf{F}^{-1}$ tends to give rather large error bars. At the other extreme, picking $\mathbf{M}_{\alpha\beta} \propto \delta_{\alpha\beta}/\mathbf{F}_{\alpha\alpha}$ (with the proportionality constant fixed by our normalization criterion) leads to the smallest possible error bars [164], at the expense of broader window functions. In our application of OQEs in the following sections, we will pick an intermediate choice for $\mathbf{M}$, one that is carefully tailored to avoid the leakage of foreground power from low k modes to high k modes.

3.5.2 Application of OQE

Here we describe the specifics of our application of the OQE formalism to measure the power spectrum. Doing so requires defining various quantities such as $\mathbf{x}$, $\mathbf{C}$, $\mathbf{Q}_\alpha$ for our analysis pipeline.

First, we consider $\mathbf{x}$, which represents the data in our experiment. Our data set consists of visibilities as a function of frequency and time for each baseline in the array. In our analysis, we group the baselines into three groups of redundant baselines (described in Section 3.2), in the sense that within each group there are multiple copies of the same baseline. In the description that follows, we first estimate the power spectrum separately for each group. Power spectrum estimates obtained from the different redundant groups are then combined in a set of averaging and bootstrapping steps described in Section 3.5.4. Note that because our data have been fringe-rate filtered in the manner described in Section 3.3.5, we may reap all the benefits of coherently integrating in time simply by estimating the power spectrum for every instant in the LST-binned data before averaging over the time-steps within the LST-binned day [128].

For the next portion of our discussion, consider only the data within a single redundant group. Within each group there are not only multiple identical copies of the same baseline, but in addition (as discussed in Section 3.3.3), our pipeline also constructs two LST-binned data sets, one from binning all even-numbered days in our observations, and the other from all odd-numbered days. Thus, we have not a single data vector, but a whole family of them, indexed by baseline (i) and odd versus even days (r). Separating the data out into independent subgroups allows one to estimate cross-power spectra rather than auto-power spectra in order to avoid the noise bias, as discussed in the previous section. The data vectors take the form

$$\mathbf{x}^{ri}(t) = \begin{pmatrix} V^{ri}(\nu_1, t) \\ V^{ri}(\nu_2, t) \\ \vdots \end{pmatrix}, \qquad (3.20)$$

where $V^{ri}(\nu, t)$ is the visibility at frequency ν at time t. Each data vector is 20 elements long, being comprised of 20 channels of a visibility spectrum spanning 10 MHz of bandwidth centered on 151.5 MHz.

Having formed the data vectors, the next step in Equation (3.12) is to weight the data by their inverse covariance. To do so, we of course require the covariance matrix $\mathbf{C}$, which by definition, is the ensemble average of $\mathbf{x}\mathbf{x}^\dagger$, namely $\mathbf{C} = \langle \mathbf{x}\mathbf{x}^\dagger \rangle$. Unfortunately, in our case the covariance is difficult to model from first principles, and we must resort to an empirically estimated $\mathbf{C}$. We make this estimation by taking the time average of the quantity $\mathbf{x}\mathbf{x}^\dagger$ over 8.5 hr of LST, estimating a different covariance matrix for each baseline and for odd versus even days. While an empirical determination of the covariance is advantageous in that it captures features that are difficult to model from first principles, it carries the risk of cosmological signal loss [162]. We will discuss and quantify this signal loss in Section 3.5.3.

To gain some intuition for the action of $\mathbf{C}^{-1}$ on our data, let us examine the combination

$$\mathbf{z}^{ri} = (\mathbf{C}^{ri})^{-1}\mathbf{x}^{ri} \tag{3.21}$$

for select baselines. This is a crucial step in the analysis since it suppresses coherent frequency structures (such as those that might arise from residual foregrounds). Note that the inverse covariance weighting employed here differs from that in P14, in that P14 modeled and included covariances between different baselines, whereas in our current treatment we only consider covariances between different frequency channels. Figure 3-13 compares the effects of applying the inverse covariance matrix to a data vector that contains foregrounds (and thus contains highly correlated frequency structures) to one in which foregrounds have been suppressed by the wideband delay filter described in Section 3.3.3. In the figure, the top row corresponds to the data vector $\mathbf{x}^{ri}$ for three selected baselines in the form of a waterfall plot of visibilities, with frequency on the horizontal axis and time on the vertical axis. The middle section shows the empirical estimate of the covariance by taking the outer product of $\mathbf{x}$ with itself and averaging over the time axis. Finally, the last row shows the results of inverse covariance weighting the data, namely $\mathbf{z}^{ri}$. In every row, the foreground-dominated data are shown in the left half of the figure, while the foreground-suppressed data are shown in the right half.

Consider the foreground-dominated $\mathbf{x}^{ri}$ in Figure 3-13, and their corresponding covariance matrices. The strongest modes that are present in the data are the eigenmodes of the covariance matrix with the largest eigenvalues. Figure 3-14 shows the full eigenvalue spectrum and the four strongest eigenmodes. For the foreground-dominated data, one sees that the eigenvalue spectrum is dominated by the first few modes, and the corresponding eigenmodes are rather smooth, highly suggestive

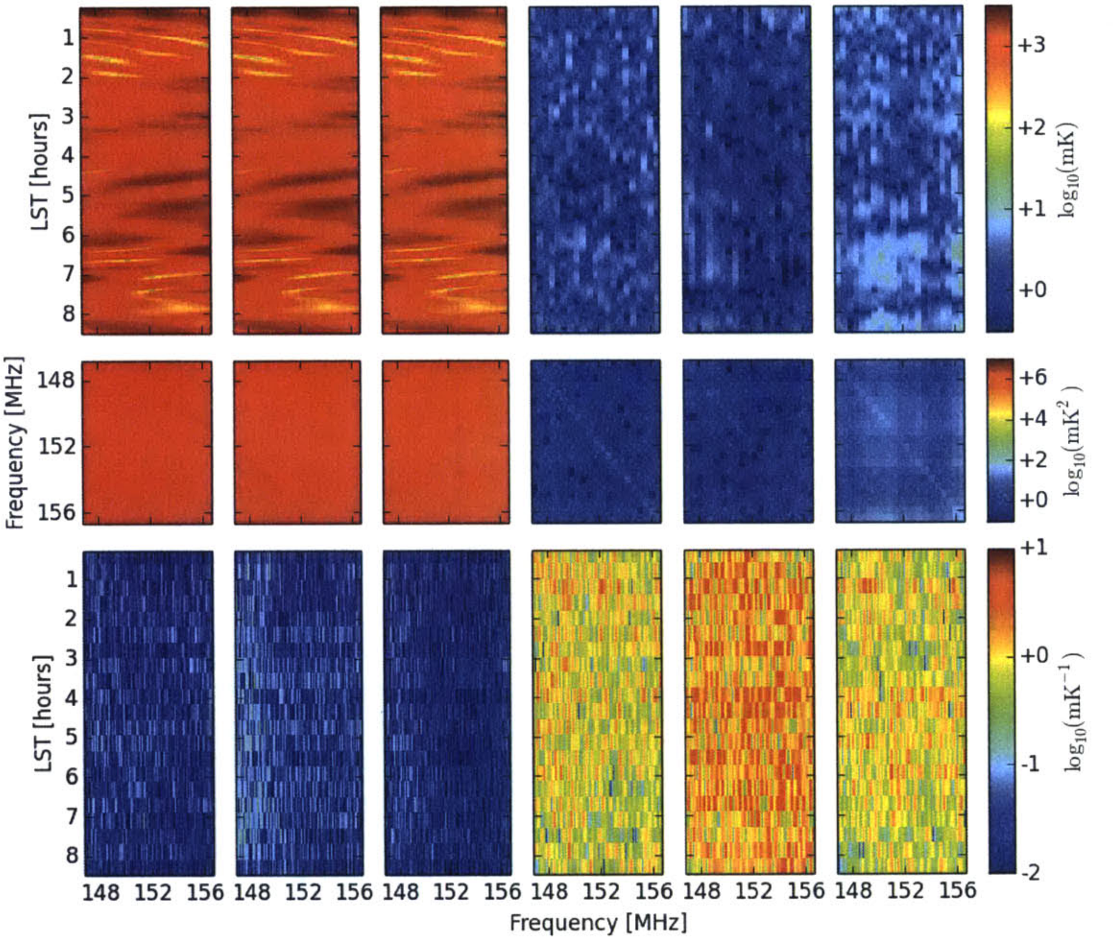

Figure 3-13: Visibilities before (top row) and after (bottom row) inverse covariance weighting. Signal covariance (middle row) is estimated empirically, averaging over LST. The three left/right columns show visibilities from three different baselines in a redundant group before/after delay filtering, respectively.

of smooth spectrum foreground sources. The application of the inverse covariance weighting down-weights these eigenmodes, revealing waterfall plots in the bottom row of Figure 3-13 that look more noise-dominated. With the foreground-suppressed portion (right half) of Figure 3-13, the initial $\mathbf{x}^{ri}$ vectors already appear noise dominated (which is corroborated by the relatively noisy form of the eigenvalue spectra in Figure 3-14). The final $\mathbf{z}^{ri}$ vectors remain noise-like, although some smooth structure (perhaps from residual foregrounds) has still been removed, and finer scale noise has been up-weighted.

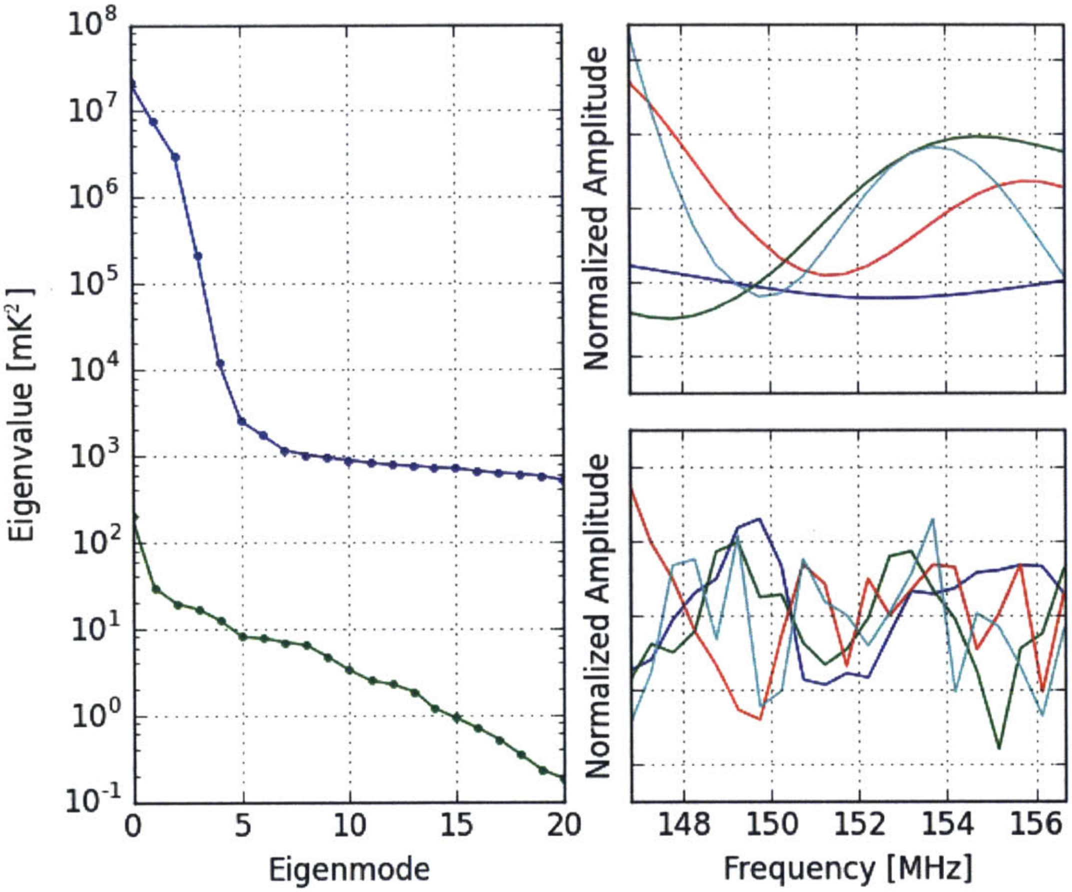

Figure 3-14: Eigenvalue spectrum of covariance matrices (left) empirically estimated from visibilities before (blue) and after (green) delay filtering. The four strongest eigenmodes of the filtered/unfiltered data are plotted on the top/bottom panels on the right, respectively.

With intuition established for the behavior of $\mathbf{C}^{-1}$, we may group our identical baselines into five different sets and average together $\mathbf{z}^{ri}$ vectors for baselines within the same set. That is, we form

$$\mathbf{z}_A^r = \sum_{i \in A} (\mathbf{C}^{ri})^{-1} \mathbf{x}^{ri}, \tag{3.22}$$

where A ranges from 1 to 5 and indexes the baseline set. At this point, we have 10 weighted data vectors $\mathbf{z}$ (5 baseline sets, each of which has an even day and odd day version) for every LST-binned time-step. As discussed in the previous section, instrumental noise bias may be avoided by forming cross-power spectra rather than auto-power spectra. Generalizing Equation (3.13) to our present case where we have 10 different data vectors, we have

$$\hat{\mathbf{q}}_\alpha = \sum_{\substack{A,B,r,s \\ r \neq s, A \neq B}} \mathbf{z}_A^{r\dagger} \mathbf{Q}_\alpha \mathbf{z}_B^s, \tag{3.23}$$

so that auto-power contributions from identical baseline groups or identical even/odd indices never appear. Residual foreground bias will remain in Equation (3.23), but in order to avoid possible signal loss from an overly aggressive foreground bias removal scheme, we conservatively allow the foreground bias to remain. Since foreground power will necessarily be positive, residual foregrounds will only serve to raise our final upper limits.

In order to implement Equation (3.23), it is necessary to derive a form for $\mathbf{Q}_\alpha \equiv \partial \mathbf{C} / \partial p_\alpha$. To do so, we follow the delay spectrum technique of P12a, where it was shown that

$$P(\mathbf{k}_{t\tau}) \approx \left(\frac{\lambda^2}{2k_B} \right)^2 \frac{X^2 Y}{\Omega B} \langle \tilde{V}_i(t,\tau) \tilde{V}_j^*(t,\tau) \rangle, \tag{3.24}$$

where $V_i(t,\tau)$ is the delay transform of baseline visibilities given by Equation (3.8), X and Y are the constants that convert from angles and frequency to the co-moving coordinate, respectively, Ω is the power squared beam (see Appendix B of P14), B is the bandwidth, λ is the spectral wavelength, and k_B is Boltzmann's constant. This suggests that in order to estimate the power spectrum from visibilities, one only needs to Fourier transform along the frequency axis (converting the spectrum into a delay spectrum) before squaring and multiplying by a scalar. Thus, the role of $\mathbf{Q}_\alpha$ in Equation (3.23) is to perform a frequency Fourier transform on each copy of $\mathbf{z}$. It is therefore a separable matrix of the form $\mathbf{Q}_\alpha = \mathbf{m}_\alpha \mathbf{m}_\alpha^\dagger$, where $\mathbf{m}_\alpha$ is a complex sinusoid of a specific frequency corresponding to delay mode α. We may thus write

$$\hat{\mathbf{q}}_\alpha = \sum_{\substack{A,B,r,s \\ r\neq s, A\neq B}} \mathbf{z}_A^{r\dagger} \mathbf{m}_\alpha \mathbf{m}_\alpha^\dagger \mathbf{z}_B^s. \tag{3.25}$$

With an explicit form for $\mathbf{Q}_\alpha$, one now also has the necessary ingredients to compute the Fisher matrix using Equation (3.17).

Having computed the $\hat{\mathbf{q}}_\alpha$s, we group our results into a vector $\hat{\mathbf{q}}$. This vector of unnormalized bandpowers is then normalized to form our final estimates of the power spectrum $\mathbf{p}$. As noted above, the normalization occurs by the $\mathbf{M}$ matrix in Equation (3.14), and can be any matrix of our desire. Even though the choices of the normalization matrices described above have certain good properties, e.g. small error bars or no leakage, we opt for a different choice of window function, as follows. We first reorder the elements in $\hat{\mathbf{q}}$ (and therefore in $\mathbf{F}$, $\mathbf{M}$, and $\hat{\mathbf{p}}$ for consistency) so that the k-modes are listed in ascending order, from low k to high k, with the exception that we place the highest k bin third after the lowest two k bins. (The reason for this exception will be made apparent shortly). We then take the Cholesky decomposition of the Fisher matrix, such that $\mathbf{F} = \mathbf{L}\mathbf{L}^\dagger$, where $\mathbf{L}$ is a lower triangular matrix. Following that, we pick $\mathbf{M} = \mathbf{D}\mathbf{L}^{-1}$, where $\mathbf{D}$ is a diagonal matrix chosen to adhere to the normalization constraint that $\mathbf{W} = \mathbf{M}\mathbf{F}$ has rows that sum to unity. In this case, the window function matrix becomes, $\mathbf{W} = \mathbf{D}\mathbf{L}^\dagger$. This means that $\mathbf{W}$ is upper triangular, and with our ordering scheme, has the consequence of allowing power to leak from high to low k, but not vice versa. Since our k axis is (to a good approximation) proportional to the delay axis, foregrounds preferentially appear at low k because their spectra are smooth. Reducing leakage from low k to high k thus mitigates leakage of foregrounds into the cleaner, more noise-dominated regions.

Additionally, our placement of the highest k bin as the third element in our reordering of $\hat{\mathbf{p}}$ prevents leakage from this edge bin that will contain aliased power. Figure 3-15 shows the resulting window functions.

Our choice of normalization matrix also has the attractive property of eliminating error correlations between bandpower estimates. Using Equation (3.19), we have that

$$\mathbf{\Sigma} = \mathbf{D}\mathbf{L}^{-1}\mathbf{L}\mathbf{L}^{\dagger}\mathbf{L}^{-\dagger}\mathbf{D} = \mathbf{D}^2. \tag{3.26}$$

The error covariance matrix on the bandpowers is thus diagonal, which implies that our final data points are uncorrelated with one another. This stands in contrast to the power-spectrum estimator used in P14, where the Blackmann–Harris taper function induced correlated errors between neighboring data points.

3.5.3 Covariance Matrix and Signal Loss

We now discuss some of the subtleties associated with empirically estimating the covariance matrix from the data. Again, the covariance matrix is defined as the ensemble average of the outer product of a vector with itself, i.e.,

$$\mathbf{C} = \langle \mathbf{x}\mathbf{x}^{\dagger} \rangle, \tag{3.27}$$

where $\mathbf{x}$ is the data (column) vector used in the analysis. In our analysis, we do not have *a priori* knowledge of the covariance matrix. and thus we must resort to empirical estimates [39]. As we have alluded to above, we replace the ensemble average with a time average that runs from 0 to 8:30 LST hours.

Since the OQE method for power spectrum estimation requires the inversion of $\mathbf{C}$, it is crucial that our empirically estimated covariance be a full rank matrix. With our data consisting of visibilities over 20 frequency channels, the covariance matrix is a 20×20 matrix. Thus, a necessary condition for our estimate to be full rank is for there to be at least 20 independent time samples in our average. As noted in Section 3.3.5 the fringe-rate filter used corresponds to averaging time samples for 31

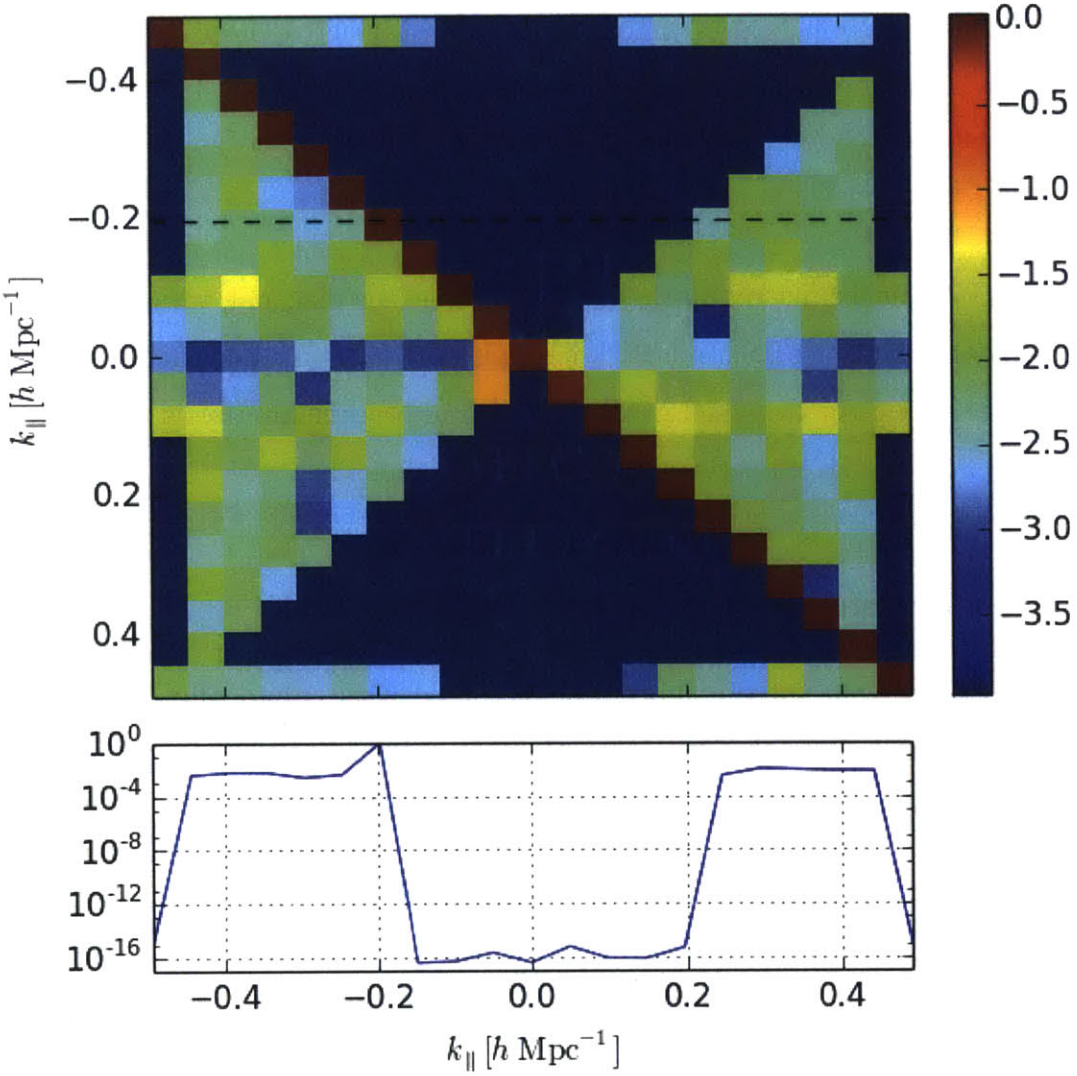

Figure 3-15: The window function matrix $\mathbf{W}$, as defined in Equation (3.15). The i^{th} row corresponds to the window function used in the estimate of the power spectrum for the i^{th} k-mode. Color scale indicates $\log_{10}\mathbf{W}$. The inset plot illustrates the window function along the dashed line in the upper panel. As described in Section 3.5.2, $\mathbf{M}$ in Equation (3.16) has been chosen so that each window function peaks at the waveband while achieving a high degree of isolation from at lower k-modes that are likely to be biased by foregrounds.

minutes. Over the LST range used in this analysis, this corresponds to roughly 20 statistically independent modes in our data after fringe-rate filtering. We therefore have just enough samples for our empirical estimate, and in practice, our covariance matrices are invertible and allow OQE techniques to be implemented.

Another potential problem that occurs from empirically estimating covariances is that it leads to models of the covariance matrix that over-fit the noise. In this scenario, the covariance matrix tells us that there may be modes in the data that should be down-weighted, for example, but if the empirical covariance estimates are dominated by noise, these may just be random fluctuations that need not be down-weighted. Said differently, the weighting of the data by the inverse covariance is heavily influenced by the noise in the estimate of the covariance matrix and thus has the ability to down-weight valid high-variance samples. Over-fitting the noise in this manner carries with it the possibility of cosmological signal loss. This seems to contradict the conventionally recognized feature of OQEs as lossless estimators of the power spectrum [164]. However, the standard proofs of this property assume that statistics such as $\mathbf{C}$ are known *a priori*, which is an assumption that we are violating with our empirical estimates.

In order to deal with possible signal loss, we perform simulations of our analysis pipeline, deriving correction factors that must be applied to our final constraints. We simulate visibilities for Gaussian temperature field with a flat amplitude in $P(k)$ that rotates with the sky, which is fringe-rate filtered in the same way as the data for our fiducial baselines. This signal is processed through our pipeline, and the output power spectrum compared to the input power spectrum, for various levels of input signal amplitude. We repeat this for 40 sky realizations at each signal level. Figure 3-16 shows the resultant signal loss associated with estimating the covariance matrix from the data. Error bars were obtained through bootstrapping.

As a function of the increasing input amplitude of the simulated power spectra, we find that the ratio of output power to input power decreases, which we interpret as signal loss through the use of our empirical OQE of the power spectrum. However, since the transfer function through this analysis is an invertible function, we can cor-

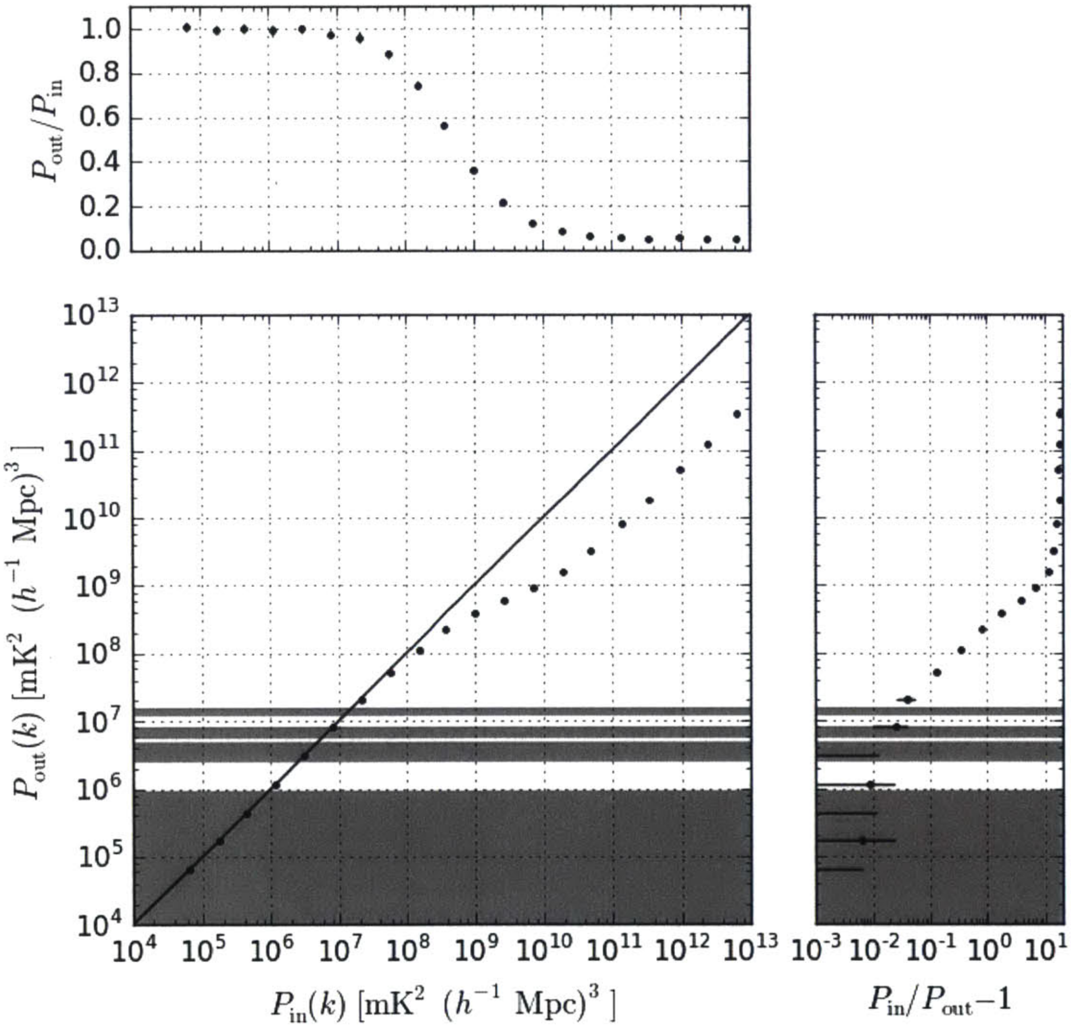

Figure 3-16: Recovered power spectrum signal as a function of injected signal amplitude. Shaded regions indicate the range in measured amplitude of power spectrum modes in Figure 3-18. Error bars indicate 95% confidence intervals as determined from the Monte Carlo simulations described in Section 3.5.3. Because the recovered signal amplitude is a monotonic function of the injected signal amplitude, it is possible to invert the effects of signal loss in the measured power spectrum values to infer the true signal amplitude on the sky. Over the range of powers measured, the maximum correction factor P_{in}/P_{out} is less than 1.02 at 97.5% confidence. The transition to significantly higher correction factors at larger signal amplitudes occurs as the injected signal dominates over the foreground modes present in estimates of the data covariance.

rect for the transfer by using the output value to infer a signal loss that is then divided out to obtain the original input signal level. In Figure 3-16, we see that deviations from unity signal transfer begin at power spectrum amplitudes of $10^7 \mathrm{mK}^2(h^{-1}\,\mathrm{Mpc})^3$. For the range of output power spectrum amplitudes in our final estimate of the 21 cm power spectrum (Figure 3-18), we show that signal loss is $< 2\%$ at 95% confidence.

Table 3.1: SIGNAL LOSS VERSUS ANALYSIS STAGE

Analysis Stage	Typical Loss	Maximum Loss
Bandpass Calibration	$< 2 \times 10^{-7}\%$	3.0%
Delay Filtering	$1.5 \times 10^{-3}\%$	4.8%
Fringe-rate Filtering	28.1%	28.1%
Quadratic Estimator	$< 2.0\%$	89.0%
Median of Modes	30.7%	30.7%

As shown in Table 3.1, the signal loss we characterize for quadratic estimation of the power spectrum band powers is tabulated along with the signal loss associated with each other potentially lossy analysis stage (see Figure 3-3). We correct for the signal loss in each stage by multiplying the final power spectrum results by the typical loss for each stage, except for modes within the horizon limit and immediately adjacent to the horizon limit, where we apply the maximum signal loss correction to be conservative.

3.5.4 Bootstrapped Averaging and Errors

When estimating our power spectra via OQEs, we generate multiple samples of the power spectrum in order to apply the bootstrap method to calculate our error bars. In detail, the power spectrum estimation scheme proposed above requires averaging at several points in the pipeline:

1. Visibilities are averaged into five baseline groups after inverse covariance weighting (see Equation (3.22))

2. Power spectrum estimates from each of the three redundant baseline types (described in Section 3.2) are averaged together.

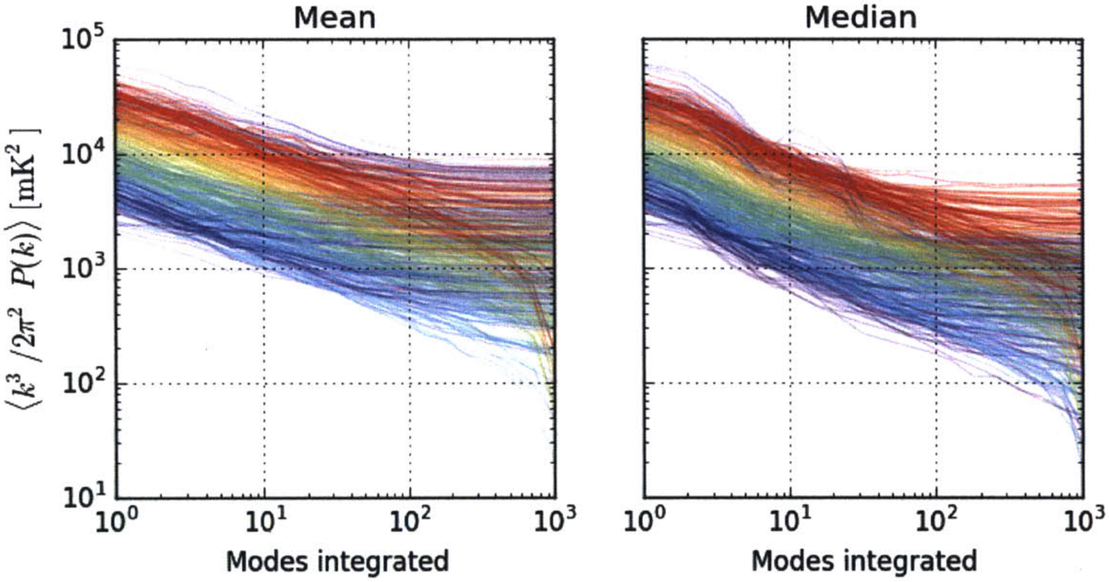

Figure 3-17: Absolute value of the cumulative mean (left) and median (right), as a function of number of modes of the power spectrum band power for $k_{\parallel}$ modes ranging from -0.49 (red) to $0.44h$ Mpc^{-1} (violet). Here, modes are defined as samples from different redundant baseline groups and LSTs. This Allen variance plot shows modes averaging down as the square root of number of modes combined until a signal floor is reached. The difference in behavior between the mean and median is an indication of outliers in the distribution of values, likely as a result of foreground contamination. We use the median in the estimation of the power spectrum in Figure 3-18, along with a correction factor compensating for the difference between the mean and median in estimating variance.

