Thermal Susceptibility Study of the Canadian Hydrogen Intensity Mapping Experiment Pathfinder Instruments

by

Stephanie A. N. Gerbrandt

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
BACHELOR OF SCIENCE

in

Faculty of Science

(Astronomy)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

May 2015

© Stephanie A. N. Gerbrandt 2015
Abstract

The Canadian Hydrogen Intensity Mapping Experiment (CHIME) will study Baryon Acoustic Oscillations (BAO) in the redshift range when the expansion of the Universe began to accelerate due to Dark Energy’s dominating influence. CHIME will measure the Hubble parameter, $H(z)$, and constrain the equation of state parameter, $w$, of Dark Energy. These measurements are critical in furthering our understanding of the expansion history of the Universe and Dark Energy.

CHIME will observe the faint cosmological signal and map expansion from its location at the Dominion Radio Astrophysical Observatory (DRAO) near Penticton, BC. In order to make the required measurements, the CHIME telescope requires an accurate calibration plan. Of the many components of the overall calibration plan, this paper specifically addresses the system gain calibration, with respect to exploring the relationship between system gain and temperature in the development of a thermal model. The model presented here describes the relationship to first order and lays the foundation for more detailed study. Findings provide insight into system gain configuration and facilitate subsequent development of the thermal model.

The results presented here are critical steps in the system gain calibration, contributing to a successful overall calibration plan that will ultimately lead to reliable data from which new science results will emerge. The analysis of these data has the potential to lead to an increased understanding of the expansion of the Universe and the nature of Dark Energy.
# Table of Contents

Abstract ................................................. ii

Table of Contents ......................................... iii

List of Tables ........................................... v

List of Figures ........................................... vi

Acknowledgements ......................................... viii

1 Introduction ............................................ 1
   1.1 Motivation ........................................... 1
   1.2 Theory and Background ............................... 3
       1.2.1 Acceleration and Dark Energy: from theory to observation 3
       1.2.2 Baryon Acoustic Oscillations ......................... 7
       1.2.3 Intensity mapping ................................ 10
   1.3 CHIME ............................................. 12
       1.3.1 The telescope ................................... 13
       1.3.2 Calibration and the Pathfinder analogue signal chain 15

2 System Gain Calibration ................................. 19
   2.1 Noise source derived gain .......................... 19
   2.2 Temperature dependent gain ......................... 22

3 Investigation and Results ............................... 24
   3.1 Processing and preparing the data .................. 24
   3.2 Temperature and system gain ......................... 25
       3.2.1 A first order model ............................. 28
       3.2.2 Time lag model ................................ 35
       3.2.3 Thermal susceptibility ......................... 40
   3.3 Thermal model applied to simulated visibility data 53

4 Discussion and future study .......................... 56

5 Conclusions ............................................ 58

Bibliography ............................................. 60
# Table of Contents

Appendix

A Additional Plots ............................................ 63
List of Tables

1.1 Table of parameters for the CHIME telescope . . . . . . . . . . 13

3.1 Table of average and median thermal fit values for 8 channels. . 28
List of Figures

1.1 Timeline of the Universe ........................................... 2
1.2 Evolution of matter, radiation, and Dark Energy .......... 5
1.3 Hotspots and coldspots ............................................ 9
1.4 Tracking expansion ................................................ 10
1.5 Diagram of the CHIME telescope ................................. 14
1.6 CHIME feed antenna ............................................... 15
1.7 Diagram of the Pathfinder analogue signal chain .......... 18

3.1 Noise source derived gain solution and thermometry data ... 26
3.2 Waterfall plots of noise source derived gain solution ...... 27
3.3 Time stream fit showing phase error ........................... 29
3.4 RMS of residuals for gain phase ................................. 30
3.5 Reduction in RMS of gain phase ................................. 31
3.6 Time stream fit showing fractional amplitude error ......... 33
3.7 RMS of residuals for gain amplitude ............................ 34
3.8 Scatter plots of fits to gain phase ............................... 36
3.9 Time lag and chi-square statistics for fits to gain phase ... 38
3.10 Time lag and chi-square statistics for fits to gain amplitude ... 39
3.11 Thermal susceptibility for gain phase .......................... 41
3.12 Time streams investigating fits to gain phase ............... 42
3.13 Time streams investigating fits to gain phase ............... 43
3.14 Thermal susceptibility for gain amplitude .................... 45
3.15 Time streams investigating fits to gain amplitude .......... 46
3.16 Time streams investigating fits to gain amplitude .......... 47
3.17 Scatter plots with fits to gain phase at low to mid-frequencies .... 49
3.18 Scatter plots with fits to gain phase at mid- to high frequencies .... 50
3.19 Fits (only) to gain phase at low to mid-frequencies ...... 51
3.20 Fits (only) to gain phase at mid- to high frequencies ...... 52
3.21 Map comparison of thermal model and noise source derived gains 54
3.22 Difference map of thermal model and noise source derived gains 55

A.1 Waterfall plots of gain phase (ch04-07) ....................... 64
A.2 Waterfall plots of gain phase (ch12-15) ....................... 65
A.3 Waterfall plots of gain phase (ch01-03) ....................... 66
A.4 Waterfall plots of gain phase (ch08-11) ....................... 67
A.5 Time stream fits to gain phase showing time lag ............. 68
A.6 Time stream fits to gain amplitude showing time lag ........ 69
A.7 Chi-square values for fits to gain phase (ch04-07) .......... 70
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.8</td>
<td>Chi-square values for fits to gain phase (ch12-15)</td>
<td>71</td>
</tr>
<tr>
<td>A.9</td>
<td>Chi-square values for fits to gain amplitude (ch04-07)</td>
<td>72</td>
</tr>
<tr>
<td>A.10</td>
<td>Chi-square values for fits to gain amplitude (ch12-15)</td>
<td>73</td>
</tr>
<tr>
<td>A.11</td>
<td>Thermal susceptibility for gain phase for 15 channels</td>
<td>74</td>
</tr>
<tr>
<td>A.12</td>
<td>Thermal susceptibility for gain amplitude for 15 channels</td>
<td>75</td>
</tr>
<tr>
<td>A.13</td>
<td>Scatter plots with fits to gain amplitude at low to mid-frequencies</td>
<td>76</td>
</tr>
<tr>
<td>A.14</td>
<td>Scatter plots with fits to gain amplitude at mid- to high frequencies</td>
<td>77</td>
</tr>
<tr>
<td>A.15</td>
<td>Fits (only) to gain amplitude at low to mid-frequencies</td>
<td>78</td>
</tr>
<tr>
<td>A.16</td>
<td>Fits (only) to gain amplitude at mid- to high frequencies</td>
<td>79</td>
</tr>
<tr>
<td>A.17</td>
<td>Map comparison of thermal model to noise source derived gains</td>
<td>80</td>
</tr>
<tr>
<td>A.18</td>
<td>Map comparison of thermal model to noise source derived gains</td>
<td>81</td>
</tr>
</tbody>
</table>
Acknowledgements

Many thanks are in order to both Dr. Gary Hinshaw and Dr. Adam Hincks for their guidance along the way and their many invaluable insights into the process of research. My thanks also goes out to Juan Mena for providing the gain solution data and to Dr. Richard Shaw for preparing the maps for a test of the thermal model. And a heartfelt thank you to everyone in the CHIME Collaboration and in the experimental cosmology lab for being more than willing to answer my questions and include me in your team!
He stretches out the heavens like a canopy,
and spreads them out like a tent to live in.

Isaiah 40:22
Chapter 1

Introduction

1.1 Motivation

A keen curiosity for the Early Universe has compelled scientists to design increasingly novel instruments to explore it more fully. The COBE (COsmic Background Explorer) experiment [1], the Wilkinson Microwave Anisotropy Probe (WMAP) experiment [2], and the Planck satellite [3], have provided increasingly refined data for the study of the Early Universe, in particular for the study of the Cosmic Microwave Background (CMB).

The CMB is comprised of blackbody radiation in the form of a 2D temperature map that is dated to 379,000 years after the birth of the Universe at a redshift of $z \approx 1100$ [4], when the Universe was $\approx 1/1100$ of its present size. While the CMB data from COBE and WMAP shape our understanding of the Early Universe, and Planck data [5] and the subsequent Planck data releases [3] further refine our knowledge, the results of these studies of the CMB raise questions about the nature of physical laws in our Universe. Two of the most pressing big picture questions in cosmology today relate to two periods of accelerated expansion in cosmic history. The questions are: (1) at early times, did an infinitesimally small period of accelerated expansion called inflation in fact occur?; and (2) what is the mysterious anti-gravity component, called Dark Energy, that is driving late time accelerated expansion? The brief period of inflation and the period of Dark Energy’s domination both necessitate negative pressure that drives exotic states of matter of which we currently have a limited understanding.

One of the fundamental goals in cosmology is to study and characterize Dark Energy. A more complete understanding of the Universe could be tied to understanding Dark Energy: the study of cosmic inflation and the period of accelerated expansion occurring on large scales ($\sim 100$ Mpc) today could reveal that they are somehow related; studying cosmic acceleration could have profound influence on our understanding of gravity; or, understanding the cause of cosmic acceleration could reveal long-range forces not yet discovered [6]. It has been determined that Dark Energy comprises $\approx 68\%$ [3] of our Universe and as such, measurements of its influence are of particular observational interest to cosmologists.

As labelled in Figure 1.1, it is estimated that Dark Energy became dominant much later than the CMB, at the time when expansion transitioned from decelerating to accelerating. Given that Dark Energy comprises the majority of the energy density of the Universe, characterizing it by studying its influence on cosmic expansion at the stage when its influence began to dominate will
1.1. Motivation

Figure 1.1: Diagram showing the timeline of the Universe. Note the position of the surface of last scattering, the afterglow pattern in the diagram at a redshift of $\sim 1100$. CHIME’s intent is to study Dark Energy’s increasing influence on the expansion history of the Universe, and will therefore observe in the redshift range of $0.8 < z < 2.5$, the much more recent time when expansion transitioned from decelerating to accelerating and thus, when Dark Energy’s influence on expansion became dominant. Image credit: NASA/WMAP Science Team.
further our understanding of the origins of the Universe and its early physics.

In order to obtain the data to study Dark Energy, we require a precisely calibrated telescope. The overall calibration plan involves an accounting for and removing of the system gain changes that occur in time in the telescope’s analogue signal chain, an accounting for and removing of the telescope’s interpretation of the sky signal, and a determination and removal of the bright Galactic foreground signal. The ultimate aim in all these calibrations is to obtain true sky data with sufficient signal-to-noise in order to detect the faint cosmological signal.

For the system gain calibration in particular, real-time corrections of image quality are required. The blurring of an image that results from the convolution of a telescope beam with the sky signal is further distorted by any additional, unaccounted for, and instrument-induced gain changes. Without a precise system gain calibration, it is impossible to accurately detect a signal and collect data to study the influence of Dark Energy. Thus, the system gain must be studied and analyzed in detail, in order to develop an accurate method for determining its effects on the sky data with minimal instrumental impact to the telescope site.

This paper presents statistics and in-depth study of the relationship between instrument temperature and system gain, in order to provide the evidence that will help determine if temperature is indeed a reliable predictor of gain. In the context of an overall calibration plan, a well-determined system gain calibration that minimizes radio frequency interference (RFI) can be combined with other essential instrument and data processing related calibrations, in order to obtain true sky data that allows for the detection and subsequent analysis of the cosmological signal. Studying accurately calibrated data can lead to new science results that have great potential for revolutionary impact on our understanding of the nature of cosmic expansion and its connection to Dark Energy.

1.2 Theory and Background

1.2.1 Acceleration and Dark Energy: from theory to observation

The first evidence that the expansion of our Universe was accelerating [7] both excited and confounded cosmologists. When the High-Z Supernova Search Team [7] and the Supernova Cosmology Project [8] independently used Type 1a Supernova events (SNe) to accurately determine large distances, they found that these distances were larger and the SNe were dimmer than expected. The teams confirmed that the only possible reason for the increasing velocities of objects resulting in their larger than expected distances was that the Universe itself was expanding at an accelerating rate. However, the mysterious anti-gravity component responsible for accelerating expansion does not fit in with the general understanding of gravity.

