

THE EPOCH OF REIONIZATION:  
FOREGROUNDS AND CALIBRATION WITH PAPER

Daniel C. Jacobs

A DISSERTATION  
in  
Physics and Astronomy

Presented to the Faculties of the University of Pennsylvania in Partial  
Fulfillment of the Requirements for the Degree of Doctor of Philosophy

2011

---

James Aguirre,  
Supervisor of Dissertation

---

Charlie Johnson, Graduate Group Chair

Dissertation Committee:

Adam Lidz, Assistant Professor  
Gary Bernstein, Professor  
Bhuvnesh Jain, Professor  
Joe Kroll, Professor

THE EPOCH OF REIONIZATION:  
FOREGROUNDS AND CALIBRATION WITH PAPER

COPYRIGHT

2011

Daniel C. Jacobs

This work is copyright under the  
Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License

To view a copy of this license, visit  
<http://creativecommons.org/licenses/by-nc-sa/3.0>

This thesis has benefited beyond measure from the help, advice and pointed criticism of many people who are actually not too numerous to thank. Naturally none of this would have been possible without team PAPER: Aaron Parsons, Jonnie Pober, David Moore, Rich Bradley, Nicole Gugliucci, Pat Klima, and Erin Benoit. My South African friends Carel van der Merwe and Jason Manley were exquisite hosts in all capacities as were invaluable interns Tandeka, Charity, Freedom & Gerard. Ngiyabonga kakhulu! PAPER was the brain-child of Don Backer, an amazing person who took a chance with me. He was an honest and gentle role model who is sorely missed.

During the course of this work I benefited immensely from specific contributions by others that are worth calling out. David Moore overcame his natural enmity with computers to write a very nice CASA data converter, Jonnie Pober suggested the faceting that resulted in a fabulous all-sky map and Aaron Parsons invented, then carefully explained delay transforms and spearheaded the (ongoing) sensitivity calculations. In places of subtle genius their hand is clear; any mistakes are my own.

Thank you to my committee for taking on yet another student and setting aside part of your brain for radio astronomy. Though for James Aguirre I must reserve the largest measure of my gratitude. Thank you for taking me on and then trusting me to follow my family back across the country to finish this work in New Mexico.

Though the west would present its challenges. This thesis would have been unfinished without the brave people who fought the Las Conchas forest fire threatening Los Alamos while I finished the last chapter. Thanks for saving my town.

During the past few years my family really shone. Sister Becky, you cheered without flagging. Mom, thanks for the commiserations. There were many times when you said exactly the right thing. You won the PhD race! Thanks for helping me to the finish line. Dad, you are kind, gentle and wise. Thanks for the phone calls, they were some of the finest procrastination a son could ask for.

While help came from many quarters, only two people suffered. This thesis is dedicated to two amazing women: wife Karen and daughter Zoe. You've seen me at my worst and I've see you at your best. I love you both very much and promise that this doctoral thesis is my last. Now, on to the next adventure!

# ABSTRACT

## THE EPOCH OF REIONIZATION: FOREGROUNDS AND CALIBRATION WITH PAPER

Daniel C. Jacobs

Supervisor: James Aguirre

Over the last 20 years we have learned that the contents of the universe are split into 76% Dark Energy and 24% Matter, 17% of which is ordinary matter. Of the ordinary matter the bulk is hydrogen which forms the raw material for building stars. The universe began 14 Billion years ago with an expanding space-time and quickly began. After about 300,000 years this all cooled enough for the plasma to recombine into neutral hydrogen gas and release photons which we eventually observe redshifted into the radio; the Cosmic Microwave Background (CMB). Nearly half a billion years passed before the slow process of gravitational collapse would lead to the formation of the first galaxies and the (re) ionization of the ubiquitous hydrogen. This Epoch of Reionization (EoR) is the next major unexplored cosmological milestone. At the current time the space between galaxies is almost completely ionized, therefore we know that the universe must have undergone a global phase transition. The nature of the ionizing sources, whether young galaxies or accreting massive black holes is unknown. Neither do we know when this reionization occurred or how long it took.

Models suggest that we can detect fluctuations in the 21cm hydrogen emission line as ionization proceeds and high contrast ionized holes are carved in the neutral hydrogen. Detecting these fluctuations is one of the few direct probes of the reionization process but is a difficult task requiring a new generation of low frequency radio telescopes. Motivated by the breadth of unknowns, the Precision Array for Probing the Epoch of Reionization (PAPER) has been slowly building in complexity while folding the results of observations back into improving the design and operation of the telescope. As part of this process, this thesis analyzes early observations to explore three major areas of concern in detecting EoR: contamination by foreground sources, calibration stability and limiting sensitivity. Catalogs produced from this early data show good agreement with previous measurements. We conclude that the calibration is stable and sensitivity floors are close to the expected theoretical levels.

# Contents

<b>Abstract</b>	<b>iv</b>
<b>Contents</b>	<b>v</b>
<b>List of Tables</b>	<b>vii</b>
<b>List of Figures</b>	<b>viii</b>
<b>Preface</b>	<b>x</b>
<b>1 The Epoch of Reionization</b>	<b>1</b>
1.1 In the beginning...	1
1.2 Observing the EoR	2
1.3 Theory	3
1.4 High z HI observing	6
1.4.1 Comparison to CMB	6
1.4.2 The Fourier Domain	10
1.4.3 Noise Power Spectrum	17
1.4.4 SNR	17
1.4.5 Foreground Power Spectrum	19
1.4.6 Foreground Observations	22
1.5 Conclusion	23
<b>2 PAPER</b>	<b>25</b>
2.1 Design	25
2.2 Deployments	28
2.3 Data Processing	33
<b>3 Foregrounds</b>	<b>37</b>
3.1 Survey of past measurements	37
3.2 Constructing a Sky Model	37
3.3 New 145-MHz measurements of Southern sky sources.	41
3.3.1 Observations and Data Reduction	41
3.3.2 Catalog Construction and Flux Calibration	47

3.3.3	Discussion and Conclusions . . . . .	54
<b>4</b>	<b>Calibration</b>	<b>56</b>
4.1	Complex Gain Calibration . . . . .	57
4.2	Pipeline . . . . .	58
4.3	System Temperature . . . . .	63
4.3.1	Tsys model . . . . .	63
4.3.2	Observed system temp . . . . .	66
4.4	Conclusion . . . . .	69
<b>5</b>	<b>Sensitivity</b>	<b>71</b>
5.1	Theory . . . . .	71
5.2	Computing the Power Spectrum . . . . .	71
5.3	Integrating Power Spectrum . . . . .	73
5.4	Cross Talk . . . . .	78
5.5	Conclusion . . . . .	82
<b>6</b>	<b>Conclusion</b>	<b>85</b>
	<b>Bibliography</b>	<b>87</b>

# List of Tables

2.1	PAPER deployments . . . . .	29
2.2	PSA32 Observations . . . . .	34
2.3	A comparison between AIPY and CASA . . . . .	36
3.1	Low frequency surveys . . . . .	38
3.2	PAPER fluxes for 480 MRC sources and matching Culgoora fluxes (where available) . . . . .	47

# List of Figures

1.1	Global evolution of the CMB, gas and Kinetic temperatures (via Pritchard & Loeb (2008)) . . . . .	7
1.2	A slice through a model of 21cm brightness temperature showing evolution with redshift. . . . .	8
1.3	A summary of the EoR power spectrum: models, sensitivity, and data . .	9
1.4	Basic interferometer operation: The delay in arrival time between wavefronts corresponds with a peak in the spectral Fourier transform "delay space". . . . .	11
1.5	Snapshot approximation: Over a ten minute snapshot the $w$ term of the baseline vector is much smaller than $u$ and $v$ terms. . . . .	13
1.6	A power spectrum schematic: foregrounds, noise, EoR . . . . .	16
1.7	The relationship between foregrounds and EoR power. . . . .	18
1.8	The dimensions of the power spectrum measured by PAPER (adapted from Morales & Wyithe (2010)). . . . .	20
2.1	Labeled photo of a PAPER station . . . . .	26
2.2	Schematic diagram of PAPER . . . . .	27
2.3	A perpendicular cut through the PAPER primary beam at 150MHz . . . .	30
2.4	Aerial photo and configuration map of PAPER South Africa 32 element imaging array . . . . .	31
2.5	$uv$ distribution of PSA32 imaging configuration . . . . .	31
2.6	Point-spread function (PSF) of the PSA32 "dirty beam". . . . .	32
2.7	RFI survey of PAPER South Africa Site . . . . .	33
2.8	RFI flagging in South Africa . . . . .	34
3.1	Comparison between NVSS catalog values and NVSS points as given in Helmboldt. . . . .	42
3.2	Distribution of delay solution used for PSA32 catalog . . . . .	43
3.3	PSA32 sky coverage. . . . .	45
3.4	Declination and LST survey coverage . . . . .	46
3.5	A PSA32 wide bandwidth image centered at 22h-30d. . . . .	48
3.6	PAPER additions to two interesting spectra. . . . .	49
3.7	Catalog flux-scale compared with other quantities . . . . .	50

3.8	Section of PSA32 sky survey image near 5h 0d, showing corruption from Crab low in the beam . . . . .	51
3.9	Catalog fluxscale map. . . . .	52
3.10	PAPER's flux scale distribution . . . . .	53
4.1	Effect of delay calibration on imaging. . . . .	58
4.2	Key regions of the sky and the de Oliveira-Costa/Haslam smooth sky model. . . . .	59
4.3	Pictor A field. . . . .	60
4.4	EoR fields with CASA calibration and imaging. . . . .	61
4.5	Fornax A. Resolved as a double lobed radio galaxy. . . . .	63
4.6	Bandpass phase solution modeled as a single delay and a phase offset. . . . .	64
4.7	Delay solutions over 11 days are remarkably stable. . . . .	65
4.8	Relative amplitude solutions over 11 days are also very stable. . . . .	66
4.9	Centaurus A as imaged by PAPER 160MHz compared with new ATCA+Parkes mosaic . . . . .	67
4.10	The system temperate spectrum and time dependence: model and data. Deviation of the correlation temperature from the system temperature. . . . .	70
5.1	The delay spectrum of a typical but steep spectrum source with some RFI flagging is well isolated by the polyphase filter bank/CLEAN combination. . . . .	74
5.2	A model of foregrounds in a wide-band delay transform. . . . .	75
5.3	Integrating noise on a simulated correlation. . . . .	78
5.4	The single baseline, unit variance power spectrum and its integration properties in minimally processed data. . . . .	79
5.5	Cross-talk is stable. . . . .	81
5.6	Removing a long time average is effective. . . . .	83
5.7	Scaled power spectrum after cross-talk removal. . . . .	84

# Preface

Sometime about 14 Billion years ago the universe began with a hot big bang. We now know that the universe was endowed with a certain amount of energy which is divided between dark energy, dark matter, photons and baryons most of which was hydrogen plasma. After 300,000 years or so this plasma cooled enough for the electrons to recombine with the protons and release the photons of the Cosmic Microwave Background. After the first billion years the first stars, galaxies and black holes had formed into the objects we recognize today.

Observations of the CMB have verified this cosmological picture while deep integrations with optical and infrared telescopes have pushed closer to the birth of stars and galaxies. Despite these efforts, much remains unknown about the first billion years of evolution. In particular, we know very little about the first stars and galaxies. Were they massive and bright? Where they numerous but dim? Where the first galaxies anything at all like we see in more recent times? We see tantalizing hints of early galactic evolution at later times but, occasionally these facts are at odds. The period between the CMB era and the earliest observed galaxies is completely uncharted territory.

One of the last truly global events is the (re)ionization of hydrogen. It is certain to have happened, we observe that the bulk of the hydrogen is ionized to very high redshift yet must have been neutral for the CMB to propagate. Hydrogen emits a narrow spectral line at 21cm which we observe redshifted to several meters, near the commercial FM radio band. This transition is theoretically detectable with a sensitive telescope operating between 100 to 200 MHz. Ideally this telescope would image the gas at narrow redshift slices and so get a complete 3D image cube. This is out of reach of current technology, but even a relatively modest telescope may measure the power spectrum which is made bright and distinctive by the high contrast ionized holes in the neutral gas. Detection of hydrogen as it undergoes this process would be the highest redshift yet seen, a significant discovery. It is an ideal probe of early star and galaxy formation processes and an ideal compliment to traditional stellar astronomy.

In the near term we are limited to a detection of the power spectrum. This statistical measure allows us to combine measurements from multiple locations on the sky into one statistical measure of fluctuations. However, the optimal telescope design for measuring the power spectrum is still an open question. There is disagreement over the best way to achieve the required sensitivity though most approaches use an interferometer with a large number of elements. An interferometer directly measures the spatial Fourier modes

of the sky, making it an ideal instrument for a power spectrum measurement, but existing telescopes do not have enough elements to achieve the desired sensitivity. The Precision Array for Probing the Epoch of Reionization (PAPER) is one of several radio interferometers now under construction with the goal of detecting the power spectrum of hydrogen undergoing reionization. It is the only interferometer built solely for this purpose.

The design of the instrument is dictated by the need to cover a wide range of redshifts and therefore a wide bandwidth, but also achieve a high sensitivity and therefore a large number of antennae. This results in a challenging amount of data and a very wide bandwidth. To simplify calibration and minimize instrumental effects, the antenna must have a smooth spectral and spatial response. When combined with cost constraints these requirements result in a small element with a very wide field of view that breaks a number of common simplifying assumptions used in radio interferometry.

All of these problems are scaled by the ever increasing size of the array with the correlation of  $N$  antenna scaling everything by  $N^2$ . This work explores the data from 32 antennae. Only a year later we are now collecting 64 antenna data which is larger by a factor of four. Assuming no improvement in analysis tools the fraction of data we are able to explore will decrease by the same amount.

Radio astronomy has a long tradition of imaging. Despite our almost exclusive interest in the power spectrum, there are still several good reasons to image as well. If the data are faulty, error-prone or mis-calibrated this is very difficult to tell by direct examination, partly because of the sheer volume of correlation measurements and partly because of the unintuitive nature of interferometric measurements. Successful imaging is a powerful argument for instrumental stability.

However, in an array that is starved for sensitivity, imaging is in direct competition with measuring the power spectrum. In an interferometer the positions of the elements determine the Fourier modes measured. A power spectrum measurement must first measure each Fourier mode to good sensitivity which necessitates an arrangement that minimizes the number of independent modes sampled while an imaging array is optimal when it maximizes the number of modes sampled. These two requirements are in tension. We believe that we must observe in both imaging and power spectrum configurations to characterize the instrument and foregrounds and to achieve a power spectrum detection.

The large number of elements, the large bandwidth and necessity of balancing imaging with measuring the power spectrum are some of the difficulties with which we must contend. We can not nor need not fully investigate each problem with the same level of detail. The PAPER experiment approach is to investigate and solve problems as they come up and save money by only solving problems that need to be solved. These lessons will eventually inform the building of much larger telescopes where such an approach would not be possible. In this thesis I embrace this approach by investigating recent PAPER observations performing the first level of checks for problems that would prevent us from reaching design sensitivity. In this thesis I try to answer this question using early observations from the PAPER instrument taken during its extended construction period.

The PAPER collaboration is a group of around 12 people: three professors and assorted

students and engineers. This small group operates mainly out of the US with around two deployments to South Africa per year. Beginning in 2007 PAPER has yearly increased the size of the array by a factor of 2 with a goal of 128 elements in 2012.

PAPER's copper pipe dipole antennae, amplifiers, plastic pipe reflector frames and correlator are fabricated in the US and shipped to South Africa where we assemble them into an interferometer. Throughout the past three years I have had many opportunities to work with the instrument in the field. In addition to several outings to the test array in Green Bank, WV I was a key member of two of the past four deployments to South Africa. In October 2009 I was first on site to break ground and work the assembly and deployment effort for the first 16 dipoles. I led the effort to calibrate the positions of the elements using precision GPS surveying equipment and held primary responsibility for array position configuration. As a member of a three person team I worked a second deployment in May 2010 to commission the now 32 element array and make the first observations.

During this time it was common for more data to be produced than could be analyzed in detail. This thesis goes some small way to rectifying this situation by providing both an image and power spectrum analysis of the two problems most likely to affect an EoR detection: foregrounds and sensitivity.

The foregrounds are addressed by imaging the sky and extracting a catalog. To show that this is the output of a stable instrument, I compare the catalog fluxes and positions to those previously recorded and find a reasonable match. Not only did this prove the instrument was more stable than was previously thought, it also provided the first new measurements of many of these sources in almost 50 years. In the power spectrum domain I show via simulations and real data that if foregrounds are smooth as expected then they can be isolated from the region of power spectrum we would like to measure. As part of this I explain in more detail that the power spectrum measured by PAPER is almost entirely in the spectral domain i.e. the Fourier transform of the frequency spectrum. Careful transformation along this axis is key to foreground isolation, but unstable calibration or poor sampling along the spectral axis can swamp the measurement.

To get a better look at the calibration I brought the data into another analysis pipeline where time dependent solutions, among other things, were possible. To do this I worked closely with the Nation Radio Astronomy Observatory scientists and software engineers and organized the effort within PAPER including setting up several project workshops at the VLA operations and science center in Socorro, New Mexico. The result of this analysis was the first proof that the calibration heretofore assumed to be stable, actually was. As an additional validation of this pipeline I also generated several images of interest to observers.

With the foreground properties, necessary stability and spectral domain techniques established I then turned to the sensitivity in the power spectrum domain. To achieve the necessary sensitivity level, two things are required. First the measurement noise must be at the predicted level and second this noise must integrate stably down over the course of many nights of observing. The system temperature is partly caused by the Galaxy,

and thus has a spectral slope and is time variable. In addition, the variance in the cross-correlations is not necessarily proportional to the system temperature. Indeed this turned out to be the case for the upper half of the band. I do not theorize about this discovery; more investigation is warranted. It is sufficient to say that it might be indicative of a kind of noise that does not integrate down as thermal noise.

For this reason I explore the ability of just a single baseline to integrate down as it should. With only rudimentary data processing the noise integrates down with some remaining residual. Though, If we hope to integrate for 100 times longer more work is left to be done in identifying the limiting factors.. There are enough hints and approximations made that this can probably be achieved in the current data but with new and better data coming soon from the latest 64 antenna deployment our focus will most likely shift in that direction.

This catalog and the work published in this thesis is the first and possibly last look at 32 antenna PAPER data. It is a snapshot in time of the project and a rare glimpse into a project moving quickly towards its goal. Hopefully it will go some small way to providing the interested observer a better idea of where we are, the promise of the data and maybe some hints about where we are going.

Danny Jacobs  
Philadelphia, PA  
2011

# Chapter 1

## The Epoch of Reionization

### 1.1 In the beginning...

The story of the Epoch of Reionization begins, as all stories must, with the Big Bang. Here was the beginning of time as we know it; a universe filled with copious amounts of ionized Hydrogen (and Helium) plasma, black body photons and hugely inflated quantum density fluctuations. All of which were extremely hot and embedded in an expanding space-time. At these temperatures and densities the Hydrogen plasma was in thermal equilibrium with the photons; the free electrons scattered the photons making space opaque. Time passed and the Universe cooled. After about 300,000 years the number of photons above 13.6eV drops below the number of baryons and Hydrogen began to capture electrons. The plasma was neutralized and photons were free to proceed a few of which were eventually observed by us as the Cosmic Microwave Background (CMB; Loeb & Barkana (2001)).

Freed of its connection to photons, the Hydrogen (HI) pursued its own gravitational interests. Density fluctuations began to slowly accrete gas into what would eventually be clusters of galaxies. At the same time the hydrogen began to radiate, including a very long lived transition emitting photons 21cm long, which unfortunately appear to us, redshifted<sup>1</sup> as they are by intervening spacetime expansion, to have the unreasonable wavelength of 230 meters. At these wavelengths the ionosphere is completely opaque, the IGM is in free-free absorption and above redshift of 150 the HI emission is invisible against the CMB. Save these few radio waves and other similar atomic lines, no radiation is thought to have been generated for the next 500 million years until the birth of the first stars.

Thus began the time period known colloquially as the Dark Ages which was ended by an enlightenment of stars and AGN beginning around a redshift of 20, lasted around 700 million years and eventually resulted in the complete RE-ionization of Hydrogen.

But of course the Dark Ages were not dark. By redshift 20, the observed HI wavelength is only 4 meters, a challenging but not impossible observation. Here we will explore the

---

<sup>1</sup>Redshift ( $z$ ) is the inverse of the expanding Universe's scale factor ( $z = 1/a$ ) where  $a$  goes from zero at the Big Bang to 1 now, which can be directly measured by the apparent stretch of a photon's wavelength ( $z = \lambda/\lambda_0 - 1$ )

global evolution during the beginning of the end for HI. We will find that the stars drive the HI to have a distinct global spectrum ending in an Epoch of Reionization (EoR). We know HI is currently ionized but models give a range of redshifts at which reionization could have reached 50%. Establishing the redshift at which the universe is 50% ionized ( $z_{reion}$ ) is a goal of current EoR experiments (Furlanetto et al., 2006; Morales & Wyithe, 2010).

Constraining  $z_{reion}$  could answer many questions about the origin of stars, galaxies and massive black holes. Is a hierarchical model of galaxy formation correct? How does feedback affect star formation? What kind of fossils might remain today? Many of these questions can be constrained by single observable: **When did the IGM re-ionize?**

## 1.2 Observing the EoR

Evidence from absorption of quasars and the CMB indicates that the EoR most likely occurred between redshifts 6 and 14 (Furlanetto et al., 2006; Fan et al., 2006), but even this broad limit assumes a fairly simple "instantaneous" model of the transition. The Gunn-Peterson troughs caused by Ly $\alpha$  absorption of quasar spectra provide a sound lower limit of  $z \approx 6$ . However the Ly $\alpha$  line quickly saturates at relatively small neutral fractions (Fan et al., 2006). This effective limit in redshift space is compounded by the lack of decent statistics; we are limited in the number of pierce points by the number of known quasars at high redshift. Most recently observations of quasar ULASJ112001.48+064124.3 at  $z=7.085$  ( $\nu_{21} = 175\text{MHz}$ ) have been found to be consistent with a neutral fraction of 0.1 or greater (Mortlock et al., 2011).

In concert with other cosmological measurements, polarized Thompson scattering of the CMB provide an optical depth of free electrons, a measure of the age of our ionized IGM. The measure is consistent with a instantaneous re-ionization process sometime between  $9 < z < 14$  or a gradual, multi-stage process with a first stage at  $12 < z < 17$  and a second near  $z \approx 7$ . Like the quasar and other absorption estimates the best constraint provided by the CMB is on the end of re-ionization. Instantaneous ionization at  $z = 6$  has been ruled out to 99.9% confidence level (Dunkley et al., 2009).

Galaxy surveys, particularly with the Hubble Space Telescope, have begun to produce estimates of star formation rate and photon escape fraction at higher redshifts. These few measurements of total radiation and star formation history at high redshift are also dominated by strong selection limitations and other errors. Together their estimates of ionizing radiation suggest that there was probably enough Lyman-continuum radiation from stars to ionize the the IGM (Robertson et al., 2010). Deeper observations might reduce the uncertainty in this measurement. A good sample would require a much deeper survey in the near-IR by *James Webb Space Telescope* or a 30m class ground-based telescope. However these observations cannot by themselves establish the stars as the cause of reionization, nor will they be able to easily probe higher redshifts. Galaxy surveys in a universe where gaseous Hydrogen is the dominant baryonic matter are only part of the story.

Unlike stars and AGN, HI 21cm is a direct probe of the re-ionization process. The spin-

flip transition radiates a very narrow-band spectral line that allows precise determination of relative velocity, which for these distances is dominated by the Hubble flow, giving us a precise distance measure. Because of its low optical depth, this 3D probe traces both mass and temperature via its intensity but also provides a high-contrast probe of the ionization process and has been recognized as a very potent observable for everything from cosmology to galaxy formation and IGM astrophysics.

Direct detection of HI during reionization remains elusive. To-date only relatively unlikely scenarios have been ruled out. The single antenna EDGES experiment (Bowman & Rogers, 2010) has been able to eliminate "fast" reionization  $dz < 0.05$  at good confidence while Paciga et al. (2011) have made GMRT observations (see Fig 1.3) that rule out a fairly unlikely cold reionization (described below).

In the absence of any detection of high redshift hydrogen we are limited to best guesses from partly analytic and partly numerical simulations that track density, temperature and ionization fronts. Broadly, these models agree that as more objects radiate UV photons, ionization regions will increase in size eventually percolating through all space leaving only small islands of neutral hydrogen in deep, galactic-scale, gravity wells. These models provide a target sensitivity for detection efforts as illustrated in Fig 1.3. Most models agree to within a factor of 2 on  $z_{reion}$  and predict a somewhat wider spread of amplitudes.

Several EoR experiments are currently operating, including the Murchison Widefield Array (MWA; Lonsdale et al. 2009), the Low Frequency Array (LOFAR; Röttgering et al. 2006), the Giant Metre-wave Radio Telescope (GMRT; Paciga et al. 2011), and the Precision Array for Probing the Epoch of Reionization (PAPER; Parsons et al. 2010). Here we will focus on PAPER, an experimental meter wave interferometer under construction in South Africa.

The first detection of 21cm radiation in  $6 \leq z \leq 13$  will put a date on ionization of various size scales by constraining the power spectrum amplitude and redshift Bittner & Loeb (2011). Later experiments will measure the shape of the power spectrum from which we can learn about the matter and velocity distribution, as well as details about the ionization process and cosmological initial conditions (Lidz et al., 2007). In particular the deviations from spherical symmetry can constrain the initial power spectrum to put limits on inflation (Bowman et al., 2007). Imaging the spectral line signal is more challenging yet and is forecasted to start to be possible with arrays 10x the size of those currently under construction while full 3D imaging needs 100x, a enormous scale usually referred to as a Square Kilometer Array (SKA).

## 1.3 Theory

Like the CMB, much can be gleaned from the global spectrum of HI as it evolves. The brightness temperature of the 21cm ( $T_b$ ) line depends on couplings between the population of the two spin states

$$\frac{n_2}{n_1} = 3 \exp \left[ \frac{T_\star}{T_s} \right]$$

and available energy sources. The HI emission is viewed in contrast to the CMB photon field

$$T_b \simeq 29\text{mK} \left( \frac{1+z}{10} \right) \left( \frac{T_s - T_{CMB}}{T_s} \right) (1 + \delta_m) \chi_H$$

and is modulated by the gas density  $\delta_m$  and the ionization fraction  $\chi_H$ .

In the absence of energy sources like stars the temperature of the HI line is limited to coupling with either the cosmological matter ( $T_k$ ) or photons ( $T_{CMB}$ ). The temperatures of these reservoirs drop as the universe expands until the birth of the first UV and X-ray sources begins to significantly heat the gas. Eventually it succumbs to ionization.

The evolution of HI temperature has been analytically calculated by Pritchard & Loeb (2008) as shown in Figure 1.1. After recombination ( $z \sim 1100$ ) there were enough free electrons left over to couple the gas to the CMB photons ( $T_s = T_{CMB}$ ) but by  $z \sim 300$  these were absorbed by the increasingly cool gas. At this early time the matter density is still high enough to collisionally couple the spin state to the gas kinetic temperature (which is colder than the CMB by a factor of  $(1+z)^{-1}$ ) paradoxically putting the line into absorption with the CMB ( $T_s \sim T_M < T_{CMB}$ ). Eventually, probably around  $z \sim 70$ , the gas became too rarified to collisionally couple and the HI was once again dominated by the black body CMB photons ( $T_b = T_{CMB}$ ).

This state of things continued until the very first stars began to radiate ( $z \sim 20 - 30$ ). First the X-Rays and then UV photons from these early objects pump the 21cm transition via the Wouthuysen Field Effect (WFE) into states that are more sensitive to the gas temperature which for a time means returning to the cooler gas temperature ( $T_s = T_M$ ). As time progresses the growing number of radiation sources raises the gas temperature into the emission regime  $T_s \gg T_{CMB}$  before finally beginning the ionization process where the ionization fraction quickly grows to unity and the differential brightness temperature quickly drops to zero. Note that if the heating is so fast that the gas transitions nearly instantaneously from cold to ionized (so-called "cold reionization") the amplitude of the differential brightness will be much larger (100mK of absorption instead of 30mK of emission). This is the scenario probably ruled out by Paciga et al. (2011).

Figure 1.1 depicts this global history for various models of star formation. Driven primarily by relatively simple cosmological scale physics, the early portions are fairly well constrained by current cosmology. The various models tend to agree. However there is a wide range of reionization end-points ( $s_{reion}$ ). As we can see from the variety of end-points, measurement of the end of reionization would provide the strongest constraint on these models.

The emission causing the heating and ionization is thought to be emitted by both stars and quasars. Low mass stars provide abundant but soft spectrum UV while quasars are rare but hard sources capable of faster rates of ionization. Each class of ionizing sources

traces out its own history in a poorly understood relationship with the underlying dark and baryonic matter. Star formation depends on the abundance of metals in the early IGM as well as the thermal-kinetic flows of matter into and around dense regions. Quasar formation depends on the formation and evolution of massive black holes. These multiple interacting timelines act together to drive the temperature and ionization state of each point in the IGM through through the phase transition at  $z_{reion}$ .

While the higher redshift global spectrum is simpler to predict, the later period of re-ionization is richer in information for the same reason that it is difficult to model. The EoR band is also the highest redshift that can be observed from the ground where the ionosphere is still transparent and Radio Frequency Interference (RFI) can be avoided by observing from a remote location. Despite these relative advantages many difficulties must still be overcome, beginning with an initial detection of high redshift HI.

Early EoR experiments do not have the sensitivity to detect or image localized HI emission but must combine observations of multiple regions into a single measurement of the emission power spectrum. The first generation of these experiments is further limited to providing constraints on portions of this power spectrum, for example by looking for a peak in HI variance predicted to occur as re-ionization reaches the halfway point (Bittner & Loeb, 2011).

Generically speaking the power spectrum of a re-ionization model tells a simple story. A simulation by Matt McQuinn (Figure 1.2; McQuinn personal communication, 2010) that includes evolution, tells the story. Before ionization begins the power spectrum is simply the average temperature times the density. A power law distribution, increasing toward smaller scales. Ionization begins with small bubbles forming around, rare massive objects, adding power at large  $k$  which will percolate to larger scales as the bubbles expand. By the time ionization fraction reaches 50% most ionization regions will overlap and the variance will peak. After the halfway point the IGM will become a series of shrinking HI islands in a sea of HII. As they shrink power will move back to smaller scales but with an increasingly diminishing average level.

To put limits on the set of possible re-ionization histories with these early observations we must have in hand models suitable for comparison to measurement that sample a wide range of possible scenarios. Sampling both in angular and frequency space, early experiments will measure scales of 0.1 to 100 Mpc and measure how the power spectrum evolves over the half-billion years of first star formation ( $20 > z > 6$ ). Theoretical efforts have focused on gross estimates with analytical methods (Loeb & Barkana, 2001; Pritchard & Loeb, 11), detailed fully numerical and semi-numerical combinations thereof.<sup>2</sup> Each flavor samples a continuum between precision and statistical significance. Relevant to our current observations are their predictions of the evolution of the 21cm brightness temperature power spectrum. Analytical models of the power spectrum easily predict over all relevant scales but are limited in their ability to constrain non-linear effects such as the shape of HII regions or velocity perturbations. To be statistically significant numerical models must span a region much larger than the largest HII zone but have resolution small enough to

---

<sup>2</sup>Unless noted otherwise all distances will be given in co-moving coordinates

identify sources of radiation. Simulations that most accurately solve the full hydrodynamics of the IGM and propagation of the ionizing radiation are not quite currently technically feasible on these scales (Zahn et al., 2010). Current work has focused on semi-numerical methods that compromise between the twin desires of generating many simulations and increasing their accuracy (Zahn et al., 2010; Santos et al., 2010; Mesinger et al., 2011) such as the simulation by McQuinn shown in Figure 1.2. Spanning a size of 1300Mpc and including evolution over a redshift range from 12 to 7 this simulation represents the state of the art and approaches the size scales measured by PAPER. However there is only one. Most simulation work has focused on comparing results from different methods; and in consequence have made efforts to use the same initial conditions and physical processes.

Of course these difficulties are moderated by the need to predict on the scale of the limited sensitivity of early experiments. Even with the limited variability within simulations, there is enough spread in possible amplitudes to make a rough estimate of the possible constraints an experiment could offer. As can be seen from Figure 1.3, PAPER will have the ability to constrain a fraction of current models.