3. Power spectrum estimates from each LST are averaged together.

With the bootstrapping technique, we do not directly perform these averages. Instead, one draws random samples within the three-dimensional parameter space specified above, with replacement, until one has as many random samples as there are total number of parameter space points. These random samples are then propagated through the power spectrum pipeline and averaged together as though they were the original data. This forms a single estimate (a "bootstrap") of $P(\mathbf{k})$. Repeating random draws allows one to quantify the inherent scatter—and hence the error bars—in our estimate of $P(\mathbf{k})$. When plotting $\Delta^2(k) \equiv k^3 P(k)/2\pi^2$ instead of $P(\mathbf{k})$, we bin power falling in $+k$ and $-k$, and so we additionally randomize the inclusion of positive and negative k bins.

We compute a total of 400 bootstraps. In combining independent samples for our final power spectrum estimate, we elect to use the median, rather than the mean, of the samples. One can see the behavior of both statistics in Figure 3-17, where we show how the absolute value of $\Delta^2(k)$ integrates down as more independent samples are included in the mean and median. In this plot, one can see modes integrating down consistent with a noise-dominated power spectrum until they bottom out on a signal. In the noise-dominated regime, the mean and the median behave similarly. However, we see that the median routinely continues to integrate down as noise for longer. This is an indication that the mean is skewed by outlier modes, suggesting variations beyond thermal noise. The magnitude of the difference is also not consistent with the Rayleigh distribution expected of a cosmological power spectrum limited by cosmic variance. For a Rayleigh distribution, the median is $\ln 2 \sim 0.69$ times the mean. Instead, we interpret the discrepancy as a sign of contributions from foregrounds, which are neither isotropic nor Gaussian distributed. Since median provides better rejection of outliers in the distribution that might arise from residual foreground power, we choose to use the median statistic to combine measurements across multiple modes. As listed in Table 3.1, we apply a $1/\ln 2$ correction factor to our power spectrum estimates to infer the mean from the median of a Rayleigh distribution.

3.6 Results

3.6.1 Power Spectrum Constraints

To summarize the previous section, we follow the power spectrum analysis procedure outlined in Section 3.5.2, we incoherently combine independent power spectrum measurements made at different times and with different baseline groups using the median statistic. As described in Section 3.5.4, we bootstrap over all of these independent measurements, as well as over the selection of baselines included in the power spectrum analysis for each baseline group, in order to estimate the error bars on the spherically averaged power spectrum $P(k)$, where positive and negative $k_{\|}$ measurements are kept separate for diagnostic purposes. In the estimation of the dimensionless power spectrum $\Delta^2(k) \equiv k^3 P(k)/2\pi^2$, the folding of $\pm k_{\|}$ is handled along with the rest of the bootstrapping over independent modes. Finally, the measured values for $P(k)$ and $\Delta^2(k)$ are corrected for signal loss through all stages of analysis, as summarized in Table 3.1.

The final results are plotted in Figure 3-18. For the first two modes outside of the horizon where $\Delta^2(k)$ is measured, we have clear detections. We attribute these to foreground leakage from inside the horizon related to the convolution kernels in Equation (3.8) (either from the chromaticity of the antenna response, or from the inherent spectrum of the foregrounds themselves). Somewhat more difficult to interpret are the 2.4σ excess at $k \approx 0.30h\ \mathrm{Mpc}^{-1}$ and the 2.9σ excess at $k \approx 0.44h\ \mathrm{Mpc}^{-1}$. Having two such outliers is unlikely to be chance.

In examining the effects on the power spectrum of omitting various stages of analysis (see Figure 3-19), we see a pronounced excess in the green curve corresponding to the omission of crosstalk removal in fringe-rate filtering. While the signal is heavily attenuated in the filtering step, it remains a possibility that the remaining detections are associated with instrumental crosstalk. We do note, however, that the qualitative shape of the excess in the crosstalk-removed data does not appear to match that of the crosstalk-containing data.

Another likely possibility is that the signal might be associated with foregrounds.

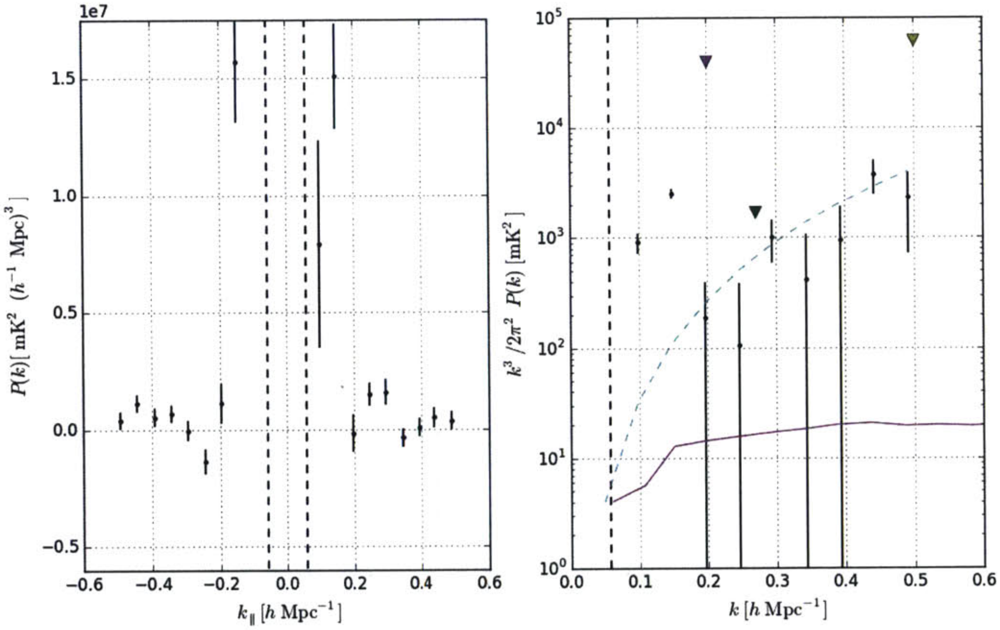

Figure 3-18: Measured power spectrum (black dots with 2σ error bars) at $z = 8.4$ resulting from a 135 day observation with PAPER-64. The dashed vertical lines at $0.6h$ Mpc^{-1} show the bounds of the delay filter described in Section 3.3.3. The predicted 2σ upper limit in the absence of the a celestial signal is shown in dashed cyan, assuming $T_{\text{sys}} = 500K$. The triangles indicate 2 σ upper limits from GMRT [115] (yellow) at $z = 8.6$, MWA [38] at $z = 9.5$ (magenta), and the previous PAPER upper limit (P14) at $z = 7.7$ (green). The magenta curve shows a predicted model 21 cm power spectrum at 50% ionization [81].

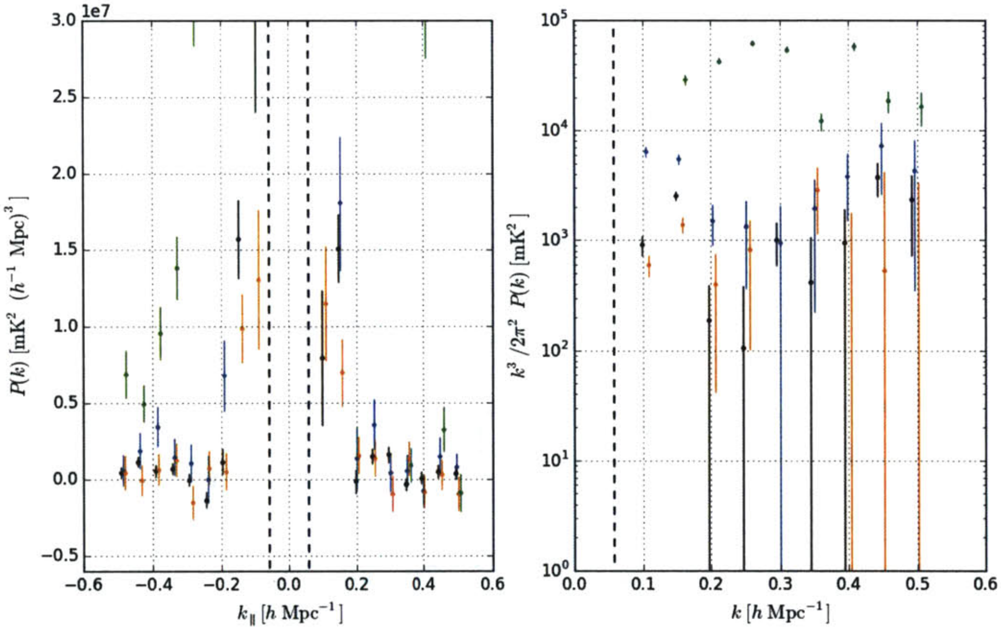

Figure 3-19: Diagnostic power spectra in the style of Figure 3-18 illustrating the impact of various analysis stages. The blue power spectrum uses the P14 fringe-rate filter combined with crosstalk removal. Green illustrates the result using the improved fringe-rate filter, but without crosstalk removal. A power spectrum derived without the application of OMNICAL is shown in orange. Black includes improved fringe-rate filtering, crosstalk removal, and OMNICAL calibration; it is the same power spectrum shown in Figure 3-18.

Foregrounds, which are not generally isotropically distributed on the sky, are likely to be affected by the spatial filtering associated with fringe-rate filtering, whereas a statistically isotropic signal is not. Indeed, we see that excesses in many modes measured with using the P14-stype time-domain filtering (blue in Figure 3-19) decrease significantly using the improved fringe-rate filter. As discussed in [128], the normalization applied to Ω_{eff} for fringe-rate filtering correctly compensates for the effect of this filtering on power-spectral measurements of a statistically isotropic Gaussian sky signal. We can surmise from any significant change in amplitude of the excess under fringe-rate filtering that it arises from emission that violates these assumptions. We conclude, therefore, that this excess is unlikely to be cosmic reionization, and is more likely the result of non-Gaussian foregrounds. As discussed earlier, one possible culprit is polarization leakage [76, 77, 105], although further work will be necessary to confirm this. The interpretation of the signal as polarization leakage is, however, rather high to be consistent with recent measurements in Stokes Q presented in [104], where the leakage is constrained to be < 100 mK^2 for all k.

That the excesses at $k \approx 0.30$ and $0.44h$ Mpc^{-1} are relatively unaffected by the filtering could be an indication that they are more isotropically distributed, but more likely, it may mean that the simply arise closer to the center of the primary beam where they are down-weighted less. Both excesses appear to be significantly affected by omitting OMNICAL calibration (orange in Figure 3-19). This could be interpreted as indicating the excess is a modulation induced by frequency structure in the calibration solution. However, OMNICAL is constrained to prohibit structure common to all baselines, so a more likely interpretation is that this faint feature decorrelates without the precision of redundant calibration. To determine the nature of these particular excesses, further work will be necessary.

In order to aggregate the information presented in the power spectrum into a single upper limit, we fit a flat $\Delta^2(k)$ model to measurements in the range $0.15 < k < 0.5h$ Mpc^{-1}. We use a uniform prior of amplitudes between –5000 and 5000 mK^2, and assume measurement errors are Gaussian. Figure 3-20 shows the posterior distribution of the fit. From this distribution, we determine a mean of (18.9 mK)^2

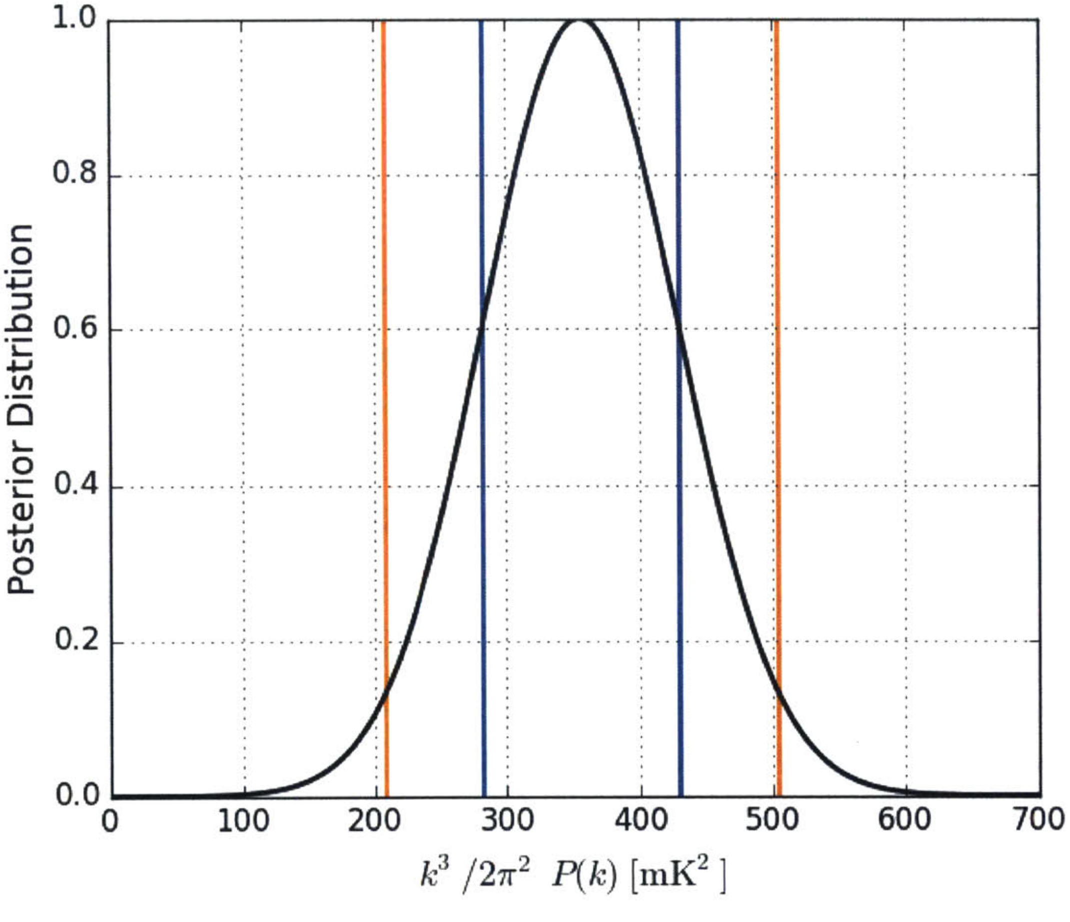

Figure 3-20: Posterior distribution of power spectrum amplitude for a flat $\Delta^2(k)$ power spectrum over $0.15 < k < 0.5h$ Mpc^{-1} (solid black), assuming Gaussian error bars. The blue and orange vertical lines correspond to the 1σ and 2σ bounds, respectively.

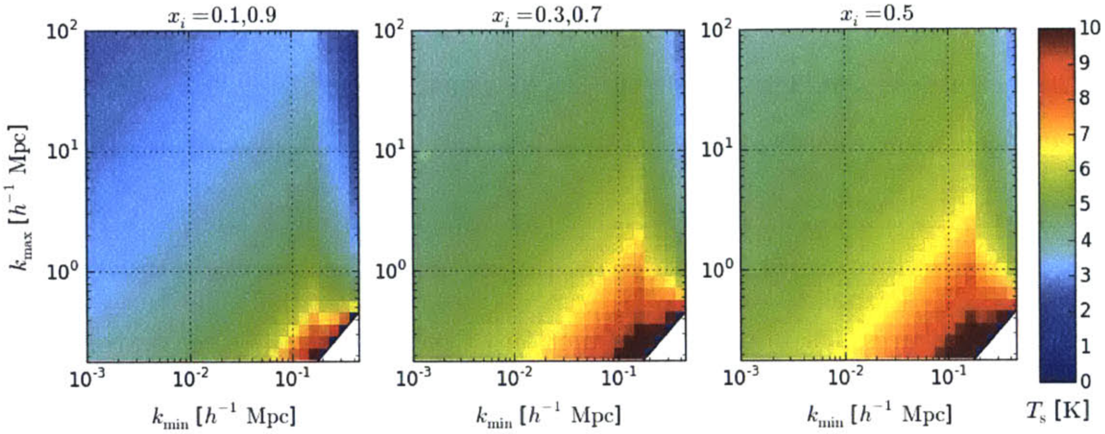

Figure 3-21: Constraints on the 21cm spin temperature at $z = 8.4$, assuming the patchy reionization model in Equations (3.31) and (3.33), which hold in the limit that $T_s < T_{\rm CMB}$.

and a 2σ upper limit of $(22.4\,\mathrm{mK})^2$. The measured mean is inconsistent with zero at the 4.7σ level, indicating that we are detecting a clear power spectrum excess at $k > 0.15h$ Mpc^{-1}.

We suspect that the excess in our measured power spectrum is likely caused by crosstalk and foregrounds. We therefore suggest ignoring the lower bound on the power spectrum amplitude as not being of relevance for the cosmological signal. On the other hand, since foreground power is necessarily positive, the 2σ upper limit of $(22.4\,\mathrm{mK})^2$ at $z = 8.4$, continues to serve as a conservative upper limit. This significantly improves over the previous best upper limit of $(41\ \mathrm{mK})^2$ at $z = 7.7$ reported in P14. As we show below and in greater detail in [140], this limit begins to have implications for the heating of the IGM prior to the completion of reionization.

3.6.2 Spin Temperature Constraints

In this section, we examine the implication of the measured upper limits on 21cm emission in Figure 3-18 on the spin temperature of the 21cm line at $z = 8.4$. In a forthcoming paper [140], we conduct a thorough analysis of the constraints that can be put on the IGM using a simulation-based framework. As a complement to that more thorough analysis, we focus here on a simpler parameterization of the shape of the 21cm power spectrum signal.

The brightness temperature of the 21cm signal, δT_b, arising from the contrast between the cosmic microwave background, T_γ, and the spin temperature, T_s, is given by

$$\delta T_b = \frac{T_\mathrm{s} - T_\gamma}{1+z}(1 - e^{-\tau}) \approx \frac{T_\mathrm{s} - T_\gamma}{1+z}\tau, \tag{3.28}$$

where temperatures are implicitly a function of redshift z, and the approximation holds for low optical depth, τ. The optical depth is given by [193]

$$\tau = \frac{3c^3 \hbar A_{10} n_\mathrm{HI}}{16 k \nu_0^2 T_\mathrm{s} H(z)} \tag{3.29}$$

where A_{10} is the Einstein A coefficient for the 21cm transition, n_HI is the density of the neutral hydrogen, $H(z)$ is the Hubble constant, x_{HI} is the neutral fraction of hydrogen, δ is the local baryon overdensity, ν_0 is the rest frequency of the 21cm transition, and the remainder are the usual constants. Plugging in the cosmological parameters from [134], we get

$$\delta T_b \approx T_0\, x_\mathrm{HI}\, (1+\delta)\, \xi, \tag{3.30}$$

where $\xi \equiv 1 - T_\gamma/T_\mathrm{s}$ and $T_0 \equiv 26.7\,\mathrm{mK}\sqrt{(1+z)/10}$.

If the spin temperature is larger than T_γ, we get the 21 cm signal in emission with respect to the CMB, and $\xi \sim 1$. However, if T_s is less than T_γ, δT_b is negative and ξ can potentially become large.

As in P14, we consider a "weak heating" scenario in which T_s is coupled to the gas temperature via the Wouthuysen-Field effect [187, 49, 67], but little heating has taken place prior to reionization, so that $T_\mathrm{s} < T_\gamma$. In this scenario, because we have assumed little heating, we can approximate ξ as having negligible spatial dependence, and therefore $T_0^2\xi^2$ becomes a simple multiplicative scalar to the 21cm power spectrum:

$$\Delta^2_{21}(k) = T_0^2 \xi^2(z) \Delta^2_i(k), \tag{3.31}$$

where $\Delta_i^2(k)$ is the dimensionless HI power spectrum.

As shown in P14, the maximum value of the prefactor in Equation (3.31) is given by a no-heating scenario where the spin temperature follows the kinetic gas temperature, which is held in equilibrium with the CMB via Compton scattering until $z_{\rm dec} \approx 150$ [54] and then cools adiabatically as $(1+z)^2$. In this case, ξ is given by

$$\xi = 1 - \frac{1+z_{\rm dec}}{1+z} \approx -\frac{150}{1+z}. \tag{3.32}$$

At $z = 8.4$, this corresponds to a minimum bound on the spin temperature of $T_{\rm s} > 1.5$ K.

We can now flip this argument around and, for a measured upper bound on $\Delta_{21}^2(k)$, we can use models for $\Delta_i^2(k)$ in Equation (3.31) to place a bound on $T_{\rm s}$. We consider a class of "patchy" reionization models (P12a;P14) which approximates the ionization power spectrum as flat between minimum and maximum bubble sizes, $k_{\rm min}$ and $k_{\rm max}$, respectively:

$$\Delta_i^2(k) = (x_{\rm HI} - x_{\rm HI}^2)/\ln{(k_{\rm max}/k_{\rm min})}. \tag{3.33}$$

For combinations of $k_{\rm min}$ and $k_{\rm max}$, we determine the minimum spin temperature implied by the 2σ 21 cm power spectrum upper limits shown in Figure 3-18. Figure 3-21 shows the results of these bounds for neutral fractions of $x_{\rm HI} = 0.1$, 0.3, 0.5, 0.7, and 0.9. In almost all cases (excepting $x_{\rm HI} = 0.1, 0.9$ for $k_{\rm min} < 0.1h$ Mpc^{-1}), we find that $T_{\rm s} \gtrsim 3$ K, indicating that our measurements are inconsistent with the spin temperature being coupled to a kinetic temperature governed strictly by adiabatic expansion.

Our results become more interesting in the range of $k_{\rm min} \sim 0.1$ and $k_{\rm max} \sim 30$ representative of fiducial simulations [191, 81]. For neutral fractions of 0.3, 0.5, and 0.7, we find that $T_{\rm s} \gtrsim 4$ K. [140] improves on these results by using a simulation-based framework, rather than relying on coarse parametrizations of the power spectrum shape. They compare the limits they find to the amount of heating possible given the currently observed star formation rates in high-redshift galaxy populations [17, 96] and assumptions about the relationship between star formation rates and X-ray luminosities [54, 143, 47]. Assuming the midpoint of reionization lies close to $z = 8.4$ (a reasonable assumption given that [134] suggests a midpoint of $z = 8.8$), both the bounds found in this paper and [140] show evidence for heating that places constraints on the possible values for the star formation rate/X-ray luminosity correlation given certain models of the star formation rate density redshift evolution. We refer the reader to [140] for a detailed examination of these results.

3.7 Discussion

The improvement in our results over those in P14 are the result of four major advances:

1. the expansion of PAPER to 64 antennas doubled our instrument's power spectrum sensitivity,
2. using OMNICAL for redundant calibration significantly improved the clustering of measurements over the previous implementation of LOGCAL used in P14,
3. fringe-rate filtering further improved power spectrum sensitivity by ~50% and suppressed systematics associated with foregrounds low in the primary beam, and
4. moving from a lossless quadratic estimator targeting difference modes in redundant measurements to an OQE (with carefully calibrated signal loss) significantly reduced contamination from residual foregrounds.

Figure 3-19 illustrates the effect of some of these advances on the final power spectrum. Other important advances include the use of the median statistic to reduce the impact of non-Gaussian outliers in power-spectral measurements, and the use of a Cholesky decomposition of the Fisher information matrix to help reduce leakage from highly contaminated modes within the wedge.

These new techniques and improvements to calibration have reduced the measured bias in nearly all wavebands by an order of magnitude or more. The use of OMNICAL to accurately calibrate the relative complex gains of the antennas has shown to be a major improvement to the data-reduction pipeline. The accuracy and improvement of this calibration brings redundant baselines into impressive agreement with one another (see Figures 3-4 and 3-10), and provides important diagnostic information for monitoring the health of the array, flagging RFI events, and otherwise assessing data quality. Fringe-rate filtering, which is described in greater depth in [128], is also proving to be a flexible and powerful tool for controlling direction-dependent gains and improving sensitivity.

As sensitivity improves, it will be possible to determine more accurately than [104] what the actual level of polarized emission, and thus leakage, may be. Independent fringe-rate filtering of the XX and YY polarizations prior to summation has the potential to better match these polarization beams and further suppress the leakage signal if the polarized signal turns out to be significant.

The end result is a major step forward, both for PAPER and for the field of 21cm cosmology. While we have not yet made a detection of the 21cm cosmological signal, our limits are now within the range of some of the brighter models. As discussed in [140], another order-of-magnitude improvement in sensitivity will make 21cm measurements highly constraining.

3.8 Conclusions

We present new upper limits on the 21 cm reionization power spectrum at $z = 8.4$, showing a factor of $\sim$4 improvement over the previous best result (P14). We find a 2σ upper limit of $(22.4\,\mathrm{mK})^2$ by fitting a flat power spectrum in a k range from $0.15 < k < 0.5\,h\ \mathrm{Mpc}^{-1}$ to the dimensionless power spectrum, $\Delta^2(k)$, measured by the PAPER instrument. We coarsely show that these upper limits imply a minimum spin temperature for hydrogen in the IGM. Although these limits are dependent on the model chosen for the power spectrum, we use a patchy reionization model to show that limits of $T_s > 4\,\mathrm{K}$ are fairly generic for models with ionization fractions between 0.3 and 0.7. A more detailed analysis of the implied constraints on spin temperature using semi-analytic reionization/heating simulations is presented in a forthcoming paper [140].

The power spectrum results that we present continue to be based on the delay-spectrum approach to foreground avoidance presented in P12b and first applied in P14. The application of a delay filter over a wide bandwidth continues to be one of the most powerful techniques yet demonstrated for managing bright smooth-spectrum foregrounds. In this paper, we extend the analysis in P14 with improved fringe-rate filtering, improved redundant calibration with OMNICAL, and with an OQE that, while not perfectly lossless, is more adept at down-weighting residual foregrounds. The combined effect of these improvements leaves a power-spectral measurement that is not consistent with zero at the 4.7σ-level, which we expect is a result of contamination from crosstalk and foregrounds. With the expansion of PAPER to 64 antennas, the extended 135 day observing campaign, and the added sensitivity benefits of fringe-rate filtering, combined with the optimization of antenna positions in PAPER for highly redundant measurements, this thermal noise limit is beginning to enter the realm of constraining realistic models of reionization.

Forthcoming from PAPER will be two seasons of observation with a 128-element array. Following the same analysis as presented here, that data set is expected to improve over the PAPER-64 sensitivity by a factor of $\sim$4 (in mK^2), with the potential for another boost to sensitivity should the new 16-m baselines provided in the PAPER-128 array configuration prove to be usable. There also remains the potential for further improvements to sensitivity through the use of longer baselines, if foregrounds can be managed effectively. As has been done recently for PAPER-32 [74, 104], future work will also extend PAPER-64 analysis to a range of redshifts and examine the power spectrum of polarized emission.

With recent breakthroughs in foreground management, the sensitivity limitations of current experiments are becoming clear. Although collecting area is vital, as discussed in [139], the impact of collecting area depends critically on the interplay of array configuration with foregrounds. Despite a large spread in collecting areas between PAPER, the MWA, and LOFAR, in the limit that foreground avoidance is the only viable strategy, these arrays all deliver, at best, comparable low-significance detections of fiducial models of reionization. To move beyond simple detection, next-generation instruments must deliver much more collecting area with very compact arrays.

The Hydrogen Epoch of Reionization Array (HERA) and the low frequency Square Kilometre Array (SKA-Low) are next generation experiments that aim to make significant detections of the 21 cm power spectrum and begin characterizing it. SKA-Low has secured pre-construction funding for a facility in western Australia. HERA was recently granted funding for its first phase under the National Science Foundation's *Mid-Scale Innovations Program.* HERA uses a close packing of 14-m diameter dishes designed to minimize the width of the delay-space kernel $\tilde{A}_\tau$ in Equation (3.8). Sensitivity forecasts for a 331-element HERA array and SKA-Low show that they can deliver detections of the 21cm reionization signal at a significance of 39σ and 21σ, respectively, using the same the conservative foreground avoidance strategy employed in this paper [139]. HERA is the natural successor to PAPER, combining a proven experimental strategy with the sensitivity to deliver results that will be truly transformative for understanding of our cosmic dawn.

Part III

Novel Imaging and The New Global Sky Model

Chapter 4

Low Frequency Mapmaking with Compact Interferometers: A MITEoR Northern Sky Map from 128 MHz to 175 MHz

The content of this chapter will be submitted the Monthly Notices of the Royal Astronomical Society *in April 2016. The authors are: Haoxuan Zheng, Max Tegmark, Joshua S. Dillon, Adrian Liu, Abraham R. Neben, Shana M. Tribiano, Victor Buza, Aaron Ewall-Wice, Hrant Gharibyan, Jack Hickish,6 Eben Kunz, Jon Losh, Andrew Lutomirski, Scott Morrison, Sruthi Narayanan, Ashley Perko, Devon Rosner, Nevada Sanchez, Katelin Schutz, Michael Valdez, Hung-I Yang, Kristian Zarb Adami, Ioana Zelko, and Kevin Zheng.*

4.1 Introduction

Mapping neutral hydrogen throughout our universe via its redshifted 21 cm line offers a unique opportunity to probe the cosmic "dark ages", the formation of the first luminous objects, and the Epoch of Reionization (EoR). In recent years a number of low-frequency radio interferometers designed to probe the EoR have been successfully deployed, such as the Low Frequency Array (LOFAR; [75]), the Murchison Widefield Array (MWA; [172]), the Donald C. Backer Precision Array for Probing the Epoch of Reionization (PAPER; [125]), the 21 cm Array (21CMA; [188]), and the Giant Metrewave Radio Telescope (GMRT; [115]). Unfortunately, the cosmological 21 cm signal is so faint that none of the current experiments have detected it yet, although increasingly stringent upper limits have recently been placed [116, 38, 127, 74, 4]. Fortunately, the next generation Hydrogen Epoch of Reionization Array (HERA; [60]) is already being commissioned and larger future arrays, such as the Square Kilometre Array (SKA; [100]), are in the planning stages.

Mapping diffuse structure is important for EoR science. A major challenge in the field is that foreground contamination from our Galaxy and extragalactic sources is perhaps four orders of magnitude larger in brightness temperature than the cosmological hydrogen signal [34, 38, 127, 4]. Many first generation experiments therefore focus on a foreground-free region of Fourier space, giving up on any sensitivity to the foreground-contaminated regions [126, 138, 89, 90, 39]. To access those regions and regain the associated sensitivity, one must accurately model and remove such contamination, a challenge that requires even greater sensitivity as well as more accurate calibration and beam modeling than the current state-of-the-art in radio astronomy (see Furlanetto et al. [54], Morales and Wyithe [108] for reviews). Moreover, an accurate sky model is important for calibrating these low-frequency arrays. For non-redundant arrays—arrays with few or no identical baselines—such as the MWA, modeling the diffuse structure is necessary for calibrating short baselines [141, 7]. For redundant arrays such as PAPER, MITEoR, and HERA [60], although they can apply redundant calibration which solves for per antenna calibration gains without any sky models [194, 4], they do need a model of the diffuse structure to lock the overall amplitude of their measurements [73].

The Global Sky Model (GSM; [34]) is currently the best model available for diffuse Galactic emission at EoR frequencies. It has been widely used in the EoR community as a foreground model. The sky maps used in the GSM that are close to the EoR frequency (100 MHz to 200 MHz) are the 1999 MRAO+JMUAR map [5, 92] at 45 MHz and 3.6° resolution, and the 1982 Haslam map [62, 63, 149] at 408 MHz and 0.8° resolution. Other notable sky maps include the 1970 Parkes maps [80] near the equatorial plane at 85 MHz and 150 MHz, with 3.5° and 2.2° resolution, respectively. However, both the frequency and sky coverage of these maps are insufficient to achieve the precision necessary for 21 cm cosmology.

All the maps listed above were made using steerable single dish antennas. The sensitivity necessary for EoR science requires a large collecting area, for which steerable single dish radio telescopes become prohibitively expensive [166]. As a result, the aforementioned EoR experiments have all opted for interferometry, combining a large number of independent antenna elements which are (except for GMRT) individually more affordable. Since these low-frequency instruments are designed for EoR science, they differ in many ways from traditional radio interferometers. For example, PAPER's array layout maximizes redundancy for optimal sensitivity and calibration [123], which makes it unsuitable for standard interferometric imaging. MWA, on the other hand, has a minimally redundant layout suitable for imaging, and customized algorithms have been developed to optimize its imaging performance such as Sullivan et al. [161] and Offringa et al. [114]. Some progress has been made to create maps of large-scale structure [183], they have relied on traditional algorithms designed for narrow-field imaging.

Fundamentally, interferometric imaging is a linear inversion problem, since the visibility data are linearly related to the sky temperature: the vector of measured data equals the vector of sky temperature times some huge matrix, with added noise. A number of map-making frameworks using this matrix approach have been proposed in recent years. For example, Shaw et al. [158] proposed a technique on spherical harmonics designed for drift scan instruments. Ghosh et al. [57] focused mostly on imaging fields of view on the scale of a few degrees. While Dillon et al. [40] develops much of the same formalism that we use, they focus less on high-fidelity mapmaking and more on propagating mapmaking effects into 21 cm power spectrum estimation. The "brute force" matrix inversion and full-sky imaging approach we pursue is desirable since it is optimal in data reduction and conceptually simple, but is very computationally expensive. In terms of computations, in order to have 1° pixel size, the entire sky consists of about 5×10^4 pixels and thus requires the inversion of a 5×10^4 by 5×10^4 matrix. While not computationally prohibitive, it is desirable to minimize the number of such inversions. Full-sky inversion has only becomes more relevant recently with the deployment of low-frequency radio interferometers with very short baselines (on the order of a few wavelengths), which are necessary to probe the structures on large angular scales.