In cosmology, the expansion of the Universe is described by the Hubble
1.2. Theory and Background

parameter, $H$:

$$H = \frac{\dot{a}}{a},$$  \hspace{0.5cm} (1.1)

where $a$ is the scale factor, which describes how distances in a homogeneous, isotropic Universe expand and contract with time.

The rate of growth of the scale factor can be derived for a sphere of constant density and radius $r$, yielding the Friedmann equation in its most general form:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3c^2} \sum \varepsilon(t) - \frac{\kappa c^2}{R_0^2 a(t)^2}$$  \hspace{0.5cm} (1.2)

where $\kappa$ is the curvature, $R_0$ is the radius of curvature, and a sum is taken over the energy density, $\varepsilon(t)$, for each component in the Universe.

Accelerating expansion can be determined similarly by using Newtonian mechanics and the equation of motion for a test mass under the gravitational pull of a sphere of constant density and radius $r$, yielding:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3c^2} (\varepsilon + 3P) + \Lambda$$  \hspace{0.5cm} (1.3)

where $\Lambda$ is the cosmological constant describing Dark Energy’s influence, and pressure, $P$, is defined in terms of a fixed equation of state parameter, $w$, and energy density, $\varepsilon$, for each component:

$$P = w\rho.$$  \hspace{0.5cm} (1.4)

Equation 1.3 describes expansion for scales $> 100$ Mpc and illustrates that when the cosmological constant dominates, expansion is accelerating, with $\ddot{a} > 0$. For the cosmological constant to dominate, the equation requires $w < -1/3$. Most recent measurements yield $w \sim -1$ [3], which is consistent with what is expected for a cosmological constant. One of the goals of Dark Energy experiments such as CHIME is to measure deviations of $w$ from $-1$ [9].

Moreover, the fundamental relationship between the wavelength of electromagnetic radiation, redshift, and the scale factor facilitates observations and underpins our understanding of the cosmos. Redshift – the change in an observed wavelength’s spectral line relative to its rest wavelength – is related to the scale factor in the following way:

$$z = \frac{\Delta \lambda}{\lambda} = \frac{\lambda_{\text{obs}} - \lambda_{\text{emit}}}{\lambda_{\text{emit}}} = \frac{a_{\text{obs}}(t)}{a_{\text{emit}}(t)} - 1$$  \hspace{0.5cm} (1.5)

For the present-day scale factor, $a_{\text{obs}}(t_0) = 1$, Equation 1.5 then becomes $1 + z = a^{-1}$. Because of this relationship, measurements at different redshifts are measurements of expansion over time.

The time evolution of the scale factor is dependent on the mass-energy composition in the Universe. Each component of the Universe – matter, radiation, the cosmological constant, and curvature – can be described in terms of density parameters:

$$\Omega = \frac{\rho}{\rho_{\text{crit}}},$$  \hspace{0.5cm} (1.6)
1.2. Theory and Background

Figure 1.2: Evolution of matter, radiation, and dark energy densities as a function of redshift. The cyan band denotes an equation of state parameter, \( w \), of \(-1\). As the Universe expands, it changes from radiation-dominated to matter-dominated, and eventually to the cosmological constant-dominated era of today, transitioning in energy density from \( \varepsilon \propto a^{-4} \) to \( \varepsilon \propto a^{-3} \) to \( \varepsilon \propto a^{0} \). Image credit: [6].

dimensionless ratios that compare the mass (energy) density of a component and the critical mass (energy) density as determined for a flat, static Universe, where

\[
\rho_{\text{crit}} = \frac{\varepsilon_{\text{crit}}}{c^2} = \frac{3H^2}{8\pi G}. 
\] (1.7)

For the present day, \( H \) can be replaced by \( H_0 \). \( H_0 \) is the Hubble constant, also known as the constant of proportionality between the proper distance to a galaxy and its velocity. The law of expansion of the Universe was first discovered by Edwin Hubble through measuring distances to nearby galaxies. The Hubble constant is determined as the slope of the relation between the distance to galaxies and their recession velocities, a relationship known as Hubble’s Law: \( v = H_0D \). Observations of 600 Cepheid variable stars using the Hubble Space Telescope (HST) have yielded a Hubble constant of \( \sim 73.8 \text{ km s}^{-1} \text{ Mpc}^{-1} \) [10]. A more recent determination of the Hubble constant by \textit{Planck} yielded a value of \( \sim 67.6 \text{ km s}^{-1} \text{ Mpc}^{-1} \) [3].

Each component of the Universe will have a different effect on the expansion of the Universe: while the matter component brings objects closer together through gravitational interaction on smaller scales, the much larger component of Dark Energy, which is attributed to vacuum energy, is currently dominating in its efforts to pull objects apart on very large scales. In the standard \( \Lambda \text{CDM} \) model, the density of matter is \( \Omega_M \sim 0.31 \) and the density of
1.2. Theory and Background

The cosmological constant (vacuum) is \( \Omega_\Lambda \sim 0.69 \) [3]. The effects of each component can be understood using the equation of state shown in Equation 1.4 together with the fluid equation as derived from the first law of thermodynamics:

\[
\dot{\varepsilon} + 3\frac{\dot{a}}{a}(\varepsilon + P) = 0. 
\] (1.8)

The time derivative of energy density and the scale factor can be separated and integrated, in order to derive the relationship between energy density, scale factor, and the fixed equation of state parameter, \( w \), for each component. Integration yields the following relation:

\[
\varepsilon_w(a) = \varepsilon_{w,0}a^{-3(1+w)}, \quad \text{(1.9)}
\]

where \( \varepsilon_{w,0} \) is the energy density of a particular component at present. For matter, \( w = 0 \); for radiation, \( w = 1/3 \); and for \( \Lambda \), \( w = -1 \) [3].

Given the fixed value of the equation of state parameter for each component, the power of the scale factor, \( a \), can be determined for each component using Equation 1.9: for radiation, \( \varepsilon \propto a^{-4} \); for matter, \( \varepsilon \propto a^{-3} \); and for \( \varepsilon \propto a^{0} \), respectively. By dividing through by the current critical density, the Friedmann equation can be rewritten in terms of the density parameters for radiation, matter, and \( \Lambda \). Assuming a flat Universe (\( \kappa = 0 \)) as described by the \( \Lambda \)CDM model, and substituting the energy densities into the Friedmann equation, it is then rewritten:

\[
H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \Omega_{m,0}a^{-3} + \Omega_{\Lambda,0}a^{0} + \Omega_{r,0}a^{-4} \quad \text{(1.10)}
\]

The Friedmann equation can be further rewritten in terms of \( z \) using Equation 1.5, yielding:

\[
\frac{H(z)^2}{H_0^2} = \Omega_{m,0}(1 + z)^3 + \Omega_{\Lambda,0} + \Omega_{r,0}(1 + z)^4 \quad \text{(1.11)}
\]

The time evolution of the scale factor is controlled by the dominant energy form and its constant, \( w \): \( a(t) \propto t^{2/3(1+w)} \). Figure 1.2 shows the evolution of matter, radiation, and Dark Energy as a function of redshift, showing the redshift at the time of radiation domination at \( z \gtrsim 3000 \), matter domination at \( 0.5 \lesssim z \lesssim 3000 \), and \( \Lambda \) (cosmological constant) domination at \( z \lesssim 0.5 \), which correspond to the energy densities determined using Equation 1.9. As the Universe expands, it changes from radiation-dominated to matter-dominated, and eventually to the cosmological constant-dominated era of today. Since the scale factor changes as a function of time in the era of domination for each component, it is determined by solving the Friedmann equation separately for each component: for the radiation dominated era, \( a(t) \propto t^{1/2} \); for matter dominated era, \( a(t) \propto t^{2/3} \), and for the Dark Energy dominated era, \( a(t) \propto e^{Ht} \).

One can determine the scale factor at which the energy densities of matter and \( \Lambda \) were equal by equating the energy densities of matter and \( \Lambda \) and dividing through by the critical density, and using the density parameter values
1.2. Theory and Background

determined from the $\Lambda$CDM model, yielding:

$$a_{m,\Lambda} = \left( \frac{\Omega_{m,0}}{\Omega_{\Lambda,0}} \right)^{1/3} = \left( \frac{0.3}{0.7} \right) \approx 0.75.$$ \hspace{1cm} (1.12)

The simple relationship between the scale factor and redshift of Equation 1.5 then yields the relatively recent redshift of matter-$\Lambda$ equality, $z_{m,\Lambda} \approx 0.33$.

The time of Dark Energy’s transition to dominance occurred earlier and one can also determine the scale factor of the transition to be:

$$a_{tr} = \left( \frac{\Omega_{m,0}}{2\Omega_{\Lambda,0}} \right)^{1/3} = \left( \frac{0.3}{2(0.7)} \right) \approx 0.60,$$ \hspace{1cm} (1.13)

which corresponds to a redshift of $z_{tr} \approx 0.67$. The effect of Dark Energy on Hubble expansion is dominant at $z \lesssim 1$ and becomes negligible at $z \gtrsim 2$ [11]. The CHIME redshift range of $0.8 < z < 2.5$ is chosen to probe this intriguing time in the history of the Universe when Dark Energy began to govern at large scales. The measurements of the Hubble parameter, $H(z)$, and equation of state parameter, $w$, constrained by the current variety of independent surveys – including supernovae surveys [7] [8], galaxy surveys [12] [13], CMB temperature anisotropies [2] [3], and polarization and weak lensing [3] – are further constrained by measuring $H(z)$ in the redshift range where Dark Energy began to dominate. Thus, it is anticipated that measurements of the Hubble parameter encompassing this transition will provide further insight into the nature of Dark Energy.

1.2.2 Baryon Acoustic Oscillations

Long before Dark Energy became dominant, quantum fluctuations seeded by the initial cosmic event expanded to exponential scales in a very short period of time called inflation (labelled in Figure 1.1), which is estimated to have lasted from $10^{-36}$ seconds to between $10^{-33}$ to $10^{-32}$ seconds after the Big Bang. At the end of inflation, quantum fluctuations had instigated sound waves, termed Baryon Acoustic Oscillations (BAO) that traveled in an expanding and cooling Universe as long as they were supported by the ionized plasma.

The extremely high temperatures during the infinitesimally short period of inflation precluded any possibility for atoms to exist on their own. Matter was tightly bound to photons, since at ultra high temperatures the typical energy of a photon was considerably larger than the ionization energy of a hydrogen atom. The coupling of matter and radiation during this era resulted in an opaque Universe [14]. Every time a proton tried to form an atom with an electron, a photon bumped into it and prevented the combination, ionizing the would-be atom. For the next 379,000 years, the Universe was a sea of frequently colliding particles through which BAO propagated during the expansion and cooling of the plasma. The proton-electron plasma (baryons) and photons behaved as a “baryon-photon fluid”, the baryons generating the inertia and the photons sustaining the pressure to propagate them [15]. However, the expansion and cooling led to a point in time when the photons suddenly
lost the energy that they required to keep would-be atoms from forming. The decoupling of radiation and matter in this way (a consequence of the lower temperatures) allowed for electrons and protons to combine to form neutral hydrogen atoms for the first time. This time in the history of the Universe is termed the Epoch of Recombination [14].

Without a significant number of free electrons, the photons were finally set free to stream through the Universe [4], making it to our detectors today. At this point there was not enough pressure support available to sustain the propagation of BAO, since photons were no longer coupled to matter, and so the Universe suddenly changed from ionized and opaque to neutral and transparent. Since the low energy of CMB photons prevented them from ionizing hydrogen, and the density of free electrons had sharply dropped, the decoupling of radiation and matter was a sudden process [4]. This sudden freezing of pressure waves occurred 379,000 years after inflation, and the CMB is an imprint of these frozen waves. The CMB is also referred to as the surface of “last scattering”, since at this time a typical CMB photon was last scattered from an electron before reaching observers today [4]. Note the location of the last scattering surface in Figure 1.1.

One can think of a single BAO as the ripples produced by throwing a rock into a pond and the rock itself as a quantum density fluctuation imparted by inflation. When a rock is thrown into a pond, energy transferred to the water causes circular waves to propagate outwards. There are two possible scenarios in this analogy. One case is when many rocks are thrown in at different times, resulting in ripple circles with different radii. The second case is where many rocks are thrown in at the same time, causing each of them to have a ripple circle with a common radius, such that the average of their ripple circles is easily seen. Like in the latter case, in the CMB data (monopole and dipole contributions removed), the average of many temperature fluctuations – “hotspots” and “coldspots” on the scale of $\delta T/T \sim 10^{-5}$ – produces a prominent ring structure which is shown in Figure 1.3.

