Enabling the Next Generation of 21 cm Cosmology: Precision Analysis Techniques for Radio Cosmology Observations

Ruby Byrne

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2021

Reading Committee:

Miguel Morales, Chair

Matthew McQuinn

Gray Rybka

Program Authorized to Offer Degree:
Physics

University of Washington

Abstract

Enabling the Next Generation of 21 cm Cosmology: Precision Analysis Techniques for Radio Cosmology Observations

Ruby Byrne

Chair of the Supervisory Committee:
Professor Miguel Morales
Physics

The 21 cm emission line from neutral hydrogen offers a promising avenue for mapping the early universe. In particular, this emission could probe the Epoch of Reionization, or EoR, a phase in cosmological history when early stars and galaxies ionized the intergalactic medium. Measuring the 21 cm signal from the EoR would provide great insight into the structure of the early universe, revealing the dynamics of the universe's first galaxies and constraining the standard model of cosmology.

Detecting the EoR will require extremely precise measurements. The signal is very faint and is obscured by intervening emission, or foregrounds, that are 4-5 orders of magnitude brighter. Thus far, analyses have not achieved the precision required to measure the EoR signal. Calibration precision is a dominant limitation of these analyses.

This thesis focuses on understanding and mitigating calibration error for interferometric measurement of the EoR power spectrum signal. We propose a novel approach to calibration, called unified calibration, that is based on Bayesian statistics. We show in simulation that unified calibration outperforms the established techniques in the field.

Additionally, we present results that expand and improve the calibration sky model. Precision calibration requires an accurate and complete model of the sky signal, and we examine the impact of sky model error on calibration performance. We describe the development of a new polarized imaging pipeline that we then use to map diffuse emission at 182 MHz across 11,000 square degrees of the Southern Hemisphere sky. The map is presented in all four Stokes parameters and contains emission on scales of 1° to 9°. We discuss the implementation of this diffuse map in calibration and explore its effect on calibration performance.

TABLE OF CONTENTS

LIST OF FIGURES

GLOSSARY

21 CM COSMOLOGY: The observational study of the history and structure of the universe using the 21 cm emission line from neutral hydrogen.

ARRAY: An interferometric telescope composed of many antennas.

BASELINE: A pair of antennas that are correlated to produce a visibility, can also refer to the vector displacement between a pair of antennas.

CALIBRATION: Process of empirically calculating antennas' responses to the sky signal by comparing measured data to the expected signal.

COMPACT SOURCE: Source of radio emission on the sky (e.g. a radio galaxy) that is small in angular extent.

CMB: Cosmic Microwave Background, electromagnetic radiation in the microwave spectrum dating from the early universe. The CMB dates back to recombination, the instant in cosmological history when neutral atoms first formed and the universe first became transparent.

DIFFUSE EMISSION: Emission on the sky with large angular extent.

EOR: Epoch of Reionization, the period in cosmological history in which the hydrogen in the IGM ionized.

EOR WINDOW: Region of power spectrum space that is not inherently dominated by foreground emission, where there is sensitivity to the 21 cm signal from the EoR.

EPPSILON: Error Propagated Power Spectrum with Interleaved Observed Noise (often stylized ϵPPSILON), a data analysis software package that performs power spectrum estimation of imaged interferometric data.

FHD: Fast Holographic Deconvolution, a software package that performs radio interferometric data analysis and imaging.

FOREGROUND WEDGE: Wedge-shaped region of power spectrum space representing modes dominated by foreground emission.

FOREGROUNDS: Radio emission, predominantly from galactic synchrotron emission, that lies between the observer and the cosmological 21 cm signal.

GAIN: Parameter fit in calibration that encodes the response of an antenna.

GRIDDING: A step in the interferometric imaging process where visibility measurements are reconstructed in the *uv* plane.

IGM: Intergalactic Medium, baryonic matter in the universe (primarily hydrogen) that lies outside of galaxies.

IMAGING: The process of analyzing interferometric data to produce images of the sky.

INTERFEROMETRY: An astronomical technique that measures electromagnetic waves by correlating signals of antennas in an array.

POINT SOURCE: A class of compact source that is sufficiently small to be unresolved by the instrument. For analysis purposes, these sources can be approximated as having no angular extent.

PSF: Point spread function, an array's image reconstruction of a point source.

REDSHIFT: A measure of cosmological time, or equivalently astronomical distance from the Earth, represented as the apparent Doppler shift due to the universe's expansion.

RFI: Radio Frequency Interference, human-made signals, predominantly telecommunications signals from Earth-based and satellite sources, that interfere with radio astronomy measurements.

UV PLANE: Plane spanned by baseline vectors, equal to the Fourier dual of the sky.

VISIBILITY: The raw data in radio interferometry, derived by correlating voltage signals between two antennas.

ZENITH: The point on the sky directly overhead.

ACKNOWLEDGMENTS

As a graduate student I have been fortunate to belong to a vibrant scientific community. None of the work in this thesis would have been possible without the support of my colleagues, collaborators, and community members. While I can't possibly thank everyone who contributed to my graduate research, I would like to highlight some of the figures who were most influential to me.

To my advisor Miguel Morales, thank you for your constant guidance, for sharing your deep knowledge of radio interferometry and the process of scientific research, and for unequivocally supporting me in my career goals even when neither of us knew what those goals were. I learn so much through our conversations, both about the profound questions of the origin of the universe and the mundane problems of a missing minus sign.

I am extremely fortunate to have been co-advised by Bryna Hazelton. Since my first day in the group, Bryna has guided my work with extreme patience and brilliant insight. She makes the hardest problems look easy and is a true inspiration to me.

Thank you to all my fellow graduate group members with whom I've worked closely: Nichole Barry, Mike Wilensky, Pyxie Starr, and Dara Storer. You are all excellent collaborators, and thank you to Nichole, Mike, and Pyxie for being wonderful companions in adventure in Australia, South Africa, Japan, and Italy. In particular, I thank Nichole for spending countless hours training me in data analysis. As a postdoc, she continues to be a valued mentor to me.

I am grateful to all my Australian collaborators, particularly to Cath Trott and Rachel Webster for hosting me in their research groups. I thank many others who guide and support me in my career, in particular Jonathan Pober, Danny Jacobs, Adrian Liu, and James

Aguirre. Ian Sullivan has provided me with invaluable help in all things FHD and is always a pleasure to collaborate with.

I thank my thesis committee for supporting and evaluating my progress toward my PhD: Miguel Morales, Matt McQuinn, Gray Rybka, Sarah Tuttle, and Paul Wiggins. I acknowledge Ann Nelson, who served on my committee until her death in 2019. The UW physics community feels her absence deeply.

Thank you to all my fellow union members and organizers at UAW Local 4121, the UW academic student employees and postdocs union, who fought and continue to fight for workplace protections and fair compensation. UW works because we do! Also, thank you to all other unions who stand in solidarity with us and who refused to cross the picket line during our 2018 strike.

My career in physics was launched by my undergraduate professors at Macalester College. Thank you to Tonnis ter Veldhuis, Jim Doyle, and James Heyman for cultivating my interest in physics. I thank Karen Saxe for envisioning a future for me in science before I could myself. Her encouragement gave me the vision to pursue graduate school. I will be forever indebted to Jim Hoppe, who played a pivotal role in my undergraduate success.

The field of physics is plagued with discrimination, including overt and covert racism and sexism. I thank all of my colleagues who challenge this precedent every day, and I join them in striving for a more equitable future for our field. I am particularly inspired by the writing of Chanda Prescod-Weinstein, who models practicing physics with courage and integrity.

I owe everything to my wonderful friends and family—you know who you are. I love you so much.

The majority of the work in this thesis was performed in Seattle, Washington, located in the unceded territory of the Coast Salish and Duwamish people. The site of the MWA telescope is traditional land of the Wajarri Yamatji people.

Chapter 1

THE 21 CM SIGNAL AS A WINDOW INTO THE EARLY UNIVERSE

Human societies throughout history have engaged in cosmological scholarship, seeking to understand the origin, structure, and evolution of the universe as we know it. In this pursuit, the night sky has offered an incredible wealth of knowledge. From our vantage point on Earth, humans have peered into space to gain clues about the universe's distant past. This thesis represents one small contribution to this age-old tradition.

This thesis relates to 21 cm cosmology, a branch of cosmology research that uses the signature 21 cm emission line from neutral hydrogen to probe the evolution of the universe. In particular, this work is concerned with study of the Epoch of Reionization, or EoR, a stage in cosmological history that occurred over 13 billion years ago. We expect that measurements of the 21 cm signal from the EoR could give us great insights into nature of the early universe and the origins of stars and galaxies.

In the past decade, the promise of measuring this signal has driven the development of several large radio telescopes. However, the measurement has proved elusive. Science teams from across the world have established progressively more constraining limits on the EoR's 21 cm signal but have not achieved a detection. In this thesis we explore the limitations of previous analyses and present tools that will enable the next generation of EoR experiments.

This chapter offers an introduction to the EoR and 21 cm cosmology. Chapter 2 describes established analysis techniques in detail. Chapter 3 discusses the interaction between sky model error and calibration, one of the dominant sources of error in previous experiments. The remainder of the thesis presents new analysis techniques that can deliver more precise measurements of the cosmological 21 cm signal.

1.1 An Overview of the History of the Universe

In the past century, astronomical observations have led us to an understanding of the universe's origins called Big Bang cosmology. We understand that the universe began 13.8 billion years ago in an event called the Big Bang. The early universe was extremely hot and dense. In the first microseconds after the Big Bang the universe expanded extremely rapidly through a process called inflation. This expansion slowed significantly after these first fractions of a second, and the fundamental particles that comprise the building blocks of all matter in the universe formed: photons, leptons, quarks, and hadrons. These building blocks ultimately evolved into the galaxies, stars, and planets of the modern universe, and the universe continues to expand to this day.

For the first hundreds of thousands of years the universe was a dense, opaque plasma. This changed nearly instantaneously 370,000 years after the Big Bang, at a redshift of $z = 1100$.[1] At this point the universe had expanded and cooled sufficiently to enable electrons and protons to combine into hydrogen atoms, an event called recombination. As a result, the universe became transparent for the first time as photons decoupled from matter. These photons are visible today as the Cosmic Microwave Background (CMB).

Observations of the CMB have provided significant insight into the structure and evolution of the early universe. The CMB was first measured in 1965 with the Holmdel Horn Antenna microwave telescope in New Jersey (Penzias and Wilson, 1965). Follow-up measurements were performed with satellite experiments including the Soviet RELIKT-1 (Strukov and Skulachev, 1984), the Cosmic Background Explorer (COBE; Bennett et al. 1996), the Wilkinson Microwave Anisotropy Probe (WMAP; Bennett et al. 2013), and the Planck Satellite (Aghanim et al., 2020). These measurements revealed that the CMB comprises highly isotropic blackbody radiation with small temperature variations. The results lent credence to the Big Bang model and provided constraints on cosmological parameters including the

---

[1] Redshift, or z, is a metric of cosmological time. z represents the factor by which radiation's wavelength increases as a result of the universe's expansion. Radiation from distant sources, which was emitted long in the past, is redshifted significantly. Larger values of z represent events further in the past.

age of the universe, its rate of expansion, and the density of baryonic and dark matter in the universe.

Following recombination and the emission of the CMB, the universe entered a phase called the Dark Ages, aptly named for the absence of bright luminous emission. During this period the universe was filled primarily with neutral hydrogen as well as some helium and trace heavier elements. At first this matter was nearly uniformly distributed throughout the universe, but small overdensities—the same inhomogeneities visible as temperature fluctuations in the CMB—began to condense under the force of gravity. Several hundred million years after the Big Bang these regions collapsed sufficiently to form early stars and galaxies, ending the Dark Ages and ushering in the Cosmic Dawn.

These early stars and galaxies put into motion another cosmological transition. The radiation from these luminous objects heated the surrounding gas, called the intergalactic medium or IGM. Moreover, the radiation was sufficiently energetic to ionize hydrogen in the IGM. Bubbles of ionized hydrogen formed around the early galaxies. In time, these bubbles multiplied and grew until the IGM was completely ionized. This process is called reionization, and the extended period in which this occurred is called the Epoch of Reionization (EoR).

There is much we don't understand about the EoR. CMB experiments have offered some constraint on the EoR, with WMAP and Planck results estimating that the EoR occurred around $z \approx 7.4 - 11.7$ (Bennett et al., 2013; Aghanim et al., 2020). On the other side, observations of high redshift galaxies suggest that reionization concluded by $z \approx 6$ (Fan et al., 2006b; Choudhury et al., 2015). However, the redshift and duration of reionization remain poorly constrained. In addition, we don't know what radiation sources drove reionization: Was reionization dominated by stars or black holes (active galactic nuclei, or AGN)? Were there many small ionized bubbles or fewer large bubbles? Did more dense regions ionize earlier or later than less dense regions? Answering these questions would aid in our understanding of early star and galaxy formation and further constrain important cosmological parameters such as the temperature, density, and expansion rate of the early universe.

With the culmination of the EoR the modern universe evolved. Currently the IGM is fully ionized. Through precise measurements of the night sky, we can probe a period in the distant past when this was not the case, offering us elusive clues about the history of our universe.

1.2 The 21 cm Signal as a Probe of Cosmological Evolution

One of the most promising probes of the EoR is the 21 cm emission line from neutral hydrogen (Zaldarriaga et al., 2004; Bowman et al., 2006; Fan et al., 2006a; McQuinn et al., 2006; Santos and Cooray, 2006; Valdés et al., 2006; McQuinn et al., 2007; Bowman et al., 2008; Mao et al., 2008; Pritchard and Loeb, 2008; Barkana, 2009; Morales and Wyithe, 2010; Pober et al., 2014). This emission has long been used by astronomers to map regions of neutral hydrogen within galaxies. More recently, cosmologists have turned their attention to 21 cm emission at high redshift. Mapping this signal can reveal the density and distribution of neutral hydrogen in the IGM before and during the EoR, providing a direct measurement of the reionization process.

The 21 cm emission line emerges from neutral hydrogen's hyperfine, or spin-flip, transition. This transition relates to the alignment between the electron and proton spins, which interact as a result of the induced magnetic dipoles of each particle. Quantum mechanically, the spin alignment states occupy a higher energy triplet state and a lower energy singlet state (Feynman, 1963). The energy transition between these two states is a tiny 5.9×10^{-6} eV, corresponding to a wavelength of 21 cm.

The hyperfine transition is known as a "forbidden" transition and is very rare, with a lifetime of $\sim 10^7$ years. As a result, the emission line is extremely narrow. This means that the observed frequency can be precisely mapped to redshift. Measuring the cosmological 21 cm emission could therefore produce a 3-D tomographic map of neutral hydrogen throughout an enormous volume of the universe.

The 21 cm signal is very faint at high redshift, and detection requires an exquisitely sensitive telescope. Furthermore, the measurement is contaminated by bright intervening emission, called foregrounds. This emission is dominated by galactic synchrotron radiation and is a staggering 4-5 orders of magnitude brighter than the 21 cm EoR signal. Mitigating foregrounds' contamination is among the principal challenges of 21 cm cosmology.

There are three classes of 21 cm measurements. Global experiments aim to measure the total intensity of 21 cm emission at a given redshift, averaged across the entire sky. This gives insight into the average density and temperature of neutral hydrogen in a given cosmological epoch[2]. One such experiment, called the Experiment to Detect the Global EoR Signature (EDGES), reported a detection of the global 21 cm signal at 78 MHz, corresponding to the cosmological Dark Ages (Bowman et al., 2007). The result suggests that the early universe could have been much colder than expected, indicating a potential discrepancy in the standard model of cosmology.

While the global measurement is a powerful tool, it has limited ability to probe the EoR because it does not contain any information about the 21 cm signal's angular structure. It therefore cannot determine the size and distribution of the neutral hydrogen regions. Direct imaging of the 21 cm signal would measure the intensity of the signal in each direction of the sky. While this may ultimately be feasible, current techniques do not enable the sensitivity required to directly image the EoR. Instead, this thesis focuses on the third class of 21 cm measurements: the power spectrum. The power spectrum is a statistical measurement of the 21 cm signal's angular scales. It takes advantage of the signal's isotropy—the fact that its statistical properties are directionally invariant—to achieve improved sensitivity over direct imaging.

1.3 The Instruments

The redshifted 21 cm line from the EoR appears in the low-frequency radio spectrum, at wavelengths of about $1-2$ meters and frequencies of $\sim 100-200$ MHz. A number of large radio arrays have established upper limits on the power spectrum of the EoR signal,

---

[2] More specifically, the intensity of the 21 cm signal depends on the "spin temperature," a measure of the fraction of hydrogen atoms in the excited hyperfine energy state. As the cosmological signal is backlit by the CMB, the observed intensity refers to the contrast with the CMB signal and appears negative in certain cosmological epochs. For further discussion of the signal intensity, see Furlanetto et al. 2006.

including the Giant Metrewave Radio Telescope (GMRT) in India (Pen et al., 2009; Paciga et al., 2011, 2013), the Low-Frequency Array (LOFAR) in the Netherlands (Mertens et al., 2020), the Murchison Widefield Array (MWA) in Australia (Beardsley et al., 2016; Barry et al., 2019b; Li et al., 2019; Trott et al., 2020), and the Donald C. Backer Precision Array for Probing the EoR (PAPER) in South Africa (Ali et al., 2015; Kolopanis et al., 2019). A forthcoming limit produced with the Hydrogen Epoch of Reionization Array (HERA) in South Africa is currently in preparation (Kern et al., in prep.).

While these instruments have diverse designs, they all have some commonalities. All are radio interferometers consisting of arrays of antennas, where signals from the antennas are correlated to produce visibility measurements (see Chapter 2 for a description of the measurement process). To achieve the sensitivity required to constrain the EoR the arrays must be quite large, with dozens or even hundreds of antenna elements. All are built in remote locations to protect against human-made radio emission (called Radio Frequency Interference, or RFI), a significant contaminant of their measurements. To further eliminate contamination, their sites are designated as “radio quiet zones”, with restrictions on the use of electronics such as cell phones, Wi-Fi, digital cameras, and microwave ovens in the vicinity of the telescopes.

These radio telescopes can be broadly categorized in two classes related to the layout of antennas in the arrays. Regular arrays place antennas on a rectangular or hexagonal grid. Non-regular arrays have irregularly spaced antennas, often with a pseudo-random configuration. The distinction between regular and non-regular array layouts has vast implications for the instruments’ capabilities and the analysis techniques used, and there is no clear consensus on which array type offers better prospects for detecting the EoR signal. Exploring the tradeoffs between these two classes of arrays is one of the key aims of Chapter 3.

Much of the work presented in this thesis applies to any EoR experiment—and even to 21 cm cosmology experiments beyond EoR science. However, this work was produced in close partnership with two telescope collaborations, the MWA and HERA, and Chapter 7 presents results from data taken with the MWA. In this section we present a brief overview of each of these two instruments.

1.3.1 The MWA

The Murchison Widefield Array (MWA) is located in the remote outback of Western Australia. The MWA enables a diverse range of astronomy research, and EoR science is one of its key science goals.

Each of the MWA's antennas, or tiles, is composed of 16 individual dual-polarization dipole elements arranged in a 4 × 4 grid, pictured in Figure 1.1. A beamformer combines signals from the 16 dipoles. The tiles are electronically steerable: the beamformer inserts an analog delay into the signal from particular dipoles to direct the tile's response toward different parts of the sky. This allows the telescope to target sources away from zenith and enables it to track sources as they move overhead. Beneath each tile is a mesh ground screen. At long wavelengths this mesh behaves like a mirror, amplifying the signal and concealing irregularities in the ground substrate that could distort the measurement.

Construction of the MWA Phase I completed in 2012, and in 2017 the telescope entered its Phase II after undergoing a major upgrade. The initial Phase I array had 128 antennas in a non-regular, pseudo-random configuration with a maximum extent of about 3 km (Beardsley et al., 2012; Tingay et al., 2013). The configuration of the Phase I array is plotted in Figure 1.2. The Phase II upgrade (Wayth et al., 2018) doubled the array size, bringing the total number of antennas to 256, of which 128 can be correlated at once. These new antennas fall into two categories. Far-flung outrigger antennas expanded the array extent to 5 km, while two hexagonal sub-arrays were added near the core of the array. Because not all antennas can be correlated at once, the MWA Phase II operates in one of two modes. An extended mode correlates the outrigger antennas and omits the hexagonal sub-arrays, improving the array's resolution for a broad variety of astronomy research applications. The compact mode, primarily used for EoR science, sacrifices resolution for the improved sensitivity from the array's core and the hexagonal sub-arrays. Figure 1.3 depicts the antenna layout of the Phase II compact configuration.

Figure 1.1: Photos of the MWA radio telescope. (Caption continued.)

Figure 1.1: Photos of the MWA radio telescope. On the top is an example of an MWA antenna, called a tile, composed of 16 dual-polarization dipole elements arranged over a mesh ground screen. The telescope consists of 256 such tiles in an array that spans 5 km. On the bottom is a photo of a science team involved in constructing the MWA's Phase II upgrade; I am pictured second from the right. Other team members pictured are, from left to right, Adam Lanman, Wenyang Li, Joshua Kerrigan, Mia Walker, Kim Steele, David Emrich, and Nichole Barry. Photos are credited to the MWA collaboration, Curtin University, and Greg Rowbotham.

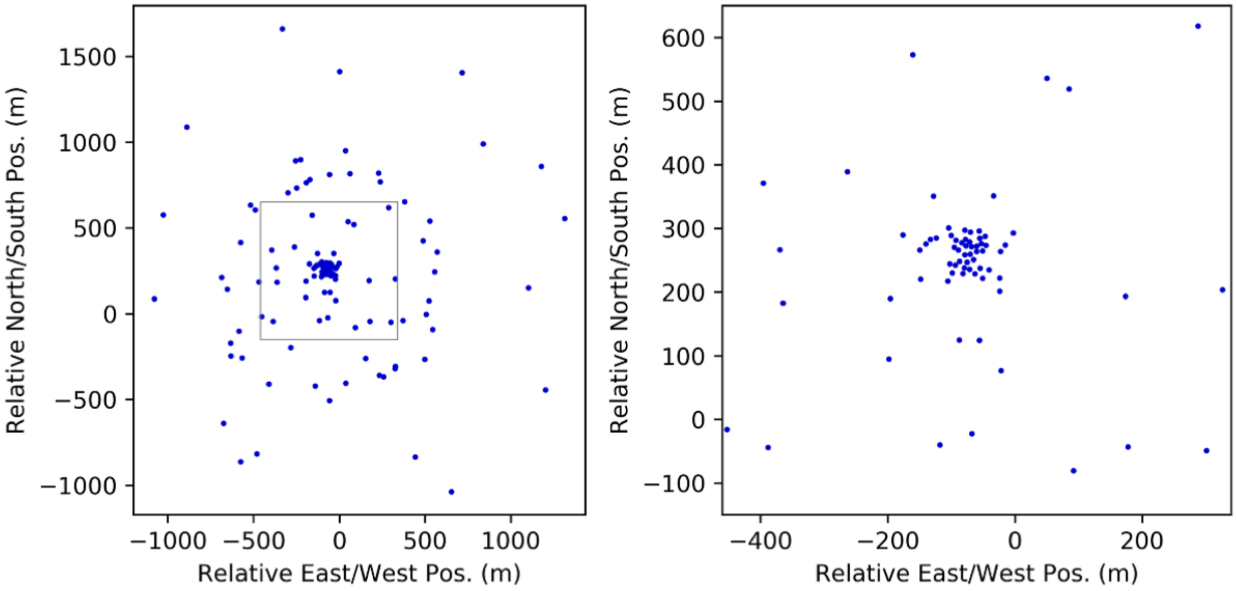


Figure 1.2: Plot of the antenna positions of the MWA Phase I. The MWA Phase I is a non-regular array with antennas arranged in a pseudo-random configuration. The plot on the left depicts the full array, and the plot on the right zooms in on the array's compact core (outlined as a grey square on the left).

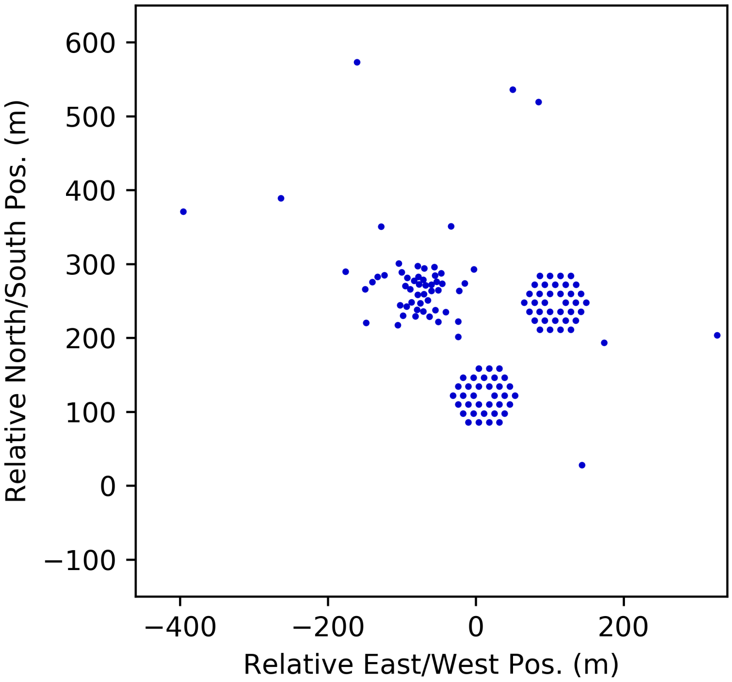


Figure 1.3: Plot of the antenna positions of the MWA Phase II in its compact configuration. The Phase II upgrade to the telescope added two hexagonal sub-arrays near the array's core. In this configuration, the MWA is a hybrid array with both regular and non-regular components.

1.3.2 HERA

The Hydrogen Epoch of Reionization Array (HERA), pictured in Figure 1.4, is located in the South African Karoo. It a second-generation instrument built on the same site as the decommissioned Donald C. Backer Precision Array for Probing the EoR (PAPER). In contrast to the MWA, HERA is purpose-built for EoR science.

When complete, HERA will consist of 350 fully-correlated dish antennas. The dishes are 14 meters in diameter, constructed of PVC and wire mesh. As with the MWA's ground screens, the mesh reflects radio waves just as a mirror reflects visible light. Dual-polarization feeds, called Vivaldi feeds, are suspended above each dish (Fagnoni et al., 2020). All but 30 antennas are configured in a close-packed configuration. The configuration, plotted in Figure 1.5, takes the form of a modified hexagon with three offset sections.

Unlike the MWA, HERA's antennas are not steerable. This means that HERA cannot track sources as they move overhead. Instead, it operates as what is called a drift-scan telescope, detecting sources as they move through the instrument's stationary response on the sky.

1.4 The Challenge of 21 cm Cosmology

In the past decade there has been considerable investment in 21 cm cosmology and EoR science by researchers around the world. The promise of measuring the EoR power spectrum has driven construction of several large radio arrays, and there has been substantial progress in the development of array hardware and data analysis software aimed at enabling to this measurement. Even so, we have not achieved a measurement of the EoR signal.

The predicted EoR power spectrum signal is very faint, but this alone cannot explain why a detection remains elusive. Existing instruments have large collecting areas and achieve the sensitivity required to detect the EoR. Instead, the limitations in EoR experiments are due to systematic error in the measurement. These errors are at a level that swamps the faint cosmological signal. Furthermore, they are not mitigated by gathering more data or

13.09.2019

Figure 1.4: Photos of the HERA telescope. (Caption continued.)

Figure 1.4: Photos of the HERA telescope. The top photo depicts part of the array, with close-packed 14 meter dishes in a hexagonal configuration. When complete, the array will have 350 dish antennas. Feeds are suspended over each dish with cables. On the bottom, I raise a feed. Photos are credited to Scott Dynes (top) and Pyxie Starr (bottom).

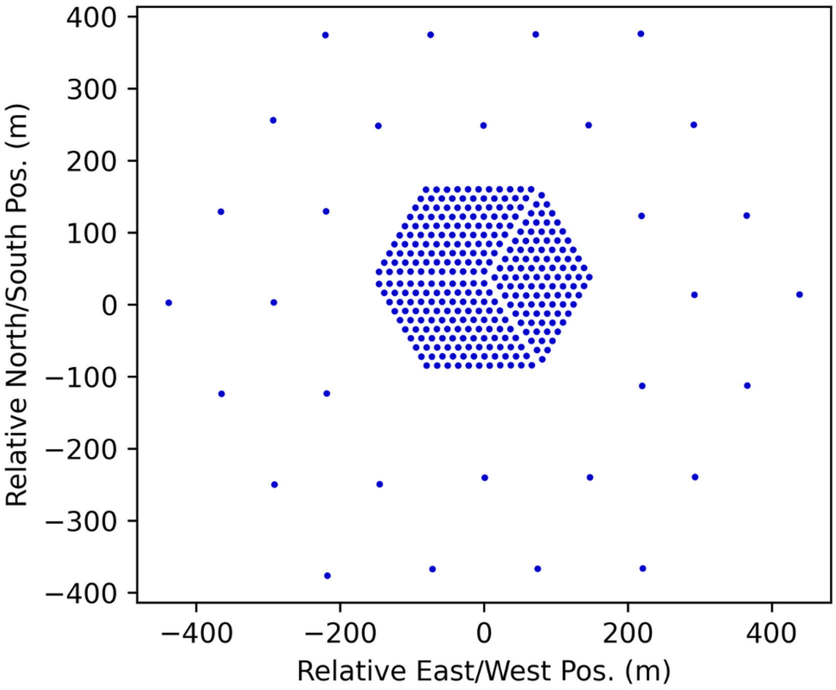


Figure 1.5: Plot of the HERA antenna positions. HERA is a regular array with antennas positioned in a close-packed modified hexagonal configuration. 30 outrigger antennas improve the array's resolution.

further expanding the instruments' collecting area. Progress in the field therefore requires mitigation of these sources of systematic error and development of new analysis techniques that can deliver the needed precision.

One of the key systematic errors in EoR analyses stems from foreground contamination. As noted above, foregrounds are 4-5 order of magnitude brighter than the EoR signal. In principle, the foreground and cosmological signals can be separated based on their distinct spectral properties. Because foregrounds are emitted through synchrotron processes, they are very spectrally smooth, whereas the EoR signal has substantial spectral structure. However, current analysis techniques limit this separability. Because the foreground signal is so bright, even low-level signal mixing precludes detection of the cosmological signal.

Calibration error is one of the dominant error mechanisms producing foreground contamination and is the focus of this thesis. At the same time, we note that calibration error is not the only limitation of current analyses. RFI contamination is a significant concern, with recent work highlighting the deleterious effect of faint RFI on cosmological measurements (Wilensky et al., 2020). Measurement precision can also be limited by the dynamic range of the analysis, which must be able to precisely reconstruct both the bright foreground power and the faint cosmological signal. Recent analysis advances have tackled aliasing effects that limit dynamic range (Barry et al., 2019b). Nevertheless, calibration error is a persistent and predominant limitation in 21 cm analyses. The calibration techniques used to produce the field's current limits will not enable a detection of the EoR signal.

Calibration calculates the response of each antenna in an array. This fits variations in the sensitivity of the antennas, the precise timing of each antenna signal, and the frequency dependence of each antenna's response. Errors in the frequency response are of particular importance for 21 cm cosmology because they introduce frequency structure into the intrinsically spectrally smooth foregrounds. The foregrounds thereby can't be separated from the spectrally varying cosmological signal. As a result, calibration precision is critical for enabling an uncontaminated measurement of the EoR signal, and the development and characterization of precision calibration techniques is an active area of research (Wieringa, 1992; Pen et al., 2009; Liu et al., 2010; Kazemi et al., 2011; Kazemi and Yatawatta, 2013; Grobler et al., 2014; Newburgh et al., 2014; Salvini and Wijnholds, 2014; Zheng et al., 2014; Yatawatta, 2015; Barry et al., 2016; Berger et al., 2016; Dillon and Parsons, 2016; Ewall-Wice et al., 2016; Grobler et al., 2016; Patil et al., 2016; van Weeren et al., 2016; Wijnholds et al., 2016; Ollier et al., 2017; Dillon et al., 2018; Grobler et al., 2018; Joseph et al., 2018; Li et al., 2018; Tasse et al., 2018; Kohn et al., 2019; Li et al., 2019; Orosz et al., 2019; Albert et al., 2020; Dillon et al., 2020; Joseph et al., 2020; Kern et al., 2020; Mertens et al., 2020; Sob et al., 2020).

This thesis presents major advances in precision calibration for 21 cm cosmology. It tackles the problem of calibration error from two angles. First, it explores the statistical underpinnings of calibration, examining the assumptions of established calibration techniques and proposing a new calibration framework called unified calibration, presented in Chapter 4. Unified calibration represents a highly flexible, physically-motivated approach to calibration based on Bayesian statistics. In Chapter 5 we show that in simulation unified calibration performs better than the other calibration approaches widely used in the field.

All precision calibration approaches require a good *a priori* model of the sky signal. The second avenue by which this thesis improves calibration's capabilities is by expanding upon existing sky models through mapping diffuse emission. This emission has characteristic angular scales of several degrees and is highly polarized. We present the map and discuss its implementation for precision calibration in Chapter 7.

The development of unified calibration and mapping of polarized diffuse emission confront two of the major limitations of previous EoR analyses. In the course of their development, we also gained a more robust understanding of the mechanisms of calibration error (Chapter 3) and developed a high-fidelity polarized imaging pipeline for widefield radio telescopes (Chapter 6). Together, these advances pave the way towards a detection of the EoR power spectrum.

Chapter 2

INTERFEROMETRIC MEASUREMENT OF THE COSMOLOGICAL 21 CM SIGNAL

This chapter presents an overview of the analysis techniques used to process interferometric data and estimate the power spectrum of the cosmological 21 cm signal. The field of 21 cm cosmology employs a broad diversity of analysis approaches, and in this chapter we focus on the approach taken by the Fast Holographic Deconvolution[1] (FHD) and Error Propagated Power Spectrum with Interleaved Observed Noise[2] (ϵPPSILON) software packages (Sullivan et al., 2012; Jacobs et al., 2016; Barry et al., 2019a). Analyses presented throughout this paper use these two packages, where FHD performs interferometric imaging and ϵPPSILON produces power spectrum estimates from those images.

Precision data analysis requires an excellent understanding of the instrumental process, so in §2.1 we discuss the properties of a radio interferometer's measurement and define variables and expressions used throughout this thesis. §2.2 describes how to transform the data into images of the sky, following FHD's analysis approach. §2.3 then provides an overview of the power spectrum estimation pipeline as implemented in the ϵPPSILON software package. §2.4 introduces interferometric calibration, which is a primary focus of the work presented in this thesis.

2.1 The Visibility Measurement

The fundamental unit of measurement for an interferometric array is the visibility. A visibility is formed by correlating (multiplying and time-averaging) the voltage signals of two

---

[1] `https://github.com/EoRImaging/FHD`

[2] `https://github.com/EoRImaging/eppsilon`

antennas. The pairs of antennas are called baselines; the term “baseline” also refers to the vector displacement between the antennas. Visibilities comprise a complex value per baseline, polarization mode, frequency channel, and time step. Because baselines are formed between all pairs of antennas in an array (and even between an antenna and itself — these are called the “autocorrelation visibilities”), the number of visibilities for an array of N antennas scales as $(N^2+N)/2$. Consequently, data volumes explode for large-N arrays, posing significant challenges for data storage and analysis.

Each antenna has a complex response pattern on the sky, called the antenna beam. We denote the beam for antenna j as $D_j(\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a two-element vector representing the angular position on the sky. We can write the visibility derived by correlating signals from antennas j and k as

$$v_{jk} = \left\langle \left[\int_{-\infty}^{\infty} d^2\boldsymbol{\theta}\, D_j(\boldsymbol{\theta})\, e^{2\pi i \boldsymbol{\theta}\cdot\boldsymbol{r}_j/\lambda} E(\boldsymbol{\theta}) \right] \left[\int_{-\infty}^{\infty} d^2\boldsymbol{\theta}\, D_k(\boldsymbol{\theta})\, e^{2\pi i \boldsymbol{\theta}\cdot\boldsymbol{r}_k/\lambda} E(\boldsymbol{\theta}) \right]^* \right\rangle . \tag{2.1}$$

Here $\boldsymbol{r}_j$ and $\boldsymbol{r}_k$ denote the positions of antennas j and k, respectively, λ is the wavelength of emission, and $E(\boldsymbol{\theta})$ represents the electric field on the sky. The asterisk * represents the complex conjugate, $\cdot$ denotes the vector dot product, and the angle brackets $\langle\rangle$ indicate the time average. Here we assume the electric field is unpolarized; in Chapter 6 we extend this formalism to describe polarized emission. This expression assumes a planar array, such that $\boldsymbol{r}$ can be represented as a two-element vector. It also assumes the far-field limit, such that the emission sources are far from the array receiving elements. Many arrays are so-called “homogeneous” arrays with near-identical antennas. For those arrays, we can assume that the antenna beams are equal: $D_j(\boldsymbol{\theta}) \approx D_k(\boldsymbol{\theta})$. For generality, we do not make that assumption here.

We can rewrite this equation as

$$v_{jk} = \int_{-\infty}^{\infty} d^2\boldsymbol{\theta} \int_{-\infty}^{\infty} d^2\boldsymbol{\theta}'\, D_j(\boldsymbol{\theta}) D_k^*(\boldsymbol{\theta}')\, e^{2\pi i (\boldsymbol{\theta}\cdot\boldsymbol{r}_j - \boldsymbol{\theta}'\cdot\boldsymbol{r}_k)/\lambda} \left\langle E(\boldsymbol{\theta}) E^*(\boldsymbol{\theta}') \right\rangle . \tag{2.2}$$

The electric field at different positions on the sky is uncorrelated, so $\left\langle E(\boldsymbol{\theta}) E^*(\boldsymbol{\theta}') \right\rangle = 0$ when

$\boldsymbol{\theta} \neq \boldsymbol{\theta}'$. We therefore get that

$$v_{jk} = \int_{-\infty}^{\infty} d^2\boldsymbol{\theta}\, D_j(\boldsymbol{\theta}) D_k^*(\boldsymbol{\theta})\, e^{2\pi i \boldsymbol{\theta}\cdot(\boldsymbol{r}_j - \boldsymbol{r}_k)/\lambda} \left\langle |E(\boldsymbol{\theta})|^2 \right\rangle. \tag{2.3}$$

This relationship is at the core of radio interferometric analysis.

To further simplify this expression, we define new quantities. The sky brightness is given by

$$S(\boldsymbol{\theta}) = \left\langle |E(\boldsymbol{\theta})|^2 \right\rangle, \tag{2.4}$$

the beam power response of the baseline formed by antennas j and k is

$$B_{jk}(\boldsymbol{\theta}) = D_j(\boldsymbol{\theta}) D_k^*(\boldsymbol{\theta}), \tag{2.5}$$

and

$$\boldsymbol{x}_{jk} = (\boldsymbol{r}_j - \boldsymbol{r}_k)/\lambda. \tag{2.6}$$

Here $\boldsymbol{x}_{jk}$ is simply the baseline vector for antennas j and k in units of wavelengths.[3] The two elements of the baseline vectors are often denoted u and v, and the plane spanned by these vectors is called the uv plane.

Rewriting Equation 2.3 in terms of these quantities, we get

$$v_{jk} = \int_{-\infty}^{\infty} d^2\boldsymbol{\theta}\, B_{jk}(\boldsymbol{\theta}) S(\boldsymbol{\theta})\, e^{2\pi i \boldsymbol{\theta}\cdot \boldsymbol{x}_{jk}}. \tag{2.7}$$

This is sometimes further simplified to

$$v_{jk} = \int_{-\infty}^{\infty} d^2\boldsymbol{\theta}\, T_{jk}(\boldsymbol{\theta}) S(\boldsymbol{\theta}), \tag{2.8}$$

where the quantity

$$T_{jk}(\boldsymbol{\theta}) = B_{jk}(\boldsymbol{\theta}) e^{2\pi i \boldsymbol{\theta}\cdot \boldsymbol{x}_{jk}} \tag{2.9}$$

is the baseline transfer function and encodes the full baseline response (Eastwood et al., 2018). Equation 2.8 is often recast as a matrix operation

$$\boldsymbol{v} = \mathbf{T}(\boldsymbol{v}, \boldsymbol{\theta}) \boldsymbol{S}(\boldsymbol{\theta}), \tag{2.10}$$

---

[3] The notation in this thesis departs from the usual convention in the field, which represents the baseline vectors with the variable $\boldsymbol{u}$. In this thesis, $\boldsymbol{u}$ refers to calibration parameters (see §2.4.2 and Chapter 4).

where $\boldsymbol{S}(\boldsymbol{\theta})$ is a vector of length equal to the number of pixels on the sky and $\mathbf{T}(\boldsymbol{v}, \boldsymbol{\theta})$ is a matrix that maps between the sky and the visibility measurements. The challenge of interferometric data analysis estimating $\boldsymbol{S}(\boldsymbol{\theta})$ from the visibilities.

$\mathbf{T}(\boldsymbol{v}, \boldsymbol{\theta})$ is not invertible. There are a number of sophisticated interferometric imaging algorithms that produce estimates of $\boldsymbol{S}(\boldsymbol{\theta})$, and in the subsequent section we discuss the approach used throughout this thesis. In general, interferometric imaging makes use of the Fourier relationship between the sky signal and the baseline measurements.

Returning to Equation 2.7, we note that the integral describes a Fourier transform. We define the Fourier transform of the sky as

$$\tilde{S}(\boldsymbol{x}) = \mathcal{FT}\left[S(\boldsymbol{\theta})\right] = \int_{-\infty}^{\infty} d^2\boldsymbol{\theta}\, S(\boldsymbol{\theta})\, e^{2\pi i \boldsymbol{\theta}\cdot\boldsymbol{x}}. \tag{2.11}$$

Here $\mathcal{FT}$ denotes the Fourier transform operation and the tilde represents the Fourier transformed quantity. Via the convolution theorem, Equation 2.7 becomes

$$v_{jk} = [\tilde{B}_{jk} * \tilde{S}](\boldsymbol{x}_{jk}) = \int_{-\infty}^{\infty} d^2\boldsymbol{x}\, \tilde{B}_{jk}(\boldsymbol{x}_{jk} - \boldsymbol{x})\tilde{S}(\boldsymbol{x}), \tag{2.12}$$

where $*$ denotes the convolution.

We can express this function in terms of the baseline transfer function. From Equation 2.9, we get that

$$\tilde{T}_{jk}(\boldsymbol{x}) = \int_{-\infty}^{\infty} d^2\boldsymbol{\theta} B_{jk}(\boldsymbol{\theta}) e^{2\pi i \boldsymbol{\theta}\cdot(\boldsymbol{x}+\boldsymbol{x}_{jk})}, \tag{2.13}$$

so

$$\tilde{T}_{jk}(\boldsymbol{x} - \boldsymbol{x}_{jk}) = \int_{-\infty}^{\infty} d^2\boldsymbol{\theta} B_{jk}(\boldsymbol{\theta}) e^{2\pi i \boldsymbol{\theta}\cdot\boldsymbol{x}} \tag{2.14}$$

and $\tilde{T}_{jk}(\boldsymbol{x} - \boldsymbol{x}_{jk}) = \tilde{B}_{jk}(\boldsymbol{x})$. We then get that

$$v_{jk} = \int_{-\infty}^{\infty} d^2\boldsymbol{x}\, \tilde{T}_{jk}(-\boldsymbol{x})\tilde{S}(\boldsymbol{x}). \tag{2.15}$$

2.2 Interferometric Imaging

The fidelity of reconstructed interferometric images is highly dependent on the the instrument's measurement coverage. Small arrays with few antennas can only sample a small number of uv modes, and the resulting images will have significant imaging artifacts. Large arrays, particularly arrays with antennas in pseudo-random configurations, can measure and reconstruct more uv modes and produce more accurate images (Beardsley et al., 2012). The point spread function (PSF) describes the reconstruction of a single point-like emission source and reflects the measurement artifacts that appear in a reconstructed image.

In this section, we discuss interferometric imaging with the FHD software package (Sullivan et al., 2012; Barry et al., 2019a). FHD's analysis approach follows the "optimal mapmaking" formalism developed in Bhatnagar et al. 2008 and Morales and Matejek 2009 and based on the techniques developed for CMB analyses explored in Tegmark et al. 1998. All analyses presented in this thesis were produced using FHD.

Interferometric imaging generally involves "gridding" visibilities to the uv plane. Gridding consists of placing the measured visibility values in the uv plane at the locations sampled by each measurement. The visibilies are smeared over a region of uv space with a gridding kernel. Optimal mapmaking is a unique imaging approach that uses the beam response as the gridding kernel. Barry et al. 2019b showed that a modified gridding kernel that slightly differs from the beam response can produce better cosmological measurements, however that work is beyond the scope of this thesis.

FHD supports three uv weighting schemes, called "natural" weighting, "optimal" weighting, and "uniform" weighting. (Note that FHD's "natural" and "uniform" weighting are *not* equivalent to those of the widely-used radio imaging software Common Astronomy Software Applications, or CASA.[4]) Here we document each of those weightings.

2.2.1 Natural Weighting

Under FHD's natural weighting scheme, the sky signal is estimated as follows:

$$\hat{S}(\boldsymbol{\theta}) = \frac{1}{C|B_{\text{avg}}(\boldsymbol{\theta})|^2} \mathcal{F}\mathcal{T}^{-1} \left[\sum_{jk} \tilde{B}^*_{jk}(\boldsymbol{x}_{jk} - \boldsymbol{x})\, v_{jk} \right] . \tag{2.16}$$

---

[4] `https://casa.nrao.edu/`

Here $\mathcal{FT}^{-1}$ denotes the inverse Fourier transform operator and $\hat{S}(\boldsymbol{\theta})$ denotes the estimate of the sky signal $S(\boldsymbol{\theta})$. $B_{\mathrm{avg}}(\boldsymbol{\theta})$ represents the average baseline response, averaged across all baselines. C is a constant factor that can account for the measurement coverage in the *uv* plane. We discuss this constant in detail in §2.2.4.

Note that here $B_{\mathrm{avg}}(\boldsymbol{\theta})$ and $B_{jk}(\boldsymbol{\theta})$ correspond to estimates of the true beam. Any physical beam model experiences errors relative to the true instrumental response, but precise analyses require that those errors are small. As a result, beam modeling is an active area of research.

In Equation 2.16, the bracketed expression is simply the gridded visibilities. Each visibility is multiplied by the baseline response and added to the *uv* plane, which FHD typically pixelates at half wavelength spacing. The result is Fourier transformed to produce an image of the sky weighted by two factors of the beam.

Natural weighting is perhaps the simplest analysis approach, but it is not the most accurate. The gridded *uv* plane represents the sum of all visibility contributions, not the average. The *uv* plane values are therefore highly dependent on the instrument's measurement coverage.

2.2.2 Optimal Weighting

Under optimal weighting, we divide the gridded *uv* plane by the weights to produce a weighted average of the *uv* plane. The weights are equivalent gridding all visibilities as 1. The resulting estimate of the sky is given by

$$\hat{S}(\boldsymbol{\theta}) = \frac{1}{C B_{\mathrm{avg}}(\boldsymbol{\theta})} \mathcal{FT}^{-1} \left[\frac{\sum_{jk} \tilde{B}^*_{jk}(\boldsymbol{x}_{jk} - \boldsymbol{x})\, v_{jk}}{\sum_{jk} \tilde{B}^*_{jk}(\boldsymbol{x}_{jk} - \boldsymbol{x})} \right]. \tag{2.17}$$

Note that we set the bracketed expression to zero for *uv* locations where the denominator is zero.

Regions of the *uv* plane where the weights are zero correspond to holes in the instrument's *uv* coverage. Those modes are not measured and cannot be reconstructed. This can produce artifacts in the reconstructed image.

2.2.3 Uniform Weighting

Uniform weighting is an analysis technique designed to reduce imaging artifacts from holes in the instrumental *uv* coverage. Under uniform weighting, we produce a simple average of the *uv* plane, in contrast to the weighted average produced in optimal weighting.

Under uniform weighting, the estimate of the sky signal is

$$\hat{S}(\boldsymbol{\theta}) = \frac{1}{C|B_{\text{avg}}(\boldsymbol{\theta})|^2}\mathcal{FT}^{-1}\left[\frac{\sum_{jk}\tilde{B}^*_{jk}(\boldsymbol{x}_{jk}-\boldsymbol{x})\,v_{jk}}{W(\boldsymbol{x})}\right], \tag{2.18}$$

where $W(\boldsymbol{x})$ is a *uv* weighting function equal to the number of visiblity measurements that contribute to each *uv* pixel. More specifically, we can define a function $P_{jk}(\boldsymbol{x})$ for each baseline $\{j, k\}$ that corresponds to the baseline's footprint in the *uv* plane:

$$P_{jk}(\boldsymbol{x}) = \begin{cases} 1, & B_{jk}(\boldsymbol{x}) \neq 0 \\ 0, & B_{jk}(\boldsymbol{x}) = 0 \end{cases}. \tag{2.19}$$

Then the weighting function is given by

$$W(\boldsymbol{x}) = \sum_{jk} P_{jk}(\boldsymbol{x}). \tag{2.20}$$

As with optimal weighting, we set the bracketed expression in Equation 2.18 to zero for *uv* locations where the $W(\boldsymbol{x}) = 0$.

2.2.4 Normalization Factor

Each weighting scheme presented in Equations 2.16, 2.17, and 2.18 contains a normalization factor C. This factor is set to normalize the reconstructed images appropriately for various science applications. Interferometric images have one of two normalization schemes: point source normalization and diffuse normalization. The normalization schemes are appropriate for measuring different classes of emission.

Point Source Normalization

Point source normalization is appropriate when imaging bright, compact sources. These sources are sufficiently small as to be unresolved by the instrument and are therefore well-approximated as point sources.

Compact sources produce power throughout the *uv* plane. Gaps in the instrument's measurement of the *uv* plane mean that the sources cannot be fully reconstructed. Point source normalization ensures that bright, compact sources have the correct peak normalization by accounting for measurement gaps in the *uv* plane that suppress the sources' reconstructed intensities. The normalization produces images where point sources have correct normalization in units of flux density, generally reported as janskys (Jy)[5].

The normalization factor C is calculated by considering the measurement of a single point source of unit intensity at the pointing center. This source yields visibility measurements of 1 for every baseline. We can imagine imaging those visibilities and calculating the source's intensity to derive the normalization factor.

For natural weighting (§2.2.1) we get

$$C = \frac{1}{|B_{\text{avg}}(\boldsymbol{\theta}=0)|^2} \mathcal{FT}^{-1}\left[\sum_{jk} \tilde{B}^*_{jk}(\boldsymbol{x}_{jk}-\boldsymbol{x})\right](\boldsymbol{\theta}=0). \tag{2.21}$$

Here $\boldsymbol{\theta}=0$ represents the pointing center. $\mathcal{FT}^{-1}\left[f(\boldsymbol{x})\right](\boldsymbol{\theta}=0)$ denotes transforming the quantity $f(\boldsymbol{x})$ from the *uv* plane to the sky and taking its value at the pointing center. For optimal weighting (§2.2.2) this becomes

$$C = \frac{1}{B_{\text{avg}}(\boldsymbol{\theta}=0)} \mathcal{FT}^{-1}\left[\frac{\sum_{jk} \tilde{B}^*_{jk}(\boldsymbol{x}_{jk}-\boldsymbol{x})}{\sum_{jk} \tilde{B}^*_{jk}(\boldsymbol{x}_{jk}-\boldsymbol{x})}\right](\boldsymbol{\theta}=0), \tag{2.22}$$

and for uniform weighting (§2.2.3) we get

$$C = \frac{1}{|B_{\text{avg}}(\boldsymbol{\theta}=0)|^2} \mathcal{FT}^{-1}\left[\frac{\sum_{jk} \tilde{B}^*_{jk}(\boldsymbol{x}_{jk}-\boldsymbol{x})}{W(\boldsymbol{x})}\right](\boldsymbol{\theta}=0), \tag{2.23}$$

---

[5] In SI units, one jansky is equal to 10^{-26} watts per square meter per hertz (1 Jy = 10^{-26} W m^{-2} Hz^{-1}).

where $W(\boldsymbol{x})$ is given by Equation 2.20. In Equations 2.22 and 2.23 the bracketed quantity is taken to be zero when the denominator is equal to zero.

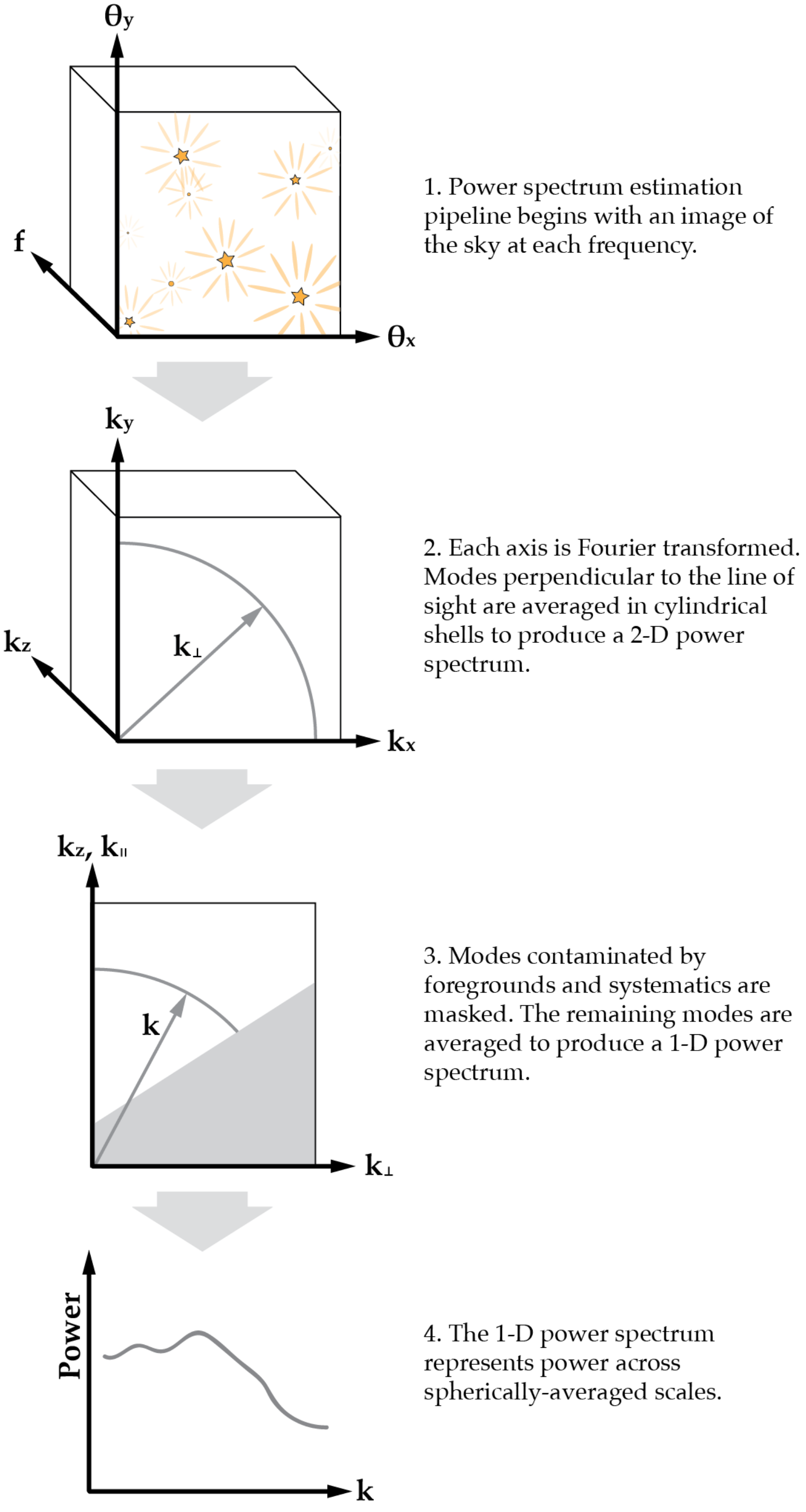
Diffuse Normalization

Diffuse normalization sets $C = 1$ for all weighting schemes. It is used when imaging large-scale structure. Under diffuse normalization, we retain the appropriate normalization of the emission's angular power spectrum. In other words, Fourier transforming the sky signal yields and estimate of the true *uv* plane. This reconstructed *uv* plane will be accurate only where the instrument measures. If the *uv* plane is fully sampled, diffuse normalization delivers an image that is properly normalized in surface brightness units such as janskys per sterradian (Jy/sr) or Kelvin.

2.3 The Power Spectrum Measurement

The goal of 21 cm cosmology is to measure the size and distribution of neutral hydrogen in the early universe. We aim to map three-dimensional regions of neutral hydrogen by measuring its extent along the two axes perpendicular to the line-of-sight, corresponding to directions on the sky, and the line-of-sight axis, corresponding to redshift or the observed frequency of the signal.

It is infeasible to directly image the faint 21 cm cosmological signal, so we instead intend to make a statistical detection of the signal's power spectrum. The power spectrum is derived by, in essence, Fourier transforming all three measurement axes and averaging the resulting power spectrum cube in spherical shells. The ϵPPSILON software package (Jacobs et al., 2016; Barry et al., 2019a) calculates a fully error-propagated measurement of the cosmological power spectrum from imaged interferometric data. ϵPPSILON's implementation follows the imaging power spectrum approach, and we note that this is distinct from the delay power spectrum approach, another method of power spectrum estimation widely used in the field (Morales et al. 2019 explores the distinctions between these two approaches). In this section we provide a qualitative description of the ϵPPSILON power spectrum pipeline.



1. Power spectrum estimation pipeline begins with an image of the sky at each frequency.

2. Each axis is Fourier transformed. Modes perpendicular to the line of sight are averaged in cylindrical shells to produce a 2-D power spectrum.

3. Modes contaminated by foregrounds and systematics are masked. The remaining modes are averaged to produce a 1-D power spectrum.

4. The 1-D power spectrum represents power across spherically-averaged scales.

Figure 2.1: Graphic representation of the power spectrum estimation pipeline.

Figure 2.1 provides a graphical representation of the imaging power spectrum pipeline. The pipeline begins with a 3-D image cube, consisting of a 2-D image of the sky across frequencies. The image cube could be produced from a single observation or many combined observations.

First, we Fourier transform the image cube along all three axes and square to produce a 3-D power spectrum cube. The Fourier dual quantities are converted into cosmological units of inverse comoving megaparsecs (Mpc^{-1}). Power spectrum modes in this thesis are reported in units of $h\,\mathrm{Mpc}^{-1}$, where $h = (3000\,\mathrm{Mpc})H_0/c$ is a dimensionless cosmological constant. Here H_0 is the Hubble constant in the present epoch and c is the speed of light; in our analysis, we let $h = 0.71$. For more information about the conversion to cosmological units, see Hogg 1999 and Morales and Hewitt 2004. The Fourier transforms of the two angular axes are denoted k_x and k_y; the Fourier transform of the frequency axis is denoted k_z or equivalently k-parallel or $k_{||}$ because it corresponds to modes parallel to the line of sight.

Next, we can imagine averaging k_x and k_y together in cylindrical shells. We call this averaged angular power spectrum k-perpendicular or $k_\perp$ because it represents modes perpendicular to the line of sight. Together $k_\perp$ and $k_{||}$ form a 2-D power spectrum. Plots of this power spectrum are important analysis diagnostics, and throughout this thesis we will refer to 2-D power spectrum plots to explore and compare analyses.

The 2-D power spectrum is key to separating the cosmological signal from the bright astrophysical foregrounds. Because of its origin as synchrotron radiation, foreground emission is spectrally very smooth. This is in contrast to the 21 cm signal from the EoR, which we expect to vary across frequency as a result of variable emission intensity along the line of sight. Foregrounds therefore intrinsically occupy the lowest line of sight power spectrum mode, $k_{||} = 0$, while the cosmological signal extends to high $k_{||}$ modes. If the foreground signal were perfectly reconstructed, excluding the $k_{||} = 0$ mode from analysis would eliminate foreground contamination of the cosmological signal.

However, this model is complicated by the instrument's frequency-dependent response. A baseline's response to emission from an off-axis source depends on the frequency of ra-

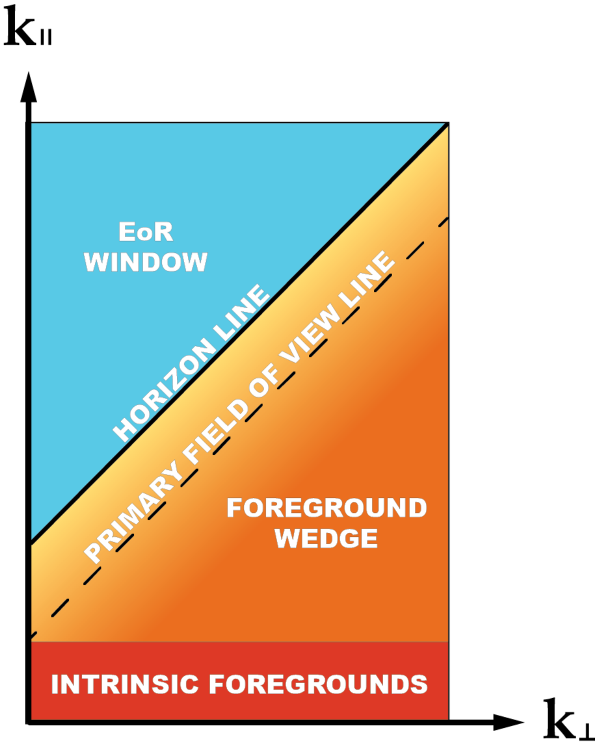


Figure 2.2: Schematic of the 2-D power spectrum. The horizontal axis represents $k_\perp$, power spectrum modes perpendicular to the line of sight. Large angular scales on the sky, measured by an array's short baselines, appear on the left side of the plot, while small angular scales, measured by long baselines, appear on the right side. The vertical axis represents $k_\|$, the Fourier transform of frequency. For the narrow 21 cm emission line, $k_\|$ maps to the power spectrum of structure parallel to the line of sight. Foreground emission is intrinsically spectrally smooth and occupies the lowest $k_\|$ mode. However, the instrument's frequency-dependent response couples foreground power into higher $k_\|$ modes occupying the "foreground wedge." Foreground emission on the horizon couples into a maximum $k_\|$ mode at the solid "horizon line", while emission from the edge of the primary field of view extends only to the dashed "primary field of view" line. The "EoR window" refers to power spectrum modes outside of the foreground wedge, where we are sensitive to the cosmological 21 cm signal.

diation. Because of information loss in the measurement process—owing to the antennas' finite size and discrete sampling of the *uv* plane—this frequency dependence cannot be fully reconstructed and mitigated. As a result, the foregrounds' intrinsically spectrally smooth power couples into higher line of sight modes, $k_\parallel > 0$. Longer baselines' responses are more sensitive to frequency variations and couple foreground power into larger $k_\parallel$ modes than short baselines, producing a wedge-shaped feature in the 2-D power spectrum called the foreground wedge and explored in depth in the literature (Furlanetto et al., 2006; Datta et al., 2010; Morales and Wyithe, 2010; Vedantham et al., 2012; Morales et al., 2012; Parsons et al., 2012; Trott et al., 2012; Dillon et al., 2013; Hazelton et al., 2013; Thyagarajan et al., 2013, 2015). Figure 2.2 presents a schematic of this feature.