In this work, we propose a practical "brute-force" matrix-based method for imaging diffuse structure using interferometers. We pixelize the temperature distribution in the entire sky into a vector $\boldsymbol{s}$, and flatten all of the complex visibility data from all baselines and all times (as much as a whole sidereal day) into another vector $\boldsymbol{v}$. These two vectors are related by a set of linear equations described by a large matrix $\mathbf{A}$, determined by the primary beam and the fringe patterns of various baselines. We then deploy familiar tools for solving linear equations such as regularization and Wiener filtering to obtain a solution for the entire sky. This algorithm can potentially be applied to any interferometer, and it is especially advantageous for imaging diffuse structure rather than compact sources. The algorithm naturally synthesizes data from different times, without requiring any approximations such as a flat sky, a co-planar array, or a "w-term". It can therefore fill in "missing modes" that are otherwise undetermined using each snapshot individually. Furthermore, it allows for the optimal combination of data from different instruments in visibility space, rather than image space. Lastly, even for arrays with very high angular resolution for which the matrix inversion is impossible, this method can be used to calibrate its short baselines without needing an external sky model.

The rest of this paper is structured as follows. In Section 4.2, we describe our method in detail. In Section 4.3, we quantify the accuracy of the method using simulations for both MITEoR's maximally redundant layout and MWA's minimally redundant layout to provide intuitive understanding of its performance. In Section 4.4, we apply the algorithm to our MITEoR data set to perform absolute calibration, and produce a northern sky map with about 2° resolution centered around 150 MHz.

We use the obtained map to compute spectral indices throughout our frequency range, as well as spectral indices when compared to 85 MHz and 408 MHz maps. To estimate the accuracy of our map, we compare it with both the 1970 Parkes map at 150 MHz and the prediction of the GSM at 150 MHz.

4.2 Wide Field Interferometric Imaging

4.2.1 Framework

We start with the interferometer equation for the visibility across the ith and jth antennas, following the conventions in Tegmark and Zaldarriaga [166]:

$$v_{ij} = \int T_s(\boldsymbol{k}) B(\boldsymbol{k}) e^{i\boldsymbol{k}\cdot\boldsymbol{r}_{ij}} d\Omega, \tag{4.1}$$

where T_s is the sky temperature, B is the primary beam strength, $\boldsymbol{r}_{ij}$ is the baseline vector between the ith and jth antennae, and $\boldsymbol{k}$ is the wave vector of the electromagnetic radiation whose amplitude $|\boldsymbol{k}| = \frac{2\pi}{\lambda}$ and whose direction $\hat{\boldsymbol{k}}$ points towards the observer [169]. Radio interferometry typically takes advantage of this equation by first performing a coordinate transformation on Eq. (4.1) from $\boldsymbol{k}$ on the spherical surface to its projection on the xy-plane:

$$v_{ij} = v(\boldsymbol{u}_{ij}) = \int \frac{T_s(\boldsymbol{q}) B(\boldsymbol{q})}{\sqrt{1 - |\boldsymbol{q}|^2}} e^{i2\pi\boldsymbol{q}\cdot\boldsymbol{u}_{ij}} d\boldsymbol{q}, \tag{4.2}$$

where $\boldsymbol{q} = \frac{\lambda}{2\pi}(k_x, k_y)$, and $\boldsymbol{u}_{ij} = \frac{\boldsymbol{r}_{ij}}{\lambda}$ [166]. By treating the visibilities, v_{ij}, and the beam-weighted sky image, $\frac{T_s(\boldsymbol{q})B(\boldsymbol{q})}{\sqrt{1-|\boldsymbol{q}|^2}}$, as a Fourier pair and performing 2D Fourier transforms on measured v_{ij}'s, radio interferometers have obtained high quality images of the sky with tremendous success. However, this Fourier approach comes with one important limitation. Generally speaking, the beam-weighted sky image is band-limited to the unit circle $|\boldsymbol{q}| \leq 1$, so by the Nyquist theorem, one has to have uv sampling finer than half a wavelength to avoid aliasing in the image. In reality, it is difficult to have any baselines shorter than half a wavelength due to the physical size of the antennas. What is more, since the shortest baseline has to be longer than the diameter of the antenna, the largest angular scale available is always smaller than the primary beam width, making aliasing inevitable. Mature techniques such as anti-aliasing filters [163] have been developed to solve this problem, but they are typically tailored for resolving compact structures rather than a diffuse background. The traditional *uv*-plane approach is therefore not ideal for imaging larger angular scales than the primary beam width.

To overcome this challenge, we follow Dillon et al. [40] and attack Eq. (4.1) from a different angle. By discretizing the integral over sky angles into a sum over sky pixels indexed by n, and including measured visibilities from all times (such as 8 hours during a night's drift scan), Eq. (4.1) becomes

$$v_{ijt} = \sum_n T_s(\boldsymbol{k}_n) B(\boldsymbol{k}_n, t) e^{i\boldsymbol{k}_n \cdot \boldsymbol{r}_{ij}(t)} \Delta\Omega, \tag{4.3}$$

where $\Delta\Omega$ is the pixel angular size, and we express all quantities in equatorial coordinates. Here the sky, T_s, is static, and $B(\boldsymbol{k}_i, t)$ and $\boldsymbol{r}_{ij}(t)$ change due to Earth's rotation and possibly the instrument's re-pointing such as in MWA's case. We then flatten visibilities from all times and baselines into one vector $\boldsymbol{v}$, flatten $T_s(\boldsymbol{k}_n)$ for all discretized directions in the sky into $\boldsymbol{s}$, and package the rotating beam and baseline information into a big matrix $\mathbf{A}$. Eq. (4.3) now takes the form of a system of linear equations

$$\boldsymbol{v} = \mathbf{A}\boldsymbol{s} + \boldsymbol{n}, \tag{4.4}$$

where we have now taken into account visibility noise $\mathbf{n}$ with mean zero and covariance $\langle \boldsymbol{n}\boldsymbol{n}^t \rangle \equiv \mathbf{N}$. Since the sky temperature is always real whereas $\mathbf{A}$ and $\boldsymbol{v}$ are complex, we "realize" the system by appending the imaginary part of $\mathbf{A}$ and $\boldsymbol{v}$ after their real parts, and double the noise variance. Also, it is important to use double precision for matrix operations throughout this work[1].

---

[1] Although neither the data nor the beam pattern are understood to any level near the numerical precision of single point floating numbers, and our simulation results do not rely on such precise knowledge (we calculate $\mathbf{A}$ using single precision, and the simulated data have noise), we found that using single precision during matrix operations has a significant and negative impact on the quality

To optimally estimate $\boldsymbol{s}$, we use the minimum variance estimator [165]

$$\hat{\boldsymbol{s}} = (\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A})^{-1}\mathbf{A}^t\mathbf{N}^{-1}\boldsymbol{v}, \tag{4.5}$$

for which the error covariance matrix is

$$\boldsymbol{\Sigma} \equiv \langle(\hat{\boldsymbol{s}} - \boldsymbol{s})(\hat{\boldsymbol{s}} - \boldsymbol{s})^t\rangle = (\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A})^{-1}. \tag{4.6}$$

While this is an elegant set-up with a straightforward solution, there are two major technical difficulties with this approach: the size of $\mathbf{A}$ and the invertibility of $\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}$. We address these topics in detail in the next two sections.

4.2.2 Constructing the A-Matrix

To obtain an intuitive understanding of the linear system described by Eq. (4.4), we express the nth column of $\mathbf{A}$ according to Eq. (4.3), and change the coordinate system back to a rotating sky with a static beam and baselines:

$$\begin{aligned} a_{ijt}^{(n)} &= B(\boldsymbol{k}_n, t)e^{i\boldsymbol{k}_n\cdot\boldsymbol{r}_{ij}(t)}\Delta\Omega \\ &= B(\boldsymbol{k}_n)e^{i\boldsymbol{k}_n(t)\cdot\boldsymbol{r}_{ij}}\Delta\Omega. \end{aligned} \tag{4.7}$$

For a given baseline pair ij, $a_{ijt}^{(n)}$ is the visibility over time on this baseline produced by the nth pixel in the sky if it had unit temperature. Stacking all baseline pairs together, the nth column of $\mathbf{A}$ is simply the set of visibilities the instrument would have measured if the sky only consisted of a single nth pixel of unit flux, with 0 in all other pixels. Because the visibilities we measure are a linear superposition of contributions from all directions in the sky [40], finding an optimal solution to Eq. (4.4) is simply asking how much flux is needed in each pixel of the sky in order to jointly produce the visibilities we actually measured.

With this intuition in mind, we turn to the topic of how to pixelize the sky. Since

---

of our results.

$\mathbf{A}$ is not sparse, computing Eq. (4.5) will inevitably involve some form of inversion of $\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}$, whose computational cost scales as n_{pix}^3. We would therefore like the number of pixels to be as small as possible. On the other hand, in order to preserve accuracy in discretizing Eq. (4.1) to Eq. (4.3), $\Delta\Omega$ needs to be smaller than square of the angular resolution of the instrument, where the angular resolution is roughly $\frac{\lambda}{d_{\text{max}}}$, and d_{max} is the longest baseline length. When we pixelize the sky using the HEALPIX convention [59], we thus need $n_{\text{side}} > \frac{d_{\text{max}}}{\lambda}$. For MITEoR, whose longest baseline is only about 15 wavelengths, we choose $n_{\text{side}} = 64$, which translates to about 4.8×10^4 pixels, which is about the largest size that can be comfortably manipulated on a personal computer. To obtain higher resolution maps for other instruments with even longer baselines, one may choose to use non-uniform pixelization, which we discuss in Appendix 4.B. It is worth pointing out that resolution alone does not decide the number of pixels needed; rather, what matters is the ratio of the primary beam width to the pixel width. Thus, the pixel number may not be very large if one wishes to image a small patch of sky at higher resolution with a narrow-beam instrument, depending on how quickly the sidelobes of the beam fall off away from its center.

4.2.3 Regularization and Point Spread Functions

For interferometric data sets, $\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}$ can be poorly conditioned and numerically non-invertible because the instrument is insensitive to certain linear combinations of the sky, so we insert a regularization matrix $\mathbf{R}$ (which we assume to be symmetric throughout this work) into Eq. (4.5):

$$\hat{\boldsymbol{s}}^{\mathbf{R}} = (\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A} + \mathbf{R})^{-1}\mathbf{A}^t\mathbf{N}^{-1}\boldsymbol{v}. \tag{4.8}$$

We then substitute Eq. (4.4) into Eq. (4.8), which leads to

$$\begin{aligned}\hat{\boldsymbol{s}}^{\mathbf{R}} =&(\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}+\mathbf{R})^{-1}\mathbf{A}^t\mathbf{N}^{-1}(\mathbf{A}\boldsymbol{s}+\boldsymbol{n})\\ =&\mathbf{P}\boldsymbol{s}+(\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}+\mathbf{R})^{-1}\mathbf{A}^t\mathbf{N}^{-1}\boldsymbol{n}, \qquad (4.9)\\ \langle\hat{\boldsymbol{s}}^{\mathbf{R}}\rangle =&\mathbf{P}\boldsymbol{s} \qquad (4.10)\end{aligned}$$

where we define the point spread matrix $\mathbf{P}\equiv(\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}+\mathbf{R})^{-1}\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}$. The matrix $\mathbf{P}$ is not the identity matrix due to the insertion of $\mathbf{R}$, so each column of $\mathbf{P}$ acts as a point spread function (PSF) for the corresponding pixel in the true sky map $\boldsymbol{s}$. The introduction of regularization thus represents a comprise between the goal of completely removing the PSF and producing a map with favorable noise properties. The result is, in the language of traditional radio astronomy, a position-dependent synthesized beam. For each of the PSF, we can calculate the effective full width half maximum (FWHM) using

$$\theta_{\mathrm{FWHM}}=\sqrt{\pi^{-1}n_{\frac{1}{2}}\Delta\Omega}, \qquad (4.11)$$

where $n_{\frac{1}{2}}$ is the number of pixels in the PSF whose absolute value is above half of the maximum pixel, and $\Omega_{n_{\mathrm{side}}}$ is the angular area for each pixel. We use θ_{FWHM} to represent angular resolution throughout this work.

It is also worth noting that the sum of each row of $\mathbf{P}$ is a damping factor for the solution. For a fictitious uniform sky, if each row of $\mathbf{P}$ does not sum to unity, then the resulting solution will not retain the same uniform amplitude as the original fictitious sky. Thus, it is desirable to renormalize each row of $\mathbf{P}$ to sum to unity.

With the introduction of $\mathbf{R}$, the error covariance is now

$$\begin{aligned}\boldsymbol{\Sigma}^{\mathbf{R}} =&\langle(\hat{\boldsymbol{s}}^{\mathbf{R}}-\mathbf{P}\boldsymbol{s})(\hat{\boldsymbol{s}}^{\mathbf{R}}-\mathbf{P}\boldsymbol{s})^t\rangle\\ =&(\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}+\mathbf{R})^{-1}(\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A})(\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}+\mathbf{R})^{-1}\\ =&\mathbf{P}(\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}+\mathbf{R})^{-1}. \qquad (4.12)\end{aligned}$$

To quantify how well the predicted error in Eq. (4.12) accounts for actual noise we obtain in either our simulations or maps using real data, we define the normalized error δ at the ith pixel as

$$\delta_i = \frac{|\hat{s}_i^{\mathbf{R}} - (\mathbf{Ps})_i|}{\sqrt{\Sigma_{ii}}}. \tag{4.13}$$

Intuitively, δ represents the discrepancy between the recovered map and the ground truth in units of the expected noise level, and is thus expected to center around 1 in absence of any systematic effects.

Now we discuss the choice of $\mathbf{R}$. Inserting $\mathbf{R}$ is equivalent to having prior knowledge of the sky, where the Bayesian prior for $\mathbf{s}$ has mean $\langle \mathbf{s} \rangle = \boldsymbol{s}^p = 0$ and covariance matrix $\langle (\boldsymbol{s} - \boldsymbol{s}^p)(\boldsymbol{s} - \boldsymbol{s}^p)^t \rangle = \mathbf{R}^{-1}$. If one has a sky model with well-characterized error properties (not the case for this work), there is a natural choice of $\mathbf{R}$ which we discuss in the next section. In the absence of such a model, however, the choice of $\mathbf{R}$ depends on the array layout, noise properties of the visibilities, as well as the trade-off between angular resolution and noise (we will demonstrate these qin much more detail through simulation in Section 4.3). The simplest form of $\mathbf{R}$ is a uniform regularization matrix proportional to the identity: $\mathbf{R} = \epsilon^2 \mathbf{I}$. ϵ^{-1} has the same units as $\boldsymbol{s}$, and can be compared to the noise level in the map solution: smaller ϵ^{-1} suppresses noisy modes more strongly since it will dominate the diagonal of $\boldsymbol{\Sigma}^{\mathbf{R}}$, but it also hurts angular resolution as it introduces wider point spread functions by bringing $\mathbf{P}$ farther away from the identity matrix. Therefore, ϵ is a tunable parameter deciding the trade-off between resolution and noise. In our simulations in Section 4.3 and map making using MITEoR data in Section 4.4, we show various choices of $\mathbf{R}$, and we leave investigations of the optimal choice of $\mathbf{R}$ for future work.

4.2.4 Wiener Filtering and Incorporating Prior Knowledge

If we have prior knowledge of the sky such as previous measurements, we would like to optimally combine the existing sky map $\boldsymbol{s}^p$, with our visibility measurements (this is not carried in our simulations or our final map product). We accomplish this by shifting our focus from $\boldsymbol{s}$ to $\boldsymbol{s} - \boldsymbol{s}^p$, and Eq. (4.5) becomes

$$\hat{\boldsymbol{s}} - \boldsymbol{s}^p = (\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A})^{-1}\mathbf{A}^t\mathbf{N}^{-1}(\boldsymbol{v} - \mathbf{A}\boldsymbol{s}^p). \tag{4.14}$$

We then quantify the uncertainty in our prior knowledge, by defining the covariance matrix $\mathbf{S} \equiv \langle(\boldsymbol{s}-\boldsymbol{s}^p)(\boldsymbol{s}-\boldsymbol{s}^p)^t\rangle$, and we use $\mathbf{R}=\mathbf{S}^{-1}$ as the regularization matrix:

$$\hat{\boldsymbol{s}}^{\mathbf{R}} - \boldsymbol{s}^p = (\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}+\mathbf{S}^{-1})^{-1}\mathbf{A}^t\mathbf{N}^{-1}(\boldsymbol{v}-\mathbf{A}\boldsymbol{s}^p). \tag{4.15}$$

Our regularized estimate for $\boldsymbol{s}$ then becomes

$$\hat{\boldsymbol{s}}^{\mathbf{R}} = (\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}+\mathbf{S}^{-1})^{-1}\mathbf{A}^t\mathbf{N}^{-1}(\boldsymbol{v}-\mathbf{A}\boldsymbol{s}^p)+\boldsymbol{s}^p. \tag{4.16}$$

There are three ways of understanding our choice of regularization $\mathbf{R}=\mathbf{S}^{-1}$. First, using the identity

$$(\mathbf{X}^{-1}+\mathbf{Y}^{-1})^{-1} = \mathbf{X}(\mathbf{X}+\mathbf{Y})^{-1}\mathbf{Y} \tag{4.17}$$

for invertible[2] square matrices $\mathbf{X}$ and $\mathbf{Y}$, one can show that Eq. (4.15) is equivalent to applying a Wiener filter $\mathbf{W}$ to Eq. (4.14):

$$\hat{\boldsymbol{s}}^{\mathbf{R}} - \boldsymbol{s}^p = \mathbf{W}(\hat{\boldsymbol{s}}-\boldsymbol{s}^p), \tag{4.18}$$

where $\mathbf{W}=\mathbf{S}(\boldsymbol{\Sigma}+\mathbf{S})^{-1}$, with $\boldsymbol{\Sigma}$ defined in Eq. (4.6).

Secondly, by using Eq. (4.17), one can show that Eq. (4.16) is equivalent to an inverse variance weighted average of the unregularized $\hat{\boldsymbol{s}}$ and the prior knowledge $\boldsymbol{s}^p$:

$$\hat{\boldsymbol{s}}^{\mathbf{R}} = \mathbf{S}(\boldsymbol{\Sigma}+\mathbf{S})^{-1}\hat{\boldsymbol{s}} + \boldsymbol{\Sigma}(\boldsymbol{\Sigma}+\mathbf{S})^{-1}\boldsymbol{s}^p, \tag{4.19}$$

where $\boldsymbol{\Sigma}$ and $\mathbf{S}$ are the noise covariance matrices for the unregularized $\hat{\boldsymbol{s}}$ and the prior knowledge $\boldsymbol{s}^p$, correspondingly. It is reassuring to see from Eq. (4.19) that $\hat{\boldsymbol{s}}^{\mathbf{R}} \to \boldsymbol{s}^p$ when $\boldsymbol{\Sigma} \gg \mathbf{S}$, and vice versa.

Lastly, Eq. (4.16) is equivalent to combining visibility data and a previous sky map through appending $\boldsymbol{s}^p$ to $\boldsymbol{v}$ and the identity matrix $\mathbf{I}$ to $\mathbf{A}$ in Eq. (4.4), and

---

[2] We have mentioned that $\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}$ is usually numerically uninvertible, but as long as one excludes pixels never above the horizon, it should be formally invertible, with some eigenvalues very close to zero.

solving it using Eq. (4.5) without any regularization.

4.2.5 Further Generalization

Throughout this section, we have, for clarity, limited our discussion to data sets in the form of baseline by time at a given frequency from one instrument. We generalize this to incorporate data sets from multiple frequencies and even multiple instruments. To synthesize multiple instruments at the same frequency, we simply append all the flattened data vectors together, and stack their corresponding **A** matrices in the same order. We then solve for the sky in one step using Eq. (4.5) or Eq. (4.8). We demonstrate this through simulation in Section 4.3.3.

To synthesize multiple frequencies, we assume that the sky map only differs by an overall scaling factor throughout a given frequency range, and the accuracy approximation depends on how wide the frequency range is. If at each frequency ν we have

$$\boldsymbol{v}_\nu = \mathbf{A}_\nu \boldsymbol{s}_\nu \tag{4.20}$$

together with

$$\boldsymbol{s}_\nu = f(\nu)\boldsymbol{s}, \tag{4.21}$$

we obtain

$$\boldsymbol{v}_\nu = f(\nu)\mathbf{A}_\nu \boldsymbol{s} = \mathbf{A}'_\nu \boldsymbol{s}. \tag{4.22}$$

If we know $f(\nu)$, we can simply stack $\boldsymbol{v}_\nu$'s and $\mathbf{A}'_\nu$'s and solve for $\boldsymbol{s}$. In reality, although we do not know $f(\nu)$ very accurately, we can iterate this, where we start with an estimated $f_0(\nu)$ and at each iteration reevaluate $f(\nu)$ given by the best-fit ratio between $\boldsymbol{v}_\nu$ and $\mathbf{A}_\nu \boldsymbol{s}$ from the previous iteration. As a result, we can empirically obtain $f(\nu)$ in addition to $\boldsymbol{s}$. We demonstrate this in detail in Section 4.4.

4.3 Simulations

In this section, we perform simulations for both MITEoR and the MWA to demonstrate the algorithms we have described in the previous section. The distinct array layouts of MWA and MITEoR complement each other in demonstrating various aspects of our algorithm. The simulations are all based on just one night of observation on each instrument. To simulate visibilities, we pixelize the GSM using to the HEALPIX resolution $n_{\text{side}} = 128$, but when we solve for the sky we use $n_{\text{side}} = 32$ (a pixel size of 2°), so that we can quantify any potential errors introduced by having coarse pixels. The noise is simulated as Gaussian noise independent across baselines, frequencies and times, whose amplitude is set to the simulated autocorrelation over $\sqrt{\Delta\nu\Delta t}$, where $\Delta\nu$ is the bandwidth for each frequency bin, and Δt is the integration time. For MITEoR, the noise variance is also scaled down by the redundancy factor (the number of antenna pairs sharing a baseline type) for each baseline type. We then follow Sections 4.2.1-4.2.3 to compute maps and their error properties.

4.3.1 MITEoR Simulation

MITEoR is a highly redundant array. It consists of 64 dual-polarization antennas on a square grid with 3 m spacing. For each polarization, it has 2,016 baselines with 112 unique baseline types. For this simulation, we include the shortest 102 out of its 112 unique baseline types, with baseline lengths between 3 m and 25.5 m. The primary beam model is numerically simulated as described in Zheng et al. [194], and its FWHM is about 40° near 150 MHz. We simulate for a single night's observation in the local sidereal time (LST) range between 12 hours and 24 hours to resemble the LST coverage in Section 4.4, at 150 MHz, with 144 seconds integration, and 0.75 MHz frequency bin width. The **A**-matrix size for this set-up is $122,400 \times 9,725$. We choose $\epsilon^2\mathbf{I}$ as the regularization matrix, where $\epsilon^{-1} = 100\,\text{K}$ (The result is not sensitive to the choice of ϵ within an order of magnitude, as shown later in Fig. 4-2.).

The results we obtain are shown in the first column of Fig. 4-1, in which the median noise is 14.9 K. We computed the position-dependent FWHM using the PSF matrix, and found that 96% of the pixels have FWHM less than the pixel size, which means that the resolution is limited by pixelization rather than the PSF. Note that Cygnus A (Cyg A) and Cassiopeia A (Cas A) cast noticeable "ringing" due to their extreme brightness and the effect of the PSF. Due to the coarseness of the pixelization in this section, we defer a demonstration of point source removal, as applied to Cyg A and Cas A, to Section 4.4. Lastly, the normalized residual map shows that the errors in the recovered map are well described by the noise properties described by the noise covariance matrix $\mathbf{\Sigma}^{\mathbf{R}}$, with the exception of those regions with sharp features, such as pixels near Cyg A, Cas A, and the Galactic center. This is caused by the coarse pixelization, and can be remedied by decreasing the pixel size.

4.3.2 MWA Simulation

MWA is a minimally redundant array, with 8,128 different baselines. For this simulation, we use the shortest 195 cross-correlation baselines, with baseline lengths between 7.7 m and 25.5 m (referred to as MWAcore from hereon). We limit the baseline lengths in order to use the same $n_{\text{side}} = 32$ pixelization as the MITEoR case, as longer baselines requires a much finer pixelization. The primary beam model is obtained by calculating the analytic expression for phased array of short dipoles, and the FWHM is about 14° around 150 MHz. We simulate for a single night's observation in the LST range between 12 hours and 24 hours, at 150 MHz, with 144 seconds integration, and 0.64 MHz frequency bin width. The beam remains in the zenith-pointing drift scan mode for the entire observation, which is not how the MWA typically operates (but doable). The $\mathbf{A}$-matrix size for this set-up is $234,000 \times 9,785$. We choose $\epsilon^2\mathbf{I}$ as the regularization matrix, where $\epsilon^{-1} = 300\,\text{K}$. The results we obtain are shown in the second column of Fig. 4-1. The median noise in the map is 47.8 K, and the higher noise lines follow the trajectory of the nulls between the main lobe and the first side lobe, as well as between the first and second side lobes. In terms of the angular resolution, 89% of the sky is limited by pixelization.

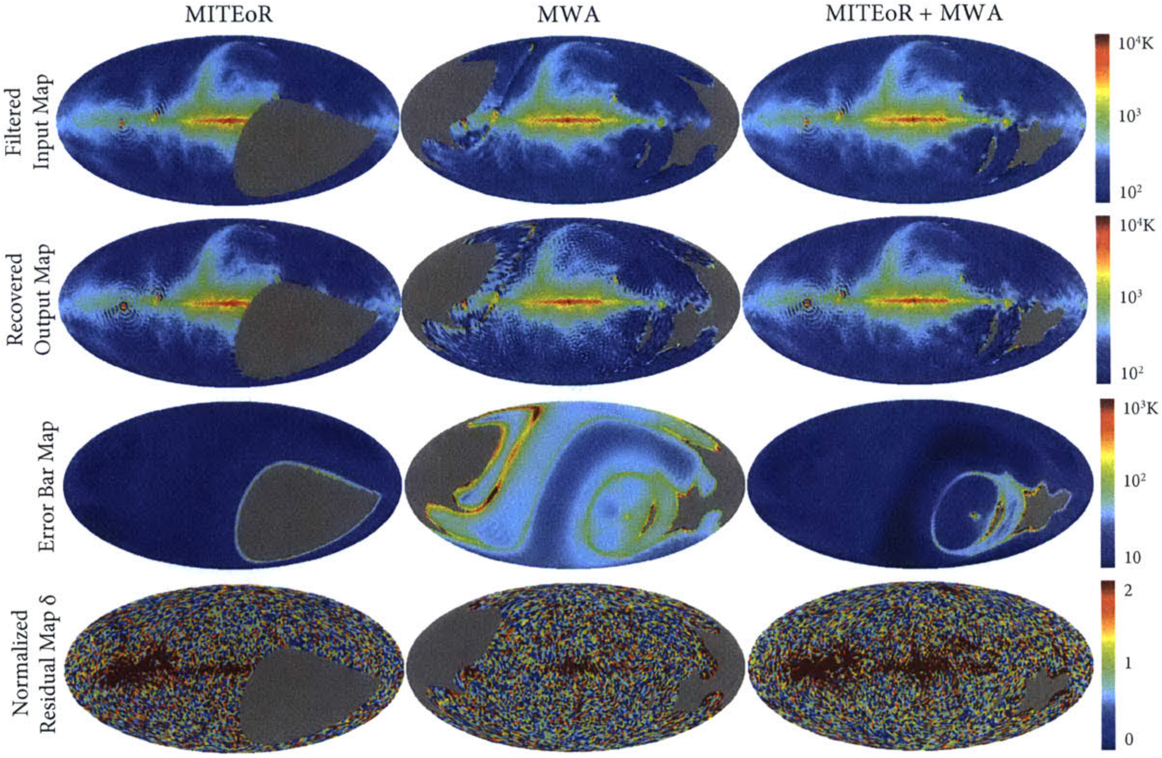

Figure 4-1: Simulated results for MITEoR (left), the MWAcore (middle), and the two combined (right). Filtered input GSM maps (top) are generated by applying the PSF matrix $\mathbf{P}$ to the GSM with $n_{\text{side}} = 32$. Grey areas represent directions that are either never above the horizon or have negligible sensitivity. Recovered output maps are obtained by applying Eq. (4.8) to simulated noisy visibilities (assuming one night's observation and one frequency channel with $< 1\,\text{MHz}$ bandwidth), using uniform diagonal regularization matrices with ϵ^{-1} of $100\,\text{K}$, $300\,\text{K}$, and $100\,\text{K}$ respectively. The error bar maps are obtained by plotting the square roots of the diagonal entries in $\mathbf{\Sigma^R}$, and the color scale is one order of magnitude smaller than the output maps. Lastly, the normalized residual δ maps represent the ratio of the actual error in our maps to the error bars, as defined in Eq. (4.13), and their values center around 1 as we expect.

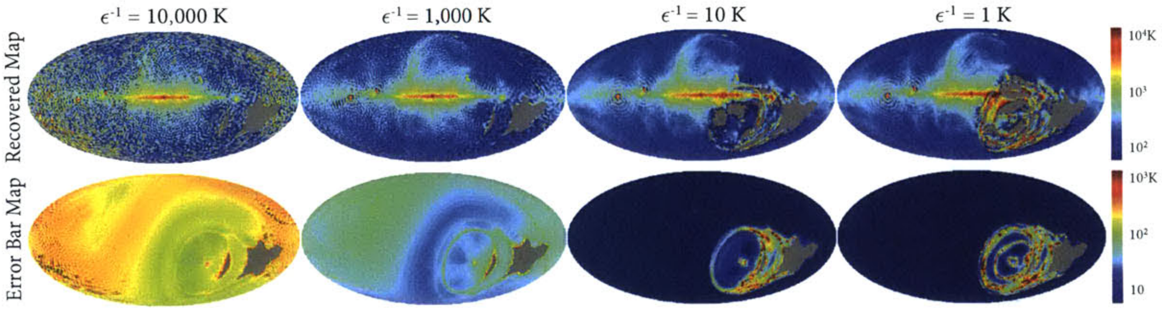

Figure 4-2: Output maps recovered in simulation and their error bars ($\sqrt{\mathbf{\Sigma}^{\mathbf{R}}_{ii}}$), using the same set of data but different regularization matrices. As ϵ increases over 5 orders of magnitude ($\epsilon^{-1} = 100\,\mathrm{K}$ is shown in the third column of Fig. 4-1), the noise is significantly suppressed, but the properties of PSF matrix become less desirable: we start to see negative temperatures near the null regions of the MWA beam. This is because stronger regularization suppresses the "missing modes" more heavily, and as the large scale structures near the MWA's side lobe region gets suppressed, more of that region becomes negative. In addition, strong regularization creates worse "ringing" near Cyg A and Cas A, and even near the north pole spur in the 1 K case.

4.3.3 Simulation Discussion

By combining the MWAcore visibilities and the MITEoR visibilities, and using a uniform regularization $\epsilon^{-1} = 100\,\mathrm{K}$, we obtain another map shown in the third column of Fig. 4-1, whose median noise is 14.6 K. In this case, 97% of the pixels' resolution is limited by pixelization. To demonstrate the effect of tuning the parameter ϵ in the regularization matrix, we show a few different choices of ϵ in Fig. 4-2. As the regularization varies from two orders of magnitude too weak to two orders of magnitude too strong, overall noise level decreases, but stronger PSFs start to overly suppress the less sensitive regions near the nulls of the MWA beam and make the temperature negative, and create worse "ringing" around Cyg A and Cas A.

As shown in Fig. 4-1, with just one night's data, our algorithm can produce high quality diffuse structure maps using either the MWAcore or MITEoR. With the MWAcore's set-up where the shortest baseline is longer than 7 wavelengths, the algorithm successfully determines large scale modes such as the overall amplitude of the map. On the other end, the longest baselines we include are 13 wavelengths long, which naively translates to about 5° resolution, but the matrix approach recovers angular scales smaller than 2°, as both simulations show that the resolution is limited by the 2° pixel size. As we will show in Section 4.4 using 1° pixels, the true resolving power of these baselines are between 1° and 2°.

Comparing the noise maps of the MWAcore and MITEoR, where MITEoR's 14.9 K median noise level is slightly lower than MWAcore's 47.8 K, we see that the overall noise level is not very sensitive to the number of baseline types, the baselines' length distribution, or the primary beam shape. Although the collecting area of MWA tiles is 16 times larger than each bow-tie antenna used in MITEoR, it is offset by the fact that there are effectively about 2000 baselines used in the MITEoR simulation (with 112 unique baseline types) compared to the MWAcore's 195, so the noise levels are comparable in both cases. In contrast to the overall noise level, the spatial patterns of noise do depend heavily on the primary beam pattern, which is not surprising, since the primary beam pattern heavily influences the instrument's sensitivity to different parts of the sky.

The δ maps show that the error properties are well characterized by $\mathbf{\Sigma^R}$, with the majority of the δ values less 2. We see that although the visibilities are simulated using the GSM with $n_{\text{side}} = 128$, the crude pixelization of $n_{\text{side}} = 32$ is not introducing significant errors, other than in regions near Cyg A, Cas A, or the Galactic center. The importance of pixelization errors depends on the baseline lengths included, as well as the amount of noise in the visibilities. We find that pixelization errors become significant if we increase the baseline length threshold by another 50% to about 40 m. We also expect the pixelization error to become dominant if the visibilities have much lower noise, such as when they are averaged over 100 nights as opposed to a single night used here.

