1-1-2014

Polarization and its Discontents: How Polarized Foregrounds Affect 21cm Epoch of Reionization Measurements

David F. Moore
University of Pennsylvania, damo@sas.upenn.edu

Polarization and its Discontents: How Polarized Foregrounds Affect 21cm Epoch of Reionization Measurements

Abstract
As the first luminous objects began to form, they heated their surrounding medium, ionizing it. This event is the most recent cosmic phase transition, and occured during what is called the Epoch of Reionization (EoR). The ionization history of the intergalactic medium can be directly measured by 21cm emission from the hyperfine transition of hydrogen. Measurments of the 21cm signal from the EoR can yield information about those first luminous objects and help complete our understanding of cosmic history. Today, we measure the 21cm EoR signal in radio frequencies.

Excavating the 21cm EoR signal from beneath the bright foregrounds present at meter wavelengths requires pristine characterization of all foregrounds. We discuss how spectrally smooth foregrounds are isolated to particular regions of the 21cm EoR power spectrum, but Faraday-rotated, polarized sources can contaminate all regions, even those typically reserved for the 21cm EoR signal. To estimate the level of contamination we can expect from polarized foregrounds, we create a physically motivated simulation of the polarized sky at these wavelengths. These simulations imply that polarized foregrounds will contaminate the power spectrum at levels much higher than the 21cm signal.

To confirm the theories we develop in simulation, we turn to data taken with the Donald C. Backer Precision Array to Probe the Epoch of Reionization (PAPER), an array of antennae operating from 100 to 200 MHz in the Karoo desert of South Africa. Using data taken during a six month deployment with PAPER elements configured into an 8 Ã? 4 grid, we measure the power spectrum of all four Stokes parameters. The measured Q power spectrum exceeds its simulated values, allowing us to constrain the input parameters to the simulations. In particular, we are able to limit the mean polarized fraction of sources to 2.2 Ã? 10^{-3} , a factor of ten lower than existing measurements at 1.4 GHz, on which we based the simulations.

Finally, we present three new tools for characterizing polarized foregrounds.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Physics & Astronomy

First Advisor
James E. Aguirre

Subject Categories
Astrophysics and Astronomy | Physics

POLARIZATION AND ITS DISCONTENTS:

HOW POLARIZED FOREGROUNDS AFFECT 21CM EPOCH OF REIONIZATION MEASUREMENTS

David F. Moore

A DISSERTATION

in

Physics and Astronomy

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

2014

Supervisor of Dissertation

James E. Aguirre
Assistant Professor, Physics and Astronomy

Graduate Group Chairperson

Marija Drndić
Professor, Physics and Astronomy

Dissertation Comittee:

Gary Bernstein, Professor, Physics and Astronomy
Adam Lidz, Assistant Professor, Physics and Astronomy
Elliot Lipeles, Associate Professor, Physics and Astronomy
Masao Sako, Associate Professor, Physics and Astronomy

POLARIZATION AND ITS DISCONTENTS:

HOW POLARIZED FOREGROUNDS AFFECT 21CM EPOCH OF REIONIZATION MEASUREMENTS

David F. Moore

To GD.

Acknowledgements

Without the help and work of an immense number of people, this thesis would never have been made. I know most theses in physics say this, but it's true: I have a new appreciation of the phrase "standing on the shoulders of giants" from having worked with you all.

A very incomplete list of specfic thank you's:

The PAPER Team — I feel lucky to have worked alongside each one of you. To the guys who designed the instrument, Rich and the digital team at UCB in particular, I'm in awe. For PAPER to be as stable and reliable as it is — what an impressive feat! To everyone who went on deployments with me, it's been a pleasure getting to know you, and I'll miss working with you guys. William, thank you in particular for making South Africa such a friendly and hospitable place for me.

Aaron — So much of this thesis builds on your ideas, and I've learned so much from you. If there were one person without whom none of this would have happend, it would be you.

Danny — You've been such a close comrade through this whole process. Thank you for answering my dumb questions, for listening to me gripe, for helping me bang my head against code, and for getting me through my "natural enmity with computers," as you called it. I have no words...

James — Thank you for giving me the freedom I've needed to get here: for putting me in uncomfortable situations that have allowed me to grow (who would have thought that I could build a working computer cluster!), for giving me the resources to learn and grow as a student and as a scientist, and for linking me to such a large network of people from whom I've learned so much.

My Parents — Y'all are a constant source of unquestioning support. Especially in the last year, this has been crucial to my mental health, and to helping me find the gumption to finish this. In many ways, this thesis is yours, because without the twenty-some-odd years of hard work and encouragement you've given me, I'd never have made it this far. I'm lucky to have you guys.

Lucia — If there is a hero to my epic of reionization, it's you. I don't think I could ever write enough to thank you for how much you've given me throughout this whole ordeal. I think the best show of gratitude I could give you is a promise never to get another Ph.D. I love you more than words can say, and I don't even want to think about where I'd be right now if it weren't for you.

To show gratitude to, and to indicate the involvement of everyone who contributed to the making of this thesis, I'll always use the word "we" when referring to the authors of this thesis. I'll try only to break from this convention in footnotes and other comments from the peanut gallery.

ABSTRACT

POLARIZATION AND ITS DISCONTENTS:
HOW POLARIZED FOREGROUNDS AFFECT 21CM EPOCH OF REIONIZATION MEASUREMENTS

David F. Moore

James E. Aguirre

As the first luminous objects began to form, they heated their surrounding medium, ionizing it. This event is the most recent cosmic phase transition, and occured during what is called the Epoch of Reionization (EoR). The ionization history of the intergalactic medium can be directly measured by 21cm emission from the hyperfine transition of hydrogen. Measurments of the 21cm signal from the EoR can yield information about those first luminous objects and help complete our understanding of cosmic history. Today, we measure the 21cm EoR signal in radio frequencies.

Excavating the 21cm EoR signal from beneath the bright foregrounds present at meter wavelengths requires pristine characterization of all foregrounds. We discuss how spectrally smooth foregrounds are isolated to particular regions of the 21cm EoR power spectrum, but Faraday-rotated, polarized sources can contaminate all regions, even those typically reserved for the 21cm EoR signal. To estimate the level of contamination we can expect from polarized foregrounds, we create a physically motivated simulation of the polarized sky at these wavelengths. These simulations imply that polarized foregrounds will contaminate the power spectrum at levels much higher than the 21cm signal.

To confirm the theories we develop in simulation, we turn to data taken with the Donald C. Backer Precision Array to Probe the Epoch of Reionization (PAPER), an array of antennae operating from 100 to 200 MHz in the Karoo desert of South Africa. Using data taken during a six month deployment with PAPER elements configured into an 8×4 grid,

we measure the power spectrum of all four Stokes parameters. The measured Q power spectrum exceeds its simulated values, allowing us to constrain the input parameters to the simulations. In particular, we are able to limit the mean polarized fraction of sources to 2.2×10^{-3}, a factor of ten lower than existing measurements at 1.4 GHz, on which we based the simulations.

Finally, we present three new tools for characterizing polarized foregrounds.

Contents

List of Tables

List of Figures

Part I

Preliminaries

Chapter 1

The Epic of Reionization

1.1 The Early Universe

13.7 billion years ago, when the universe was only about 380,000 years old, the cosmic radiation background (CMB)[1] cooled to the point where its photons could no longer dissociate hydrogen, allowing the first stable atoms to form. This point in cosmic time is often called the surface of last scattering or the epoch of recombination, and is the earliest event we can measure, since before this time, the scattering rate of photons with early protons and electrons was so great as to render the universe opaque.

The relic abundance of photons from this period (the CMB) has provided a pristine picture of the initial conditions of our universe, and has provided much of the evidence for our current understanding of cosmology. It is truly remarkable how simple these initial conditions were — the standard model of cosmology, "vanilla" ΛCDM can be fully characterized by only eight numbers! Today, such a simplistic representation of our current universe

[1] I hesitate to call it the Cosmic Microwave Background (which is what CMB stands for) when talking about it during recombination. The "M" in CMB represents the fact that today, its brightness peaks at microwave frequencies. During recombination, it peaked in the UV. Do we call it the CUVB then? Hence, "cosmic radiation background."

would be absurd — with the growth of structure also came the growth of complexity. One of the main goals of the study of cosmology is to answer the question, how did the complex universe we see today arise?

To begin answering this question, we turn to the universe as it was just after recombination. Baryonic matter in the universe mostly consisted of neutral hydrogen. Quantum fluctuations in the radiation background yielded fluctuations in the density of matter. Fluctuations in the matter field allowed the first structures to form. The increased gravitational potential in the overdensities drew in the surrounding matter, making the overdensities more dense, and the underdensities less dense. Eventually, overdensities above some threshold density collapsed into galactic haloes — self-supporting structures held together by their own gravitational potential.

Within these haloes, the first stars and galaxies formed. It is in these stars and galaxies that we are interested for this thesis. As we will see in the next section, few Lyman alpha (Lyα) photons, the most abundant source of radiation from HI escaped these early times without being scattered. Without looking to other tracers of the HI, we know very little about this period of time. Because of this effect, some call this period "the cosmic dark ages." Others call it "cosmic dawn," because it was at this time that the first luminous objects originated.

The universe we see today is much different than it was during the cosmic dark ages. Stars have processed hydrogen and helium into heavier elements, allowing for a more complex chemistry. Galaxies have merged and evolved into much more complex structures than their relatively simple progenitors. Notably, UV emission from early luminous objects ionized the field of neutral hydrogen surrounding them, the intergalactic medium (IGM). Today, we measure a highly ionized IGM, but we know that without a neutral IGM, the CMB could not have arisen. The period of time when the IGM transitioned from neutral to ionized, called the Epoch of Reionization (EoR) captivates researchers, as it is the most recent cosmic phase transition.

By characterizing the ionization history of the IGM during the EoR, we can begin to solve some of the mysteries of the early universe. In what environment did the first luminous objects arise? What were they? When did the universe begin to look as it does today? These questions allows us to fill the gaps in our understanding between relatively thorough knowledge of the universe at early and late times.

1.2 How Do We Measure The Earliest Galaxies?

Now, we turn our attention to observations of the IGM during the EoR, focusing on both its timing and the nature of reionizing galaxies. Not only will this discussion elucidate the methods of detecting the signature of the first stars and galaxies, but it will also show the limits of our understanding of the EoR.

There are three main methods to measure HI during the EoR: the absorption of Lyα emission from high-redshift quasars, the scattering of CMB photons off of free electrons after reionization, and 21cm emission from the hyperfine transition of HI. We will discuss these three measurements in turn, focusing on the advantages and limitations of each. We will argue that the measurement most likely to detect signatures of reionization is the 21cm power spectrum, on which we will focus for the remainder of this thesis.

1.2.1 High Redshift Lyman Alpha Emitters

Lyα emission from high-redshift quasars is both relatively abundant and relatively bright, so it would seem to be an excellent candidate for detecting the signature of HI during the EoR. Lyα photons have a high cross-section to HI and will quickly become absorbed in a neutral IGM. As a Lyα photon from a quasar passes through a cloud of HI, it becomes absorbed, creating a dip in the quasar spectrum. Gunn and Peterson (25) investigated this effect in their seminal paper, predicting that, for a quasar in a highly neutral interstellar medium, all emission blue-wards of Lyα in the quasar's rest frame would be nearly completely absorbed.

Furthermore, they suggest using the optical depth of Lyα to that quasar yields a measure of the neutral hydrogen fraction, integrated along the line of sight.

To date, several measurements have been made of these so-called Gunn-Peterson troughs (4, 20, e.g.). Figure 1.1 shows measurements from Becker et al. (4) of several quasar spectra, demonstrating the increasing optical depth of the Lyα line with redshift. This indicates the increase of the neutral fraction with increasing redshift. Measurements like this one indicate the presence of HI in the IGM as late as $z \sim 6$. Measuring the global neutral fraction through this method is uncertain, as each measurement of x_{HI} depends on the line of sight to the quasar, and cannot measure a global quantity (82).

Gunn and Peterson (25), in a nearly identical calculation to the one presented in Section 1.3, show that the optical depth of Lyα through neutral hydrogen is, to an order of magnitude, $\tau_\nu \sim 10^4 x_{HI}$, where x_{HI} is the neutral fraction of hydrogen. This poses the difficult problem that nearly all Lyα emission passing through an IGM with $x_{HI} \gtrsim 10^{-4}$ is absorbed. Thus, these measurements can only access the late stages of the EoR, and will prove to be difficult to use in measuring x_{HI} during peak reionization. Lyα emitters can also be measured without spectroscopy. Objects with an excess brightness in one photometric band which could correspond to highly-redshifted Lyα are called "Lyman break" galaxies. Lyman break galaxies can be used to measure the luminosity function — and thus the mass function — of galaxies as a function of redshift Bouwens et al. (9, e.g.).

Figure 1.2 shows the luminosity function for several Lyman break galaxies in a series of redshift bins. The luminosity function seems to steepen with age, indicating fewer, older luminous galaxies; hence, galaxies during the EoR were most likely relatively small. There are two major sources of uncertainty in this measurement, though. First, since the redshift Lyman break galaxies are measured photometrically, there is a high uncertainty in their redshift. This uncertainty may be caused by a (photometric) degeneracy between redshift 6 Lyα and a redshift 2 OIII line, allowing for only $\sim 85\%$ confidence in the redshifts measured for these galaxies. Second, since only the brightest galaxies from this epoch can be measured,

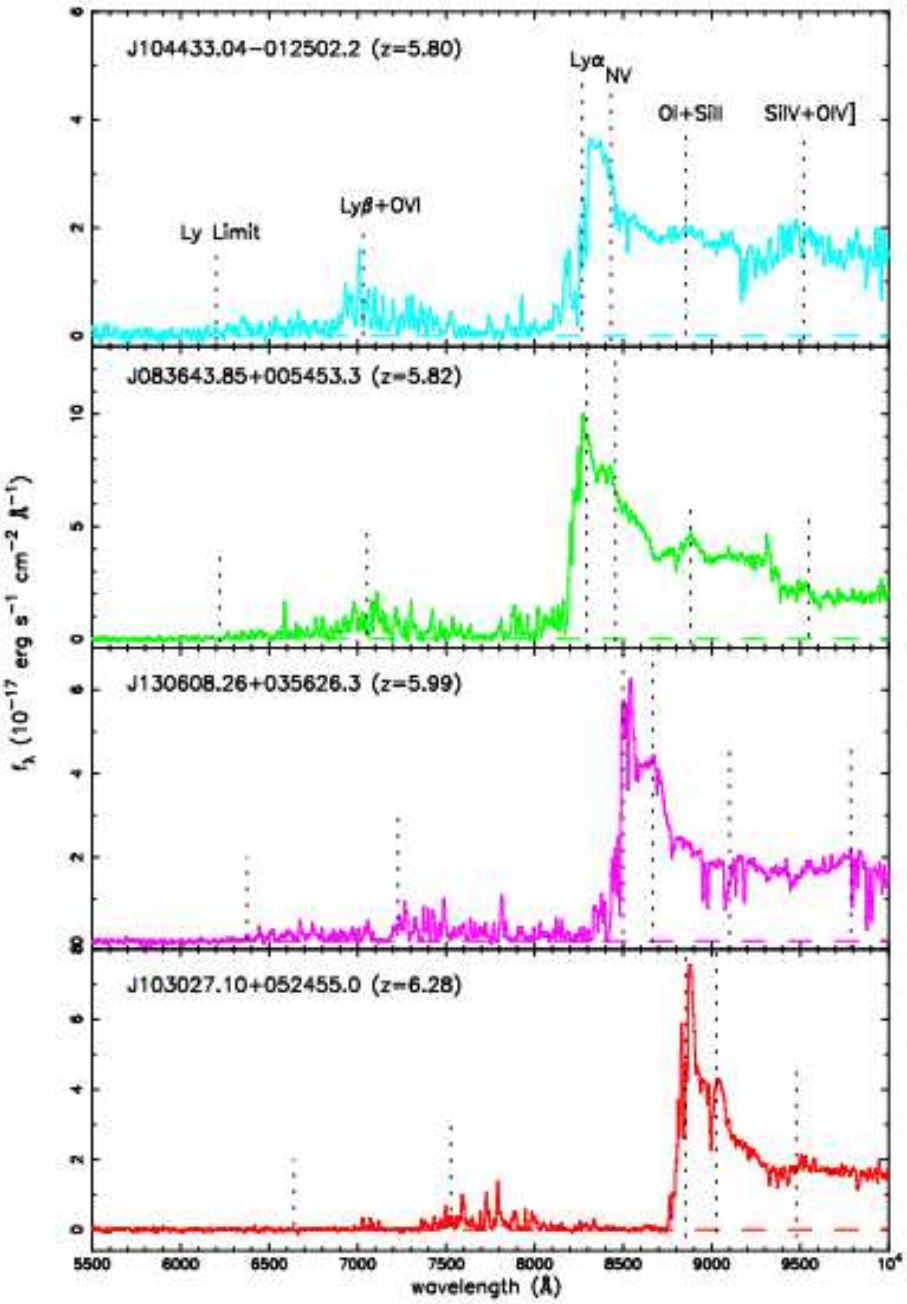

Figure 1.1: Quasar spectra at redshift 5.80 (cyan), 5.82 (green), 5.99 (magenta), and 6.28 (red), showing the near-total absorption of Ly-α in the high-redshift IGM. This so-called Gunn-Peterson trough indicates the presence of neutral hydrogen in the quasar environment. Figure taken from Becker et al. (4).

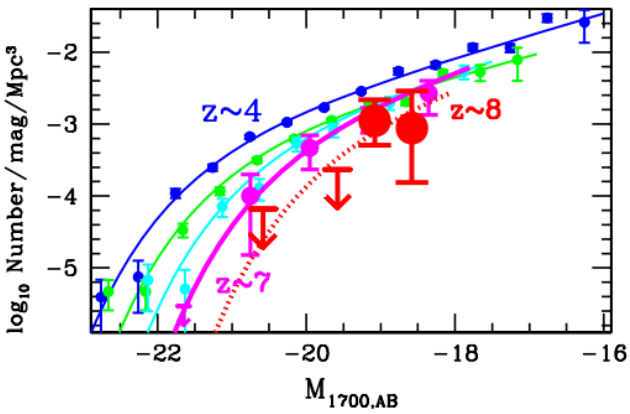

Figure 1.2: Luminosity function of Lyman break galaxies, in several redshift bins. These data indicate a steepening of the luminosity function with age. Figure taken from Bouwens et al. (9).

one must extrapolate measurements of the luminosity function to lower brightness, creating a high degree of uncertainty. Nonetheless, these measurements indicate that reionization was most likely driven by low-mass, low-brightness galaxies.

1.2.2 Hints from the CMB

Free electrons present after reionization will scatter CMB photons, suppressing the overall amplitude of the temperature power spectrum by a factor of $\exp\{-\tau_{ri}\}$, where τ_{ri} is the optical depth of a CMB photon through those free electrons (94). The global ionization history can be estimated by τ_{ri}, and in fact, τ_{ri} is one of the free parameters of vanilla ΛCDM (43, 77, e.g.), but there is a clear degeneracy between τ_{ri} and the overall scaling of the TT power spectrum.

This degeneracy is broken by analyzing the polarization of the CMB. If a CMB photon Thompson-scatters off of an electron from within a quadrupolar temperature anisotropy, the CMB becomes linearly polarized (this is one generator of E-modes). This type of scat-

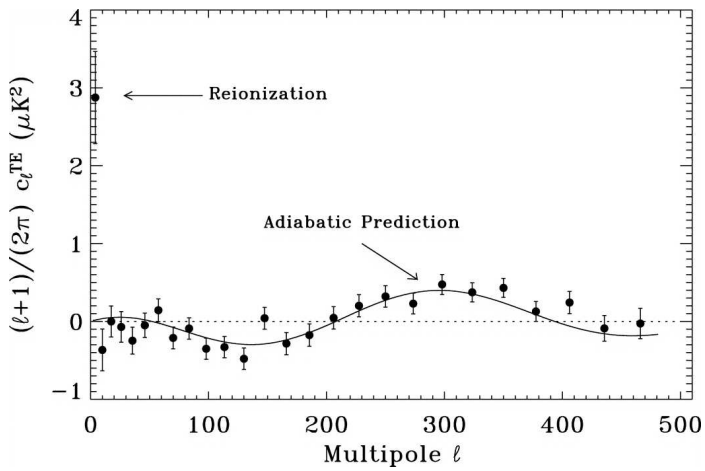

Figure 1.3: TE cross-correlation from WMAP, showing the first detection of the effects of reionziation on the CMB. Thompson scattering of CMB photons off of free electrons produced during reionization creates polarized power on scales of $\ell \sim 2\sqrt{\tau_{ri}}$, where τ_{ri} is the optical depth of a CMB photon through reionization. Figure from Kogut et al. (42).

tering during reionization contributes power to the EE spectrum, generating a peak whose amplitude is proportional to τ_{ri}, and whose location is around $\ell \sim 2\sqrt{\tau_{ri}}$ (93). Figure 1.3 shows the first detection of this effect, seen in the TE cross-correlation using the Wilkinson Microwave Anisotropy Probe (WMAP).

Extracting the ionization history from τ_{ri} is highly model dependent, since τ_{ri} is an integral quantity, summing information from all times since last scattering. Typically, researchers will assume simple models for the ionization history and model the optical depth as $\tau_{ri} \propto \int_0^{z_{recomb}} x_e(z)(1+z)^{-1}(dl/dz)\ \mathrm{d}z$, where z_{recomb} is the redshift of recombination, $x_e(z)$ is the ionization history, and dl/dz is the cosmological line-element at redshift z. Popular models for $x_{HI}(z)$ include "instantaneous reionization," in which $x_{HI}(z)$ is 1 until the redshift of reionization z_{reion}, and 0 afterwards, and a model including a sustained reionization, $x_{HI}(z) = \tan^{-1}\{(z - z_{reion})/\Delta z\}$, where Δz is the duration of reionization.

A final constraint on the ionization history of the IGM that can be drawn from the

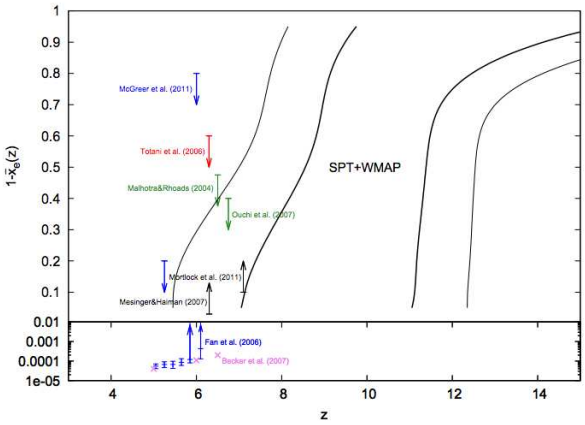

Figure 1.4: Summary of measurements of $x_{HI}(z) \approx 1 - \bar{x}_e(z)$ including information from both the CMB data and Lyα emitters, from Zahn et al. (92). Contours show the 1- and 2σ confidence intervals of $x_{HI}(z)$ taken from data from WMAP and the South Pole Telescope (SPT). Points, upper, and lower limits show measurements from Lyα emitters (citations shown in the figure).

CMB is due to the kinetic Sunyaev-Zel'dovic effect. CMB photons will Doppler-shift as they scatter off of moving electrons (79). The flow of free electrons generated by ionized bubbles will generated this Doppler shift in the CMB, and will add structure to the TT power spectrum (92). Measurements of this type are both heavily model dependent and vulnerable to systematic errors due to imperfect foreground removal. Figure 1.4 shows inferences of the ionization history from measurements of this effect from Zahn et al. (92), alongside measurements from CMB polarization and Lyα emitters.

1.2.3 Hyperfine Transition of Neutral Hydrogen

We now turn our attention to the third method of detecting the signature of the EoR, emission from the hyperfine transition of neutral hydrogen. The hyperfine transition occurs via an interaction between the spins of the electron and proton in an HI atom. The singlet state, when the spins are symmetric under interchange, has a slightly lower energy level than the triplet state, when they are anti-symmetric under interchange. This effect is usually summarized by a spin flip in the electron, though this only describes two of the three triplet states. This transition yields a photon with a wavelength of 21cm.

The hyperfine transition may be much more useful for detecting HI during the EoR than the previous two measurements. As we will show in Section 1.3, all lines of sight to neutral hydrogen at high redshift are optically thin, so emission from all times throughout the EoR is accessible. It also reaches Earth at radio frequencies, so ground-based experiments are sufficient to measure it. By contrast, redshift 10 Lyα reaches us in the infrared, to which the atmosphere is opaque. Also, since 21cm emission directly detects HI, information about the EoR can be directly inferred, with no dependence on a model, as in CMB analysis.

There are two downsides to using the 21cm line to measure neutral hydrogen. First, the hyperfine transition is a forbidden transition, and has a mean lifetime of around 10^7 years. The effect of this small transition rate is relatively dim emission from neutral hydrogen clouds. Second, the 21cm line from the EoR redshifts into meter wavelengths at which

galactic synchrotron emission dominates all astrophysical radiation. To give an estimate of the relative strengths of these two processes, the temperature of 21cm EoR emission is around 30 mK, and the temperature of galactic synchrotron emission is around 1000 K. That foreground emission is $\sim 10^5$ times brighter than the target signal necessitates a highly accurate foreground removal or avoidance scheme.

Despite the difficulties facing its measurement, the benefits of 21cm tomogrophy seem to outweigh the challenges. Again, the universe is optically thin to 21cm emission, and detecting HI during the EoR requires no model. Hence, for the duration of this thesis, we will focus on 21cm emission from neutral hydrogen, and inspect is usefulness for detecting the signature of the first stars and galaxies.

1.3 Spin Temperature

We begin our investigation into the utility of the hyperfine transition of HI as a probe of the EoR by solving the radiative transfer equation for a simple model of a neutral IGM. This is done in an attempt to gain intuition about the processes which generate 21cm emission, and to allow for a discussion of the environments in which it is produced. Much of this discussion will follow Furlanetto et al. (22) and class notes from Adam Lidz, but specific results will be individually cited.

The relative occupancy of hyperfine states in hydrogen gas can be characterized by the spin temperature T_s, defined in a Saha-like equation

$$\frac{n_1}{n_0} = 3 \exp\left\{\frac{T_*}{T_s}\right\}. \tag{1.1}$$

Here, n_1 and n_0 are the density of hydrogen in the triplet and singlet states, respectively; the factor of three represents the threefold degeneracy of the triplet state to the singlet state, and T_* is the line-temperature, defined such that $k_B T_* = h\nu_{21}$. For 21cm emission, $T_* = 68$ mK.

The intensity of 21cm emission can be calculated via the radiative transfer equation,

$$\frac{dI_\nu}{d\tau} = -I_\nu + s_\nu. \tag{1.2}$$

With the usual notation, I_ν is the intensity of 21cm emission; τ_ν is the optical depth, here a proxy for distance along the line of sight; and $s_\nu = j_\nu/\alpha_\nu$ is the source function.

By an argument of dimensional analysis ($\mathrm{d}E = j_\nu \mathrm{d}V \mathrm{d}\Omega \mathrm{d}\nu \mathrm{d}t$), we can write the emission coefficient in terms of the density of hydrogen atoms in the triplet state n_1, the line-profile of the emission line $\Phi(\nu) \approx \delta(\nu - \nu_{21})$, and the Einstein coefficient for the transition rate A_{10}:

$$j_\nu = \frac{h\nu}{4\pi} n_1 A_{10} \Phi(\nu). \tag{1.3}$$

In a similar fashion, we can write the absorption coefficient in terms of the Einstein coefficient for photoabsorption, B_{01}, and stimulated emission B_{10}:

$$\alpha_\nu = \frac{h\nu}{4\pi} \Phi(\nu)(n_0 B_{01} - n_1 B_{10}). \tag{1.4}$$

We can write all Einstein coefficients in terms of the emission rate A_{10} to write α_ν in the more managable form,

$$\alpha_\nu = 3 n_0 A_{10} \frac{\lambda^2}{8\pi} \Phi(\nu) \left(1 - e^{T_*/T_s}\right). \tag{1.5}$$

Finally, we write the source function, easing the burden by assuming that $T_s \gg T_*$:

$$s_\nu = \frac{j_\nu}{\alpha_\nu} = \frac{2k_B}{\lambda^2} T_s. \tag{1.6}$$

In the Rayleigh-Jeans limit, the spin temperature is the source function for the temperature of 21cm emission within a cloud of hydrogen gas.

Solutions to the radiative transfer equation (always in the Rayleigh-Jeans limit) take the form

$$T = T_{ex}(1 - e^{-\tau_\nu}) + T_{bg} e^{-\tau_\nu}, \tag{1.7}$$

where T_{ex} is the temperature associated with the source function s_ν, and T_{bg} is the temperature of background radiation — in our case, the temperature of the CMB at redshift z. In the limit where the optical depth is small, which we will justify shortly, the solution to the radiative transfer equation becomes

$$T = T_s\tau_\nu + T_{bg}(1-\tau_\nu), \tag{1.8}$$

where we take advantage of the source function's being the spin temperature.

The optical depth, τ_ν, can be calculated by integrating the absorption coefficient along the line of sight. This can be written in terms of the more cosmologically interesting quantities: the neutral fraction of hydrogen x_{HI}, the matter overdensity field δ, the density of Hydrogen n_H, the Hubble parameter $H(z)$, and the peculiar velocity of the Hydrogen cloud per unit length along the line of sight, $dv_{||}/dr_{||}$.

$$\tau_\nu = \int \alpha_\nu \, \mathrm{d}s = 9.2\times10^{-3}(1+\delta)(1+z)^{3/2}\frac{x_{HI}}{T_s}\left[\frac{H(z)/(1+z)}{dv_{||}/dr_{||}}\right] \tag{1.9}$$

The fiducial value of 10^{-2} justifies the assertion that $\tau_\nu \ll 1$, which we used to derive Equation 1.8. Such a small optical depth also provides a justification for preferring the hyperfine transition over Lyman-α as a probe of the EoR.

We can give an approximate solution for the contrast in brightness of 21cm emission to that of the CMB by inserting Equation 1.9 into Equation 1.8 and rearranging terms. Noting that $T_{bg} = T_{CMB}(1+z)$, where T_{CMB} is the current temperature of the CMB (2.4 K), we write the brightness contrast as

$$\begin{aligned}\delta T \equiv & \frac{T_S - T_{CMB}(1+z)}{(1+z)} \\ \approx & 9 \text{ mK } x_{HI}(1+\delta)(1+z)^{1/2}\left[1-\frac{T_{CMB}(1+z)}{T_s}\right]\left[\frac{H(z)/(1+z)}{dv_{||}/dr_{||}}\right]\end{aligned} \tag{1.10}$$

The point of writing this rather tedious calculation is to elucidate the processes that generate 21cm radiation, and how we may detect it. There are three salient points: first, we

detect the brightness contrast between 21cm radiation and the CMB; second, the spin temperature provides the source of that contrast, so the brightness contrast yields information on the astrophysical processes which drive the spin temerature; and third, the brightness contrast is proportional to the neutral fraction, so the cosmic ionization history can be derived from measurements of the brightness contrast over many redshifts. While cosmological parameters may be derived from the brightness contrast (53), there is also a wealth of rich astrophysics to be gleaned from the evolution of the spin temperature.

Thorough discussions of the evolution of the spin temperature may be found in Furlanetto et al. (22) and Pritchard and Loeb (71), which we summarize here. First, we will discuss the evolution of the spin temperature's global average, and then discuss the fluctuations.

There are two main physical processes which drive the evolution of the spin temperature. The first is collisional excitation of HI — as an HI atom scatters off of another species of particle (usually another HI atom), some of the kinetic energy of the collision is transferred to excite the hyperfine structure of the atom. This couples the spin temperature field to the underlying baryon density field (δ in Equation 1.10). This coupling depends on the density of the baryon field, which varies due to the growth of structure and the expansion of the universe. It also depends on the kinetic energy of that field, typically represented by the kinetic temperature of the gas.

The second process that drives the evolution of the spin temperature is a coupling to the underlying UV radiation field. A UV photon will excite the electron in an HI atom from the 1S state in the hyperfine singlet state into the 2P state. When the electron decays back into the 1S state, it may not return to the hyperfine singlet state, but to the hyperfine triplet state. This is called the Wouthuysen-Field effect (21, 91). This effect introduces astrophysics into the spin temperature by allowing it to be affected by UV emitters — early stars, galaxies, quasars, and small black holes. Because of this, the spin temperature is an excellent probe of the early universe.

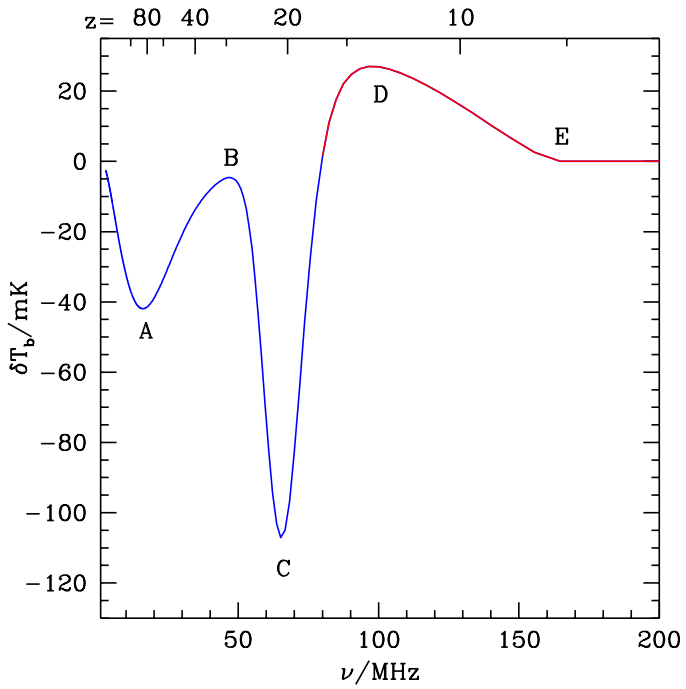

Figure 1.5: Mean brightness contrast versus redshift and frequency for a fiducial model of reionization. "Turning points" are labelled and explained in the text. The blue region shows where the brightness contrast is seen in absorption; red, in emission. Figure credit: Harker et al. (31).

The interplay of these two processes can be summarized by the "turning points" of the global spin temerature (31, 50, 54, e.g.). These are the points in time at which the time-derivative of the brightness contrast is zero, i.e. the spin temperature transitions from increasing to decreasing. Figure 1.5 shows a sample model of the evolution of the global spin temperature, to which we refer to explain the "turning points." (A) At around a redshift of 80, the spin temperature becomes uncoupled to the gas kinetic temperature, mostly due to the baryonic density growing smaller by cosmic expansion. It recouples to the CMB temperature, increasing the spin temperature. (B) As the first stars begin emitting UV photons, sometime around redshift 50, they decrease the spin temperature through the Wouthuysen-Field effect. (C) X-ray heating of the IGM, probably from accreting black holes, raises the spin temperature at around redshift 20. (D) The spin temperature begins to fall due to the start of reionization around redshift 13. (E) reionization ends — since the neutral fraction is nearly zero by redshift 6, the brightness contrast also falls to zero.

Information from the EoR can be obtained from the global signal between redshifts of 13-6, but a much richer story can be told by looking at fluctuations in the spin temperatue during the EoR. These fluctuations are characterized by a power spectrum, described in the next section. The power spectrum of the spin temperature will clearly depend on the underlying baryonic power spectrum — it depends explicitly on the baryon density field (Equation 1.10. But the power spectrum of spin temperature fluctuations is much more complex, due in part to the growth of ionized structures surrounding UV emitters. An over-simplified explanation is that "bubbles" of ionized hydrogen within the neutral hydrogen field arise as an X-ray or UV emitter ionizes its surrounding gas. As these bubbles first form, they introduce small-scale structure to the spin temperature field, steepening the power spectrum. As they grow, they flatten the slope of the power spectrum. As they merge, they introduce small-scale structure, causing the power spectrum to re-steepen towards high wavenumber. Finally, as reionization comes to completion, the amplitude of the power spectrum goes to zero.

Measuring both the mean spin temperature and fluctuations in the spin temperature field will allow us to discover both the timing of reionization and the nature of reionizing sources. This will give us a clearer picture of cosmic history and potentially uncover new physics.

1.4 Characterizing Fluctuations of the Spin Temperature

We begin by defining the correlation function to a temperature field $T(\mathbf{x})$:

$$\xi(\mathbf{r}) = \frac{1}{\mathbb{V}} \int T(\mathbf{x}) T^*(\mathbf{x} - \mathbf{r}) \, \mathrm{d}^3 x, \tag{1.11}$$

where $\mathbb{V}$ is the cosmological volume over which the field is sampled, and the limits of the integral extend over $\mathbb{V}$. This is more appropriately viewed as a power spectrum, which is simply the Fourier transform of the correlation function

$$P(\mathbf{k}) = \int \xi(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}} \, \mathrm{d}^3 r, \tag{1.12}$$

where $\mathbf{k}$ is the wave-number, typically measured in $h\mathrm{Mpc}^{-1}$.[2] Due to the convolution theorem, this is simply

$$P(\mathbf{k}) = \left| \widetilde{T}(\mathbf{k}) \right|^2, \tag{1.13}$$

where we have defined the Fourier-transformed temperature field as

$$\widetilde{T}(\mathbf{k}) = \frac{1}{\mathbb{V}} \int T(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} \, \mathrm{d}^3 x. \tag{1.14}$$

The power spectrum defined here is the goal of our measurements and the quantity which most easily allows us to compare different models of reionization and track the evolution of neutral hydrogen in the intergalactic medium at different scales.

[2] Note the difference in Fourier convention between the theorist's power spectrum and that of an interferometer.

Isotropy demands that $P(\mathbf{k})$ is rotationally invariant, which implies that the power spectrum is only dependent on the magnitude of the k-mode, i.e. $P(\vec{k}) = P(k)$. Because of this, it is customary to spherically average the power spectrum with log-spaced bins, defining the quantity $\Delta^2(k)$, often called the "dimensionless" power spectrum[3] as

$$\Delta^2(k) \equiv \frac{1}{(2\pi)^3} \int P(\mathbf{k}) k^3 \, \mathrm{d}\log k \, \mathrm{d}\Omega = \frac{k^3}{2\pi^2} P(k) \tag{1.15}$$

It is our task to define a method by which we can detect $P(\mathbf{k})$. We begin by inspecting the visibility (Equation 2.8) and comparing it with the expressions in Equations 1.12 and 1.14.

In the flat sky limit, neglecting calibration terms (including the primary beam which will be discussed later), the visibility reads

$$V(u, v, \nu) = \int I(l, m, \nu) e^{-2\pi i(ul+vm)} \, \mathrm{d}l \, \mathrm{d}m. \tag{1.16}$$

As we will discuss in Section 2.1, this is the two-dimensional Fourier transform over direction cosines l and m.

First, we note that the specific intensity of the observation, $I(l, m, \nu)$, is directly proportional to the spin temperature field $T(\mathbf{x})$. The constant of proportionality and the k-modes sampled by an observation are set by the limits of the observation, and will be discussed.

If we can assume that the measured range in l or m, call it $\Delta\theta$ is small, then the comoving distance subtended by $\Delta\theta$ can be written as a linear scaling, $r_{com} \approx X\Delta\theta$. We note that X is simply the comoving distance to redshift z, and can be found by integrating $c/H(z')$ with respect to z' from redshift 0 to z. In general, this expression is complicated, but for redshifts 5-15, this can be written in terms of a simple power law.

$$X = 6.5 \times 10^3 \left(\frac{\Omega_m}{0.27}\right)^{-1/2} \left(\frac{1+z}{10}\right)^{0.2} \; h^{-1}\mathrm{Mpc\ rad}^{-1}, \tag{1.17}$$

[3]Theorists tend to normalize the power spectrum by the global spin temperature, making this a truly dimensionless quantity. Since the global spin temperature is as interesting and as unknown as the power spectrum, we do not use this normalization. Our values of $\Delta^2(k)$ will have units of temperature squared.

where Ω_m is the cosmic matter density. We use WMAP7 values for cosmological paramters (43) to derive this number. This follows the expression in Furlanetto et al. (22).

