Title
Sensitive Detection of CMB B-Mode Polarization : : Instrumentation and Systematics

Permalink
https://escholarship.org/uc/item/5fp5w0z0

Author
Moyerman, Stephanie

Publication Date
2013-09-11

Peer reviewed|Thesis/dissertation
Sensitive Detection of CMB B-Mode Polarization: Instrumentation and Systematics

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Physics with Specialization Computational Science by

Stephanie Moyerman

Committee in charge:

Professor Brian Keating, Chair
Professor Eric Fullerton
Professor Kim Griest
Professor Hans Paar
Professor Gabriel Rebeiz

2013
The dissertation of Stephanie Moyerman is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

______________________________________________________________

______________________________________________________________

______________________________________________________________

______________________________________________________________

______________________________________________________________ Chair

University of California, San Diego

2013
DEDICATION

For Kelsey
For telling me to, “Sack up and Finish.”
I love you.
EPIGRAPH

It’s gotta be at least 5σ for facebook.

—Bryan Steinbach
TABLE OF CONTENTS

Signature Page .................................................. iii
Dedication ......................................................... iv
Epigraph ............................................................ v
Table of Contents ................................................ vi
List of Figures .................................................... ix
List of Tables ...................................................... xv
Acknowledgements ................................................. xvi
Vita and Publications ........................................... xix
Abstract of the Dissertation ..................................... xxi

Chapter 1  Background and Motivation ............................ 1
1.1 Modern Cosmology ........................................... 1
  1.1.1 The Big Bang Model ...................................... 6
  1.1.2 Problems with the Standard Big Bang ................. 7
  1.1.3 Inflation ................................................. 8
1.2 The Cosmic Microwave Background ......................... 9
  1.2.1 CMB Temperature Anisotropy ......................... 10
  1.2.2 Temperature Power Spectrum ........................ 12
  1.2.3 Acoustic Oscillations ................................. 12
  1.2.4 CMB Polarization Anisotropy ......................... 15
  1.2.5 Polarization Power Spectrum ......................... 17
  1.2.6 Causes of Polarization ................................. 19
1.3 CMB Measurements .......................................... 27
  1.3.1 Difficulty of Measurement ............................. 27
  1.3.2 State of the Field ..................................... 28
  1.3.3 Future CMB Studies ................................... 30

Chapter 2  Faraday Rotation Modulators .......................... 31
2.1 Motivation .................................................... 31
2.2 FRM Design for Use in BICEP ............................... 33
2.3 Instrument Characterization .................................. 36
  2.3.1 Rotation Angle and Bias Signal .................... 36
  2.3.2 Instrumental Polarization ............................ 38
  2.3.3 Rotation Angle Calibration ......................... 38
2.4 Galactic Observations ...................................... 40
  2.4.1 Observing Strategy .................................... 40
  2.4.2 Timestream Processing ................................ 42
2.5 Map Results ............................................... 61
  2.5.1 Noise and Systematics .................................. 67
  2.5.2 Non-FRM Pixels ........................................ 74
2.6 Results from Unbiased FRMs .............................. 77
2.7 Conclusions ................................................ 78
2.8 Acknowledgements ......................................... 79

Chapter 3 POLARBEAR Pointing ............................... 80
  3.1 An overview of POLARBEAR ............................... 80
    3.1.1 POLARBEAR Design .................................. 81
    3.1.2 Detectors and Focal Plane ............................. 84
    3.1.3 Inyo Mountains Engineering Run ...................... 85
    3.1.4 POLARBEAR Current Status .......................... 87
  3.2 Pointing Model ........................................... 90
  3.3 Cedar Flat Results ....................................... 92
    3.3.1 Optical Pointing ..................................... 93
    3.3.2 Data Analysis ....................................... 96
    3.3.3 Radio Pointing ...................................... 99
  3.4 Chile Results ............................................. 104
    3.4.1 Observations ...................................... 104
    3.4.2 Analysis and Results ................................ 105
  3.5 Acknowledgements ......................................... 112

Chapter 4 Detector Technology .................................. 113
  4.1 NIS Junction Background .................................. 113
    4.1.1 Superconductivity ................................... 113
    4.1.2 Corrections to BCS Theory ............................. 115
    4.1.3 NIS Tunneling Theory ................................ 116
    4.1.4 Cooling Power ....................................... 125
  4.2 TES Bolometers ........................................... 133
    4.2.1 Introduction ...................................... 134
    4.2.2 Bolometer Noise .................................... 137
    4.2.3 Noise Limitations .................................... 140
  4.3 UCSD Fabrication .......................................... 142
    4.3.1 POLARBEAR Pixel Fabrication ......................... 143
    4.3.2 Seating Wafers for POLARBEAR-2 ..................... 150
    4.3.3 AlMn Bolometers .................................... 154
  4.4 Payload Cooling ........................................... 160
    4.4.1 NIST Junctions .................................... 160
    4.4.2 Design Motivation ................................... 162
**LIST OF FIGURES**

| Figure 1.1: | The original Hubble Diagram | 2 |
| Figure 1.2: | Pictorial explanation of the expanding universe | 3 |
| Figure 1.3: | The COBE FIRAS measurement of the CMB blackbody spectrum | 10 |
| Figure 1.4: | Full sky CMB anisotropy maps from (a) COBE DMR and (b) WMAP | 11 |
| Figure 1.5: | WMAP measurement of the CMB temperature power spectrum with ΛCDM best-fit model plotted | 13 |
| Figure 1.6: | Sensitivity of CMB temperature power spectrum to cosmological parameter variation | 16 |
| Figure 1.7: | Cartoon depiction of $E$ and $B$-mode polarizations | 18 |
| Figure 1.8: | Pictorial representation of Thomson scattering from a local quadrupole moment | 19 |
| Figure 1.9: | Depiction of the quadrupole moment resultant from scalar perturbations | 20 |
| Figure 1.10: | $E$-mode polarization generated from scalar perturbations of a single Fourier mode | 21 |
| Figure 1.11: | Both $E$ and $B$-mode polarization resultant from a single Fourier mode of a gravitational wave | 22 |
| Figure 1.12: | The temperature and $E$-mode cross correlation power spectrum, $E$-mode power spectrum, and $B$-mode power spectrum as measured by numerous CMB experiments | 23 |
| Figure 1.13: | A depiction of (a) the effects of gravitation lensing on temperature, $E$, and $B$-modes and (b) the resultant $E$ and $B$-mode power spectra | 26 |
| Figure 2.1: | Cross-Sectional schematic of a FRM. | 34 |
| Figure 2.2: | (a) The BICEP focal plane for the 2006 observing season and (b) the relative locations of the FRMs within the optical path. | 35 |
| Figure 2.3: | Raw detector response of a PSB pair and calculated rotation angle of the FRM shown for a single FRM calibration run. | 37 |
| Figure 2.4: | Dielectric sheet calibration for FRM150A. The top panel shows the 1 Hz FRM modulated signal. The bottom panel shows the split demodulated timestreams at ±45 degrees. | 39 |
| Figure 2.5: | 100 GHz FDS Model 8 Galactic dust emission prediction [1] plotted in celestial coordinates on a logarithmic scale. The outlined regions show the areas of integration for the shallow and deep FRM observations. | 42 |
| Figure 2.6: | Figure showing the effects of deconvolution on the modulated timestreams. | 44 |
| Figure 2.7: | Plots of the bias cleaning process. | 47 |
| Figure 2.8: | Illustration of time domain demodulation. | 48 |
Figure 2.9: FRM demodulation test schemes showing (a) a demodulated white noise signal with a break in the normal periodicity of the bias and (b) a test case with input temperature, polarization, and white noise.

Figure 2.10: Four panels showing masks of the Galactic plane region increasing in diameter from left to right, top to bottom.

Figure 2.11: Summed detector timestreams for FRM deep integration Galactic half-scans with different filtering methods.

Figure 2.12: Instrumental polarization measurements from the elevation nods for a single bolometer.

Figure 2.13: A plot showing the results of the summed (left half) and differenced (right half) demodulated timestreams for each of the individual FRM bolometers. The correlation between the two indicates IP from the FRMs and not another mechanism.

Figure 2.14: A plot summarizing the instrumental polarization values found for each bolometer during each observation.

Figure 2.15: Pixel-pixel scatter plots of the Galactic FRM maps versus the nominal BICEP maps with identical masking and filtering.

Figure 2.16: Integration time and temperature maps from the shallow FRM observation on April 21, 2006.

Figure 2.17: FRM and BICEP maps of the deep integration region.

Figure 2.18: A schematic of the difference between filtering the nominal BICEP Galactic half-scans and the FRM observation half-scans.

Figure 2.19: Masks used for processing BICEP data to mimic the FRM observations.

Figure 2.20: $T$, $Q$, and $U$ differenced data from individual detectors within a single PSB pair.

Figure 2.21: Maps of the difference between the FRM and non-FRM BICEP observations.

Figure 2.22: Plot of rms map noise for the scan-direction difference maps as a function of the portion of total integration time used.

Figure 2.23: Histograms of the noise at 100 GHz (left) and 150 GHz (right) in the unsmoothed scan direction jackknife maps for the FRM deep observations. Gaussian fits using the derived mean and standard deviation at each frequency are shown for comparison.

Figure 2.24: Maps accumulated from the timestreams of the non-FRM pixels during FRM observations.

Figure 2.25: Maps accumulated from the timestreams of the non-FRM pixels during FRM observations.

Figure 2.26: Average power spectral distributions calculated for the pair-sum and pair-differenced FRM PSBs at 100 and 150 GHz during Galactic observations in 2006 while the FRMs remained unbiased.
Figure 3.1: Projection of POLARBEAR limits on the B-mode power spectra after one observing season. The inflation model plots with $r=0.1$ and $r = 0.025$ are also shown for comparison. Figure courtesy of Nathan Miller.