Given the result that just one night's data from MWAcore's zenith pointing scan can determine more than half of the sky to better than 50 K precision, MWA has the potential to make a high quality (< 10 K noise through multiple nights' observations) southern sky map with multiple beam pointings and multiple nights throughout the year to fill the sky. In addition, since MWAcore only includes a small fraction of MWA's baselines, a much higher angular resolution (less than a degree) is also achievable with finer pixelizations.

4.4 New Sky Map

As we have demonstrated in the previous section, the new imaging method works well on simulated data from both MITEoR and MWAcore. In this section, we apply the method to real data collected by MITEoR, and produce a Northern sky map at 150 MHz.

4.4.1 MITEoR Instrument and Data Reduction

As mentioned above, MITEoR [194] is a compact radio interferometer with 64 dual-pol antennas, deployed in July 2013 in The Forks, Maine (latitude 45.3°). The antennas are identical to individual MWA bow-tie antennas (without beam forming as tiles), and the full width half maximum (FWHM) of the primary beam is about 40° throughout our frequency range. The data used in this work are collected with an 8 by 8 square grid array configuration with 3 m spacings. The 8 bit correlator cross-correlates all 128 antenna-polarizations (each bow-tie antenna has two polarizations as outputs), with integration time of 2.7 seconds, instantaneous bandwidth of 12.5 MHz (tunable between 125 and 185 MHz), and frequency bin width of 50 kHz. The data used in this work were collected through 7 observing sessions, as shown in the left panel of Fig. 4-3.

We first perform redundant calibration on the raw data, using our redundant calibration pipeline described in Zheng et al. [194], with further improvements described in Appendix 4.A. Redundant calibration compresses the data in the baseline direction from 2016 cross-correlation visibilities per snapshot to 112 unique baseline types, and automatically flags bad antennas, baselines, frequencies, and time stamps from the data. It is worth noting that redundant calibration uses only the self-consistency between redundant baselines, and does not use any sky model. After redundant calibration, we further compress the data to 0.75 MHz frequency bin width by averaging every 15 frequency bins. We then average over the time direction in 2 minute intervals. We empirically estimate noise during the time averaging step by performing linear fitting over 2 minutes of data and calculating the residual power.

At this stage, we have a data cube of 4 polarizations by 75 frequency bins by 240 time steps by 112 baselines. Since we have not used any sky information, the data is not yet absolute-calibrated, meaning that for each of the $4 \times 75 = 300$ different polarization-frequency data sets, we have 3 undetermined numbers: one overall amplitude, and two re-phasing degrees of freedom. These numbers cannot be determined without performing absolute calibration (as opposed to redundant calibration) with a sky model, which we describe in the next two sections.

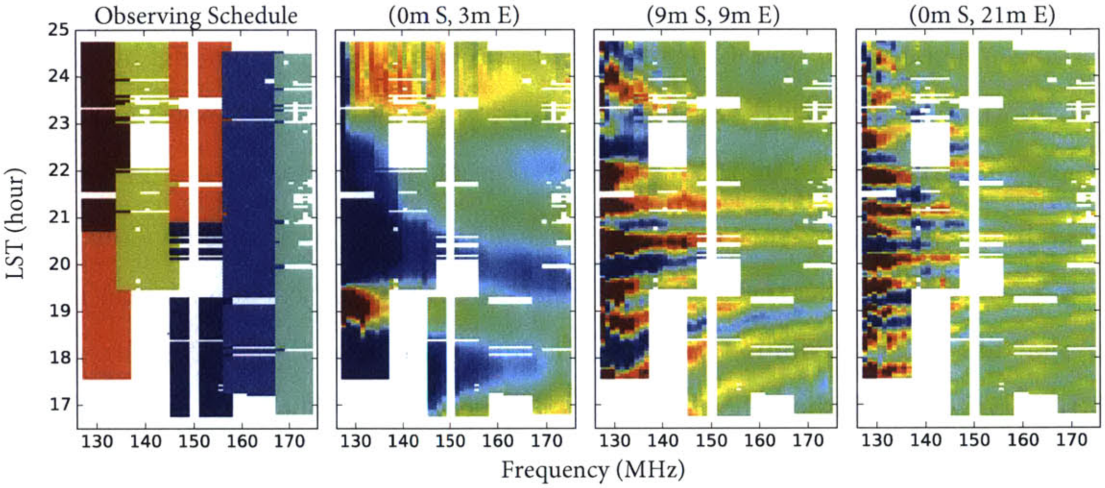

Figure 4-3: MITEoR's observing schedule on the left and three sets of visibilities in the MITEoR data product on the right. Each of the six colors represent one of the nights between July 27th and August 2nd, 2013. Midnight corresponds to LST at roughly 21 hours. The real part of visibilities over time and frequency on three baselines of very different lengths are shown here. The white gaps in frequency and time are RFI events automatically flagged by the redundant calibration pipeline.

4.4.1.1 Absolute Amplitude Calibration

The goal of absolute calibration is to determine three numbers (an overall amplitude and two re-phasing degrees of freedom) for every time by baseline data set, where each data set consists of more than 10^4 visibility measurements. This is a drastically overdetermined system given a sky model. In this section, we discuss how we determine the overall amplitude. Since we intend to use the map obtained in this work to improve the GSM in a future work, we choose not to use the GSM as our sky model. Rather, we use Cyg A and Cas A as our calibrators. After we obtain a map in Section 4.4.2, we will lock the amplitude of our map to the Parkes map at 150 MHZ [80], so the amplitude calibration and its error will only affect our spectral index results, not the amplitude of the map.

Our amplitude calibration is based on extrapolating the frequency-dependent secular decreasing flux models of Cas A [178] and the spectrum of Cyg A from Vinyaǐkin [179]. Cyg A is an ultra-luminous, jet-powered, radio-loud galaxy. For our Cyg A calibration spectrum, we use the model in Eq. 6a in Vinyaǐkin [179] of a transparent source, with a power law spectrum with spectral index α, observed through the absorbing ionized gas of our Galaxy. The data point used in the model at 152 MHz has a reported 3% error in Parker [119], and propagating the model parameters' error bars leads to a maximum of 3.2% error in our frequency range.

The frequency dependence of the decreasing flux of Cas A has been widely studied (see [65] and references, therein), and we adopt the empirical model in Vinyaikin [178] that fits to the accumulated published data taken from 1961 to 2011 from about 12 MHz - 93 GHz, including their most recent observations. The spectrum of Cas A is evaluated in frequency range 125-175 MHz, using the fitted function in Eqs. 15 and 16 in Vinyaikin [178]. With this spectrum modeled at epoch 2015.5, and the model for the frequency dependent secular decrease in Eqs. 9 and 10 of Vinyaikin [178], we evaluate the spectrum for Cas A during MITEoR observations in August 2013, approximately two years earlier. The largest source of error in the radio spectrum model of Cas A comes from our lack of complete understanding of the behavior of this supernova remnant. In an evaluation of possible periodic deviations, Helmboldt and Kassim [65] identify 4 possible modes ranging from 3 to 24 years, in a slightly lower frequency range of interest, $38 - 80$ MHz, contributing to flux deviations from a secular decrease of up to 15%.

We calibrate in the LST range between 19 and 23 hours, during which Cyg A's elevation ranges from 59° to 85°, and Cas A from 36° to 65°. To minimize error introduced by diffuse structures, we only include for calibration baselines longer than

8.6 wavelengths, which are the longest 9 baselines at the lowest frequency and 34 at the highest frequency. After fitting using Cas A and Cyg A, we find about 15% residual on the visibilities. Since the errors introduced by the Galactic plane are averaged down over different baseline types, we estimate our amplitude calibration to go down by a factor equal to the square root of the number of baselines used. Thus, our absolute calibration has an overall error of about 5% at the lowest frequency and 2.5% at the highest frequency, relative to the calibrators.

4.4.1.2 Absolute Phase Calibration

As discussed in Zheng et al. [194], the two re-phasing degrees of freedom (or re-phasing degeneracies) in the visibility space correspond to shifting the beam-weighted sky image $\frac{T_s(\boldsymbol{q})B(\boldsymbol{q})}{\sqrt{1-|\boldsymbol{q}|^2}}$, and cannot be determined using only the self consistency of visibility data without a sky model. However, this is only true for isolated snapshots in time. For instruments with large fields of view such as MITEoR, rotation of the sky does not exactly translate into shifting the beam-weighted sky image in the projected $\mathbf{q}$-plane, so a constant shift of the beam-weighted sky image caused by constant re-phasing cannot be consistent with a rotating sky. Therefore, we can use a procedure conceptually similar to self-cal to determine the re-phasing: we first image using the visibility without correcting for the re-phasing degrees of freedom, then use the image we obtain to solve for the re-phasing, and iterate until convergence. In theory, this can be done without any prior sky model, but since each iteration can be computationally expensive, we use the GSM to provide the initial re-phasing solution, and start iterating from there. It is worth noting that, at a given frequency and polarization, unlike self-cal which is solving for, say, 128 antenna calibration parameters at every time stamp, here we are only solving for 2 numbers for an entire observing session, so this iterative algorithm has negligible impact on the validity of Eq. (4.12).

In addition, even for instruments whose array layout prevents making usable images, this "self-cal" approach can be applied to remove re-phasing degeneracies. Rather than inserting regularization matrices to make $\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}$ invertible, if the goal is to calibrate out the re-phasing degeneracies, we can simply use a pseudo-inverse for $\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}$. For example, for a redundant array with no short baselines, a pseudo-inverse will remove large scale structures in the image, but this has no effect when the resulting image is used as a model to simulate visibilities on those long baselines, so "self-cal" should work just as well.

4.4.1.3 Cross-talk Removal

We define cross-talk as additive offsets on visibilities that are proportional to the amplitude of auto-correlation, with a small but constant proportionality coefficient. In theory, the cross-talk terms can be solved for in an iterative fashion similar to how we determine the re-phasing degeneracies in the previous section. However, due to the level of thermal noise and systematics present in our data set, cross-talk on our shortest baselines is highly degenerate with having a bright stripe in the trajectory of our local zenith. Thus for this work, we use the GSM to perform cross-talk removal. We use the GSM to simulate all the visibilities we measure, and for each visibility time series, we solve for the best fit using the GSM model visibility and the auto-correlation, and subtract the auto-correlation component from our data. Thus, for each visibility time series over a few hours, we use the GSM to fit and remove one degree of freedom corresponding to cross-talk.

4.4.2 Northern Sky Map Combining Multiple MITEoR Frequencies

We apply the algorithm described in Section 4.2 on the MITEoR data to obtain our Northern sky map. We have shown in Section 4.3 that the MITEoR data can in principle make high quality maps at individual frequency bins, but due to the systematics present in our data, which we discuss more in the next section, we are not able to make high quality maps using each individual frequency alone. Since in our frequency range the diffuse emission is dominantly synchrotron, which follows a smooth power law, we use techniques described in Section 4.2.5 to combine multiple

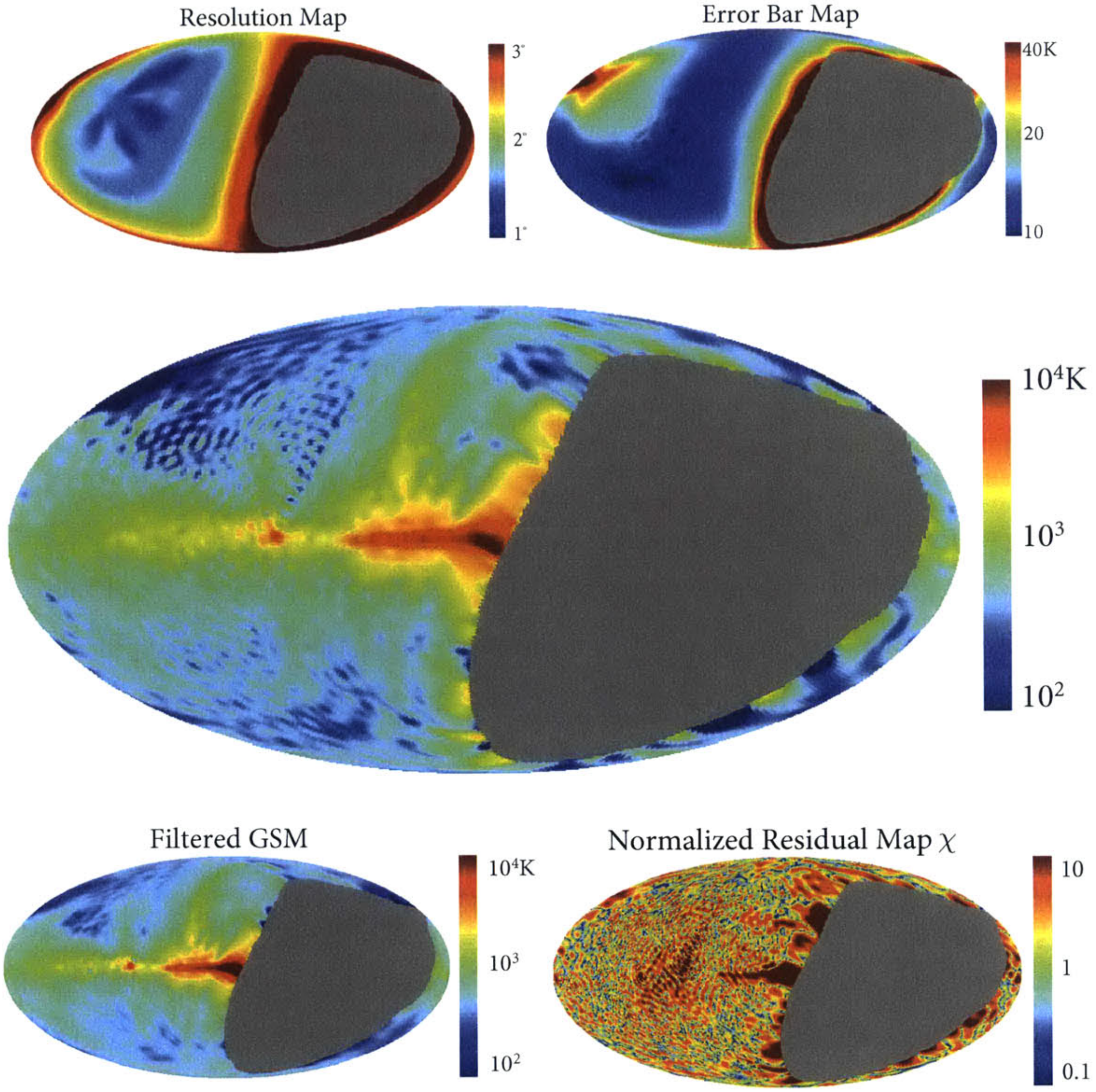

Figure 4-4: Top left: angular resolution map obtained from FWHM of the point spread functions. Top right: error bar map obtained from $\sqrt{\Sigma_{ii}}$ based on empirically estimated noise covariance $\mathbf{N}$. Mid: the northern sky map at 150 MHz, averaged from 128.5 MHz to 174.5 MHz, with Cas A and Cyg A removed using CLEAN. Bottom left: The GSM at 150 MHz, with the PSF matrix applied. Bottom right: δ as defined in Eq. (4.13), which represents the ratio between the difference between the maps and our map's error bars. The median δ is 2.16.

frequencies to form a single map with an overall spectral index, as well as the beam-averaged spectral index as a function of time. We pixelize the sky to HEALPIX $n_{\text{side}} = 64$. Since the size of the **A**-matrix is proportional to the number of frequency bins, and including the entire data set amkes the size of our **A**-matrix too large, we include only one out of every 5 frequency bins throughout the frequency range of 128.5 MHz to 174.5 MHz. This forms an **A**-matrix of size roughly 6×10^5 by 4×10^4. Since **A** has an order of magnitude more rows than columns, computing $\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}$ is the speed bottleneck and takes more than 2 days when parallelized on a single CPU. In comparison, inverting $\mathbf{A}^t\mathbf{N}^{-1}\mathbf{A}$ takes about 3 hours.

In order to calculate the relative amplitude at different frequencies, to solve for the re-phasing degeneracies in the data, and to empirically estimate the level of noise and systematics in each data set, we iterate the process described in Section 4.2.3 until the amplitudes converge to within 0.1% and re-phasings within 0.1°. At each iteration, we calculate visibilities using the solution from the previous iteration as a model, and for each frequency we fit for the re-phasing, the relative amplitude, as well as the overall error RMS. Both the best fit amplitude and the error RMS are used to re-weigh the noise covariance matrices for each frequency. When iterating, we prioritize the map's accuracy in modeling visibilities over its noise properties, so we choose a weaker regularization of $\epsilon^{-1} = 1500$ K. For the final map we choose a stronger regularization with $\epsilon^{-1} = 300$ K to obtain lower noise in the map at a cost of lower resolution in the noisy areas. The map's overall amplitude is locked to the Parkes map at 150 MHz [80] using the overlapping region. Cyg A, Cas A, and their "ringing" are removed using the CLEAN algorithm [68]. The map we obtain together with its angular resolution and error bars are shown in Fig. 4-4.

4.4.3 Error Analysis

Fig. 4-4 shows that the map we obtain agree very well with the prediction of the GSM at 150 MHz. In this section, we discuss the errors in our map and their possible causes in more detail. In terms of overall amplitude, the map's overall amplitude is locked to the Parkes 150 MHz equatorial map [80], which has 20 K uncertainty in zero level and 4% in temperature scale. In order to compare detailed structures in our result with the Parkes map and the GSM, we calculate the normalized error δ maps, shown in Fig. 4-4. The median δ compared to the filtered GSM is 2.16, and 2.79 to the unfiltered Parkes map, which are slight higher than what one might expect.

There are a few factors that make the median δ higher than 1. Firstly, our modeling of our instrument is not perfect, leading to error in our **A**-matrix. As we investigated in detail in Zheng et al. [194], our beam model has up to 10% error in some directions. Our empirically estimated visibility errors are dominated by slowly varying modes, and beam mis-modeling is the most likely cause. Another cause of error is the averaging over frequency, which assumes constant spectral index over the sky. As we show in much more detail in the next section, the spectral index changes by as much as 0.5 from the Galactic plane to out-of-plane regions, so this introduces an error for the edge frequencies on the order of $(175\,\mathrm{MHz}/150\,\mathrm{MHz})^{0.5} - 1 = 8\%$. Moreover, as we have seen in our simulation results in Fig. 4-1, pixelization can also cause un-modeled error near high temperature regions such as the point sources (Cas A and Cyg A) and the Galactic center. Lastly, neither the GSM nor the Parkes map are the ground truth: the GSM at 150 MHz is essentially an interpolation product between a 45 MHz sky map and a 408 MHz sky map, with an estimated relative error of 10% [34]. For the equatorial Parkes map, due to its 2.2° resolution, we cannot apply the PSF matrix to it before comparing it to our result. The lack of PSF on the Parkes map leads to higher error in the comparison, which we think is what makes the median δ higher for the unfiltered Parkes map than for the GSM.

4.4.4 Spectral Index Results

4.4.4.1 Spectral Indices from 128 MHz to 175 MHz

In the iterative process to compute the 150 MHz map, we also obtain relative amplitudes between all the data sets at different frequencies. At each frequency, we calculate visibilities using the 150 MHz map we obtained as a model, and compute the relative amplitude by comparing them to our data. We then perform a linear fit

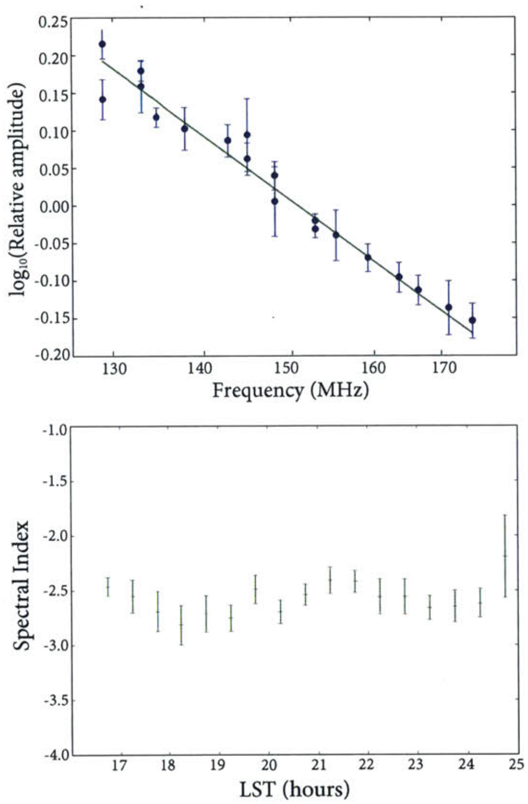

Figure 4-5: Map-averaged overall spectral index fit (top) and beam-averaged spectral index over LST (bottom). The relative amplitudes and error bars in the top plot are obtained from the iterative procedure described in Section 4.4.4. For the bottom plot, each LST is observed on 5 different frequencies on 5 different nights (see Fig. 4-3), so we can obtain a spectral index at each LST by fitting the overall amplitude over frequency with a power law. The error bars are 1σ, empirically estimated using the residuals of the power law fits.

in the log(amplitude)-log(frequency) space, and compute an overall spectral index of -2.73 ± 0.11, as shown in Fig. 4-5. The error bar is calculated using the residuals in the fitting process. In addition to this overall spectral index averaged over the entire data set, we also perform the same procedure on subsets of the data, and fit for spectral indices for every half an hour of LST. The resulting time series of spectral indices varies smoothly between -2.4 and -2.8, as shown in Fig. 4-5.

In comparison to our spectral index result, the Experiment to Detect the Global EoR Signature (EDGES; [21]) measured spectral index of -2.5 ± 0.1 [153], centering at an out-of-plane region at declination $-26.5°$ and right ascension 2 h. Our overall spectral index agrees with the spectral index obtained by EDGES, with a 1.5σ

difference. However, the difference is likely not due to statistical variation alone. The overall spectral index we present is averaged over the northern sky, so we are observing a different patch of sky compared to EDGES. Unlike EDGES, our result is influenced by both Cyg A and Cas A[3]. These strong radio sources have spectral indices of -2.7 and -2.8, respectively [119, 179, 178], so they would shift our results towards a steeper spectral index.

4.4.4.2 Spectral Indices from 85 MHz to 408 MHz

In addition to spectral indices in the EoR frequency range computed using the MITEoR data alone, we also calculate maps of spectral indices by comparing our MITEoR map to the Parkes map at 85 MHz [80] and the Haslam map at 408 MHz [62, 63, 149]. We compute per-pixel spectral index maps for all three pairs of these three maps, as shown in Fig. 4-6.

The medians of the spectral indices shown in Fig. 4-6 are $-2.60 \pm 0.29 \pm 0.07$, $-2.43 \pm 0.18 \pm 0.04$, and -2.50 ± 0.07, respectively. The first error bars come from the spread in spectral indices over the sky, and the second error bars come from 4% absolute calibration uncertainty of MITEoR. This is a very weak indication that the spectral index softens over the range from 85 MHz to 408 MHz. For comparison, Platania et al. [135] presented an overall spectral index of -2.695 ± 0.120 between the 408 MHz [62, 63, 149], 1.4 GHz [148, 146, 147], and 2.3 GHz [78] maps, which also agrees well with an earlier study at these frequencies in Giardino et al. [58].

There are two spatial features worth noting in these maps. First, the Galactic plane has softer spectral indices than the out-of-plane regions. The median spectral indices within 5° of the Galactic plane for the three map pairs are -2.27, -2.37, and -2.37, respectively. Softer spectral indices in the Galactic plane are also observed at higher frequencies in Platania et al. [135], whose spectral index map comes from three maps above the EoR frequency range, as mentioned above.

In addition to softer indices in the Galactic plane, there are regions that clearly

---

[3] The point sources are present in our spectral index results because our spectral index results are obtained in the visibility space, whereas the CLEAN algorithm that removed the point sources is performed in the image space.

deviate from the median near the Galactic poles, such as the blue regions in the Parkes vs MITEoR map, and the red regions in the MITEoR vs Haslam map. Since such departure is not seen in the Parkes vs Haslam map, this suggests that the MITEoR map is about 50 K lower in the Galactic pole regions than what the Parkes and Haslam maps jointly predict. There are two possible causes for this. The first is that the 50 K deficiency in the MITEoR map is due to systematic errors. However, the temperatures in these regions are about 240 K, so neither the 4% absolute calibration uncertainty nor the $\sim$ 15 K error bars can fully explain the 50 K difference. Another possible cause is that the MITEoR map has a higher dynamic range than the Haslam and Parkes maps, so that it recovers more details at the low end of the temperature range compared to these maps. We leave a more careful investigation of this 50 K discrepancy to a future study.

4.5 Summary and Outlook

We have presented a new method for mapping diffuse sky emission using interferometric data. We have demonstrated its effectiveness through simulations for both MITEoR and the MWAcore, where we obtained maps with better than 50 K noise and better than 2° resolution for both instruments. We applied this method on the MITEoR data set collected in July 2013, which was absolutely calibrated using Cyg A and Cas A. We obtained a northern sky map averaged from 128 MHz to 175 MHz, with around 2° angular resolution, 5% uncertainty in its overall amplitude, and better than 100 K noise. We also obtained an overall spectral index of -2.69 ± 0.11, and beam-averaged spectral indices that vary over LST between -2.4 and -2.8. Both the MITEoR visibility data and the 150 MHz sky map are publicly available at `github.com/jeffzhen/MITEoR`.

As this is our first application of this new method, there are many aspects of it that we are excited to investigate in future work. Throughout this work, we have focused on regularization matrices that are multiples of the identity matrix. However, since the sensitivity varies across the sky, especially in the case of MWAcore, a regularization

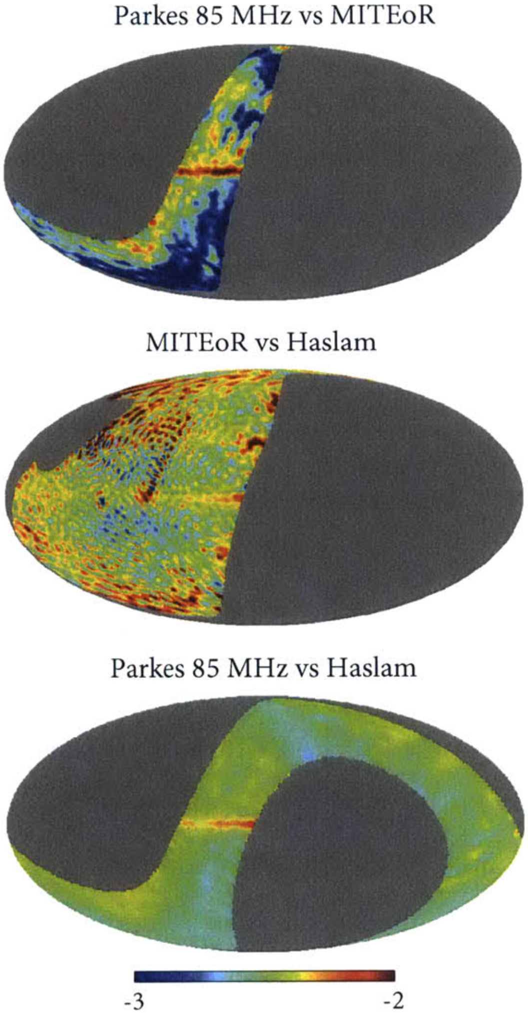

Figure 4-6: Spectral index maps between the Parkes map at 85 MHz, the MITEoR map at 150 MHz, and the Haslam map at 408 MHz. The MITEoR map is masked for regions with FWHM above 2.5° or error bar above 20 K. The median spectral indices in these maps are -2.60, -2.43, and -2.50, respectively.

matrix whose strength varies with sensitivity may achieve a better balance between noise suppression and PSF. It is also interesting to study what the optimal array layout is for imaging diffuse structure, along the lines of Dillon and Parsons [36]. Since the Earth rotates in the East-West direction, we expect the optimal array layout to be very different in the E-W direction than the N-S direction, perhaps similar to that of PAPER or CHIME [158]. Moreover, it is interesting to investigate the effectiveness of this algorithm for instruments with much narrower primary beams, such as HERA. Lastly, it is valuable to perform further study in the effectiveness of this method for the purpose of calibration, such as calibrating the shorter baselines of MWA, which complements the existing calibration algorithms that focus on point source models for very long baselines.

4.A Appendix: Improvements to Redundant Calibration

Zheng et al. [194] first demonstrated the precision and speed of redundant calibration, which has since been applied to PAPER data analysis in the latest EoR power spectrum upper limits [4]. In this work we use the same core algorithms as described in Zheng et al. [194], with improvements that make the algorithm much easier to use. In particular, we describe the improvement to rough calibration, which is the first step in the redundant calibration process. A similar algorithm has been discussed and used in Parsons et al. [127]. The completed redundant calibration software is publicly available at `https://github.com/jeffzhen/omnical`.

As described in Zheng et al. [194], redundant calibration is a three step process: rough calibration, log calibration, and linear calibration. While log calibration and linear calibration do not rely on any sky information, they require rough calibration to get started. Zheng et al. [194] described a rough calibration algorithm that requires a sky model, which makes the whole pipeline rather cumbersome, especially when dealing with data sets from a new season or a new instrument. In this section we describe a new rough calibration algorithm that can be directly applied to data without any preprocessing, thus making the entire redundant calibration pipeline sky independent. The computational complexity is proportional to the number of antennas.

Here we describe our algorithm using an 4 by 4 array on a rectangular grid as shown in Table 4.1, but the algorithm is easily generalized to other redundant configurations. At a given time and frequency, we start with $16 \times 15/2 = 120$ visibilities, v_{ij}, whose phases are γ_{ij}, where i, j are antenna numbers. The goal of rough calibration is to obtain antenna calibration phases, ϕ_i, for all 16 antennas. The simple equation that describes phase calibration is

$$\gamma_{ij} = -\phi_i + \phi_j + \theta_{i-j} \tag{4.23}$$

where θ_{i-j} is the true phase shared by all redundant baselines that share the same baseline type with γ_{ij}, and we omit 2π wrapping for this section.

We start by taking advantage of the three degree of freedom in phase degeneracies that are intrinsic to redundant calibration, namely an overall phase to all antenna phases, and two phases corresponding to rephasing the array. Due to these degeneracies, we are free to declare that $\phi_1 = 0$, $\theta_{1-2} = \gamma_{1-2}$, and $\theta_{1-5} = \gamma_{1-5}$ [4]. This then gives us

$$\phi_1 = \phi_2 = \phi_5 = 0. \tag{4.24}$$

With the first five phases $\phi_1, \phi_2, \phi_5, \theta_{1-2}, \theta_{1-5}$ all determined, we can now proceed to solve all 16 antenna phases. By applying Eq. (4.23) on baseline $\gamma_{2,3}$, we can solve for ϕ_3 since ϕ_2 and $\theta_{2-3} = \theta_{1-2}$ are known. We can repeat the process to obtain ϕ_4. Similarly, we can obtain ϕ_9 using $\gamma_{5,9}$, and extend that to obtain ϕ_{13}. Now we see that after determining the first five phases, solving for all the antennas is simply a matter of traversing the entire array and visiting every antenna with either one of the two baseline types we picked, in this case θ_{1-2} and θ_{1-5}.

As shown in Table 4.1, the only baselines we used in this example are the arrows

---

[4] We choose θ_{1-2} and θ_{1-5} because they are the two shortest non-parallel baseline types.

$$\begin{tikzcd}
\smile_1 \arrow[r] \arrow[d] & \smile_2 \arrow[r] \arrow[d] & \smile_3 \arrow[r] \arrow[d] & \smile_4 \arrow[d] \\
\smile_5 \arrow[d] & \smile_6 \arrow[d] & \smile_7 \arrow[d] & \smile_8 \arrow[d] \\
\smile_9 \arrow[d] & \smile_{10} \arrow[d] & \smile_{11} \arrow[d] & \smile_{12} \arrow[d] \\
\smile_{13} & \smile_{14} & \smile_{15} & \smile_{16}
\end{tikzcd}$$

Table 4.1: A 4 by 4 antenna array on a regular grid. Each $\smile_n$ represents an antenna, and each arrow represents a baseline used in the rough calibration algorithm.

in the table, so the computational complexity is $\mathcal{O}(N)$ where N is the number of antennas. If noise is a concern, one can increase the number of baseline types used to visit each antenna, so each phase calibration is sampled multiple times using multiple baseline types. For both MITEoR and PAPER 32 element data [127], we found that using just the two most redundant baseline types suffices.

It's important to note that while this rough calibration is much faster than log calibration and linear calibration, it is not advisable to re-run it on every data set before log calibration. As we pointed out above, rough calibration makes an arbitrary decision on the phase degeneracies, which later requires absolute calibration (with an accurate sky model) to determine. If each time step (typically a few seconds) uses a different rough calibration, we will have to run absolute calibration separately on every time step, which is very challenging as we may not always have good calibrators in the sky. It is much easier to run rough calibration only once at a single time stamp, and use it for a long period of data, such as one night or even a whole season. This way the entire time window will share one constant set of phase degeneracies, which are easy to determine in absolute calibration. Similarly, because log calibration and linear calibration also make arbitrary choices of these same re-phasing degrees of freedom, it's important to always project out any re-phasing components from the calibration phase solutions.

4.B Appendix: Dynamic Pixelization

Multiplications and inversions of 10^5 by 10^5 matrices can currently be done in a matter of days, but since the computational times grows as $\mathcal{O}(\Delta\Omega^{-3})$, improving the resolution can be very computationally demanding. To reduce the matrix size. we can resort to a more intelligent pixelization scheme, which we call dynamic pixelization.