Having given a prescription to convert a measured quantity into the two transverse cosmological distances, we now turn to the line of sight distance. Taking advantage of the one-to-one mapping from the redshift of a 21cm line to the distance to it, we write the distance spanned by bandwidth $\Delta\nu$ as $\Delta D \approx (dr_{com}/dz)(dz/d\nu)\Delta\nu$. Defining the slope $\Delta D \equiv Y\Delta\nu$, we find another simple, linear, scaling relation:

$$Y = 1.7 \times 10^{-2} \left(\frac{1+z}{10}\right)^{1/2} \left(\frac{\Omega_m}{0.27}\right)^{-1/2} h^{-1}\text{Mpc GHz}^{-1}. \tag{1.18}$$

This discussion can be summarized into three main points:

1. The sky intensity $I(l, m, \nu)$ is proportional to the spin-temperature field.
2. Angles on the sky can be converted into transverse k-modes.
3. The frequency dimension measures line-of-sight k-modes.

We have given approximate proportionality constants between measured l, m, ν coordinates and the cosmological x, y, z, and a more detailed discussion of power spectral inference will be given in Section 2.4.

1.5 Observational Prospects

Now that we have discussed the scientific goals of 21cm EoR observations, we turn to two state of the art measurements. These two datasets are taken with the PAPER experiment, the primary instrument used for the result of this thesis, described in detail in Chapter 3. One of these measurements is the power spectrum of foregrounds; the other is an upper limit on the EoR power spectrum, Combined, they provide context for the work in this thesis, and help elucidate some of the observational challenges.

1.5.1 Foregrounds

Perhaps the biggest hurdle to measuring the 21cm EoR power spectrum is mitigating the contributions of foregrounds to the power spectrum. As a general rule, the synchrotron emission from our galaxy and extragalactic radio sources are around 10^4 to 10^5 times brighter than the expected level of the spin temperature at peak reionization. To give representative values, the expected brightness temperature of a neutral hydrogen bubble at redshift of 10 is about 30 mK (Equation 1.10), but the contributions of foregrounds fall in the hundreds of degrees Kelvin.

There are two main strategies for excavating the 21cm signal from underneath the foregrounds, model subraction, and avoidance. Modelling and subtracting source spectra allows observers to access the underlying EoR power spectrum, but requires a high level of precision in both the model spectrum and instrumental calibration terms (10, 16, e.g.). While some strategies require precise imaging of sources, many have simply removed the few brightest principal components from their spectra (51, e.g.). This method has been used with some success (19, 61), but as Paciga et al. (61) demonstrated, blind subtraction of principal components can also remove components of the EoR power spectrum, severely decreasing sensitivity.

The second strategy for dealing with foregrounds is through avoidance. Synchrotron emission generally follows a spectral power law, and thus will accumulate in the lowest bins of $k_{||}$, while the 21cm EoR signal falls in higher $k_{||}$ modes as well. Hence, foregrounds may be avoided by focusing analysis only on the highest $k_{||}$ modes.

A number of studies show evidence of a foreground "wedge" in the $k_\perp$-$k_{||}$ plane for smooth-spectrum foregrounds (49, 58, 66, e.g.). This wedge is due to the spectral response of an interferometer, and as we will show in Section 2.3, is set by the geometry of an interferometric array. Figure 1.6 shows the first observational confirmation of the "wedge." The region above and to the left of the orange lines in Figure 1.6 is relatively free of contamination from foregrounds, and can be designated as the primary target for observations.

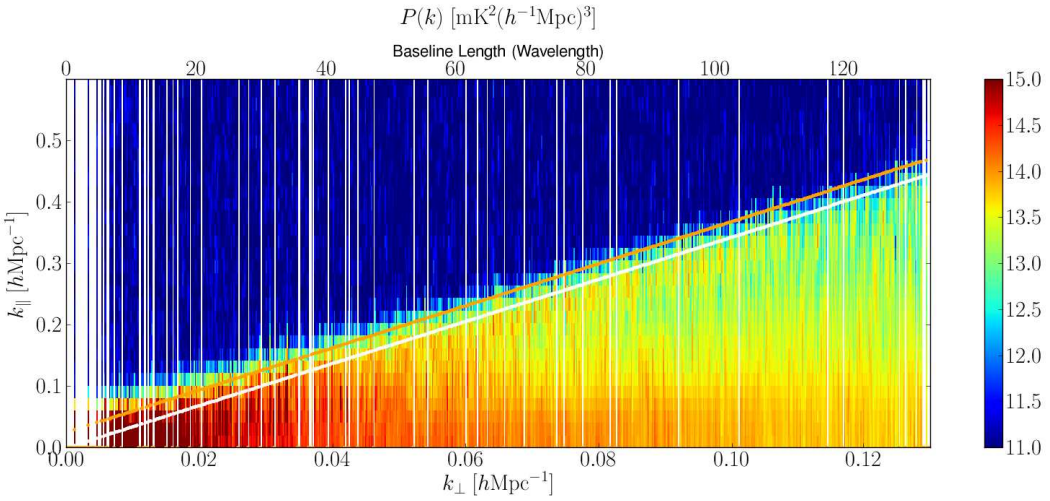

Figure 1.6: Power spectrum of foregrounds, in mK^2 (h^3Mpc^{-3}). The data presented were taken for four hours during the PSA64 season (Table 3.1). Smooth spectrum foreground emission is contained within a foreground "wedge," delimited by the baseline length. A white line shows the baseline length, and an orange line shows the baseline length plus a 50 ns buffer. This buffer encloses the foreground emission, convolved by the kernel of the power law spectra typical to radio sources. Figure taken directly from Pober et al. (69).

A negative consequence of restricting analysis to high $k_{||}$ modes is the relative levels of the 21cm EoR power spectrum and uncertainty due to thermal noise. Typically, observers target the spherically averaged power spectrum ($\Delta^2(k)$ in Equation 1.15), in which the 21cm EoR signal is relatively flat, and uncertainty due to thermal noise rises as k^3. Thus, observers target the lower $k_{||}$ modes of the upper triangle of Figure 1.6. This effect provides the impetus for using short baselines for EoR analysis, since short baselines probe the lowest values of $k_\perp$, and restrict foregrounds to the smallest region in $k_{||}$.

Future observations and possible detections will likely use a combination of both strategies. Pober et al. (68) show that the most highly sensitive modes to the power spectrum exist within the foreground wedge, and the best prospects for detection are those in which inter-horizon modes can be accessed.

1.5.2 Current Upper Limits to the Power Spectrum

To date, there are two prominent upper limits to the 21cm EoR power spectrum: Parsons et al. (65), and Paciga et al. (61). Both of these measurements have overcome significant hurdles in foreground removal and avoidance (respectively), and show that great progress is being made to the detection of the 21cm EoR power spectrum. Since this thesis focuses on data taken with the PAPER array (Chapter 3), we will focus on the Parsons result, taken with the PAPER instrument.

Figure 1.7 shows the power spectrum at a redshift of 7.7, measured in the EoR2011 Season (Table 3.1), alongside a fiducial model of the power spectrum (48). While uncertainties from thermal noise, residual foregrounds, and other instrumental systematics prevent a detection of the power spectrum, this measurement can constrain the brightness contrast of the spin temperature to below $(41\ \mathrm{mK})^2$ at $k = 0.27\ h\mathrm{Mpc}^{-1}$. This value allows for one of the first physical constraints to the history of the IGM, since it can rule out reionization scenarios in which the IGM cools adiabatically, with no heating from early ionizing

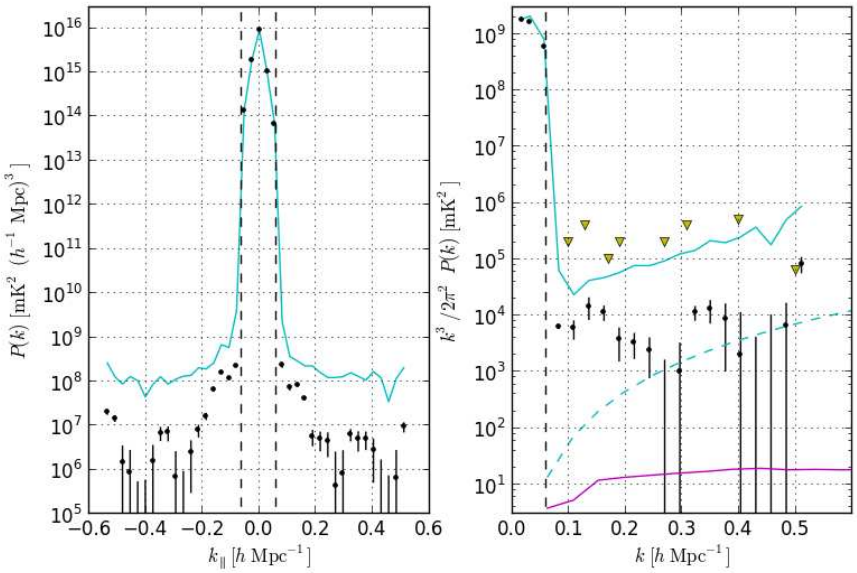

Figure 1.7: (Left Panel): Power spectrum measurements from the EoR2011 observing season. (Right Panel): Spherically averaged power spectrum taken from that same data. Dashed, vertical lines show the horizon limit for the 16λ baselines used for these measurements. Points and error bars show covariance-removed data, and 95% confidence intervals (See Section 6.1 and 6.2.1 for details). Cyan lines show the 2σ upper limit of data without the covariance removal applied. Dashed, cyan lines show the level of thermal fluctuations, assuming $T_{sys} = 550$ K. The magenta line shows a fiducial 21cm EoR model, taken from Lidz et al. (48). Yellow triangles show the 2σ upper limits presented in Paciga et al. (61). Figure taken directly from Parsons et al. (65).

sources. Essentially, restricting the brightness temperature to below 41 mK shows that X-ray heating of the IGM contributed to the bulk of reionization. Scenarios in which the IGM cools adiabatically are highly unlikely, but this measurement is one of the first observational confirmations of that statement.

Chapter 2

Interferometry

2.1 A Thought Experiment

In this section, we will introduce interferometry, and derive an expression for the visibility from first principles. While this material is used commonly enough and predates many references typically given, it is nonetheless useful to provide a reference for a more thorough discussion of this material, which serve as the sources for it. A more thorough and complete discussion of interferometry can be found in two books: Taylor et al. (81) and Thompson et al. (83).

Suppose we have two telescopes — label them A and B — which are separated by some baseline vector $\vec{b}$. Next, suppose that a single plane wave at frequency ν is indendent on these telescopes A and B, with a propagation direction $-\hat{s}$. For now, let's assume that these two telescopes are evenly sensitive to radiation coming from all directions — we can add that complication later. We will also neglect projection effects of the electric field vector on telescopes A and B until later — we're only concerned with the phase information for now.

We define the phase of the electric field so that at telescope A, it takes a value

$$E_A = E_0 \exp\left\{-2\pi i \nu t\right\}, \tag{2.1}$$

which sets the value of the electric field at telescope B to be

$$E_B = E_0 \exp\left\{-2\pi i\nu[t + (\vec{b}/c)\cdot\hat{s}]\right\}. \tag{2.2}$$

The extra term in the phase, $(\vec{b}/c)\cdot\hat{s}$, represents the time-difference of a wavefront's arrival on telescope A and B, and is thus defined as the delay. If we correlate the signals from the two telescopes in time, we extract the delay thus:

$$\langle E_A E_B^* \rangle_t = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} E_A(t) E_B(t)\ \mathrm{d}t \tag{2.3}$$

$$= |E_0|^2 \exp\left\{-2\pi i\nu(\vec{b}/c)\cdot\hat{s}\right\}. \tag{2.4}$$

The next stage of complication is to allow plane waves to come from all directions. In order to represent the response of telescopes A and B to the entire celestial sphere, we integrate over the sphere, and allow $|E_0|^2 \propto I$, the intensity of the incident emission, to vary as a function of direction:

$$\langle E_A E_B^* \rangle_t = \int I(\hat{s}) \exp\left\{-2\pi i\nu(\vec{b}/c)\cdot\hat{s}\right\}\ \mathrm{d}\Omega. \tag{2.5}$$

We can leave the details of the projection of the sphere onto our antennae for later, but this equation reveals the two fundamental aspects of interferometry:

1. There is a Fourier relationship between the intensity of celestial emission, $I(\hat{s})$, and the inteferometric response $\langle E_A E_B^* \rangle_t$.

2. The dual variable to sky position $\hat{s}$ (in the Fourier sense) is the baseline vector between two antennae, measured in units of wavelength.

Before attempting to create an image from the interferometer's response, we will add two final complications. First, we define a coordinate system. The antenna locations are defined in a topocentric coordinate system (u, v, w), with u the local easting, v is the local northing, and w is pointed towards zenith. This coordinate system is fixed to the earth at

the location of the observer. The upper half of the celestial sphere is characterized by the coordinates (l, m), where at zenith, l points in the direction of u. The relevant terms of equation 2.5 are the measure,

$$\mathrm{d}\Omega = \frac{\mathrm{d}l\mathrm{d}m}{\sqrt{1-l^2-m^2}}, \tag{2.6}$$

and the delay,

$$(\vec{b}/\lambda)\cdot\hat{s} = ul + vm + w(1-\sqrt{1-l^2-m^2}). \tag{2.7}$$

It is typical to assume that the array of antennae is coplanar, so we can set $w = 0$, and define baselines as being in the uv-plane. It is also typical to use the flat-sky approximation for these coordinates, setting the measure $\mathrm{d}\Omega \simeq \mathrm{d}l\ \mathrm{d}m$, but we will forgo this approximation.

The penultimate complication we will add (the final complication, polarization, gets its own section) is the spatial response pattern of an antenna, $A(l, m)$, called the primary beam. The primary beam attenuates the signal received by the electric field, modifying $E_i \rightarrow A_i E_i$. We assume that all antennae are identical, so we can combine the product $A_i A_j$ into a single primary beam A, which attenuates the intensity of incident radiation, rather than the electric field.

Finally, we can name the interferometer's response, and define the visibility (a word studiously avoided until now) as

$$V(u, v, \nu) = \int A(l, m) I(l, m) e^{-2\pi i(ul+vm)} \frac{\mathrm{d}l\ \mathrm{d}m}{\sqrt{1-l^2-m^2}}. \tag{2.8}$$

Images may be recreated from the visibilities by an inverse Fourier transforms over all visibilties in an array:

$$\frac{A(l, m)\widetilde{I}(l, m)}{\sqrt{1-l^2-m^2}} = \int \text{Ш}(u, v) V(u, v) e^{+2\pi i\nu(ul+vm)}\ \mathrm{d}u\ \mathrm{d}v. \tag{2.9}$$

A couple of items to note in Equation 2.9: first, the reconstructed image $\widetilde{I}$ is attenuated by the beam-response pattern and the measure[4] Second (and more importantly) the sampling

[4]Typically, the factor of $(1-l^2-m^2)^{-1/2}$ is absorbed into the beam response $A(l, m)$. Henceforth, we will follow this convention.

function $\text{Ш}(u, v)$, which is defined as being 1 in the uv points that an array samples and 0 elsewhere, prohibits the full Fourier spectrum from being included in the reconstructed image. The effect is that the true image I is convolved with what is defined as the dirty beam, the Fourier-transform of the sampling function. Because $\text{Ш}(u, v)$ contains zeros, a complete deconvolution of the reconstructed image from the dirty beam is impossible.

2.2 Polarimetry

Here, we add the final complication to our model of the visibility (Equation 2.8), which will be the subject of this thesis: polarization. To understand the interferometer's response to a polarized signal, we need first to lay some groundwork defining our coordinate systems.

As mentioned in the previous section, the uv-plane is defined as being fixed to the earth, but the polarization vector of a source is obviously fixed to the celestial sphere. To account for this, we usually define the uv-plane in what we call the topocentric coordinate system (East, North, Up), and we define the source position in the equatorial coordinate system, fixed to the sky, ((X, Y, Z), where $X^2 + Y^2 + Z^2 = 1$). The three-coordinate equatorial system can be converted into the more familiar right-ascension α and declination δ via a similar transformation to convert from Cartesian to spherical coordinates:[5]

$$\tan \alpha = \frac{Y}{X}, \quad \sin \delta = Z. \tag{2.10}$$

The projection of the topocentric uv-plane into equatorial coordinates is the projection of the uv-plane onto the tangent plane of the celestial sphere. Since the hemisphere available to an observer is dependent on the observer's location on the earth and the local sidereal time for the observer, the projection matrix $\mathbf{R}$ is a function of location and time. We write $\mathbf{R}$

[5] The differences are due to the definitions of the two coordinates — δ is defined to be zero at the equator, rather than the pole ($\delta = \pi/2 - \theta$), and α is defined to be left-handed, to track the earth's rotation, from the point of view of an observer looking up ($\alpha = -\phi$).

in terms of the declination of the pointing δ and the hour angle of the pointing $H = LST - \alpha$ as

$$\mathbf{R} = \begin{pmatrix} \sin H & -\sin\delta\cos H & \cos\delta\cos H \\ \cos H & \sin\delta\sin H & -\cos\delta\sin H \\ 0 & \cos\delta & \sin\delta \end{pmatrix}, \tag{2.11}$$

which is defined to act on topocentric vectors (the baseline $\vec{b}$, for instance) on the right as

$$\vec{b}_{eq} = \vec{b}_{top} \cdot \mathbf{R}, \tag{2.12}$$

where $\vec{b}_{eq}$ is the baseline vector represented in equatorial coordinates, and $\vec{b}$ is the baseline vector represented in topocentric coordinates. The projection matrix $\mathbf{R}$ is defined to act on baselines from the right to facilitate interpretation of the product $\vec{b} \cdot \hat{s}$.[6]

Having defined the equatorial and topocentric coordinate systems, and described how a vector from one projects onto the other, let us now discuss the physical process of a polarization vector projecting onto an interferometer's dipole. In doing so, we will introduce three concepts: the Kronecker product, the Stokes parameters, and parallactic rotation.

The polarization vector of a source is fixed in the celestial sphere, making equatorial coordinates a natural choice of basis. Since the propagation direction of the E-field of any source is radially inward, toward the observer, right-ascension and declination are better suited to describe these vectors. An interferometer, and any telescope in general, projects these vectors onto its own local frame of reference. If we choose topocentric coordinates (x, y), colinear with the previously defined (u, v),[7] then we can write this projection as a simple rotation

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} \equiv \mathbf{P}' \cdot \begin{pmatrix} E_\alpha \\ E_\delta \end{pmatrix} = \begin{pmatrix} \cos\psi & \sin\psi \\ -\sin\psi & \cos\psi \end{pmatrix} \begin{pmatrix} E_\alpha \\ E_\delta \end{pmatrix}, \tag{2.13}$$

[6] Am I standing still and the sky rotates around me, or is the sky fixed and I'm rotating through it? Either interpretation is correct, depending on your choice of frame of reference. The source vector $\hat{s}$ can be written in the topocentric frame as $\mathbf{R} \cdot \hat{s}$, so in either frame of reference, the term $\vec{b} \cdot \mathbf{R} \cdot \hat{s}$, being a scalar, is coordinate-independent, and thus is always constant. This is highly analogous to the Schrödinger and Heisenberg pictures of quantum mechanics.

[7] Generally, u and v are the Fourier-dual variables to sky coordinates l and m, measured in units of wavelength. x and y are the physical easting and northing, measured in units of length.

where $E_{x,y}$ are the components of the electric field in the xy coordinate system, $E_{\alpha,\delta}$ are the components in the equatorial frame, and ψ is the parallactic angle, defined in terms of latitude λ, hour angle of observation H, and the declination of observation δ as

$$\tan\psi = \frac{\cos\lambda\sin H}{\sin\lambda\cos\delta - \cos\lambda\sin\delta\cos H} \tag{2.14}$$

Figure 2.1 shows the parallactic rotation of a plus sign.

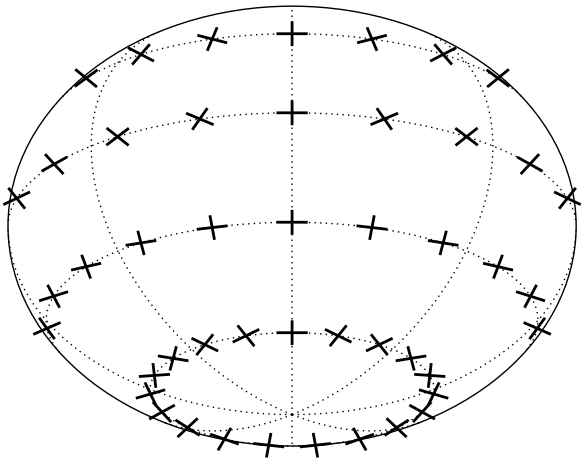

Figure 2.1: Parallactic rotation of a plus sign. Lines of constant declination at $\delta = 30,\ 0,\ -30,\ -60$ and lines of constant hour angle $H = 0,\ 3h,\ 6h,\ 12h,\ 15h,\ 18h,\ 21h$ are shown with dotted lines. The latitude is chosen to be that of the PAPER array ($-30°43'17.5'$). Hour angle 0 is chosen to be the meridian.

The next step in propagating a celestial signal through an interferometer is applying any instrumental gains to each component of the electric field. In general, this can be represented by a 2×2, complex matrix ($\mathbf{G}$), all of whose elements are non-zero. There are three major components to setting the elements of $\mathbf{G}$: the overall amplification of the

signal, the electrical delay of the signal with respect to an array-wide average, and any instrumental polarization terms, most of which can arise from improperly aligned feeds or inter-signal crosstalk. Each signal recieves each of these calibration parameters, so there exist N_{ant} copies of the matrix $\mathbf{G}$, for an array with N_{ant} antennae. Finally, we separate the direction-dependent terms and call them the primary beam $A(l,m)$. For the purposes of simplicity[8], we will omit any discussion of cross-polarizing terms and represent the matrix for the i^{th} antenna's matrix $\mathbf{G}_i$ in terms of signal gain for each polarization α, g_i^α and an electrical delay for each polarization τ_i^α, and the α polarization's primary beam:

$$\mathbf{G}_i = \begin{pmatrix} g_i^x A_x(l,m) e^{-2\pi i \nu \tau_i^x} & 0 \\ 0 & g_i^y A_y(l,m) e^{-2\pi i \nu \tau_i^y} \end{pmatrix}. \tag{2.15}$$

This operation is applied to the signal at each antenna, making the electric field measured from the i^{th} antenna at the point of correlation

$$\vec{E}_i = \mathbf{G}_i \cdot \mathbf{P} \cdot \vec{E}_{\alpha\delta}. \tag{2.16}$$

At last, we have discussed the necessary precursors, and can begin correlation! An interferometer whose dipoles are aligned along the x- and y-axes (equivalently the u- and v-axes) correlates each component of the E-field with both itself and the other, totaling to four polarization products. This operation can be represented by the Kronecker outer product between two matrices[9], which takes an $m \times n$ matrix $\mathbf{A}$ and a $p \times q$ matrix $\mathbf{B}$ and computes the $mp \times nq$ matrix

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & \dots & a_{1n}\mathbf{B} \\ \vdots & \ddots & \vdots \\ a_{1m}\mathbf{B} & \dots & a_{mn}\mathbf{B} \end{pmatrix}, \tag{2.17}$$

[8] As we will discuss in Chapter 6, these cross-terms are measured to be negligibly small for our purposes.

[9] There is another convention, defining an outer product in which $(\mathbf{A} \otimes \mathbf{B}^\dagger)_{ij} = A_i B_j^*$. Using this notation, the native, linear polarization products are simply an expansion of the Stokes parameters times the Pauli matrices and unity. While this definition of an outer product leads to mathematically elegant results, and is in theory equivalent to our choice of an outer product, I personally find the calculations to be quite cumbersome and will not use it.

where a_{ij} is the $(i,j)^{\text{th}}$ element of the matrix $\mathbf{A}$, and each $p \times q$ block (represented by $\mathbf{B}$) contains the elements of matrix $\mathbf{B}$. As an example, the Kronecker product of the E field (in the (x,y) representation) is

$$\vec{E}_{xy} \otimes \vec{E}^*_{xy} = \begin{pmatrix} E_x E_x^* \\ E_x E_y^* \\ E_y E_x^* \\ E_y E_y^* \end{pmatrix}, \tag{2.18}$$

which weighted by the primary beam, and integrated over time and space, is a visibility. A useful property of the Kronecker product is the mixed-product property, which states that

$$(\mathbf{AB}) \otimes (\mathbf{CD}) = (\mathbf{A} \otimes \mathbf{C}) \cdot (\mathbf{B} \otimes \mathbf{D}). \tag{2.19}$$

So, if we define the 4×1 vector of visibilities measured between antennae i and j as $\mathcal{V} \equiv \langle \vec{E}_i \otimes \vec{E}^*_j \rangle$, then

$$\mathcal{V} = \left\langle (\mathbf{G}_i \mathbf{P} \vec{E}_{\alpha\delta}) \otimes (\mathbf{G}_j \mathbf{P} \vec{E}_{\alpha\delta})^* \right\rangle = (\mathbf{G}_i \otimes \mathbf{G}^*_j)(\mathbf{P} \otimes \mathbf{P}) \left\langle \vec{E}_{\alpha\delta} \otimes \vec{E}_{\alpha\delta} \right\rangle \tag{2.20}$$

We can tackle this expression term-by-term, starting on the right. Equation 2.18 gives an expression for the outer product of the two linear components of an electric field, but a much more useful basis can be found. This basis, whose components are called the Stokes parameters, is defined for linearly polarized components of the electric field thus:

$$I = |E_\alpha|^2 + |E_\delta|^2 \tag{2.21}$$

$$Q = |E_\alpha|^2 - |E_\delta|^2 \tag{2.22}$$

$$U = 2\mathfrak{Re}\left\{E_\alpha E^*_\delta\right\} = E_\alpha E^*_\delta + E_\delta E^*_\alpha \tag{2.23}$$

$$V = 2\mathfrak{Im}\left\{E_\alpha E^*_\delta\right\} = -i(E_\alpha E^*_\delta - E_\delta E^*_\alpha). \tag{2.24}$$

This basis conveniently represents the power of the electric fields in terms of its total intensity (I), the power contained in each component of a basis of two, orthogonal, linear polarizations

(Q and U), and the power contained in circular polarizations (V). The rotation can be represented by the matrix $\overleftrightarrow{S}$, defined as

$$\begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -i & i & 0 \end{pmatrix} \begin{pmatrix} |E_\alpha|^2 \\ E_\alpha E_\delta^* \\ E_\delta E_\alpha^* \\ |E_\delta|^2 \end{pmatrix} = \overleftrightarrow{S} \begin{pmatrix} |E_\alpha|^2 \\ E_\alpha E_\delta^* \\ E_\delta E_\alpha^* \\ |E_\delta|^2 \end{pmatrix} \tag{2.25}$$

For completeness, we present the inverse of $\overleftrightarrow{S}$, and note that were it not for the normalization which requires I to contain the *total* intensity of the electric field, $\overleftrightarrow{S}$ would be Hermitian:

$$\overleftrightarrow{S}^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & i \\ 0 & 0 & 1 & -i \\ 1 & -1 & 0 & 0 \end{pmatrix}. \tag{2.26}$$

This definition allows us to write Equation 2.20 in terms of the Stokes parameters:

$$\mathcal{V} = (\mathbf{G}_i \otimes \mathbf{G}_j^*)(\mathbf{P} \otimes \mathbf{P}) \overleftrightarrow{S}^{-1} \left\langle \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} \right\rangle, \tag{2.27}$$

which allows us to write the parallactic rotation of the Stokes parameters in a convenient form:

$$\overleftrightarrow{P} = (\mathbf{P} \otimes \mathbf{P}) \overleftrightarrow{S}^{-1} = \frac{1}{2} \begin{pmatrix} 1 & \cos 2\psi & \sin 2\psi & 0 \\ 0 & -\sin 2\psi & \cos 2\psi & i \\ 0 & -\sin 2\psi & \cos 2\psi & -i \\ 1 & -\cos 2\psi & -\sin 2\psi & 0 \end{pmatrix}, \tag{2.28}$$

which looks confusing until we make one observation. The Stokes parameters defined on the celestial sphere (in the α, δ basis) are different from those observed (in the x, y basis). If we define topocentric, observed Stokes parameters, ($I' = |E_x|^2 + |E_y|^2$, etc.), we find that observed Q' and U' are rotated from the celestial Q and U by an angle 2ψ, while I and V

remain fixed. We can represent this transformation as

$$\begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix} = \vec{\vec{S}}(\mathbf{P}^{-1} \otimes \mathbf{P}^{-1})\vec{\vec{S}}^{-1} \begin{pmatrix} I' \\ Q' \\ U' \\ V' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\psi & -\sin 2\psi & 0 \\ 0 & \sin 2\psi & \cos 2\psi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} I' \\ Q' \\ U' \\ V' \end{pmatrix}, \tag{2.29}$$

which aligns with our intuition that Q and U rotate between frames, but I and V remain constant.

Finally, we can write the parallactic rotation from the point of view from the feeds as

$$\vec{\vec{P}}' = (\mathbf{P} \otimes \mathbf{P})\vec{\vec{S}}^{-1}\vec{\vec{S}}(\mathbf{P}^{-1} \otimes \mathbf{P}^{-1})\vec{\vec{S}}^{-1} = \vec{\vec{S}}^{-1}, \tag{2.30}$$

where we have taken advantage of another property of the Kronecker outer product, that $\mathbf{A}^{-1} \otimes \mathbf{B}^{-1} = (\mathbf{A} \otimes \mathbf{B})^{-1}$. A curious reader at this point will ask why we bothered defining the visibilities in terms of the equatorially defined Stokes parameters – since the interferometer measures in topocentrically defined coordinates, why not use those? The answer is because of the time-dependence of parallactic rotation. A visibility measured at some time with some pointing is not equal to the same visibility with the same pointing at a later time. We've written the matrices and carefully defined each rotation and projection to define visibilities in a fixed coordinate system, and furthermore, to develop intuition about the rotations between Stokes parameters and their projection onto the xy plane.

We can use Figure 2.1 to visualize the parallactic rotation of Q and U, that is, the rotation from Q', U' to Q, U. In this figure, all symbols represent Q, since they align with lines of constant right ascension and declination. However, as the + sign "rises" and "sets," increasing or decreasing in hour angle, it rotates into $\times$ (and back again). The + and $\times$ in Figure 2.1 represent Q' and U', respectively.

To complete this brief discussion of interferometric polarimetry, we present the full equations for a visibility measured using linear feeds:

$$\mathcal{V} = \left\langle (\mathbf{G}_i \mathbf{P} \vec{E}_{\alpha\delta}) \otimes (\mathbf{G}_j \mathbf{P} \vec{E}_{\alpha\delta})^* \right\rangle, \tag{2.31}$$

or, writing each component explicitly in terms of topocentric Stokes parameters:

$$\mathcal{V}_{xx} = \int A_{xx}(\hat{s}) g_i^x g_j^{x*} e^{-2\pi i \nu(\tau_i^x - \tau_j^x)} \left[I'(\hat{s}) + Q'(\hat{s})\right] e^{-2\pi i \nu(\vec{b}/c)\cdot\hat{s}} \, \mathrm{d}\Omega \tag{2.32}$$

$$\mathcal{V}_{xy} = \int A_{xy}(\hat{s}) g_i^x g_j^{y*} e^{-2\pi i \nu(\tau_i^x - \tau_j^y)} \left[U'(\hat{s}) + iV'(\hat{s})\right] e^{-2\pi i \nu(\vec{b}/c)\cdot\hat{s}} \, \mathrm{d}\Omega \tag{2.33}$$

$$\mathcal{V}_{yx} = \int A_{yx}(\hat{s}) g_i^y g_j^{x*} e^{-2\pi i \nu(\tau_i^y - \tau_j^x)} \left[U'(\hat{s}) - iV'(\hat{s})\right] e^{-2\pi i \nu(\vec{b}/c)\cdot\hat{s}} \, \mathrm{d}\Omega \tag{2.34}$$

$$\mathcal{V}_{yy} = \int A_{yy}(\hat{s}) g_i^y g_j^{y*} e^{-2\pi i \nu(\tau_i^y - \tau_j^y)} \left[I'(\hat{s}) - Q'(\hat{s})\right] e^{-2\pi i \nu(\vec{b}/c)\cdot\hat{s}} \, \mathrm{d}\Omega \tag{2.35}$$

It is often convenient to rotate the linearly-polarized visibilities as the linearly-polarized images are rotated into Stokes parameters. This does not exactly represent the true Fourier-transformed Stokes parameters, but it does provide a useful approximation. Section 4.3 will discuss one of the negative consequences of such a rotation. For completeness, we provide the definition of Stokes visibilities:

$$\begin{pmatrix} \mathcal{V}_I \\ \mathcal{V}_Q \\ \mathcal{V}_U \\ \mathcal{V}_V \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -i & i & 0 \end{pmatrix} \begin{pmatrix} \mathcal{V}_{xx} \\ \mathcal{V}_{xy} \\ \mathcal{V}_{yx} \\ \mathcal{V}_{yy} \end{pmatrix} \tag{2.36}$$

2.3 The Time and Frequency Dependence of Visibilities

In this section, we go through a thorough investigation of the $\vec{b} \cdot \hat{s}$ term in Equation 2.8. The discussion here will largely follow Parsons and Backer 2009 (63) and Appendix A of Parsons, et al. 2014 (65). It will also serve as a precursor to Section 2.4 and be used as a reference for discussing the compression techniques in Section 6.1.

We begin by expanding $(\vec{b}/c) \cdot \hat{s}$ in the bases defined in Section 2.2, using the equatorial to topocentric rotation matrix $\mathbf{R}$. We note that in topocentric coordinates, $\hat{s} = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}^T$,

since the uv-plane is defined to be tangent to the celestial sphere at $\hat{s}$. Written in this basis,

$$\begin{aligned}\vec{b}\cdot\hat{s} &= \begin{pmatrix} b_x & b_y & b_z \end{pmatrix} \begin{pmatrix} \sin H & -\sin\delta\cos H & \cos\delta\cos H \\ \cos H & \sin\delta\sin H & -\cos\delta\sin H \\ 0 & \cos\delta & \sin\delta \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ &= b_x\cos\delta\cos H - b_y\cos\delta\sin H + b_z\sin\delta, \end{aligned} \tag{2.37}$$

where H is the hour angle of the pointing, and δ is the declination. We notice that the phase of a visibility is $2\pi i\nu(\vec{b}/c)\cdot\hat{s}$, and define a time-constant of the visibility, $\tau_g = (\vec{b}/c)\cdot\hat{s}$, called the geometric delay. This is simply the time delay of the signal between the two antennae in a basline. This argument indicates that the geometric delay is bound between the positive and negative baseline lengths, $-|\vec{b}| \leq \tau_g \leq |\vec{b}|$, whose values are realized when $\hat{s}$ aligns with the basline vector, i.e. the interferometer is pointing along the bore sight of its baseline.

Neglecting calibration terms, we may write the visibility (Equation 2.8) in terms of τ_g:

$$V = \int A(\hat{s},\nu)I(\hat{s},\nu)e^{-2\pi i\nu\tau_g}\ \mathrm{d}\Omega. \tag{2.38}$$

Noting the similariry of Equation 2.38 to a Fourier transform with respect to ν, we define the delay transform as the inverse Fourier transform of a visibility with respect to frequency:

$$\widetilde{V}(\tau) = \int Ve^{+2\pi i\nu\tau}\ \mathrm{d}\nu = \int A(\hat{s},\nu)I(\hat{s},\nu)e^{-2\pi i\nu(\tau_g-\tau)}\ \mathrm{d}\Omega\ \mathrm{d}\nu, \tag{2.39}$$

the convolution of the beam-weighted image with a delta-function kernel, peaked at τ_g. Written more explicitly in terms of the convolution operation $\star$, we have

$$\widetilde{V} = \widetilde{A}(\tau)\star\widetilde{I}(\tau)\star\delta\left(\tau-\tau_g\right). \tag{2.40}$$

To build intuition about this tranform, we can neglect the primary beam A, and assume that the only incident radiation is due to a flat-spectrum point source, i.e. $I(\hat{s},\nu) = I_o\delta(\hat{s}-\hat{s}_0)$. This assumption reduces Equation 2.39 simply to $\widetilde{V}(\tau) = I_0\delta(\tau-\tau_g(\hat{s}_0))$, exhibiting the important property of the delay spectrum, that it isolates smooth-spectrum, point sources in a space accessible to all baslines individually. The isolation of smooth-spectrum point

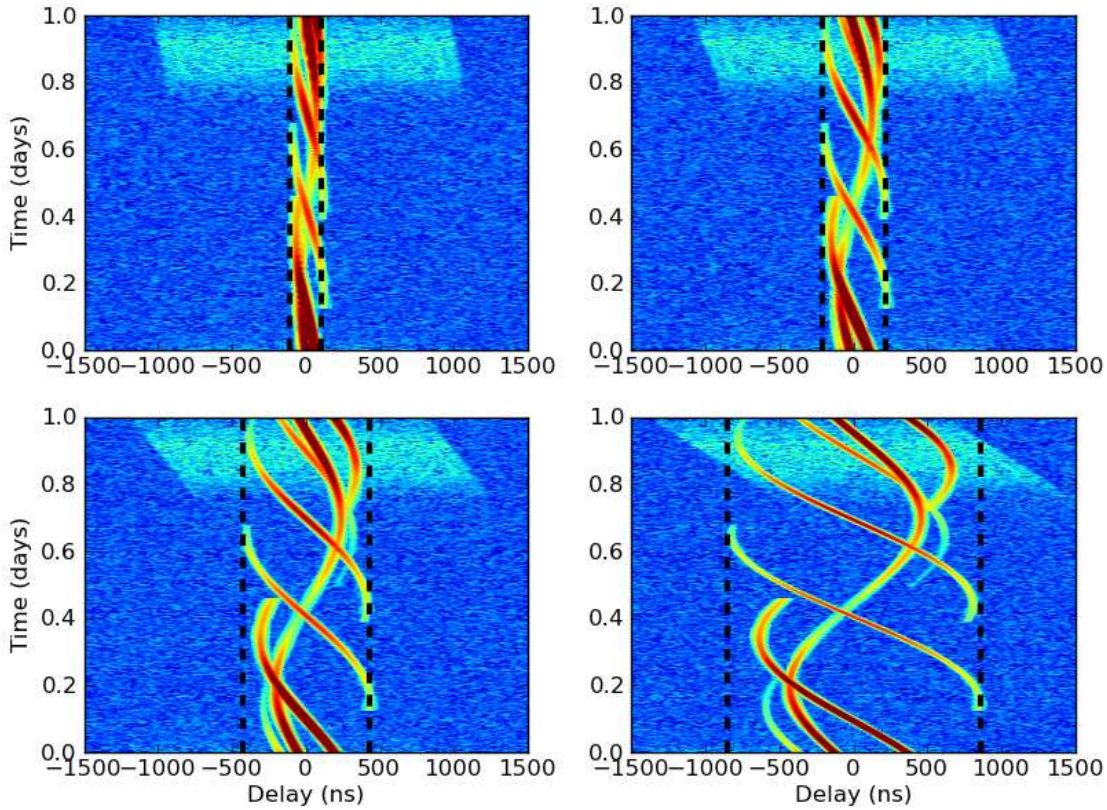

Figure 2.2: Delay spectra of baslines of different lengths of five simulated sources. Four of these sources have smooth, power law spectra, and one (which "turns on" at 0.8 days) with a non-smooth spectrum. Horizon limits for the four baslines are shown by black, dashed lines — these correspond to the baseline lengths of 32 m (top left), 64m (top right), 128 m (bottom left), and 256 m (bottom right). Visibilities are computed over the PAPER band, spanning frequencies of 100 MHz to 200 MHz. Color scale denotes the flux, with red showing the brightest sources and blue showing the dimmest, but the absolute scaling of the flux scale is arbitrary. Figure credit: Parsons et al. (66).

soures is demonstrated in Figure 2.2 — this figure also demonstrates the restriction of smooth-spectrum emission to within the horizon limit of a baseline.