Figure 3.2: Image of the Huan Tran Telescope at the James Ax Observatory in 2012 with important components labeled. Incident light is reflected off the primary mirror, onto the secondary mirror (not seen in the image), into the cryogenic receiver, and onto the focal plane. Figure courtesy of Zigmund Kermish.

Figure 3.3: Cross section of the POLARBEAR receiver with optical ray trace paths shown.

Figure 3.4: (a) An image of the full POLARBEAR focal plane as seen from the top (b) A single hex wafer plus vertically integrated readout (c) Microscope image of a single POLARBEAR pixel with important components labeled. Image courtesy of Zigmund Kermish.

Figure 3.5: (a) Co-added map of Jupiter used to characterize beam Gaussianity, ellipticity, and beam-width. b) Temperature and polarization map of Tau-A from 2 hours of observing data with 12 pixels. Data for both plots were obtained during the 2010 Inyo Mountains, California engineering run.

Figure 3.6: Geographical location of POLARBEAR and an image of the assembled telescope in the Atacama Desert.

Figure 3.7: (a) Fit for pixel locations in the POLARBEAR focal plane and (b) co-added beams for each device wafer.

Figure 3.8: Noise power spectrum of the POLARBEAR array summed and differenced timestreams.

Figure 3.9: POLARBEAR’s (a) current CMB observing patches, (b) preliminary CMB temperature map for a single patch, and (c) total observation and CMB observation time since first light in April 2012.

Figure 3.10: POLARBEAR beam broadening caused by a 2 arcminute pointing error in azimuth and elevation.

Figure 3.11: The main components of the star camera used for optical pointing of POLARBEAR.

Figure 3.12: The star camera mounted to the telescope during assembly in Chile, with me in the foreground.

Figure 3.13: Star camera data taken for a single source.

Figure 3.14: Fits and residuals for the star camera pointing model derived from the data taken on June 27th, 2010.

Figure 3.15: Raw (a) elevation, (b) azimuth, and (c) bolometer timestreams for a single Saturn scan.

Figure 3.16: Radio pointing beam map and pointing center fit for a single Saturn observation.
Figure 3.17: (a) Radio pointing fit for a single Saturn scan set from rising till setting and (b) the corresponding residuals. 103
Figure 3.18: (a) Radio pointing fit using the second iteration of the beam fits for all the pointing data taken in Cedar Flat and (b) the resultant residuals. 104
Figure 3.19: Single pixel and co-added maps illustrating the effects of masking on radio pointing observations using RCW38 data from October 2012. 106
Figure 3.20: (a) The pointing model fits and (b) residuals for the first iteration of radio pointing data from June 20th, 2012 until April 2013 in the Chilean deployment of POLARBEAR. 107
Figure 3.21: (a) The pointing model fits, (b) residuals, and (c) vector residuals in az/el space for the second iteration of radio pointing data from June 20th until the present in the Chilean deployment of POLARBEAR. 109
Figure 3.22: Pointing error systematics in azimuth and elevation from June 20, 2012 until April 2013 investigated via (a) histograms of the residuals and (b) scatter plots of the residuals as a function of time. 110
Figure 3.23: (a) Source coverage for scans taken from late April 2012 until June 20th, 2012 and (b) the pointing model fits to the data. 111
Figure 3.24: Coadded map of Jupiter for all pixels and all observations taken from late April 2012 until April 1, 2013. 112

Figure 4.1: Plot of the normalized superconducting density of states. 115
Figure 4.2: Measured superconducting band gap dependence on temperature for a number of metals. 116
Figure 4.3: Plots of the BCS density of states with Dynes Smearing in (a) normal and (b) logarithmic units. 117
Figure 4.4: Current-voltage curves of the first superconducting tunnel junction taken from [2]. 118
Figure 4.5: Schematic of an NIS tunnel junction. 120
Figure 4.6: Body of IV curves for an NIS junction at different temperatures. 122
Figure 4.7: Energy band diagrams of an NIS junction at $T = 0$ with increasing bias voltages. 124
Figure 4.8: Energy band diagrams of an NIS junction at $0 < T < T_c$ with increasing bias voltages. 125
Figure 4.9: Plots of (a) normal metal cooling power from a NIS junction and (b) optimization of superconducting temperature for maximal normal metal cooling power. 127
Figure 4.10: NIS junction IV curves for different values of Dynes parameters that approximate heavy subgap leakage. 128
Figure 4.11: Image of a SINIS tunnel junction structure. 129
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.12</td>
<td>SINIS junction cooling power and the different contributions to the power loading.</td>
</tr>
<tr>
<td>4.13</td>
<td>Two different types of suspended membranes: (a) traditional membrane and (b) suspended brick membrane.</td>
</tr>
<tr>
<td>4.14</td>
<td>Schematic of a bolometer.</td>
</tr>
<tr>
<td>4.15</td>
<td>The superconducting transition of niobium illustrating thermistor operation and negative electrothermal feedback.</td>
</tr>
<tr>
<td>4.16</td>
<td>POLARBEAR pixel with labeled components fabricated in UCSD’s nano3 facility.</td>
</tr>
<tr>
<td>4.17</td>
<td>The (a) double slot dipole antenna used in POLARBEAR pixels and (b) a cartoon of the slot dipole antenna coupled to a silicon lenslet.</td>
</tr>
<tr>
<td>4.18</td>
<td>A cartoon diagram of the sequential steps in the crossover fabrication process as viewed from above.</td>
</tr>
<tr>
<td>4.19</td>
<td>Two microscope images of the POLARBEAR TES bolometer made by the author.</td>
</tr>
<tr>
<td>4.20</td>
<td>The required materials and thicknesses for a POLARBEAR pixel.</td>
</tr>
<tr>
<td>4.21</td>
<td>Parallel rays incident upon (a) an elliptical surface and (b) an extended hemisphere.</td>
</tr>
<tr>
<td>4.22</td>
<td>Photographs of the fabrication process for POLARBEAR-2 seating wafers.</td>
</tr>
<tr>
<td>4.23</td>
<td>X-Ray fluorescence spectra for the AlMn sputtering target from ACI Alloys.</td>
</tr>
<tr>
<td>4.24</td>
<td>(a) Resistance versus temperature measurements for different doping concentrations of AlMn and (b) transition temperatures and normal state resistances derived from the data.</td>
</tr>
<tr>
<td>4.25</td>
<td>Electronic transport tests taken on AlMn dark bolometers fabricated at UCSD.</td>
</tr>
<tr>
<td>4.26</td>
<td>(a) SEM image of NIST membrane coolers and (b) a cross sectional schematic of a single NIS junction coupled to a suspension leg.</td>
</tr>
<tr>
<td>4.27</td>
<td>CAD rendering of (a) the cooling platform assembly and (b) the Aluminum endcaps on the stage.</td>
</tr>
<tr>
<td>4.28</td>
<td>The completed cooling platform shown (a) assembled and by itself, (b) held in place by the Aluminum support bracket, and (c) being tensioned during assembly.</td>
</tr>
<tr>
<td>4.29</td>
<td>Total thermal loading as well as the contribution from all the individual parasitic heat leaks on the NIS junction cooled platform.</td>
</tr>
<tr>
<td>4.30</td>
<td>Stage temperature as a function of time calculated from the net cooling power and the thermal conductivity of the stage.</td>
</tr>
<tr>
<td>4.31</td>
<td>An image of the completed stage mounting in the heat shield.</td>
</tr>
</tbody>
</table>
Figure 4.32: The modular platform for attaching NIS junctions to the stage shown (a) in an image with the entire assembly and (b) as a CAD rendering with Aluminum end braces in place. 172

Figure 4.33: (a) A CAD rendering of the heat switch in the shield and stage assembly and (b) a close-up photograph of the heat switch and Cu foil that connects the switch thermally to the 300 mK shield. 174

Figure 4.34: Images showing the integration of the electronic circuitry into the cooling platform package. 176

Figure 4.35: Photographs of the stage testing performed after assembly of the stage. 177

Figure 4.36: The platform assembly inside of the heat shield loaded into the cryostat at NIST. 180

Figure 5.1: Schematic of NIS tunnel junction fabrication process. 185

Figure 5.2: Energy dispersive X-ray spectra taken immediately before and after the Al₂O₃ barrier was deposited using ALD. 187

Figure 5.3: NIS junction IV curves as a function of temperature along with theoretical curves based on the measured stage temperature. 188

Figure 5.4: Best fit temperatures for the NIS junction versus a calibrated Cernox on the refrigerator head. 189

Figure 5.5: Surface roughness plots of the Au film before any treatment for the tunnel barrier done by (a) scanning electron microscopy and (b) phase sensitive interferometry. 191

Figure A.1: Microscopy images of the sequential steps in the fabrication process for AlMn NIS tunnel junctions. 196

Figure A.2: IV curve of an AlMn SINIS refrigerator at 220 mK. 197

Figure B.1: Sample box that houses the junctions on the cryogenic stage. 199

Figure B.2: Room temperature electronics for NIS junction testing. 200
LIST OF TABLES

Table 2.1: DSC calibration results for all six PSB pairs that were modulated by the FRMs. ................................. 40
Table 2.2: Details of each FRM observation, ordered sequentially. .......... 41
Table 2.3: The rms map noise for each of the three difference maps analyzed for the deep FRM observations. ........... 68
Table 3.1: A description of the physical meaning of the individual parameters that contribute to the modeled pointing error. ............... 91
ACKNOWLEDGEMENTS

This is, perhaps, the most daunting section of my thesis that I will write. I am deathly afraid of leaving off any of the important people in my life, so please forgive me if I do. I’ll start at the beginning. To my family, thank you for all the love and support during these years. My parents, who told me when I was young that, “I could do anything I put my mind to.” Although I’m not sure I believe that now, it did set an interesting tone for my thesis work and probably my entire life. To my siblings, the Asian Rice Cracker Association, I couldn’t have asked for a more varied, intelligent, fun-loving group of people to grow up with. Thanks for showing me that there are so many things outside of the lab that I should get to know and ensuring that I grew up in an environment where my aspergers couldn’t take full form. I’m a reasonable human being (sometimes) because of you guys.