Though exploration of different re-ionization scenarios has been limited, several classes of scenarios have emerged as coarse testable areas. These are divided into "early" and "late" which are hypothesized to coincide with hard spectrum Quasars and softer small cool stars, respectively. Furthermore, a Quasar dominated spectrum would manifest as an "outward in" percolation from rare regions in contrast to a star dominated epoch of reionization "inside out" transition with more regions acting as nucleation sites for HII bubble growth (Zaldarriaga et al., 2004).

Constraints on these histories will require detection and constraint of the power spectrum at multiple scales and redshifts. A reasonable expectation for PAPER, given existing models, is that it will eliminate late (low-redshift) models that predict the most power on small scales as ionization reaches the halfway point.

## 1.4 High z HI observing

### 1.4.1 Comparison to CMB

Our first goal, then, is detecting any emission from high redshift HI. Ideally we'll do this by localizing excess power in both space (a range of  $k$  modes) and time (a redshift range). The obstacles to this detection are formidable both instrumentally, and observationally. Often parallels between the CMB and EoR are drawn that perhaps give a false sense of security. Both are global cosmological radio signals at high redshift. Whether one is observing the global temperature as a function of redshift (as in Fig 1.1), density fluctuations (Fig. 1.2) or the power spectrum (Fig. 1.3) all are at a brightness dictated by the global HI temperature. Since fluctuations are caused by ionization, their amplitude is proportional to the overall amplitude of the HI signal. In the CMB case the global signal was predicted and discovered independently many years before more precise measurements could be done. Though many attempts were made to measure the spectrum of this signal, 25+ years

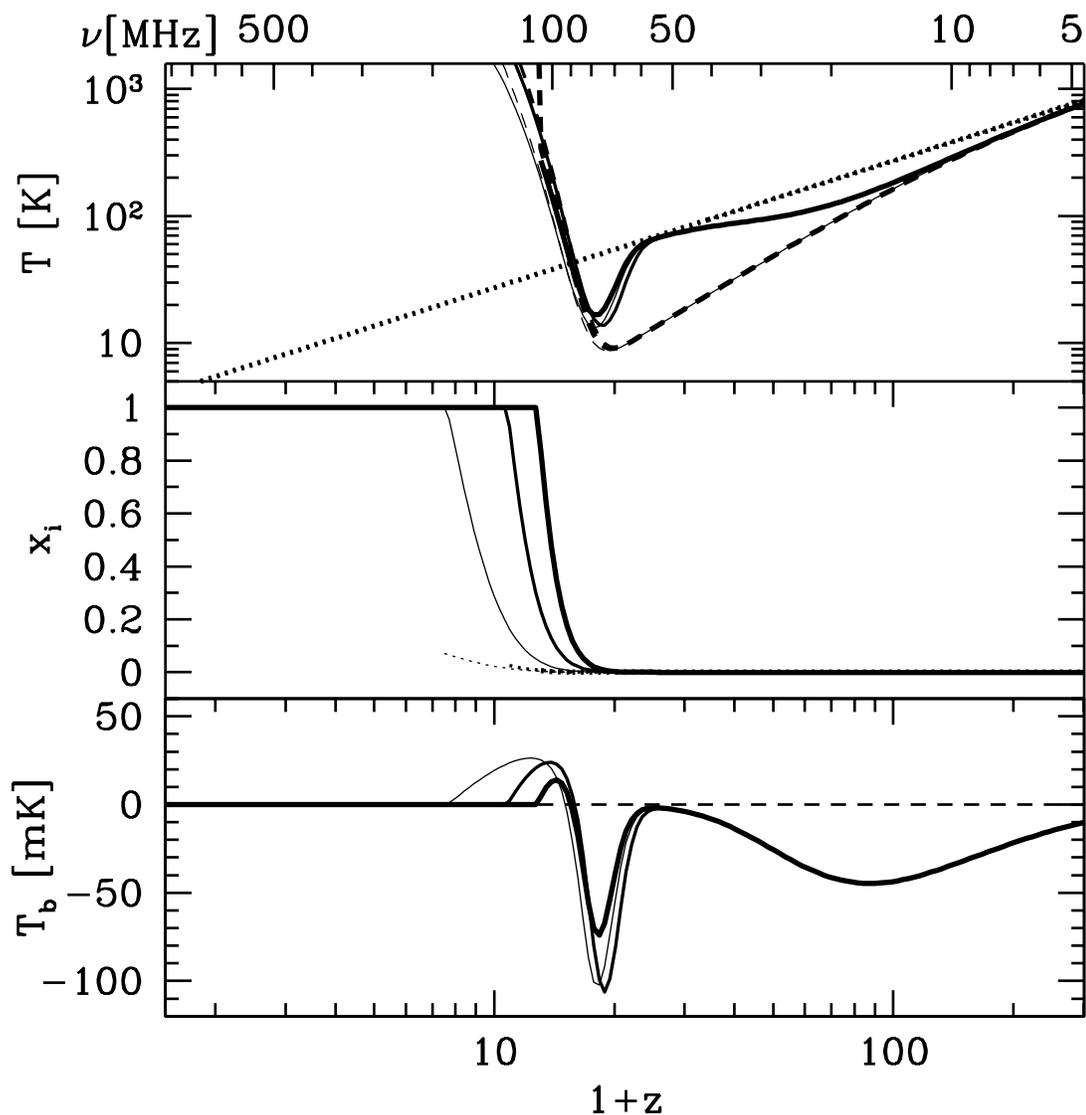


Figure 1.1: *Top panel:* Global evolution of the CMB temperature  $T_{\text{CMB}}$  (dotted curve), the gas kinetic temperature  $T_K$  (dashed curve), and the spin temperature  $T_S$  (solid curve). *Middle panel:* Evolution of the gas fraction in ionized regions  $x_i$  (solid curve) and the ionized fraction outside these regions (due to diffuse X-rays)  $x_e$  (dotted curve). *Bottom panel:* Evolution of mean 21 cm brightness temperature  $T_b$ . Each panel plots several models of star formation in solid lines of varying thickness. Driven primarily by relatively simple cosmological scale physics, the early portions are fairly well constrained by current cosmology. As we can see from the variety of end-points in the bottom panel, the makeup of ionization sources driving reionization less well understood. Measurement of the end of reionization would provide the strongest constraint on the variety of star formation models included here. (via Pritchard & Loeb (2008))

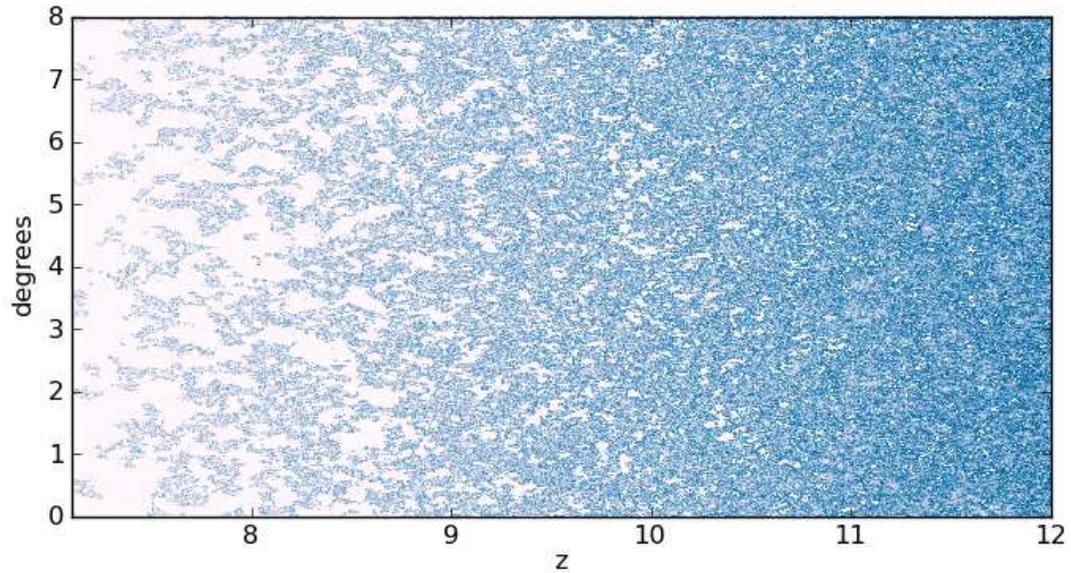


Figure 1.2: Here we have a radial  $0.02^\circ$  (3Mpc)-thick slice of a 1300Mpc wide simulation of 21 cm brightness temperature of HI, accounting for evolution in ionization fraction (McQuinn 2010, personal communication). At high redshift ( $z = 12$ ; 106 MHz), brightness temperature tracks gas density. As density fluctuations grow, UV photon production outpaces recombination, and regions become ionized (white). At the end of the era ( $z = 7$ ; 177 MHz) only the rarest regions have any remaining neutral hydrogen. Large simulation volumes and continuous redshift coverage from  $z = 6$  to  $z = 12$  are particularly relevant to PAPER.

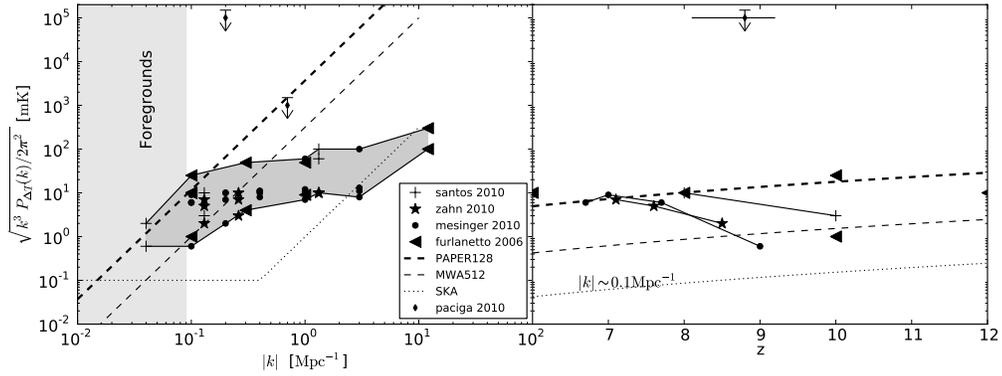


Figure 1.3: A summary showing the range of various recent models, recent measurements and instrumental sensitivities. On the left points are given regardless of redshift, though all are actively undergoing reionization with about 50% neutral fraction. For similar model parameters, various techniques agree well. All predict that peak amplitude occurs at 50% neutral fraction, but disagree on the exact redshift where this might occur. The Furlanetto et al. (2006) models are meant to provide a range of values in both  $z$  and  $k$  but are parametric in nature and are thus not connected chronologically. The two measurements by Paciga et al. (2011) (diamonds) are upper limits. The PAPER sensitivity curve assumes a maximally redundant grid-type 128 element at  $z = 9$  (Parsons et al., 2011). MWA and SKA sensitivity curves from Furlanetto et al. (2006), their Figure 30.

elapsed between the discovery by Penzias & Wilson (1965) and precision measurement of the black body spectrum by COBE. Operating in space, COBE’s FIRAS instrument compared the sky temperature to a precision calibration source in a relatively noise-free band to measure the spectrum to one part in  $10^5$ . In the case of the EoR neither the signal level nor the transition band-width are well constrained. EoR experiments with no knowledge of the amplitude can not with certainty estimate signal to noise ratio making the design sensitivity a matter of guesswork. Foregrounds are at least five orders of magnitude above the signal and RFI is epidemic<sup>3</sup>. Finally EoR (a 3D signal which varies with time) is fully twice as many dimensions as the 2D CMB. It fills all more of the spectrum with its signal leaving fewer data points with which to make an uncontaminated measure of the foregrounds. Surely the goal is worthy but the road is long, longer maybe than we might expect if we are comparing to the CMB.

Given that we cannot confidently set the parameters of the experiment with safe precision we must prosecute our search more carefully. In addition to our ignorance of the target amplitude we must overcome serious observational challenges about which we are also ignorant. Astronomy in this frequency range is complicated by large physical size (1 to 3 meters), the wide field of view that comes from using small cheap elements and the large number of elements necessary to achieve the requisite sensitivity. These telescopes must correlate thousands of channels over hundreds of elements; a difficult technical challenge. The large Field of View (FoV) and  $>100\%$  fractional bandwidth strain or break many of the interferometrists simplifying assumptions. Yet the instrument must be exceptionally precise to distinguish between EoR and 100,000 times brighter foregrounds. Confirmation of EoR fluctuations will require exquisite understanding of both instrumental and foreground effects. This suggests that a careful program of sky model and instrument improvement are essential elements of a path to detection.

## 1.4.2 The Fourier Domain

### Interferometric measurement

An interferometer measures the correlation between the electric fields measured by a pair of antennae separated by distance (or *baseline*) vector  $\vec{b}$ . On a quiet night, these electric fields are dominated by astronomical radiation  $I(\hat{s}, \nu)$  from direction  $\hat{s}$ , the wavefronts proceeding regularly across the array. In the limit of parallel wavefronts from a single distant source the correlation is given by the field power multiplied by the complex phase rotation the wave undergoes as it propagates the additional geometric distance ( $\hat{s} \cdot \vec{b}$ ) between the two antennae  $i$  and  $j$  (See Fig 1.4.

$$V_{ij} = I(\hat{s}, \nu) e^{-2\pi i \hat{s} \cdot \vec{b} \nu / c} \quad (1.1)$$

---

<sup>3</sup>The observation that EoR is might be a more difficult observation than the CMB was originally made by Furlanetto et. al. in their 2006 review.

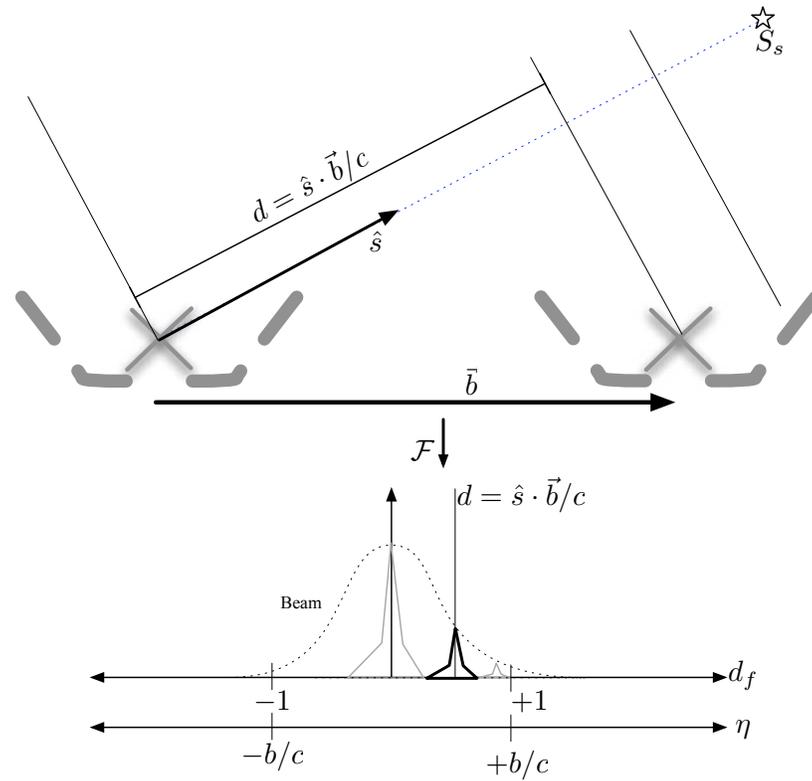


Figure 1.4: An interferometer correlates signals between two antennae (Eq. 1.1). The delay in arrival time between wavefronts corresponds with a peak in the spectral Fourier transform "delay space". See §5.2.

Of course the sky is full of sources, both continuous and discrete so we must integrate over the entire sky to find the total correlated power.

$$V_{ij} = \int I(\hat{s}, \nu) \exp[-2i\pi \hat{s} \cdot \vec{b}\nu/c] d\Omega \quad (1.2)$$

The baseline is often written in wavelengths  $\vec{u} = \vec{b}/\lambda = (u, v, w)$ , and the sky vector in cosines  $\hat{s} = (l, m, \sqrt{1-l^2-m^2})$  where  $l = \cos(\phi) \cos(\theta)$  and  $m = \sin(\phi) \cos(\theta)$  if  $\theta$  and  $\phi$  are elevation and azimuth, respectively. Another way of reckoning the correlation phase is that it is the geometric delay, the extra travel time, experienced by the light between the two antennae,  $d = \hat{s} \cdot \vec{b}/c$

$$V_{ij} = \int I(l, m, \nu) \exp[-2i\pi(ul + vm + w\sqrt{1-l^2-m^2})] dl dm / \sqrt{1-l^2-m^2} \quad (1.3)$$

Where  $I$  is the intensity at sky position cosines  $l, m$  and  $u, v, w$  are the relative coordinates of the two antennae  $i$  and  $j$ , measured in wavelengths. These are the coordinates in the frame of the sky, and thus rotate with the earth once every sidereal day. The integral is the sum of the correlated electric fields from all points on the sky where the field due to each point on the sky has a different geometric delay  $d = (\hat{s} \cdot \vec{u})$ .

At the moment of correlation the electric field has been amplified by each antenna differently. Each antenna has an overall amplitude calibration  $a(\nu)$  which changes counts to volts, and also has a relative phase  $\phi(\nu)$ . This phase is dominated by an electrical delay  $d$ ; a phase changing linearly with frequency ( $\phi = d\nu$ ).

$$g_i(\nu) = a(\nu)e^{i\phi(\nu)} \approx a(\nu)e^{id\nu} \quad (1.4)$$

Finally, the antenna has its beam pattern, a direction dependent gain  $A(\hat{s}, \nu)$  to go inside the integral, which we will assume to be similar for each antenna. Combining all of these we get a complete relation between the sky and the output of the interferometer.

$$V_{ij}^o = g_i g_j \int A^2(l, m, \nu) I(l, m, \nu) \frac{e^{[-2i\pi(ul+vm+w\sqrt{1-l^2-m^2})]}}{\sqrt{1-l^2-m^2}} dl dm \quad (1.5)$$

commonly referred to as the "visibility".

Given a model or measurements of the sky, beam, and baseline vectors we can integrate the right hand side, to get a model visibility  $V_{ij}^m$  which is related via complex gains to the observed visibility and can be written as a matrix equation,

$$\mathbb{V}^o = \mathbb{G} \mathbb{V}^m \mathbb{G} \quad (1.6)$$

Where the rows and columns of  $\mathbb{V}_{ij}$  are the antenna correlations, while the diagonal elements of  $\mathbb{G}$  are the gains. From here there are a variety of methods available for producing the model  $\mathbb{V}^m$  and solving for  $\mathbb{G}$ . This is necessarily an iterative process, where increas-

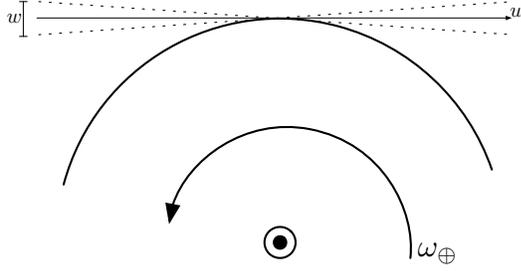


Figure 1.5: For a flat array, the  $w$  component is instantaneously zero and over a longer period of time is approximately zero. This period of time we refer to as the "snap-shot" which for PAPER is about 10 minutes. For illustrative reasons we have exaggerated the angles by a factor of 10.

ingly accurate knowledge of the sky is traded for a better gain solution.

As we will see, with certain simplifying assumptions measuring the correlation of the sky gives us access to both the flux distribution and the power spectrum with only a Fourier transform along the appropriate axes. Next we will quickly explore these approximations, use the results to relate the theoretical power spectrum to our measurement and finally estimate the necessary sensitivity level and the stability required to meet it.

### To the power spectrum

Even without the complication of instrumental calibration, the peculiar quantity measured by the interferometer itself is, at first glance, of limited use. The complicated interaction of the round sky, the flat array, and the change in effective baseline length with time and frequency makes deconvolution of  $I$  troublesome. However, a careful examination of the magnitude of the terms in the baseline vector, sky direction dot product  $\hat{s} \cdot \vec{u}$  allows us to make several simplifying approximations.

First, consider the output of our transit interferometer. Each sample measures a slightly different pointing  $\hat{s}(t)$  but we would like to image or compute power spectra with many samples toward a common pointing  $\hat{s}_0$ . To do this we can rotate to a coordinate system of the sky, where the the pointing stays the same but the array rotates.

$$\hat{s}(t) \cdot \vec{u} = (\hat{s}_0 + \delta\hat{s}) \cdot \vec{u}(t) \quad (1.7)$$

Our array is laid out on a uniform graded surface and is approximately flat. When  $\hat{s}(t) = \hat{s}_0$  the  $w$  or vertical component of the baseline vector is zero. After  $dt$  seconds the  $w$  term has increased by  $\cos(\delta) \sin(dt\omega_{\oplus})$ . At the north pole ( $\delta = 90^\circ$ )  $w$  is always zero. As an extreme example consider a 300 meter East-West baseline (the maximum length possible with PAPER) on the equator (see Fig. 1.5. Neglecting the  $w$ -term as the pointing rotates through  $\sim 3\%$  of the PAPER field of view, over 10 minutes, we incur a 4% error in our estimate of the phase.

Thus in this "snapshot" mode we are free to drop  $w$ , making the "flat-array" approxi-

mation

$$V_{ij} = \int A^2(l, m) / \sqrt{1 - l^2 - m^2} I(l, m) \exp[-2i\pi(ul + vm)] dldm \quad (1.8)$$

Looking ahead at our antenna beam pattern (Fig. 2.3), we see that the beam width of  $40^\circ$  is much less than the size of the sky, the  $\sqrt{1 - l^2 - m^2}$  component. In other words we may approximate that

$$A^2(l, m) \approx \frac{A^2(l, m)}{\sqrt{1 - l^2 - m^2}} \quad (1.9)$$

which at  $20^\circ$  amounts to an error of 4% and is known as the "flat sky" approximation. Together these two approximations have linearized our measurement equation

$$V_{ij} = \int A(l, m) I(l, m) \exp[-2i\pi(ul + vm)] dldm \quad (1.10)$$

which is now directly measuring the 2D Fourier transform of the sky. The image of the apparent sky  $A * I$  is just a Fourier transform away! Of course we aren't interested in the image. We want the full three dimensional Fourier transform, where the spectral domain gives us distance via the Hubble relation. Returning to our notation from Equation 1.2

$$V_{ij} = \int A^2(l, m) I(l, m, \nu) \exp[-2i\pi(b_x l + b_y m)\nu/c] dldm \quad (1.11)$$

The  $uv\nu$  are coordinates the space defined by the Fourier sky plane extended in the redshift direction. The  $uv$  coordinates are frequency dependent; a single cross-correlation spectrum (baseline) samples  $uv\nu$  space at a radial slant. Strictly this would mean that a baseline does not exactly lie "along" the line of sight  $k$  mode ( $k_{\parallel}$ ), however it is close. The largest range over which evolution could be considered static at these redshifts is usually assumed (Furlanetto et al., 2006; Morales & Wyithe, 2010) to be about  $dz$  1/2 or 6 to 8 MHz. Assuming the worst case of 8MHz of bandwidth at 110 MHz observed on our 300 meter baseline ignoring the wavelength dependence of a baseline length incurs a 7% error in the phase, somewhat larger than but reasonably commensurate with the flat-sky approximation. Ignoring the frequency dependence of a baseline is not significantly different than making the usual flat-sky approximating. Thus, as we illustrate in Fig. 1.7, we may extend our approximation to the  $z$  dimension.

Using this "Flat Space" approximation we set the frequency dependent baseline  $\vec{b}\nu$  to have its mean value  $\vec{u} = \vec{b}\nu_0$ . We are now able to take the Fourier transform in the frequency direction

$$\tilde{V}(u, v, \eta)_{ij} = \int \int I(l, m, \nu) \exp[-2i\pi\nu\eta] d\nu \exp[-2\pi i l u + m v] dldm \quad (1.12)$$

Over a narrow bandwidth this gives us a single line of sight skewer through the 3D power

spectrum.<sup>4</sup>. Substituting the measured units  $(u, v, \eta)$  for the physical wavenumber vector  $\vec{k} = (\vec{k}_\perp, k_\parallel)$ , converting to temperature units and squaring we find the relation between power spectrum and

$$\hat{V}_{ij}^2(\eta) \approx \left(\frac{2k_B}{\lambda^2}\right)^2 \frac{\mathbb{V}}{X^2Y} \hat{P}(\vec{k}_{ij}). \quad (1.13)$$

Where  $\left(\frac{2k_B}{\lambda^2}\right)^2$  converts from Kelvins to Janskys, X and Y convert angle and frequency to distance, and V specifies the volume ( $\text{Mpc}^3$ ) integrated to get  $\hat{P}(\vec{k})$ . In the case of our observation this is proportional to the volume of space probed by the product of field of view  $\Omega$  and bandwidth  $B$ . Here we convert to cosmological coordinates using the approximate relation used by Furlanetto et al. (2006) which is consistent with the WMAP5 cosmological parameters (Dunkley et al., 2009)

$$X \approx 1.9 \left(\frac{1+z}{10}\right)^{0.2} h^{-1} \frac{\text{Mpc}}{\text{arcmin}} \quad (1.14)$$

$$Y \approx 17 \left(\frac{1+z}{10}\right)^{\frac{1}{2}} \left(\frac{\Omega_m h^2}{0.15}\right)^{-\frac{1}{2}} \frac{\text{Mpc}}{\text{MHz}}, \quad (1.15)$$

where for  $\Omega_m = 0.26$  the volume conversion becomes

$$X^2Y \approx 540 \left(\frac{1+z}{10}\right)^{0.9} \frac{h^{-3} \text{Mpc}^3}{\text{sr} \cdot \text{Hz}} \quad (1.16)$$

. The power spectrum quantity most commonly computed by theorists is the total power in a radially logarithmic bin of a symmetric power density  $\hat{P}(\vec{k})$

$$\hat{\Delta}^2(k) \equiv \frac{k^3}{2\pi^2} \hat{P}(\vec{k}) \quad (1.17)$$

Switching to this representation, our power spectrum becomes

$$\hat{V}_{ij}^2(\eta) \approx \left(\frac{2k_B}{\lambda^2}\right)^2 \frac{\mathbb{V}}{X^2Y} \frac{2\pi^2}{k_{ij}^3} \hat{\Delta}^2(k_{ij}). \quad (1.18)$$

$\Delta^2(k)$  could be any of the many recent predictions, including those above. As noted above, there is still some disagreement among models of the 21cm power spectrum owing to uncertainties about the timing of reionization and the strength of star formation. Given this uncertainty we won't belabor the selection of a prediction beyond the fact that all are in rough agreement with (at most) 10mK of power at  $k > 0.1\text{Mpc}^{-1}$  occurring at some redshift between 7 and 11. Most importantly it provides a way to relate the noise in a visibility  $V_N$  to the power spectrum noise level  $\Delta_N^2$ .

---

<sup>4</sup>Portions of this section are submitted as Parsons et al. (2011)

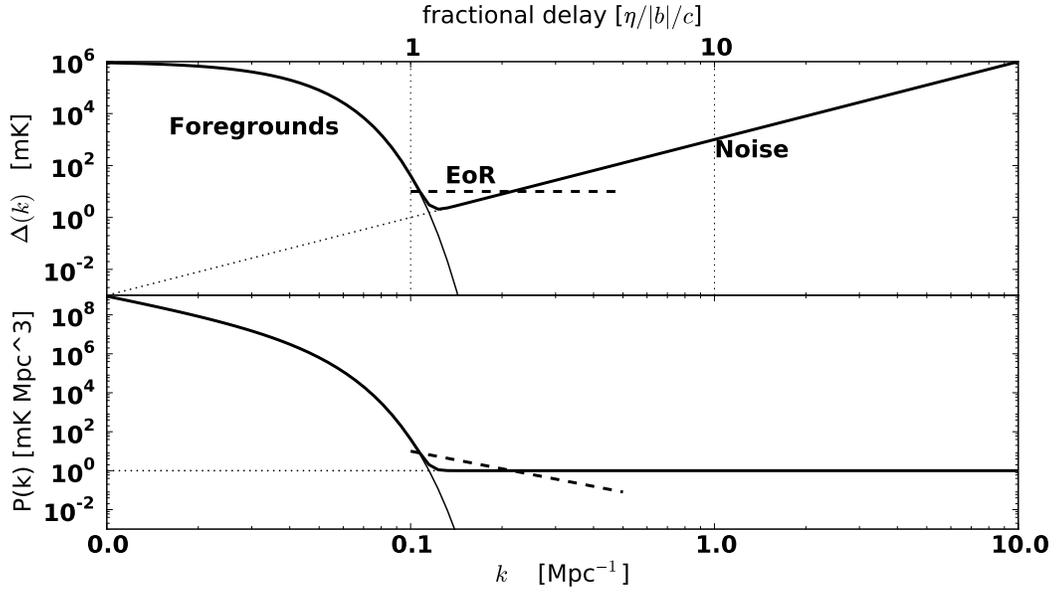


Figure 1.6: The observed power spectrum is composed of three components as shown in this approximate illustration of an ideal power spectrum noise curve. Smooth spectrum components will occur at fractional delays of  $|d_f| < 1$  and be attenuated by an approximately gaussian beam of width  $40^\circ$ . Here we assume a uniform distribution of sources over a range of fluxes that are attenuated by the primary beam. At higher  $k$  white noise will increase geometrically as  $k^3$  with the optimum detection point at the intersection of the two. Thus as baseline length  $|b|$  decreases the noise floor drops as  $|b|^3$ . The relative amplitudes in this plot are approximate and for illustration purposes only. While the EoR power spectrum makes the most sense on a narrow bandwidth, the foregrounds are most effectively filtered on a much wider bandwidth (see Fig. 1.7) and §5.2 for details.

### 1.4.3 Noise Power Spectrum

An interferometer with field of view  $\Omega$  will have a noise component  $V_N$  with amplitude

$$|V(\nu)_{ij,N}| = \left( \frac{2k_B}{\lambda^2} \right) T_{N,\text{rms}} \Omega \quad (1.19)$$

integrating over bandwidth to get the delay spectrum just adds a bandwidth term

$$\tilde{V}(\eta)_{ij,N} = \left( \frac{2k_B}{\lambda^2} \right) T_{N,\text{rms}} \Omega B \quad (1.20)$$

$T_{N,\text{rms}}$  is the temperature after integrating over bandwidth  $B$ , time  $t$ , and two polarizations  $T_{N,\text{rms}} = T_{\text{sys}}/2Bt$ . Combining Eqs 1.20 and 1.18 we arrive at an estimate of the noise contribution to the power spectrum *for a single baseline*

$$\Delta_N^2(k) \approx X^2 Y \frac{k^3}{2\pi^2} \frac{\Omega}{2t} T_{\text{sys}}^2, \quad (1.21)$$

### 1.4.4 SNR

The sensitivity of the entire array depends on the sensitivity of a single baseline, the number of independent samples of each  $k$  mode (or  $uv$  pixel) and the number of modes sampled. Here we derive the sensitivity of single baseline and relate the net sensitivity given the PAPER configuration found in a recent study to have the highest SNR on a single  $\vec{k}$  mode (Parsons et al., 2011).