Once we understand that the geometric delay of a source changes with time,[10] we can apply a similar technique in the time-domain. We define the fringe-rate transform of a visibility thus:

$$\widehat{V}(\nu, f) = \int V(\nu, t) e^{-2\pi i f t} \, \mathrm{d}t \tag{2.41}$$

Assuming again that we may neglect calibration terms and may only focus on a single, smooth-spectrum, point source, we find the fringe-rate transform isolates sources in the similar manner as a delay transform:

$$\begin{aligned} \widehat{V}(\nu, f) &= \int \left[\int I_0 \delta(\hat{s} - \hat{s}_0) e^{-2\pi i \tau_g \nu} \, \mathrm{d}\Omega \right] e^{-2\pi i f t} \, \mathrm{d}t \\ &\propto \int e^{-2\pi i (\tau_g \nu + f t)} \, \mathrm{d}t \approx \delta \left(f + \frac{d\tau_g}{dt} \nu \right), \end{aligned} \tag{2.42}$$

where τ_g is evaluated at $\hat{s}_0$ after the integration with respect to the sky coordinates is performed. This allows us to immediately read off the fringe-rate of a source,

$$f_0(\hat{s}) = \nu \frac{d\tau_g}{dt} = -\omega_\oplus \cos\delta \left((\nu/c) b_x \cos H + (\nu/c) b_y \sin H \right), \tag{2.43}$$

where $\omega_\oplus = -dH/dt$, the angular frequency of the earth's rotation.

Notice that $b_x \cos H + b_y \sin H$ is simply the east-west portion of the basline (in topocentric coordinates), and the linear velocity of the earth's rotation at latitude δ is proportional to $\omega_\oplus \cos\delta$ — the fringe-rate is simply the dot product of the basline with the angular velocity of the earth!

[10] Many sources ((83), e.g.) call this the fringe. We will accept this nomenclature, reserving the term "delay rate" for the time-dependence of the geometric delay, rather than the visibility. Hence, a fringe-rate transform will be the Fourier transform w.r.t. time of a visibility, but the delay-rate transform will be that of a delay-transformed visiblity. The difference is subtle, and won't appear in this work, but nonetheless, it is great enough to warrant two different names for the two different transforms.

To complete the analogy with the delay transform, we note that the fringe rate of a source is limited to $-(\nu/c)b_E \cos\delta \leq f_0/\omega_\oplus \leq (\nu/c)b_E$, where b_E is the topocentric, east-west component of the basline.

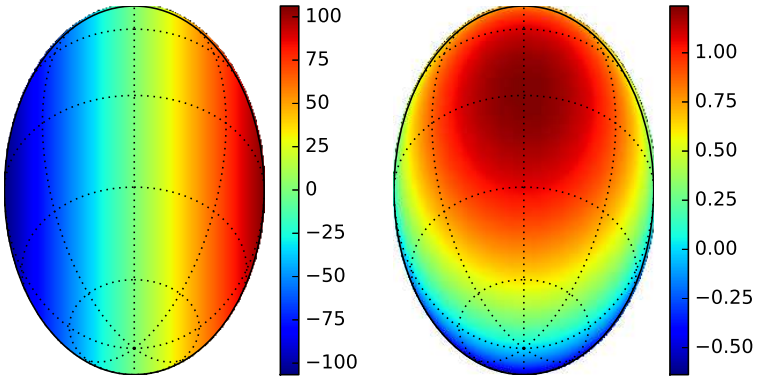

Figure 2.3: (Left Panel) Delay, in nanoseconds, of a 32m (106ns), east-west baseline at the latitude $-31°43'17.5''$. The delay ranges from $\pm|\vec{b}|$ and is constant along the north-south axis. (Right panel) Fringe rate of that same baseline, in mHz.

Figures 2.3 and 2.4 show the limits of the delay and fringe rate transform. Figure 2.3 shows the mapping of delay and fringe rate onto the sky, and Figure 2.4 shows evidence for the horizon limits in PAPER data.

2.4 The Delay Spectrum

At this point, we can discuss what may be the secondmost important tool used in this thesis — interferometry being the first — the delay spectrum method to power spectrum estimation. The development of the delay spectrum approach was first presented in Parsons et al. (66), and has since been implemented in a number of papers (39, 56, 65).

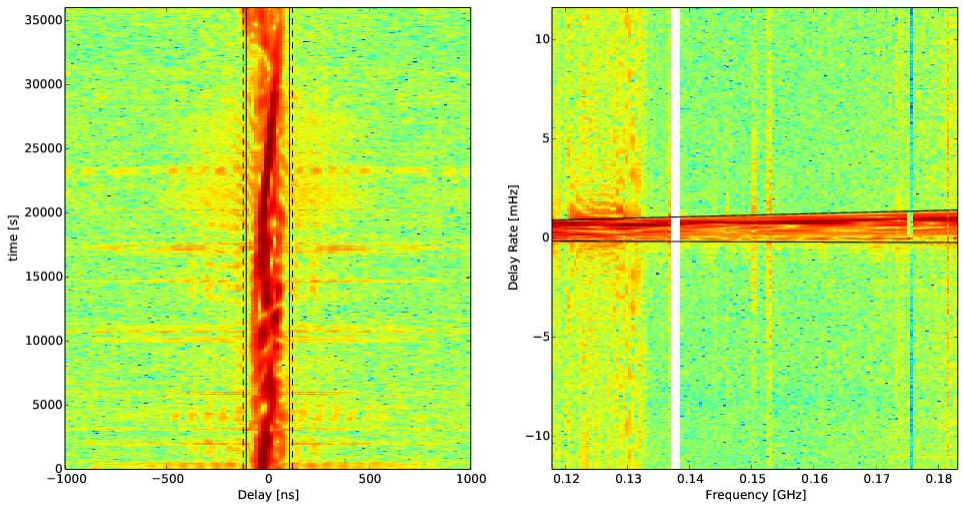

Figure 2.4: (Left Panel) Delay spectrum of one day's worth of PAPER data vs. time. A solid black lines showing the horizon limit of the 30 m baseline used to take this data. (Right Panel) Delay rate spectrum of the same data vs. frequency. Again, black lines denote the horizon limits of the fringe rate. In both plot, the colorscale denotes flux, with red being the brightest, but the absolute flux scale is arbitrary.

When I began work on this thesis, I would have considered the delay spectrum a novel approach (and it was!), and I would have presented it in constrast to what was then the "standard method" (10, 22, 48, e.g.). Since its inception, however, the three major players on the 21cm EoR stage have all diverged in their methods, each further from early approaches than the last, so I will present the delay spectrum approach from first principles.

To summarize the process, we will attempt to measure the three-dimensional power spectrum of a temperature field (Section 1.4) from some function of the visibility equation (Secion 2.1). Calculating power spectra directly from visibilities is a method that was first used[11] to compute the two-dimensional C_ℓ spectrum for the CMB using the DASI experiment (44) using a full covariance analysis of the visibilities (87). We aim to extend some of their work into the third dimension in the sense that we will be using visibilities as direct tracers of a power spectrum, but will wildly simplify the analysis with approximations of the primary beam and the signal. The two main points are these:

1. Transverse wavemodes $k_\perp \propto \ell$ are measured by the Fourier transform along the two dimensions on the sky. Since the visibility natively measures these modes, it natively measures the Fourier-transformed intensity or temperature field of the sky.

2. The line-of-sight k-modes, $k_{||}$, may be measured by the Fourier-dual varaible to frequency, since we compute a one-to-one mapping of frequency to cosmological distance.

We'll start with the flat-sky approximation of the visibility:

$$\mathcal{V}(u,v,\nu) = \int A(l,m,\nu) I(l,m,\nu) e^{-2\pi i(ul+vm)} \, \mathrm{d}l \, \mathrm{d}m. \tag{2.44}$$

Guided by the the one-to-one mapping between cosmological distance and frequency, we can compute the Fourier transform with respect to frequency of the visibility (i.e. the delay transform of Section 2.3) as

$$\widetilde{\mathcal{V}}(u,v,\eta) = \int A(l,m,\nu) I(l,m,\nu) e^{-2\pi i(ul+vm+\eta\nu)} \, \mathrm{d}l \, \mathrm{d}m \, \mathrm{d}\nu. \tag{2.45}$$

[11] And in my opinion, perfected.

If we restric our efforts to a single baseline, this cannot exactly represent a three-dimensional Fourier transform. The frequency-dependence of u and v sampled by a single baseline prohibits a totally independent transform along the ν axis. The kernel of this transform would be given by the expression

$$K(\eta|\vec{b}) = \int \tilde{A}\left(u - \nu\frac{b_x}{c}, v - \nu\frac{b_y}{c}, \nu\right) e^{-2\pi i\nu\eta}\ \mathrm{d}\nu, \tag{2.46}$$

where the integral is computed over the bandwidth of the observation and $\tilde{A}$ is the Fourier transform of the primary beam over l and m. This kernel approaches a delta function as $\nu(b_x/c)$ approaches 0, or to first order in the magnitude of $\vec{b}$, as $\Delta\nu|b|/c \ll 1$. Intuitively, the baseline length cannot change appreciably across the bandwidth, and more formally, the phase along the components of $k_\perp$ must remain coherent along the frequency direction. Figure 2.5 shows the path of a baseline along the $k_\perp$-ν plane alongside a fringe in $k_\perp$. In this figure, as baslines cross the fringes in blue, the baseline does not integrate over $\vec{k}$ in phase, and the delay transform loses its correspondence with a Fourier transform in frequency.

Moving forward with the assumption that our baselines are small enough that a delay transform approximates a Fourier transform along the frequency axis, we convert the units of a visibility to temperature, always in the Rayleigh-Jeans limit, as

$$\widetilde{\mathcal{V}}(u,v,\eta) \approx \frac{2k_B}{\lambda^2}\int A(l,m,\nu)T(l,m,\nu)e^{-2\pi i(ul+vm+\eta\nu)}\ \mathrm{d}l\ \mathrm{d}m\ \mathrm{d}\nu. \tag{2.47}$$

where an assumption is made that the intensity-to-temperature conversion remains constant over the band. We cross-multiply two instances of the visibility to give

$$\begin{aligned}\left|\widetilde{\mathcal{V}}(u,v,\eta)\right|^2 \approx \left(\frac{2k_B}{\lambda^2}\right)^2 \int A(l,m,\nu)A(l',m',\nu')T(l,m,\nu)T(l',m',\nu')\times \\ \times e^{-2\pi i(u(l-l')+v(m-m')+\eta(\nu-\nu'))}\ \mathrm{d}l\ \mathrm{d}m\ \mathrm{d}\nu\ \mathrm{d}l'\ \mathrm{d}m'\ \mathrm{d}\nu'.\end{aligned} \tag{2.48}$$

We can assume that the primary beam is a tophat function in l, m, and ν, and that it spans an area Ω in l and m, and $\Delta\nu$ in ν. These assumptions allow us to remove the primary

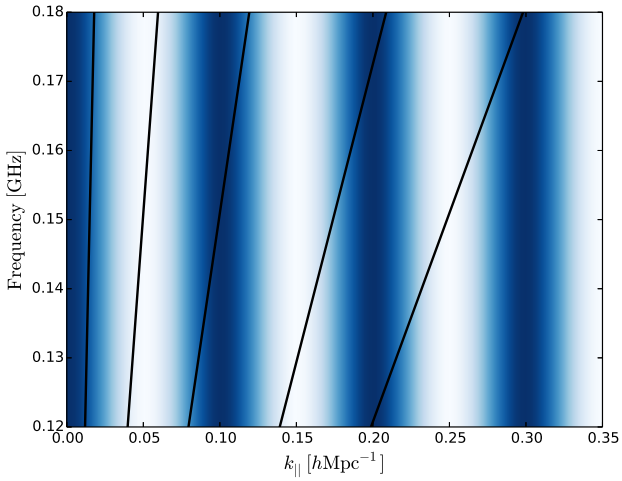

Figure 2.5: Baseline tracks through the $k_\perp$-ν plane. From left to right, the baseline lengths are 30, 100, 200, 350, and 500 m. Shown in blue is a fringe in $k_\perp$.

beam from within the integral and set limits to the integral:

$$\left|\widetilde{\mathcal{V}}(u,v,\eta)\right|^2 \approx \left(\frac{2k_B}{\lambda^2}\right)^2 \int_{(0,0,0)}^{(\sqrt{\Omega},\sqrt{\Omega},\Delta\nu)} \mathrm{d}l'\ \mathrm{d}m'\ \mathrm{d}\nu' \int_{(0,0,0)}^{(\sqrt{\Omega},\sqrt{\Omega},\Delta\nu)} \mathrm{d}l\ \mathrm{d}m\ \mathrm{d}\nu \times T(l,m,\nu)T(l',m',\nu')e^{-2\pi i(u(l-l')+v(m-m')+\eta(\nu-\nu'))}. \tag{2.49}$$

Changing variables $(l,m,\nu) \rightarrow (l_r,m_r,\nu_r) \equiv (l-l',m-m',\nu-\nu')$ and integrating over the dummy variables (l_r, e.g.) yields

$$\left|\widetilde{\mathcal{V}}(u,v,\eta)\right|^2 \approx \Omega\Delta\nu \left(\frac{2k_B}{\lambda^2}\right)^2 \int_{(-\sqrt{\Omega},-\sqrt{\Omega},-\Delta\nu)}^{(\sqrt{\Omega},\sqrt{\Omega},\Delta\nu)} \xi_{21}(l_r,m_r,\nu_r)e^{-2\pi i(ul_r+vm_r+\eta\nu_r)}\ \mathrm{d}l_r\ \mathrm{d}m_r\ \mathrm{d}\nu_r, \tag{2.50}$$

where ξ_{21} is the correlation function of the temperature field, defined in Equation 1.11. As in Parsons et al. (62), we use the subscript 21 to denote quantities derived explicitly for measuring 21cm reionization. Following Equation 1.11, we note that $\Omega\Delta\nu$ is the cosmological volume over which we are sampling in the native units of an interferometer.

At this point we invoke the interpretations of the three axes and their duals, $(u, v, \eta) \rightleftharpoons (k_x, k_y, k_z)$, that they measure the cosmological wavemodes given in Section 1.4. This allows us to change units according to the Jacobian

$$J = \begin{pmatrix} X & 0 & 0 \\ 0 & X & 0 \\ 0 & 0 & Y \end{pmatrix}, \tag{2.51}$$

where X and Y are defined such that $2\pi(ul + vm + \eta\nu) = k_x x/X + k_y y/X + k_z z/Y$ for comoving coordinates x, y, z. The specific values for X and Y are given in Equations 1.17 and 1.18, respectively. This coordinate transformation allows us to write Equation 2.50 in terms of the familiar quantities

$$\left|\widetilde{\mathcal{V}}(u, v, \eta)\right|^2 \approx \frac{\Omega\Delta\nu}{X^2 Y}\left(\frac{2k_B}{\lambda^2}\right)^2 \int_{(-X\sqrt{\Omega}, -X\sqrt{\Omega}, -Y\Delta\nu)}^{(X\sqrt{\Omega}, X\sqrt{\Omega}, Y\Delta\nu)} \xi_{21}(\mathbf{r}) e^{-i\mathbf{k}\cdot\mathbf{r}}\, \mathrm{d}^3 r \tag{2.52}$$

If we can assume that the limits of the integral span many phase-wrappings of $\mathbf{k} \cdot \mathbf{r}$ — for instance, if $\sqrt{\Omega} \ll 1/(2Xk_x)$ — then the integral becomes a power spectrum, and we can write Equation 2.52 as

$$\left|\widetilde{\mathcal{V}}(u, v, \eta)\right|^2 \approx \frac{\Omega\Delta\nu}{X^2 Y}\left(\frac{2k_B}{\lambda^2}\right)^2 P_{21}(\mathbf{k}), \tag{2.53}$$

where $\mathbf{k}$ is chosen by the (u, v, η) coordinate in the correct coordinate system. Finally, by noting that the comoving area subtended by the primary beam is $X^2\Omega \equiv D^2$ and the comoving distance spanned by the band is $Y\Delta\nu \equiv \Delta D$, then we can write the power spectrum of 21cm fluctuations in terms of the visibility and the cosmological distances measured as

$$P_{21}(\mathbf{k}) = \left(\frac{\lambda^2}{2k_B}\right)^2 \frac{(\Omega\Delta\nu)^2}{D^2\Delta D}\left|\widetilde{\mathcal{V}}(u, v, \eta)\right|^2 \tag{2.54}$$

Chapter 3

The Donald C. Backer Precision Array to Probe the Epoch of Reionization

The Donald C. Backer Precision Array to Probe the Epoch of Reionization (PAPER) was built specifically to detect the 21cm EoR power spectrum in mind. To achieve this goal, many design choices violate the assumptions made in normal interferometric analysis — the two most pertinent of these are the flat sky approximation, and the assumption that one source dominates the field of view. This chapter will describe the instrument and discuss the impetus for many characteristics of the array. Much of this material is discussed in Parsons et al. (64), but several improvements have been made since 2009, when that paper was written.

3.1 Instrument Design

PAPER is located in the Karoo desert, in the Northern Cape province of South Africa, around 60km west of Carnarvon, at latitude and longitude of $30°43'17.5''$S, $21°25'41.9''$E.

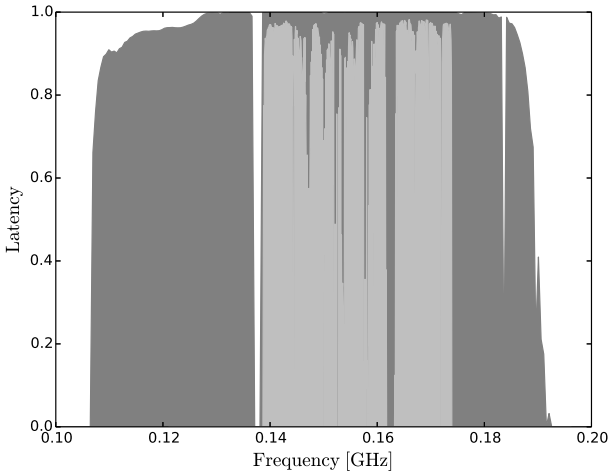

Figure 3.1: Comparison of RFI environments in South Africa (dark grey) and Green Bank, West Virginia (light grey). The value shown is the probability that a datum survives RFI excision.

The array sits approximately 1km east of the KAT7 array, and roughly 0.5km south of the precursor to the Square Kilometer Array, MeerKAT.[12] Such a remote location is chosen to isolate PAPER from human-generated radio frequency interference (RFI). Figure 3.1 compares the latency of RFI with the site in South Africa to that of the radio-quiet zone in Green Bank, West Virginia, where we operate a test array. Though the South African site has a pristine RFI environment, certain frequencies must be discarded at all times — most notably, for the Orbcomm constellation of tracking satellites at 137 MHz, and for the International Space Station at 150 MHz.

Figure 3.2: Photograph of a PAPER element, showing the groundscreen and the sleeved dipole.

Celestial signal enters through a dipole hoisted above a wire-mesh groundscreen. Four mesh flaps positioned at roughly a 45° angle above the groundscreen increase the response of the dipoles towards zenith. The dipole is placed within a sleeve, which broadens the response over a wider range of frequencies, and allows for better impedance matching of the element over such a wide band. All electrical elements of the antenna are isolated from one

[12] www.ska.ac.za

another by plastic fittings, which are transparent to radio frequencies. Figure 3.2 shows a photograph of a PAPER element.

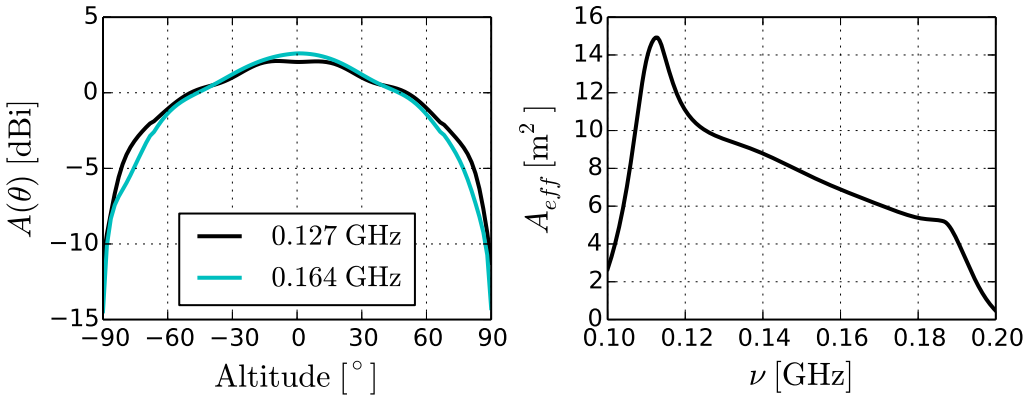

Figure 3.3: (Left panel) East-west cuts through zenith of the PAPER beam, at 127 MHz (cyan) and 164 MHz (black). In these units, isotropic emission would have 0 dBi at all angles. These are the two central frequencies of the results in Chapter 6. (Right panel) Effective area of an antenna as a function of frequency.

The groundscreen is designed to have a relatively low effective area so that the first nulls of the dipole response occur below the horizon. Since the location of these nulls is highly frequency-dependent, they can introduce systematic errors to the high $k_{||}$ modes reserved for EoR analysis. Such a small effective area results in a large field of view. Figure 3.3 shows an axial cut through the beam at two frequencies and the effective area A_{eff} of the element, defined as $A_{eff}\Omega = \lambda^2$, where Ω is the field of view.

The first stage of amplification occurs at the dipole, boosting the signal by 60 dB. Signal then propagates along 75Ω, coaxial cable[13]. The length of all coaxial cables is set to be roughly equal to minimize the differences of signal travel time along these cables, which lessens the calibration burden. These coaxial cables pass into an RFI-shielded enclosure

[13]The same coming out of the back of your TV!

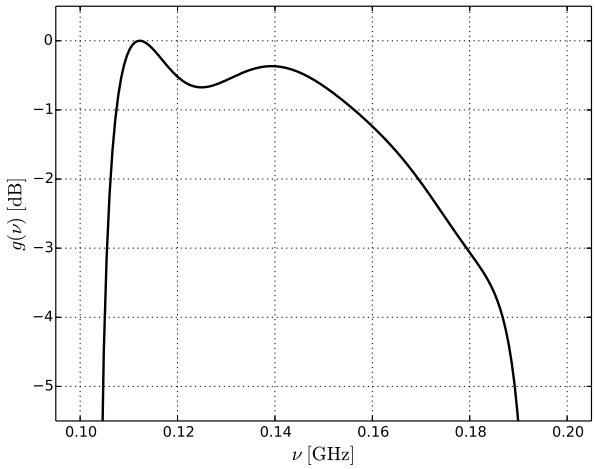

Figure 3.4: Model of the PAPER bandpass, normalized to peak at one.

containing receivers and the digital equipment. A final amplification stage (40 dB), located inside the RFI-enclosure, corrects for signal loss along the cables, and an analog bandpass filter is applied. This bandpass is designed to have a flat frequency response from 120 MHz to 180 MHz, and attenuate signal in the FM band below around 107 MHz and also attenuate signal at 200 MHz, which contains the total power signal. A plot of the bandpass (gain as a function of frequency) is shown in Figure 3.4.

Next, the signal is digitized with a sampling rate of 100 MHz. It is then passed through a F-engine, which computes the Fourier transform using a four-tap, polyphase filter bank. Since the sampling rate is below the frequencies allowed by the bandpass filter, we measure an aliased copy of the frequency spectrum, in the second Nyquist zone. Because the frequency channels are aliased and reversed, the DC signal is contained in the 200 MHz frequency bin. The integration time of each Fourier transform is set to be around 10 seconds. The Fourier-transformed signal is then distributed over 10GbE to the X-engine, which cross-multiplies each signal and computes the visibility for each antenna pair. The visibilities are finally

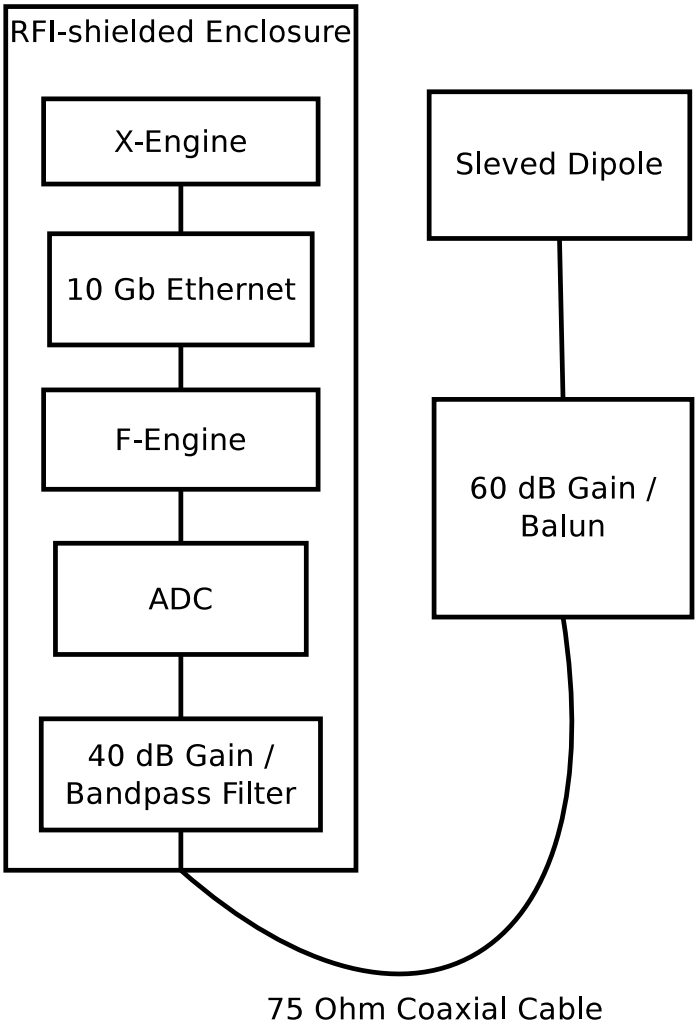

Figure 3.5: Flow chart of analog and digital systems.

written in MIRIAD format (76) and stored for analysis. Figure 3.5 shows a flowchart which summarizes the propagation of signal through the analog and digital systems.

For the EoR2012 and EoR 2013 campaigns described in Table 3.1, there is a final process after digitization and correlation. Since the data rate goes as the number of baselines (roughly, the number of antennae squared) campaigns with many-element arrays will have untenably large data rates. The extreme case is the EoR2013 campaign with 128 antennae. With 1024 frequency channels per 10 second integration for each of 10^4 baselines, this campaign has a data rate of over 200 Mbps. The total data volume of a 120 day season at this rate would be 127 TB, which, with storage costs of $150/TB, costs $19,000 just to keep.

To mitigate the "big data" problems of the latter observing seasons, we implement a data compression system. Data is piped over 1GbE into a large RAID array, and once the night's data is taken, a small cluster of compute nodes performs a low-pass filter and decimation on the data, outputting a smaller, more readily analyzed dataset. The algorithm for the compression process is described in detail in Section 6.1. This reduces the data rate (and the data volume) by roughly a factor of twenty.

Once the data is compressed, it is stored on small, portable disks and shipped to its final resting place, a computer cluster in Philadelphia. From this location, it can be processed using a larger cluster (`folio.sas.upenn.edu`) and served elsewhere in the United States. Figure 3.6 shows a map of this process.

3.2 Deployments of the PAPER Array

Table 3.1 summarizes PAPER's five major campaigns since I joined the group in 2009. The first two, PSA32 and PSA64 span an entire year, but due to lack of internet connectivity to site, consisted of two week-long seasons, separated by about six months. These two campaigns were designed to characterize foregrounds, so the configuration of antenna was designed to maximize *uv*-coverage. Data from PSA32 yielded a new catalog of sources in the southern hemisphere (37).

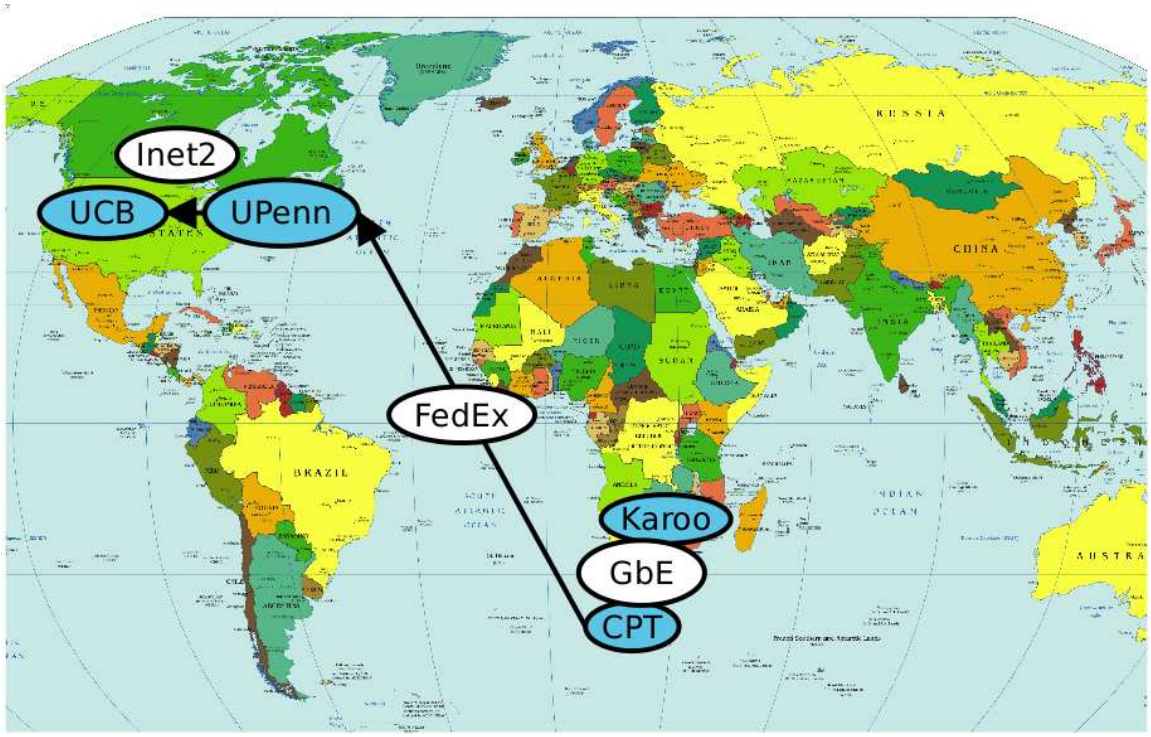

Figure 3.6: Map of the data flow. Data is taken on site in the Karoo desert, and shipped to Philadelphia via Cape Town (CPT). Data may then be served from Philadelphia to other institutions in the United States, such as the University of California (UCB).

	Julian Date	N_{ant}	N_{corr}	Configuration	Publications
PSA32	2455460-469, 538-544	32	32	Imaging	(37)
PSA64	2455743-747, 817-822	64	64	Imaging	(38, 69, 78)
EoR2011	2455903-6006	32	64	8×4 Grid	(39, 56, 65)
EoR2012	2456261-383	64	128	8×8 Grid	—
EoR2013	2456620-637	128	512	16×7 Grid, 16 Outliers	—

Table 3.1: PAPER Campaigns from 2009 through 2014. N_{ant} is the number of antennae, and N_{corr} is the number of inputs to the correlator.

In the PSA64 campaign, we doubled both the number of antennae and the number of inputs to the correlator. Single-polarization data provided confusion-limited images which yielded one of the first measurements of the foreground "wedge" (69), described in Section 1.5, and shown in Figure 1.6. It also provided precision measurements of the flux of several calibration sources (38), and a detailed analysis of the spectral structure of Centaurus A (78). All four polarization products were correlated on subarrays of 32 antennae to give data for full-Stokes images.

In 2011, because we had reached the confusion limit in our images, we reconfigured the array into a grid pattern. Parsons et al. (62) show the sensitivity benefits of an antenna configuration which maximizes the number of redundant baselines. The argument can be summarized by noting that averaging redundant baselines reduces the variance of $T(\mathbf{k})$ by $1/n$, where n is the number of measurements, but averaging the power spectrum of baselines in the same annulus of constant k reduces the variance of $P(k)$ by $1/n$. Hence, redundant baselines reduce the uncertainty in the power spectrum by $1/n$, while non-redundant spacings reduce it by $1/\sqrt{n}$. With this fact in mind, and also considering the necessity of characterizing the polarized power spectrum (55), we arranged a subset of 32 antennae into an 8×4 grid, and took data for about six months. This campaign was our first long integration designed to measure the 21cm EoR power spectrum. It yeilded three papers: the first limits on X-ray heating of the IGM (65, Figure 1.7), sensitivity limits on multiple redshift bins (39), and a characterization of polarized foregrounds (56), the main result of this thesis.

Since the EoR2011 campaign, we have doubled the number of inputs to the correlator each year, and in 2013, we increased the number of antennae to 128. In this last campaign, we placed a small subset of antennae along the perimeter of a 300m circle, centered on the grid. This allows us to increase the imaging power of the array, allowing for more accurate foreground characterization and mitigation. Work is currently underway to analyze the data from these campaigns, but to date, results have yet to be published.

Part II

The Problem of Polarization

Chapter 4

Initial Thoughts

4.1 Polarization at Meter Wavelengths

Nearly all celestial emission at meter wavelengths comes from the synchrotron radiation of electrons. Synchrotron emission is natively polarized, so it stands to reason that all emission at radio frequencies is polarized. These polarized, radio sources yield a wealth of astrophysics in their own right, but since we are interested in detecting the 21cm EoR power spectrum, we focus on polarized sources' impact on its detection.

This section will discuss the provenance of polarized sources, give a summary of the current understanding of the polarized sky at meter wavelengths, and then discuss how they can affect efforts to measure the 21cm EoR power spectrum.

4.1.1 Why is Synchrotron Emission Polarized?

To discuss the cause of polarized synchrotron emission, we review how synchrotron emission comes about. This discussion will summarize the appropriate chapters of Wilson et al. (89) and Rybicki and Lightman (74), but also loosely follows the original paper, Westfold (86)[14].

[14] In my opinion, an interesting read in how it differs from modern astrophysical calculations

We begin by recalling a few basics of Larmour precession. As an electron takes its helical path about a magnetic field, it emits a dipole radiation pattern perpendicular to both its acceleration (pointed toward the center of its helix) and the magnetic field. In the nonrelativistic limit, this is simply cyclotron emission. As the speed of the electron approaches c, the Larmour dipole pattern elongates in the direction of the electron's motion, forming a thin beam of emission preceding the electron.

An observer off bore-sight of the magnetic field will see a series of pulses in time from a single synchrotron electron — these pulses in the delta-train arrive at the Lorentz-boosted cyclotron frequency,

$$\nu_c \propto \gamma^3 \omega_B \approx \left[1 - \left(\frac{v}{c}\right)^2\right] \frac{eB}{m_e c}, \tag{4.1}$$

where ν_c is the observed frequency of the synchrotron emission, $\omega_B \equiv eB/\gamma mc$ is the cyclotron frequency, and γ is the Lorentz factor. Hence, each synchrotron electron emits radiation whose frequency is a function of that electron's energy.

We extend this notion to an ensemble of electrons. Since each contributes emission whose frequency is proportional to the energy of the electron, the total spectrum will be proportional to the energy spectrum of the ensemble of electrons.

The same notion can be extended to the polarization properties of synchrotron emission. One can write the full Larmour formula for the relativistic ensemble of electrons (as in the three sources cited), and find that the polarized fraction of synchrotron emission, p, exactly written in terms of modified Hankel functions (86), can be written for power-law energy distributions for the electrons, $N(E) \propto E^{-n}$, simply as

$$p = \frac{n+1}{n+7/3}. \tag{4.2}$$

The direction of polarization is perpendicular to both the magnetic field and the line of sight.

Finding n in Equation 4.2 from the spectral slope of the galaxy implies a polarized fraction of around 75%. This is a couple orders of magnitude higher that what is measured, which we will discuss later in this section.

4.1.2 Faraday Rotation

The next most important process in radio polarimetry is the rotation of the polarization vector through an ionized, magnetized plasma.[15]

An easy way to understand Faraday rotation is through birefringence. If a polarized wavefront is incident on a Faraday screen, a thin layer of electrons with a constant magnetic field through them, then one circular polarization — the one which opposes cyclotron motion of electrons in that magnetic field — is slowed with respect to the other.

Integrating through the Faraday screen, we find that the phase difference between the two polarizations can be written like a column density:

$$\Delta\varphi = \frac{e^3}{(m_e c^2)^2}\lambda^2 \int n_e(s) B_{||}(s)\ \mathrm{d}s \equiv \lambda^2 \Phi, \tag{4.3}$$

which defines the rotation measure Φ. In this expression, m_e and e are the mass and charge of the electron, $n_e(s)$ is the electron density along the line of sight, and $B_{||}$ is the component of the magnetic field that aligns with the line of sight. The integral extends from the observer to the emitting source. Since a Faraday screen rotates the polarization vector of a field, it affects Stokes parameters Q and U by rotating them by $2\Phi\lambda^2$:

$$(Q + iU)_{meas} = e^{-2i\Phi\lambda^2}(Q + iU)_{int}, \tag{4.4}$$

where the subscripts *meas* and *inc* refer to the measured and incident polarization states, respectively.

[15] Faraday rotation got its start in the optical, though we mostly see it in the radio since we measure the E-field directly in radio. Faraday rotation is much like interferometry in that sense (remember Michelson and Morley).

Measurements of the Faraday depth of sources can be used to characterize galactic magnetic fields (3, e.g.). To measure the rotation measure of sources, we can either fit a line in λ^2 to the polarization angle, or exploit the Fourier relationship between Φ and λ^2 in Equation 4.4 (11). A method to characterize the Faraday depths rotation the polarization vectors of a source is the topic of discussion in Chapter 7. Faraday rotation of sources is also mechanism which gives polarized sources spectral structure, which has led to the studies in the remainder of this thesis.

4.1.3 Methods of Depolarization

We will show in Section 6.3 that the mean polarized fraction of point sources at 150 MHz is less than 3×10^{-3}, a far cry from the 75% predicted from the spectral slope of our own galaxy.[16] The question for us to answer now is why there is such a disparity between these two polarized fractions. The remainder of this section will discuss methods of depolarization, which we split into two groups: instrumental, and intrinsic.

The first form of instrumental depolarization, called beam depolarization, arises from the minimum angular resolution element of an array. In general, galactic magnetic fields can be turbulent, so the polarization angle varies with position. A large synthesized beam sums many of these randomized polarization vectors, yielding a large amount of depolarization. The extent of this attenuation is dependent on the size of the synthesized beam, the angular distribution, and alignment of polarization vectors at a pointing.

The second method of instrumental polarization comes from integrating in frequency, and is called bandwidth depolarization. Since all polarized emission passing through magnetic fields undergoes Faraday rotation, it all gains spectral structure given by Equation 4.4. As we measure this emission, we must integrate over some bandwidth $\Delta\nu$. The rotation within

[16] This is the primary result of this thesis.