I have been exceptionally blessed during graduate school with great lab-mates and friends. Without Dr. Nathan Stebor in the lab I would have worked ten times as hard and accomplished 1/5 as much; Nate you’ve been the guiding force behind most of my decisions as well as the majority of my daytime drinking. Thank you. Dave Boettger provided a constant wall for me to bounce ideas off of and read all my crappy papers more times than everyone else combined. I’ve managed to ask our new lab addition Kam Arnold more science and life questions in these past 5 months than I’ve asked during the rest of graduate school. To my army of undergraduates - Andrew Crooks, Steve Choi, Lisa Krayer, James Feng, and Chris Aleman - I owe you countless thanks for all the work you’ve done for me. I hope an army as good as you guys awaits me at Intel. Brandon Wilson has been a constant source of smiles in the lab for me over the past year and has introduced me to the phrase, “bag of dicks,” which I will be using in my personal and professional life from now until forever. There are many more people to thank from the lab: Matt Atlas, Darcy Barron, Fred Matsuda, Chang Feng, Jon Kaufman, and Praween whose last name will go unwritten. Outside of UCSD I’ve been blessed by help from many smart minds at Berkeley (Zigmund Kermish, Bryan Steinbach, and Aritoki Suzuki to name a few), NIST Boulder (Peter Lowell and Joel Ullom), and other institutions (John Kovac, Denis Barkats, and the remainder
of the BICEP and POLARBEAR teams).

I have to express my gratitude outside the lab to a great many others. The UCSD craft center, may she rest in peace, was my safe haven on campus for years. I don’t think I would have made it through grad school without my glassblowing addiction keeping me sane. To all the members of the craft center - Joyce Rooks, Patty Yockey, Kathleen Mitchell, Clay Logan, Buzz Blodgett, and Ron Carlson - who laughed at me when I said I did cosmological astrophysics, thanks. Special thanks to my glassblowing partner Jennie Brennan, who listened to be bitch endlessly about the trials and tribulations of graduate work. To the judo guys - Ross, Jason, and Daniel - who provided many laughs and a physical outlet for my frustration. To Jonathan, my rock, who supported (and continues to support) me in my endeavors, who showed me how to take calculated risks, who engaged me in all weird manner of intellectual conversation, and, most of all, who showed me that getting a PhD in physics is possible, endless thanks. And finally, to Kelsey, my everything, for your everlasting belief that I’m somehow special and your endless support and patience. This would not have been possible without you.

On the advisor front, huge thanks to Jim Eckert, the only research advisor who could have turned a juvenile delinquent into an even worse adult delinquent with a strong passion for experimental physics. Thanks for setting my on my way. And to Eric Fullerton, although I didn’t finish under your guidance you were always a great guy and helped me find the path that I most wanted. And, finally, to Brian Keating, who always pushed to get the best out of me, even when I wasn’t sure what that was.

Technical Acknowledgements:


The introduction to Chapter 3 is a reprint of material as it appears in B. Keating, S. Moyerman, et. al., Ultra High Energy Cosmology with POLAR-
BEAR, DPF 2011 conference proceedings, arXiv 1110.2101v1 (2011). The dissertation/thesis author was the primary author of this paper.

Chapter 5 is a reprint of material as it appears in S. Moyerman, et. al., Atomic layer deposition of superconducting tunnel barriers, Submitted for Publication to Superconductor Science and Technology (2013). The dissertation/thesis author was the primary investigator and author of this paper.
VITA

2006 B. S. in Mathematics, Harvey Mudd College, Claremont
2006 B. S. in Physics, Harvey Mudd College, Claremont
2009 M. S. in Physics, University of California, San Diego
2013 Ph. D. in Physics, University of California, San Diego

PUBLICATIONS

B. Wilson, S. Moyerman, et. al., A cryogenic stage for user-supplied payloads with Normal-Insulator-Superconductor tunnel junctions, Cryogenics In Preparation (2013)

S. Moyerman, et. al., Atomic layer deposition of superconducting tunnel barriers, Submitted for Publication to Superconductor Science and Technology (2013)


C. Bordel, et. al., Fe spin reorientation across the metamagnetic transition in strained FeRh thin films, Phys. Rev. Lett. 109, 117201 (2012)


ABSTRACT OF THE DISSERTATION

Sensitive Detection of CMB B-Mode Polarization: Instrumentation and Systematics

by

Stephanie Moyerman

Doctor of Philosophy in Physics with Specialization Computational Science

University of California, San Diego, 2013

Professor Brian Keating, Chair

Numerous experiments in the last two decades have shown that the cosmic microwave background (CMB) is a powerful cosmological probe. The temperature anisotropy of the CMB has now been mapped to exquisite precision by many experiments, yielding tight constraints on the standard ΛCDM cosmological model. Many current and upcoming experiments focus on measuring CMB polarization, in particular the $B$-mode polarization, which potentially encodes information from long before the epoch of matter-radiation decoupling. However, the magnitude of the inflationary $B$-mode signal is constrained by an upper limit of tens of nK, which represents a massive experimental challenge. Foreground contamination and systematic effects, among other factors, further increase the difficulty of detection.
A measurement of this signal therefore requires the development of dedicated telescopes with exquisite control of systematics and large kilo-pixel arrays of background limited detectors. This thesis describes my work on Cosmic Microwave Background polarization studies. Specifically, it describes my data analysis efforts on two CMB polarization telescopes, BICEP and POLARBEAR, my contribution to hardware efforts on POLARBEAR, and my design and fabrication work on next generation detector arrays.
Chapter 1

Background and Motivation

“Let’s start at the very beginning, a very good place to start.” (The Sound of Music)

This chapter offers the reader a brief overview of modern cosmology with special focus on the CMB and, more specifically, CMB polarization. A brief description of the state of the field as well as current and future goals are discussed towards the end. Obstacles to experimental progress and to future experiments are highlighted as the motivation for this thesis.

1.1 Modern Cosmology

The foundation of modern cosmological models is the Cosmological Principle (CP), which states that the universe is everywhere homogeneous and isotropic. Homogeneity refers to spatial invariance; the universe appears the same from every location. Isotropy is an invariance with respect to direction; the universe appears the same when observed in any direction from one fixed vantage point. Although not true locally, the CP appears to be true on length scales above approximately 400 megaparsecs.

One of the two cornerstones in the birth of modern cosmology came in 1929, when the astronomer Edwin Hubble derived his famous Hubble’s Law

\[ \vec{v} = H \vec{r} \]  

(1.1)
where $\vec{v}$ is the object’s velocity, $H$ is now known as the Hubble constant (a misnomer - Hubble’s constant actually varies in time), and $\vec{r}$ is the distance to the object. Simply put, Hubble’s Law states that all objects in the universe are receding at a rate that is proportional to their distance from the observer. Figure 1.1 shows a plot of Hubble’s original data, the motivating data for Hubble’s Law.

![Figure 1.1: The original Hubble diagram used to measure the relationship between velocity and distance. Image taken from [3]](image)

Hubble’s findings had great implications for cosmology. At a glance, the data appear to support the hypothesis that the Earth occupies a very special place in the universe from which all objects recede. Since the Copernican model of our solar system had already eliminated this possibility, the only other explanation was the expansion of space itself. This would yield an observed recession of all objects not only from Earth but from every location in the universe. In this way, Hubble’s Law led directly to the theory of the expanding universe. The cause of this expansion is still being actively sought out today.

A common analogy is often made using tacks in a rubber sheet. As the sheet is pulled apart, the distance between the tacks expands as shown in Figure 1.16. From the position of any one tack, the distance to all other tacks has increased by the scale factor, $a$. The expansion makes it appear as though every other object is receding.
Figure 1.2: Pictorial representation of tacks in an expanding sheet, or non-interacting objects in an expanding universe. As time passes the grid expands, and an observer at any location sees all other objects receding. The distance between any two objects is given by \( r \times a(t) \). For convenience, \( a(t) = 1 \) today.

The other cornerstone in the birth of modern cosmology is Einstein’s theory of General Relativity (GR), which describes the spatial curvature of the universe as a function of its constituents. A solution to the equations of GR that is compatible with Hubble’s findings and the cosmological principle is the Friedmann-Robertson-Walker (FRW) metric

\[
ds^2 = -c^2 dt^2 + a^2(t) \left[ \frac{dr^2}{1 - kr^2} + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 \right],
\]

where \( c \) is the speed of light in free space, \( k \) describes the curvature of space, and \( r \) is known as the co-moving coordinate. As mentioned previously, the time evolution of the scale factor, \( a(t) \), is a dynamical question that is governed by the GR equations. The Hubble parameter can be described by the scale factor as

\[
H(t) = \frac{\dot{a}}{a}
\]

where the dot represents the derivative with respect to time.

Using the FRW metric, the measurable distance \( R \) between any two points can be found by aligning the axes such that \( d\theta = d\phi = 0 \), setting \( dt = 0 \), corresponding to a snapshot in time, and integrating with respect to \( r \)

\[
R = a(t) \int_0^r \frac{dr}{\sqrt{1 - kr^2}}.
\]
This gives us a method for quantifying time and space that is independent of our chosen coordinate system. If $k = 0$ the FRW implies a flat universe, whereas if $k = 1$ or $k = -1$ we have a spherical or hyperbolic universe, respectively.