To calculate the sensitivity given in Eq. 1.21 we need only estimate the total integration time for a single baseline measuring a single  $k$  mode. Sampling an angular size  $1/|u|$  our baseline will observe a fraction  $1/|u|/\sqrt{\Omega}$  of the sky in one earth rotation  $t_\oplus$  for a coherence time of  $t_\oplus/|u|\sqrt{\Omega}$ . For a fiducial baseline of 20 wavelengths (133ns) this corresponds to about 10 minutes, during which time we may average visibilities *before* squaring them which after 120 nights will reach:

$$\begin{aligned} \Delta_N^2(k) \approx & 2.8 \times 10^4 \left[ \frac{k}{0.1h \text{ Mpc}^{-1}} \right]^3 \left[ \frac{\Omega}{0.76 \text{ sr}} \right]^{\frac{3}{2}} \\ & \times \left[ \frac{T_{\text{sys}}}{500 \text{ K}} \right]^2 \left[ \frac{120 \text{ days}}{t_{\text{days}}} \right] \left[ \frac{|\vec{u}|}{20} \right] \text{ mK}^2, \end{aligned} \quad (1.22)$$

A sensitivity giving us an SNR of (at best)  $\approx 10^{-2}$ . Naively this means we would need 100 more baselines all measuring the same  $\vec{k}$  mode to reach  $\text{SNR} \sim 1$ . Of course our interferometer measures  $N(N-1)/2$  baselines, which in general might be independent  $\vec{k}$  modes. If this is the only baseline sampling this particular  $k$  mode our only further option is to average the *squares* of the power on other  $\vec{k}$ s of the same length but different

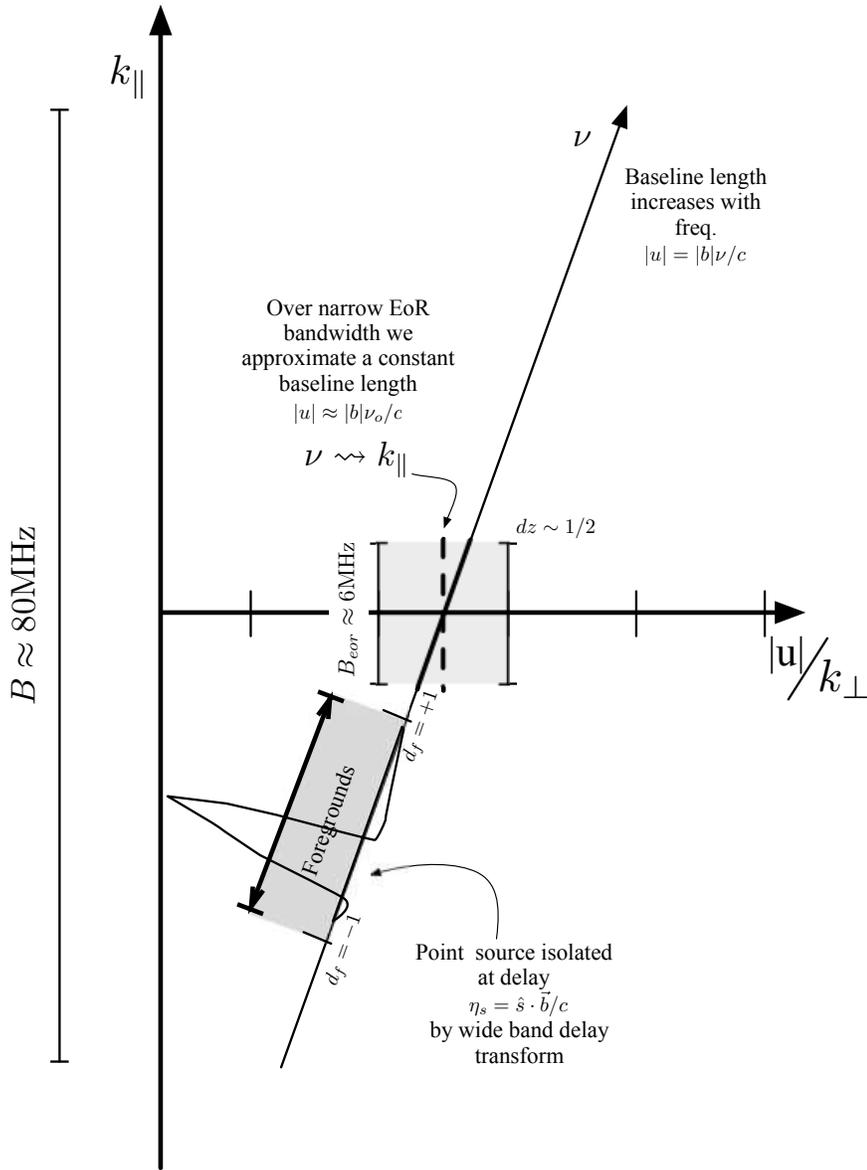


Figure 1.7: A single baseline at frequency  $\nu$  samples a baseline of length  $|u|$  corresponding to the  $k_{\perp}$  component of the 3D wavenode vector. Though  $|u|$  increases with frequency the change over a small, evolution-free or "co-temporal" redshift range is small and can be ignored. Over wide bandwidths where foreground sources are spectrally smooth the Fourier transform in the frequency dimension is close to a delta function in "delay" space.

angle. As was pointed out by Halverson (2002), in the limit where we do not achieve an  $SNR \geq 1$  on each mode, our noise will only decrease as  $1/N^{1/2}$  as we average independent  $\vec{k}$ s whereas the noise in power spectra with  $SNR \geq 1$  samples will average as  $1/N$ . Thus it is in our best interest to get as close as possible to that 100 measurements per  $\vec{k}$ -mode number before combining power on different  $\vec{k}$ s.

To do this with a limited number of antennae we must select an array that maximizes the number of similar baselines. After comparing the redundancy of several configurations, Parsons et al. (2011) selected a grid of 11 columns separated by 20 wavelengths (10m @150MHz) having 12 densely packed rows. This configuration, given the above single baseline sensitivity, will achieve a theoretical sensitivity of  $\Delta_N^2(k) \approx 33\text{mK}^2$  at  $k = 0.1h\text{Mpc}^{-1}$  or an SNR of  $\approx 3$ . In summary: For each baseline we must be able to average the same 10 minutes of sky for 120 days decreasing as  $1/t$  the entire way to  $\approx 1/N = 5 \times 10^{-5}$ , or  $\approx 45\text{dB}$ . In Chapter 5 we will test PAPER's ability to integrate as  $1/N$  and examine the importance of various instrumental effects and our ability to ameliorate these problems in post-processing. An integration this deep is challenging, particularly with a relatively simple instrument, however, as we've seen in our derivation of the power spectrum sensitivity, the geometrically flat nature of the EoR power spectrum means that the most sensitive part of the power spectrum is also the closest to the foregrounds which also merit careful study.

## 1.4.5 Foreground Power Spectrum

Noticing that noise drops steeply with  $k$  we might be forgiven for looking at Fig. 1.3 and assuming that the most sensitive part of the power spectrum is at the very lowest  $k$ s. But this assumption is only valid in the absence of foregrounds. Consider the relative dimensions of the  $\vec{k}$ s sampled by the interferometer as shown in Figure 1.8. In the high SNR array postulated above that focuses only a few  $uv$  pixels we are almost by definition sampling  $k_{\parallel}$  exclusively. To first order we may identify  $k$  with the  $\eta$  Fourier co-domain frequency with  $k_{\perp}$  acting as a constant offset. Thus we are concerned almost exclusively with the spectral behavior of foregrounds.

Galactic and extragalactic radio sources in the foreground are almost exclusively due to synchrotron radiation which is by nature spectrally smooth and therefore brightest at low  $k$ . They also span a bandwidth much wider than our EoR band. As illustrated in Figure 1.7 and described in §5.2 the "Flat Space" approximation (Eq. 1.12) does not apply. Rather than fight it we may use the linear dependence of the phase on frequency. The Fourier transform of a smooth point source in the frequency direction is nearly a delta function in  $\eta$  space at delays corresponding to the geometric delay  $d_s = \hat{s} \cdot \vec{b}/c$  which is geometrically limited to the physical length of the baseline; the **fractional delay** is less than unity ( $|d_s|/|b|/c < 1$ ). Naturally the foregrounds are not perfectly flat nor is the bandwidth infinite so in practice the width of a source in delay space is increased somewhat. These practical questions are addressed briefly in §5.2 via some simple simulations and carefully constructed Fourier transforms but for now it is sufficient to note that foregrounds occupy

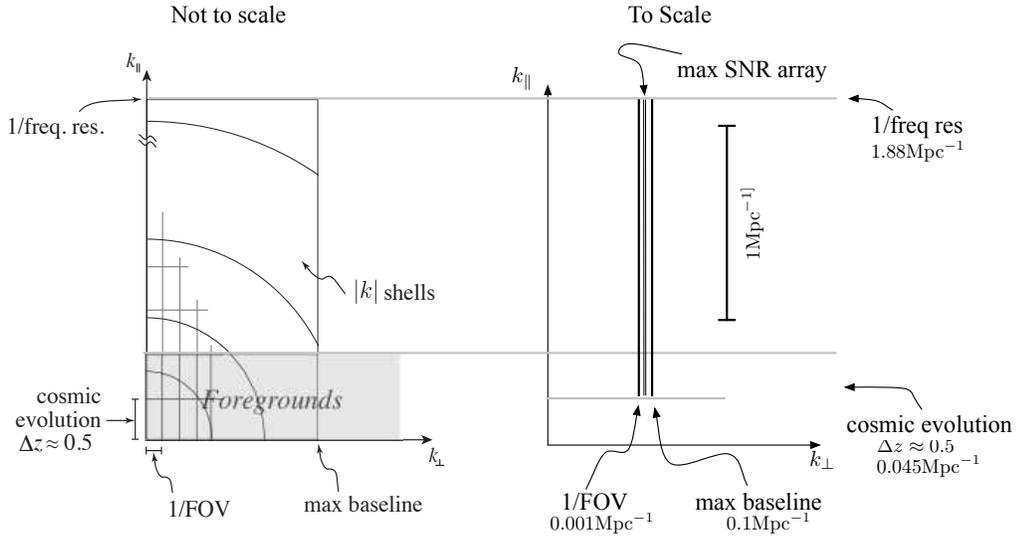


Figure 1.8: The dimensions of the power spectrum sampled by PAPER. As we see from the schematic by Morales & Wyithe (2010) (left), baseline length and field of view set the ultimate limits on  $k_{\perp}$  while bandwidth (limited by evolution) and frequency resolution set the limits on  $k_{\parallel}$ . On the right we see that for PAPER these limits (vertical black lines) sample mostly in the line-of-sight  $k_{\parallel}$  direction. The highly redundant, high SNR array measures an even narrower range of baselines (vertical grey lines) and almost by definition samples  $k_{\parallel}$  exclusively.

a range of  $k$  that could theoretically drop off very steeply going to higher  $k$ .

Finally we have a clear picture of the power spectrum playing field. The foregrounds dominate the low delay modes of power spectrum but shouldn't spill beyond a fractional delay of 1, while the steep  $k^3$  ( $\eta^3$ ) dependence of the noise takes over at high delays. The minimum in the sum of these two terms (shown in Fig 1.6) is the area of optimum sensitivity. As baseline length  $|b|$  decreases the noise floor drops as  $|b|^3$ . As we can see in Fig 1.3 EoR power does decrease with decreasing  $k$  but at a much slower rate than our noise. Therefore we should choose as short of baselines as possible to achieve a minimum floor and maximum SNR. This sharp dependence of the foreground location in delay also highlights possible problem areas. The Fourier transform to delay space must be carefully done to maximize dynamic range and delay calibration (§5.2) and determination of the  $\eta$  zero point, must be precise and stable (§4).

In this way we have a very simple and mostly correct estimate of the 3D power spectrum. Several features of this method make it particularly valuable in these early experiments. Firstly, we can observe and understand the instrumental and foreground affects at the single baseline level. As an experimental telescope PAPER is constantly recommissioning new hardware and increasing the complexity of the instrument. New errors both digital and analog must be identified and flagged at many levels. Being able to flag based on the power spectra of single baselines is a powerful tool. Second, it opens up a better way to filter foregrounds.

The most promising approach to removing foregrounds is to identify and subtract a smooth spectral component. A standard technique used to separate continuum from spectral line observations. As foregrounds are assumed to be both wide-bandwidth and smooth, we should characterize and subtract them over as wide a bandwidth as possible before we compute power spectra in the much smaller redshift range. Removing foregrounds while leaving enough signal to make a detection has been shown to be possible in simulation (Morales, 2005; Bowman et al., 2008; Morales & Wyithe, 2010) and was the principle technique in a recent detection attempt (Paciga et al., 2011). In all of these efforts, observations are calibrated and then gridded onto a  $uvv$  cube before the smooth component was removed. As we've seen, each baseline describes a slanted sampling path through this space which is small over EoR bandwidths but is large over the entire redshift range of interest and thus will span many  $uv$  samples. In some cases the size the array is chosen such that the density of  $uv$  plane samples is high enough that gaps are minimized. With a completely sampled  $uvv$  cube one may simply fit and subtract a smooth model. However this means that a single spectrum in the  $uvv$  cube will be the combination of many different baselines each with a slightly different calibration and the occasional gap in coverage. The net effect of this process is that the number of samples and the calibration accuracy will vary along the frequency axis. This variability will increase the error in fitting the smooth foreground component. Errors which will propagate in turn into the power spectrum and our EoR detection. By considering each baseline without gridding we can more carefully characterize and eliminate foregrounds.

## 1.4.6 Foreground Observations

As observed in §1.4.5 the (hopefully) smooth foregrounds are well isolated at low  $k$  by carefully constructed delay transforms as in Fig. 1.3 allowing us to filter on wide bandwidth before zooming in on a narrow EoR band. However, much remains unknown about the sky at 2 meters. The extreme brightness (greater than 5 orders of magnitude in the best case) puts tough requirements on instrumental and filter dynamic range while unexpected power at high  $k$  or unstable delay calibration can easily contaminate the power spectrum.

The brightest foregrounds also serve as calibrators. Calibration is thought by some to be the most important element of an EoR experiment (Datta et al., 2009). Under the right circumstances, subtraction of a point source with an incorrect gain model could scatter power to high  $k$ , (Datta et al., 2010) though the precision needed is still the subject of much debate. The effect has been minimally explored though is occasionally worried about (Morales & Wyithe, 2010). In all cases these treatments have focused on "frequency dependent sidelobes" an effect that comes from gridding in the  $uvv$  domain before removing foregrounds rather than in the delay domain (see §1.4.2 and §5.2) where electromagnetic linearity is embraced rather than fought. Regardless of method most simulations of foreground removal assume perfectly calibrated data and all bright sources removed to the few mJy level (Bowman et al., 2008; Jelić et al., 2008).

Another reason to refine the calibration is to reduce errors in measurement caused not by errant signal but incorrect estimation of complex gain introducing a 'calibration noise', the magnitude of which depends on the stability of the instrument. Calibration noise is the ultimate limiting factor, and has been recognized as such by eg Furlanetto et al. (2006) by way of recognizing a fundamental upper limit on the possible integration time beyond which small drifts and instabilities make any increase in precision, via integration, impossible.

Other elements of unknown relative significance include polarization and RFI. Measurements of polarized foregrounds with Westerbork have revealed evidence for a complex, frequency dependent polarized signal arising from the interplay of Faraday rotating plasmas, and polarized galactic and extragalactic emission (de Bruyn et al., 2009; Bernardi et al., 2010). Should this rapidly and non-linearly rotating polarization leak into the stokes I, there is a possibility of several mK of high- $k$  spectral variation (Jelić et al., 2010), though the contribution to the power spectrum itself has not been explored. Having both narrow bandwidth and unpredictable manifestation, RFI represents a great threat to redshift domain power spectra. RFI at the remote South African and Western Australian sites is quite low, with interference from satellites and airplanes being the most common element. However the RFI levels at the levels required to achieve mK sensitivity have not been measured. Existing RFI surveys have a sensitivity floor of  $\approx 200 \text{ dB F} [W/m^2/Hz]$  ( $10^6 \text{ Jy}$ ) or about a million times brighter than most foreground sources (Furlanetto et al., 2006; of South Africa, 2005).

For all of these reasons we wish to arrive at an accurate model of the sky: first to

calibrate and then to discriminate from EoR. Chapter 3<sup>5</sup> is another step towards this goal. In §3.2 we describe existing measurements and examine their relative merits. §3.3 then introduces a new Southern hemisphere catalog at 150MHz as observed by PAPER. Many of these sources have yet to be measured in this band and will be used extensively by PAPER and other EoR experiments to improve sky models and calibration.

## 1.5 Conclusion

As the SNR calculation clearly demonstrates, even under the best of circumstances PAPER (and EoR experiments in general) has a challenging detection ahead. Some of the issues that could arise have been mentioned already: noise sources (eg  $T_{\text{sys}}$ , calibration, rfi), cross-talk, unknown spectral foreground features, unknown spectral features in the beam or the just plain unknown. We are attempting to measure a variance of 10mK on top of a noise of 500K, a dynamic range of at least a million very close to foregrounds that are another  $10^5$  brighter than the noise. We are digging deeper than ever before and should therefore expect the unexpected.

In this work we address one broad question: Is the instrument stable enough for an integration this long? By asking three related questions:

1. **Does PAPER reliably measure the sky with little calibration?** This is key to assessing our ability to isolate foregrounds in delay space but also validates the instrument and the data pipeline to zeroth order. In Chapter 3 we use a small amount of data (two nights) to image the entire southern sky, construct a catalog, and compare our flux measurements with past data.
2. **Is the calibration assumed to be stable, actually stable?** In Chapter 4 we answer this question by building the foundation for an alternate data pipeline capable of time and frequency dependent calibration. We demonstrate the pipeline by calibrating 11 nights of observation of a single  $\vec{k}$  mode, showing that it is stable. We also note the ancillary benefits of this second toolbox by imaging several regions including the field most suitable for EoR and the spectacular Centaurus A radio galaxy. Finally we use this calibration to compute the system temperature and compare with a likely model.
3. **Are the properties of the instrument noise and stability such that we can reach our desired single baseline sensitivity?** In Chapter 5 we test whether our 11 nights of data integrate as  $1/N$  on a single  $\vec{k}$  mode.

By cataloging our early imaging results, we find that even given a crude calibration and imaging process we do an adequate (as good as past measurements) job measuring broadband source fluxes. We even produced a new catalog of the most low frequency measurements in the southern sky in the bargain. Furthermore, the problems that limit our accuracy

---

<sup>5</sup>§3.3 is an expanded version of Jacobs et al. (2011).

appear to be primarily related to the limitations of the data processing not the instrument itself.

The ability of the instrument to integrate to the necessary level, while not directly confirmed (due to lack of data) is consistent with the available data. We show, for the first time, that the system temperature agrees with models and that a single baseline integrates down as expected.

# Chapter 2

## PAPER

### 2.1 Design

Many of the problems thought to confound 21cm EoR detection and measurement are still unexplored. PAPER is an experimental interferometer built with the intention of approaching these problems carefully. Beginning as a set of 4 dipoles in a Green Bank, WV field, it has steadily grown over four years to a 32 element array in South Africa, soon to increase to 128. Prior to this work a series of commissioning observations of an 8 antenna configuration in Green Bank were documented by Parsons et al. (2010), hereafter PGB8.

In Figure 2.2 we see a block diagram of the intentionally simple "paper clips and a correlator" design. Crossed dipoles are mounted in matching "crossed-trough" ground screens (see Fig. 2.1). Crossed dipole antennae are sandwiched between aluminum disks that act as a sleeve to broaden the frequency response to span 100 to 200MHz and a beam that is about 40° wide at the half-power point and has its first null directed almost 180 degrees from zenith. A detailed model has been constructed using "computer simulation technology" (CST) electromagnetic modeling software which includes the sleeve, reflector and ground plane (See Fig.2.3) and is used where necessary in this work.

Directly attached to the antenna is an "active-balun" which provides 60dB of gain before the unbalanced signal is transmitted over non-buried coaxial cable to a central RFI tight container. Inside the container the signal is amplified again, and filtered to the desired bandwidth. The resulting signal is digitized at 100MHz, Fourier transformed by an IBOB<sup>1</sup> Field Programmable Gate Array (FPGA)-based "F-engine" and distributed over a commercial ethernet switch to a grid of FPGAs in the standardized ROACH<sup>2</sup> platform for cross-multiplication in an "X-engine".

Keeping the system low cost while simple to use and calibrate has been the focus of early design iterations. Single dipoles are used (in place of a phased array) to keep the frequency and sky variability of the primary beam to a minimum. The pre-amplifiers also act as baluns to low-cost unbalanced tv coax transmission lines. The second stage only

---

<sup>1</sup>Interconnect Break-Out Board

<sup>2</sup>Reconfigurable Open Architecture Computing Hardware

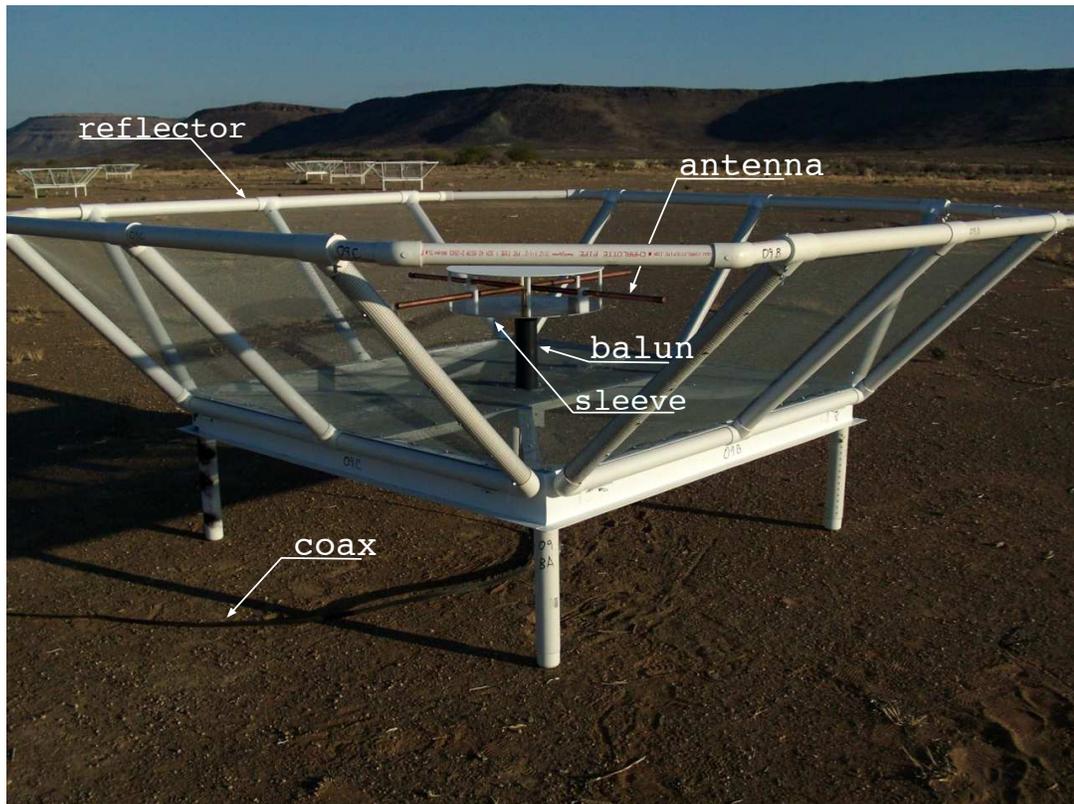
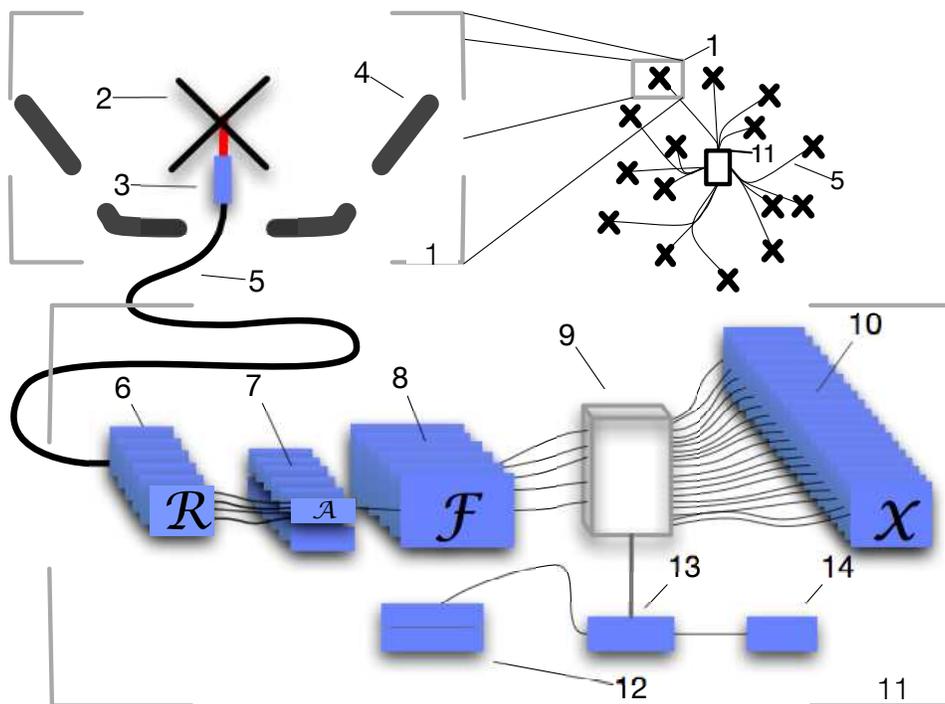


Figure 2.1: A PAPER station is designed to be very wide bandwidth and field-of-view. The crossed-dipoles are sandwiched between steel plates that act as frequency broadening sleeves, while the trough reflectors limit sensitivity at low elevations without introducing nulls (see Fig. 2.3). Within the mast is an "active" balun which provides 60dB of amplification as well as a match to the commercial 75Ohm coax which is left un-buried to allow for simple repositioning of elements.



1. Dual polarization receiving element,
2. linear crossed broadband dipole,
3. amplifying balun,
4. square trough reflector,
5. TV coax cable (above ground),
6. filter and amplifier,
7. Quad - Analog to Digital Converter (ADC)
8. IBOB Fourier Transform Engine,
9. 10Ge ethernet switch performing "few-to-many,  $N$  to  $N^2$  operation"
10. ROACH cross-multiply engine,
11. RFI tight 40' container,
12. RAID storage,
13. Data receiving server,
14. Visualization,

Figure 2.2: An overview schematic of PAPER, a reconfigurable interferometric array of dipoles operating between 100 and 200 MHz. Signals are received at a large number of stations (32 in 2010, 128 in 2012) and correlated by a reconfigurable FPGA-based correlator that uses commercially available network switches in place of a custom backplane and is easily scalable to take advantage of faster hardware to correlate more channels or antennae. Using this approach PAPER has deployed a new correlator, arguably one of the most difficult aspects of radio interferometer construction, once a year for three years in a row with each new system correlating twice as many inputs as the previous.

filters and amplifies; analog-to-digital conversion is done at the baseband. The correlator is implemented on a standardized FPGA platform developed by team members in collaboration with other astronomy users within the Collaboration for Astronomical Signal Processing Electronics Research (CASPER<sup>3</sup>). The combination of open libraries and development tools invented by this group has enabled PAPER to rapidly design and deploy a correlator to take advantage of constantly improving hardware. The use of standard ethernet switches and protocols to distribute data to the X-engines gives the framework the flexibility to change the design of the array in number of antennae or channels. In this way PAPER has deployed a new correlator, arguably one of the most difficult aspects of radio interferometer construction, once a year for three years in a row. Each new system used faster hardware to correlate double the inputs as the previous model.

By leaving out many conventional aspects of radio interferometer design we bring the cost to a practical level. Items like calibration sources, phase switches, buried cables and custom correlators have all been used in the past, often for good reasons but significantly raise cost and complexity. So while we have stripped the idea of the interferometer to the bare essentials we must analyze the results from early experiments with an eye towards the quality of data and the necessity of adding back the most critical of these elements. Particularly relevant in this work are the effects of self-interference, feedback, cross-talk and calibration stability. The relative ranking of these systematics on the power spectrum sensitivity will be the crucial factor in determining the most profitable experimental direction for PAPER and other EoR telescopes

## 2.2 Deployments

PAPER has progressed as a series of experimental deployments (listed in Table 2.1) focused on increasing the sensitivity of the array both by improving stability, reducing noise and adding elements. Early experiments with digital correlation and data reduction used a series of four element arrays deployed in Green Bank, WV and Western Australia. These deployments demonstrated the quick time-to-science possible with cheap antennae, a simple front-end and the CASPER FPGA development process. In a second series of deployments an 8 antenna array in Green Bank was used to validate the concept of calibration and imaging and to make early observations of northern foregrounds.

PAPER broke ground on a new site at the SKA candidate site in the Karoo desert of South Africa in October 2009 and began recording with a 32 element array in May 2010. This setup used a single EW polarization to record correlation spectra between 120 and 180MHz at 48kHz resolution and an integration time of 5.37 seconds. Four incomplete days of commissioning data were recorded in May followed by almost 10 days of uninterupted recording in September 2010 (see Table 2.2 for details).

These first observations are intended to characterize foregrounds and make initial estimates of instrumental issues ahead of EoR observations. An optimal configuration for

---

<sup>3</sup><https://casper.berkeley.edu/>

Table 2.1. PAPER deployments

Location	Ant. Number	Date	Configuration
Green Bank, WV	4	2007	Cross
Western Australia	4	2007	Cross
Green Bank, WV	8	2008	Imaging (Circular <sup>a</sup> )
Green Bank, WV	16	2009	Imaging (Circular)
Green Bank, WV	32	2010	Imaging
South Africa	32	2010	Imaging
<i>South Africa</i>	<i>64</i>	<i>2011</i>	<i>Imaging (Future)</i>
<i>South Africa</i>	<i>128</i>	<i>2012-</i>	<i>Imaging/Power spectrum (Future)</i>

<sup>a</sup>Antennae arranged non-uniformly around a circle of radius 100m.

this goal is to emphasize all of the above criteria as well as uniformity of UV coverage. A configuration support matrix

$$\mathcal{C} = \sum_a \delta(x_i - x_{A_i}) \delta(x_j - x_{A_j})$$

$\mathcal{C}$  which achieves this will maximize the energy function

$$E(\mathcal{C}) = a|\mathcal{C}| + b|\mathcal{C} * \mathcal{C}| + c|(\mathcal{C} * \mathcal{C}) * (\mathcal{C} * \mathcal{C})|$$

Where the first term <sup>4</sup> is proportional to the distance of elements from the center, the second distance from each other, and the third distance of UV points from each other. This formalism has the useful property of defining a parameter space for various optimizations. For example, when b dominates a and b one gets a centrally condensed uv distribution at the expense of uv plane uniqueness. Here we have chosen to minimize the redundancy of the uv points by selecting a relatively large and positive c. The resulting configuration, uv distribution and psf are shown in Figures 2.3, 2.5, and 2.6 respectively. The typical dynamic range, the ratio of psf primary to brightest secondary lobe, of this configuration is 100:1, making it quite suitable for imaging. For EoR however, this configuration is only optimal if the SNR for each uv pixel is greater than 1 (Halverson, 2002). Recent SNR calculations indicate that this is not the case and in fact a large and *negative* c, a maximal uv redundancy configuration is most likely to yield a power spectrum detection (Parsons et al., 2011).

<sup>4</sup> $\mathcal{C}$  is a matrix with ones at the rows and columns corresponding to positions of the antennae. The autocorrelation  $\mathcal{C} * \mathcal{C}$  is then the matrix corresponding to the uv distribution. The “length” operation  $|\mathcal{M}|$  is defined as the average distance of all the points from the center,  $|\mathcal{M}| = \sum_{ij} \sqrt{x_i^2 + x_j^2} M_{ij} / \sum_{ij} M_{ij}$ .

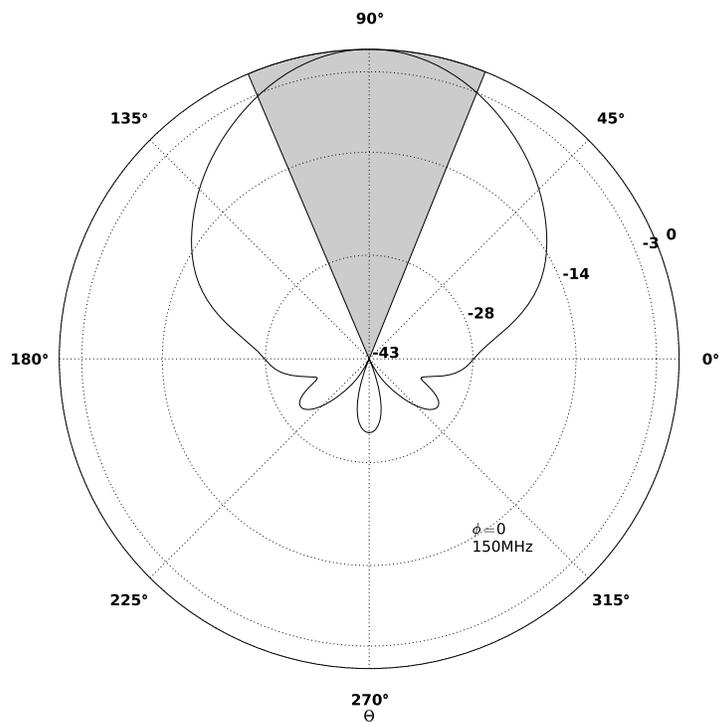


Figure 2.3: A perpendicular (the dipole is into the page) cut through a model of the PAPER primary beam at 150MHz. Note the 45°FWHM Field of View (grey) and sub-horizon nulls. At that level all angle cuts are identical. The region most likely to vary with angle and frequency, the first null, occurs below the horizon.