2-D power spectrum modes that lie outside the foreground wedge are called the "EoR window." These modes are sensitive to the cosmological signal and are the only modes we use to produce 1-D power spectrum measurements. We estimate the cosmological power spectrum my masking the 2-D power spectrum modes contaminated by foreground emission and averaging the remaining modes in rings across $k_\perp$ and $k_\parallel$. At this stage, we may mask additional modes in the 2-D power spectrum that are contaminated by systematic error. For example, frequency structure in the instrumental response, such as the MWA's coarse band structure, can render particular $k_\parallel$ modes unusable (Beardsley et al., 2016; Barry et al., 2019b; Li et al., 2019). We also exclude poorly-measured $k_\perp$ modes where our specific array layout does not provide good *uv* plane measurement coverage. (Note that, in practice, ϵppsilon calculates the 1-D power spectrum directly from the 3-D power spectrum cube with appropriate masking. However, the masked contaminated modes are more naturally defined in the 2-D power spectrum space described here.)

By excising sources of measurement contamination in the 2-D power spectrum, we can produce a 1-D power spectrum measurement that is sensitive to the faint cosmological signal. A measurement of this cosmological power spectrum would illuminate the distribution of neutral hydrogen in a particular cosmological epoch, or redshift range, as dictated by the frequency range of the input image cube. However, to date low-level measurement error has precluded detection of this signal. The challenge of the field is to mitigate even extremely faint sources of contamination of the EoR window power spectrum modes, enabling detection of the EoR's power spectrum.

2.4 Interferometric Calibration

Calibration is the process of constraining the instrument's response to a sky signal. It is a critical aspect of all radio interferometry, but it is especially important for 21 cm cosmology because of the measurement's precision requirements. Even low-level calibration errors can quickly preclude detection of the cosmological signal.

In general, calibration consists of determining the antenna gains, $\boldsymbol{g}$. The gains consist of a complex number per antenna, frequency channel, time step, and polarization mode. The amplitude of the gain corresponds to the sensitivity of the antenna: in the convention used in this thesis, large gain amplitudes indicate that the antenna amplifies the incident signal with respect to the expected response and small gain amplitudes indicate that the antenna suppresses the signal. The complex phase corresponds to an antenna's precise timing, and offsets in the phase of the gains indicate early or delayed wavefront detection.

The gains are related to the visibilities via the measurement equation:

$$v_{jk} = g_j g_k^* u_{jk} + n_{jk} \tag{2.24}$$

(Hamaker et al., 1996). Here v_{jk} is the visibility formed by correlating signals from antennas j and k, g_j is the gain of antenna j, and g_k is the gain of antenna k. u_{jk} is the theoretical true sky visibility, and n_{jk} is the noise. Here we have excluded explicit frequency-, time-, and polarization-dependence. For further discussion of frequency-dependent calibration see §4.4. §6.4 discusses calibration's generalization to the fully polarized case.

Calibration is the process of calculating $\boldsymbol{g}$ from the data and an estimation of the true visibilities $\boldsymbol{u}$. While there are many approaches to calibration, most calibration implementations to date can be broadly classified in one of two categories: sky-based calibration and redundant calibration. In this section we introduce those two classes of calibration. Chapter

4 expands upon these calibration methods by presenting a new calibration approach called unified calibration. Unified calibration enables a physically-motivated and highly flexible calibration framework that combines aspects of sky-based and redundant calibration.

2.4.1 Sky-Based Calibration

Sky-based calibration estimates $\boldsymbol{u}$ from an estimate of the sky signal and instrumental response. The sky model generally consists of an extensive catalog of compact radio sources and may also include models of diffuse structure. We propagate this sky model through a simulation pipeline that includes detailed models of the instrumental response to produce model visibilities $\boldsymbol{m}$, and we then assume that $\boldsymbol{u} \approx \boldsymbol{m}$.

From the measurement equation, we then get that $v_{jk} \approx g_j g_k^* m_{jk}$. The gains are calculated from a cost function that quantifies the discrepancy between these values:

$$L(\boldsymbol{g}) = \sum_{jk} |v_{jk} - g_j g_k^* m_{jk}|^2. \tag{2.25}$$

The fit gain values $\hat{\boldsymbol{g}}$ are the values of $\boldsymbol{g}$ that minimize $L(\boldsymbol{g})$. Note that this cost function is often called a chi-squared or χ^2; for generality, we use the symbol L. In Chapter 4 we present a generalized calibration formalism for which L is not necessarily equivalent to a chi-squared.

We can calculate the minimum of $L(\boldsymbol{g})$ analytically by taking derivatives with respect to the gains. We find that, excluding the autocorrelation antennas (using only data $\boldsymbol{v}_{jk}$ where $j \neq k$),

$$\frac{\partial L(\boldsymbol{g})}{\partial \operatorname{Re}(g_k)} = 4 \sum_{j \neq k} \left[\operatorname{Re}(g_k)|g_j|^2 |m_{kj}|^2 - \operatorname{Re}(v_{kj}^* g_j^* m_{kj}) \right] \tag{2.26}$$

and

$$\frac{\partial L(\boldsymbol{g})}{\partial \operatorname{Im}(g_k)} = 4 \sum_{j \neq k} \left[\operatorname{Im}(g_k)|g_j|^2 |m_{kj}|^2 + \operatorname{Im}(v_{kj}^* g_j^* m_{kj}) \right]. \tag{2.27}$$

The extremum condition states that $\left.\frac{\partial L(\boldsymbol{g})}{\partial \operatorname{Re}(g_k)}\right|_{\boldsymbol{g}=\hat{\boldsymbol{g}}} = 0$ and $\left.\frac{\partial L(\boldsymbol{g})}{\partial \operatorname{Im}(g_k)}\right|_{\boldsymbol{g}=\hat{\boldsymbol{g}}} = 0$ for all antennas k. We thereby get that

$$\hat{g}_k = \frac{\sum_{j \neq k} \hat{g}_j m_{jk} v_{jk}^*}{\sum_{j \neq k} |\hat{g}_j|^2 |m_{jk}|^2}. \tag{2.28}$$

For an array with N antennas, this represents a system of N coupled nonlinear equations. Calibration consists of solving this system of equations, typically with an iterative method such as gradient descent (Salvini and Wijnholds, 2014).

Minimizing Equation 2.25 yields degenerate solutions. We note that adding a constant phase term Δ to all gains, such that $\boldsymbol{g} \rightarrow \boldsymbol{g}e^{i\Delta}$, leaves $L(\boldsymbol{g})$ unchanged for any real value of Δ. The overall phase of the gains, which corresponds to the absolute timing of incident radiation on the array, is therefore a degenerate parameter under this calibration approach and must be constrained in some other way. (See §4.1.2 for a more formal discussion of calibration degeneracies.)

A simple way to constrain the overall phase of the gains is to use a reference antenna. One then requires that $\mathrm{Arg}[\hat{g}_{\mathrm{ref}}] = 0$ where $\hat{g}_{\mathrm{ref}}$ is the maximum-likelihood gain of the reference antenna and Arg denotes the complex phase. However, this means that systematic errors in the phase of the reference antenna affect all antennas. More sophisticated calibration approaches mitigate calibration errors in the reference antenna to better constrain the overall phase of the gains (Barry et al., 2019b; Li et al., 2019) or use an external time-variable signal such as a pulsar (Pen et al., 2009).

Sky-based calibration is susceptible to error if the model visiblities $\boldsymbol{m}$ are incorrect. This arises when the sky model is inaccurate or incomplete (Barry et al., 2016; Ewall-Wice et al., 2016), an effect we explore in Chapter 3. Errors also occur when the instrument model that maps the sky model to the model visiblities $\boldsymbol{m}$ is incorrect. In practice, all sky and instrument models have limited accuracy, and the challenge is to suppress these errors to levels that are acceptable for the EoR measurement.

2.4.2 Redundant Calibration

Redundant calibration is a calibration approach applicable to arrays with antennas in a regular grid pattern (Wieringa, 1992; Liu et al., 2010; Dillon and Parsons, 2016; Dillon et al., 2018; Grobler et al., 2018; Kern et al., 2020; Dillon et al., 2020). These regular arrays have many repeated baselines measurements, and these repeated measurements can be leveraged to further constrain the estimate of $\boldsymbol{u}$.

In redundant calibration, we assume that $v_{jk} \approx g_j g_k^* u_a$ where j and k index antennas and a represents the redundant baseline set that contains baseline $\{j, k\}$. We then get a cost function of the form

$$L(\boldsymbol{g}, \boldsymbol{u}) = \sum_a \sum_{\{j,k\} \text{ in } a} |v_{jk} - g_j g_k^* u_a|^2. \tag{2.29}$$

Here $\sum_{\{j,k\} \text{ in } a}$ represents the sum over all baselines $\{j, k\}$ that belong to redundant set a. Calibration consists of calculating the values $\hat{\boldsymbol{g}}$ and $\hat{\boldsymbol{u}}$ that minimize $L(\boldsymbol{g}, \boldsymbol{u})$. Although $\hat{\boldsymbol{u}}$ are fit in calibration, only $\hat{\boldsymbol{g}}$ are applied to the data to produce calibrated visibilities.

As for sky-based calibration, we can derive analytic expressions for the fit calibration parameters $\hat{\boldsymbol{g}}$ and $\hat{\boldsymbol{u}}$ by calculating derivatives of $L(\boldsymbol{g}, \boldsymbol{u})$ from Equation 2.29. We find that

$$\frac{\partial L(\boldsymbol{g}, \boldsymbol{u})}{\partial \operatorname{Re}(g_k)} = 4 \sum_{j \neq k} \left[\operatorname{Re}(g_k) |g_j|^2 |u_{kj}|^2 - \operatorname{Re}(v_{kj}^* g_j^* u_{kj}) \right] \tag{2.30}$$

and

$$\frac{\partial L(\boldsymbol{g}, \boldsymbol{u})}{\partial \operatorname{Im}(g_k)} = 4 \sum_{j \neq k} \left[\operatorname{Im}(g_k) |g_j|^2 |u_{kj}|^2 + \operatorname{Im}(v_{kj}^* g_j^* u_{kj}) \right]. \tag{2.31}$$

Solving the extremum conditions $\left.\frac{\partial L(\boldsymbol{g},\boldsymbol{u})}{\partial \operatorname{Re}(g_k)}\right|_{\boldsymbol{g}=\hat{\boldsymbol{g}},\boldsymbol{u}=\hat{\boldsymbol{u}}} = 0$ and $\left.\frac{\partial L(\boldsymbol{g},\boldsymbol{u})}{\partial \operatorname{Im}(g_k)}\right|_{\boldsymbol{g}=\hat{\boldsymbol{g}},\boldsymbol{u}=\hat{\boldsymbol{u}}} = 0$, we find that calibration solutions must satisfy

$$\hat{g}_k = \frac{\sum_{j \neq k} \hat{g}_j \hat{u}_{jk} v_{jk}^*}{\sum_{j \neq k} |\hat{g}_j|^2 |\hat{u}_{jk}|^2}, \tag{2.32}$$

which is simply Equation 2.28 where we have replaced the model visibilities $\boldsymbol{m}$ with the fit visibilities $\hat{\boldsymbol{u}}$.

Taking derivatives with respect to the visibility terms $\boldsymbol{u}$, we get

$$\frac{\partial L(\boldsymbol{g}, \boldsymbol{u})}{\partial \operatorname{Re}(u_a)} = 2 \sum_{\{j,k\} \text{ in } a} \left[|g_j|^2 |g_k|^2 \operatorname{Re}(u_a) - \operatorname{Re}(v_{jk}^* g_j g_k^*) \right] \tag{2.33}$$

and

$$\frac{\partial L(\boldsymbol{g}, \boldsymbol{u})}{\partial \operatorname{Im}(u_a)} = 2 \sum_{\{j,k\} \text{ in } a} \left[|g_j|^2 |g_k|^2 \operatorname{Im}(u_a) + \operatorname{Im}(v_{jk}^* g_j g_k^*) \right]. \tag{2.34}$$

Solving $\left.\frac{\partial L(\boldsymbol{g},\boldsymbol{u})}{\partial\,\mathrm{Re}(u_a)}\right|_{\boldsymbol{g}=\hat{\boldsymbol{g}},\boldsymbol{u}=\hat{\boldsymbol{u}}} = 0$ and $\left.\frac{\partial L(\boldsymbol{g},\boldsymbol{u})}{\partial\,\mathrm{Im}(u_a)}\right|_{\boldsymbol{g}=\hat{\boldsymbol{g}},\boldsymbol{u}=\hat{\boldsymbol{u}}} = 0$, we get

$$\hat{u}_a = \frac{\sum_{\{j,k\}\text{ in }a} v_{jk}\hat{g}_j^*\hat{g}_k}{\sum_{\{j,k\}\text{ in }a} |\hat{g}_j|^2|\hat{g}_k|^2}. \tag{2.35}$$

The calibration solutions that minimize Equation 2.29 experience the same degeneracy in the overall phase of the gains seen in sky-based calibration: the transformation $\boldsymbol{g} \rightarrow \boldsymbol{g}e^{i\Delta}$ leaves $L(\boldsymbol{g},\boldsymbol{u})$ unchanged. However, redundant calibration also includes additional degeneracies. The solutions are degenerate in the overall amplitude of the gains: consider that $L(\boldsymbol{g},\boldsymbol{u})$ is invariant under the transformations $\boldsymbol{g} \rightarrow A\boldsymbol{g}$ and $\boldsymbol{u} \rightarrow \frac{1}{A^2}\boldsymbol{g}$. Errors in the overall gain amplitude A make the sky appear artificially bright or dim.

Finally, the solutions that minimize Equation 2.29 are degenerate in the phase gradient of the gains. Assuming a co-planar array, $L(\boldsymbol{g},\boldsymbol{u})$ is invariant under the transformations $g_j \rightarrow g_j e^{i(\Delta_x x_j + \Delta_y y_j)}$ and $u_a \rightarrow u_a e^{-i(\Delta_x x_a + \Delta_y y_a)}$. Here (x_j, y_j) are the coordinates of antenna j and $(x_a, y_a) = (x_j - x_k, y_j - y_k)$, where baseline $\{j,k\}$ belongs to redundant set a. Errors in the phase gradient parameters Δ_x and Δ_y shift the sky image such that sources appear offset from their true positions.

In total, calibration solutions that minimize Equation 2.29 are degenerate in a minimum of four parameters per time step, frequency, and polarization: the overall amplitude of the gains, A; the overall phase of the gains, Δ; and the x- and y- components of the phase gradient, Δ_x and Δ_y. Additional degeneracies arise in special cases. For example, arrays that contain multiple redundant sub-arrays can have more than four degenerate parameters. Each of the hexagonal sub-arrays in the MWA Phase II has an independent overall phase degeneracy (Li et al., 2018).

In light of these degeneracies, redundant calibration must include an additional calibration step, called absolute calibration, that constrains these parameters. Minimizing Equation 2.29 alone is called relative calibration (Zheng et al., 2014; Byrne et al., 2019; Kern et al., 2020), and it does not yield physical calibration solutions that can be applied to data. The relatively calibrated gains that have not had absolute calibration applied are sometimes called $\hat{\boldsymbol{h}}$ to distinguish them from the true gains $\hat{\boldsymbol{g}}$. There are several approaches to absolute calibration (see Chapter 3 and Appendix A), but all are similar to sky-based calibration in that they constrain the overall amplitude and phase gradient terms (A, Δ_x, and Δ_y) from model visiblities $\boldsymbol{m}$ derived from a sky model propagated through an visibility simulator. The overall phase of the gains Δ is constrained in absolute calibration using the same techniques used for sky-based calibration and described above in §2.4.1.

Redundant calibration assumes that the measurements from redundant baselines are identical. Antenna position errors or response inhomogeneities—for example, if the beam response from one antenna has a different shape than that of another—break this assumption and introduce errors in the relative calibration step of redundant calibration. Furthermore, absolute calibration is susceptible to errors from inaccurate or incomplete sky and instrument models, much like sky-based calibration. We explore the effect of sky model error on redundant calibration in Chapter 3.

2.4.3 Direction-Dependent Calibration

In both the sky-based and redundant calibration approaches described above, we assume that the shape of the antenna responses across the sky are well-modeled *a priori*. This is called direction-independent calibration. In practice, unmodeled direction-dependent effects can degrade interferometric performance. Direction-dependent calibration is a more flexible alternative that can capture uncertainties in the spatial structure of the antenna responses (Kazemi et al., 2011, 2013; Patil et al., 2016; van Weeren et al., 2016; Tasse et al., 2018; Albert et al., 2020; Mertens et al., 2020). Direction-dependent calibration is largely beyond the scope of this thesis, but for completeness we summarize it in this section.

In direction-dependent calibration, the gains can take on different values for different directions on the sky. In typical implementations, the sky is faceted into discrete sections, and a gain is calculated for each antenna and each sky facet (van Weeren et al., 2016).

Direction-dependent calibration can be a good calibration approach when antenna responses are difficult to model. The direction-dependent gains can fit response shapes that are not well-modeled and vary antenna-to-antenna. Additionally, direction-dependent calibration can be important for modeling ionospheric effects. This is particularly critical for large arrays where antennas have different lines-of-sight through the ionosphere. Refraction through the ionosphere imposes direction-dependent phase distortions that vary in time as ionospheric conditions change (Jordan et al., 2017; Albert et al., 2020).

Direction-dependent, per-frequency calibration increases the calibration degrees-of-freedom in a way that can suppress diffuse foreground emission and the cosmological signal itself, leading to signal loss. One can protect against this signal loss by reducing calibration degrees-of-freedom across frequency (Patil et al., 2016). With any calibration approach, one must quantify and mitigate potential signal loss stemming from coupling of the calibration parameters and the cosmological signal.

2.4.4 Calibration and the Power Spectrum Measurement

Frequency-dependent calibration errors pose a major challenge for 21 cm cosmology. In particular, fast frequency structure in the calibrated gains can contaminate the cosmological signal by limiting the separability of the cosmological signal and the astronomical foregrounds.

Gain-based calibration, as described by the measurement equation in Equation 2.24, assumes that the instrumental response is well-modeled as a multiplicative factor on the true sky signal. Reproducing Equation 2.24 with explicit frequency dependence gives

$$v_j(f) \approx G_j(f) u_j(f) \tag{2.36}$$

where now j indexes baselines. $v_j(f)$ is the measured visibility and $u_j(f)$ is the theoretical true sky signal from baseline j at frequency f. $G_j(f)$ is the instrumental gain, represented here as a single variable and equal to the product of the antenna gains. If calibration is error-free, we can divide the data by the gains to recover the true sky signal: the reconstructed signal $\hat{u}_j(f) = u_j(f)$.

We now consider the Fourier transform of the true sky signal $u_j(f)$ across frequency, which we will denote $\tilde{u}_j(\eta)$. Here η is the Fourier dual of frequency, or delay, with units of time. The foreground signal is compact in this space. Because foregrounds are very spectrally smooth, their signal predominantly lies in the $\eta = 0$ mode. Some foreground signal appears in modes $|\eta| > 0$ due to the foreground wedge effect (see §2.3 and Figure 2.2), but $\tilde{u}_j(\eta)$ is free from foreground contamination at large $|\eta|$, modes that lie outside the foreground wedge.

From the convolution theorem we can rewrite Equation 2.36 in terms of $\tilde{u}_j(\eta)$:

$$\tilde{v}_j(\eta) \approx \tilde{G}_j(\eta) * \tilde{u}_j(\eta). \tag{2.37}$$

Here $*$ denotes the convolution and $\tilde{v}_j(\eta)$ and $\tilde{G}_j(\eta)$ are the frequency Fourier transforms of the visibilities and gains, respectively. We can now consider the case that the gains calculated in calibration have an error that contributes to a particular delay mode η_0. For that mode, $\tilde{G}_j(\eta_0) \to \tilde{G}_j(\eta_0) + \epsilon_j(\eta_0)$, where $\epsilon_j(\eta_0)$ is the error on the gains. Because of the convolution relationship in Equation 2.37, the calibration error at $\eta = \eta_0$ convolves the sky power. The bright foreground power at $\eta = 0$ is replicated, scaled by the amplitude of the gain error, at $\eta = \eta_0$. Similarly, foreground wedge power is coupled into modes near η_0. This contaminates a range of delay modes, preventing detection of the cosmological signal in the EoR window. This issue is compounded when calibration error affects multiple delay modes.

For this reason, it is critical that calibration is essentially free from errors with fast frequency structure. Because the foregrounds are up to 5 orders of magnitude brighter than the cosmological signal, errors in the bandpass calibration on delay modes in the EoR window must be below 1 part in 10^5 (Barry et al., 2016). These calibration errors can take two forms. First of all, we require that calibration does not introduce new fast frequency structure by overfitting. Overfitting could correspond to fitting thermal noise, although this would be mitigated by extending the integration time. More importantly, errors in the calibration sky model introduce fast frequency structure that does not average down in time, and these errors can be fit in calibration (see Chapter 3). Secondly, we require that calibration accurately captures any true frequency structure in the instrumental response. Either class of error will couple foreground power into the EoR window and affect our ability to detect the faint cosmological signal.

Chapter 3

CALIBRATION IN THE PRESENCE OF SKY MODEL ERROR

Errors in the sky model used in calibration introduces errors in the calibration solutions at a level that is important for EoR science and other cosmology measurements. In this chapter we explore the effects of sky model error on calibration performance for each sky-based calibration and redundant calibration. This chapter is largely based on Byrne et al. 2019.

3.1 Sky Model Error and Sky-Based Calibration

Foreground emission is very spectrally smooth, and the sky models used in calibration therefore also have minimal, if any, spectral structure. However, interferometers inherently have a frequency-dependent response. Visibilities simulated from a spectrally smooth sky model will nonetheless have spectral structure. This mechanism gives rise to the foreground wedge, discussed in §2.3 and pictured in Figure 2.2.

It follows that errors in the sky model are frequency-dependent. In the presence of sky model error, calibration consists of fitting the data to a model with spectrally varying error. The resulting calibration solutions will likewise have frequency-dependent error. For typically antenna-based calibration, these errors are not contained to the foreground wedge region of 2-D power spectrum space. Calibration couples sky model errors into power spectrum modes that lie in the EoR window, contaminating the region of power spectrum space in which we expect sensitivity to the EoR.

Barry et al. 2016 explored this effect in simulation. The authors simulated data from the KGS point source catalog (Carroll et al., 2016) using FHD's simulation pipeline. They then performed sky-based calibration on this data, using model visibilities $\boldsymbol{m}$ simulated from an incomplete catalog. The catalog used for the calibration sky model omitted KGS's faintest sources, including only 4,000 of the full catalog's 6,950 sources.

Omitting faint sources from the calibration sky model represents just one class of sky model error. Other sources of error include incorrect source intensities or locations, mismodeling of the structure of extended and diffuse emission, the inclusion of false sources, or the omission of one or more bright sources. However, simulating missing faint sources is a simple and realistic test case that probes the effects of sky model error on calibration performance. In practice, all source catalogs used in calibration have a completeness threshold below which sources are too faint to be modeled, and precision calibration must contend with errors from these faint missing sources.

The simulations in Barry et al. 2016 demonstrate that sky model incompleteness in sky-based calibration produces substantial spectral error and contaminates the EoR window. For typical per-antenna, per-frequency calibration, the level of error simulated overwhelms the expected EoR signal. Barry et al. 2016 showed that this error could be mitigated somewhat by reducing the calibration degrees of freedom. For example, if the antenna and signal chain are known to have smooth frequency responses, fitting the gains' bandpasses with a low-order polynomial reduces calibration error. If antenna responses are quite homogeneous, we can average the gains across antennas. However, this requires that antennas are manufactured to exacting specifications. In recent years this work has driven the design of next-generation EoR experiments such as the Hydrogen Epoch of Reionization Array (HERA), for which there has been significant effort to control the uniformity and spectral features of the antenna responses. Nevertheless, even these newer arrays necessitate per-antenna bandpass calibration, making them vulnerable to the class of calibration errors explored in Barry et al. 2016.

3.2 Sky Model Error and Redundant Calibration

In contrast to sky-based calibration, redundant calibration fits a minority of calibration parameters from a sky model, instead imposing agreement between measurements from redundant baselines (see §2.4.2). Even so, redundant calibration relies on a sky model in the absolute calibration step. As in sky-based calibration, sky model errors produce frequency-dependent calibration errors in redundant calibration. In the remainder of this chapter, we extend Barry et al. 2016 to redundant calibration, exploring the impact of sky model incompleteness on redundant calibration solutions.

We explore this effect by simulating sky model errors for a redundant array of 331 antennas, arranged in a regular hexagonal layout with minimum antenna spacings of 15 m (see Figure 3.1). The simulations use the MWA antenna zenith-pointed beam model in the 167-198 MHz frequency band. One beam model is used for all frequencies across this band to eliminate errors from frequency-dependent beam modulation. For simplicity, we consider only one polarization; all results in this chapter use simulated data from East-West dipole antennas. Visibilities are created for a 2-minute observation with FHD (Sullivan et al., 2012) and are based on the GLEAM catalog (Hurley-Walker et al., 2017) at the "EoR-0" field (centered on Right Ascension 0.00 h and Declination $-27°$) for a total of 51,821 simulated sources with a minimum flux density of 10 mJy. Following the techniques developed in Barry et al. 2016, we then create a calibration catalog from the 4,000 brightest sources in apparent flux density (minimum flux density 89 mJy). By calibrating on only a subset of the simulated catalog, we represent the fact that calibration catalogs are realistically incomplete. The missing sources in the calibration catalog introduce errors in the calibration solutions.

To perform absolute calibration, we first implement sky-based calibration with FHD. This minimizes $L(\boldsymbol{g})$ from Equation 2.25, where the "measured visibilities" $\boldsymbol{v}$ are simulated from 51,821 sources and the "modeled visibilites" $\boldsymbol{m}$ are simulated from 4,000. This gives per-antenna, per-frequency gain solutions $\hat{\boldsymbol{g}}^{\text{sky}}(f)$. Here the superscript "sky" denotes that these are the gains from sky-based calibration. In simulation, the true gains are equal to 1. Deviations from 1 correspond to errors in $\hat{\boldsymbol{g}}^{\text{sky}}(f)$ due to the missing sources in the sky model.

Barry et al. 2016 demonstrated that sky-based calibration gains $\hat{\boldsymbol{g}}^{\text{sky}}(f)$ have frequency-dependent errors when the sky model is missing sources. We are interested in the extent to

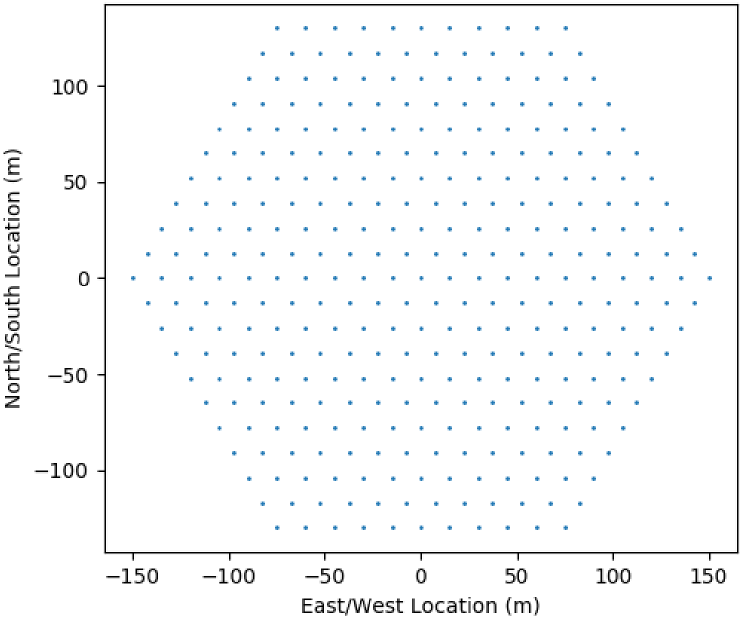


Figure 3.1: Representation of the hexagonal array layout used in simulation. The array has 331 antennas and minimum antenna spacings of 15 m.

which these errors affect the absolute calibration step of redundant calibration. We calculate the overall amplitude of the gains to be

$$\hat{A}(f) = \frac{1}{N_{\text{ant}}} \sum_{j=0}^{N_{\text{ant}}} \left| \hat{g_j}^{\text{sky}}(f) \right|, \tag{3.1}$$

where N_{ant} is the number of antennas. The results are plotted in Figure 3.2 as a function of frequency. Deviations from $\hat{A} = 1$ are calibration errors due to the incompleteness of the sky model. These errors are frequency-dependent. When applied to data, they introduce frequency structure to the intrinsically spectrally smooth foregrounds, coupling their power into the EoR window and obscuring the EoR signal.

We note that the overall amplitude of the gains is biased low for all frequencies. This stems from decoherence between the data and the model visiblities in the presence of sky model error. We explore this effect in Appendix C.

In this simulation we do not consider errors in the overall phase Δ, which is degenerate

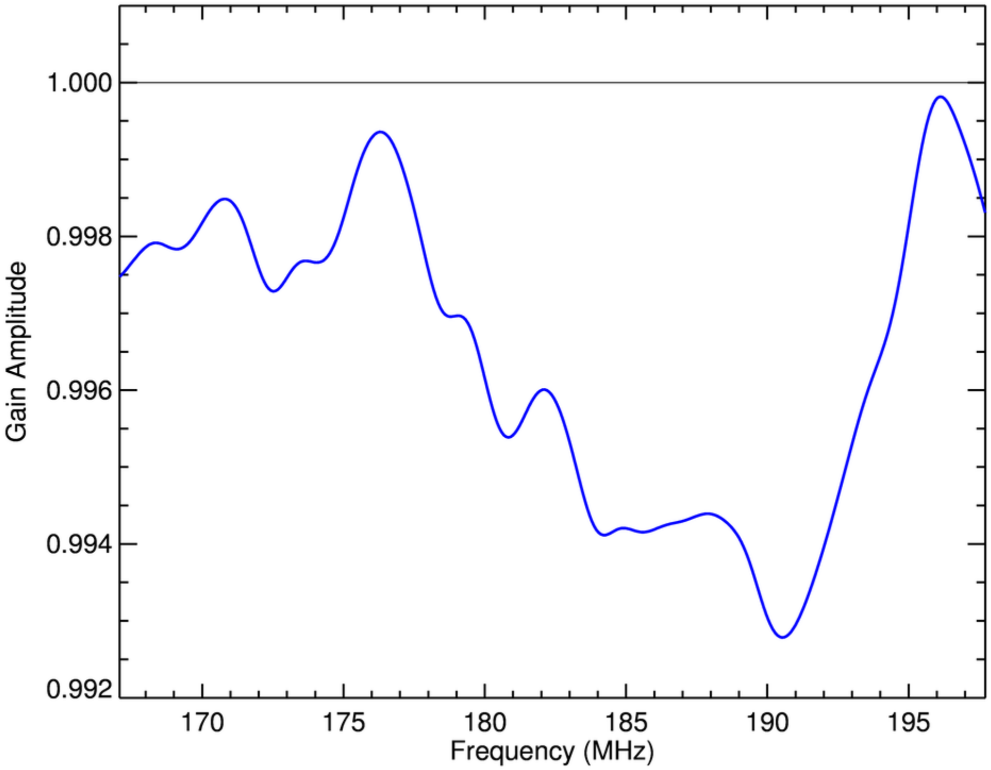


Figure 3.2: Plot of the average amplitude of the per-antenna gains as a function of frequency, $\hat{A}(f)$, for the hexagonal array in Figure 3.1. The visibilities are calibrated to an incomplete catalog that includes only a subset of the simulated sources. Deviations from 1 correspond to calibration errors due to an incomplete sky model. Since the average gain amplitude is a degenerate parameter in the relative calibration step of redundant calibration, these errors persist in redundant calibration even in the limit of perfect redundancy. The specific features of the errors in this parameter depend on the locations and flux densities of the sources missing from the sky model. In this case, those missing sources are faint sources in the EoR-0 field, as described by the GLEAM catalog (Hurley-Walker et al., 2017).

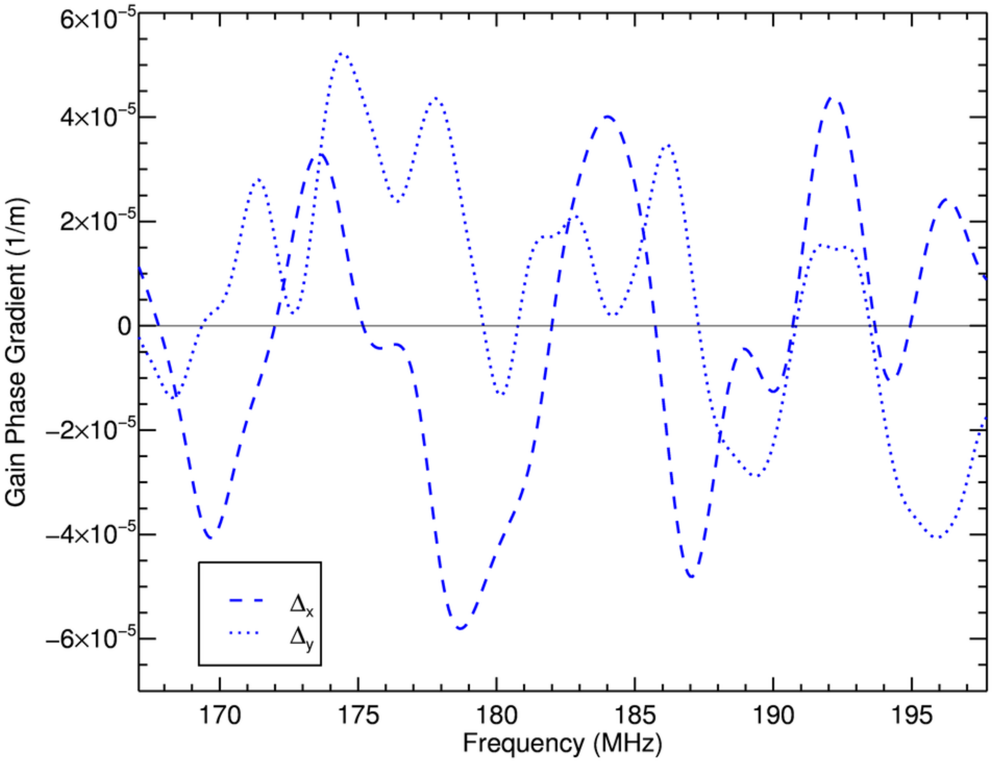


Figure 3.3: Plot of the gain complex phase gradient fit parameters as a function of frequency, $\hat{\Delta}_x(f)$ and $\hat{\Delta}_y(f)$, for the simulation described in Figure 3.2. Here x refers to the East-West direction and y refers to the North-South direction. Deviations from 0 correspond to calibration errors due to an incomplete sky model. Like the average gain amplitude, these two phase gradient parameters are degenerate in relative calibration, so these errors persist in redundant calibration even in the limit of perfect redundancy.

under both sky-based and redundant calibration (see §2.4.1 and §2.4.2). Instead, we require that the average phase of the gains is zero:

$$\frac{1}{N_{\mathrm{ant}}} \sum_{j=0}^{N_{\mathrm{ant}}} \mathrm{Arg}\left[\hat{g}_j{}^{\mathrm{sky}}(f)\right] = 0 \tag{3.2}$$

where $\mathrm{Arg}[\hat{g}_j{}^{\mathrm{sky}}(f)]$ is the complex phase of $\hat{g}_j{}^{\mathrm{sky}}(f)$. We fit the phase gradient parameters by minimizing the expression

$$\chi^2(f) = \sum_{j=0}^{N_{\mathrm{ant}}} \left[\mathrm{Arg}[\hat{g}_j{}^{\mathrm{sky}}(f)] - \Delta_x(f) x_j - \Delta_y(f) y_j\right]^2 \tag{3.3}$$

by varying Δ_x and Δ_y. We assume that $\mathrm{Arg}[\hat{\boldsymbol{g}}^{\mathrm{sky}}(f)] \ll 2\pi$ and therefore do not have to account for the branch cut in the complex plane. The fit values $\hat{\Delta}_x$ and $\hat{\Delta}_y$ are plotted in Figure 3.3. Here deviations from 0 are calibration errors due to the incompleteness of the sky model. As with the overall amplitude, the phase gradient parameter errors are frequency-dependent and therefore couple foreground power into the EoR window.

We note that calculating the absolute calibration parameters directly from the sky-based calibrated gains is a non-standard approach to redundant calibration. In practice, sky-based calibration is rarely performed in tandem with redundant calibration. In Appendix A, we compare the results in this section to those derived from another absolute calibration approach, showing that the differences between the two approaches are negligible.

To demonstrate contamination of the EoR window from the errors in the average gain amplitudes plotted in Figure 3.2 and the gain phase gradient fit terms plotted in Figure 3.3, we produce 2-D power spectra with the ϵPPSILON software package (Jacobs et al., 2016; Barry et al., 2019a). Figure 3.4 gives 2-D PS for three separate simulated arrays. The leftmost column of the figure corresponds to the hexagonal array pictured in Figure 3.1. The middle and rightmost columns correspond to additional array configurations discussed in §3.3. The power spectrum of the sky model used in calibration has been subtracted, producing “residual” power spectra.

In the top left, Figure 3.4(a) gives the 2-D power spectum of simulated visibilities calibrated with absolute calibration errors. The gains applied to these data are $\hat{g}_j = \hat{A} e^{\hat{\Delta}_x x_j + \hat{\Delta}_y y_j}$,

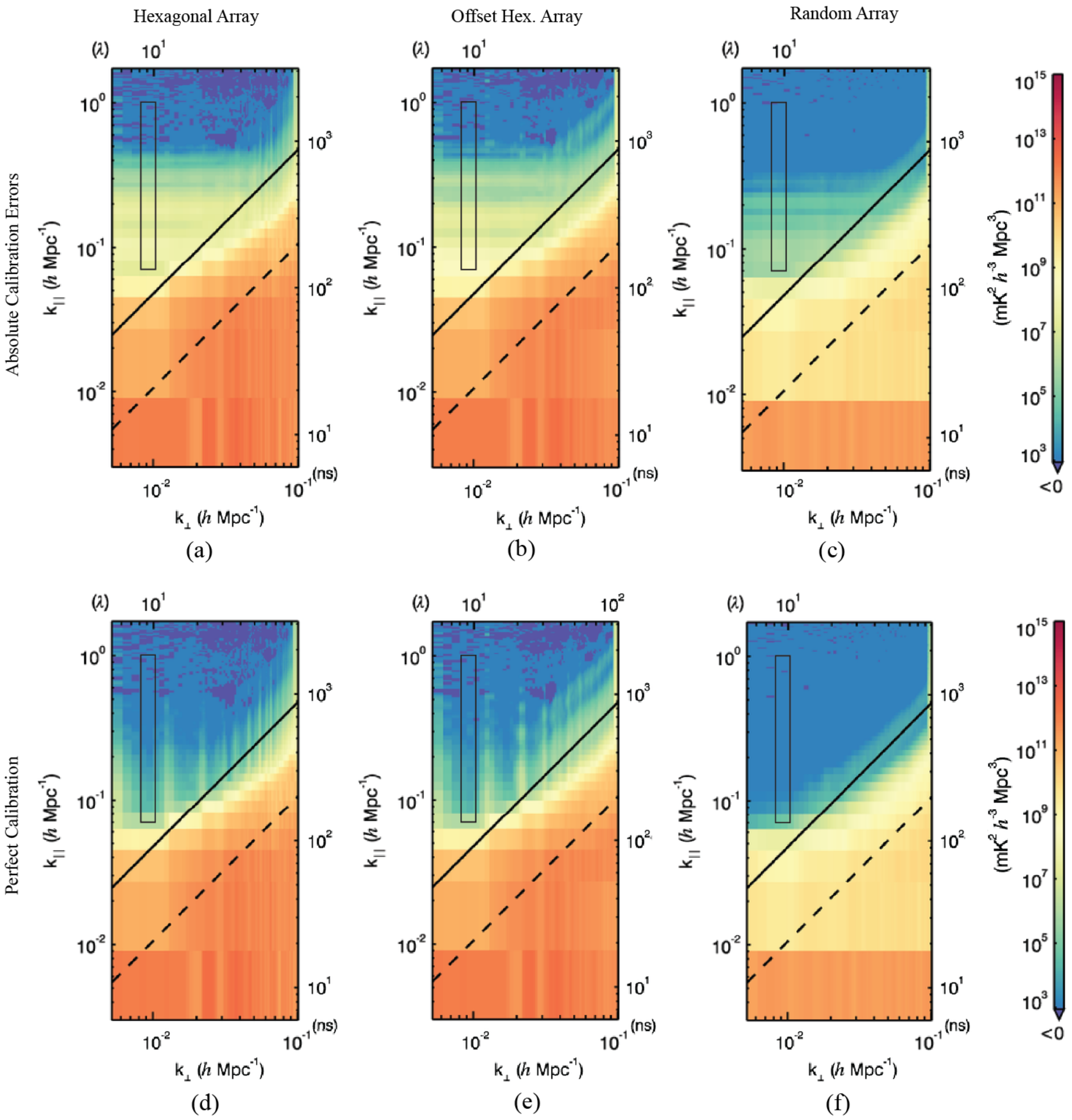


Figure 3.4: 2-D power spectra illustrating simulated calibration errors. (Caption continued.)

Figure 3.4: 2-D "residual" PS of simulated data calibrated to an incomplete sky model. The plots in the top row have been calibrated with errors in the absolute calibration parameters A, the overall gain amplitude, and Δ_x and Δ_y, the gain phase gradient parameters. There are no errors in relative calibration. This corresponds to a regime of perfect array redundancy and no thermal noise, where the only calibration errors emerge from absolute calibration to an incomplete model. The plots in the bottom row have perfect calibration. Notice that errors in the absolute calibration parameters causes power bleed into higher line-of-sight modes. The leftmost column are simulations of the hexagonal array pictured in Figures 3.1 and 3.6(a) and discussed in §3.2. The middle column are simulations from the offset hexagonal array in Figure 3.6(b) and the rightmost column are simulations from the random array pictured in Figure 3.6(c), both discussed in §3.3. Line-of-sight modes ($k_\parallel$) are plotted on the vertical axis and modes perpendicular to the line-of-sight ($k_\perp$) on the horizontal axis. The high power in the lowest line-of-sight ($k_\parallel = 0$) mode represents the intrinsic foregrounds. The red-orange wedge across the lower right part of the spectrum is the "foreground wedge" and comes from the chromatic instrument response, which mixes the intrinsic foregrounds with higher line-of-sight modes (Datta et al., 2010; Morales et al., 2012; Vedantham et al., 2012; Parsons et al., 2012; Trott et al., 2012; Hazelton et al., 2013; Thyagarajan et al., 2013; Dillon et al., 2013; Thyagarajan et al., 2015). The solid and dashed diagonal lines are the "horizon" and "primary field of view" lines, respectively. These denote contamination limits based on sources' off-axis positions. The vertical streaks visible in the upper half of Figures (d) and (e) are regions of low uv coverage. The black rectangular outlines in each plot denote the values that contribute to the 1-D plots in Figures 3.5 and 3.8.

where the parameters $\hat{A}$, $\hat{\Delta}_x$, and $\hat{\Delta}_y$ take the values calculated from Equations 3.1 and 3.3 and plotted in Figures 3.2 and 3.3. Here relative calibration is perfect: calibration errors come from the absolute calibration parameters only. This represents the best possible redundant calibration achievable with the 4,000-source incomplete sky model in the limit of unrealistically perfect redundancy and the absence of noise. Below, Figure 3.4(d) gives the 2-D power spectrum of the same simulated data without calibration, or with $\hat{\boldsymbol{g}} = 1$ for all antennas. In simulation this is equivalent to perfect calibration, or calibration to a complete sky model.

Figures 3.4(a) and 3.4(d) illustrate inherent power spectrum features. First of all, the high power in the lowest line-of-sight mode ($k_{\parallel} = 0$) represents the intrinsic foregrounds. Foreground emission is extremely spectrally smooth; in this simulation, it is perfectly flat. The red-orange wedge across the lower right part of the spectrum is the "foreground wedge" and comes from the chromatic instrument response, which mixes the intrinsic foregrounds with higher line-of-sight modes (Vedantham et al., 2012; Morales et al., 2012; Trott et al., 2012; Parsons et al., 2012; Hazelton et al., 2013). The vertical streaks in the upper half of the spectrum are an intrinsic power spectrum feature resulting from regions of low *uv* coverage (Figure 3.7(d) shows the *uv* coverage of this array). The periodicity of these streaks emerges from the regular layout of antennas in the hexagonal array. Averaging over sufficiently long time intervals can mitigate this effect by leveraging the Earth's rotation to fill in areas of low *uv* coverage through a process called "*uv* rotation."

Comparing the uncalibrated power spectrum simulation in Figure 3.4(d) to the power spectrum simulation with absolute calibration errors in Figure 3.4(a) shows that missing sources in the sky model introduce errors that cause power leakage into high line-of-sight power spectrum modes. Even with perfect relative calibration, the frequency-dependent errors in the absolute calibration parameters plotted in Figures 3.2 and 3.3 introduce frequency structure into the spectrally smooth foregrounds. This frequency structure results in foreground power leakage into the power spectrum modes in the EoR-sensitive window, obscuring the faint EoR signal.

The power leakage in the EoR window falls off at large $k_{\|}$ values. The maximum contaminated $k_{\|}$ mode is proportional to the length of the array's longest baseline. For a widefield array sensitive to emission at the horizon, power leakage occurs at a maximum mode of b/c, where b is the length of the longest baseline used in calibration and c is the speed of light. Converting to cosmological units (Morales and Hewitt, 2004), we expect calibration errors to produce power spectrum contamination for the hexagonal array pictured in Figure 3.1 on modes $k_{\|} \lesssim 0.58\ h\mathrm{Mpc}^{-1}$. Limiting calibration to short baselines can restrict foreground leakage to low $k_{\|}$ modes, freeing a larger region of the EoR window from contamination (Ewall-Wice et al., 2016). However, it is critical that the calibration model can accurately model visibilities from the baselines used in calibration. Relying on short baselines for calibration would require a highly accurate and complete model of diffuse foreground emission, since these short baselines are sensitive to large-scale structure on the sky. In the absence of a diffuse foreground emission model, it is advantageous to calibrate to *long* baselines only (Patil et al., 2016).

To quantify the power leakage in the EoR window due to absolute calibration errors from an incomplete sky model, we first subtract the perfect calibration power spectrum in Figure 3.4(d) from the power spectrum with absolute calibration errors in Figure 3.4(a). We then select a slice of the EoR window region of the 2-D power spectrum that spans $k_{\|} = 0.07 - 1.0\ h\mathrm{Mpc}^{-1}$ and $k_{\perp} = 8.15 \times 10^{-3} - 1.015 \times 10^{-2}\ h\mathrm{Mpc}^{-1}$. This slice is centered on the mode measured by the shortest 15 m baselines in the array and is a characteristic representation of EoR window contamination that avoids regions of low uv coverage. The black rectangular outlines in the power spectra in Figure 3.4 delimit the region. Finally, we average over the power in this slice to produce a 1-D representation of power leakage in the EoR window as a function of power spectrum mode $|k|$. We plot the result in blue in Figure 3.5. The black line in Figure 3.5 is a fiducial EoR signal (Furlanetto et al., 2006).

Figure 3.5 demonstrates that absolute calibration errors from sky model incompleteness can contaminate the power spectrum measurement to an extent that dwarfs the EoR signal. The power spectrum modes plotted here are characteristic of the EoR window, the region

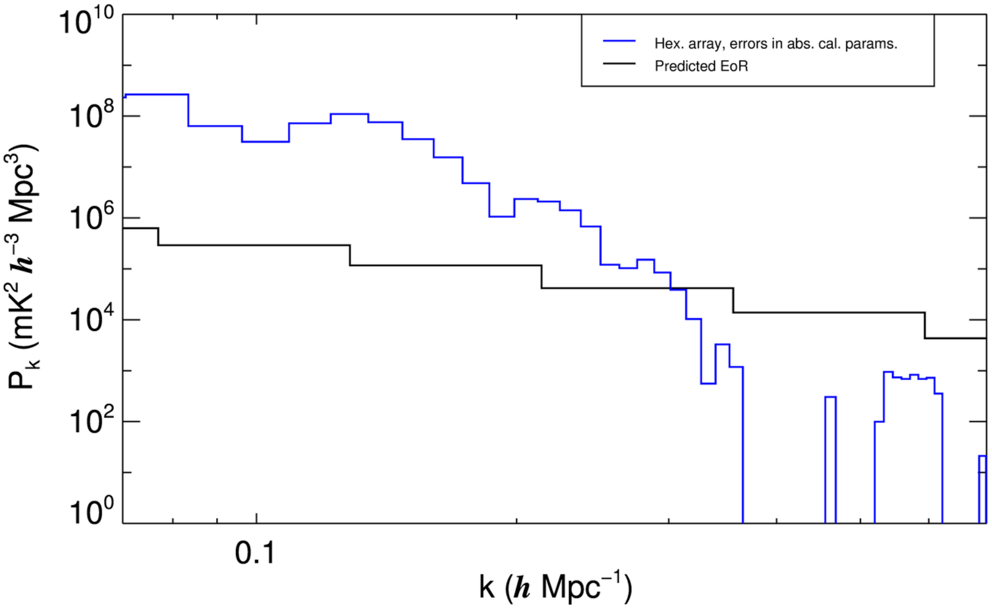


Figure 3.5: 1-D representation of power leakage in the EoR window due to errors in absolute calibration from sky model incompleteness. The simulations are based on the hexagonal array pictured in Figure 3.1. The blue line represents the difference in EoR window power between simulations with errors in absolute calibration and those with perfect calibration. The black rectangular outlines in Figure 3.4 indicate the 2-D power spectrum modes that contribute to the blue line. The black line is the predicted EoR signal. Excess power in the EoR window due to absolute calibration errors overwhelm the predicted EoR signal on most of the power spectrum modes plotted here.

of the power spectrum that is sensitive to the EoR. An EoR detection is not possible in modes in which foreground power leakage exceeds the EoR signal. The power leakage shown here is one error budget contribution corresponding to the limit of perfect array redundancy. Errors in relative calibration and thermal noise are not included in this simulation and will compound power leakage in the EoR window. An EoR detection will require mitigation of frequency-dependent absolute calibration errors to keep the total error budget within the required tolerances.

3.3 Impact of Array Layouts on Calibration Errors

The magnitude of absolute calibration errors due to sky model incompleteness depends on array layout. Errors are driven by the frequency-dependent instrument PSF, which couples to sources that are missing from the calibration model. Arrays with good PSFs consolidate power in the true source locations, reducing the amount of power in the frequency-dependent source sidelobes.

Pseudo-random arrays sample the *uv* plane more uniformly than redundant arrays. In choosing an array layout, the benefits of redundant calibration must be weighed against the trade-offs associated with redundant arrays' degraded *uv* coverage. In this section we compare calibration errors for different classes of array layouts, some of which could support redundant calibration. We compare errors in the absolute calibration parameters due to calibration model incompleteness for each of these arrays.

To illustrate the impact of *uv* coverage on absolute calibration errors, we compare simulations of three array configurations, each with 331 antennas and with similar radial antenna distributions. The first is the simple hexagonal array discussed in §3.2 and pictured in Figures 3.1 and 3.6(a). Next, we consider a hexagonal array divided into three sub-arrays offset by one third of the minimum antenna spacing (see Figure 3.6(b)), inspired by HERA's configuration (DeBoer et al., 2017). This array configuration enables redundant calibration while offering better *uv* coverage than a simple hexagonal array. Finally, we simulate the randomized array pictured in Figure 3.6(c). To create this array, we calculate the radial baseline density of the hexagonal array from Figure 3.6(a) and randomly select 331 radial distances from that density distribution. We then randomly choose an azimuthal position for each antenna, requiring a minimum spacing of 5 m between antennas. For each the offset hexagonal array and the randomized array, we repeat the simulations described in §3.2: we simulate visibilities from a 51,821-source catalog, calibrate to a 4,000-source sky model with the FHD software pipeline, and use those calibration solutions to calculate absolute calibration parameters from Equations 3.1 and 3.3.

Figure 3.7 shows the average gain amplitudes $\hat{A}$ for each the simple hexagonal array (blue), the offset hexagonal array (red) and the randomized array pictured in Figure 3.6(c) (bold orange). The thin orange lines correspond to nine additional realizations of randomized arrays, illustrating the degree of variability expected across realizations. Figure 3.7 indicates that randomized arrays have significantly smaller variations in $\hat{A}$ than either the hexagonal array or the offset hexagonal array, while the offset hexagonal array has slightly less $\hat{A}$ variation than the simple hexagonal array. Similarly, but not pictured here, the gain phase gradient parameters $\hat{\Delta}_x$ and $\hat{\Delta}_y$ have smaller variations for arrays with better *uv* coverage.

To understand this effect, it is helpful to consider the per-antenna sky-based calibration solutions alongside their average values. Figure 3.6 plots the per-antenna gain amplitudes in gray for the simple hexagonal array (left column), the offset hexagonal array (middle column), and the randomized array (right column). The average gain amplitudes $\hat{A}$ are over-plotted in color. The third row plots the per-antenna and averaged gain amplitudes as a function of frequency. The bottom row gives the PS representation of the gain amplitudes by taking their Fourier Transform and squaring. This PS representation highlights the magnitude of the frequency structure in the gain amplitude errors and shows the power spectrum modes that are contaminated by these errors.

Although the randomized array's average gain amplitudes are less variable than those of either the simple or offset hexagonal arrays, its per-antenna gain amplitudes are actually more variable. This is because the error contributions for each antenna in a highly redundant array are correlated. An individual antenna's calibration solutions depend on

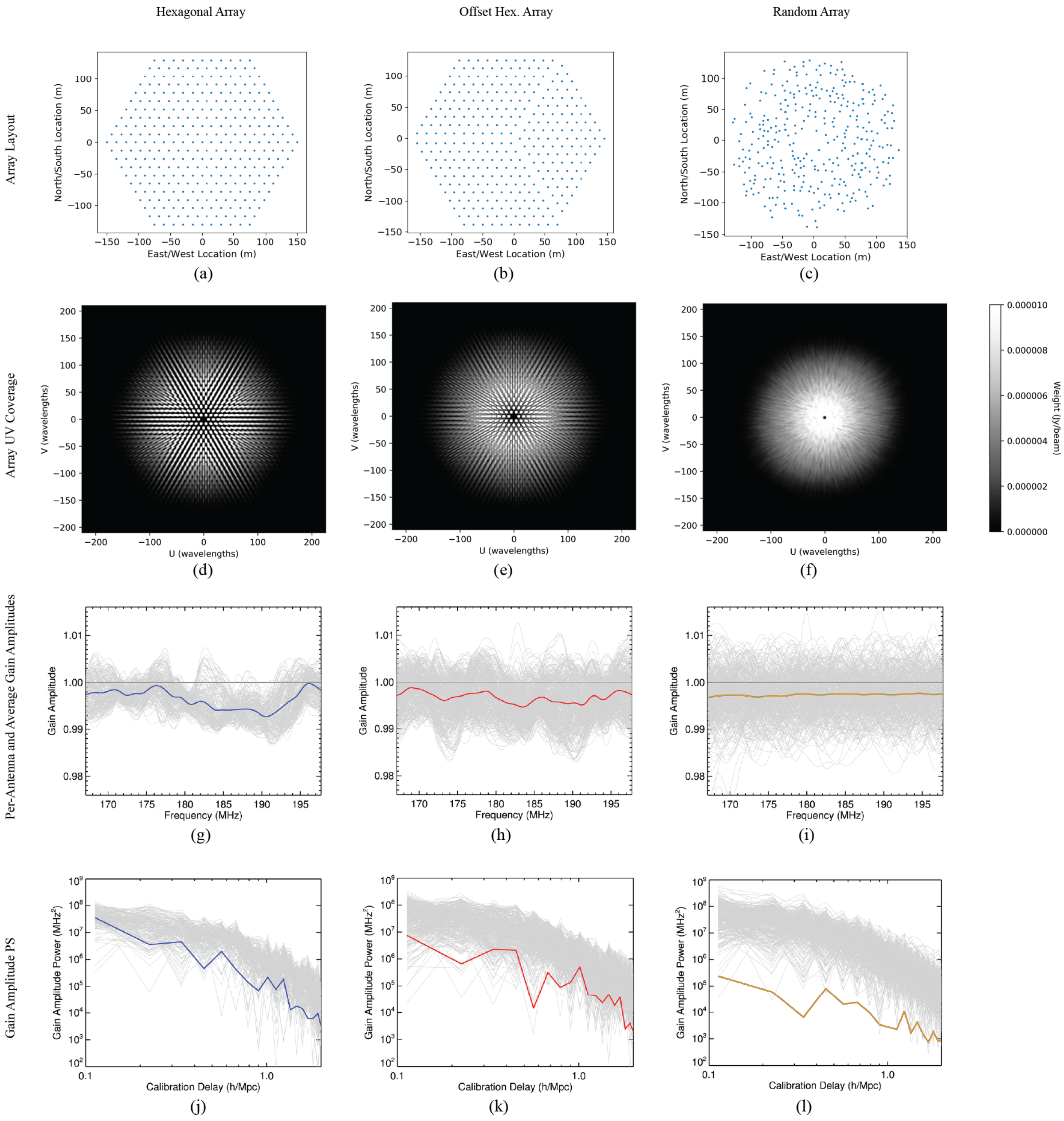


Figure 3.6: Comparison of gain amplitude errors. (Caption continued.)

Figure 3.6: Comparison of the gain amplitude errors for three array configurations. The top row shows the antenna positions for each configurations of 331 antennas: from left to right, a simple hexagonal array (also pictured in Figure 3.1), an offset hexagonal array, and a randomized array. The antenna locations in the randomized array are chosen to approximate the radial antenna distribution of the simple hexagonal array. The second row shows plots of the *uv* weights for each array configuration integrated across the 167-198 MHz frequency band and across a 2-minute observation. These plots are analogous to the measured response in the *uv* plane to a 1 Jy source at zenith. They illustrate the *uv* coverage and smoothness of each of the three arrays. The third row plots gain amplitudes as a function of frequency for each respective array. The grey lines represent the per-antenna gain amplitudes and the colored lines (also plotted in Figure 3.7) denote their average. The bottom row gives the PS (Fourier Transform squared) representation of the per-antenna and average gain amplitudes. Note that the random array has greater per-antenna gain amplitude variation than the other two array layouts but smaller average gain amplitude variation.

that antenna's PSF, i.e. the PSF from all baselines that include the antenna. The hexagonal arrays have more uniform antenna PSFs than the random array and therefore have smaller per-antenna gain amplitude variations. However, modeled visibilities for redundant baselines experience exactly the same errors from missing sources in the calibration catalog. Thus errors in per-antenna calibration solutions will be correlated across any antennas that contribute to redundant baselines. When calculating the average calibration solutions across an array, errors average coherently in highly redundant arrays and incoherently in random arrays, leading to larger error variations in the absolute calibration parameters for the highly redundant arrays.

The impact of these errors in the absolute calibration parameters is apparent in the 2-D power spectrum. Figure 3.4 presents 2-D power spectra for simulations of each the simple hexagonal array (left column), the offset hexagonal array (center column), and the random array realization pictured in Figure 3.6(c) (right column). As described in §3.2, the top row of the figure shows the PS of simulated data calibrated with errors in the absolute calibration parameters. The bottom row shows the same data with perfect calibration. In each of these figures, the PS of the calibration model has been subtracted, leaving "residual" PS.

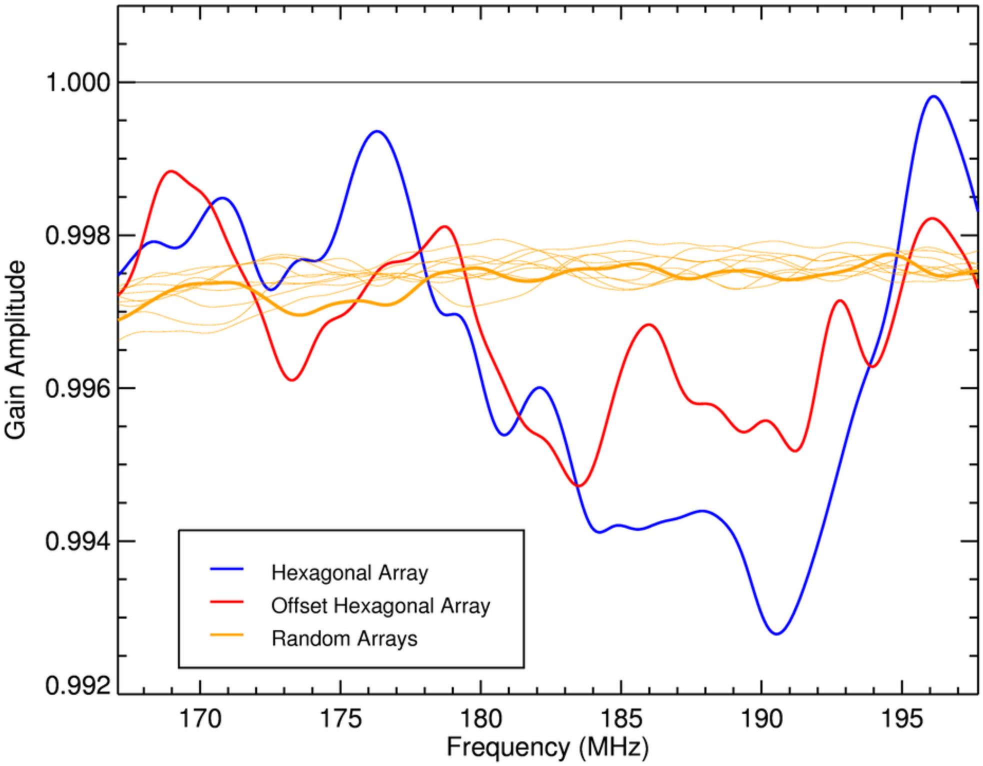


Figure 3.7: Plot of the average gain amplitudes as a function of frequency, $\hat{A}(f)$, for the hexagonal array in Figure 3.6(a) (blue), the offset hexagonal array pictured in Figure 3.6(d) (red), and ten realizations of random arrays (orange). The bold orange line corresponds to the random array pictured in Figure 3.6(c), and the additional faint orange lines illustrate the degree of variability across randomized realizations of the array configuration. The blue line is plotted alone in Figure 3.2. This figure illustrates that the random arrays have less variation in the average gain amplitude errors than the other two array configurations. The random arrays therefore exhibit less foreground power leakage in the EoR window region of the 2-D PS from absolute calibration errors.