Having uniform pixelization can be wasteful. For example, near the edge of our observable region, fine pixelization brings little advantage since our sensitivity is very low. In contrast, regions with strong flux demand much finer pixelization, such as the Galactic center and strong point sources such as Cas A and Cyg A. We use a recursive algorithm to generate a pixelization scheme that accommodates all of the above considerations. We start with a map predicted by the GSM[5], and a uniform but coarse HEALPIX nested pixelization with $n_{\text{side}} \sim n_\lambda$ for the longest baseline[6]. We then calculate a beam weighted sky map s_{bweight} given by

$$s_{\text{bweight}}(\hat{\boldsymbol{k}}) = \int_t s(\hat{\boldsymbol{k}})B(\hat{\boldsymbol{k}}, t)dt. \tag{4.25}$$

Lastly, we recursively divide each pixel into 4 sub-pixels according to HEALPIX nest scheme, until the standard deviation among the s_{bweight} values at these 4 sub-pixels are below a threshold value σ. Intuitively, since each pixel corresponds to a column in the **A**-matrix, the algorithm only splits each column into 4 when the resulting 4 columns will be sufficiently different in fitting the data.

We decide the threshold σ by numerically searching for the optimal choice. As larger σ translates into fewer pixels but less pixel precision, we choose the largest σ such that the simulated pixelization error is less than 1% in power compared to the noise in the visibility data. Fig. 4-7 shows an example pixelization scheme for MITEoR data at 160 MHz, where the final pixel count is less than a quarter of uniform pixelization, leading to a factor of 64 speed-up in computation time, while the

---

[5] Note that using the GSM for pixelization does not automatically mix GSM information into our solution.

[6] This relation is derived from $\frac{4\pi}{n_{\text{pix}}} \sim \left(\frac{1}{n_\lambda}\right)^2$.

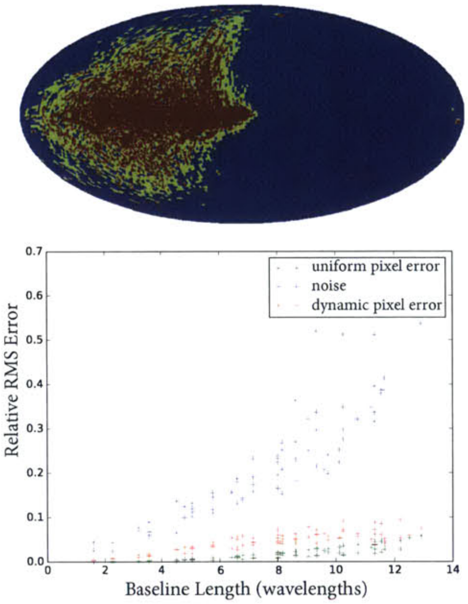

Figure 4-7: The upper plot shows a dynamic pixelization scheme, where the blue, green, and red regions represent n_{side} of 32, 64, and 128, respectively. This scheme contains 48900 total pixels, 24.9% of the pixel count in a uniform $n_{\text{side}} = 128$ pixelization. The lower plot shows the errors in the simulated visibilities compared to those simulated with a $n_{\text{side}} = 512$ pixelization, and we see that the dynamic pixelization error is a factor of two higher than uniform pixelization, and it is negligible compared to the noise level assuming one night's observation. The dynamic pixelization in this case saves computation time by a factor of 64, and the increase in error is negligible unless we average close to a hundred nights of data.

pixelization error remains negligible compared to noise in the data.

Chapter 5

An Improved Model of Diffuse Galactic Radio Emission from 10 MHz to 5 THz

The content of this chapter will be submitted the Monthly Notices of the Royal Astronomical Society *in April 2016. The authors are: Haoxuan Zheng, Max Tegmark, Joshua S. Dillon, Adrian Liu, and Abraham R. Neben.*

5.1 Introduction

Modeling diffuse Galactic radio emission has received great interest in cosmology. Cosmological signals, such as the cosmic microwave background (CMB) and redshifted 21 cm emission, have to pass through our Galaxy before reaching us, so they are inevitably mixed with foreground contamination from Galactic emission. For CMB experiments, the unpolarized CMB signal dominates over foregrounds in out-of-plane directions, though the situation for polarized CMB signal is about one to two orders of magnitude worse (for reviews, see de Oliveira-Costa and Tegmark [33] and references therein). For the redshifted 21 cm signals from the Epoch of Reionization (EoR), the foregrounds are thought to be four or more orders of magnitudes higher than the 21 cm signal (see Furlanetto et al. [54], Morales and Wyithe [108] for reviews, and

Dillon et al. [38], Parsons et al. [127], Ali et al. [4] for the latest 21 cm power spectrum upper limits), which makes foreground modeling even more important [32, 176, 126, 110, 170, 174, 136, 171, 111, 141, 138, 89, 90, 39].

In the original GSM [34], the authors carried out an exhaustive survey of existing sky maps from 10 MHz to 100 GHz, and performed a Principal Component Analysis (PCA) on 11 of the highest quality maps. They found that using just the top three principal components could explain between 90% and 99% of the variations in all the maps, depending on frequency and sky direction. By interpolating these three components over frequency, the GSM can therefore model the diffuse Galactic emission anywhere between 10 MHz and 100 GHz.

The considerable flexibility and power of the GSM software has resulted in its widespread application. Given that foregrounds are not only ubiquitous in cosmological surveys, but also interesting probes of astrophysical phenomena in their own right, it is unsurprising that the GSM has had broad influence within the astrophysics community. Astrophysical research using the GSM has ranged, for example, from studies of inflation with the CMB (e.g. Armitage-Caplan et al. [6]) to interstellar turbulence (e.g. Iacobelli et al. [69]) and from EoR power spectrum limits (e.g. Ali et al. [4]) to the discovery of a Fast Radio Burst (e.g. Burke-Spolaor and Bannister [24]).

However, there are some notable limitations to the GSM. In terms of frequency coverage, there is a lack of direct observation in the EoR frequency range. The closest maps in frequency are a 45 MHz map with about 3.6° resolution [5, 92] and a 408 MHz map with $48'$ resolution [62, 63, 149], and for the high resolution GSM version, the entire model is locked to the 408 MHz map. In terms of sky coverage, the PCA algorithm is performed on a rather small region covered by all 11 maps, mostly in out-of-plane directions, so the resulting principal components may suffer from biased frequency dependencies skewed towards that of the small region covered. In terms of accuracy, each of the 11 data sets are normalized beforehand so that they are treated equally, regardless of their relative accuracy. Thus, the model may suffer unnecessarily from noise and systematics in the maps of lesser quality. Lastly, since the PCA produces orthogonal principal components and orthogonal principal maps, these components are not likely to correspond to actual physical processes such as synchrotron or dust emission. Actual emission mechanisms have similar spatial structures (such as strong emission in the Galactic plane) and thus are not mutually orthogonal.

In this work, we present a new GSM-building method that naturally extends the original PCA algorithm. We use PCA as the initial step to obtain crude estimates of the principal components and their corresponding maps, and we iterate between the components and the maps to find the best fit to all of the data available. This method allows us to include 29 sky maps in the frequency range 10 MHz to 5 THz (including the Parkes maps at 85 MHz and 150 MHz [80]), which share no completely common sky coverage. This method can be further extended in the future to incorporate error information from each map into the fitting process. Furthermore, we use a blind component separation technique to recombine the orthogonal components into ones that are physically interpretable as different emission mechanisms. Both our blind spectra and blind component maps agree remarkably well with existing physical models.

The remainder of this paper is structured as follows. In Section 5.2, we describe our improved GSM-building method. In Section 5.3, we describe the 29 sky maps included in this work. Lastly, in Section 5.4, we present the improved GSM from 10 MHz to 5 THz with six components in Section 5.4.1, estimate the new GSM's predictive accuracy in Section 5.4.2, use a blind approach to recombine those components into physically meaningful contributions in Section 5.4.3, and compare our blind spectra and component maps with existing results in the literature in Sec. 5.4.4 and 5.4.5.

5.2 Iterative Algorithm for Building a Global Sky Model

5.2.1 Framework

We describe the Global Sky Model in the form of two matrices, an $n_{\text{pix}} \times n_{\text{c}}$ map matrix $\mathbf{M}$ and an $n_{\text{c}} \times n_{\text{f}}$ normalized spectrum matrix $\mathbf{S}$, where n_{pix} is the number of pixels in each sky map, n_{c} is the number of components, and n_{f} is the number of frequencies for which we have maps. Furthermore, we encapsulate all the sky map data into an $n_{\text{pix}} \times n_{\text{f}}$ matrix $\mathbf{D}$, where all maps are normalized to the same level at each frequency.[1] The Global Sky Model then models the sky by

$$\mathbf{D} \approx \mathbf{MS}. \tag{5.1}$$

Thus to construct a Global Sky Model is to find the pair of $\mathbf{M}$ and $\mathbf{S}$ that minimizes the cost function

$$||\mathbf{D} - \mathbf{MS}||^2. \tag{5.2}$$

In general, because both components in the product $\mathbf{MS}$ are unknown, we have degeneracies in the form of an invertible $n_{\text{c}} \times n_{\text{c}}$ matrix $\mathbf{\Psi}$. For any solution $\mathbf{M}$ and $\mathbf{S}$, an alternative solution $\mathbf{M}' = \mathbf{M}\mathbf{\Psi}^{-1}$ and $\mathbf{S}' = \mathbf{\Psi}\mathbf{S}$ will produce an identical prediction for the data $\mathbf{D}$.

To use the GSM to predict sky maps at previously unmeasured frequencies or sky regions, one interpolates at the desired frequency both the normalized spectra in $\mathbf{S}$ and the overall normalization. For the rest of this work, we choose linear interpolation for the normalized spectra over the log of frequency, and linear interpolation for the log of normalization over the log of frequency. Because normalized spectra are interpolated linearly, the choice of degeneracy matrix $\mathbf{\Psi}$ will have no effect on the predictions of the GSM.

---

[1] Due to incomplete and different sky coverages between maps, a naive normalization will not weigh the maps properly. We discuss our normalization in more detail in Section 5.2.3.

5.2.2 PCA Algorithm

We first describe the original PCA algorithm in de Oliveira-Costa et al. [34], before generalizing it to our iterative algorithm. Given the data matrix **D**, one first performs an eigen-decomposition of $\mathbf{D}^t\mathbf{D}$:

$$\mathbf{D}^t\mathbf{D} = \mathbf{C}^t\mathbf{\Lambda}\mathbf{C}, \tag{5.3}$$

where **C** is an $n_\mathrm{f} \times n_\mathrm{f}$ unitary matrix with eigenvectors as its rows, and $\mathbf{\Lambda}$ is an $n_\mathrm{f} \times n_\mathrm{f}$ diagonal matrix with eigenvalues on its diagonal. If the sky can be described by n_c components where $n_\mathrm{c} < n_\mathrm{f}$, then $\mathbf{\Lambda}$ has only n_c non-zero eigenvalues on its diagonal[2], so

$$\mathbf{D}^t\mathbf{D} = \widetilde{\mathbf{C}}^t\widetilde{\mathbf{\Lambda}}\widetilde{\mathbf{C}}, \tag{5.4}$$

where $\widetilde{\mathbf{C}}$ and $\widetilde{\mathbf{\Lambda}}$ are the $n_\mathrm{c} \times n_\mathrm{c}$ part of their non-tilde counterparts corresponding to the non-zero eigenvalues. One then takes the principal components $\widetilde{\mathbf{C}}$ and solves for the best **M** that satisfies

$$\mathbf{M}\widetilde{\mathbf{C}} \approx \mathbf{D}. \tag{5.5}$$

Comparing Eq. (5.1) and Eq. (5.5), we see that the PCA algorithm obtains one solution to Eq. (5.1), with $\mathbf{S} = \widetilde{\mathbf{C}}$. However, in practice, the sky maps have various sky coverages with potentially very few overlapping pixels, so rather than using the full **D** for Eq. (5.3), the original GSM uses a $\mathbf{D}^*$ that consists of a subset of the rows in **D** that corresponds to the common pixels covered at all frequencies. Therefore, the process of obtaining $\widetilde{\mathbf{C}}$ is ignoring the majority of information in **D**, and the solutions are thus not minimizing the cost function in Eq. (5.2). Furthermore, if one wishes to include more data sets, such as we do in this work, there are no overlapping regions that cover all of the frequency range, so it is difficult to apply the original PCA algorithm as-is.

---

[2] In practice, no eigenvalues are perfect zeros, and we will discuss our choice of n_c in more details in Section 5.4.1.

5.2.3 Iterative Algorithm

To overcome the challenges discussed above, we extend the PCA algorithm by iterating on the results obtained from it. We start by temporarily excluding the data sets with smallest sky coverage one by one, until the remaining data sets have more than 5% common sky coverage. We then use the 5% common pixels to obtain n_c, $\mathbf{M}^{(0)}$, and $\widetilde{\mathbf{C}}$ following the PCA algorithm described in the previous section, where $\mathbf{M}^{(0)}$ denotes our starting $\mathbf{M}$ at the 0th iteration. Once we obtain $\mathbf{M}^{(0)}$, we fold the temporarily excluded frequencies back into $\mathbf{D}$ and start iterating. At the ith iteration we compute

$$\mathbf{S}^{(i)} = (1-\eta)\mathbf{S}^{(i-1)} + \eta(\mathbf{M}^{(i-1)t}\mathbf{M}^{(i-1)})^{-1}\mathbf{M}^{(i-1)t}\mathbf{D}, \tag{5.6}$$

$$\mathbf{M}^{(i)} = (1-\eta)\mathbf{M}^{(i-1)} + \eta(\mathbf{S}^{(i)}\mathbf{S}^{(i)t})^{-1}\mathbf{S}^{(i)}\mathbf{D}^t, \tag{5.7}$$

where $0 < \eta \leq 1$ is a step size. Intuitively, at each iteration we are computing the least square solution $\mathbf{S}^{(i)}$ to Eq. (5.1), treating $\mathbf{M}^{(i-1)}$ as the truth, and vice versa. We keep iterating until the cost function decreases by less than 0.01%.

Due to incomplete sky coverage, there are two further tweaks to the above equations. Firstly, we do not directly compute the above equations in matrix form. Rather, we compute $\mathbf{S}^{(i)}$ column by column, with each column representing one frequency, and we modify $\mathbf{M}^{(i-1)}$ to exclude the pixels not covered at that frequency. Similarly for Eq. (5.7), each column corresponds to a pixel, and we modify $\mathbf{S}^{(i)}$ to exclude frequencies that do not cover that pixel. Thus, at every iteration, the algorithm incorporates all pixel data in $\mathbf{D}$, regardless of the sky coverage at each frequency or the frequency coverage of each pixel. In addition, since the maps have different sky coverages, simply normalizing valid pixels in each map will over-weigh the maps covering low temperature out-of-plane regions compared to the ones that only cover the Galactic center. Thus, we re-normalize the data matrix $\mathbf{D}$ at every iteration by dividing each map by the norm (root-sum-square) of its best fit map $\mathbf{M}^{(i)}\mathbf{S}^{(i)}$.

5.2.4 Incorporating More General Data Formats

Eq. (5.1) assumes that all data sets contained in **D** are sky maps with the same pixelization and angular resolution, which is what we focus on in this work. However, this comes with the drawback of forcing us to work with the lowest common resolution for all maps, thus discarding the high frequency information in higher quality maps. Furthermore, many low-frequency interferometers such as MWA [172] and PAPER [125] have produced high quality data, but because they sample sparsely in the Fourier domain, it is challenging to reduce those measurements into pixelized sky maps. Fortunately, we can generalize our algorithm and allow all data sets that are linearly related to the pixelized sky maps. The data set at frequency index f can be described by

$$D_{if} = \sum_{jn} B^f_{ij} M_{jn} S_{nf}, \tag{5.8}$$

which is a generalized version of Eq. (5.1). At any given frequency index f, $\mathbf{B}^f$ is a known matrix describing the linear relationship between the pixelized sky map and the data set. For pixelized sky-maps, $\mathbf{B}^f$ is the identity matrix and Eq. (5.8) reduces to Eq. (5.1). For low resolution data sets, $\mathbf{B}^f$ contains the point spread function for all pixels in the sky, which is typically the antenna beam pattern. For drift-scanning interferometer data sets, $\mathbf{B}^f$ contains the primary beam pattern as well as the rotating fringe patterns for all baselines, and D_{fi} at f is a flattened list of visibilities over both time and baselines. Given **B** and **D**, one can iteratively solve for **S** and **M**. It is worth noting that while including the **B**-matrix does not significantly increase the amount of computations for obtaining **S** at each iteration, it makes the computation of **M** much more demanding. This is because one can no longer use Eq. (5.7) to compute **M** pixel by pixel, as the **B** mixes pixels in the sky. One now needs to operate matrices with $n_{\text{pix}} \times n_{\text{c}}$ rows rather than n_{c} rows as in Eq. (5.7). For a HEALPIX map [59] with $n_{\text{side}} = 64$ corresponding to pixel size of 1° and 6 principal components, the matrix size is about 3×10^5 on each side, which demands significant computing resources beyond the scope of this work.

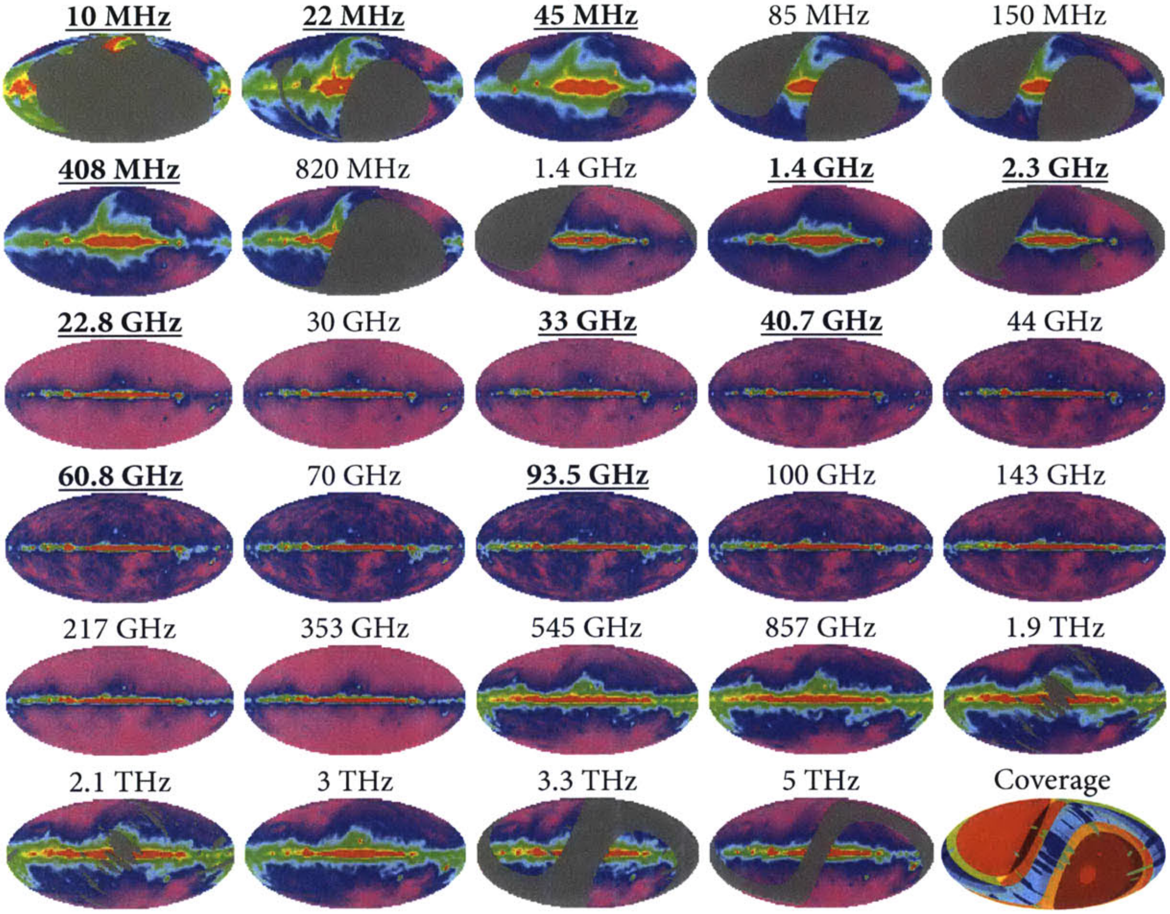

Figure 5-1: 29 sky maps used in this work from 10 MHz to 5 THz, plotted on arcsinh scale in Mollweide projection. The 11 bold and underscored frequencies are the ones included in the original GSM. The last panel shows the 120 different frequency coverage regions, each represented by a different color, and none of which contains all 29 frequencies. The overall amplitudes of the maps are shown in Fig. 5-2.

5.3 Sky Survey Data Sets

Table 5.1 and Fig. 5-1 lists all sky maps we use in this work, ranging from 10 MHz to 5 THz. We have included all sky maps in this frequency range that have angular resolutions better than 5° and substantial sky coverage. All data sets are publicly available online[3] except the ones at 45 MHz and 2.33 GHz. We manually mask out the ecliptic plane in the 3.33 THz AKARI data and the 5 THz IRIS data, and exclude IRIS data above 5 THz, in order to remove zodiacal contamination. Unlike

---

[3] MPIfR: `www3.mpifr-bonn.mpg.de/survey.html`,
LAMBDA: `lambda.gsfc.nasa.gov/product/`,
Planck: `pla.esac.esa.int/pla`.

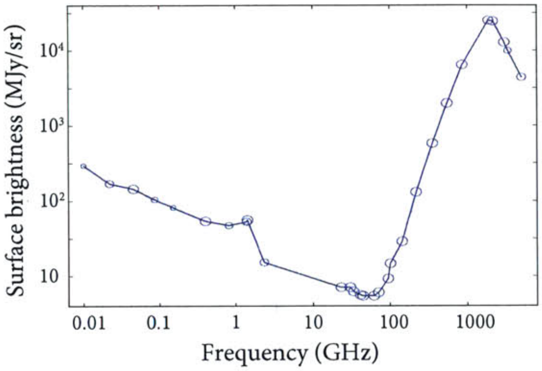

Figure 5-2: The overall amplitudes of the 29 sky maps included in this work. The sizes of the circles represent the sky coverage of each map. The amplitudes can be thought of as the root-mean-square of the input maps, with some iterative adjustments explained in Section 5.2.3.

the original GSM, which used physical models to pre-remove the CMB anisotropy from the WMAP and Plack maps, we choose to let the data speak for itself and not pre-remove the CMB anisotropy. We smooth all maps to 5°, and remove another 3° from edge areas in incomplete maps. All map are pixelized onto a HEALPIX grid with $n_{\text{side}} = 64$, and all units are converted to MJy/sr. These 29 maps form regions with 120 different combinations of map overlap, with no common region between all maps (see the bottom right plot in Fig. 5-1), so the improved GSM-making method is necessary to combine all these data sets.

5.4 Results: the Improved Global Sky Model

5.4.1 Orthogonal Components Result

We apply the algorithms described in Section 5.2 on maps described in Section 5.3. To obtain the initial 5% common coverage, maps at 10 MHz, 85 MHz, 150 MHz, and 5 THz are temporarily excluded. In the iteration process, it takes 12 iterations with step size $\eta = 1$ to converge. We also tried a smaller $\eta = 0.2$, and obtained the same result with 4 times more iterations. After the first convergence, we mask out the 1% of the pixels with the highest errors in order to eliminate point sources, and re-iterate for another 8 iterations to arrive at the final convergence. The new GSM

Project/Instrument	ν(GHz)	Area	Resolution	Ref.
DRAO, CAN	0.01	N	$2.6 \times 1.9^\circ$	[28]
DRAO, CAN	0.022	N	$1.1 \times 1.7^\circ$	[152]
MRAO+JMUAR	0.045	F	$\sim 3.6^\circ$	[5, 92]
Parkes	0.085	E	$3.8 \times 3.5^\circ$	[80]
Parkes	0.15	E	2.2°	[80]
GER, AUS, ENG	0.408	F	$48'$	[62, 63, 149]
Dwingeloo, NLD	0.82	N	1.2°	[10]
CHIPASS	1.39	S	$14.4'$	[26]
Stokert, Villa Elisa	1.42	F	$36'$	[148, 146, 147]
Rhodes/HartRAO	2.33	S	$20'$	[78]
WMAP	22.8	F	$49'$	[66]
Planck	30	F	$32'$	[133]
WMAP	33	F	$37'$	[66]
WMAP	40.7	F	$29'$	[66]
Planck	44	F	$24'$	[133]
WMAP	60.8	F	$20'$	[66]
Planck	70	F	$14'$	[133]
WMAP	93.5	F	$13'$	[66]
Planck	100	F	$10'$	[133]
Planck	143	F	$7'$	[133]
Planck	217	F	$5'$	[133]
Planck	353	F	$5'$	[133]
Planck	545	F	$5'$	[133]
Planck	857	F	$5'$	[133]
AKARI	1875	P	$1.5'$	[41]
AKARI	2143	P	$1.5'$	[41]
IRAS (IRIS)	3000	F	$4.3'$	[103]
AKARI	3333	P	$1.5'$	[41]
IRAS (IRIS)	5000	P	$4'$	[103]

Table 5.1: List of sky maps we use in our multi-frequency modeling. F: full sky; S: southern sky; N: northern sky; E: equatorial plane; P: partial sky. CHIPASS at 1.39 GHzhas a bandwidth of 64 MHz, so its frequency largely overlaps with the Stokert+Villa Elisa map at 1.42 GHz.

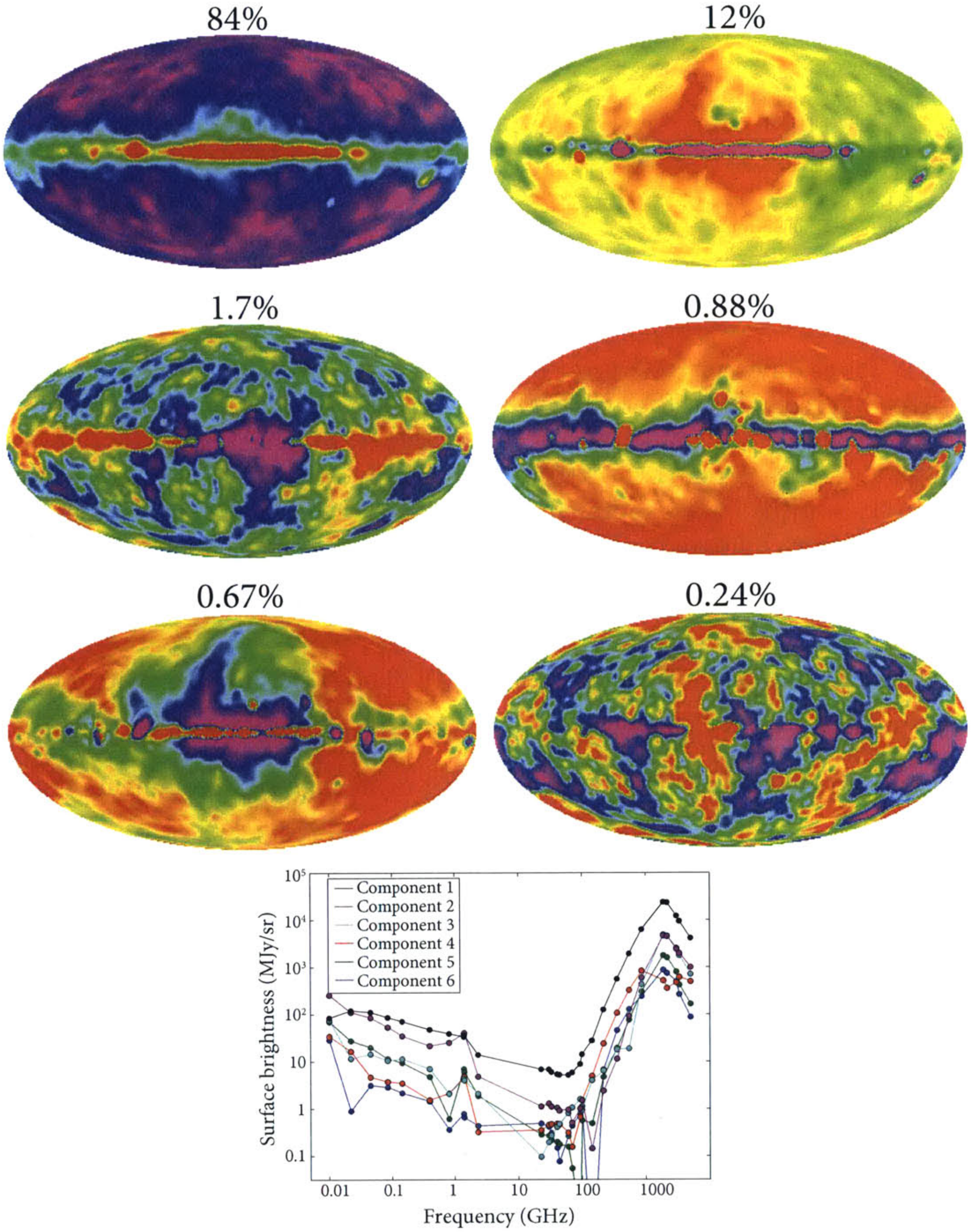

Figure 5-3: The 6 orthogonal components (top) and their spectra (bottom). The maps are plotted on arcsinh scales. The percentage on each component represents the fraction of all variations among the 29 input maps explained by that component. Here the spectra are not directly **S**, but the rows of **S** multiplied by the overall amplitudes of the input maps shown in Fig. 5-2.

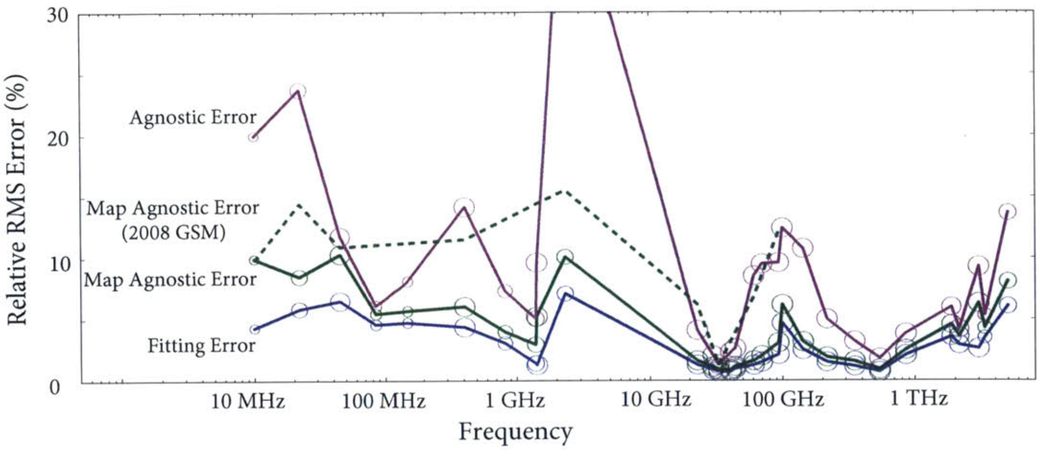

Figure 5-4: Three different RMS error percentage estimations for our GSM, compared to the error quoted with the original GSM. Agnostic error is an overly pessimistic upper bound to the GSM's predictive accuracy, whereas the map agnostic error is the lower bound. Fitting error is the fitting residual in each map not explained by the GSM. The sizes of the circles represent the sky coverage of each map. The improved GSM has less error compared to the original GSM across the entire frequency range.

with six components is shown in Fig. 5-3. We choose the number of components to be six based on the predictive power of the resulting GSM, which we discuss in more detail in Section 5.4.2. We find that using more than six components does not noticeably increase the predictive power of the resulting GSM. We plan to perform a more rigorous model selection to determine the optimal number of components in a future work that incorporate error information from each input map.

5.4.2 Error Analysis

To assess the predictive power of the GSM, we compute three types of Root Mean Square (RMS) errors at each frequency. The first is the fitting residual, which is the difference between the complete GSM model and the input sky maps. This fitting error is typically 1% for CMB frequencies and around 5% everywhere else, as shown in blue in Fig. 5-4. This is a combination of the measurement errors in the input maps and weak emission mechanisms comparable to such errors, which we are unable to capture. The fitting error represent the lower bound of the predictive accuracy of our GSM.

The second type of error is the map "agnostic error", where at a given frequency, we use the same normalized spectra but do not use the input map at that frequency to calculate the principal component maps. We then calculate the RMS of the difference between the resulting model and the actual sky map. As shown in green in Fig. 5-4, this error is typically 1.5 to 2 times larger than the fitting error. Among the three errors discussed in this section, this error is the most probable estimate of the predictive accuracy of the GSM, and is also used in the original GSM work.

The third type of error is a conservative upper-bound to the predictive accuracy of the GSM. Here we pretend to have zero knowledge of each map and exclude it from the algorithm from the outset, and produce a new GSM that is agnostic of that map. We then compare the excluded map with the predicted map using the new agnostic GSM, and calculate the RMS of the difference after an overall renormalization. The renormalization is typically less than 15%. This is an overly pessimistic estimate of the predictive power of the GSM, because we are not in the regime of an over-abundance of sky maps, especially at frequencies below the CMB. The only two full-sky maps below CMB (at 408 MHz and 1.4 GHz) serve as "cornerstones" of the GSM, so the agnostic error caused by removing them completely (and thus crippling the GSM) is not a good estimate of the complete GSM's predictive power at an unexplored frequency such as 300 MHz.