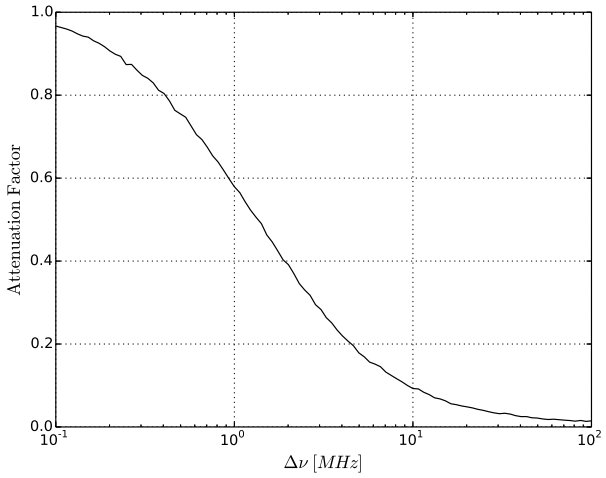

Figure 4.1: Attenuation factor from Equation 4.6 versus bandwidth for a band centered at 150 MHz, and a distribution of sources drawn from the Oppermann et al. (60) maps.

the band yields an attenuation given by

$$P_{meas} = \int W(\nu) P(\nu)\ \mathrm{d}\nu, \tag{4.5}$$

where $P(\nu)$ is the unaveraged power, P_{meas} is the measured power within the band, and $W(\nu)$ is the window function of the band. If we assume the window is a tophat function centered at frequency ν_0, with a width $\Delta\nu$, and if we also assume that the intrinsic power follows a flat-spectrum polarized source, $P(\nu) = P_0 \exp\{-2i\lambda^2\Phi\}$, then we can expect an attenuation rate of

$$\frac{P_{meas}}{P_0} = \int \mathbb{P}(\Phi) \left[\int_{\nu_0 - \Delta\nu/2}^{\nu_0 + \Delta\nu/2} e^{-2i\lambda^2\Phi}\ \mathrm{d}\nu \right]\ \mathrm{d}\Phi, \tag{4.6}$$

where $\mathbb{P}(\Phi)$ is the probability distribution of Faraday depths. Figure 4.1 shows the level of bandwidth depolarization from Equation 4.6 with a distribution of Faraday depths drawn from the Oppermann et al. (60) maps. Using the channel widths we use for the results

in Chapter 6, we expect nearly all polarization to be preserved by this mechanism. Using the entire bandwidth available to PAPER, around 80 MHz, attenuates the signal to around 10% of its intrinsic power. This effect strengthens at lower frequencies, since the $2\Phi\lambda^2$ phase-wrapping grows faster as the frequency decreases.

These two methods of depolarization are instrumental, meaning they can be mitigated by the correct design choices in an instrument. The next two forms of depolarization are intrinsic, meaning they arise by galactic physics. No data analysis methods or instrument design can remove these sources of depolarization.

The primary form of intrinsic depolarization is due to turbulent magnetic fields. The expression linking spectral slope with polarization fraction (Equation 4.2) assumes a constant magnetic field driving the electron's acceleration. This assumption is not true in general, but rather, galactic magnetic fields usually vary as a function of position, causing emission from different parts of the magnetic field to have different polarization angles. This effect causes the polarization in a line of sight to have emission from many different photons with different polarization angles. This amounts to an attenuation of polarized power. In the extreme case, a totally spatially random magnetic field, the net polarization is zero.

The secondary form of intrinsic depolarization is due to a polarized emitter's position inside a magnetized, ionized plasma. This effect is discussed at length in Jelić et al. (40), and we will briefly discuss it here. If an emitter lies within the plasma which Faraday-rotates its emission, different photons will travel through different lengths of the plasma, which by Equation 4.3 gives each a different Faraday depth. Many photons with different rotation measures developed this way may scatter into one line of sight, yielding a signal containing many Faraday-rotated components. Once this line of sight is summed in the instrument, the ensemble average of the polarized signal is attenuated by a factor proportional to $\exp\{-4\Delta\Phi\lambda^4\}$, where $\Delta\Phi^2$ is the variance of Faraday depths available in this line of sight (89).

These methods of depolarization, both instrumental and intrinsic, are discussed in Gaensler et al. (23), and Landecker et al. (46), to give two examples. The latter synthesizes information regarding depolarization into the notion of a polarization horizon. The polarization horizon is the maximum distance of a polarized source measurable by a given instrument — polarized emission produced beyond this horizon will be subjected to a level of instrumental polarization so great as to attenuate the signal below any reasonable detection threshold. Instruments like PAPER have polarization horizons on the order of around 10-20 kpc (5).

4.2 Recent Observations

The polarized sky at meter wavelengths is relatively uncharted territory, especially in the southern hemisphere.[17] Significant efforts have been made to characterize the polarized intensity of individual sources, the power spectrum of polarized emission, and the structure of magnetic fields of both our galaxy and extragalactic objects. Here, we will briefly describe the status of polarimetry at the time of writing this thesis. Many of these measurements were taken at 1.4 GHz, and we must extrapolate their properties down to 150 MHz.

There are two nearly complete surveys of the polarized sky: the NRAO VLA Sky Survey (NVSS, 14) and a survey using the Dominion Radio Astrophysical Observatory of Canada (DRAO, 90), both at 1.4 GHz. The former constructed a catalog of polarized point sources, alongside their rotation measures. The latter, with its shorter baselines, was more able to measure the extended polarized emission mostly due to our own galaxy. Both the Very Large Array (VLA) and DRAO are located in the northern hemisphere, so much of the southern sky blocked by the Earth. Figure 4.2 shows the map of polarized intensity derived from the DRAO survey(90). The polarized intensity in this map does not match well with unpolarized intensity (see Haslam et al. (33) or de Oliveira-Costa et al. (18) for examples).

[17] *coeli incogniti*

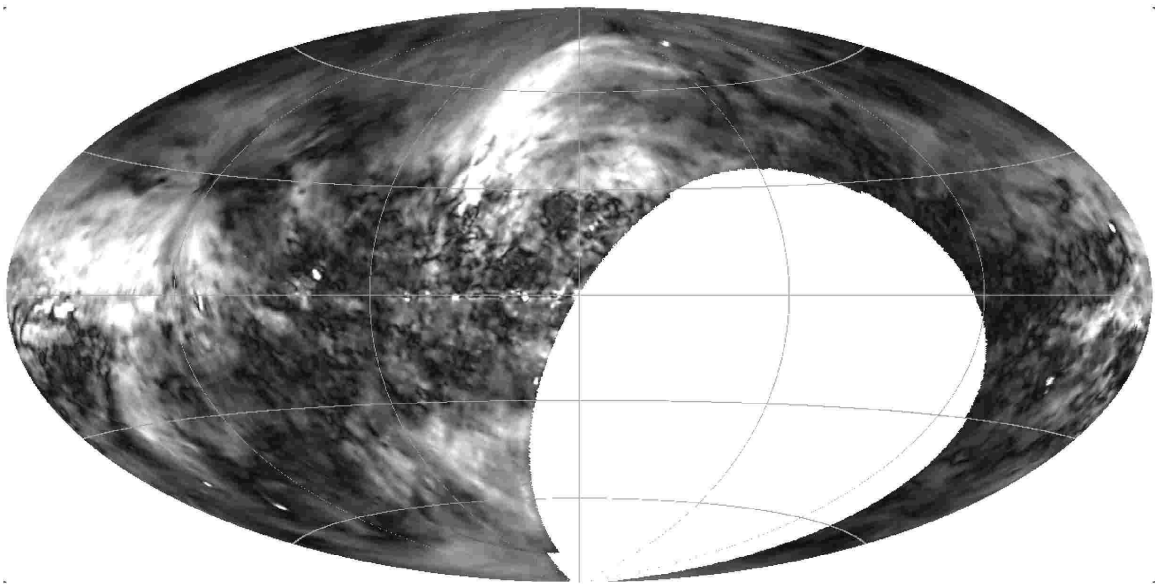

Figure 4.2: Polarized intensity in galactic coordinates from the DRAO survey (90). We present this to show the relative levels of polarized emission rather than the exact intensity. Lighter regions have a higher polarized intensity than darker regions.

For example, the north galactic spur is highly polarized in these maps[18], but the galactic plane is unpolarized due to turbulent magnetic fields.

A more recent measurement from Bernardi et al. (8) surveys 2400 square degrees, a stripe centered at the PAPER latitude, so within the PAPER field of view. These observations, alongside Wolleben et al. (90) and other measurements (40, e.g.), indicate that qualitatively, the majority of polarized emission is contained in diffuse structures. Bernardi et al. (8) does measure one point source, PMN J0351-2744, whose flux at 184 MHz is 320 Jy, and whose rotation measure is $+33$ m^{-2}. The discovery of this single source will be used to constrain the mean polarized flux at these frequencies in the analysis of Section 6.3.

To better characterize the relative levels of diffuse and point-like polarized emission, groups have measured its angular power spectrum. The first relevant upper limits were provided by Pen et al. (67) using the Giant Metrewave Radio Telescope (GMRT). They found, for spherical harmonic multipoles $200 \leq \ell \leq 5000$, and upper limit of around $C_\ell \lesssim 100$ mK^2. More recent work by Bernardi et al. (7) detected polarized power at the same level at $\ell \lesssim 1000$ using the Westerbork Synthesis Radio Telescope (WSRT), with no significant detection of power above ℓ of 1000. Bernardi et al. (7) did not detect emission directly attributable to polarized point sources.

Figure 4.3 gives a summary of the low-frequency measurements of polarized power spectra. In addition to the Pen et al. (67) and Bernardi et al. (7) measurements, It also shows the angular power spectrum of Haslam et al. (33), scaled by a mean polarized fraction of 0.3%, which roughly causes it to agree with the two measurements. This scaling requires a large degree of depolarization of the synchrotron emission from an ordered magnetic field — the assumption we used in calculating the expression for the expected polarization fraction for synchrotron radiation earlier in this section (Equation 4.2.

A recent simulation (41) attempted to further constrain the problem by fully simulating a full-Stokes realization of galactic synchrotron emission over a $10^\circ \times 10^\circ$ field of view. They

[18] Mostly due to its close proximity to us

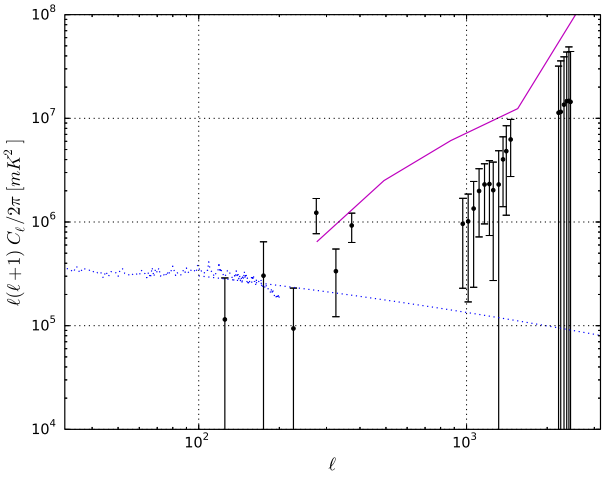

Figure 4.3: Recent measurements of low-frequency polarized power spectra. The thick black points show the Bernardi et al. (7) measurements of a field around 3C196, and the solid, magenta line shows the upper limit of Pen et al. (67). The Haslam map at 408 MHz (33), scaled by a polarization fraction of 0.3% is shown by thin, blue points, and a power-law extrapolation is shown with a dotted line above $\ell = 200$. This fraction was chosen to agree with the low-ℓ points in the Bernardi measurement. At high-ℓ, the upper limits do not constrain the level of polarized emission.

present a realistic spectrum of the mean temperature of polarized emission, but do not extend their analysis into the power spectrum. They also predict a polarization fraction from diffuse emission much higher than the limited available measurements allow.

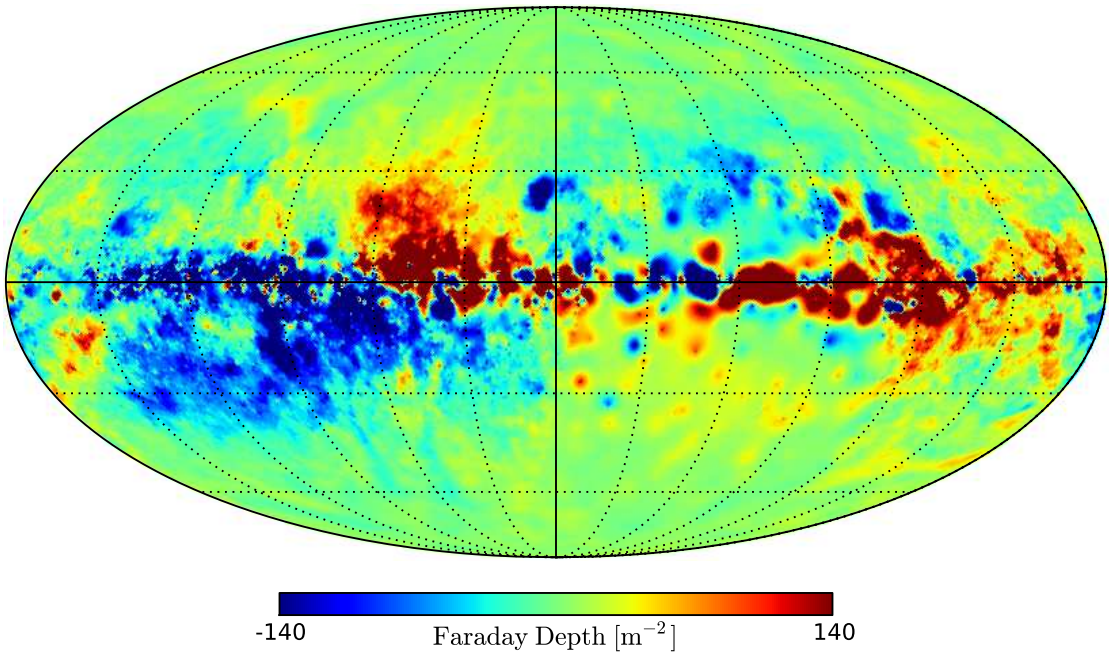

Figure 4.4: Map of Faraday depths, from Oppermann et al. (60), shown in galactic coordinates.

Finally, we turn to measurements of Faraday depth. The most comprehensive study is a meta-analysis by Oppermann et al. (60), which synthesizes a full-sky map of Faraday depth (Figure 4.4) from existing measurements, mostly from the NVSS polarization survey (14, 80). The quadrupole pattern of rotation measure indicates a large-scale ordering of the galactic magnetic field (45, 72), though turbulence of the magnetic field within the plane diminishes this effect. A recent study (59) finds that the dominant contribution to the

rotation measure of a source comes from magnetic fields within our galaxy.

4.3 Beam Leakage

Since the 21cm EoR power spectrum is largely unpolarized, we ask the question, why care about polarized emission? In short, Faraday-rotated, polarized sources will exhibit the spectral structure reserved solely for the EoR and noise (Section 1.5). Any $Q \rightarrow I$ leakage will damage prospects of making a clean detection of the 21cm EoR power spectrum. Hence, we begin our investigation of $Q \rightarrow I$ leakage.

The two most prominent ways in which polarized sky emission can leak into an interferometric estimate of Stokes I are leakages due to non-orthogonal and rotated feeds, and beam ellipticity — an asymmetry in the two linear polarizations of a primary beam which causes unpolarized signals to appear polarized, and vice versa. The first is a well-understood question, discussed at length in the series of papers by Hammaker, Bregman, and Sault (27, 28, 29, 30, 75). This type of leakage can be corrected by the proper linear combination of visibilities. Hence, we will omit discussion of this, and focus entirely on the latter issue. Unlike misaligned feeds, Beam leakage cannot be calibrated away. To begin, we will examine the contents of a visibility and relate them to the intrinsic Stokes parameters.

We begin by expanding the terms for the I and Q visibilities, from Equation 2.36, noting that they are not exact representations of the Stokes parameters they approximate:

$$\mathcal{V}_I = \mathcal{V}_{xx} + \mathcal{V}_{yy} = \int (A_{xx} + A_{yy})\, I e^{-2\pi i \vec{b}\cdot\hat{s}}\, \mathrm{d}\Omega + \int (A_{xx} - A_{yy})\, Q e^{-2\pi i \vec{b}\cdot\hat{s}}\, \mathrm{d}\Omega; \tag{4.7}$$

$$\mathcal{V}_Q = \mathcal{V}_{xx} - \mathcal{V}_{yy} = \int (A_{xx} + A_{yy})\, Q e^{-2\pi i \vec{b}\cdot\hat{s}}\, \mathrm{d}\Omega + \int (A_{xx} - A_{yy})\, I e^{-2\pi i \vec{b}\cdot\hat{s}}\, \mathrm{d}\Omega. \tag{4.8}$$

$\mathcal{V}_I$ is the Fourier transform of I, weighted by the sum of the xx and yy beams (call it A_+), and that of Q, weighted by the differenced beam (A_-). $\mathcal{V}_Q$ is symmetric to $\mathcal{V}_I$. Naïve addition of the Stokes visibilities clearly produces a mechanism for $Q \rightarrow I$ leakage.

If we allowed ourselves the ability to image visiblities, we could simply add a linear combintation of images to acheive "pure" Stokes parameters, but if we restrict ourselves to visibilities, this process, now a convolution, is impossible without a densely-sampled uv-plane. The requirement of densely-sampled uv-plane is often not satisfied, especially for the array used for the duration of this thesis certainly falls into this category. More discussion on the benefits and consequences of this effect can be found in Section 4.4.

Having established this particular mechanism for $I \rightarrow Q$ leakage, we now ask how much of Q's power is contained in the power of $\mathcal{V}_I$, which yeilds an estimate of the fraction of our power spectra are corrupted by beam leakage.

To find this estimate, let us first make a three assumptions:

1. The flux contained in I and Q are Gaussian, random fields.
2. The fields I and Q are uncorrelated.
3. The varaiance of Q is some fraction p^2 that of I, and $p \ll 1$.

Despite the obvious exceptions to these three assumptions, we proceed with our estimate of the leakage power, sacrificing accuracy for an analytic solution, and thus, intuition. The second assumption allows us to separate the contributions from I and Q when squaring $\mathcal{V}_I$, allowing us to write

$$|\mathcal{V}_I|^2 = \int |A_+|^2 \, \mathrm{d}\Omega \star P_i + \int |A_-|^2 \, \mathrm{d}\Omega \star P_Q, \tag{4.9}$$

where $P_I(P_Q)$ is the power spectrum of $I(Q)$, and $\star$ denotes a convolution. The first assumption sets P_I and P_Q to be constant, since the power spectrum of a Gaussian, random field is flat, and the third assumption allows us to write the simple expression for the power spectrum of $\mathcal{V}_I$,

$$|\mathcal{V}_I|^2 \propto \int |A_+|^2 \, \mathrm{d}\Omega + p^2 \int |A_-|^2 \, \mathrm{d}\Omega; \tag{4.10}$$

and similarly for Q,

$$|\mathcal{V}_Q|^2 \propto \int |A_-|^2 \, \mathrm{d}\Omega + p^2 \int |A_+|^2 \, \mathrm{d}\Omega. \tag{4.11}$$

We now ask what the ratio $|\mathcal{V}_Q|^2/|\mathcal{V}_Q|^2$ is — this allows us to answer the question, what fraction of what I measure in Q is present in what I measure for I. Defining $\mathcal{A}_\pm$ to be the power spectrum of $A_\pm$, we can write this ratio as

$$\frac{|\mathcal{V}_Q|^2}{|\mathcal{V}_I|^2} = \frac{\mathcal{A}_+ + p^2\mathcal{A}_-}{p^2\mathcal{A}_+ + \mathcal{A}_-} \approx \frac{\mathcal{A}_-}{\mathcal{A}_+}, \tag{4.12}$$

where the approximation is the first-order Taylor expansion of the ratio in p, which has been measured to be much smaller than one at higher frequencies (85), and been measured to decrease at lower frequencies (6, e.g.).

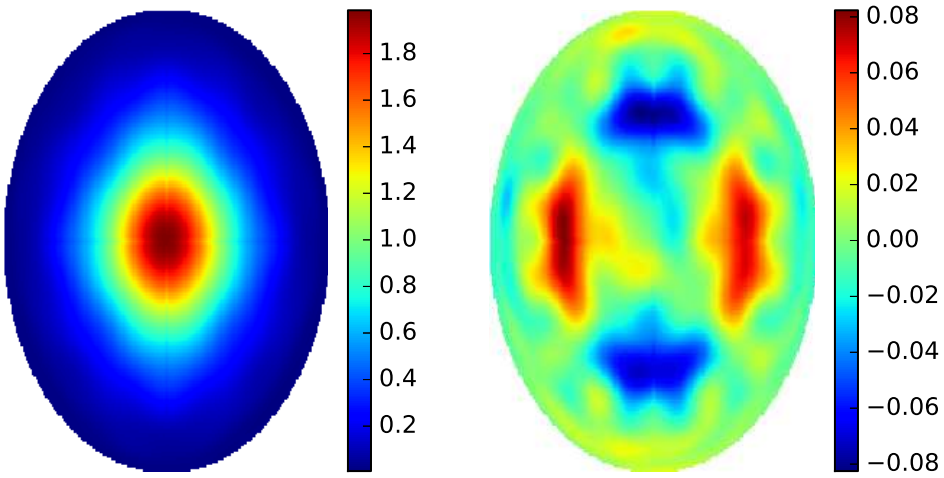

Figure 4.5: (Right Panel) Summed PAPER beam (70), $A_{xx} + A_{yy}$, at 164 MHz, normalized to peak at two. (Left Panel) Differenced PAPER beam, $A_{xx} - A_{yy}$, where each component A_p is normalized to peak at one.

We can apply these parameters to the measured PAPER beam (70), whose sum and difference are shown in Figure 4.5. We find the metric for leakage at 164 MHz to be

$$\frac{\mathcal{A}_-}{\mathcal{A}_+} = 2.1 \times 10^{-3}, \tag{4.13}$$

— this is roughly the fractional level of contamination we'd expect in the I power spectrum. In fact, this metric can be a function of frequency — the PAPER beam is most matched at 150 MHz, and we would expect the leakage to be least at that frequency, and increase as we approach the edges of the band. The measured values, shown in Figure 4.6, confirm our intuition, showing that the minimum leakage occurs at roughly the central frequency of 156 MHz, where the PAPER antenna's impedence matching is optimized.

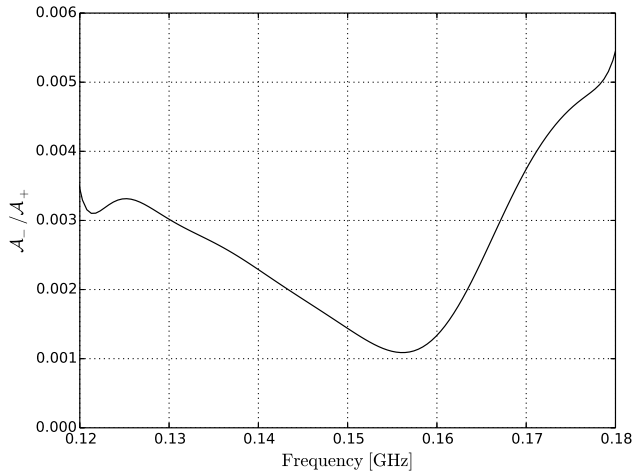

Figure 4.6: Fractional beam leakage, defined by Equation 4.12, as a function of frequency for the PAPER beam. This metric of leakage takes its minumum value at 156 MHz, near where the PAPER antenna is optimized for impedence matching.

4.4 Sparse uv Sampling and Wide-field Polarimetry

One advantage of the delay spectrum approach (Section 2.4) is that it relaxes the requirement of gridding in the uv-plane. Each baseline is assigned a position in the uv-plane ab initio, and visibilities from similar baselines may be coherently added without imaging. This allows for sparse sampling in the uv-plane without damaging effects from sidelobes or missing data, problems other methods may experience. Since the delay spectrum rotates a power-spectrum estimate into the native coordinate system of an interferometer, there are no inherently missing frequency-data. Parsons et al. (62) present the sensitivity benefits of a sparse, redundant array configuration, but other techniques aim to uniformly sample the (u, v, ν) cube, mitigating the systematic effects of computing a Fourier Transform across unevenly sampled data.

An obvious disadvantage of having sparesely-sampled data is poor imaging. Not only does sparse sampling provide a highly-irregular synthesized beam, but it also limits the available information for a full reconstruction of the image. Without adjacent uv-samples, a full, accurate deconvolution by a wide beam simply has inssufficient information. As seen in Section 4.3, the inability to correct for beam effects will provide a significant source of systematic error via polarized leakage.

By choosing to wield the full power of the delay spectrum approach and redundant sampling, an observer is forced to add visibilities with no beam weighting. The beam information supplied by adjacent uv-samples simply does not exist, and without transforming into the image plane, is unrecoverable. Hence, the imperative to investigate the implications of a lack of beam-weighting, the naïve construction of the I visibility, arises.

Together, redundant sampling and the delay spectrum approach give a 21cm EoR experiment incentive to add raw visibilities, subjecting it to potential leakage. An elliptical primary beam givees a mechanism whereby polarized emission can corrupt an estimate of the total power. To what degree does polarized emission corrupt an estimate of the 21cm EoR signal?

4.5 The Power Spectrum of a Single, Polarized Source

We begin our investigation of the effects of polarized foregrounds on the 21cm EoR signal by examining the power spectrum of a single source at zenith, whose signal has the structure of a single Faraday screen. In doing so, we can develop an intuition for the rotation measures that affect cosmologically interesting k modes of the power spectrum. By looking at what is effectively the impulse response of a Faraday screen on the power spectrum, it will be easier later to interpret a more complicated model.

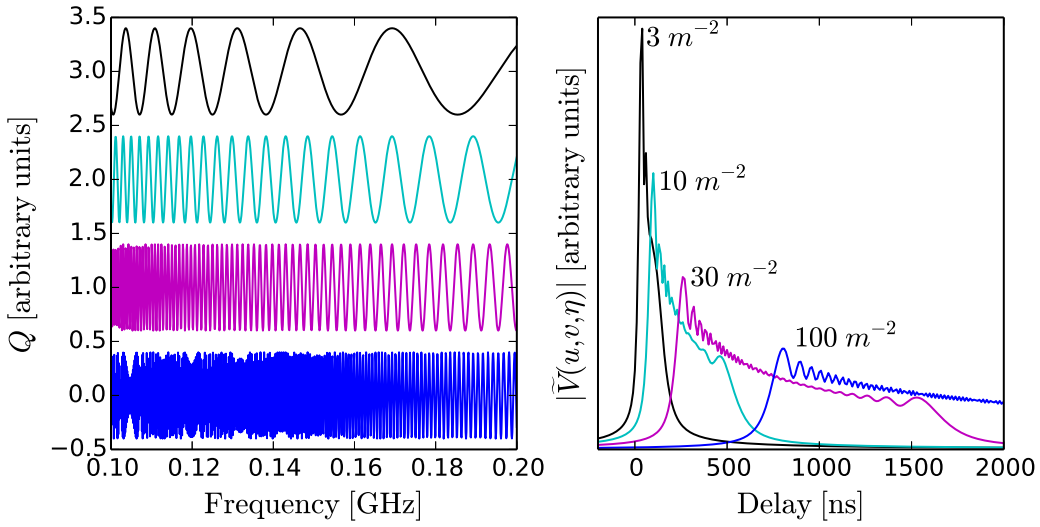

Figure 4.7: (Left Panel) Simulated visibilities with Faraday rotation, whose spectra can be written as $\exp\{-2i\Phi\lambda^2\}$. The four visibilities show four rotation measures: black, $\Phi = 3$ m^{-2}; cyan, $\Phi = 10$ m^2; magenta, $\Phi = 30$ m^{-2}; and blue, $\Phi = 100$ m^{-2}. (Right Panel) Amplitudes of the corresponding, delay-transformed visibilities.

The left-hand panel of Figure 4.7 shows the real component of visibilities containing a few linearly polarized sources behind Faraday screens, $S(\nu) = \exp\{-2i\Phi\lambda^2\}$, where Φ is the rotation measure of the screen. Each spectrum is normalized to contain one arbitrary

unit of flux, and is located at zenith (delay of zero). Note that at the highest Φ shown, the spectrum is not critically sampled at the lowest frequencies. This is due to the uneven sampling of λ^2 across the band: as $\Delta\lambda^2 \approx d\lambda^2/d\nu\ \Delta\nu \propto \Delta\nu/\nu^3$ increases, the sensitivity to large rotation measures decreases. A more thorough discussion of this effect can be found in Chapter 7.

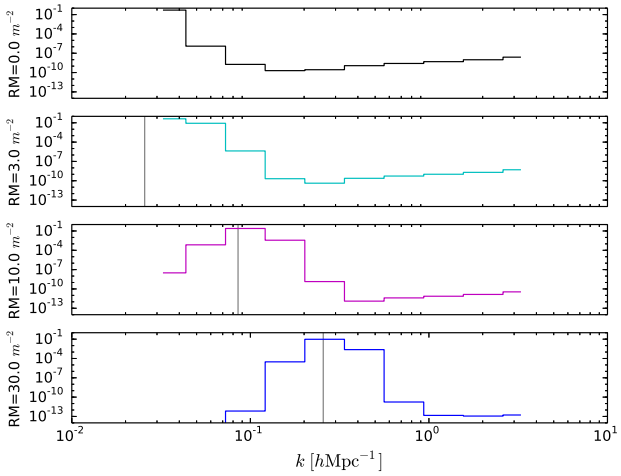

Figure 4.8: $k^2P(k)/2\pi^2$ for the four visibilities in Figure 4.7. The k with maximum contamination for each rotation measure (Equation 4.17) is shown with a gray, horizontal line.

The right-hand panel of Figure 4.7 shows the Fourier Transform over frequency of the spectra in the left-hand panel. While this does not exactly represent the delay spectrum of a visibility — there is no beam-weighting, and no $\exp\{-2\pi i\vec{b}\cdot\hat{s}\}$ component, which essentially defines the delay spectrum — we interpret it as the delay structure introduced by a polarized source behind a Faraday screen. The results of these transforms over a subband representing a cosmological measurement are shown in Figure 4.8. The most important feature of this plot is this: there is a single k-mode associated with each rotation measure at each redshift.

We can construct an analytic estimate of this in the following manner.[19]

First, we approximate the cosmological k-mode sampled as $\tau \approx k_{||}\ dr_{||}/d\nu$. Next, we recall the cosmological scaling from frequency into $h\mathrm{Mpc}^{-1}$,

$$\frac{dr_{||}}{d\nu} = \frac{dr_{||}}{dz}\frac{dz}{d\nu} = -\frac{c(1+z)}{H(z)\nu}, \tag{4.14}$$

where $H(z)$ is the Hubble parameter at redshift z.

Finally, we find the $k_{||}$,Φ pair which maximizes the product of a delay mode and a rotation measure mode,

$$\int e^{-2\pi i(\nu\tau - \Phi\lambda^2/\pi)}\ \mathrm{d}\nu \equiv \int e^{-i(\varphi_k - \varphi_\Phi)}\ \mathrm{d}\nu. \tag{4.15}$$

This occurs when $0 = \varphi_k - \varphi_\Phi$. Differentiating with respect to λ^2, since by one convention, this defines the rotation measure, and applying the chain rule several times, we arrive at the conclusion

$$\begin{aligned} 0 &= \frac{\partial\varphi_k}{\partial\lambda^2} - \frac{\partial\varphi_\Phi}{\partial\lambda^2} = \frac{\partial\varphi_k}{\partial r_{||}}\frac{\partial r_{||}}{\partial z}\frac{\partial z}{\partial\nu}\frac{\partial\nu}{\partial\lambda^2} - 2\Phi \\ &= k_{||}\left(-\frac{c(1+z)}{H(z)\nu}\right)\left(-\frac{1}{2}\frac{c}{\lambda^3}\right) - 2\Phi, \end{aligned} \tag{4.16}$$

which yeilds the value of $k_{||}$ that Φ modes most approximate:

$$k_{||} = \frac{4}{c}\frac{H(z)}{(1+z)}\Phi\lambda^2. \tag{4.17}$$

This factor differs from Moore et al. (55), which was derived by simply setting $2\pi\tau\nu = 2\Phi\lambda^2$ by a factor of two, and by the expression derived in Appendix C of Pen et al. (67) by a factor of c^{-2}, which has a simple error in the derivation of $dr_{||}/d\nu$ which propagated through their calculations.

[19]There is an error in the published version of this paper (55) which omits a factor of two. We present the correct calculation here. Another peer-reviewed, published paper (67) also incorrectly calculates this scaling. Thanks to Gianni Bernardi for pointing this out and working through this with me.

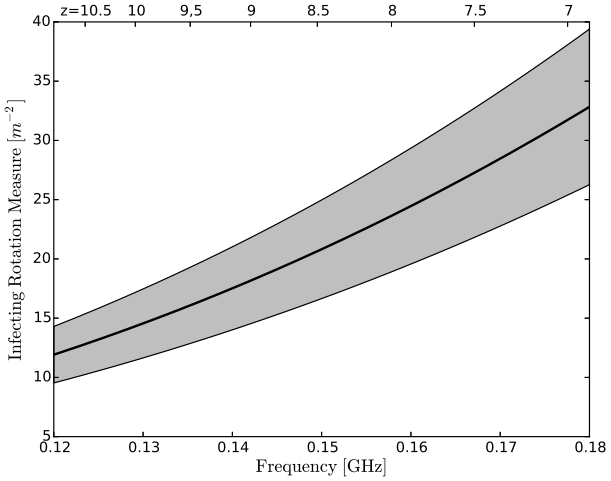

Figure 4.9: The rotation measure which infects $k \approx 0.25\ h\mathrm{Mpc}^{-1}$ is shown in black, with the rotation measures maximally infecting $0.2\ h\mathrm{Mpc}^{-1} \leq k \leq 0.3\ h\mathrm{Mpc}^{-1}$ are shown in grey. This is plotted as a function of frequency and redshift to highlight the fact that lower rotation measures affect the same k at lower frequencies or redshifts.

Figure 4.9 shows the most-infecting rotation measure for wave-numbers 0.2 $h\mathrm{Mpc}^{-1} \leq k_{||} \leq 0.3\ h\mathrm{Mpc}^{-1}$, which abut the horizon for many of PAPER's short baselines. These are the k-modes which both minimize thermal noise and avoid foreground signals, so for that reason, they are critical to the initial detection of a 21cm EoR power spectrum.

Chapter 5

Simulations

To better grasp the effects of $Q \rightarrow I$ leakage into the 21cm signal, we generate several random realizations of the sky, each consisting of many polarized point sources. Each source passes through a Faraday screen with some rotation measure, chosen from a distribution based on current measurements. Next, we simulate that source for a single baseline. Finally, we calculate the power spectrum measured by that visibility. Only one visibility needs to be simulated, because the delay-spectrum approach makes use of the fact that each baseline measures the 21cm EoR with a range of k-modes determened by the baseline length, orientation, and bandwidth.

Rather than creating an exact simulation of the physical sky, we create a simulation whose statistical properties are physically motivated. This choice reflects a desire for simple, easliy tunable parameters for the simulation. In that same spirit, we model all sources simply as point sources with a Poisson angular distribution. The simulation's primary concern with the spectral information of polarized foregrounds allows us to justify neglecting the angular terms. This is equivalent to assuming for all the relevant k-modes, $k_{||} \gg k_{\perp}$. Emphasis on the $k_{||}$, or line-of-sight, spectral modes also motivates our decision to model the sky as nuemerous point sources. For a more detailed discussion of diffuse polarized emission, we

Survey Name	Frequency	Survey Area	Full Stokes?	Citation
VLSS	74 MHz	$\delta > -30°$	No	Cohen et al. (12)
6C	151 MHz	0.82 sr	No	Hales et al. (26)
NVSS	1.4 GHz	$\delta > -30°$	Yes	Condon et al. (14)

Table 5.1: Summary of three, low-frequency surveys.

direct the reader to Jelić et al. (41).

5.1 Parameterizing the Polarized Sky

Source Positions

Source positions are distributed uniformly over the sphere. A single source's altitude θ is drawn from a distribution in which $\cos\theta$ is uniform on $[0, 1]$. A source's azimuthal angle ϕ is drawn independently from $\cos\theta$, uniform on $[0, 2\pi)$. This choice of source position distributions conserves number of sources per unit area across the sky, and is equivalent to drawing both direction cosines, $(l, m) = (\sin\theta\cos\phi, \sin\theta\sin\phi)$, from a uniform distribution on $[-1, 1]$.

Flux Counts

In order to achieve realistic source fluxes and source counts, we base the distributions from which we draw various parameters on previous radio surveys. For the source fluxes, we aim to agree with the VLA Large Sky Survey (12, henceforth, called VLSS) and the Sixth Cambridge Survey (26, henceforth, called 6C) surveys, summarized alongside a full polarization survey in Table 5.1. We can take 6C source counts at face value, since it was measured in the PAPER band at 151 MHz. However, we must extrapolate the VLSS source counts from

the observed 74 MHz into the paper band. We perform this extrapolation, following Cohen et al. (13), by applying a spectral index of -0.79 to the amplitude of the source counts.

Above some limiting flux S_{min}, the differential number counts in (dN/dS) found in the 6C survey may be characterized by a double power law, turning over at some knee flux S_0:

$$\frac{dN}{dS} = \begin{cases} 4000\ S_0^{-0.76} S^{-1.75}\ \mathrm{Jy}^{-1}\mathrm{sr}^{-1} & S_{min} \leq S < S_0 \\ 4000\ S^{-2.81}\ \mathrm{Jy}^{-1}\mathrm{sr}^{-1} & S_0 \leq S \end{cases}. \tag{5.1}$$

Following the 6C results, we choose the turning point S_0 to be 0.88 Jy. The number of sources simulated (20,531) is chosen by the size of the PAPER beam at 151 MHz (0.76 sr) and a flux range over which to integrate. In general, this operation can be expressed by the integral

$$N = \Omega \int_{S_{min}}^{S_{max}} \frac{dN}{dS}\ \mathrm{d}S, \tag{5.2}$$

where S_{min} and S_{max} are the limiting fluxes of the survey, and Ω is the survey's field of view. We choose to include sources in between 100 mJy and 10 Jy. This choice provides a reasonable dynamic range of sources. Below the lower limit, the 6C sources are unreliable due to signal-to-noise issues and confusion[20], and above 10 Jy, we expect that sources may be easily identified and removed.

By setting our lower cutoff too high, are we omitting much of the power contained in our measurement? If we were to blindly extrapolate the 6C source counts below the lower limit of the catalog, we would add a negligible amount of power. Integrating $S^2\ dN/dS$ down to some minimum flux estimates the contribution of the sources above that flux to the total variance of the flux. Performing this operation to the 6C source counts, we find that we are including $\sim 70\%$ of the total variance. Extending the minimum flux would indeed add more power to the simulation, but it would not drastically alter these results.

[20] Below a certain flux, I expect to find more than one source per resolution element — this uncertainty in the number of sources per pointing adds to the variance of my measurement.

The VLSS source counts follow a single power law, given by

$$\frac{dN}{dS} = 4865\ S^{-2.3}\ \mathrm{Jy}^{-1}\mathrm{sr}^{-1}, \tag{5.3}$$

where the spectral index of -0.79 has been applied. For these, we choose minimum and maximum fluxes of 0.8 Jy and 100 Jy, respectively, rejecting sources well below the lower limit of the catalogue, and providing a reasonable dynamic range for the included sources. Integrating over the PAPER beam provides 11,262 sources.

Qualitatively, the source counts for these two surveys differ in two ways. The 6C survey yields more, dimmer sources, where the VLSS survey yields fewer, brighter sources. By examining the difference in polarized power from these two source counts, we may answer the question "Is $Q \rightarrow I$ leakage due mostly to a few, bright sources, or is it due to a forest of unresolved, dim sources?"

It is worth noting the robustness of these two source counts with respect to independent measurements — both agree with the results of a recent survey from the Murchison Widefield Array (88), an instrument similar in many regards to PAPER.

Spectral Indices

All sources are assigned a spectral index, which is drawn from a normal distribution with mean -0.8, and standard deviation 0.1. This roughly agrees with the findings of Helmboldt et al. (34).