As in §3.2, we produce 1-D plots of power leakage in the EoR window from the slices of the 2-D PS space delimited by the black rectangular outlines in Figure 3.4. We subtract the perfect calibration PS from the PS that includes absolute calibration errors to isolate the power leakage that comes from the calibration errors. Next, we average the differences to represent power leakage as function of $|k|$. These results are plotted as the solid colored lines in Figure 3.8. The blue line corresponds to the hexagonal array and is also plotted in Figure 3.5. The red line corresponds to the offset hexagonal array, and the solid orange line corresponds to the random array from Figure 3.6(c). The black line is the fiducial EoR signal.

We also plot a dashed orange line corresponding to a simulation of the random array calibrated with a traditional per-antenna, per-frequency sky-based calibration scheme. Under traditional sky-based calibration, sky model incompleteness introduces errors not only in the absolute calibration parameters but in all per-antenna, per-frequency calibration parameters. The discrepancy between the dashed and solid orange lines shows the magnitude of EoR window contamination from relative calibration errors for the random array.

3.4 Discussion

Any calibration scheme that involves a sky model must contend with frequency-dependent calibration errors from sky model incompleteness. This includes redundant calibration techniques that eliminate sky model dependence in the relative calibration step but nonetheless require a sky model for absolute calibration. Error mitigation techniques must target and suppress frequency-dependent calibration errors. These techniques include developing near-perfect sky models, calibrating to short baselines, building instruments with extremely good *uv* coverage, and manufacturing antennas with very smooth bandpasses. The success of next-generation 21-cm cosmology experiments such as HERA and the SKA is contingent on their ability to sufficiently mitigate calibration errors.

Developing highly complete and accurate sky models is an active area of research (Carroll et al., 2016; Hurley-Walker et al., 2017). While the importance of sky model completeness

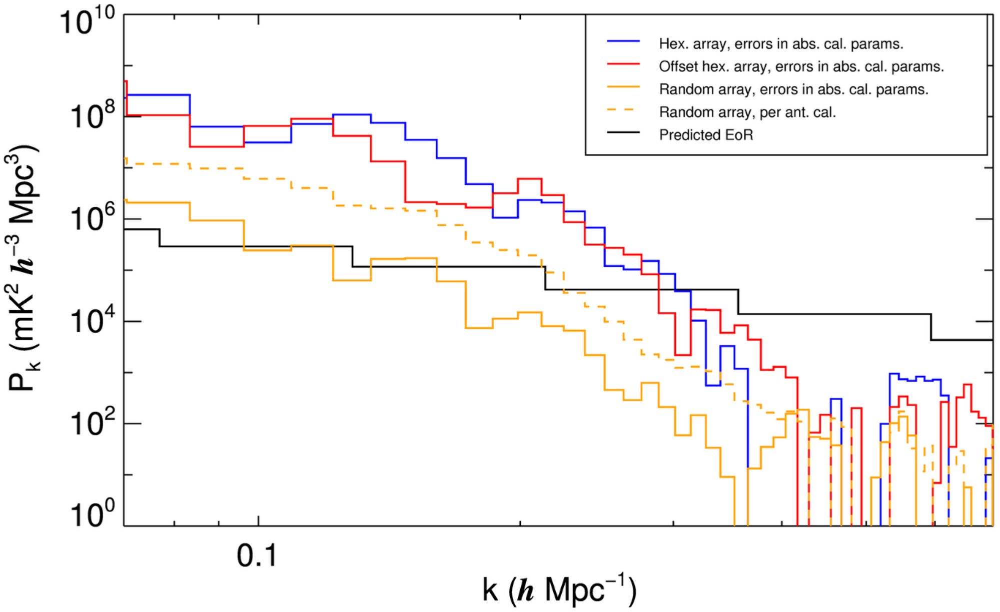


Figure 3.8: 1-D representation of power leakage from absolute calibration errors in the region of the EoR window indicated by the black rectangular outlines in Figure 3.4. The solid blue, red, and orange lines are the difference in EoR window power between simulations with errors in absolute calibration and perfect calibration for each the simple hexagonal array (Figure 3.6(a)), the offset hexagonal array (Figure 3.6(b)), and the random array (Figure 3.6(c)). The blue line is also plotted in Figure 3.5. The dashed yellow line represents the power leakage in the EoR window when, instead of allowing errors in only the absolute calibration parameters, we have implemented traditional sky-based calibration and allowed errors in every per-antenna, per-frequency calibration parameter. The black line is the predicted EoR signal.

has long been recognized in the context of sky-based calibration, redundant calibration also benefits from better sky modeling. Current efforts to image diffuse foreground structure will also enable better calibration of compact arrays that are sensitive to large-scale structure. While realistic sky models can never achieve perfect accuracy and completeness, better sky models minimize errors from the sky model incompleteness.

In §3.2 we explain that foreground power leakage from calibration errors falls off at a $k_{\parallel}$ threshold determined by the maximum baseline extent of the array. By calibrating to short baselines only, it may be possible to restrict contamination of the EoR window to low $k_{\parallel}$ modes (Ewall-Wice et al., 2016). These short-baseline calibration schemes are sensitive to the accuracy and completeness of the diffuse foreground model. To characterize these errors, future work must extend the simulations described in this chapter to include diffuse emission.

In §3.3 we show that errors from an incomplete sky model are reduced for arrays with more uniform *uv* coverage. For redundant arrays, improving *uv* coverage may mean using an array layout like the offset hexagon from Figure 3.6(d) (DeBoer et al., 2017) instead of a simple hexagon like that from Figure 3.6(a) or adopting a hybrid array configuration with both redundant and non-redundant components such as the MWA Phase II (Wayth et al., 2018). Non-redundant arrays can have more uniform *uv* coverage than highly redundant arrays; the benefits of redundancy must be weighed against the greater errors in absolute calibration that result from poor *uv* coverage. Furthermore, arrays with sufficiently uniform antenna responses such that calibration benefits from averaging across antennas do not gain an advantage from redundant calibration. Averaging antennas eliminates the relative calibration degrees of freedom such that calibration consists of the absolute calibration step only. In that regime, array redundancy has an advantage if PS sensitivity, rather than calibration systematics, is a principal concern (Parsons et al., 2012).

Instruments with smooth spectral responses are the gold standard of precision calibration. Like antenna uniformity, spectral smoothness can be used as a prior on calibration solutions, prohibiting calibration errors from introducing power into the high $k_{\parallel}$ modes sensitive to the

EoR (Trott et al., 2016; de Lera Acedo et al., 2017; Trott and Wayth, 2017). However, if this prior is imposed on an instrument that is not inherently spectrally smooth calibration will not fit the true spectral features of the instrument response and they will contaminate the PS. To avoid contamination of the EoR signal, any spectral features in the antenna and receiver system faster than ~ 8 MHz must be smaller than $\sim 10^{-5}$ (Barry et al., 2016).

Chapter 4

THE UNIFIED CALIBRATION FRAMEWORK

As discussed in §2.4, interferometric calibration approaches can typically be categorized into two classes, sky-based calibration and redundant calibration. Sky-based calibration uses models of the sky and instrument to simulate data. Calibration then consists of fitting measurements to the simulation. In contrast, redundant calibration relies on highly regular arrays with many redundant baselines measuring the same sky signal (Wieringa, 1992; Liu et al., 2010; Dillon and Parsons, 2016; Dillon et al., 2018; Grobler et al., 2018; Kern et al., 2020; Dillon et al., 2020). Calibration then fits redundant measurements to each other, aiming for consistency between the measurements. Many diverse approaches fall into one of these two classes of calibration.

As the field of 21 cm cosmology pushes the limits of precision interferometric calibration it has become increasingly important that calibration frameworks mitigate error while capturing all instrumental systematics. Sky-based calibration approaches assume very good *a priori* models of the sky and instrument; model errors quickly degrade calibration and can preclude detection of the cosmological signal (Barry et al., 2016; Ewall-Wice et al., 2016). Redundant calibration is likewise vulnerable to sky model errors (see Chapter 3 and Byrne et al. 2019) and experiences further calibration errors from instrument non-redundancies. Antenna position offsets and beam response inhomogeneities break redundant calibration's assumption of baseline redundancy and produce calibration errors (Joseph et al., 2018; Li et al., 2018; Orosz et al., 2019). The field requires novel calibration approaches that are resilient to sky and instrument model errors, can capture array non-redundancies, and mitigate contamination of the cosmological signal.

Combined calibration approaches integrate aspects of sky-based and redundant calibration. The MWA Phase II is a hybrid array that incorporates regular hexagonal sub-arrays within an otherwise pseudo-random array (Wayth et al., 2018). Li et al. 2018, Li et al. 2019, and Zhang et al. 2020 combine redundant calibration of the sub-arrays with sky-based calibration for the remainder of the array. Sievers 2017 proposes a calibration algorithm that relaxes redundancy requirements for redundant calibration of a regular array. This chapter, based on Byrne et al. 2021, expands on these ideas to present a fully unified calibration framework that is statistically rigorous, highly flexible, and physically-motivated.

§4.1 introduces a calibration formalism based on a Bayesian statistical approach. In §4.2 we re-derive simple sky-based and redundant calibration frameworks with a focus on delineating the implicit assumptions of those approaches. In §4.3 we present a novel calibration framework that unifies redundant and sky-based calibration, and in §4.4 we extend this framework to encompass frequency-dependent calibration. In the subsequent chapter (Chapter 5) we explore some of the implications of unified calibration in simulation.

4.1 Statistical Formalism

In this section we present statistical underpinnings of the calibration problem. We use Bayes' theorem to formulate a likelihood function with a statistical prior. Calibration consists of maximizing this likelihood function to calculate the instrumental response.

4.1.1 Bayesian Statement of the Calibration Problem

Calibration can be interpreted as a model fitting problem that fits tunable parameters to data. The fitting procedure asks, "what are the most likely calibration parameters given the data?" If we define our visibility data as $\boldsymbol{v}$ and our tunable calibration parameters as $\boldsymbol{\xi}$, we can write the probability of $\boldsymbol{\xi}$ given $\boldsymbol{v}$ as $P(\boldsymbol{\xi}|\boldsymbol{v})$. This quantity is also called the likelihood function. We can use the likelihood function to calculate maximum likelihood parameters $\hat{\boldsymbol{\xi}}$ that maximize $P(\boldsymbol{\xi}|\boldsymbol{v})$.

As in most model fitting problems, there is no single right answer for the form of the likelihood. One must construct a model that appropriately represents the data and mitigates systematics that contaminate the measurement. As precision calibration is crucial to the success of 21cm cosmology, choosing a well-motivated form for the likelihood function is extremely important. This chapter describes a general calibration framework and explores different physically motivated likelihood functions that could improve calibration performance for 21cm cosmology.

From Bayes' theorem we can write the likelihood function as

$$P(\boldsymbol{\xi}|\boldsymbol{v}) = \frac{P(\boldsymbol{v}|\boldsymbol{\xi})P(\boldsymbol{\xi})}{P(\boldsymbol{v})} \tag{4.1}$$

where $P(\boldsymbol{v}|\boldsymbol{\xi})$ is the probability of the data $\boldsymbol{v}$ given model parameters $\boldsymbol{\xi}$. $P(\boldsymbol{\xi})$ is the prior probability, i.e. the probability of $\boldsymbol{\xi}$ independent of the data. $P(\boldsymbol{v})$ is the marginal likelihood of $\boldsymbol{v}$ and is constant across all models. We can therefore simplify the problem by considering a proportionality

$$P(\boldsymbol{\xi}|\boldsymbol{v}) \propto P(\boldsymbol{v}|\boldsymbol{\xi})P(\boldsymbol{\xi}) \tag{4.2}$$

and maximizing the right hand side.

Maximizing the likelihood function $P(\boldsymbol{\xi}|\boldsymbol{v})$ is equivalent to minimizing the negative log-likelihood. Calibration algorithms typical perform the latter procedure rather than explicitly maximizing $P(\boldsymbol{\xi}|\boldsymbol{v})$. We describe the negative log-likelihood as

$$L(\boldsymbol{\xi}) = -C_1 \log[P(\boldsymbol{\xi}|\boldsymbol{v})] + C_2, \tag{4.3}$$

where C_1 is an arbitrary positive constant and C_2 is an arbitrary constant of either sign. Neither constant affects the result $\hat{\boldsymbol{\xi}}$ achieved by minimizing $L(\boldsymbol{\xi})$, and we therefore choose C_1 and C_2 to simplify the form of $L(\boldsymbol{\xi})$. For the case that $P(\boldsymbol{\xi}|\boldsymbol{v})$ takes the form of a Gaussian distribution, the negative log-likelihood is equivalent to the least-squares cost function, also called the chi-squared and denoted χ^2. For the sake of generality we do not assume that the likelihood is Gaussian. $L(\boldsymbol{\xi})$ can, but does not necessarily, take the form of a chi-squared.

This formulation of the likelihood is completely general and $\boldsymbol{\xi}$ can take any form. For example, traditional direction-independent calibration defines $\boldsymbol{\xi}$ to be per-antenna gains;

direction-dependent calibration allows these gains to assume different values at different positions on the sky. Other calibration approaches such as redundant calibration may introduce additional tunable model parameters that parameterize uncertainties on the sky model. As with the gains, calibration tunes these additional parameters to the values that maximize the likelihood function.

$\boldsymbol{\xi}$ models the data according to some general function $\boldsymbol{\zeta}(\boldsymbol{\xi})$. Our calibration model assumes that this function can reconstruct the expectation value of the data for some values of the tunable parameters:

$$\langle \boldsymbol{v} \rangle = \boldsymbol{\zeta}(\hat{\boldsymbol{\xi}}) \tag{4.4}$$

where $\langle \boldsymbol{v} \rangle$ is the expectation value of the data.

A specific calibration approach is defined by the following choices:

1. A choice of the tunable model parameters $\boldsymbol{\xi}$. These parameterize the instrument model and could parameterize the sky model as well.
2. A choice of the form of the visibility model as a function of the tunable parameters $\boldsymbol{\zeta}(\boldsymbol{\xi})$.
3. A choice of the form of $P(\boldsymbol{v}|\boldsymbol{\xi})$ that describes the noise properties of the data.
4. A choice of the prior on the tunable parameters $P(\boldsymbol{\xi})$. This could be a flat prior, such that $P(\boldsymbol{\xi})$ is a constant, or it could have a functional form where $\hat{\boldsymbol{\xi}}$ favors an expectation value.

Once these choices are defined, calibration consists of fitting maximum-likelihood parameters $\hat{\boldsymbol{\xi}}$. A final choice determines how these parameters are applied to the data to produce calibrated results. In some calibration frameworks only a subset of the fitted parameters get applied to the data. The rest of the fitted parameters are internal to the calibration optimization procedure. In general, calibration parameters related to the instrument response are involved in transforming uncalibrated data into calibrated data. Calibration parameters related to the sky model are internal to calibration.

All calibration frameworks described in this chapter assume that the thermal noise properties of the data are Gaussian. The probability distribution $P(\boldsymbol{v}|\boldsymbol{\xi})$ from item (3) above therefore takes the form of a Gaussian distribution. While this is the most common choice, other calibration approaches relax the assumption of Gaussianity (Kazemi and Yatawatta, 2013; Ollier et al., 2017; Sob et al., 2020). The calibration frameworks in this chapter can be easily extended to non-Gaussian distributions.

4.1.2 Quantifying Constraint of Calibration Solutions

A good parameterization of the model leads to highly constrained tunable parameters. The constraint on the tunable calibration parameters $\boldsymbol{\xi}$ is quantified by the Fisher information.

The Fisher information matrix $\mathbf{I}$ is given by the negative Hessian of the log-likelihood, where the elements of $\mathbf{I}$ are

$$I_{jk}(\boldsymbol{\xi}) = -\frac{\partial^2 \log[P(\boldsymbol{\xi}|\boldsymbol{v})]}{\partial\xi_j \partial\xi_k}. \tag{4.5}$$

From Equation 4.3 we can also express the Fisher information in terms of the negative log-likelihood, $L(\boldsymbol{\xi})$:

$$I_{jk}(\boldsymbol{\xi}) = \frac{\partial^2 L(\boldsymbol{\xi})}{\partial\xi_j \partial\xi_k}. \tag{4.6}$$

One can interpret the Fisher information as a measure of the curvature of the $L(\boldsymbol{\xi})$ hypersurface. At the maximum-likelihood point, where $\boldsymbol{\xi} = \hat{\boldsymbol{\xi}}$, $L(\boldsymbol{\xi})$ experiences a minimum. We therefore expect that all elements of the Fisher information matrix are non-negative:

$$I_{jk}(\hat{\boldsymbol{\xi}}) \geq 0. \tag{4.7}$$

Large values of this Fisher information indicate that $L(\boldsymbol{\xi})$ experiences a sharp minimum and the tunable parameters are highly constrained. Conversely, small Fisher information values mean the model is relatively agnostic to the parameter values and that the tunable parameters are not well-constrained.

If $I_{jk}(\hat{\boldsymbol{\xi}}) = 0$ for all k then the parameter ξ_j is completely unconstrained by the model. It follows that the calibration solutions are degenerate. $L(\boldsymbol{\xi})$ does not experience a unique minimum but is rather minimized for any value of ξ_j. In this case, ξ_j is a degenerate parameter.

It is not always evident when calibration solutions are degenerate. Degenerate parameters can consist of linear combinations of the tunable parameters $\boldsymbol{\xi}$. We can calculate the number of degenerate parameters by taking the rank of the Fisher information matrix at the maximum-likelihood point. The calibration solutions are degenerate if $\mathbf{I}(\hat{\boldsymbol{\xi}})$ is singular. The number of degenerate parameters is equal to the dimensionality of the null space of $\mathbf{I}(\hat{\boldsymbol{\xi}})$; the degenerate parameters are the eigenvectors that span the null space.

Calibration degeneracies are a major challenge in 21cm cosmology. All degeneracies must be constrained to yield physical calibration solutions. Often, calibration consists of two stages where an initial calibration framework yields degenerate solutions and second step constrains those degeneracies (Liu et al., 2010; Zheng et al., 2014; Byrne et al., 2019; Kern et al., 2020).

4.2 Traditional Direction-Independent Calibration Approaches

In this section we revisit the traditional calibration approaches introduced in §2.4—sky-based calibration (§2.4.1) and redundant calibration (§2.4.2)—and describe how they relate to the calibration framework presented in §4.1.

Both these calibration frameworks assume that the likelihood function is separable in frequency, time, and polarization. This means that each frequency channel, observation interval, and polarization mode can be calibrated independently, and we therefore omit explicit frequency, time, and polarization dependence. For a discussion of frequency calibration see §4.4; for a discussion of polarized calibration see §6.4.

4.2.1 Sky-Based Calibration

Traditional sky-based calibration, introduced in §2.4.1, makes the following choices:

1. It parameterizes the tunable instrument calibration parameters as a complex gain per antenna (and implicitly per polarization, frequency, and time step). $\boldsymbol{\xi} = \boldsymbol{g}$ where $\boldsymbol{g}$ has length equal to the number of antennas.

2. It models the data as
$$\zeta_{ab} = g_a g_b^* m_{ab} \tag{4.8}$$
where indices a and b index antennas. The combined indices ab indicate the index of the visibility formed by correlating signals from antennas a and b. $\boldsymbol{m}$ are model visibilities developed with a sky model and instrument simulator. The sky model typically consists of a point-source catalog, sometimes in conjunction with a diffuse foreground emission map.

3. It describes $P(\boldsymbol{v}|\boldsymbol{\xi})$ as a Gaussian probability distribution:
$$P(\boldsymbol{v}|\boldsymbol{\xi}) \propto e^{-\frac{1}{2}[\boldsymbol{v}-\boldsymbol{\zeta}(\boldsymbol{\xi})]^\dagger \mathbf{C}_\mathrm{T}^{-1}[\boldsymbol{v}-\boldsymbol{\zeta}(\boldsymbol{\xi})]}, \tag{4.9}$$
where the † symbol denotes the conjugate transpose. Here $\mathbf{C}_\mathrm{T}$ is the thermal covariance matrix, given by
$$\mathbf{C}_\mathrm{T} = \langle (\boldsymbol{v} - \langle \boldsymbol{v} \rangle)(\boldsymbol{v} - \langle \boldsymbol{v} \rangle)^\dagger \rangle. \tag{4.10}$$
Furthermore, it assumes the visibilities are independent such that $\mathbf{C}_\mathrm{T}$ is diagonal. The probability function therefore takes the form
$$P(\boldsymbol{v}|\boldsymbol{\xi}) \propto e^{-\frac{1}{2}\sum_j \frac{1}{\sigma_{\mathrm{T}j}^2} \left| v_j - \zeta_j(\boldsymbol{\xi}) \right|^2} \tag{4.11}$$
where $\sigma_{\mathrm{T}j}^2$ are the diagonal elements of $\mathbf{C}_\mathrm{T}$.

4. It uses a flat prior such that $P(\boldsymbol{\xi}) = P(\boldsymbol{g})$ is a constant.

With these choices we can write the likelihood function for traditional sky-based calibration as
$$P(\boldsymbol{g}|\boldsymbol{v}) \propto e^{-\frac{1}{2}\sum_{ab} \frac{1}{\sigma_{\mathrm{T}ab}^2} \left| v_{ab} - g_a g_b^* m_{ab} \right|^2}. \tag{4.12}$$

Maximizing this likelihood function is equivalent to minimizing the negative log-likelihood given by[1]

$$L(\boldsymbol{g}) = \sum_{ab} \frac{1}{\sigma_{\mathrm{T}ab}^2} \left| v_{ab} - g_a g_b^* m_{ab} \right|^2 . \tag{4.13}$$

This is equivalent to Equation 2.25 when $\sigma_{\mathrm{T}ab}^2$ is constant across baselines.

As discussed in §2.4.1, the resulting calibration solutions are degenerate. They have one degenerate parameter (implicitly per frequency, time, and polarization) that corresponds to the overall phase of the gains.

Traditional sky-based calibration assumes excellent knowledge of the sky and requires a highly complete sky model. It also assumes that the antenna responses are known up to a complex multiplicative factor. In other words, it assumes that the data can be fully modeled by the expression in Equation 4.8. This is clearly an inaccurate assumption. In reality, sky models are incomplete and inaccurate. These errors can propagate through the calibration process and degrade the calibration solutions (Grobler et al., 2014; Barry et al., 2016; Patil et al., 2016; Ewall-Wice et al., 2016; Joseph et al., 2020).

To account for sky model errors, we can replace the model visibilities with tunable calibration parameters $\boldsymbol{u}$. Now $\boldsymbol{\xi} = \{\boldsymbol{g}, \boldsymbol{u}\}$ and

$$\zeta_{ab} = g_a g_b^* u_{ab}. \tag{4.14}$$

The likelihood function is now given by

$$P(\boldsymbol{g}, \boldsymbol{u}|\boldsymbol{v}) \propto e^{-\frac{1}{2} \sum_{ab} \frac{1}{\sigma_{Tab}^2} \left| v_{ab} - g_a g_b^* u_{ab} \right|^2} P(\boldsymbol{u}), \tag{4.15}$$

where $P(\boldsymbol{u})$ is the prior on the tunable visibility parameters. Traditional sky-based calibration assumes that $\boldsymbol{u}$ is known exactly. Equation 4.15 converges to Equation 4.12 when

$$P(\boldsymbol{u}) \propto \delta(\boldsymbol{u} - \boldsymbol{m}), \tag{4.16}$$

---

[1] Note that this quantity is actually twice the negative log-likelihood: $L(\boldsymbol{g}) = -2\log[P(\boldsymbol{g}|\boldsymbol{v})]$. However, per Equation 4.3 the coefficient is an arbitrary constant. For conciseness, throughout this chapter we will refer to this quantity as the “negative log-likelihood.”

where δ denotes the Dirac delta function. One expects that a more accurate calibration framework would not require so stringent a prior on $\boldsymbol{u}$. In §4.3 we explore the implications of relaxing this assumption.

4.2.2 Redundant Calibration

Traditional redundant calibration (Wieringa, 1992), presented in §2.4.2, uses the following choices:

1. Like traditional sky-based calibration, traditional redundant calibration involves a tunable complex gain per antenna. However, it additionally assumes that the visibilities are not well-modeled and includes them as tunable sky model parameters. Redundant calibration works with highly regular arrays. In its traditional form, it assumes that baselines within a redundant set measure the same sky signal. The number of free visibility parameters is therefore equal to the number of redundant baseline sets. $\boldsymbol{\xi} = \{\boldsymbol{g}, \boldsymbol{u}\}$ where $\boldsymbol{u}$ correspond to the visibilities from redundant baseline sets. The number of parameters to fit in calibration is equal to the number of antennas plus the number of redundant baseline sets.

2. It models the data as
$$\zeta_{ab} = g_a g_b^* A_{ab\,j} u_j. \tag{4.17}$$
Here a and b index antennas and j indexes redundant baseline sets. $A_{ab\,j}$ represents elements of a matrix $\mathbf{A}$ that maps $\boldsymbol{u}$ to the full set of visibilities. Its elements are given by
$$A_{ab\,j} = \begin{cases} 1, & \text{if baseline } \{a, b\} \text{ belongs to set } j \\ 0, & \text{otherwise} \end{cases}. \tag{4.18}$$

3. Like traditional sky-based calibration, traditional redundant calibration describes $P(\boldsymbol{v}|\boldsymbol{\xi})$ as an independent Gaussian probability distribution as given by Equation 4.11.

4. It uses a flat prior on the tunable calibration parameters such that $P(\boldsymbol{\xi}) = P(\boldsymbol{g}, \boldsymbol{u})$ is a constant.

Under these assumptions, the likelihood function becomes

$$P(\boldsymbol{g}, \boldsymbol{u}|\boldsymbol{v}) \propto e^{-\frac{1}{2}\sum_{ab}\sum_j \frac{1}{\sigma^2_{\mathrm{T}ab}}\left|\boldsymbol{v}_{ab} - g_a g_b^* A_{ab\,j} u_j\right|^2}. \tag{4.19}$$

It follows that the negative log-likelihood is

$$L(\boldsymbol{g}, \boldsymbol{u}) = \sum_{ab}\sum_j \frac{1}{\sigma^2_{\mathrm{T}ab}}\left|v_{ab} - g_a g_b^* A_{ab,j} u_j\right|^2. \tag{4.20}$$

As noted in §2.4.2, the additional tunable calibration parameters in traditional redundant calibration, as compared to traditional sky-based calibration, introduce additional degeneracies. Not only are the calibration solutions degenerate in the overall phase of the gains but they are also degenerate in the overall gain amplitude and the gradient of the gains' complex phase across the array (Liu et al., 2010). These degeneracies mean that redundant calibration must be combined with absolute calibration to yield physical calibration solutions, where absolute calibration sets the values of the degenerate parameters. Minimizing Equation 4.20 is called relative calibration to distinguish it from true redundant calibration, which must include both relative and absolute calibration steps (Zheng et al., 2014; Byrne et al., 2019; Kern et al., 2020).

The traditional redundant calibration framework assumes perfect redundancy of baselines within a redundant set. In order for different baselines within a redundant set to measure the same sky signal, antenna positions must lie on a perfect grid and each antennas' response must be identical up to a multiplicative gain. In practice, antenna position errors and response inhomogeneities degrade array redundancy (Joseph et al., 2018; Li et al., 2018; Orosz et al., 2019). In §4.3.5-§4.3.8 we propose new calibration frameworks that account for imperfect array redundancy.

Furthermore, traditional redundant calibration neglects covariance between redundant baseline sets. This is an appropriate assumption for some redundant arrays, such as the

Donald C. Backer Precision Array for Probing the EoR (PAPER) (Parsons et al., 2010; Ali et al., 2015) and the hexagonal sub-arrays in the MWA Phase II configuration (Wayth et al., 2018), but compact redundant arrays such as the HERA (DeBoer et al., 2017), the Hydrogen Intensity Real-Time Analysis Experiment (HIRAX) (Newburgh et al., 2016), and the Canadian Hydrogen Observatory and Radio-transient Detector (CHORD) (Vanderlinde et al., 2019) can have appreciable covariance between baselines with centers at small *uv* separations. In §4.3.6-§4.3.8 we relax the assumption that baselines in different redundant baseline sets have zero covariance.

Redundant calibration can be an attractive alternative to sky-based calibration because it has reduced reliance on the sky model. One might think that redundant calibration can calibrate interferometric data without any prior knowledge of the sky — after all, Equation 4.20 doesn't require a sky model at all! However, this neglects the role of the sky model in absolute calibration. Although the promise of redundant calibration is to reduce the impact of sky model errors on calibration solutions, it is inaccurate to assume one can calibrate without *any* prior knowledge of the sky. Furthermore, it is possible that introducing a sky model into the relative calibration step of redundant calibration as a prior on the fit visibility values could improve calibration performance. In §4.3 we explore calibration approaches that do just that.

4.3 A General Framework for Direction-Independent Calibration

In this section we describe a novel and highly general calibration framework. This framework allows for extensions to traditional sky-based and redundant calibration that relax some of the non-physical assumptions of those approaches. The framework is inherently direction-independent. Future work could extend these techniques to direction-dependent calibration approaches (see §2.4.3). In this section, this calibration framework assumes per-frequency, per-time, and per-polarization calibration. §4.4 extends the framework to frequency-dependent calibration. See §6.4 for a discussion of fully polarized calibration techniques.

4.3.1 General Framework

In its most general form, this calibration framework makes the following assumptions:

1. It involves tunable complex gains per antenna (and implicitly per polarization, frequency, and time) $\boldsymbol{g}$. The gains are not direction-dependent, reflecting that this calibration framework is inherently direction-independent. In addition, this calibration framework involves tunable sky parameters $\boldsymbol{u}$. The specific form of $\boldsymbol{u}$ depends on the class of calibration used. $\boldsymbol{u}$ can be interpreted as visibilities, pixels in the *uv* plane, source flux densities, or something else altogether. In general, $\boldsymbol{u}$ represents aspects of the sky model, possibly together with aspects of the instrument model, that are fit in calibration. The tunable calibration parameters are $\boldsymbol{\xi} = \{\boldsymbol{g}, \boldsymbol{u}\}$.

2. We model the data as

   $$\zeta_{ab} = g_a g_b^* \sum_j A_{ab\,j} u_j. \tag{4.21}$$

   Here $A_{ab\,j}$ are elements of a matrix $\mathbf{A}$ that maps the sky parameters $\boldsymbol{u}$ to model visibilities. In more compact matrix multiplication notation, we can write

   $$\boldsymbol{\zeta} = \mathbf{G}(\boldsymbol{g})\mathbf{A}\boldsymbol{u}, \tag{4.22}$$

   where we have defined $\mathbf{G}(\boldsymbol{g})$ as a diagonal matrix with elements $g_a g_b^*$. In the case that $\boldsymbol{u}$ represents visibilities, $\mathbf{A}$ is simply the identity matrix: $\mathbf{A} = \mathbb{1}$.

3. We describe $P(\boldsymbol{v}|\boldsymbol{\xi})$ as an independent Gaussian probability distribution of the form of Equation 4.11.

4. We use a Gaussian prior on $\boldsymbol{u}$ such that

   $$P(\boldsymbol{u}) \propto e^{-\frac{1}{2}(\boldsymbol{u}-\boldsymbol{m})^\dagger \mathbf{C}_{\mathrm{M}}^{-1}(\boldsymbol{u}-\boldsymbol{m})} \tag{4.23}$$

   where $\boldsymbol{m} = \langle \boldsymbol{u} \rangle$. $\mathbf{C}_{\mathrm{M}} = \mathrm{cov}[\boldsymbol{u}, \boldsymbol{u}^\dagger]$ is an invertible matrix that encodes the covariances between elements of $\boldsymbol{u}$. Furthermore, this calibration framework assumes that $\boldsymbol{g}$ and

$\boldsymbol{u}$ are independent and uses a flat prior on $\boldsymbol{g}$:

$$P(\boldsymbol{\xi}) = P(\boldsymbol{g}, \boldsymbol{u}) = P(\boldsymbol{g})P(\boldsymbol{u}) \propto P(\boldsymbol{u}). \tag{4.24}$$

Under this general framework, the likelihood function is given by

$$P(\boldsymbol{g}, \boldsymbol{u}|\boldsymbol{v}) \propto e^{-\frac{1}{2}[\boldsymbol{v}-\mathbf{G}(\boldsymbol{g})\mathbf{A}\boldsymbol{u}]^{\dagger}\mathbf{C}_{\mathrm{T}}^{-1}[\boldsymbol{v}-\mathbf{G}(\boldsymbol{g})\mathbf{A}\boldsymbol{u}]}\, e^{-\frac{1}{2}(\boldsymbol{u}-\boldsymbol{m})^{\dagger}\mathbf{C}_{\mathrm{M}}^{-1}(\boldsymbol{u}-\boldsymbol{m})} \tag{4.25}$$

and $L(\boldsymbol{g}, \boldsymbol{u})$ is given by

$$L(\boldsymbol{g}, \boldsymbol{u}) = [\boldsymbol{v} - \mathbf{G}(\boldsymbol{g})\mathbf{A}\boldsymbol{u}]^{\dagger}\mathbf{C}_{\mathrm{T}}^{-1}[\boldsymbol{v} - \mathbf{G}(\boldsymbol{g})\mathbf{A}\boldsymbol{u}] + (\boldsymbol{u} - \boldsymbol{m})^{\dagger}\mathbf{C}_{\mathrm{M}}^{-1}(\boldsymbol{u} - \boldsymbol{m}). \tag{4.26}$$

$\mathbf{C}_{\mathrm{T}}$ is diagonal because the thermal noise on the visibilities is assumed to be independent.

The calibration framework described here is highly abstracted. It represents a generalized treatment can be applied to practical calibration problems. In §4.3.2-§4.3.8 we delineate examples of calibration approaches that emerge from this framework. These are just some of the possible avenues for future exploration.

4.3.2 Sky-Based Calibration with Partial Sky Model Knowledge

Traditional sky-based calibration assumes that the model of the visibilities represents the true signal up to a multiplicative complex gain. Of course this is not true. Model visibilities are susceptible to errors from incomplete knowledge of the sky and instrument. We can make sky-based calibration more resilient to sky model errors by including tunable visibility parameters $\boldsymbol{u}$ in calibration.

A dominant source of model visibility error stems from faint missing sources in the sky model catalog. We assume many sources exist with intensities below the sensitivity limit of the catalog. Each of these sources produces a fringe pattern in *uv* space. The signals from different sources add incoherently. In the limit of many sources randomly distributed on the sky, with flux densities randomly drawn from a power-law distribution, the central limit theorem dictates that the combined signal approaches Gaussian random noise in the *uv* plane (Tegmark et al., 1998). This suggests that the true visibilities are Gaussian-distributed around their model values and motivates using a Gaussian prior $P(\boldsymbol{u})$ given by Equation 4.23. This assumption breaks down if model visibility errors are instead dominated by a small number of bright sources that are either missing or mis-modeled in the sky model catalog. §4.3.3 presents a calibration framework that assumes correctly modeled source positions but uncertain source intensities.

If the baseline measurements sample discrete regions of uv space we further expect that the visibilities are independent. This means that $\mathbf{C}_\mathrm{M}$ is diagonal. Compact or redundant arrays can have significant baseline overlap in uv space and the assumption of independent visibilities does not hold. In §4.3.4-§4.3.8 we explore additional calibration frameworks that account for visibility covariance.

Under these assumptions, the prior on the tunable visibility parameters $\boldsymbol{u}$ takes the form

$$P(\boldsymbol{u}) \propto e^{-\frac{1}{2}\sum_j \frac{1}{\sigma^2_{\mathrm{M}j}}|u_j - m_j|^2}, \tag{4.27}$$

where $\sigma^2_{\mathrm{M}j}$ are the diagonal elements of $\mathbf{C}_\mathrm{M}$. We therefore get $L(\boldsymbol{g}, \boldsymbol{u})$ of the form

$$L(\boldsymbol{g}, \boldsymbol{u}) = \sum_{ab} \frac{1}{\sigma^2_{\mathrm{T}ab}} \left|v_{ab} - g_a g_b^* u_{ab}\right|^2 + \sum_j \frac{1}{\sigma^2_{\mathrm{M}j}} |u_j - m_j|^2 \tag{4.28}$$

and fit for calibration parameters $\boldsymbol{\xi} = \{\boldsymbol{g}, \boldsymbol{u}\}$. Note that this converges to traditional sky calibration (Equation 4.13) in the limit that $\sigma^2_{\mathrm{M}j} \to 0$.

The addition of new tunable calibration parameters $\boldsymbol{u}$ expands the number of tunable parameters as compared to traditional sky-based calibration. Traditional sky-based calibration fits one complex parameter per antenna for a total of N_ant complex parameters, where N_ant refers to the number of antennas. Minimizing Equation 4.28 involves fitting an additional number of parameters equal to the number of independent baseline measurements. As autocorrelation visibilities are typically omitted from calibration, this corresponds to $(N^2_\mathrm{ant} - N_\mathrm{ant})/2$ additional complex parameters. Calibration runtimes are highly dependent on the number of tunable parameters with the precise relationship depending on the implementation details.

We note that while this calibration approach involves many tunable parameters, overfitting is not a concern because the tunable visibility parameters $\boldsymbol{u}$ are constrained by their prior $P(\boldsymbol{u})$. While their maximum-likelihood values may fit some degree of thermal noise, this noise does not impact the gains $\boldsymbol{g}$ to a greater extent than it would in a traditional sky-based calibration approach.

One common sky-based calibration technique is to calibrate only on selected baselines where the sky model is trusted. For example, this could involve calibrating only on baselines longer than 50 wavelengths in order to eliminate the short baselines that are poorly modeled by a point source catalog. We can replicate this technique from the formalism described in Equation 4.28 by taking the limits

$$\sigma^2_{\mathrm{M}j} \rightarrow \begin{cases} 0 & \text{if baseline } j \text{ is included in calibration} \\ \infty & \text{if baseline } j \text{ is excluded from calibration} \end{cases} . \tag{4.29}$$

We now find that $u_j \rightarrow m_j$ for those baselines that are included in calibration: the fitted visibilities for those baselines are constrained to match the model visibilities. On the other hand, u_j for baselines that are excluded from calibration are completely unconstrained. They will therefore take on values such that $g_a g_b^* u_{ab} = v_{ab}$ for any values of the gains g_a and g_b.

While this calibration framework can replicate binary baseline selection it need not completely include or exclude baselines. σ^2_{M} quantifies the uncertainty on the model visibilities and can take any values. For example, it can be adjusted to be a function of baseline length to represent different levels of model confidence on different angular scales (Ewall-Wice et al., 2016). Rather than completely eliminating short baselines from calibration, one could instead selectively downweight them, gradually increasing the value of $\sigma^2_{\mathrm{M}j}$ on subsequently shorter and shorter baselines. σ^2_{M} could also be calculated empirically by measuring the agreement of the data and sky model as a function of baseline length.

4.3.3 Sky-Based Calibration with Uncertain Source Intensitites

Some calibration approaches in the literature consider the case of a point source sky model in which source intensities are not well-constrained (Mitchell et al., 2008; Sievers, 2017). In this case, it could be advantageous to redefine $\boldsymbol{u}$ to be the source intensities rather than visibilities. Rather than assuming that sky model errors are dominated by many unmodeled sources with random positions, as in §4.3.2, this framework assumes that the source catalog is highly complete with well-modeled source positions. It could work well when the dominant sky model errors stem from errors in the modeled flux density of known sources. Under these assumptions, we can define $L(\boldsymbol{g}, \boldsymbol{u})$ to take the form

$$L(\boldsymbol{g}, \boldsymbol{u}) = [\boldsymbol{v} - \mathbf{G}(\boldsymbol{g})\mathbf{A}\boldsymbol{u}]^\dagger \mathbf{C}_\mathrm{T}^{-1} [\boldsymbol{v} - \mathbf{G}(\boldsymbol{g})\mathbf{A}\boldsymbol{u}] + \sum_j \frac{1}{\sigma^2_{\mathrm{M}j}} |u_j - m_j|^2, \tag{4.30}$$

where $m_j = \langle u_j \rangle$ is the expected flux density of source j and $\sigma^2_{\mathrm{M},j}$ quantifies the uncertainty of that intensity. The matrix $\mathbf{A}$ models the instrument response. The element A_{jk} is equal to the contribution to visibility j from a source of unity intensity at the position of source k.

Precision calibration for 21cm cosmology often uses tens of thousands of calibrator sources in the model. Allowing the intensities of all sources to vary is computationally infeasible. Instead, a more realistic approach would be to take $\sigma^2_{\mathrm{M}j} \to 0$, such that $u_j \to m_j$, for the majority of sources j. Then $\sigma^2_{\mathrm{M}j} > 0$ only for a select subset of troublesome sources that one expects are inaccurately modeled.

This calibration approach is similar to direction-dependent calibration (§2.4.3) in that it can account for mis-modeled apparent source intensities that vary across the sky. However, we classify this approach as direction-independent because the fitted gains — the calibration parameters that are actually applied to the data — are not direction-dependent. Importantly, this calibration framework does not fit a per-antenna intensity variation across the sky.

4.3.4 Redundant Calibration With a Sky Model and Calibration of Hybrid (Redundant and Non-Redundant) Arrays

As explained in §4.2.2, traditional redundant calibration does not require a sky model or an explicit instrument model in the relative calibration step. However, this means that it yields degenerate solutions. These degeneracies must be broken through an absolute calibration step that fits the degenerate parameters to model visibilities.

An extension to traditional redundant calibration incorporates model visibilities in $L(\boldsymbol{g}, \boldsymbol{u})$. This eliminates the need for a separate absolute calibration step and instead performs relative and absolute calibration simultaneously. It can also constrain the tunable visibility parameters $\boldsymbol{u}$ to approximate a known sky model.

As in §4.2.2 we define $\boldsymbol{u}$ to be the independent visibilities with length equal to the number of redundant baseline sets. We assume that different redundant baseline sets are independent and that, as in §4.3.2, the errors on the model *uv* plane are Gaussian distributed. Under these assumptions, we expand the traditional redundant calibration framework given by Equation 4.20 by including a Bayesian prior on $\boldsymbol{u}$:

$$L(\boldsymbol{g}, \boldsymbol{u}) = [\boldsymbol{v} - \mathbf{G}(\boldsymbol{g})\mathbf{A}\boldsymbol{u}]^\dagger \mathbf{C}_\mathrm{T}^{-1} [\boldsymbol{v} - \mathbf{G}(\boldsymbol{g})\mathbf{A}\boldsymbol{u}] + \sum_j \frac{1}{\sigma_{\mathrm{M}j}^2} |u_j - m_j|^2. \tag{4.31}$$

$\mathbf{A}$ maps the independent visibilites to the full set of visibilities and is given by Equation 4.18.

At this point it is useful to note that Equation 4.31 can be equivalently derived by considering the full basis of visibilities. Instead of defining $\boldsymbol{u}$ as the visibilities from redundant baseline sets, we instead consider $\boldsymbol{u}_\mathrm{orig}$, with length equal to the total number of visibilities. The associated matrix $\mathbf{A}_\mathrm{orig} = \mathbb{1}$. The visibility covariance matrix $\mathbf{C}_{\mathrm{M\,orig}} = \mathrm{cov}[\boldsymbol{u}_\mathrm{orig}, \boldsymbol{u}_\mathrm{orig}^\dagger]$. If we assume perfect redundancy, such that baselines within a redundant set measure the same signal, then $\mathrm{cov}[\boldsymbol{u}_{\mathrm{orig}\,j}, \boldsymbol{u}^*_{\mathrm{orig}\,k}] = \mathrm{var}[\boldsymbol{u}_{\mathrm{orig}\,j}]$ when baselines j and k belong to the same redundant baseline set. If we further assume independence between visibilities from different redundant baseline sets then $\mathrm{cov}[\boldsymbol{u}_{\mathrm{orig}\,j}, \boldsymbol{u}^*_{\mathrm{orig}\,k}] = 0$ when baselines j and k belong to different redundant baseline sets. We thereby derive a block-diagonal matrix $\mathbf{C}_{\mathrm{M\,orig}}$ where the number of blocks corresponds to the number of redundant baseline sets.

However, we would encounter a problem deriving $L(\boldsymbol{g}, \boldsymbol{u}_\mathrm{orig})$ by plugging these quantities into Equation 4.26. $\mathbf{C}_{\mathrm{M\,orig}}$ is singular: it has rank equal to the number of redundant baseline sets. This belies that $\boldsymbol{u}_\mathrm{orig}$ is not a proper basis for calibration. It must be remapped into a new basis, which we derive by applying Singular Value Decomposition (SVD) to $\mathbf{C}_{\mathrm{M\,orig}}$ (see Appendix B). This generates a new calibration basis $\boldsymbol{u}$ where $\boldsymbol{u}_\mathrm{orig} = \mathbf{A}\boldsymbol{u}$ and $\mathbf{A}$ is given by Equation 4.18. We thereby recover the calibration framework given in Equation 4.31.

The calibration framework described in this section unifies the "absolute" and "relative" steps of redundant calibration. Like traditional sky-based calibration (§4.2.1), it is degenerate only in the overall phase of the gains. Yet, like traditional redundant calibration (§4.2.2), it incorporates information from baseline redundancy.

This framework involves $N_{\mathrm{ant}} + N_{\mathrm{red}}$ tunable calibration parameters, where N_{red} refers to the number of redundant baseline sets. This is equivalent to the number of tunable parameters used in traditional redundant calibration.

The framework also has interesting implications for hybrid arrays, i.e. arrays with some redundant and some non-redundant elements. Phase II of the MWA, for example, contains two hexagonal sub-arrays that support redundant calibration. These redundant sub-arrays are contained within a larger pseudo-random array. Thus far, calibrating a hybrid array has required separate calibration operations for the redundant and non-redundant elements (Li et al., 2018, 2019). In order to exploit the redundancy of the sub-arrays one had to relatively calibrate those antennas with traditional redundant calibration techniques, therefore using baseline measurements only from within the sub-arrays. The calibration framework described in this section allows for calibration of the redundant sub-array antennas using all baseline measurements. This incorporates information from all baselines involving that antenna and can increase the calibration signal-to-noise.

To explore the scope of this calibration approach we describe two opposing limits of the calibration model. The first limit takes $\boldsymbol{\sigma}_{\mathrm{M}}^2 \ll \boldsymbol{\sigma}_{\mathrm{T}}^2$, where $\boldsymbol{\sigma}_{\mathrm{T}}^2$ are elements of the diagonal matrix $\mathbf{C}_{\mathrm{T}}$. In this limit we assume that the model visibilities $\boldsymbol{m}$ are well-known and require that $\boldsymbol{u} \approx \boldsymbol{m}$. Extending this limit such that $\boldsymbol{\sigma}_{\mathrm{M}}^2 \to 0$ enforces that $\boldsymbol{u} \to \boldsymbol{m}$. Equation 4.31 is thereby equal Equation 4.13, and we recover traditional sky-based calibration.

In the opposing limit $\boldsymbol{\sigma}_{\mathrm{M}}^2 \gg \boldsymbol{\sigma}_{\mathrm{T}}^2$. The model visibilities $\boldsymbol{m}$ will then have a negligible contribution to the relative calibration parameters, i.e. the calibration parameters constrained by minimizing Equation 4.20. Instead, the model visibilities will fit the absolute calibration parameters only. This limit corresponds to a unified redundant calibration framework where absolute and relative calibration are encompassed in a single minimization problem. We note that extending this limit such that $\boldsymbol{\sigma}^2_{\mathrm{M}} \to \infty$ eliminates the second term in Equation 4.31. We thereby recover Equation 4.20 and traditional redundant calibration.

In practice, neither of these limits are physically-motivated. $\boldsymbol{m}$ cannot perfectly model the visibilities; on the other hand, it can give some information about the relative calibration solutions. The strength of this calibration approach is that it can inhabit the middle ground of *some* confidence in the model visibilities. As in §4.3.2, we could set $\boldsymbol{\sigma}^2_{\mathrm{M}}$ empirically by measuring discrepancies between the data and model. In Chapter 5 we present simulation results that explore the implications of working in an intermediate regime between traditional redundant and sky-based calibration.

4.3.5 Redundant Calibration Accounting for Unmodeled Imperfect Redundancy

As explained in §4.3.4, we can describe traditional redundant calibration with a block-diagonal $\mathbf{C}_{\mathrm{M}}$ when the basis of sky parameters $\boldsymbol{u}$ corresponds to the full set of visibilities (denoted $\mathbf{C}_{\mathrm{M\,orig}}$ in §4.3.4). Traditional redundant calibration assumes $\mathrm{cov}[u_j, u_k^*] = \mathrm{var}[u_j]$ for visibilities u_j and u_k if baselines j and k belong to the same redundant set.

However, it is reasonable to assume that small antenna position errors and beam response inhomogeneities make $\mathrm{cov}[u_j, u_k^*] < \mathrm{var}[u_j]$ when $j \neq k$. We can account for this by applying a suppression term to the off-diagonal elements of $\mathbf{C}_{\mathrm{M}}$. This suppression factor could be calculated empirically by measuring the covariance between visibilities from redundant baselines as the array measures different fields on the sky. It could also be baseline- or antenna-dependent. For example, if one antenna is known to have a particularly irregular beam response, the baseline covariance terms for baselines that include that antenna could be preferentially suppressed.

In general, suppressing the off-diagonal elements of a block-diagonal $\mathbf{C}_{\mathrm{M}}$ renders it invertible (in theory at least; in practice inverting $\mathbf{C}_{\mathrm{M}}$ could be computationally prohibitive). We may no longer need to recast $\boldsymbol{u}$ to a reduced basis such as the set of independent visibilities or the Singular Value Decomposition (SVD) basis described in Appendix B. We note that under this calibration approach the tunable visibility parameters $\boldsymbol{u}$ will be highly correlated.

The maximum likelihood values $\hat{\boldsymbol{u}}$ associated with redundant baselines will, by construction, have similar values. For this reason, we cannot use agreement of $\hat{\boldsymbol{u}}$ for redundant baselines as evidence of the baselines' true degree of redundancy.

That said, considering the SVD basis can be enlightening for understanding the effect of suppressing the off-diagonal elements of $\mathbf{C}_{\mathrm{M}}$ by a small amount. The additional calibration degrees-of-freedom introduced correspond to differences in $\boldsymbol{u}$ between visibility measurements within a redundant baseline set. The variance on these differences is small. Provided the model visibilities $\boldsymbol{m}$ are the same for all visibilities from a redundant set, the differences between visibility measurements in a redundant set are constrained to be near-zero.

Suppressing all off-diagonal elements of $\mathbf{C}_{\mathrm{M}}$ increases the number of tunable calibration parameters to $(N_{\mathrm{ant}}^2 + N_{\mathrm{ant}})/2$, as compared to $N_{\mathrm{ant}} + N_{\mathrm{red}}$ parameters for traditional redundant calibration. For large arrays this could substantially increase calibration runtimes.

4.3.6 Redundant Calibration with Modeled Imperfect Redundancy

§4.3.5 describes a calibration framework that accounts for unmodeled antenna position errors and beam response inhomogeneities. However, often antenna positions and beam responses can be measured to greater accuracy than they can be controlled. For example, an antenna in a redundant array may have a known position offset from its ideal position. In this case, the covariance matrix $\mathbf{C}_{\mathrm{M}}$ can be calculated from the modeled *uv* responses of the antennas.

In §2.1 we represent the *uv* response of baseline j as $\tilde{T}_j(-\boldsymbol{x})$, where $\boldsymbol{x}$ is the *uv* position vector. These baseline responses can be modeled from beam simulators or direct measurements and are readily available from instrument simulators such as FHD (Sullivan et al., 2012; Barry et al., 2019a), PYUVSIM[2] (Lanman et al., 2019), OSKAR[3] (Mort et al., 2010), the Precision Radio Interferometry Simulator[4] (PRISIM), or CASA (see Jagannathan et al. 2017 for an example of using CASA with a fully-polarized primary beam model). From Equation

---

[2] `https://github.com/RadioAstronomySoftwareGroup/pyuvsim`

[3] `https://github.com/OxfordSKA/OSKAR`

[4] `https://github.com/nithyanandan/PRISim`

2.15, we get that the visibility u_j relates to the true uv plane $\tilde{S}(\boldsymbol{x})$ via

$$u_j = \int_{-\infty}^{\infty} d^2\boldsymbol{x}\, \tilde{T}_j(-\boldsymbol{x})\tilde{S}(\boldsymbol{x}). \tag{4.32}$$

If we assume no uncertainty on the *uv* response models, the elements of the covariance matrix $\mathbf{C}_\mathrm{M}$ are given by

$$C_{\mathrm{M}\,jk} = \mathrm{cov}[u_j, u_k^*] = \int_{-\infty}^{\infty} d^2\boldsymbol{x} \int_{-\infty}^{\infty} d^2\boldsymbol{x}'\, \tilde{T}_j(-\boldsymbol{x})\tilde{T}_k^*(-\boldsymbol{x}')\, \mathrm{cov}[\tilde{S}(\boldsymbol{x}), \tilde{S}^*(\boldsymbol{x}')]. \tag{4.33}$$

As in §4.3.2, we assume that sky model errors are dominated by many faint sources missing from the model catalog. We expect the resulting error in the *uv* plane to be Gaussian-distributed and fully described by a variance σ_M^2 at each point. We further assume that points in the *uv* plane are independent:

$$\mathrm{cov}[\tilde{S}(\boldsymbol{x}), \tilde{S}^*(\boldsymbol{x}')] = \begin{cases} 0 & \text{when } \boldsymbol{x} \neq \boldsymbol{x}' \\ \sigma_\mathrm{M}^2 & \text{when } \boldsymbol{x} = \boldsymbol{x}' \end{cases}. \tag{4.34}$$

Under these assumptions, we get that

$$C_{\mathrm{M}\,jk} = \sigma_\mathrm{M}^2 \int_{-\infty}^{\infty} d^2\boldsymbol{x}\, \tilde{T}_j(\boldsymbol{x})\tilde{T}_k^*(\boldsymbol{x}). \tag{4.35}$$

This method of calculating elements of the covariance matrix allows us to accurately represent all amounts of modeled baseline covariance in the calibration model. Redundant baselines have highly overlapping *uv* responses. In a closely packed array, baselines that are not constructed to be redundant may nonetheless overlap somewhat in *uv* coverage; this overlap introduces nonzero covariance in their measurements. Traditional redundant calibration assumes that $\tilde{T}_j(\boldsymbol{x}) = \tilde{T}_k(\boldsymbol{x})$ when baselines j and k belong to the same redundant baseline set. It also assumes that when j and k do not belong to the same redundant baseline set their *uv* coverage does not overlap at all: $\int_{-\infty}^{\infty} d^2\boldsymbol{x}\, \tilde{T}_j(\boldsymbol{x})\tilde{T}_k^*(\boldsymbol{x}) = 0$. If we instead calculate a covariance matrix from Equation 4.35 we can relax those assumptions and build a more physically-motivated calibration model that accounts for array non-redundancies and can incorporate covariances from baseline overlap in closely packed arrays such as HERA (DeBoer et al., 2017) and HIRAX (Newburgh et al., 2016).

In §4.3.2 we noted that a sky model may be more trusted for some baselines than others. The assumption in Equation 4.35 of uniform variance across the uv plane therefore must be relaxed. If we instead assume that the variance of the uv plane is approximately constant at scales equal to the size of a baseline response, we can rewrite Equation 4.35 as

$$C_{\mathrm{M}\,jk} = \sigma_{\mathrm{M}\,j}\sigma_{\mathrm{M}\,k} \int_{-\infty}^{\infty} d^2\boldsymbol{x}\, \tilde{T}_j(\boldsymbol{x})\tilde{T}_k^*(\boldsymbol{x}). \tag{4.36}$$

Here $\sigma_{\mathrm{M}\,j}^2$ is simply the variance of u_j; $\sigma_{\mathrm{M}\,j}$ is the standard deviation. As noted in §4.3.2, one could determine values of $\boldsymbol{\sigma}_{\mathrm{M}}$ empirically by comparing the data to the model. Poorly-modeled power spectrum modes could be downweighted in calibration by increasing $\sigma_{\mathrm{M}\,j}$ for the associated baseline lengths. A baseline j could be excluded from calibration altogether by taking $\sigma_{\mathrm{M}\,j} \to \infty$.

We can combine the calibration approach described in this section with that of §4.3.5. One could expect that further non-redundancies exist beyond what is encoded in the baseline response model $\tilde{T}_j(\boldsymbol{x})$. In that case, one could represent non-modeled antenna position or beam response errors by suppressing the off-diagonal elements of $\mathbf{C}_{\mathrm{M}}$. As in §4.3.5, this suppression factor could be estimated from empirical measurements of the covariance of visibilities. Calculating visibility covariances across observations of different parts of the sky could help validate or adjust a covariance matrix constructed from Equation 4.35.

As in §4.3.5, under the calibration framework presented in this section the number of tunable calibration parameters lies between $N_{\mathrm{ant}} + N_{\mathrm{red}}$ and $(N_{\mathrm{ant}}^2 + N_{\mathrm{ant}})/2$, with the exact number depending on whether any baselines are modeled as fully redundant. The number of tunable parameters can be reduced by eliminating the most constrained calibration degrees-of-freedom. This is analogous to deciding that the most redundant baselines can be approximated as fully redundant. Formally, this is accomplished by performing Principal Value Decomposition (PVD) on the covariance matrix $\mathbf{C}_{\mathrm{M}}$. We calculate the SVD basis according to Appendix B. Of the resulting parameters, we can choose not to calibrate the parameters with the smallest variances.

4.3.7 Covariant Calibration of Compact Non-Redundant Arrays

The calibration formalism presented in §4.3.6 is fully generalizable to a compact non-redundant array. The framework allows us to incorporate information from baseline covariance to constrain measurements across a non-redundant array.

For example, Phase I of the MWA (Tingay et al., 2013) has no redundant baselines. As an imaging array, it was built to be pseudo-random with maximal *uv* coverage, in a sense making it as non-redundant as possible (Beardsley et al., 2012). At the same time, it has a highly compact core and is *uv* complete out to about 50 wavelengths. This high *uv* coverage means that baselines measurements are not independent. Instead, many baselines are highly covariant with one another.

We can incorporate this covariance into the calibration framework by constructing $\mathbf{C}_{\mathrm{M}}$ from Equation 4.35. Fit visibility values $\boldsymbol{u}$ would then be constrained not only by their modeled values $\boldsymbol{m}$ but also by the visiblities from overlapping baselines in the *uv* plane.

In this way the ideas behind redundant calibration can be extended to non-redundant compact arrays. Baselines need not be highly redundant to exhibit appreciable covariance, and this covariance can be used to constrain calibration solutions. These calibration techniques have applications for many classes of interferometers, not just highly redundant arrays.

4.3.8 uv-Space Calibration

Traditional redundant calibration interprets the tunable sky parameters $\boldsymbol{u}$ to be visibilities (denoted $\boldsymbol{u}_{\mathrm{orig}}$ in this section), however one could instead reinterpret $\boldsymbol{u}$ as pixels of the *uv* plane. This is a natural parameter space for describing the model and means that $\mathbf{C}_{\mathrm{M}}$ does not depend on the instrument response. Instead, $\mathbf{A}$ encodes degridding by mapping the pixels of the *uv* plane to the visibilitites.

In §4.3.6 we describe the instrument response model as a continuous function in the *uv* plane, but in practice the *uv* plane is discretized. We can define a covariance matrix $\mathbf{C}_{\mathrm{M\,orig}}$

by rewriting Equation 4.35 for a discrete *uv* plane:

$$\mathbf{C}_{\mathrm{M\,orig}} = \sigma_{\mathrm{M}}^2 \mathbf{T}\mathbf{T}^\dagger \tag{4.37}$$

where $\mathbf{T}$ is a rectangular matrix that maps *uv* pixels to visibilities. $\mathbf{T}$ has a number of columns equal to the number of pixels in the *uv* plane and a number of rows equal to the number of visibilities. The product $\mathbf{T}\mathbf{T}^\dagger$ is equivalent to the holographic mapping function described in Sullivan et al. 2012.

We define a new set of sky parameters $\boldsymbol{u}$ that correspond to pixels in the *uv* plane. $\mathbf{T}$ defines the mapping between $\boldsymbol{u}$ and $\boldsymbol{u}_{\mathrm{orig}}$:

$$\boldsymbol{u}_{\mathrm{orig}} = \mathbf{T}\boldsymbol{u}. \tag{4.38}$$

We now get that

$$\begin{aligned}\mathbf{C}_{\mathrm{M\,orig}} &= \langle \boldsymbol{u}_{\mathrm{orig}} \boldsymbol{u}_{\mathrm{orig}}^\dagger \rangle - \langle \boldsymbol{u}_{\mathrm{orig}} \rangle \langle \boldsymbol{u}_{\mathrm{orig}}^\dagger \rangle \\ &= \mathbf{T}(\langle \boldsymbol{u}\boldsymbol{u}^\dagger \rangle - \langle \boldsymbol{u} \rangle \langle \boldsymbol{u}^\dagger \rangle)\mathbf{T}^\dagger \\ &= \sigma_{\mathrm{M}}^2 \mathbf{T}\mathbf{T}^\dagger,\end{aligned} \tag{4.39}$$

so the new covariance matrix is

$$\mathbf{C}_{\mathrm{M}} = \langle \boldsymbol{u}\boldsymbol{u}^\dagger \rangle - \langle \boldsymbol{u} \rangle \langle \boldsymbol{u}^\dagger \rangle = \sigma_{\mathrm{M}}^2 \mathbb{1}. \tag{4.40}$$

The fact that this covariance matrix is diagonal highlights the assumption in §4.3.6 of independent *uv* pixels.