The peaks in the agnostic error curves that rise above 10% fall into three categories. The first category includes the ones near spectral lines, where the high errors are caused by having a spike at one end of the interpolation. One such example is the strongest peak at 2.3 GHz, whose error mostly comes from interpolating between the peak at 1.4 GHz and the lowest WMAP frequency. Another example is the peak around the 100 GHz CO line. The second category includes the ones at the edge of our frequency range, namely 10 MHz and 5 THz. They have high errors due to extrapolation. The third category includes ones that interpolate between incomplete and typically non-overlapping maps. One example is the 408 MHz Haslam map, which is interpolated between the 150 MHz equatorial map and the 820 MHz northern sky map. Another example is the 3 THz IRIS map that is interpolated between two AKARI maps with poor equatorial coverage.

In summary, for a map predicted by the GSM, we expect the RMS accuracy to be above the fitting error, and below the completely agnostic error. This is typically between 5% and 15% for most frequencies, and for CMB frequencies and frequencies near 500, GHz, the accuracy can be as good as 2%. In addition, we expect an overall amplitude offset of less than 15%. The improved GSM is seen to be consistently more accurate than the original GSM across the entire frequency range.

5.4.3 Blind Component Separation

Since the component maps we obtain (Fig. 5-3) are mutually orthogonal by construction, they are linear combinations of underlying emission mechanisms such as synchrotron and thermal dust. As discussed in Section 5.2.1, we have the freedom to apply any 6×6 invertible matrix $\boldsymbol{\Psi}$ and its inverse to the matrices $\mathbf{S}$ and $\mathbf{M}$ without changing our model, $\mathbf{MS}$. Applying $\boldsymbol{\Psi}$ will not change any linear interpolation results either. In this section we discuss our two-step automatic algorithm to determine $\boldsymbol{\Psi}$ and blindly extract the physical contributions. This algorithm focuses on manipulating the normalized spectrum matrix $\mathbf{S}$. The only two pieces of information that we use in order to determine $\boldsymbol{\Psi}$ are: 1) the absence of monopole in CMB; and 2) compact frequency support for each normalized spectrum, meaning that the each component dominates a limited frequency range. Our algorithm is thus blind to any existing models for known emission mechanisms.

We start with a $\boldsymbol{\Psi}$ that equals the identity matrix. The first step tries to search for the smallest frequency range outside which there is only five components, meaning that the sixth component must be limited to said frequency range. To do this, we enumerate all possible frequency ranges, and for each frequency range, we remove the columns in $\mathbf{S}$ that correspond to that range to form $\mathbf{S}^*$. We then calculate the eigenvalues and eigenvectors for $\mathbf{S}^*\mathbf{S}^{*t}$. For the smallest frequency range whose

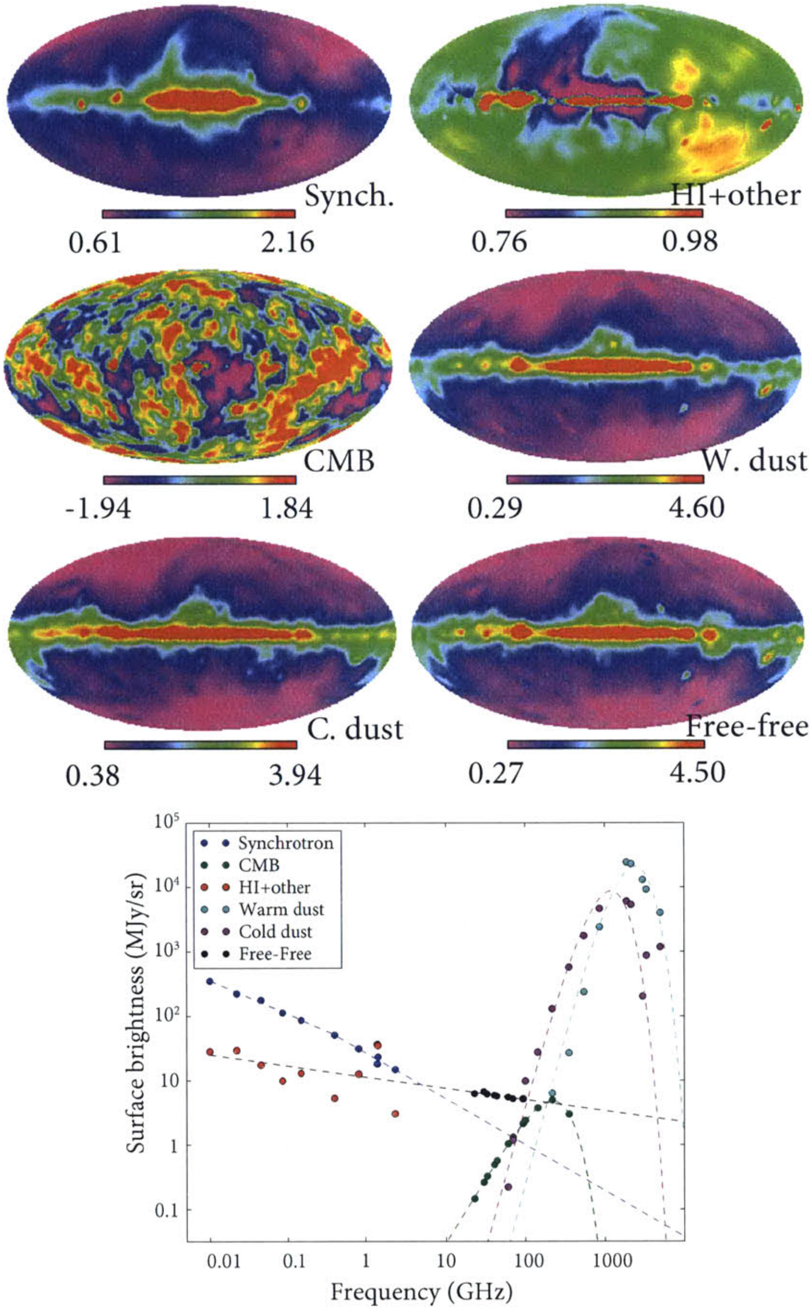

Figure 5-5: Top: six component maps from blind component separation. Bottom: six component spectra (solid points) and their best fits (dashed lines). For plotting purpose, each component map is normalized to have a median of 1 and plotted on arcsinh scales, where the color range corresponds to 2 and 98 percentiles in each map. In the bottom panel, we only plot the two most dominant component spectra at each frequency to reduce visual clutter. The best fit parameters are listed in Table 5.2.

smallest eigenmode explains less than 1% of all variations in $\mathbf{S}^*$,

$$\frac{\lambda_{\min}}{\mathrm{tr}(\mathbf{S}^*\mathbf{S}^{*t})} < 1\%, \tag{5.9}$$

we multiply $\boldsymbol{\Psi}$ by its eigenvectors $\mathbf{C}^*$. This way, the last row of $\boldsymbol{\Psi}\mathbf{S}$, which corresponds to the smallest eigenvalue, is our first separated component. If there is no frequency range that satisfies this criteria, we keep doubling the 1% threshold until one valid range appears. In practice, the synchrotron component is separated first with 2% threshold. We then repeat the above procedure, and separate the CMB component at 4%. Lastly, the "HI" component is separated at 8%. The first step ends since we cap the threshold at 10%.

Before we proceed to the second step, we attempt to clean the foreground contamination in the CMB component found in the first step. To do this we find the $\boldsymbol{\Psi}$ that minimizes the power in the CMB component map. This is the only procedure in the entire algorithm where we are working with the component map rather than normalized spectra, also the only procedure that use physical knowledge, namely the lack of monopole in the CMB.

The automatic procedure mentioned above is not able to separate the last three modes, which suggests that they cover largely overlapping frequency ranges. For our second step, we demand that the last three components each dominate one of the three frequency ranges: 10 GHz to 100 GHz, 100 GHz to 1 THz, and above 1 THz. To do this, we find the $\boldsymbol{\Psi}$ that minimizes each component's normalized spectrum outside its designated range. The results we obtain are shown in solid dots in Fig. 5-5.

5.4.4 Fitting the Blind Components

As shown in Fig. 5-5, five of the six components obtained in our blind separation can be related to known physical processes: synchrotron, free-free, CMB, and two thermal dust species at different temperatures. In this section, we take the known spectrum models of these mechanisms, such as a power law for synchrotron, and find the best parameters for them that fit our blind spectra. We first compute six spectra by taking the new normalized spectrum matrix $\mathbf{\Psi S}$ and multiplying back the normalization (shown in Fig. 5-2). We then fit each spectrum using its corresponding physical model: synchrotron and free-free as power laws with spectral indices β, thermal dust as blackbody spectra multiplied by power laws with temperatures T and spectral indices β, and CMB anisotropy as the first order Taylor expansion of the blackbody spectrum with temperature T. Note that we do not change our blind spectra based on these models.

The fits are performed in log(frequency)-log(surface brightness) space, and input variances are empirically estimated using the root mean square of best-fit residuals. We limit the fitting range for each component to the frequency range it dominates, as shown in Fig. 5-5. For power law fits, we perform a standard linear regression to obtain the best-fits and error bars. For the dusts and CMB components, we first numerically search for the best fit, and then perform Monte-Carlo on the input spectra using the empirical variances to estimate the error bars.

Although our component separation is blind to any of these model spectra, the best-fit parameters for all five of our blind spectra are seen to agree very well with existing literature values to within 2σ, as shown in Table 5.2. For all the non-CMB models, the error bars on our best-fit parameters are comparable, if not better, than the current best values in the literature. In addition to the spectra, their component maps also agree well with existing spatial templates, which we discuss in the next section.

5.4.5 High Resolution GSM

In addition to the 5° resolution GSM we have discussed, we also produce a high resolution version using the same normalized spectra as the 5° version. Thanks to the successful physical component separation as shown in Fig. 5-5, we are able to adopt two different resolutions throughout the frequency range. For frequencies above 10 GHz, we take those WMAP, Planck, AKARI, and IRIS maps whose resolution is better than 24 arcminutes, smooth them to 24 arcminutes, and solve for the CMB, the free-free, and the two dust component maps. Then, for frequencies below 10 GHz, we

Component	Parameter	Best-fit for blind spectra	Literature value	Ref.
Synchrotron	spectral index β below Haslam	-2.519 ± 0.018	-2.5 ± 0.1	[153]
	spectral index β above Haslam	-2.715 ± 0.082	-2.695 ± 0.120	[135]
Free-free	spectral index β	-2.175 ± 0.032	-2.1 ± 0.03	[159]
CMB	temperature T	$2.748 \pm 0.016\,\mathrm{K}$	$2.72548 \pm 0.00057\,\mathrm{K}$	[52]
Warm dust	temperature T	$20.0 \pm 3.3\,\mathrm{K}$	$16.2\,\mathrm{K}$	[51]
	spectral index β	2.78 ± 0.70	2.70	
Cold dust	temperature T	$10.4 \pm 1.2\,\mathrm{K}$	$9.4\,\mathrm{K}$	[51]
	spectral index β	2.54 ± 0.51	1.67	

Table 5.2: List of model parameters for our blind components compared to existing literature values. All parameters that fit our blind components agree with existing literature values within 2σ. The error bar in Platania et al. [135] represents spatial dispersion rather than statistical error.

take the 408 MHz and 1.42 GHz maps, smooth the latter to 48 arcminutes, smooth and remove the four component maps obtained in the high frequency band, and fit for the synchrotron and "HI" maps. Overall, the high resolution product has 48 arcminute angular resolution at frequencies below 10 GHz, and 24 arcminute angular resolution above 10 GHz, shown in Fig. 5-6. Note that while the high resolution component maps share the same spectra and large scale features as their low resolution counterparts, directly smoothing them to 5° will not produce exactly the low resolution versions, because the high resolution component maps are calculated using only a subset of the 29 input maps.

Our blind high resolution component maps agree remarkably well with existing maps and spatial templates. Our synchrotron, CMB anisotropy, free-free, and cold dust maps share most of the features seen in their counterparts in both the WMAP 9 year results and Planck 2015 results (see e.g. Fig. 19 in Bennett et al. [9] and Fig. 16 in Planck Collaboration et al. [133]). Free-free emission is known to closely trace Hα emission [50], and our free-free map shows all the key features visible in the composite Hα map presented in Finkbeiner [50]. Moreover, our cold dust map share all the key spatial features of the 94 GHz dust map presented in Finkbeiner et al. [51].

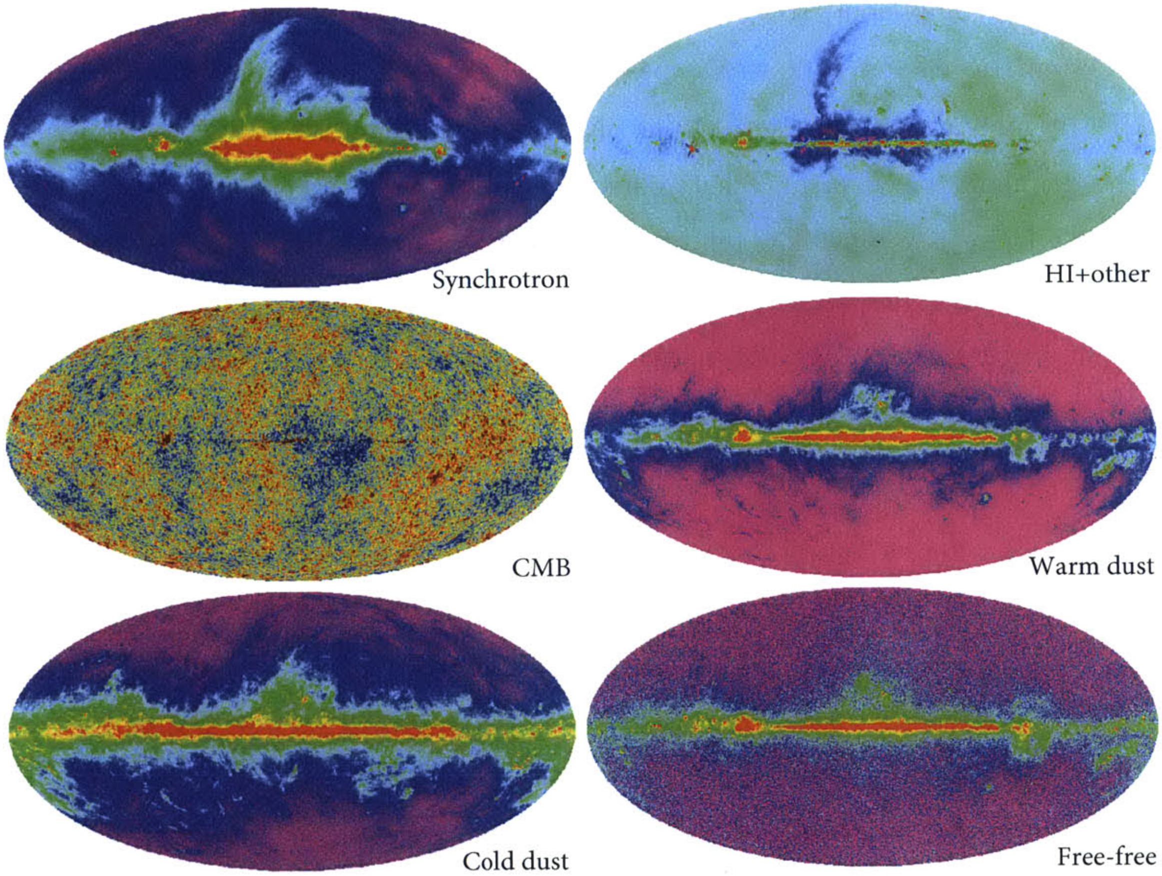

Figure 5-6: High resolution version of the six component maps. The five non-CMB maps follow the same color scales as those in Fig. 5-5, whereas the CMB component is plotted on a linear scale. The top two components have 48′ angular resolution, and the other four 24′.

5.4.6 Further Discussion

The blind spectra can be fit very well in the frequency ranges they dominate, as shown in Fig. 5-5 and Table 5.2. On the flip side, however, because our blind separation approach minimizes each component's normalized spectrum outside the frequency range it dominates, this approach cannot recover very well the parts of the spectrum outside its dominant range. One such example is the free-free spectrum, which is well fit by a power law with spectral index of -2.175 ± 0.032 near the CMB frequencies, but due to the minimization procedure, it largely disappears in the MHz range where synchrotron dominates. It is still contained in the GSM, but absorbed into the synchrotron and the HI+other components.

The sixth component, HI+other, is likely a combination of multiple mechanisms and systematic effects. Its strong peak near 1.4 GHz is well agreed by both the CHIPASS map and the Stockert+Villa Elisa map, and suggests local 21 cm emission as the dominant mechanism. In its component map, however, we notice two more features. There is a blue (low temperature) region whose shape resembles the synchrotron map, so this might indicate another weak synchrotron component with a different spectral index. We also notice a dipole component that is aligned with the equatorial poles, with striping artifacts especially visible near the northern equatorial pole, which suggests contribution from scanning systematics in the low frequency maps.

5.5 Summary and Outlook

We have presented an improved algorithm that builds upon the original PCA algorithm, allowing us to include many more incomplete sky maps across a larger frequency range. We have presented an improved GSM, with expected predictive accuracy between 5% and 15% for most frequencies, and around 2% near CMB frequencies, with overall amplitude offsets less than 15%. We have also presented a blind component separation technique, which identifies five physical components that agree very well with existing physical models. Lastly, we also create a high resolution GSM, with $48'$ resolution at frequencies below 10 GHz, and $24'$ resolution above 10 GHz.

There are two ways to further improve the GSM in the future: adding more data sets and improving the algorithm. On the data side, many high quality maps will become available in the near future, such as CBASS [79, 70], SPASS [27], GMIMS [186], and GEM [168]. Some of these upcoming maps will fill in the gap between 2.3 GHz and 23 GHz. In addition, in a separate paper [195], we present a new imaging method that allows existing low-frequency interferometers such as the MWA to produce high quality foreground maps that cover nearly the full sky. Last but not least, we aim to produce a polarized GSM using existing and upcoming polarized sky maps.

On the algorithmic side, we would like to include error information for all sky maps in the form of the maps' error covariance matrices into Eq. (5.6) and Eq. (5.7). This would allow us to properly weigh each map rather than resorting to renomalizing all maps and treating them all equally. In addition, this will allow us to perform rigorous model selection in order to decide the optimal number of principal components. Last but not least, we are interested in improving the blind physical component separation procedure to learn more about the physical mechanisms that correspond to the components found in the GSM.

With the help of upcoming high-quality sky maps, the improved GSM will not only serve as a powerful predictive model for diffuse Galactic emission, but will also have the potential to uncover new physical mechanisms that contribute to this emission.

Chapter 6

Conclusion

I began this thesis in Chapter 1 explaining the tremendous progress in precision cosmology made during the past two decades, and explained how 21 cm cosmology had the potential to provide even more knowledge of our universe in the decades to come. I then gave a brief introduction to traditional radio interferometry, which was being adapted by the 21 cm community for the new science objective of detecting redshifted 21 cm signals. I then explained the challenges the new field must overcome in order to make progress in the short term and reach the full potential of 21 cm tomography in the long term - challenges including calibration, foreground contamination, and computational cost.

In Part I, *Building MITEoR and OMNICAL*, I described my work leading the MIT effort to build a new radio interferometer: the MIT Epoch of Reionization (MITEoR) experiment. I demonstrated many new instrument design ideas and algorithms through this experiment, one of the most notable being redundant calibration algorithms. I showed that redundant calibration algorithms are able to rapidly perform calibration without the need of a sky model, and that they can achieve optimal precision (Fig. 2-11). I also found that the calibration parameters fluctuated on minute timescales (Fig. 2-12). This capability has influenced the design of the next generation 21 cm instrument HERA [60], which will have 384 dish elements arranged in a maximally redundant hexagonal layout and take full advantage of redundant calibration.

In Part II, *Latest Epoch of Reionization Science Results*, I applied the redundant calibration algorithms to the latest data at the time, collected by the Precision Array for Probing the Epoch of Reionization (PAPER), thereby dramatically improving the quality of the PAPER data. With the data quality improved by a combination of this calibration, other data processing enhancements, and more measurements, the PAPER collaboration obtained the most stringent upper limits to the 21 cm power spectrum to date, only one order of magnitude away from a detection (Fig. 3-18). With these upper bounds on the power spectrum, a crude constraint of spin temperature $T_S > 4\,\mathrm{K}$ at $z = 8.4$ was obtained, as shown in Fig. 3-21. Pober et al. [140] carried out a much more thorough study using these power spectrum upper bounds, and found for cold reionization a lower bound of $T_S > 10\,\mathrm{K}$ assuming Plancks reionization midpoint measurement [134] was current. This is only the beginning of 21 cm measurements teaching us about astrophysics, and with PAPERs latest 128 element data set and the HERA 19 element data set currently being analyzed, we are likely to make the first detection of 21 cm power spectrum very soon, and learn much more about both astrophysics and cosmology [93, 139, 88].

Finally, in Part III, *Novel Imaging and The New Global Sky Model*, I presented two new algorithms optimized for 21 cm cosmology. The first was a new imaging method that did not use the conventional Fourier approach. Rather, the new method took on Eq. (2.3) as a linear system of equations, and used precise mathematical tools developed by the CMB community to make high precision, large field of view images of the radio sky. I showed that this imaging method can produce high quality maps of the diffuse Galactic emission using both redundant arrays such as MITEoR and non-redundant array such as the MWA (Fig. 4-1). By applying this new method on the MITEoR data described in Part I, I produced the first Northern sky map at 150 MHz at about 2° resolution and about 15 K uncertainty (Fig. 4-4). I also calculated the averaged spectral index in the MITEoR frequency range to be -2.73 ± 0.11, which agrees well with the EDGES result of -2.5 ± 0.1 measured in a different sky region [153]. By comparing the MITEoR 150 MHz map to the Parkes map at 85 MHz [80] and the Haslam map at 408 MHz [62, 63, 149], I produced spectral index maps for the

Northern sky and found that the Galactic plane has a significantly shallower spectral index around -2.3 compared to out of plane regions with around -2.5. By applying this imaging method to future data sets from instruments such as MWA and HERA, we should be able to make very high quality foreground maps at the EoR frequency to help foreground removal efforts in 21 cm cosmology.

The second algorithm I built was an upgrade to the Global Sky Model (GSM) of diffuse galactic radio emission, which combined information from sky maps from many different frequencies, and whose uses included foreground modeling for CMB and 21cm cosmology. My new GSM algorithm improved both the precision and the flexibility of the original, allowing me to combine many more datasets in a more accurate fashion. The new GSM I obtained had a much wider frequency range and better precision across the entire frequency range (Fig. 5-4).

I also developed a blind method for extracting physical emission mechanisms from the GSM components, making no assumptions about physical emission mechanisms (synchrotron, free-free, dust, etc). Remarkably, this blind method automatically finds five components that have previously only been found "by hand", which we identify with synchrotron, free-free, cold dust, warm dust, and the CMB, with maps and spectra agreeing with previous work but in many cases with smaller error bars, as shown in Table 5.2 and Fig. 5-6. For example, I found the synchrotron spectral index to be -2.52 ± 0.02 and -2.72 ± 0.08 below and above the Haslam frequency, the free-free spectral index -2.18 ± 0.03, and two component dust temperatures $20.0 \pm 3.3\,\mathrm{K}$ and $10.4 \pm 1.2\,\mathrm{K}$. My blind method even independently rediscovered the CMB and constrained its temperature to be $2.75 \pm 0.02\,\mathrm{K}$. Upcoming high-quality data sets should enable future GSMs to not only serve as powerful predictive tools for unprobed frequencies, but also to uncover potential new physical emission mechanisms. The improved GSM is available online at `github.com/jeffzhen/gsm2016`.

I hope that my work in this thesis has helped move us closer to shedding light on the "dark ages" of our universe.

Bibliography

[1] M. Abramowitz and I.A. Stegun. *Handbook of mathematical functions: with formulas, graphs, and mathematical tables.* Applied mathematics series. Dover Publications, 1964. ISBN 9780486612720. URL `http://books.google.com/books?id=hLqv8pHoO0AC`.

[2] P. A. R. Ade, R. W. Aikin, D. Barkats, S. J. Benton, C. A. Bischoff, J. J. Bock, J. A. Brevik, I. Buder, E. Bullock, C. D. Dowell, L. Duband, J. P. Filippini, S. Fliescher, S. R. Golwala, and M. Halpern. Detection of B-Mode Polarization at Degree Angular Scales by BICEP2. *Physical Review Letters*, 112(24):241101, June 2014. doi: 10.1103/PhysRevLett.112.241101.

[3] S. S. Ali, S. Bharadwaj, and J. N. Chengalur. Foregrounds for redshifted 21-cm studies of reionization: Giant Meter Wave Radio Telescope 153-MHz observations. MNRAS, 385:2166–2174, April 2008. doi: 10.1111/j.1365-2966.2008.12984.x.

[4] Z. S. Ali, A. R. Parsons, H. Zheng, J. C. Pober, A. Liu, J. E. Aguirre, R. F. Bradley, G. Bernardi, C. L. Carilli, C. Cheng, D. R. DeBoer, M. R. Dexter, J. Grobbelaar, J. Horrell, D. C. Jacobs, P. Klima, D. H. E. MacMahon, M. Maree, D. F. Moore, N. Razavi, I. I. Stefan, W. P. Walbrugh, and A. Walker. PAPER-64 Constraints on Reionization: The 21 cm Power Spectrum at z = 8.4. ApJ, 809:61, August 2015. doi: 10.1088/0004-637X/809/1/61.

[5] H. Alvarez, J. Aparici, J. May, and F. Olmos. A 45-MHz continuum survey of the southern hemisphere. A&AS, 124:315, August 1997. doi: 10.1051/aas:1997196.

[6] C. Armitage-Caplan, J. Dunkley, H. K. Eriksen, and C. Dickinson. Impact on the tensor-to-scalar ratio of incorrect Galactic foreground modelling. MNRAS, 424:1914–1924, August 2012. doi: 10.1111/j.1365-2966.2012.21314.x.

[7] N. Barry, B. Hazelton, I. Sullivan, M. F. Morales, and J. C. Pober. Calibration Requirements for Detecting the 21 cm Epoch of Reionization Power Spectrum and Implications for the SKA. *ArXiv e-prints*, March 2016.

[8] N. Battaglia, A. Natarajan, H. Trac, R. Cen, and A. Loeb. Reionization on Large Scales. III. Predictions for Low-l Cosmic Microwave Background Polariza-

tion and High-l Kinetic Sunyaev-Zel'dovich Observables. ApJ, 776:83, October 2013. doi: 10.1088/0004-637X/776/2/83.

[9] C. L. Bennett, D. Larson, J. L. Weiland, N. Jarosik, G. Hinshaw, N. Odegard, K. M. Smith, R. S. Hill, B. Gold, M. Halpern, E. Komatsu, M. R. Nolta, L. Page, D. N. Spergel, E. Wollack, J. Dunkley, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. L. Wright. Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results. ApJS, 208:20, October 2013. doi: 10.1088/0067-0049/208/2/20.

[10] E. M. Berkhuijsen. A Survey of the Continuum Radiation at 820 MHz between Declinations -7° and +85°. A Study of the Galactic Radiation and the Degree of Polarization with Special Reference to the Loops and Spurs. A&A, 14:359, October 1971.

[11] G. Bernardi, A. G. de Bruyn, M. A. Brentjens, B. Ciardi, G. Harker, V. Jelić, L. V. E. Koopmans, P. Labropoulos, A. Offringa, V. N. Pandey, J. Schaye, R. M. Thomas, S. Yatawatta, and S. Zaroubi. Foregrounds for observations of the cosmological 21 cm line. I. First Westerbork measurements of Galactic emission at 150 MHz in a low latitude field. A&A, 500:965–979, June 2009. doi: 10.1051/0004-6361/200911627.

[12] G. Bernardi, A. G. de Bruyn, G. Harker, M. A. Brentjens, B. Ciardi, V. Jelić, L. V. E. Koopmans, P. Labropoulos, A. Offringa, V. N. Pandey, J. Schaye, R. M. Thomas, S. Yatawatta, and S. Zaroubi. Foregrounds for observations of the cosmological 21 cm line. II. Westerbork observations of the fields around 3C 196 and the North Celestial Pole. A&A, 522:A67+, November 2010. doi: 10.1051/0004-6361/200913420.

[13] G. Bernardi, M. McQuinn, and L. J. Greenhill. Foreground Model and Antenna Calibration Errors in the Measurement of the Sky-averaged λ21 cm Signal at z$\approx$ 20. ApJ, 799:90, January 2015. doi: 10.1088/0004-637X/799/1/90.

[14] J. S. Bolton, M. G. Haehnelt, S. J. Warren, P. C. Hewett, D. J. Mortlock, B. P. Venemans, R. G. McMahon, and C. Simpson. How neutral is the intergalactic medium surrounding the redshift z = 7.085 quasar ULAS J1120+0641? MNRAS, 416:L70–L74, September 2011. doi: 10.1111/j.1745-3933.2011.01100.x.

[15] S. E. I. Bosman and G. D. Becker. Re-examining the case for neutral gas near the redshift 7 quasar ULAS J1120+0641. *ArXiv e-prints: 1505.06880*, May 2015.

[16] R. J. Bouwens, G. D. Illingworth, P. A. Oesch, M. Stiavelli, P. van Dokkum, M. Trenti, D. Magee, I. Labbé, M. Franx, C. M. Carollo, and V. Gonzalez. Discovery of z $\approx$ 8 Galaxies in the Hubble Ultra Deep Field from Ultra-Deep WFC3/IR Observations. ApJ, 709:L133–L137, February 2010. doi: 10.1088/2041-8205/709/2/L133.

[17] R. J. Bouwens, L. Bradley, A. Zitrin, D. Coe, M. Franx, W. Zheng, R. Smit, O. Host, M. Postman, L. Moustakas, I. Labbé, M. Carrasco, A. Molino, M. Donahue, D. D. Kelson, M. Meneghetti, N. Benítez, D. Lemze, K. Umetsu, T. Broadhurst, J. Moustakas, P. Rosati, S. Jouvel, M. Bartelmann, H. Ford, G. Graves, C. Grillo, L. Infante, Y. Jimenez-Teja, O. Lahav, D. Maoz, E. Medezinski, P. Melchior, J. Merten, M. Nonino, S. Ogaz, and S. Seitz. A Census of Star-forming Galaxies in the Z $\approx$ 9-10 Universe based on HST+Spitzer Observations over 19 Clash Clusters: Three Candidate Z $\approx$ 9-10 Galaxies and Improved Constraints on the Star Formation Rate Density at Z $\approx$ 9.2. ApJ, 795:126, November 2014. doi: 10.1088/0004-637X/795/2/126.

[18] J. Bowman and A. E. E. Rogers. VHF-band RFI in Geographically Remote Areas. In *RFI Mitigation Workshop*, May 2010.

[19] J. D. Bowman and A. E. E. Rogers. A lower limit of $\Delta z>0.06$ for the duration of the reionization epoch. Nature, 468:796–798, December 2010. doi: 10.1038/nature09601.

[20] J. D. Bowman, M. F. Morales, and J. N. Hewitt. The Sensitivity of First-Generation Epoch of Reionization Observatories and Their Potential for Differentiating Theoretical Power Spectra. *ApJ*, 638:20–26, February 2006. doi: 10.1086/498703.

[21] 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. *ApJ*, 676:1–9, March 2008. doi: 10.1086/528675.

[22] J. D. Bowman, M. F. Morales, and J. N. Hewitt. Foreground Contamination in Interferometric Measurements of the Redshifted 21 cm Power Spectrum. ApJ, 695:183–199, April 2009. doi: 10.1088/0004-637X/695/1/183.

[23] R. Braun. Understanding synthesis imaging dynamic range. A&A, 551:A91, March 2013. doi: 10.1051/0004-6361/201220257.

[24] S. Burke-Spolaor and K. W. Bannister. The Galactic Position Dependence of Fast Radio Bursts and the Discovery of FRB011025. ApJ, 792:19, September 2014. doi: 10.1088/0004-637X/792/1/19.

[25] 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, February 2012. doi: 10.1016/j.asr.2011.10.014.

[26] M. R. Calabretta, L. Staveley-Smith, and D. G. Barnes. A New 1.4 GHz Radio Continuum Map of the Sky South of Declination +25°. PASA, 31:e007, January 2014. doi: 10.1017/pasa.2013.36.

[27] E. Carretti, B. Gaensler, L. Staveley-Smith, M. Haverkorn, M. Kesteven, S. Cortiglioni, G. Bernardi, and S. Poppi. S-band Polarization All Sky Survey (S-PASS). ATNF Proposal, April 2009.

[28] J. L. Caswell. A map of the northern sky at 10 MHz. MNRAS, 177:601–616, December 1976.

[29] E. Chapman, F. B. Abdalla, J. Bobin, J.-L. Starck, G. Harker, V. Jelić, P. Labropoulos, S. Zaroubi, M. A. Brentjens, A. G. de Bruyn, and L. V. E. Koopmans. The scale of the problem: recovering images of reionization with Generalized Morphological Component Analysis. MNRAS, 429:165–176, February 2013. doi: 10.1093/mnras/sts333.

[30] M.A. Clark, PC La Plante, and L.J. Greenhill. Accelerating radio astronomy cross-correlation with graphics processing units. *International Journal of High Performance Computing Applications*, 27(2):178–192, 2013. doi: 10.1177/1094342012444794. URL `http://hpc.sagepub.com/content/27/2/178.abstract`.

[31] Edward Collett. *Field guide to polarization*. SPIE Press, Bellingham, Wash, 2005. ISBN 0819458686.