Polarized Fractions

Instead of drawing polarized sources from a measured polarized flux distribution, we simply down-weight the total intensity by some polarized fraction (Π), chosen to reflect the studies of Tucci and Toffolatti (85). We sample the Π from a log-normal distribution whose mean is 2.01% and whose standard deviation is log(3.845%). Because the log-normal distribution

has no upper bound, and it is unreasonable to find sources with a high polarized fraction[21], we truncate the distribution at 30%. As we will investigate in Section 6.3, this upper limit is considerably higher than what has been measured at 150 MHz. Following the aforementioned study, we do not impose any correlation between source flux and polarization fraction.

It has been noted that, among other effects, bandwidth depolarization causes the polarized fraction to decrease at lower frequencies (47). This, alongside the GMRT measurements (67), indicates that these distributions, taken at 1.4 GHz, may overestimate the distribution at 150 MHz. We neglect this instinct that we are overestimating the polarized flux, taking the 1.4 GHz measurements at face value, since the mean polarization fraction can be thought of as a scale factor to the overall power spectrum.

Polarization Angles

The polarization angle of each source is chosen to be uniformly sampled on $[0, \pi)$, which assumes no correlation in the polarization angles of individual extragalactic sources. Section 5.3 will investigate the validity of this claim.

Rotation Measures

We draw our distribution of rotation measures on the map presented in Oppermann et al. (60). To mimic the effects of depolarization due to a finite spacial resolution (47), we apply a low-pass filter to the rotation measure map. Projecting the map into a spherical harmonic basis, we keep only those modes below the resolution of our simulated instrument. In the case of this simulation, we choose to keep only $\ell \leq 100 = 2\pi|\vec{b}/\lambda|$. This averages the polarization vectors in much the same way as a synthesized beam, and its effect is to essentially remove outliers in the rotation measure distribution, to which instruments like PAPER may not be sensitive. We then randomly draw rotation measures from the empirical cumulative distribution function of rotation measures, computed from the filtered,

[21] In fact, it is impossible to measure $\Pi > 1$!

Oppermann et al. (60) maps. Aside from low-pass filtering, no spatial information from the data is used. Section 5.3 briefly discusses the negligible consequences of spatially correlating rotation measures.

Histograms of the empirical distributions of rotation measure, polarized fraction, and source counts may be found in Figure 5.1. Over-plotted on all are the distributions from which they are drawn. Figure 5.2 shows the empirical distribution of polarized flux, using the NVSS and 6C surveys — these distributions should be convolutions of the power law source counts and the log-normal polarized fraction. These distributions qualitatively agree with the total power source counts: the 6C survey produces dimmer sources, while NVSS produces fewer.

We calculate visibilities for a single 30m, east-west baseline, corresponding to the most common spacing in the maximum-redundancy PAPER array (62, 65). The choice of baseline orientation is arbitrary, and since we are only modelling point sources, the choice of baseline length will only set the horizon limit of the power spectrum. Since the delay affected by a rotation measure is independent of a choice of baseline (Equation 4.17), choosing a relatively short baseline length will isolate smooth-spectrum foregrounds at lower τ and highlight Faraday leakage.

The full measurement equation for the visibility with linear polarization p ($\mathcal{V}_p$) used in this simulation is

$$\mathcal{V}_p = \sum_{j=1}^{N_{src}} A_p(l_j, m_j, \nu) S_j^{150} \left(\frac{150\text{ MHz}}{\nu}\right)^{\alpha_j} e^{-2\pi i \nu (u l_j + v m_j)} \left(1 \pm \Pi_j e^{-2i(\Phi_j \lambda^2 + \chi_j)}\right), \quad (5.4)$$

where each source j is assigned a flux (S_j), a polarized fraction (Π_j), a spectral index (α_j), a position (l_j, m_j), rotation measure (Φ_j), polarization angle (χ_j), and is weighted by the model primary beam in that linear polarization (A_p). To include both I and Q emission, xx visibilities receive the $+$, while yy visibilities recieve the $-$. They are then summed and differenced to yield $\mathcal{V}_I$ and $\mathcal{V}_Q$. A sample Q visibility is shown in Figure 5.3.

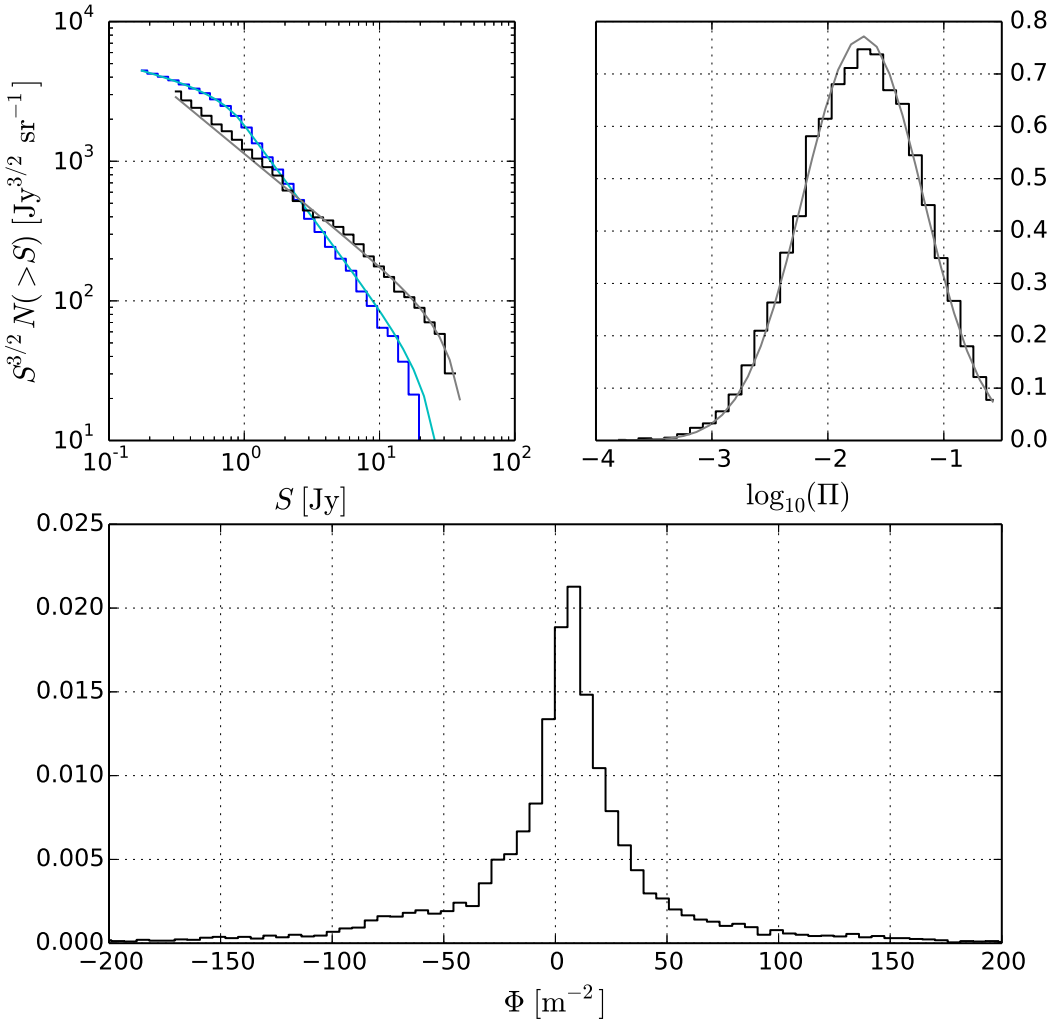

Figure 5.1: Distributions of simulated parameters. (Top Left) Euclidean normalized source counts for the sources produced in simulations A (blue) and B (black). Over plotted in cyan and gray, respectively, are the analytical distributions from which they are drawn. (Top Right) Distribution of polarized fractions used, with the log-normal distribution over-plotted in gray. (Bottom) Empirical distribution of rotation measures, generated from a spatially low-pass-filtered map (60).

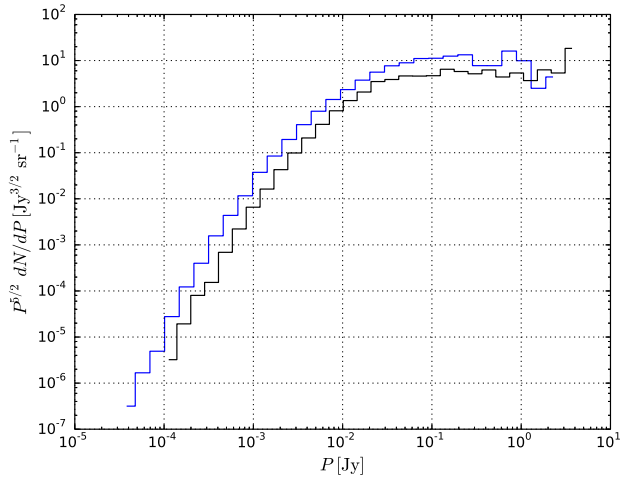

Figure 5.2: Euclidean-normalized, differential source counts for polarized flux in simulations A (blue) and B (black).

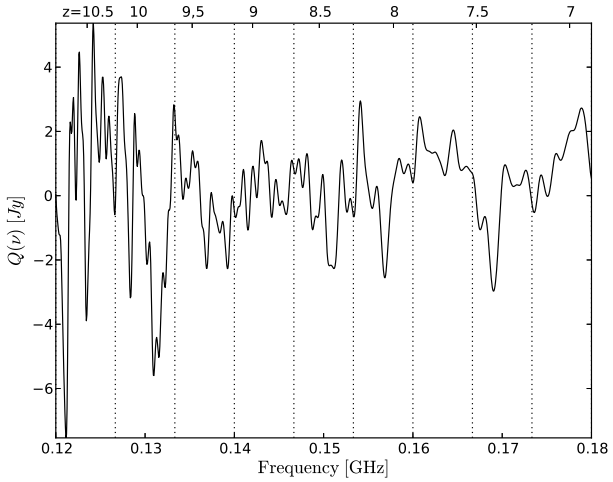

Figure 5.3: The real part of a sample Q visibility, given by Equation 5.4, and generated using the parameters shown in Figure 5.1.

Label	Source Counts	N_{src}	Rotation Measure Distribution
A	6C	20,531	Oppermann
B	NVSS	11,262	Oppermann
C	6C	20,531	2×Oppermann

Table 5.2: Three simulation treatments used in this section.

We choose not to include the parallactic rotation of Q and U (see Section 2.2), implying that the Q we label for this simulation is fixed to topocentric, azimuth and altitude coordinates. This choice clarifies equations and allows for an ease of understanding which would be obfuscated by writing equatorially defined Q and U.

Finally, we take advantage of our built-in tunable parameters and present three treatments of the simulation, summarized in Table 5.2. Simulation A serves as a baseline measurement, with reliable $6C$ source counts and the conservative, low-pass filtered Oppermann rotation measure distribution. Simulation B uses NVSS source counts, asking is $Q \rightarrow I$ leakage is dominated by a few, bright sources, rather than the forest of dim sources in 6C. Simulation C uses the 6C source counts, but doubles the Oppermann rotation measures, asking how large rotation measures affect $Q \rightarrow I$ leakage.

5.2 Results

Figure 5.4 the power spectra of several renderings of simulations A, B, and C. We interpret the power spectrum of $\mathcal{V}_I$ outside the horizon as the amount of polarized leakage corrupting the EoR signal ($Q \rightarrow I$ leakage), and the power spectrum of $\mathcal{V}_Q$ is our best representation of the polarized signal. These plots show the median power in each k bin for 100 realizations of the simulation, with error bars showing the 1-σ extend of the bandpowers for these realizations. These power spectra confirm the prediction made in Section 4.5 that λ^2 phase wrapping extends the foreground cutoff (62, 69, e.g.) to higher delay bins, corrupting some

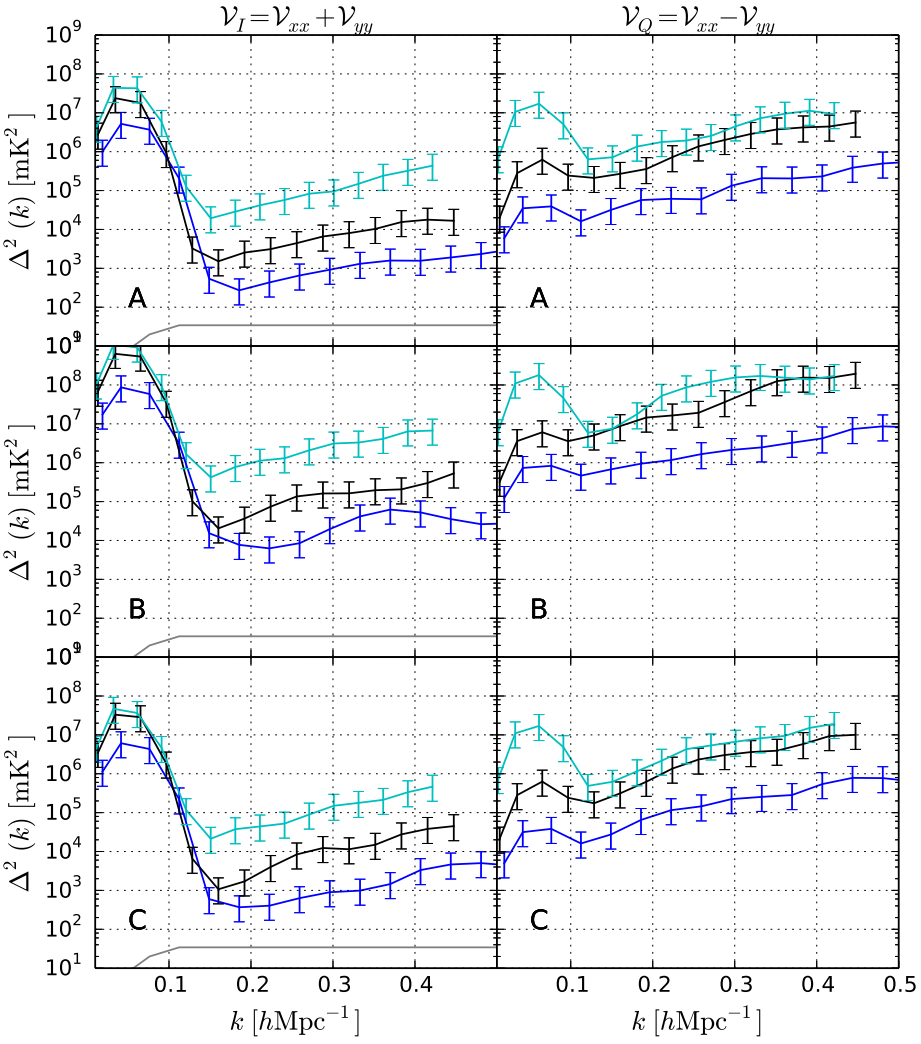

Figure 5.4: Power spectra for the three treatments of the simulation discussed in Section 5.1. From top to bottom, the rows are treatments A, B, and C. The left column shows the I power spectrum, highlighting $Q \rightarrow I$ leakage. The right column shows the Q power spectrum. Three redshift bins are shown in all plots: $z = 11.13$ (cyan), $z = 9.77$ (black), and $z = 7.05$ (blue). Error bars show 95% confidence intervals of several iterations of the simulation. For a point of reference in the left column, a fiducial EoR model (48) is plotted in grey.

of the most sensitive regions of k space for 21cm EoR analysis. They also demonstrate the prediction that high-redshift bins will be most affected.

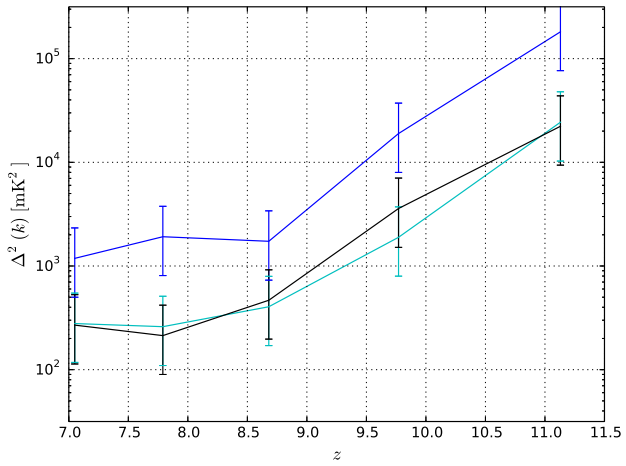

Figure 5.5: A (cyan), B (blue), C (black)

The severity of the leakage can be inferred from the power in the most EoR-sensitive k-bins which lie outside the horizon for small baselines ($0.2\ h\text{Mpc}^{-1} \leq k \leq 0.3\ h\text{Mpc}^{-1}$). Figure 5.5 shows $\Delta^2(k)$ in these bins as a function of redshift. The leaked power ranges in the hundreds of mK^2 to thousands, increasing from high frequency/low redshift to low frequency/high redshift. These simulations are about an order of magnitude above the level of the expected 21cm signal (48). If we may take this simulation as an accurate prediction of the low-frequency sky's polarized emission, these results imply that naïvely adding $\mathcal{V}_{xx}$ and $\mathcal{V}_{yy}$, formed with an approximately 10% asymmetric primary beam, incorporates enough bias from polarized leakage to completely obscure the 21cm signal. The levels of leakage in our simulations demand a strategy to model and remove point sources.

We note that simply forming $\mathcal{V}_I$ will also remove a negligible component of the EoR signal via the same mechanism. In a sense, the $Q \rightarrow I$ leakage can be thought of as a

rotation of power between the two Stokes parameters. Hence, for precision measurements of the EoR signal, this simple estimate may not be ideal. However, the effects of the $I \rightarrow Q$ leakage is small (compare the levels of $\mathcal{V}_Q$ and high k-modes of $\mathcal{V}_I$) and should not provide a significant hiderance to detection.

Were a power spectrum computed from only one linearly polarized visibiltity (xx, for instance), all polarized power would corrupt the measurement. We have chosen to suppress the polarized leakage by adding the linearly-polarized visibilities xx and yy. The leakage is dependent on the difference of the two beams (see discussion in Section 4.3), and by having beams that are at most 10% different suppresses the signal by around two to three orders of magnitude. Correcting for the beam-weighting in the image domain can further suppress the leakage, but errors in the beam model will introduce leakage in much the same manner. Hence, the constraint of having to suppress polarized leakage by four orders of magnitude causes the need for an accurate primary beam model to around the 1% level in the case of imaging, or symmetric at the 1% level if visibilities are used directly.

We conlude this discussion by noting the large variance in simulted power. The results shown are the mean bandpowers in $\Delta^2(k)$ for several realizations of the simulation. Taking so many realizations into account essentially maps out the posterior distribution of the $\Delta^2(k)$ bandpowers. The 2σ width covers nearly two orders of magnitude, which indicates that the level of $Q \rightarrow I$ leakage is highly sensitive to the exact parameters drawn in any realization. The actual level of leakage measured will thus be highly dependent on a choice of field, and on the actual distributions of polarized fluxes and rotation measures.

5.3 Consistency Tests

Now that we have presented the power spectra, we must check two things: first, that the two-dimensional C_ℓ power spectrum generated by the simulations agrees with current measurements (6), and second, that the assumption that uncorrelated polarization angles is valid.

5.3.1 Two-Dimensional Power Spectrum and Diffuse Emission

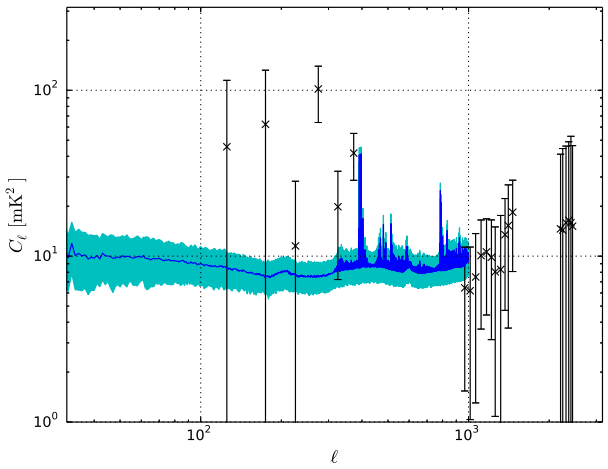

Figure 5.6: Black points show the C_ℓ power spectrum from Bernardi et al. (6). The blue line shows the mean C_ℓ of several simulations. The shaded, cyan region shows 2σ limits of the distributions for each bin in ℓ. This shows the consistency between our simulations and recent measurements. Treatment A is used for these simulations.

Figure 5.6 shows the distribution of two-dimensional power spectrum over several simulations, plotted alongside the C_ℓ measurements from Bernardi et al. (6). We see qualitatively that our simulation well obeys the upper limits imposed by the Bernardi measurement. This agreement helps validate our results.

The estimates of power in Section 5.2 are dependent on the relative strengths of diffuse, polarized emission and polarized point sources. We have taken care to agree with measurements of all polarized emission, but those measurements are uncertain above $\ell \sim 300$. We interpret them as an upper limit. In the limiting case where diffuse emission is the only component to the polarized sky, this leakage could be suppressed by measuring with a longer baseline, which in turn measures a lower ℓ or $k_\perp$. We have chosen a 30m baseline, which

corresponds to $\ell \approx 200$. This choice of baseline length is relatively short for interferometers at these wavelengths, but falls at the high end of the Bernardi et al. (6) measurements.

Including additional diffuse emission in the simulation would certainly increase the total power in the simulation for low ℓ, but the frequency structure would remain qualitatively the same as point sources. As we will show in the following section, the correlation of rotation measures and polarization angles that could be introduced by an extended structure will not significantly affect the power spectrum. For this reason, we can consider the polarized sky as having two components with nearly identical footprints in the line-of-sight direction: diffuse and point-like. Both components will exhibit similar frequency structure, so choice of baseline length will set the relative weightings of these components. Bernardi et al. (6) briefly discuss some of the implications of their measurement of extended structure to the three-dimensional power spectrum in their conclusion, which agrees with our analysis of point-like structure. We will discuss the qualitative differences between diffuse and point-like emission in Section 6.3.3.

5.3.2 Correlating Polarization Vectors

The analysis of Section 5.2 neglects known spatial correlations of the rotation measure distribution (45). Furthermore, the random drawing of polarization angles could have a cancelling effect on the visibilities. This neglect could potentially suppress our estimation of polarized leakage into the power spectrum.

To investigate these possible effects, we choose rotation measures from the Oppermann map (60), with a pointing center at the Galactic south pole — a reasonable field for EoR analysis. We then set all polarization angles to zero, maximally correlating polarization vectors, while still including information of the polarized sky. All other simulation parameters are identical to treatment A of Table 5.2.

Figure 5.7 compares the results of this treatment with simulation A from Table 5.2. The power spectrum of this treatment agreees with simulation A at all redshifts and values of k,

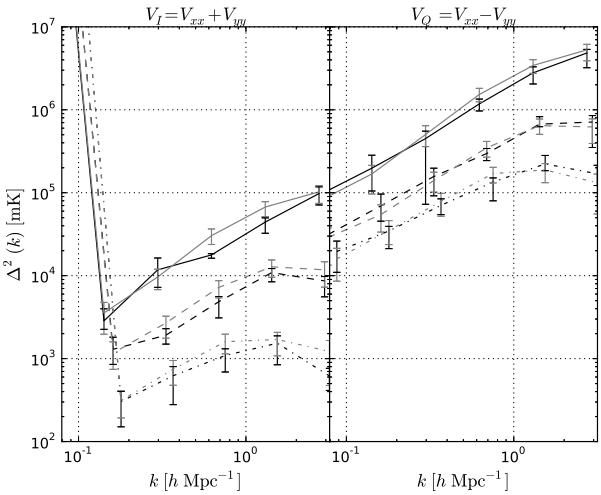

Figure 5.7: A comparison of power spectrum measurements for a treatment of the simulation with correlated polarization angles (black), and those from treatment A (gray). As in Figure 5.4, the left panel shows the I power spectrum, and the right panel shows the Q power spectrum. Three redshift bins are shown, each denoted with a different line style: 9.73 (solid), 8.33 (dashed), and 7.25 (dot-dashed). The results of simulation A with this simulation show that correlating polarization vectors does not affect the power spectrum.

for both polarizations. This agreement indicates that the spatial correlation of polarization vectors do not significantly affect the power spectrum. Thus, the assumption in Section 5.1 are spatially uncorrelated does not affect the results of these simulations.

5.4 Mitigating Leakage

Section 5.2 predicts an excess polarized signal due only to point sources of around 10^4 mK^2 at $k \sim 0.15$ hMpc^{-1} for most treatments of the simulation. While the exact levels of these predictions may be subject to some error, the need certainly arises for some removal scheme. This removal must suppress power from polarized foregrounds by around four orders of magnitude in the power spectrum.

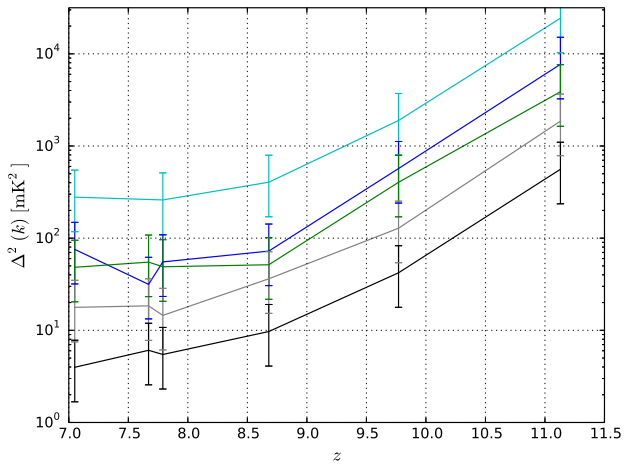

Figure 5.8

To investigate the effects of modelling and removing polarized sources, we rerun the simulation, excluding the brightest polarized sources. Figure 5.8 shows the median value of several simulations of the k-bin nearest 0.25 hMpc^{-1}, having removed the brightest 1000,

2000, 5000, and 10,000 sources. These limits in numbers of sources correspond to unpolarized flux-limits of 1300, 900, 460, and 240 mJy, respectively. Polarized flux limits are roughly 2% of these. We remove these sources from treatment A of the simulation, which includes around 21,000 sources. Despite having removed nearly one-third of the sources, the leaked power still exceeds 10 mK^2, the expected level of the 21cm EoR power spectrum.

To remove enough flux to consistently fall below the expected EoR signal, we need to remove a large majority of polarized point sources. We recomputed the simulations with a lower minimum flux (60 mJy), expecting a similar result, but found that we increased the power in this k-bin by only one or two mK^2. For total power to fall below 10 mK^2, more sources required removal. We exclude further investigation of this analysis for three reasons. First, current measurements do not constrain dN/dS to the levels necessary to accurately model such low-flux sources. Second, including lower-flux sources does not significantly affect the result that the result that the expected polarized power spectrum will be of the order of 10^4-10^6 mK^2. Third, the variance in power from one simulation to the next was large enough that the two treatments of the simulation — even with 10,000 sources removed — could not be considered significantly different.

The onerous levels of source-removal suggest that a different mitigation scheme be considered. Future instruments may take polarization into consideration in their design. Leakage can be mitigated with more circular beams, and circular feeds avoid the $Q \rightarrow I$ leakage entirely. Even with existing data, rotation measure synthesis (11) could potentially provide the ability to separate sources with distinct rotation measure structure to be separated from the EoR signal.

Chapter 6

Power Spectra

To directly measure the level of polarized power described and simulated in the previous two sections, we turn to data taken during the EoR2011 observing campaign defined in Table 3.1. While the primary objective of this campaign was to measure the mostly unpolarized signature of neutral Hydrogen in the early universe (39, 65), it could not reach the sensitivity levels required to characterize or even detect the power spectrum. Hence, it is a useful study for characterizing foregrounds. All four polarization products were correlated specifically to characterize the level of polarized power which could corrupt 21cm EoR power spectrum measurements.

As as brief reminder of Section 3.2 and Table 3.1, this data was taken during Winter of 2011 and Spring of 2012, spanning eighty-two nights of observations. PAPER's configuration for this season was in an 8×4 grid. Since the antennae are arranged in a redundant grid, we can label the subsets of redundant baselines by their grid spacings. For instance, a baseline composed of two adjacent antennae in the same row can be written $(0, 1)$. Similarly, a baseline composed of two adjacent antennae in the same column can be written as $(1, 0)$. For the results presented here, and in the two sister papers to this work, (39, 65), only baseline types $(0, 1)$, $(1, 1)$, and $(-1, 1)$ are considered. The row-spacings in this configuration were

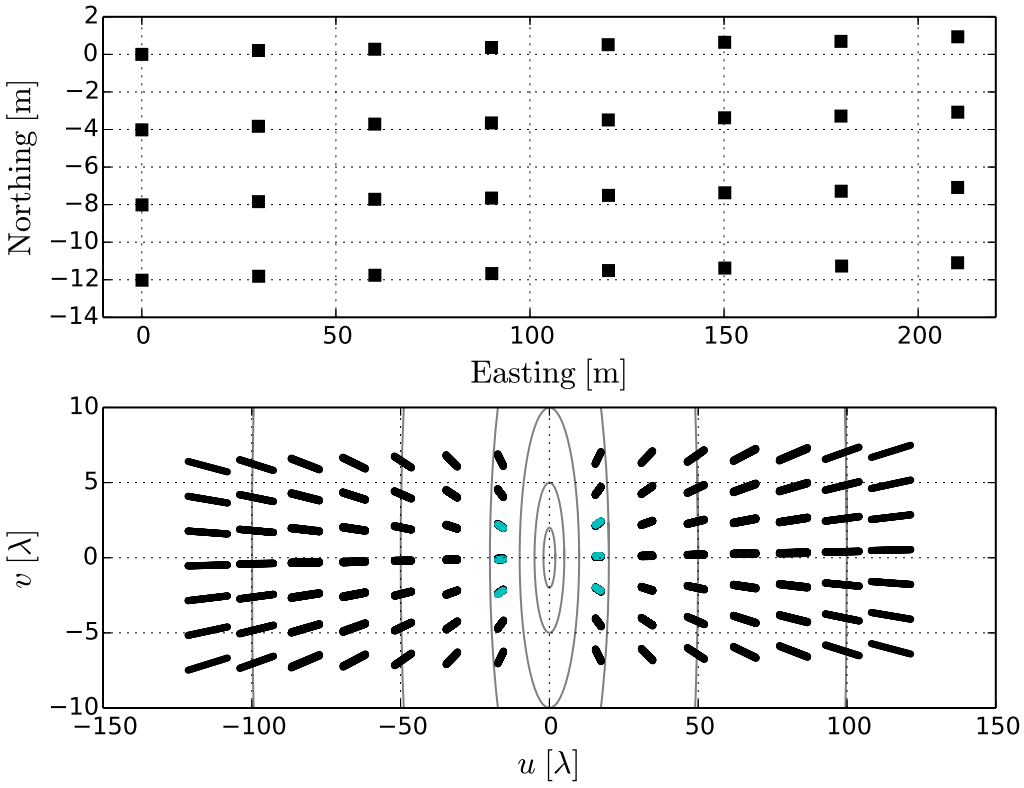

Figure 6.1: (Top Panel) Antenna positions, referenced to the top, left antenna. (Bottom Panel) uv-coverage for the entire array in black. The uv-coverage of the subset of antennae used for this analysis is shown in cyan. To give a sense of scale between the u- and v-axes, concentric circles with radii of 2λ, 5λ, 10λ, 20λ, 50λ, and 100λ are shown. Since the power spectrum is computed for each integration, there is no Earth-rotation synthesis.

30m, chosen to reduce the extent of the foreground wedge without incurring the any antenna-to-antenna cross talk. The column spacings were 4m, maximizing the redundancy between baselines $(0, 1)$ and $(1, 1)$ or $(-1, 1)$. Figure 6.1 shows a map of the antenna spacing, as well as the uv-coverage of the array. The four-degree offset of the columns from true north is due to the projection of the tangent plane to the Earth at the array location to UTM plane 34.

Data was taken continuously from 6pm SAST until 6am SAST each night during this campaign. To remove effects of the sun, we only consider data when the sun is below -5° in altitude. We restrict the number of nights due to some systematic errors which corrupt the data after April 1, 2012 — Julian Date 2456018. A catastrophic event occurred on this date,[22] and most of the data taken after it was unusable. 82 nights of data survived quality checks and are used for this analysis.

We focus our efforts on the range in LST from 1h00m until 8h00m, which maximizes the total integration time available, but minimizes the effects of systematics. Since PAPER is a drift-scan array, this sets both the pointing and the field of view. Figure 6.2 shows a map of effective integration time per pointing, defined as

$$t_{eff}(\alpha, \delta) = \sum_i t_{int} A(\alpha, \delta, t_i) \tag{6.1}$$

where the sum extends over the each integration in the season. This metric is defined to give the total integration time when integrated over position on the sphere. The total field of view surveyed is 2.39 sr.

Finally, we restrict the final analysis to two bands, though the full range of frequencies is used throughout much of the analysis. We label the lower one Band I, and the upper, Band II. Band II corresponds with that used for the results in Parsons et al. (65), and Band I is chosen to correspond to the lowest band in Jacobs et al. (39). Many of the observational parameters defining these two bands are presented in Table 6.1.

[22] Because I observed this data remotely, I can only guess, but my money is on a lightning strike.

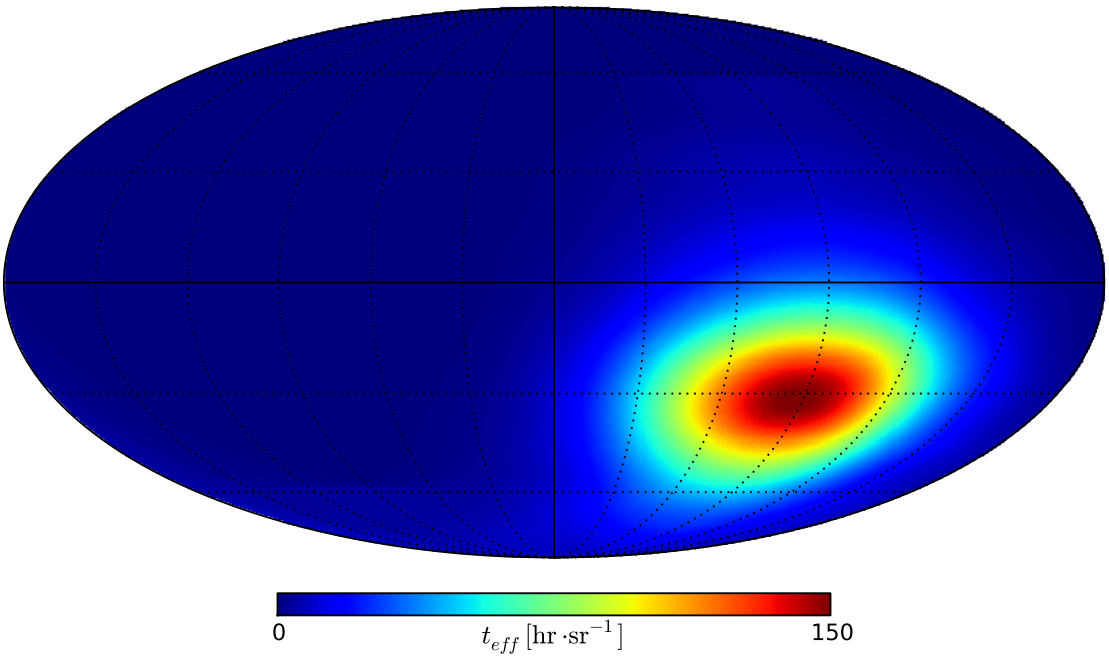

Figure 6.2: Effective integration time per pointing (Equation 6.1), as a function of position on the sphere.

Band	ν_0 [MHz]	$\Delta\nu$ [MHz]	z	A_{eff} [m^2]	T_{sys} [K]	$\mathcal{A}_-/\mathcal{A}_+$
I	126	7.9	10.3	4.47	836	3.3×10^{-3}
II	164	9.4	7.66	5.80	505	2.2×10^{-2}

Table 6.1: Observational parameters for the two sets of power spectra presented. Given are the central frequency ν_0, the effective bandwidth $\Delta\nu$, the central redshift of observation z, the effective area of the antennae A_{eff}, and the ratio which paramterizes $Q \to I$ leakage, $\mathcal{A}_-/\mathcal{A}_+$ (Section 4.3).

This Chapter outlines the processing and analysis of the polarization properties of this data. Section 6.1 describes the analysis and quality checks this data underwent; Section 6.2 presents the power spectra of these data, and finally, Section 6.3 gives the physical properties of polarized point sources that these data imply.

6.1 Data Processing

6.1.1 RFI Excision

We begin with an excision of RFI from the raw data, a three step process. First, we flag known frequency channels containing nearly constant RFI — for example, the 137 MHz bin contains the continuous signal from a constellation of communications satellites. Next, we difference the data in time and frequency, flagging the data which produces 6σ outliers. Finally, we remove a fiducial foreground model, the process of finding this model is described in Section 6.1.5, and flag 4σ outliers of the residuals. The flags generated from this process were used in the production of Figure 3.1. A single set of flags is generated for all times and frequencies for the entire array for each night of data taking.

6.1.2 Compression

The volume of raw data generated in the EoR2011 season exceeds 10 TB, which is unwieldy for the level of computation required. While the relatively short integration times of 10s and relatively narrow channel widths of 50 kHz are useful for reducing the attrition of data due to RFI excision, these rates highly oversample both the frequency structure of foregrounds and EoR signal and the temporal structure of anything tied to the sky. To remedy the abundance of oversampled — and thus redundant — data, we employ a compression technique, first described in Parsons et al. (65), which critically samples the data in both time and frequency.

This compression algorithm hinges on two results from Section 2.3:

1. The delay of a smooth-spectrum point source is restricted to the range $\tau \leq |b|/c$, where $|b|$ is the baseline length.

2. Similarly, the fringe-rate of a source is restricted to the range $-(\nu b_E/c)\omega_\oplus \cos\delta_0 \leq f \leq (\nu b_E/c)\omega_\oplus$, where b_E is the east-west component of the baseline, $\omega_\oplus$ is the angular speed of the Earth's rotation, and δ_0 is the latitude of the array.

These two properties of drift-scan interferometers allow us to set limits on the fringe rate and delay at which celestial emission can enter the signal — this in turn allows us to set minimum integration times and channel widths which preserve that emission.

To set this minimum sampling rate in time, we inspect the maximum fringe rate, $(\nu b_E/c)\omega_\oplus$. We define the delay rate as the frequency-integrated fringe rate — this allows us to simultaneously compute this alongside the delay. Hence, the maximum delay rate allowed by celestial emission is $(b_E/c)\omega_\oplus$. The maximum delay rate across the entire PAPER array occurs in the 210m east-west baselines (between the leftmost and rightmost columns in Figure 6.1): 9.6 mHz. The Nyquist-Shannon sampling theorem dictates that an sampling time of 33s can completely describe this structure.

Though the maximum delay can be set as low as the horizon, we intend to preserve supra-horizon modes containing high-$k_{||}$ EoR modes. However, we achieve maximum sensitivity to these modes on the shortest baselines, and use the longest only for foreground characterization. Hence, we set the limit in delay to the horizon limit of the longest baseline in the array. This requires a sampling rate in frequency of 713 kHz. Setting the maximum delay allows supra-horizon modes to enter into the visibilities of short-baselines, including cosmological modes up to 0.38 $h\mathrm{Mpc}^{-1}$.

Figure 6.3 shows the extent of the skypass filters — defined as one on the intervals that contain emission (shown in the preceding discussion) and zero elsewhere. The skypass filters are shown atop the delay/delay rate transform of a visibility, confirming the claims about foreground signal's extent in these directions made in the text.