Plugging this new parameterization of $\boldsymbol{u}$ into $L(\boldsymbol{g}, \boldsymbol{u})$ gives

$$L(\boldsymbol{g}, \boldsymbol{u}) = [\boldsymbol{v} - \mathbf{G}(\boldsymbol{g})\mathbf{T}\boldsymbol{u}]^\dagger \mathbf{C}_{\mathrm{T}}^{-1} [\boldsymbol{v} - \mathbf{G}(\boldsymbol{g})\mathbf{T}\boldsymbol{u}] + \frac{1}{\sigma_{\mathrm{M}}^2} \sum_j |u_j - m_j|^2, \tag{4.41}$$

where $\boldsymbol{m}$ is the model of the *uv* plane and j indexes over *uv* plane pixels. We can also relax the assumption that $\sigma_{\mathrm{M}\,j}^2 = \sigma_{\mathrm{M}}^2$ for all *uv* pixels j and rewrite Equation 4.41 as

$$L(\boldsymbol{g}, \boldsymbol{u}) = [\boldsymbol{v} - \mathbf{G}(\boldsymbol{g})\mathbf{T}\boldsymbol{u}]^\dagger \mathbf{C}_{\mathrm{T}}^{-1} [\boldsymbol{v} - \mathbf{G}(\boldsymbol{g})\mathbf{T}\boldsymbol{u}] + \sum_j \frac{1}{\sigma_{\mathrm{M}\,j}^2} |u_j - m_j|^2. \tag{4.42}$$

Now σ_{M}^2 can represent variations in the quality of the sky model across the uv plane. A point source sky model that omits diffuse emission could have σ_{M}^2 increase for shorter baselines. Regions of uv space could be excluded from calibration altogether by taking the limit $\sigma_{\mathrm{M}\,j}^2 \to \infty$ for those pixels j. One could calculate σ_{M}^2 empirically by comparing the model of the uv plane to measurements.

We call this new calibration framework "uv-space calibration." It offers an alternative formulation of the calibration framework described in §4.3.6. Parameterizing $\boldsymbol{u}$ as uv pixels rather than visibilities highlights the implicit assumption of statistical independence of the pixels by making $\mathbf{C}_{\mathrm{M}}$ diagonal. Here $\boldsymbol{m}$ depends on the sky model only and is calculated without an instrument simulator. The instrument model is moved from the second to the first term of $L(\boldsymbol{g}, \boldsymbol{u})$, where $\mathbf{T}$ encodes the baseline responses to the uv plane. Baseline covariances need not be calculated explicitly. Instead, this calibration framework implicitly constrains visiblities from fully- and partially-redundant baselines from their uv plane overlap.

uv-space calibration is a natural framework for unified calibration that models all baseline covariances. As with the calibration framework described in §4.3.6, it can account for small non-redundancies of a redundant array due to antenna position offsets and beam inhomogeneities. Like in §4.3.7, it models covariances of non-redundant baselines stemming from baseline overlaps in the uv plane. Note that uv-space calibration incorporates modeled covariances only; it cannot include unmodeled imperfect redundancy as in §4.3.5. Finally, like the calibration framework described in §4.3.2, uv-space calibration can incorporate variable sky model uncertainties as a function of uv position to represent sky model incompleteness and preferentially fit calibration solutions to well-modeled uv modes.

4.4 Frequency Calibration

In §4.3 we assumed that the calibration likelihood function is separable in frequency:

$$P(\boldsymbol{\xi}|\boldsymbol{v}) = \prod_f P_f[\boldsymbol{\xi}(f)|\boldsymbol{v}(f)] \tag{4.43}$$

where f denotes the frequency channel. A likelihood function of this form allows for each frequency to be calibrated separately but precludes the calibration model from incorporating cross-frequency covariance information. In this section we describe calibration extensions that break the assumption of separability of the likelihood function in frequency.

4.4.1 Parameterizing Gains Across Frequency

One approach to eliminating frequency-dependent calibration errors is to fit calibration solutions across frequency after the fact (Barry et al., 2019b; Li et al., 2019). With this approach, calibration consists of two steps. First, one calculates per-frequency calibration solutions. Next, the gains are adjusted to remove false frequency structure and fit the allowable degrees of freedom determined by a model of the instrument bandpass response.

A more optimal way of calibrating with a modeled bandpass is to fit the parameters of the bandpass directly (Yatawatta, 2015; Mertens et al., 2020). We can model the gains as functions of tunable parameters $\boldsymbol{\gamma}_a$ for each antenna a such that

$$g_a(f) = g_a(f, \boldsymbol{\gamma}_a). \tag{4.44}$$

Here $\boldsymbol{\gamma}_a$ could include the amplitudes and positions of known features in the bandpass or any other parameters related to fitting the instrument's bandpass response. If the bandpass is parameterized as a low-order polynomial, $\boldsymbol{\gamma}_a$ could be the polynomial coefficients:

$$g_a(f, \boldsymbol{\gamma}_a) = \sum_{n=0}^{n_{\max}} \gamma_{a\,n} f^n, \tag{4.45}$$

where $n_{\max}$ is the maximum mode in the low-order polynomial fit.

Because the parameters $\boldsymbol{\gamma}_a$ are not per-frequency, calibration can no longer be parallelized across frequency channels. Instead, the algorithm must access data from across a large frequency range at once. This increases the calibration memory requirements by a factor equal to the number of frequency channels, which can be computationally infeasible. Novel calibration algorithms allow for distributed optimization without sacrificing cross-frequency performance (Yatawatta, 2015).

An extension to traditional sky-based calibration (Equation 4.13) that parameterizes the gains across frequencies has the following features:

1. We define tunable calibration parameters corresponding to parameterizations of the antenna gains across frequency. $\boldsymbol{\xi} = \{\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \dots\}$ where for antenna a the gain $\boldsymbol{g}_a(f) = \boldsymbol{g}_a(f, \boldsymbol{\gamma}_a)$.

2. We model the data as

$$\zeta_{ab}(f) = g_a(f, \boldsymbol{\gamma}_a) g_b^*(f, \boldsymbol{\gamma}_b) m_{ab}(f) \tag{4.46}$$

   where indices a and b index antennas.

3. As in previous sections, we describe $P(\boldsymbol{v}|\boldsymbol{\xi})$ as an independent Gaussian probability distribution.

4. We use a flat prior such that $P(\boldsymbol{\xi}) = P(\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \dots)$ is a constant.

This version of sky-based calibration has a negative log-likelihood of the form

$$L(\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \dots) = \sum_f \sum_{ab} \frac{1}{\sigma^2_{\mathrm{T}\,ab}(f)} \left| v_{ab}(f) - g_a(f, \boldsymbol{\gamma}_a) g_b^*(f, \boldsymbol{\gamma}_b) m_{ab}(f) \right|^2 . \tag{4.47}$$

Calibration consists of minimizing this quantity by varying the parameters $\boldsymbol{\gamma}_a$ for each antenna a.

This calibration approach allows for variable weighting across frequency. For example, frequency channels with greater noise contamination could be downweighted by increasing $\sigma^2_{\mathrm{T}}(f)$ with respect to the uncontaminated channels. Frequency channels can be removed from calibration altogether by taking the limit $\sigma^2_{\mathrm{T}}(f) \to \infty$. If the calibration parameters are chosen to span frequencies then they will interpolate the gains across the masked frequency channels.

This approach to fitting the gains is not only applicable to sky-based calibration. One can choose to parameterize the gains across frequency in any calibration framework.

4.4.2 Capturing Redundancy Across Frequency

Calibrating across frequencies may allow redundant calibration to capture cross-frequency visibility covariances (Dillon & Parsons, private communication). It is possible that baselines of vastly different physical lengths are nonetheless highly covariant at different frequencies. A long baseline could measure the same sky modes at a low frequency that a short baseline measures at a higher frequency. This has interesting implications for redundant calibration. As in §4.4.1, this calibration approach would pose computational challenges as it precludes parallelization across frequency channels.

To illustrate cross-frequency redundant calibration we consider an extension to redundant calibration as described in §4.2.2. Under this extension, calibration includes the following assumptions:

1. As in §4.2 and §4.3, we parameterize the gains as per-antenna, per-frequency, and direction-independent. (Alternatively, one could combine this calibration approach with the cross-frequency gain parameterizations described in §4.4.1. However, for simplicity we will describe a per-frequency gain parameterization.) We parameterize the sky model with per-frequency visibilities $\boldsymbol{u}(f)$. The tunable calibration parameters are $\boldsymbol{\xi} = \{\boldsymbol{g}(f), \boldsymbol{u}(f)\}$.

2. As in §4.3, we model the data as

$$\zeta_{ab}(f) = g_a(f) g_b^*(f) \sum_j A_{ab\,j} u_j(f) \tag{4.48}$$

   where $A_{ab\,j}$ are elements of some matrix $\mathbf{A}$ that maps $\boldsymbol{u}$ to visibilities.

3. As before, $P(\boldsymbol{v}|\boldsymbol{\xi})$ is an independent Gaussian probability distribution (Equation 4.11).

4. We use a Gaussian prior on $\boldsymbol{u}$ and a flat prior on $\boldsymbol{g}$ such that

$$P(\boldsymbol{\xi}) \propto P(\boldsymbol{u}) \propto \prod_{f_1 f_2} e^{-\frac{1}{2}[\boldsymbol{u}(f_1)-\boldsymbol{m}(f_1)]^\dagger \mathbf{C}_{\mathrm{M}}{}^{-1}_{f_1 f_2}[\boldsymbol{u}(f_2)-\boldsymbol{m}(f_2)]}. \tag{4.49}$$

$\mathbf{C}_\mathrm{M}$ now encodes covariances between every visibility at every frequency, with elements given by

$$C_{\mathrm{M}\,jkf_1f_2} = \mathrm{cov}[u_j(f_1), u_k^*(f_2)] \tag{4.50}$$

where j and k index baselines and f_1 and f_2 index frequency channels.

$L(\boldsymbol{g}, \boldsymbol{u})$ from Equation 4.26 now takes the form

$$\begin{aligned} L(\boldsymbol{g}, \boldsymbol{u}) = & \sum_f \sum_{ab} \sum_j \frac{1}{\sigma^2_{\mathrm{T}\,ab}(f)} \left| v_{ab}(f) - g_a(f) g_b^*(f) A_{ab\,j} u_j(f) \right|^2 \\ & + \sum_{f_1 f_2} \sum_{jk} [u_j(f_1) - m_j(f_1)]^* C_{\mathrm{M}\,jkf_1f_2}^{-1} [u_k(f_2) - m_k(f_2)]. \end{aligned} \tag{4.51}$$

This converges to traditional redundant calibration in the limit that $\mathbf{C}_{\mathrm{M}\,jkf_1f_2} \to 0$ for $f_1 \neq f_2$.

Assuming a stable sky as a function of frequency, a physically-motivated construction of $\mathbf{C}_\mathrm{M}$ could follow the approach described in §4.3.6. From Equation 4.35 we get

$$C_{\mathrm{M}\,jkf_1f_2} = \int_{-\infty}^{\infty} d^2\boldsymbol{x} \int_{-\infty}^{\infty} d^2\boldsymbol{x}' \, \tilde{T}_j(f_1, \boldsymbol{x}) \tilde{T}_k^*(f_2, \boldsymbol{x}') \, \mathrm{cov}[\tilde{S}(f_1, \boldsymbol{x}), \tilde{S}^*(f_2, \boldsymbol{x}')], \tag{4.52}$$

where $\tilde{T}_j(f, \boldsymbol{x})$ is the *uv* response of baseline j and $\tilde{S}(f, \boldsymbol{x})$ is the *uv* plane at frequency f. If, as in §4.3.6, we assume that different points in the *uv* plane are independent and that the variance across the *uv* plane is constant, we get that

$$C_{\mathrm{M}\,jkf_1f_2} = C_{\mathrm{S}\,f_1f_2} \int_{-\infty}^{\infty} d^2\boldsymbol{x} \, \tilde{T}_j(f_1, \boldsymbol{x}) \tilde{T}_k^*(f_2, \boldsymbol{x}). \tag{4.53}$$

Here $C_{\mathrm{S}\,ff}$ is the variance of *uv* pixels at frequency f; $C_{\mathrm{S}\,f_1f_2}$ is the covariance of *uv* pixels at frequencies f_1 and f_2.

It follows from Equation 4.53 that the construction of a physically-motivated cross-frequency $\mathbf{C}_\mathrm{M}$ requires accurate modeling of the beam responses at all frequencies in question. In practice, beam responses vary considerably as a function of frequency. While baselines of different physical lengths may sample the same *uv* locations at different frequencies, we cannot expect their beam responses to be the same. Therefore, redundancy is degraded across frequencies. In other words, we expect the largest elements of $C_{\mathrm{M}\,f_1f_2}$ to occur when $f_1 = f_2$.

If we assume the sky is constant across our frequency range then ${C_{\mathrm{S}}}_{f_1 f_2} = {C_{\mathrm{S}}}_{f_1 f_1}$ at all frequencies. However, in practice the sky has some frequency-dependence. We therefore expect that a more accurate model sets ${C_{\mathrm{S}}}_{f_1 f_2} < {C_{\mathrm{S}}}_{f_1 f_1}$ when $f_1 \neq f_2$. This further suppressed the elements of $\mathbf{C}_{\mathrm{M}}$ that correspond to covariances of visibilities at different frequencies.

Incorporating nonzero cross-frequency visibility covariances in calibration could be a powerful tool for constraining the frequency structure of the instrument response. However, beam shape variation across antennas and frequencies, along with frequency-dependence in the sky signal, suppress the covariance of visibilities at different frequencies. A realistic calibration model must account for these effects.

4.4.3 Delay-Dependent Weighting

The calibration approaches in §4.4.1 and §4.4.2 allow for frequency-dependent weighting through the parameter ${\sigma^2_{\mathrm{T}}}_{ab}(f)$. However, in practice it is useful to have delay-dependent weighting, where delay refers to the Fourier dual of frequency.

Due to the foreground wedge effect, described in §2.3, errors in the calibration sky model lie in delay modes $\eta \leq b/c$, where b is the baseline length in meters and c is the speed of light. Calibration mixes these errors across baselines, introducing contamination into modes $\eta \leq b_{\mathrm{max}}/c$ where b_{max} is the length of the longest baseline used in calibration. For this reason, Ewall-Wice et al. 2016 proposes using only short baselines in calibration.

However, delay-dependent weighting can restrict calibration from fitting features that appear only in the foreground wedge region of power spectrum space. Here we describe a delay-weighted variation of traditional sky based calibration. This approach uses the following assumptions:

1. Like traditional sky-based calibration, it parameters the tunable calibration parameters as a complex gain per antenna and per frequency $\boldsymbol{\xi} = \boldsymbol{g}(f)$.

2. Again like traditional sky-based calibration, it models the data as $\zeta_{ab}(f) = g_a(f) g_b^*(f) m_{ab}(f)$.

3. $P(\boldsymbol{v}|\boldsymbol{\xi})$ is no longer an independent probability distribution. Instead, delay-dependent weighting is imposed through a matrix that defines effective covariances between frequency channels.

4. Like traditional sky-based calibration, it uses a flat prior for the gains.

To simplify notation, we define a new expression $r_{ab}(f)$ to be the sky-based calibration cost function of a single baseline $\{a, b\}$:

$$r_{ab}(f) = v_{ab}(f) - g_a(f)g_b^*(f)m_{ab}(f). \tag{4.54}$$

If we assume there is no baseline-dependent weighting, the negative log-likelihood is simply

$$L(\boldsymbol{g}) = \sum_f \sum_{ab} |r_{ab}(f)|^2. \tag{4.55}$$

For simplicity, we can approximate the sum over frequencies as a continuous integral:

$$L(\boldsymbol{g}) = \sum_{ab} \int df |r_{ab}(f)|^2. \tag{4.56}$$

From the Plancherel theorem, this is equivalent to

$$L(\boldsymbol{g}) = \sum_{ab} \int d\eta \, |\tilde{r}_{ab}(\eta)|^2, \tag{4.57}$$

where $\tilde{r}_{ab}(\eta)$ is the Fourier dual of $r_{ab}(f)$.

At this point we can add a delay-dependent weighting function $\tilde{W}_{ab}(\eta)$:

$$L(\boldsymbol{g}) = \sum_{ab} \int d\eta \, \tilde{W}_{ab}(\eta) |\tilde{r}_{ab}(\eta)|^2, \tag{4.58}$$

This weighting function could affect only long baselines and mask out modes within the foreground wedge. For example, the weighting function could be

$$\tilde{W}_{ab}(\eta) = \begin{cases} 0, & \text{if } b_0/c \leq \eta \leq b_{ab}/c \text{ and } b_{ab} > b_0 \\ 1, & \text{otherwise} \end{cases}, \tag{4.59}$$

where b_{ab} is the length of baseline $\{a, b\}$ and b_0 is the shortest weighted baseline. This weighting would mask fast delay modes in the foreground wedge.

We can now transform the negative log-likelihood back into the frequency domain. Expressing the Fourier transform explicitly, we get

$$L(\boldsymbol{g}) = \sum_{ab} \int d\eta\, \tilde{W}_{ab}(\eta) \left| \int df e^{2\pi i \eta f} r_{ab}(f) \right|^2, \tag{4.60}$$

or, expanding,

$$L(\boldsymbol{g}) = \sum_{ab} \int d\eta\, \tilde{W}_{ab}(\eta) \int df \int df' e^{2\pi i \eta (f - f')} r_{ab}(f) r^*_{ab}(f'). \tag{4.61}$$

Rearranging terms, we get

$$L(\boldsymbol{g}) = \sum_{ab} \int df \int df' \left[\int d\eta\, \tilde{W}_{ab}(\eta) e^{-2\pi i \eta (f' - f)} \right] r_{ab}(f) r^*_{ab}(f'). \tag{4.62}$$

The bracketed term is just the inverse Fourier transform of the weighting function:

$$L(\boldsymbol{g}) = \sum_{ab} \int df \int df' W_{ab}(f' - f) r_{ab}(f) r^*_{ab}(f'). \tag{4.63}$$

Converting back into a discrete sum, we get

$$L(\boldsymbol{g}) = \sum_{ab} \sum_{ff'} W_{ab\,ff'} \left[v_{ab}(f) - g_a(f) g^*_b(f) m_{ab}(f) \right] \left[v_{ab}(f') - g_a(f') g^*_b(f') m_{ab}(f') \right]^*. \tag{4.64}$$

$W_{ab\,ff'}$ are elements of a matrix that mixes different frequency channels.

Chapter 5

UNIFIED CALIBRATION IN SIMULATION

In this chapter we explore the implications of the unified calibration approach presented in Chapter 4 through simulation. Unified calibration is a broad framework for developing calibration approaches. Here we use a simplified simulation to explore just one variant of unified calibration related to calibrating a perfectly redundant array. The simulation results demonstrate and approach to implementing unified calibration and illustrate that a unified calibration approach that combines redundant calibration with a sky model prior can improve calibration performance.

5.1 Simulation Setup

The simulation is based on an idealized redundant array of 36 antennas configured as a 6×6 square grid. The antennas are simulated as uniform circular apertures with diameters of 14 m and the array is close-packed: antenna centers are separated by 14 m. Figure 5.1 provides a schematic of the array configuration. We simulate visibilities from the GLEAM catalog Hurley-Walker et al. (2017) with the pyuvsim simulation package Lanman et al. (2019). The simulation corresponds to a zenith observation of the MWA's "EoR-0" field, a field studied extensively by the MWA's EoR team centered at Right Ascension 0 hours and Declination -27°.

We calibrate according to the framework presented in Equation 4.26. We assume perfect array redundancy and take $\boldsymbol{u}$ to correspond to visibilities from redundant baseline sets. $\mathbf{A}$ therefore takes the form of Equation 4.18.

In contrast to traditional redundant calibration, this calibration implementation accounts for non-zero correlations between baselines from different redundant sets. As we model

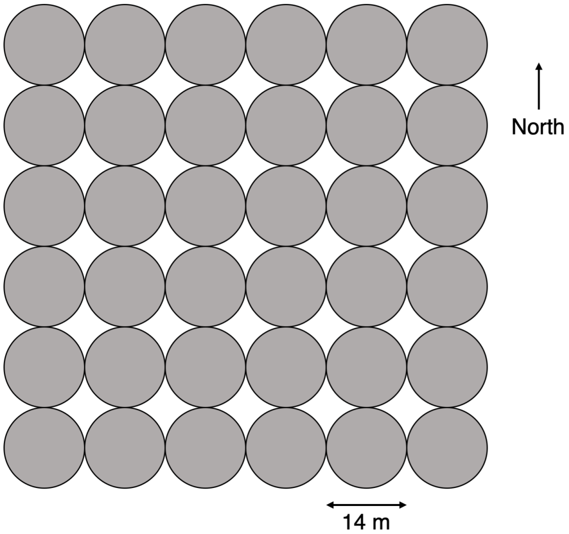


Figure 5.1: Schematic of the array layout used in simulation. The antennas are represented as uniform circular apertures with diameters of 14 m. The array comprises 36 antennas in a close-packed regular 6 × 6 grid.

antennas as circular apertures, each antenna response is given by

$$D(\boldsymbol{r}) \propto \begin{cases} 1, & \text{for } |\boldsymbol{r}| < a/2 \\ 0, & \text{otherwise} \end{cases}, \tag{5.1}$$

where $\boldsymbol{r}$ is the vector position on the ground from the antenna center and a is the antenna diameter (14 m in our simulation). The baseline response is evaluated by convolving the antenna responses:

$$B(\boldsymbol{x}) \propto \int_{-\infty}^{\infty} d^2\boldsymbol{r}\, D(\boldsymbol{r}) D^*(\boldsymbol{x} - \boldsymbol{r}), \tag{5.2}$$

where $\boldsymbol{x}$ is the vector distance from the baseline center in the uv plane. Evaluating this integral and normalizing such that the integrated baseline response is equal to 1, we get that

$$B(\boldsymbol{x}) = \frac{8}{\pi^2 a^2} \left[\cos^{-1}\left(\frac{|\boldsymbol{x}|}{a}\right) - \frac{|\boldsymbol{x}|}{a} \sqrt{1 - \left(\frac{|\boldsymbol{x}|}{a}\right)^2} \right] \tag{5.3}$$

for $|\boldsymbol{x}| \leq a$.

Following Equation 4.35, we define the baseline correlation matrix $\mathbf{C}_\mathrm{R}$ with elements

$$C_{\mathrm{R}\,jk} = \mathrm{corr}[\boldsymbol{u}_j, \boldsymbol{u}_k^*] = \frac{\int_{-\infty}^{\infty} d^2\boldsymbol{x}\, B(\boldsymbol{x})B(\boldsymbol{x} - \Delta\boldsymbol{x}_{jk})}{\int_{-\infty}^{\infty} d^2\boldsymbol{x}\, B^2(\boldsymbol{x})} \tag{5.4}$$

where $\Delta\boldsymbol{x}_{jk}$ is the uv separation between the centers of baselines j and k. From the definition of the correlation, the matrix is normalized such that the diagonal elements $C_{\mathrm{R}\,jj} = 1$. For our array, the closest non-redundant baselines are separated by 14 m. Numerically evaluating the integrals in Equation 5.4, we find that those baselines have correlations of 0.1617. The next closest baseline separations are 19.80 m; those baselines have correlations of 0.0176. Subsequently more distant baselines have no uv overlap and therefore no correlation. We define the model covariance matrix from these correlation values: $\mathbf{C}_\mathrm{M} = \sigma_\mathrm{M}^2 \mathbf{C}_\mathrm{R}$, where σ_M^2 quantifies the variance on the sky model errors. The inverse quantity $\mathbf{C}_\mathrm{M}^{-1} = \frac{1}{\sigma_\mathrm{M}^2}\mathbf{C}_\mathrm{R}^{-1}$. $\mathbf{C}_\mathrm{R}$ is full-rank; we invert it with the NUMPY linear algebra package.

Combining these elements, we produce $L(\boldsymbol{g}, \boldsymbol{u})$ for calibration of the form

$$L(\boldsymbol{g}, \boldsymbol{u}) = \frac{1}{\sigma_\mathrm{T}^2} \sum_{ab} \sum_{j} \left| v_{ab} - g_a g_b^* A_{ab\,j} u_j \right|^2 + \frac{1}{\sigma_\mathrm{M}^2} (\boldsymbol{u} - \boldsymbol{m})^\dagger \mathbf{C}_\mathrm{R}^{-1} (\boldsymbol{u} - \boldsymbol{m}). \tag{5.5}$$

As explained in §4.3, in the limit that $\sigma_\mathrm{M}^2 \to \infty$ the second term of $L(\boldsymbol{g}, \boldsymbol{u})$ disappears and we recover traditional redundant calibration. In the opposite limit that $\sigma_\mathrm{M}^2 \to 0$, $\boldsymbol{u} \to \boldsymbol{m}$ and we recover traditional sky-based calibration. We can also operate between these limits by setting σ_M^2 to a finite value that represents our confidence in the model visibilities $\boldsymbol{m}$.

Unified calibration has one degenerate parameter, corresponding to the overall phase of the gains. We constrain this term by requiring that the average complex phase of the gains is zero.

To simulate thermal noise, we inject an independent random signal into the each of the 630 simulated visibilities $\boldsymbol{v}$. The injected noise is drawn from a complex Gaussian distribution with a variance of 0.04 Jy^2; reflecting that, we set $\sigma_\mathrm{T}^2 = 0.04$ Jy^2. We explore two distinct classes of sky model errors. In §5.2 we introduce random error into the model visibilities $\boldsymbol{m}$, whereas in §5.3 we introduce errors in $\boldsymbol{m}$ by omitting faint sources from the sky model. In both cases the variance on the model visibility errors is 0.16 Jy^2, so we set $\sigma_\mathrm{M}^2 = 0.16$ Jy^2.

5.2 Simulating Random Model Error

In this simulation, we represent error on the model visibilities $\boldsymbol{m}$ by injecting an independent random signal into each of the 60 unique baseline measurements. The error is drawn from a complex Gaussian distribution with a variance of 0.16 Jy^2. This technique allows us to explore the implications of sky model error across a statistical ensemble of error realizations. In §5.3 we examine a more physical case in which error stems from faint missing sources in the sky model.

We implement calibration by minimizing Equation 5.5 with a SCIPY optimization routine. We perform 10,000 trials, corresponding to 100 independent realizations of each thermal noise and sky model error. Histograms of the calibrated gains $\hat{\boldsymbol{g}}$ over all 10,000 trials are plotted in the center column of Figure 5.2. The true gains $\hat{\boldsymbol{g}} = 1$, so deviations from 1 represent errors in the fit gains.

In addition, we calibrate the same noisy data and sky models using traditional redundant and sky-based calibration. The results from redundant calibration are plotted in the left column of Figure 5.2 while the results from sky-based calibration are plotted in the right column.

Of the three classes of calibration, we find that unified calibration performs best, producing the least gain error. The distribution of fit gains $\hat{\boldsymbol{g}}$ is more compact for unified calibration than for either redundant or sky-based calibration. Furthermore, its peak is better localized to $\hat{\boldsymbol{g}} = 1$. In redundant and sky-based calibration, the gain amplitudes are systematically biased low. In redundant calibration this feature results from the absolute calibration step, which constrains the overall amplitude of the gains. The bias is consistent with results in Chapter 3 and stems from decoherence of the data and model visibilities in the presence of model errors. We explore this effect in more depth in Appendix C.

The three rows in Figure 5.2 depict the same underlying histograms with different overlaid contours. The contours in the top row highlight the distribution of the full set of 10,000 calibration trials. The solid white contour and the dashed white contour enclose an estimated

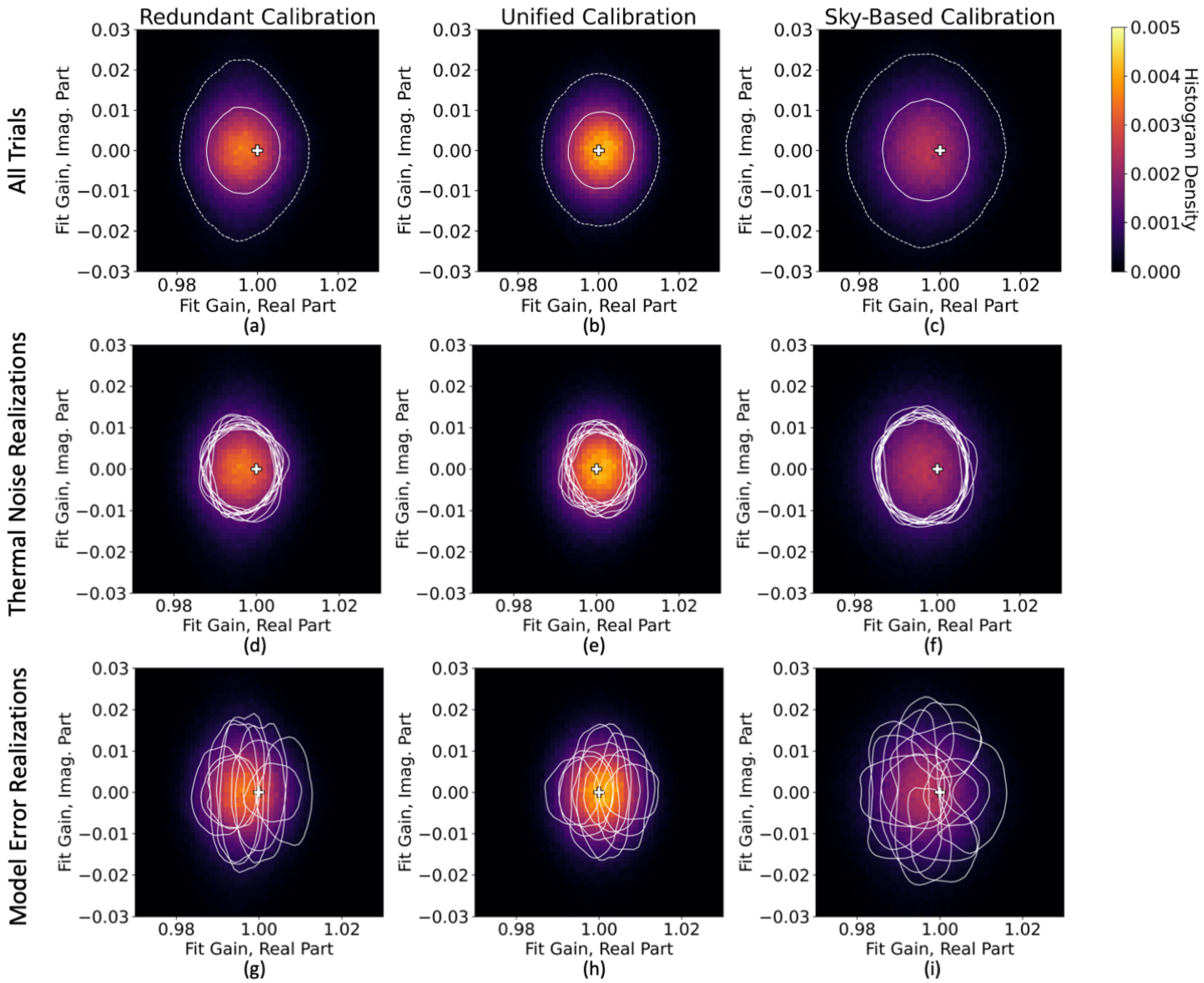


Figure 5.2: Gains calculated from a calibration simulation with random sky model error. (Caption continued.)

Figure 5.2: 2-D histograms of the fit gains $\hat{\boldsymbol{g}}$ from a simulation of the array depicted in Figure 5.1 in the presence of random thermal noise and sky model error. The true gains are 1 (marked by a white plus sign), and deviations from 1 correspond to errors in the fit values. The three columns show the results of three different classes of calibration: traditional redundant calibration on the left, unified calibration implemented by minimizing Equation 5.5 in the middle, and traditional sky-based calibration on the right. The histograms depict 10,000 calibration trials corresponding to 100 separate realizations of injected thermal noise and 100 realizations of sky model error. Each row depicts identical histograms with different contours overlaid. For the top row, the white solid and dashed contours enclose an estimated 50% and 90% of the full distribution, respectively. We see that unified calibration performs better than either traditional redundant or sky-based calibration as the histogram distribution is better localized to the correct value of $\hat{\boldsymbol{g}} = 1$. In the middle row, each of the 10 white contours corresponds to a single realization of thermal noise. Each contour encloses 50% of the distribution of fit gains from that noise realization. In the bottom row, each of the 10 contours corresponds to a single realization of model error, again enclosing 50% of the distribution from that realization. This demonstrates that the distribution of fit gains is highly sensitive to the sky model error.

50% and 90% of the distribution, respectively, as calculated from a Gaussian Kernel Density Estimator (KDE).

In the second row of Figure 5.2 each contour describes the distribution of fit gains derived from a single realization of thermal noise. The contours enclose an estimated 50% of the distribution calculated across all 100 realizations of model error. We plot 10 such contours, randomly selected from the 100 total realizations of thermal noise. While there is some variation between the contours, the variation is small. Each realization of thermal noise produces a distribution of fit gain errors representative of the overall distribution.

In contrast, in the bottom row of Figure 5.2 each contour corresponds a single realization of sky model error. These contours are calculated to enclose 50% of the distribution of fit gains across all 100 realizations of thermal noise. We see that independent draws of sky model error produce very different gain distributions. This is because model error is coherent across measurements within a redundant baseline set. As regular arrays have few independent measurements of the sky signal, sky model error affecting just one baseline

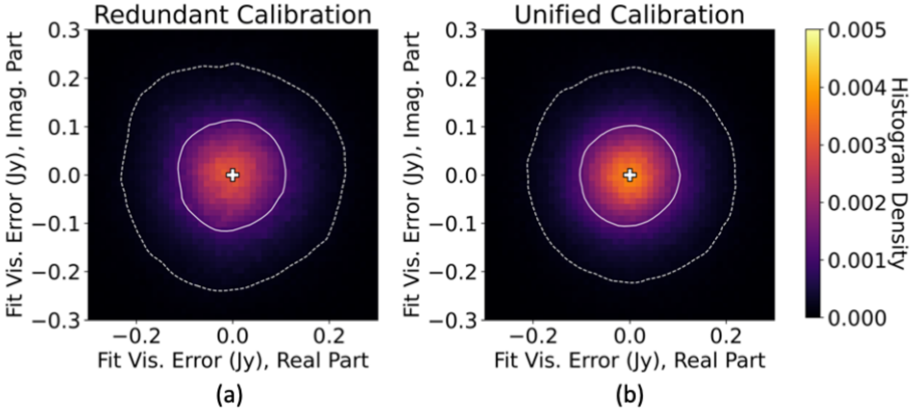


Figure 5.3: Complex error on fit visibility parameters $\hat{\boldsymbol{u}}$ from the simulation presented in Figure 5.2. (a) presents the results from traditional redundant calibration and (b) presents the results from unified calibration. Traditional sky-based calibration does not fit visibility parameters. The white plus signs mark the position of zero error, and the white solid and dashed contours enclose an estimated 50% and 90% of the data, respectively. Unified calibration produces slightly more precise values of $\hat{\boldsymbol{u}}$ than traditional redundant calibration.

mode can impact calibration for a large fraction of the measured visibilities. Calibration is typically implemented for observations of a single field on the sky, so in practice we work with a single realization of sky model error. Unlike with thermal noise, the impact of sky model error on calibration solutions will not average down across long data integrations.

We can likewise compare the fit visibility values $\hat{\boldsymbol{u}}$ between traditional redundant calibration and unified calibration. Figure 5.3 presents histograms of the error in $\hat{\boldsymbol{u}}$. Figure 5.3(a) corresponds to traditional redundant calibration and Figure 5.3(b) corresponds to unified calibration. Traditional sky-based calibration does not fit visibility values. Figure 5.3 demonstrates that unified calibration produces slightly less fit visibility error than redundant calibration, as the distribution is marginally more compact. That said, the errors in the fit gains $\hat{\boldsymbol{g}}$ are more relevant for evaluating calibration performance because the gains are applied to the data in calibration.

These simulations illustrate that, in a simplified case, the calibration framework presented in §4.3 improves upon traditional calibration approaches. Traditional sky-based calibration does not fit visibility values $\boldsymbol{u}$ and instead only uses the model visibilities $\boldsymbol{m}$, which are susceptible to errors from the sky model. Unified and redundant calibration approaches introduce additional calibration degrees-of-freedom through $\boldsymbol{u}$ and thereby allow calibration to fit out sky model errors.

Redundant calibration uses the model visibilities $\boldsymbol{m}$ for absolute calibration only. It fits the relative calibration parameters entirely from the data and imposing agreement between redundant measurements. However, thermal noise introduces error into the relative calibration parameters. Unified calibration uses the model visibilities as a prior on calibration solutions. The prior is weighted by the measured uncertainty on the sky model. This further constrains redundant calibration's relative calibration parameters and mitigates error from thermal noise.

The calibration simulations presented in this section illustrate one implementation style of the calibration framework described in §4.3. The simulated instrument is perfectly redundant and we assume no errors in the instrument model. Instead, calibration errors result only from error in the model visibilities $\boldsymbol{m}$ and from thermal noise. The results in Figure 5.2 demonstrate that, for our simplified simulation, using a unified calibration framework based on a physically-motivated instrument model produces better calibration results than traditional methods. In the next section, we demonstrate that unified calibration similarly improves calibration performance in the case that sky model errors stem from faint missing sources.

5.3 Simulating Model Incompleteness

In §5.2 we explored calibration across an ensemble of random sky model errors. As a more physically-motivated simulation, we explore the case that sky model errors stem from missing faint missing sources. This technique follows approaches used in Barry et al. (2016) and Byrne et al. (2019) and represents the fact that real source catalogs always have a completeness limit dictated by the measurement sensitivity.

To represent this source of model error we once again simulate visibilities from the

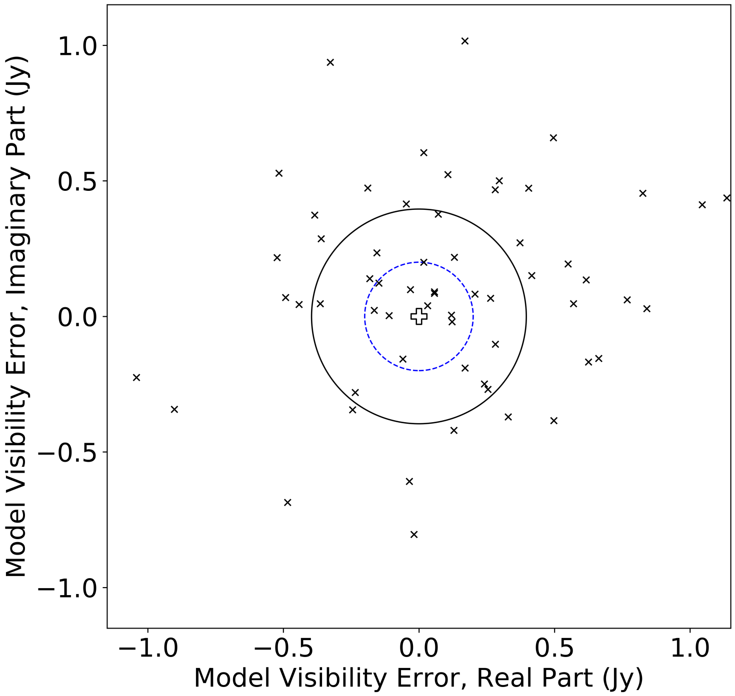


Figure 5.4: Scatter plot of the errors in the simulated model visiblities $\boldsymbol{m}$ (denoted with black 'X' symbols). Data visibilities are simulated with the full GLEAM catalog and have an average amplitude of 5.58 Jy; model visibilities are simulated from only the GLEAM sources brighter than 100 mJy. The errors in the model visibilities stem from these missing faint sources. The white plus sign marks the position of zero error. The black solid contour has a radius equal to the standard deviation of the model visibility errors. The blue dashed contour has a radius equal to the standard deviation of the complex Gaussian distribution from which we simulate thermal noise.

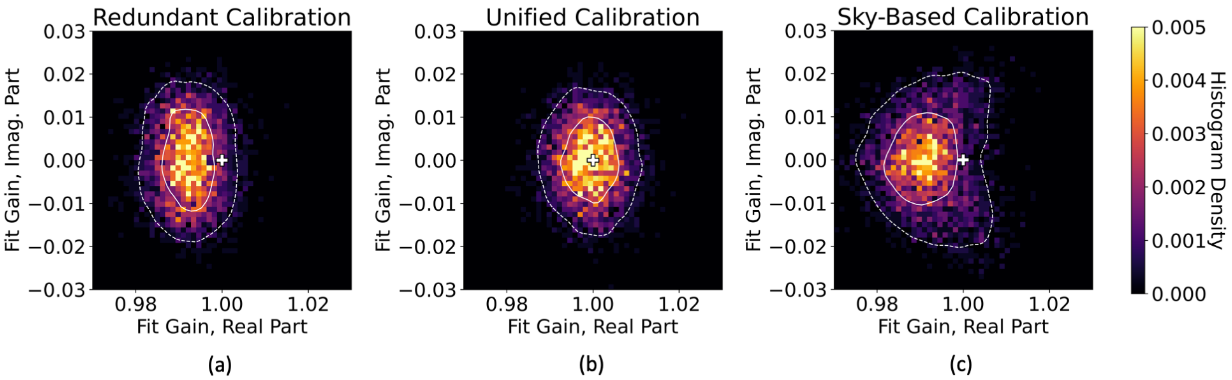


Figure 5.5: 2-D histograms of the fit gains $\hat{\boldsymbol{g}}$ from a simulation of the array depicted in Figure 5.1. Sky model error is represented by excluding faint sources in the simulated catalog. The true gains are 1 (marked by a white plus sign), and deviations from 1 correspond to errors in the fit values. The three plots show the results of three different classes of calibration: (a) redundant calibration, (b) unified calibration, and (c) sky-based calibration. The white solid and dashed contours enclose an estimated 50% and 90% of the data, respectively. The histogram distribution in (b) is better localized to the zero error point than either (a) or (c), indicating that unified calibration delivers better calibration solutions than either traditional redundant or sky-based calibration.

GLEAM catalog, this time omitting sources with flux densities less than 100 mJy. We take these simulated visibilities to be the model visibilities $\boldsymbol{m}$. The missing faint sources introduce disparities between $\boldsymbol{m}$ and the visibilities simulated with the full catalog (plotted as black ‘X’ symbols in Figure 5.4). The variance of these disparities is 0.16 Jy^2, compared to the average visibility amplitude of 5.58 Jy. The black circular contour in Figure 5.4 has a radius equal to the standard deviation of the errors.

As in §5.3, we calibrate across 100 independent realizations of injected thermal noise drawn from a complex Gaussian distribution. The resulting calibrated gains $\hat{\boldsymbol{g}}$ are presented as histograms in Figure 5.5. Figure 5.5a corresponds to traditional redundant calibration, Figure 5.5b corresponds to unified calibration, and Figure 5.5c corresponds to sky-based calibration. As in Figure 5.2, we find that unified calibration performs better than each redundant or sky-based calibration. The fit gain distribution is most compact for unified

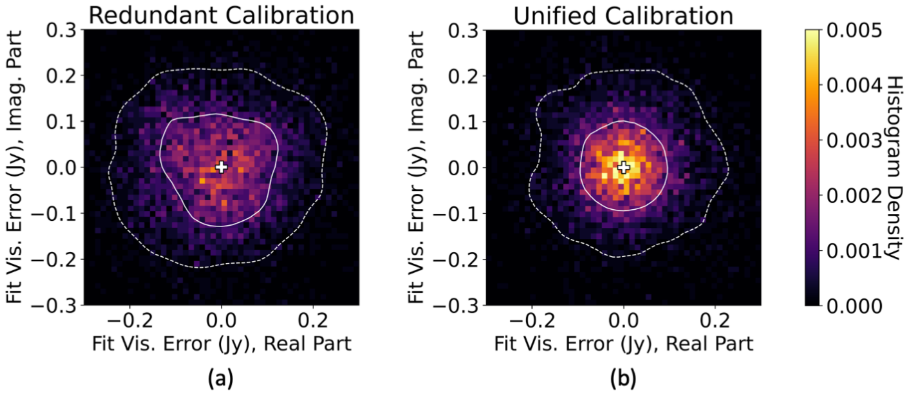


Figure 5.6: Complex error on the fit visibility parameters $\hat{\boldsymbol{u}}$ from the simulation presented in Figure 5.5 for redundant calibration (a) and unified calibration (b). The white solid and dashed contours enclose an estimated 50% and 90% of the data, respectively.

calibration, and redundant and sky-based calibration produce gains with amplitudes that are biased low.

In Figure 5.6 we plot histograms of the error on the fit visibility values $\hat{\boldsymbol{u}}$ for redundant and unified calibration. As in Figure 5.3, we find that unified calibration produces less fit visibility error than redundant calibration.

The results of this simulation indicate that, just as unified calibration improved calibration in the presence of the random model error simulated in §5.2, it can reduce calibration error in the more physical scenario of missing faint sources in the sky model. The unified calibration framework described by Equation 5.5 assumes that the error on the model visibilities $\boldsymbol{m}$ is Gaussian. This could be a poor approximation for some classes of sky model errors, for example, if the sky model error is dominated by mismodeling of a single bright source. Further work could develop more accurate models of sky model error based on specific knowledge of the sky. However, we expect that this assumption is appropriate for the case of sky model error from many faint missing sources distributed across the sky (and Figure 5.4 gives no indication that the errors on $\boldsymbol{m}$ are dramatically non-Gaussian). The results in this section show that the calibration framework represented by Equation 5.5 can improve upon traditional calibration methods in the presence of missing sky model sources.

Chapter 6

THE FHD POLARIZED IMAGING PIPELINE

In the Chapter 4 we presented a radical rethinking of the calibration problem, developing a new statistical framework for calibration called unified calibration. In Chapter 5, we showed that this new calibration approach can deliver better calibration performance than the techniques widely used in the field. However, none of the calibration approaches presented eliminate the need for a highly complete and accurate sky model.

One of the major limitations of the sky models typically used in calibration is that they do not include large-scale, diffuse structure. In the next chapter, Chapter 7, we present a new map of diffuse emission designed to improve calibration sky models in the Southern Hemisphere. Diffuse radio emission is distinct from compact sources because it appears highly polarized, and capturing this polarization is an important component of achieving precision sky modeling. Before embarking on this project, the FHD data analysis software did not support polarized imaging. We built polarimetry capabilities into FHD, using an innovative technique that fully accounts for widefield projection effects.

In this chapter, we present FHD's polarized imaging techniques. First, in §6.1 we define aspects of radio astronomy's polarization formalism: the Stokes parameters, the coherency, the Jones matrix, and the Mueller matrix. §6.2 introduces the instrumental basis, a mathematical construct that is a key component of FHD's polarized imaging approach. §6.3 revisits interferometric imaging, first discussed in §2.2, and extend it to the fully polarized case. §6.4 discusses FHD's polarized calibration.

6.1 Polarization Formalism

6.1.1 The Stokes Parameters

Polarized emission is often described with respect to the Stokes parameters I, Q, U, and V (Stokes, 1851). Stokes I corresponds to the total intensity, Stokes Q and U correspond to linear polarization, and Stokes V corresponds to circular polarization. The Stokes parameters in a given direction on the sky $\boldsymbol{\theta}$ are related to the coherency vector $\boldsymbol{S}(\boldsymbol{\theta})$ via the relationship

$$\begin{bmatrix} I(\boldsymbol{\theta}) \\ Q(\boldsymbol{\theta}) \\ U(\boldsymbol{\theta}) \\ V(\boldsymbol{\theta}) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & i & -i \end{bmatrix} \boldsymbol{S}(\boldsymbol{\theta}). \tag{6.1}$$

In turn, the coherency vector is related to the electric field by

$$\boldsymbol{S}(\boldsymbol{\theta}) = \begin{bmatrix} \langle |E_{\mathrm{R}}(\boldsymbol{\theta})|^2 \rangle \\ \langle |E_{\mathrm{D}}(\boldsymbol{\theta})|^2 \rangle \\ \langle E_{\mathrm{R}}(\boldsymbol{\theta}) E^*_{\mathrm{D}}(\boldsymbol{\theta}) \rangle \\ \langle E^*_{\mathrm{R}}(\boldsymbol{\theta}) E_{\mathrm{D}}(\boldsymbol{\theta}) \rangle \end{bmatrix}, \tag{6.2}$$

where $E_{\mathrm{R}}(\boldsymbol{\theta})$ and $E_{\mathrm{D}}(\boldsymbol{\theta})$ are the components of the electric field aligned with the RA and Dec directions on the sky. The angled brackets $\langle\rangle$ denote the time average.

Note that here we have decided to define the Stokes parameters with respect to the RA/Dec coordinate system. This is an arbitrary choice, and it is equally valid to define polarization modes with respect to another orthogonal basis. However, because the Stokes parameters are basis-dependent, it is critical that any polarized measurements specify the basis used. The poles of the RA/Dec basis mean that the Stokes parameters are undefined at the North and South poles.

6.1.2 The Jones Matrix

The Jones matrix is a 2 × 2 complex matrix that represents the polarized antenna response to the sky (Jones, 1941). The matrix transforms between the true electric field on the sky and the electric field measured by the instrument (Hamaker et al., 1996; Sault et al., 1996; Hamaker and Bregman, 1996; Hamaker, 2000, 2006; Ord et al., 2010):

$$\begin{bmatrix} \epsilon_{j\mathrm{P}}(\boldsymbol{\theta}) \\ \epsilon_{j\mathrm{Q}}(\boldsymbol{\theta}) \end{bmatrix} = \mathbf{J}_j^{\mathrm{ZA}}(\boldsymbol{\theta}) \begin{bmatrix} E_{\mathrm{Z}}(\boldsymbol{\theta}) \\ E_{\mathrm{A}}(\boldsymbol{\theta}) \end{bmatrix}. \tag{6.3}$$

Here $E_{\mathrm{Z}}(\boldsymbol{\theta})$ and $E_{\mathrm{A}}(\boldsymbol{\theta})$ are components of the electric field aligned with the zenith angle and azimuth directions, respectively. $\epsilon_{j\mathrm{P}}(\boldsymbol{\theta})$ and $\epsilon_{j\mathrm{Q}}(\boldsymbol{\theta})$ represent the contribution of emission from sky direction $\boldsymbol{\theta}$ to the measurements made by antenna j. The subscripts P and Q correspond to the two instrumental polarization modes. $\mathbf{J}_j^{\mathrm{ZA}}(\boldsymbol{\theta})$ is the Jones matrix for antenna j; the superscript ZA indicates that it corresponds to the zenith angle/azimuth basis on the sky, the usual basis convention for reporting the Jones matrix.

This formalism is fully general in that it assumes dual-polarization antenna but makes no further assumptions about the antenna type. The two antenna polarizations could be orthogonal but need not be. They could measure linearly or circularly polarized modes. Note that here we have defined a per-antenna Jones matrix that depends on the antenna index j: antennas need not have homogeneous polarized responses. The Jones matrix is also direction-dependent, and we require a 2 × 2 Jones matrix for each location on the sky $\boldsymbol{\theta}$.

Equation 6.3 presents the Jones matrix in the zenith angle/azimuth coordinate system, but in Equation 6.2 we define the coherency with respect to the RA/Dec basis. Our analysis is therefore simplified if we define a Jones matrix $\mathbf{J}_j^{\mathrm{RD}}(\boldsymbol{\theta})$ with respect to RA/Dec. The transformation between bases involves rotating by the direction-dependent parallactic angle $\phi(\boldsymbol{\theta})$:

$$\begin{bmatrix} \boldsymbol{e}_{Z}(\boldsymbol{\theta}) \\ \boldsymbol{e}_{A}(\boldsymbol{\theta}) \end{bmatrix} = \begin{bmatrix} \sin[\phi(\boldsymbol{\theta})] & -\cos[\phi(\boldsymbol{\theta})] \\ -\cos[\phi(\boldsymbol{\theta})] & -\sin[\phi(\boldsymbol{\theta})] \end{bmatrix} \begin{bmatrix} \boldsymbol{e}_{\mathrm{R}}(\boldsymbol{\theta}) \\ \boldsymbol{e}_{\mathrm{D}}(\boldsymbol{\theta}) \end{bmatrix}. \tag{6.4}$$

Here $\boldsymbol{e}_{\mathrm{R}}(\boldsymbol{\theta})$, $\boldsymbol{e}_{\mathrm{D}}(\boldsymbol{\theta})$, $\boldsymbol{e}_{Z}(\boldsymbol{\theta})$, and $\boldsymbol{e}_{A}(\boldsymbol{\theta})$ are unit vectors in the RA, Dec, zenith angle, and azimuth directions, respectively. Appendix D presents a derivation of the parallactic angle. We then get that

$$\mathbf{J}_j^{\mathrm{RD}}(\boldsymbol{\theta}) = \mathbf{J}_j^{\mathrm{ZA}}(\boldsymbol{\theta}) \begin{bmatrix} \sin[\phi(\boldsymbol{\theta})] & -\cos[\phi(\boldsymbol{\theta})] \\ -\cos[\phi(\boldsymbol{\theta})] & -\sin[\phi(\boldsymbol{\theta})] \end{bmatrix} \tag{6.5}$$

where

$$\begin{bmatrix} \epsilon_{j\mathrm{P}}(\boldsymbol{\theta}) \\ \epsilon_{j\mathrm{Q}}(\boldsymbol{\theta}) \end{bmatrix} = \mathbf{J}_j^{\mathrm{RD}}(\boldsymbol{\theta}) \begin{bmatrix} E_{\mathrm{R}}(\boldsymbol{\theta}) \\ E_{\mathrm{D}}(\boldsymbol{\theta}) \end{bmatrix}. \tag{6.6}$$

6.1.3 The Mueller Matrix

The Mueller matrix is a 4×4 complex matrix formed by the Kronecker product of two Jones matrices:

$$\mathbf{M}_{jk}(\boldsymbol{\theta}) = \mathbf{J}_j^{\mathrm{RD}}(\boldsymbol{\theta}) \otimes \mathbf{J}_k^{\mathrm{RD}*}(\boldsymbol{\theta}). \tag{6.7}$$

Here we choose to define the Mueller matrix with respect to the RA/Dec basis on the sky to align with the convention in Equation 6.2. We can then write the visibilities formed from correlating antennas j and k in terms of the Mueller matrix:

$$\boldsymbol{v}_{jk} = \int_{-\infty}^{\infty} d^2\boldsymbol{\theta}\, \mathbf{M}_{jk}(\boldsymbol{\theta}) \boldsymbol{S}(\boldsymbol{\theta})\, e^{2\pi i \boldsymbol{\theta} \cdot \boldsymbol{x}_{jk}}. \tag{6.8}$$

Here $\boldsymbol{v}_{jk}$ is a 4-element vector with components

$$\boldsymbol{v}_{jk} = \begin{bmatrix} v_{jk\mathrm{PP}} \\ v_{jk\mathrm{QQ}} \\ v_{jk\mathrm{PQ}} \\ v_{jk\mathrm{QP}} \end{bmatrix} \tag{6.9}$$

where, for example, $v_{jk\mathrm{PQ}}$ represents the visibility formed by correlating the P-polarized signal from antenna j with the Q-polarized signal from antenna k. Equation 6.8 is the polarized analogue to Equation 2.7.

6.2 The Instrumental Basis

The Jones matrix can be decomposed into a product of two matrices:

$$\mathbf{J}_j^{\mathrm{RD}}(\boldsymbol{\theta}) = \mathbf{D}_j(\boldsymbol{\theta})\mathbf{K}_j(\boldsymbol{\theta}). \tag{6.10}$$

$\mathbf{D}_j(\boldsymbol{\theta})$ is a diagonal matrix that encodes the amplitude of the instrumental response:

$$\mathbf{D}_j(\boldsymbol{\theta}) = \begin{bmatrix} \sqrt{|J_{\mathrm{RP}j}(\boldsymbol{\theta})|^2 + |J_{\mathrm{DP}j}(\boldsymbol{\theta})|^2} & 0 \\ 0 & \sqrt{|J_{\mathrm{RQ}j}(\boldsymbol{\theta})|^2 + |J_{\mathrm{DQ}j}(\boldsymbol{\theta})|^2} \end{bmatrix} \tag{6.11}$$

where the elements of the Jones matrix are given by

$$\mathbf{J}_j^{\mathrm{RD}}(\boldsymbol{\theta}) = \begin{bmatrix} J_{\mathrm{RP}j}(\boldsymbol{\theta}) & J_{\mathrm{DP}j}(\boldsymbol{\theta}) \\ J_{\mathrm{RQ}j}(\boldsymbol{\theta}) & J_{\mathrm{DQ}j}(\boldsymbol{\theta}) \end{bmatrix}. \tag{6.12}$$

The two elements of $\mathbf{D}_j(\boldsymbol{\theta})$ give the sensitivity of each the P and Q polarizations of antenna j to unpolarized emission from location $\boldsymbol{\theta}$ on the sky.

$\mathbf{K}_j(\boldsymbol{\theta})$ captures the instrument's response to polarized emission. If $\mathbf{K}_j(\boldsymbol{\theta})$ is identical for all antennas (such as for homogeneous arrays), then $\mathbf{K}_j(\boldsymbol{\theta}) = \mathbf{K}(\boldsymbol{\theta})$ for all antennas j. We can then define a new "instrumental basis" on the sky, where $\mathbf{K}(\boldsymbol{\theta})$ transforms between the usual RA/Dec basis and our new instrumental basis:

$$\begin{bmatrix} \boldsymbol{e}_{\mathrm{P}}(\boldsymbol{\theta}) \\ \boldsymbol{e}_{\mathrm{Q}}(\boldsymbol{\theta}) \end{bmatrix} = \mathbf{K}(\boldsymbol{\theta}) \begin{bmatrix} \boldsymbol{e}_{\mathrm{R}}(\boldsymbol{\theta}) \\ \boldsymbol{e}_{\mathrm{D}}(\boldsymbol{\theta}) \end{bmatrix}. \tag{6.13}$$

Here $\boldsymbol{e}_{\mathrm{P}}(\boldsymbol{\theta})$ and $\boldsymbol{e}_{\mathrm{Q}}(\boldsymbol{\theta})$ are unit vectors aligned with the polarization directions that produce the maximal response from the P and Q polarizations of each antenna. $\boldsymbol{e}_{\mathrm{P}}(\boldsymbol{\theta})$ and $\boldsymbol{e}_{\mathrm{Q}}(\boldsymbol{\theta})$ are generally not orthogonal so $\mathbf{K}(\boldsymbol{\theta})$ is not a unitary matrix.

We can express the Mueller matrix in terms of this Jones matrix decomposition:

$$\mathbf{M}_{jk}(\boldsymbol{\theta}) = \mathbf{B}_{jk}(\boldsymbol{\theta})\mathbf{L}(\boldsymbol{\theta}) \tag{6.14}$$

where

$$\mathbf{L}(\boldsymbol{\theta}) = \mathbf{K}(\boldsymbol{\theta}) \otimes \mathbf{K}^*(\boldsymbol{\theta}) \tag{6.15}$$

and

$$\mathbf{B}_{jk}(\boldsymbol{\theta}) = \mathbf{D}_j(\boldsymbol{\theta}) \otimes \mathbf{D}_k(\boldsymbol{\theta}). \tag{6.16}$$

Here $\mathbf{B}_{jk}(\boldsymbol{\theta})$ is a diagonal 4×4 matrix that defines the amplitude of the response of baseline $\{j, k\}$ (note that in Equation 6.11 we have defined $\mathbf{D}(\boldsymbol{\theta})$ to be real).

We define a new coherency vector in the instrumental basis

$$\boldsymbol{S}_{\text{inst}}(\boldsymbol{\theta}) = \begin{bmatrix} \langle |E_{\text{P}}(\boldsymbol{\theta})|^2 \rangle \\ \langle |E_{\text{Q}}(\boldsymbol{\theta})|^2 \rangle \\ \langle E_{\text{P}}(\boldsymbol{\theta}) E^*_{\text{Q}}(\boldsymbol{\theta}) \rangle \\ \langle E^*_{\text{P}}(\boldsymbol{\theta}) E_{\text{Q}}(\boldsymbol{\theta}) \rangle \end{bmatrix} = \mathbf{L}(\boldsymbol{\theta}) \boldsymbol{S}(\boldsymbol{\theta}) \tag{6.17}$$

where $E_{\text{P}}(\boldsymbol{\theta})$ and $E_{\text{Q}}(\boldsymbol{\theta})$ are the components of the electric field aligned with unit vectors $\boldsymbol{e}_{\text{P}}(\boldsymbol{\theta})$ and $\boldsymbol{e}_{\text{Q}}(\boldsymbol{\theta})$, respectively. The instrumental coherency $\boldsymbol{S}_{\text{inst}}(\boldsymbol{\theta})$ is the related to the "pseudo-Stokes" polarization modes, defined as

$$\begin{bmatrix} I_{\text{pseudo}}(\boldsymbol{\theta}) \\ Q_{\text{pseudo}}(\boldsymbol{\theta}) \\ U_{\text{pseudo}}(\boldsymbol{\theta}) \\ V_{\text{pseudo}}(\boldsymbol{\theta}) \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & i & -i \end{bmatrix} \boldsymbol{S}_{\text{inst}}(\boldsymbol{\theta}). \tag{6.18}$$

Pseudo-Stokes should not be confused with true Stokes modes, which must be defined with respect to an orthogonal basis.