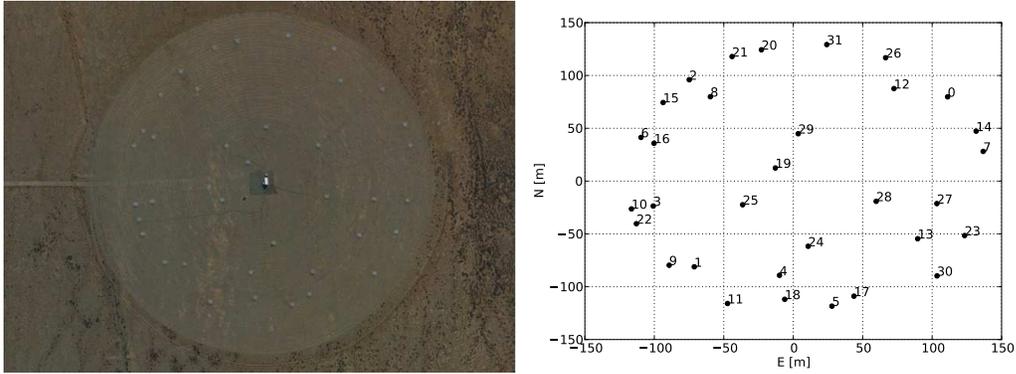


Figure 2.4: Aerial photo (left) and diagram (right) of the 32 element array in South African Karoo. The elements are arranged in a minimum redundancy configuration optimized to maximize uniqueness of  $uv$  spacings as well as distance from the central container. Several extra stations used for other experiments are also visible. Signals from the antennae are transmitted via television coaxial cable to the central electronics hut which is a shipping container fitted to be RFI tight to 110 dB by Commtest.<sup>a</sup> Similar containers are being used by LOFAR and MeerKat. The road leads west to the nearby KAT7 test array.

<sup>a</sup>[http://www.lofar.org/wiki/lib/exe/fetch.php?media=public2008-06-10\\_doc\\_specs\\_lofar\\_container.pdf](http://www.lofar.org/wiki/lib/exe/fetch.php?media=public2008-06-10_doc_specs_lofar_container.pdf)

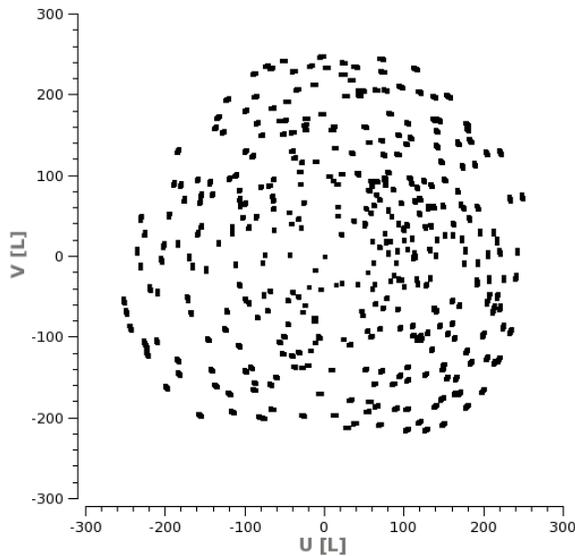


Figure 2.5: The  $uv$  distribution of the PSA32 imaging configuration shown in Fig. 2.4. The configuration is optimized to sample as many unique  $uv$  cells as possible. This is optimal for imaging and sky-based calibration but is not for EoR detection unless each  $uv$  cell can achieve  $\text{SNR} \sim 1$ . For EoR the array might be arranged in a grid, to put as many samples as possible into a few  $uv$  cells.

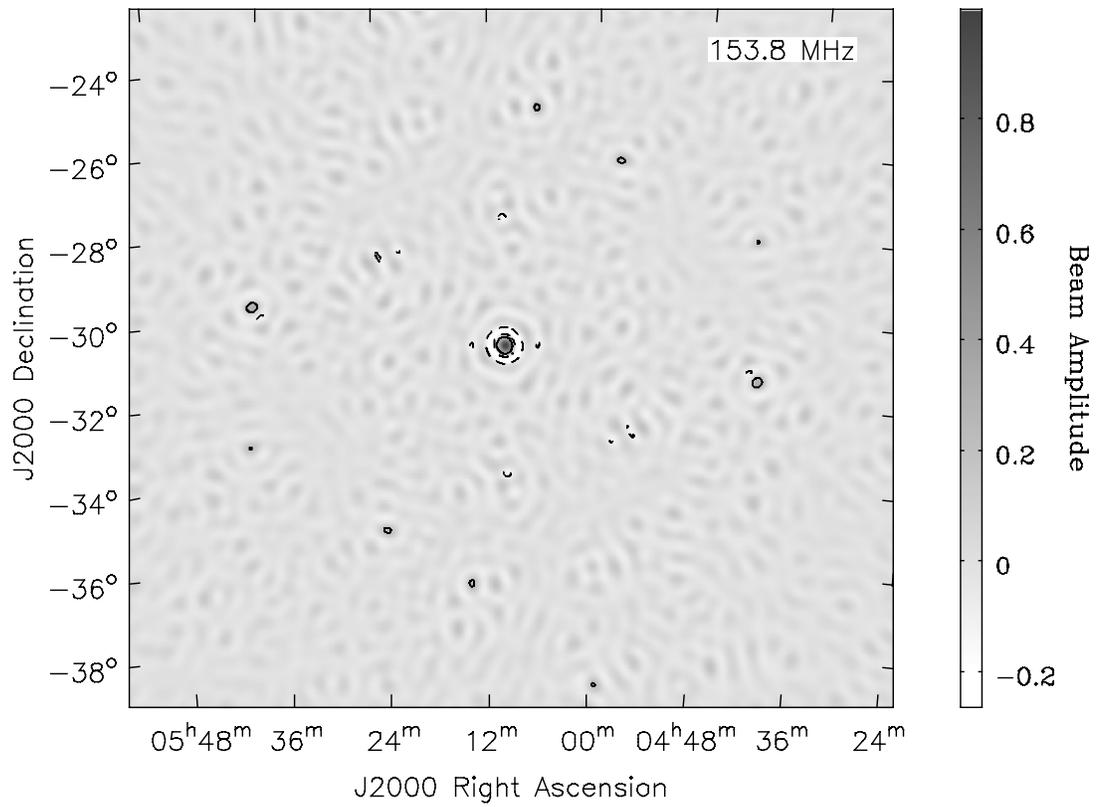


Figure 2.6: The zenith point-spread-function (psf) of the PSA32 dirty beam. Contours are -0.2 (dashed) and 0.2 (solid). Radial weights, increasing linearly with distance from the center of the  $uv$  plane, have been applied to emphasize the primary lobe. Points with radial  $uv$  distance  $< 25\lambda$  are not included here. After 10 minutes and 6MHz bandwidth the highest secondary sidelobe, is at the 30% level.

Given the high quality of the RFI environment at the Karoo site (see 2.7), RFI-flagging was limited to flagging of a few satellite bands, as well as any visibilities with amplitudes  $2\sigma$  above the mean amplitude as in PGB8 (see 2.8). Less than one percent of the data are flagged in this way.

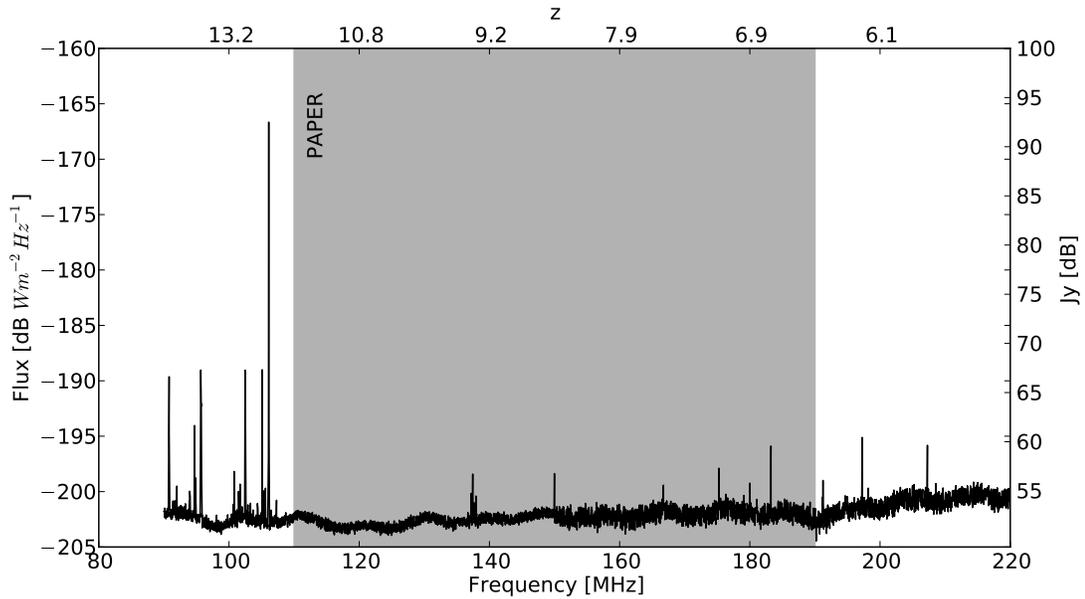


Figure 2.7: Spectrum from an initial survey of the South Africa Karoo site. The spectrum suffers little interference save those features omnipresent on earth: FM radio between 87.5 and 108MHz, airplanes at 125.3MHz, Orbcomm satellite constellation at 137MHz and amateur radio satellites at 150MHz. The uneven baseline and change in instrument noise at 150MHz are instrumental errors. (of South Africa, 2005)

## 2.3 Data Processing

Data have been reduced in two parallel pipelines throughout much of the foregoing. The work described in §3 was done with a custom pipeline built on the Astronomical Interferometry in PYthon (AIPY) package. Other parts (§4 and §5) were completed in the Common Astronomy Software Applications (CASA) environment<sup>5</sup>. Table 2.3 provides a rough guide to the differences between the AIPY and CASA environments. In most cases these tools provide orthogonal feature sets and are in fact most powerful when combined in the python environment, then when considered separately.

<sup>5</sup><http://casa.nrao.edu/><sup>6</sup>

<sup>6</sup>The National Radio Astronomy Observatory is a facility of the National Science Foundation operated under cooperative agreement by Associated Universities, Inc.

Table 2.2. PSA32 Observations

	Dates [UT 2010]	Length [hrs]	LST range
Catalog (§3)	May 19 13:11 - May 20 04:50	15	06h00m - 16h00m
	Sep 15 16:48 - Sep 16 04:04	12	16h00m - 06h00m
Calibration (§4) Sensitivity (§5)	Sep 18 03:51 - Sep 29 04:14	38.5	03h45m - 07h15m

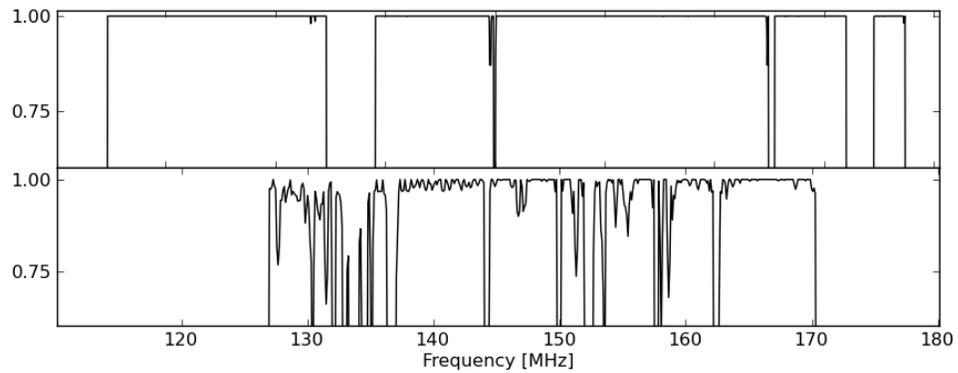


Figure 2.8: Fraction of data remaining after RFI flagging (manual flagging of satellite bands and automatic flagging of amplitudes above  $2\sigma$  in a 10 minute period. South Africa site (top) in comparison with the test array at NRAO's radio quiet Green Bank, WV site (bottom) See Parsons et al. (2010) (hereafter PGB8) for details about the Green Bank array.

Standard methods common to all analyses are RFI and artifact flagging. RFI is flagged by removing points beyond  $n\sigma$  where  $\sigma$  is calculated for each baseline over all channels and many integrations (usually about 10 minutes). In several cases artifacts due to errors in the correlator were flagged either manually or manually as the case could be generalized. Beyond these initial steps, the main elements in both pipelines are: modeling, calibration, and turning correlations into images. As an optimum pipeline is still under development and different elements of this work used different methods (exploration of this subject is a dominant theme of this work) we will elaborate on the details as needed.

Table 2.3. A comparison between AIPY and CASA

Feature	AIPY	CASA
Imaging Deconvolution	Image plane	Cotton-Schwab
FoV	<b>full 180°</b>	<i>80° limit<sup>a</sup></i>
W-projection	Supported	Supported
Faceting	<i>Not supported</i>	<b>Supported</b>
MFS <sup>b</sup>	Linear only	Non-linear
Calibration	Fitting	Least-Squares
Time Dependence	<i>Not supported</i>	<b>Supported</b>
Speed	<i>Slower</i>	<b>Faster</b>
Antenna Positions	<b>Supported</b>	<i>Not supported</i>
Assumed parameters	amplitude+delay	full complex bandpass
Flagging	Statistical/Manual	Statistical/Manual
Plotting	<i>Scripted</i>	<b>Full Gui</b>
Gridding	<b>scriptable</b>	<i>not scriptable</i>
Access to internals	<b>Complete</b>	<i>Very limited</i>
HealPix	<b>Built in support</b>	<i>Not supported</i>
Catalogs	<b>Fully supported</b>	<i>Not supported</i>
Modeling	Partially supported	Completely supported
Catalog sources	Completely supported	Mostly supported
Image	<i>Not supported</i>	<b>Completely supported</b>
Use model	Python module	Complete environment

<sup>a</sup>This limitation is in the CASA software architecture and is not a fundamental algorithmic problem. As of this writing (2011) the issue remains open.

<sup>b</sup>Multi-Frequency Synthesis. Linear MFS ignores the possibility of spectral variability and uses each different channel as a different sample in the same  $uv$  plane, sometimes drastically increasing the  $uv$  coverage. Non-linear MFS adds the possibility of spectral variation by solving for multiple Taylor coefficient images Rau (2010).

# Chapter 3

## Foregrounds

### 3.1 Survey of past measurements

Several recent attempts have been made to synthesize what is known about the radio sky from existing surveys. de Oliveira-Costa et al. (2008) compile measurements from 10 MHz to 90 GHz and uses them to create a global sky model of extended emission (sizes greater than two degrees), which at 150MHz is dominated by the all-sky 408-MHz map by Haslam et al. (1982). Discrete sources are listed in many catalogs. Those used for comparison here are listed in Table 3.1. Meta-catalogs such as a small but detailed compilation of bright sources by Helmboldt (2008) and the large SPECFIND meta-survey by Vollmer et al. (2010) are of particular interest.

The literature grows increasingly sparse at lower frequencies and in the southern hemisphere; indeed, *no* blind survey has been reported below 408 MHz in the south. The SPECFIND cross-identification catalog identifies 6000 unique sources between 100 and 200 MHz with  $\delta > 0^\circ$ , but fewer than 1000 low frequency sources with declination  $\delta < 0^\circ$ . The best information in the south ( $\delta < -30^\circ$ ) below 408 MHz is provided by the Molonglo 4 Jy Survey (MS4; Burgess & Hunstead (2006)), which uses the Molonglo Reference Catalog (MRC) (complete to 1 Jy), Culgoora observations by Slee (1995) at 160 MHz and 80 MHz, and new Molonglo observations at 408 and 800 MHz to estimate the 178-MHz flux of bright sources in a sample similar to the northern 3CRR survey (Laing et al., 1983). Since most blind southern sky surveys have been done at wavelengths shorter than 30cm (NVSS, SUMSS, Parkes —see Table 3.1), construction of a sky model at lower frequencies must be an extrapolation based on assumptions of spectral shape and our ability to account for instrumental differences.

### 3.2 Constructing a Sky Model

A good sky model is crucial for calibration and modeling EoR foregrounds. To be of any use during actual observations it must be both *accurate* and *complete*. Each source it

Table 3.1. Low frequency surveys

Name	Res	Freq [Mhz]	Dec Limits	Flux Limit [Jy]	Ref
MRC	2'	408	18.5 > $d$ > -85	1	Large et al. (1981)
MS4	2'	178 <sup>b</sup> ,408	-30 > $d$ > -85	4	Burgess & Hunstead (2006)
Culgoora	1.85',3.7'	160,80	32 > $d$ > -50	2	Slee (1995)
3CR(R) <sup>a</sup>	6'	178	75 > $d$ > -50	5	Bennett (1962) (Laing et al., 1983)
6C <sup>a</sup>	4.2'	151	> 30	0.3	Baldwin et al. (1985)
7C <sup>a</sup>	1.17'	151	> 26	0.2	Gower et al. (1967)
VLSS <sup>a</sup>	80''	74	> -30	0.4	Lane et al. (2008)
NVSS <sup>a</sup>	45''	1400	> -40	2.5e-3	Condon et al. (1998)
SUMSS	45''	843	< -30	6x10 <sup>-3</sup>	Mauch et al. (2003)
Parkes <sup>c</sup>	90'	178	> -10	?	Wright & Otrupcek (1990)
PAPER	26'	145	10 >	10	Jacobs et al. (2011)

<sup>a</sup> Sources included via the SPECFIND meta-catalog (Vollmer et al. 2010) .

<sup>b</sup>178MHz fluxes in MS4 are estimates based on MRC, Culgoora and other measurements.

<sup>c</sup>This survey was performed originally at 408 and 2700MHz. 20% of the sources were later observed at 178MHz, though only 13 sources (0.2% of the entire survey) were observed at 178MHz at declinations  $\delta < -10^\circ$ . The resolution of these measurements is unclear. Here we give the diffraction limit of the Parkes dish at 178MHz.

catalogs must be an accurate prediction of the flux on the sky at the observed bandwidth and resolution. Second it must be complete; it must not miss anything or include false positives. As we will see, sometimes even bright sources violate assumptions made during cataloging and are thus left out of the sample. Even if errors are avoided, a catalog is only truly complete at one frequency. Extrapolation in frequency without a good understanding of the spectrum will introduce a bias away from samples that change unexpectedly in frequency. As we will see, unexpected spectral behavior can happen for any number of reasons.

Many of the problems inherent in building a sky model can be seen simply by comparing between catalogs. Catalog comparisons are hindered both by spectral/temporal variation of sources and by differences in the angular resolution of observations. Spectra at low frequencies tend to be dominated by non-thermal sources like synchrotron emission. While this emission tends to be well-described by a power-law<sup>1</sup> in frequency, at lower frequencies several effects complicate comparisons between bands. Chief among these is synchrotron self-absorption, which is most often in evidence at the lowest frequencies, but may also produce spectral features between 100 and 200 MHz (Helmboldt et al., 2008). Another complication arises from sources not present in the deep high-frequency catalogs appearing at lower frequencies due to their exceptionally steep spectral indices. A source steep enough to appear in these first PAPER images ( $> 10Jy$ ) and be absent in the MRC sample ( $> 1Jy$ ) would need a very steep spectral index  $< -2.2$ . Out of over 70000 sources in NVSS only two or three sources are known to have spectral indices this steep (van Weeren et al., 2009). The 4Jy sample on the other hand would only include objects with spectral index of -0.92 or shallower and thus be incomplete and biased towards shallow spectrum sources. For these reasons, it is desirable to use reliable spectral slopes (when available) to estimate flux in the band of interest which might be far from the originally measured band.

Gathering of spectral slopes and assessing reliability is one aspect of the more general problem of building a sky model at a given frequency from many measurements by different instruments at different frequencies. Spectral variation can be intrinsic, but it can also be introduced by instrumental effects. Higher resolution will not only distinguish between sources that coarser observations would conflate (source confusion) but will also be able to resolve sources that coarser surveys would have attributed to an unresolved background (confusion background). This latter effect not only drastically increases the source number density, but also reduces the overall upward bias from the unresolved background. In images with high side-lobes flux measurements of nearby sources could be biased either up or down depending on the phase of the side-lobe (side-lobe confusion). Finally as most interferometers are sensitive to a limited range of size scales, higher resolution instruments will also "resolve out" structure on larger scales. All of these effects must be taken into consideration when comparing two catalogs at different frequencies and resolutions.

In principle the differences between catalogs might become tractable given enough overlapping measurements with a variety of instruments, however the paucity of low fre-

---

<sup>1</sup> $S(\nu) = S_0 \left(\frac{\nu}{\nu_0}\right)^\alpha$  where the flux ( $S$ ) at frequency ( $\nu$ ) scales as the power ( $\alpha$ ) of the frequency.

quency southern hemisphere data makes a detailed study of the accuracy of these catalogs difficult. We can study the some of the pitfalls of catalog comparison by proxy by comparing the far more numerous Northern hemisphere catalogs.

Helmboldt et al. (2008) have compiled a large number of measurements over the entire radio band for 386 bright 74MHz VLSS sources and provided spectral fits, suitable for interpolation and extrapolation. Among all the data collected, a few catalogs rate status as "anchor" data points, they are the only measurement used in their band. A key anchor point is the NVSS at 1400-MHz. Several (25 or 6%) of the VLSS sources were given a single flux measurement, but noted to have multiple components. In order to build a sky model at 1400-MHz one needs some prescription for performing an effective smoothing to match resolutions. In this case from 40" of NVSS to 80" of VLSS. To do this Helmboldt et al searched NVSS separately for each VLSS component and summed the results. This works at higher frequencies if each component is resolved and well above the noise. Note that this is not a correction for source confusion. Apparent conjunction of two bright sources at this level is statistically improbable. In NVSS there is less than one source brighter than 1Jy per 8000 VLSS beams ( $1\text{sq}^\circ$ ).<sup>2</sup> Thus most confused multiple component sources would be resolved sources, such as compact supernova remnants or double lobed jets. Should flux be evenly distributed the total flux increase would be 50-100%. This method also has the potential for making large errors. If each component does not resolve into a single source at the higher frequency, the search for that dim confused source could return the bright primary again. Consider a two component source composed of a steep spectrum component and a shallow, each having a similar amplitude at 74MHz but a large difference in amplitude at 1400MHz. Both component searches will find the same bright source and estimate the 1400 low resolution flux at exactly two times its actual value. The same outcome will also result from the VLSS catalog identifying side-lobes or distorted psf as multiple components.

We can evaluate this method by comparing Helmboldt's summed 1400MHz fluxes with the actual NVSS fluxes (Fig. 3.1). To make this comparison we have searched for each NVSS source within 40" of the VLSS source in the Helmboldt catalog and compared the listed fluxes. We can see several cases where this operation increases the NVSS flux by 100% or more (the dashed line). At least two of the errors (J1339+385 and J1843+795) are a clear doubling, the increase is exactly a factor of 2, as if the measurements have been included twice. Others are not so clear; of the 25 multiple component sources 65% have received corrections >10%. Sources with only one component are not affected.

Over the entire catalog the error is small (affecting less than 1% of the sources). In each source, however, the error is large and high on the flux scale (100%). After a more detailed check (as we have done here) the errors could be corrected but this method does not scale well to inclusion of many catalogs. A typical strategy is to exclude, or "flag", all sources having more than one component or high spectral rms, or known to be variable;

---

<sup>2</sup>In making comparisons between bands like this one needs to mentally scale by the inverse ratio of frequencies, 1400MHz sources will in general be about 10 times dimmer than 150MHz sources. A 1Jy NVSS source would most likely be a 10Jy source for PAPER and a 20Jy source in VLSS. A dimmish but distinguishable object visible even with the limited methods used in this section.

whatever the problematic category may be. Flagging is a standard approach to the problem of errant data, for example RFI in radio data or cosmic rays in CCD images but implicitly assumes that more repeated samples will fill the gap. When there is only one catalog at that band, the gap remains. This kind of random incompleteness occurring at even high flux levels is essentially an error in our model of the sky of the same amplitude (100%) as a doubling. These kinds of errors propagate when a catalog with sources flagged is used as finding survey at a new band (eg the Culgoora survey) or in a joint solution for a spectral model (eg SPECFIND).

The only way to make a complete and accurate catalog from old data at a new resolution is to generate a new catalog from the original images, smoothed to the resolution of interest. Given the data volumes and limited availability of original images this is a tremendous task. In our case it is simpler to simply observe using our own instrument.

### 3.3 New 145-MHz measurements of Southern sky sources.

In this Section, we present new flux measurements of 480 sources at 145 MHz using PAPER in the southern hemisphere. These measurements cover the largest area of the southern sky yet surveyed in the EoR band.

#### 3.3.1 Observations and Data Reduction

During May and September 2010, we recorded commissioning data with PSA32 in two separate campaigns. Between the May and September data-taking, a number of small improvements to the correlator were made, but all other hardware remained unchanged. The data presented here are from May 19 and Sep 15 2010 (see Table 2.2), both periods being predominantly between sunset and sunrise. Only the linear EW dipoles of each antenna were correlated. Visibilities were integrated and recorded every 5.37 seconds. The frequency resolution was 96 kHz in May and 45 kHz in September. The separation in LST between the two observing epochs, along with PAPER’s wide primary beam, make these two observations sufficient to map the entire sky below  $\delta < 10^\circ$ .

Data editing, calibration and imaging were performed using an AIPY based pipeline (see §2.3). The first analysis step was to obtain a phase calibration. Because of the wide field-of-view (FOV) of the antennas, the data are dominated by bright sources that are sometimes far from the zenith. During the May observation, the brightest source visible was Cen A<sup>3</sup>, while the brightest source during September observations was Pic A. These sources are bright enough to perform single-baseline fringe fitting. Data observed within ten minutes of the transit of these sources were used to derive a phase calibration by fitting a time and frequency visibility model to the data using a conjugate-gradient solver (Parsons et al., 2010)<sup>4</sup>. Phase terms in the calibration are dominated by cable and correlator delays;

---

<sup>3</sup>While Cen A is resolved, the central point source dominates the smooth structure by several orders of magnitude.

<sup>4</sup>Hereafter referred to as PGB8

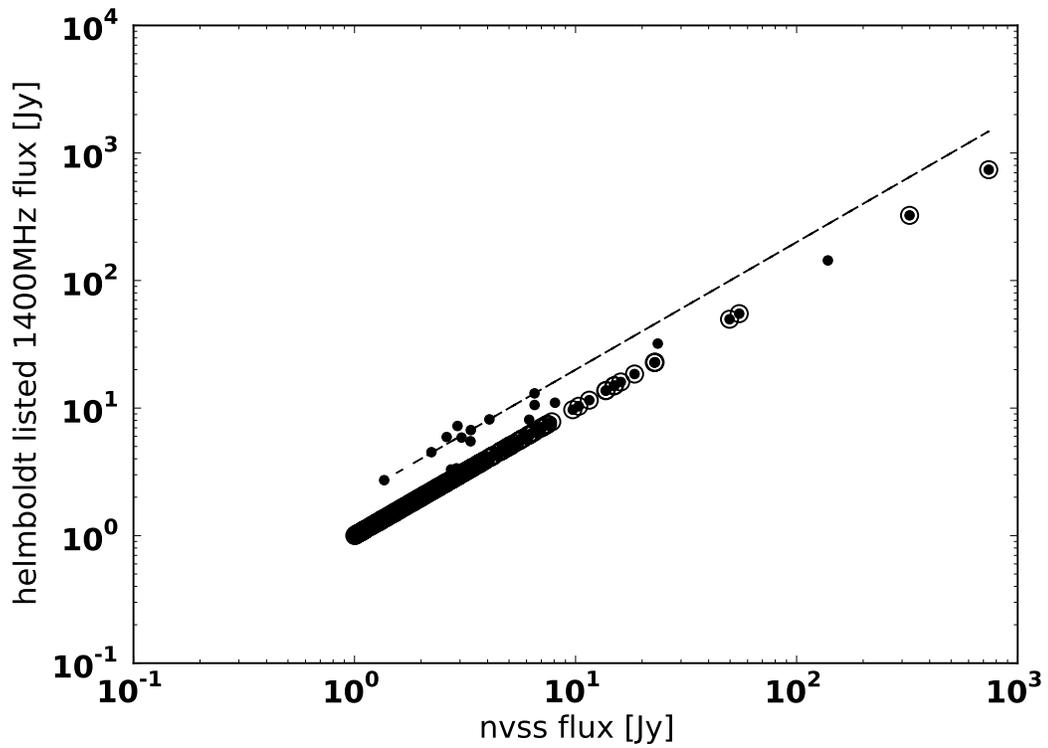


Figure 3.1: A comparison between 1400 MHz "anchor points" in the Helmboldt compilation and the NVSS flux (dots). The Helmboldt fluxes are an attempt to smooth the NVSS 40' resolution catalog to match the 80' VLSS resolution. Some of the differences appear to be exact doublings over the catalog values which is suggestive of an error in component matching. A subset of the sources noted by NVSS to have a single component (circles) has no such errors. In general a catalog may emphasize completeness or accuracy. Achieving both simultaneously is both difficult to achieve and to verify.

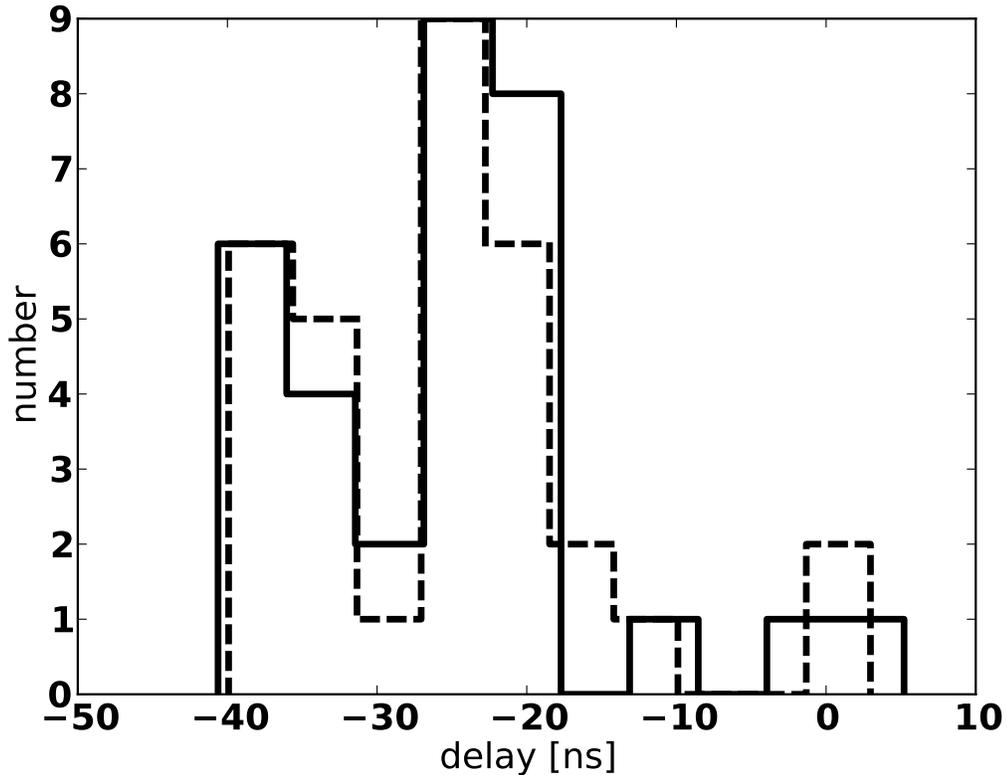


Figure 3.2: Distribution of delay solution used for PSA32 catalog. The May solution [5h<RA<12h] (dashed) was used as the starting point for the September solution [12h<RA<5h] (solid). Only small changes were necessary over this 4 month period.

these have been found to be quite stable. Thus for the analysis here, the phase calibration (shown in Fig 3.2) derived from these two 10-minute observations is applied to each night's entire observation.