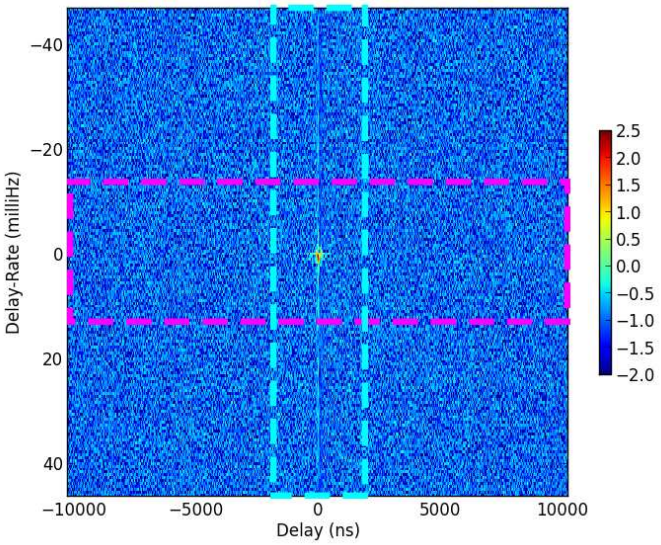

Figure 6.3: Delay / Delay rate transform of one days' worth of raw PAPER visibilities from a 30m baseline. The relatively fine sampling in frequency and time result in large ranges of delay and delay rate (respectively). A dashed, cyan box shows the skypass filter in delay, and the magenta, in delay rate. The fluxscale, in $\log_{10}$(Jy), is shown on the right. The skypass filters are designed to preserve all smooth-spectrum, celestial emission for the entire PAPER array, with baselines ranging from 30m (shown) up to 300m. For this short baseline, sky emission is contained within a relatively small range around 0 delay and 0 delay rate. Figure credit: Parsons et al. (65)

One serious hurdle to overcome in this compression process is the spectral and temporal structure introduced by nonuniform RFI flagging. To accommodate the scattering of signal into high delay/delay rate bins, we first deconvolve the data by the sampling function using a variation on the CLEAN algorithm, discussed in Section 6.1.5. The difference between this implementation of the algorithm and the foreground-removal strategy discussed in Section 6.1.5 is that we add the CLEAN components back into the residual spectra. This is the most computationally costly step.

Once the CLEAN deconvolution has been performed, we simply decimate the data, resampling the data at the rates described in the preceding paragraphs. While we could sample each baseline type with its own integration time and channel width, we set the limits based on the longest baselines — this both ensures a conservative application of this new procedure and allows for ease of data analysis and storage.

We implement the compression algorithm each night on the data using a 35 node computer cluster located on site. This allows us to perform all preprocessing steps up until this point, including compression, in real time as the data is taken. This algorithm reduces both the data rate and data volume by a nearly factor of twenty, reducing storage costs and required computational power.

6.1.3 Crosstalk Removal

For our purposes, crosstalk may be defined as a additive offset to the visibilities, which is stable on long timescales. To remove crosstalk, we simply subtract the nightly average of each baseline from each integration of that baseline.

6.1.4 Calibration

Calibration is a two step process. First, we solve for the antenna-based gains and delays which enforce redundancy among redundant baselines. This procedure is described in greater detail in Section 8.1. We treat the xx and yy polarizations of the array separately in this

analysis, linking the two calibrations with a cross-polarization delay and the assumption that all calibration terms are antenna-dependent. Next, we solve for the remaining four calibration terms — an overall flux scale for the x and y polarizations, and the delay of fiducial baselines — by fitting visibilities to a model of Pictor A (38).

We compute the calibration parameters using a relatively small amount of data — for two hours when Pictor A is overhead during a single day. We then apply these calibration terms to the entire seasons' data. Jacobs (36) and Parsons et al. (65) have shown that calibration terms remain constant for long timescales, and we take advantage of the remarkable stability of PAPER's calibration[23] to ease the computational burden of calibration. We could solve for all calibration terms on smaller timescales but we have found that this only causes slight improvements to the variance in our data. In practice, the errors caused by such a cavalier calibration effort can be absorbed into the uncertainty of the data, causing a roughly 5% increase in T_{sys}.

6.1.5 Foreground Removal

The final step before averaging multiple days is to remove a foreground model from the raw visibilities. Rather than removing a number of previously-identified sources from the data — this leaves us vulnerable to calibration errors as well as errors in the primary beam model — we employ a non-parametric method to remove foregrounds modelled on each visibility itself. Since most foreground sources are smooth spectrum and can be modelled as point sources, we model them as delta functions in delay (See Section 2.3). Rather than fitting for antenna gains and delays, a primary beam model, and the source flux at each frequency, we only fit for a single parameter: the flux in a given delay mode of a visibility.

To fit these fluxes, we use a version of the CLEAN algorithm (35), reduced to one dimension, and extended to allow complex flux values. We find the peak in the delay

[23]To me, that PAPER's calibration terms are stable for these long timescales is one of the more magical things about the instrument. I've even successfully applied calibration terms from one observing season to another!

spectrum of a single integration for one baseline, and subtract the kernel of the sampling function, weighted by the flux of that peak, from the delay spectrum. We iterate this process until the variance of the residual spectrum is 10^{-8} times that of the original spectrum. We restrict the algorithm to peaks found within the horizon limits, described in Section 2.3, which enforces that smooth-spectrum foreground sources be removed. This procedure both removes foregrounds and deconvolves from the spectral sampling function created from the flagging of RFI.

In the limit where all spectral bins contain data, this is simply a notch filter which nulls inter-horizon delay modes.

6.1.6 Averaging Multiple Days

As a final excision of spurious signals (most likely due to RFI), for each day, we flag outlying measurements in each bin of LST and frequency. We use measurements of T_{sys} outlined in the Section 6.1.8 to estimate the variance in each bin, flagging 3σ outliers.

If the data followed a complex normal distribution, consistent with pure, thermal noise, then we would expect this procedure to flag one measurement per frequency/LST bin, causing a slight miscalculation of statistics post flagging. Most notably, this causes an underestimate in the variance of the power spectrum. To counteract this effect, we calculate the ratio of the variance of a normal distribution truncated at $\pm 3\sigma$ to the variance of its parent distribution (97.3%). Henceforth, we will increase all errors in the power spectrum by a factor of $1.03 \approx 1/97.3\%$ to accommodate for this error.

We compute the mean of the RFI-removed data for each bin of LST and frequency, creating a dataset comprised of a single, fiducial day. We continue analysis on this averaged dataset.

6.1.7 Final Processing

After visibilities are averaged in LST, a final round of crosstalk removal is performed. Again, we simply subtract the daily average from the data. Much of the crosstalk lies beneath the sensitivity level of only one day's worth of data, and appears in the LST-averaged dataset. Recomputing the mean with the increased sensitivity of an averaged dataset allows for a more accurate removal.

In the penultimate processing step, we pass the data through a second low-pass filter in time. Sections 2.3 and 6.1.2 describe the celestial limits of the fringe rate for drift-scan arrays (f) as $b_E\omega_\oplus \cos\delta_0 \leq f \leq b_E\omega_\oplus$, where b_E is the east-west component of the baseline, $\omega_\oplus$ is the angular velocity of the Earth's rotation, and δ_0 is the latitude of the array. We filter the data in time using a boxcar filter, defined as one on $0 \leq f \leq b_E\omega_\oplus$ and zero elsewhere. While this filter does null some celestial emission (roughly the area between the south celestial pole and the horizon), its effect is small, since the primary beam heavily attenuates these areas of the sky. We null these fringe rates as an additional step of cross-talk removal.

Finally, we rotate the linearly polarized visibilities into Stokes visibilities, defined in Equation 2.36.

6.1.8 System Temperature

Alongside the calculation of statistics for binning in LST and frequency, we take advantage of the nightly redundancy as a check on the data. Since PAPER is a tracking array, measurements taken at the same LST on different nights should be totally redundant. This redundancy allows us to measure the system temperature via fluctuations in signal in the same LST bin from day to day.

First, we compute the variance in each frequency and LST bin over all nights of data ($\sigma^2_{J_y}(\nu, t)$), and convert this variance into a measurement of the system temperature T_{sys}. This measurement is totally independent of the following power spectrum analysis, and can be used to quantify the level of systematic and statistical uncertainty in the power spectra.

It compliments measurements of T_{sys} in Parsons et al. (65) and Jacobs et al. (39). The variance computed in each LST/frequency bin is converted into a system temperature in the usual fashion:

$$T_{sys}(\nu, t) = \frac{A_{eff}}{k_B} \frac{\sigma_{Jy}}{\sqrt{2\Delta\nu t_{int}}}, \tag{6.2}$$

where A_{eff} is the effective area of the antenna (see Figure 3.3), k_B is the Boltzmann constant, $\Delta\nu$ is the channel width, and t_{int} is the integration time of the LST bin.

Figure 6.4 shows the measured system temperature for each frequency and LST bin collected during the EoR2011 observing season. To further summarize our data's variance, we can average $T_{sys}(\nu, t)$ over the time- and frequency-axes. The frequency-averaged system temperature is computed as

$$\langle T_{sys} \rangle(t) \equiv \frac{\int_{\Delta\nu} W(\nu) T_{sys}(\nu, t)\ \mathrm{d}\nu}{\int_{\Delta\nu} W(\nu)\ \mathrm{d}\nu}, \tag{6.3}$$

where $W(\nu)$ is the spectral window function, and the integral is computed over the frequency band $\Delta\nu$. For our analysis, we use a Blackman-Harris window function (32), chosen to maximally suppress sidelobe levels. A similar expression may be written for the time-axis, where our window function is simply the number of redundant samples in each frequency channel.

Figure 6.5 shows the system temperature averaged over frequency and LST ranges used to compute the power spectra. T_{sys}, averaged in both frequency and time for both bands are reported in Table 6.1

6.2 Power Spectra

Now that the data are processed and averaged, we begin computing power spectra by a two-step process. First, we remove off-diagonal covariances from the correlation matrix of two baselines — this results in a power spectrum for each integration time of the fiducial,

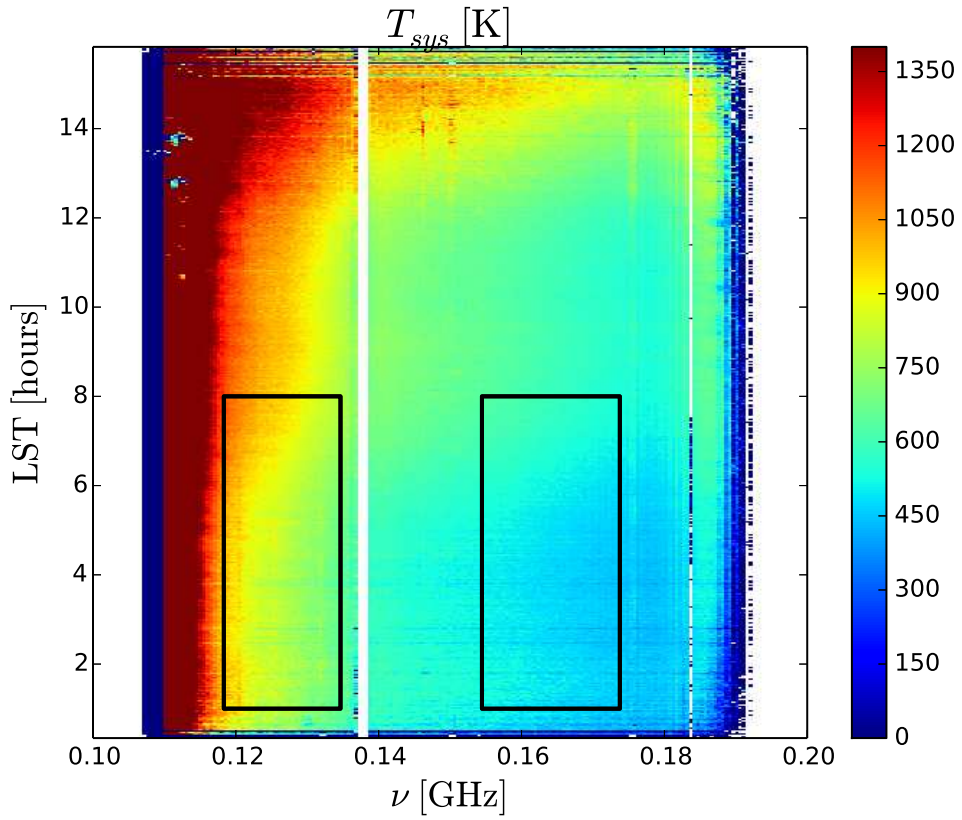

Figure 6.4: System temperature in Kelvin as a function of LST and frequency ν, calculated by Equation 6.2. Black boxes enclose the range in LST and ν used to compute the power spectra.

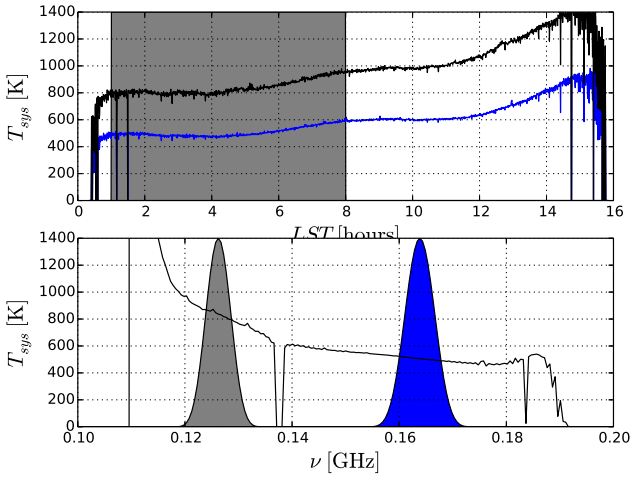

Figure 6.5: (Top Panel) band-averaged system temperature (Equation 6.3) as a function of LST for Bands I and II in black and blue, respectively. The shaded grey region indicates the range in LST used to compute the power spectra. (Bottom Panel) Time-averaged system temperature, averaged over LST 1h00m until 8h00m. The shaded grey and blue regions show the spectral window functions for Bands I and II, respectively.

averaged day and for each baseline pair. Next, we bootstrap multiple times and baseline pairs to characterize the statistics of the distributions of power spectra.

6.2.1 Covariance Removal

Even after having undergone several layers of RFI excision and crosstalk removal, the data still show large covariances between delay bins. These covariances dominate the averaged power spectrum despite varying wildly between baseline pairs. We remove them via the covariance removal strategy described in Appendix C of Parsons et al. (65).

This strategy essentially diagonalizes the average covariance matrix between all baseline pairs. Since we expect all redundant measurements to see the same sky, and we expect all k-bins of the power spectrum to be independent, then we expect sky signal to appear in the diagonal elements of the covariance matrix between two delay-transformed visibilities. By measuring the full covariance matrix of all baseline pairs, we can estimate the instrumental systematics which would leak signal from one k-bin to another. By inverting the mean covariance matrix of all baseline pairs and dotting this into a delay-transformed visibility, we remove our best guess at the covariances in that visibility.

Once we estimate and remove covariances from the delay-transformed visibilities, we may proceed with power spectrum estimation. We direct the reader to Parsons et al. (65) for a more detailed discussion.

6.2.2 Results

The covariance removal described in the previous section projects the delay-transformed visibilities into a basis in which the covariance between two redundant baselines is diagonal, and then computes the power spectrum from the projected delay spectra. This procedure produces an estimate to the power spectrum for each LST bin and baseline type. To measure the uncertainties in the time-dependent power spectra, we bootstrap over groups of

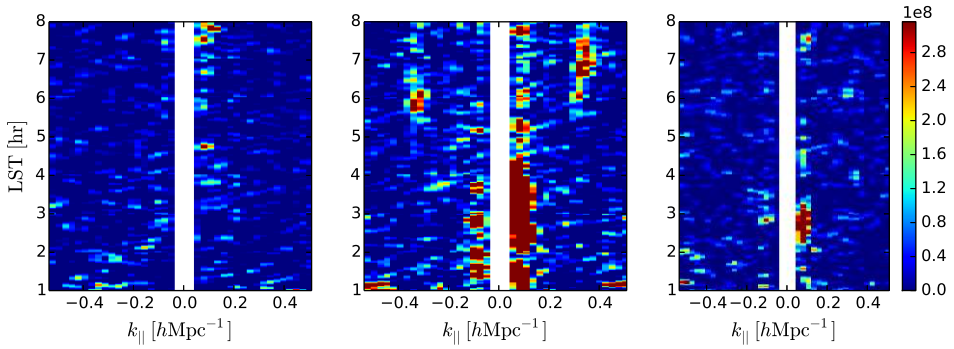

Figure 6.6: Power spectra, in units of mK^2 $(h\text{Mpc}^{-1})^3$ shown for each $k_{||}$ and LST measured. $k_{||}$ modes within the horizon for the 30m baselines used are masked. The left panel show I, the middle panel, $P_P = P_Q + P_U$, and the right panel shows P_V. An excess of power in P_P below LST 4h30m could indicate polarized emission. The excess at $k_{||} \approx 0.35$ $h\text{Mpc}^{-1}$ between LST 6h00m and 8h00m could be generated by a 170 mJy source with rotation measure between 42 and 59 m^{-2}.

redundant baselines. Figure 6.6 shows the linearly polarized power spectra from unpolarized emission (I), linearly polarized emission ($P = Q + iU$), and circularly polarized emission (V). as functions of LST, computed via this bootstrapping.

There are two features in P_P worth noting. First is the excess of emission at $0 \lesssim k_{||} \lesssim 0.2$ $h\text{Mpc}^{-1}$, between right ascension 1h00m and 4h30m. That the excess survives the LST averaging over 82 days indicates that it is fixed to the sky, and that it exceeds its corresponding k-bins in P_I indicates that it is polarized. Furthermore, it roughly corresponds with the diffuse, polarized power shown in Bernardi et al. (8), which takes its minimum value at around LST of 5h00m.

The second feature of Figure 6.6 that we will comment on is the track of excess power from fight ascension 6h00m to 8h00m, at $k_{||} \sim 0.35$ $h\text{Mpc}^{-1}$. This type of excess power could be generated by a polarized point source. This stripe satisfies the two criteria put forth to indicate the excess at lower right ascension could be polarized emission, and it appears to

be localized in $k_{||}$, a feature of polarized point sources behind a single Faraday screen.

What properties of a polarized point source would be necessary to generate the excess shown? To answer this question, we describe the source spectrum with three parameters: a flux S, a geometrical delay τ_g, and a rotation measure Φ. The power spectrum of this point sources $P_1(k)$, taken from Equation 2.54 becomes

$$P_1(k) \approx \left(\frac{\lambda^2}{2k_B}\right)^2 \frac{X^2Y}{\Omega} S^2 \delta\left(k_{||} - \frac{dk_{||}}{d\eta}\tau_g - k_{leak}(\Phi)\right), \tag{6.4}$$

where $dk/d\eta$ is the linear conversion from delay into $k_{||}$ ($1/Y$ of Equation 1.18), and $k_{leak}(\Phi)$ is the $k_{||}$ mode most infected by rotation measure Φ (Equation 4.17). All other terms agree with Equation 2.54.The bandwidth in the denominator of Equation 2.54 is removed to properly normalize the delta function. Note the important result from Section 4.5 that a rotation measure component to a spectrum *adds* to the normal position of a source in $k_{||}$.

The potential source peaks in power at right ascension of 6h52m at a $k_{||}$ value of 0.345 $h\text{Mpc}^{-1}$. Its power peaks at 4.62×10^8 $\text{mK}^2(h^3\text{Mpc}^{-3})$. A point source whose right ascension is between 6h00m and 8h00m, with an apparent flux of 173 mJy, whose rotation measure is between 42 and 59 m^{-2} satisfies the requirements to generate this type of emission. To confirm the existence of such a source would require follow-up observations with an imaging array, but the excess power at that location in the P_P spectrum is well-described by such a source.

Figure 6.7 shows the power spectra of the four Stokes parameters in Band I, and Figure 6.8 sohws that of Band II. Sensitivity limits using the T_{sys} computed in Section 6.1.8 and the sensitivity calculations of Parsons et al. (62) and Pober et al. (68) are shown in dashed, cyan lines. The noise level of these power spectra are computed by examining the spread of bootstrapped power spectra, where we bootstrap over both redundant baselines and LST samples. These noise levels are considerably higher than the predictions from the T_{sys} of Section 6.1.8. The increase in T_{sys} is most likely due to calibration errors and systematics not targeted by the covariance removal.

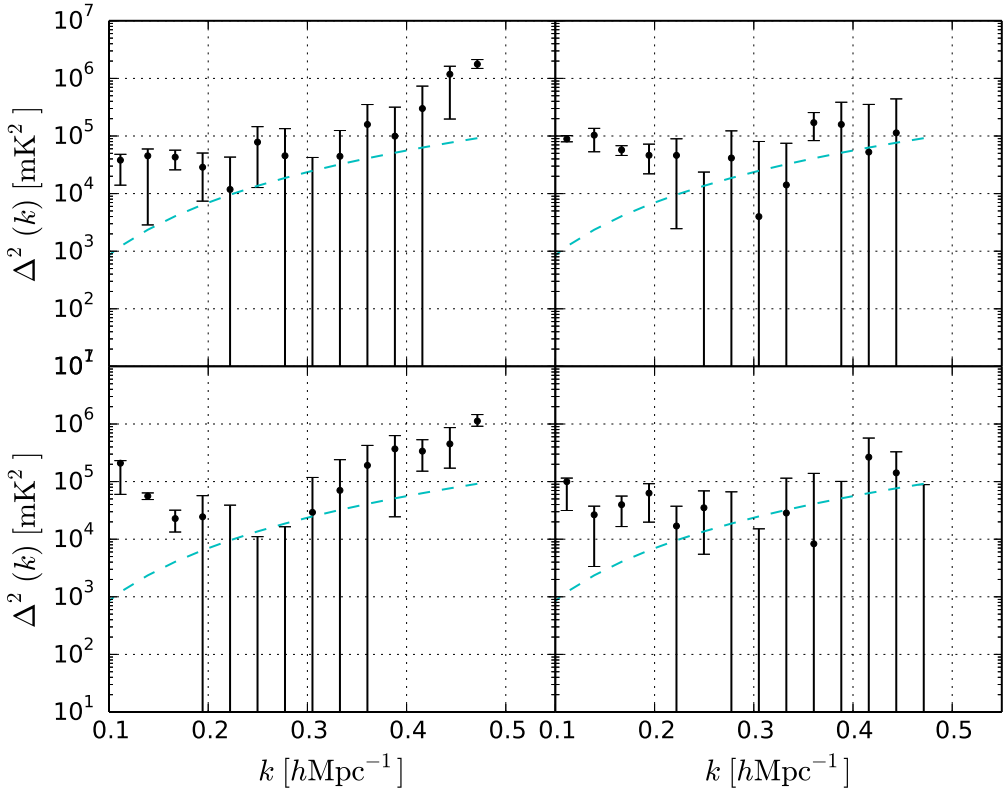

Figure 6.7: Spherically averaged power spectra for the four Stokes parameters: I in the top left panel. Q in the top right, U in the lower left, and V in the lower right. The data from Band I are shown. Error basrs show the 98% confidence intervals derived from bootstrapping over all samples in LST and all redundant baselines. Dashed, cyan lines show the theoretical level of thermal fluctuations, with T_{sys} calculated in Section 6.1.8.

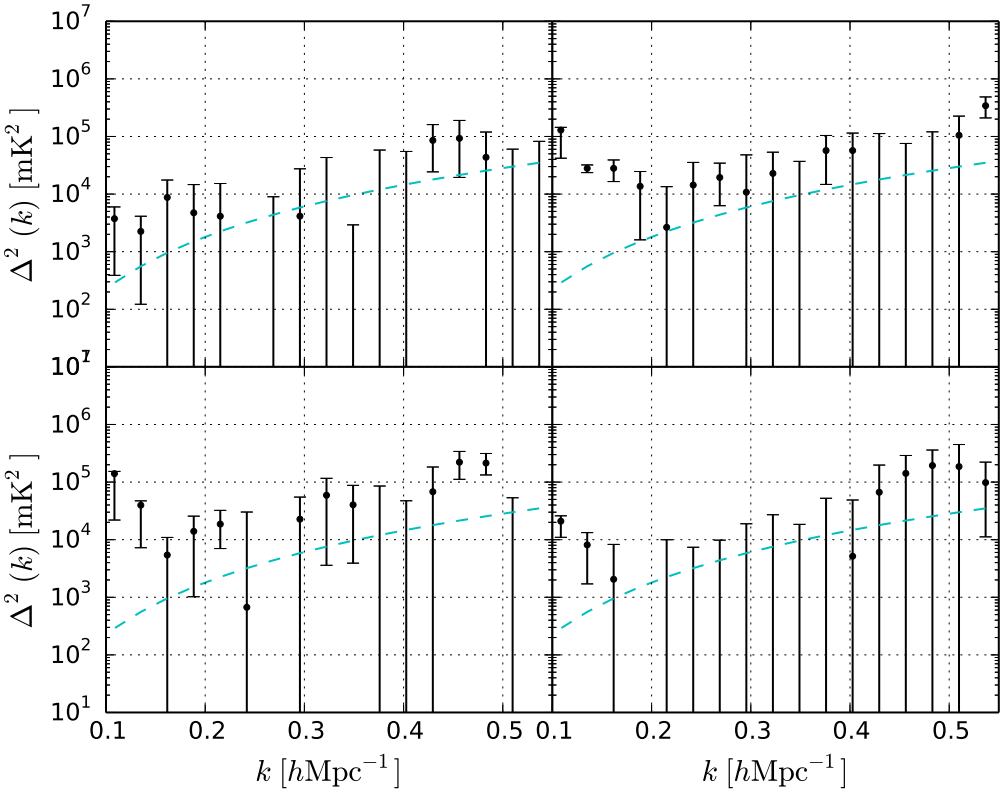

Figure 6.8: Same as figure 6.7, for Band II.

The level of leakage predicted from the arguments of Section 4.3 show that in the lowest $k_{||}$ bins, the I power spectrum cannot be dominated by $Q \rightarrow I$ leakage. The levels of polarized leakage in P_I, to an order of magnitude, are 10^3 mK^2 in Band I, and 10^2 mK^2 in Band II. These levels are well below the systematics which dominate the lowest $k_{||}$ bins of the I power spectrum, and are also well below the levels of $Q \rightarrow I$ leakage predicted by the simulations in Section 5.2. We will investigate this further in Section 6.3.

6.2.3 Ionospheric Effects

Daily changes in the Faraday depth of the Earth's ionosphere could potentially attenuate polarized signal. As the total electron content (TEC) varies, it modulates the incoming polarized signal by some Faraday depth that which is a function of both the local TEC of that time, and the strength of the Earth's magnetic field. Though we assume visibilities are redundant in LST, they do have slight variations due to the variable TEC of the ionosphere. Thus, averaging in LST could result in some attenuation of signal.

To quantify this, we first assume that the ionospheric TEC is constant over the PAPER beam, and assume that from day to day, the Faraday depth of the ionosphere is a random variable, based on the TEC measurements of Datta et al. (17). For now, we neglect the day-to-day correlations, though we can check the effects of any correlation later.

We begin by writing the LST-averaged visibility as the Faraday depth-weighted sum of otherwise redundant visibilities:

$$V' = \frac{1}{N} \sum_i e^{-2\Phi_i \lambda^2} V, \tag{6.5}$$

where Φ_i is the ionospheric Faraday depth from day i and V is the redundant component of the visibilities. Using this expression, we compute the magnitude of the rotated visibilities, which is proportional to the power spectrum,

$$P' = |V'|^2 = \frac{1}{N^2} \left(\sum_{i,j} e^{-2i(\Phi_i - \Phi_j)\lambda^2} \right) |V|^2. \tag{6.6}$$

When $i = j$, the term in parentheses becomes one, and the i, j component of thes sum is the conjugate of the j, i component. This allows us to rewrite the sum in terms of the $i = j$ component, and the $j > i$ components, now written as cosines:

$$\sum_{i,j} e^{-2i(\Phi_i - \Phi_j)\lambda^2} = N + 2\sum_{i>j} \cos\left\{2(\Phi_i - \Phi_j)\lambda^2\right\}. \tag{6.7}$$

In the limit where all values of Φ_i are equal, the second term becomes $N(N-1)/2$, the number of i, j pairs with $i > j$. This produces the desired result that with no daily fluctuations in ionospheric Faraday depth, there is no effect on the signal. In the limit of totally uncorrelated data (i.e. $\langle \Phi_i \Phi_j \rangle \propto \delta_{ij}$), this term takes its minimum value, maximally attenuating the measured power spectrum. Hence, considering uncorrelated Φ_i gives the worst-case scenario.

To estimate the level of ionospheric attenuation, we estimate the attenuation factor in Equation 6.7. Using typical TEC values of 6×10^{16} m^2 and a typical value for the Earth's magnetic field at the PAPER site[24], we calculate a rotation measure for each day, and estimate the attenuation factor, and estimate the attenuation factor for eighty-two days, as the data was averaged over that period of time.

This procedure yields a distribution of values for the attenuation factor which peaks at 88%. Considering that we have neglected day-to-day correlations of the TEC, which increases this factor, decreasing the level of attenuation, we assume that attenuation due to the ionosphere is negligible and do not adjust our results for it.

6.3 Updated Polarization Fractions

6.3.1 Scaling the Simulations

Figure 6.9 compares the measured Q and I power spectra to those simulated in Chapter 5. Since the measured values consistently disagree with the simulations, we can constrain the

[24] `http://www.ngdc.noaa.gov/geomag/magfield.shtml`

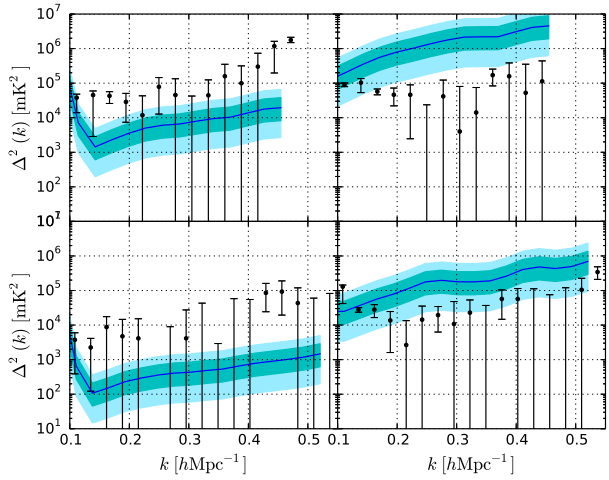

Figure 6.9: Top row: measured and simulated power spectra for I (left) and Q (right). Measured power spectra are in black, simulated values are in blue (median value), cyan (68% confidence interval), and light cyan (95% confidence interval). Simulations are generated as in Chapter 5, with a mean polarized fraction of 2.01%. Bottom row: Same as the top row, for Band II.

input parameters to the simulations, beginning with a simple scaling relation:

$$P_k = x^2 S_k, \tag{6.8}$$

where P_k is the measured Q power spectrum in the k^{th} bin, S_k is the simulated power spectrum in that bin, and x^2 is the scale factor between the two. We choose to use a scale factor of x^2 rather than x in order to facilitate the interpretation of x as an adjustment to the mean polarized fraction of point sources, as we will soon discuss.

For the duration of this section, we will approximate the measured power spectrum $\hat{P}_k$ as normal, random variables,

$$\hat{P}_k \sim \mathcal{N}(P_k, \sigma_k^2), \tag{6.9}$$

with mean P_k and variance σ_k^2, derived from the distribution of bootstrapped power spectra.

We find the scale factor x that best fits Equation 6.8, and then interpret its physical meaning. The likelihood of drawing a simulated power spectrum $\mathcal{S}$ by a factor x^2 given the measured data $\mathcal{D}$ is

$$P(x, \mathcal{S}|\mathcal{D}) \propto \exp\left\{-\frac{1}{2}\sum_k \frac{|x^2 S_k - P_k|^2}{\sigma_k^2}\right\}, \tag{6.10}$$

where the sum extends over all available values of k.

By marginalizing over $\mathcal{S}$, we find the likelihood of x:

$$P(x|\mathcal{D}) \propto \int P(x, \mathcal{S}|\mathcal{D}) P(\mathcal{S})\ \mathrm{d}\mathcal{S}. \tag{6.11}$$

Here, $P(\mathcal{S})$ is the joint probability of all k-bins of the simulation, i.e. $P(S_0, \ldots, S_n)$ for k-bins labelled 0 to n, and $\mathrm{d}\mathcal{S}$ denotes the n values of $\mathcal{S}$ over which we integrate. We compute the integral in Equation 6.11 by the Monte Carlo technique, sampling $\mathcal{S}$ from different instances of the simulation. This encapsulates both the probability distribution functions of each S_k and the covariances between k-bins in $\mathcal{S}$. To insulate the result from potentially damaging effects of the foreground removal (Section 6.1.5), we do not consider k-bins within the horizon in this integral.

To find the most likely value of x which would produce the measured power spectrum, we turn to Bayes theorem, $P(\mathcal{D}|x) \propto P(x)P(x|\mathcal{D})$, where $P(\mathcal{D}|x)$ is the posterior distribution of $\mathcal{D}$ and $P(x)$ is our prior on x. Since x^2 is a scale factor, we choose to use Jeffrey's prior in x^2, which sets $P(x) \propto 1/x$.

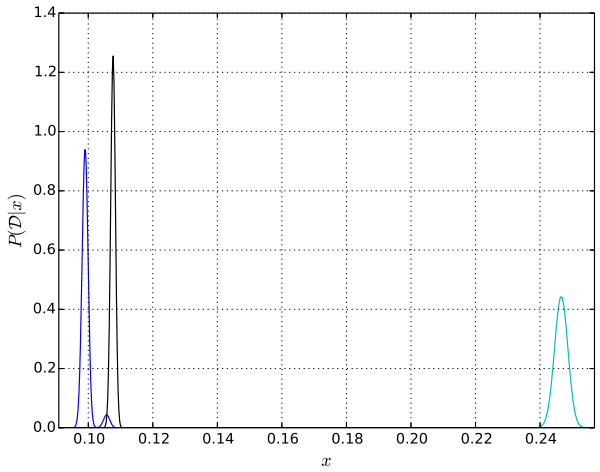

Figure 6.10: Posterior distributions for the data, given scale factor x. Blue shows that from Band I; cyan, from Band II; and black shows the joint posterior from both bands. Moments of these distributions are summarized in Table 6.2.

Measurements from the different bands can be summarized into a joint posterior by simply computing the product of the posterior of each band. This assumes that each band is independent, a reasonable assumption given the high level of noise in the measured power spectra. Figure 6.10 shows the posterior distributions of $\mathcal{D}$ given x for Bands I and II, alongside the joint posterior. The moments of the three distributions are summarized in Table 6.2.

Band	$\bar{x}$	σ_x	Implied Mean Polarized Fraction
I	0.99	0.002	2.0×10^{-3}
II	0.247	0.002	5.0×10^{-3}
Both	0.108	0.001	2.2×10^{-3}

Table 6.2: Moments of $P(\mathcal{D}|x)$.

6.3.2 Why is x Related to the Polarized Fraction?

As mentioned in the previous section, we interpret the scale factor x as an adjustment to the mean polarized fraction of point sources. The simulations parameterize each point source with a polarized fraction p, an unpolarized flux f, and a rotation measure Φ. All sources are given a spectrum $pf \exp\{-2i\Phi\lambda^2\}$, where λ^2 is the squared wavelength. This is a simplified account of Equation 5.4, but encapsulates the relevant quantities for this discussion. Since the source counts are well measured at these frequencies (26), and rotation measures are also well-measured (60) and independent of frequency, we regard the distributions of these two quantities to be fixed. Hence, any constraints we place on these simulations can be considered as updates to the distribution of polarized fraction.

The amplitude of the power spectrum in the simulations (S_k) can be expressed in terms of the source fluxes f_i, and the polarized fractions p_i.,

$$S_k \sim \sum_{i,j} p_i p_j f_i f_j \equiv \bar{p}^2 \sum_{i,j} \pi_i \pi_j f_i f_j, \tag{6.12}$$

where we have defined $\pi_i \equiv p_i/\bar{p}$ as the ratio of a single polarization fraction p_i to the mean, $\bar{p}$. Hence, the simulated power spectra are proportional to the mean polarized fraction squared, $\bar{p}^2$, and we can interpret the scale factor x as the fractional change in the mean polarized fraction.

Table 6.2 gives the implied mean polarized fraction of point sources for the two bands and the joint posterior. In these simulations, we drew polarized fractions from a distribution

with a mean of around 2%, and now we can set a limit about an order of magnitude lower.

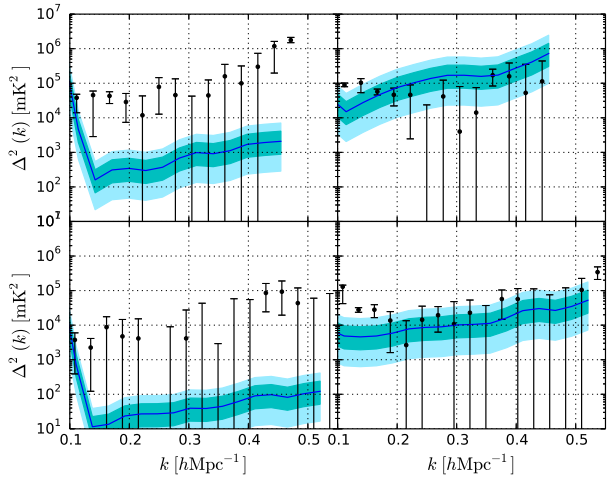

Figure 6.11: Same as Figure 6.9, but using a polarized fraction distribution scaled by $x = 0.108$, the maximum likelihood value of x using both bands.

Figure 6.9 shows updated simulations using the implied polarization fraction of the joint posterior from Figure 6.10: 2.2×10^{-3}. The simulated $Q \rightarrow I$ leakage now lies between 10 and 100 mK^2 in Band II, around the expected level of the 21cm EoR power spectrum at redshift 7.

6.3.3 On the Applicability of the Simulations

We now turn our attention to a qualitative discussion of the diffuse emission found in Bernardi et al. (8, B13) and Jelić et al. (40, J14), and how applicable the simulations are to the polarized emission found in those measurements, which is largely characterized as being diffuse and as having low rotation measures. We will argue that the simulations may apply to both these two measurements, and also that the simulations are a valid point of comparison to the measurements made in Section 6.2. This discussion builds on Chap-

ter 5, and justifies the use of the simulation as a tool for understanding the power spectra presented in Section 6.2.

Both B13 and J14 show diffuse, weakly polarized emission found at relatively low rotation measure ($|\Phi| \lesssim 25$ m^{-2}. These measurements differ from the input sources of the simulations in two ways: in the choice of rotation measures included, and in the spatial correlation of power. We will discuss these in turn.

The simulations sample rotation measures from the entire Oppermann et al. (60) map, rather than restricting to a particular field of view. Two effects may arise from such a generality. First, the inclusion of large rotation measures could scatter power to larger k in the simulations than in reality, and second, uncorrelated polarization vectors in the simulation could have a depolarizing effect. This concern was addressed in Section 5.3 by computing the simulation with all rotation measures drawn from a pointing at the galactic south pole (coincidentally, the B13 field), with maximally correlated polarization vectors. The result of this test was identical to the random drawing (Figure 5.7). The reproduction of power spectra between the two simulations indicates that rotation measure is not a dominant factor in determining the shape or the amplitude of the power spectrum of polarized emission.

The rotation measures in B13 and J14 are considerably lower than those sampled in the simulations. Since there is overlap between our field and the B13 field, our measurements include lower rotation measures than sampled in the simulation as well. A comparison with our field and the Oppermann et al. (60) maps shows that this is due simply to our choice of pointing. Again, the results of Section 5.3 show that this does not significantly affect the power spectrum.

There is another more subtle difference between the Faraday depths of diffuse emission and point sources. Since we expect diffuse emission mostly to be generated from within our galaxy, we expect it to be emitting from within the magnetized, ionized plasma that rotates its polarization vector. As discussed in Jelić et al. (41), this creates both a depolarizing effect and structure in the rotation measure spectrum of the source.

The simulations account for this by distributing sources on smaller scales than the resolution element of the array.[25] Since each source is assigned an independent rotation measure, and many sources are placed within the inverse baseline length ($\theta \sim 1/u$), then the visibility averages over many rotation measures and polarization angles per pointing. This has the same effect as a polarized source emitting from within an ionized, magnetized plasma — different lines of sight summed within the same resolution element of an array produce a complex Faraday depth spectrum, and they also add incoherently.