FHD's polarized analysis pipeline consists of reconstructing the instrumental coherency $\boldsymbol{S}_{\text{inst}}(\boldsymbol{\theta})$ and then transforming into the true coherency $\boldsymbol{S}(\boldsymbol{\theta})$ by inverting Equation 6.17. We then calculate the Stokes polarization parameters from Equation 6.1.

6.3 Imaging

FHD's imaging pipeline follows the optimal mapmaking formalism (Tegmark et al., 1998; Bhatnagar et al., 2008; Morales and Matejek, 2009). In this approach, visibilities are gridded to the uv plane using the instrumental response kernel. FHD grids in the instrumental basis, forming uv planes corresponding to each instrumental polarization mode. We typically pixelate the uv plane at half wavelength spacing.

For example, gridding the PQ-polarized visibilities produces the following reconstructed uv plane:

$$\hat{\widetilde{S}}_{\mathrm{app\,PQ}}(\boldsymbol{x}) = \frac{\sum_{jk} \widetilde{B}_{jk\,\mathrm{PQ}}(\boldsymbol{x}_{jk} - \boldsymbol{x})\, v_{jk\,\mathrm{PQ}}}{\sum_{jk} \widetilde{B}_{jk\,\mathrm{PQ}}(\boldsymbol{x}_{jk} - \boldsymbol{x})}. \tag{6.19}$$

Here $S_{\mathrm{app\,PQ}}(\boldsymbol{\theta})$ is the apparent sky in the PQ polarization, where the apparent sky is windowed by the instrumental beam response. The tilde symbol $\widetilde{}$ denotes the Fourier transformed quantity: $\widetilde{S}_{\mathrm{app\,PQ}}(\boldsymbol{x}) = \mathcal{FT}\left[S_{\mathrm{app\,PQ}}(\boldsymbol{\theta})\right]$. $\boldsymbol{x}$ denotes the uv plane coordinate. The hat symbol $\hat{}$ indicates that this is the reconstructed estimate of the apparent uv plane. $B_{jk\,\mathrm{PQ}}(\boldsymbol{\theta})$ is the element of the diagonal matrix $\mathbf{B}_{jk}(\boldsymbol{\theta})$ that corresponds to the PQ polarization; once again, $\widetilde{B}_{jk\,\mathrm{PQ}}(\boldsymbol{x})$ is the uv plane transformation of that quantity. j and k index antennas, and $\sum_{jk}$ denotes the sum over all baselines.

The denominator in Equation 6.19 is known as the weights. If the instrument does not have complete uv plane measurement coverage, the weights will be equal to zero in certain uv plane locations. We set the expression in Equation 6.19 equal zero where the weights are zero.

Transforming this gridded uv plane into sky coordinates gives

$$\hat{S}_{\mathrm{app\,PQ}}(\boldsymbol{\theta}) = \mathcal{FT}^{-1}\left[\frac{\sum_{jk} \widetilde{B}_{jk\,\mathrm{PQ}}(\boldsymbol{x}_{jk} - \boldsymbol{x})\, v_{jk\,\mathrm{PQ}}}{\sum_{jk} \widetilde{B}_{jk\,\mathrm{PQ}}(\boldsymbol{x}_{jk} - \boldsymbol{x})}\right]. \tag{6.20}$$

This apparent sky map can then be converted into an estimate of the true sky, in the instrumental PQ polarization, by undoing the beam weighting:

$$\hat{S}_{\mathrm{inst\,PQ}}(\boldsymbol{\theta}) = \frac{1}{B_{\mathrm{avg\,PQ}}(\boldsymbol{\theta})}\mathcal{FT}^{-1}\left[\frac{\sum_{jk} \widetilde{B}_{jk\,\mathrm{PQ}}(\boldsymbol{x}_{jk} - \boldsymbol{x})\, v_{jk\,\mathrm{PQ}}}{\sum_{jk} \widetilde{B}_{jk\,\mathrm{PQ}}(\boldsymbol{x}_{jk} - \boldsymbol{x})}\right]. \tag{6.21}$$

Here $\hat{S}_{\mathrm{inst\,PQ}}(\boldsymbol{\theta})$ is an element of the instrumental coherency estimate, as defined in Equation 6.17. $B_{\mathrm{avg\,PQ}}(\boldsymbol{\theta})$ is the average PQ instrumental response, averaged across baselines. For a homogeneous array, $B_{jk\,\mathrm{PQ}}(\boldsymbol{\theta}) = B_{\mathrm{avg\,PQ}}(\boldsymbol{\theta}) = B_{\mathrm{PQ}}(\boldsymbol{\theta})$ for all baselines $\{j, k\}$.

The PP-, QQ- and QP-polarized visiblities are likewise gridded to their respective *uv* planes and transformed to produce an estimate of the instrumental coherency $\hat{\boldsymbol{S}}_{\text{inst}}(\boldsymbol{\theta})$. This polarized image is then converted into the true coherency via Equation 6.17 and subsequently into Stokes polarization parameters via Equation 6.1. Note that because the instrumental beam response is polarization-dependent, we cannot transform the apparent sky estimate $\hat{\boldsymbol{S}}_{\text{app}}(\boldsymbol{\theta})$ into the true coherency or Stokes polarization basis. The apparent sky is intrinsically defined in the instrumental polarization basis.

A unique aspect of this polarization imaging pipeline is that each visibility is gridded to just one *uv* plane. PQ-polarized visiblities contribute to the PQ-polarized *uv* plane only. This is a result of the fact that we have defined $\mathbf{B}_{jk}(\boldsymbol{\theta})$ to be diagonal. $\mathbf{B}_{jk}(\boldsymbol{\theta})$ has four nonzero elements, corresponding to four gridding operations. As gridding is the most computationally intensive step in the imaging pipeline, this allows for significant computational savings over other polarized imaging pipelines.

Note that here we have described polarized imaging using FHD's optimal weighting scheme and diffuse normalization, as described in §2.2. FHD fully supports polarized imaging with natural or uniform weighting and with point source normalization. We focus on optimal weighting and diffuse normalization here because they were used to produce the polarized diffuse map presented in Chapter 7.

6.4 Polarized Calibration

Direction-independent polarized calibration is based on the polarized measurement equation:

$$\boldsymbol{v}_{jk} = \mathbf{G}_{jk}\boldsymbol{u}_{jk} + \boldsymbol{n}_{jk}. \tag{6.22}$$

Here $\boldsymbol{v}_{jk}$ is the 4-element measured visibility vector given by Equation 6.9, $\boldsymbol{u}_{jk}$ is the theoretical true visibility vector, and $\boldsymbol{n}_{jk}$ is the noise on the measurement. $\mathbf{G}_{jk}$ is a 4×4 gain matrix. The gain matrix serves as a direction-independent modification of the Mueller matrix. Empirically fitting the gain terms constrains uncertainties in the modeled Mueller matrix, helping it to better align with the true Mueller matrix that governs the instrumental response.

6.4.1 Calibration Parameterization

In its most general form, G_{jk} is given by

$$\mathsf{G}_{jk} = \begin{bmatrix} g_{j\mathrm{PP}}\, g^*_{k\mathrm{PP}} & g_{j\mathrm{PQ}}\, g^*_{k\mathrm{PQ}} & g_{j\mathrm{PP}}\, g^*_{k\mathrm{PQ}} & g_{j\mathrm{PQ}}\, g^*_{k\mathrm{PP}} \\ g_{j\mathrm{QP}}\, g^*_{k\mathrm{QP}} & g_{j\mathrm{QQ}}\, g^*_{k\mathrm{QQ}} & g_{j\mathrm{QP}}\, g^*_{k\mathrm{QQ}} & g_{j\mathrm{QQ}}\, g^*_{k\mathrm{QP}} \\ g_{j\mathrm{PP}}\, g^*_{k\mathrm{QP}} & g_{j\mathrm{PQ}}\, g^*_{k\mathrm{QQ}} & g_{j\mathrm{PP}}\, g^*_{k\mathrm{QQ}} & g_{j\mathrm{PQ}}\, g^*_{k\mathrm{QP}} \\ g_{j\mathrm{QP}}\, g^*_{k\mathrm{PP}} & g_{j\mathrm{QQ}}\, g^*_{k\mathrm{PQ}} & g_{j\mathrm{QP}}\, g^*_{k\mathrm{PQ}} & g_{j\mathrm{QQ}}\, g^*_{k\mathrm{PP}} \end{bmatrix} . \tag{6.23}$$

Here $g_{j\mathrm{PP}}$ and $g_{j\mathrm{QQ}}$ are the complex gains of the P and Q polarizations of antenna j, respectively. $g_{j\mathrm{PQ}}$ and $g_{j\mathrm{QP}}$ are the cross gains. $g_{j\mathrm{PQ}}$, for example, denotes the degree to which signal we expect to appear only in the Q polarization of antenna j also appears in the P polarization.

For sky-based calibration, we approximate the true sky visibilities $\boldsymbol{u}$ with model visibilities $\boldsymbol{m}$ derived from a sky model propagated through a fully polarized instrument simulator. The calibration solutions are then calculated by minimizing a cost function given by

$$L(\boldsymbol{g}) = \sum_{jk} \left(\boldsymbol{v}_{jk} - \mathsf{G}_{jk}\boldsymbol{m}_{jk}\right)^{\dagger} \left(\boldsymbol{v}_{jk} - \mathsf{G}_{jk}\boldsymbol{m}_{jk}\right) . \tag{6.24}$$

Here we omit all frequency dependence. In practice, precision calibration approaches fit the gains' bandpass structure (see §4.4 for a discussion of frequency-dependent calibration).

If an instrument experiences minimal polarization cross-talk, as is the case for the MWA, we can safely set all the cross gains to zero. G_{jk} is then diagonal, and we can denote each antenna gain with a single polarization index: $g_{j\mathrm{P}}$ and $g_{j\mathrm{Q}}$. The calibration cost function then becomes

$$L(\boldsymbol{g}) = \sum_{jk} \sum_{ab} \left| v_{jk\,ab} - g_{ja} g^*_{kb} m_{jk\,ab} \right|^2 \tag{6.25}$$

where a and b each index the two instrumental polarization modes $\{\mathrm{P}, \mathrm{Q}\}$. FHD's polarized calibration currently does not support nonzero cross gains.

6.4.2 Constraining the Cross Polarization Phase

FHD initially did not use the cross polarization visibilities $\boldsymbol{v}_{\mathrm{PQ}}$ and $\boldsymbol{v}_{\mathrm{QP}}$ in calibration. Excluding the cross polarization visibilities makes calibration separable in polarization. Two independent cost functions are minimized, corresponding to the two instrumental polarizations:

$$L(\boldsymbol{g}_{\mathrm{P}}) = \sum_{jk} \left| v_{jk\,\mathrm{PP}} - g_{j\mathrm{P}} g^*_{k\mathrm{P}} m_{jk\,\mathrm{PP}} \right|^2 \tag{6.26}$$

and

$$L(\boldsymbol{g}_{\mathrm{Q}}) = \sum_{jk} \left| v_{jk\,\mathrm{QQ}} - g_{j\mathrm{Q}} g^*_{k\mathrm{Q}} m_{jk\,\mathrm{QQ}} \right|^2 . \tag{6.27}$$

However, calibrating in this way introduces a new calibration degeneracy, corresponding to the overall phase difference between the P and Q gains. We can see this degeneracy by noting that the transformation $\boldsymbol{g}_{\mathrm{P}} \to \boldsymbol{g}_{\mathrm{P}}\, e^{-i\Delta/2}$ and $\boldsymbol{g}_{\mathrm{Q}} \to \boldsymbol{g}_{\mathrm{Q}}\, e^{i\Delta/2}$ does not affect the calibration solutions. We call the parameter Δ the "cross polarization phase."

The cross polarization phase is not degenerate when we calibrate with the cross polarization visibilities, as in Equation 6.24 or 6.25 — provided these cross visibilities are nonzero. Sault et al. (1996) and Bernardi et al. (2013) assert that Δ can only be constrained by calibrating to a polarized source. However, in the widefield limit the nonorthogonality of the instrumental polarization basis couples appreciable unpolarized source power into the cross polarization visibilities. This allows for constraint of the cross polarization phase even while using a fully unpolarized sky model.

When adapting FHD to perform fully polarized calibration, we chose to largely retain the original calibration pipeline and to supplement it with an additional step to constrain the cross polarization phase. As a result, all calibration parameters other than Δ are fit from the single polarization visibilities $\boldsymbol{v}_{\mathrm{PP}}$ and $\boldsymbol{v}_{\mathrm{QQ}}$. Since the original calibration pipeline constrains the relative phase of the gains across frequency (Beardsley et al., 2016; Barry et al., 2019a,b; Li et al., 2019), the degenerate cross polarization phase amounts to a single parameter across all frequencies. We then fit the cross polarization phase by plugging the solutions into the fully polarized cost function given by Equation 6.25. We let $\boldsymbol{g}_{\mathrm{P}} = \hat{\boldsymbol{g}}_{\mathrm{P}}\, e^{-i\Delta/2}$

and $\boldsymbol{g}_{\mathrm{Q}} = \hat{\boldsymbol{g}}_{\mathrm{Q}}\, e^{i\Delta/2}$, where $\hat{\boldsymbol{g}}$ are the gains calibrated up the the cross polarization phase. This gives

$$L(\Delta) = \sum_{jk} \left[\left| v_{jk\,\mathrm{PQ}} - e^{-i\Delta} \hat{g}_{j\mathrm{P}}\, \hat{g}^*_{k\mathrm{Q}}\, m_{jk\,\mathrm{PQ}} \right|^2 + \left| v_{jk\,\mathrm{QP}} - e^{i\Delta} \hat{g}_{j\mathrm{Q}}\, \hat{g}^*_{k\mathrm{P}}\, m_{jk\,\mathrm{QP}} \right|^2 \right]. \tag{6.28}$$

We can calculate the cross polarization phase analytically by finding the value $\hat{\Delta}$ that minimizes $L(\Delta)$. We get that

$$\tan \hat{\Delta} = \mathrm{Arg} \left[\sum_{jk} \left(v^*_{jk\,\mathrm{PQ}}\, \hat{g}_{j\mathrm{P}}\, \hat{g}^*_{k\mathrm{Q}}\, m_{jk\,\mathrm{PQ}} + v_{jk\,\mathrm{QP}}\, \hat{g}^*_{j\mathrm{Q}}\, \hat{g}_{k\mathrm{P}}\, m^*_{jk\,\mathrm{QP}} \right) \right]. \tag{6.29}$$

Chapter 7

MAPPING POLARIZED DIFFUSE EMISSION

The low-frequency radio sky is bright at large scales, yet our understanding of this diffuse emission is limited. Mapping this emission can illuminate the structure of the Milky Way's interstellar medium (ISM), and measurement of its polarization modes probes galactic magnetic fields (McKee and Ostriker, 2007; Jelić et al., 2010; Bernardi et al., 2013; Lenc et al., 2016). Furthermore, diffuse mapping can aid in the precision calibration of radio cosmology instruments, enabling deeper limits on the highly redshifted 21 cm signal.

Motivated by the extreme calibration precision required for low-frequency radio cosmology (Barry et al., 2016; Trott and Wayth, 2016; Patil et al., 2016; Byrne et al., 2019), we present a high-fidelity map of broadband diffuse emission at 182 MHz across much of the Southern Hemisphere sky. This chapter presents preliminary results from Byrne et. al, in review. The map covers 11,000 sq. deg., including approximate Right Ascensions (RAs) of -3.5 to 9.5 h and Declinations (Decs) of -62° to 10°. It represents angular scales of approximately 1° to 9° in all four Stokes parameters. Compact sources have been deconvolved from the data, meaning that the map can be combined with existing compact source catalogs to produce an expanded and improved sky model for calibration.

The map is produced from data from the Murchison Widefield Array (MWA), a radio interferometer in the Western Australian outback, in its Phase I configuration (Tingay et al., 2013). This instrument is particularly well-suited to diffuse mapping because of it high *uv* coverage at short baselines and good resolution from long baselines. The resolution allows for localizing and removing compact sources, and the density of short baselines allows for high sensitivity measurements of diffuse emission. Furthermore, across a large frequency continuum the short baselines sample nearly every *uv* mode shorter than 50 wavelengths. This enables us to produce maps of diffuse emission in physical units of surface brightness without deconvolving. The maps can be transferred to other instruments for visibility simulation.

We quantify the map's accuracy by measuring the discrepancy between visibilities simulated from the map and true data from an MWA observation. We find that the diffuse map reconstructs a significant fraction of visibility power on short baselines. The map improves short baseline modeling to accuracy levels nearly equivalent to those of long baseline modeling with state of the art compact source models.

Existing low-frequency diffuse maps include Haslam et al. 1981, 1982; Bennett et al. 2003; de Oliveira-Costa et al. 2008; Remazeilles et al. 2015; Zheng et al. 2017; Eastwood et al. 2018; and Wolleben et al. 2019. These maps have minimal, if any, compact source subtraction, meaning that they cannot be combined with deep catalogs of compact sources. All but Wolleben et al. 2019 are unpolarized, omitting the bright linearly polarized emission seen on large angular scales. As a result, these maps do not have the precision required to achieve highly sensitive cosmological measurements. In contrast, we present a map with deep compact source removal and accurate polarization reconstruction in the widefield limit.

We quantify the map's accuracy by measuring the discrepancy between visibilities simulated from the map and true data from an MWA observation. We find that the diffuse map reconstructs a significant fraction of visibility power on short baselines. The map improves short baseline modeling to accuracy levels nearly equivalent to those of long baseline modeling with state of the art compact source sky models.

We expect that the map presented here will improve high-precision calibration for 21 cm cosmology experiments with not only the MWA itself but also other Southern Hemisphere interferometers such as the Hydrogen Epoch of Reionization Array (HERA) in South Africa (DeBoer et al., 2017) and future instruments such as the Square Kilometre Array (SKA). HERA is highly compact, dominated by short baseline measurements, and is therefore particularly sensitive to diffuse structure in calibration. Improved diffuse modeling could also enable new short-baseline calibration approaches that mitigate systematic errors (Ewall-Wice et al., 2016).

In this chapter §7.1 describes the data processing procedure and §7.2 presents the resulting polarized diffuse map. §7.3 defines the angular scales over which we have confidence in the results, §7.4 quantifies the effect of the map on improving visibility simulation for an MWA observation, and §7.5 presents a discussion of how to use this diffuse map for visibility simulation.

7.1 Methods

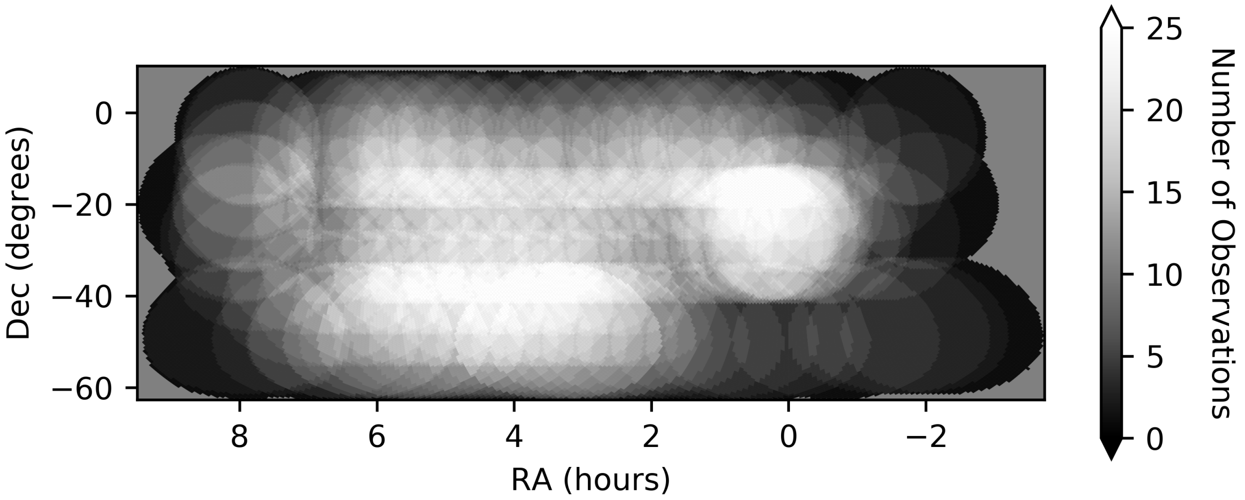
In this section we describe the analysis process and techniques used to achieve a high-fidelity polarized map of diffuse emission. The bulk of the analysis is performed with the FHD software package (Sullivan et al., 2012; Barry et al., 2019a). We analyze all data in the Amazon Web Services (AWS) cloud using the cloud computing pipeline described in Byrne and Jacobs 2021.

7.1.1 The Data

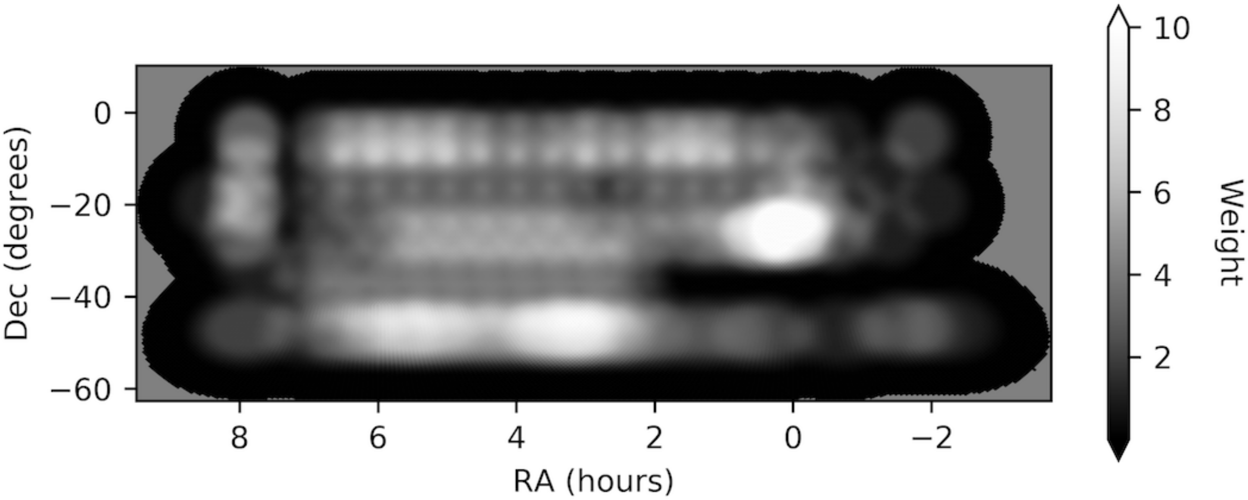
We present a map comprising 5.8 hours of data from the MWA Phase I, taken November 5-11, 2015. The data consists of 173 individual observations that are approximately 2 minutes (112 seconds) in length. This time interval is sufficiently short to enable each observation to be processed as a fringe stopped snapshot image.

The observational fields are chosen to overlap across the entire survey field, plotted in Figure 7.1(a). Each field is observed with a variety of elevation angles to maximize uv rotation and reduce sidelobe confusion.

The observations cover frequencies of 167-198 MHz and are processed as continuum images averaged across the full 31 MHz frequency range. This further fills the uv plane, reducing sidelobe confusion and providing an accurate measurement of diffuse power on all uv modes. However, it also means that polarized emission with substantial frequency evolution will depolarize through frequency averaging. We therefore expect to only measure polarized emission with low rotation measure (RM) magnitudes.



(a)



(b)

Figure 7.1: Plots of the observational coverage of the map. Observations are chosen to overlap each field on the sky. (a) depicts the number of observations that contribute to each field. The observations are averaged with a weighting function given by Equation 7.1. (b) gives the total weight in each field on the sky, given this weighting scheme. We use an especially large sample of observations of the "EoR-0" field centered at RA 0 h, Dec. −27° because it is of particular interest to the MWA's EoR team and displays strong linearly polarized signal. Just south of that field, at a declination of about −50°, the map has a dearth of observational coverage owing to scheduling limitations during the data acquisition.

Data pre-processing is performed with the COTTER package[1] (Offringa et al., 2015). This reduces the data volume by frequency-averaging to 80 kHz and time-averaging to 2 seconds. We perform Radio-Frequency Interference (RFI) excision in two steps. Initial RFI flagging is performed with AOFLAGGER[2], which is contained within the COTTER pre-processing package. We subsequently perform additional RFI flagging with SSINS[3] (Wilensky et al., 2019).

7.1.2 Calibration

We calibrate with FHD's sky-based calibration pipeline, leveraging the precision calibration techniques developed for EoR science (Barry et al., 2016; Beardsley et al., 2016; Barry et al., 2019b,a; Li et al., 2019).

The calibration sky model is based on the GLEAM catalog (Hurley-Walker et al., 2017). As GLEAM omits some of the brightest sources in the field, we supplement the catalog with additional source models. We use point source models for 3C161, 3C409, and Cassiopeia A and multi-component extended source models for Centaurus A, Hera A, Hydra A, Pictor A, and Virgo A (White, private communication). We additionally include an extended source model for Fornax A produced with FHD (Carroll, private communication). All extended source models use point-like sub-components to represent the sources' extended structure.

We use FHD to simulate visibilities for calibration from this supplemented GLEAM catalog. We calibrate with all baselines longer than 50 wavelengths. Shorter baselines are poorly modeled by the compact source catalog, and we omit them from calibration to avoid bias in our map estimation.

FHD's calibration calculates a complex gain per antenna polarization. We denote the two antenna polarization modes p and q, such that calibration calculates gains $\boldsymbol{g}_p$ and $\boldsymbol{g}_q$ for each antenna, frequency interval, and time step. The MWA has well-isolated signal paths and experiences minimal cross-talk between the instrumental polarizations. We therefore do not

---

[1] `https://github.com/MWATelescope/cotter`

[2] `https://sourceforge.net/projects/aoflagger/`

[3] `https://github.com/mwilensky768/SSINS`

fit calibration parameters that mix p and q (Hamaker et al., 1996; Sault et al., 1996).

Previous unpolarized analyses with FHD calibrate using only the single-polarization visibilities, i.e. $\boldsymbol{v}_{pp}$ and $\boldsymbol{v}_{qq}$ (Barry et al., 2019a). This does not enable fully polarized imaging because it leaves the phase relationship between $\boldsymbol{g}_p$ and $\boldsymbol{g}_q$ unconstrained. We therefore include an additional calibration step in which we constrain this degeneracy with the cross-polarization visibilities $\boldsymbol{v}_{pq}$ and $\boldsymbol{v}_{qp}$. We fit all other calibration degrees-of-freedom with only the single-polarization visibilities. For further discussion of polarized calibration, see Sault et al. 1996; Lenc et al. 2017; Gehlot et al. 2018; Dillon et al. 2018; Byrne et al. 2021.

Calibration constrains the observations' overall flux scale. The resulting calibrated images are therefore normalized to the flux scale of the GLEAM catalog, which is in turn referenced to the Very Large Array Low-frequency Sky Survey Redux (VLSSr; Lane et al. 2014), the Molonglo Reference Catalogue (MRC; Large et al. 1981), the National Radio Astronomy Observatory Very Large Array Sky Survey (NVSS; Condon et al. 1998), and Baars et al. 1977 (Hurley-Walker et al., 2017).

7.1.3 Imaging and Compact Source Removal

FHD performs high-fidelity widefield imaging based on the optimal mapmaking formalism (Bhatnagar et al., 2008; Morales and Matejek, 2009). It uses the instrumental response kernel to grid visibilities to a *uv* plane, which we choose to pixelate at half-wavelength spacing. This naturally yields an accurate horizon-to-horizon image reconstruction without requiring faceting across the instrumental field-of-view, enabling measurement of large-scale structure spanning several angular degrees.

To enable this work, we updated FHD to perform fully-polarized imaging, as described in Chapter 6. Images are reconstructed in the non-orthogonal "instrumental" polarization basis, aligned with the polarization vectors of maximal instrumental response. We then use the instrument's polarized response model, or Jones matrix (Jones, 1941), to translate the instrumental polarization modes to Stokes polarization modes defined in Equation 6.1 and in Hamaker et al. 1996; Sault et al. 1996; Hamaker and Bregman 1996; Hamaker 2000, 2006;

Ord et al. 2010.

We crop each image at a diameter of 30° to limit the images to regions of high sensitivity. Next, we perform a data quality cut based on calibration success. We omit observations with clear calibration systematics, usually apparent as high Stokes V power across the image, from further analysis. The effect shows some time correlation but is not field-dependent: observations of the same field experience variable calibration success. We expect that poor calibration solutions could result from unmitigated RFI contamination or ionospheric activity (Jordan et al., 2017).

Next, we remove compact sources with FHD's deconvolution algorithm (Sullivan et al., 2012). We expect compact sources to have negligible intrinsic polarization (Ord et al., 2010; Bernardi et al., 2013), and we therefore restrict our source models to Stokes I only in this step. Deconvolving compact sources with FHD produces better removal of extended sources and mitigation of source intensity errors from beam response mismodeling than simply modeling and subtracting the calibration catalog. We remove about 10,000 sources from each observation.

7.1.4 Correcting for Polarization Leakage from Beam Modeling Errors

Errors in the polarized instrumental response model lead to polarization mode-mixing (Ord et al., 2010; Sutinjo et al., 2015). Our observations are dominated by power from unpolarized point-like sources (Bernardi et al., 2013), and we see evidence of beam modeling errors in the form of significant point source power imaged in Stokes Q and U. As FHD's deconvolution algorithm removes sources in Stokes I only, this power persists through source removal and contaminates the Stokes Q and U data products. We do not see high levels of Stokes I to V leakage.

To mitigate residual point source power in Stokes Q and U, we follow techniques similar to those presented in Lenc et al. 2017 to fit and subtract polarization leakage. We apply the correction to compact sources only and do not adjust the measured diffuse emission based on this fit.

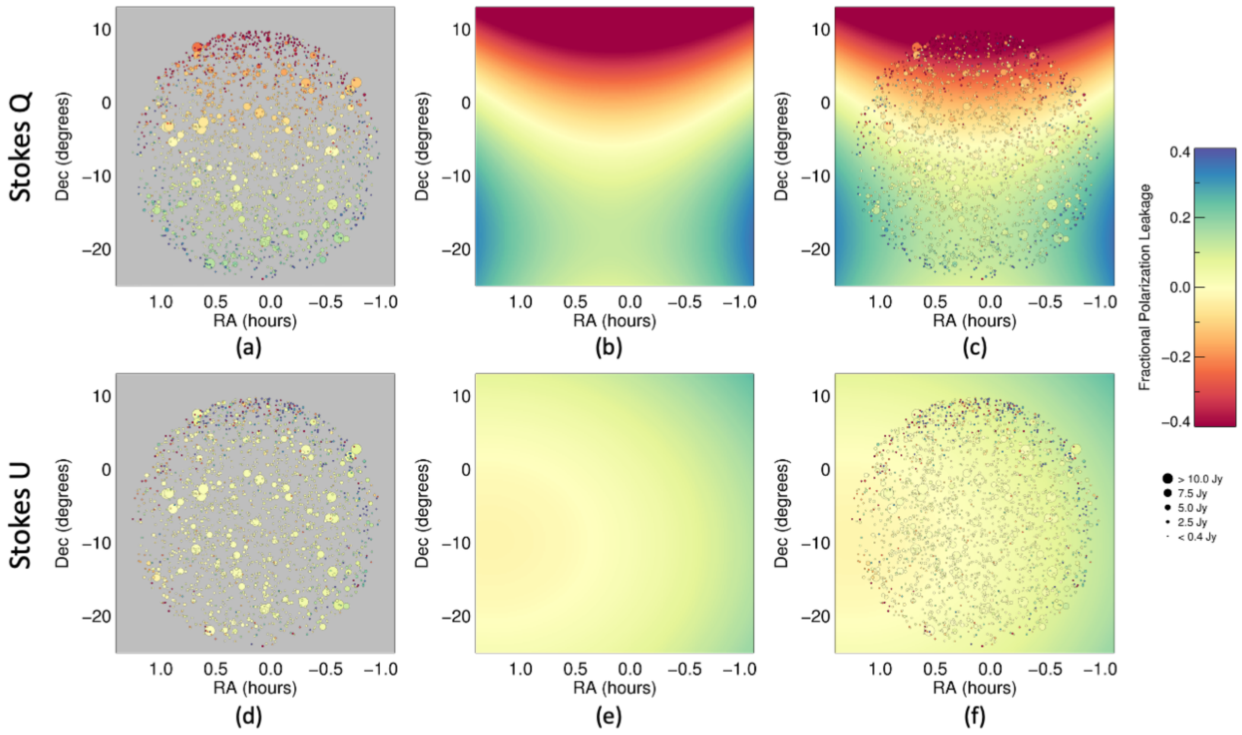


Figure 7.2: Errors in the polarized instrumental response model lead to significant observed compact source power in Stokes Q (top row) and U (bottom row). In the left column we plot the observed Stokes Q and U polarization fraction of the 2,000 brightest removed sources for a single observation. The point size corresponds to the source intensity, and its color corresponds to the apparent polarization fraction. In the center column we display a second-order polynomial surface fit that we use to mitigate residual point source power in Stokes Q and U. In the right column we combine the sources and surface fit.

For every observation, we calculate the Stokes Q and U polarization components at each of the locations of the 2,000 brightest sources removed. We then convert that measurement into an apparent Stokes Q and U polarization fraction. We use these values to fit a second-order polynomial surface across the observation field-of-view; the fit is weighted by source intensity, preferentially fitting to the brightest observed sources. Figure 7.2 presents the result of this fit for an example observation.

We use the fit polynomial surface for each observation to modify the catalogs of sources removed with FHD's deconvolution. Each source is modified to reflect its apparent polarization mode according to the surface fit, up to 40% polarization fraction each in Stokes Q and U. We then model visiblities from those catalogs and subtract them from the data. This technique effectively removes residual point source power from Stokes Q and U, yielding cleaner polarized maps that better highlight diffuse emission.

At this point we could choose to use the same fit polynomial surface to modify the diffuse maps themselves. However, this approach would make the diffuse maps highly sensitive to small errors in the polarization leakage fit. The second-order polynomial surfaces represent structure on angular scales equivalent to those of the diffuse emission, and errors could significantly bias our diffuse measurement. Furthermore, as we capture Stokes I to Q and U leakage only, this modification would not capture other forms of polarization leakage. By averaging across observations taken with different instrumental pointings, we expect some mitigation of the effects of polarized beam model errors in the measured diffuse power.

7.1.5 Normalization and Combining Observations

After removing unpolarized point sources and mitigating point source polarization leakage in Stokes Q and U, we re-image each observation with short baselines only, omitting baselines longer than 50 wavelengths. This re-imaging accomplishes two things. First, it highlights diffuse emission by mapping large-scale structure only. Secondly, it limits the imaged data to regions of the *uv* plane where the MWA samples virtually every *uv* point (see Figure 7.3).

Imaging in this regime of near-complete *uv* sampling is known as deconvolution-free imag-

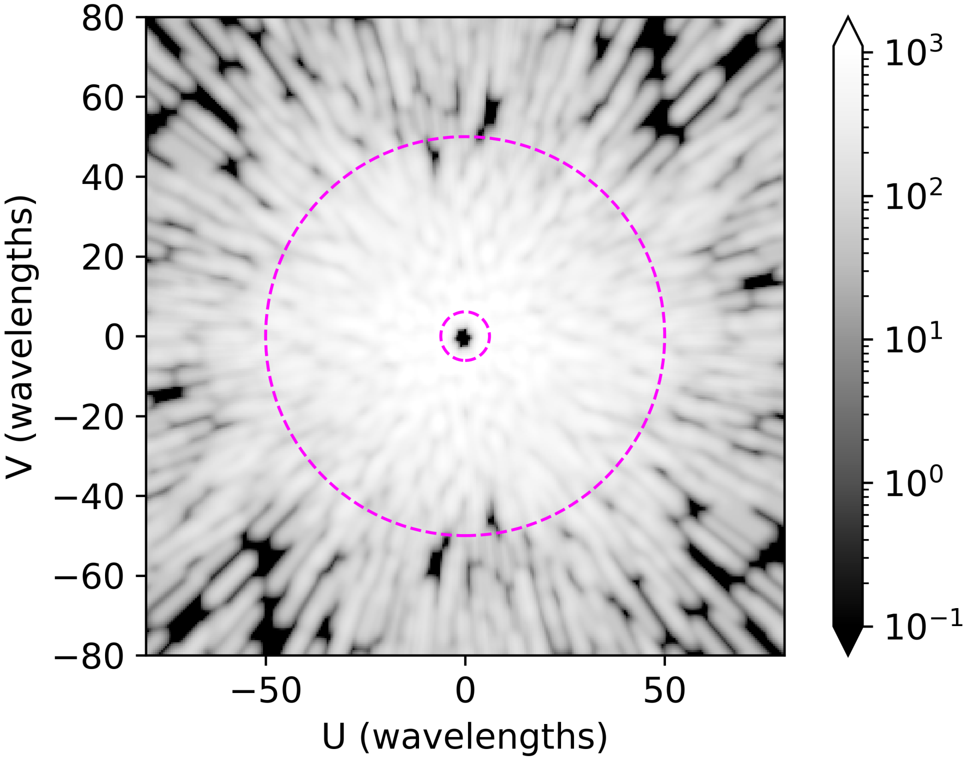


Figure 7.3: Plot of the *uv* measurement coverage for a zenith observation from the MWA Phase I for the 167-198 MHz frequency range. The large dashed magenta contour has a radius of 50 wavelengths, represented the longest baselines used in the diffuse map presented in this paper. Across this 31 MHz frequency continuum, the MWA samples nearly every *uv* point below 50 wavelengths. The small dashed magenta contour has a radius of 6.1 wavelengths, representing the extent of our confidence in the diffuse map at large angular scales. In the optimal mapmaking formalism, this measurement coverage is known as the "*uv* weights" (see Barry et al. 2019a for further discussion).

ing because it enables faithful reconstruction of the sky signal without deconvolving. This is a necessary condition for accurately imaging diffuse emission. Deconvolution inherently requires that the signal in sparse, with fewer components than independent measurements. For example, the deconvolution algorithm we use to remove compact sources, described above in §7.1.3, assumes that the signal is localized at discrete points on the sky. Diffuse emission is not sparse and produces an independent measurement at all *uv* modes. A high-fidelity measurement of diffuse emission therefore requires a measurement of nearly all *uv* modes.

For the MWA, achieving near-complete *uv* sampling below 50 wavelengths requires imaging across a frequency continuum. The *uv* plane is not well-sampled at any single frequency within our 31 MHz band. However, combining measurements across the full frequency range fills the *uv* plane and offers improved *uv* measurement coverage.

By reconstructing diffuse emission in the deconvolution-free regime, we produce images with well-defined normalization on angular scales of 1° to 9° (see §7.3). This normalization is independent of the instrumental beam response (Byrne, thesis in prep.), and the resulting maps can thereby be transferred to other instruments for visibility modeling. The images are produced with FHD's "optimal" weighting scheme, discussed in §2.2.2. We present the resulting diffuse map in surface brightness units of Janskys per steradian (Jy/sr).

Next, we average all observations to produce a single polarized map across the observed fields. The observations are averaged with a position-dependent tapered cosine (or Tukey) weighting $W_n(\boldsymbol{\theta})$ given by

$$W_n(\boldsymbol{\theta}) = \begin{cases} 1, & \text{for } |\boldsymbol{\theta} - \boldsymbol{\theta}_n| < 4^\circ \\ \cos\left(\frac{\pi}{12^\circ}|\boldsymbol{\theta} - \boldsymbol{\theta}_n| - \frac{\pi}{3}\right), & \text{for } 4^\circ \leq |\boldsymbol{\theta} - \boldsymbol{\theta}_n| \leq 10^\circ \\ 0, & \text{for } |\boldsymbol{\theta} - \boldsymbol{\theta}_n| > 10^\circ \end{cases} . \tag{7.1}$$

Here n indexes the observation and $\boldsymbol{\theta}$ represents the position on the sky. $\boldsymbol{\theta}_n$ denotes the position of the beam center of observation n. This weighting reflects our greater confidence near the beam center of each observation. Figure 7.1(b) depicts the total weight of each field across the map.

7.1.6 Correcting for Ionospheric Faraday Rotation

Propagation through intervening magnetized plasma induces Faraday rotation in linearly polarized emission. In particular, the ionosphere induces time- and direction-dependent Faraday rotation. This rotation mixes the Stokes Q and U polarization modes of measured radiation and causes depolarization across the measured frequency continuum and observations in the combined map (Ord et al., 2010). To mitigate these effects, we measure and correct for ionospheric Faraday rotation.

We assume that ionospheric Total Electron Content (TEC) is constant across an observation's field-of-view. Gradients in the TEC induce positional offsets in compact sources (Loi et al., 2015; Jordan et al., 2017), but we can safely ignore these effects when imaging diffuse structure (Lenc et al., 2016). We therefore describe Faraday rotation from the ionosphere with a single RM per observation, plotted in Figure 7.4. The electric field rotation angle β relates the the RM R according to the relationship

$$\beta = R\lambda^2, \tag{7.2}$$

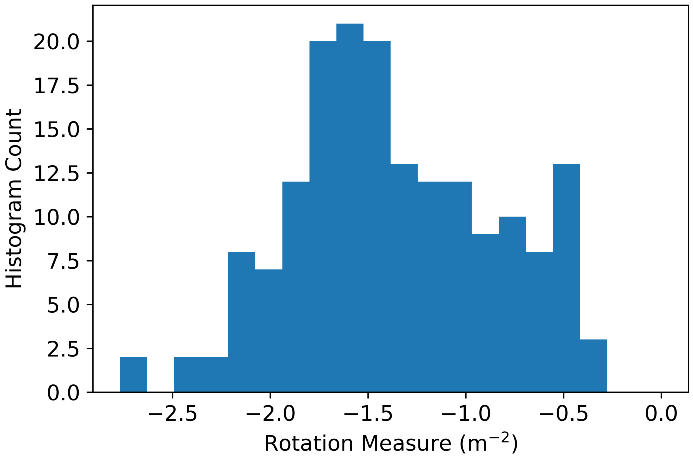
where λ denotes the wavelength.

We estimate the ionospheric RM values with RMEXTRACT[4], which calculates the ionospheric TEC along each observation's line-of-sight from Global Position System (GPS) and World Magnetic Model (WMM) data (Mevius, 2018). We then correct the continuum Stokes Q and U images for each observation according to the calculation in Appendix E. This correction consists of unwrapping the effective rotation angle given by Equation E.6 and boosting the image power by a factor given by Equation E.7, which accounts for depolarization across the full frequency continuum.

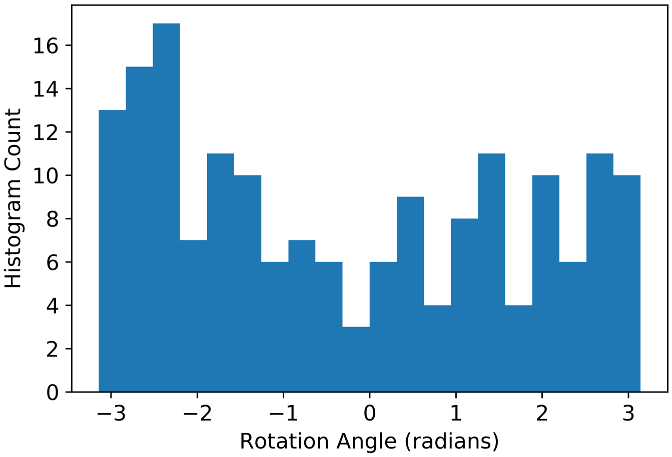
We further correct for Faraday rotation near the EoR-0 field centered at RA 0 h, Dec. −27° using an ionospheric self-calibration approach based on that described in Lenc et al. 2017. For each observation n, we calculate an empirical correction $\Delta\beta_n$ to the effective

---

[4] `https://github.com/lofar-astron/RMextract`



(a)



(b)

Figure 7.4: Histograms of the applied ionospheric Faraday rotation metrics. The RM is calculated with the RMEXTRACT tool. Observations of the EoR-0 field include an additional empirical correction, plotted in Figure 7.5. A single RM is calculated for each of the 173 2-minute observations, plotted in (a). Each RM corresponds to a rotation angle between the Stokes Q and U parameters, as defined by Equation 7.2 for a single frequency and Equation E.6 for a frequency continuum. (b) histograms the effective rotation angles applied to the data. We find significant Faraday rotation, and all rotation angles are represented in the data.

rotation angle between Stokes Q and U that maximizes the agreement of that observation to the average map. We minimize the cost function between the average Stokes Q and U maps, $Q_{\mathrm{avg}}(\boldsymbol{\theta})$ and $U_{\mathrm{avg}}(\boldsymbol{\theta})$, and the individual observation's Stokes Q and U maps, $Q_n(\boldsymbol{\theta})$ and $U_n(\boldsymbol{\theta})$:

$$\begin{aligned}\chi^2(\Delta\beta_n) = \sum_{\boldsymbol{\theta}} \Big(&\left[Q_{\mathrm{avg}}(\boldsymbol{\theta}) - Q_n(\boldsymbol{\theta})\cos(2\Delta\beta_n) - U_n(\boldsymbol{\theta})\sin(2\Delta\beta_n)\right]^2 \\ &+ \left[U_{\mathrm{avg}}(\boldsymbol{\theta}) + Q_n(\boldsymbol{\theta})\sin(2\Delta\beta_n) - U_n(\boldsymbol{\theta})\cos(2\Delta\beta_n)\right]^2 \Big).\end{aligned} \tag{7.3}$$

It follows that

$$\tan(2\Delta\beta_n) = \frac{\sum_{\boldsymbol{\theta}} \left[Q_{\mathrm{avg}}(\boldsymbol{\theta})U_n(\boldsymbol{\theta}) - U_{\mathrm{avg}}(\boldsymbol{\theta})Q_n(\boldsymbol{\theta})\right]}{\sum_{\boldsymbol{\theta}} \left[Q_{\mathrm{avg}}(\boldsymbol{\theta})Q_n(\boldsymbol{\theta}) + U_{\mathrm{avg}}(\boldsymbol{\theta})U_n(\boldsymbol{\theta})\right]}. \tag{7.4}$$

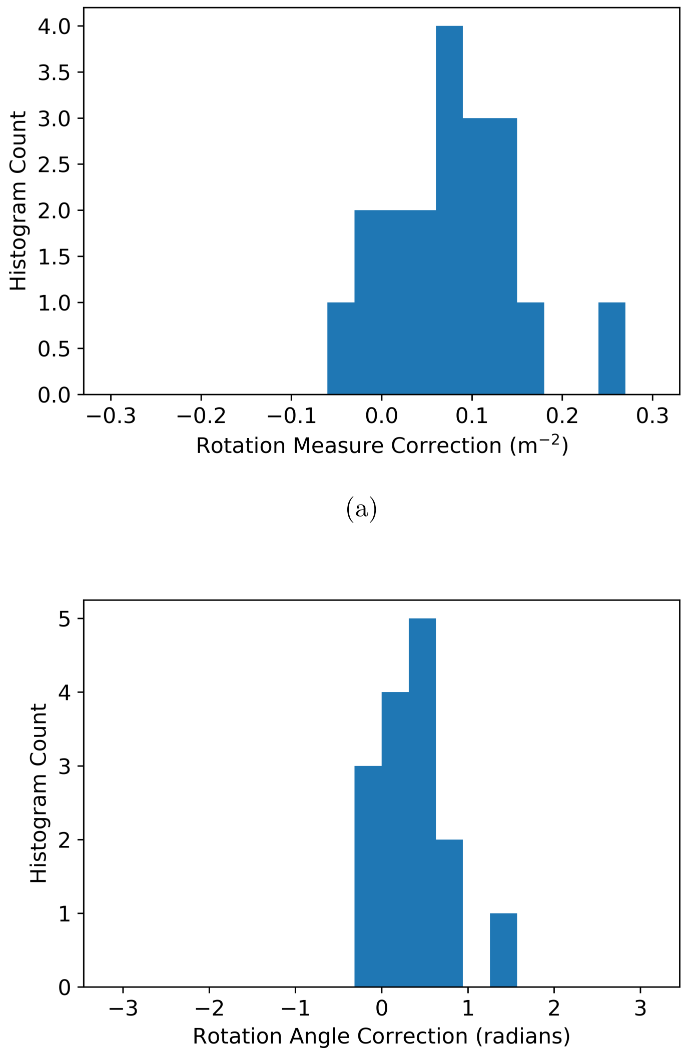
After calculating this correction to the effective rotation angle we numerically invert Equation E.6 to calculate the associated RM. This conversion is not one-to-one as many RM values correspond to the same effective rotation angle. We choose the closest solution to the initial RM estimate before self-calibration. The results are plotted in Figure 7.5.

This self-calibration approach works well in fields with a strong linearly polarized signal. We therefore apply this correction only to 19 observations in the EoR-0 field where we measure particularly strong polarized emission. We do not apply ionospheric self-calibration to the map outside of EoR-0.

After correcting for ionospheric Faraday rotation, we repeat the averaging procedure described in §7.1.5 to combine the corrected observations. The results are presented below in §7.2 and plotted in Figures 7.6-7.10.

7.2 Results

Figures 7.6-7.9 present the polarized mapped diffuse emission. Figure 7.6 presents the unpolarized (Stokes I) emission, Figure 7.7 presents the Stokes Q linearly polarized emission, Figure 7.8 presents the Stokes U linearly polarized emission, and Figure 7.9 presents the Stokes V circularly polarized emission. The Stokes parameters are defined by Equation 6.1. The map is presented as a pixelated image with the HEALPix equal-area pixelation scheme (Gorski et al., 2005).



(b)

Figure 7.5: We estimate an empirical self-calibration correction to the RMs of 19 observations near the EoR-0 field. (a) histograms the change in the RM of those observations from this correction. The correction is small compared to the typical RM magnitudes calculated from RMExtract (note that the horizontal axis in (a) has a significantly smaller extent than that of Figure 7.4(a)). (b) presents the associated change in the rotation angles from the RM correction.

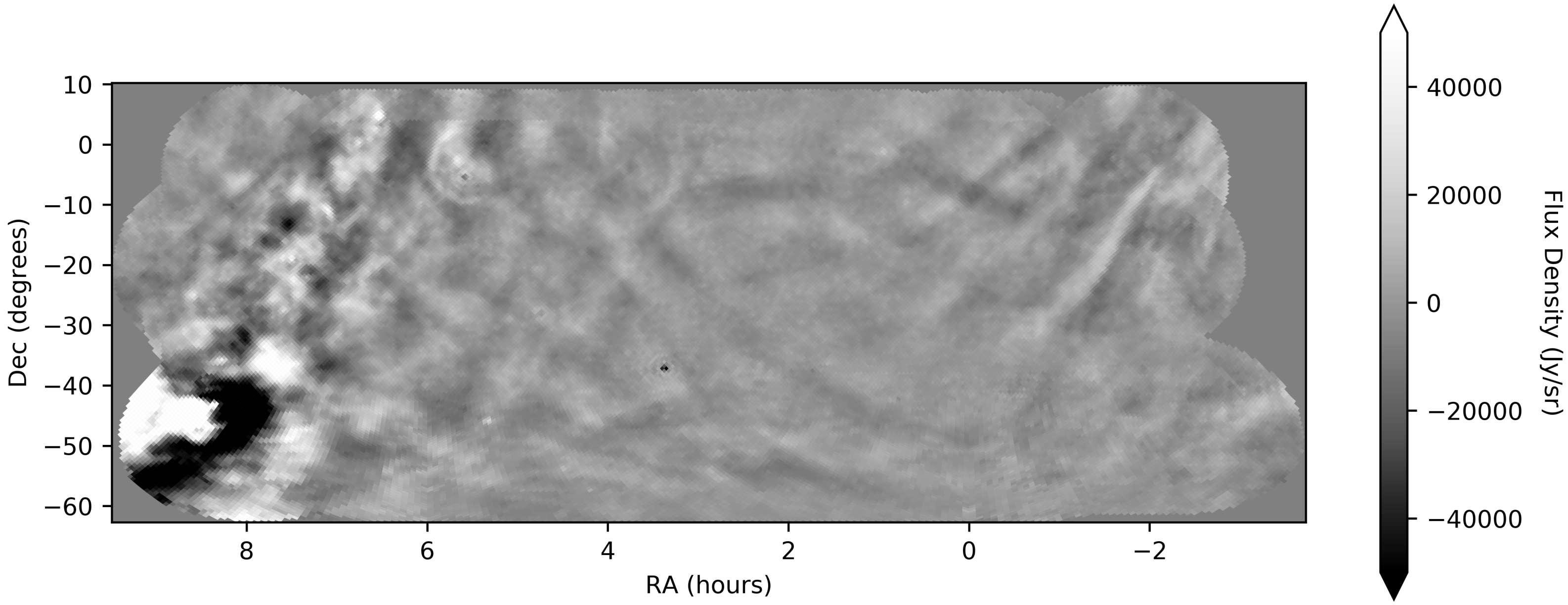


Figure 7.6: Unpolarized (Stokes I) diffuse emission.

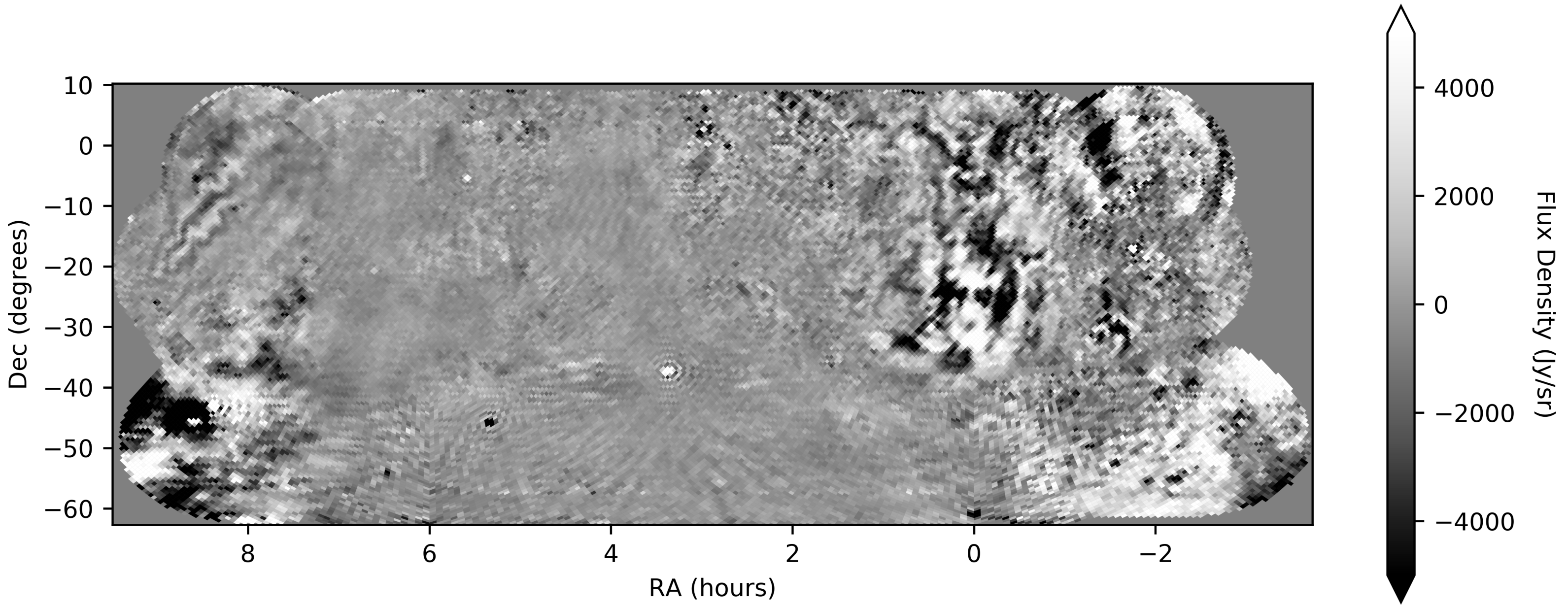


Figure 7.7: Stokes Q linearly polarized diffuse emission.

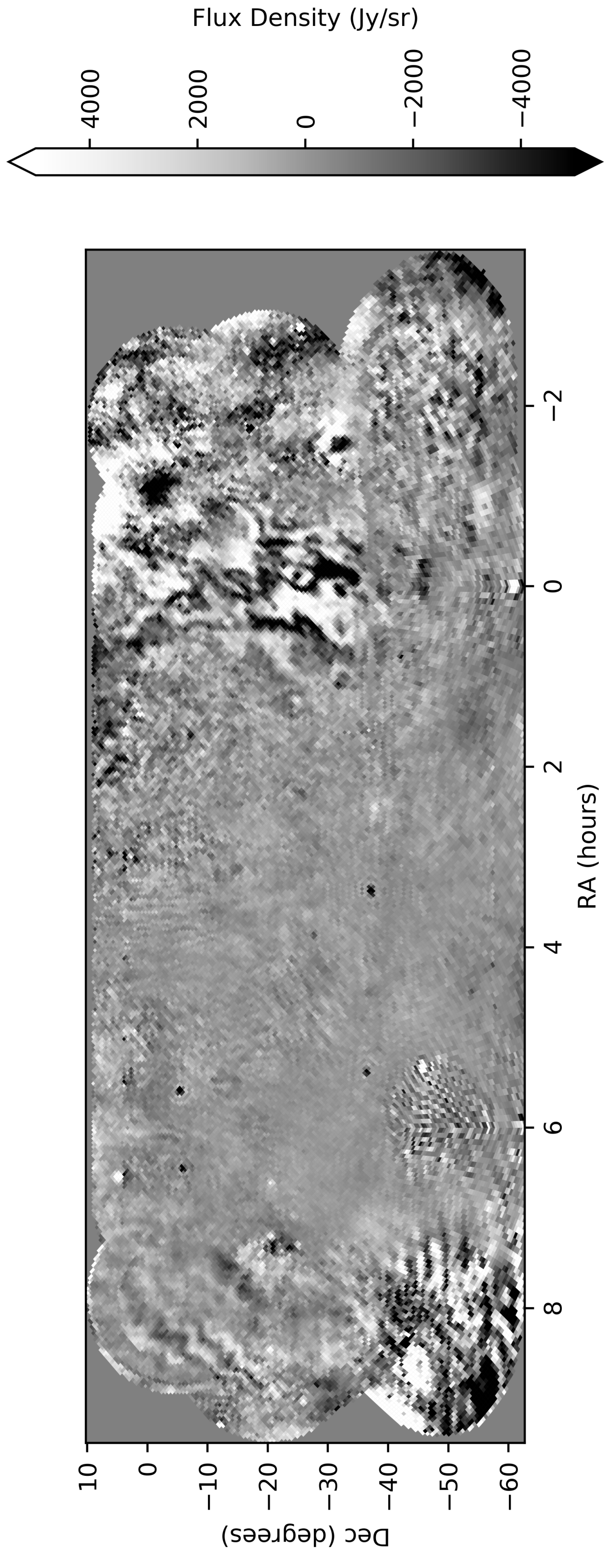


Figure 7.8: Stokes U linearly polarized diffuse emission.

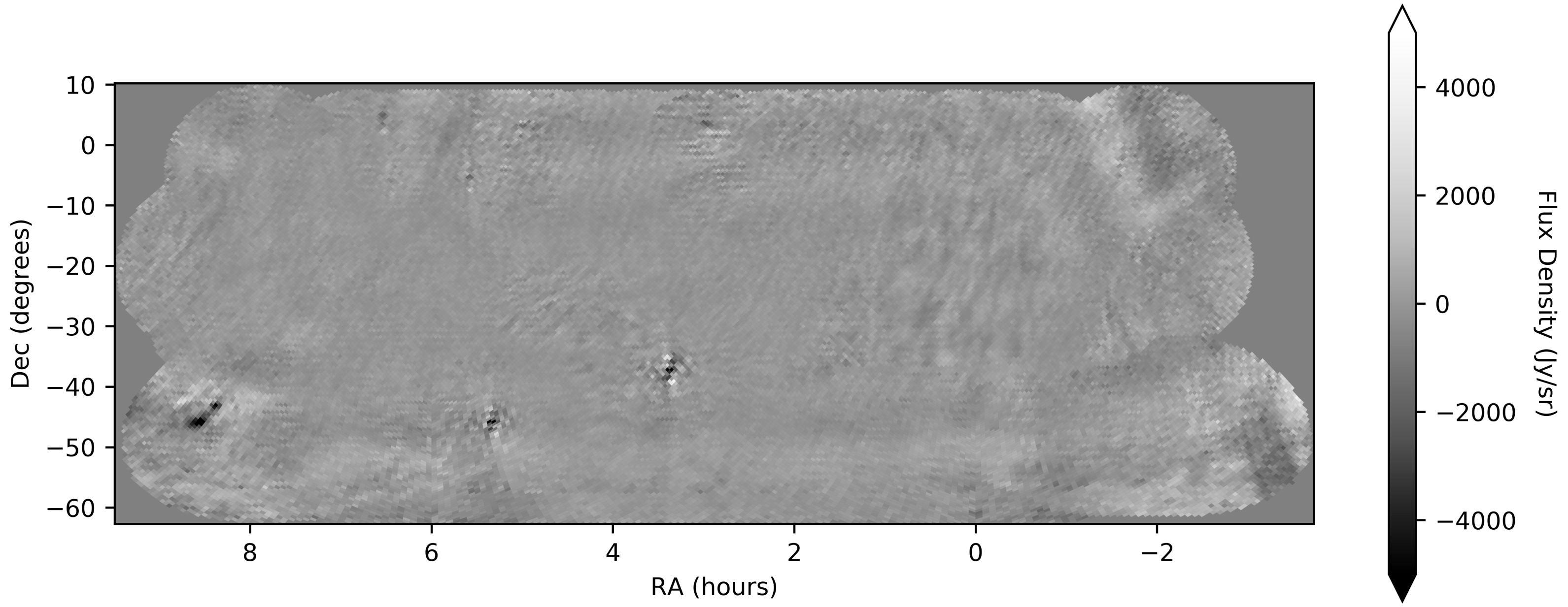


Figure 7.9: Stokes V circularly polarized diffuse emission.