The Einstein Equation of General Relativity is given by

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{\Lambda g_{\mu\nu}}{c^2} + \frac{8\pi G T_{\mu\nu}}{c^4},$$

(1.5)

where $R_{\mu\nu}$ is the Riemann Tensor, $R$ is the Ricci Scalar, $g_{\mu\nu}$ is the metric, $\Lambda$ is a cosmological constant defined by the vacuum energy, $G$ is the gravitational constant, and $T_{\mu\nu}$ is the stress-energy tensor. Using a simplified model of the universe, the stress-energy tensor can be written as [4]

$$T_{\mu\nu} = (p + \rho)u_{\mu}u_{\nu} - pg_{\mu\nu}$$

(1.6)

where $p$ is pressure, $\rho$ is total energy density, $u$ is the four velocity. Conservation of energy in GR dictates that the covariant derivative of the energy momentum tensor must vanish

$$\nabla_{\mu}T^{\mu\nu} = 0.$$  

(1.7)

Combining this constraint with Equations 1.2 and 1.6 yields the result

$$\dot{\rho} + 3H(\rho + p) = 0.$$  

(1.8)

To solve for the evolution of the energy density, we must derive a relation between $\rho$ and $p$. Using Weyl’s postulate we can treat universe as an ideal fluid in which $p = \alpha \rho$. The constant $\alpha$ is determined by the source of the energy with which we are concerned. Substituting this into Equation 1.8 admits the solution

$$\rho \propto a^{-3(1+\alpha)}$$

(1.9)

after integration. Many cosmological sources obey the above relation. A matter dominated universe has $\alpha = 0$. For an equivalent amount of matter, the energy density decreases proportionally to the volume, scaling as $a^{-3}$. For a radiation dominated universe $\alpha = 1/3$. As the scale factor increases the radiation is shifted to longer wavelength, which decreases the energy proportionally. Combining this with a decrease in energy based on volumetric expansion, we expect that radiation
energy density scales as $a^{-4}$. The vacuum energy of space should be constant, proportional to $a^0$ or $\alpha = -1/3$.

Using the energy momentum tensor, FRW metric, and Einstein equations, one can derive the Friedmann equations in their general form:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p) + \frac{\Lambda}{3}$$  \hspace{1cm} (1.10)

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G\rho}{3} - \frac{k}{a^2} + \frac{\Lambda}{3}$$  \hspace{1cm} (1.11)

Here, the speed of light, $c$, has been set to unity.

Currently we live in a vacuum energy dominated universe with $p = 0$ and $\rho_T = \rho + \frac{\Lambda}{8\pi G}$. Plugging these into the Friedmann equations gives

$$\frac{2\dot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} = 0$$  \hspace{1cm} (1.12)

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{k}{a^2} = \frac{8\pi G\rho_T}{3}.$$  \hspace{1cm} (1.13)

Equation 1.12 is often referred to as the acceleration equation. Again, the curvature of our universe depends on the constant $k$.

The Einstein DeSitter Model is achieved when we set $k = 0$ above. Substituting this value into Equation 1.13 gives

$$\left(\frac{\dot{a}}{a}\right)^2 = H(t)^2 = \frac{8\pi G\rho_T}{3},$$  \hspace{1cm} (1.14)

which can be used to solve for the critical density as a function of time

$$\rho_C = \frac{3H_0^2}{8\pi G}.$$  \hspace{1cm} (1.15)

If the density of the universe is greater than this then the universe will be closed; the mass will be great enough such that gravitational collapse will occur as time increases. Conversely, if the density is smaller than $\rho_c$ then the universe is open; the gravitational force is insufficient to stop the universe from accelerating out to infinity. A universe with this critical density will be flat. The mass density in such a universe is fine-tuned such that the universe will expand forever and barely escape gravitational forces. Based on the observed value of the Hubble constant, this critical density equates to roughly six hydrogen atoms per cubic meter today.
1.1.1 The Big Bang Model

The model of the universe described above is often referred to as the standard or Hot Big Bang. As the universe expands the temperature \( T \) evolves as

\[
T \propto a^{-3(\gamma - 1)},
\]

(1.16)

where \( \gamma \) is the ratio of specific heats and \( a \) is the scale factor. For the early radiation dominated universe, we have \( \gamma = 4/3 \), which gives \( T \propto a^{-1} \). Extrapolating the expansion of the universe all the way back in time to \( a = 0 \), one infers that the universe started from a point of infinite temperature and zero size. Current models of physics break down at energy scales of approximately \( 10^{19} \) GeV (the Planck scale), allowing for massive speculation about the causal mechanisms of the early universe with energies above this value. It is not until energies below this scale have been reached that our known quantum mechanics, statistical mechanics, and General Relativity can be accurately employed. Equation 1.16 makes it obvious that from this point, as time evolves forward and the scale factor increases, two things will happen: the universe will expand and cool.

Until approximately \( 10^{-6} \) seconds, all matter in the universe existed as free quarks and leptons. At \( 10^{-6} \) seconds, quarks condensed under the strong interaction to form baryons, such as neutrons and protons, in a process known as baryogenesis. Several minutes later the universe cooled below the threshold for nuclear fusion, and light nuclei, such as deuterons and alpha particles, formed. The theory describing this phenomenon is known as Big Bang nucleosynthesis, which accurately predicts the abundances of light elements found in the universe today. This agreement is one of the pillars of success on which the Big Bang model stands.

For the next approximately 400,000 years, the universe was still sufficiently hot to prohibit electrons from binding to nuclei, and matter existed in a highly ionized state. During this time, the matter was tightly coupled to the photons through Thomson scattering from charged particles. Towards the end of this time period the universe cooled sufficiently so that nuclei and electrons could combine to

\footnote{Here, the standard Big Bang is meant not to include the theory of inflation, which will be discussed in later sections}
form neutral atoms. This time period is known as recombination. Recombination marks a time when the universe turned from opaque to transparent. The cross-section for photon scattering off bound, uncharged atoms is much smaller than the scattering cross-section of ionized atoms and so photons could now propagate freely through space without scattering. It is these same photons that we observe today as the Cosmic Microwave Background (CMB), a relic of the early universe at just 0.3% of its current age.

Recombination and last scattering occurred at a cosmological redshift of approximately 1100, when the universe was approximately 3000 K. As the redshift decreased over time, the wavelength, the temperature decreased proportionally. Today, at cosmological redshift zero, the CMB has cooled to a temperature of approximately 2.7 K. The CMB is a robust prediction of the expanding universe, and its discovery by Penzias and Wilson in 1965 [5] is yet another pillar on which the Big Bang theory stands.

1.1.2 Problems with the Standard Big Bang

The Hot Big Bang accounts for many of the overarching features of the modern universe. Big Bang nucleosynthesis accurately predicts the relative abundance of light elements, such as Hydrogen, Lithium, Beryllium, and Helium, in the universe today. It also predicts the existence and temperature of the CMB and the expansion of the universe, both of which have been observed with great precision.

Unfortunately, for several reasons, the Big Bang has been shown to provide an incomplete picture of the origins of our universe. The following is meant to highlight its shortcomings: the horizon and flatness problems.

The Horizon Problem

The horizon problem describes the observed homogeneity of the universe on large angular scales. The cosmic microwave background provides an image of the last scattering surface and is remarkably uniform over the whole sky to approximately a few parts in $10^5$. However, the horizon size at the surface of last scattering was approximately 2 degrees on the sky today, so areas with greater
separation would have been causally disconnected at the time of last scattering and would never have had the chance to thermally equilibrate. The uniformity of the CMB thus presents a severe problem for the standard Big Bang model, as it has no mechanism to explain why causally disconnected regions are in thermal equilibrium.

**Flatness**

The flatness problem is an example of a fine-tuning problem, where the current universe could only result from a very finely tuned set of initial conditions. The results of the recent Planck satellite confirm that the universe is spatially flat to within a margin of error of 0.4% [6]. Defining \( \Omega(t) = \rho(t)/\rho_c(t) \) as the total density of the constituents of the universe and knowing that our currently observed universe is flat, we require that our parameter \( \Omega = 1 \). However, using the ratios of baryonic matter, radiation, and dark matter to extrapolate back to our early universe, we find that

\[
|\Omega(t) - 1| = 10^{-16} \tag{1.17}
\]

and

\[
|\Omega(t) - 1| = 10^{-60} \tag{1.18}
\]

at the times of recombination and the Planck epoch, respectively. The numbers seem arbitrary and exceptionally fine tuned. Small deviations from these numbers would result in a universe that collapsed or “blew up” shortly after the Big Bang, and yet we still exist to observe the flat universe today. The Big Bang does not leave open the possibility for a flat universe across time without precise fine tuning.

**1.1.3 Inflation**

It was the monopole problem, the lack of measured magnetic monopoles, that led A. Guth to first postulate the theory of inflation in 1980 [7]. Inflation describes the superluminal exponential expansion of space-time, driven by a negative vacuum energy in the early universe. The inflationary epoch begins at approximately \( 10^{-36} \) seconds after the Big Bang and ends quickly, before the universe is
10^{-32} seconds old. In order for the causal horizon to match our observations today, the universe must have expanded by at least a factor of $10^{25}$ during this short time.

Although Guth was focusing specifically on the monopole problem, the theory of inflation solves the horizon and flatness problems as well. Inflation provides a mechanism for seemingly disconnected regions today to have been in causal contact in the past. Essentially, before inflation, the entire universe existed in a small, causally connected region. Inflation also smoothes local curvatures and inhomogeneities in space. The rapid expansion of space time suppresses location deviations from flatness, alleviating the fine tuning problem.