Next, we address the spatial correlation of emission. The simulations assume an isotropic placement of point sources. Projecting the fringe onto the simulated sky selects the modes correlated on the baseline scale (in our case, 3°), so the simulation represents any power correlated on those 3° scales — that we model it as a series of point sources in many ways is irrelevant, since it does not affect this fact. Figure 5.6 shows the agreement of these simulations with existing angular power spectral measurements, which indicates that the level of the power correlated on 3° scales in the simulations agrees with real, diffuse emission.

Since the simulations produce angular power spectra that agree with measurements, and address spectral differences between polarized point sources and diffuse emission, we take the simulations as a good reference for all types of polarized emission. Hence, for the 15λ baselines we simulate and measure, any adjustments to the simulation can be considered as adjustments to our understanding of the polarized sky.

6.3.4 Comparison to Other Measurements

As mentioned before, most measurements of the polarized sky at meter wavelengths detect large amounts of diffuse polarized emission, compared to the relatively few point sources. Bernardi et al. (8, abbreviated in this section as B13), in a 2400 square degree, detect a

[25] I'm trying very hard not to use the word "synthesized beam" here. We are looking at the power spectrum on a single baseline — there is no synthesis, so a "synthesized beam" doesn't really make sense.

single point source, PMN J0351-2744, whose polarized flux is 320 mJy. Since this was the only source detected in this survey, they claimed that the polarized fraction of sources must be bounded at 2%.

We ask if our measurements agree with the B13 detection and upper bound. First, we compute the probability of detecting a source whose polarized fraction is greater than 2% in the context of our measurements. We scale the log-normal distribution of polarization fractions in Tucci and Toffolatti (85) by the maximum likelihood value of x from the joint distribution of both bands (Table 6.2). Integrating this distribution above 2% yields the probability of detecting a polarized point source above 2%. That probability is 5%. While this implies that 2% cannot be counted as a strict upper limit, as B13 imply, it does roughly agree with their sstatement that detecting sources this polarized is unlikely. Thus, we can relax their strict upper limit to a 2σ upper bound in the polarization fraction.

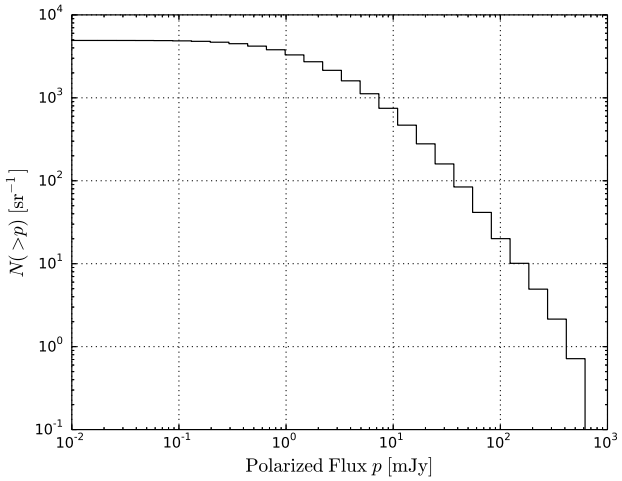

Figure 6.12: Number counts of polarized sources, from a simulation with mean polarized fraction of 2.2×10^{-3}, derived from the power spectra in Figures 6.7 and 6.8. This is a convolution of the unpolarized source counts (26), and the polarized fraction distribution (85), scaled by the maximum-likelihood value of x (Table 6.2).

Second, we ask if our updated simulations can produce the occurrence of sources like PMN J0351-2744. Figure 6.12 shows the simulated, integrate source counts from the updated simulation. These source counts imply that one source with a polarized flux of 320 mJy occurs roughly every 1700 square degrees. These number counts are in close agreement with the detection of one source of this strength in 2400 square degrees.

This data does have the sensitivity to detect sources like PMN J0351-2744, but its location amidst other polarized emission provides difficulty isolating it, as can be seen in Figure 6.6.

6.3.5 Closing Remarks

Though PAPER in its grid configuration is incapable of creating the high dynamic-range images needed to isolate polarized point sources, there are several hints in the data indicating the presence of polarized foregrounds. The power we described in the previous sections is consistent with the general properties of diffuse, polarized emission described by other measurements (6, 8, 40, 67, e.g.). Follow-up observations with arrays more suited for imaging will be necessary to fully detect and characterize this emission.

Even with the much reduced polarized fraction inferred from this emission, the implied level of $Q \rightarrow I$ leakage exceeds the expected level of the 21cm EoR power spectrum (48, 57, e.g.). This excess presents a challenge for ongoing and future observations. There are two mitigation strategies. First, polarized point sources may be identified and subtracted, as in Geil et al. (24). Subtracting to the requisite levels will require highly accurate models for an unreasonable number of sources, as we showed in Section 5.4. The second mitigation strategy will involve the design of future instruments, limiting $Q \rightarrow I$ leakage. Engineering a symmetric primary beam could limit instrumental polarization, and is one of the drivers of the design of the HERA array.

Part III

Polarimetric Tools

Chapter 7

The Discrete Rotation Measure Transform

7.1 Background and Justification

One of the key science goals for the Square Kilometre Array is to investigate the nature of galactic and cosmic magnetism (1, e.g.). A powerful probe for characterizing those magnetic fields is the Faraday rotation of polarized emission, discussed in Section 4.1. As a brief reminder, the polarization vector of emission passing through an ionized, magnetized plasma will incur a polarization-dependent phase rotation given by

$$(Q + iU)_{meas} = e^{-2i\Phi\lambda^2}(Q + iU)_{inc}, \tag{7.1}$$

where Q and U are the Stokes parameters, λ is the wavelength of the emission, the subscripts *inc* and *meas* denote the incident and measured polarization angles, and Φ is the rotation measure, defined as

$$\Phi = \lambda^2 \frac{e^3}{(m_e c^2)^2} \int B_{||}(s) n_e(s)\ \mathrm{d}s. \tag{7.2}$$

The rotation measure is the line of sight component of the magnetic field times the number density of electrons, integrated along the line of sight to the emitting source.

Brentjens and de Bruyn (11, abbreviated to BdB) present a novel approach to isolating the rotation measure structure of a polarized source: rotation measure synthesis. Instead of the usual route of fitting a quadratic function in λ^2 to the polarization angle, taken in several older studies (80, e.g.), BdB take advantage of the Fourier relationship between Φ and λ^2 in Equation 7.1. They prescribe a method of disentangling the Φ structure, analogous to a Fourier transform. In the continuum limit, this transform is written as

$$\widetilde{F}(\Phi) = \int P(\lambda^2) e^{-2i\Phi\lambda^2} \, \mathrm{d}\lambda^2, \tag{7.3}$$

where $P(\lambda^2) \equiv Q(\lambda^2) + iU(\lambda^2)$ is the total polarized flux, and $\widetilde{F}(\Phi)$ is the rotation measure transform (RMT) of the spectrum.

In general, low frequency arrays measure spectra that are spaced constantly in frequency. This type of sampling, while relatively easy to produce, prohibits the estimation of Equation 7.3 by a conventional discrete Fourier transform (DFT). BdB and others (2, e.g.) suggest working in the limit where the spacings in λ^2 remain roughly constant over the spectrum, in which case, a valid approximation of the RMT can be written as the sum

$$F_{approx}(\Phi) \approx \left(\sum_i w_i \right)^{-1} \sum_i w_i F_i e^{-2i\Phi\lambda_i^2}, \tag{7.4}$$

where the weights w_i can be taken all to be 1. Unfortunately, this approximation breaks down at low frequencies, in which λ^2 rapidly changes across a normal observing bandwidth. To give a sense of the rapidity of this change, we give a simple example band which may be representative of those measured with instruments like PAPER, LoFAR (73), GMRT, and the MWA (84). That band extends from 140 to 180 MHz, with 100 evenly spaced channels. The difference in observed wavelength-squared will vary from 0.012 m^2 (in the highest frequency bin) to 0.026 m^2 (in the lowest) — around a factor of two!

The breakdown of the assumption that evenly spaced frequencies are evenly spaced in λ^2 leads us to revise the standard prescription for the RMT. This assumption is implicit in

the setting of w_i to one, since it implies that the measure of each frequency bin in Equation 7.4 is equal.

7.2 The Mechanics of the DRMT

Now, we will discuss the computational details of estimating the integral 7.4. This discussion builds upon and is intended to complement the work done in BdB, in which rotation measure synthesis is described, and a prescription for its computation is given. We will review the discussion of BdB and add a discussion of the effects of discretizing the RMT, creating (in an analogy to the DFT) a discrete rotation measure transform (DRMT)

We begin the prescription for a DRMT by assigning λ^2 values to each frequency bin of data. Following BdB, we estimate the average value of λ^2 in the i^{th} frequency bin (ν_i) in a spectrum with evenly spaced frequencies, whose spacing is $\Delta\nu$. We define the discrete values of the squared wavelength λ_i^2 as

$$\lambda_i^2 = \frac{1}{\Delta\nu}\int_{\nu_i-\Delta\nu/2}^{\nu_i+\Delta\nu/2}\frac{c^2}{\nu^2}\,\mathrm{d}\nu = \frac{c^2}{\nu_i\Delta\nu}\left[\left(1-\frac{1}{2}\frac{\Delta\nu}{\nu_i}\right)^{-1}-\left(1+\frac{1}{2}\frac{\Delta\nu}{\nu_i}\right)^{-1}\right]\approx\left(\frac{c}{\nu_i}\right)^2 \tag{7.5}$$

where c is the speed of light. To approximate the band-averaged value of λ^2, we have used the assumption that $\Delta\nu \ll \nu_i$.[26] Since frequencies are spaced evenly, λ_i^2 cannot be; ergo, there is no set of rotation measures Φ_j which allow the DRMT to evenly sample the unit circle, a crucial property of the choice of frequencies in the DFT. Formally, there are no solutions to the equation

$$2(\lambda_i^2-\lambda_0^2)\Phi_j = 2\pi\frac{ij}{N} \tag{7.6}$$

which are independent of i.[27] Here i runs on the interval $[0, N)$ and $j \in [-N/2, N/2)$. We will discuss the implications of the lack of solutions to Equation 7.6 throughout this section,

[26] That this is different than the value cited in BdB, $\lambda_i^2 = (c/\nu)^2(1+3\Delta\nu/4\nu_i)$, due to our only approximating to first order in $\Delta\nu/\nu$. I personally haven't found any justification for going to a higher order.

[27] As a reminder, the frequencies of a DFT are chosen such that $2\pi(t_i - t_0)\nu_j = 2\pi(ij/N)$. This allows for an even sampling of the unit circle on $[-\pi, \pi)$ and allows for a complete basis of frequencies.

but the first and most obvious of which prohibits a straightforward set of rotation measures to sample.

To find an optimal set of rotation measures to sample, we must first approximate the spacing between adjacent frequency bins and then choose a set of rotation measures accordingly. To do this, we simply Taylor expand the spacing in λ^2 as

$$\Delta\lambda_i^2 \approx \Delta\nu \left| \frac{\partial \lambda^2}{\partial \nu} \right|_{\nu=\nu_i} = 2 \left(\frac{c}{\nu_i} \right)^2 \left(\frac{\Delta\nu}{\nu_i} \right), \tag{7.7}$$

and the approximating Equation 7.6 for appropriate values of Φ.

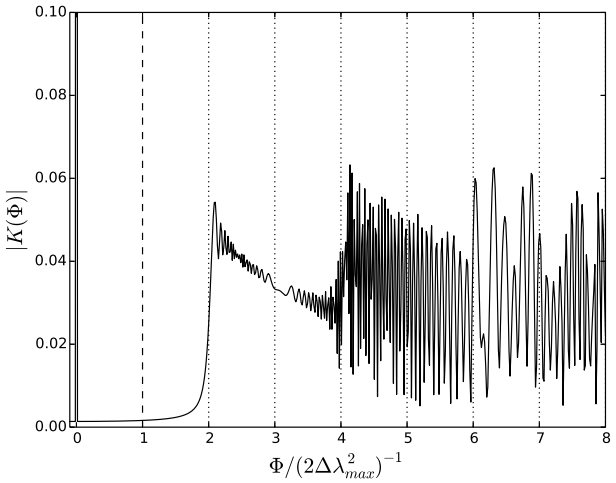

Figure 7.1: Magnitude of the kernel of the DRMT $K(\Phi)$, sampled to a maximum value $\Phi_{max} = (2\Delta\lambda^2_{min})^{-1}$. Vertical lines correspond to the harmonics of $\Phi_{min} = (2\Delta\lambda^2_{max})^{-1}$, with the dashed vertical line corresponding to Φ_{min} and the dotted lines corresponding to higher harmonics. This plot demonstrates the pitfalls of choosing $\Delta\Phi$ too large, since it shows the kernel structure induced from oversampling the lowest frequencies in a band.

Next, we choose spacings in $\Delta\Phi$ to fully sample the spectrum in rotation measure. Setting $\Delta\Phi < 1/2N\Delta\lambda^2_{min}$ prohibits complete reconstruction (i.e. aliasing) of the high-frequency channels, since this choice of $\Delta\Phi$ is less than the maximum Nyquist frequency of

the spectrum. On the other hand, setting $\Delta\Phi > 1/2N\Delta\lambda^2_{max}$ over-samples the Φ spectrum at the lowest frequencies, adding structure to the kernel of the DRMT at the harmonics of $1/2N\Delta\lambda^2_i$, demonstrated in Figure 7.1.

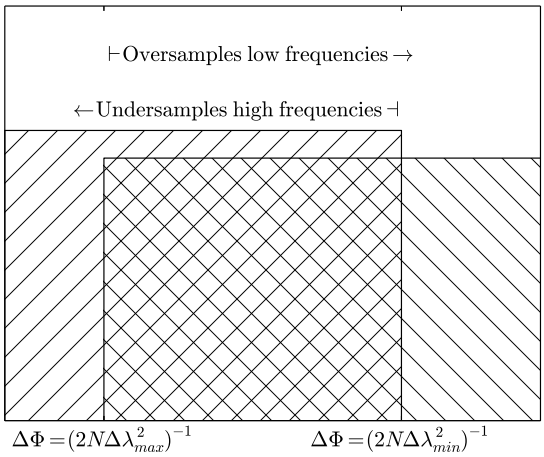

Figure 7.2: A cartoon to demonstrate the range of acceptable choices for $\Delta\Phi$. The region on the $\Delta\Phi$-axis hashed to the upper-right represents the range $\Delta\Phi < (2N\Delta\lambda^2_{min})^{-1}$ which under-samples the high frequencies. The region hashed to the lower-right depicts the region $\Delta\Phi > (2N\Delta\lambda^2_{max})^{-1}$, which over-samples the high frequencies.

Figure 7.2 gives a graphic depiction of the argment for choosing $\Delta\Phi$. We avoid the doubly-excluded region $(2N\Delta\lambda^2_{max})^{-1} < \Delta\Phi < (2N\Delta\lambda^2_{min})^{-1}$, and choose to undersample rather than to introduce unwanted structure. Finally, we choose a sampling in rotation measure

$$\Delta\Phi = \frac{1}{N\Delta\lambda^2_{max}}. \tag{7.8}$$

We use this choice of spacing in rotation measure to construct the rotation measure spectrum which we sample:

$$\Phi_j = j\Delta\Phi, \tag{7.9}$$

where the index $j \in [N/2, N/2)$ mirrors that of the DFT.

Finally, we set the phase-offset

$$\lambda_0^2 \equiv \frac{1}{N}\sum_i \lambda_i^2 \tag{7.10}$$

to reduce the phase error between the $i = 0$ and $i = N$ frequency bins.

Now, we can approximate the integral in Equation 7.3 as a matrix equation:

$$\begin{aligned}\widetilde{F}_j &\approx \frac{1}{\sum_i \Delta\lambda_i^2}\sum_i F_i e^{-2i\Phi_j(\lambda_i-\lambda_0^2)}\Delta\lambda_i^2 \\ &\equiv \frac{1}{N}\sum_i F_i e^{-2i\Phi_j(\lambda_i^2-\lambda_0^2)}\omega_i,\end{aligned} \tag{7.11}$$

where $\widetilde{F}_j$ is the DRMT of F at Φ_j, F_i are the spectrum values, and λ_i^2 are the bin-averaged values of the square wavelength (Equation 7.5). Equation 7.11 also defines the weights

$$\omega_i = \frac{N\Delta\lambda_i^2}{\sum_i \Delta\lambda_i^2} \tag{7.12}$$

which are the ratios of the bin-widths in λ^2 to the total bandwidth of the spectrum. We notat that this is a particular case of Equation 7.4 whose weights are chosen to most accurately approximate the measure of the rotation measure transform (Equation 7.3).

The matrix $\mathbf{W}$, defined as

$$W_{ij} = e^{-2i\Phi_j(\lambda_i^2-\lambda_0^2)}\omega_i \tag{7.13}$$

allows us to write Equation 7.11 in matrix form as $\widetilde{\mathbf{F}} = \mathbf{W}\cdot\mathbf{F}$. This highlights the computational benefit that we need only calculate the matrix $\mathbf{W}$ one for each class of spectra $\mathbf{F}$.

7.2.1 Example Spectra

As a proof of concept, we simulate a mock spectrum containing five "point sources," each with a randomly selected flux and rotation measure, shown in the top panel of Figure 7.3.

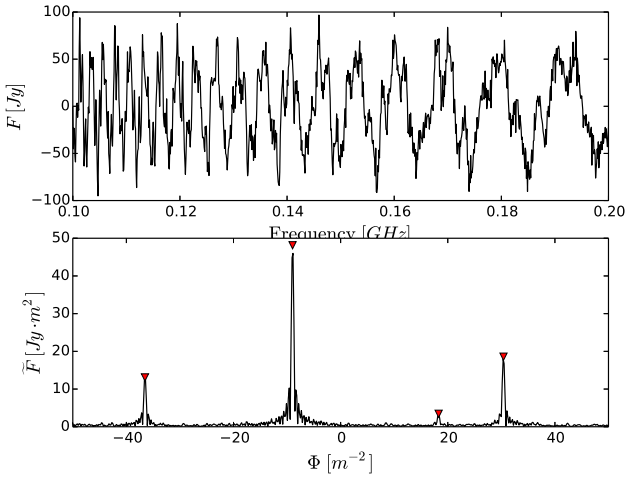

Figure 7.3: (Top Panel) Simulated spectrum with five "sources" with randomly chosen fluxes and rotation measures. Each channel is injected with 10Jy noise. (Bottom Panel) The magnitude of the DRMT of the spectrum in the top panel in black, with red triangles corresponding to the simulated Φ / flux pairs overlaid. The range of Φ values shown is restricted to $|\Phi| \leq 50$ m^{-2} to better show the kernel of the DRMT.

The simulated spectrum S_i is calculated as

$$S_i = \sum_{j=1}^{5} A_j e^{-2i\Phi_j \lambda_i^2} + n_i, \tag{7.14}$$

where A_j are the simulated fluxes, drawn from a chi-squared distribution with two degrees of freedom, Φ_j are the simulated rotation measures, drawn from a normal distribution whose mean is zero and width is 25 m^{-2}, n_i is the injected noise, drawn from a mean-zero normal distribution with a width of 10 Jy, and the spectrum is calculated on frequencies ν_i (and their corresponding square wavelengths λ_i^2), which are evenly spaced on 0.1 to 0.2 GHz, with 1000 channels.

Figure 7.3 displays excellent isolation of each Φ component and demonstrates the recovery of an input model. Normally, a physical source will not exhibit such a complex rotation measure structure, but this simulation is designed to highlight the strengths of the DRMT, rather than be physically representative.

7.2.2 Inverse Transform

In general, the transformation matrix $\mathbf{W}$ (Equation 7.13) is singular, prohibiting an exact expression for the inverse DRMT. Hence, we must supply an approximation for the inverse rotation measure transform,

$$F(\lambda^2) = \frac{1}{\pi} \int \widetilde{F}(\Phi) e^{+2i\Phi\lambda^2} \, \mathrm{d}\Phi \tag{7.15}$$

which should follow the same conventions as the rotation measure transform. The matrix

$$\widetilde{\mathbf{W}}_{ij} = \frac{1}{\pi} e^{+2i\Phi_i \lambda_j^2} \Delta\Phi = \frac{1}{\pi} \frac{\sum_k \Delta\lambda_k^2}{N \Delta\lambda_{max}^2} e^{+2i\Phi_i(\lambda_j^2 - \lambda_0^2)}, \tag{7.16}$$

defined such that $\mathbf{F} \approx \widetilde{\mathbf{W}} \cdot \widetilde{\mathbf{F}}$ is an approximation to Equation 7.15. A direct consequence of the singularity of $\mathbf{W}$ is that the product

$$\sum_j W_{ij} \widetilde{W}_{jk} \approx \frac{\omega_i}{\omega_{max}} \delta_{ik} \tag{7.17}$$

is not the identity matrix, but converges to the identity matrix in the limit where the spacing in λ^2 becomes even. In this special case, the DRMT reduces to the DFT, as expected. We also note that in this special case, the DRMT is unitary.

7.2.3 Noise Characteristics

Parseval's theorem states that power must be conserved between a function and its Fourier transform — the rotation measure transform is no different. Rewritten to adhere to the conventions of rotation measure synthesis, Parseval's theorem requires

$$\int |F(\lambda^2)|^2 \, \mathrm{d}\lambda^2 = \frac{1}{\pi} \int |\widetilde{F}(\Phi)|^2 \, \mathrm{d}\Phi, \tag{7.18}$$

where the factor of π^{-1} accommodates the abnormal Fourier convention of rotation measure synthesis. With the substitutions $\lambda^2 \rightarrow x$ and $\Phi \rightarrow \pi k$, Equation 7.18 reduces to the conventional notation for a Fourier transform defined as $\tilde{f}(k) = \int f(x) \exp\{-2\pi i k x\} \, \mathrm{d}x$.

Because the transfer matrix for the DRMT in Equation 7.13 is non-unitary ($\mathbf{W}^\dagger \neq \mathbf{W}^{-1}$), the DRMT will not satisfy Parseval's theorem exactly. However, just as we approximate $\mathbf{W}^{-1}$ in Equation 7.16, we can approximate the noise characteristics of the DRMT by enforcing Parseval's theorem in the limit where the λ_i^2 are uniformly spaced.

We begin our approximation of Parseval's theorem by discretizing the right-hand side of Equation 7.18, and expanding $\widetilde{F}$ in terms of the DRMT of the original spectrum:

$$\sum_i |\widetilde{F}_i|^2 = N \sum_i |F_i|^2 \, \omega_i^2. \tag{7.19}$$

In the limit $\omega_i \rightarrow 1$, this expression behaves as expected and gives the usual statement of Parseval's theorem for a DFT. Otherwise, we can approximate Equation 7.19 by providing an upper limit,

$$\sum_i |\widetilde{F}_i|^2 \leq N \omega_{max}^2 \sum_i |F_i|^2 \,, \tag{7.20}$$

where ω_{max} is the maximum fractional bin width for λ^2 spectrum, corresponding with the lowest frequency bin.

The relationship in Equation 7.20 allows us to estimate the noise properties of the DRMT in terms of the original spectrum. If a spectrum has uniform noise with variance $\sigma^2_{\lambda^2}$, then the DRMT of that wignal will also have uniform noise with variance $\sigma^2_\Phi \leq N\omega^2_{max}\sigma^2_{\lambda^2}$.

Though the noise level of the DRMT is uniform, the noise in each Φ sample will not necessarily be independent. Since the kernel of the DRMT,

$$K_j = \sum_i e^{-2i\Phi_j\lambda_i^2}\omega_i \tag{7.21}$$

only reduces to a delta function when the λ_i^2 are uniformly spaced, the noise will have some covariance given by the matrix $C_{ij} = K_i K_j^*$. As the spacings in λ^2 become uniform ($\omega_i \to 1$), the covariance will vanish in all but its off-diagonal terms ($C_{ij} \to \delta_{ij}$), since the DRMT reduces to a DFT.

7.3 Comparison with the Discrete Fourier Transform

Using the tools developed in the previous section, we investigate the validity of setting $w_i = 1$ in Equation 7.4. This is equivalent to examining the validity of the approximation that the spacings in λ^2 may be approximated as even.

There are two effects that may arise from setting $w_i = 1$. First, the samples in Φ may not be independent, that is, the uneven sampling of λ^2 widens the native sinc window of the DFT. Second, because the data are improperly weighted, the low-frequency samples, where the λ^2 sampling function is least dense is weighted unnaturally low. This causes the DFT to underestimate the noise in the RMT.

The effects of these errors may be parameterized by the measure of the DRMT, ω_i, defined in Equation 7.12. This quantity is maximized in the lowest frequency bins (Equation 7.7). We can use this value, ω_{max} as a measure of the uniformity of spacings in λ^2. This

value can be parameterized by the fractional bandwidth of the measurement (f_ν), defined for spectra ranging in frequency from $\nu_0(1 - f_\nu/2)$ to $\nu_0(1 + f_\nu/2)$. The maximum deviation from even spacings can be written in terms of f_ν as

$$N\omega_{max} = \frac{N\Delta\lambda^2_{max}}{\sum_i \Delta\lambda^2_i} = \frac{(1 - f_\nu/2)^2}{(1 - f_\nu/2)}, \tag{7.22}$$

where we employ Equation 7.7 to approximate the spacings $\Delta\lambda^2$ and noticed that $\sum_i \Delta\lambda^2_i = \lambda^2_{max} - \lambda^2_{min}$. As $f_\nu \to 0$, $N\omega_{max} \to 1$, and the DRMT will reduce to a DFT, validating the assumption made to set w_i in Equation 7.4 to 1. Allowing ten percent tolerance on $N\omega_{max}$ allows all bands with $f_\nu < 0.07$ to be well approximated by the DFT.

Figure 7.4 shows the kernel (Equation 7.21) for the transform over three bands with varying fractional bandwidths. Each spectrum is computed with $\nu_0 = 150$ MHz and $N = 1000$ channels, though these numbers do not affect the level of disagreement between the DRMT and the standard method. As $f_\nu \to 0$, the kernels of these two methods converge to a delta function, the kernel of a true DFT.

We will compare the use of ω_i set by Equation 7.12 to setting $w_i = 1$ by two metrics — the with of the kernel, and the noise equivalent bandwidth. We will compute these as functions of the fractional bandwidth f_ν, noting that for the limiting case, when $f_\nu = 0$, the DRMT reduces to the DFT. The first metric, the FWHM, demonstrates the transforms ability to isolate distinct rotation measure structures. In a true DFT, the FWHM of a transform is always one bin — this is the simple statement that the kernel of a DFT in frequency space is a delta function. Away from that limit, the lower the FWHM, the less covariance between neighboring Φ modes a transform allows. The second is a measure of the statistical uncertainty in the transformed measurement. For a true DFT, this is always one, but in the regime with unevenly spaced λ^2 samples, its meaning is more complicated. In general, higher values of the noise-equivalent bandwidth indicate lower levels of uncertainty in the measurement.

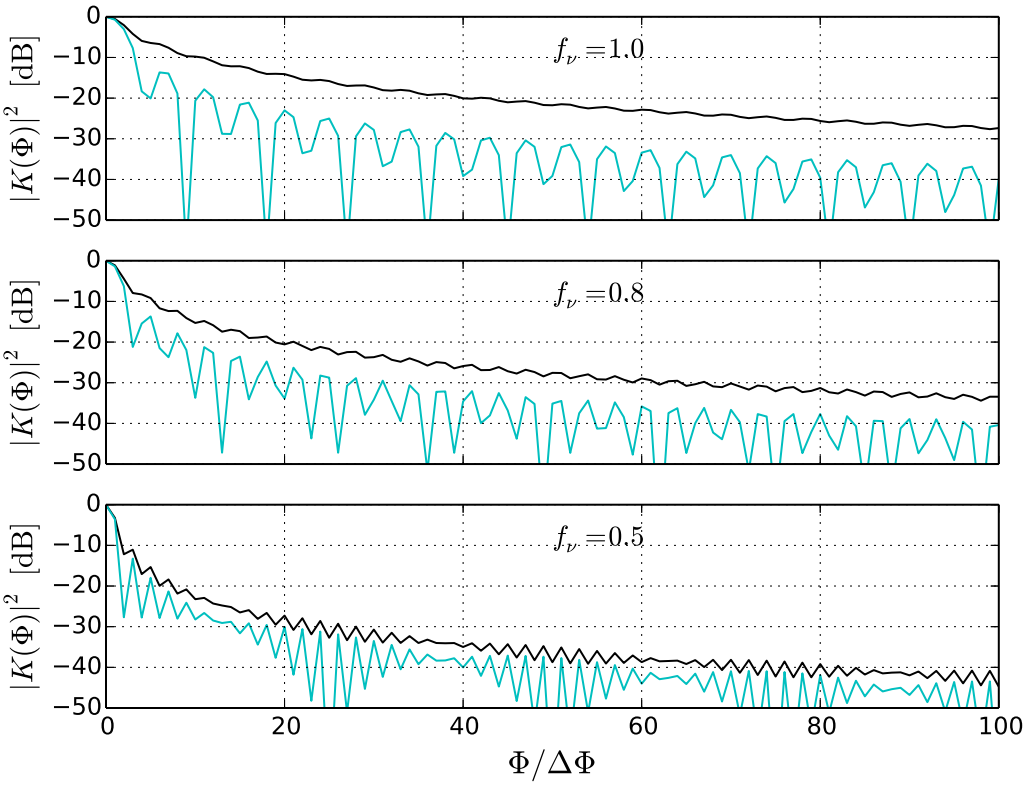

Figure 7.4: A comparison of the two methods of the RMT. Each plot shows the squared magnitude of the kernel of the DRMT (cyan) and the approximate DFT (black). The three panels show three representative bands with central frequency $\nu_0 = 150$MHz and different fractional bandwidths f_ν. (Top) $f_\nu = 1$ (Middle) $f_\nu = 0.8$ (Bottom) $f_\nu = 0.5$. Only 100 bins of the kernel are shown to better show structure near the central peak.

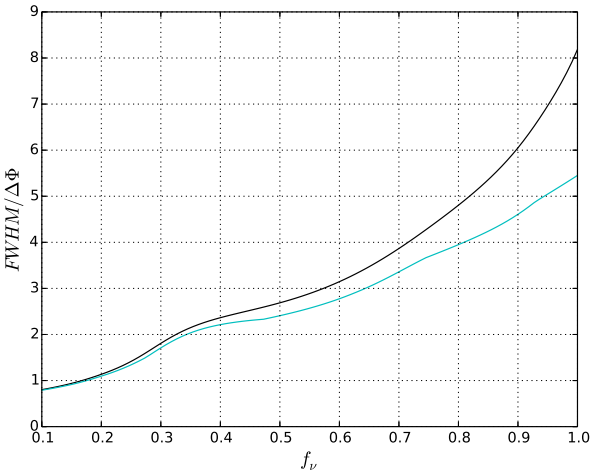

Figure 7.5: Full width at half maximum (FWHM) of the kernel of the DRMT (cyan) and the kernel of the approximate DFT (black). The FWHM is found by interpolating the value where kernel approaches 0.5 — this interpolation accounts for both the non-integer values, and the deviation from one as $f_\nu \to 0$.

To further examine the kernels of these two transforms, we plot the full width at half maximum (FWHM) of the kernels as functions of f_ν. We interpolate the kernel to find the precise value of the FWHM — this produces non-integer values. As $f_\nu \to 0$, we approach one, the value taken by a DFT. The slight offset from one is due to the interpolation. As $f_\nu \to 1$, the FWHM of the DRMT becomes 2/3 that of the approximate DFT.

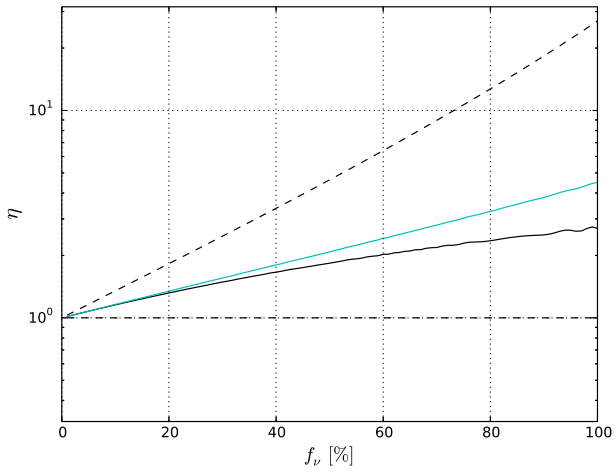

Figure 7.6: Noise equivalent bandwidth (Equation 7.23 of the DRMT is shown in cyan, of the approximate DFT in black. That the bandwidth exceeds one reflects the choice of $\Delta\Phi$ from the largest λ^2 bin. Dashed, black lines show the allowable range in η, for any choice of $\Delta\Phi$.

Another measure of the transform is the noise-equivalent bandwidth of the kernel, defined for a kernel K_i as

$$\eta = \frac{|\sum_i K_i|^2}{\sum_i |K_i|^2} \tag{7.23}$$

This quantity depends on the choice of Φ_j. As we discussed in Section 7.2, the available range in Φ is well determined in the limit of evenly spaced λ^2, but outside of that limit, the available range is not constrained. We choose to sample Φ based on the largest spacings in λ^2, but this neglects the higher Φ modes which are measured in the highest frequency

bins. This effect allows the noise-equivalent bandwidth η to exceed one, which typically is forbidden. Figure 7.6 shows η as a function of f_ν, alongside the allowed ranges of that quantity. The DRMT shows a factor of 1.7 increase in this quantity when compared to the approximate DFT.

The DRMT outperforms the approximate DFT in both metrics, indicating that the proper choice of a metric (w_i in Equation 7.4) can lead to more precise measurements of the rotation measure structures of polarized spectra.

Chapter 8

New Methods for Polarization Calibration

This chapter presents two novel approaches to polarization calibration, designed mainly for redundant, widefield arrays like PAPER. The two authoritative sources in interferometric polarization calibration are the series of papers by Hamacker, Bregman, and Sault (27, 28, 29, 30, 75) and Chapter three of Thompson et al. (83). They present a calibration scheme in which instrumental polarization is represented in terms of a rotation matrix $\mathbf{D}_i$, comprised of off-diagonal terms of the gain equation presented in Section 2.2, in Equation 2.15. In the Jones formalism, the measured electric field can be described by the equation

$$\begin{pmatrix} E_x \\ E_y \end{pmatrix} = \begin{pmatrix} g_{ix} & 0 \\ 0 & g_{iy} \end{pmatrix} \begin{pmatrix} 1 & d_{iy} \\ -d_{ix} & 1 \end{pmatrix} \begin{pmatrix} \cos\psi & -\sin\psi \\ \sin\psi & \cos\psi \end{pmatrix} \begin{pmatrix} E_\alpha \\ E_\delta \end{pmatrix} \equiv \mathbf{G}_i \cdot \mathbf{D}_i \cdot \mathbf{P} \cdot \vec{E}, \qquad (8.1)$$

where all terms agree with Equation 2.15. The new, off-diagonal terms d_{ix}, e.g. may be written in term of the three Euler angles of the feed of an antenna (83), but in general, instrumental polarization arises in the analog electronics, and cannot be represented so simply.

The cited sources give prescriptions for solving for the $4N_{ant}$ complex calibration terms as

functions of both time and frequency, but these methods require two criteria which PAPER does not meet:

1. The primary beam must be dominated by a single or few calibration sources. With PAPER's nearly 1 sr beam, there will almost certainly be a plethora of sources with calibration-level sensitivities.

2. The uv-coverage of the array must be filled to sufficiently isolate point sources, creating a clarity of image which PAPER in its redundant configuration cannot provide.

With these two criteria in mind, we must find non-traditional calibration methods. We present two: the first of which relies on the redundancy of identical baselines, and the second synthesizes images of only a few pointings on the sphere.

8.1 Polarization Calibration in Redundant Arrays

Since the field of view of PAPER cannot be dominated by a single source, we turn to broader, more statistical measures to calibrate against. The first is that a handful of calibrator sources, Pictor A, and Fornax A to name two, are only weakly polarized, allowing their constituent unpolarized signals to dominate the emission in $\mathcal{V}_{xx}$ and $\mathcal{V}_{yy}$. This allows us to calibrate the xx and yy polarizations with only an unpolarized model. Once the antenna-dependent gains are solved for, we apply those to $\mathcal{V}_{xy}$ and $\mathcal{V}_{yx}$.

This calibration scheme is incomplete, leaving an uncalibrated phase between the two cross-polarized visibilities. To solve for this phase difference, we make a final assumption, that $V = 0$, and solve for the phase which minimizes $\mathcal{V}_V \equiv \mathcal{V}_{xy} - \mathcal{V}_{yx}$. In other words, the assumption that $V = 0$ allows us to assume that $\mathcal{V}_{xy}$ and $\mathcal{V}_{yx}$ are redundant.

To understand the process of calibrating for the cross-polarization phase difference, we review redundant calibration in general, focusing on PAPER's implementation in particular. We extend this into polarization, explaining the impetus for solving for a single cross-polarization calibration term.

If the signal in two visibilities labeled 12 and 34 is redundant and calibration terms are antenna-based, then we can solve the equation

$$g_1 g_2^* \mathcal{V}_{12} = g_3 g_4^* \mathcal{V}_{34} \tag{8.2}$$

by minimizing the value

$$\chi^2 = \sum_{(i,j),(k,l)\in\mathcal{R}} \left| g_i g_j^* \mathcal{V}_{ij} - g_k g_l^* \mathcal{V}_{kl} \right|^2, \tag{8.3}$$

where $\mathcal{R}$ is the set of redundant baselines in an array. A full description of redundant calibration can be found in Liu et al. (52), and a clever implementation of it can be found in Zheng et al. (95).

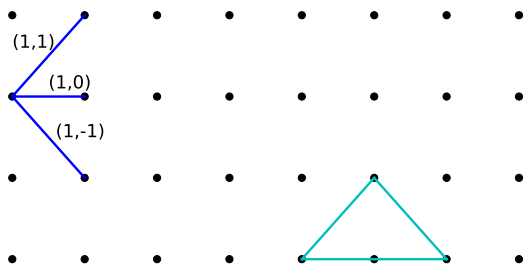

Figure 8.1

A full minimization of Equation 8.1 requires using all redundant baselines in an array, a computationally costly task. We restrict ourselves to three redundant types, a subset of $\mathcal{R}$, labelled by their grid-spacings as $(0,1)$,$(\pm 1,1)$ in Figure 8.1, which shows a cartoon map of the PAPER array.

We are restricted in our choice of a subspace of $\mathcal{R}$. Consider the subspace of all baselines with grid spacing $(0,1)$. Each row of the array may be calibrated relative to itself, but there are no terms in which link different rows. If we add all baselines with $(1,1)$ to this

subset, then we can link rows, but the solutions for the top-left and bottom-left antennae are under-constrained. Adding the $(-1, 1)$ baseline type fixes this problem and illuminates the two related criteria for the subspace of $\mathcal{R}$ necessary to fully solve for redundant calibration: the subset $\mathcal{R}'$ must extend through all antennae twice, and it also must allow for closure quantities to be calculated. As a reminder, the phase of all baselines in a closed loop of antennae must add to zero — this is called a closure quantity. A general rule is to draw closed loops with elements of $\mathcal{R}'$, shown in Figure 8.1 as a cyan triangle made of two $(0, 1)$ baselines, and one each of $(1, 1)$ and $(-1, 1)$.

Having chosen a subset of $\mathcal{R}$, we further simplify the procedure by linearizing Equation 8.1, computing the logarithm of the visibilities. Finally, we model the phase as a line (an electrical delay τ_i), representing Equation 8.1 as

$$\log\left(\frac{\mathcal{V}_{ij}}{\mathcal{V}_{kl}}\right) = \log g_i + \log g_j + \log g_k + \log g_l + 2\pi i\nu(\tau_i - \tau_j - \tau_k + \tau_l). \tag{8.4}$$

This equation sacrifices an unbiased, optimal solution for computational ease. The real part of the log of the ratio $\mathcal{V}_{ij}/\mathcal{V}_{kl}$ is simply a function of the antenna gains, and the imaginary part is simply a function of the electrical delays.