Brightness fluctuations of the photons from the CMB are seen as temperature anisotropies, and correspond to variations in the density of modes at a given instant in time [15]. They are the evidence of oscillations instigated by overdensities in the Early Universe. The oscillations can be likened to a driven harmonic oscillator in the baryon-photon fluid: the driving force being gravity, the inertia provided by the baryons, and the restoring force provided by the photon pressure [15]. Overdense modes appear brighter and hotter; underdense modes appear colder. The tiny variance of the CMB reveals the very significant fact about the time before it formed, that many quantum density fluctuations all began simultaneously propagating through the plasma following inflation, which is consistent with the picture of a tiny inflation era that created them.

The distance spanned by the pressure waves (BAO) from their initial propagation outward to when they stopped is termed the “sound horizon”, or the BAO scale, and is calculated to be 150 Mpc across, a characteristic comoving scale. Since the acoustic waves are “frozen in” after the Epoch of Recombina-
1.2. Theory and Background

Figure 1.3: Full sky map of the CMB anisotropy (temperature only, polarization not shown) is shown in mollweide projection (3D map projected onto a 2D surface) on the right, with $0.2^\circ$ angular resolution. The averaged circular shape and size of hotspots (top left panel) and coldspots (bottom left panel) represent a temperature fluctuation [16] in the map. The prominent ring structure results from an average of all the hotspots and an average of all the coldspots from the entire map and indicates that all the quantum density fluctuations imparted by inflation occurred simultaneously. The map is a picture of the end of the Epoch of Recombination after expansion and cooling of the Universe, revealing a cool Universe with an average CMB photon temperature of $\sim 2.7$ K. Image credit: WMAP Collaboration.
1.2. Theory and Background

Figure 1.4: Diagram (not to scale) illustrating the concept of tracking of expansion forward through time since the CMB, as shown from right to left. The CMB map in the rightmost image dates from 13.7 Gyr ago, the middle galaxy map dates from 5.5 Gyr ago, and the leftmost galaxy map dates from 3.8 Gyr ago. Expansion can be charted over time by charting the BAO scale; the BAO scale in turn, can be tracked by measuring diffuse gas from unresolved galaxies on very large scales. Image credit: Eric Huff, the SDSS-III team, and the South Pole Telescope team. Graphic by Zosia Rostomian. http://newscenter.lbl.gov/2012/03/30/boss-first-results/#sthash.bLyua2Uf.dpuf

Figure 1.4: Diagram (not to scale) illustrating the concept of tracking of expansion forward through time since the CMB, as shown from right to left. The CMB map in the rightmost image dates from 13.7 Gyr ago, the middle galaxy map dates from 5.5 Gyr ago, and the leftmost galaxy map dates from 3.8 Gyr ago. Expansion can be charted over time by charting the BAO scale; the BAO scale in turn, can be tracked by measuring diffuse gas from unresolved galaxies on very large scales. Image credit: Eric Huff, the SDSS-III team, and the South Pole Telescope team. Graphic by Zosia Rostomian. http://newscenter.lbl.gov/2012/03/30/boss-first-results/#sthash.bLyua2Uf.dpuf

tion, they have left an imprint on the matter distribution of the Universe and thus the sound horizon can be used as a “standard ruler” in cosmology [9], an object of fixed intrinsic size. This BAO scale is a basic Dark Energy probe of distance [17]: by measuring it in the redshift range of transition from matter to Dark Energy domination, one can infer the evolution of the equation of state for Dark Energy [11] and thus procure insight into the nature of Dark Energy.

1.2.3 Intensity mapping

Starting from the CMB, expansion can be tracked by measuring the BAO scale forward in time across a range of cosmological redshifts, as illustrated in Figure 1.4. Given the relationship between the wavelength of radiation and redshift, and the relationship of redshift to the scale factor and shown in Equation 1.5, measurements of expansion can be determined by observing at different redshifts.

Spectroscopic galaxy redshift surveys have been successful in using BAO to probe the expansion history of the Universe. Galaxy surveys such as the
1.2. Theory and Background

Sloan Digital Sky Survey (SDSS) search for the BAO signal by studying the distribution of matter and finding whether the sound horizon separates a large number of galaxies. The WiggleZ Dark Energy Survey [18] and Sloan Digital Sky Survey-Luminous Red Galaxies (SDSS-LRG) [19] independently establish a flat $\Lambda$CDM model. Analysis of the Baryon Oscillation Spectroscopic Survey (BOSS) data, as part of the SDSS, also finds continued support for a flat Universe with a cosmological constant [12]. Galaxy correlation functions measured from WiggleZ, the 6 degree Field Galaxy Survey (6dFGS), and SDSS-LRG show independent evidence for the baryon acoustic peak [18].

While the combined power spectrum and correlation function analyses provide a powerful estimate of the cosmological distance scale, as well as an independent measure of the cosmological parameters, the two-point galaxy correlation statistic requires a substantial amount of data to obtain precision. Spectroscopic galaxy redshift surveys require millions of galaxies [20] and countless observation hours, requiring that spectra and a redshift be determined for each galaxy. The surveys also necessitate expensive space-based telescopes, since the visible wavelengths are redshifted to infrared wavelengths and are thus largely blocked by Earth’s atmosphere [20].

In comparison, intensity mapping is a measure of the collective emission of many galaxies at low-resolution [21] with no need for individual detections. Intensity mapping presents an independent approach to capture the BAO scale that makes use of the 21 cm spin-flip transition line of neutral hydrogen (HI) and maximizes volume [9]. In a hydrogen atom, the parallel spin state for the electron and proton emits a photon of 21 cm wavelength. Although this high energy spin state is a rare occurrence for hydrogen (occurring once every 100 Myr), the abundance of hydrogen in the Universe ensures that there are plenty 21 cm photons to study.

The BAO scale can be determined by measuring the intensity of the abundant HI gas on large scales. Ubiquitous, equal-sized density enhancements left over from the propagating acoustic wave fronts [21] that were frozen in at the surface of last scattering are observed as overdensities (see Figure 1.4) of diffuse HI gas of galaxies along the BAO “ring-like” structure at different redshifts.

The 21 cm HI line is an isolated transition at frequencies less than the rest frame frequency and so HI density vs. redshift can be inferred directly from HI intensity vs. frequency [21], effectively measuring the 3D density of matter as a function of redshift along the line of sight [22]. Since brightness measurements are proportional to the density of HI, 21cm brightness data can be stacked to see the BAO scale. Both the angular and redshift space BAO wavelength are measured at each wavelength, tracing out expansion history [21]. Thus, observing the large scale HI emission from unresolved galaxies allows one to obtain an enormous volume of data in a relatively short period of time.

Since Dark Energy influences the expansion rate of the Universe, it also affects the development of structure when density perturbations are growing [6]. Therefore, measuring the HI as a tracer of structure at different redshifts also constrains the Hubble parameter, which can illuminate our understanding
1.3 CHIME

of Dark Energy’s influence.

In order to measure structure, cosmological distances must be understood. Objects of known size with measurable angular diameters can be studied and their distances recovered. In a homogeneous, isotropic Universe, the Friedmann-Roberston-Walker metric for a photon’s path on a null geodesic provides the definition of the proper distance, $\ell$, as measured today:

$$\ell = c \int_{t_e}^{t_0} \frac{dt}{a(t)}$$

(1.14)

To help understand distances in an expanding universe, one can think of the Universe as a grid. The size of objects moving with a hypothetical grid are measured in grid units, or a comoving distance, $\chi$. Any proper distance, $\ell$, is defined as the comoving distance times the scale factor. As the grid expands radially and transversely, the evolution of the measured angular diameter as a function of redshift can be determined using the familiar small angle formula and substituting the proper distance with the product of comoving distance and the scale factor:

$$\theta(z) = \frac{\ell}{d_A} = \frac{a[\chi_{BAO}]}{d_A}(z)$$

(1.15)

where $d_A$ is the angular diameter distance and $\chi_{BAO}$ is the comoving BAO scale. The angular diameter distance, $d_A$, in turn, can be determined with the following integral:

$$d_A \propto \int_0^z \frac{dz'}{H(z')}$$

(1.16)

where $H(z')$ is the Hubble parameter as a function of redshift. $H(z)$ can be extracted at different redshifts, using intensity mapping as an observational tool. In this manner, its value can be constrained, together with other independent surveys, as discussed in Section 1.2.1.

1.3 CHIME

As a complement to the early picture of the Universe from the CMB, the Canadian Hydrogen Intensity Mapping Experiment (CHIME) located at the Dominion Radio Astrophysical Observatory (DRAO) near Penticton, BC, will probe a more recent time in the Universe, producing full sky maps to measure BAO in the redshift range of 0.8 < $z$ < 2.5 (corresponding to a bandwidth of 400-800 MHz).

CHIME data will contribute to the effort of constraining the Hubble parameter and equation of state parameter to a higher degree of accuracy in order to understand Dark Energy’s role in the accelerating expansion of the Universe.
1.3. CHIME

1.3.1 The telescope

The physical properties of BAO drive the main physical requirements of the CHIME telescope: angular resolution, sensitivity, and collecting volume. Various parameters describing the CHIME telescope are included in Table 1.1.

<table>
<thead>
<tr>
<th>Location</th>
<th>DRAO (49°19'N, 119°37'W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of cylinders</td>
<td>4</td>
</tr>
<tr>
<td>No. of feeds</td>
<td>256 cylinder⁻¹</td>
</tr>
<tr>
<td>Redshift range</td>
<td>2.5 - 0.8</td>
</tr>
<tr>
<td>Bandwidth</td>
<td>400-800 MHz</td>
</tr>
<tr>
<td>Frequency resolution</td>
<td>0.39 MHz</td>
</tr>
<tr>
<td>E-W F.O.V</td>
<td>2.5° - 1.3°</td>
</tr>
<tr>
<td>N-S F.O.V</td>
<td>~ 90° about zenith</td>
</tr>
<tr>
<td>Angular resolution</td>
<td>0.52° - 0.26°</td>
</tr>
<tr>
<td>Spatial (map) resolution</td>
<td>10 - 50 Mpc</td>
</tr>
</tbody>
</table>

Table 1.1: Some features of the CHIME telescope.

Angular resolution

To sufficiently resolve the BAO sound horizon, the telescope must see 1/10 of the BAO scale, which constrains it to have a low resolution range between 0.26° and 0.52° [16], which in turn, constrains the physical dimensions of the telescope to be at least 80m × 100m, representing a substantial collecting area [22].

Figure 1.5 shows the CHIME telescope, which consists of an array of contiguous North-South oriented cylindrical dishes, with 256 detectors along each focal line. The detector is a dual-polarized feed antenna with North-South and East-West polarizations, and is shown in Figure 1.6. Although the parabolic shape of each cylinder is designed to obtain sufficient resolution in the East-West direction, the CHIME beams cannot be resolved in the North-South direction using the geometry of the cylinders; instead, resolution is obtained from the correlation and averaging of multiple feed pairings termed visibilities, the instantaneous correlations between pairs of feeds. The number of feed pairings is determined by the resolution requirement for mapping BAO. The requirement for sensitivity to brightness intensity sets the required number of detectors (or feed antennas) per cylinder to 256. With two polarizations per feed, a sophisticated correlator is required to correlate and average the raw data from many channels. Since the variations in BAO hotspots and coldspots are fluctuations on the scale of $\delta T \sim 10^{-5}$ K [4], even using a fast telescope tracking rate (where every feed sees half the Celestial Sphere each day), the data will have to be collected over at least a two year period in order to obtain the required sensitivity.
1.3. CHIME

- **400-800MHz band**
- **21cm from z ~ 0.8 - 2.5**
- **Resolution: 1MHz, 13-26**
- **3rd BAO peak resolved**
- **Drift scan, no moving parts**
- **20,000 deg**
- **1280 Dual-polarization feeds**
- **Cosmic-variance-limited survey**

**Figure 1.5:** Diagram of the full CHIME experiment. 256 feeds are positioned equidistant from each other along each cylinder’s focal line. The telescope is approximately 80×100m in area; 5 cylinders are shown, but it has recently been redesigned for 4 cylinders. On the far left of the diagram, a person indicates the overall scale of the telescope. Image credit: CHIME Collaboration.