In this way, inflation resolves the flatness and horizon problems that appeared as gaping holes in the standard model. Despite this and all other evidence in its favor (discussed in the next section), inflation remains largely speculative; to date, none of the unique predictions of inflation have been observed. The observational data in favor of inflation is currently degenerate among competing theories; a plethora of inflationary models, as well as more recently proposed substitutes for inflation, are all viable candidates. Fortunately, there is one non-degenerate data set that can be used not only to validate inflationary theory, but also to characterize the energy scale that drove inflation: the B-mode polarization of the Cosmic Microwave Background.

\section{1.2 The Cosmic Microwave Background}

The CMB is the oldest observable electromagnetic relic of the early universe - a nearly uniform glow of radiation that permeates all of space. In the 47 years since Penzias and Wilson’s initial discovery, steady progress in instrumental sensitivity has allowed researchers to characterize the CMB with much greater depth. In the early 1990s, the far infrared spectrometer (FIRAS) on board the COBE satellite measured the spectrum of the CMB to unprecedented precision [8], confirming its blackbody nature and thermal origins with a best fit blackbody spectrum at 2.725 K. This is shown in Figure 1.3, where the error bars have been exaggerated to $400\sigma$. The CMB is the most precisely measured blackbody to date.
Figure 1.3: COBE FIRAS measurement of the CMB blackbody with 400σ error bars. The best fit temperature is 2.725 K. Data and error bars taken from [8].

1.2.1 CMB Temperature Anisotropy

The CMB is nearly perfectly uniform across the sky, but temperature fluctuations do exist at the level of a few parts in $10^5$. The differential microwave radiometer (DMR) on COBE was the first measurement of CMB anisotropy at 7° resolution, and is shown in Figure 1.4(a). At this point, an arms race of CMB ground and balloon-based experiments set out to characterize the anisotropy of the CMB [9, 10, 11, 12, 13] with great success. The standard for many years has been the measurements of the Wilkinson Microwave Anisotropy Probe (WMAP), which recently measured full sky CMB anisotropy at a resolution of 13 arcminute resolution (Figure 1.4) [14]. The standard has shifted even more recently to the newly released Planck data [15], which has measured the full sky CMB anisotropy with an increased resolution of 7.1 arcminutes at 143 GHz.
Figure 1.4: Full sky CMB anisotropy measurements by (a) COBE DMR and (b) WMAP. The average CMB temperature and dipole terms have been subtracted from the maps. Emission from the Galaxy can be seen as the red stripe across the middle of both maps. Images courtesy of NASA Space Science Team.
1.2.2 Temperature Power Spectrum

The small fluctuations of the CMB temperature field can be represented by \( \Theta(\hat{n}) = \Delta T/T \) where \( T \) is the 2.725 K monopole moment of the CMB and \( \hat{n} \) is the direction on the sky. The multipole moments of the field can be written as

\[
\Theta_{lm} = \int dn Y_{lm}^*(\hat{n}) \Theta(\hat{n}),
\]

(1.19)

where \( Y_{lm} \) are the spherical harmonics. If the field is Gaussian, then the moments are fully characterized by their power spectrum

\[
\langle \Theta_{lm}^* \Theta_{lm} \rangle = \delta_{ll'} \delta_{mm'} C_l.
\]

(1.20)

The complete characterization by Fourier analysis makes it common for CMB anisotropy to be studied in Fourier space. The index \( l \) corresponds to an angular scale of \( \theta = 2\pi/l \), so angular size decreases with increasing values of \( l \). The power spectrum is typically displayed using the power per logarithmic interval as

\[
\Delta^2_l = \frac{l(l + 1)}{2\pi} C_l.
\]

(1.21)

A measurement of the angular power spectra from WMAP is given in Figure 1.5. The peaks and troughs indicate that there are many correlated hot and cold temperature variations at specific angular scales.

1.2.3 Acoustic Oscillations

What causes the shape and peaks in the observed temperature power spectrum? As explained in the previous section, the photons and baryons were tightly coupled at the time of last scattering, so the shape and structure in the power spectrum can be explained by examining inhomogeneities in this coupled fluid. It is easy to explain the peak and trough structure, known as acoustic oscillations, with some simple fluid mechanics. Neglecting the dynamical effects of gravity and baryons, the temperature perturbations in Fourier space obey the continuity equation

\[
\dot{\Theta}(k) = -\frac{1}{3} k v_\gamma,
\]

(1.22)
where $\dot{\Omega}$ is the derivative of $\Omega$ with respect to conformal time and $v_\gamma$ is the scalar velocity of the photon fluid, and the Euler equation of momentum conservation

$$
\dot{v}_\gamma = k \Theta(k).
$$

(1.23)

The small size of the temperature perturbations causes different Fourier modes to evolve independently for each value of $k$. For the remainder of this treatment we use $\Omega = \Theta(k)$ for brevity.

Differentiating Equation 1.22 and combining with Equation 1.23 yields the wave equation

$$
\ddot{\Theta} + c_s^2 k^2 \Theta = 0,
$$

(1.24)

where $c_s = \sqrt{p/\rho}$ is the speed of sound in the fluid. Initial perturbations to the fluid experience pressure gradients as a restoring force, causing them to oscillate at the speed of sound until recombination. The temperature oscillations are the manifestations of the heating and cooling of a fluid under compression and rarification by an acoustic wave. At recombination, this gives a temperature distribution of

$$
\Theta(\eta_*) = \Theta(0) \cos(k s_*),
$$

(1.25)
where $\eta$ is the conformal time, $s$ is the sounds horizon, and the subscript $^*$ denotes evaluation at recombination. In the limit of very small $k$ (i.e., very large angular scale waves), the perturbation is unchanged from the initial conditions. Smaller scale modes oscillate, and modes that are caught in a maximum or minimum at recombination result in peaks in the power spectrum. The peaks follow the simple harmonic relationship $k_n = n\pi/s^*$ where $n$ is an integer. In real space, this corresponds to a series of coherent peaks with an approximate $1^\circ$ spacing.

The treatment above begs the obvious question, what caused these temperature perturbations that lead to the CMB temperature peak and trough structure? Again, the answer lies with inflation. Inflation theories postulate the existence of a scalar field that drives the exponential expansion of the universe. Quantum fluctuations in the scalar field are expanded during the inflationary epoch and introduce variations to the spatial curvature of the universe. In the diagonalized Newtonian gauge, the spatial curvature is given by $\delta g_{ij} = 2a^2\Phi\delta_{ij}$ and the time fluctuation by $\delta g_{tt} = 2\Psi$. Typically, we set $\Psi \approx \Phi$ in magnitude.

Using general relativity, the time-time fluctuation in the metric gives a temporal shift of $\delta t/t = \Psi$. CMB temperature varies as the inverse of the scale factor and the evolution of the scale factor with respect to time and is given by

$$a \propto t^{2/3(1+\rho/\rho_c)}.$$  

Combining this information gives

$$\Theta = -\frac{\delta a}{a} = -\frac{2}{3}\left(1 + \frac{\rho}{\rho_c}\right)^{-1}\frac{\delta t}{t},$$

where one observes that the temperature fluctuations are directly dependent on the perturbation $\Psi$.

In this simplified treatment of acoustic oscillations, we have neglected the effects of gravity, baryons, and damping, each of which alter the power spectrum in unique ways. Gravity offers a competing force, $k\Psi$ to the pressure gradients $k\Theta$, giving rise to an effective oscillating temperature of $\Theta + \Psi$. This gives the inverse relationship between density and temperature; overdense regions are actually colder because photons lose energy climbing out of potential gravitational wells and vice versa for underdense regions. The addition of baryons has to main
effects: drag and restorative forces. Baryons yield extra inertia for the pressure and potential gradients to overcome, decreasing the sound speed of the fluid. The extra gravity provided by the baryons also increases the potential wells, leading to a boost in compression over rarification. Correspondingly, odd numbered peaks in the power spectrum have an elevated power over their even numbered counterparts. Finally, damping is caused by shear viscosity and heat conduction in the fluid. Inhomogeneities are damped by a factor of order $e^{-k^2\eta/\dot{\tau}}$ where $\tau$ is the Thomson optical depth. The damping scale $k_d$ is of order $\sqrt{\tau/\eta}$. Numerical integration shows that at last scattering $k_d s_* \approx 10$, which suppresses power spectrum beyond the third peak.

The CMB power spectrum is sensitive to variations in parameters in the standard cosmological model. For instance, increasing or decreasing the matter density shifts the relative heights of the odd numbered peaks with respect to the evens. Fitting the CMB power spectrum allows cosmologists to constrain the parameters of the standard model and provides a wealth of information about the early universe, such as geometry, components of the energy, dark matter, etc. Figure 1.6 highlights some of the parameter sensitivities in the CMB power spectrum.

As is evidenced from Figure 1.6, cosmological parameters exhibit a certain degree of degeneracy in the CMB temperature power spectrum. Although constraints can be placed on individual cosmological parameters, these constraints are interdependent. The temperature power spectrum appears to agree with inflationary predictions, but inflation is not uniquely supported and still remains unproven. What information remains in the CMB that allows us to gather more information about our early universe and robustly test inflation?

1.2.4 CMB Polarization Anisotropy

With temperature well characterized, many experimentalists have turned to exploring CMB polarization, a much fainter signal that potentially encodes new information about the early universe.