If the source is not carefully removed, the phase-calibrator sidelobes severely limit imaging dynamic range. To do this we use an efficient source-removal technique that filters data by removing the corresponding region of delay/delay-rate (DDR) space from each baseline (Parsons & Backer, 2009). This technique filters a source from each baseline by nulling data having a delay and fringe-rate corresponding to the desired sky location. The net effect is to remove a large fraction of the filtered source *without* having to construct a multi-component image-domain model. For the May data, the point-source component of Cen A (estimated flux  $\sim 5000$  Jy) is filtered; for September, we have removed Pic A and For A (400 and 150 Jys, respectively).

In this quiet environment, instrumental effects became dominant. A troublesome instrumental effect in many interferometric instruments is common-mode interference, inter-

fering signals common to two or more inputs and sky signals crossing antenna boundaries within the analog system. Both of these kinds of cross-talk are removed by subtracting a 4 hour long running average as described in PGB8 and investigated in detail in §5.4.

Map-making is done in three stages: snapshot-imaging, mosaicking each night and finally summing into a single calibrated map.

Images are made in 10-minute zenith-phased “snapshots”; this is a sufficiently short time that the affect (change in apparent flux) of sources moving through the primary beam is negligible as is the  $w$  component of the baseline (Fig. 1.5). Visibility data are gridded into the  $uv$ -plane using linear multi-frequency synthesis (G. B. Taylor, C. L. Carilli, & R. A. Perley, 1999) and  $w$ -projection (Cornwell et al., 2008). To this  $uv$ -gridded data we apply radial weighting—increasing radially in the  $uv$ -plane—to emphasize point sources. Gridded data are Fourier transformed to produce a snapshot image  $114^\circ$  wide, with an effective synthesized beam width of  $26'$ . These facets are then deconvolved in the image plane by the dirty beam (Fig. 2.6) using the Högbom CLEAN algorithm (Högbom, 1974). Image-domain deconvolution is limited in its ability to reconstruct the flux, particularly in the wide-field case (Rau et al., 2009). Thus, the CLEANing is not fully effective, and the images contain artifacts from the side-lobes of sources far from the phase center.

Some snapshots at the very beginning or end of a run are contaminated by strong side-lobes from the rising or setting Sun. Two or three ( $\approx 5\%$  of the total) are flagged in each run. The rest of the snapshots made over the course of a night are divided by a model of the primary beam (see Fig. 2.3) and averaged onto a HEALPIX grid (Górski et al., 2005) with  $7'$  pixels (NSIDE=512), to create two maps — one for each epoch. A typical pixel has weighted contributions from approximately 2 snapshots. In the mosaic average each contribution is weighted by the square of the primary beam model, reflecting our lack of confidence in deconvolution far from the beam center. The total weights are also a good model of the survey coverage (Fig. 3.3) and a good estimate for the total number of samples in each LST bin (Fig. 3.4). Here we can see the effect of slightly shorter snapshots in the September (LSTs 16h-6h) set increasing the relative sample density as well as the effect of flagging facets visible most clearly near LST= 6.

Each map is flux-calibrated to a bright source near the phase-calibrator using a flux taken from the Culgoora catalog. The May map is flux-calibrated to 1422-297 at 21 Jy and the September map is calibrated to 0521-365 at 72 Jy. Both sources chosen for brightness, proximity to zenith as well as nearness to that night’s phase calibrator. Once on the same flux-scale, the two epoch maps are summed together into a single map, weighted by the number of snapshot contributions. The final product covers 36000 square degrees at  $\delta < 10^\circ$ , with an effective resolution of  $26'$ . Though dynamic range varies across the image, sources brighter than 10Jys are consistently distinguishable. We will limit our analysis here to sources likely to be above this 10Jy side-lobe confusion limit.

The limitations of these reduction steps, as well as instrumental artifacts, impact image fidelity. Final images include residual cross-talk and errors due to delay-filtering, which necessarily removes flux from multiple points on the map. The absence of time-dependent calibration, the limitation to image-plane deconvolution and uncertainty in the beam model

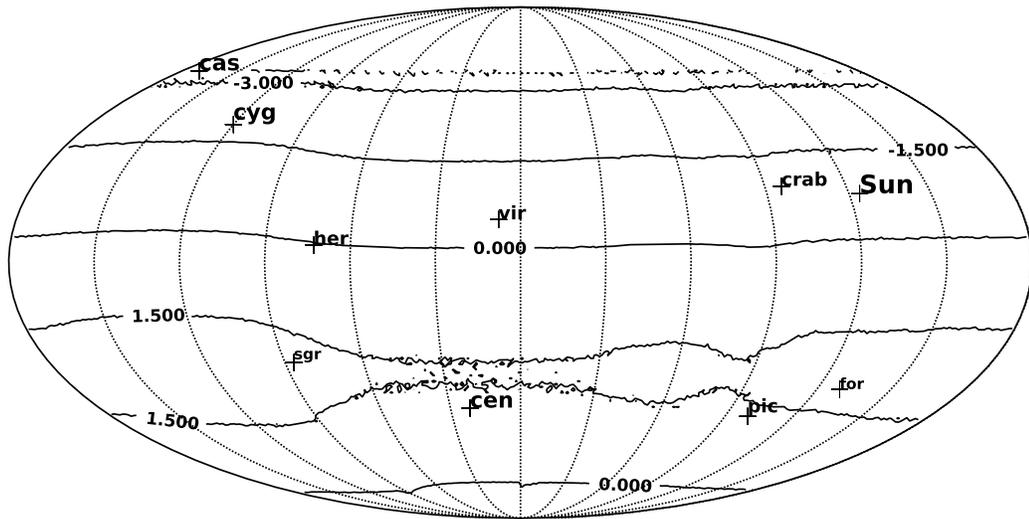


Figure 3.3: During the averaging process many snapshots were summed weighted by the square of the primary beam model and then divided by the total weight. This total weight also provides a visual map of survey coverage.. Right Ascension begins with 0 on the right, increasing to the left. The contours are units of total beam-model squared with unity occurring at zenith pointing towards  $\delta = -30.7^\circ$ , where the weights are also the number of included snapshots (see Fig. 3.4)

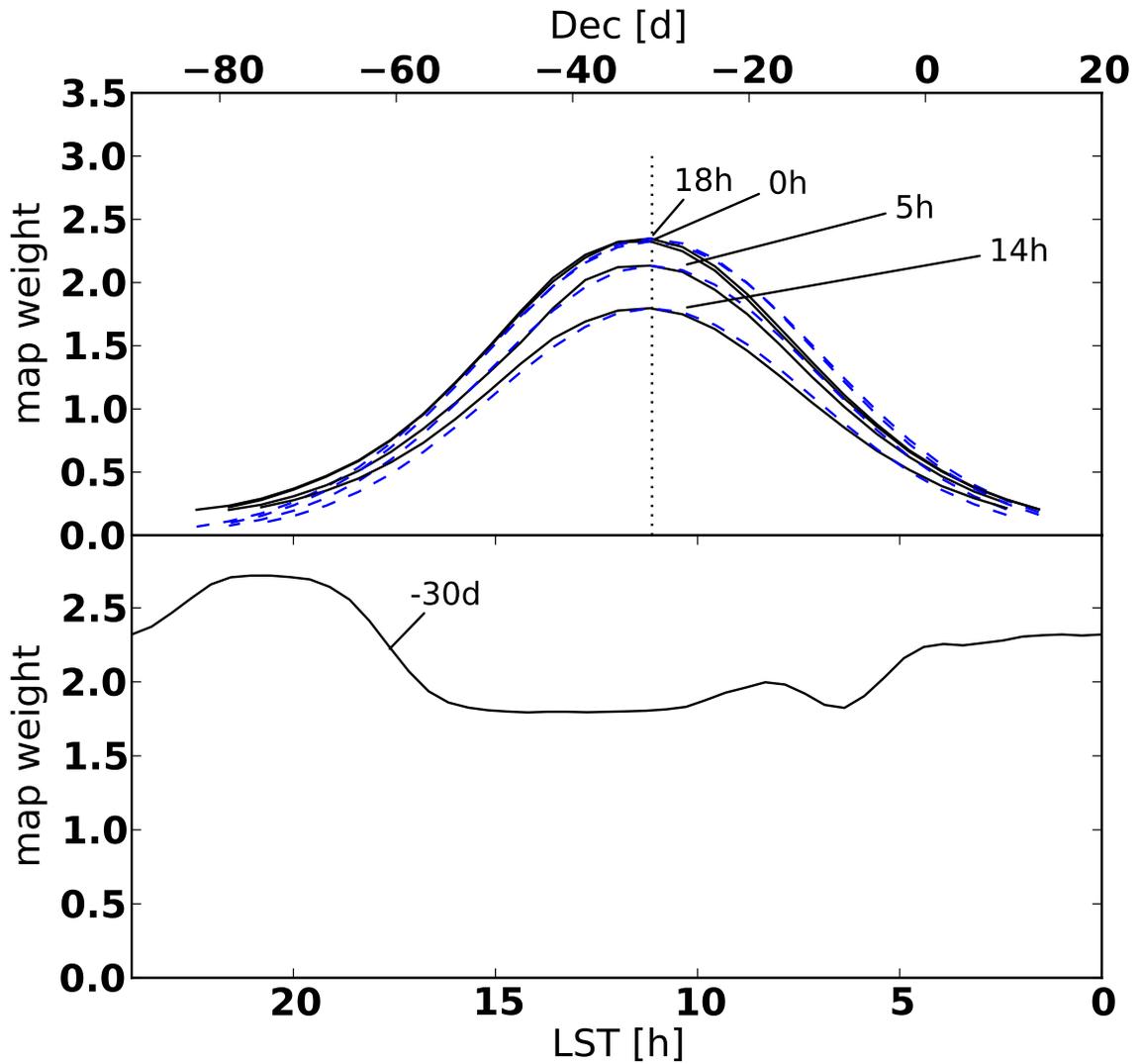


Figure 3.4: Longitudinal and latitudinal slices through the survey coverage map shown in Fig 3.3. The longitudinal slices (top) show the declination coverage, the latitude= $-30^\circ$  (bottom) slice shows the number of snapshots included in each LST bin. The snapshot length differed slightly between the two data periods resulting in a slightly shorter snapshot period and an increased number of samples per LST during the 16h-6h period with one dip at 5h where three snapshots were flagged due to high rms.

Table 3.2. PAPER fluxes for 480 MRC sources and matching Culgoora fluxes (where available)

Ra	Dec	Name	S145	rms	MRC_sep	Cul	S(160)	S(80)	SpIndex
0.60	-83.14	0003-833	18.9	4.7	0.17				
0.88	-17.50	0000-177	11.6	1.2	0.11	0000-177	11.8	22	-0.9
1.37	-56.54	0003-567	12.0	2.2	0.17				
1.52	-42.61	0003-428	9.6	1.3	0.16	0003-428	11.9	11	0.11
1.58	-0.07	0003-003	25.5	2.6	0.12	0003-003	16.8	27	-0.68
2.11	-6.05	0005-062	11.2	1.8	0.23	0005-062	6.9	10	-0.54
2.51	-44.50	0007-446	14.3	1.2	0.16	0007-446	17.0	26	-0.61
3.27	-42.11	0008-421	1.5	1.4	0.41				

Note. — PAPER Southern Sky catalog generated by searching for sources in the Molonglo Reference Catalog above 4 Jy and below  $+10^\circ$  Declination. Beginning on the left, the columns list: Right Ascension and Declination in degrees, MRC name, calibrated PAPER flux [Jy], rms around source [Jy] and angular separation in degrees from MRC location. Included for reference are Culgoora 160MHz, 80MHz and spectral indices fluxes where available. Complete table available in the online edition of ApJL (Jacobs et al., 2011) and at <http://arxiv.org/abs/1105.1367>.

also affect the accuracy of the map.

Successful future work in foreground mapping and EoR detection will depend on our ability to rank the relative importance of these issues. This is true not only within the PAPER project, but also between similar projects. For these reasons we establish an accuracy baseline by measuring and comparing the fluxes of many sources to catalog values.

### 3.3.2 Catalog Construction and Flux Calibration

We have used the entire sky below  $\delta < 10^\circ$  to find fluxes corresponding to 480 MRC sources above 4 Jy — selection criteria similar to those used by Burgess & Hunstead (2006) to generate the MS4 sample. This sample is intended here to construct a sample with an approximate 150-MHz lower limit of 10Jys, the sidelobe confusion limit of our map while acknowledging that we will be limited to flatter spectrum ( $\alpha > -0.92$ ) sources, leaving any steeper sources unmeasured.

The PAPER flux is identified as the brightest pixel within  $30'$  of the MRC source. They are listed along with separation distances and local RMS in Table 3.2; 90% of the sources identified are within one beam-width (see Figure 3.10). In the following we will explore the relative completeness and accuracy of this catalog.

The accuracy of the PAPER image and of these fluxes can be evaluated both superfi-

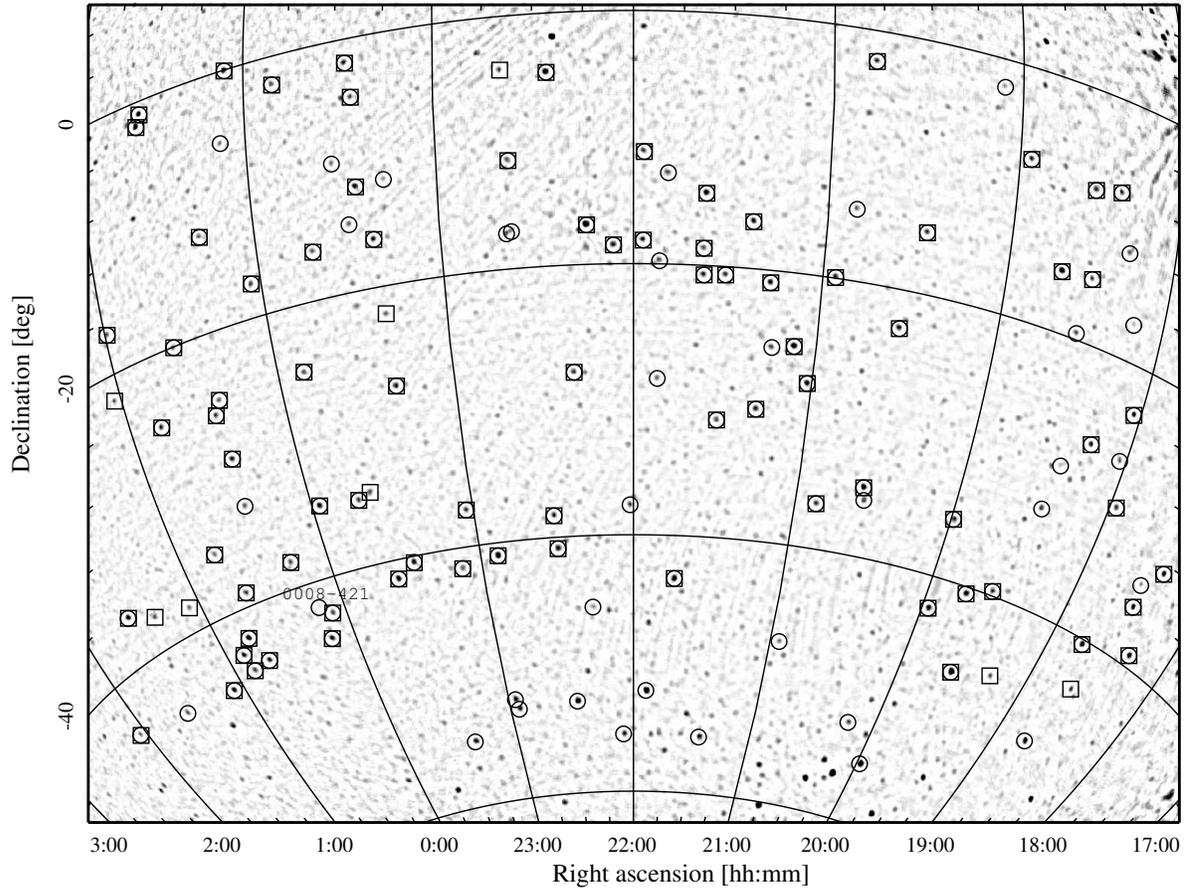


Figure 3.5: A low foreground region of the PSA32 field, centered at  $23^{\text{h}} -31^{\circ}$ . Pixels above 99% of the flux scale, approximately 1 Jy, are shown in black. MRC sources above 4 Jy are shown in circles. PAPER fluxes for these sources are given in Table 3.2. All MRC markers, save one, have a corresponding source at this flux level. The 160MHz Culgoora survey, used to evaluate the flux scale, is shown in squares. This image has an effective integration time of 30 minutes, a bandwidth of 70MHz, a field rms of 0.4 Jy, and a peak to field dynamic range of 120.

cially in the image plane and numerically by comparing to past measurements. By comparing the MRC catalog directly to the image we can ascertain the relative completeness of the MRC sample. In Figure 3.5 we overlay MRC markers from our 4-Jy subsample onto a 4800 square-degree sub-image and see that at the 4-Jy flux-level MRC is not one-to-one but does agree with the map on all but one source. As shown in Figure 3.6 this source manifests a rare high-frequency turn-over. In the case of the Culgoora catalog, there are no such differences. In places where Culgoora shows a bright source that is not in the 408 MHz sample, PAPER also finds a bright source. Sources as dim as 10Jy do not have a matching MRC marker. Together these facts suggest that these sources are steeper than  $\alpha \approx -0.9$  and were excluded by our 4Jy MRC cut, as expected. Thus our complete flux-limited sample of MRC at 408MHz becomes incomplete at 145MHz. Given the apparently large fraction of steep spectrum sources we should note that lowering the 408-MHz flux limit would give us the opportunity to add these apparently bright 150-MHz sources to our

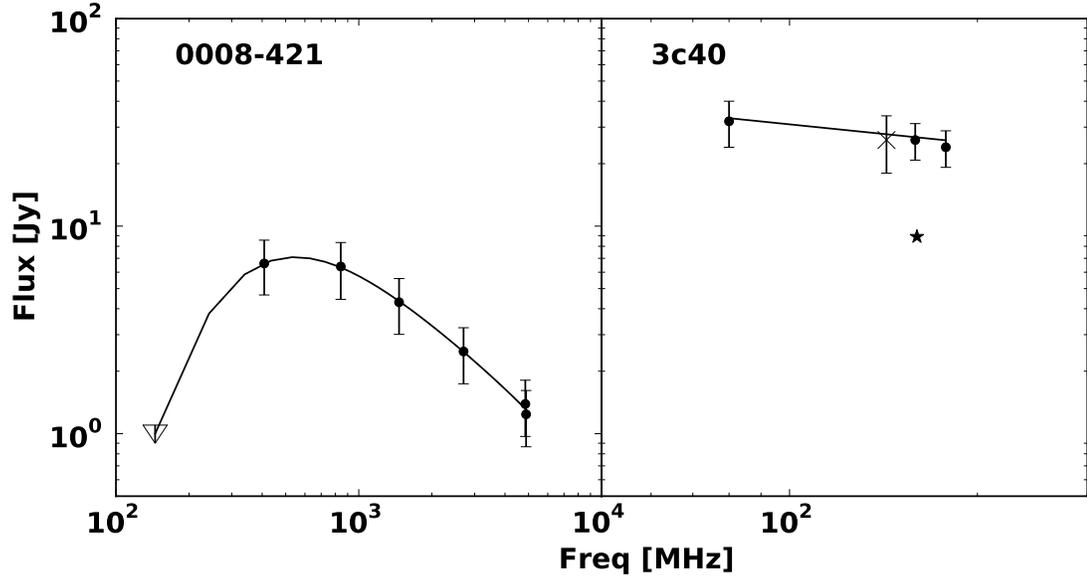


Figure 3.6: Left: Spectrum of 0008-421, the only MRC source in Figure 3.5 without a PAPER counterpart. A rare example of spectral turnover at around 500 MHz. Right: Spectrum of 3c40 (0123-016) with an 'x' for PAPER flux and star for Culgoora 160-MHz. All circles taken from SPECFIND.

catalog, at the expense of adding a large number of sources below or 10Jy detection limit.

To assess the accuracy of the flux measurements we compare with 332 sources also found in the Culgoora catalog and 225 found in MS4. The accuracy of the PAPER measurements will be limited by the image dynamic range, as discussed above, as well as the relatively wide bandwidth of the PAPER correlator. However these errors must be set against the error in the catalog comparison. As discussed above, the presence of multiple components or extended structure in sources hampers comparisons between observations with different resolutions, while the presence of self-absorption or other spectral structure impedes comparison between catalogs generated at different frequencies. To set the scale of these effects we inter-compare several catalogs with measurements near the PAPER band.

A simple comparison metric is the per-source flux-scale; the ratio of fluxes between two sets. Accounting for spectral slope, the flux-scale would have a nominal value of one, with a certain amount of spread encompassing all the sources of error in flux determination and catalog comparison. The assumption here is that flux-scale is a reasonable estimate of instrument and reduction precision, but also reflects spectral or confusion errors in catalog comparison. We see that flux scale depends weakly on spectral index, and local image rms (Fig. 3.7), but is particularly useful in identifying spatial errors (Figures 3.8 and 3.9). In examining the spatial distribution of flux-scale errors we see that the worst errors are highly

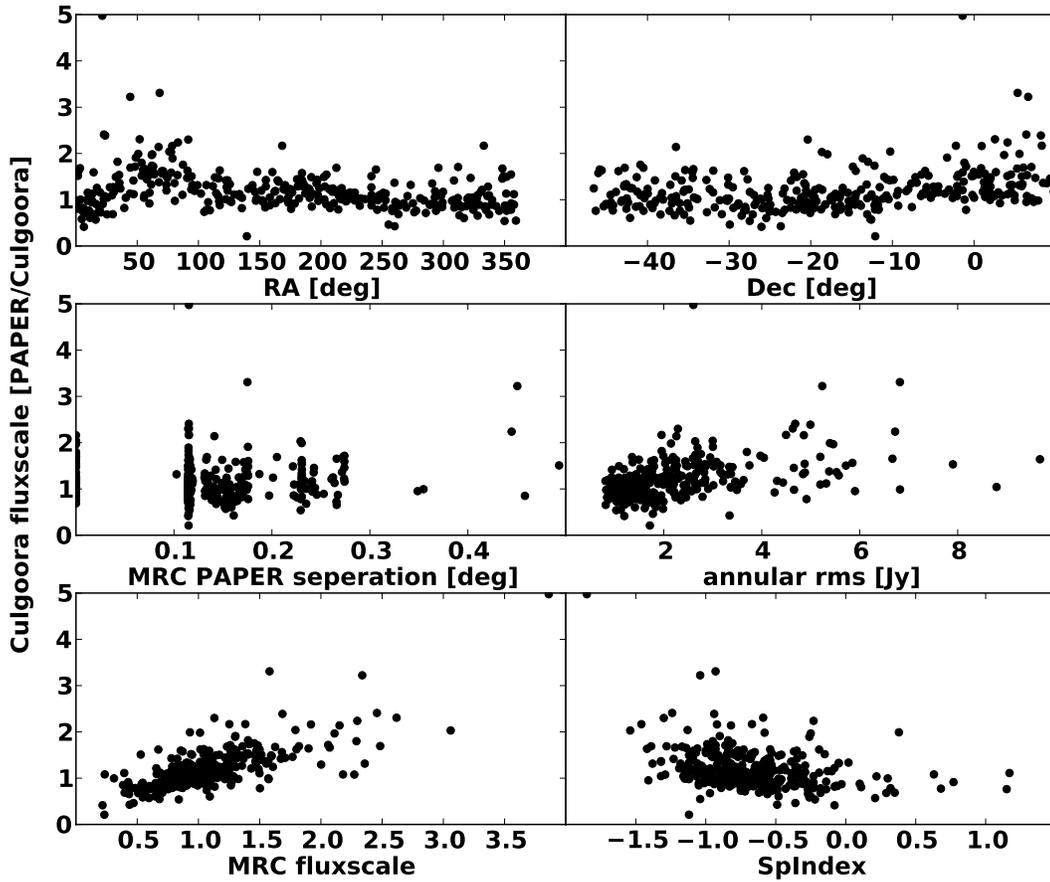


Figure 3.7: Here we assess the use of the flux-scale [ratio of PAPER to Culgoora flux] as a single estimate of source measurement quality that encompasses spectral and confusion errors in the reference catalog as well as measurement error. As shown above the flux-scale is appropriately correlated with the MRC flux-scale (correlation coefficient of 0.74) and weakly correlated with both the local image rms (0.44) and Culgoora spectral index ( $-0.35$ ). Flux-scale is apparently unconnected with the precision of the location (center-left), though there are hints of a position dependence in Right Ascension and Declination dependence which we follow-up in Fig 3.9

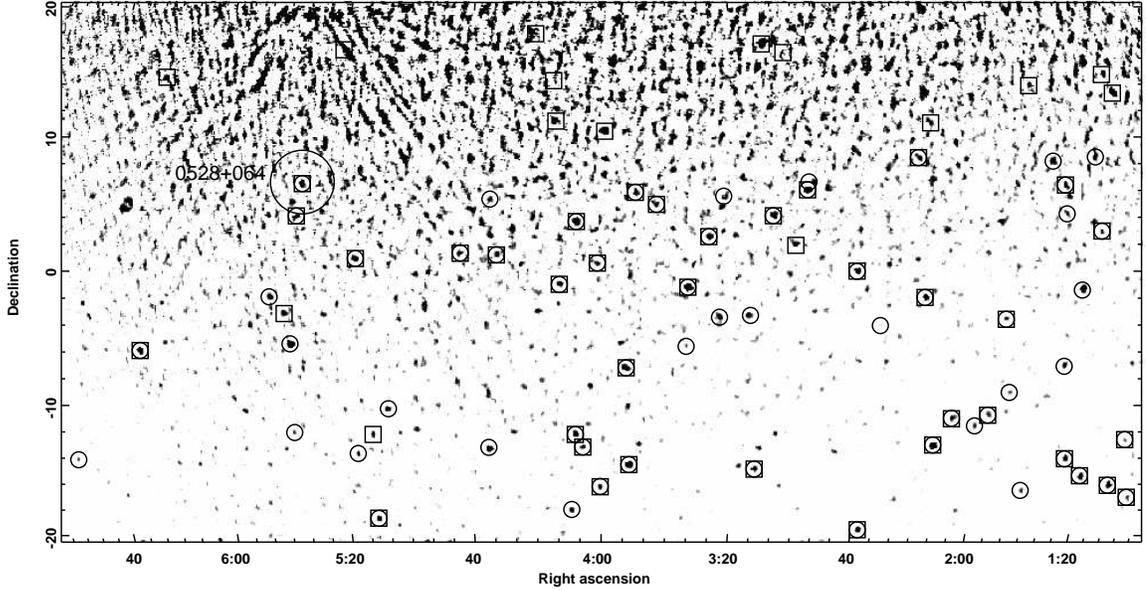


Figure 3.8: Region of sky with a bright source (Crab) low in the beam just off the top edge at  $+38^\circ$  elevation. Crab’s high angle from zenith proves too much for image plane deconvolution which depends on the PSF not changing across the beam, or the flux of the source staying constant during deconvolution. As we see in Fig 3.9, sources in this area have a higher flux-scale (brighter flux compared to catalog) and higher than average rms. For comparison purposes the high rms source 0528+064 is annotated here and in Fig. 3.9. For illustration purposes the color-scale undergoes a steep transition. Regions colored black are above 4Jy, white pixels are below. Source markers as in Fig. 3.5.

spatially correlated. A large fraction of the flux over-estimates occur at high declination, where the number of samples is geometrically diluted and image-plane deconvolution is more error-prone. The declination span of the psf increases with decreasing elevation  $a$  as  $\text{cosec}(a)$  which is a 55% error at  $20^\circ$  of elevation. This effect is most noticeable when bright sources are present at low declinations. The largest concentration of flux-scale errors occurs near an LST of 5h at high declination (Fig. 3.8). Here the number of samples is as at a minimum while side-lobes from a bright source low in the beam (Crab at  $+38^\circ$  elevation) are very high. A second region of high error occurs near Hercules A (also marked in Fig. 3.9). At  $+55^\circ$  elevation Hercules is high enough to be imaged, though side-lobes from psf distortion do distort the nearby flux-scale and raise the rms of flux measurements in the region.

With the proviso that the flux-scale includes imaging and catalog comparison errors we may evaluate the basic properties of PAPER’s flux-scale as compared to the flux-scales of similar catalogs. In Figure 3.10 we have plotted the distributions for the PAPER/Culgoora and PAPER/MRC<sup>5</sup> flux-scales. To estimate the catalog comparison error we have cal-

<sup>5</sup>Where fluxes were measured at frequencies outside of the PAPER band (110-180 MHz), we have scaled

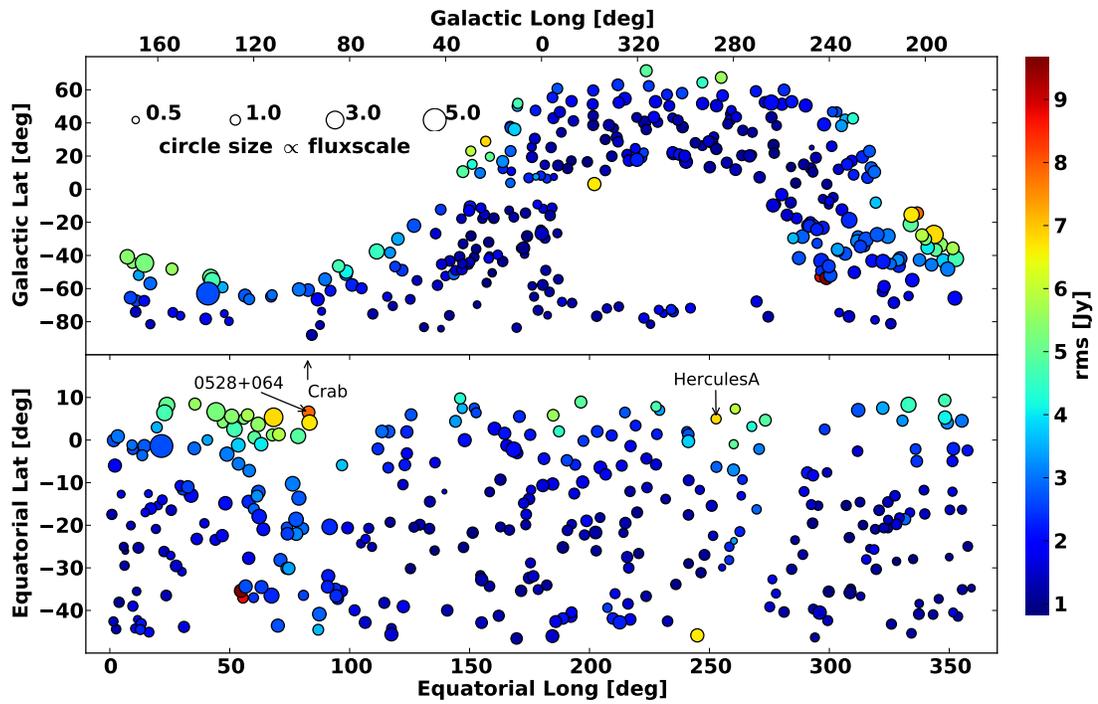


Figure 3.9: Spatial distribution of flux-scale and rms in Galactic (top) and Equatorial (bottom) coordinates. Point size indicates flux scale as shown in the bottom key while image rms in an annulus between 1 and 3 degrees radius around the source is related by color. The area of high flux-scale appears to be correlated with high rms in a high declination region near the galactic center.