**Sensitivity**

Since CHIME is a spatial interferometer, it preserves phase data along with total intensity data. A plane wave arriving from the sky at an array of detectors will not retain the same phase after being recorded by the instruments. However, the interferometer corrects for this wavefront distortion by reconstructing the wave in phase and amplitude. The telescope also relies on the principles of spherical harmonics and Fourier transforms. A map of the sky is achieved by recording the differences in phase from incoming electromagnetic waves arriving at pairs of feeds located at various physical separations. Each pair of signals gives a corresponding Fourier component of the image on the sky, with the differences in phase allowing for the determination of source direction. Thus, with complex data - unique pairs of intensity and direction data - one can create a map of the sky. To map the Celestial sphere, one must first transform to a spherical harmonic basis in frequency-space, where data in each direction can be weighted and summed, and then inverse transformed and averaged, in order to generate a map of the sky [23].

**Volume**

The CHIME telescope requires a volume large enough to obtain many independent BAO-scale fluctuations [22] in order to calculate a robust statistical average of the BAO scale. Since CHIME is a transit telescope (it has no moving parts) and is situated in the Northern Hemisphere, it can map the northern sky in each sidereal day of Earth’s rotation. The large volume of data required to precisely measure the faint BAO signal is obtained by both observing half the sky transiting zenith each day and by observing in CHIME’s redshift range.
1.3.2 Calibration and the Pathfinder analogue signal chain

The CHIME Pathfinder is a prototype instrument – a smaller version of the full CHIME project and used for developing calibration, testing hardware, and performing data analysis. The Pathfinder is comprised of two parabolic cylinders of total dimensions $\sim 40 \times 40$ m. A diagram of the Pathfinder is shown in Figure 1.7.

One of the main purposes of Pathfinder testing is to determine an appropriate calibration plan, allowing for the determination of true sky values from measured voltages [16]. The precise calibration of the CHIME telescope is vital to obtaining true sky data in order to map expansion and characterize Dark Energy. However, the Galactic foreground synchrotron emission drowns the faint BAO signal by a factor of $\sim 10^5$ and thus represents the most significant challenge to calibration and post-calibration data-processing. Therefore, the foreground removal appropriately drives the calibration requirements and necessitates an effective and precise calibration plan. Since the telescope mixes frequency and spatial structure [16], in order to remove the extremely bright foreground signal and measure the faint BAO scale, the CHIME beam must be known to 0.1% (together with the system gain for each feed which must be known to 1%) [24]. Given that the fluctuations of the large-scale structure hide the variation of gas density of BAO, a Fourier analysis of large regions of the sky is required [9]. Since the foregrounds are spectrally smooth, they can be subtracted spectrally.

There are many aspects of calibration that must be addressed, although three calibration types are significant for the overall calibration plan:

- *System Gain calibration*. Measured sky data are separated from gain changes in time in the analogue signal chain, which thus far has been determined by using a noise source injection. This calibration is the focus of this paper;

Figure 1.6: Image of CHIME feed, a dual-polarized antenna. Image credit: CHIME Collaboration.
1.3. CHIME

- **Polarization calibration.** Since the telescope beam is convolved with the sky data, in order to obtain the true sky data the beam must be characterized and removed from the sky measurements. True sky data are obtained by using pulsars to remove the telescope’s interpretation of the measurements; and

- **Beam calibration.** This is a data-processing challenge. It is critical that the beam pattern as a function of frequency be understood, in order to accurately remove foreground.

The analogue signal chain, as shown in Figure 1.7, begins with the telescope’s cylinders, which collect the faint radio signal and direct it to the antenna feeds along the focal line. Detecting the faint signal requires that low-noise amplifiers (LNAs) be placed immediately after each of the antenna feeds. After initial amplification, the signal travels for \(~60\) m of cable and along the way is attenuated by some unquantifiable amount. The signal is then further amplified in the filter low-noise amplifiers (FLAs) located in the Radio Frequency (RF) room, and makes its way to the correlator and analogue-to-digital converter (ADC). All the gain changes in time experienced by the signal along the analogue chain are termed system gain. Without performing a system gain calibration, it is impossible to quantify the changing system gain along its path from antenna feed to correlator.

The system gain can be determined by using a noise source injection: a known, blinking signal is used to separate the sky signal from the gain changes in time. The noise source is a key component in the system configuration. It is a blinking circuit that consists of three components: (1) a diode (operating in reverse-bias mode to create random signal when biased to its breakdown voltage), (2) amplifiers that are continuously powered on to maintain stability, and (3) a radio frequency (RF) switch that turns the signal to the noise source antennas on and off at regular intervals. The noise source itself is positioned in the Faraday screened RF room, as shown in Figure 1.7, and its signal is split three ways: it is simultaneously injected into two noise antennas, one centred at the vertex of each cylinder, and is also directly injected into the correlator to serve as the reference signal.

In the Pathfinder configuration, each cylinder is fitted with a cassette of four feeds along the focal line. The feeds detect the noise source and sky signal and amplify them using LNAs before they are sent via coaxial cable to the FLAs and the correlator and ADC in the RF Room. A delay cable for the noise source reference signal ensures that the reference signal and sky signals are received simultaneously by the correlator. The correlator is a Field Programmable Gate Array (FPGA) that outputs a matrix of data: the auto-correlate visibility data on the diagonal, and cross-correlate visibility data on the off-diagonal. Figure 1.7 outlines a simple 3-channel example of the Pathfinder; however, the gain data used for this study were obtained from a Pathfinder configuration of 16 channel inputs into the correlator. In addition, the configuration used to obtain the gain data for this study had the noise source antenna positioned at the end of the cylinders.
A stable, blinking noise source is an excellent calibrator: it provides a relative calibration for each channel, allowing a comparison of the response of sky channel with a reference channel. The “on/off” switching of the noise source permits the separation of data containing sky signal, noise source, and gain changes in time (the “on” signal) from data containing just the sky signal and gain changes in time (the “off” signal). When taking the difference of the complex “on/off” matrices, the sky signal falls out, leaving only the noise source signal and the gain changes in time. Since the output of the noise source is a known signal (it has a dedicated reference channel in the correlator), the amount of attenuation, amplification, etc. that is experienced in time along the analogue signal chain – what is termed a gain solution – can be calculated by diagonalizing the matrix using a singular value decomposition (SVD). Thus, the noise source is a useful tool to determine the otherwise unquantifiable gain changes in time that both the noise signal and sky signal simultaneously experience on the path from the antenna feeds to the correlator and ADC.

Once a gain solution, $g(t)$, is determined through using the configuration of noise source reference and sky channels, it can be removed from the sky map, thus recovering a map of sky data with the system gain removed. Further calibration must then be performed to deconvolve the telescope beam from the sky map. Given that the foreground is so incredibly bright, the system gain calibration is a crucial component of the overall calibration plan that will eventually lead to successful foreground removal and BAO detection.
Figure 1.7: Diagram illustrating the analogue signal chain configuration used in determining the Pathfinder system gain calibration. The noise source signal is inputted into a splitter, and the splitter has three outputs: one leading directly to the correlator and two outputs going to the noise source antennas, one to each cylinder. When the noise source is switched “on”, the correlator receives noise source signal and sky signal combined with gain changes in time along the analogue signal chain, the noise source serving as a reference signal. When the noise source is switched “off”, the correlator receives sky signal only, together with gain changes in time experienced along the analogue signal chain. The correlator calculates the difference of these two datasets – the complex matrix “on” and complex matrix “off” signals – and determines a complex gain matrix which includes the reference noise source signal. Because the reference noise source signal is a known signal, a diagonalization of the matrix with a singular value decomposition (SVD) operation will determine what would otherwise be an unquantifiable complex system gain. This calculation is performed numerous times, for every on-off cycle of the noise source switching, which occurs faster than the sky changes in time. The delay cable for the noise source reference signal ensures that all three signals to the correlator arrive simultaneously. This configuration is used to determine the noise source derived gain solution and involves a dedicated channel for the noise source in the correlator. Such a configuration can be generalized to many more channels.

Image Credit: CHIME Collaboration.
Chapter 2

System Gain Calibration

2.1 Noise source derived gain

The dual polarized feeds detect the sky signal as an electric field. The electric field is a complex vector quantity that can be described by the following equation:

\[ \mathbf{E} = E_0 e^{i\phi} \]  \hspace{1cm} (2.1)

The phase and amplitude portions are considered when determining a system gain solution. To make a map of the sky, one requires complex values, describing unique points of intensity (brightness temperature) and direction on the sky. The ADC digitizes voltages as complex gain values – data pairs of intensity and angle \( \phi \) from the positive real axis to the electric field vector – and records them as data acquisition units. In the case of the gain solution data studied in this investigation, each data point consists of two components: an intensity value and associated phase delay. These values must also be projected onto a sphere: one must transform to frequency-space, perform the summations for each point on the sky, and then inverse transform back to obtain a map of the sky.

There are several methods used to determine the system gain through time so that it can be removed from the data in real-time. One reliable method to use is a noise source injection. CHIME has been using a noise source injection to test calibration methods to determine system gain.

Since the sky data and noise source data are both detections of electric field amplitudes, for ease of distinction between sky and noise source data, the electric field amplitude from the sky data is denoted with \( s \), and the electric field amplitude from the noise source is denoted with \( n \) in all equations. A simple example is outlined here, although it can be generalized to many more channels, into an N+1 by N+1 correlation matrix. Working with a simple three channel example, the alternating signals of noise source and sky data inputted into the correlator (in units of counts) can be written as follows:

Channel 0: \([n, 0]\)
Channel \(i\): \(g_i[s_i + n, s_i]\),

where the subscript \(i\) denotes Channels 1 or 2, respectively, and \(g_i\) represents the system gain contributed by the analogue signal chain in the path from feed to correlator/ADC. In all channels, the first term in brackets represents the noise source “on” signal and the second represents the noise source “off” signal.
2.1. Noise source derived gain

A correlation matrix can be built up from the cross-correlated terms from the three channels: three amplitudes are inputted into the correlator which then outputs a 3x3 digitized product for both “on” and “off” segments of each cycle. The following three matrices represent the “on”, “off”, and difference of “on-off” segments, respectively:

\[
\begin{bmatrix}
|n|^2 & n|g_1|^*(s_1 + n)^* & n|g_2|^*(s_2 + n)^* \\
g_1(s_1 + n)n^* & |g_1|^2|s_1 + n|^2 & g_1(s_1 + n)g_2^*(s_2 + n)^* \\
g_2(s_2 + n)n^* & g_2(s_2 + n)g_1^*(s_1 + n)^* & |g_2|^2|s_2 + n|^2 \\
\end{bmatrix}
\]  

(2.2)

\[
\begin{bmatrix}
0 & 0 & 0 \\
0 & |g_1|^2|s_1|^2 & g_1s_1g_2^* s_2^* \\
0 & g_2s_2g_1^* s_1^* & |g_2|^2|s_2|^2 \\
\end{bmatrix}
\]

(2.3)

\[
\Delta = \begin{bmatrix}
|n|^2 & n|g_1|^*(s_1 + n)^* & n|g_2|^*(s_2 + n)^* \\
g_1(s_1 + n)n^* & |g_1|^2|s_1 + n|^2 - |g_1|^2|s_1|^2 & g_1(s_1 + n)g_2^*(s_2 + n)^* - g_1s_1g_2^* s_2^* \\
g_2(s_2 + n)n^* & g_2(s_2 + n)g_1^*(s_1 + n)^* - g_2s_2g_1^* s_1^* & |g_2|^2|s_2 + n|^2 - |g_2|^2|s_2|^2 \\
\end{bmatrix}
\]

(2.4)

Using this three channel example, it is evident that when taking the difference of “on” and “off” segments, element-by-element, terms that include both the noise source and a sky channel can be omitted, since the noise source is not correlated to the sky data. In all cases, the incoherence of the noise source channel with any on-sky channel can be explained in the same way as the incoherence of any two on-sky sources: a sky source and the noise source can never be correlated since they do not oscillate in phase over time and thus their oscillations will always cancel out. The remaining complex matrix then contains only the noise source and gain terms:

\[
\Delta = \begin{bmatrix}
|n|^2 & |n|^2g_1^* & |n|^2g_2^* \\
g_1|n|^2 & |g_1|^2|n|^2 & g_1g_2^*|n|^2 \\
g_2|n|^2 & g_2g_1^*|n|^2 & |g_2|^2|n|^2 \\
\end{bmatrix}
\]

(2.5)

Since the noise source amplitude is present in all elements and is a known value (it has a dedicated reference channel in the correlator), it can therefore be pulled out as a constant, revealing that the matrix consists of gain terms only. The matrix can also be written in general form as the amplitude of the noise source multiplied by the outer product of the gains:

\[
\Delta = |n|^2 g_i \otimes g_j = |n|^2 g_i g_j^* 
\]

(2.6)

\(\Delta\) represents a single complex relative gain solution determined from one “on/off” cycle; however, gain values are constantly changing in time from one cycle to the next. Therefore, it is important to establish a relative gain solution over time, \(g(t)\). As long as the gain is constant within the cycle where it is determined – within the \(\sim 1\) second interval that the noise source switches “on” and “off” [25] – such that the blinking occurs faster than the gain drifts from cycle to cycle and faster than the sky signal changes, then the gain determined
2.1. Noise source derived gain

in each cycle can be taken as an accurate representation of what is occurring in the analogue signal chain.