[32] A. Datta, J. D. Bowman, and C. L. Carilli. Bright Source Subtraction Requirements for Redshifted 21 cm Measurements. ApJ, 724:526–538, November 2010. doi: 10.1088/0004-637X/724/1/526.

[33] A. de Oliveira-Costa and M. Tegmark, editors. *Microwave Foregrounds*, volume 181 of *Astronomical Society of the Pacific Conference Series*, 1999.

[34] 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. MNRAS, 388:247–260, July 2008. doi: 10.1111/j.1365-2966.2008.13376.x.

[35] J. S. Dillon. It's Always Darkest Before the Cosmic Dawn: Early Results from Novel Tools and Telescopes for 21 cm Cosmology. *ArXiv e-prints*, June 2015.

[36] J. S. Dillon and A. R. Parsons. Redundant Array Configurations for 21 cm Cosmology. *ArXiv e-prints*, February 2016.

[37] J. S. Dillon, A. Liu, and M. Tegmark. A fast method for power spectrum and foreground analysis for 21 cm cosmology. Phys. Rev. D, 87(4):043005, February 2013. doi: 10.1103/PhysRevD.87.043005.

[38] J. S. Dillon, A. Liu, C. L. Williams, J. N. Hewitt, M. Tegmark, E. H. Morgan, A. M. Levine, M. F. Morales, S. J. Tingay, G. Bernardi, J. D. Bowman, F. H. Briggs, R. C. Cappallo, D. Emrich, D. A. Mitchell, D. Oberoi, T. Prabu, R. Wayth, and R. L. Webster. Overcoming real-world obstacles in 21 cm power

spectrum estimation: A method demonstration and results from early Murchison Widefield Array data. Phys. Rev. D, 89(2):023002, January 2014. doi: 10.1103/PhysRevD.89.023002.

[39] J. S. Dillon, A. R. Neben, J. N. Hewitt, M. Tegmark, N. Barry, A. P. Beardsley, J. D. Bowman, F. Briggs, P. Carroll, A. de Oliveira-Costa, A. Ewall-Wice, L. Feng, L. J. Greenhill, B. J. Hazelton, L. Hernquist, N. Hurley-Walker, D. C. Jacobs, H. S. Kim, P. Kittiwisit, E. Lenc, J. Line, A. Loeb, B. McKinley, D. A. Mitchell, M. F. Morales, A. R. Offringa, S. Paul, B. Pindor, J. C. Pober, P. Procopio, J. Riding, S. Sethi, N. Udaya Shankar, R. Subrahmanyan, I. Sullivan, N. Thyagarajan, S. J. Tingay, C. Trott, R. B. Wayth, R. L. Webster, S. Wyithe, G. Bernardi, R. J. Cappallo, A. A. Deshpande, M. Johnston-Hollitt, D. L. Kaplan, C. J. Lonsdale, S. R. McWhirter, E. Morgan, D. Oberoi, S. M. Ord, T. Prabu, K. S. Srivani, A. Williams, and C. L. Williams. Empirical Covariance Modeling for 21 cm Power Spectrum Estimation: A Method Demonstration and New Limits from Early Murchison Widefield Array 128-Tile Data. *ArXiv e-prints: 1506.01026*, June 2015.

[40] J. S. Dillon, M. Tegmark, A. Liu, A. Ewall-Wice, J. N. Hewitt, M. F. Morales, A. R. Neben, A. R. Parsons, and H. Zheng. Mapmaking for precision 21 cm cosmology. Phys. Rev. D, 91(2):023002, January 2015. doi: 10.1103/PhysRevD.91.023002.

[41] Y. Doi, S. Takita, T. Ootsubo, K. Arimatsu, M. Tanaka, Y. Kitamura, M. Kawada, S. Matsuura, T. Nakagawa, T. Morishima, M. Hattori, S. Komugi, G. J. White, N. Ikeda, D. Kato, Y. Chinone, M. Etxaluze, and E. F. Cypriano. The AKARI far-infrared all-sky survey maps. PASJ, 67:50, June 2015. doi: 10.1093/pasj/psv022.

[42] S. W. Ellingson, G. B. Taylor, J. Craig, J. Hartman, J. Dowell, C. N. Wolfe, T. E. Clarke, B. C. Hicks, N. E. Kassim, P. S. Ray, L. J. Rickard, F. K. Schinzel, and K. W. Weiler. The LWA1 Radio Telescope. *IEEE Transactions on Antennas and Propagation*, 61:2540–2549, May 2013. doi: 10.1109/TAP.2013.2242826.

[43] R. S. Ellis, R. J. McLure, J. S. Dunlop, B. E. Robertson, Y. Ono, M. A. Schenker, A. Koekemoer, R. A. A. Bowler, M. Ouchi, A. B. Rogers, E. Curtis-Lake, E. Schneider, S. Charlot, D. P. Stark, S. R. Furlanetto, and M. Cirasuolo. The Abundance of Star-forming Galaxies in the Redshift Range 8.5-12: New Results from the 2012 Hubble Ultra Deep Field Campaign. ApJ, 763:L7, January 2013. doi: 10.1088/2041-8205/763/1/L7.

[44] H. I. Ewen and E. M. Purcell. Observation of a line in the galactic radio spectrum: Radiation from galactic hydrogen at 1,420 mc./sec. *Nature*, 168 (4270):356, September 1951. doi: 10.1038/168356a0. URL `http://dx.doi.org/10.1038/168356a0`.

[45] A. L. Faisst, P. Capak, C. M. Carollo, C. Scarlata, and N. Scoville. Spectroscopic Observation of Lyα Emitters at z $\approx$ 7.7 and Implications on Re-ionization. ApJ, 788:87, June 2014. doi: 10.1088/0004-637X/788/1/87.

[46] X. Fan, C. L. Carilli, and B. Keating. Observational Constraints on Cosmic Reionization. ARA&A, 44:415–462, September 2006. doi: 10.1146/annurev.astro.44.051905.092514PDF:http://arjournals.annualreviews.org/doi/pdf/10.1146/annurev.astro.44.051905.092514.

[47] A. Fialkov, R. Barkana, and E. Visbal. The observable signature of late heating of the Universe during cosmic reionization. Nature, 506:197–199, February 2014. doi: 10.1038/nature12999.

[48] G. B. Field. Excitation of the Hydrogen 21-CM Line. *Proceedings of the IRE*, 46:240–250, January 1958. doi: 10.1109/JRPROC.1958.286741.

[49] G. B. Field. *Proc. IRE*, 46:240, 1958.

[50] D. P. Finkbeiner. A Full-Sky Hα Template for Microwave Foreground Prediction. ApJS, 146:407–415, June 2003. doi: 10.1086/374411.

[51] D. P. Finkbeiner, M. Davis, and D. J. Schlegel. Extrapolation of Galactic Dust Emission at 100 Microns to Cosmic Microwave Background Radiation Frequencies Using FIRAS. ApJ, 524:867–886, October 1999. doi: 10.1086/307852.

[52] D. J. Fixsen. The Temperature of the Cosmic Microwave Background. ApJ, 707:916–920, December 2009. doi: 10.1088/0004-637X/707/2/916.

[53] G. Foster, J. Hickish, A. Magro, D. Price, and K. Zarb Adami. Implementation of a direct-imaging and FX correlator for the BEST-2 array. MNRAS, 439: 3180–3188, April 2014. doi: 10.1093/mnras/stu188.

[54] S. R. Furlanetto, S. P. Oh, and F. H. Briggs. Cosmology at low frequencies: The 21 cm transition and the high-redshift Universe. Phys. Rep., 433:181–301, October 2006. doi: 10.1016/j.physrep.2006.08.002.

[55] E. M. George, C. L. Reichardt, K. A. Aird, B. A. Benson, L. E. Bleem, J. E. Carlstrom, C. L. Chang, H. Cho, T. M. Crawford, A. T. Crites, T. de Haan, M. A. Dobbs, J. Dudley, N. W. Halverson, N. L. Harrington, G. P. Holder, W. L. Holzapfel, Z. Hou, J. D. Hrubes, R. Keisler, L. Knox, A. T. Lee, E. M. Leitch, M. Lueker, D. Luong-Van, J. J. McMahon, J. Mehl, S. S. Meyer, M. Millea, L. M. Mocanu, J. J. Mohr, T. E. Montroy, S. Padin, T. Plagge, C. Pryke, J. E. Ruhl, K. K. Schaffer, L. Shaw, E. Shirokoff, H. G. Spieler, Z. Staniszewski, A. A. Stark, K. T. Story, A. van Engelen, K. Vanderlinde, J. D. Vieira, R. Williamson, and O. Zahn. A measurement of secondary cosmic microwave background anisotropies from the 2500-square-degree SPT-SZ survey. *ArXiv e-prints: 1408.3161*, August 2014.

[56] A. Ghosh, S. Bharadwaj, S. S. Ali, and J. N. Chengalur. Improved foreground removal in GMRT 610 MHz observations towards redshifted 21-cm tomography. MNRAS, 418:2584–2589, December 2011. doi: 10.1111/j.1365-2966.2011.19649.x.

[57] A. Ghosh, L. V. E. Koopmans, E. Chapman, and V. Jelić. A Bayesian analysis of redshifted 21-cm H I signal and foregrounds: simulations for LOFAR. MNRAS, 452:1587–1600, September 2015. doi: 10.1093/mnras/stv1355.

[58] G. Giardino, A. J. Banday, K. M. Górski, K. Bennett, J. L. Jonas, and J. Tauber. Towards a model of full-sky Galactic synchrotron intensity and linear polarisation: A re-analysis of the Parkes data. A&A, 387:82–97, May 2002. doi: 10.1051/0004-6361:20020285.

[59] K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann. HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. ApJ, 622: 759–771, April 2005. doi: 10.1086/427976.

[60] L. J. Greenhill and G. Bernardi. HI Epoch of Reionization Arrays. *ArXiv e-prints*, January 2012.

[61] G. Harker, S. Zaroubi, G. Bernardi, M. A. Brentjens, A. G. de Bruyn, B. Ciardi, V. Jelić, L. V. E. Koopmans, P. Labropoulos, G. Mellema, A. Offringa, V. N. Pandey, J. Schaye, R. M. Thomas, and S. Yatawatta. Non-parametric foreground subtraction for 21-cm epoch of reionization experiments. MNRAS, 397:1138–1152, August 2009. doi: 10.1111/j.1365-2966.2009.15081.x.

[62] C. G. T. Haslam, U. Klein, C. J. Salter, H. Stoffel, W. E. Wilson, M. N. Cleary, D. J. Cooke, and P. Thomasson. A 408 MHz all-sky continuum survey. I - Observations at southern declinations and for the North Polar region. A&A, 100:209–219, July 1981.

[63] C. G. T. Haslam, C. J. Salter, H. Stoffel, and W. E. Wilson. A 408 MHz all-sky continuum survey. II - The atlas of contour maps. A&AS, 47:1, January 1982.

[64] B. J. Hazelton, M. F. Morales, and I. S. Sullivan. The Fundamental Multibaseline Mode-mixing Foreground in 21 cm Epoch of Reionization Observations. ApJ, 770:156, June 2013. doi: 10.1088/0004-637X/770/2/156.

[65] J. F. Helmboldt and N. E. Kassim. The Evolution of Cassiopeia A at Low Radio Frequencies. AJ, 138:838–844, September 2009. doi: 10.1088/0004-6256/138/3/838.

[66] G. Hinshaw, J. L. Weiland, R. S. Hill, N. Odegard, D. Larson, C. L. Bennett, J. Dunkley, B. Gold, M. R. Greason, N. Jarosik, E. Komatsu, M. R. Nolta, L. Page, D. N. Spergel, E. Wollack, M. Halpern, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, and E. L. Wright. Five-Year Wilkinson Microwave

Anisotropy Probe Observations: Data Processing, Sky Maps, and Basic Results. ApJS, 180:225–245, February 2009. doi: 10.1088/0067-0049/180/2/225.

[67] C. M. Hirata. Wouthuysen-Field coupling strength and application to high-redshift 21-cm radiation. MNRAS, 367:259–274, March 2006. doi: 10.1111/j.1365-2966.2005.09949.x.

[68] J. A. Högbom. Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines. A&AS, 15:417, June 1974.

[69] M. Iacobelli, M. Haverkorn, E. Orrú, R. F. Pizzo, J. Anderson, and et al. Studying Galactic interstellar turbulence through fluctuations in synchrotron emission. First LOFAR Galactic foreground detection. A&A, 558:A72, October 2013. doi: 10.1051/0004-6361/201322013.

[70] M. O. Irfan, C. Dickinson, R. D. Davies, C. Copley, R. J. Davis, P. G. Ferreira, C. M. Holler, J. L. Jonas, M. E. Jones, O. G. King, J. P. Leahy, J. Leech, E. M. Leitch, S. J. C. Muchovej, T. J. Pearson, M. W. Peel, A. C. S. Readhead, M. A. Stevenson, D. Sutton, A. C. Taylor, and J. Zuntz. C-Band All-Sky Survey: a first look at the Galaxy. MNRAS, 448:3572–3586, April 2015. doi: 10.1093/mnras/stv212.

[71] D. C. Jacobs, J. E. Aguirre, A. R. Parsons, J. C. Pober, R. F. Bradley, C. L. Carilli, N. E. Gugliucci, J. R. Manley, C. van der Merwe, D. F. Moore, and C. R. Parashare. New 145 MHz Source Measurements by PAPER in the Southern Sky. ApJ, 734:L34, June 2011. doi: 10.1088/2041-8205/734/2/L34.

[72] D. C. Jacobs, J. Bowman, and J. E. Aguirre. The Precision and Accuracy of Early Epoch of Reionization Foreground Models: Comparing MWA and PAPER 32-antenna Source Catalogs. ApJ, 769:5, May 2013. doi: 10.1088/0004-637X/769/1/5.

[73] D. C. Jacobs, 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, D. H. E. MacMahon, J. R. Manley, D. F. Moore, J. C. Pober, I. I. Stefan, and W. P. Walbrugh. A Flux Scale for Southern Hemisphere 21 cm Epoch of Reionization Experiments. ApJ, 776:108, October 2013. doi: 10.1088/0004-637X/776/2/108.

[74] D. C. Jacobs, J. C. Pober, A. R. Parsons, J. E. Aguirre, Z. S. 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. Multiredshift Limits on the 21 cm Power Spectrum from PAPER. ApJ, 801:51, March 2015. doi: 10.1088/0004-637X/801/1/51.

[75] V. Jelić, S. Zaroubi, P. Labropoulos, R. M. Thomas, G. Bernardi, M. A. Brentjens, A. G. de Bruyn, B. Ciardi, G. Harker, L. V. E. Koopmans, V. N. Pandey, J. Schaye, and S. Yatawatta. Foreground simulations for the LOFAR-epoch

of reionization experiment. *MNRAS*, 389:1319–1335, September 2008. doi: 10.1111/j.1365-2966.2008.13634.x.

[76] V. Jelić, S. Zaroubi, P. Labropoulos, G. Bernardi, A. G. de Bruyn, and L. V. E. Koopmans. Realistic simulations of the Galactic polarized foreground: consequences for 21-cm reionization detection experiments. MNRAS, 409:1647–1659, December 2010. doi: 10.1111/j.1365-2966.2010.17407.x.

[77] 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. A&A, 568:A101, August 2014. doi: 10.1051/0004-6361/201423998.

[78] J. L. Jonas, E. E. Baart, and G. D. Nicolson. The Rhodes/HartRAO 2326-MHz radio continuum survey. MNRAS, 297:977–989, July 1998. doi: 10.1046/j.1365-8711.1998.01367.x.

[79] Oliver G. King, Charles Copley, Rod Davies, Richard Davis, Clive Dickinson, Yaser A. Hafez, Christian Holler, Jaya John John, Justin L. Jonas, Michael E. Jones, J. Patrick Leahy, Stephen J. C. Muchovej, Timothy J. Pearson, Anthony C. S. Readhead, Matthew A. Stevenson, and Angela C. Taylor. The c-band all-sky survey: instrument design, status, and first-look data. volume 7741, pages 77411I–77411I–10, 2010. doi: 10.1117/12.858011. URL `http://dx.doi.org/10.1117/12.858011`.

[80] T. L. Landecker and R. Wielebinski. The Galactic Metre Wave Radiation: A two-frequency survey between declinations +25^o and -25^o and the preparation of a map of the whole sky. *Australian Journal of Physics Astrophysical Supplement*, 16:1, 1970.

[81] A. Lidz, O. Zahn, M. McQuinn, M. Zaldarriaga, and L. Hernquist. Detecting the Rise and Fall of 21 cm Fluctuations with the Murchison Widefield Array. ApJ, 680:962–974, June 2008. doi: 10.1086/587618.

[82] A. Lidz, O. Zahn, S. R. Furlanetto, M. McQuinn, L. Hernquist, and M. Zaldarriaga. Probing Reionization with the 21 cm Galaxy Cross-Power Spectrum. ApJ, 690:252–266, January 2009. doi: 10.1088/0004-637X/690/1/252.

[83] A. Liu and M. Tegmark. A method for 21 cm power spectrum estimation in the presence of foregrounds. Phys. Rev. D, 83(10):103006, May 2011. doi: 10.1103/PhysRevD.83.103006.

[84] A. Liu and M. Tegmark. How well can we measure and understand foregrounds with 21-cm experiments? MNRAS, 419:3491–3504, February 2012. doi: 10.1111/j.1365-2966.2011.19989.x.

[85] A. Liu, M. Tegmark, J. Bowman, J. Hewitt, and M. Zaldarriaga. An improved method for 21-cm foreground removal. MNRAS, 398:401–406, September 2009. doi: 10.1111/j.1365-2966.2009.15156.x.

[86] A. Liu, M. Tegmark, and M. Zaldarriaga. Will point sources spoil 21-cm tomography? MNRAS, 394:1575–1587, April 2009. doi: 10.1111/j.1365-2966.2009.14426.x.

[87] A. Liu, M. Tegmark, S. Morrison, A. Lutomirski, and M. Zaldarriaga. Precision calibration of radio interferometers using redundant baselines. MNRAS, 408:1029–1050, October 2010. doi: 10.1111/j.1365-2966.2010.17174.x.

[88] A. Liu, J. R. Pritchard, R. Allison, A. R. Parsons, U. Seljak, and B. D. Sherwin. Eliminating the optical depth nuisance from the CMB with 21 cm cosmology. Phys. Rev. D, 93(4):043013, February 2016. doi: 10.1103/PhysRevD.93.043013.

[89] Adrian Liu, Aaron R. Parsons, and Cathryn M. Trott. Epoch of reionization window. i. mathematical formalism. *Phys. Rev. D*, 90:023018, Jul 2014. doi: 10.1103/PhysRevD.90.023018. URL `http://link.aps.org/doi/10.1103/PhysRevD.90.023018`.

[90] Adrian Liu, Aaron R. Parsons, and Cathryn M. Trott. Epoch of reionization window. ii. statistical methods for foreground wedge reduction. *Phys. Rev. D*, 90:023019, Jul 2014. doi: 10.1103/PhysRevD.90.023019. URL `http://link.aps.org/doi/10.1103/PhysRevD.90.023019`.

[91] C. J. Lonsdale, R. J. Cappallo, M. F. Morales, F. H. Briggs, L. Benkevitch, J. D. Bowman, J. D. Bunton, S. Burns, B. E. Corey, L. Desouza, S. S. Doeleman, M. Derome, A. Deshpande, M. R. Gopala, L. J. Greenhill, D. E. Herne, J. N. Hewitt, and et al. The Murchison Widefield Array: Design Overview. *IEEE Proceedings*, 97:1497–1506, August 2009. doi: 10.1109/JPROC.2009.2017564.

[92] K. Maeda, H. Alvarez, J. Aparici, J. May, and P. Reich. A 45-MHz continuum survey of the northern hemisphere. A&AS, 140:145–154, December 1999. doi: 10.1051/aas:1999413.

[93] Y. Mao, M. Tegmark, M. McQuinn, M. Zaldarriaga, and O. Zahn. How accurately can 21cm tomography constrain cosmology? Phys. Rev. D, 78(2): 023529–+, July 2008. doi: 10.1103/PhysRevD.78.023529.

[94] V. R. Marthi and J. Chengalur. Non-linear redundancy calibration. MNRAS, 437:524–531, January 2014. doi: 10.1093/mnras/stt1902.

[95] 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 $\approx$ 0.8 in Cross-correlation. ApJ, 763:L20, January 2013. doi: 10.1088/2041-8205/763/1/L20.

[96] D. J. McLeod, R. J. McLure, J. S. Dunlop, B. E. Robertson, R. S. Ellis, and T. T. Targett. New redshift z $\approx$ 9 galaxies in the Hubble Frontier Fields: Implications for early evolution of the UV luminosity density. *ArXiv e-prints*, December 2014.

[97] M. McQuinn and R. M. O'Leary. The Impact of the Supersonic Baryon-Dark Matter Velocity Difference on the z $\approx$ 20 21 cm Background. ApJ, 760:3, November 2012. doi: 10.1088/0004-637X/760/1/3.

[98] M. McQuinn, O. Zahn, M. Zaldarriaga, L. Hernquist, and S. R. Furlanetto. Cosmological Parameter Estimation Using 21 cm Radiation from the Epoch of Reionization. *ApJ*, 653:815–834, December 2006. doi: 10.1086/505167.

[99] M. McQuinn, A. Lidz, O. Zahn, S. Dutta, L. Hernquist, and M. Zaldarriaga. The morphology of HII regions during reionization. MNRAS, 377:1043–1063, May 2007. doi: 10.1111/j.1365-2966.2007.11489.x.

[100] G. Mellema, L. V. E. Koopmans, F. A. Abdalla, G. Bernardi, B. Ciardi, S. Daiboo, A. G. de Bruyn, K. K. Datta, H. Falcke, A. Ferrara, I. T. Iliev, F. Iocco, V. Jelić, H. Jensen, R. Joseph, P. Labroupoulos, A. Meiksin, A. Mesinger, A. R. Offringa, V. N. Pandey, J. R. Pritchard, M. G. Santos, D. J. Schwarz, B. Semelin, H. Vedantham, S. Yatawatta, and S. Zaroubi. Reionization and the Cosmic Dawn with the Square Kilometre Array. *Experimental Astronomy*, 36: 235–318, August 2013. doi: 10.1007/s10686-013-9334-5.

[101] A. Mesinger, M. McQuinn, and D. N. Spergel. The kinetic Sunyaev-Zel'dovich signal from inhomogeneous reionization: a parameter space study. MNRAS, 422:1403–1417, May 2012. doi: 10.1111/j.1365-2966.2012.20713.x.

[102] A. Mesinger, A. Ferrara, and D. S. Spiegel. Signatures of X-rays in the early Universe. MNRAS, 431:621–637, May 2013. doi: 10.1093/mnras/stt198.

[103] M.-A. Miville-Deschênes and G. Lagache. IRIS: A New Generation of IRAS Maps. ApJS, 157:302–323, April 2005. doi: 10.1086/427938.

[104] D. C. Moore, J. Aguirre, A. Parsons, Z. Ali, R. F. Bradley, C. L. Carilli, D. R. DeBoer, M. Dexter, N. Gugliucci, D. C. Jacobs, P. Klima, A. Liu, D. MacMahon, J. Manley, J. Pober, I. Stefan, and W. Walbrugh. New Limits on Polarized Power Spectra at 126 and 164 MHz: Relevance to Epoch of Reionization Measurements. *ArXiv e-prints: 1502.05072*, February 2015.

[105] D. F. Moore, J. E. Aguirre, A. R. Parsons, D. C. Jacobs, and J. C. Pober. The Effects of Polarized Foregrounds on 21 cm Epoch of Reionization Power Spectrum Measurements. ApJ, 769:154, June 2013. doi: 10.1088/0004-637X/769/2/154.

[106] M. F. Morales. Power Spectrum Sensitivity and the Design of Epoch of Reionization Observatories. *ApJ*, 619:678–683, February 2005. doi: 10.1086/426730.

[107] M. F. Morales. Enabling Next-Generation Dark Energy and Epoch of Reionization Radio Observatories with the MOFF Correlator. PASP, 123:1265–1272, November 2011. doi: 10.1086/663092.

[108] M. F. Morales and J. S. B. Wyithe. Reionization and Cosmology with 21-cm Fluctuations. ARA&A, 48:127–171, September 2010. doi: 10.1146/annurev-astro-081309-130936.

[109] M. F. Morales, J. D. Bowman, and J. N. Hewitt. Improving Foreground Subtraction in Statistical Observations of 21 cm Emission from the Epoch of Reionization. ApJ, 648:767–773, September 2006. doi: 10.1086/506135.

[110] M. F. Morales, B. Hazelton, I. Sullivan, and A. Beardsley. Four Fundamental Foreground Power Spectrum Shapes for 21 cm Cosmology Observations. ApJ, 752:137, June 2012. doi: 10.1088/0004-637X/752/2/137.

[111] T. J. Mozdzen, J. D. Bowman, R. A. Monsalve, and A. E. E. Rogers. Limits on foreground subtraction from chromatic beam effects in global redshifted 21 cm measurements. MNRAS, 455:3890–3900, February 2016. doi: 10.1093/mnras/stv2601.

[112] P. Noorishad, S. J. Wijnholds, A. van Ardenne, and J. M. van der Hulst. Redundancy calibration of phased-array stations. A&A, 545:A108, September 2012. doi: 10.1051/0004-6361/201219087.

[113] P. A. Oesch, R. J. Bouwens, G. D. Illingworth, I. Labbé, M. Franx, P. G. van Dokkum, M. Trenti, M. Stiavelli, V. Gonzalez, and D. Magee. Probing the Dawn of Galaxies at z $\approx$ 9-12: New Constraints from HUDF12/XDF and CANDELS data. ApJ, 773:75, August 2013. doi: 10.1088/0004-637X/773/1/75.

[114] A. R. Offringa, B. McKinley, N. Hurley-Walker, F. H. Briggs, R. B. Wayth, D. L. Kaplan, M. E. Bell, L. Feng, A. R. Neben, J. D. Hughes, J. Rhee, T. Murphy, N. D. R. Bhat, G. Bernardi, J. D. Bowman, R. J. Cappallo, B. E. Corey, A. A. Deshpande, D. Emrich, A. Ewall-Wice, B. M. Gaensler, R. Goeke, L. J. Greenhill, B. J. Hazelton, L. Hindson, M. Johnston-Hollitt, D. C. Jacobs, J. C. Kasper, E. Kratzenberg, E. Lenc, C. J. Lonsdale, M. J. Lynch, S. R. McWhirter, D. A. Mitchell, M. F. Morales, E. Morgan, N. Kudryavtseva, D. Oberoi, S. M. Ord, B. Pindor, P. Procopio, T. Prabu, J. Riding, D. A. Roshi, N. U. Shankar, K. S. Srivani, R. Subrahmanyan, S. J. Tingay, M. Waterson, R. L. Webster, A. R. Whitney, A. Williams, and C. L. Williams. WSCLEAN: an implementation of a fast, generic wide-field imager for radio astronomy. MNRAS, 444: 606–619, October 2014. doi: 10.1093/mnras/stu1368.

[115] G. Paciga, T.-C. Chang, Y. Gupta, R. Nityanada, J. Odegova, U.-L. Pen, J. B. Peterson, J. Roy, and K. Sigurdson. The GMRT Epoch of Reionization experiment: a new upper limit on the neutral hydrogen power spectrum at $z \approx 8.6$. MNRAS, 413:1174–1183, May 2011. doi: 10.1111/j.1365-2966.2011.18208.x.

[116] 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. MNRAS, 433:639–647, July 2013. doi: 10.1093/mnras/stt753.

[117] L. Page, G. Hinshaw, E. Komatsu, M. R. Nolta, D. N. Spergel, C. L. Bennett, C. Barnes, R. Bean, O. Doré, J. Dunkley, M. Halpern, R. S. Hill, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, H. V. Peiris, G. S. Tucker, L. Verde, J. L. Weiland, E. Wollack, and E. L. Wright. Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Polarization Analysis. ApJS, 170:335–376, June 2007. doi: 10.1086/513699.

[118] H. Park, P. R. Shapiro, E. Komatsu, I. T. Iliev, K. Ahn, and G. Mellema. The Kinetic Sunyaev-Zel'dovich Effect as a Probe of the Physics of Cosmic Reionization: The Effect of Self-regulated Reionization. ApJ, 769:93, June 2013. doi: 10.1088/0004-637X/769/2/93.

[119] E. A. Parker. Precise measurement of the flux densities of the radio sources Cas A and Cyg A at metre wavelengths. MNRAS, 138:407, 1968. doi: 10.1093/mnras/138.4.407.

[120] A. Parsons. The symmetric group in data permutation, with applications to high-bandwidth pipelined fft architectures. *Signal Processing Letters, IEEE*, 16 (6):477–480, 2009. ISSN 1070-9908. doi: 10.1109/LSP.2009.2016836.

[121] A. Parsons, D. Backer, C. Chang, D. Chapman, H. Chen, P. Crescini, C. de Jesus, C. Dick, P. Droz, D. MacMahon, K. Meder, J. Mock, V. Nagpal, B. Nikolic, A. Parsa, B. Richards, A. Siemion, J. Wawrzynek, D. Werthimer,

and M. Wright. Petaop/second fpga signal processing for seti and radio astronomy. In *Signals, Systems and Computers, 2006. ACSSC '06. Fortieth Asilomar Conference on*, pages 2031–2035, 2006. doi: 10.1109/ACSSC.2006.355123.

[122] A. Parsons, D. Backer, A. Siemion, H. Chen, D. Werthimer, P. Droz, T. Filiba, J. Manley, P. McMahon, A. Parsa, D. MacMahon, and M. Wright. A Scalable Correlator Architecture Based on Modular FPGA Hardware, Reuseable Gateware, and Data Packetization. PASP, 120:1207–1221, November 2008. doi: 10.1086/593053.

[123] A. Parsons, J. Pober, M. McQuinn, D. Jacobs, and J. Aguirre. A Sensitivity and Array-configuration Study for Measuring the Power Spectrum of 21 cm Emission from Reionization. ApJ, 753:81, July 2012. doi: 10.1088/0004-637X/753/1/81.

[124] A. R. Parsons and D. C. Backer. Calibration of Low-Frequency, Wide-Field Radio Interferometers Using Delay/Delay-Rate Filtering. AJ, 138:219–226, July 2009. doi: 10.1088/0004-6256/138/1/219.

[125] A. R. Parsons, D. C. Backer, G. S. Foster, M. C. H. Wright, R. F. Bradley, N. E. Gugliucci, C. R. Parashare, E. E. Benoit, J. E. Aguirre, D. C. Jacobs, C. L. Carilli, D. Herne, M. J. Lynch, J. R. Manley, and D. J. Werthimer. The Precision Array for Probing the Epoch of Re-ionization: Eight Station Results. AJ, 139:1468–1480, April 2010. doi: 10.1088/0004-6256/139/4/1468.

[126] A. R. Parsons, J. C. Pober, J. E. Aguirre, C. L. Carilli, D. C. Jacobs, and D. F. Moore. A Per-baseline, Delay-spectrum Technique for Accessing the 21 cm Cosmic Reionization Signature. ApJ, 756:165, September 2012. doi: 10.1088/0004-637X/756/2/165.

[127] A. R. Parsons, A. Liu, J. E. Aguirre, Z. S. Ali, R. F. Bradley, C. L. Carilli, D. R. DeBoer, M. R. Dexter, N. E. Gugliucci, D. C. Jacobs, P. Klima, D. H. E. MacMahon, J. R. Manley, D. F. Moore, J. C. Pober, I. I. Stefan, and W. P. Walbrugh. New Limits on 21 cm Epoch of Reionization from PAPER-32 Consistent with an X-Ray Heated Intergalactic Medium at z = 7.7. ApJ, 788:106, June 2014. doi: 10.1088/0004-637X/788/2/106.

[128] A. R. Parsons, A. Liu, Z. S. Ali, and C. Cheng. Optimized Beam Sculpting with Generalized Fringe-Rate Filters. *ArXiv e-prints: 1503.05564*, March 2015.

[129] N. Patra, R. Subrahmanyan, S. Sethi, N. Udaya Shankar, and A. Raghunathan. Saras Measurement of the Radio Background At Long Wavelengths. ApJ, 801:138, March 2015. doi: 10.1088/0004-637X/801/2/138.

[130] J. B. Peterson, U.-L. Pen, and X.-P. Wu. The Primeval Structure Telescope. *Modern Physics Letters A*, 19:1001–1008, 2004. doi: 10.1142/S0217732304014288.

[131] N. Petrovic and S. P. Oh. Systematic effects of foreground removal in 21-cm surveys of reionization. MNRAS, 413:2103–2120, May 2011. doi: 10.1111/j.1365-2966.2011.18276.x.

[132] Planck Collaboration, P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al. Planck 2013 results. XVI. Cosmological parameters. A&A, 571: A16, November 2014. doi: 10.1051/0004-6361/201321591.

[133] Planck Collaboration, R. Adam, P. A. R. Ade, N. Aghanim, Y. Akrami, M. I. R. Alves, M. Arnaud, F. Arroja, J. Aumont, C. Baccigalupi, and et al. Planck 2015 results. I. Overview of products and scientific results. *ArXiv e-prints*, February 2015.

[134] Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al. Planck 2015 results. XIII. Cosmological parameters. *ArXiv e-prints*, February 2015.

[135] P. Platania, C. Burigana, D. Maino, E. Caserini, M. Bersanelli, B. Cappellini, and A. Mennella. Full sky study of diffuse Galactic emission at decimeter wavelengths. A&A, 410:847–863, November 2003. doi: 10.1051/0004-6361:20031125.