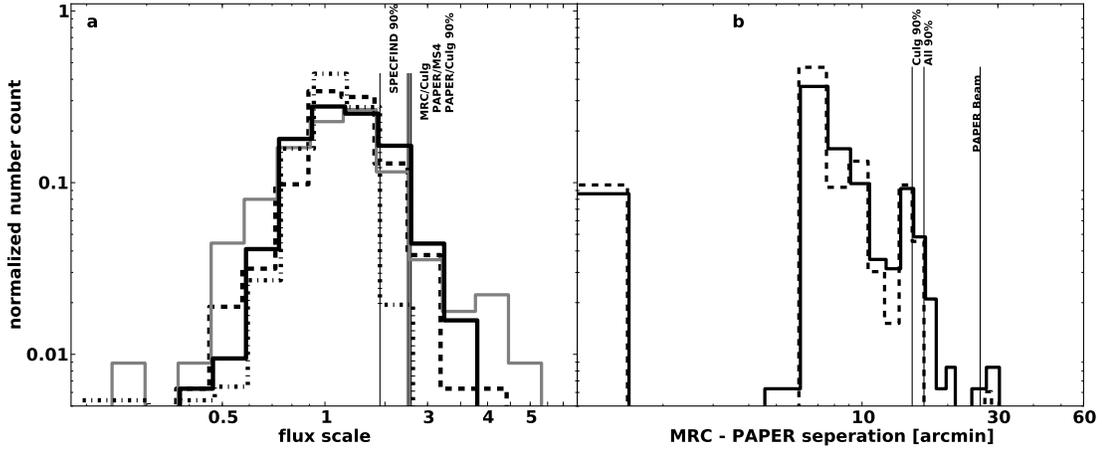


Figure 3.10: a) PAPER’s flux scale —the ratio of PAPER fluxes to Culgoora fluxes at 160MHz (black, solid) and to MS4 fluxes interpolated to 178MHz (grey, solid). PAPER/Culgoora and PAPER/MS4 with FWHM of 1 and 0.7 have distributions similar to the MRC/Culgoora flux ratio (thick, dashed) with a FWHMs of 0.5. A similar comparison between 178 and 151 MHz objects co-identified by SPECFIND (dot-dashed) shows that a somewhat tighter agreement (0.5 FWHM with fewer large outliers) is possible if cross-matching is done while accounting for morphology and instrumentation. b) Distance in degrees between MRC position and identified PAPER peak. All MRC sources within  $-85^\circ < \delta < 10^\circ$  have been paired with a PAPER peak within the plotted range. PAPER positions are defined by the centers of HEALpix pixels that are  $7'$  on a side leading to quantization effects; 90% of sources are within one PAPER beam (thin solid vertical line).

culated the MRC/Culgoora flux-scale, as well as between all sources in SPECFIND at 151-MHz and those at 178-MHz. The SPECFIND comparison has the advantage of more accurate cross-identification between sources. In addition, most of the measurements in these bands were done by the Cambridge Low Frequency Telescope (6 and 7C) and the 3CR all of which are known to be in good agreement (Bennett, 1962; Gower et al., 1967; Baldwin et al., 1985). Even so, a noticeable (although narrower) spread of flux-scales is apparent (Figure 3.10). In the SPECFIND comparison 90% of sources are below a flux-scale of 1.5, while in all other comparisons the 90% level occurs at 1.75. The distribution of the MRC/Culgoora flux comparison has a spread similar to the distribution of PAPER’s fluxes relative to each of these catalogs. The MRC sources have been cross-identified with the Culgoora using the same algorithm as the MRC-PAPER cross-identification. Thus the MRC-Culgoora comparison would be more likely to have similar catalog comparison errors and in fact does have a similar distribution of flux-scales. This similarity implies that

the flux using a spectral index of  $\alpha = -1$  (the average index at these frequencies; (Slee, 1995; Helmboldt et al., 2008; Bennett, 1962)

systematic errors will not be easily distinguished by catalog comparison. As an example consider the most extreme flux-scale outlier 0123-016 (3c40). PAPER observes a flux of 26 Jy while Culgoora only 8.9 Jy. Culgoora notes this source to have multiple components and measures 32 Jy in the 80-MHz band. When we add 3C and 3CR to the spectrum we see a consistent picture of a source around 30 Jy as shown in Figure 3.6 with the 160-MHz point the only in disagreement. Catalog discrepancies of *this* type are rare but there are many types.

### 3.3.3 Discussion and Conclusions

The Epoch of Reionization signal will be faint; detection will require precise calibration as well as deep foreground removal. Self-calibration of a wide-field instrument requires both *complete* and *accurate* knowledge of sources covering a large fraction of the sky. This calibration may then be verified by comparing multiple measurements, ideally between telescopes. Use of catalogs of past measurements are limited by confusion and errors from extrapolation.

To facilitate such comparison and fill the gap we have published (Jacobs et al., 2011) the first catalog derived from early PAPER data. Despite suffering from various systematics related to instrument response and analysis methodology this catalog shows qualitatively good agreement with other measurements. In the process we have demonstrated our ability to map more than half of the sky with two days of observation. This represents a major advance in survey speed and is made possible by the width of PAPER’s primary beam, the bandwidth of PAPER’s correlator, and the use of w-projection.

A number of improvements to the instrument and data processing methodology are currently underway. Cross-talk can be mitigated by the addition of one-way RF coupling and phase switching (as described in Rohlfs & Wilson (1996)). Both are likely to reduce excess correlations though their effectiveness are still being evaluated.

Algorithmic improvements are also available in the CASA environment. As we’ll discuss below, tests of Cotton CLEAN, faceting combined with w-projection, time-dependent calibration, and spectroscopic imaging have been favorable; this system will be used to produce higher dynamic range maps suitable for a blind survey and spectroscopic exploration. Finally, the dynamic range is limited by the instantaneous *uv*-coverage of the 32-element antenna configuration. Future deployments with larger number of antennas will result in additional improvements to the snapshot *uv*-coverage and imaging dynamic range.

Implementing these instrumentation and processing improvements will help produce images of even higher fidelity that will reach the sub-Jy confusion limit. From such images, it will be possible to construct a complete blind catalog using source fitting and photometric analysis. This more precise catalog will merit a stricter comparison with previous catalogs that more carefully accounts for extended structure, confusion and spectral slope.

Compared to the current array, the final PAPER South Africa array will have four times as many elements and should have 16 times the dynamic range. Here we have

imaged and cataloged what is essentially a dirty image of the sky and already found good agreement. The final PAPER telescope will be capable of spectral imaging the 110 to 180-MHz night sky to the confusion limit once a day. Although the radio sky was first explored at meter wavelengths, much remains unknown about the spectral and temporal behaviour of sources in this frequency band. These early PAPER 32-element commissioning data are already demonstrating a reliable level of accuracy that are limited primarily by well-known problems. We can reasonably expect future PAPER data to add substantially to our understanding of the sky at meter wavelengths.

We have combined observations separated by many months, correlator upgrades and power cycles to generate a map of the entire sky below  $+10^\circ$  Declination using only minimal calibration to do so. That fluxes measured in this image compare well with previous measurements indicates that PAPER is stable over a long period of time. While this is an encouraging result, what is not clear is the degree to which stability is tested by imaging. To more directly answer this question we must calibrate as a function of time.

# Chapter 4

## Calibration

In this chapter we explore some of the assumptions made in Chapter 3 and in doing so introduce elements of another pipeline with which we can check our results. Using this pipeline we will check the stability of the calibration both directly by comparing nightly solutions and indirectly by imaging. Finally we will use this calibration to estimate the system temperature and compare with a model.

In Chapter 3 we assumed gain with a phase dominated by a delay; a phase that changes linearly with frequency, which is stable over time. We found in §1.4.2 that over narrow bandwidths the delay domain translates almost directly into the most useful power spectrum axis and over wide bandwidths allows us to build a model of foregrounds. However the combination of steeply dropping foregrounds and steeply rising noise (Fig. 1.6) means that errors in delay ( $\Delta\eta$ ) effectively raise the noise floor as  $\Delta\eta^3$ . Thus we would like to show that delay is either a) intrinsically stable or b) easily calibrated.

To begin, in §4.1 and 4.2 we will transition from our AIPY solver to a faster, but more limited solver implemented in CASA. To eliminate confusion between sky and instrumental variability we focus on a very short time around a bright source, from which we can easily derive high SNR solutions. In §4.2 we will test the quality of the solution through a series of images. As outlined in Fig 4.2 we include both the calibration field and the nearby EoR fields where  $T_{sys}$  and foregrounds are at minimum. Finally we include an image of Centaurus A, a spectacular nearby radio loud galaxy as an example of PAPER's imaging power.

While a stable delay is important to minimizing foreground contamination the system temperature must also be as low as possible. With a projected required integration time of 120 days on a field that is available for only 165 days per season a mere 38% increase in the system temperature can push observations into a second season where there is increased risk of long-term instabilities and an increased operational cost. To accurately plan our long integrations and understand our detection significance in §4.3.1 we develop a model of the expected system temperature and in §4.3.2 make two somewhat independent estimates of the measured temperature.

## 4.1 Complex Gain Calibration

In principle the measurement equation (Eq. 1.5), relating the observed visibility to the actual sky has no single solution. Strictly speaking, if nothing is known about the sky a solution may in fact be impossible to find. Only by incrementally adding or assuming some knowledge of the sky can we eliminate unknowns. Even given a decent model of the true visibility which must include the sky, primary beam, and antenna positions we must still solve Eq. 1.6 for which no linear solutions exist. However, with each approximation or addition of information we reduce the number of calculations necessary to reach a solution. There is, of course, no single correct method to arrive at the solution. The method, the level of approximation and the estimation of error are all contingent on the sophistication of the analysis tools and that most imprecise of variables, user discretion.

In Chapter 3 our solution method was as simple as possible. Parameters were found by solving the full measurement equation with a generic amoeba-type conjugate gradient (CG) optimization routine. At each step the rms difference between data and model was computed at several points in parameter space (the amoeba feet) randomly chosen around a starting value locations. Using all of these points a new point was chosen along the rms gradient around which the gradient is again estimated for a new step. The process is repeated until the rms stops decreasing.

This forward modeling approach as implemented in AIPY is well suited to solving the measurement equation when very little is known and has two important advantages. First, it is very simple to implement and can be built directly off of the basic visibility simulator and free, open-source minimization routines. Second, as each step re-computes the *phase* of the measurement equation we are able to fit for more terms like antenna positions or source location. However there are down sides to this system. First, since the complete model visibility is computed with each iteration a step is quite costly and scales as square number of antennae  $\mathcal{O}(N^2)$ . Second, since we have an analytic form for the model, we could derive the gradient at any point and take an optimal step and eliminate the costly amoeba feet computations which must scale as the approximate number of per-antenna gains for an overall scaling of  $\mathcal{O}(N^3)$ . This scaling translates to very long execution times when solving large numbers antennae. Entire array solutions are only feasible in a reasonable amount of time by parallelizing the process, spreading the simulation over many nodes.

In contrast, the gain calibration system in CASA limits itself to solving only the terms outside the integral (Eq. 1.6). The gains *outside* the integral are solved given a model visibility  $\mathbb{V}^m$  that is only computed once. As this is still a non-linear problem we must still employ iterative methods. However in CASA these are fully optimized for the known slope of the measurement equation and are therefor solved with fewer computations per step in parameter space. This solver is more limited in its output, it is highly automated and streamlined; it is capable of providing time and frequency dependent solutions as well as an estimate of the solution signal-to-noise ratio (SNR). While this is not a true error-bar, the SNR provides a good indication about the *relative* quality of a solution and provides a basis for flagging low quality solutions. Regardless of the manner in which a solution is

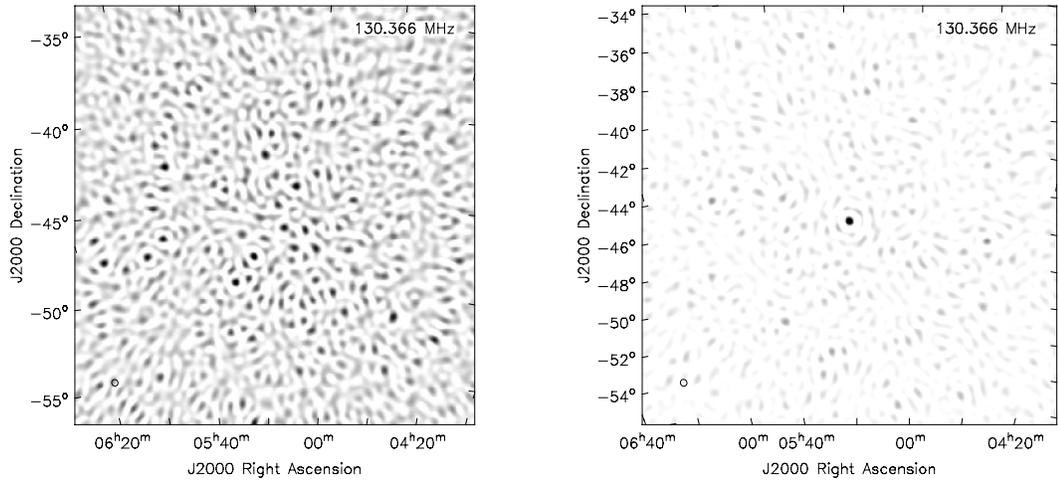


Figure 4.1: Without delay calibration, each antenna perceives a different location for Pictor A (left) as extra system delays are interpreted geometrically. Delay calibration uses a model of the source to remove these components. If successful the many apparent sources will converge on a single point (right).

reached its quality will still depend on the accuracy of the input model so estimating the solution accuracy is a desirable feature. Since there is only a single computation of the model visibilities we can estimate the scaling with antenna count as a single step in the CG/AIPY method of order  $\mathcal{O}(N^2)$  but with significantly quicker time per iteration.

In limiting ourselves to the complex gains, we avoid repeated computation of the model visibilities and reduce solution time by at least a factor of order  $\mathcal{O}(N)$  or a  $\approx 32\times$  for the  $N = 32$  antenna array under study here. In practice, a parallelized CG method takes an hour on a 64 node cluster while the linear-least-squares takes minutes on a single compute node or a factor of several thousand.

## 4.2 Pipeline

In chapters 4 and 5 we explore PAPER’s stability examining many observations of the same field. The data set spans the 3 hours surrounding 5h30 over 11 days between 18 Sept and 29 Sept (see Table 2.2). The first steps of the data processing pipeline are very similar to those in Ch. 3. RFI is flagged at the  $2\sigma$  level followed by a spectral averaging to 192-kHz. Cross-talk is removed by subtracting an average over the full three hours (see 5.4 below), followed by another RFI flagging run. At this point the data are converted from the AIPY miriad format into the CASA measurement set. The principle difference between the two data sets is that in the measurement set we must phase the drift scanning visibilities to a single location, in this case the median LST of each 10 minute long observation.

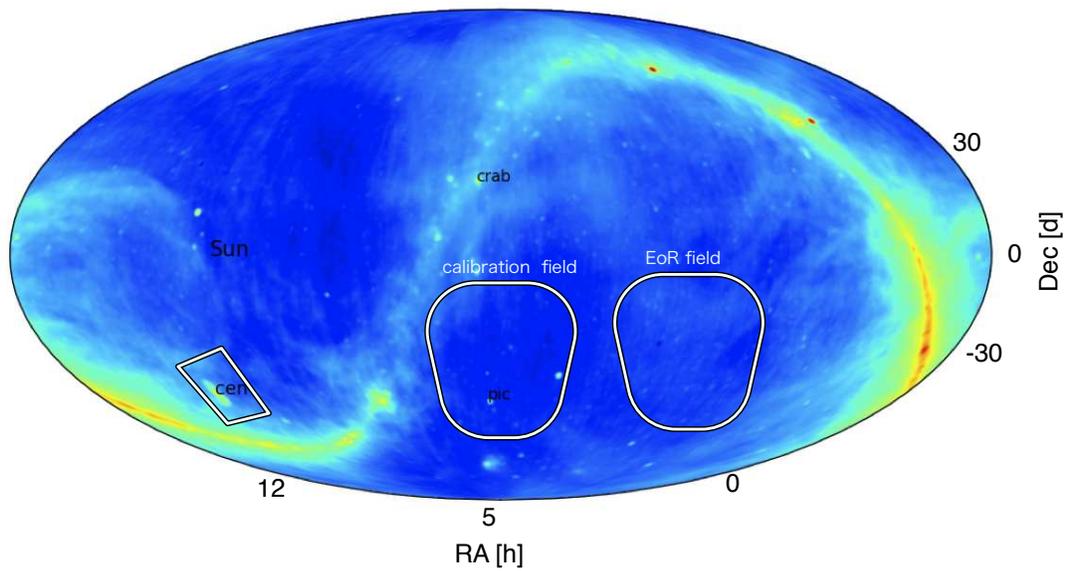


Figure 4.2: Key regions of the sky and the de Oliveira-Costa sky model. In this part of the spectrum the sky model is dominated by the Haslam map at 408-MHz. With a bright source and minimal complicated structure, the Pictor field is ideal for calibration (Fig. 4.3). Its proximity to the cool foreground region identified as optimal for reionization (Fig. 4.4) is also a plus. Centaurus A the brightest and most compact complex object in the southern sky is an ideal imaging candidate (Fig. 4.9).

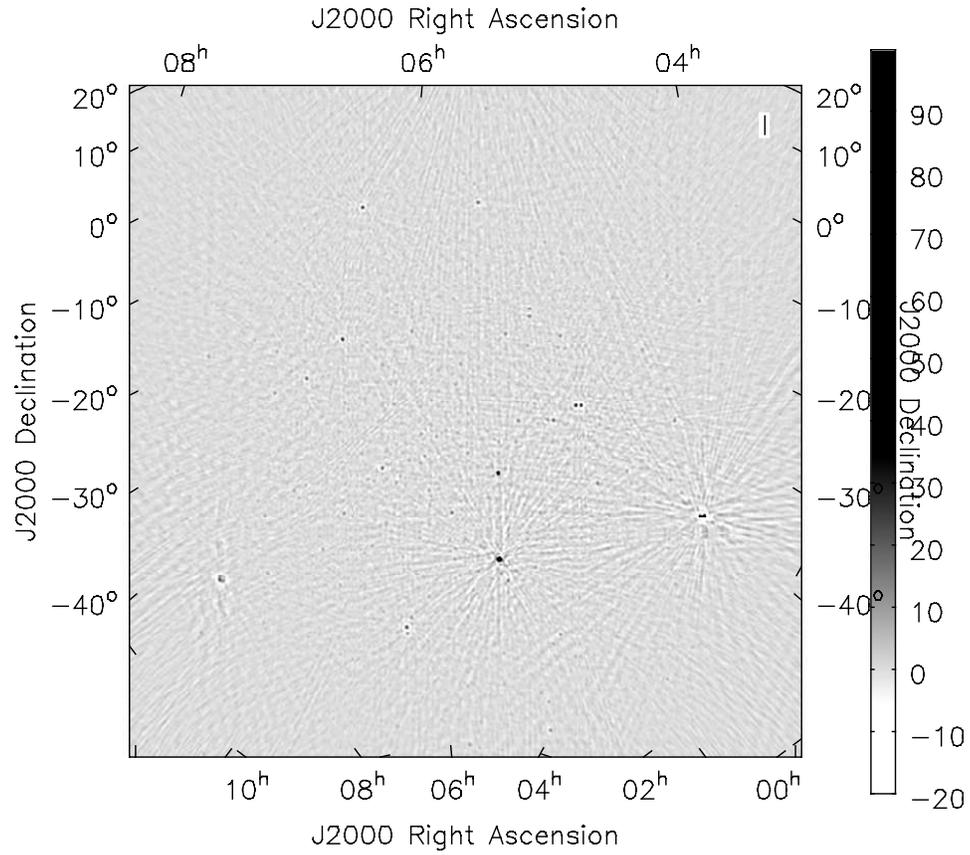


Figure 4.3: A wide bandwidth (120-180MHz) image of the Pictor A field using CASA. After delay calibration and a stack of 6MHz bands were imaged and self-calibrated in several rounds and the results averaged. Sources down to 1Jy are distinguishable.

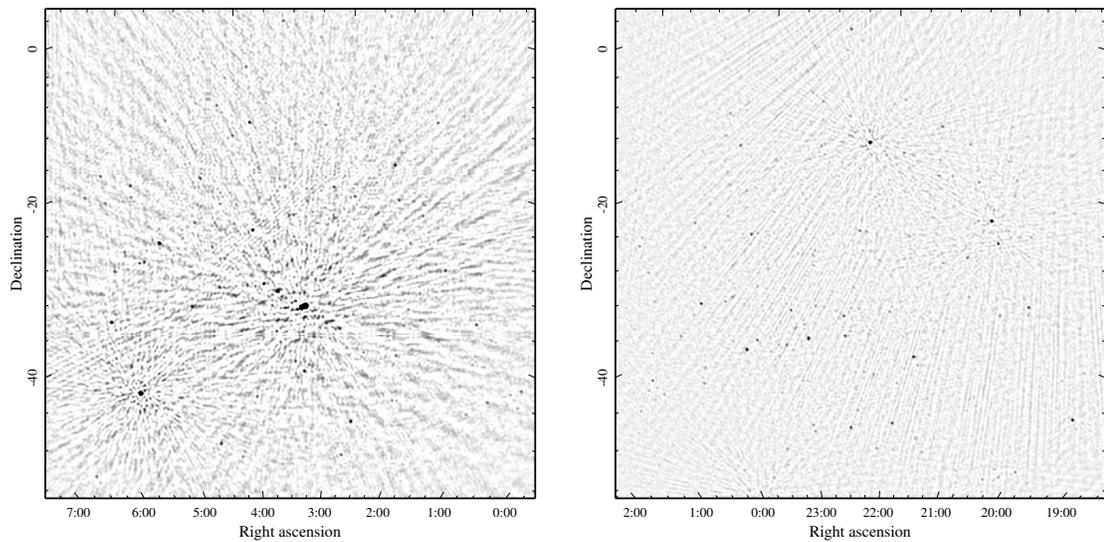


Figure 4.4: There are two possible regions best suited to EoR. On the left we see the field where  $T_{sys}$  is at a minimum near LST=3h. This region is not far from our calibration region. Pictor is still visible at the bottom left and Fornax A is directly overhead, scattering power due to its resolved structure (see Figure 4.5). The bandwidth of this image was limited to 10MHz to limit sidelobes due to Fornax's complicated spectral structure. On the right we have a 40MHz-wide image of the region centered on LST=22h (similar to Fig. 3.5). Both images are on the same scale with white pixels at 3 Jy or lower and black at 7 Jy or higher. Despite having a slightly higher  $T_{sys}$ , the field at 22h would seem to have to have preferable foregrounds. Both would seem to validate the Pictor field calibration. It is both stable and portable.

With our data cleaned of RFI and common-mode, we pursue a calibration solution by generating a model of the sky and solving for frequency dependent gains. In these observations Pic A (5h19m -45d46m) is the brightest apparent source reaching 380Jy at  $15^\circ$  from zenith. To further control variation in this calibration test we limit our data selection to measurements within 10 minutes of Pic transiting giving us what should be 11 identical observations of the Pictor field. For each observation we generate a visibility model using the same source positions, fluxes and beam amplitude used for delay calibration in Ch. 3 (in this case the five most apparently bright sources with altitudes above  $15^\circ$ ) and solve for a complex gain solution on each channel of each antenna using the CASA *bandpass* task. As in the Ch. 3 calibration we then found an average delay solution (Eq. 1.4) for each 10 minute observation. The phase angle  $\phi(\nu)$  output by the solver is wrapped to  $\pm\pi$  and must be unwrapped by removing discontinuities of magnitude  $2\pi$ . After unwrapping the complex phase we fit a linear model where the slope is the physical delay of each antenna (Fig. 4.6).

RFI or flagging can introduce discontinuities in this phase spectrum solution which in turn can interfere in the unwrapping process as demonstrated in the right pane of Fig 4.6. Furthermore we should note that to account for the gross shape of the bandpass filter we have fit a fourth order polynomial to the amplitude of the gain solution and observe that the amplitude portion of the solution appears to have less spectral stability than the angle. Despite these errors, as we can see in Figures 4.7 and 4.8, both the delay and the amplitude are largely stable from day to day.

Equipped with what appears to be a reliable and stable gain solution, we can verify its broad affect by imaging our calibration field 4.3. Point sources in this field include Pic, 0521-364 and Fornax A which has the unusual property of being a resolved double-lobed radio galaxy at 170-MHz and a slightly extended point source at 120MHz. A textbook example apparent "source confusion" (see §3.2) happening within our own band! Possessing little complex foreground but a very bright point source and situated on the edge of our primary EoR field (see Fig. 4.2) this region is an ideal location for calibration.

While the gain, particularly the delay, appears to be stable from night to night, we do not yet know whether we have derived a *correct* solution. Solution residuals might be large but stable. A truly useful gain solution will be successful (no multiple images ala Fig. 4.1) in other parts of the sky as well. In Figure 4.4 we see two possible EoR fields with this calibration applied. The minimum  $T_{sys}$  region centered on 3h30m,  $-31^\circ$ , is not far from our calibration region. Pictor is still visible at the bottom left and Fornax A is directly overhead, scattering power due to its resolved structure (see Figure 4.5). The bandwidth of this image was limited to 10MHz to limit sidelobes due to Fornax's complicated spectral structure. The sidelobes are worse than in Fig. 4.3 because we have not performed self-cal. As we can see, the situation is not so dire in our "low foreground" region from chapter 3. Here we have a 40MHz-wide image of the region centered on LST=22h (similar to Fig. 3.5) where the same calibration has been used. Both images are on the same scale with white pixels at 3 Jy or lower and black at 7 Jy or higher. Despite having a slightly higher  $T_{sys}$ , the field at 22h would seem to have to have less troublesome foregrounds. Both

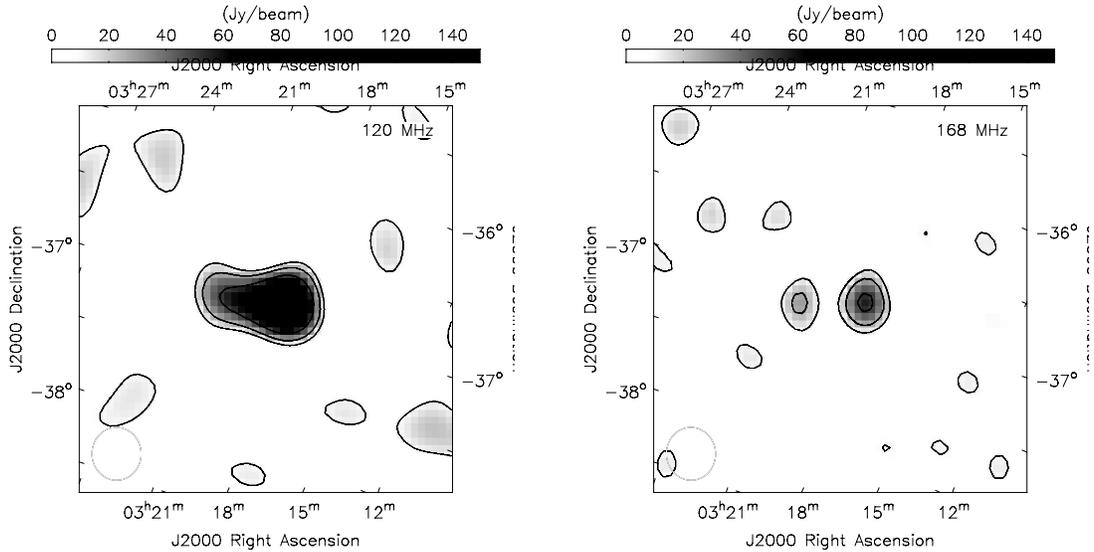


Figure 4.5: Fornax A. A slightly extended point source at 125MHz (left) resolves into a double lobed radio galaxy at 170MHz. A textbook example apparent "source confusion" (see §3.2) happening within our own band!

would seem to validate the Pictor field calibration. It is both stable and portable.

## 4.3 System Temperature

### 4.3.1 T<sub>sys</sub> model

As in any faint detection experiment the system temperature is the ultimate limit on the instantaneous sensitivity and sets the required integration time. PAPER (and any other wide field of view instrument) will have a component of system temperature due to smooth bright galactic emission that spans the entire field of view in addition to any stable system based component.

$$T(t, \nu)_{sys} = T(t, \nu)_{sky} + T_{rx} \quad (4.1)$$

While the sky component will vary with time and have a power law spectrum, we expect the the receiver component to be comparatively stable and flat spectrum. Realistically there will be variability due to ambient temperature changes, but as this is difficult to separate from gain variability we will not address it here.

A full sky map at 408-MHz by Haslam et al. (1982) has been combined with scarce other observations by de Oliveira-Costa et al. (2008) into a principle component model that predicts sky temperatures at 150-MHz between in a range from  $\approx 1,800K$  at galactic center to  $\approx 80K$  at high galactic latitudes with a fairly consistent spectral index of -2.55.

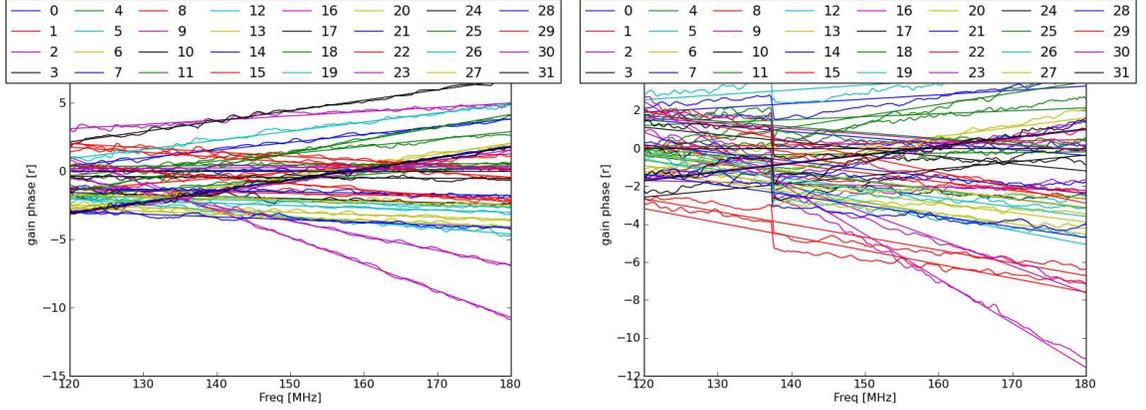


Figure 4.6: Using a single source model and the CASA *bandpass* task we solve the linear gain equation (Eq. 1.6), find a phase angle as a function of frequency, and fit a linear delay and offset model.

Roughly this spectral index means the temperature at 100 MHz is higher than 200MHz by a factor of 4.

We estimate the added temperature due to the sky  $T_{sky}$ , by integrating the sky model multiplied by the primary beam  $\mathcal{A}$  which is pointing at right ascension or local sidereal time  $t$ .

$$T(t, \nu)_{sky} = \frac{\int A(t, \hat{s}) * T(\hat{s})_{sky}(\hat{s}) ds^2}{\int A(t, \hat{s}) ds^2} \left( \frac{\nu}{\nu_0} \right)^\alpha \quad (4.2)$$

As we see in Fig 4.10 the theoretical temperature due to galactic emission peaks at 600K at 18 hours (near galactic center) and goes through a minimum at 3.6 hours (galactic coordinates  $l = 228^\circ b = -54^\circ$ ). In consideration of this unavoidable noise we have chosen to minimize cost by designing our amplification system to have a system temperature comparable to the colder part of the sky. Current laboratory measurements estimate the system temperature of the analog system at 120K. Limiting the net system temperature to around 200K.

This region of very low galactic temperature ( $23 < RA < 5$ ) is useful both for its low noise, and its minimal smooth foregrounds. It spans the 6 LST hour period used in our fiducial sensitivity calculation. In the South these LST bins zenith at night during a period beginning in June (with LST=23 hours) and ending in February (with LST = 5 hours), each pointing being available for between 172 days starting in June (longer nights during southern winter) to 160 days for bins ending in February (summer).

With our first order calibration and beam model in hand, we are now able to compare this model with our data.