The gain calibration performed on the sky data from each channel (or each feed input, $F_i$) can be written as follows:

$$F_i = g_i g_n (s_i + \alpha_i(t)n)$$ (2.7)

where $F_i$ is the feed input, $s_i$ is the sky data, $n$ is the noise source data, $g_i$ and $g_n$ refer to the gains associated with sky data obtained from feed $i$ and from noise source data, respectively, and $\alpha_i$ is a constant that is the result of the geometric effect of the position of feed $i$ relative to the position of the noise source.

The visibilities – raw data from the pairing of feed channels – can be expressed in terms of the gain calibration coefficients as follows:

$$V_{ij} = \langle F_i F_j^* \rangle$$ (2.8)

$$= |g_n|^2 |g_i g_j^* s_i s_j^* + |g_n|^2 g_i g_j^* \alpha_i \alpha_j n|^2$$, (2.9)

where the terms containing both the sky data and noise source data are uncorrelated and thus fall out of the equation. The visibilities can then be written as a difference of “on-off” cycles:

$$\Delta V_{ij} = V_{ij}|_{on} - V_{ij}|_{off}$$ (2.10)

$$= g_i g_j \alpha_i \alpha_j |E_0|^2$$ (2.11)

Taking the difference of the “on/off” cycles is critical to the gain calibration: it means that only the channel connected to the noise source (Channel 0) is a fixed source. Since by Equation 2.6, the matrix is defined as the outer product of a vector with itself, it is by definition a rank 1 matrix. With data from only one channel, the resulting matrix is rank 1: it has only one linearly independent column vector and all other column vectors associated with other channels are multiples of it. If the correlation matrix is not rank 1, it would be very difficult to determine the system gain. However, since the sky signal is removed using this method, it is expected that the largest gain will be in the noise source channel and so it is required that only the noise source channel provide data against which an evaluation of the system gain can be made.

To determine the system gain from $\Delta$, the correlation matrix is diagonalized by performing an SVD using the following equation:

$$A_{m \times n} = U_{m \times m} S_{m \times n} V_{n \times n}^\dagger$$, (2.12)

where $U_{m \times m}$ and $V_{n \times n}^\dagger$ are unitary matrices and $S_{m \times n}$ is the diagonal matrix consisting of the singular values. For the situation in which the noise source is the only fixed source, the largest value (the only non-zero value) in the resulting singular value matrix and its corresponding left singular vector represent the system gain for one cycle. In a rank 1 matrix, only one singular value is nonzero, and this largest singular value together with its corresponding left singular vector represent the gain of the system for each cycle.
2.2 Temperature dependent gain

To test the quality of the gain calibration, the ratio of the largest singular value to the next largest singular value in the singular matrix can be computed. Determining this ratio for all “on/off” cycles will help establish the quality of the overall gain solution over time. As a diagnostic, a good gain calibration is expected to yield a large ratio, indicating that the system is closer to one degree of freedom (i.e., that the matrix obtained is in fact rank 1). Smaller values indicate that the second largest eigenvalue is comparable to the first largest, and therefore the matrix cannot be rank 1.

With the system gain determined, the data can be calibrated by removing the calculated gain from the measured sky data (visibility data). An element-by-element inverse operation can be performed as follows:

\[
\langle F_i F_j^\dagger \rangle_{\text{true}} = \frac{V_{ij}}{g_i g_j^{\ast}} = |g_n|^2 \langle s_i s_j^\ast \rangle + |g_n|^2 |\alpha_j \alpha_j^\ast| n|^2,
\]

where the gain changes in time as seen in Equation 2.9 have been removed. Equation 2.13 is applied to the visibility data for each “on/off” cycle in which a gain solution was obtained.

Once the system gain calibration is determined, further calibrations can be performed, such as the polarization calibration used to remove the effects of the telescope beam convolved with the sky. These calibrations are essential to determining the beam calibration required to remove the Galactic foreground signal. Thus, the system gain calibration is an essential step in the overall calibration plan.

2.2 Temperature dependent gain

Although the system gain can be reliably determined using a noise source injection, there are a number of disadvantages to using the noise source. A noise source introduces unwanted radio frequency interference (RFI) in the DRAO radio quiet zone and also adds unwanted noise to the data and complexity to the signal chain, requiring a dedicated channel into the correlator which would otherwise be used to collect data. Therefore, it is preferable to find an alternate method of determining a gain solution, in order to minimize the use of the noise source and possibly find a substitute for it.

One such alternate method of determining a gain solution is to use a thermal model. Since the LNAs are located outdoors, it is expected that their gain will vary with ambient outdoor temperature. Using a thermal model, on-site RFI can be reduced and the signal chain simplified.

In order to verify that a thermal model is viable, both gain and temperature data must be studied to find a relationship between them. If the difference between a noise source derived gain model and a thermal model falls within the surface accuracy requirements for the project, the thermal model would be deemed acceptable to be used as an alternative method of determining system gain. In addition, in taking advantage of the periodicity in the sky temperature signal (due to the rotation of the Earth), a successfully determined temperature-based gain solution for the CHIME transit telescope can
be used to shrink daily deviations seen in the periodic sky signal which are
due to scatter produced by system gain. A good gain model will allow for
repeatable sky signals from day to day.

The principal goal of this investigation is to determine if temperature can
reliably predict system gain. Since the first stage system amplifiers (LNAs) are
located outdoors, it is expected that ambient, outdoor temperature is one of
the largest contributors to relative gain variations. (Noise source amplifiers are
excluded from this study, since the noise source “on” segment will be common
to all channels, and therefore the noise source signal cancels out of the ratio of
pairs of feed signals). A subsidiary goal for this study is to better understand
the configuration of the noise source injection method of determining system
gain.

Given the expected correlation between gain and temperature, the follow-
ing linear model is studied as a first prediction for the relationship between
gain and temperature:

\[
g(T) = g_0 + \frac{\partial g}{\partial T}(T - T_0),
\]

where \(g(T)\) is the gain solution determined using the noise source, \(T\) represents
the corresponding temperature measurements from thermistors fitted to each
metal box encasing each LNA, \(T_0\) is a fiducial temperature value of the system,
and \(\frac{\partial g}{\partial T}\) is the slope in the relationship of \(g\) and \(T\) which is measured in order
to confirm the relationship.
Chapter 3

Investigation and Results

To study the relationship between temperature and gain, complex gain solution data calculated from the Pass0E dataset (data collected September 16-20, 2014) were obtained alongside temperature data from the housekeeping data of corresponding days. The temperature sensors (thermistors) were mounted to the LNA boxes and measured ambient outdoor temperature as a function of time while gain calibration data were simultaneously collected.

A cassette of four feeds is mounted at the focal line of each cylinder, each feed having two polarizations. The correlator receives a signal from each feed polarization; therefore, a total of eight feed antennas necessitates a 16-channel correlator to correlate feed pairs and average the data. The correlator receives complex sky data from each channel (intensity and direction on the sky). When each channel is multiplied by itself, auto-visibility data are obtained; cross-visibility data are determined by multiplying each channel by every other channel. In the Pathfinder configuration for Pass0E, the gain solution was calculated from the auto-visibility data only. Eight of the 16 channels were fitted with temperature sensors. It is important to note that the noise source antenna at this time was positioned on the end of the cylinders and that the reliability of the gain solution was not verified.

3.1 Processing and preparing the data

The data used in this study are a calculated complex gain solution based on sky data collected during noise source operation. To generate the plots, the complex data are separated into phase and amplitude. Temperature data were interpolated such that for each value of the gain solution, there is an associated measured value of temperature. Temperature data were subsequently fitted separately to the phase and amplitude parts of the gain solutions based on Equation 2.14.

Prior to proceeding with the investigation, RFI frequencies were identified and masked. The RFI-contaminated frequencies were determined using corresponding visibility data from September 16-20, 2014 and masked using a CHIME-developed python function called rfi_pipeline. The timestamps of solar transits were determined using the CHIME-developed ephemeris tools in the python ch_util package. The solar contamination was then excised and excluded from the temperature fits to gain. An iterative method was employed to excise the contamination: the number of timestamps on both sides of the transit were successively increased until the spikes in time stream data surrounding the transit time were removed. This method resulted in 15
3.2 Temperature and system gain

Temperature data are fitted to the linear model of Equation 2.14, and various statistics are explored to evaluate if the linear model is a good predictor of system gain. It is important to note that the calculated gain solution data represent changes in gain for the entire signal chain, and not merely for the LNAs themselves; however, the outdoor location of the LNAs implies that the most significant thermal effect on system gain would stem from ambient temperatures affecting the gain of the LNAs.

Various plots show the relationship between gain and temperature. These plots include:

- Time stream plots of the noise source gain solution and raw temperature data, as well as waterfall plots of the noise source gain solution for each channel as an indication of the behaviour of the LNAs;
- Time stream plots of temperature fits to system gain over the three-day period;
- Root-mean-square (RMS) plots showing scatter over the bandwidth;
- Scatter plots showing temperature fits to system gain;
- Chi-square plots for determining a typical time lag;
- Thermal susceptibility plots, showing how gain varies with temperature in the bandwidth; and
- Maps of gain solutions applied to simulated sky data.

A time stream of gain phase calculated from the noise source derived solution for 8 channels at a single frequency is shown in the top panel of Figure 3.2; a corresponding time stream for temperature changes over the same three day period is shown in the bottom panel. A comparison of the panels in Figure 3.2 reveals that ambient temperature variations clearly have a strong influence on the system gain, suggesting that there is likely a linear relationship between gain and temperature. Waterfall plots of gain phase and gain amplitude for a single channel are shown in Figure 3.2 and indicate that the LNAs are uniformly behaved. Additional waterfall plots for the remaining seven channels are included in Appendix A.
3.2. Temperature and system gain

Figure 3.1: The top panel shows a time stream of the calculated gain phase from the gain solution data for a three day period for 8 LNAs at a single frequency, and the bottom panel shows the time stream of temperature from the same LNAs over the same time period. The gain data appear to be linearly related to the temperature variations.

Figure 3.1: The top panel shows a time stream of the calculated gain phase from the gain solution data for a three day period for 8 LNAs at a single frequency, and the bottom panel shows the time stream of temperature from the same LNAs over the same time period. The gain data appear to be linearly related to the temperature variations.
3.2. Temperature and system gain

Figure 3.2: The top panel shows a waterfall plot of the calculated gain phase from the gain solution data for a three day period for channel 12, and the bottom panel shows the calculated gain amplitude for the same channel over the same time period. RFI-contaminated frequencies and solar transits have been removed. Additional waterfall plots are included in Appendix A. A comparison of the plots of different channels reveals that the LNAs behave uniformly across frequencies over the three-day period.
3.2. Temperature and system gain

3.2.1 A first order model

The top panel of Figure 3.3 shows a representative plot of the time stream of gain phase data and the linear fit of temperature to gain phase, for a single LNA at a single frequency. Prior to calculating gain phase and gain amplitude, complex gain data are first normalized by the noise source reference channel, in order to cancel out any drift in the gain. Gain amplitude data have also been subsequently normalized by the average gain amplitude. Fits were calculated for all channels at all frequencies and are tabulated as averages in Table 3.1 for both phase and amplitude. The average fit of temperature to gain phase was \( \sim 0.9 \text{ deg K}^{-1} \) and the average fit of temperature to gain amplitude was \( \sim -0.5 \% \text{ K}^{-1} \). The direction of the amplitude fit value suggests reasonable agreement with lab tests of LNAs in a previous study, which found that LNA gain decreases linearly with increasing temperature [26].