Any electromagnetic field can be quantified fully by the Stokes parameters: $I, Q, U,$ and $V$, where $I$ is the intensity (temperature), $V$ describes the circular
Figure 1.6: Changes to the behavior of the temperature power spectrum to four different cosmological parameters: (a) curvature defined by total mass, (b) dark energy, (c) baryon density, and (d) matter density. Figure taken from [16].
polarization, and $Q$ and $U$ describe the linear polarization. CMB polarization is generated through Thomson scattering, which results only in linear polarization, and so the circular polarization $V$ of the CMB is assumed to be zero. The Stokes parameters are given by

\begin{align}
I &= \langle E_x^2 \rangle + \langle E_y^2 \rangle \\
Q &= \langle E_x^2 \rangle - \langle E_y^2 \rangle \\
U &= \langle 2E_x E_y \sin(\theta_x - \theta_y) \rangle 
\end{align}

(1.28)

where $E_x$ and $E_y$ are the electric field magnitudes along orthogonal directions and $\theta_x - \theta_y$ is the phase difference between the two.

The linear polarization components $Q$ and $U$ form a useful mathematical basis for decomposing CMB polarization and are often used as the bases for analyzing CMB data. However, cosmologists have found that description of CMB polarization using the scalar $E$-modes and tensorial $B$-modes, linear combinations of $Q$ and $U$, is a much more useful construct. In the small-scale limit in Fourier space, the wavevector $k$ assigns the direction from which polarization is measured. The polarization aligned along both axes orthogonal to the wave vector are $E$-modes, whereas the polarization aligned at $+45^\circ$ and $-45^\circ$ from the wave vector are the $B$-modes. This is shown pictorially in Figure 1.7(a).

Another way of thinking about $E$ and $B$-modes is in terms of a curl or handedness. When observing polarization from the vantage point of the propagating plane wave, $E$-modes appear the same regardless of orientation along the propagating wave; they are reflectively symmetric. $B$-modes, on the other hand, are not reflectively symmetric; traveling against the propagating wave-vector switches their orientation. For this reason, it is often said that $E$-modes are curl-free whereas $B$-modes possess a curl or handedness. Examples of $E$ and $B$-mode polarization patterns are given in Figure 1.7(b).

### 1.2.5 Polarization Power Spectrum

Just as temperature can be expanded on a sphere using ordinary spherical harmonics, the polarization tensor can be described by the tensor harmonics with
Figure 1.7: Depiction of $E$ and $B$-mode polarization. (a) As the wavevector $k$ propagates along the $z$-axis, polarization in the $x-y$ plane aligned along and orthogonal to the wavevector is designated $E$ and polarization at $\pm 45^\circ$ is designated $B$. The $E$-mode polarization is symmetric when viewed from both $\pm z$, whereas the $B$-mode is not. This is highlighted by polarization patterns in (b). The $E$ mode is symmetric upon reflection about the vertical axis. The $B$-mode is not.

The resultant $E$ and $B$-mode power spectra are described by

$$
\langle E_{l m}^* E_{l' m'} \rangle = \delta_{ll'} \delta_{mm'} C_{l}^{EE}
$$

$$
\langle B_{l m}^* B_{l' m'} \rangle = \delta_{ll'} \delta_{mm'} C_{l}^{BB}
$$

CMB polarization has two requirements: free electrons for Thomson scattering and a quadrupolar temperature distribution at recombination. A free electron surrounded by a bath of photons with an isotropic or dipole distribution of temperature will scatter photons with no net polarization. However, if the electron is surrounded by a quadrupolar distribution of photon temperature, the scattered beam results in a net linear polarization as shown in Figure 1.8. Over time, many random scatterings would tend to cancel any overall polarization, so the problem of understanding CMB polarization reduces to understanding the quadrupolar moment at the last scattering surface.
Figure 1.8: A depiction of Thomson scattering from an electron surrounded by a local quadrupole moment. Warmer radiation (blue) incident from the $\hat{x}$ direction and colder radiation (red) incident from $\hat{y}$ direction combine to give a net scattered polarization (blue and red) in the $\hat{z}$ direction. Figure taken from [17].

1.2.6 Causes of Polarization

There are three main mechanisms that can cause the quadrupolar moment that results in a net polarization pattern of the CMB: scalar, vector, and tensor perturbations. Scalar perturbations are the same modes that cause the acoustic oscillations in CMB temperature; they are the only fluctuations in the fluid density caused by gravitational instability. The quadrupolar moment is established in the photon temperature distribution when gradients in the photon velocity from two neighboring crests (or locally underdense regions) meet a trough (or locally overdense region). In the image, an observer in the trough would observe hotter photons flowing towards them from top and bottom whereas colder photons surround them in the plane. An observer in a local crest would see a similar result, though the direction of the photon velocities would be reversed. This quadrupole pattern is represented by the spherical harmonic $Y_{l=2,m=0}$ as shown in Figure 1.9.

Scalar perturbations have an intrinsic direction associated with them based on the direction of the propagating wave vector $k$. Thomson scattering then results in polarization that is perpendicular or parallel to the this vector, resulting in an
Figure 1.9: Velocity gradients from hot to cold result in a local quadrupole moment represented by the spherical harmonic \( m = 0 \). The gradients are set by scalar perturbations. Figure taken from [17].

The strength and shape of the \( E \)-mode polarization pattern from scalar modes can be predicted by considering the monopole (\( \Theta_0 \)), dipole (\( v_\gamma \)), and quadrupole moments. The continuity equation (Equation 1.22) shows that the dipole moment is proportional to the first derivative of the monopole. Since the monopole is given by Equation 1.25, taking the derivative shows that the monopole and dipole moments are \( \pi/2 \) out of phase from each other at last scattering. The quadrupole moment that generates scalar mode polarization is of order \( kv_\gamma / \dot{\tau} \), and so is proportional to the dipole by the factor \( k / \dot{\tau} \approx 10 \). Thus, the \( E \)-mode power spectrum resultant from scalar modes (the quadrupolar term) is expected to be out of phase with the temperature anisotropy (\( \Theta \)) and reduced in magnitude by at least a factor of 10. The cross-correlation of \( T \) and \( E \) exhibits oscillations at twice the frequency of either the \( TT \) or \( EE \) auto-correlation.

Vector perturbations are vortical motions in matter caused by cosmological defects. Vorticity is damped as the universe expands and, consequently, vector
Figure 1.10: Scalar perturbations resulting in polarization for a single Fourier mode. The wavevector (indicated in green) is oriented along the same direction as the density perturbations, and all scattering occurs at a $90^\circ$ angle to this direction. Consequently, only $E$-mode polarization results. Figure taken from [18].

Tensor perturbations are gravitational waves. A gravitational plane wave induces a quadrupolar moment by stretching or shearing space time in the plane of the perturbation. As the wave propagates through the crests and troughs at the last scattering surface, a circle of test particles at these locations becomes stretched and distorted along opposite axes, as shown in Figure 1.11. The main angular variation is perpendicular to the wavevector, resulting in a polarization with both $E$ and $B$ components in nearly equal amounts.

Tensor perturbations are uniquely predicted by inflationary theory and are the only source of $B$-mode polarization at large angular scales. Gravitational waves created during inflation are immediately stretched beyond the causal horizon and become “frozen” in time. When a tensor perturbation re-enters the causal horizon it is damped in amplitude by the expansion of the universe and decays rapidly. Consequently, the $B$-mode signal should peak at length scales corresponding to modes entering the horizon at recombination and should be suppressed at all other length scales.

Figure 1.12 shows the temperature and $E$-mode cross correlation spectrum as well as the $E$-mode and $B$-mode autocorrelation spectra. Both scalar and tensor perturbations contribute to the features in the temperature and $E$-mode
**Figure 1.11**: Illustration of $E$ and $B$-mode polarization resulting from a single Fourier mode of a gravitational wave. In the top panel the polarization orientation is aligned with the gravitational wave and hence temperature gradient, giving $E$-mode polarization. In the bottom panel the polarization orientation is at $45^\circ$ to the gravitational wave, giving $B$-mode polarization. Figure taken from [18].
Figure 1.12: From top to bottom: temperature and $E$-mode cross correlation power spectrum, $E$-mode power spectrum, and $B$-mode power spectrum. The data points shown are experimental data and the solid lines represent the standard cosmological model. The $B$-mode power spectrum is simulated with a value of $r = 0.1$. Figure taken from [19].
power spectra. Because they are generated by the same mechanism, the \(E\)-mode structure is well predicted by the standard cosmological model generated by fitting the temperature data. And, although the \(E\)-mode stands as confirmation that the working model is correct and has imposed tighter constraints on cosmological parameters, very little new science has been gleaned from its measurements.

Although tensor perturbations contribute to both the temperature and \(E\)-mode spectra, the contribution is orders of magnitude smaller than that of scalar perturbations and is therefore degenerate with other model parameters. The as-yet undetected primordial \(B\)-mode signal, by comparison, is \textit{uniquely} determined by tensor perturbations at large angular scales. A measurement of the \(B\)-mode signal would be monumental. It would stand as the first unique and non-degenerate evidence of inflationary theory, earning it the moniker the “smoking gun” of inflation.

A measurement of the \(B\)-mode would not only confirm that inflation took place, but also serve as a direct probe of the energy scale of the tensor perturbations and inflation itself. A wide range of inflationary models have been postulated, many of which would generate measurable \(B\)-mode signals. The strength of the \(B\)-mode signal is generally parameterized by the tensor-to-scalar ratio \(r\), which is linked to the energy scale of inflation by

\[
E \approx 10^{16} \text{GeV} \left( \frac{r}{0.01} \right)^{1/4}
\]

where \(E\) is the energy scale of inflation. Characterizing the \(B\)-mode probes energies at the GUT scale, one trillion times higher than earthbound particle accelerators. Many models of inflation predict a value of \(r\) that would be measurable in the \(B\)-mode spectrum, so a detection of primordial \(B\) modes would greatly limit the number of viable models of inflation.