[136] J. C. Pober. The impact of foregrounds on redshift space distortion measurements with the highly redshifted 21-cm line. MNRAS, 447:1705–1712, February 2015. doi: 10.1093/mnras/stu2575.

[137] J. C. Pober, A. R. Parsons, D. C. Jacobs, J. E. Aguirre, R. F. Bradley, C. L. Carilli, N. E. Gugliucci, D. F. Moore, and C. R. Parashare. A Technique for Primary Beam Calibration of Drift-scanning, Wide-field Antenna Elements. AJ, 143:53, February 2012. doi: 10.1088/0004-6256/143/2/53.

[138] 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. ApJ, 768:L36, May 2013. doi: 10.1088/2041-8205/768/2/L36.

[139] J. C. Pober, A. Liu, J. S. Dillon, J. E. Aguirre, J. D. Bowman, R. F. Bradley, C. L. Carilli, D. R. DeBoer, J. N. Hewitt, D. C. Jacobs, M. McQuinn, M. F. Morales, A. R. Parsons, M. Tegmark, and D. J. Werthimer. What Next-generation 21 cm Power Spectrum Measurements can Teach us About the Epoch of Reionization. ApJ, 782:66, February 2014. doi: 10.1088/0004-637X/782/2/66.

[140] J. C. Pober, Z. S. Ali, A. R. Parsons, M. McQuinn, J. E. Aguirre, G. Bernardi, R. F. Bradley, C. L. Carilli, C. Cheng, D. R. DeBoer, M. R. Dexter, S. R.

Furlanetto, J. Grobbelaar, J. Horrell, D. C. Jacobs, P. J. Klima, S. A. Kohn, A. Liu, D. H. E. MacMahon, M. Maree, A. Mesinger, D. F. Moore, N. Razavi-Ghods, I. I. Stefan, W. P. Walbrugh, A. Walker, and H. Zheng. PAPER-64 Constraints On Reionization II: The Temperature Of The z=8.4 Intergalactic Medium. *ArXiv e-prints: 1503.00045*, February 2015.

[141] J. C. Pober, B. J. Hazelton, A. P. Beardsley, N. A. Barry, Z. E. Martinot, I. S. Sullivan, M. F. Morales, M. E. Bell, G. Bernardi, N. D. R. Bhat, J. D. Bowman, F. Briggs, R. J. Cappallo, P. Carroll, B. E. Corey, A. de Oliveira-Costa, A. A. Deshpande, J. S. Dillon, D. Emrich, A. M. Ewall-Wice, L. Feng, R. Goeke, L. J. Greenhill, J. N. Hewitt, L. Hindson, N. Hurley-Walker, D. C. Jacobs, M. Johnston-Hollitt, D. L. Kaplan, J. C. Kasper, H.-S. Kim, P. Kittiwisit, E. Kratzenberg, N. Kudryavtseva, E. Lenc, J. Line, A. Loeb, C. J. Lonsdale, M. J. Lynch, B. McKinley, S. R. McWhirter, D. A. Mitchell, E. Morgan, A. R. Neben, D. Oberoi, A. R. Offringa, S. M. Ord, S. Paul, B. Pindor, T. Prabu, P. Procopio, J. Riding, A. E. E. Rogers, A. Roshi, S. K. Sethi, N. Udaya Shankar, K. S. Srivani, R. Subrahmanyan, M. Tegmark, N. Thyagarajan, S. J. Tingay, C. M. Trott, M. Waterson, R. B. Wayth, R. L. Webster, A. R. Whitney, A. Williams, C. L. Williams, and J. S. B. Wyithe. The Importance of Wide-field Foreground Removal for 21 cm Cosmology: A Demonstration with Early MWA Epoch of Reionization Observations. ApJ, 819:8, March 2016. doi: 10.3847/0004-637X/819/1/8.

[142] M. Presley, A. Liu, and A. Parsons. Measuring the Cosmological 21 cm Monopole with an Interferometer. *ArXiv e-prints: 1501.01633*, January 2015.

[143] J. R. Pritchard and A. Loeb. Evolution of the 21cm signal throughout cosmic history. Phys. Rev. D, 78(10):103511, November 2008. doi: 10.1103/PhysRevD.78.103511.

[144] J. R. Pritchard and A. Loeb. 21 cm cosmology in the 21st century. *Reports on Progress in Physics*, 75(8):086901, August 2012. doi: 10.1088/0034-4885/75/8/086901.

[145] U. Rau, S. Bhatnagar, M. A. Voronkov, and T. J. Cornwell. Advances in Calibration and Imaging Techniques in Radio Interferometry. *IEEE Proceedings*, 97:1472–1481, August 2009. doi: 10.1109/JPROC.2009.2014853.

[146] P. Reich and W. Reich. A radio continuum survey of the northern sky at 1420 MHz. II. A&AS, 63:205–288, February 1986.

[147] P. Reich, J. C. Testori, and W. Reich. A radio continuum survey of the southern sky at 1420 MHz. The atlas of contour maps. A&A, 376:861–877, September 2001. doi: 10.1051/0004-6361:20011000.

[148] W. Reich. A radio continuum survey of the northern sky at 1420 MHz. I. A&AS, 48:219–297, May 1982.

[149] M. Remazeilles, C. Dickinson, A. J. Banday, M.-A. Bigot-Sazy, and T. Ghosh. An improved source-subtracted and destriped 408-MHz all-sky map. MNRAS, 451:4311–4327, August 2015. doi: 10.1093/mnras/stv1274.

[150] B. E. Robertson, S. R. Furlanetto, E. Schneider, S. Charlot, R. S. Ellis, D. P. Stark, R. J. McLure, J. S. Dunlop, A. Koekemoer, M. A. Schenker, M. Ouchi, Y. Ono, E. Curtis-Lake, A. B. Rogers, R. A. A. Bowler, and M. Cirasuolo. New Constraints on Cosmic Reionization from the 2012 Hubble Ultra Deep Field Campaign. ApJ, 768:71, May 2013. doi: 10.1088/0004-637X/768/1/71.

[151] B. E. Robertson, R. S. Ellis, S. R. Furlanetto, and J. S. Dunlop. Cosmic Reionization and Early Star-forming Galaxies: A Joint Analysis of New Constraints from Planck and the Hubble Space Telescope. ApJ, 802:L19, April 2015. doi: 10.1088/2041-8205/802/2/L19.

[152] R. S. Roger, C. H. Costain, T. L. Landecker, and C. M. Swerdlyk. The radio emission from the Galaxy at 22 MHz. A&AS, 137:7–19, May 1999. doi: 10.1051/aas:1999239.

[153] A. E. E. Rogers and J. D. Bowman. Spectral Index of the Diffuse Radio Background Measured from 100 to 200 MHz. AJ, 136:641–648, August 2008. doi: 10.1088/0004-6256/136/2/641.

[154] N. J. Sanchez. On the instrumentation of the Omniscope. Master's thesis, Massachusetts Institute of Technology, 2011. URL `http://hdl.handle.net/1721.1/66807`.

[155] M. G. Santos, A. Cooray, and L. Knox. Multifrequency Analysis of 21 Centimeter Fluctuations from the Era of Reionization. ApJ, 625:575–587, June 2005. doi: 10.1086/429857.

[156] M. A. Schenker, B. E. Robertson, R. S. Ellis, Y. Ono, R. J. McLure, J. S. Dunlop, A. Koekemoer, R. A. A. Bowler, M. Ouchi, E. Curtis-Lake, A. B. Rogers, E. Schneider, S. Charlot, D. P. Stark, S. R. Furlanetto, and M. Cirasuolo. The UV Luminosity Function of Star-forming Galaxies via Dropout Selection at Redshifts z $\approx$ 7 and 8 from the 2012 Ultra Deep Field Campaign. ApJ, 768:196, May 2013. doi: 10.1088/0004-637X/768/2/196.

[157] F. R. Schwab. Relaxing the isoplanatism assumption in self-calibration; applications to low-frequency radio interferometry. AJ, 89:1076–1081, July 1984. doi: 10.1086/113605.

[158] J. R. Shaw, K. Sigurdson, U.-L. Pen, A. Stebbins, and M. Sitwell. All-sky Interferometry with Spherical Harmonic Transit Telescopes. ApJ, 781:57, February 2014. doi: 10.1088/0004-637X/781/2/57.

[159] G. F. Smoot. Galactic Free-free and H-alpha Emission. *ArXiv Astrophysics e-prints*, January 1998.

[160] 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. PASA, 32:e004, February 2015. doi: 10.1017/pasa.2015.3.

[161] I. S. Sullivan, M. F. Morales, B. J. Hazelton, W. Arcus, D. Barnes, G. Bernardi, F. H. Briggs, J. D. Bowman, J. D. Bunton, R. J. Cappallo, B. E. Corey, A. Deshpande, L. deSouza, D. Emrich, B. M. Gaensler, R. Goeke, L. J. Greenhill, D. Herne, J. N. Hewitt, M. Johnston-Hollitt, D. L. Kaplan, J. C. Kasper, B. B. Kincaid, R. Koenig, E. Kratzenberg, C. J. Lonsdale, M. J. Lynch, S. R. McWhirter, D. A. Mitchell, E. Morgan, D. Oberoi, S. M. Ord, J. Pathikulangara, T. Prabu, R. A. Remillard, A. E. E. Rogers, A. Roshi, J. E. Salah, R. J. Sault, N. Udaya Shankar, K. S. Srivani, J. Stevens, R. Subrahmanyan, S. J. Tingay, R. B. Wayth, M. Waterson, R. L. Webster, A. R. Whitney, A. Williams, C. L. Williams, and J. S. B. Wyithe. Fast Holographic Deconvolution: A New Technique for Precision Radio Interferometry. ApJ, 759:17, November 2012. doi: 10.1088/0004-637X/759/1/17.

[162] Eric R. Switzer and Adrian Liu. Erasing the variable: Empirical foreground discovery for global 21 cm spectrum experiments. *The Astrophysical Journal*, 793(2):102, 2014. URL `http://stacks.iop.org/0004-637X/793/i=2/a=102`.

[163] G. B. Taylor, C. L. Carilli, and R. A. Perley, editors. *Synthesis Imaging in Radio Astronomy II*, volume 180 of *Astronomical Society of the Pacific Conference Series*, 1999.

[164] M. Tegmark. How to measure CMB power spectra without losing information. Phys. Rev. D, 55:5895–5907, May 1997. doi: 10.1103/PhysRevD.55.5895.

[165] M. Tegmark. How to Make Maps from Cosmic Microwave Background Data without Losing Information. ApJ, 480:L87–L90, May 1997. doi: 10.1086/310631.

[166] M. Tegmark and M. Zaldarriaga. Fast Fourier transform telescope. Phys. Rev. D, 79(8):083530, April 2009. doi: 10.1103/PhysRevD.79.083530.

[167] M. Tegmark and M. Zaldarriaga. Omniscopes: Large area telescope arrays with only N log N computational cost. Phys. Rev. D, 82(10):103501, November 2010. doi: 10.1103/PhysRevD.82.103501.

[168] C. Tello, T. Villela, S. Torres, M. Bersanelli, G. F. Smoot, I. S. Ferreira, A. Cingoz, J. Lamb, D. Barbosa, D. Perez-Becker, S. Ricciardi, J. A. Currivan, P. Platania, and D. Maino. The 2.3 GHz continuum survey of the GEM project. *ArXiv e-prints*, December 2007.

[169] A. R. Thompson, J. M. Moran, and G. W. Swenson, Jr. *Interferometry and Synthesis in Radio Astronomy, 2nd Edition.* New York, Wiley-Interscience, 2001. 692 p., 2001.

[170] N. Thyagarajan, N. Udaya Shankar, R. Subrahmanyan, W. Arcus, G. Bernardi, J. D. Bowman, F. Briggs, J. D. Bunton, R. J. Cappallo, B. E. Corey, L. deSouza, D. Emrich, B. M. Gaensler, R. F. Goeke, L. J. Greenhill, B. J. Hazelton, D. Herne, J. N. Hewitt, M. Johnston-Hollitt, D. L. Kaplan, J. C. Kasper, B. B. Kincaid, R. Koenig, E. Kratzenberg, C. J. Lonsdale, M. J. Lynch, S. R. McWhirter, D. A. Mitchell, M. F. Morales, E. H. Morgan, D. Oberoi, S. M. Ord, J. Pathikulangara, R. A. Remillard, A. E. E. Rogers, D. Anish Roshi, J. E. Salah, R. J. Sault, K. S. Srivani, J. B. Stevens, P. Thiagaraj, S. J. Tingay, R. B. Wayth, M. Waterson, R. L. Webster, A. R. Whitney, A. J. Williams, C. L. Williams, and J. S. B. Wyithe. A Study of Fundamental Limitations to Statistical Detection of Redshifted H I from the Epoch of Reionization. ApJ, 776:6, October 2013. doi: 10.1088/0004-637X/776/1/6.

[171] N. Thyagarajan, D. C. Jacobs, J. D. Bowman, N. Barry, A. P. Beardsley, G. Bernardi, F. Briggs, R. J. Cappallo, P. Carroll, B. E. Corey, A. de Oliveira-Costa, J. S. Dillon, D. Emrich, A. Ewall-Wice, L. Feng, R. Goeke, L. J. Greenhill, B. J. Hazelton, J. N. Hewitt, N. Hurley-Walker, M. Johnston-Hollitt, D. L. Kaplan, J. C. Kasper, H.-S. Kim, P. Kittiwisit, E. Kratzenberg, E. Lenc, J. Line, A. Loeb, C. J. Lonsdale, M. J. Lynch, B. McKinley, S. R. McWhirter, D. A. Mitchell, M. F. Morales, E. Morgan, A. R. Neben, D. Oberoi, A. R. Offringa, S. M. Ord, S. Paul, B. Pindor, J. C. Pober, T. Prabu, P. Procopio, J. Riding, A. E. E. Rogers, A. Roshi, N. Udaya Shankar, S. K. Sethi, K. S. Srivani, R. Subrahmanyan, I. S. Sullivan, M. Tegmark, S. J. Tingay, C. M. Trott, M. Waterson, R. B. Wayth, R. L. Webster, A. R. Whitney, A. Williams, C. L. Williams, C. Wu, and J. S. B. Wyithe. Foregrounds in Wide-field Redshifted 21 cm Power Spectra. ApJ, 804:14, May 2015. doi: 10.1088/0004-637X/804/1/14.

[172] S. J. Tingay, R. Goeke, J. D. Bowman, D. Emrich, S. M. Ord, D. A. Mitchell, M. F. Morales, T. Booler, B. Crosse, R. B. Wayth, C. J. Lonsdale, S. Tremblay, D. Pallot, T. Colegate, A. Wicenec, N. Kudryavtseva, W. Arcus, and et al. The Murchison Widefield Array: The Square Kilometre Array Precursor at Low Radio Frequencies. PASA, 30:e007, June 2013. doi: 10.1017/pasa.2012.007.

[173] T. Treu, K. B. Schmidt, M. Trenti, L. D. Bradley, and M. Stiavelli. The Changing Lyα Optical Depth in the Range $6 < z < 9$ from the MOSFIRE Spectroscopy of Y-dropouts. ApJ, 775:L29, September 2013. doi: 10.1088/2041-8205/775/1/L29.

[174] C. M. Trott, R. B. Wayth, and S. J. Tingay. The Impact of Point-source Subtraction Residuals on 21 cm Epoch of Reionization Estimation. ApJ, 757:101, September 2012. doi: 10.1088/0004-637X/757/1/101.

[175] M. P. van Haarlem, M. W. Wise, A. W. Gunst, G. Heald, J. P. McKean, J. W. T. Hessels, A. G. de Bruyn, R. Nijboer, J. Swinbank, R. Fallows, M. Brentjens, A. Nelles, R. Beck, H. Falcke, R. Fender, J. Hörandel, L. V. E. Koopmans, G. Mann, G. Miley, H. Röttgering, B. W. Stappers, R. A. M. J. Wijers, S. Zaroubi, M. van den Akker, A. Alexov, J. Anderson, K. Anderson, A. van Ardenne, M. Arts, A. Asgekar, I. M. Avruch, F. Batejat, L. Bähren, M. E. Bell, M. R. Bell, I. van Bemmel, P. Bennema, M. J. Bentum, G. Bernardi, P. Best, L. Bîrzan, A. Bonafede, A.-J. Boonstra, R. Braun, J. Bregman, F. Breitling, R. H. van de Brink, J. Broderick, P. C. Broekema, W. N. Brouw, M. Brüggen, H. R. Butcher, W. van Cappellen, B. Ciardi, T. Coenen, J. Conway, A. Coolen, A. Corstanje, S. Damstra, O. Davies, A. T. Deller, R.-J. Dettmar, G. van Diepen, K. Dijkstra, P. Donker, A. Doorduin, J. Dromer, M. Drost, A. van Duin, J. Eislöffel, J. van Enst, C. Ferrari, W. Frieswijk, H. Gankema, M. A. Garrett, F. de Gasparin, M. Gerbers, E. de Geus, J.-M. Grießmeier, T. Grit, P. Gruppen, J. P. Hamaker, T. Hassall, M. Hoeft, H. A. Holties, A. Horneffer, A. van der Horst, A. van Houwelingen, A. Huijgen, M. Iacobelli, H. Intema, N. Jackson, V. Jelic, A. de Jong, D. Kant, A. Karastergiou, A. Koers, H. Kollen, V. I. Kondratiev, E. Kooistra, Y. Koopman, A. Koster, M. Kuniyoshi, M. Kramer, G. Kuper, P. Lambropoulos, C. Law, J. van Leeuwen, J. Lemaitre, M. Loose, P. Maat, G. Macario, S. Markoff, J. Masters, R. A. McFadden, D. McKay-Bukowski, H. Meijering, H. Meulman, M. Mevius, R. Millenaar, J. C. A. Miller-Jones, R. N. Mohan, J. D. Mol, J. Morawietz, R. Morganti, D. D. Mulcahy, E. Mulder, H. Munk, L. Nieuwenhuis, R. van Nieuwpoort, J. E. Noordam, M. Norden, A. Noutsos, A. R. Offringa, H. Olofsson, A. Omar, E. Orrú, R. Overeem, H. Paas, M. Pandey-Pommier, V. N. Pandey, R. Pizzo, A. Polatidis, D. Rafferty, S. Rawlings, W. Reich, J.-P. de Reijer, J. Reitsma, G. A. Renting, P. Riemers, E. Rol, J. W. Romein, J. Roosjen, M. Ruiter, A. Scaife, K. van der Schaaf, B. Scheers, P. Schellart, A. Schoenmakers, G. Schoonderbeek, M. Serylak, A. Shulevski, J. Sluman, O. Smirnov, C. Sobey, H. Spreeuw, M. Steinmetz, C. G. M. Sterks, H.-J. Stiepel, K. Stuurwold, M. Tagger, Y. Tang, C. Tasse, I. Thomas, S. Thoudam, M. C. Toribio, B. van der Tol, O. Usov, M. van Veelen, A.-J. van der Veen, S. ter Veen, J. P. W. Verbiest, R. Vermeulen, N. Vermaas, C. Vocks, C. Vogt, M. de Vos, E. van der Wal, R. van Weeren, H. Weggemans, P. Weltevrede, S. White, S. J. Wijnholds, T. Wilhelmsson, O. Wucknitz, S. Yatawatta, P. Zarka, A. Zensus, and J. van Zwieten. LOFAR: The LOw-Frequency ARray. A&A, 556:A2, August 2013. doi: 10.1051/0004-6361/201220873.

[176] H. Vedantham, N. Udaya Shankar, and R. Subrahmanyan. Imaging the Epoch of Reionization: Limitations from Foreground Confusion and Imaging Algorithms. ApJ, 745:176, February 2012. doi: 10.1088/0004-637X/745/2/176.

[177] H. K. Vedantham, L. V. E. Koopmans, A. G. de Bruyn, S. J. Wijnholds, M. Brentjens, F. B. Abdalla, K. M. B. Asad, G. Bernardi, S. Bus, E. Chapman, B. Ciardi, S. Daiboo, E. R. Fernandez, A. Ghosh, G. Harker, V. Jelic, H. Jensen,

S. Kazemi, P. Lambropoulos, O. Martinez-Rubi, G. Mellema, M. Mevius, A. R. Offringa, V. N. Pandey, A. H. Patil, R. M. Thomas, V. Veligatla, S. Yatawatta, S. Zaroubi, J. Anderson, A. Asgekar, M. E. Bell, M. J. Bentum, P. Best, A. Bonafede, F. Breitling, J. Broderick, M. Brüggen, H. R. Butcher, A. Corstanje, F. de Gasperin, E. de Geus, A. Deller, 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, T. E. Hassall, G. Heald, M. Hoeft, J. Hörandel, M. Iacobelli, E. Juette, V. I. Kondratiev, M. Kuniyoshi, G. Kuper, G. Mann, S. Markoff, R. McFadden, D. McKay-Bukowski, J. P. McKean, D. D. Mulcahy, H. Munk, A. Nelles, M. J. Norden, E. Orru, M. Pandey-Pommier, R. Pizzo, A. G. Polatidis, W. Reich, A. Renting, H. Röttgering, D. Schwarz, A. Shulevski, O. Smirnov, B. W. Stappers, M. Steinmetz, J. Swinbank, M. Tagger, Y. Tang, C. Tasse, S. ter Veen, S. Thoudam, C. Toribio, C. Vocks, M. W. Wise, O. Wucknitz, and P. Zarka. Lunar occultation of the diffuse radio sky: LOFAR measurements between 35 and 80 MHz. MNRAS, 450:2291–2305, July 2015. doi: 10.1093/mnras/stv746.

[178] E. N. Vinyaikin. Frequency dependence of the evolution of the radio emission of the supernova remnant Cas A. *Astronomy Reports*, 58:626–639, September 2014. doi: 10.1134/S1063772914090078.

[179] E. N. Vinyaǐkin. Secular decrease of the radio flux density of Cassiopeia A determined from long-term observations. Observations at 38, 81.5, and 151.5 MHz. *Astronomy Reports*, 50:143–151, February 2006. doi: 10.1134/S1063772906020077.

[180] E. Visbal, R. Barkana, A. Fialkov, D. Tseliakhovich, and C. M. Hirata. The signature of the first stars in atomic hydrogen at redshift 20. Nature, 487:70–73, July 2012. doi: 10.1038/nature11177.

[181] T. C. Voytek, A. Natarajan, J. M. Jáuregui García, J. B. Peterson, and O. López-Cruz. Probing the Dark Ages at z $\approx$ 20: The SCI-HI 21 cm All-sky Spectrum Experiment. ApJ, 782:L9, February 2014. doi: 10.1088/2041-8205/782/1/L9.

[182] X. Wang, M. Tegmark, M. G. Santos, and L. Knox. 21 cm Tomography with Foregrounds. ApJ, 650:529–537, October 2006. doi: 10.1086/506597.

[183] R. B. Wayth, E. Lenc, M. E. Bell, J. R. Callingham, K. S. Dwarakanath, T. M. O. Franzen, B.-Q. For, B. Gaensler, P. Hancock, L. Hindson, N. Hurley-Walker, C. A. Jackson, M. Johnston-Hollitt, A. D. Kapińska, B. McKinley, J. Morgan, A. R. Offringa, P. Procopio, L. Staveley-Smith, C. Wu, Q. Zheng, C. M. Trott, G. Bernardi, J. D. Bowman, F. Briggs, R. J. Cappallo, B. E. Corey, A. A. Deshpande, D. Emrich, R. Goeke, L. J. Greenhill, B. J. Hazelton, D. L. Kaplan, J. C. Kasper, E. Kratzenberg, C. J. Lonsdale, M. J. Lynch, S. R. McWhirter, D. A. Mitchell, M. F. Morales, E. Morgan, D. Oberoi, S. M. Ord,

T. Prabu, A. E. E. Rogers, A. Roshi, N. U. Shankar, K. S. Srivani, R. Subrahmanyan, S. J. Tingay, M. Waterson, R. L. Webster, A. R. Whitney, A. Williams, and C. L. Williams. GLEAM: The GaLactic and Extragalactic All-Sky MWA Survey. PASA, 32:e025, June 2015. doi: 10.1017/pasa.2015.26.

[184] N Wiener. The interpolation, extrapolation and smoothing of stationary time series. Report of the Services 19, Massachusetts Institute of Technology, February 1942. Research Project DIC-6037.

[185] M. H. Wieringa. An investigation of the telescope based calibration methods 'redundancy' and 'self-cal'. *Experimental Astronomy*, 2:203–225, 1992. doi: 10.1007/BF00420576.

[186] M. Wolleben, T. L. Landecker, E. Carretti, J. M. Dickey, A. Fletcher, B. M. Gaensler, J. L. Han, M. Haverkorn, J. P. Leahy, N. M. McClure-Griffiths, D. McConnell, W. Reich, and A. R. Taylor. GMIMS: the Global Magneto-Ionic Medium Survey. In K. G. Strassmeier, A. G. Kosovichev, and J. E. Beckman, editors, *Cosmic Magnetic Fields: From Planets, to Stars and Galaxies*, volume 259 of *IAU Symposium*, pages 89–90, April 2009. doi: 10.1017/S1743921309030117.

[187] S. A. Wouthuysen. On the excitation mechanism of the 21-cm (radio-frequency) interstellar hydrogen emission line. AJ, 57:31–32, 1952. doi: 10.1086/106661.

[188] X. Wu. Probing the Epoch of Reionization with 21CMA: Status and Prospects. In *American Astronomical Society Meeting Abstracts #213*, volume 41 of *Bulletin of the American Astronomical Society*, page 226.05, January 2009.

[189] S. Yatawatta, A. G. de Bruyn, M. A. Brentjens, P. Labropoulos, V. N. Pandey, S. Kazemi, S. Zaroubi, L. V. E. Koopmans, A. R. Offringa, V. Jelić, O. Martinez Rubi, V. Veligatla, S. J. Wijnholds, W. N. Brouw, G. Bernardi, B. Ciardi, S. Daiboo, G. Harker, G. Mellema, J. Schaye, R. Thomas, H. Vedantham, E. Chapman, F. B. Abdalla, A. Alexov, J. Anderson, I. M. Avruch, F. Batejat, M. E. Bell, M. R. Bell, M. Bentum, P. Best, A. Bonafede, J. Bregman, F. Breitling, R. H. van de Brink, J. W. Broderick, M. Brüggen, J. Conway, F. de Gasperin, E. de Geus, S. Duscha, H. Falcke, R. A. Fallows, C. Ferrari, W. Frieswijk, M. A. Garrett, J. M. Griessmeier, A. W. Gunst, T. E. Hassall, J. W. T. Hessels, M. Hoeft, M. Iacobelli, E. Juette, A. Karastergiou, V. I. Kondratiev, M. Kramer, M. Kuniyoshi, G. Kuper, J. van Leeuwen, P. Maat, G. Mann, J. P. McKean, M. Mevius, J. D. Mol, H. Munk, R. Nijboer, J. E. Noordam, M. J. Norden, E. Orru, H. Paas, M. Pandey-Pommier, R. Pizzo, A. G. Polatidis, W. Reich, H. J. A. Röttgering, J. Sluman, O. Smirnov, B. Stappers, M. Steinmetz, M. Tagger, Y. Tang, C. Tasse, S. ter Veen, R. Vermeulen, R. J. van Weeren, M. Wise, O. Wucknitz, and P. Zarka. Initial deep LOFAR observations of epoch of reionization windows. I. The north celestial pole. A&A, 550: A136, February 2013. doi: 10.1051/0004-6361/201220874.

[190] D. G. York, J. Adelman, J. E. Anderson, Jr., S. F. Anderson, J. Annis, N. A. Bahcall, J. A. Bakken, R. Barkhouser, S. Bastian, E. Berman, W. N. Boroski, S. Bracker, C. Briegel, J. W. Briggs, J. Brinkmann, R. Brunner, S. Burles, L. Carey, M. A. Carr, F. J. Castander, B. Chen, P. L. Colestock, A. J. Connolly, J. H. Crocker, I. Csabai, P. C. Czarapata, J. E. Davis, M. Doi, T. Dombeck, D. Eisenstein, N. Ellman, B. R. Elms, M. L. Evans, X. Fan, G. R. Federwitz, L. Fiscelli, S. Friedman, J. A. Frieman, M. Fukugita, B. Gillespie, J. E. Gunn, V. K. Gurbani, E. de Haas, M. Haldeman, F. H. Harris, J. Hayes, T. M. Heckman, G. S. Hennessy, R. B. Hindsley, S. Holm, D. J. Holmgren, C.-h. Huang, C. Hull, D. Husby, S.-I. Ichikawa, T. Ichikawa, Ž. Ivezić, S. Kent, R. S. J. Kim, E. Kinney, M. Klaene, A. N. Kleinman, S. Kleinman, G. R. Knapp, J. Korienek, R. G. Kron, P. Z. Kunszt, D. Q. Lamb, B. Lee, R. F. Leger, S. Limmongkol, C. Lindenmeyer, D. C. Long, C. Loomis, J. Loveday, R. Lucinio, R. H. Lupton, B. MacKinnon, E. J. Mannery, P. M. Mantsch, B. Margon, P. McGehee, T. A. McKay, A. Meiksin, A. Merelli, D. G. Monet, J. A. Munn, V. K. Narayanan, T. Nash, E. Neilsen, R. Neswold, H. J. Newberg, R. C. Nichol, T. Nicinski, M. Nonino, N. Okada, S. Okamura, J. P. Ostriker, R. Owen, A. G. Pauls, J. Peoples, R. L. Peterson, D. Petravick, J. R. Pier, A. Pope, R. Pordes, A. Prosapio, R. Rechenmacher, T. R. Quinn, G. T. Richards, M. W. Richmond, C. H. Rivetta, C. M. Rockosi, K. Ruthmansdorfer, D. Sandford, D. J. Schlegel, D. P. Schneider, M. Sekiguchi, G. Sergey, K. Shimasaku, W. A. Siegmund, S. Smee, J. A. Smith, S. Snedden, R. Stone, C. Stoughton, M. A. Strauss, C. Stubbs, M. SubbaRao, A. S. Szalay, I. Szapudi, G. P. Szokoly, A. R. Thakar, C. Tremonti, D. L. Tucker, A. Uomoto, D. Vanden Berk, M. S. Vogeley, P. Waddell, S.-i. Wang, M. Watanabe, D. H. Weinberg, B. Yanny, N. Yasuda, and SDSS Collaboration. The Sloan Digital Sky Survey: Technical Summary. AJ, 120: 1579–1587, September 2000. doi: 10.1086/301513.

[191] O. Zahn, A. Lidz, M. McQuinn, S. Dutta, L. Hernquist, M. Zaldarriaga, and S. R. Furlanetto. Simulations and Analytic Calculations of Bubble Growth during Hydrogen Reionization. ApJ, 654:12–26, January 2007. doi: 10.1086/509597.

[192] O. Zahn, C. L. Reichardt, L. Shaw, A. Lidz, K. A. Aird, B. A. Benson, L. E. Bleem, J. E. Carlstrom, C. L. Chang, H. M. Cho, T. M. Crawford, A. T. Crites, T. de Haan, M. A. Dobbs, O. Doré, J. Dudley, E. M. George, N. W. Halverson, G. P. Holder, W. L. Holzapfel, S. Hoover, Z. Hou, J. D. Hrubes, M. Joy, R. Keisler, L. Knox, A. T. Lee, E. M. Leitch, M. Lueker, D. Luong-Van, J. J. McMahon, J. Mehl, S. S. Meyer, M. Millea, J. J. Mohr, T. E. Montroy, T. Natoli, S. Padin, T. Plagge, C. Pryke, J. E. Ruhl, K. K. Schaffer, E. Shirokoff, H. G. Spieler, Z. Staniszewski, A. A. Stark, K. Story, A. van Engelen, K. Vanderlinde, J. D. Vieira, and R. Williamson. Cosmic Microwave Background Constraints on the Duration and Timing of Reionization from the South Pole Telescope. ApJ, 756:65, September 2012. doi: 10.1088/0004-637X/756/1/65.

[193] M. Zaldarriaga, S. R. Furlanetto, and L. Hernquist. 21 Centimeter Fluctuations

from Cosmic Gas at High Redshifts. ApJ, 608:622–635, June 2004. doi: 10.1086/386327.

[194] H. Zheng, M. Tegmark, V. Buza, J. S. Dillon, H. Gharibyan, J. Hickish, E. Kunz, A. Liu, J. Losh, A. Lutomirski, S. Morrison, S. Narayanan, A. Perko, D. Rosner, N. Sanchez, K. Schutz, S. M. Tribiano, M. Valdez, H. Yang, K. Z. Adami, I. Zelko, K. Zheng, R. P. Armstrong, R. F. Bradley, M. R. Dexter, A. Ewall-Wice, A. Magro, M. Matejek, E. Morgan, A. R. Neben, Q. Pan, R. F. Penna, C. M. Peterson, M. Su, J. Villasenor, C. L. Williams, and Y. Zhu. MITEoR: a scalable interferometer for precision 21 cm cosmology. MNRAS, 445:1084–1103, December 2014. doi: 10.1093/mnras/stu1773.

[195] H. Zheng, M. Tegmark, J. S. Dillon, A. Liu, A. R. Neben, S. M. Tribiano, V. Buza, A. Ewall-Wice, H. Gharibyan, J. Hickish, E. Kunz, J. Losh, A. Lutomirski, S. Morrison, S. Narayanan, A. Perko, D. Rosner, N. Sanchez, K. Schutz, M. Valdez, H. Yang, K. Z. Adami, I. Zelko, and K. Zheng. Low Frequency Mapmaking with Compact Interferometers: A MITEoR Northern Sky Map from 128 MHz to 175 MHz. 2016.