<table>
<thead>
<tr>
<th></th>
<th>Phase [deg K(^{-1})]</th>
<th>Amplitude [% K(^{-1})]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>fit error</td>
<td>fit error</td>
</tr>
<tr>
<td>average</td>
<td>0.892 0.021</td>
<td>-0.479 0.458</td>
</tr>
<tr>
<td>median</td>
<td>0.863 0.016</td>
<td>-0.357 0.012</td>
</tr>
</tbody>
</table>

Table 3.1: Average and median thermal fit values (slopes, \( \delta g/\delta T \)) and errors, for gain phase and gain amplitude for 8 channels with thermometry data. Uncertainties are obtained through summing in quadrature all the fit errors at each frequency.

The bottom panel of Figure 3.3 shows the phase error – essentially the residuals of the gain phase and the fit to gain phase – plotted in white over the time stream. The plot shows a phase error of \( \sim \pm 2^\circ \), indicative of how close the temperature fit matches the gain phase, with a scatter of 0 being the ideal residual if the noise source derived gain solution is taken to be ideal.

Figure 3.4 shows the root-mean-square (RMS) of the phase error for 8 LNAs, which indicates the scatter in the ability of the thermal fit to match the gain phase. A feasible model, which can be used to minimize the use of the noise source, requires \( \lesssim 10^9 \) scatter in phase. Since extra phase will translate to extra path length, a phase error on the order of \( 10^9 \) translates into \( \sim 1 \text{ cm} \) RMS fluctuation in surface area of the collecting cylinder. Figure 3.4 indicates a scatter in phase of \( \sim 1^\circ \) in the lower end of the spectrum (higher frequencies have more anomalies) and therefore highlights the significant result that the thermal model is clearly within the \( \sim 10^9 \) scatter in phase that is acceptable as a viable model to determine system gain. Thus, this thermal model can be used to minimize the use of the noise source, provided the gain solution data used to develop the thermal model are reliable.

One can estimate the effectiveness of the thermal fit by dividing the scatter in the ability of the thermal model to match the noise source derived gain by the scatter in the noise source derived gain solution on its own. Figure 3.5 is
3.2. Temperature and system gain

Figure 3.3: The top panel shows a time stream of temperature fitted to gain phase over three days. The bottom panel shows the residuals of the gain phase and fit data. The $\sim \pm 2^\circ$ phase error indicates the variation between the thermal model and the gain phase.
3.2. Temperature and system gain

Figure 3.4: The RMS of residuals for 8 LNAs shows phase scatter of $\sim 1^\circ$ in the lower frequency range, which is well within the acceptable range for a viable model to determine system gain.
3.2. Temperature and system gain

Figure 3.5: The plots show a comparison of the RMS of the residuals and the RMS of the gain phase for 8 LNAs. A reduction in RMS of a factor of \(\sim 3\) can be obtained by using a thermal model.
3.2. Temperature and system gain

a comparison between the RMS of the residuals as obtained from Figure 3.4 ($\sim 1^\circ$) and the RMS of gain phase ($\sim 3^\circ$) for all 8 channels with thermometry data. The panels of Figure 3.5 confirm that the scatter in phase is reduced by a factor of $\sim 3$ when implementing a thermal model. This result provides further incentive to explore the use of a thermal model.

In addition, a representative plot of the time stream of the thermal fit to the gain amplitude is shown in Figure 3.6, with the bottom panel showing the fractional amplitude variation superimposed in white, as determined from the residuals of the thermal fit and gain amplitude. The plot indicates a $\sim \pm 2\%$ variation about the mean.

The scatter in the ability of the thermal fit to match the gain amplitude is indicated in Figure 3.7, showing the RMS of the residuals of fits to gain amplitude for 8 LNAs across the bandwidth. For most channels, the degree of scatter in gain amplitude is on the order of $\sim 1\%$ at lower frequencies. These plots studying the thermal fits to gain amplitude serve simply as a guide, since a defined standard to test if the thermal fits to gain amplitude fall within acceptable limits does not exist.
3.2. Temperature and system gain

Figure 3.6: The top panel shows a time stream of temperature fitted to gain amplitude over three days. The bottom panel shows the difference (fractional amplitude error) between the gain and fit data. The fractional amplitude error is an indication of how much the thermal model varies from the gain solution, which is estimated to be a $\sim \pm 2\%$ variation.
Figure 3.7: The RMS of residuals for 8 LNAs show deviation in fractional gain amplitude (intensity) as a percentage of the total gain amplitude. Gain amplitude values are divided by the average gain amplitude. The scatter in the ability of the thermal model to match the gain amplitude is estimated at $\sim 1\%$ in the lower frequency range.
3.2. Temperature and system gain

3.2.2 Time lag model

A time difference between gain response and temperature is expected, due to the effects of changing temperature on materials. The LNA circuits are encased in metal boxes and the thermistors are fitted to the exterior of the boxes. The LNA boxes have a larger heat capacity than the thermistors, and will take more time to heat up and cool down, which could result in delayed gain change in the noise source derived gain data as compared to the temperature readings. Given the difference in heat capacity of the metal boxes versus that of the thermistors, it seems reasonable to expect a time lag in the noise source gain data. It was therefore expected that the time stream plots would indicate that the temperature fit leads the gain.

Figure 3.3 shows a representative time stream of the temperature fit to gain phase, for a single LNA at a single frequency. The plot indicates a time lag due to temperature, which could ultimately distort CHIME’s sky maps if not corrected. However, the time lag as shown in Figure 3.3 is peculiar – note that the gain phase actually leads the temperature fit to gain phase at most points in the three day period. In fact, for most channels and at most frequencies, the temperature fit rarely leads the noise source gain solution. Additional time stream plots for a representative channel at a sampling of frequencies are included in Figures A.5 and A.6, showing with few exceptions that temperature fits generally lag both gain phase and gain amplitude.

The time lag effect is also present in the scatter plot of Figure 3.8, evidenced in the loop-like structure separating rising and falling points. The top panel of Figure 3.8 shows a scatter plot for one day only, and the bottom panel shows data for each of the full three days (24 hour periods from noon to noon). The fit for all times in both panels, however, includes all the data from the three 24-hour periods, as well as data from hours prior to and following these three full days. The time component is captured in the gradation of markers from noon (coloured circle) to midnight (darkened markers) and then to noon again (coloured triangle), indicating the direction of phase angle change with changing temperature.

Interestingly, the separation between rising and falling points in the scatter plots, which is induced by the time lag effect, serendipitously reveals more information about the signal chain. As shown in the top panel of Figure 3.8, oscillations can be seen in the scatter plots for all channels and most frequencies. These oscillations are speculated to be the evidence of temperature fluctuations in the FLAs, which are located in the RF room and are subject to an indoor air-conditioned mechanical system. Incidentally, if the effect of the mechanical system on the FLAs can be linked to the oscillating structure seen in the scatter plots, it would suggest the importance of considering other aspects of the signal chain when developing a model to determine system gain. Additional scatter plots are included in Figures 3.17, 3.18, A.13, and A.14, showing the temperature fits to gain phase and gain amplitude at several frequencies for a representative channel on a single day.
3.2. Temperature and system gain

Figure 3.8: Thermal fits to gain phase for a single LNA at a single frequency. The top panel shows data for a single day (24 hours, noon to noon) and the bottom panel for 3 days. The black line indicates a phase fit for all times. Note that data before noon on the 16th and data after noon on the 20th are not shown for clarity, but are included in the fit for all times. To indicate direction of phase angle change, coloured circle markers indicate noon starting the 24 hour period, gradating to darker markers, and then gradating back to coloured triangle markers at noon of the following day.
3.2. Temperature and system gain

In order to characterize the time lag, time stream data were investigated across all channels and frequencies. An initial estimate of a time lag of about an hour was made by inspection of the time stream plots. To determine a typical time lag, an algorithm was developed as follows. An array of lag times, 5 minutes apart, was created and the temperature data were shifted in 5 minute increments. At each shift of the temperature data, a fit was made to the noise source derived gain data and a corresponding chi-square statistic was calculated. This method was performed for an array of time shifts from 0 to 180 minutes for all channels at a sampling of frequencies.

The top panels of Figures 3.9 and 3.10 are time streams of temperature data shifted by 91 minutes prior to calculating fits to gain phase and gain amplitude, respectively. The bottom panels of Figure 3.9 and 3.10 show the chi-square statistics for a single channel at 9 frequencies for all time lags between 0 and 180 minutes for gain phase and gain amplitude, respectively. The curves gently slope toward a time lag where the chi-square value is a minimum, representing the significant result that gain can be predicted to first order based on temperatures. For each channel, the shift that resulted in the minimum chi-square statistic for each of a sample of 9 frequencies, is an indicator of where gain and temperature best match and is thus the best choice for a typical time lag. The chi-square plots of Figures 3.9 and 3.10 indicate that the most typical time lag estimate is ~45 min to 1 hour for both gain and amplitude. Additional plots showing the chi-square statistics of the time lags for the other channels with thermometry data are shown in Figures A.7 and A.8 for gain phase and in Figures A.9 and A.10 for gain amplitude.
3.2. Temperature and system gain

Figure 3.9: The top panel is a time stream of a temperature fit to gain phase for channel 12 at a single frequency, showing temperature data fitted after a shift of 91 minutes. The bottom panel shows chi-square values at each time lag interval for several frequencies in channel 12, obtained after fitting temperature to gain phase at each time shift of 5 minutes, up to 3 hours. The best estimate of time lag across the sampling of frequencies is $\sim 45$ min to 1 hour.
3.2. Temperature and system gain

Figure 3.10: The top panel is a time stream of a temperature fit to gain amplitude for channel 12 at a single frequency, showing temperature data fitted after a shift of 91 minutes. The bottom panel shows chi-square values at each time lag interval for several frequencies in channel 12, obtained after fitting temperature to gain amplitude at each time shift of 5 minutes, up to 3 hours. From the chi-square plot, the best estimate of time lag is $\sim 45$ min to 1 hour.
3.2. Temperature and system gain

3.2.3 Thermal susceptibility

An investigation of thermal susceptibility establishes an understanding of how susceptible system gain is to changes in temperature across the bandwidth. Figures 3.11 and 3.14 show susceptibility plots for gain phase and gain amplitude for the eight LNAs of this study. Since the LNAs are intrinsically the same and exposed to the same temperatures, they are expected to behave in a similar manner. They should change together, such that temperature fits to gain phase and to gain amplitude for all LNAs should fall within the same family of curves. The LNAs coupled to the North-South polarization of the feed antennas exhibit the same family of curves. Similarly, the LNAs coupled to the East-West polarization of the feed antennas share a similar curve structure. However, the family of curves across polarizations are distinctly different. This ostensible discrepancy in frequency structure across polarizations is also apparent when performing temperature fits by proxy to the gain data from the LNAs which had no thermometry data, but were coupled to a polarization in the same way as those that did have thermometry data. The susceptibility plots for all 15 channels are shown in Figures A.11 and A.12 for both gain phase and gain amplitude.

The apparent polarization dependence in the frequency structure is a surprising result, since polarization properties belong to the feed antennas and not to the LNAs. The LNAs are merely coupled to the antennas. There is no physical reason why the gain data should be exhibiting polarization properties. Therefore, a comparison was made of the regions of the susceptibility curves where the two curve families differed and where they were similar. Frequency ranges at the “misbehaving” frequencies were chosen – frequencies where the curve families differed most significantly from each other – and were compared to “behaving” frequencies, where the curve families were similar. Misbehaving frequencies between 600 - 650 MHz in the time streams of the noise source derived gain were studied and the frequencies where anomalies were present were identified. Spikes in the time stream data served as an alert to the need to phase wrap the data such that complementary angles were determined from angles greater than $2\pi$. After applying a phase wrapping algorithm to ensure the angles calculated ranged between $-180^\circ$ and $+180^\circ$ across the bandwidth, the susceptibility structure remained discrepant for families of curves across the two polarizations.