**Gravitational Lensing**

While tensor and scalar perturbations generate primordial CMB signal, gravitational lensing alters CMB structure through non-primordial mechanisms. Intervening structure affects the path of CMB photons via gravitational interaction, distorting the primordial CMB signal. The temperature power spectrum
suggests that intervening structure with a potential well high enough to alter CMB structure has a comoving diameter of \( \approx 150 \) megaparsecs, and that this potential will deflect CMB photons by an angle of \( 1.2 \times 10^{-4} \) radians. The distance to the last scattering surface is approximately 14,000 megaparsecs, meaning CMB photons are deflected approximately 100 times on their random walk to earth. This gives an average deflection angle of \( \sqrt{100} \times 1.2 \times 10^{-4} \) radians \( \approx 3 \) arcminutes \((l \approx 2000)\). At such angular scales, the lensing signal dominates the primordial CMB signal.

Although the lensing signal peaks at small angular scales, the signal itself is correlated on scales corresponding to half the last scattering distance, approximately \( \approx 1.2^\circ \). The effect distorts the CMB temperature power spectrum by widening the peaks at the percent level.

Gravitational lensing has a much more dramatic effect on CMB polarization than on temperature. Gravitational potentials along the line of sight shear polarization, mixing the stronger \( E \)-mode signal into \( B \)-modes (Figure 1.13(a)). For relatively small values of \( r \), the gravitational lensing signal will be dominant to the primordial \( B \)-modes. Fortunately, the two signals peak at different angular scales, making it possible to separate and fit the lensing signal. A template can then be subtracted from the \( B \)-modes and the gravitational wave signal can be revealed. Figure 1.13(b) shows the lensed \( E \) mode power spectra, the primordial \( B \)-mode power spectra for two values of \( r \), and the lensed \( B \)-mode power spectra. For the smaller value of \( r \), it is obvious that the lensed signal dominates at the majority of angular scales. Note that because the \( B \)-mode lensing signal results from the original \( E \)-mode signal, the peaks of the two are correlated at the same angular scales.

Although the lensing \( B \)-modes act as a contaminant for inflationary studies, they are rich in interesting physics of their own. The \( B \)-mode lensing signal is influenced by the sum of the neutrino masses, because more massive neutrinos suppress the formation of large-scale structure and lower the amplitude of the lensing signal. Because the CMB is the oldest electromagnetic relic on the universe, it is the oldest possible tracer of lensing and large scale structure. By combining
Figure 1.13: (a) The effects of gravitational lensing on temperature, $E$-modes, and $B$-modes. Lensing blurs the temperature signal on all scales and lenses $E$-modes into $B$-modes as shown. Figure taken from [20]. (b) Theoretical $E$ and $B$-mode power spectra with shaded 95% confidence intervals if $r < 0.1$, based on a compilation of 2005 CMB and large scale structure data. For low $r$, the lensing signal dominates the tensor perturbations. Figure taken from [21].
CMB lensing with optical and other lensing studies one can probe the evolution of structure and the equation of state of dark energy over time. Finally, $B$-mode lensing is a “clean” signal. Gravitational lensing of temperature is just one of many factors that accumulate into the observed temperature power spectrum. Conversely, $B$-mode lensing is the only effect that contributes to the $B$-mode signal at small angular scales; it is completely non-degenerate.

### 1.3 CMB Measurements

One question remains: If $B$-mode science encodes this much cosmological information, why does it remain undetected?

#### 1.3.1 Difficulty of Measurement

Detection of the sought-after $B$-mode signal presents many difficulties. First, the inflationary $B$-mode amplitude is extremely small in comparison to the sky temperature, the CMB’s temperature, and even CMB temperature and $E$-mode anisotropy. Currently, the most restrictive limits, derived from temperature anisotropy, place an upper limit on the expected $B$-mode signal of less than 100 nK [22, 23, 6].

Second, polarized emission from galactic and extra-galactic objects acts as a foreground contaminant for CMB polarization measurements. Even in the case of optimistic predictions for the signal amplitude (i.e. inflationary models with higher values of $r$), models suggest that foreground contamination could be an order of magnitude larger than the $B$-mode signal at the largest angular scales [24]. Further, the three main contributing mechanisms that generate galactic polarization (free-free, synchrotron, and dust emission) vary independently as a function of frequency, making the removal of galactic foregrounds a complex issue.

Finally, in the absence of systematic effects, any CMB measurement is ultimately limited by two factors: cosmic variance and instrumental sensitivity. Cosmic variance stems from the fact that, from our fixed vantage point in time and space, can only observe one realization of all the possible observable universes.
In the case of the CMB, we are limited by the fact that only one observable spectrum is available from earth. The total uncertainty, $\Delta C_{l}^{TT}$, can be written as

$$\Delta C_{l}^{TT} = \sqrt{\frac{2}{(2l + 1)f_{\text{sky}}}}(C_{l}^{TT} + w_{T}^{-1}W_{l}^{-1})$$

(1.31)

where the first term in parenthesis is the cosmic variance and the second is instrumental sensitivity [25]. The variable $w_{T}$ is the temperature sensitivity weight function and is given by

$$w_{T}^{-1} = \frac{4\pi f_{\text{sky}} s^2}{t_{\text{obs}}}$$

(1.32)

and

$$W_{l}^{-1} = e^{-l(l+1)\sigma_{b}^2}.$$ 

(1.33)

In these equations, $t_{\text{obs}}$ is the observing time, $s$ is the total array sensitivity, $f_{\text{sky}}$ is the fraction of total sky covered, and $\sigma_{b}$ is the instrument beam size. The obvious obstacle for ground based telescopes is that $f_{\text{sky}}$ is limited when observing from a fixed location on earth, which imposes a minimum on the noise contribution due to cosmic variance. Obviously then, in the quest to detect the $B$-mode signal, any ground-based telescope must be designed such that the instrumental noise term is as small as possible. Given that the amplitude of the $B$-mode signal is unknown, detector arrays with unprecedented sensitivity (able to detect $\mu$K temperature fluctuations each second) and careful optimization of integration time must be used.

1.3.2 State of the Field

The past few decades have seen remarkable progress in characterizing the CMB. In terms of a broad timescale, the progress in CMB measurements has been almost unbelievable. The existence of the CMB was not even postulated until 1948 by Alpher and Herman [26], who predicted that a uniform radiation would permeate all of space with a temperature of approximately 5K. Some 17 years later, Penzias and Wilson scooped all other groups in hot pursuit of the CMB signal (and by accident!) [5], measuring the CMB monopole from a radio telescope at Bell Laboratories. Although many groups toiled to measure the predicted CMB
blackbody, ultimately it was the COBE satellite in 1990 that definitely characterized it using the FIRAS instrument [8]. Two years later COBE also detected the CMB temperature anisotropy using the onboard DMR instrument [27]. Although 27 years seems like a relatively long timescale, the sensitivity required to measure temperature anisotropy is several orders of magnitude higher than measuring the monopole moment. CMB anisotropy is on the order of 0.001% of the CMB’s temperature, which illustrates how much instrumental sensitivity increased in this short period of time.

The detection of CMB anisotropy set forth a true detector arms race, spurring relatively small university-led research efforts in balloon-borne and earth-bound CMB telescopes. The first acoustic peak of the CMB was measured in 2000 by two balloon-borne experiments: MAXIMA [12] and BOOMERanG [11]. In the years since, the temperature anisotropy of the CMB has been mapped to exquisite precision by many experiments [28, 29, 30, 31, 32, 33, 34], and most recently the standard has been set by the newly released satellite Planck data set [15]. As a result of this data, tight constraints have been derived for the standard $\Lambda$CDM cosmological model [35, 36, 37, 22, 23, 6]. Satellites and ground based experiments have mapped the acoustic peaks to very high precision all the way out through the CMB damping tail, and have characterized the temperature lensing [38, 39, 40] and Sunyaev-Zeldovich effect [41, 42, 43, 44, 45].

CMB polarization anisotropy, which is orders of magnitude smaller than CMB temperature anisotropy, has not yet been characterized to the same precision. The $E$-mode polarization was first detected by DASI in 2002 [46] and since the first detection, multiple experiments have measured the $E$-mode and temperature cross $E$-mode power spectrum [47, 48, 49, 50, 51, 52, 14, 53, 19, 54] as shown in Figure 1.12. Both the inflationary $B$-mode signal and the $B$-mode lensing signal are undetected, with an upper bound from $B$-mode studies alone of $r < 0.72$ [19]. The recent results from the Planck experiment give an upper limit of $r < 0.11$ (at 95% confidence) using measurements of temperature and other independent astrophysical datasets [6].
1.3.3 Future CMB Studies

To detect the $B$-mode polarization signals (from both lensing and from inflation) requires sensitive detectors, high angular resolution, and superb control over systematic effects among other considerations. The remainder of this thesis describes my graduate work developing and analyzing the use of new and novel CMB measurement technology. Chapter 2 is an expanded version of my Astrophysical Journal paper describing the design and use of Faraday Rotation Modulators in the BICEP telescope. Chapter 3 describes my work on the ground based CMB telescope POLARBEAR. The introduction for this chapter is taken largely from a conference proceeding that I authored for the American Physical Society Division of Particles and Fields meeting in 2011. Chapter 4 describes my work at UCSD developing novel CMB technology: superconducting tunnel junction refrigerators and superconducting transition edge sensor bolometers. Chapter 5 is a reprint of a paper submitted to Superconductor Science and Technology.
2.1 Motivation

As mentioned in the previous chapter, the detection of the sought-after $B$-mode signal presents many difficulties. Exquisite control of systematic and instrumental effects, down to the tens of nK level, will be required before a detection of $B$-modes can be claimed with confidence. One approach to mitigating systematic errors is to modulate incoming polarization, thereby shifting the signal to higher frequencies and away from lower frequency systematic contaminants. Rapid modulation eliminates concerns about time varying thermal, optical, or electrical fluctuations that often change on much longer time scales. In essence, modulation speed can be used to tune the signal band of the instrument, otherwise set by a combination of scan speed and beam size, placing the signal away from microphonic lines and low frequency $1/f$ noise. The lack of limitation on scan speed expands the parameter space for observation strategies. Modulation also mitigates polarized systematic effects that are introduced by optical elements between the modulator and the detector.