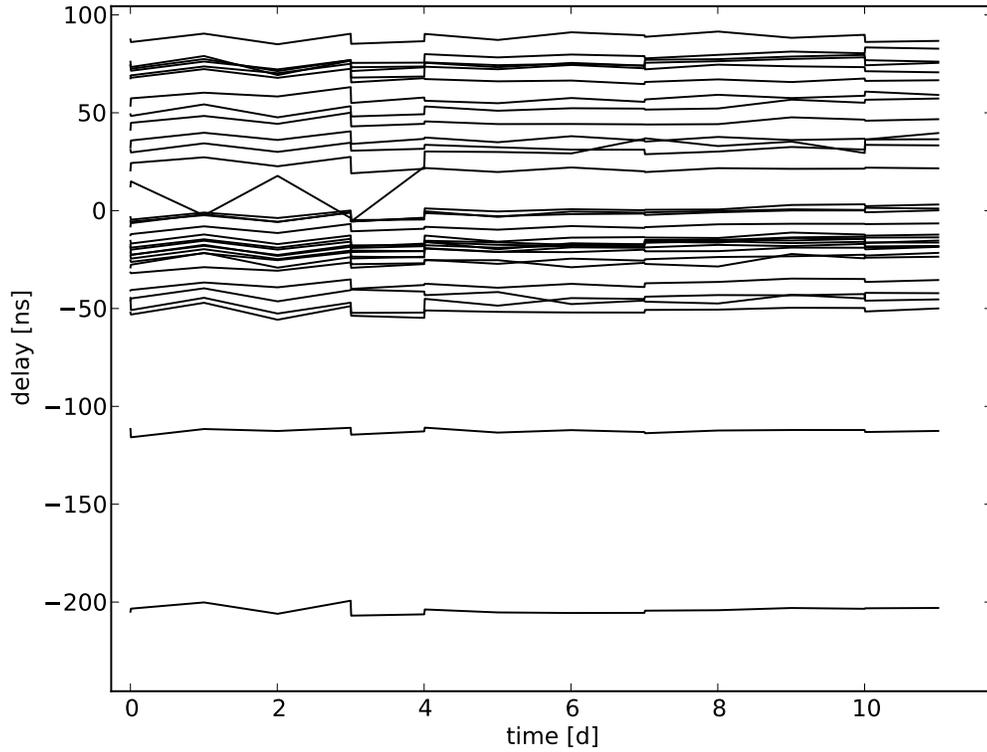


Figure 4.7: Delay solutions found using the CASA *bandpass* task and a single source model of Pictor A at transit. These solutions use only Pic A and by excluding other bright but more complicated sources like Fornax we are making an error in the model visibility and hence the delay solution. This is most clearly evident in the two solutions at 120 and 200ns which represent many meters of delay that was not seen in the Chapter 3 delay solution (Fig 3.2). However, we make the same error every night, and can therefore interpret any variation in the solutions as being primarily instrumental. This variation is minimal. The dip at day 3 is due to a small change between the two samples on that day that happens to affect the reference antenna and thus moves all solutions by the same amount. In general, the instrumental variation we see is much less than the difference between solutions.

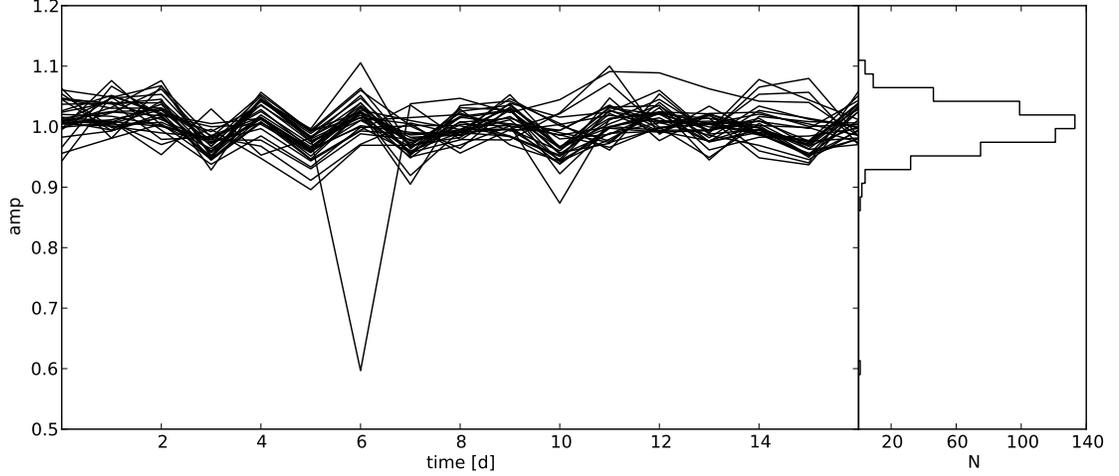


Figure 4.8: Relative amplitudes found by calibrating to Pictor A at transit. Here we see the average over the bandpass for solutions with  $SNR > 3$ .

### 4.3.2 Observed system temp

There are several equivalent methods for estimating system temperature from a set of visibilities. Let us use the autocorrelations to estimate the system temperature during transit of our EoR field (LST=3h30m) and then compare with the observed temperature of the crosses.

The autocorrelation measures the power from a single antenna and is a direct estimate of system temperature.

$$V_{ii} = \frac{g_i^2}{K}(T_{sky} + T_{rx}) \quad (4.3)$$

$$= \frac{g_i^2}{K}T_{sys} \quad (4.4)$$

where  $K = \eta_e A / (2k_B)$  is the conversion from Kelvins to Jys and the receiver front end contributes an additional  $T_{rx}$  which is thought to be approximately constant and spectrally flat. The effective antenna area  $A$  is approximately related to the integral of the beam primary beam  $\mathcal{A}(\hat{s})$  by the diffraction equation

$$A = \frac{\lambda^2}{\Omega} \quad (4.5)$$

$$= \frac{\lambda^2}{\int \mathcal{A}(\hat{s}) ds^2} \quad (4.6)$$

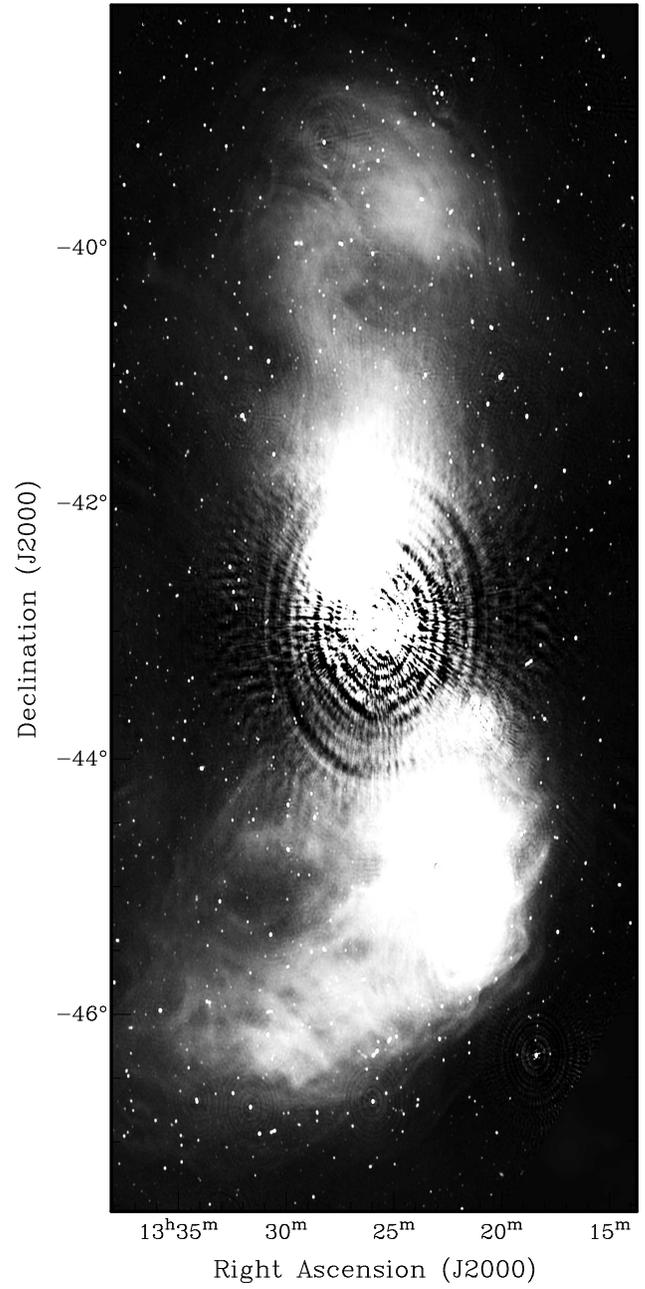
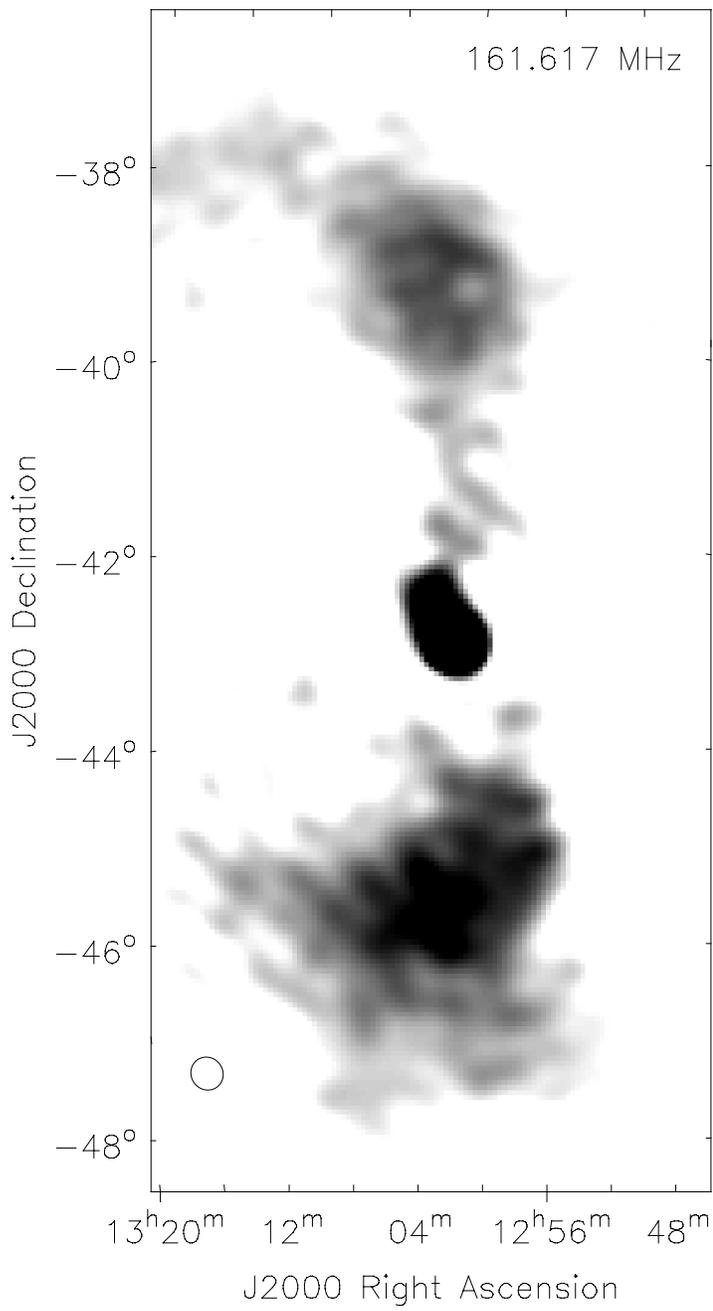


Figure 4.9: Left: Centaurus A as imaged by PAPER at 160MHz. Right: As imaged with ATCA and Parkes (1400MHz) by Feain et al. (2011), reprinted by permission.

The efficiency ( $\eta_e$ ) is included for consistency with other derivations (Thompson, 1999) but amounts to an overall normalization for  $\Omega$  which can roughly be considered to be related to the width of the beam. In fact we found it most expedient to consider this relatively unknown factor (regardless of its physical significance) as an overall scaling of the galactic temperature model. Writing  $K = A/(2k_B)$  we may absorb the  $\eta_e$  into  $T_{rx}$  and rewrite equation 4.1

$$T(t, \nu)_{sys} = 1/\eta_e T(t, \nu_0)_{sky} \left( \frac{\nu}{\nu_0} \right)^\alpha + T_{rx} \quad (4.7)$$

Given the measured  $T(t, \nu)_{sys}$  shown in thick black in Figure 4.10 and the galactic model (dotted line in same Figure) we find a linear least squares solution for  $\eta_e$  and  $T_{rx}$  of ( $1/\eta_e = 0.86, T_{rx} = 206\text{K}$ ). If we interpret the  $\eta_e$  as an actual efficiency and a correction to the beam model the actual receiver temperature is closer to  $T_{rx}/\eta_e \approx 180\text{K}$ . This implies an overall uncertainty of about 15% in the system temperature or about  $1 - \sqrt{1/\eta_e} \approx 7\%$  in the width of the beam.

Of course other things can contribute excess noise, most notably the RFI seen most clearly as the spikes at 137 and 150MHz. Neither is the rms of the cross-correlations, which is ultimately what we must overcome by integration, entirely due to the uncorrelated noise at each antenna. As we can see by computing the rms of the channel differenced correlations.

Each channel of visibility measurement  $v$  is the product of the signal + noise  $s_i + n_i$ , averaged over  $N = df * dt$  samples where  $df$  is the channel width and  $dt$  is the length of the integration

$$\langle v_{ij} \rangle_N = \langle (s_i + n_i)(s_j + n_j) \rangle_N \quad (4.8)$$

Working out the cross-product and assuming the noise obeys the central limit theorem, each component will decrease as  $\sqrt{N}$

$$\langle v_{ij} \rangle_N = \langle s_i s_j \rangle + \langle s_i n_j \rangle / \sqrt{N} + \langle s_j n_i \rangle / \sqrt{N} + \langle n_i n_j \rangle / \sqrt{N} \quad (4.9)$$

Differencing adjacent channels ( $p, q$ )

$$\Delta v_{ij} = \langle v_{ij} \rangle_N^p - \langle v_{ij} \rangle_N^q \quad (4.10)$$

$$= \langle s_i s_j \rangle^p - \langle s_i s_j \rangle^q \rightarrow 0 \quad (4.11)$$

$$+ \left( \frac{\langle s_i n_j \rangle^p}{\sqrt{N}} + \frac{\langle s_j n_i \rangle^p}{\sqrt{N}} \right) - \left( \frac{\langle s_i n_j \rangle^q}{\sqrt{N}} + \frac{\langle s_j n_i \rangle^q}{\sqrt{N}} \right) \quad (4.12)$$

$$+ \frac{\langle n_i n_j \rangle^p}{\sqrt{N}} - \frac{\langle n_i n_j \rangle^q}{\sqrt{N}} \quad (4.13)$$

we find that the signal terms (4.11) cancel exactly.

Squaring and averaging the result over many time and baseline samples and ignoring

cross terms

$$\langle |\Delta \langle v_{ij} \rangle_N|^2 \rangle_{ij,t} = 2 \langle \langle n^2 \rangle \rangle^2 / N \quad (4.14)$$

$$+ \mathcal{O}(\langle (\langle sn \rangle^p - \langle sn \rangle^q)^2 / N \rangle) \quad (4.15)$$

we find a noise term (4.14) and terms proportional to the derivative of the noise-signal correlation (4.3.2). These require careful attention as they can contribute to our measurement error.

In the signal dominated case, the difference approaches zero, as the signal correlations did in Eq. 4.13. In the noise dominated case the noise-signal correlation will be at a minimum where we may also safely ignore these terms. In the intermediate, SNR 1 case, the correlation  $\langle sn \rangle$  is at a maximum and would presumably contribute additional noise. In the following let us assume that we are not in this regime goes to zero

$$\langle |\Delta \langle v_{ij} \rangle|^2 \rangle_{ij,t} \approx 2 \langle n^2 \rangle / N \quad (4.16)$$

that depends only on  $T_{sys}$ .

$$\frac{\sqrt{\langle |\Delta V_{ij}| \rangle_{ij,t} \sqrt{df dt} K}}{\sqrt{2} g_i g_j} = T_{rms} \approx T_{sys} \quad (4.17)$$

Where we have substituted  $V_{ij}$  for the accurate but cumbersome  $\langle v_{ij} \rangle_N$  and replaced  $N$  with the explicit number of samples in a channel. We plot the resulting measurement as the thick grey line in Figure 4.10.

At low frequencies, where  $T_{sys}$  is higher, the rms agrees well. At higher frequencies the agreement is not so good. Whether this is due to our SNR approximation or our channel differencing to remove the signal is not clear.

## 4.4 Conclusion

In this Chapter we have found that the crucial delay calibration is stable over many repeated observations, incorporating new gain solving tools to do so. These gain tools are really a small part of the larger CASA toolset which we have experimented with as the beginnings of an alternate pipeline. We demonstrate our calibration and use of advanced imaging tools by imaging regions of interest including Centaurus A and an EoR Field where foregrounds and  $T_{sys}$  are minimal. Using observations over a 24 hour period we have modeled the receiver temperature and compared with data, an analysis summarized in Figure 4.10. Using the auto correlations we compute the system temperature over a full day which we found is broadly consistent with our model beam, a model of the sky, and a system temperature of 180K. We then found that the auto-correlation spectrum at our EoR field runs from 300 at 120MHz to 200K at 180MHz. We then noted that the noise on cross-correlations is only approximately equal to the system temperature. We then found

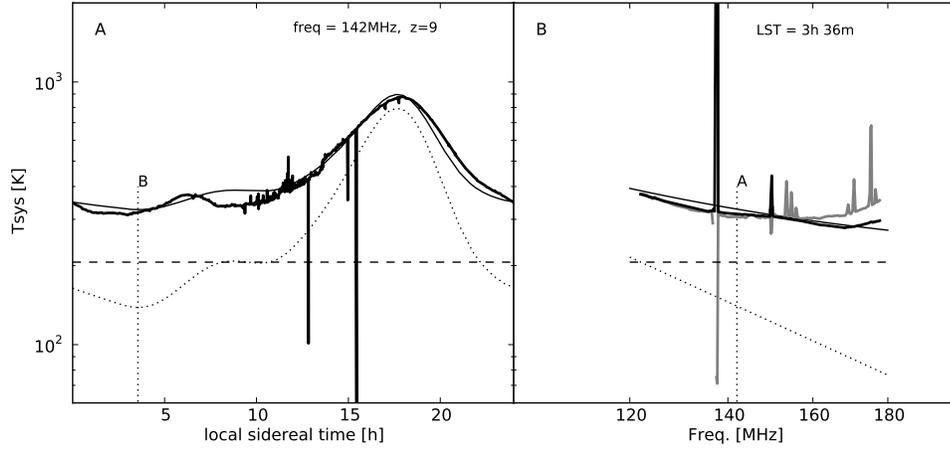


Figure 4.10: System temperature is due to a combination of sky (dotted) and receiver noise (dashed).  $T_{sys}$  is estimated by the auto-correlations (thick black) and cross-correlation rms difference (thick grey). We estimate the sky temperature by convolving (Eq 4.2) the primary beam model  $\mathcal{A}$  with the sky model by de Oliveira-Costa et al. (2008) and assume a spectral index of -2.55. The overall efficiency of the beam which sets the scale of the sky temperature remains uncertain as does the exact receiver temperature. Here we have used both the spectrum and the time series to jointly solve  $T_{sys} = 1/\eta_e T_{sky}(t)(f/f_o)^{-2.55} + T_{rx}$  for  $\eta_e = 1.14$  and  $T_{rx} = 206\text{K}$ . The resulting model is shown in thin, solid. From this we can see that our beam is slightly narrower ( $\sim 7\%$ ) than modeled and that there is a 100K excess of variance frequencies above 150MHz. Whether this is due to our SNR approximation or our channel differencing to remove the signal is not clear.

that the noise on the cross-correlations is consistent with  $T_{sys}$  below 150MHz but diverges to higher temperatures above this frequency.

In even the worst case the temperature never goes above  $\sim 300\text{K}$ , which is little more than half the 500K assumed in our sensitivity estimate. If this is stable over many days it could decrease the required integration time by a factor of 2. Of course this assumes that the noise is stable and obeys the central limit theorem by integrating down, an assumption we will explore in Ch. 5.

# Chapter 5

## Sensitivity

In this chapter we investigate the matter of instrument stability and its ability to reach a useful sensitivity level. In Chapter 1 we found that each  $\vec{k}$  must be sampled to an SNR of 1 before multiple modes of the same length can be profitably combined. Working backwards from a given configuration and hypothesized EoR amplitude we found in Parsons et al. (2011) that each baseline must integrate down as noise for approximately 120 days for a precision of one part in ten thousand. In this Chapter we make a first assessment of the ability of the instrument to reach this level. In Chapter 3 we found that a single delay and amplitude calibration worked well enough to combine two days of data separated by 5 months. It was accurate enough to allow us to integrate coherently to make flux measurements consistent with catalogs. In Chapter 4 we verified that these delays are stable from day to day. Here we will assess stability on a shorter time-scale but with more data. While we lack the 120 days of observation required to reach the  $SNR \sim 1$  level, in §5.1 we devise a simple test to assess the degree to which our noise integrates down and in §5.3 and §5.4 we apply this test to our 11 day data set at two points in the processing pipeline noting the effects on the integrating properties at various delays.

### 5.1 Theory

### 5.2 Computing the Power Spectrum

In §1.4.2 we noted that the epoch of reionization is thought to be "co-temporal" over a redshift range of  $dz \sim 0.5$  which at the redshifts of interest is a mere 6MHz out of the available 80MHz. This dichotomy lead quite naturally to two useful approximations. Over wide bandwidths the linear dependence of the correlation phase on frequency means that the Fourier transform of a visibility to "delay" space will isolate the foregrounds to within the range of physically possible delays, ie the light travel time across the baseline. Over the relatively short bandwidth relevant to reionization the opposite is true. Over 6MHz the change in phase with frequency is small enough to ignore. In this case the Fourier/delay transform of each baseline is a sample of the 3D power spectrum as we

showed schematically in Figures 1.7 and 1.6. Here we will recapitulate the derivation in more detail, for even more detail see Parsons (2008).

Recall our flat-space, flat-array approximation of the measurement equation

$$V_{ij} = \int I(l, m, \nu) e^{(-2\pi i \hat{s} \cdot \vec{b} \nu / c)} dl dm \quad (5.1)$$

Let us model the sky as being a single source of flux  $S$  at position  $\hat{s}$

$$V = S \exp[-2i\pi \hat{s} \cdot \vec{b} \nu / c] \quad (5.2)$$

The Fourier transform along the frequency axis is (in the infinite case) a delta function

$$\tilde{V} = S \delta(\eta - \hat{s} \cdot \vec{b} / c) \quad (5.3)$$

and we recall that the dot product of the baseline with the source direction is simply the geometric delay ( $d$ ) of the source. This source might occur anywhere along the delay spectrum corresponding with delays *less* than the light-travel length of a baseline which could be anywhere between 5 and 300m or 16 and 1000ns light travel time. A good fiducial baseline length for EoR is around 133ns (20 wavelengths), though throughout our analysis we will favor the fractional delay ( $d_f$ ), where the baseline length in question has been divided out. Over an infinite bandwidth an ideal collection of point sources would occupy fractional delays between -1 and 1 with those at the limits significantly attenuated by antenna pattern. (See simulations in §5.2).

Of course a source with non-flat spectral shape will necessarily have a wider delay spectrum. Most sources are thought to be primarily smooth and thus limited to only a relatively small number of delay bins. A worst-case scenario might be a steep spectrum source with strong absorption that peaks near the center of our 100-MHz-wide band but is well below the noise near the edge. Modeling this as a gaussian of width 50MHz we see that it corresponds with a delay smearing of  $\sim 1/50\text{MHz} \sim 20\text{ns}$ . Compared with a typical baseline length of 200ns, and calibration errors of 1ns, we see that even pathological sources will most likely lie well within delays shorter than the physical baseline length.

This technique might be used to good effect by first using the wide bandwidth to isolate and filter the bright foregrounds before computing the power spectrum on a more narrow frequency range. It is essentially identical to the DDR filtering used in our foreground imaging (§3.3.1) but widened to include the entire sky. However, the steepness of the noise power spectrum places strict requirements on the sharpness or PSF of such a filter. In this work we are not yet concerned with precision measurements in the SNR "sweet spot" but rather the statistical properties of the noise dominated portion of the spectrum. At higher delays a wide PSF may still scatter power as we can see by simulating the delay space spectrum of a single source at various elevations with a square bandpass.

The PSF of a square bandpass of width  $B$  is a sinc function ( $\sin(\eta B) / \eta B$ ) of width  $1/B$  and introduces ripples sometimes known as Gibbs oscillation which mix power between many channels. The solution to this limited bandpass is to first multiply the spec-

trum by a sinc function so that the resulting psf is square. This is still imperfect because the sinc function goes to zero at infinity but our bandpass is not infinite, so we multiply by an additional weighting function that smoothly approaches zero at the band edge. This collection of weightings is known as Polyphase Filter Bank (PFB) and is also used by the "F engine" of the correlator. Addition of RFI flagging further distorts the PSF and generalizes this problem to one of deconvolution of the true delay spectrum from the convolved PSF, a process for which CLEAN is well suited. Using CLEAN to compute the optimal delay transform in the presence of RFI is discussed at length in Parsons & Backer (2009) which we will only repeat in outline: CLEAN iteratively finds the brightest channel in the delay spectrum, subtracts a fraction of this power convolved with the PSF and repeats until the residual rms reaches a minimum threshold. As we can see from our simple simulation in Figure 5.1, combination of the two methods proves quite effective at achieving a high dynamic range and minimizing the degree to which smooth foreground sources leak into higher delays.

Of course the sky will have many sources, which will necessarily limit the effectiveness of CLEAN leading to leakage of power to "non-sky" ( $|d_f| > 1$ ) delays. In Figure 5.2 we see a simulation of a single baseline (8-12 in Fig 2.4) observing our EoR field. Here we have simulated a visibility spectrum for the complete MRC catalog (see §3.1) to  $\sim 1$  Jy with a 1mK noise floor sufficient to constrain reionization models and typical flagging of the satellite channels. The resulting delay spectrum is foreground dominated within  $|d_f| \lesssim 1.5$  a promising result for a young technique and enough to assure us a noise power spectrum  $|d_f| \gg 2$  uncontaminated by foregrounds.

### 5.3 Integrating Power Spectrum

Let us assume that the power spectrum is dominated by a flat Gaussian noise component and a stable foreground component composed of spectrally smooth sources. As described in §§ 1.4.5 and 5.2 these will dominate the power spectrum at delays shorter than the length of the baseline ( $|d_f| < 1$ ). At higher delays the amplitude of the noise power spectrum will decrease by  $1/N$  as  $N$  samples are added. Non-gaussian noise will continue to integrate down but at a much slower rate. The presence of signal power due to cross-talk, scattered foregrounds or even EoR will be signaled by a complete leveling-off of the power as more samples are added.

When looking for EoR we will look at bandwidths of only a few MHz over which evolution will be minimal. However at this early stage we should be careful to separate the foreground and noise dominated regimes in order to get a better handle on the noise. So for the following we will use the entire bandwidth (80MHz) which will put more delay bins between noise and foregrounds and enable better CLEANing of the foregrounds. In the future we envision this being the first step to filtering foregrounds which change little across the band.

As discussed in §1.4.2 the interferometer measures the correlation of the electric fields

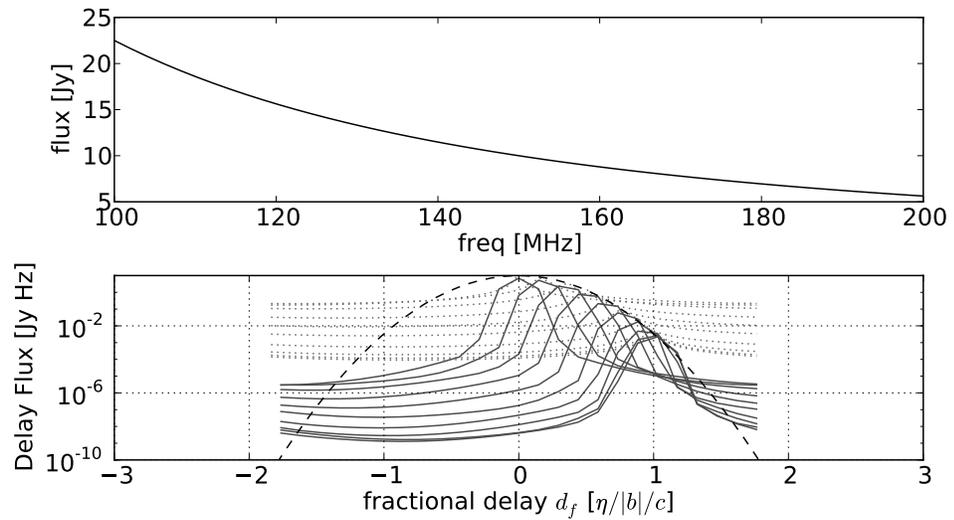


Figure 5.1: The delay spectrum (bottom) of a steep spectrum ( $\alpha = -2$ ) source (top) over a range of elevations. The limited band and spectral slope combine to scatter power to high delays (dotted). Using a polyphase filter bank to weight the spectrum eliminates the sharp transition at the band edge and drastically increases the dynamic range of the transform (solid). In both cases the beam response (dashed) smoothly approaches zero near the horizon. With its earth pointing nulls the beam only affects the source amplitude without injecting any spectral features/delay smearing due to secondary lobe response. Here we give the delay spectrum in fractional delay units, where the delay is divided by the length of the baseline.

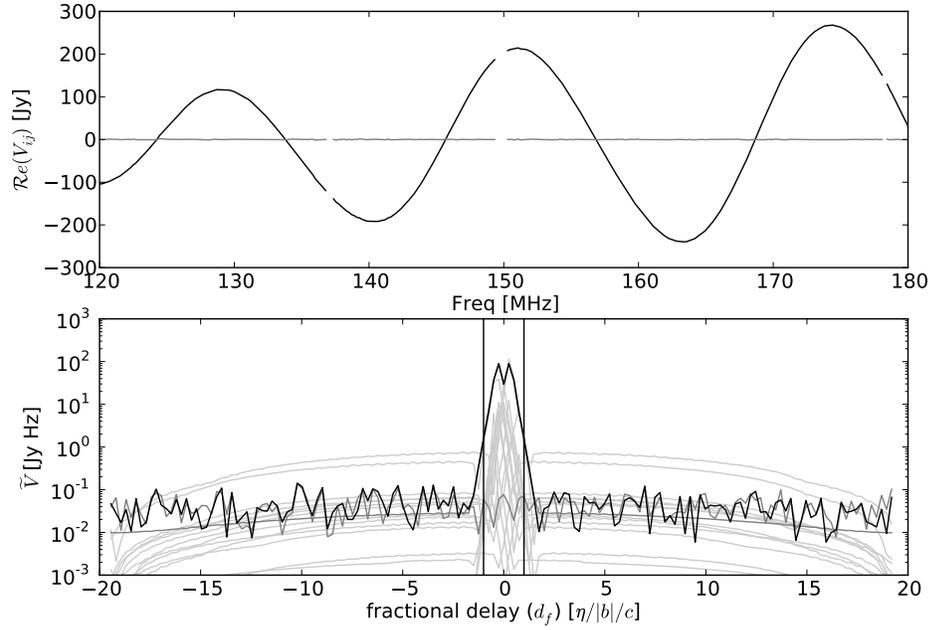


Figure 5.2: Here we have a model of the sky for baseline 8-12 assuming 120 days of perfect integration towards 3h30m -31d. Here we have simulated a visibility spectrum for the complete MRC catalog (see §3.1) to  $\sim 1$  Jy with a 1mK noise floor sufficient to constrain reionization models and typical flagging of the satellite channels. The delay spectra of the components of the model are shown in grey while the transform of the sum is shown in black. The net delay spectrum is consistent with noise above  $1.5 \lesssim |d_f|$ , a promising result for a young technique and enough to assure us a noise power spectrum  $|d_f| \gg 2$  uncontaminated by foregrounds.

between two antennae.

$$V_{ij} = E_i(t) \star E_j(t) \quad (5.4)$$

which is computed by the correlator with the FX <sup>1</sup> method

$$V_{ij} = \tilde{E}(\nu)_i \star \tilde{E}(\nu)_j \quad (5.5)$$

where  $N_t$  samples of E are used to find the Fourier transform

$$\tilde{E}(\nu) = \sum_t^{N_t} E(t) e^{-2i\pi t\nu} \quad (5.6)$$

The electric field from each antenna consists of a signal that is correlated with other antennae, and a noise part that is not. The noise will have zero mean and an rms amplitude given by the system temperature. As we have seen, over a good fraction of our band, the correlation noise is equal to the system temperature. The correlation of the noise will also have zero mean and a proportional rms (ala Eq. 4.16)

$$\text{rms}(V_n) = \text{rms}(E_n) N_t \sqrt{2} \quad (5.7)$$

as will the Fourier Transform of the visibility spectrum (the delay transform), which is just the inverse of the original transform in the correlator, though we have now associate the time variable  $t$  with the geometric delay  $\eta$

$$\tilde{V}_n(\eta)_{ij} = \frac{1}{N_t} \sum_{\nu}^{N_t} V_n(\nu) e^{-2i\pi\eta\nu} \quad (5.8)$$

the square of which is an estimate of the power  $P$  on the Fourier mode  $\vec{k}$  corresponding to  $(u, v, \eta)$

$$P_n = |\tilde{V}_n|^2 \quad (5.9)$$

where the subscript  $n$  is a placeholder for repeated sample index and could be used to index time or redundant baselines. Here we will limit to a single baseline and, for now, drop the antenna  $(ij)$  index. Where the mean is proportional to  $T_{sys}$  as given by the noise spectrum in Equation 1.20 and is spectrally flat in delay.