In analyzing the time streams of data at both misbehaving and behaving frequencies, it was noted that generally the time streams for different polarizations of the same antenna exhibited similar fits and similar structure, as shown in Figures 3.12 and 3.13, regardless of the differing polarization structure indicated in the susceptibility plots. However, at some frequencies, such as the misbehaving frequency shown in Figure 3.12, the time stream of the LNA coupled to the North-South polarization of the antenna also indicates an instability in the polarization from day 2 to day 3, as seen in the top panel. The time stream of the LNA coupled to the East-West polarization of the
Figure 3.11: Thermal susceptibility of gain phase from 8 channels, 4 channels per panel. The top panel shows the LNAs coupled to the North-South antenna polarization and the bottom panel the LNAs coupled to the East-West antenna polarization. Note that the families of curves are significantly different across polarizations at the higher frequencies.
3.2. Temperature and system gain

Figure 3.12: The time streams are part of the investigation of thermal fits to gain phase at misbehaving frequencies. The top panel is a time stream for an LNA coupled to the North-South polarization of an antenna, and the bottom panel a time stream for an LNA coupled to the East-West polarization of the same antenna. The plots show similar fits and structure across polarizations, which is representative of most channels for most frequencies; however, the top panel also reveals an instability in the North-South polarization, seen from day 2 to day 3.
3.2. Temperature and system gain

Figure 3.13: The time streams are part of the investigation of thermal fits to gain phase at behaving frequencies. The top panel is a time stream for an LNA coupled to the North-South polarization of an antenna, and the bottom panel a time stream for an LNA coupled to the East-West polarization of the same antenna. The plots show similar fits and structure across polarizations, which is representative of most channels for most frequencies.
3.2. Temperature and system gain

antenna shown in the bottom panel shows less instability. Comparing residuals of the gain phase and temperature fit for each of these LNAs coupled to different polarizations at this misbehaving frequency would reveal very different slopes. At the behaving frequency shown in Figure 3.13, there are fewer differences in the fit across polarization. Thus, the residuals of the gain phase and temperature fit for each of these LNAs coupled to different polarizations would reveal less variation in slope across polarization than at the misbehaving frequency. This observation further suggests a non-temperature related effect that is generating different structure and different fits across polarizations at the same frequencies.

In addition, susceptibility plots of thermal fits to gain amplitude are shown in Figure 3.14. Note that the difference in frequency structure across polarizations for the fits to gain amplitude is even more pronounced than for the fits to gain phase. Misbehaving frequencies across polarization that were investigated in the time stream data for gain amplitude include a range between 600 and 650 MHz and between 700 and 750 MHz. The top panel of Figure 3.15 shows a time stream of the thermal fit to gain amplitude at a misbehaving frequency for an LNA connected to the North-South polarization of an antenna and the bottom panel shows a plot for an LNA connected to the East-West polarization of the same antenna. Figure 3.16 shows timestream plots at a behaving frequency. In both misbehaving and behaving frequencies, the structure and fits to gain amplitude are similar and do not generally exhibit any drift of polarization as shown in the time stream of the thermal fit to gain phase in the top panel of Figure 3.12. For both phase and amplitude, the structure of the gain and the thermal fit values are similar across polarization for both behaving and misbehaving frequencies for most LNAs. Therefore, the apparent polarization dependence observed across the bandwidth is likely a real, physical effect captured in the gain solution data.
3.2. Temperature and system gain

Figure 3.14: Thermal susceptibility of gain amplitude from 8 channels, 4 channels per panel. The top panel shows the LNAs coupled to the North-South antenna polarization and the bottom panel the LNAs coupled to the East-West antenna polarization. Note that the families of curves are significantly different across polarizations at the higher frequencies.
3.2. Temperature and system gain

Figure 3.15: The time streams are part of an investigation of thermal fits to gain amplitude at misbehaving frequencies. The top panel is a time stream for an LNA coupled to the North-South polarization of an antenna, and the bottom panel a time stream for an LNA coupled to the East-West polarization of the same antenna. The plots show similar fits and structure across polarizations, which is representative of most channels for most frequencies.
3.2. Temperature and system gain

Figure 3.16: The time streams are part of an investigation of thermal fits to gain amplitude at behaving frequencies. The top panel is a time stream for an LNA coupled to the North-South polarization of an antenna, and the bottom panel a time stream for an LNA coupled to the East-West polarization of the same antenna. The plots show similar fits and structure across polarizations, which is representative of most channels for most frequencies.
3.2. Temperature and system gain

Frequency dependent gain

Examining the thermal fits across the bandwidth of the susceptibility plots of Figures 3.11 and 3.14, it is evident that the fits at low frequency are better behaved than those at higher frequencies. For both gain phase and gain amplitude there appears to be more noise at higher frequencies. Excess scatter at higher frequencies is also seen in Figures 3.8, 3.17, and 3.18. These plots are representative of the behaviour seen in most channels for temperature fits to gain phase. On the whole, the scatter plots are “cleaner” at lower frequencies with excess scatter appearing at higher frequencies. This discrepancy between lower and higher frequencies may be indicative of a lack of gain required to reliably see the noise source in the data at higher frequencies. A similar trend of increasing noise at higher frequencies is evident in the scatter plots for gain amplitude in Figures A.13 and A.14.

Figures 3.19 and 3.20 show slopes of the fits to gain phase for the 8 channels at various frequencies over the 3-day period. Confirming the trend of increased noise at higher frequencies as evidenced in the gain versus temperature plots, the fit slopes of Figures 3.19 and 3.20 become increasingly scattered at higher frequencies, while remaining behaved at lower frequencies. Similar behaviour is evidenced in the fits to gain amplitude shown in Figures A.15 and A.16. The trend of noisier data at higher frequencies may be contributing to the discrepancy in shape between families of curves in the thermal susceptibility plots – an effect which could be superimposed over the apparent polarization dependence in the frequency structure, as discussed in Section 3.2.3.
3.2. Temperature and system gain

Figure 3.17: The panels show data for a single LNA on a single day, with a fit for all times indicated by the black line. Frequency increases from the low to mid-range from (a) to (d). It appears that on the whole, the data are noisier at higher frequencies than at lower frequencies, making it more difficult to see the loop-like structure as discussed in Section 3.2.2. Compare to Figure 3.18, to see behaviour across the bandwidth.
3.2. Temperature and system gain

Figure 3.18: The panels show data for a single LNA on a single day, with a fit for all times indicated by the black line. Frequency increases from the mid- to high range from (a) to (d). It appears that on the whole, the data are noisier at higher frequencies than at lower frequencies, making it more difficult to see the loop-like structure as discussed in Section 3.2.2. Compare to Figure 3.17, to see behaviour across the bandwidth. Noisier data at high frequencies may have some influence on why the fits at higher frequencies are worse behaved than those at lower frequencies, as shown in the thermal susceptibility plots.
3.2. Temperature and system gain

Figure 3.19: The panels show temperature fits for gain phase data from 3 days for all channels at a sampling of frequencies, increasing from low to mid-range from (a) to (d). For most low frequencies the fits are well-behaved, i.e. the LNAs have similar fit slopes. However, for most high frequencies, there is greater variation in fit slopes. This may be evidence of noisier data at the higher frequencies, which may contribute to the distinction between behaving fits at low frequencies and misbehaving fits at higher frequencies seen in the thermal susceptibility plots. See Figure 3.20 to see the behaviour of fits across the bandwidth.
3.2. Temperature and system gain

Figure 3.20: The panels show temperature fits for gain phase data from 3 days for all channels at a sampling of frequencies, increasing from mid- to high range from (a) to (d). For most low frequencies the fits are well-behaved, i.e. the LNAs have similar fit slopes. However, for most high frequencies, there is greater variation in fit slopes. This may be evidence of noisier data at the higher frequencies, which may contribute to the distinction between behaving fits at low frequencies and misbehaving fits at higher frequencies, as seen in the thermal susceptibility plots. When comparing with Figure 3.19 to see the behaviour of fits across the bandwidth, it is clear that higher frequencies are noisier.
3.3 Thermal model applied to simulated visibility data

The goal of a system gain solution is ultimately to obtain real-time corrections of image quality. By generating two identical maps of simulated sky data and applying the noise source derived gain solution from the Pass0E data used in this study to one map and the thermal model gain solution to the other, one can facilitate a comparison of the two methods of determining system gain. If the noise source derived gain solution is taken to be the ideal model, comparing the two maps can provide an understanding of the effectiveness of the thermal model to match the noise source derived model.

Figure 3.21 shows sky maps to which the noise source derived gain solution and thermal model have been applied. Each map is a “dirty” image, meaning that the telescope beam is convolved with the sky data and no deconvolution has yet been performed, thus leaving artifacts in the map. Note the position of the Sun as the bright line on the left-hand side of the plots. Figure 3.21 is determined from the xy polarization; additional maps are included in Figures A.17 and Figure A.18, generated from the x only polarization and from the y only polarization, respectively.

A difference map of the two maps in Figure 3.21 is shown in Figure 3.22. Figure 3.22 gives an indication of how well the thermal model performs in determining a system gain solution, provided the noise source derived gain solution is reliable. It is clear from the maps that the thermal model is a model that can well predict system gain.
3.3. Thermal model applied to simulated visibility data

Figure 3.21: The top image is a map in molleweide projection of the noise source derived gains (from the Pass0E dataset) at 706.1 MHz using xy polarization, as applied to simulated sky data. The bottom image is a map of the thermal model using xy polarization, as applied to the same simulated sky data. Note the position of the Sun on the lefthand side of the maps. If the noise source gains have been shown to be reliable, a comparison of the two maps qualitatively shows the effectiveness of the thermal model in predicting system gain.
3.3. Thermal model applied to simulated visibility data

Figure 3.22: The image is a map in molleweide projection of the difference between the maps in Figure 3.21. If the noise source gains have been shown to be reliable, the smoothness of the map illustrates the minimal difference between the two models, and thus qualitatively demonstrates the effectiveness of the thermal model in predicting system gain.
Chapter 4

Discussion and future study

The results of this study are informative for understanding the advantages and limitations of a thermal model. The determination of a peculiar time lag, presence of an unexpected polarization dependence, and discovery of excess noise at higher frequencies can initiate thoughtful discussion to further develop the first order thermal model.

To understand the peculiar time lag in the temperature fits, it is worth noting the physical configuration of the LNAs and thermistors. The LNAs are housed in the bulkhead along each cylinder’s focal line and the thermistors are fitted to the exterior of the metal boxes encasing the LNAs. Although both the metal boxes and thermistors are exposed to the outdoor temperatures while inside the bulkhead, they are not exposed to the Sun or direct wind. By contrast, the feed antennas are exposed to sunlight and wind; as a result, it is likely that through their electrical coupling to the LNAs, there is a thermal path created between the antennas and LNAs. The relatively small thermal mass of the LNAs allows them to respond quickly to changes in temperature. Given that the thermistors are fitted to the exterior of the LNA boxes which are located inside the bulkhead, their temperature readings may not reflect the same temperature that the LNAs are experiencing as a result of a thermal path induced by the antennas’ direct exposure to the elements. This physical configuration when taking measurements may be responsible for the peculiar feature in the time stream plots showing temperature fits lagging gain. Future studies of the thermal model would benefit from placing the thermistors inside the LNA boxes or even fitting them onto the antennas themselves, to help obtain an improved fit of temperature to the noise source derived gain data.