As a final simplifying step, we choose a fiducial baseline from each type to serve as the denominator of the ratio in Equation 8.4. This eases the computational burden from order N_{ant}^4, the number of baselines squared, to order N_{ant}^2. This step essentially reduces the size of the matrix representing the addition of gains (or delays) from $N_{ant} \times N_{ant}^4$ to $N_{ant} \times N_{ant}^2$ — the system of equations is still over-constrained, but now we have less work to do.

Thus we solve for calibration parameters which force the array to be redundant. There are two terms per polarization which cannot be calibrated in this way, though. The first is obvious: an array-wide flux scale, setting the calibration of the xx-polarization, say, to the sky. We solve for these by assuming that I dominates the signal in xx and yy, and setting the flux to a model of Pictor A (38). The second, less obvious terms remaining are the delays of the fiducial baselines we chose, which we decompose into a fiducial east-west

and a fiducial north-south baseline. These final terms are also found by fitting a model of Pictor A to all of the now-redundant visibilities.

Now, both the xx and the yy polarizations are calibrated. If we assume that the gains and delays are antenna-based, we can apply them to the xy and yx polarizations. This leaves one final term: a cross polarization delay. Up until this point in the redundant calibration process, each polarization of the array has been treated independently. To set the x and y delays to the same reference, we must add information to our calibration schema.

To solve for the cross-polarization delay, we assume that $V = 0$ at these frequencies, so we can treat the xy and yx visibilities as redundant. Then, as before, we solve for the delay which minimizes $\mathcal{V}_V$,

$$\mathfrak{Im}\left\{\log \mathcal{V}_{xy} - \log \mathcal{V}_{yx}\right\} = 2\pi i \tau_{xy} \nu. \tag{8.5}$$

Since all baselines are calibrated to be redundant at this point, we use all available data to solve for τ_{xy} and apply it to fully calibrate the redundant array.

It should be noted that this method is similar to that presented in Cotton (15), with two differences. First, Cotton (15) suggests maximizing the sum on the left-hand side of Equation 8.5, which pushes all available signal into $\mathcal{V}_U$, allowing some to remain in $\mathcal{V}_V$. In general, one cannot make the assumption that $V = 0$, but at the low frequencies measured by PAPER, no circularly polarized emission has been measured to date. Second, Cotton (15) uses this method on a per-baseline basis, not assuming a redundant array, which we clearly do. The use of multiple baselines to solve for a single calibration term increases the signal-to-noise of the measurements of that calibration term.

8.2 Beamforming

We now discuss a method for calibrating the off-diagonal gain terms. Typical measurements require a single calibration source to dominate emission in a field of view. With a wide field

imager, this requirement can never be met. We discuss a method to artificially restrict the field of view, only imaging a few points on the sky at once.

We begin by modelling a calibrator as a point source at position $\hat{s}$ with Stokes parameters I, Q, U, and V. We employ Equations 2.32, 2.33, 2.34, and 2.35 to represent a visibility containing only this source as a matrix equation,

$$\begin{pmatrix} \mathcal{V}_{xx} \\ \mathcal{V}_{xy} \\ \mathcal{V}_{yx} \\ \mathcal{V}_{yy} \end{pmatrix} = e^{-2\pi i(\vec{b}_{ij}\cdot\hat{s})}\times \begin{pmatrix} A_{xx} & & & \\ & A_{xy} & & \\ & & A_{yx} & \\ & & & A_{yy} \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & -i \\ 0 & 0 & 1 & i \\ 1 & -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\psi & -\sin 2\psi & 0 \\ 0 & \sin 2\psi & \cos 2\psi & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} I \\ Q \\ U \\ V \end{pmatrix}. \tag{8.6}$$

This equation has neglected gain terms not associated with the primary beam. In contrast with the discussion in Section 2.2, we allow the gain matrix $\overleftrightarrow{G}_{ij}$ to have sixteen, independent components, rather than forcing it to be diagonal. Furthermore, we write Equation 8.6 in terms of a vector the source's stokes parameters $\vec{M}$, a vector of visibilities $\vec{\mathcal{V}}_{ij,t}$, a diagonal matrix containing the elements of the beam, $\overleftrightarrow{A}_t$, and a transfer matrix $\overleftrightarrow{W}_t$, and the fringe, $\exp\{-2\pi i(\vec{b}\cdot\hat{s})\}$:

$$\vec{\mathcal{V}}_{ij,t} = e^{-2\pi i(\vec{b}_{ij}\cdot\hat{s})}\overleftrightarrow{G}_{ij}\cdot\overleftrightarrow{A}_t\cdot\overleftrightarrow{W}_t\cdot\vec{M}. \tag{8.7}$$

We explicitly label the time-dependent quantities with subscript t. Modelling multiple point sources is as simple as summing over different models M, with different transfer matrices and beams,

$$\vec{\mathcal{V}}_{ij,t} = \overleftrightarrow{G}_{ij}\cdot\sum_s e^{-2\pi i(\vec{b}_{ij}\cdot\hat{s})}\overleftrightarrow{A}_{t,s}\cdot\overleftrightarrow{W}_{t,s}\cdot\vec{M}_s. \tag{8.8}$$

Our task is to solve for the sixteen components of $\overleftrightarrow{G}_{ij}$. We do this by assuming that $\vec{\mathcal{V}}_{ij}$ is comprised of only our model, a sum of point sources, and thermal noise. This allows us

to solve for the components of $\vec{\vec{G}}_{ij}$ in the least squares sense, minimizing

$$\chi^2 = \sum_t \left| \vec{V}_{ij} - \vec{\vec{G}}_{ij} \sum_s e^{-2\pi i (\vec{b}_{ij} \cdot \hat{s})} \vec{\vec{A}}_{t,s} \cdot \vec{\vec{W}}_{t,s} \cdot \vec{M}_s \right|^2 . \tag{8.9}$$

For simplicity, we will drop the subscripts i and j, which until now have denoted the visibility's baseline label. We do this noting that this calibration solution is baseline-independent. In an additional measure of notational simplicity, we redefine the model visibility $\vec{M}'_t \equiv \sum_s \exp\{-2\pi i(\vec{b} \cdot \hat{s})\} \vec{\vec{A}}_{t,s} \cdot \vec{\vec{W}}_{t,s} \cdot \vec{M}_s$ as the time-dependent model visibility. These simplifications reduce the chi-squared expression to

$$\chi^2 = \sum_t \left| \vec{V}_t - \vec{\vec{G}} \cdot \vec{M}'_t \right|^2 . \tag{8.10}$$

The values of $\vec{\vec{G}}$ which minimize χ^2, found by setting $d\chi^2/dG^*_{\alpha\beta} = 0$ is

$$\vec{\vec{G}} = \left(\sum_t \vec{M}'_t \otimes \vec{M}'^{\dagger}_t \right)^{-1} \cdot \left(\sum_t \vec{V}_t \otimes \vec{M}'^{\dagger}_t \right) , \tag{8.11}$$

where $\otimes$ represents the Kronecker outer product, defined in Section 2.2.

In order for $\sum_t \vec{M}'_t \otimes \vec{M}'^{\dagger}_t$ to be non-singular, a full Stokes model of a polarized source must be included. Otherwise, one may approximate the inverse of that matrix by assuming that model Stokes parameters Q, U, and V are much smaller than I.

Including multiple sources in the model $\vec{M}'_t$ will increase the accuracy of the calibration. By including many sources, we model both the sources in question and sidelobes from nearby sources — this provides accurate source spectra for each pointing. With redundant arrays like PAPER, a single pointing on the sky may include emission from multiple sources. Modelling this effect will clearly increase the accuracy of both the calibration and the measurement of source spectra.

To date, no calibrators are sufficiently accurately measured for use in this method, so this method has yet to be implemented. Once a more accurate model of the polarized sky

in the southern hemisphere is made, future instruments will be able to use this method to calibrate off-diagonal polarization calibration terms without the computationally costly and oftentimes uncertain imaging deconvolution.

Chapter 9

Conclusion

We have demonstrated how crucial foreground characterization will be for uncovering signal from the 21cm EoR. For smooth-spectrum sources — the overwhelming majority of radio sources — power is isolated within a "wedge" in the $k_\perp$-$k_{||}$ plane, which allows us to undertake foreground avoidance. Faraday-rotated, polarized point sources disobey this rule, and generally will not be restricted to low values of $k_{||}$.

We synthesized state-of-the-art measurements of the polarized sky in radio frequencies to construct a simulation of polarized point sources. This allows us to simulate the level of contamination of polarized sources to the 21cm EoR power spectrum. This simulation predicted that the expected level of polarized foregrounds, leaked into the unpolarized power spectrum, will far exceed the levels of reasonable models for the 21cm EoR power spectrum. The simulations presented were based mostly on measurements within the PAPER band, but some extrapolations from higher frequencies were required. In particular, no comprehensive measurements of the polarized fraction of point sources had been measured at meter wavelengths, so we drew on measurements at 1.4 GHz.

To test the predictions made with the simulations, we constructed the most sensitive polarized power spectra at these frequencies ever made. These power spectra were made

from about six months of data from the PAPER array. Comparing the measured Q power spectrum to the simulated, we find that the simulations overestimate the levels of power. This yields a revised estimate of polarized leakage into the unpolarized, 21cm EoR power spectrum — the revision is lower, but still at the level of leading models.

Because of the disagreement between the measured power spectra and the simulated ones, we update the input parameters to the simulation. Since the distributions of source counts and rotation measures are relatively accurately measured, we focus on the distribution of polarized fractions. By a Bayesian analysis, we find that the data exclude our assumptions to high significance, and prefer a mean polarized fraction of 2.2×10^{-3}, a factor of ten lower than measured at 1.4 GHz. This new distribution of polarized fractions qualitatively agrees with the few recent measurements of the polarized sky at meter wavelengths.

These new measurements show the importance of characterizing polarized foregrounds. Future observations will require mitigation strategies for instrumental, polarized leakage in order to detect and characterize the 21cm EoR power spectrum. Mitigation strategies will incorporate both source detection and the updated design of instruments. Measuring polarized sources at these frequencies will require new observations from arrays more suited for imaging, as well as updated techniques for polarimetry. Instruments will also need to limit beam leakage by uniformly illuminating their dishes, creating a primary beam which is symmetric about 90° rotations. These tactics will limit polarized leakage and more smoothly pave the way to measuring the ionization history of the IGM through the 21cm EoR power spectrum.

Glossary of Symbols and Abbreviations

α	Right ascension, page 28
δ	Declination, page 28
$\Delta\nu$	Bandwidth, page 42
$\Delta^2(k)$	Spherically-averaged power spectrum, page 18
η	Fourier-dual variable to frequency, page 42
ν	Frequency, page 25
Ω	Angular extent of the primary beam, page 42
$\omega_\oplus$	Angular frequency of the Earth's rotation, page 38
Ω_m	Cosmic matter density in units of the critical density $\rho_c = 3H^2/8\pi G$, page 19
Φ	Rotation measure, page 57
ψ	Parallactic angle, page 30
τ	Delay, page 31
τ_g	Geometric delay, page 36
$\xi(\mathbf{r})$	Correlation function of spin temperature field, page 17
$\text{Ш}(u,v)$	Sampling function in the uv-plane, page 27
$A(l,m)$	Primary beam, page 27
A_{eff}	Effective area of an antenna, page 48
$\vec{b}$	Baseline (in meters), page 25
$\mathcal{R}$	Set of redundant baselines, page 141
$\mathcal{D}$	Set of all bandpowers in a measured power spectrum, page 115
f_ν	Fractional bandwidth, page 134
g	Antenna gain, page 31
H	Hour angle, page 29
$H(z)$	Hubble parameter, page 13
p	Polarized fraction, page 56
$P(\mathbf{k})$	Power Spectrum of Spin Temperature, page 17
P_k	Power spectrum in the k^{th} bin, page 115
$\overline{S}$	Stokes rotation matrix, page 33
$\hat{s}$	Unit pointing vector, page 25
$\mathcal{S}$	Set of all bandpowers in a simulated power spectrum, page 115
S_k	Simulated power spectrum in the k^{th} bin, page 115

T_s	Spin Temperature, page 11
T_{sys}	System Temperature, page 103
$\mathcal{V}$	Visibility, calibration terms included, page 32
$\widehat{V}(\nu, f)$	Fringe-rate transformed visibility, page 38
$\widetilde{V}(\tau)$	Delay-transformed visibility, page 36
$V(u, v, \nu)$	Visibility, page 27
$W(\nu)$	Window function in frequency, page 59
x	Scale factor between the simulated and measured power spectra, page 115
$\otimes$	Kronecker outer product, page 31

Abbreviations

CMB	Cosmic Microwave Background Radiation, page 2
DFT	Discrete Fourier transform, page 125
DRAO	Dominion Radio Astrophysical Observatory of Canada, page 61
DRMT	Discrete Rotation Measure Transform, page 126
EoR	Epoch of Reionization, page 3
GMRT	Giant Metrewave Radio Telescope, page 63
IGM	Intergalactic Medium, page 3
LST	Local sidereal time, page 28
Lyα	Lyman alpha, page 3
NVSS	NRAO VLA Sky Survey, page 61
PAPER	The Donald C. Backer Precision Array to Probe the Epoch of Reionization, page 45
RFI	Radio Frequency Interference, page 47
RMT	Rotation measure transform, page 125
6C	Sixth Cambridge Survey, page 77
TEC	Total Electron Content of the ionosphere, page 112
VLA	Very Large Array, page 61
VLSS	VLA Large Sky Survey, page 77
WMAP	Wilkinson Microwave Anisotropy Probe, page 8
WSRT	Westerbork Synthesis Radio Telescope, page 63

References

[1] Beck, R., Anderson, J., Heald, G., Horneffer, A., Iacobelli, M., Köhler, J., Mulcahy, D., Pizzo, R., et al. The LOFAR view of cosmic magnetism. Astronomische Nachrichten, **334**:548–557 (2013). doi:10.1002/asna.201311894. 124

[2] Beck, R., Frick, P., Stepanov, R., and Sokoloff, D. Recognizing magnetic structures by present and future radio telescopes with Faraday rotation measure synthesis. A&A, **543**:A113 (2012). doi:10.1051/0004-6361/201219094. 125

[3] Beck, R. and Wielebinski, R. Magnetic Fields in Galaxies, page 641 (2013). doi:10.1007/978-94-007-5612-0_13. 58

[4] Becker, R. H., Fan, X., White, R. L., Strauss, M. A., Narayanan, V. K., Lupton, R. H., Gunn, J. E., Annis, J., et al. Evidence for Reionization at z˜6: Detection of a Gunn-Peterson Trough in a z=6.28 Quasar. AJ, **122**:2850–2857 (2001). doi:10.1086/324231. xiii, 5, 6

[5] Bernardi, G., Carretti, E., Fabbri, R., Sbarra, C., Poppi, S., Cortiglioni, S., and Jonas, J. L. A polarized synchrotron template for cosmic microwave background polarization experiments based on WMAP data. MNRAS, **351**:436–446 (2004). doi:10.1111/j.1365-2966.2004.07797.x. 61

[6] Bernardi, G., de Bruyn, A. G., Brentjens, M. A., Ciardi, B., Harker, G., Jelić, V., Koopmans, L. V. E., Labropoulos, P., et al. 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 (2009). doi:10.1051/0004-6361/200911627. 68, 87, 88, 89, 122

[7] Bernardi, G., de Bruyn, A. G., Harker, G., Brentjens, M. A., Ciardi, B., Jelić, V., Koopmans, L. V. E., Labropoulos, P., et al. 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 (2010). doi:10.1051/0004-6361/200913420. 63, 64

[8] Bernardi, G., Greenhill, L. J., Mitchell, D. A., Ord, S. M., Hazelton, B. J., Gaensler, B. M., de Oliveira-Costa, A., Morales, M. F., et al. A 189 MHz, 2400 deg^2 Polarization Survey with the Murchison Widefield Array 32-element Prototype. ApJ, **771**:105 (2013). doi:10.1088/0004-637X/771/2/105. 63, 108, 118, 120, 122

[9] Bouwens, R. J., Illingworth, G. D., Oesch, P. A., Labbé, I., Trenti, M., van Dokkum, P., Franx, M., Stiavelli, M., et al. Ultraviolet Luminosity Functions from 132 z ˜ 7 and z ˜ 8 Lyman-break Galaxies in the Ultra-deep HUDF09 and Wide-area Early Release Science WFC3/IR Observations. ApJ, **737**:90 (2011). doi:10.1088/0004-637X/737/2/90. xiii, 5, 7

[10] Bowman, J. D., Morales, M. F., and Hewitt, J. N. Foreground Contamination in Interferometric Measurements of the Redshifted 21 cm Power Spectrum. ApJ, **695**:183–199 (2009). doi:10.1088/0004-637X/695/1/183. 20, 41

[11] Brentjens, M. A. and de Bruyn, A. G. Faraday rotation measure synthesis. A&A, **441**:1217–1228 (2005). doi:10.1051/0004-6361:20052990. 58, 92, 125

[12] Cohen, A. S., Lane, W. M., Cotton, W. D., Kassim, N. E., Lazio, T. J. W., Perley, R. A., Condon, J. J., and Erickson, W. C. The VLA Low-Frequency Sky Survey. AJ, **134**:1245–1262 (2007). doi:10.1086/520719. 77

[13] Cohen, A. S., Röttgering, H. J. A., Jarvis, M. J., Kassim, N. E., and Lazio, T. J. W. A Deep, High-Resolution Survey at 74 MHz. ApJ Supp., **150**:417–430 (2004). doi:10.1086/380783. 78

[14] Condon, J. J., Cotton, W. D., Greisen, E. W., Yin, Q. F., Perley, R. A., Taylor, G. B., and Broderick, J. J. The NRAO VLA Sky Survey. AJ, **115**:1693–1716 (1998). doi:10.1086/300337. 61, 65, 77

[15] Cotton, W. D. a New Method for Cross Polarized Delay Calibration of Radio Interferometers. Journal of Astronomical Instrumentation, **1**:1250001 (2012). doi:10.1142/S2251171712500018. 143

[16] Datta, A., Bowman, J. D., and Carilli, C. L. Bright Source Subtraction Requirements for Redshifted 21 cm Measurements. ApJ, **724**:526–538 (2010). doi:10.1088/0004-637X/724/1/526. 20

[17] Datta, A., Bradley, R., Burns, J. O., Harker, G., Komjathy, A., and Lazio, T. J. W. Effects Of The Ionosphere On Ground-Based Detection Of The Global 21 CM Signal From The Cosmic Dawn And The Dark Ages. ArXiv e-prints (2014). 112

[18] de Oliveira-Costa, A., Tegmark, M., Gaensler, B. M., Jonas, J., Landecker, T. L., and Reich, P. A model of diffuse Galactic radio emission from 10 MHz to 100 GHz. MNRAS, **388**:247–260 (2008). doi:10.1111/j.1365-2966.2008.13376.x. 61

[19] Dillon, J. S., Liu, A., Williams, C. L., Hewitt, J. N., Tegmark, M., Morgan, E. H., Levine, A. M., Morales, M. F., et al. 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 (2014). doi:10.1103/PhysRevD.89.023002. 20

[20] Fan, X., Strauss, M. A., Becker, R. H., White, R. L., Gunn, J. E., Knapp, G. R., Richards, G. T., Schneider, D. P., et al. Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with z˜6 Quasars. II. A Sample of 19 Quasars. AJ, **132**:117–136 (2006). doi:10.1086/504836. 5

[21] Field, G. B. Excitation of the Hydrogen 21-CM Line. Proceedings of the IRE, **46**:240–250 (1958). doi:10.1109/JRPROC.1958.286741. 14

[22] Furlanetto, S. R., Oh, S. P., and Briggs, F. H. Cosmology at low frequencies: The 21 cm transition and the high-redshift Universe. PhysRep, **433**:181–301 (2006). doi:10.1016/j.physrep.2006.08.002. 11, 14, 19, 41

[23] Gaensler, B. M., Dickey, J. M., McClure-Griffiths, N. M., Green, A. J., Wieringa, M. H., and Haynes, R. F. Radio Polarization from the Inner Galaxy at Arcminute Resolution. ApJ, **549**:959–978 (2001). doi:10.1086/319468. 61

[24] Geil, P. M., Gaensler, B. M., and Wyithe, J. S. B. Polarized foreground removal at low radio frequencies using rotation measure synthesis: uncovering the signature of hydrogen reionization. MNRAS, **418**:516–535 (2011). doi:10.1111/j.1365-2966.2011.19509.x. 122

[25] Gunn, J. E. and Peterson, B. A. On the Density of Neutral Hydrogen in Intergalactic Space. ApJ, **142**:1633–1641 (1965). doi:10.1086/148444. 4, 5

[26] Hales, S. E. G., Baldwin, J. E., and Warner, P. J. The 6C survey of radio sources. II - The zone delta = 30-51 deg, alpha = 08h30m-17h30m. MNRAS, **234**:919–936 (1988). 77, 117, 121

[27] Hamaker, J. P. Understanding radio polarimetry. IV. The full-coherency analogue of scalar self-calibration: Self-alignment, dynamic range and polarimetric fidelity. A&A Supp., **143**:515–534 (2000). doi:10.1051/aas:2000337. 66, 139

[28] Hamaker, J. P. Understanding radio polarimetry. V. Making matrix self-calibration work: processing of a simulated observation. A&A, **456**:395–404 (2006). doi:10.1051/0004-6361:20065145. 66, 139

[29] Hamaker, J. P. and Bregman, J. D. Understanding radio polarimetry. III. Interpreting the IAU/IEEE definitions of the Stokes parameters. A&A Supp., **117**:161–165 (1996). 66, 139

[30] Hamaker, J. P., Bregman, J. D., and Sault, R. J. Understanding radio polarimetry. I. Mathematical foundations. A&A Supp., **117**:137–147 (1996). 66, 139

[31] Harker, G. J. A., Pritchard, J. R., Burns, J. O., and Bowman, J. D. An MCMC approach to extracting the global 21-cm signal during the cosmic dawn from sky-averaged radio observations. MNRAS, **419**:1070–1084 (2012). doi:10.1111/j.1365-2966.2011.19766.x. xiii, 15, 16

[32] Harris, F. J. On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform. IEEE Proceedings, **66**:51–83 (1978). 104

[33] Haslam, C. G. T., Salter, C. J., Stoffel, H., and Wilson, W. E. A 408 MHz all-sky continuum survey. II - The atlas of contour maps. A&A Supp., **47**:1 (1982). 61, 63, 64

[34] Helmboldt, J. F., Kassim, N. E., Cohen, A. S., Lane, W. M., and Lazio, T. J. Radio Frequency Spectra of 388 Bright 74 MHz Sources. ApJ Supp., **174**:313–336 (2008). doi:10.1086/521829. 79

[35] Högbom, J. A. Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines. A&A Supp., **15**:417 (1974). 101

[36] Jacobs, D. C. The Epoch of Reionization: Foregrounds and calibration with PAPER. Ph.D. thesis, University of Pennsylvania (2011). 101

[37] Jacobs, D. C., Aguirre, J. E., Parsons, A. R., Pober, J. C., Bradley, R. F., Carilli, C. L., Gugliucci, N. E., Manley, J. R., et al. New 145 MHz Source Measurements by PAPER in the Southern Sky. ApJ Letters, **734**:L34 (2011). doi:10.1088/2041-8205/734/2/L34. 51, 52

[38] Jacobs, D. C., Parsons, A. R., Aguirre, J. E., Ali, Z., Bowman, J., Bradley, R. F., Carilli, C. L., DeBoer, D. R., et al. A Flux Scale for Southern Hemisphere 21 cm Epoch of Reionization Experiments. ApJ, **776**:108 (2013). doi:10.1088/0004-637X/776/2/108. 52, 53, 101, 142

[39] Jacobs, D. C., Pober, J. C., Parsons, A. R., Aguirre, J. E., Ali, Z., Bowman, J., Bradley, R. F., Carilli, C. L., et al. Multi-redshift limits on the 21cm power spectrum from PAPER. ArXiv e-prints (2014). 39, 52, 53, 93, 95, 104

[40] Jelić, V., de Bruyn, A. G., Mevius, M., Abdalla, F. B., Asad, K. M. B., Bernardi, G., Brentjens, M. A., Bus, S., et al. Initial LOFAR observations of epoch of reionization windows. II. Diffuse polarized emission in the ELAIS-N1 field. A&A, **568**:A101 (2014). doi:10.1051/0004-6361/201423998. 60, 63, 118, 122

[41] Jelić, V., Zaroubi, S., Labropoulos, P., Bernardi, G., de Bruyn, A. G., and Koopmans, L. V. E. Realistic simulations of the Galactic polarized foreground: consequences for 21-cm reionization detection experiments. MNRAS, **409**:1647–1659 (2010). doi:10.1111/j.1365-2966.2010.17407.x. 63, 77, 119

[42] Kogut, A., Spergel, D. N., Barnes, C., Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Limon, M., et al. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Temperature-Polarization Correlation. ApJ Supp., **148**:161–173 (2003). doi:10.1086/377219. xiii, 8

[43] Komatsu, E., Dunkley, J., Nolta, M. R., Bennett, C. L., Gold, B., Hinshaw, G., Jarosik, N., Larson, D., et al. Five-Year Wilkinson Microwave Anisotropy Probe Observations: Cosmological Interpretation. ApJ Supp., **180**:330–376 (2009). doi:10.1088/0067-0049/180/2/330. 7, 19

[44] Kovac, J. M., Leitch, E. M., Pryke, C., Carlstrom, J. E., Halverson, N. W., and Holzapfel, W. L. Detection of polarization in the cosmic microwave background using DASI. Nature, **420**:772–787 (2002). doi:10.1038/nature01269. 41

[45] Kronberg, P. P. and Newton-McGee, K. J. Remarkable Symmetries in the Milky Way Disc's Magnetic Field. PASA, **28**:171–176 (2011). doi:10.1071/AS10045. 65, 89

[46] Landecker, T. L., Dewdney, P. E., Burgess, T. A., Gray, A. D., Higgs, L. A., Hoffmann, A. P., Hovey, G. J., Karpa, D. R., et al. The synthesis telescope at the Dominion Radio Astrophysical Observatory. A&A Supp., **145**:509–524 (2000). doi:10.1051/aas:2000257. 61

[47] Law, C. J., Gaensler, B. M., Bower, G. C., Backer, D. C., Bauermeister, A., Croft, S., Forster, R., Gutierrez-Kraybill, C., et al. Spectropolarimetry with the Allen Telescope Array: Faraday Rotation Toward Bright Polarized Radio Galaxies. ApJ, **728**:57 (2011). doi:10.1088/0004-637X/728/1/57. 80

[48] Lidz, A., Zahn, O., McQuinn, M., Zaldarriaga, M., and Hernquist, L. Detecting the Rise and Fall of 21 cm Fluctuations with the Murchison Widefield Array. ApJ, **680**:962–974 (2008). doi:10.1086/587618. 22, 23, 41, 85, 86, 122

[49] Liu, A., Parsons, A. R., and Trott, C. M. Epoch of reionization window. I. Mathematical formalism. Phys. Rev. D, **90**(2):023018 (2014). doi:10.1103/PhysRevD.90.023018. 20

[50] Liu, A., Pritchard, J. R., Tegmark, M., and Loeb, A. Global 21 cm signal experiments: A designer's guide. Phys. Rev. D, **87**(4):043002 (2013). doi:10.1103/PhysRevD.87.043002. 16

[51] Liu, A., Tegmark, M., Bowman, J., Hewitt, J., and Zaldarriaga, M. An improved method for 21-cm foreground removal. MNRAS, **398**:401–406 (2009). doi:10.1111/j.1365-2966.2009.15156.x. 20

[52] Liu, A., Tegmark, M., Morrison, S., Lutomirski, A., and Zaldarriaga, M. Precision calibration of radio interferometers using redundant baselines. MNRAS, **408**:1029–1050 (2010). doi:10.1111/j.1365-2966.2010.17174.x. 141

[53] McQuinn, M., Zahn, O., Zaldarriaga, M., Hernquist, L., and Furlanetto, S. R. Cosmological Parameter Estimation Using 21 cm Radiation from the Epoch of Reionization. ApJ, **653**:815–834 (2006). doi:10.1086/505167. 14

[54] Mirocha, J., Harker, G. J. A., and Burns, J. O. Interpreting the Global 21 cm Signal from High Redshifts. I. Model-independent constraints. ApJ, **777**:118 (2013). doi:10.1088/0004-637X/777/2/118. 16

[55] Moore, D. F., Aguirre, J. E., Parsons, A. R., Jacobs, D. C., and Pober, J. C. The Effects of Polarized Foregrounds on 21 cm Epoch of Reionization Power Spectrum Measurements. ApJ, **769**:154 (2013). doi:10.1088/0004-637X/769/2/154. 53, 73

[56] Moore, D. F. et al. New Measurements of Polarized Power Spectra at 126 and 164 MHz. in prep. (2014). 39, 52, 53

[57] Morales, M. F., Bowman, J. D., and Hewitt, J. N. Improving Foreground Subtraction in Statistical Observations of 21 cm Emission from the Epoch of Reionization. ApJ, **648**:767–773 (2006). doi:10.1086/506135. 122

[58] Morales, M. F., Hazelton, B., Sullivan, I., and Beardsley, A. Four Fundamental Foreground Power Spectrum Shapes for 21cm Cosmology Observations. ApJ, **752**:137 (2012). doi:10.1088/0004-637X/752/2/137. 20

[59] Oppermann, N., Junklewitz, H., Greiner, M., Ensslin, T. A., Akahori, T., Carretti, E., Gaensler, B. M., Goobar, A., et al. Estimating extragalactic Faraday rotation. ArXiv e-prints (2014). 65

[60] Oppermann, N., Junklewitz, H., Robbers, G., Bell, M. R., Ensslin, T. A., Bonafede, A., Braun, R., Brown, J. C., et al. An improved map of the Galactic Faraday sky. A&A, **542**:A93 (2012). doi:10.1051/0004-6361/201118526. xiv, 59, 65, 80, 81, 82, 89, 117, 119

[61] Paciga, G., Albert, J. G., Bandura, K., Chang, T.-C., Gupta, Y., Hirata, C., Odegova, J., Pen, U.-L., et al. A simulation-calibrated limit on the H I power spectrum from the GMRT Epoch of Reionization experiment. MNRAS, **433**:639–647 (2013). doi:10.1093/mnras/stt753. 20, 22, 23

[62] Parsons, A., Pober, J., McQuinn, M., Jacobs, D., and Aguirre, J. A Sensitivity and Array-configuration Study for Measuring the Power Spectrum of 21 cm Emission from Reionization. ApJ, **753**:81 (2012). doi:10.1088/0004-637X/753/1/81. 43, 53, 70, 81, 84, 109

[63] Parsons, A. R. and Backer, D. C. Calibration of Low-Frequency, Wide-Field Radio Interferometers Using Delay/Delay-Rate Filtering. AJ, **138**:219–226 (2009). doi:10.1088/0004-6256/138/1/219. 35

[64] Parsons, A. R., Backer, D. C., Foster, G. S., Wright, M. C. H., Bradley, R. F., Gugliucci, N. E., Parashare, C. R., Benoit, E. E., et al. The Precision Array for Probing the Epoch of Re-ionization: Eight Station Results. **139**:1468–1480 (2010). 45

[65] Parsons, A. R., Liu, A., Aguirre, J. E., Ali, Z. S., Bradley, R. F., Carilli, C. L., DeBoer, D. R., Dexter, M. R., et al. New Limits on 21cm Epoch of Reionization from PAPER-32 Consistent with an X-Ray Heated Intergalactic Medium at z = 7.7. ApJ, **788**:106 (2014). doi:10.1088/0004-637X/788/2/106. xiii, xiv, 22, 23, 35, 39, 52, 53, 81, 93, 95, 97, 99, 101, 104, 107

[66] Parsons, A. R., Pober, J. C., Aguirre, J. E., Carilli, C. L., Jacobs, D. C., and Moore, D. F. A Per-baseline, Delay-spectrum Technique for Accessing the 21 cm Cosmic Reionization Signature. ApJ, **756**:165 (2012). doi:10.1088/0004-637X/756/2/165. xiii, 20, 37, 39

[67] Pen, U. L., Chang, T. C., Hirata, C. M., Peterson, J. B., Roy, J., Gupta, Y., Odegova, J., and Sigurdson, K. The GMRT EoR experiment: limits on polarized sky brightness at 150 MHz. MNRAS, **399**:181–194 (2009). doi:10.1111/j.1365-2966.2009.14980.x. 63, 64, 73, 80, 122

[68] Pober, J. C., Liu, A., Dillon, J. S., Aguirre, J. E., Bowman, J. D., Bradley, R. F., Carilli, C. L., DeBoer, D. R., et al. What Next-generation 21 cm Power Spectrum Measurements can Teach us About the Epoch of Reionization. ApJ, **782**:66 (2014). doi:10.1088/0004-637X/782/2/66. 22, 109

[69] Pober, J. C., Parsons, A. R., Aguirre, J. E., Ali, Z., Bradley, R. F., Carilli, C. L., DeBoer, D., Dexter, M., et al. Opening the 21 cm Epoch of Reionization Window: Measurements of Foreground Isolation with PAPER. ApJ Letters, **768**:L36 (2013). doi:10.1088/2041-8205/768/2/L36. xiii, 21, 52, 53, 84

[70] Pober, J. C., Parsons, A. R., Jacobs, D. C., Aguirre, J. E., Bradley, R. F., Carilli, C. L., Gugliucci, N. E., Moore, D. F., et al. A Technique for Primary Beam Calibration of Drift-scanning, Wide-field Antenna Elements. AJ, **143**:53 (2012). doi:10.1088/0004-6256/143/2/53. 68, 69

[71] Pritchard, J. R. and Loeb, A. 21 cm cosmology in the 21st century. Reports on Progress in Physics, **75**(8):086901 (2012). doi:10.1088/0034-4885/75/8/086901. 14

[72] Pshirkov, M. S., Tinyakov, P. G., Kronberg, P. P., and Newton-McGee, K. J. Deriving the Global Structure of the Galactic Magnetic Field from Faraday Rotation Measures of Extragalactic Sources. ApJ, **738**:192 (2011). doi:10.1088/0004-637X/738/2/192. 65

[73] Röttgering, H. LOFAR, a new low frequency radio telescope. NAR, **47**:405–409 (2003). doi:10.1016/S1387-6473(03)00057-5. 125

[74] Rybicki, G. B. and Lightman, A. P. Radiative processes in astrophysics (1979). 55

[75] Sault, R. J., Hamaker, J. P., and Bregman, J. D. Understanding radio polarimetry. II. Instrumental calibration of an interferometer array. A&A Supp., **117**:149–159 (1996). 66, 139

[76] Sault, R. J., Teuben, P. J., and Wright, M. C. H. A Retrospective View of MIRIAD. In R. A. Shaw, H. E. Payne, and J. J. E. Hayes, editors, Astronomical Data Analysis Software and Systems IV, volume 77 of Astronomical Society of the Pacific Conference Series, page 433 (1995). 51

[77] Spergel, D. N., Verde, L., Peiris, H. V., Komatsu, E., Nolta, M. R., Bennett, C. L., Halpern, M., Hinshaw, G., et al. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters. ApJ Supp., **148**:175–194 (2003). doi:10.1086/377226. 7

[78] Stefan, I. I., Carilli, C. L., Green, D. A., Ali, Z., Aguirre, J. E., Bradley, R. F., DeBoer, D., Dexter, M., et al. Imaging on PAPER: Centaurus A at 148 MHz. MNRAS, **432**:1285–1293 (2013). doi:10.1093/mnras/stt548. 52, 53

[79] Sunyaev, R. A. and Zeldovich, I. B. Microwave background radiation as a probe of the contemporary structure and history of the universe. ARAA, **18**:537–560 (1980). doi:10.1146/annurev.aa.18.090180.002541. 10

[80] Taylor, A. R., Stil, J. M., and Sunstrum, C. A Rotation Measure Image of the Sky. ApJ, **702**:1230–1236 (2009). doi:10.1088/0004-637X/702/2/1230. 65, 125

[81] Taylor, G. B., Carilli, C. L., and Perley, R. A., editors. Synthesis Imaging in Radio Astronomy II, volume 180 of Astronomical Society of the Pacific Conference Series (1999). 25

[82] Taylor, J. and Lidz, A. What do observations of the Lyman α fraction tell us about reionization? MNRAS, **437**:2542–2553 (2014). doi:10.1093/mnras/stt2067. 5

[83] Thompson, A. R., Moran, J. M., and Swenson, G. W., Jr. Interferometry and Synthesis in Radio Astronomy, 2nd Edition (2001). 25, 38, 139

[84] Tingay, S. J., Goeke, R., Bowman, J. D., Emrich, D., Ord, S. M., Mitchell, D. A., Morales, M. F., Booler, T., et al. The Murchison Widefield Array: the Square Kilometre Array Precursor at low radio frequencies. ArXiv e-prints (2012). 125

[85] Tucci, M. and Toffolatti, L. The Impact of Polarized Extragalactic Radio Sources on the Detection of CMB Anisotropies in Polarization. Advances in Astronomy, **2012**:624987 (2012). doi:10.1155/2012/624987. 68, 79, 121

[86] Westfold, K. C. The Polarization of Synchrotron Radiation. ApJ, **130**:241 (1959). doi:10.1086/146713. 55, 56

[87] White, M., Carlstrom, J. E., Dragovan, M., and Holzapfel, W. L. Interferometric Observation of Cosmic Microwave Background Anisotropies. ApJ, **514**:12–24 (1999). doi:10.1086/306911. 41

[88] Williams, C. L., Hewitt, J. N., Levine, A. M., de Oliveira-Costa, A., Bowman, J. D., Briggs, F. H., Gaensler, B. M., Hernquist, L. L., et al. Low-frequency Imaging of Fields at High Galactic Latitude with the Murchison Widefield Array 32 Element Prototype. ApJ, **755**:47 (2012). doi:10.1088/0004-637X/755/1/47. 79

[89] Wilson, T. L., Rohlfs, K., and Hüttemeister, S. Tools of Radio Astronomy. Springer-Verlag (2009). doi:10.1007/978-3-540-85122-6. 55, 60

[90] Wolleben, M., Landecker, T. L., Reich, W., and Wielebinski, R. An absolutely calibrated survey of polarized emission from the northern sky at 1.4 GHz. A&A, **448**:411–424 (2006). doi:10.1051/0004-6361:20053851. xiv, 61, 62, 63

[91] Wouthuysen, S. A. On the excitation mechanism of the 21 cm interstellar hydrogen emission line. Physica, **18**:75–76 (1952). doi:10.1016/S0031-8914(52)80143-0. 14

[92] Zahn, O., Reichardt, C. L., Shaw, L., Lidz, A., Aird, K. A., Benson, B. A., Bleem, L. E., Carlstrom, J. E., et al. Cosmic Microwave Background Constraints on the Duration and Timing of Reionization from the South Pole Telescope. ApJ, **756**:65 (2012). doi:10.1088/0004-637X/756/1/65. xiii, 9, 10

[93] Zaldarriaga, M. Polarization of the microwave background in reionized models. Phys. Rev. D, **55**:1822–1829 (1997). doi:10.1103/PhysRevD.55.1822. 8

[94] Zaldarriaga, M., Spergel, D. N., and Seljak, U. Microwave Background Constraints on Cosmological Parameters. ApJ, **488**:1–13 (1997). 7

[95] Zheng, H., Tegmark, M., Buza, V., Dillon, J. S., Gharibyan, H., Hickish, J., Kunz, E., Liu, A., et al. MITEoR: a scalable interferometer for precision 21 cm cosmology. MNRAS, **445**:1084–1103 (2014). doi:10.1093/mnras/stu1773. 141