For a noise dominated delay  $d$ , adding more samples (summing over  $n$ ) will drop the magnitude as  $1/N$

$$P_N^2 = \left| \frac{\sum_{n=1}^N V(d)_n}{N} \right|^2 \quad (5.10)$$

$$= P_n^2 / N \quad (5.11)$$

---

<sup>1</sup>Fourier transform and cross-multiply

Of course  $P_N$  is also a random variable with amplitude randomly determined by *which*  $N$  samples we use for the average. We find a more robust estimate is the expectation value of  $P_N$  averaged over  $M$  draws of  $N$  samples.<sup>2</sup>

$$\langle P_N \rangle_M = \frac{\sum_{m=1}^M P_{N_m}}{M} \quad (5.12)$$

where  $N_m$  is the  $m$ th randomly chosen set of  $N$  samples. For each of these  $m$  averages we can also compute the variance of our power spectrum measurement.

$$\sigma(P_N)^2 = \frac{\sum_{n=1}^M |P_{N_m} - \langle P_N \rangle_M|^2}{M} \quad (5.13)$$

Let us verify this chain of reasoning and validate our statistical intuition by performing a simple simulation. Our interferometer measures the correlation between two electric fields, which are zero mean, and in the case of noise, uncorrelated. Let us draw two sets of 2000 electric field measurements ( $E_i, E_j$ ) from a unit width, zero-mean normal distribution, compute the FX correlation, followed by the delay transform, and finally the expectation (Eq. 5.12) averaging of  $N$  samples in a single delay bin as described in Eq. 5.11. As we see in Fig. 5.3, the power does indeed decrease exactly as expected all the way to  $N = 2000$ . As we examine real data let us decouple from the question of the magnitude of the instantaneous sensitivity which as discussed in §4.3.2 depends on  $T_{sys}$  and gain calibration by simply dividing out by the  $N=1$  or instantaneous power to get a unit variance power  $S_N$

$$S_N = \frac{\langle P_N \rangle}{P_1} \quad (5.14)$$

Let us now choose a single well behaved baseline and examine how various delay bins integrate down. In particular we are interested in how bins at key delays like those corresponding with the longest possible sky delays ( $d_f = 1$ ) and those very high, noise dominated delays far from any smooth spectrum foregrounds ( $d_f \gg 1$ ). For comparison we will also add the zero delay bin corresponding to power at the zenith, and several other fractional delays in useful regions. Baseline 8-12 is a 40m ( $20\lambda$ ) east-west baseline bisecting the north half of the array (see Fig. 2.4). The same baseline was used as our fiducial baseline in §1.4.2 and is approximately of the same scale and orientation as the row spacing in the high SNR grid configuration being contemplated for PAPER's EoR observations.

Not only are we curious about PAPER's ability to integrate down under optimal conditions, we should also examine the relative significance of the various issues explored above. Let us examine the degree to which various processing steps described in Chapters 3 and 4 can affect the noise statistics.

---

<sup>2</sup>Randomly chosen samples found by drawing without replacement. In practice, the entire set is randomly shuffled before summing over the first  $N$  samples.

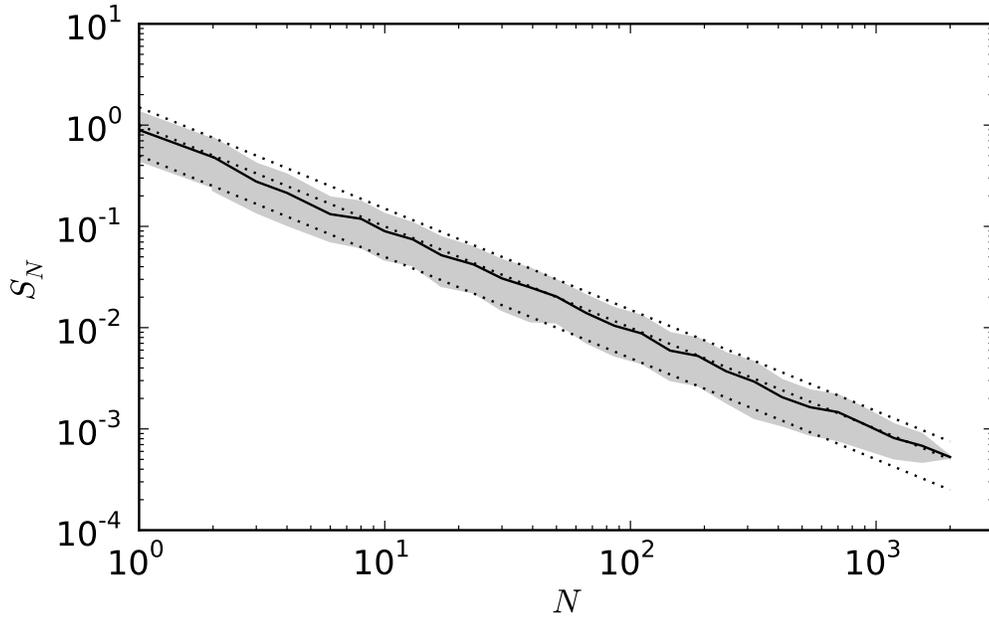


Figure 5.3: Integrating the power spectrum of a noise dominated visibility. The power (black) and error bars (grey) track  $1/N \pm \frac{N}{2}$  (dotted) as expected.

Consider the Pictor calibration field observations used to study calibration in the previous chapter. These first few observations should get us to within an order of magnitude of our  $10^{-4}$  relative sensitivity level. We will begin by considering data that have had one run of RFI removal and have been averaged to 129kHz.

As we can see in Figure 5.4 all delay bins appear to be noise dominated when only a small number of samples are included. After 6 minutes the the brightest delays begin to stabilize on the foregrounds and after an hour, the rest of the foreground channels begin to deviate from  $1/N$ . In the same figure we also see the complete power spectrum which clearly suffers from elevated power out to almost double maximum sky delay  $d_f = 1$ . To see why this is the case we must now consider more carefully the properties of cross-talk.

## 5.4 Cross Talk

To understand why we have hit a floor so early lets reconsider our simplified correlation of two signals from equation 4.8

$$\langle v_{ij} \rangle = \langle (s_i + n_i)(s_j + n_j) \rangle \quad (5.15)$$

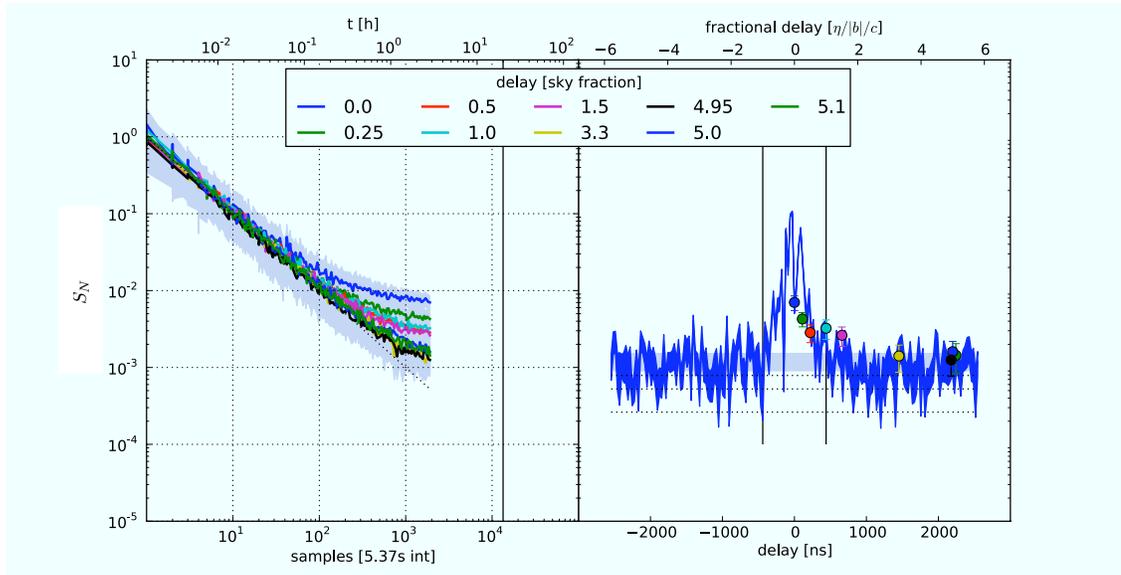


Figure 5.4: The unit variance power spectrum for LST=5h22m (Pictor field) on baseline 8,12 as a function of sample count for selected delay bins (left) and the complete spectrum with the  $1\sigma$  bootstrapped error region in blue and standard deviation in grey. Even at fractional delays as high as 1.5 the power spectrum has leveled off by the 11th day by which time even the highest delay bins appear to be deviating from a steady  $1/N$  decrease (dotted lines). Our sensitivity goal of 120 days shown as the solid line on the left. If the power spectrum at fractional delays above unity is consistent with noise the bulk of the samples (and the gray bar) will lie between the dotted lines indicating  $1/N \pm 1/2N$ .

Only the signals are correlated with each other, so given a long enough integration or high enough SNR the correlated power should be a perfect estimate of the theoretical sky signal.

$$\langle v_{ij} \rangle = \langle s_j s_j \rangle \quad (5.16)$$

As mentioned briefly in §3.3.1 a common problem at many radio wavelengths is cross-talk. When this happens, a signal from antenna i becomes contaminated by a certain amount of signal from antenna j.

$$\langle v_{ij} \rangle = \langle (s_i + n_i + \epsilon_{ji}(s_j + n_j))(s_j + n_j + \epsilon_{ij}(s_i + n_i)) \rangle \quad (5.17)$$

$$= \langle s_i s_j \rangle + \epsilon_{ji}(\langle s_j^2 \rangle + \langle n_j^2 \rangle) + \epsilon_{ij}(\langle s_i^2 \rangle + \langle n_i^2 \rangle) \quad (5.18)$$

$$= \langle s_i s_j \rangle + \epsilon_{ji} \langle v_{jj} \rangle + \epsilon_{ij} \langle v_{ii} \rangle \quad (5.19)$$

This extra term is an additive offset in the complex plane. Assuming the coupling factor  $\epsilon$  is stable, it would have a constant phase and an amplitude that varies proportionally to the autocorrelation which changes slowly with time (see Fig. 4.10).

When the signal crosses the boundary between antennae, perhaps between cables or through power lines it acquires some characteristic delay  $\delta_{ij}$ . The amplitude of the correlation depends both on the autocorrelation signal strength as well as the coupling efficiency  $\alpha(\nu)_{ij}$

$$\epsilon_{ij} = \alpha(\nu)_{ij} e^{i\delta_{ij}\nu} \quad (5.20)$$

Essentially the cross-talk can be modeled as a per-baseline gain response to the auto correlation which looks like offset in the complex plane. In many respects it is similar to the "baseline" of a single dish or spectral line observation and is removed in the same way, with a temporal low pass filter. The amplitude is proportional to the autocorrelation, which mostly varies slowly as the galaxy passes overhead except in the few cases of narrow-band RFI which must be flagged in any case.

As the true visibility is, by definition (Eq. 1.5), zero mean when integrated over a time longer than it takes for the phase to undergo a complete rotation, the residual after such an average is a good estimate of cross-talk contamination. As we can see from the autocorrelation time series in Figure 4.10, the cross-talk should be very slowly varying and repeat nightly. Indeed, after a 3.5 hour average our fiducial baseline has an average complex value (shown in Fig. 5.5) that is quite stable from night to night.

Cross-talk is most clearly visible in waterfall plots of the phase and amplitude as shown in Figure 5.6 where the eye easily picks out the constant offset and increased amplitude over the quickly varying source fringes. The effectiveness of our simple average subtraction is also clearly evident in the same figure.

Returning to the power spectrum, we can see that cross-talk removal has significantly improved outcomes here as well. Comparing Figures 5.7 and 5.4 we see that all fractional delays above unity are now much more consistent with noise with the greatest improvements at fractional delays between one and two. Despite these improvements there is still a slight deviation from linear decrease in power which grows in significance as we add in

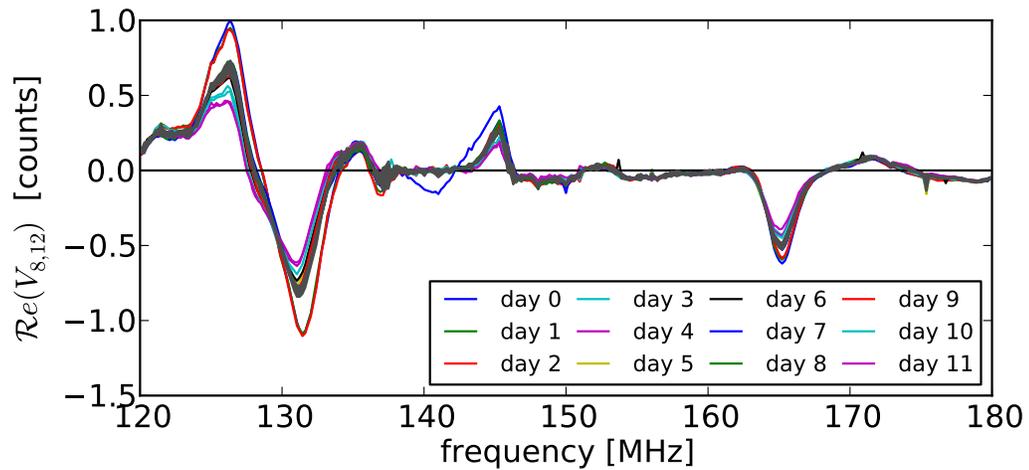


Figure 5.5: Averaging the complex visibilities over 3.5 hours we see that many channels have a stable offset in the complex plane. Here we see the real part of this average on each of our 11 days (colors). For clarity the rms error bars on each day have been omitted. The average of these measurements and the quadratic average of their error bars is shown in black with grey error. Some channels are clearly experiencing a larger, coherent, offset with a linear slope consistent with a delay-like coupling phase. Others appear to be relatively free of contamination with an average consistent with zero. The variation from day to day, excepting Day 0, is almost entirely due to gain drifts which at this stage of the pipeline have not been calibrated.

the last few data points.

## 5.5 Conclusion

As we saw in Chapter 4 the variance of the crosses does not always obey  $T_{rms} = T_{sys}$ . If this is due to terms of order  $\langle sn \rangle / \sqrt{N}$  we would expect to see this continue to integrate down as  $1/\sqrt{N}$  while the spectrum below 150-MHz would be unaffected. Steps to mitigate this will be coupled with a deeper exploration of the cross-correlation temperature by exploring the spectral and temporal behavior of  $S_N$ . The other possibility is residual in the cross-talk subtraction as the autocorrelation changes slightly over 3.5 hours. We know that the cross-talk can have an effect at high delays. Future work will focus on more careful removal of cross-talk by building a model using the autocorrelations. If it is cross-talk residuals  $S_N$  will flatten out while noise-signal cross terms will continue to integrate down, albeit more slowly.

In this chapter we explored the ability of a single baseline to integrate down. To do this we carefully implemented the delay transform to isolate power at low delays, which we verified with simulation. We then showed and verified via simulation that noise dominated correlations will integrate down as  $1/N$ . We then explored the degree to which this was true in the high delay/"noise dominated" portions of the power spectrum. By averaging the amplitude of the noise-dominated delay spectrum over an increasing number of samples we showed that minimally processed data only average down for a short time. We then took a closer look at cross-talk and found that it was primarily in the form of a complex offset proportional to the autocorrelations which we removed by subtracting a long time average. Post-cross-talk removal the rms of the correlation integrated to a much smaller value, though a slight departure from  $1/N$  was still in evidence. A crude extrapolation of the trend would imply that a much longer integration (895 days instead of 120) would be required to reach the  $10^{-4}$  level. Naturally, this result requires further study particularly by extending to more baselines.

After subtracting this nightly average we see that some of the high delay channels which formerly did not integrate down coherently, now do so (Fig. 5.7). While all high delay channels are now achieving a lower noise floor, several above a fractional delay of unity are beginning to show signs of a shallower slope, possibly indicating a departure from gaussian noise, or sky signal being scattered from lower delays by RFI flagging.

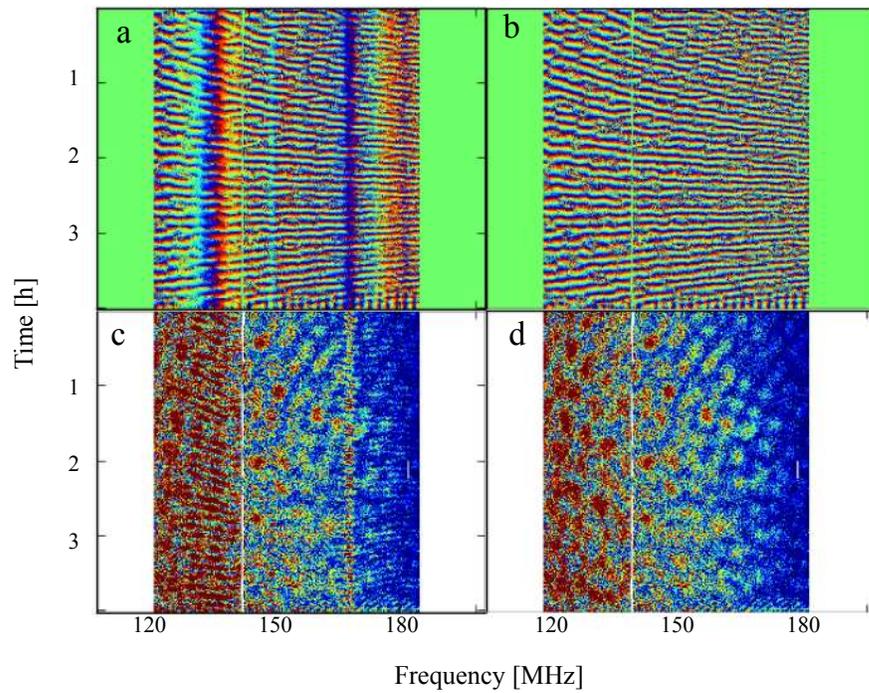


Figure 5.6: Amplitude and (left column) phase waterfalls (right column) before (top) and after (bottom) subtracting a 3.5 hour average. The eye easily picks out the constant phase offset and increased amplitude over the quickly varying source fringes (top). The effectiveness of this average subtraction operation is remarkably effective.

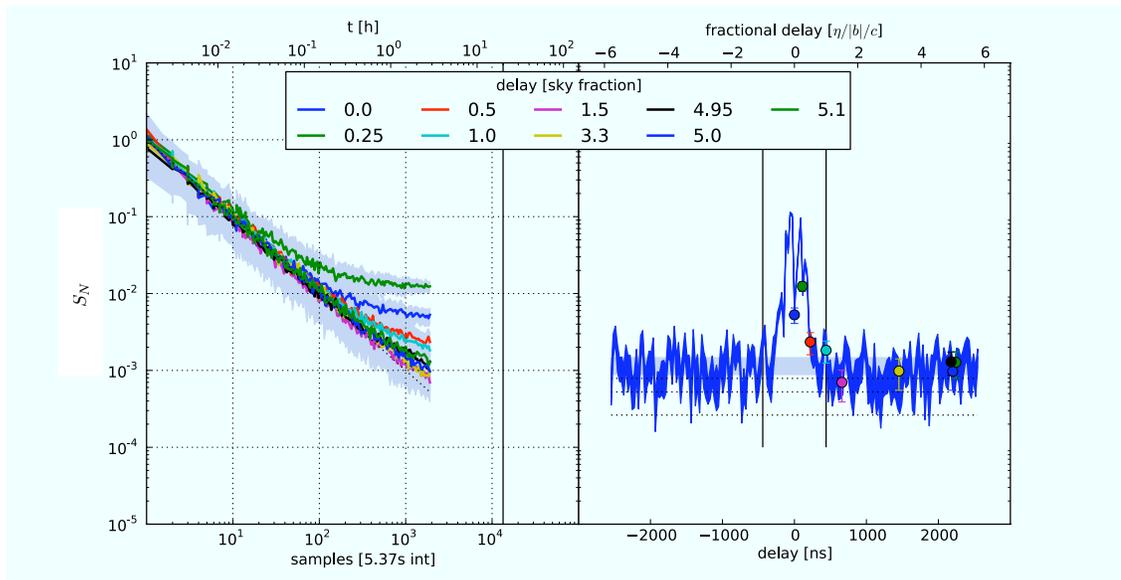


Figure 5.7: The scaled power spectrum after cross-talk removal, with axes and colors as defined in 5.7. Comparing with our previous power spectrum we see that all delays above fractional delay unity (the black lines) are now consistent with noise which is almost as low as it could be. Despite the slight elevation above the desired sensitivity we see on the left that the power level in the noise dominated bins is still integrating down albeit at a slightly slower rate. Compare the power spectrum on the right with our cartoon in Fig. 1.6.

# Chapter 6

## Conclusion

The Epoch of Reionization is in many ways the next great astronomical frontier. While the physics leading up to the event is global, the reionization of hydrogen is very likely due to stars and AGN. It is the handoff between global cosmology and local astrophysics. However astronomy at these redshifts is non-trivial. Only a handful of quasars and galaxies above  $z \sim 7$  while nothing is yet known about the hydrogen. Experiments are currently focused on measuring the power spectrum of 21cm radiation as the gas is ionized and experiences a brief period of high variance as HII regions spread through the IGM while others look for the global signal. Both seek  $z_{reion}$ , the redshift at which the IGM was  $\sim 50\%$  ionized.

As we discovered in Chapter 1 current measurements are orders of magnitude above even the most optimistic simulations but experiments (like PAPER or the MWA) are projected to have useful sensitivity levels and the ability to constrain a larger fraction of models. In the foregoing we asked one broad question: Is PAPER stable enough to integrate down to the predicted sensitivities?

We approached this question from several angles. First we noted (in Chapter 3) that imaging is a good metric of calibration and proceeded to use a small amount of data (two nights) to image the entire sky below  $+10^\circ$  declination. Extracting the fluxes of many known sources we found that their values compared well with known catalog values while noting that catalog comparison is an error-prone process. This catalog is publicly available. Next we asked if the calibration assumptions we made in this work were justified. In Chapter 4 we explored the time variability of calibration over an 11 day period and found that though the accuracy of the solution depends heavily on the sky model, there was no evidence of destructive variation in delay or amplitude calibration. We then used this calibration to experiment with imaging and to get a first look at PAPER's system temperature. The system temperature, as measured by the autocorrelations, is fairly well modeled by the known sky temperature distribution, primary beam model, and a slightly higher than expected receiver temperature. The temperature of the rms cross-correlation, the temperature that actually matters to the EoR power spectrum, also agrees with this model at low frequencies but above the middle of the measured band begins to diverge to higher values. Whether this somewhat surprising fact is due to residual signal-noise correlation or

perhaps some kind of RFI can only more investigation can say. In any case, the actual rms temperature is lower by almost half than that used to compute our sensitivity. Finally (Chapter 5) we explored the degree to which the noise "integrates down". Will we, after 120 days, realize rms levels that are smaller by  $1/(N_{\text{samples}})$ ? Though we lack 120 nights of data we can verify that after 11 nights, we are on the right track. To illustrate the importance of this question we compare the degree to which a single baseline integrates down before and after cross-talk removal. The best possible sensitivity after cross-talk filtering still retained a slight deviation from optimal sensitivity. Whether this is due to residual RFI, cross-talk or the statistical properties of noise on cross-correlated signals remains an open question.

Taken together we have learned that PAPER is quite stable with a reasonable system temperature. It is subject to cross-talk which might be manageable in post-processing but could be better handled at the instrumental level. Slight inconsistencies in the system temperature and sensitivity need further exploration. The ultimate effect on the sensitivity level is not clear but the large improvement simple processing steps like cross-talk removal can make suggest that a more comprehensive attack on the problem will earn that last bit of sensitivity and make a detection of hydrogen during the Epoch of Reionization.

# Bibliography

- Baldwin, J. E., Boysen, R. C., Hales, S. E. G., Jennings, J. E., Waggett, P. C., Warner, P. J., & Wilson, D. M. A. 1985, MNRAS, 217, 717
- Bennett, A. S. 1962, MNRAS, 125, 75
- Bernardi, G. et al. 2010, A&A, 522, A67+
- Bittner, J. M. & Loeb, A. 2011, J. Cosm. & Astr. Physics, 4, 38
- Bowman, J., Morales, M., & Hewitt, J. 2007, The Astrophysical Journal
- Bowman, J. D., Morales, M. F., & Hewitt, J. N. 2008, eprint arXiv, 0807, 3956
- Bowman, J. D. & Rogers, A. E. E. 2010, Nature, 468, 796
- Burgess, A. M. & Hunstead, R. W. 2006, AJ, 131, 100
- Condon, J. J., Cotton, W. D., Greisen, E. W., Yin, Q. F., Perley, R. A., Taylor, G. B., & Broderick, J. J. 1998, AJ, 115, 1693
- Cornwell, T. J., Golap, K., & Bhatnagar, S. 2008, IEEE Journal of Selected Topics in Signal Processing, 2, 647
- Datta, A., Bhatnagar, S., & Carilli, C. L. 2009, The Astrophysical Journal, 703, 1851
- Datta, A., Bowman, J. D., & Carilli, C. L. 2010, eprint arXiv, 1005, 4071, 12 pages, 19 Figures, submitted to ApJ
- de Bruyn, A. G., Bernardi, G., & The Lofar Eor-Team. 2009, in Astronomical Society of the Pacific Conference Series, Vol. 407, The Low-Frequency Radio Universe, ed. D. J. Saikia, D. A. Green, Y. Gupta, & T. Venturi, 3–+
- de Oliveira-Costa, A., Tegmark, M., Gaensler, B. M., Jonas, J., Landecker, T. L., & Reich, P. 2008, MNRAS, 388, 247
- Dunkley, J. et al. 2009, The Astrophysical Journal Supplement, 180, 306
- Fan, X., Carilli, C. L., & Keating, B. 2006, ARA&A, 44, 415

- Feain, I. et al. 2011, ArXiv e-prints
- Furlanetto, S. R., Oh, S. P., & Briggs, F. H. 2006, *Phys. Rep.*, 433, 181
- G. B. Taylor, C. L. Carilli, & R. A. Perley, ed. 1999, *ASP Conf. Ser.*, Vol. 180, *Synthesis Imaging in Radio Astronomy II*
- Górski, K. M., Hivon, E., Banday, A. J., Wandelt, B. D., Hansen, F. K., Reinecke, M., & Bartelmann, M. 2005, *ApJ*, 622, 759
- Gower, J. F. R., Scott, P. F., & Wills, D. 1967, *MmRAS*, 71, 49
- Halverson, N. W. 2002, PhD thesis, CALIFORNIA INSTITUTE OF TECHNOLOGY
- Haslam, C. G. T., Salter, C. J., Stoffel, H., & Wilson, W. E. 1982, *A&AS*, 47, 1
- Helmboldt, J. F., Kassim, N. E., Cohen, A. S., Lane, W. M., & Lazio, T. J. 2008, *ApJS*, 174, 313
- Högbom, J. A. 1974, *A&AS*, 15, 417
- Jacobs, D. C. et al. 2011, *ApJ*, 734, L34
- Jelić, V., Zaroubi, S., Labropoulos, P., Bernardi, G., de Bruyn, A. G., & Koopmans, L. V. E. 2010, *MNRAS*, 409, 1647
- Jelić, V. et al. 2008, *Monthly Notices of the Royal Astronomical Society*, 389, 1319, (c) Journal compilation © 2008 RAS
- Laing, R. A., Riley, J. M., & Longair, M. S. 1983, *MNRAS*, 204, 151
- Lane, W. M., Cohen, A. S., Cotton, W. D., Perley, R. A., Condon, J. J., Lazio, T. J. W., Kassim, N. E., & Erickson, W. C. 2008, in *ASP Conf. Ser.*, Vol. 395, *Frontiers of Astrophysics: A Celebration of NRAO's 50th Anniversary*, ed. A. H. Bridle, J. J. Condon, & G. C. Hunt, 370–+
- Large, M. I., Mills, B. Y., Little, A. G., Crawford, D. F., & Sutton, J. M. 1981, *MNRAS*, 194, 693
- Lidz, A., Zahn, O., McQuinn, M., Zaldarriaga, M., & Dutta, S. 2007, *The Astrophysical Journal*
- Loeb, A. & Barkana, R. 2001, *Annual review of astronomy and astrophysics*, 39, 19
- Mauch, T., Murphy, T., Buttery, H. J., Curran, J., Hunstead, R. W., Piestrzynski, B., Robertson, J. G., & Sadler, E. M. 2003, *MNRAS*, 342, 1117
- Mesinger, A., Furlanetto, S., & Cen, R. 2011, *MNRAS*, 411, 955

- Morales, M. F. 2005, *The Astrophysical Journal*, 619, 678, (c) 2005: The American Astronomical Society
- Morales, M. F. & Wyithe, J. S. B. 2010, *Annual review of astronomy and astrophysics*, 48, 127
- Mortlock, D. J. et al. 2011, *Nature*, 474, 616
- of South Africa, N. R. F. 2005, Proposal to site the Square Kilometre Array, submitted to SKA consortium
- Paciga, G. et al. 2011, *MNRAS*, 244
- Parsons, A. e. 2008, *PASP*, 120, 1207
- Parsons, A. R. & Backer, D. C. 2009, *AJ*, 138, 219
- Parsons, A. R. et al. 2010, *AJ*, 139, 1468
- Parsons, A. R., McQuinn, M., Jacobs, D. C., Aguirre, J., & Pober, J. 2011, *ApJ*, submitted, arxiv:astro-ph/1103.2135
- Penzias, A. A. & Wilson, R. W. 1965, *ApJ*, 142, 419
- Pritchard, J. & Loeb, A. 11, *Physical Review D*, 78, 103511
- Pritchard, J. R. & Loeb, A. 2008, *Phys. Rev. D*, 78, 103511
- Rau, U. 2010, PhD thesis, New Mexico Tech
- Rau, U., Bhatnagar, S., Voronkov, M. A., & Cornwell, T. J. 2009, *IEEE Proceedings*, 97, 1472
- Robertson, B. E., Ellis, R. S., Dunlop, J. S., McLure, R. J., & Stark, D. P. 2010, *Nature*, 468, 49
- Rohlfs, K. & Wilson, T. L. 1996, *Tools of Radio Astronomy* (Springer-Verlag Berlin Heidelberg New York.)
- Santos, M. G., Ferramacho, L., Silva, M. B., Amblard, A., & Cooray, A. 2010, *Monthly Notices of the Royal Astronomical Society*, 864, (c) Journal compilation © 2010 RAS
- Slee, O. B. 1995, *Australian Journal of Physics*, 48, 143
- Thompson, A. R. 1999, in *Astronomical Society of the Pacific Conference Series*, Vol. 180, *Synthesis Imaging in Radio Astronomy II*, ed. G. B. Taylor, C. L. Carilli, & R. A. Perley, 11–+
- van Weeren, R. J., Röttgering, H. J. A., Brügger, M., & Cohen, A. 2009, *A&A*, 508, 75

Vollmer, B. et al. 2010, A&A, 511, A53+

Wright, A. & Otrupcek, R. 1990, in PKS Catalog (1990), 0–+

Zahn, O., Mesinger, A., McQuinn, M., Trac, H., Cen, R., & Hernquist, L. 2010, eprint  
arXiv:1003.3455

Zaldarriaga, M., Furlanetto, S. R., & Hernquist, L. 2004, ApJ, 608, 622