Moreover, the sloping chi-square curves used to determine the typical time lag confirm that system gain can be predicted to first order based on temperatures. However, since the curves are gently sloping, it is likely that there are subtleties in the thermal model that need to be identified and characterized, in order to describe in detail how gain changes with temperature. Some aspects to be further explored include the dependence of gain on seasonal temperature and the dependence of gain on rising temperature versus falling temperature, which may be captured by using the derivative of temperature. Further study could describe these dependencies in detail and make suggestions based on the trends observed in the data. Given that the data presented here encompass the effects of the entire signal chain, future studies would also do well to address the effects of other components along the signal chain, such as the FLAs. The oscillations superimposed on the data (as seen in the scatter plot of Figure 3.8) are likely due to the effect of the air-conditioner on the FLAs in the RF room.
Chapter 4. Discussion and future study

An understanding of the contribution of the FLAs to the gain change along the signal chain would help to characterize and refine the thermal model. These corrections to the first order model may not only help explain why the time lag is estimated at $\sim 1$ hour (which at first glance seems a considerable time difference), but may also help elucidate why the temperature fit lags the gain. Furthermore, since the structure of the gain and the thermal fit values are similar across polarization for both misbehaving and behaving frequencies, the apparent polarization dependence observed across the bandwidth in the susceptibility plots is likely a real, physical effect, perhaps influenced by the noise antenna position with respect to the feed antennas. At the time the Pass0E data were collected, it is likely that, because of its position at the end of the cylinder, the noise source antenna was broadcasting at an angle relative to the feed antennas (instead of directly into them), and thus the oscillations of the noise source signal were likely not in the same direction as the polarization that the feed antennas receive. Thus, the polarization dependent structure in the gain solution data is likely a resulting effect of the position of the noise source antenna, which alerts us to consider the importance of the location of the noise antenna relative to the feeds. The artificial polarization is a good caveat to keep in mind for future studies that use the noise source to determine a reliable gain solution against which a thermal model is developed. This result is significant because without knowing the effect of the position of the noise source, and without properly calibrating the instrument in this respect, there may be structure in the data that would mistakenly be interpreted as real. Such artificial structure would result in incorrect pointing on the sky in the case of gain phase and incorrect intensities in the case of gain amplitude.

In addition, the noisier data at high frequencies may also be contributing to the discrepancy in shape for the families of susceptibility curves across polarizations. A previous study [26] reports that the gain in the LNAs is not entirely due to ambient temperatures and shows that although LNA gain decreases with increasing temperature, this behaviour varies per frequency. The gain at lower frequencies in particular, is less influenced by changing temperature, which may explain why the thermal fits at lower frequencies are better behaved than those at higher frequencies. Increased noise at higher frequencies may be partly due to the LNAs themselves, whose gain varies with operating frequency. Characterizing and correcting for the excess noise at high frequencies will be important in future work to refine the thermal model. Finally, a comparison of the sky maps to which the thermal model and gain solution have been applied reveals the effectiveness of the thermal model at predicting system gain and emphasizes that the thermal model is worthy of further study. A future test would be to apply the thermal model to different gain solutions that have been verified. Each gain solution and corresponding thermal model can then be applied to sky data and the maps subsequently compared. Many such comparisons of the noise source derived gain and the fitted thermal model each separately applied to the same sky data would comprehensively test how well the thermal model can predict gain.
Chapter 5

Conclusions

The investigations outlined in this paper have shown that the thermal model is an acceptable first order model which requires further, detailed analysis. Its advantages over the noise source injection method lie chiefly in that it poses minimal impact to the DRAO radio quiet zone and reduces system complexity, while facilitating the mapping of the sensitive BAO signal.

To begin, the study determined that the scatter in phase of $\varphi$ for the thermal model falls well within the requirements of acceptable phase error and therefore confirms that the thermal model is a viable model that can be used to minimize the use of the noise source in the DRAO radio quiet zone. The study also confirmed the overall average of thermal fits to gain phase to be $\sim 0.9$ deg $K^{-1}$. In keeping with the findings of a previous study [26], the overall average of thermal fits to gain amplitude was determined to be $-0.5\ %\ K^{-1}$.

Moreover, the unexpected result of temperature lagging gain for most channels and at most frequencies, is likely a consequence of different thermal environments: that of the feed antennas (exposed to weather) versus that of the thermistors and metal boxes housing the LNAs (inside the bulkhead). Acknowledging the physical effects resulting from the system configuration can inform future thermal studies. The time lag itself was studied using chi-square statistics and has shown an average time lag of about an hour for most channels and frequencies, for both gain phase and gain amplitude. Although the model is acceptable to first order, the gentle sloping feature of the chi-square curves confirms that further detailing of the temperature trends is required in order to more accurately characterize the thermal model.

Similar to the unexpected time lag result, investigations of thermal susceptibility also yielded an unexpected polarization dependent frequency structure evident in the thermal susceptibility plots. This artificial structure provides insight into the behaviour of LNAs when determining a noise source derived system gain. The study suggests that consideration must be made in the positioning of the noise source antenna so as to avoid such artificial structure. Thus, in addition to exploring a thermal model to determine system gain, this study has also provided insight into the noise source injection method of determining system gain. Without correcting for noise source positioning, an incorrect noise source derived gain calibration would be obtained, which in turn would make it impossible to develop a reliable thermal model and would destroy the ability to detect the faint BAO signal.

In addition to the time lag and thermal susceptibility findings, investigations into the discrepancy between thermal fits at low- versus high frequencies were made, which suggest further study is required to characterize and account...
for excess noise at high frequencies.

Finally, assuming a reliable noise source derived gain solution for this study, comparing a sky map with the applied gain solution to the same sky map with the applied thermal model has served as a test to qualitatively confirm the relative effectiveness of the thermal model in predicting system gain. Such tests can be performed on future datasets to further refine the thermal model.

These investigations in development of a thermal model are critical steps in determining a system gain calibration and they inform the larger, overall calibration plan. The system gain distorts the map with phase delays; without removing the effects of phase delay the maps would be laden with phase errors and the sensitive BAO signal would be entirely hidden. Thus, in addition to other calibrations, accurately accounting for system gain is a critical calibration step that helps lead the way for the subsequent beam calibration, which will eventually remove the bright Galactic foreground in preparation for detecting the faint BAO signal.

The reliability of the science results obtained from the CHIME telescope is ultimately determined by the successful calibration of the instrument. Since CHIME will provide the first-ever HI map of BAO in the redshift range where the onset of Dark Energy’s domination can be studied, successfully implementing a calibration plan that begins with these critical steps will ultimately contribute to detecting the faint cosmological signal. And it is precisely the analysis of this faint signal that has the potential to enhance our understanding of the expansion of the Universe and Dark Energy.
Bibliography


Bibliography


Appendix A

Additional Plots

The following plots supplement the plots in the body of the paper, showing data from additional channels and frequencies and further confirming the results.
Figure A.1: The panels show waterfall plots of calculated gain phase for 4 channels which have thermometry data. RFI frequencies and solar transits have been removed. Comparison with waterfall plots from the other channels reveals that the LNAs behave uniformly.
Figure A.2: The panels show waterfall plots of calculated gain phase for 4 channels which have thermometry data. RFI frequencies and solar transits have been removed. Comparison with waterfall plots from the other channels reveals that the LNAs behave uniformly.
Appendix A. Additional Plots

Figure A.3: The panels show waterfall plots of calculated gain phase for 3 channels which do not have thermometry data. Note that the noise source channel is excluded. RFI frequencies and solar transits have been removed. Comparison with waterfall plots from the other channels reveals that the LNAs behave uniformly.
Figure A.4: The panels show waterfall plots of calculated gain phase for 4 channels which do not have thermometry data. RFI frequencies and solar transits have been removed. Comparison with waterfall plots from the other channels reveals that the LNAs behave uniformly.
Figure A.5: Each panel shows a time stream of temperature fitted to gain phase over three days for a single channel. The panels increase in frequency from (a) to (d). Rarely does temperature lead the gain phase in the majority of channels at all frequencies. However, for Channel 04 shown here, temperature does lead mildly in some places in the 664.8 MHz and 773.8 MHz plots.
Appendix A. Additional Plots

Figure A.6: Each panel shows a time stream of temperature fitted to gain amplitude over three days for a single channel. The panels increase in frequency from (a) to (d). Rarely does temperature lead the gain amplitude in the majority of channels at all frequencies.
Figure A.7: Chi-square values for temperature fits to gain phase at 5 minute time lags at several frequencies for channels 04-07. The typical time lag is where the chi-square is minimized and is estimated at \( \sim 45 \) min to 1 hour for most channels at most frequencies. The gentle slope is indicative of a first order model.
Figure A.8: Chi-square values for temperature fits to gain phase at 5 minute time lags at several frequencies for channels 12-15. The typical time lag is where the chi-square is minimized and is estimated at ~ 45 min to 1 hour for most channels at most frequencies. The gentle slope is indicative of a first order model.
Figure A.9: Chi-square values for temperature fits to gain amplitude at 5 minute time lags at several frequencies for channels 04-07. The typical time lag is where the chi-square is minimized and is estimated at $\sim 45$ min to 1 hour for most channels at most frequencies. The gentle slope is indicative of a first order model.
Figure A.10: Chi-square values for temperature fits to gain amplitude at 5 minute time lags at several frequencies for channels 12-15. The typical time lag is where the chi-square is minimized and is estimated at ~ 45 min to 1 hour for most channels at most frequencies. The gentle slope is indicative of a first order model.
Figure A.11: Thermal susceptibility of gain phase from 15 channels. The top panel shows 8 channels for LNAs coupled to the North-South antenna polarization and the bottom panel shows 7 channels for LNAs coupled to the East-West antenna polarization (note that the noise source channel is excluded). Temperature data from LNAs with thermometry were fitted to gain data from LNAs without thermometry (ch01-03 and ch08-11). Note that the families of curves across polarization differ most significantly at higher frequencies.
Figure A.12: Thermal susceptibility of gain amplitude from 15 channels. The top panel shows 8 channels for LNAs coupled to the North-South antenna polarization and the bottom panel shows 7 channels for LNAs coupled to the East-West antenna polarization (note that the noise source channel is excluded). Temperature data from LNAs with thermometry were fitted to gain data from LNAs without thermometry (ch01-03 and ch08-11). Note that the families of curves across polarization differ most significantly at higher frequencies.
Figure A.13: The panels show thermal fits to gain amplitude for a single LNA on a single day, with a fit for all times indicated by the black line. Frequency increases from low to mid-range from (a) to (d). It appears that on the whole, the data are noisier at higher frequencies than at lower frequencies. See Figure A.14 for a comparison across the bandwidth. Noisier data at higher frequencies may have some influence on why the fits at higher frequencies are worse behaved than those at lower frequencies, as seen in the thermal susceptibility plots.
Appendix A. Additional Plots

Figure A.14: The panels show thermal fits to gain amplitude for a single LNA on a single day, with a fit for all times indicated by the black line. Frequency increases from the mid- to high range from (a) to (d). It appears that on the whole, the data are noisier at higher frequencies than at lower frequencies. See Figure A.13 for a comparison across the bandwidth. Noisier data at higher frequencies may have some influence on why the fits at higher frequencies are worse behaved than those at lower frequencies, as seen in the thermal susceptibility plots.
Appendix A. Additional Plots

Figure A.15: The panels show temperature fits to gain amplitude data from 3 days for all channels at a sampling of frequencies, from low to mid-range. For most low frequencies the fits are well-behaved, as evidenced by the similar fit slopes across LNAs. However, for most high frequencies, there is greater variation in fit slopes. This may be evidence of noisier data at the higher frequencies, which may contribute to the distinction between behaving fits at low frequencies and misbehaving fits at higher frequencies, as seen in the thermal susceptibility plots. See Figure A.16 for a comparison of fits across the bandwidth.
Figure A.16: The panels show temperature fits to gain amplitude data from 3 days for all channels at a sampling of frequencies, from mid- to high range. For most low frequencies the fits are well-behaved, as evidenced by the similar fit slopes across LNAs. However, for most high frequencies, there is greater variation in fit slopes. This may be evidence of noisier data at the higher frequencies, which may contribute to the distinction between behaving fits at low frequencies and misbehaving fits at higher frequencies, as seen in the thermal susceptibility plots. By comparing with Figure A.15, it is evident that the slopes at higher frequencies are more varied, which is indicative of increased noise at higher frequencies.
Figure A.17: The top image is a map in molleweide projection of the noise source derived gains (from the Pass0E dataset) at 706.1 MHz using x polarization, as applied to simulated sky data. The bottom image is a map of the thermal model using x polarization, as applied to the same simulated sky data. Note the position of the Sun on the lefthand side of the maps. If the noise source gains have been shown to be reliable, a comparison of the two maps qualitatively shows the effectiveness of the thermal model in predicting system gain.
Figure A.18: The top image is a map in molleweide projection of the noise source derived gains (from the Pass0E dataset) at 706.1 MHz using y polarization, as applied to simulated sky data. The bottom image is a map of the thermal model using y polarization, as applied to the same simulated sky data. Note the position of the Sun on the lefthand side of the maps. If the noise source gains have been shown to be reliable, a comparison of the two maps qualitatively shows the effectiveness of the thermal model in predicting system gain.