Figure 7.10 replicates the map with further annotations. The map covers much of the fields of interest on the sky for EoR science with the MWA and HERA. The solid cyan contours in Figure 7.10 denote the "EoR-0" and "EoR-1" fields, regions that have been the focus of the MWA's EoR analyses (Beardsley et al., 2016; Barry et al., 2019b; Li et al., 2019; Trott et al., 2020). The contours are defined as the full width at half maximum (FWHM) of the MWA zenith-pointed beam, averaged across frequencies and polarizations. The dashed cyan lines denote the North and South extents of the FWHM of HERA's beam. Figure 7.10 also marks two notable bright soures: Fornax A, centered at RA 3.378 h, Dec. $-37.21°$; and Pictor A, centered at RA 5.330 h, Dec. $-45.78°$. The red plus sign marks the South Galactic Pole at RA 0.857 h, Dec. $-27.13°$

As the map presented here is an interferometric image, we do not measure the absolute amplitude of emission. The map is mean-zero in all polarization modes. This presents a number of limitations for our analysis. Although we can confidently measure Stokes polarization across the interferometric modes given by the array design, we cannot calculate the emission's polarization fraction in the image domain. Likewise, we cannot determine the emission's polarization angle.

Even in light of these limitations, we can conclude that the diffuse structure we measure is predominantly unpolarized. Note that the Stokes I image in Figure 7.6 displays an order of magnitude greater range than the other polarization modes in Figures 7.7-7.9. We nonetheless measure high levels of linearly polarized diffuse emission. Near RA 0 h, we see very strong polarized signal, consistent with other results in the field (Bernardi et al., 2013; Lenc et al., 2016). This structure is particularly concentrated in the MWA's EoR-0 field, suggesting that it could be a significant contaminant in EoR analyses such as Beardsley et al. 2016, Barry et al. 2019b, Li et al. 2019, and Trott et al. 2020.

We expect the diffuse circularly polarized signal to be very weak and below the sensitivity threshold of this map (Enßlin et al., 2017; Lenc et al., 2018). Indeed, Stokes V appears largely featureless in Figure 7.9. The observed features—notably at the locations of Fornax A, Pictor A, and at RA > 8 h—are imaging artifacts and not true polarized structure. These

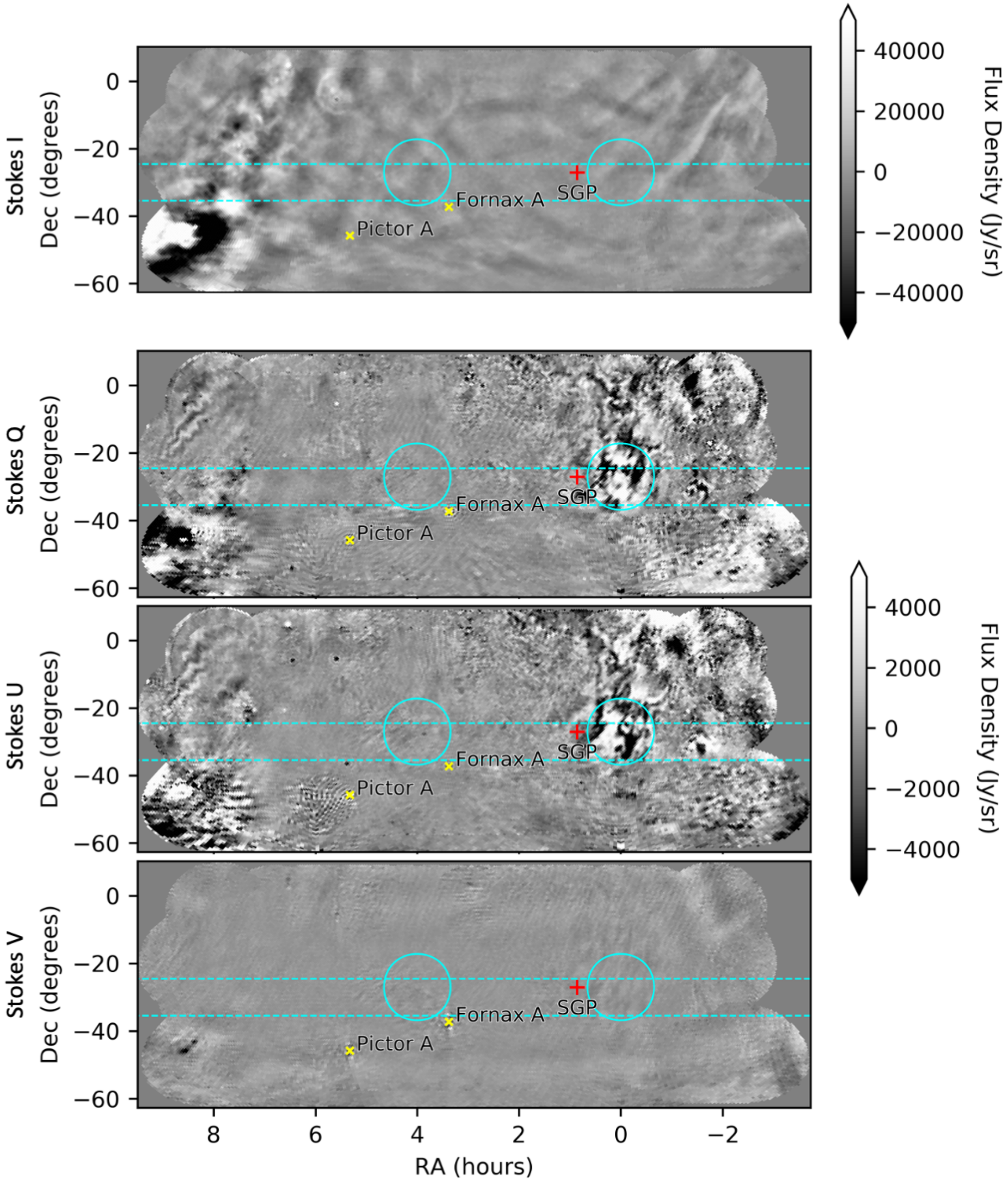


Figure 7.10: Polarized diffuse map with annotations. (Caption continued.)

Figure 7.10: Here we reproduce the diffuse map presented in Figures 7.6-7.9 with annotations. The dashed cyan lines span the FWHM of the HERA's field-of-view across its observing band, centered at a declination of −30°. The solid cyan contours represent the FWHM of the MWA beam at two primary observing fields: EoR-0 is centered at an RA of 0 hours and EoR-1 is centered at an RA of 4 hours. The EoR-0 field contains significant linearly polarized diffuse structure. We label two bright sources, Fornax A and Pictor A, with yellow "X" symbols. The red plus sign denotes the Southern Galactic Pole (SGP).

artifacts amount to poorly deconvolved sources due to limitations in calibration precision and polarized beam modeling.

We measure agreement between overlapping observations by evaluating the standard deviation of each pixel value across contributing observations, plotted in Figure 7.11. Because the map is mean-zero, we cannot directly compare the standard deviation to the map power to, for example, calculate a signal-to-noise metric. Instead, we present the standard deviation plots in Figure 7.11 alongside the maps in Figures 7.6-7.10 as a qualitative assessment of observation agreement. Features that appear in the standard deviation plots but do not correlate with diffuse structure indicate systematic error. See §7.5 for further discussion of field-dependent imaging systematics.

7.3 Defining Measured Angular Scales

In this section we discuss the scales over which we are confident this map gives accurate results. Across those scales, we expect this map to improve visibility modeling for a wide range of instruments. We show that diffuse emission is a significant component of the total sky signal at these scales.

The angular scales reconstructed by the map presented in Figures 7.6-7.10 are dictated by the MWA's baseline locations. The array's shortest baselines measure large angular scales on the sky, while longer baselines measure small angular scales.

We plot the MWA's *uv* measurement coverage in Figure 7.3. Nonzero regions of the plot correspond to points in the *uv* plane that contribute to one or more visibilities. We see that,

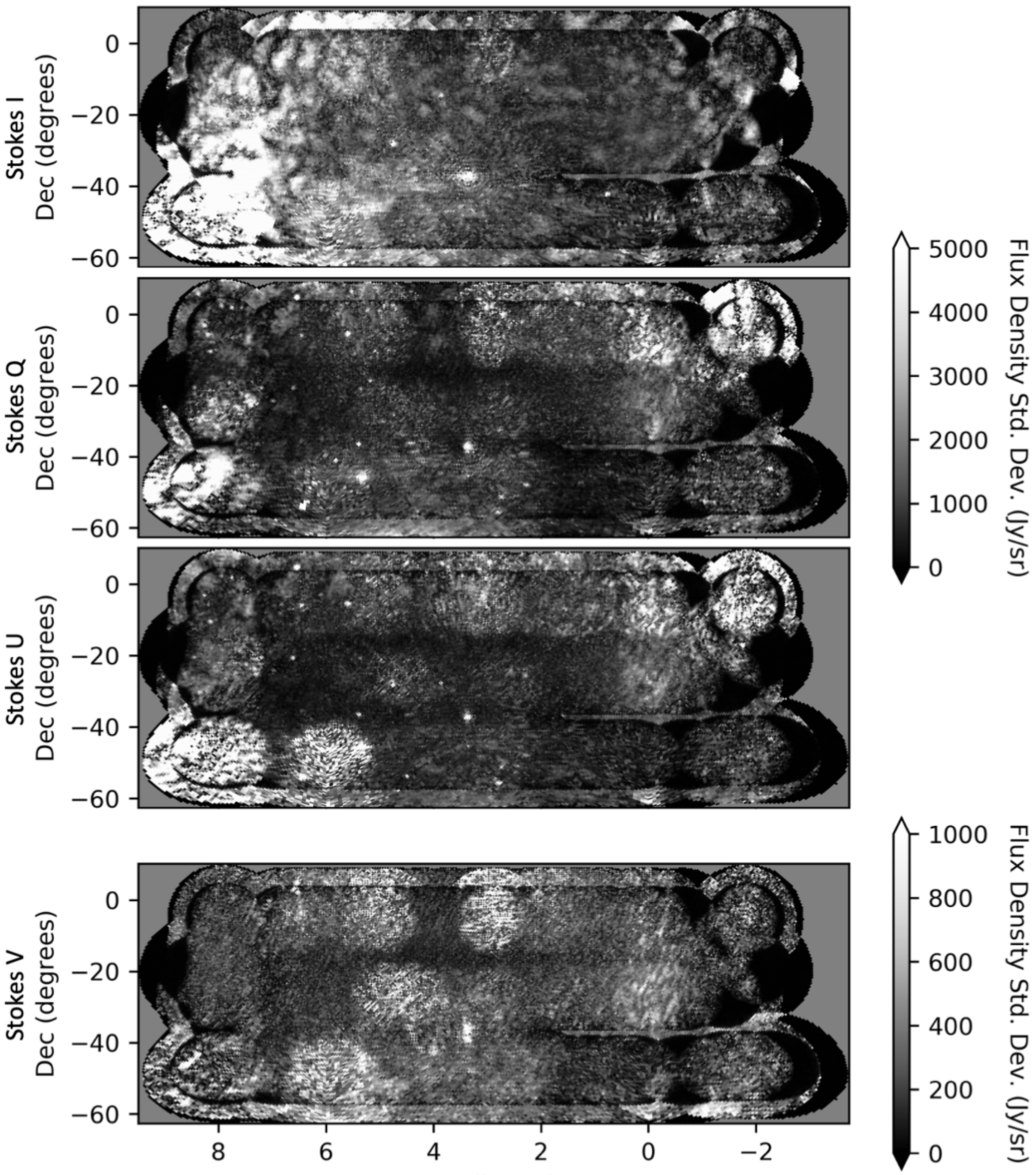


Figure 7.11: Standard deviation of mapped diffuse emission. (Caption continued.)

Figure 7.11: Plots of the standard deviation in the mapped diffuse emission across independent observations. For each Stokes parameter and map pixel we calculate the standard deviation across observations that contribute to that pixel measurement via the weighted average described in §7.1.5. High standard deviation values that do not correlate with true diffuse structure indicate systematic errors in the analysis.

below 50 wavelengths (denoted by the large dashed magenta contour in Figure 7.3), nearly every *uv* point is measured. This is a result of the MWA's many closely-spaced antennas and pseudo-random configuration (Beardsley et al., 2012). It also benefits from the large frequency continuum, as baselines measure different *uv* modes at different frequencies.

As noted in §7.1.5, we image only visibility measurements from baselines shorter than 50 wavelengths in order to image in the "deconvolution-free" regime. This means that we reconstruct a minimum angular scale of 1.1°. The map presented in this chapter does not include information on angular scales smaller than 1.1°.

The largest angular scales in the map are limited by the MWA's shortest baselines. The MWA's shortest baseline is centered at 7.7 m in length, equal to 4.7 wavelengths at 182 MHz. However, because of the size of the baseline response kernel, the MWA samples *uv* modes below the center of this shortest baseline. From Figure 7.3, we find that short baselines sample the *uv* plane down to ~ 2 wavelengths. However, all visibilities that sample the *uv* plane at 2 wavelengths derive from longer baselines. A well-measured *uv* location is measured by an assortment of baselines distributed around the point itself. We therefore expect that measurements of the *uv* plane at 2 wavelengths experience bias.

Reflecting this sampling bias near the origin of the *uv* plane, we make a conservative estimate of the *uv* modes that are well-measured. We use the threshold down to which all *uv* locations are sampled (2.2 wavelengths) and add the radius of the baseline response kernel (3.9 wavelengths at 182 MHz) to derive an expectation that the *uv* plane is well-measured above 6.1 wavelengths. This corresponds to angular scales up to 9.4°. However, as we do not filter scales above 9.4°, the map includes some information from larger angular scales that could be of limited utility in certain simulation applications (see §7.5 for further discussion of using this diffuse map in simulation).

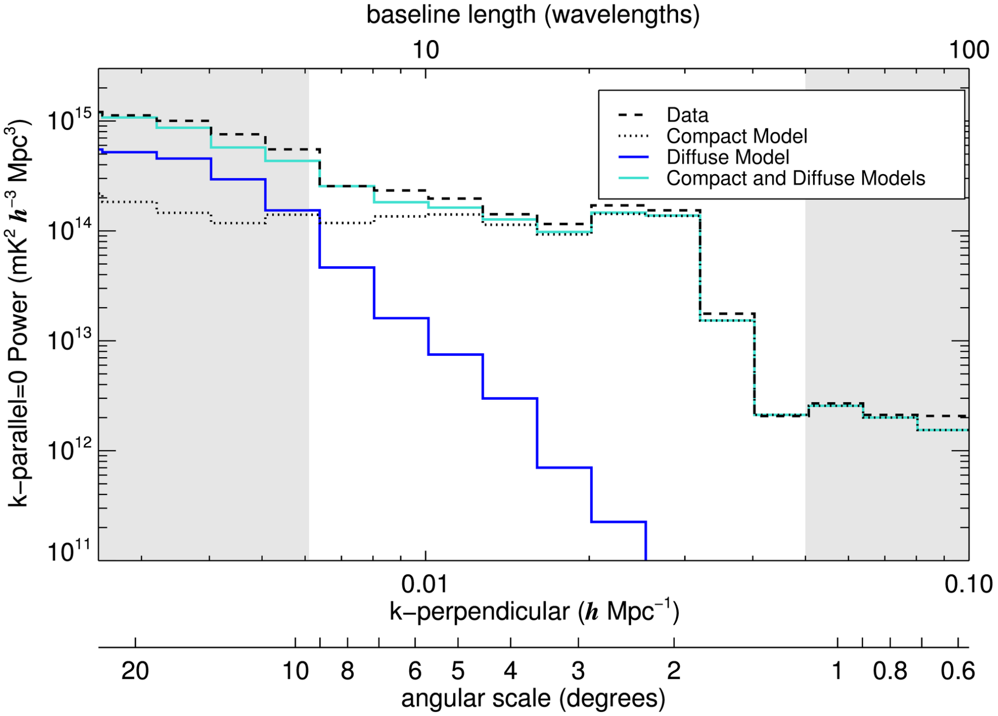
We are confident that the map presented in this paper accurately captures angular scales between 1.1° and 9.4° ($50 - 6.1$ wavelengths). The map can be used to reconstruct the *uv* plane for visibility simulation on those scales. We note that, when simulating the map's boundaries, the *uv* plane reconstruction can experience bias from aliasing effects. The map may not allow for accurate visibility simulation of observations near its boundaries.

Diffuse emission is a significant component of the total sky power at large scales. Figure 7.12 compares the diffuse map power to that of an MWA observation. The observation, one of the 173 that contribute to this map, is zenith-pointed and centered on the "EoR-0" field (centered at RA 0 hours, Dec. $-27°$ and plotted in Figure 7.10). The dashed black lines present the frequency-averaged angular power spectra of that data. The solid blue lines in Figure 7.12 represent the power spectra of the diffuse map presented in §7.2. Here the include Stokes I, Q, and U emission but omit Stokes V. The shaded grey regions of the plots represent scales outside our confidence region of 1.1° to 9.4°. Figure 7.12(a) depicts power spectra from the East-West aligned (or p) dipoles, and Figure 7.12(b) depicts power spectra from North-South aligned (or q) dipoles.

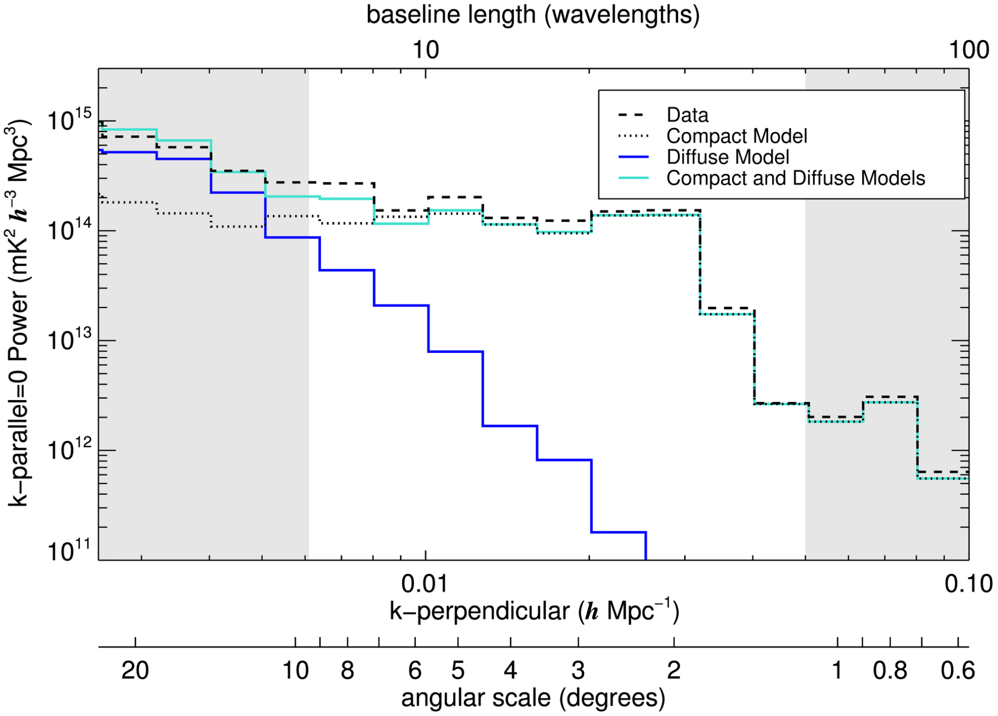
The dotted black lines correspond to the power spectra of the compact source model described in §7.1.2 and based on the GLEAM catalog (Hurley-Walker et al., 2017). We see a clear discrepancy between this model and the data at large scales, corresponding to the absence of diffuse structure in the model. The solid cyan line represents the sky model derived from combining the diffuse map and the compact source model. We find that the diffuse map recovers the power missing in the compact source model, producing a power spectrum that well-approximates that of the data.

7.4 Evaluating Map Accuracy

The diffuse map is intended to enable accurate visibility modeling for calibration. The most comprehensive validation of the diffuse map's agreement with data is to quantify the



(a) East-West polarization



(b) North-South polarization

Figure 7.12: Angular power spectrum comparison. (Caption continued.)

Figure 7.12: Comparison of the angular power spectra of data and three simulated sky models. The power spectrum represents the frequency-averaged (often denoted k-parallel $= 0$ or $k_{\|} = 0$) power at different angular scales on the sky (often denoted k-perpendicular or $k_{\perp}$). The dashed black line corresponds to a single zenith-pointed observation from the MWA of the EoR-0 field. The other three lines correspond to simulations of the same observation based on three different sky models. The dotted black line corresponds to the compact source model derived from the GLEAM catalog (Hurley-Walker et al., 2017) and commonly used in calibration (see §7.1.3). The blue line corresponds to the polarized diffuse model presented in §7.2, and the cyan line corresponds to the combined model derived by adding the compact model to the diffuse map. (a) corresponds to the East-West aligned dipoles, and (b) corresponds to the North-South aligned dipoles. We find that the diffuse map recovers power at large scales on the sky that is missing from the compact source model. The shaded gray regions represent scales on which we do not have confidence in the diffuse map. On the right side of each plot, angular scales measured by baselines longer than 50 wavelengths are omitted from the diffuse map presented in this chapter. On the left, measurements shorter than 6.1 wavelengths are expected to exhibit bias.

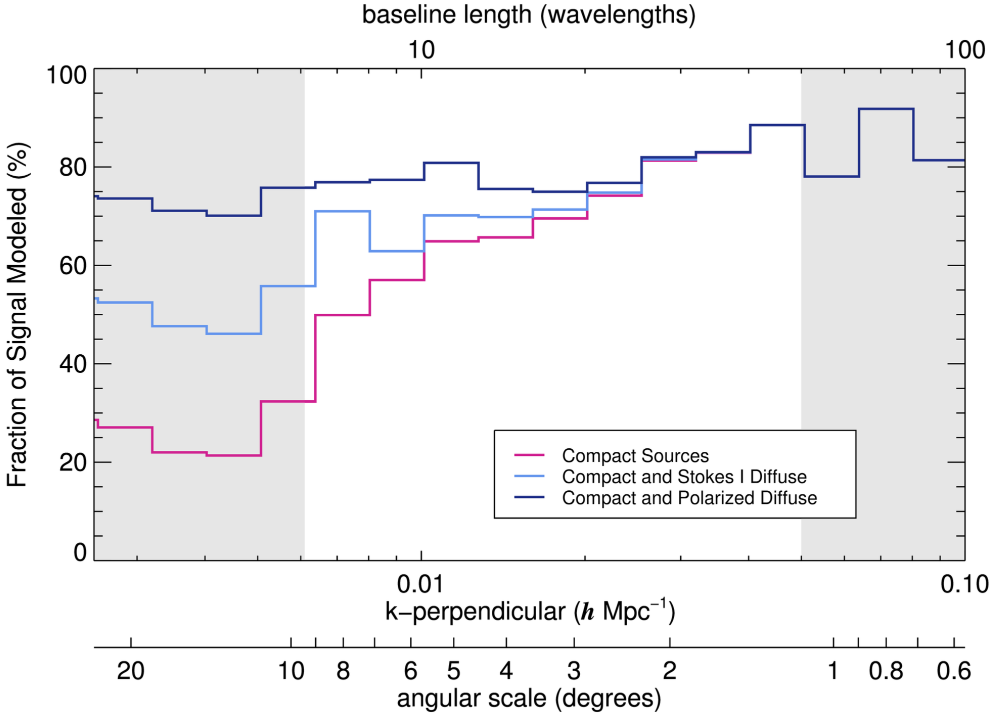
discrepancy between measured visibilities and those modeled from the map.

We define a metric of model accuracy as

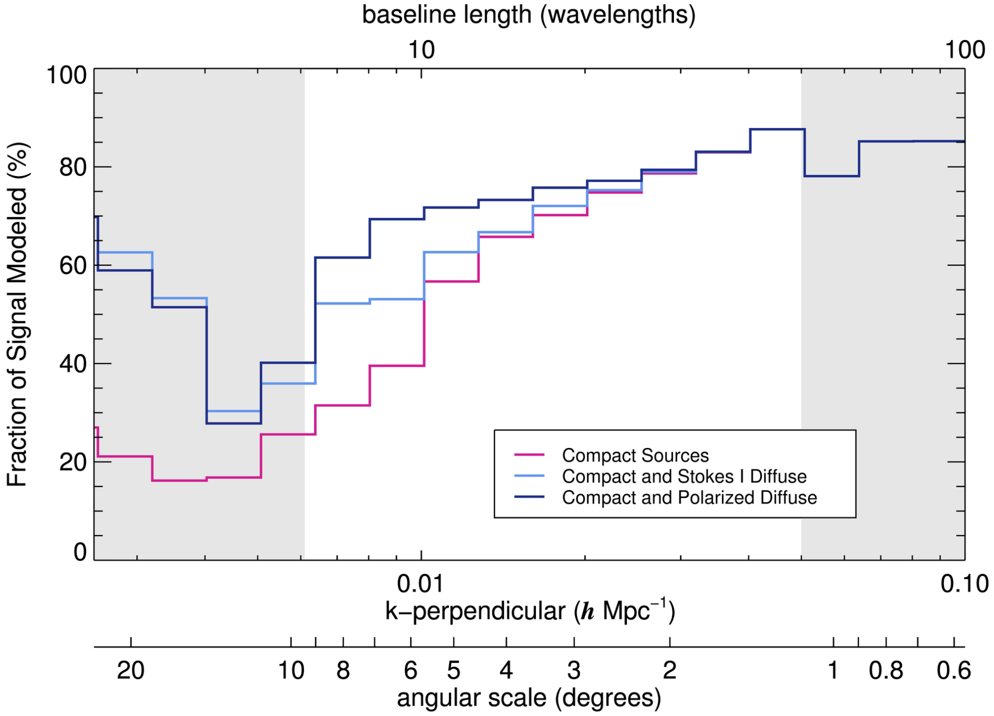
$$\text{fraction of signal modeled} = 100\% \times \left(1 - \frac{\text{RMS}\left[\text{Grid}\left(\boldsymbol{v}_{\text{d}} - \boldsymbol{v}_{\text{m}}\right)\right]}{\text{RMS}\left[\text{Grid}\left(\boldsymbol{v}_{\text{d}}\right)\right]}\right). \tag{7.5}$$

Here $\boldsymbol{v}_{\text{d}}$ are the measured visibilities and $\boldsymbol{v}_{\text{m}}$ are the modeled visibilities derived from a sky model. The differenced values $\boldsymbol{v}_{\text{d}} - \boldsymbol{v}_{\text{m}}$ capture the discrepancy between the data and model and are often called "residual" visibilities (Barry et al., 2019a; Byrne et al., 2019). "Grid" denotes the gridding operation, which transforms the visiblities into a reconstructed *uv* plane, and "RMS" denotes evaluating the RMS across *uv* pixels. We calculate the RMS in bands of baseline length to derive a metric of modeling accuracy as a function of baseline length, or angular scale. Values of 100% would indicate that the modeled visibilities exactly equal the measured visibilities; values of 0% would indicate very poor agreement, where the discrepancy between the measured and modeled visibilities has equivalent magnitude to the measured visibilities.

In Figure 7.13 we use this metric to evaluate the accuracy of modeling a single observation.



(a) East-West polarization



(b) North-South polarization

Figure 7.13: Modeling accuracy comparison. (Caption continued.)

Figure 7.13: Comparison of the modeling accuracy of one observation with three distinct sky models. The observation is the zenith-pointed and centered on the “EoR-0” field. The fraction of signal modeled is defined by Equation 7.5 and quantifies agreement between the measured and modeled visibilities as a function of baseline length. Larger values indicate better modeling accuracy. The magenta lines represent a sky model with compact sources only, the light blue lines include unpolarized diffuse emission in the sky model, and the dark blue lines include polarized diffuse emission. We find that the polarized diffuse map improves modeling accuracy of short baselines to approximately that of long baselines.

We use the same zenith-pointed observation of the “EoR-0” field plotted in Figure 7.12. The plots in Figure 7.13 present the results from three sky models: the magenta lines correspond to the compact source model discussed in §7.1.2 and based on GLEAM (Hurley-Walker et al., 2017); the light blue lines additionally include unpolarized diffuse emission, plotted in Figure 7.6; and the dark blue lines include polarized Stokes Q and U emission as well (Figures 7.7 and 7.8), which we have corrected to account for the observation’s ionospheric rotation measure, as discussed in §7.1.6. Figure 7.13(a) presents results from the East-West aligned dipoles, and Figure 7.13(b) presents results from the North-South aligned dipoles. The shaded grey regions represent scales outside our range of confidence for the diffuse map (see §7.3).

The magenta line in Figure 7.13 indicates that, for a sky model consisting of compact sources only, modeling accuracy degrades significantly at short baselines. This metric motivates calibrating with only baselines longer than 50 wavelengths, as discussed in §7.1.2. We find that the diffuse map recovers a significant fraction of the power at large scales, particularly when including polarized emission. For baselines of length 6.1-50 wavelengths, where we are confident in the accuracy of the diffuse map (see §7.3), the polarized diffuse map improves modeling accuracy to levels commensurate to those of long baselines.

Figure 7.13 shows that the diffuse map models a larger fraction of the data in the East-West polarization (Figure 7.13(a)) than in the North-South polarization (Figure 7.13(b)). This effect is most pronounced for baselines shorter than 6.1 wavelengths and may be attributable to the modeling bias discussed in §7.3. However, we also expect that this effect emerges from the diffuse map's finite extent on the sky. The "EoR-0" field is at zenith for this observation, and the bright galactic center is visible above the horizon to the west. The MWA's North-South aligned dipoles are sensitive to low-elevation emission to the east and west and could therefore detect power from the galactic plane that is not captured in our sky model. This points to the diffuse map's limitation in modeling widefield observations near the edges of the mapped region. Future work could measure the galactic plane and other regions beyond the limits of the map presented here.

7.5 Discussion

The map presented in this chapter is intended to enable accurate visibility simulations for a wide range of low-frequency radio instruments, with the particular goal of facilitating sky modeling for precision calibration of 21 cm cosmology measurements. In this section we discuss the process of implementing this map for visibility simulation and calibration.

We have removed compact sources using to the FHD deconvolution algorithm discussed in §7.1.3. To enable accurate visibility modeling, the map must be combined with a source catalog such as the modified GLEAM catalog discussed in §7.1.2 or the Long Baseline EoR Survey (LoBES; Lynch et al. in review). The success of precision calibration for 21 cm cosmology depends on highly accurate and complete modeling of pointlike and compact extended sources (Carroll et al., 2016; Line et al., 2020; Zhang et al., 2020).

The diffuse map has physical surface brightness units of Jy/sr, which can be converted to brightness temperature units such as Kelvin. Many visibility simulators represent pixelated images as a series of point sources at the pixel centers. To enable proper normalization, the map must then be converted into units of flux density by multiplying the given values in Jy/sr by the pixel area.

In order to properly model polarized emission in the widefield limit, the visibility simulator must properly convert between the Stokes polarization parameters, defined in Equation 6.1, and the instrumental response using the full direction-dependent instrumental Jones matrix. In the absence of a fully polarized simulator, only the unpolarized Stokes I emission should be used.

As demonstrated in §7.4, inclusion of Stokes Q and U emission significantly improves visibility modeling in the EoR-0 field, where we observe bright linearly polarized structure. We expect simulations of other fields to benefit less from polarized modeling. In fields with very faint polarized emission, Stokes Q and U could be dominated by imaging errors, in which case visibility simulators could benefit from omitting Stokes Q and U emission entirely. We do not expect modeling Stokes V emission to improve visibility simulations in any field.

The apparent Stokes Q and U emission is highly dependent on ionospheric conditions. If a simulation models this polarized emission, it should account for Faraday rotation from propagation through the ionosphere. At minimum, this should consist of measuring a single ionospheric RM per snapshot observation.

The diffuse map represents continuum observations across a frequency range of 167-198 MHz. It will most accurately represent observations at or near the central frequency of 182 MHz. We have not measured the spectral index or RM of this emission, and it is unclear how to interpolate the map far from this frequency.

Since the map does not capture the diffuse emission's variation across frequency, it cannot accurately model frequency-dependent structure. Emission with high RM amplitude depolarizes across this map's 31 MHz bandwidth, its polarized features reflect only structure with low RM amplitude (we estimate that emission with RM amplitude > 6 rad/m^2 depolarizes in this map). There has been significant work in the field exploring the impact of polarized emission with high RM amplitude on the cosmological measurement (Jelić et al., 2010; Moore et al., 2013; Nunhokee et al., 2017; Spinelli et al., 2019; Cunnington et al., 2021), but that emission is not captured in this map.

As discussed in §7.3, the map accurately models angular scales of $1.1^\circ - 9.4^\circ$, or *uv* modes between 6.1 and 50 wavelengths. The map does not include information from modes beyond 50 wavelengths. It can be used to accurately simulate visibilities from short baselines that only sample *uv* modes above 6.1 wavelengths. Below 6.1 wavelengths, reconstruction of the *uv* plane is susceptible to measurement bias, as discussed in §7.3. This bias can affect visibility simulations from baselines that sample *uv* modes below 6.1 wavelengths even when the center of the baseline response lies above 6.1 wavelengths. The impact of this bias is highly instrument-specific and could be minimal in certain situations. For example, instruments with large antennas have baselines that sample large swaths of the *uv* plane. Simulations of these arrays will likely be less vulnerable to this sky model measurement bias.

The diffuse map presented here has a finite extent on the sky. Unmodeled emission could introduce errors in the simulated visibilities of widefield observations that exceed the map's boundaries. Even if emission in an instrument's primary beam is well-modeled, modeling error in the far sidelobes can introduce frequency-dependent calibration errors at a level that is significant for EoR science (Barry et al., 2016; Pober et al., 2016; Byrne et al., 2019). As noted in §7.4, the bright galactic center, slightly west of the westernmost extent of the map, is of particularly concern for observations of the EoR-0 field. These errors are mitigated by ensuring that any observational fields simulated with this map are well-contained within the map extent.

As mentioned in §7.3, Fourier aliasing effects from the map boundaries can also contaminate simulated visibilities. For simulation of observations near the map boundaries, this aliasing can introduce errors in the reconstructed *uv* plane.

Imaging artifacts from poorly deconvolved compact sources can produce errors in simulated visibilities. These errors could be mitigated with improved polarized beam modeling and calibration. Of particular concern are imaging errors near Fornax A and Pictor A that appear predominantly in Stokes Q and U. In order to reduce aliasing in the reconstructed *uv* plane, we do not mask those regions. In certain simulation applications, mitigation approaches such as masking, interpolation over the affected regions, and/or omitting polarized emission from the simulation could be beneficial. The appropriate implementation of these techniques depends on the specific instrumental response and simulation procedure.

Other field-dependent systematic errors are illuminated by the standard deviation plots in Figure 7.11. These plots suggest that some observations perform relatively poorly. Most notably, an observation just east of Pictor A shows significant standard deviation structure in Stokes U and V. This is likely related to poor calibration and could potentially be improved with better modeling of the widefield polarized beam and Pictor A. We also observe high standard deviation values in regions of low observational coverage (as plotted in Figure 7.1(a)). This is apparent at the north edge of the Stokes I standard deviation plot and the top right corner of the Stokes Q and U plots. Furthermore, the high standard deviation values in the lower left corner of the Stokes Q and U in Figure 7.11 suggest Stokes I to Q and U polarization leakage in the vicinity of the Vela supernova remnant. Once again, visibility simulation could benefit from avoiding these fields, although masking runs the risk of introducing aliasing errors.

If implemented properly, we expect this diffuse map to improve calibration precision for cosmological measurements. Results presented in §7.4 indicate that this map enables accurate visibility modeling for baselines as short as 6.1 wavelengths in length. This significantly expands the baseline range available for calibration. As a result, this diffuse map can allow for better calibration signal-to-noise, enable precision calibration of highly compact arrays, and mitigate the frequency-dependent calibration errors inherent to long baseline calibration (Ewall-Wice et al., 2016). Improved calibration sky models that include this map could facilitate the next generation of EoR power spectrum results, allowing for highly sensitive measurements of the the cosmological 21 cm signal.

Chapter 8

LOOKING AHEAD: THE NEXT GENERATION OF 21 CM ANALYSES

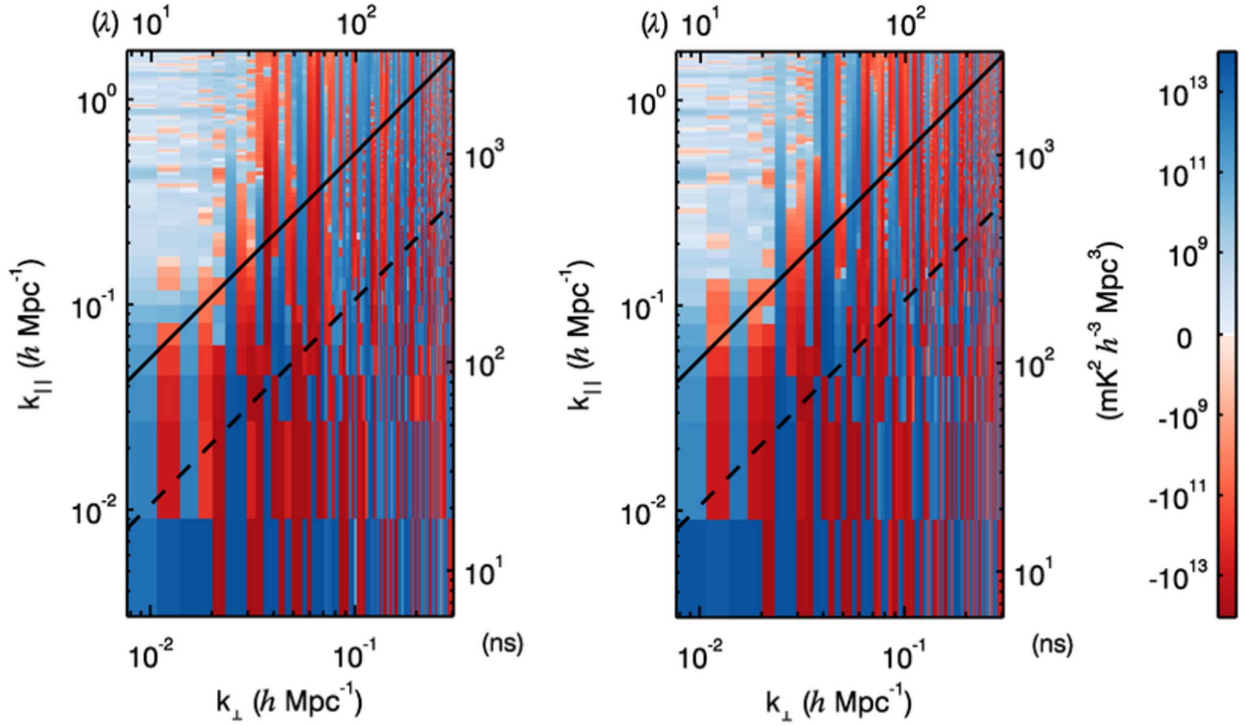
The work contained in this thesis enables the next generation of 21 cm cosmology measurements by tackling the problem of calibration error, a dominant limitation of previous 21 cm analyses. Here we discuss next steps for achieving improved calibration performance and increasing analysis precision for EoR science.

8.1 Calibration with a Map of Polarized Diffuse Emission

In Chapter 7 we present a map of polarized diffuse emission covering much of the Southern Hemisphere sky, including most of the fields of interest for EoR science with the MWA and HERA. This map was enabled by a new polarized imaging pipeline developed with FHD and described in Chapter 6.

Currently the best limits in the field were produced by calibrating to compact sources only (Barry et al., 2019b; Li et al., 2019; Trott et al., 2020). The omission of diffuse emission means that these maps can only properly model long baselines. Short baselines must therefore be omitted from calibration. The diffuse map presented in this thesis expands the calibration sky model, making it accurate over a wider range of angular scales. Calibration can then make full use of much shorter baselines. This is important for a number of reasons. First, using a larger fraction of the data improves calibration signal-to-noise. This is particularly important for compact arrays, such as HERA and the MWA Phase II, with a high fraction of short baselines. In addition, short baseline calibration improves spectral power leakage (Ewall-Wice et al., 2016).

To test the impact of this new diffuse map on calibration performance, we use it to



(a) east-west polarization

(b) north-south polarization

Figure 8.1: 2-D power spectrum of difference visibilities. The data corresponds to a single night of observations from the MWA Phase I, taken in August 2013. Visibilities calibrated with the polarized diffuse map presented in Chapter 7 with baselines of all lengths are subtracted from those calibrated to compact sources only using only baselines longer than 50 wavelengths. The overall normalization of the gain amplitudes between the two calibration approaches is constrained to be equal. The EoR window region of each plot is predominantly positive (blue), indicating that calibration with the diffuse map reduces contamination of the EoR window.

analyze one night of data from the MWA Phase I. This data was taken on August 23, 2013 and is part of the data set used to produce the EoR limits in Beardsley et al. 2016 and Barry et al. 2019b. It covers the EoR-0 field, centered at RA 0 h and Dec. -27°. We use simple per-frequency and per-antenna calibration and calibrate both with and without modeled diffuse emission.

For the first processing run, we calibrate without diffuse emission. The sky model consists of the GLEAM catalog (Hurley-Walker et al., 2017) supplemented with additional compact source models. This is the catalog used to calibrate data in the diffuse map and described in detail §7.1.2. As this sky model does not include diffuse emission, we use only baselines longer than 50 wavelengths to calibrate.

For the second processing run, we supplement the compact source catalog with the polarized diffuse map presented in Chapter 7. We include linearly polarized (Stokes Q and U) emission but omit Stokes V emission. For each two-minute observation, we estimate an ionospheric rotation measure with the RMEXTRACT package. The calibration sky model is then adjusted for ionospheric Faraday rotation to better align with the observation's apparent polarized measurement. For this run, we calibrate with baselines of all lengths. We also constrain the average amplitude of the gains for each observation and polarization to match that of the first run. This is important for enabling direct comparison between the two runs. Changes in the sky model and calibration procedure can introduce low-level overall normalization offsets that can make it difficult to compare results, and constraining the amplitude of the gains to match eliminates this effect.

Figure 8.1 presents results from these two runs. We calculate 2-D power spectra from each run and subract them, producing differenced, or jackknifed, power spectra. Here the power spectra from the second run, calibrated to the diffuse map, are subtracted from those of the first run. On the left, Figure 8.1(a) presents results from the east-west instrumental polarization and 8.1(b) presents results from the north-south instrumental polarization. For each polarization, we note that the EoR window at low $k_\perp$ is predominantly blue, indicating that calibration with the diffuse map reduces power in the EoR window. This is a promising indication that the diffuse map can improve calibration performance, reducing foregrounds' contamination of the EoR window and enabling good calibration with a wide range of baseline lengths.

The calibration test presented here is a very basic implementation of calibration. State of the art calibration approaches in the field are not per-frequency and instead impose some bandpass calibration constraint to reduce frequency-dependent calibration error by, for exampling, using the autocorrelation visibilities to constrain the instrument's bandpass response (Barry et al., 2019b; Li et al., 2019). One of the primary aims of these approaches is to overcome calibration limitations related to calibrating with long baselines only. We expect that high-performance calibration with the diffuse map presented in this thesis will likewise benefit from sophisticated bandpass calibration approaches, but these techniques may differ from those optimized for compact source calibration. Further work is necessary to determine if calibration with a diffuse map will benefit from autocorrelation calibration or other advanced bandpass calibration techniques.

8.2 New Calibration Approaches with Unified Calibration

In recent years there has been great effort towards overcoming calibration error in EoR power spectrum analyses (Liu et al. 2010; Kazemi et al. 2011; Carroll et al. 2016; Dillon et al. 2018; Barry et al. 2019b; Li et al. 2019; Sob et al. 2020; Kern et al. 2020; and many others). However, until the arrival of unified calibration, presented in Chapter 4, the paradigm in the field delineated between sky-based and redundant calibration as separate and distinct classes of calibration. Unified calibration offers a method of combining these approaches or working in an intermediate regime between the two.

In Chapter 4, we show that this intermediate regime more accurately captures the reality of the interferometric measurement. Sky-based calibration assumes very good *a priori* models of the sky and instrument response. It further assumes independent visibility measurements, neglecting visibility covariance from baseline response overlap in the *uv* plane. On the other hand, redundant calibration assumes that baselines within a redundant set are perfectly redundant both in antenna positions and response shapes, that baselines from different redundant baseline sets exhibit no covariance, and that a sky model can accurately constrain the absolute calibration parameters. The literature shows that violating these assumptions introduces calibration errors that contaminate the measurement. For example, Barry et al. 2016 shows that sky model incompleteness leads to sky-based calibration errors, Joseph et al. 2018 and Orosz et al. 2019 explore redundant calibration errors from antenna position offsets and response irregularities, and Byrne et al. 2019, replicated in large part in Chapter 3, examines the effect of sky model errors on the absolute calibration step of redundant calibration. Through unified calibration, we can relax the assumptions of each sky-based and redundant calibration and achieve a more physically-motivated calibration model. In Chapter 5, we demonstrate in a simplified simulation that this improves calibration performance.

Unified calibration situates calibration within a rigorous statistical framework. By relating the calibration problem to a Bayesian likelihood function, we can directly incorporate measured uncertainty of calibration parameters into the calibration model. This vastly expands our capability to build high performance calibrators for a wide variety of instruments.

The unified calibration framework is not a specific calibration implementation nor a prescription for optimal calibration. Instead, it represents a broad framework for how to conceptualize and define the calibration problem. In this way, it is the first radically new calibration framework since the development of redundant calibration (Wieringa, 1992; Liu et al., 2010). In the coming years, implementation of unified calibration will require significant exploration, determining the appropriate models to best represent calibration of various instruments. In a field where calibration precision is of the utmost importance, unified calibration offers a roadmap for the next generation of high performing calibration implementations.

8.3 Looking Toward an EoR Detection

This thesis is one step towards the elusive goal of measuring the EoR power spectrum. It provides tools and techniques for improving calibration performance, yet calibration remains a formidable challenge for 21 cm cosmology. How will we overcome these challenges and achieve a detection?

Precision calibration requires an excellent sky model. As shown in Chapter 3, low-level errors in the sky model used in calibration significantly contaminate the EoR window. Along with the polarized diffuse map presented in Chapter 7, the field requires continued commitment to developing accurate and complete sky models.

In the same vein, redundant calibration does not eliminate—or even substantially reduce—these sky model requirements. Regular array layouts can be beneficial for EoR analyses because their redundant measurements improve sensitivity. However, the field is currently not limited by measurement sensitivity. We show in Chapter 3 that non-regular arrays with high *uv* coverage, despite not being redundantly calibratable, deliver better spectral calibration performance than redundantly calibrated regular arrays.

In Chapter 5, we demonstrated in simulation that using a statistically rigorous, physically motivated calibration model produces measurable improvements. In the future, 21 cm analyses must adopt state of the art calibration frameworks, potentially based on unified calibration. This also requires investment in measuring and quantifying error in various aspects of the calibration model—the accuracy of the sky model across angular scales, the agreement of redundant baselines, the accuracy of antenna positions, the error associated with beam response models, and more. Incorporating these metrics into the calibration model will improve measurement precision.

Finally, controlling the uniformity and spectral features of the instrument's response continues to be of critical importance for precision calibration. The best way to eliminate errors in calibrating fast spectral structure is to eliminate the structure itself. Ultimately, advances in antenna design and signal processing could substantially reduce the degrees of freedom required in calibration, thereby eliminating avenues for calibration error.

BIBLIOGRAPHY

N. Aghanim, Y. Akrami, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak, R. Battye, K. Benabed, J. P. Bernard, M. Bersanelli, P. Bielewicz, J. J. Bock, J. R. Bond, J. Borrill, F. R. Bouchet, F. Boulanger, M. Bucher, C. Burigana, R. C. Butler, E. Calabrese, J. F. Cardoso, J. Carron, B. Casaponsa, A. Challinor, H. C. Chiang, L. P. Colombo, C. Combet, D. Contreras, B. P. Crill, F. Cuttaia, P. De Bernardis, G. De Zotti, J. Delabrouille, J. M. Delouis, F. X. Désert, E. Di Valentino, C. Dickinson, J. M. Diego, S. Donzelli, O. Doré, M. Douspis, A. Ducout, X. Dupac, G. Efstathiou, F. Elsner, T. A. Enßlin, H. K. Eriksen, E. Falgarone, Y. Fantaye, J. Fergusson, R. Fernandez-Cobos, F. Finelli, F. Forastieri, M. Frailis, E. Franceschi, A. Frolov, S. Galeotta, S. Galli, K. Ganga, R. T. Génova-Santos, M. Gerbino, T. Ghosh, J. González-Nuevo, K. M. Górski, S. Gratton, A. Gruppuso, J. E. Gudmundsson, J. Hamann, W. Handley, F. K. Hansen, G. Helou, D. Herranz, S. R. Hildebrandt, E. Hivon, Z. Huang, A. H. Jaffe, W. C. Jones, A. Karakci, E. Keihänen, R. Keskitalo, K. Kiiveri, J. Kim, T. S. Kisner, L. Knox, N. Krachmalnicoff, M. Kunz, H. Kurki-Suonio, G. Lagache, J. M. Lamarre, M. Langer, A. Lasenby, M. Lattanzi, C. R. Lawrence, M. Le Jeune, J. P. Leahy, J. Lesgourgues, F. Levrier, A. Lewis, M. Liguori, P. B. Lilje, M. Lilley, V. Lindholm, M. López-Caniego, P. M. Lubin, Y. Z. Ma, J. F. Maciás-Pérez, G. Maggio, D. Maino, N. Mandolesi, A. Mangilli, A. Marcos-Caballero, M. Maris, P. G. Martin, M. Martinelli, E. Martínez-González, S. Matarrese, N. Mauri, J. D. McEwen, P. D. Meerburg, P. R. Meinhold, A. Melchiorri, A. Mennella, M. Migliaccio, M. Millea, S. Mitra, M. A. Miville-Deschênes, D. Molinari, A. Moneti, L. Montier, G. Morgante, A. Moss, S. Mottet, M. Münchmeyer, P. Natoli, H. U. Nørgaard-Nielsen, C. A. Oxborrow, L. Pagano, D. Paoletti, B. Partridge, G. Patanchon, T. J. Pearson, M. Peel, H. V. Peiris,

F. Perrotta, V. Pettorino, F. Piacentini, L. Polastri, G. Polenta, J. L. Puget, J. P. Rachen, M. Reinecke, M. Remazeilles, C. Renault, A. Renzi, G. Rocha, C. Rosset, G. Roudier, J. A. Rubiño-Martín, B. Ruiz-Granados, L. Salvati, M. Sandri, M. Savelainen, D. Scott, E. P. Shellard, M. Shiraishi, C. Sirignano, G. Sirri, L. D. Spencer, R. Sunyaev, A. S. Suur-Uski, J. A. Tauber, D. Tavagnacco, M. Tenti, L. Terenzi, L. Toffolatti, M. Tomasi, T. Trombetti, J. Valiviita, B. Van Tent, L. Vibert, P. Vielva, F. Villa, N. Vittorio, B. D. Wandelt, I. K. Wehus, M. White, S. D. White, A. Zacchei, and A. Zonca. Planck 2018 Results: I. Overview and the Cosmological Legacy of Planck. *Astron. Astrophys.*, 641, 2020. doi: 10.1051/0004-6361/201833880.

J. G. Albert, R. J. van Weeren, H. T. Intema, and H. J. A. Röttgering. Probabilistic Direction-Dependent Ionospheric Calibration for LOFAR-HBA. *Astron. Astrophys.*, 635: A147, 2020. doi: 10.1051/0004-6361/201937424.

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

J. W. M. Baars, R. Genzel, I. I. K. Pauliny-Toth, and A. Witzel. The Absolute Spectrum of Cas A: An Accurate Flux Density Scale and a Set of Secondary Calibrators. *Astron. Astrophys.*, 500:135–142, 1977.

R. Barkana. Studying the Sources of Cosmic Reionization with 21-cm Fluctuations. *Mon. Not. R. Astron. Soc.*, 397(3):1454–1463, 2009. doi: 10.1111/j.1365-2966.2009.14929.x.

N. Barry, B. Hazelton, I. Sullivan, M. F. Morales, and J. C. Pober. Calibration Requirements for Detecting the 21 cm Epoch of Reionization Power Spectrum and Implications for the SKA. *Mon. Not. R. Astron. Soc.*, 461(3):3135–3144, 2016. doi: 10.1093/mnras/stw1380.

N. Barry, A. P. Beardsley, R. Byrne, B. Hazelton, M. F. Morales, J. C. Pober, and I. Sullivan. The FHD/ϵppsilon Epoch of Reionisation Power Spectrum Pipeline. *Publ. Astron. Soc. Aust.*, 36:E026, 2019a. doi: 10.1017/pasa.2019.21.

N. Barry, M. Wilensky, C. M. Trott, B. Pindor, A. P. Beardsley, B. J. Hazelton, I. S. Sullivan, M. F. Morales, J. C. Pober, J. Line, B. Greig, R. Byrne, A. Lanman, W. Li, C. H. Jordan, R. C. Joseph, B. McKinley, M. Rahimi, S. Yoshiura, J. D. Bowman, B. M. Gaensler, J. N. Hewitt, D. C. Jacobs, D. A. Mitchell, N. Udaya Shankar, S. K. Sethi, R. Subrahmanyan, S. J. Tingay, R. L. Webster, and J. S. B. Wyithe. Improving the Epoch of Reionization Power Spectrum Results from Murchison Widefield Array Season 1 Observations. *Astrophys. J.*, 884(1):1, 2019b. doi: 10.3847/1538-4357/ab40a8.

A. P. Beardsley, B. J. Hazelton, M. F. Morales, R. J. Capallo, R. Goeke, D. Emrich, C. J. Lonsdale, W. Arcus, D. Barnes, G. Bernardi, J. D. Bowman, J. D. Bunton, B. E. Corey, A. Deshpande, L. DeSouza, B. M. Gaensler, L. J. Greenhill, D. Herne, J. N. Hewitt, D. L. Kaplan, J. C. Kasper, B. B. Kincaid, R. Koenig, E. Kratzenberg, M. J. Lynch, S. R. McWhirter, D. A. Mitchell, E. Morgan, D. Oberoi, S. M. Ord, J. Pathikulangara, T. Prabu, R. A. Remillard, A. E. Rogers, A. Roshi, J. E. Salah, R. J. Sault, N. U. Shankar, K. S. Srivani, J. Stevens, R. Subrahmanyan, S. J. Tingay, R. B. Wayth, M. Waterson, R. L. Webster, A. R. Whitney, A. Williams, C. L. Williams, and J. S. B. Wyithe. A New Layout Optimization Technique for Interferometric Arrays, Applied to the Murchison Widefield Array. *Mon. Not. R. Astron. Soc.*, 425(3):1781–1788, 2012. doi: 10.1111/j.1365-2966.2012.20878.x.

A. P. Beardsley, B. J. Hazelton, I. S. Sullivan, P. Carroll, N. Barry, M. Rahimi, B. Pindor, C. M. Trott, J. Line, D. C. Jacobs, M. F. Morales, J. C. Pober, G. Bernardi, J. D. Bowman, M. P. Busch, F. Briggs, R. J. Cappallo, B. E. Corey, A. de Oliveira-Costa, J. S. Dillon, D. Emrich, A. Ewall-Wice, L. Feng, B. M. Gaensler, R. Goeke, L. J. Greenhill, J. N. Hewitt, N. Hurley-Walker, M. Johnston-Hollitt, D. L. Kaplan, J. C. Kasper, H. S. Kim, E. Kratzenberg, E. Lenc, A. Loeb, C. J. Lonsdale, M. J. Lynch, B. McKinley, S. R.

McWhirter, D. A. Mitchell, E. Morgan, A. R. Neben, N. Thyagarajan, D. Oberoi, A. R. Offringa, S. M. Ord, S. Paul, T. Prabu, P. Procopio, J. Riding, A. E. E. Rogers, A. Roshi, N. U. Shankar, S. K. Sethi, K. S. Srivani, R. Subrahmanyan, M. Tegmark, S. J. Tingay, M. Waterson, R. B. Wayth, R. L. Webster, A. R. Whitney, A. Williams, C. L. Williams, C. Wu, and J. S. B. Wyithe. First Season MWA EoR Power Spectrum Results at Redshift 7. *Astrophys. J.*, 833(1):102, 2016. doi: 10.3847/1538-4357/833/1/102.

C. L. Bennett, A. J. Banday, K. M. Górski, G. Hinshaw, P. Jackson, P. Keegstra, A. Kogut, G. F. Smoot, D. T. Wilkinson, and E. L. Wright. Four-Year COBE DMR Cosmic Microwave Background Observations: Maps and Basic Results. *Astrophys. J.*, 464(1), 1996. doi: 10.1086/310075.

C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, L. Page, D. N. Spergel, G. S. Tucker, E. Wollack, E. L. Wright, C. Barnes, M. R. Greason, R. S. Hill, E. Komatsu, M. R. Nolta, N. Odegard, H. V. Peiris, L. Verde, and J. L. Weiland. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Preliminary Maps and Basic Results. *Astrophys. J. Suppl. Ser.*, 148(1):1–27, 2003. doi: 10.1086/377253.

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

P. Berger, L. B. Newburgh, M. Amiri, K. Bandura, J.-F. Cliche, L. Connor, M. Deng, N. Denman, M. Dobbs, M. Fandino, A. J. Gilbert, D. Good, M. Halpern, D. Hanna, A. D. Hincks, G. Hinshaw, C. Höfer, A. M. Johnson, T. L. Landecker, K. W. Masui, J. Mena Parra, N. Oppermann, U.-L. Pen, J. B. Peterson, A. Recnik, T. Robishaw, J. R. Shaw, S. Siegel, K. Sigurdson, K. Smith, E. Storer, I. Tretyakov, K. Van Gassen, K. Vanderlinde, and D. Wiebe. Holographic Beam Mapping of the CHIME Pathfinder Array. *Ground-based Airborne Telesc. VI*, 9906:99060D, 2016. doi: 10.1117/12.2233782.

G. Bernardi, L. J. Greenhill, D. A. Mitchell, S. M. Ord, B. J. Hazelton, B. M. Gaensler, A. De Oliveira-Costa, M. F. Morales, N. U. Shankar, R. Subrahmanyan, R. B. Wayth, E. Lenc, C. L. Williams, W. Arcus, B. S. Arora, D. G. Barnes, J. D. Bowman, F. H. Briggs, J. D. Bunton, R. J. Cappallo, B. E. Corey, A. Deshpande, L. Desouza, D. Emrich, R. Goeke, D. Herne, J. N. Hewitt, M. Johnston-Hollitt, D. Kaplan, J. C. Kasper, B. B. Kincaid, R. Koenig, E. Kratzenberg, C. J. Lonsdale, M. J. Lynch, S. R. McWhirter, E. Morgan, D. Oberoi, J. Pathikulangara, T. Prabu, R. A. Remillard, A. E. Rogers, A. Roshi, J. E. Salah, R. J. Sault, K. S. Srivani, J. Stevens, S. J. Tingay, M. Waterson, R. L. Webster, A. R. Whitney, A. Williams, and J. S. Wyithe. A 189 MHz, 2400 $\deg^2$ Polarization Survey with the Murchison Widefield Array 32-Element Prototype. *Astrophys. J.*, 771(2):105, 2013. doi: 10.1088/0004-637X/771/2/105.

S. Bhatnagar, T. J. Cornwell, K. Golap, and J. M. Uson. Correcting Direction-Dependent Gains in the Deconvolution of Radio Interferometric Images. *Astron. Astrophys.*, 487(1): 419–429, 2008. doi: 10.1051/0004-6361:20079284.

J. D. Bowman, M. F. Morales, and J. N. Hewitt. The Sensitivity of First-Generation Epoch of Reionization Observatories and Their Potential for Differentiating Theoretical Power Spectra. *Astrophys. J.*, 638(1):20–26, 2006. doi: 10.1086/498703.

J. D. Bowman, M. F. Morales, and J. N. Hewitt. Constraints on Fundamental Cosmological Parameters with Upcoming Redshifted 21 cm Observations. *Astrophys. J.*, 661(1):1–9, 2007. doi: 10.1086/516560.

J. D. Bowman, A. E. E. Rogers, and J. N. Hewitt. Toward Empirical Constraints on the Global Redshifted 21 cm Brightness Temperature During the Epoch of Reionization. *Astrophys. J.*, 676(1):1–9, 2008. doi: 10.1086/528675.

R. Byrne and D. Jacobs. Development of a High Throughput Cloud-Based Data Pipeline for 21 cm Cosmology. *Astron. Comput.*, 34:100447, 2021. doi: 10.1016/j.ascom.2021.100447.

R. Byrne, M. F. Morales, B. Hazelton, W. Li, N. Barry, A. P. Beardsley, R. Joseph, J. Pober, I. Sullivan, and C. Trott. Fundamental Limitations on the Calibration of Redundant 21 cm Cosmology Instruments and Implications for HERA and the SKA. *Astrophys. J.*, 875 (1):70, 2019. doi: 10.3847/1538-4357/ab107d.

R. Byrne, M. F. Morales, B. J. Hazelton, and M. Wilensky. A Unified Calibration Framework for 21 cm Cosmology. *Mon. Not. R. Astron. Soc.*, 503(2):2457–2477, 2021. doi: 10.1093/mnras/stab647.

P. A. Carroll, J. Line, M. F. Morales, N. Barry, A. P. Beardsley, B. J. Hazelton, D. C. Jacobs, J. C. Pober, I. S. Sullivan, R. L. Webster, G. Bernardi, J. D. Bowman, F. Briggs, R. J. Cappallo, B. E. Corey, A. de Oliveira-Costa, J. S. Dillon, D. Emrich, A. Ewall-Wice, L. Feng, B. M. Gaensler, R. Goeke, L. J. Greenhill, N. Hewitt, N. Hurley-Walker, M. Johnston-Hollitt, D. L. Kaplan, J. C. Kasper, H. S. Kim, E. Kratzenberg, E. Lenc, A. Loeb, C. J. Lonsdale, M. J. Lynch, B. McKinley, S. R. McWhirter, D. A. Mitchell, E. Morgan, A. R. Neben, D. Oberoi, A. R. Offringa, S. M. Ord, S. Paul, B. Pindor, T. Prabu, P. Procopio, J. Riding, A. E. Rogers, A. Roshi, N. U. Shankar, S. K. Sethi, K. S. Srivani, R. Subrahmanyan, M. Tegmark, N. Thyagarajan, S. J. Tingay, C. M. Trott, M. Waterson, R. B. Wayth, A. R. Whitney, A. Williams, C. L. Williams, C. Wu, and J. S. Wyithe. A High Reliability Survey of Discrete Epoch of Reionization Foreground Sources in the MWA EoR0 Field. *Mon. Not. R. Astron. Soc.*, 461(4):4151–4175, 2016. doi: 10.1093/mnras/stw1599.

T. R. Choudhury, E. Puchwein, M. G. Haehnelt, and J. S. Bolton. Lyman α Emitters Gone Missing: Evidence for Late Reionization? *Mon. Not. R. Astron. Soc.*, 452(1):261–277, 2015. doi: 10.1093/mnras/stv1250.

J. J. Condon, W. D. Cotton, E. W. Greisen, Q. F. Yin, R. A. Perley, G. B. Taylor, and

J. J. Broderick. The NRAO VLA Sky Survey. *Astron. J.*, 115(5):1693–1716, 1998. doi: 10.1086/300337.

S. Cunnington, M. O. Irfan, I. P. Carucci, A. Pourtsidou, and J. Bobin. 21-cm Foregrounds and Polarization Leakage: Cleaning and Mitigation Strategies. *Mon. Not. R. Astron. Soc.*, 504(1):208–227, 2021. doi: 10.1093/mnras/stab856.

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

E. de Lera Acedo, C. M. Trott, R. B. Wayth, N. Fagnoni, G. Bernardi, B. Wakley, L. V. E. Koopmans, A. J. Faulkner, and J. G. bij de Vaate. Spectral Performance of SKA Log-Periodic Antennas I: Mitigating Spectral Artefacts in SKA1-LOW 21 cm Cosmology Experiments. *Mon. Not. R. Astron. Soc.*, 469(3):2662–2671, 2017. doi: 10.1093/mnras/stx904.

A. de Oliveira-Costa, M. Tegmark, B. M. Gaensler, J. Jonas, T. L. Landecker, and P. Reich. A Model of Diffuse Galactic Radio Emission from 10 MHz to 100 GHz. *Mon. Not. R. Astron. Soc.*, 388(1):247–260, 2008. doi: 10.1111/j.1365-2966.2008.13376.x.

D. R. DeBoer, A. R. Parsons, J. E. Aguirre, P. Alexander, Z. S. Ali, A. P. Beardsley, G. Bernardi, J. D. Bowman, R. F. Bradley, C. L. Carilli, C. Cheng, E. de Lera Acedo, J. S. Dillon, A. Ewall-Wice, G. Fadana, N. Fagnoni, R. Fritz, S. R. Furlanetto, B. Glendenning, B. Greig, J. Grobbelaar, B. J. Hazelton, J. N. Hewitt, J. Hickish, D. C. Jacobs, A. Julius, M. Kariseb, S. A. Kohn, T. Lekalake, A. Liu, A. Loots, D. MacMahon, L. Malan, C. Malgas, M. Maree, Z. Martinot, N. Mathison, E. Matsetela, A. Mesinger, M. F. Morales, A. R. Neben, N. Patra, S. Pieterse, J. C. Pober, N. Razavi-Ghods, J. Ringuette, J. Robnett, K. Rosie, R. Sell, C. Smith, A. Syce, M. Tegmark, N. Thyagarajan, P. K. G. Williams, and H. Zheng. Hydrogen Epoch of Reionization Array (HERA). *Publ. Astron. Soc. Pacific*, 129(974):045001, 2017. doi: 10.1088/1538-3873/129/974/045001.

J. S. Dillon and A. R. Parsons. Redundant Array Configurations for 21 cm Cosmology. *Astrophys. J.*, 826(2):181, 2016. doi: 10.3847/0004-637X/826/2/181.

J. S. Dillon, A. Liu, and M. Tegmark. A Fast Method for Power Spectrum and Foreground Analysis for 21 cm Cosmology. *Phys. Rev. D*, 87(4):043005, 2013. doi: 10.1103/PhysRevD.87.043005.

J. S. Dillon, S. A. Kohn, A. R. Parsons, J. E. Aguirre, Z. S. Ali, G. Bernardi, N. S. Kern, W. Li, A. Liu, C. D. Nunhokee, and J. C. Pober. Polarized Redundant-Baseline Calibration for 21 cm Cosmology Without Adding Spectral Structure. *Mon. Not. R. Astron. Soc.*, 477 (4):5670–5681, 2018. doi: 10.1093/mnras/sty1060.

J. S. Dillon, M. Lee, Z. S. Ali, A. R. Parsons, N. Orosz, C. D. Nunhokee, P. La Plante, A. P. Beardsley, N. S. Kern, Z. Abdurashidova, J. E. Aguirre, P. Alexander, Y. Balfour, G. Bernardi, T. S. Billings, J. D. Bowman, R. F. Bradley, P. Bull, J. Burba, S. Carey, C. L. Carilli, C. Cheng, D. R. DeBoer, M. Dexter, E. de Lera Acedo, J. Ely, A. Ewall-Wice, N. Fagnoni, R. Fritz, S. R. Furlanetto, K. Gale-Sides, B. Glendenning, D. Gorthi, B. Greig, J. Grobbelaar, Z. Halday, B. J. Hazelton, J. N. Hewitt, J. Hickish, D. C. Jacobs, A. Julius, J. Kerrigan, P. Kittiwisit, S. A. Kohn, M. Kolopanis, A. Lanman, T. Lekalake, D. Lewis, A. Liu, Y. Z. Ma, D. MacMahon, L. Malan, C. Malgas, M. Maree, Z. E. Martinot, E. Matsetela, A. Mesinger, M. Molewa, M. F. Morales, T. Mosiane, S. Murray, A. R. Neben, B. Nikolic, R. Pascua, N. Patra, S. Pieterse, J. C. Pober, N. Razavi-Ghods, J. Ringuette, J. Robnett, K. Rosie, M. G. Santos, P. Sims, C. Smith, A. Syce, M. Tegmark, N. Thyagarajan, P. K. G. Williams, and H. Zheng. Redundant-Baseline Calibration of the Hydrogen Epoch of Reionization Array. *Mon. Not. R. Astron. Soc.*, 499(4):5840–5861, 2020. doi: 10.1093/mnras/staa3001.

M. W. Eastwood, M. M. Anderson, R. M. Monroe, G. Hallinan, B. R. Barsdell, S. A. Bourke, M. A. Clark, S. W. Ellingson, J. Dowell, H. Garsden, L. J. Greenhill, J. M. Hartman, J. Kocz, T. J. W. Lazio, D. C. Price, F. K. Schinzel, G. B. Taylor, H. K.

Vedantham, Y. Wang, and D. P. Woody. The Radio Sky at Meter Wavelengths: m-Mode Analysis Imaging with the OVRO-LWA. *Astron. J.*, 156(1):32, 2018. doi: 10.3847/1538-3881/aac721.

T. A. Enßlin, S. Hutschenreuter, V. Vacca, and N. Oppermann. The Galaxy in Circular Polarization: All-Sky Radio Prediction, Detection Strategy, and the Charge of the Leptonic Cosmic Rays. *Phys. Rev. D*, 96(4):043021, 2017. doi: 10.1103/PhysRevD.96.043021.

A. Ewall-Wice, J. S. Dillon, A. Liu, and J. Hewitt. The Impact of Modeling Errors on Interferometer Calibration for 21 cm Power Spectra. *Mon. Not. R. Astron. Soc.*, 470(2): 1849–1870, oct 2016. doi: 10.1093/mnras/stx1221.

N. Fagnoni, E. de Lera Acedo, N. Drought, D. R. DeBoer, D. Riley, N. Razavi-Ghods, S. Carey, and A. R. Parsons. Design of the New Wideband Vivaldi Feed for the HERA Radio-Telescope Phase II. *arXiv e-prints*, 2020.

X. Fan, C. L. Carilli, and B. Keating. Observational Constraints on Cosmic Reionization. *Annu. Rev. Astron. Astrophys.*, 44(1):415–462, 2006a. doi: 10.1146/annurev.astro.44.051905.092514.

X. Fan, M. A. Strauss, R. H. Becker, R. L. White, J. E. Gunn, G. R. Knapp, G. T. Richards, D. P. Schneider, J. Brinkmann, and M. Fukugita. Constraining the Evolution of the Ionizing Background and the Epoch of Reionization with z $\sim$ 6 Quasars. II. A Sample of 19 Quasars. *Astron. J.*, 132(1):117–136, 2006b. doi: 10.1086/504836.

R. Feynman. *The Feynman Lectures on Physics Volume III: Quantum Mechanics*, chapter 12. Addison-Wesley Pub. C., 1963.

S. R. Furlanetto, S. Peng Oh, and F. H. Briggs. Cosmology at Low Frequencies: The 21 cm Transition and the High-Redshift Universe. *Phys. Rep.*, 433(4-6):181–301, 2006. doi: 10.1016/j.physrep.2006.08.002.

B. K. Gehlot, L. V. Koopmans, A. G. de Bruyn, S. Zaroubi, M. A. Brentjens, K. M. Asad, M. Hatef, V. Jelić, M. Mevius, A. R. Offringa, V. N. Pandey, and S. Yatawatta. Wide-Field LOFAR-LBA Power-Spectra Analyses: Impact of Calibration, Polarization Leakage, and Ionosphere. *Mon. Not. R. Astron. Soc.*, 478(2):1484–1501, 2018. doi: 10.1093/mnras/sty1095.

K. M. Gorski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann. HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. *Astrophys. J.*, 622(2):759–771, 2005. doi: 10.1086/427976.

T. L. Grobler, C. D. Nunhokee, O. M. Smirnov, A. J. van Zyl, and A. G. de Bruyn. Calibration Artefacts in Radio Interferometry - I. Ghost Sources in Westerbork Synthesis Radio Telescope Data. *Mon. Not. R. Astron. Soc.*, 439(4):4030–4047, 2014. doi: 10.1093/mnras/stu268.

T. L. Grobler, A. J. Stewart, S. J. Wijnholds, J. S. Kenyon, and O. M. Smirnov. Calibration Artefacts in Radio Interferometry - III. Phase-Only Calibration and Primary Beam Correction. *Mon. Not. R. Astron. Soc.*, 461(3):2975–2992, 2016. doi: 10.1093/mnras/stw1437.

T. L. Grobler, G. Bernardi, J. S. Kenyon, A. R. Parsons, and O. M. Smirnov. Redundant Interferometric Calibration as a Complex Optimization Problem. *Mon. Not. R. Astron. Soc.*, 476(2):2410–2420, 2018. doi: 10.1093/mnras/sty357.

J. P. Hamaker. Understanding Radio Polarimetry. IV. The Full-Coherency Analogue of Scalar Self-Calibration: Self-Alignment, Dynamic Range and Polarimetric Fidelity. *Astron. Astrophys. Suppl. Ser.*, 143(3):515–534, 2000. doi: 10.1051/aas:2000337.

J. P. Hamaker. Understanding Radio Polarimetry. *Astron. Astrophys.*, 456(1):395–404, 2006. doi: 10.1051/0004-6361:20065145.

J. P. Hamaker and J. D. Bregman. Understanding Radio Polarimetry. III. Interpreting the IAU/IEEE Definitions of the Stokes Parameters. *Astron. Astrophys. Suppl. Ser.*, 117(1): 161–165, 1996. doi: 10.1051/aas:1996147.

J. P. Hamaker, J. D. Bregman, and R. J. Sault. Understanding Radio Polarimetry. I. Mathematical Foundations. *Astron. Astrophys. Suppl. Ser.*, 117(1):137–147, 1996. doi: 10.1051/aas:1996146.

C. Haslam, U. Klein, C. Salter, H. Stoffel, W. Wilson, M. Cleary, D. Cooke, and P. Thomasson. A 408 MHz All-Sky Continuum Survey. I - Observations at Southern Declinations and for the North Polar Region. *Astron. Astrophys.*, 100:209–219, 1981.

C. Haslam, C. Salter, H. Stoffel, and W. Wilson. A 408 MHz All-Sky Continuum Survey. II - The Atlas of Contour Maps. *Astron. Astrophys. Suppl. Ser.*, 47:1, 1982.

B. J. Hazelton, M. F. Morales, and I. S. Sullivan. The Fundamental Multi-Baseline Mode-Mixing Foreground in 21 cm Epoch of Reionization Observations. *Astrophys. J.*, 770(2): 156, 2013. doi: 10.1088/0004-637X/770/2/156.

D. W. Hogg. Distance Measures in Cosmology. *arXiv e-prints*, 1999.

N. Hurley-Walker, J. R. Callingham, P. J. Hancock, T. M. Franzen, L. Hindson, A. D. Kapińska, J. Morgan, A. R. Offringa, R. B. Wayth, C. Wu, Q. Zheng, T. Murphy, M. E. Bell, K. S. Dwarakanath, B. For, B. M. Gaensler, M. Johnston-Hollitt, E. Lenc, P. Procopio, L. Staveley-Smith, R. Ekers, J. D. Bowman, F. Briggs, R. J. Cappallo, A. A. Deshpande, L. Greenhill, B. J. Hazelton, D. L. Kaplan, C. J. Lonsdale, S. R. McWhirter, D. A. Mitchell, M. F. Morales, E. Morgan, D. Oberoi, S. M. Ord, T. Prabu, N. Udaya Shankar, K. S. Srivani, R. Subrahmanyan, S. J. Tingay, R. L. Webster, A. Williams, and C. L. Williams. GaLactic and Extragalactic All-Sky Murchison Widefield Array (GLEAM) Survey - I. A Low-Frequency Extragalactic Catalogue. *Mon. Not. R. Astron. Soc.*, 464(1): 1146–1167, 2017. doi: 10.1093/mnras/stw2337.

D. C. Jacobs, B. J. Hazelton, C. M. Trott, J. S. Dillon, B. Pindor, I. S. Sullivan, J. C. Pober, N. Barry, A. P. Beardsley, G. Bernardi, J. D. Bowman, F. Briggs, R. J. Cappallo, P. Carroll, B. E. Corey, A. de Oliveira-Costa, D. Emrich, A. Ewall-Wice, L. Feng, B. M. Gaensler, R. Goeke, L. J. Greenhill, J. N. Hewitt, N. Hurley-Walker, M. Johnston-Hollitt, D. L. Kaplan, J. C. Kasper, H. Kim, E. Kratzenberg, E. Lenc, J. Line, A. Loeb, C. J. Lonsdale, M. J. Lynch, B. McKinley, S. R. McWhirter, D. A. Mitchell, M. F. Morales, E. Morgan, A. R. Neben, N. Thyagarajan, D. Oberoi, A. R. Offringa, S. M. Ord, S. Paul, T. Prabu, P. Procopio, J. Riding, A. E. E. Rogers, A. Roshi, N. U. Shankar, S. K. Sethi, K. S. Srivani, R. Subrahmanyan, M. Tegmark, S. J. Tingay, M. Waterson, R. B. Wayth, R. L. Webster, A. R. Whitney, A. Williams, C. L. Williams, C. Wu, and J. S. B. Wyithe. The Murchison Widefield Array 21 cm Power Spectrum Analysis Methodology. *Astrophys. J.*, 825(2):114, 2016. doi: 10.3847/0004-637X/825/2/114.

P. Jagannathan, S. Bhatnagar, U. Rau, and A. R. Taylor. Direction-Dependent Corrections in Polarimetric Radio Imaging. I. Characterizing the Effects of the Primary Beam on Full-Stokes Imaging. *Astron. J.*, 154(2):56, 2017. doi: 10.3847/1538-3881/aa77f8.

V. Jelić, S. Zaroubi, P. Labropoulos, G. Bernardi, A. G. de Bruyn, and L. V. E. Koopmans. Realistic Simulations of the Galactic Polarized Foreground: Consequences for 21-cm Reionization Detection Experiments. *Mon. Not. R. Astron. Soc.*, 409(4):1647–1659, 2010. doi: 10.1111/j.1365-2966.2010.17407.x.

R. C. Jones. A New Calculus for the Treatment of Optical Systems - I. Description and Discussion of the Calculus. *J. Opt. Soc. Am.*, 31(7):488, 1941. doi: 10.1364/josa.31.000488.

C. H. Jordan, S. Murray, C. M. Trott, R. B. Wayth, D. A. Mitchell, M. Rahimi, B. Pindor, P. Procopio, and J. Morgan. Characterization of the Ionosphere Above the Murchison Radio Observatory Using the Murchison Widefield Array. *Mon. Not. R. Astron. Soc.*, 471 (4):3974–3987, 2017. doi: 10.1093/mnras/stx1797.

R. C. Joseph, C. M. Trott, and R. B. Wayth. The Bias and Uncertainty of Redundant and

Sky-based Calibration Under Realistic Sky and Telescope Conditions. *Astron. J.*, 156(6): 285, 2018. doi: 10.3847/1538-3881/aaec0b.

R. C. Joseph, C. M. Trott, R. B. Wayth, and A. Nasirudin. Calibration and 21-cm Power Spectrum Estimation in the Presence of Antenna Beam Variations. *Mon. Not. R. Astron. Soc.*, 492(2):2017–2028, 2020. doi: 10.1093/mnras/stz3375.

S. Kazemi and S. Yatawatta. Robust Radio Interferometric Calibration Using the t-Distribution. *Mon. Not. R. Astron. Soc.*, 435(1):597–605, 2013. doi: 10.1093/mnras/stt1347.

S. Kazemi, S. Yatawatta, S. Zaroubi, P. Lampropoulos, A. G. de Bruyn, L. V. E. Koopmans, and J. Noordam. Radio Interferometric Calibration Using the SAGE Algorithm. *Mon. Not. R. Astron. Soc.*, 414(2):1656–1666, 2011. doi: 10.1111/j.1365-2966.2011.18506.x.

S. Kazemi, S. Yatawatta, and S. Zaroubi. Clustered Calibration: An Improvement to Radio Interferometric Direction-Dependent Self-Calibration. *Mon. Not. R. Astron. Soc.*, 430(2): 1457–1472, 2013. doi: 10.1093/mnras/stt018.

N. S. Kern, J. S. Dillon, A. R. Parsons, C. L. Carilli, G. Bernardi, Z. Abdurashidova, J. E. Aguirre, P. Alexander, Z. S. Ali, Y. Balfour, A. P. Beardsley, T. S. Billings, J. D. Bowman, R. F. Bradley, P. Bull, J. Burba, S. Carey, C. Cheng, D. R. DeBoer, M. Dexter, E. de Lera Acedo, J. Ely, A. Ewall-Wice, N. Fagnoni, R. Fritz, S. R. Furlanetto, K. Gale-Sides, B. Glendenning, D. Gorthi, B. Greig, J. Grobbelaar, Z. Halday, B. J. Hazelton, J. N. Hewitt, J. Hickish, D. C. Jacobs, A. Julius, J. Kerrigan, P. Kittiwisit, S. A. Kohn, M. Kolopanis, A. Lanman, P. La Plante, T. Lekalake, A. Liu, D. MacMahon, L. Malan, C. Malgas, M. Maree, Z. E. Martinot, E. Matsetela, A. Mesinger, M. Molewa, M. F. Morales, T. Mosiane, S. G. Murray, A. R. Neben, B. Nikolic, C. D. Nunhokee, N. Patra, S. Pieterse, J. C. Pober, N. Razavi-Ghods, J. Ringuette, J. Robnett, K. Rosie, P. Sims, C. Smith, A. Syce, N. Thyagarajan, P. K. G. Williams, and H. Zheng. Absolute Calibration

Strategies for the Hydrogen Epoch of Reionization Array and Their Impact on the 21 cm Power Spectrum. *Astrophys. J.*, 890(2):122, 2020. doi: 10.3847/1538-4357/ab67bc.

S. A. Kohn, J. E. Aguirre, P. La Plante, T. S. Billings, P. M. Chichura, A. F. Fortino, A. S. Igarashi, R. K. Benefo, S. Gallardo, Z. E. Martinot, C. D. Nunhokee, N. S. Kern, P. Bull, A. Liu, P. Alexander, Z. S. Ali, A. P. Beardsley, G. Bernardi, J. D. Bowman, R. F. Bradley, C. L. Carilli, C. Cheng, D. R. DeBoer, E. d. L. Acedo, J. S. Dillon, A. Ewall-Wice, G. Fadana, N. Fagnoni, R. Fritz, S. R. Furlanetto, B. Glendenning, B. Greig, J. Grobbelaar, B. J. Hazelton, J. N. Hewitt, J. Hickish, D. C. Jacobs, A. Julius, M. Kariseb, M. Kolopanis, T. Lekalake, A. Loots, D. MacMahon, L. Malan, C. Malgas, M. Maree, N. Mathison, E. Matsetela, A. Mesinger, M. F. Morales, A. R. Neben, B. Nikolic, A. R. Parsons, N. Patra, S. Pieterse, J. C. Pober, N. Razavi-Ghods, J. Ringuette, J. Robnett, K. Rosie, R. Sell, C. Smith, A. Syce, M. Tegmark, N. Thyagarajan, P. K. G. Williams, and H. Zheng. The HERA-19 Commissioning Array: Direction-Dependent Effects. *Astrophys. J.*, 882(1):58, 2019. doi: 10.3847/1538-4357/ab2f72.

M. Kolopanis, D. C. Jacobs, C. Cheng, A. R. Parsons, S. A. Kohn, J. C. Pober, J. E. Aguirre, Z. S. Ali, G. Bernardi, R. F. Bradley, C. L. Carilli, D. R. DeBoer, M. R. Dexter, J. S. Dillon, J. Kerrigan, P. Klima, A. Liu, D. H. E. MacMahon, D. F. Moore, N. Thyagarajan, C. D. Nunhokee, W. P. Walbrugh, and A. Walker. A Simplified, Lossless Reanalysis of PAPER-64. *Astrophys. J.*, 883(2), 2019. doi: 10.3847/1538-4357/ab3e3a.

W. Lane, W. Cotton, S. van Velzen, T. Clarke, N. Kassim, J. Helmboldt, T. Lazio, and A. Cohen. The Very Large Array Low-frequency Sky Survey Redux (VLSSr). *Mon. Not. R. Astron. Soc.*, 440(1):327–338, 2014. doi: 10.1093/mnras/stu256.

A. Lanman, B. Hazelton, D. Jacobs, M. Kolopanis, J. Pober, J. Aguirre, and N. Thyagarajan. pyuvsim: A Comprehensive Simulation Package for Radio Interferometers in Python. *J. Open Source Softw.*, 4(37):1234, 2019. doi: 10.21105/joss.01234.

M. I. Large, B. Y. Mills, A. G. Little, D. F. Crawford, and J. M. Sutton. The Molonglo

Reference Catalogue of Radio Sources. *Mon. Not. R. Astron. Soc.*, 194:693–704, 1981. doi: 10.1093/mnras/194.3.693.

E. Lenc, B. M. Gaensler, X. H. Sun, E. M. Sadler, A. G. Willis, N. Barry, A. P. Beardsley, M. E. Bell, G. Bernardi, J. D. Bowman, F. Briggs, J. R. Callingham, R. J. Cappallo, P. Carroll, B. E. Corey, A. de Oliveira-Costa, A. A. Deshpande, J. S. Dillon, K. S. Dwarkanath, D. Emrich, A. Ewall-Wice, L. Feng, B.-Q. For, R. Goeke, L. J. Greenhill, P. Hancock, B. J. Hazelton, J. N. Hewitt, L. Hindson, N. Hurley-Walker, M. Johnston-Hollitt, D. C. Jacobs, A. D. Kapińska, D. L. Kaplan, J. C. Kasper, H.-S. Kim, E. Kratzenberg, J. Line, A. Loeb, C. J. Lonsdale, M. J. Lynch, B. McKinley, S. R. McWhirter, D. A. Mitchell, M. F. Morales, E. Morgan, J. Morgan, T. Murphy, A. R. Neben, D. Oberoi, A. R. Offringa, S. M. Ord, S. Paul, B. Pindor, J. C. Pober, T. Prabu, P. Procopio, J. Riding, A. E. E. Rogers, A. Roshi, N. U. Shankar, S. K. Sethi, K. S. Srivani, L. Staveley-Smith, R. Subrahmanyan, I. S. Sullivan, M. Tegmark, N. Thyagarajan, S. J. Tingay, C. Trott, M. Waterson, R. B. Wayth, R. L. Webster, A. R. Whitney, A. Williams, C. L. Williams, C. Wu, J. S. B. Wyithe, and Q. Zheng. Low-Frequency Observations of Linearly Polarized Structures in the Interstellar Medium Near the South Galactic Pole. *Astrophys. J.*, 830 (1):38, 2016. doi: 10.3847/0004-637x/830/1/38.

E. Lenc, C. S. Anderson, N. Barry, J. D. Bowman, I. H. Cairns, J. S. Farnes, B. M. Gaensler, G. Heald, M. Johnston-Hollitt, D. L. Kaplan, C. R. Lynch, P. I. McCauley, D. A. Mitchell, J. Morgan, M. F. Morales, T. Murphy, A. R. Offringa, S. M. Ord, B. Pindor, C. Riseley, E. M. Sadler, C. Sobey, M. Sokolowski, I. S. Sullivan, S. P. O'Sullivan, X. H. Sun, S. E. Tremblay, C. M. Trott, and R. B. Wayth. The Challenges of Low-Frequency Radio Polarimetry: Lessons from the Murchison Widefield Array. *Publ. Astron. Soc. Aust.*, 34: E040, 2017. doi: 10.1017/pasa.2017.36.

E. Lenc, T. Murphy, C. R. Lynch, D. L. Kaplan, and S. N. Zhang. An All-Sky Survey of Circular Polarization at 200 MHz. *Mon. Not. R. Astron. Soc.*, 478(2):2835–2849, 2018. doi: 10.1093/mnras/sty1304.

W. Li, J. C. Pober, B. J. Hazelton, N. Barry, M. F. Morales, I. Sullivan, A. R. Parsons, Z. S. Ali, J. S. Dillon, A. P. Beardsley, J. D. Bowman, F. Briggs, R. Byrne, P. Carroll, B. Crosse, D. Emrich, A. Ewall-Wice, L. Feng, T. M. O. Franzen, J. N. Hewitt, L. Horsley, D. C. Jacobs, M. Johnston-Hollitt, C. Jordan, R. C. Joseph, D. L. Kaplan, D. Kenney, H. Kim, P. Kittiwisit, A. Lanman, J. Line, B. McKinley, D. A. Mitchell, S. Murray, A. Neben, A. R. Offringa, D. Pallot, S. Paul, B. Pindor, P. Procopio, M. Rahimi, J. Riding, S. K. Sethi, N. Udaya Shankar, K. Steele, R. Subrahmanian, M. Tegmark, N. Thyagarajan, S. J. Tingay, C. Trott, M. Walker, R. B. Wayth, R. L. Webster, A. Williams, C. Wu, and S. Wyithe. Comparing Redundant and Sky-model-based Interferometric Calibration: A First Look with Phase II of the MWA. *Astrophys. J.*, 863(2):170, 2018. doi: 10.3847/1538-4357/aad3c3.

W. Li, J. C. Pober, N. Barry, B. J. Hazelton, M. F. Morales, C. M. Trott, A. Lanman, M. Wilensky, I. Sullivan, A. P. Beardsley, T. Booler, J. D. Bowman, R. Byrne, B. Crosse, D. Emrich, T. M. O. Franzen, K. Hasegawa, L. Horsley, M. Johnston-Hollitt, D. C. Jacobs, C. H. Jordan, R. C. Joseph, T. Kaneuji, D. L. Kaplan, D. Kenney, K. Kubota, J. Line, C. Lynch, B. McKinley, D. A. Mitchell, S. Murray, D. Pallot, B. Pindor, M. Rahimi, J. Riding, G. Sleap, K. Steele, K. Takahashi, S. J. Tingay, M. Walker, R. B. Wayth, R. L. Webster, A. Williams, C. Wu, J. S. B. Wyithe, S. Yoshiura, and Q. Zheng. First Season MWA Phase II Epoch of Reionization Power Spectrum Results at Redshift 7. *Astrophys. J.*, 887(2):141, 2019. doi: 10.3847/1538-4357/ab55e4.

J. L. B. Line, D. A. Mitchell, B. Pindor, J. L. Riding, B. McKinley, R. L. Webster, C. M. Trott, N. Hurley-Walker, and A. R. Offringa. Modelling and Peeling Extended Sources with Shapelets: A Fornax A Case Study. *Publ. Astron. Soc. Aust.*, 37:E027, 2020. doi: 10.1017/pasa.2020.18.

A. Liu, M. Tegmark, S. Morrison, A. Lutomirski, and M. Zaldarriaga. Precision Calibration of Radio Interferometers Using Redundant Baselines. *Mon. Not. R. Astron. Soc.*, 408(2): 1029–1050, 2010. doi: 10.1111/j.1365-2966.2010.17174.x.

S. T. Loi, T. Murphy, I. H. Cairns, F. W. Menk, C. L. Waters, P. J. Erickson, C. M. Trott, N. Hurley-Walker, J. Morgan, E. Lenc, A. R. Offringa, M. E. Bell, R. D. Ekers, B. M. Gaensler, C. J. Lonsdale, L. Feng, P. J. Hancock, D. L. Kaplan, G. Bernardi, J. D. Bowman, F. Briggs, R. J. Cappallo, A. A. Deshpande, L. J. Greenhill, B. J. Hazelton, M. Johnston-Hollitt, S. R. McWhirter, D. A. Mitchell, M. F. Morales, E. Morgan, D. Oberoi, S. M. Ord, T. Prabu, N. U. Shankar, K. S. Srivani, R. Subrahmanyan, S. J. Tingay, R. B. Wayth, R. L. Webster, A. Williams, and C. L. Williams. Real-Time Imaging of Density Ducts Between the Plasmasphere and Ionosphere. *Geophys. Res. Lett.*, 42(10): 3707–3714, 2015. doi: 10.1002/2015GL063699.

Y. Mao, M. Tegmark, M. McQuinn, M. Zaldarriaga, and O. Zahn. How Accurately Can 21cm Tomography Constrain Cosmology? *Phys. Rev. D*, 78(2):23529, 2008. doi: 10.1103/PhysRevD.78.023529.

C. F. McKee and E. C. Ostriker. Theory of Star Formation. *Annu. Rev. Astron. Astrophys.*, 45:565–687, 2007. doi: 10.1146/annurev.astro.45.051806.110602.

M. McQuinn, O. Zahn, M. Zaldarriaga, L. Hernquist, and S. R. Furlanetto. Cosmological Parameter Estimation Using 21 cm Radiation from the Epoch of Reionization. *Astrophys. J.*, 653(2):815–834, 2006. doi: 10.1086/505167.

M. McQuinn, A. Lidz, O. Zahn, S. Dutta, L. Hernquist, and M. Zaldarriaga. The Morphology of HII Regions During Reionization. *Mon. Not. R. Astron. Soc.*, 377(3):1043–1063, 2007. doi: 10.1111/j.1365-2966.2007.11489.x.

F. G. Mertens, M. Mevius, L. V. E. Koopmans, A. R. Offringa, G. Mellema, S. Zaroubi, M. A. Brentjens, H. Gan, B. K. Gehlot, V. N. Pandey, A. M. Sardarabadi, H. K. Vedantham, S. Yatawatta, K. M. B. Asad, B. Ciardi, E. Chapman, S. Gazagnes, R. Ghara, A. Ghosh, S. K. Giri, I. T. Iliev, V. Jelić, R. Kooistra, R. Mondal, J. Schaye, and M. B. Silva. Improved Upper Limits on the 21 cm Signal Power Spectrum of Neutral Hydrogen at z ≈ 9.1 from LOFAR. *Mon. Not. R. Astron. Soc.*, 493(2):1662–1685, 2020. doi: 10.1093/mnras/staa327.

M. Mevius. RMextract: Ionospheric Faraday Rotation Calculator. *Astrophys. Source Code Libr.*, 2018.

D. Mitchell, L. Greenhill, R. Wayth, R. Sault, C. Lonsdale, R. Cappallo, M. Morales, and S. Ord. Real-Time Calibration of the Murchison Widefield Array. *IEEE J. Sel. Top. Signal Process.*, 2(5):707–717, 2008. doi: 10.1109/JSTSP.2008.2005327.

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

M. F. Morales and J. Hewitt. Toward Epoch of Reionization Measurements with Wide-Field Radio Observations. *Astrophys. J.*, 615(1):7–18, 2004. doi: 10.1086/424437.

M. F. Morales and M. Matejek. Software Holography: Interferometric Data Analysis for the Challenges of Next Generation Observatories. *Mon. Not. R. Astron. Soc.*, 400(4): 1814–1820, 2009. doi: 10.1111/j.1365-2966.2009.15537.x.

M. F. Morales and J. S. B. Wyithe. Reionization and Cosmology with 21-cm Fluctuations. *Annu. Rev. Astron. Astrophys.*, 48(1):127–171, 2010. doi: 10.1146/annurev-astro-081309-130936.

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

M. F. Morales, A. Beardsley, J. Pober, N. Barry, B. Hazelton, D. Jacobs, and I. Sullivan. Understanding the Diversity of 21 cm Cosmology Analyses. *Mon. Not. R. Astron. Soc.*, 483(2):2207–2216, 2019. doi: 10.1093/mnras/sty2844.

B. J. Mort, F. Dulwich, S. Salvini, K. Z. Adami, and M. E. Jones. OSKAR: Simulating Digital Beamforming for the SKA Aperture Array. In *IEEE Int. Symp. Phased Array Syst. Technol.*, pages 690–694, 2010. doi: 10.1109/ARRAY.2010.5613289.

L. B. Newburgh, G. E. Addison, M. Amiri, K. Bandura, J. R. Bond, L. Connor, J.-F. Cliche, G. Davis, M. Deng, N. Denman, M. Dobbs, M. Fandino, H. Fong, K. Gibbs, A. Gilbert, E. Griffin, M. Halpern, D. Hanna, A. D. Hincks, G. Hinshaw, C. Höfer, P. Klages, T. Landecker, K. Masui, J. M. Parra, U.-L. Pen, J. Peterson, A. Recnik, J. R. Shaw, K. Sigurdson, M. Sitwell, G. Smecher, R. Smegal, K. Vanderlinde, and D. Wiebe. Calibrating CHIME: A New Radio Interferometer to Probe Dark Energy. *Ground-based Airborne Telesc. V*, 9145:91454V, 2014. doi: 10.1117/12.2056962.

L. B. Newburgh, K. Bandura, M. A. Bucher, T.-C. Chang, H. C. Chiang, J. Cliche, R. Davé, M. Dobbs, C. Clarkson, K. M. Ganga, T. Gogo, A. Gumba, N. Gupta, M. Hilton, B. Johnstone, A. Karastergiou, M. Kunz, D. Lokhorst, R. Maartens, S. Macpherson, M. Mdlalose, K. Moodley, L. Ngwenya, J. M. Parra, J. Peterson, O. Recnik, B. Saliwanchik, M. G. Santos, J. L. Sievers, O. Smirnov, P. Stronkhorst, R. Taylor, K. Vanderlinde, G. Van Vuuren, A. Weltman, and A. Witzemann. HIRAX: A Probe of Dark Energy and Radio Transients. *Ground-based Airborne Telesc. VI*, 9906:99065X, 2016. doi: 10.1117/12.2234286.

C. Nunhokee, G. Bernardi, S. Kohn, J. Aguirre, N. Thyagarajan, J. Dillon, G. Foster, T. Grobler, J. Martinot, and A. Parsons. Constraining Polarized Foregrounds for EoR Experiments. II. Polarization Leakage Simulations in the Avoidance Scheme. *Astrophys. J.*, 848(1):47, 2017. doi: 10.3847/1538-4357/aa8b73.

A. Offringa, R. Wayth, N. Hurley-Walker, D. Kaplan, N. Barry, A. Beardsley, M. Bell, G. Bernardi, J. Bowman, F. Briggs, J. Callingham, R. Cappallo, P. Carroll, A. Deshpande, J. Dillon, K. Dwarakanath, A. Ewall-Wice, L. Feng, B.-Q. For, B. Gaensler, L. Greenhill, P. Hancock, B. Hazelton, J. Hewitt, L. Hindson, D. Jacobs, M. Johnston-Hollitt, A. Kapińska, H.-S. Kim, P. Kittiwisit, E. Lenc, J. Line, A. Loeb, C. Lonsdale, B. McKinley, S. McWhirter, D. Mitchell, M. Morales, E. Morgan, J. Morgan, A. Neben, D. Oberoi, S. Ord, S. Paul, B. Pindor, J. Pober, T. Prabu, P. Procopio, J. Riding, N. Udaya Shankar, S. Sethi, K. Srivani, L. Staveley-Smith, R. Subrahmanyan, I. Sullivan, M. Tegmark, N. Thyagarajan, S. Tingay, C. Trott, R. Webster, A. Williams, C. Williams, C. Wu, J. Wyithe, and Q. Zheng. The Low-Frequency Environment of the Murchison Widefield Array: Radio-Frequency Interference Analysis and Mitigation. *Publ. Astron. Soc. Aust.*, 32:E008, 2015. doi: 10.1017/pasa.2015.7.

V. Ollier, M. N. El Korso, R. Boyer, P. Larzabal, and M. Pesavento. Robust Calibration of Radio Interferometers in Non-Gaussian Environment. *IEEE Trans. Signal Process.*, 65 (21):5649–5660, 2017. doi: 10.1109/TSP.2017.2733496.

S. M. Ord, D. A. Mitchell, R. B. Wayth, L. J. Greenhill, G. Bernardi, S. Gleadow, R. G. Edgar, M. A. Clark, G. Allen, W. Arcus, L. Benkevitch, J. D. Bowman, F. H. Briggs, J. D. Bunton, S. Burns, R. J. Cappallo, W. A. Coles, B. E. Corey, L. DeSouza, S. S. Doeleman, M. Derome, A. Deshpande, D. Emrich, R. Goeke, M. R. Gopalakrishna, D. Herne, J. N. Hewitt, P. A. Kamini, D. L. Kaplan, J. C. Kasper, B. B. Kincaid, J. Kocz, E. Kowald, E. Kratzenberg, D. Kumar, C. J. Lonsdale, M. J. Lynch, S. R. McWhirter, S. Madhavi, M. Matejek, M. F. Morales, E. Morgan, D. Oberoi, J. Pathikulangara, T. Prabu, A. E. E. Rogers, A. Roshi, J. E. Salah, A. Schinkel, N. Udaya Shankar, K. S. Srivani, J. Stevens, S. J. Tingay, A. Vaccarella, M. Waterson, R. L. Webster, A. R. Whitney, A. Williams, and C. Williams. Interferometric Imaging with the 32 Element Murchison Wide-Field Array. *Publ. Astron. Soc. Pacific*, 122(897):1353–1366, 2010. doi: 10.1086/657160.

N. Orosz, J. S. Dillon, A. Ewall-Wice, A. R. Parsons, and N. Thyagarajan. Mitigating the Dffects of Antenna-to-Antenna Variation on Redundant-Baseline Calibration for 21 cm Cosmology. *Mon. Not. R. Astron. Soc.*, 487(1):537–549, 2019. doi: 10.1093/mnras/stz1287.

G. Paciga, T. C. Chang, Y. Gupta, R. Nityanada, J. Odegova, U. L. Pen, J. B. Peterson,

J. Roy, and K. Sigurdson. The GMRT Epoch of Reionization Experiment: A New Upper Limit on the Neutral Hydrogen Power Spectrum at z ≈ 8.6. *Mon. Not. R. Astron. Soc.*, 413(2), 2011. doi: 10.1111/j.1365-2966.2011.18208.x.

G. Paciga, J. G. Albert, K. Bandura, T.-C. Chang, Y. Gupta, C. Hirata, J. Odegova, U.-L. Pen, J. B. Peterson, J. Roy, J. R. Shaw, K. Sigurdson, and T. Voytek. A Simulation-Calibrated Limit on the HI Power Spectrum from the GMRT Epoch of Reionization Experiment. *Mon. Not. R. Astron. Soc.*, 433(1):639–647, 2013. doi: 10.1093/mnras/stt753.

A. R. Parsons, D. C. Backer, G. S. Foster, M. C. Wright, R. F. Bradley, N. E. Gugliucci, C. R. Parashare, E. E. Benoit, J. E. Aguirre, D. C. Jacobs, C. L. Carilli, D. Herne, M. J. Lynch, J. R. Manley, and D. J. Werthimer. The Precision Array for Probing the Epoch of Re-Ionization: Eight Station Results. *Astron. J.*, 139(4):1468–1480, 2010. doi: 10.1088/0004-6256/139/4/1468.

A. R. Parsons, J. C. Pober, J. E. Aguirre, C. L. Carilli, D. C. Jacobs, and D. F. Moore. A Per-Baseline, Delay-Spectrum Technique for Accessing the 21cm Cosmic Reionization Signature. *Astrophys. J.*, 756(2):165, 2012. doi: 10.1088/0004-637X/756/2/165.

A. H. Patil, S. Yatawatta, S. Zaroubi, L. V. E. Koopmans, A. G. de Bruyn, V. Jelić, B. Ciardi, I. T. Iliev, M. Mevius, V. N. Pandey, and B. K. Gehlot. Systematic Biases in Low-Frequency Radio Interferometric Data Due to Calibration: The LOFAR-EoR Case. *Mon. Not. R. Astron. Soc.*, 463(4):4317–4330, 2016. doi: 10.1093/mnras/stw2277.

U. L. Pen, T. C. Chang, C. M. Hirata, J. B. Peterson, J. Roy, Y. Gupta, J. Odegova, and K. Sigurdson. The GMRT EoR Dxperiment: Limits on Polarized Sky Brightness at 150 MHz. *Mon. Not. R. Astron. Soc.*, 399(1):181–194, 2009. doi: 10.1111/j.1365-2966.2009.14980.x.

A. Penzias and R. Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *Astrophys. J.*, 142:419–421, 1965. doi: 10.1086/148307.

J. C. Pober, A. Liu, J. S. Dillon, J. E. Aguirre, J. D. Bowman, R. F. Bradley, C. L. Carilli, D. R. DeBoer, J. N. Hewitt, D. C. Jacobs, M. McQuinn, M. F. Morales, A. R. Parsons, M. Tegmark, and D. J. Werthimer. What Next-generation 21 cm Power Spectrum Measurements can Teach us About the Epoch of Reionization. *Astrophys. J.*, 782(2):66, 2014. doi: 10.1088/0004-637X/782/2/66.

J. C. Pober, B. J. Hazelton, A. P. Beardsley, N. A. Barry, Z. E. Martinot, I. S. Sullivan, M. F. Morales, M. E. Bell, G. Bernardi, N. D. R. Bhat, J. D. Bowman, F. Briggs, R. J. Cappallo, P. Carroll, B. E. Corey, A. De Oliveira-Costa, A. A. Deshpande, J. S. Dillon, D. Emrich, A. M. Ewall-Wice, L. Feng, R. Goeke, L. J. Greenhill, J. N. Hewitt, L. Hindson, N. Hurley-Walker, D. C. Jacobs, M. Johnston-Hollitt, D. L. Kaplan, J. C. Kasper, H.-S. Kim, P. Kittiwisit, E. Kratzenberg, N. Kudryavtseva, E. Lenc, J. Line, A. Loeb, C. J. Lonsdale, M. J. Lynch, B. McKinley, S. R. McWhirter, D. A. Mitchell, E. Morgan, A. R. Neben, D. Oberoi, A. R. Offringa, S. M. Ord, S. Paul, B. Pindor, T. Prabu, P. Procopio, J. Riding, A. E. E. Rogers, A. Roshi, S. K. Sethi, N. U. Shankar, K. S. Srivani, R. Subrahmanyan, M. Tegmark, N. Thyagarajan, S. J. Tingay, C. M. Trott, M. Waterson, R. B. Wayth, R. L. Webster, A. R. Whitney, A. Williams, C. L. Williams, and J. S. B. Wyithe. The Importance of Wide-Field Foreground Removal for 21 cm Cosmology: A Demonstration with Early MWA Epoch of Reionization Observations. *Astrophys. J.*, 819 (1):8, 2016. doi: 10.3847/0004-637x/819/1/8.

J. R. Pritchard and A. Loeb. Evolution of the 21cm Signal Throughout Cosmic History. *Phys. Rev. D*, 78(10):103511, 2008. doi: 10.1103/PhysRevD.78.103511.

M. Remazeilles, C. Dickinson, A. J. Banday, M. A. Bigot-Sazy, and T. Ghosh. An Improved Source-Subtracted and Destriped 408-MHz All-Sky Map. *Mon. Not. R. Astron. Soc.*, 451 (4):4311–4327, 2015. doi: 10.1093/mnras/stv1274.

S. Salvini and S. J. Wijnholds. StEFCal - An Alternating Direction Implicit Method for Fast

Full Polarization Array Calibration. In *2014 31th URSI Gen. Assem. Sci. Symp. URSI GASS 2014*, 2014. doi: 10.1109/URSIGASS.2014.6930038.

M. G. Santos and A. Cooray. Cosmological and Astrophysical Parameter Measurements with 21-cm Anisotropies During the Era of Reionization. *Phys. Rev. D*, 74(8):83517, 2006. doi: 10.1103/PhysRevD.74.083517.

R. J. Sault, J. P. Hamaker, and J. D. Bregman. Understanding Radio Polarimetry. II. Instrumental Calibration of an Interferometer Array. *Astron. Astrophys. Suppl. Ser.*, 117 (1):149–159, 1996. doi: 10.1051/aas:1996100.

J. L. Sievers. Calibration of Quasi-Redundant Interferometers. *arXiv e-prints*, 2017.

U. M. Sob, H. L. Bester, O. M. Smirnov, J. S. Kenyon, and T. L. Grobler. Radio Interferometric Calibration Using a Complex Student's t-Distribution and Wirtinger Derivatives. *Mon. Not. R. Astron. Soc.*, 491(1):1026–1042, 2020. doi: 10.1093/mnras/stz3037.

M. Spinelli, G. Bernardi, and M. G. Santos. On the Contamination of the Global 21-cm Signal from Polarized Foregrounds. *Mon. Not. R. Astron. Soc.*, 489(3):4007–4015, 2019.

G. G. Stokes. On the Composition and Resolution of Streams of Polarized Light from Different Sources. *Trans. Cambridge Philos. Soc.*, 9:399, 1851.

I. A. Strukov and D. P. Skulachev. Deep-Space Measurements of the Microwave Background Anisotropy - First Results of the Relikt Experiment. *Sov. Astron. Lett.*, 10:1–4, 1984.

I. S. Sullivan, M. F. Morales, B. J. Hazelton, W. Arcus, D. Barnes, G. Bernardi, F. H. Briggs, J. D. Bowman, J. D. Bunton, R. J. Cappallo, B. E. Corey, A. Deshpande, L. Desouza, D. Emrich, B. M. Gaensler, R. Goeke, L. J. Greenhill, D. Herne, J. N. Hewitt, M. Johnston-Hollitt, D. L. Kaplan, J. C. Kasper, B. B. Kincaid, R. Koenig, E. Kratzenberg, C. J. Lonsdale, M. J. Lynch, S. R. McWhirter, D. A. Mitchell, E. Morgan, D. Oberoi, S. M. Ord, J. Pathikulangara, T. Prabu, R. A. Remillard, A. E. Rogers, A. Roshi, J. E. Salah,

R. J. Sault, N. U. Shankar, K. S. Srivani, J. Stevens, R. Subrahmanyan, S. J. Tingay, R. B. Wayth, M. Waterson, R. L. Webster, A. R. Whitney, A. Williams, C. L. Williams, and J. S. Wyithe. Fast Holographic Deconvolution: A New Technique for Precision Radio Interferometry. *Astrophys. J.*, 759(1):17, 2012. doi: 10.1088/0004-637X/759/1/17.

A. Sutinjo, J. O'Sullivan, E. Lenc, R. B. Wayth, S. Padhi, P. Hall, and S. J. Tingay. Understanding instrumental Stokes leakage in Murchison Widefield Array polarimetry. *Radio Sci.*, 50(1):52–65, 2015. doi: 10.1002/2014RS005517.

C. Tasse, B. Hugo, M. Mirmont, O. Smirnov, M. Atemkeng, L. Bester, M. J. Hardcastle, R. Lakhoo, S. Perkins, and T. Shimwell. Faceting for Direction-Dependent Spectral Deconvolution. *Astron. Astrophys.*, 611:A87, 2018. doi: 10.1051/0004-6361/201731474.

M. Tegmark, A. J. S. Hamilton, M. A. Strauss, M. S. Vogeley, and A. S. Szalay. Measuring the Galaxy Power Spectrum with Future Redshift Surveys. *Astrophys. J.*, 499(2):555–576, 1998. doi: 10.1086/305663.

N. Thyagarajan, N. Udaya Shankar, R. Subrahmanyan, W. Arcus, G. Bernardi, J. D. Bowman, F. Briggs, J. D. Bunton, R. J. Cappallo, B. E. Corey, L. Desouza, D. Emrich, B. M. Gaensler, R. F. Goeke, L. J. Greenhill, B. J. Hazelton, D. Herne, J. N. Hewitt, M. Johnston-Hollitt, D. L. Kaplan, J. C. Kasper, B. B. Kincaid, R. Koenig, E. Kratzenberg, C. J. Lonsdale, M. J. Lynch, S. R. McWhirter, D. A. Mitchell, M. F. Morales, E. H. Morgan, D. Oberoi, S. M. Ord, J. Pathikulangara, R. A. Remillard, A. E. Rogers, D. A. Roshi, J. E. Salah, R. J. Sault, K. S. Srivani, J. B. Stevens, P. Thiagaraj, S. J. Tingay, R. B. Wayth, M. Waterson, R. L. Webster, A. R. Whitney, A. J. Williams, C. L. Williams, and J. S. B. Wyithe. A Study of Fundamental Limitations to Statistical Detection of Redshifted HI from the Epoch of Reionization. *Astrophys. J.*, 776(1):6, 2013. doi: 10.1088/0004-637X/776/1/6.

N. Thyagarajan, D. D. C. Jacobs, J. D. Bowman, N. Barry, A. P. Beardsley, G. Bernardi, F. Briggs, R. J. Cappallo, P. Carroll, B. E. Corey, A. D. Oliveira-Costa, J. S. Dillon,

D. Emrich, A. Ewall-Wice, L. Feng, R. Goeke, L. J. Greenhill, B. J. Hazelton, J. N. Hewitt, N. Hurley-Walker, M. Johnston-Hollitt, D. L. Kaplan, J. C. Kasper, H.-S. H. S. Kim, P. Kittiwisit, E. Kratzenberg, E. Lenc, J. Line, A. Loeb, C. J. Lonsdale, M. J. Lynch, B. McKinley, S. R. McWhirter, D. A. Mitchell, M. F. Morales, E. Morgan, A. R. Neben, D. Oberoi, A. R. Offringa, S. M. Ord, S. Paul, B. Pindor, J. C. Pober, T. Prabu, P. Procopio, J. Riding, A. E. Rogers, A. Roshi, N. U. Shankar, S. S. K. Sethi, K. S. Srivani, R. Subrahmanyan, I. S. Sullivan, M. Tegmark, S. J. Tingay, C. M. Trott, M. Waterson, R. B. Wayth, R. L. Webster, A. R. Whitney, A. Williams, C. L. Williams, C. Wu, and J. S. Wyithe. Foregrounds in Wide-Field Redshifted 21 cm Power Spectra. *Astrophys. J.*, 804(1):14, 2015. doi: 10.1088/0004-637X/804/1/14.

S. J. Tingay, R. Goeke, J. D. Bowman, D. Emrich, S. M. Ord, D. A. Mitchell, M. F. Morales, T. Booler, B. Crosse, R. B. Wayth, C. J. Lonsdale, S. Tremblay, D. Pallot, T. Colegate, A. Wicenec, N. Kudryavtseva, W. Arcus, D. Barnes, G. Bernardi, F. Briggs, S. Burns, J. D. Bunton, R. J. Cappallo, B. E. Corey, A. Deshpande, L. Desouza, B. M. Gaensler, L. J. Greenhill, P. J. Hall, B. J. Hazelton, D. Herne, J. N. Hewitt, M. Johnston-Hollitt, D. L. Kaplan, J. C. Kasper, B. B. Kincaid, R. Koenig, E. Kratzenberg, M. J. Lynch, B. McKinley, S. R. McWhirter, E. Morgan, D. Oberoi, J. Pathikulangara, T. Prabu, R. A. Remillard, A. E. Rogers, A. Roshi, J. E. Salah, R. J. Sault, N. Udaya-Shankar, F. Schlagenhaufer, K. S. Srivani, J. Stevens, R. Subrahmanyan, M. Waterson, R. L. Webster, A. R. Whitney, A. Williams, C. L. Williams, and J. S. Wyithe. The Murchison Widefield Array: The Square Kilometre Array Precursor at Low Radio Frequencies. *Publ. Astron. Soc. Aust.*, 30(1):E007, 2013. doi: 10.1017/pasa.2012.007.

C. M. Trott and R. B. Wayth. Spectral Calibration Requirements of Radio Interferometers for Epoch of Reionisation Science with the SKA. *Publ. Astron. Soc. Aust.*, 33:E019, 2016. doi: 10.1017/pasa.2016.18.

C. M. Trott and R. B. Wayth. Building Models for Extended Radio Sources: Implications for Epoch of Reionisation Science. *Publ. Astron. Soc. Aust.*, 2017. doi: 10.1017/pasa.2017.57.

C. M. Trott, R. B. Wayth, and S. J. Tingay. The Impact of Point-Source Subtraction Residuals on 21cm Epoch of Reionization Estimation. *Astrophys. J.*, 757(1):101, 2012. doi: 10.1088/0004-637X/757/1/101.

C. M. Trott, B. Pindor, P. Procopio, R. B. Wayth, D. A. Mitchell, B. McKinley, S. J. Tingay, N. Barry, A. P. Beardsley, G. Bernardi, J. D. Bowman, F. Briggs, R. J. Cappallo, P. Carroll, A. de Oliveira-Costa, J. S. Dillon, A. Ewall-Wice, L. Feng, L. J. Greenhill, B. J. Hazelton, J. N. Hewitt, N. Hurley-Walker, M. Johnston-Hollitt, D. C. Jacobs, D. L. Kaplan, H. S. Kim, E. Lenc, J. Line, A. Loeb, C. J. Lonsdale, M. F. Morales, E. Morgan, A. R. Neben, N. Thyagarajan, D. Oberoi, A. R. Offringa, S. M. Ord, S. Paul, J. C. Pober, T. Prabu, J. Riding, N. U. Shankar, S. K. Sethi, K. S. Srivani, R. Subrahmanyan, I. S. Sullivan, M. Tegmark, R. L. Webster, A. Williams, C. L. Williams, C. Wu, and J. S. B. Wyithe. CHIPS: The Cosmological HI Power Spectrum Estimator. *Astrophys. J.*, 818(2): 139, 2016. doi: 10.3847/0004-637x/818/2/139.

C. M. Trott, C. H. Jordan, S. Midgley, N. Barry, B. Greig, B. Pindor, J. H. Cook, G. Sleap, S. J. Tingay, D. Ung, P. Hancock, A. Williams, J. Bowman, R. Byrne, A. Chokshi, B. J. Hazelton, K. Hasegawa, D. Jacobs, R. C. Joseph, W. Li, J. L. B. Line, C. Lynch, B. McKinley, D. A. Mitchell, M. F. Morales, M. Ouchi, J. C. Pober, M. Rahimi, K. Takahashi, R. B. Wayth, R. L. Webster, M. Wilensky, J. S. B. Wyithe, S. Yoshiura, Z. Zhang, and Q. Zheng. Deep Multiredshift Limits on Epoch of Reionization 21 cm Power Spectra from Four Seasons of Murchison Widefield Array Observations. *Mon. Not. R. Astron. Soc.*, 493(4), 2020. doi: 10.1093/mnras/staa414.

M. Valdés, B. Ciardi, A. Ferrara, M. Johnston-Hollitt, and H. Röttgering. Radio Views of Cosmic Reionization. *Mon. Not. R. Astron. Soc.*, 369(1):L66–L70, 2006. doi: 10.1111/j.1745-3933.2006.00179.x.

R. J. van Weeren, W. L. Williams, M. J. Hardcastle, T. W. Shimwell, D. A. Rafferty, J. Sabater, G. Heald, S. S. Sridhar, T. J. Dijkema, G. Brunetti, M. Brüggen, F. Andrade-

Santos, G. A. Ogrean, H. J. A. Röttgering, W. A. Dawson, W. R. Forman, F. de Gasperin, C. Jones, G. K. Miley, L. Rudnick, C. L. Sarazin, A. Bonafede, P. N. Best, L. Bîrzan, R. Cassano, K. T. Chyży, J. H. Croston, T. Ensslin, C. Ferrari, M. Hoeft, C. Horellou, M. J. Jarvis, R. P. Kraft, M. Mevius, H. T. Intema, S. S. Murray, E. Orrú, R. Pizzo, A. Simionescu, A. Stroe, S. van der Tol, and G. J. White. Lofar Facet Calibration. *Astrophys. J. Suppl. Ser.*, 223(1):2, 2016. doi: 10.3847/0067-0049/223/1/2.

K. Vanderlinde, K. Bandura, L. Belostotski, R. Bond, P. Boyle, J. Brown, H. C. Chiang, M. Dobbs, B. Gaensler, G. Hinshaw, V. Kaspi, T. Landecker, A. Liu, K. Masui, J. Mena-Parra, C. Ng, U. Pen, M. Rupen, J. Sievers, K. Smith, K. Spekkens, I. Stairs, and N. Turok. LRP 2020 Whitepaper: The Canadian Hydrogen Observatory and Radio-Transient Detector (CHORD). *Can. Long Range Plan Astron. Astrophys. White Pap.*, 2019. doi: 10.5281/zenodo.3765414.

H. Vedantham, N. Udaya Shankar, and R. Subrahmanyan. Imaging the Epoch of Reionization: Limitations from Roreground Confusion and Imaging Algorithms. *Astrophys. J.*, 745 (2):176, 2012. doi: 10.1088/0004-637X/745/2/176.

R. B. Wayth, S. J. Tingay, C. M. Trott, D. Emrich, M. Johnston-Hollitt, B. McKinley, B. M. Gaensler, A. P. Beardsley, T. Booler, B. Crosse, T. M. Franzen, L. Horsley, D. L. Kaplan, D. Kenney, M. F. Morales, D. Pallot, G. Sleap, K. Steele, M. Walker, A. Williams, C. Wu, I. H. Cairns, M. D. Filipovic, S. Johnston, T. Murphy, P. Quinn, L. Staveley-Smith, R. Webster, and J. S. Wyithe. The Phase II Murchison Widefield Array: Design Overview. *Publ. Astron. Soc. Aust.*, 35:E033, 2018. doi: 10.1017/pasa.2018.37.

M. H. Wieringa. An Investigation of the Telescope Based Calibration Methods 'Redundancy' and 'Self-Cal'. *Exp. Astron.*, 2(4):203–225, 1992. doi: 10.1007/BF00420576.

S. J. Wijnholds, T. L. Grobler, and O. M. Smirnov. Calibration Srtefacts in Radio Interferometry - II. Ghost Patterns for Irregular Arrays. *Mon. Not. R. Astron. Soc.*, 457(3): 2331–2354, 2016. doi: 10.1093/mnras/stw118.

M. J. Wilensky, M. F. Morales, B. J. Hazelton, N. Barry, R. Byrne, and S. Roy. Absolving the SSINS of Precision Interferometric Radio Data: A New Technique for Mitigating Faint Radio Frequency Interference. *Publ. Astron. Soc. Pacific*, 131(1005):114507, 2019. doi: 10.1088/1538-3873/ab3cad.

M. J. Wilensky, N. Barry, M. F. Morales, B. J. Hazelton, and R. Byrne. Quantifying Excess Power from Radio Frequency Interference in Epoch of Reionization Measurements. *Mon. Not. R. Astron. Soc.*, 498(1):265–275, 2020. doi: 10.1093/mnras/staa2442.

M. Wolleben, T. L. Landecker, E. Carretti, J. M. Dickey, A. Fletcher, N. M. McClure-Griffiths, D. McConnell, A. J. M. Thomson, A. S. Hill, B. M. Gaensler, J.-L. Han, M. Haverkorn, J. P. Leahy, W. Reich, and A. R. Taylor. The Global Magneto-Ionic Medium Survey: Polarimetry of the Southern Sky from 300 to 480 MHz. *Astron. J.*, 158 (1):44, 2019. doi: 10.3847/1538-3881/ab22b0.

S. Yatawatta. Distributed Radio Interferometric Calibration. *Mon. Not. R. Astron. Soc.*, 449(4):4506–4514, 2015. doi: 10.1093/mnras/stv596.

M. Zaldarriaga, S. R. Furlanetto, and L. Hernquist. 21 Centimeter Fluctuations from Cosmic Gas at High Redshifts. *Astrophys. J.*, 608(2):622–635, 2004. doi: 10.1086/386327.

Z. Zhang, J. C. Pober, W. Li, B. J. Hazelton, M. F. Morales, C. M. Trott, C. H. Jordan, R. C. Joseph, A. Beardsley, N. Barry, R. Byrne, S. J. Tingay, A. Chokshi, K. Hasegawa, D. C. Jacobs, A. Lanman, J. L. Line, C. Lynch, B. McKinley, D. A. Mitchell, S. Murray, B. Pindor, M. Rahimi, K. Takahashi, R. B. Wayth, R. L. Webster, M. Wilensky, S. Yoshiura, and Q. Zheng. The Impact of Tandem Redundant/Sky-Based Calibration in MWA Phase II Data Analysis. *Publ. Astron. Soc. Aust.*, 37:E045, 2020. doi: 10.1017/pasa.2020.37.

H. Zheng, M. Tegmark, V. Buza, J. S. Dillon, H. Gharibyan, J. Hickish, E. Kunz, A. Liu, J. Losh, A. Lutomirski, S. Morrison, S. Narayanan, A. Perko, D. Rosner, N. Sanchez, K. Schutz, S. M. Tribiano, M. Valdez, H. Yang, K. Zarb Adami, I. Zelko, K. Zheng,

R. P. Armstrong, R. F. Bradley, M. R. Dexter, A. Ewall-Wice, A. Magro, M. Matejek, E. Morgan, A. R. Neben, Q. Pan, R. F. Penna, C. M. Peterson, M. Su, J. Villasenor, C. L. Williams, and Y. Zhu. MITEoR: A Scalable Interferometer for Precision 21 cm Cosmology. *Mon. Not. R. Astron. Soc.*, 445(2):1084–1103, 2014. doi: 10.1093/mnras/stu1773.

H. Zheng, M. Tegmark, J. S. Dillon, D. A. Kim, A. Liu, A. R. Neben, J. Jonas, P. Reich, and W. Reich. An Improved Model of Diffuse Galactic Radio Emission from 10 MHz to 5 THz. *Mon. Not. R. Astron. Soc.*, 464(3):3486–3497, 2017. doi: 10.1093/mnras/stw2525.

Appendix A

COMPARING ABSOLUTE CALIBRATION APPROACHES

In §3.2 we calculate the absolute calibration parameters directly from the gains calculated with sky-based calibration according to Equations 3.1 and 3.3. In this section we replicate the results of §3.2 with another, more practical approach to absolute calibration that we call "optimal absolute calibration." We show that, while these two absolute calibration approaches are not exactly equal, they deliver equivalent results.

Optimal absolute calibration calculates the absolute calibration parameters by minimizing the following cost function:

$$L(A, \Delta, \Delta_x, \Delta_y) = \sum_a \sum_{\{j,k\} \text{ in } a} \left| v_{jk} - A^2 e^{-i(\Delta + \Delta_x x_a + \Delta_y y_a)} \hat{h}_j \hat{h}_k^* m_a \right|^2 . \tag{A.1}$$

Here a indexes redundant baseline sets and j and k index antennas. $\boldsymbol{m}$ denotes the model visibilities and $\hat{\boldsymbol{h}}$ denotes the relatively calibrated gains. x_a and y_a are the coordinates of the baseline vectors in redundant set a.

To isolate the effect of sky model error on redundant calibration, we simulate the case that relative calibration is perfect. In simulation, the true gains are equal to 1, so perfect relative calibration means $\hat{\boldsymbol{h}} = 1$. We also require that the overall phase of the gains $\Delta = 0$. We then get that

$$L(A, \Delta_x, \Delta_y) = \sum_a \sum_{\{j,k\} \text{ in } a} \left| v_{jk} - A^2 e^{-i(\Delta_x x_a + \Delta_y y_a)} m_a \right|^2 . \tag{A.2}$$

We can calculate the absolute calibration parameters analytically by minimizing this expression. Taking derivatives with respect to A gives

$$\frac{\partial L(A, \Delta_x, \Delta_y)}{\partial A} = 4A \sum_a \sum_{\{j,k\} \text{ in } a} \left[A^2 |m_a|^2 - \mathrm{Re}\left(v_{jk}^* e^{-i(\Delta_x x_a + \Delta_y y_a)} m_a \right) \right] . \tag{A.3}$$

Imposing the condition $\frac{\partial L(A,\Delta_x,\Delta_y)}{\partial A}\Big|_{A=\hat{A},\Delta_x=\hat{\Delta_x},\Delta_y=\hat{\Delta_y}} = 0$, we get

$$\hat{A}^2 = \frac{\sum_a \sum_{\{j,k\}\text{ in }a} \operatorname{Re}\left(v_{jk}^* e^{-i(\hat{\Delta_x} x_a + \hat{\Delta_y} y_a)} m_a\right)}{\sum_a \sum_{\{j,k\}\text{ in }a} |m_a|^2}. \tag{A.4}$$

Now taking derivatives of Equation A.2 with respect to the gain phase gradient terms, we get

$$\begin{aligned}\frac{\partial L(A,\Delta_x,\Delta_y)}{\partial \Delta_x} = 2A^2 \sum_a x_a \sum_{\{j,k\}\text{ in }a} &\Big[\sin(\Delta_x x_a + \Delta_y y_a)\operatorname{Re}(v_{jk}^* m_a) \\ &- \cos(\Delta_x x_a + \Delta_y y_a)\operatorname{Im}(v_{jk}^* m_a)\Big]\end{aligned} \tag{A.5}$$

and

$$\begin{aligned}\frac{\partial L(A,\Delta_x,\Delta_y)}{\partial \Delta_y} = 2A^2 \sum_a y_a \sum_{\{j,k\}\text{ in }a} &\Big[\sin(\Delta_x x_a + \Delta_y y_a)\operatorname{Re}(v_{jk}^* m_a) \\ &- \cos(\Delta_x x_a + \Delta_y y_a)\operatorname{Im}(v_{jk}^* m_a)\Big].\end{aligned} \tag{A.6}$$

Using the extrema conditions $\frac{\partial L(A,\Delta_x,\Delta_y)}{\partial \Delta_x}\Big|_{A=\hat{A},\Delta_x=\hat{\Delta_x},\Delta_y=\hat{\Delta_y}} = 0$ and $\frac{\partial L(A,\Delta_x,\Delta_y)}{\partial \Delta_y}\Big|_{A=\hat{A},\Delta_x=\hat{\Delta_x},\Delta_y=\hat{\Delta_y}} = 0$, we get

$$\sum_a x_a \sin(\hat{\Delta}_x x_a + \hat{\Delta}_y y_a) \sum_{\{j,k\}\text{ in }a} \operatorname{Re}(v_{jk}^* m_a) = \sum_a x_a \cos(\hat{\Delta}_x x_a + \hat{\Delta}_y y_a) \sum_{\{j,k\}\text{ in }a} \operatorname{Im}(v_{jk}^* m_a) \tag{A.7}$$

and

$$\sum_a y_a \sin(\hat{\Delta}_x x_a + \hat{\Delta}_y y_a) \sum_{\{j,k\}\text{ in }a} \operatorname{Re}(v_{jk}^* m_a) = \sum_a y_a \cos(\hat{\Delta}_x x_a + \hat{\Delta}_y y_a) \sum_{\{j,k\}\text{ in }a} \operatorname{Im}(v_{jk}^* m_a). \tag{A.8}$$

If we assume $\hat{\Delta}_x$ and $\hat{\Delta}_y$ are small, we can Taylor expand these expressions to get

$$\hat{\Delta}_x \sum_a x_a^2 \sum_{\{j,k\}\text{ in }a} \operatorname{Re}(v_{jk}^* m_a) + \hat{\Delta}_y \sum_a x_a y_a \sum_{\{j,k\}\text{ in }a} \operatorname{Re}(v_{jk}^* m_a) = \sum_a x_a \sum_{\{j,k\}\text{ in }a} \operatorname{Im}(v_{jk}^* m_a) \tag{A.9}$$

and

$$\hat{\Delta}_x \sum_a x_a y_a \sum_{\{j,k\}\text{ in }a} \operatorname{Re}(v_{jk}^* m_a) + \hat{\Delta}_y \sum_a y_a^2 \sum_{\{j,k\}\text{ in }a} \operatorname{Re}(v_{jk}^* m_a) = \sum_a y_a \sum_{\{j,k\}\text{ in }a} \operatorname{Im}(v_{jk}^* m_a). \tag{A.10}$$

This is simply a system of linear equations that we can solve to calculate $\hat{\Delta}_x$ and $\hat{\Delta}_y$.

We now replicate the absolute calibration of the simulation presented in §3.2 using Equations A.4, A.9, and A.10. The results are plotted as dashed blue lines in Figures A.1-A.3. The solid blue lines represent the solutions presented in Chapter 3 and plotted in Figures 3.2 and 3.3. We find that the two absolute calibration methods are effectively equivalent, and that discrepancies between the two methods are at least two orders-of-magnitude less than the errors in the absolute calibration parameters from sky model incompleteness.

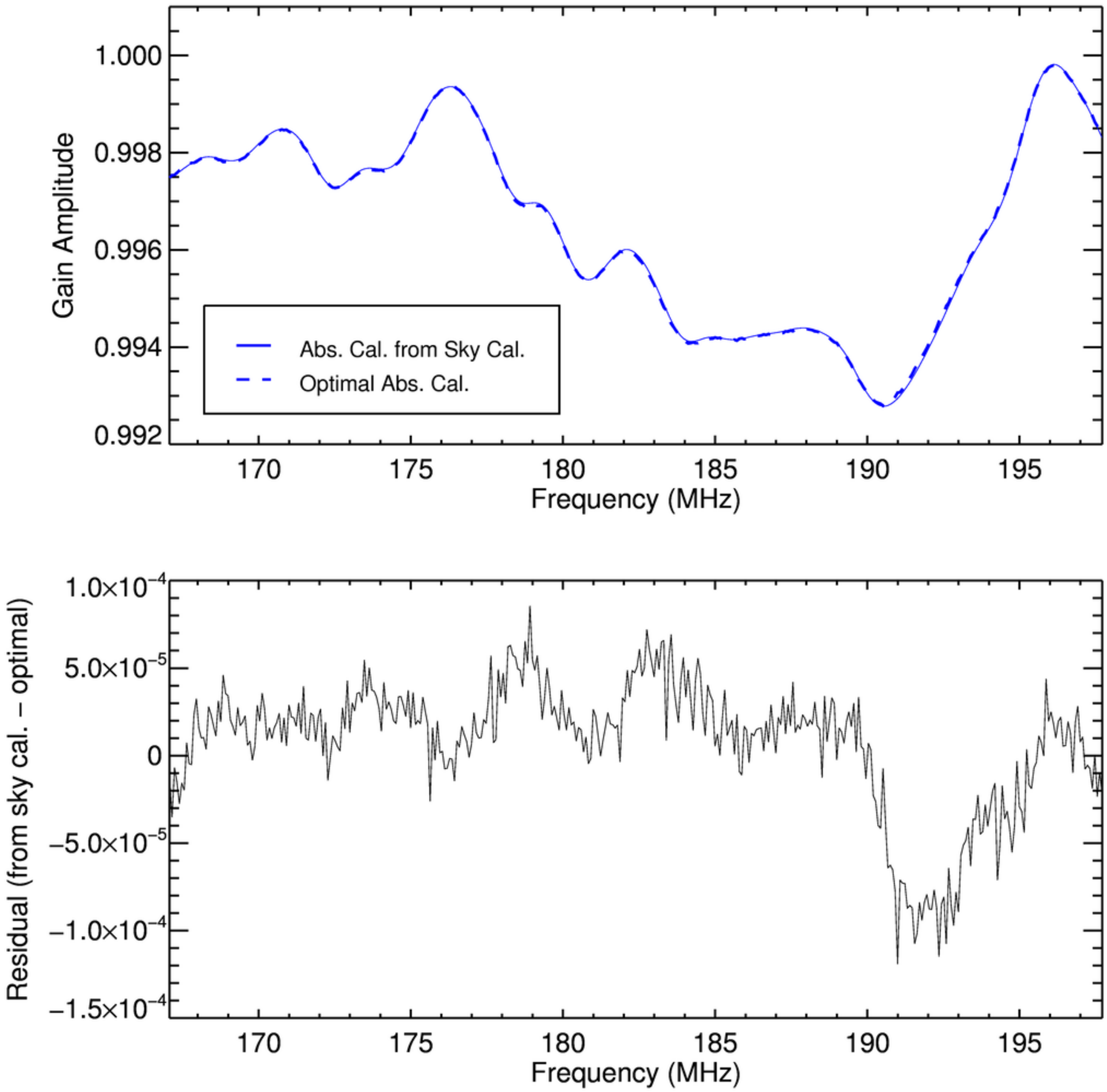


Figure A.1: Comparison of the overall gain amplitude calculated with two different absolute calibration approaches. The calculation is based on the simulation presented in §3.2, exploring the impact of missing sources in the calibration sky model. The top panel overplots the two solutions and the bottom panel plots their discrepancy. We see that the discrepancy between the two calculations' results is very small compared to the error from missing sky model sources.

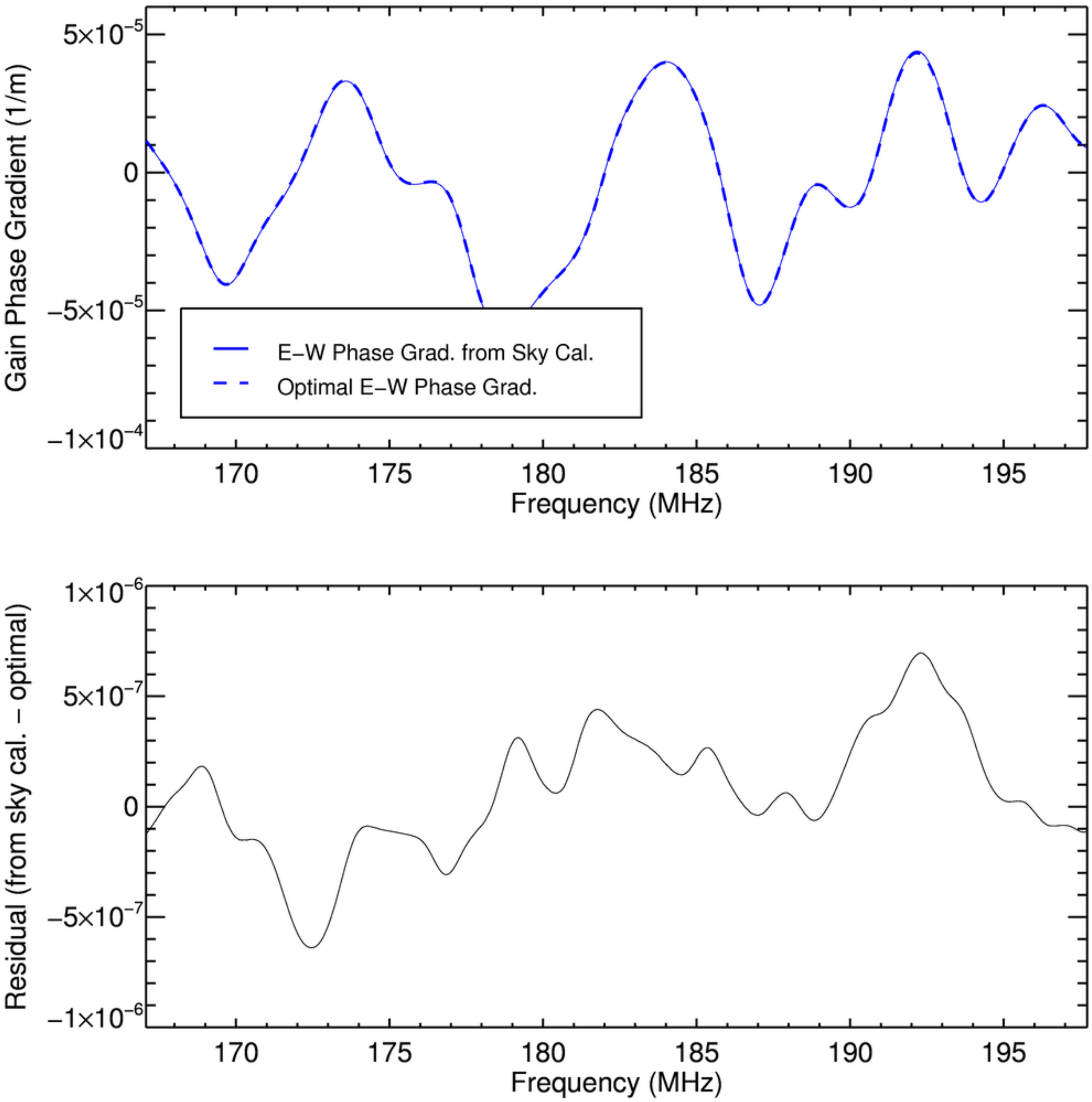


Figure A.2: Comparison of Δ_x, the east-west gain phase gradient, calculated with two different absolute calibration approaches. The top panel overplots the two solutions and the bottom panel plots their discrepancy, which is much smaller than the phase gradient's characteristic error.

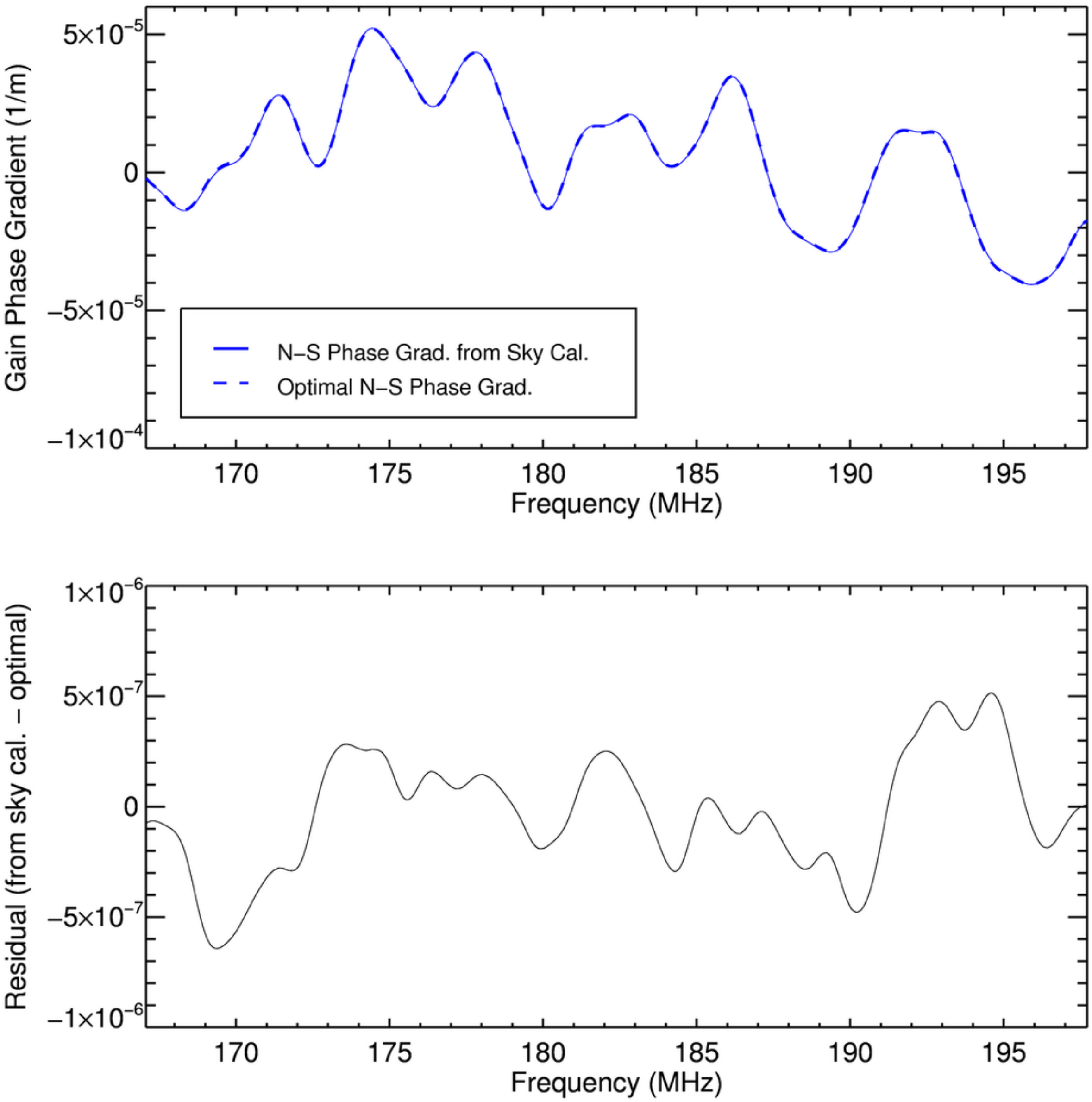


Figure A.3: Comparison of Δ_y, the north-south gain phase gradient, calculated with two different absolute calibration approaches. The top panel overplots the two solutions and the bottom panel plots their discrepancy. Once again, two methods give roughly equivalent results.

Appendix B

WORKING WITH SINGULAR COVARIANCE MATRICES

In this section we explain how to use Singular Value Decomposition (SVD) to resolve the problem of non-invertible covariance matrices by remapping the calibration parameters to a well-defined basis. This approach is fully general and can be applied to any singular covariance matrix, not just those corresponding to traditional redundant calibration as described in §4.2.2. It remaps an initial calibration parameter basis $\boldsymbol{u}_{\mathrm{orig}}$, where the associated covariance matrix $\mathbf{C}_{\mathrm{M\,orig}} = \mathrm{cov}[\boldsymbol{u}_{\mathrm{orig}}, \boldsymbol{u}_{\mathrm{orig}}^{\dagger}]$ is singular, to a new basis $\boldsymbol{u}$ with an invertible covariance matrix $\mathbf{C}_{\mathrm{M}} = \mathrm{cov}[\boldsymbol{u}, \boldsymbol{u}^{\dagger}]$.

All covariance matrices have eigenvalues $\lambda_j \geq 0$ and eigenvectors that span the space. We can therefore transform $\boldsymbol{u}_{\mathrm{orig}}$ into the basis of eigenvectors of $\mathbf{C}_{\mathrm{M\,orig}}$. We define a new vector $\boldsymbol{u}_{\mathrm{diag}}$ as $\boldsymbol{u}_{\mathrm{orig}}$ expressed in the orthogonal eigenbasis of $\mathbf{C}_{\mathrm{M\,orig}}$:

$$\boldsymbol{u}_{\mathrm{diag}} = \mathbf{E}\boldsymbol{u}_{\mathrm{orig}} \tag{B.1}$$

where

$$\mathbf{E} = \begin{bmatrix} \boldsymbol{e}_1^{\dagger} \\ \boldsymbol{e}_2^{\dagger} \\ \vdots \\ \boldsymbol{e}_{N_{\mathrm{orig}}}^{\dagger} \end{bmatrix} \tag{B.2}$$

and $\boldsymbol{e}_j$ are the orthogonal eigenvectors of $\mathbf{C}_{\mathrm{M\,orig}}$. Here N_{orig} is the length of $\boldsymbol{u}_{\mathrm{orig}}$. The associated covariance matrix is

$$\mathbf{C}_{\mathrm{M\,diag}} = \mathrm{cov}[\boldsymbol{u}_{\mathrm{diag}}, \boldsymbol{u}_{\mathrm{diag}}^{\dagger}] = \mathbf{E}\mathbf{C}_{\mathrm{M\,orig}}\mathbf{E}^{-1} = \mathrm{diag}\left(\begin{bmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_{N_{\mathrm{orig}}} \end{bmatrix}\right), \tag{B.3}$$

where λ_j are the eigenvalues of $\mathbf{C}_{\mathrm{M\,orig}}$.

If $\mathbf{C}_{\mathrm{M\,orig}}$ is singular than it will have $\lambda_j = 0$ for some j. This has physical meaning. If $\lambda_j = 0$ then $\mathrm{var}[u_{\mathrm{diag}\,j}] = 0$. It follows that those elements of $\boldsymbol{u}_{\mathrm{diag}}$ must take their expectation values: $u_{\mathrm{diag}\,j} = \langle u_{\mathrm{diag}\,j} \rangle$. We assume that $\langle \boldsymbol{u}_{\mathrm{orig}} \rangle = \boldsymbol{m}_{\mathrm{orig}}$, so likewise $\langle \boldsymbol{u}_{\mathrm{diag}} \rangle = \boldsymbol{m}_{\mathrm{diag}}$ where

$$\boldsymbol{m}_{\mathrm{diag}} = \mathbf{E}\boldsymbol{m}_{\mathrm{orig}}. \tag{B.4}$$

Therefore $u_{\mathrm{diag}\,j} = m_{\mathrm{diag}\,j}$ for all j where $\lambda_j = 0$. Those calibration parameters simply take on their modeled values; they are no longer fit during the calibration process.

We can now restrict the basis of the calibration parameters to only those that can vary in calibration. The number of independent calibration parameters is N_{red}, which is equal to the rank of $\mathbf{C}_{\mathrm{M\,orig}}$. The new basis of calibration parameters is given by

$$\boldsymbol{u} = \mathbf{E}'\boldsymbol{u}_{\mathrm{orig}}. \tag{B.5}$$

Here $\mathbf{E}'$ is

$$\mathbf{E}' = \begin{bmatrix} \boldsymbol{e}_1^\dagger \\ \boldsymbol{e}_2^\dagger \\ \vdots \\ \boldsymbol{e}_{N_{\mathrm{red}}}^\dagger \end{bmatrix}, \tag{B.6}$$

where we use only $\boldsymbol{e}_j$ for $\lambda_j \neq 0$. $\mathbf{C}_{\mathrm{M}} = \mathrm{cov}[\boldsymbol{u}, \boldsymbol{u}^\dagger]$ is now diagonal and invertible.

To develop physical intuition for this remapping we can consider the case of a perfectly redundant array as described in §4.2.2, where $\boldsymbol{u}_{\mathrm{orig}}$ correspond to the full set of visibilities. Here $\mathrm{cov}[u_{\mathrm{orig}\,j}, u^*_{\mathrm{orig}\,k}] = \mathrm{var}[u_{\mathrm{orig}\,j}]$ if baselines j and k belong to the same redundant set and $\mathrm{cov}[u_{\mathrm{orig}\,j}, u^*_{\mathrm{orig}\,k}] = 0$ otherwise. In this case, eigenvectors $\boldsymbol{e}_j$ for $\lambda_j \neq 0$ represent the average visibilities from redundant baseline sets. Eigenvectors $\boldsymbol{e}_j$ for $\lambda_j = 0$ represent visibility differences within redundant baseline sets.

$\mathbf{E}'$ is a rectangular matrix with linearly independent rows. This means that a right pseudoinverse $\mathbf{E}'^{\mathrm{p}}$ exists such that $\mathbf{E}'\mathbf{E}'^{\mathrm{p}} = \mathbb{1}$. However, it does not follow that $\mathbf{E}'^{\mathrm{p}}\mathbf{E}' = \mathbb{1}$, so inverting Equation B.5 is not straightforward. Rather, we get the somewhat unwieldy expression

$$\boldsymbol{u}_{\mathrm{orig}} = \mathbf{E}'^{\mathrm{p}}\boldsymbol{u} + \sum_{\substack{j \text{ for all} \\ \lambda_j = 0}} m_{\mathrm{diag}\,j} \frac{\boldsymbol{e}_j}{|\boldsymbol{e}_j|^2}. \tag{B.7}$$

This expression simplifies when we assume that $m_{\mathrm{diag}\,j} = 0$ when $\lambda_j = 0$ — in other words, that the model values $\boldsymbol{m}_{\mathrm{orig}}$ are orthogonal to the null space of $\mathbf{C}_{\mathrm{M\,orig}}$. We then get that

$$\boldsymbol{u}_{\mathrm{orig}} = \mathbf{E}'^{\mathrm{p}}\boldsymbol{u}. \tag{B.8}$$

This is a physically-motivated assumption. For the case of a perfectly redundant array, requiring that $\boldsymbol{m}_{\mathrm{orig}}$ is orthogonal to the null space means requiring that model visibilities from different baselines within redundant baseline sets are equal. Components of $\boldsymbol{m}_{\mathrm{orig}}$ in the null space indicate a disagreement between the instrument models used to produce $\boldsymbol{m}_{\mathrm{orig}}$ and $\mathbf{C}_{\mathrm{M\,orig}}$.

Appendix C

EXPLORING GAIN AMPLITUDE BIASES

In Chapter 5 we demonstrate that, in simulation, traditional sky-based and redundant calibration methods fit gain solutions $\hat{\boldsymbol{g}}$ that are on average biased low in their amplitudes. This is also consistent with simulation results presented in Chapter 3. Unified calibration solutions do not exhibit this effect. In this appendix we explore the gain amplitude bias using analytic techniques. We show that suppression of the average gain amplitude in redundant and sky-based calibration stems from decoherence between the measured visibilities and the model visibilities in the presence of sky model error. We compare analytic predictions to results from the simulations in Chapter 5, demonstrating agreement.

C.0.1 Sky-Based Calibration

From Chapter 2 and Equation 2.28, sky-based calibration gains are given by

$$\hat{g}_c = \frac{\sum_{b \neq c} \hat{g}_b m_{bc} v_{bc}^*}{\sum_{b \neq c} |\hat{g}_b|^2 |m_{bc}|^2}. \tag{C.1}$$

For a redundant array, we can rewrite this as

$$\hat{g}_c = \frac{\sum_\alpha \sum_{b \text{ for } \{b,c\} \in \alpha} \hat{g}_b m_\alpha v_{bc}^*}{\sum_\alpha \sum_{b \text{ for } \{b,c\} \in \alpha} |\hat{g}_b|^2 |m_\alpha|^2}, \tag{C.2}$$

where α indexes redundant baseline groups and $\sum_{b \text{ for } \{b,c\} \in \alpha}$ denotes the sum of antennas b such that baseline $\{b, c\}$ belongs to redundant group α.

We define a the variable A to be equal to the average gain amplitude,

$$A = \frac{1}{N_{\text{ant}}} \sum_c |g_c|, \tag{C.3}$$

and we parameterize the gains as $\boldsymbol{g} = A\boldsymbol{h}$. Plugging this into equation C.2, we get

$$\hat{A}^2 = \frac{1}{N_{\text{ant}}} \sum_c \frac{\left|\sum_\alpha \sum_{b \text{ for } \{b,c\} \in \alpha} v_{bc}^* \hat{h}_b m_\alpha\right|}{\sum_\alpha \sum_{b \text{ for } \{b,c\} \in \alpha} |\hat{h}_b m_\alpha|^2} \tag{C.4}$$

where N_{ant} is the number of antennas.

At this point we make a series of simplifying approximations. We allow that the true gains are equal to 1, and we can therefore make a first-order approximation of $\hat{A}$ by setting $\hat{\boldsymbol{h}} = 1$:

$$\hat{A}^2 \approx \frac{1}{N_{\text{ant}}} \sum_c \frac{\left|\sum_\alpha \sum_{b \text{ for } \{b,c\} \in \alpha} v_{bc}^* m_\alpha\right|}{\sum_\alpha |m_\alpha|^2}. \tag{C.5}$$

This can be rewritten as

$$\hat{A}^2 \approx \frac{\left|\langle \boldsymbol{v}^* \boldsymbol{m} \rangle\right|}{\langle |\boldsymbol{m}|^2 \rangle}, \tag{C.6}$$

where $\langle\rangle$ denotes the average across redundant baseline sets.

We describe the model visibilities $\boldsymbol{m}$ as a ‘true’ visibility $\boldsymbol{u}_{\text{true}}$ with error $\boldsymbol{n}_{\text{M}}$:

$$m_\alpha = u_{\text{true}\,\alpha} + n_{\text{M}\,\alpha}. \tag{C.7}$$

Likewise, we represent the data as

$$v_{bc} = u_{\text{true}\,\alpha} + n_{\text{T}\,bc} \tag{C.8}$$

where $\boldsymbol{n}_{\text{T}}$ represents the thermal noise and baseline $\{b, c\}$ belongs to redundant set α. If we assume that $\boldsymbol{n}_{\text{M}}$, $\boldsymbol{n}_{\text{T}}$, and $\boldsymbol{u}_{\text{true}}$ are uncorrelated with one another, then the average quantities

$$\langle \boldsymbol{v}^* \boldsymbol{m} \rangle = \langle |\boldsymbol{u}_{\text{true}}|^2 \rangle \tag{C.9}$$

and

$$\langle |\boldsymbol{m}|^2 \rangle = \langle |\boldsymbol{u}_{\text{true}}|^2 \rangle + \langle |\boldsymbol{n}_{\text{M}}|^2 \rangle. \tag{C.10}$$

We thereby estimate that

$$\hat{A} \approx \sqrt{\frac{\langle |\boldsymbol{u}_{\text{true}}|^2 \rangle}{\langle |\boldsymbol{u}_{\text{true}}|^2 \rangle + \langle |\boldsymbol{n}_{\text{M}}|^2 \rangle}} \leq 1. \tag{C.11}$$

Decoherence of the sky model error and thermal noise means that we expect suppression of average amplitude of the fit gains.

We can compare this approximation to the results of the simulation in §5.2. Across all trials in that simulation we calculate that $\left\langle |\boldsymbol{u}_{\rm true}|^2 \right\rangle = 38.45$ Jy^2 and $\left\langle |\boldsymbol{n}_{\rm M}|^2 \right\rangle = 0.31$ Jy^2. Plugging these into Equation C.11, we estimate $A \approx 0.9960$. The average gain amplitude from the sky-based calibration simulation is 0.9964.

C.0.2 Redundant Calibration

The overall amplitude of the gains is fit in absolute calibration. From Appendix A and Equation A.4 we get that

$$\hat{A}^2 = \frac{\sum_\alpha \sum_{\{b,c\}\in\alpha} \mathrm{Re}\left(v_{bc}^* e^{-i(\hat{\Delta}+\hat{\Delta}_x x_\alpha + \hat{\Delta}_y y_\alpha)} \hat{h}_b \hat{h}_c^* m_\alpha \right)}{\sum_\alpha \sum_{\{b,c\}\in\alpha} |\hat{h}_b|^2 |\hat{h}_c|^2 |m_\alpha|^2}. \tag{C.12}$$

where $\hat{\boldsymbol{h}}$ have been calculated from relative calibration.

As in §C.0.1, we take the first-order approximation that $\hat{\boldsymbol{h}} = 1$. We also let $\hat{\Delta} = 0$, $\hat{\Delta}_x = 0$, and $\hat{\Delta}_y = 0$. This gives

$$\hat{A}^2 \approx \frac{\sum_\alpha \sum_{\{b,c\}\in\alpha} \mathrm{Re}(v_{bc}^* m_\alpha)}{\sum_\alpha \sum_{\{b,c\}\in\alpha} |m_\alpha|^2} \approx \frac{\langle \mathrm{Re}(\boldsymbol{v}^* \boldsymbol{m}) \rangle}{\langle |\boldsymbol{m}|^2 \rangle}. \tag{C.13}$$

Using the decompositions in Equations C.7 and C.8, and once again assuming no correlation between $\boldsymbol{n}_{\rm M}$, $\boldsymbol{n}_{\rm T}$, and $\boldsymbol{u}_{\rm true}$, we get a result identical to Equation C.11:

$$\hat{A} \approx \sqrt{\frac{\langle |\boldsymbol{u}_{\rm true}|^2 \rangle}{\langle |\boldsymbol{u}_{\rm true}|^2 \rangle + \langle |\boldsymbol{n}_{\rm M}|^2 \rangle}} \leq 1. \tag{C.14}$$

We expect redundant calibration to return approximately the same average gain amplitude as sky-based calibration. Indeed, in the simulation presented in §5.2, the average gain amplitude is 0.9964 for each sky-based and redundant calibration.

C.0.3 Unified Calibration

From Equation 5.5, the unified calibration framework explored in simulation in this chapter consists of minimizing the negative log-likelihood

$$L(\boldsymbol{g},\boldsymbol{u}) = \frac{1}{\sigma_{\mathrm{T}}^2}\sum_{\alpha}\sum_{\{b,c\}\in\alpha}|v_{bc} - g_b g_c^* u_\alpha|^2 + \frac{1}{\sigma_{\mathrm{M}}^2}\sum_{\alpha}\sum_{\gamma}(C_{\mathrm{R}}^{-1})_{\alpha\gamma}(u_\alpha - m_\alpha)(u_\gamma^* - m_\gamma^*), \quad (\mathrm{C.15})$$

where α and γ index redundant baseline sets and $(C_{\mathrm{R}}^{-1})_{\alpha\gamma}$ represents elements of the inverse correlation matrix $\mathbf{C}_{\mathrm{R}}^{-1}$. For the purposes of this discussion we will approximate $\mathbf{C}_{\mathrm{R}} \approx \mathbb{1}$, neglecting the covariance between redundant baseline sets. The negative log-likelihood then simplifies to

$$L(\boldsymbol{g},\boldsymbol{u}) = \frac{1}{\sigma_{\mathrm{T}}^2}\sum_{\alpha}\sum_{\{b,c\}\in\alpha}|v_{bc} - g_b g_c^* u_\alpha|^2 + \frac{1}{\sigma_{\mathrm{M}}^2}\sum_{\alpha}|u_\alpha - m_\alpha|^2. \quad (\mathrm{C.16})$$

Solving the extremum conditions $\left.\frac{\partial L(\boldsymbol{g},\boldsymbol{u})}{\partial\,\mathrm{Re}(g_c)}\right|_{\boldsymbol{g}=\hat{\boldsymbol{g}},\boldsymbol{u}=\hat{\boldsymbol{u}}} = 0$ and $\left.\frac{\partial L(\boldsymbol{g},\boldsymbol{u})}{\partial\,\mathrm{Im}(g_c)}\right|_{\boldsymbol{g}=\hat{\boldsymbol{g}},\boldsymbol{u}=\hat{\boldsymbol{u}}} = 0$ gives

$$\hat{g}_c = \frac{\sum_\alpha \sum_{b \text{ for } \{b,c\}\in\alpha} v_{bc}^* \hat{g}_b \hat{u}_\alpha}{\sum_\alpha \sum_{b \text{ for } \{b,c\}\in\alpha} |\hat{g}_b \hat{u}_\alpha|^2}, \quad (\mathrm{C.17})$$

from which it follows that the average gain amplitude is

$$\hat{A} = \frac{1}{N_{\mathrm{ant}}}\sum_c \frac{\left|\sum_\alpha \sum_{b \text{ for } \{b,c\}\in\alpha} v_{bc}^* \hat{g}_b \hat{u}_\alpha\right|}{\sum_\alpha \sum_{b \text{ for } \{b,c\}\in\alpha} |\hat{g}_b \hat{u}_\alpha|^2}. \quad (\mathrm{C.18})$$

We have a further extremum conditions from the fit visibilities $\boldsymbol{u}$: $\left.\frac{\partial L(\boldsymbol{g},\boldsymbol{u})}{\partial\,\mathrm{Re}(u_\alpha)}\right|_{\boldsymbol{g}=\hat{\boldsymbol{g}},\boldsymbol{u}=\hat{\boldsymbol{u}}} = 0$ and $\left.\frac{\partial L(\boldsymbol{g},\boldsymbol{u})}{\partial\,\mathrm{Im}(u_\alpha)}\right|_{\boldsymbol{g}=\hat{\boldsymbol{g}},\boldsymbol{u}=\hat{\boldsymbol{u}}} = 0$. Solving gives

$$\hat{u}_\alpha = \frac{\frac{1}{\sigma_{\mathrm{T}}^2}\sum_{\{b,c\}\in\alpha} v_{bc}\hat{g}_b^*\hat{g}_c + \frac{1}{\sigma_{\mathrm{M}}^2}\boldsymbol{m}_\alpha}{\frac{1}{\sigma_{\mathrm{T}}^2}\sum_{\{b,c\}\in\alpha}|\hat{g}_b\hat{g}_c^*|^2 + \frac{1}{\sigma_{\mathrm{M}}^2}}. \quad (\mathrm{C.19})$$

If we again let $\boldsymbol{g} = A\boldsymbol{h}$ and take the first-order approximation $\hat{\boldsymbol{h}} \approx 1$, we get

$$\hat{A}^2 \approx \frac{1}{N_{\mathrm{ant}}}\sum_c \frac{\left|\sum_\alpha \sum_{b \text{ for } \{b,c\}\in\alpha} v_{bc}^* \hat{u}_\alpha\right|}{\sum_\alpha |\hat{u}_\alpha|^2} \approx \frac{|\langle \boldsymbol{v}^* \hat{\boldsymbol{u}}\rangle|}{\langle|\hat{\boldsymbol{u}}|^2\rangle} \quad (\mathrm{C.20})$$

and

$$\hat{u}_\alpha \approx \frac{\frac{1}{\sigma_\mathrm{T}^2}\hat{A}^2\sum_{\{b,c\}\in\alpha} v_{bc} + \frac{1}{\sigma_\mathrm{M}^2} m_\alpha}{\frac{1}{\sigma_\mathrm{T}^2}\hat{A}^4 N_\alpha + \frac{1}{\sigma_\mathrm{M}^2}}, \tag{C.21}$$

where N_α is the number of baselines in redundant set α.

We can explore the opposing limits that $\sigma_\mathrm{M} \to 0$ and $\sigma_\mathrm{M} \to \infty$. For $\sigma_\mathrm{M} \to 0$, Equation C.21 gives $\hat{u}_\alpha \approx m_\alpha$. We then get that

$$\hat{A}^2 \approx \frac{\left|\langle \boldsymbol{v}^* \boldsymbol{m}\rangle\right|}{\langle |\boldsymbol{m}|^2\rangle}, \tag{C.22}$$

which is identical to the solution for sky-based calibration in §C.0.1. In the opposite limit that $\sigma_\mathrm{M} \to \infty$, Equation C.21 becomes

$$\hat{u}_\alpha \approx \frac{\sum_{\{b,c\}\in\alpha} v_{bc}}{\hat{A}^2 N_\alpha}. \tag{C.23}$$

Following Equation C.8 and again assume no correlation between $\boldsymbol{n}_\mathrm{T}$ and $\boldsymbol{u}_\mathrm{true}$, we find that

$$\left|\langle \boldsymbol{v}^* \hat{\boldsymbol{u}}\rangle\right| \approx \frac{1}{\hat{A}^2}\left(\langle |\boldsymbol{u}_\mathrm{true}|^2\rangle + \langle |\boldsymbol{n}_\mathrm{T}|^2\rangle\right) \tag{C.24}$$

and

$$\langle |\hat{\boldsymbol{u}}|^2\rangle \approx \frac{1}{\hat{A}^4}\left(\langle |\boldsymbol{u}_\mathrm{true}|^2\rangle + \langle |\boldsymbol{n}_\mathrm{T}|^2\rangle\right). \tag{C.25}$$

We therefore get that $\hat{A}^2 \approx \hat{A}^2$, reflecting that the overall amplitudes of the gains is degenerate in the relative calibration step of redundant calibration.

In unified calibration σ_M is finite and nonzero. In order to evaluate the average gain amplitude, we make the approximation that N_α is constant across redundant baseline sets. This is a non-physical assumption—for the simulation in Chapter 5 the number of baselines in each redundant set varies from 1 to 30—but is useful for simplifying Equation C.20 and deriving a rough approximation of the average gain amplitude in unified calibration. Under this approximation, we evaluate

$$\left|\langle \boldsymbol{v}^* \hat{\boldsymbol{u}}\rangle\right| \approx \frac{\left|\frac{1}{\sigma_\mathrm{T}^2}\hat{A}^2\langle N\rangle\langle |\boldsymbol{v}|^2\rangle + \frac{1}{\sigma_\mathrm{M}^2}\langle \boldsymbol{v}^*\boldsymbol{m}\rangle\right|}{\frac{1}{\sigma_\mathrm{T}^2}\hat{A}^4\langle N\rangle + \frac{1}{\sigma_\mathrm{M}^2}} \approx \frac{\left(\frac{1}{\sigma_\mathrm{T}^2}\hat{A}^2\langle N\rangle + \frac{1}{\sigma_\mathrm{M}^2}\right)\langle |\boldsymbol{u}_\mathrm{true}|^2\rangle + \frac{1}{\sigma_\mathrm{T}^2}\hat{A}^2\langle N\rangle\langle |\boldsymbol{n}_\mathrm{T}|^2\rangle}{\frac{1}{\sigma_\mathrm{T}^2}\hat{A}^4\langle N\rangle + \frac{1}{\sigma_\mathrm{M}^2}} \tag{C.26}$$

and

$$\begin{aligned}\langle|\hat{\boldsymbol{u}}|^2\rangle &\approx \frac{\frac{1}{\sigma_{\mathrm{T}}^4}\hat{A}^4\langle N\rangle^2\langle|\boldsymbol{v}|^2\rangle + \frac{1}{\sigma_{\mathrm{M}}^4}\langle|\boldsymbol{m}|^2\rangle + \frac{2}{\sigma_{\mathrm{T}}^2\sigma_{\mathrm{M}}^2}\hat{A}^2\langle N\rangle\langle\mathrm{Re}(\boldsymbol{v}^*\boldsymbol{m})\rangle}{\left(\frac{1}{\sigma_{\mathrm{T}}^2}\hat{A}^4\langle N\rangle + \frac{1}{\sigma_{\mathrm{M}}^2}\right)^2} \\ &\approx \frac{\left(\frac{1}{\sigma_{\mathrm{T}}^2}\hat{A}^2\langle N\rangle + \frac{1}{\sigma_{\mathrm{M}}^2}\right)^2\langle|\boldsymbol{u}_{\mathrm{true}}|^2\rangle + \frac{1}{\sigma_{\mathrm{T}}^4}\hat{A}^4\langle N\rangle^2\langle|\boldsymbol{n}_{\mathrm{T}}|^2\rangle + \frac{1}{\sigma_{\mathrm{M}}^4}\langle|\boldsymbol{n}_{\mathrm{M}}|^2\rangle}{\left(\frac{1}{\sigma_{\mathrm{T}}^2}\hat{A}^4\langle N\rangle + \frac{1}{\sigma_{\mathrm{M}}^2}\right)^2},\end{aligned} \tag{C.27}$$

where $\langle N\rangle$ is the average number of baselines in a redundant set. Combining these, we get

$$\hat{A}^2 \approx \left(\frac{1}{\sigma_{\mathrm{T}}^2}\hat{A}^4\langle N\rangle + \frac{1}{\sigma_{\mathrm{M}}^2}\right)\frac{\left(\frac{1}{\sigma_{\mathrm{T}}^2}\hat{A}^2\langle N\rangle + \frac{1}{\sigma_{\mathrm{M}}^2}\right)\langle|\boldsymbol{u}_{\mathrm{true}}|^2\rangle + \frac{1}{\sigma_{\mathrm{T}}^2}\hat{A}^2\langle N\rangle\langle|\boldsymbol{n}_{\mathrm{T}}|^2\rangle}{\left(\frac{1}{\sigma_{\mathrm{T}}^2}\hat{A}^2\langle N\rangle + \frac{1}{\sigma_{\mathrm{M}}^2}\right)^2\langle|\boldsymbol{u}_{\mathrm{true}}|^2\rangle + \frac{1}{\sigma_{\mathrm{T}}^4}\hat{A}^4\langle N\rangle^2\langle|\boldsymbol{n}_{\mathrm{T}}|^2\rangle + \frac{1}{\sigma_{\mathrm{M}}^4}\langle|\boldsymbol{n}_{\mathrm{M}}|^2\rangle}, \tag{C.28}$$

which can be solved numerically to evaluate $\hat{A}$.

As noted in §C.0.1, in the simulation presented in §5.2 we have $\left\langle|\boldsymbol{u}_{\mathrm{true}}|^2\right\rangle = 38.45$ Jy^2 and $\left\langle|\boldsymbol{n}_{\mathrm{M}}|^2\right\rangle = 0.31$ Jy^2. We also have $\left\langle|\boldsymbol{n}_{\mathrm{T}}|^2\right\rangle = 0.08$ Jy^2, $\sigma_{\mathrm{M}} = 0.4$ Jy, $\sigma_{\mathrm{T}} = 0.2$ Jy, and $\langle N\rangle = 10.5$. Plugging in these values and numerically solving, we get $\hat{A} \approx 1.0009$. In the simulation presented in §5.2, unified calibration returns an average gain amplitude of 1.0005.

Appendix D

THE PARALLACTIC ANGLE

The parallactic angle defines the transformation between RA/Dec and azimuth/zenith angle coordinate systems. Equation 6.4 presents the coordinate transformation. In this appendix, we present a derivation of the parallactic angle.

We consider a location on the sky $\boldsymbol{\theta}$, and we define zenith as $\boldsymbol{\theta}_{\text{zen}}$. In the RA/Dec coordinate system

$$\boldsymbol{\theta}_{\text{zen}} = \alpha_{\text{zen}}\,\boldsymbol{e}_\alpha + \delta_{\text{zen}}\,\boldsymbol{e}_\delta \tag{D.1}$$

and

$$\boldsymbol{\theta} = (\alpha + \alpha_{\text{zen}})\,\boldsymbol{e}_\alpha + \delta\,\boldsymbol{e}_\delta, \tag{D.2}$$

where $\boldsymbol{e}_\alpha$ is a unit vector in the RA direction and $\boldsymbol{e}_\delta$ is a unit vector in the Dec direction.

We now define a Cartesian coordinate system with the origin at the center of the sphere and the x-axis aligned with an RA of α_{zen}. We define the radius of the sphere to be 1. The RA/Dec coordinates are defined such that at zenith RA increases in the $+y$ direction and Dec increases in the $+z$ direction. The zenith angle increases away from zenith and azimuthal angle increases clockwise from the $+y$ direction at zenith.

We can now represent $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_{\text{zen}}$ in our Cartesian coordinate system:

$$\boldsymbol{\theta} = \cos\alpha\cos\delta\,\boldsymbol{e}_x + \sin\alpha\cos\delta\,\boldsymbol{e}_y + \sin\delta\,\boldsymbol{e}_z \tag{D.3}$$

and

$$\boldsymbol{\theta}_{\text{zen}} = \cos\delta_{\text{zen}}\,\boldsymbol{e}_x + \sin\delta_{\text{zen}}\,\boldsymbol{e}_z, \tag{D.4}$$

where $\boldsymbol{e}_x$, $\boldsymbol{e}_y$, and $\boldsymbol{e}_z$ are unit vectors in the x, y, and z directions.

At $\boldsymbol{\theta}$, the azimuth/zenith angle coordinate system is aligned with the plane that includes $\boldsymbol{\theta}$, $\boldsymbol{\theta}_{\text{zen}}$, and the center of the sphere. A vector normal to this plane can be calculated by

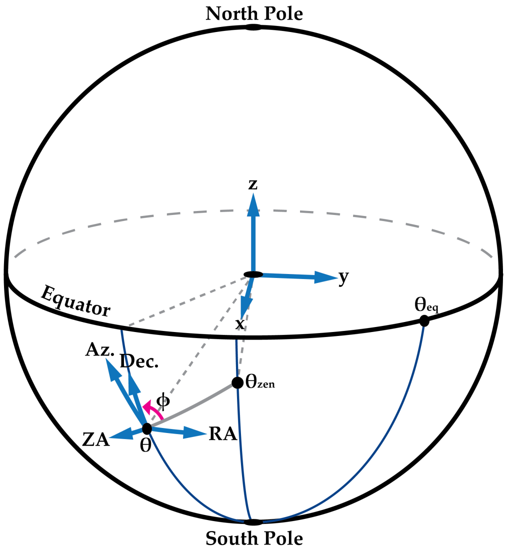


Figure D.1: Diagrammatic representation of the parallactic angle ϕ for a location on the sky $\boldsymbol{\theta}$.

taking the vector cross product between $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_{\text{zen}}$:

$$\begin{aligned}\boldsymbol{\theta}_{\text{zen}} \times \boldsymbol{\theta} = &\sin\delta_{\text{zen}}\cos\delta\sin\alpha\,\boldsymbol{e}_x \\ &+ (\sin\delta_{\text{zen}}\cos\delta\cos\alpha - \cos\delta_{\text{zen}}\sin\delta)\,\boldsymbol{e}_y \\ &+ \cos\delta_{\text{zen}}\cos\delta\sin\alpha\,\boldsymbol{e}_z\end{aligned} \tag{D.5}$$

To normalize, we find the amplitude of this vector:

$$|\boldsymbol{\theta}_{\text{zen}} \times \boldsymbol{\theta}|^2 = \cos^2\delta_{\text{zen}}(\cos\delta\tan\delta_{\text{zen}} - \sin\delta\cos\alpha)^2 + \cos^2\delta_{\text{zen}}\sin^2\alpha, \tag{D.6}$$

so the unit vector normal to the plane is

$$\begin{aligned}\frac{\boldsymbol{\theta}_{\text{zen}} \times \boldsymbol{\theta}}{|\boldsymbol{\theta}_{\text{zen}} \times \boldsymbol{\theta}|} = &\sqrt{\frac{1}{\cos^2\delta_{\text{zen}}(\cos\delta\tan\delta_{\text{zen}} - \sin\delta\cos\alpha)^2 + \cos^2\delta_{\text{zen}}\sin^2\alpha}} \\ &\times \left[\sin\delta_{\text{zen}}\cos\delta\sin\alpha\,\boldsymbol{e}_x + (\sin\delta_{\text{zen}}\cos\delta\cos\alpha - \cos\delta_{\text{zen}}\sin\Delta)\;\boldsymbol{e}_y\right. \\ &\left.+ \cos\delta_{\text{zen}}\cos\delta\sin\alpha\,\boldsymbol{e}_z\right].\end{aligned} \tag{D.7}$$

This vector is aligned in the negative azimuth direction.

At $\boldsymbol{\theta}$, the RA/Dec coordinate system is aligned with a plane that includes $\boldsymbol{\theta}$, the center of the sphere, and a point on the equator that we will denote $\boldsymbol{\theta}_{\mathrm{eq}}$:

$$\boldsymbol{\theta}_{\mathrm{eq}} = \left(\alpha + \alpha_{\mathrm{zen}} + \frac{\pi}{2}\right) \boldsymbol{e}_{\alpha}. \tag{D.8}$$

In our Cartesian coordinate system,

$$\boldsymbol{\theta}_{\mathrm{eq}} = -\sin\alpha\,\boldsymbol{e}_x + \cos\alpha\,\boldsymbol{e}_y. \tag{D.9}$$

A vector normal to this plane can be calculated by taking the vector cross product between $\boldsymbol{\theta}$ and $\boldsymbol{\theta}_{\mathrm{eq}}$:

$$\boldsymbol{\theta} \times \boldsymbol{\theta}_{\mathrm{eq}} = -\sin\delta\cos\alpha\,\boldsymbol{e}_x - \sin\delta\sin\alpha\,\boldsymbol{e}_y + \cos\delta\,\boldsymbol{e}_z. \tag{D.10}$$

Normalizing, we find that

$$|\boldsymbol{\theta} \times \boldsymbol{\theta}_{\mathrm{eq}}|^2 = 1, \tag{D.11}$$

so the unit vector normal to the plane is

$$\frac{\boldsymbol{\theta} \times \boldsymbol{\theta}_{\mathrm{eq}}}{|\boldsymbol{\theta} \times \boldsymbol{\theta}_{\mathrm{eq}}|} = -\sin\delta\cos\alpha\,\boldsymbol{e}_x - \sin\delta\sin\alpha\,\boldsymbol{e}_y + \cos\delta\,\boldsymbol{e}_z. \tag{D.12}$$

This vector is parallel to the positive Dec axis.

We now need to evaluate the angle between the unit vectors $\frac{\boldsymbol{\theta}_{\mathrm{zen}} \times \boldsymbol{\theta}}{|\boldsymbol{\theta}_{\mathrm{zen}} \times \boldsymbol{\theta}|}$ and $\frac{\boldsymbol{\theta} \times \boldsymbol{\theta}_{\mathrm{eq}}}{|\boldsymbol{\theta} \times \boldsymbol{\theta}_{\mathrm{eq}}|}$. The parallactic angle $\phi(\boldsymbol{\theta})$ is defined as the angle that rotates the negative zenith angle axis into the positive Dec axis. It follows that

$$-\sin\phi(\boldsymbol{\theta}) = \frac{\boldsymbol{\theta}_{\mathrm{zen}} \times \boldsymbol{\theta}}{|\boldsymbol{\theta}_{\mathrm{zen}} \times \boldsymbol{\theta}|} \cdot \frac{\boldsymbol{\theta} \times \boldsymbol{\theta}_{\mathrm{eq}}}{|\boldsymbol{\theta} \times \boldsymbol{\theta}_{\mathrm{eq}}|}. \tag{D.13}$$

Evaluating, we find that

$$\sin\phi(\boldsymbol{\theta}) = \frac{-\cos\delta_{\mathrm{zen}}\sin\alpha}{\sqrt{\cos^2\delta_{\mathrm{zen}}(\cos\delta\tan\delta_{\mathrm{zen}} - \sin\delta\cos\alpha)^2 + \cos^2\delta_{\mathrm{zen}}\sin^2\alpha}}, \tag{D.14}$$

and simplifying gives

$$\tan\phi(\boldsymbol{\theta}) = \frac{-\sin\alpha}{\cos\delta\tan\delta_{\mathrm{zen}} - \sin\delta\cos\alpha}. \tag{D.15}$$

$\phi(\boldsymbol{\theta})$ lies in quadrants III and IV when $\alpha > 0$ and quadrants I and II when $\alpha < 0$. We can therefore evaluate the parallactic angle with a four-quadrant inverse tangent function `arctan2` such that

$$\phi(\boldsymbol{\theta}) = \texttt{arctan2}\left[-\sin\alpha,\ \cos\delta\tan\delta_{\mathrm{zen}} - \sin\delta\cos\alpha\right]. \tag{D.16}$$

Appendix E

FARADAY ROTATION CORRECTION OVER A FREQUENCY CONTINUUM

Given an estimate of the RM R, we can use Equation 7.2 to relate observed polarized signal to true emitted signal across a frequency continuum.

At a single frequency f, observed Stokes Q and U power $Q'(f)$ and $U'(f)$ relates to true emitted power $Q(f)$ and $U(f)$ via the relationship

$$\begin{bmatrix} Q'(f) \\ U'(f) \end{bmatrix} = \begin{bmatrix} \cos[2\beta(f)] & -\sin[2\beta(f)] \\ \sin[2\beta(f)] & \cos[2\beta(f)] \end{bmatrix} \begin{bmatrix} Q(f) \\ U(f) \end{bmatrix} . \tag{E.1}$$

Faraday rotation does not affect Stokes I and V emission.

Continuum images combine signal across a frequency range $(f_{\text{min}}, f_{\text{max}})$. The observed continuum Stokes Q and U power Q'_{cont} and U'_{cont} is given by

$$\begin{bmatrix} Q'_{\text{cont}} \\ U'_{\text{cont}} \end{bmatrix} = \int_{f_{\text{min}}}^{f_{\text{max}}} \begin{bmatrix} Q'(f) \\ U'(f) \end{bmatrix} df = \int_{f_{\text{min}}}^{f_{\text{max}}} \begin{bmatrix} \cos[2\beta(f)] & -\sin[2\beta(f)] \\ \sin[2\beta(f)] & \cos[2\beta(f)] \end{bmatrix} \begin{bmatrix} Q(f) \\ U(f) \end{bmatrix} df. \tag{E.2}$$

We now assume that frequency evolution of the observed signal is dominated by Faraday rotation along the propagation path, such that the emitted signal can be approximated as frequency-invariant. Neglecting the frequency dependence of the emitted signal, we approximate the continuum emitted signal as

$$\begin{bmatrix} Q_{\text{cont}} \\ U_{\text{cont}} \end{bmatrix} = \int_{f_{\text{min}}}^{f_{\text{max}}} \begin{bmatrix} Q(f) \\ U(f) \end{bmatrix} df = (f_{\text{max}} - f_{\text{min}}) \begin{bmatrix} Q \\ U \end{bmatrix} , \tag{E.3}$$

where Q and U represent the emitted Stokes Q and U signal at any frequency within the observed frequency range. Plugging this approximation into Equation E.2, we get

$$\begin{bmatrix} Q'_{\text{cont}} \\ U'_{\text{cont}} \end{bmatrix} = \frac{1}{(f_{\max} - f_{\min})} \left(\int_{f_{\min}}^{f_{\max}} \begin{bmatrix} \cos[2\beta(f)] & -\sin[2\beta(f)] \\ \sin[2\beta(f)] & \cos[2\beta(f)] \end{bmatrix} df \right) \begin{bmatrix} Q_{\text{cont}} \\ U_{\text{cont}} \end{bmatrix}. \tag{E.4}$$

Evaluating, we find that

$$\begin{bmatrix} Q'_{\text{cont}} \\ U'_{\text{cont}} \end{bmatrix} = D \begin{bmatrix} \cos(2\beta_{\text{eff}}) & -\sin(2\beta_{\text{eff}}) \\ \sin(2\beta_{\text{eff}}) & \cos(2\beta_{\text{eff}}) \end{bmatrix} \begin{bmatrix} Q_{\text{cont}} \\ U_{\text{cont}} \end{bmatrix}. \tag{E.5}$$

Here β_{eff} is the effective rotation angle, given by

$$\beta_{\text{eff}} = \tan^{-1} \left(\frac{\int_{\lambda_{\min}}^{\lambda_{\max}} \frac{\sin(2R\lambda^2)}{\lambda^2} d\lambda}{\int_{\lambda_{\min}}^{\lambda_{\max}} \frac{\cos(2R\lambda^2)}{\lambda^2} d\lambda} \right). \tag{E.6}$$

D is a factor that accounts for the decoherence of the signal across the frequency range. It is equal to

$$D = \frac{1}{\left(\frac{1}{\lambda_{\min}} - \frac{1}{\lambda_{\max}} \right)} \sqrt{ \left(\int_{\lambda_{\min}}^{\lambda_{\max}} \frac{\cos(2R\lambda^2)}{\lambda^2} d\lambda \right)^2 + \left(\int_{\lambda_{\min}}^{\lambda_{\max}} \frac{\sin(2R\lambda^2)}{\lambda^2} d\lambda \right)^2 }. \tag{E.7}$$

The integrals evaluate to

$$\begin{aligned} \int_{\lambda_{\min}}^{\lambda_{\max}} \frac{\sin(2R\lambda^2)}{\lambda^2} d\lambda = & \frac{\sin(2R\lambda_{\min}^2)}{\lambda_{\min}} - \frac{\sin(2R\lambda_{\max}^2)}{\lambda_{\max}} \\ & - 2\sqrt{\pi R} \left[C(2\sqrt{R/\pi}\lambda_{\min}) - C(2\sqrt{R/\pi}\lambda_{\max}) \right] \end{aligned} \tag{E.8}$$

and

$$\begin{aligned} \int_{\lambda_{\min}}^{\lambda_{\max}} \frac{\cos(2R\lambda^2)}{\lambda^2} d\lambda = & \frac{\cos(2R\lambda_{\min}^2)}{\lambda_{\min}} - \frac{\cos(2R\lambda_{\max}^2)}{\lambda_{\max}} \\ & + 2\sqrt{\pi R} \left[S(2\sqrt{R/\pi}\lambda_{\min}) - S(2\sqrt{R/\pi}\lambda_{\max}) \right], \end{aligned} \tag{E.9}$$

where S and C are the Fresnel integral special functions:

$$S(z) = \int_0^z \cos(\pi t^2/2) dt \tag{E.10}$$

and

$$C(z) = \int_0^z \sin(\pi t^2/2) dt. \tag{E.11}$$

VITA

Ruby Byrne is from Portland, Oregon. From 2010–2014 she attended Macalester College in St. Paul, Minnesota, where she majored in physics and minored in mathematics. In 2015 Ruby moved to Seattle to work towards her Ph.D. at the University of Washington, where she joined Prof. Miguel Morales's radio cosmology research group. After graduating she will continue researching radio astronomy and 21 cm cosmology as a postdoc at the California Institute of Technology in Pasadena, California.