There are additional benefits to polarization modulation. The reconstruction of the Stokes $Q$ (or $U$) parameter requires two independent measurements
at different detector angles, typically done by two detectors. Mismatched detector properties, such as those caused by differential spectral response, differential pointing, and mismatched transfer functions, can result in spurious polarization. Modulation allows a single linearly polarized detector to act as an independent polarimeter, removing systematics resulting from combining mismatched detectors.

Many CMB polarimeters already employ mechanisms to modulate the incident radiation field about the optical axis of the instrument. The modulator effectively exchanges electric fields between two detectors (or a single detector in two orientations) and decomposes the radiation into two orthogonal bases. Traditional methods for modulation in millimeter wave polarimetry include physical rotation of the entire instrument about the optical axis [55, 56, 57, 58, 59, 60, 61], rotation or translation of a wiregrid polarizer [62, 63] or birefringent half-waveplate [64, 65, 66, 67, 68, 69, 70, 71], or parallactic angle rotation ("sky rotation") with respect to the fixed instrument coordinate system.

These classical modes of polarization modulation often employ bulky and complex mechanisms, where any mechanical failure would result in the complete loss of polarization modulation. Faraday Rotation Modulators (FRMs) are coupled to individual detectors, eliminating the single point failure mode and relieving the need for producing large grids and waveplates that are often limited to tens of centimeters in diameter. FRMs can be tuned individually for specific bandwidths, whereas birefringent crystal waveplates themselves are not broadband and require specialized anti-reflection coatings that are difficult to optimize for multiple frequencies.

In addition, polarization modulation via mechanical motion is limited to a low frequency range due to tolerances on vibrational and microphonic noise. Specifically, half-waveplate rotation is limited to modulation frequencies of tens of Hertz (Hz) at most and boresight rotation is limited to much less still. Parallactic angle rotation is also a slow signal modulation and is location dependent; at South Pole observatories, such as BICEP, parallactic angle modulation caused by sky rotation is simply not present.

In this chapter I describe the Faraday Rotation Modulator, a solid-state
polarization switch that employs the Faraday effect to rapidly modulate polarized light at cryogenic temperatures. The FRM is compact, works over a large frequency range, and is capable of modulation rates up to 10 kHz. Specifically, I describe the laboratory testing and Galactic observations of the FRMs in the BICEP experiment during the 2006 observing season.

2.2 FRM Design for Use in BICEP

Figure 2.1 shows a cross-sectional schematic of a FRM. Along the path of incident radiation are two alumina cones attached to either side of a ferrite cylinder, providing an impedance match to minimize reflection loss. The ferrite and cones form a “toothpick” assembly. The input and output waveguides preserve polarization orientation. The ferrite itself is surrounded by a superconducting solenoid and the entire toothpick assembly acts as a dielectric waveguide, allowing hybrid electric modes to propagate. Details of the design and construction of FRMs can be found in [72].

Polarization rotation takes place only within the cylindrical ferrite. Bias currents driven through the solenoid generate a longitudinal magnetic field within the FRM, rotating the polarized input by an angle

\[ \theta = V l B, \]

where \( V \) is the Verdet constant, a parameter describing the intrinsic polarization properties of the ferrite material, \( l \) is the length of the ferrite cylinder, and \( B \) is the strength of the applied magnetic field. Equation 2.1 is known as the second-order magneto-optical Faraday effect.

The FRMs were optimized for use in BICEP, a ground-based CMB polarimeter designed to measure the B-mode polarization of the CMB at degree angular scales. BICEP is an on-axis refracting telescope with a 0.25 m aperture capable of scanning in azimuth and elevation and rotating about the boresight. Full details of the BICEP instrument can be found in [58] and [73]. Here we provide only a brief summary.
Figure 2.1: Cross-Sectional schematic of a FRM. Polarized radiation traverses the FRM from right to left. The alumina cones serve as an impedance match and are attached to the ferrite in the center. The ferrite is supported by dielectric washers (shown in pink). A superconducting solenoid surrounds the ferrite and a corrugated waveguide surrounds the cones.

BICEP’s focal plane consists of 49 pairs of orthogonally oriented polarization sensitive bolometers (PSBs) [74] cooled to 0.25 K. During the 2006 observing season, 25 PSB pairs were tuned for a band center of 100 GHz (beam size 0.93 degrees) and the other 24 for 150 GHz (beam size 0.60 degrees). Each pixel is individually coupled to a stack of three corrugated microwave feed horns cooled to 4 K. The upper edge of the frequency pass band is defined by a series of metal mesh resonant filters placed in front of the feed horn. The lower edge is defined by the waveguide cut-off imposed by the horn itself.

The FRMs were present in the optical path of six PSB pairs around the perimeter of the focal plane: three at each of 100 GHz and 150 GHz. Figure 2.2 shows the location of the FRMs in the focal plane of BICEP along with a cross section of the FRM positioning in the optical path. For the remainder of this chapter the FRMs at 100 GHz will be referred to as 100A, 100B, and 100C and at 150 GHz as 150A, 150B, and 150C as labeled in Figure 2.2(a).
Figure 2.2: (a) The BICEP focal plane for the 2006 observing season. Each circle represents a pixel with a pair of orthogonal bolometers with polarization sensitivity axes depicted by +. The larger blue circles represent 100 GHz pairs and the smaller red circles represent 150 GHz pairs. The dashed line represents the boresight reference angle of zero. The circles that are shaded in gray in boxes around the perimeter feature FRMs. (b) The relative location of FRMs within the optical path. The FRMs are housed at 4 K between the feedhorn and bolometer enclosure. The bolometer signal is demodulated using lock-in amplification with phase referenced to the driving current.
2.3 Instrument Characterization

The FRMs were subjected to a rigorous series of lab and field tests to characterize their behavior and choose optimal operating parameters for BICEP observations. The following section describes the FRM laboratory testing and results.

2.3.1 Rotation Angle and Bias Signal

FRMs modulate mm-wave signals by rotating the direction of polarized radiation relative to the axis of polarization of the corresponding bolometers. To mitigate systematic offsets, this relative rotation must be well-calibrated and time-invariant. Measurements of the rotation angle as a function of bias current were made by placing an aperture-filling wire grid polarizer (resulting in a 100% linearly polarized signal) at the telescope window, biasing the FRM with a triangle wave signal, and measuring the voltage response of the PSB pair beneath the rotator. An example output of this technique is given in Figure 2.3.

As the FRM bias current is increased and decreased, the bolometer voltage responds with a sinusoidal wave form that “flattens” at the turnarounds when the ferrite begins to saturate magnetically. From this data, the relative rotation angle of the FRM, $\theta$, can be calculated by

$$
\theta = \frac{1}{2} \sin^{-1} \left( \frac{V - (V_{\text{max}} + V_{\text{min}})/2}{(V_{\text{max}} - V_{\text{min}})/2} \right)
$$

where $V$ is the bolometer response voltage and $V_{\text{max}}$ and $V_{\text{min}}$ are the maximum and minimum of that voltage, respectively. Figure 2.3 shows the voltage response curve as a function of $\theta$. FRMs in BICEP can continuously rotate a polarized signal through a range of approximately $\pm 80^\circ$. Maximum rotation corresponds to a bias current of approximately 300 mA.

The curves shown in red in Figure 2.3 are for increasing bias whereas the curves shown in blue are for decreasing bias. The two curves do not overlap due to magnetic hysteresis, where a change in magnetization will lag with respect to the externally applied field due to configurational internal forces. Magnetic
Figure 2.3: Raw detector response of a PSB pair and calculated rotation angle of the FRM shown for a single FRM calibration run. The red and blue curves correspond to increasing and decreasing FRM bias current respectively. FRM hysteresis can clearly be seen in both plots. The hysteresis curve separation at zero bias current was between 10° and 25°.

Hysteresis prevents assigning a one-to-one correspondence between current and rotation angle, creating a degeneracy in the bias angle for values of the bias current where hysteresis is present. Additionally, the degree of hysteresis was shown also to be a function of bias frequency, further increasing the complexity of demodulation. To avoid the complications of magnetic hysteresis the FRMs were biased with a square wave signal. This limited the FRM rotation angle to only two states, which could be uniquely determined by the magnitude and sign of the bias current.

The amplitude of the bias signal was chosen such that the FRMs would have angular modulation of 45 degrees, corresponding to a bias current of approximately ±125 mA. In the reference frame of the bolometer, this is equivalent to switching between Stokes +Q and −Q, effectively transforming a single PSB into two orthogonal detectors. The square wave bias was tuned for a frequency of 1 Hz, slightly higher than the 1/f knee for the BICEP temperature data, generally 0.5-1 Hz for a single pixel [73]. During unmodulated CMB and Galactic half-scans, BICEP shifts the 1/f knee by scanning the telescope at a scan speed of 2.8°/s.
Although the FRMs are capable of switching at frequencies orders of magnitude higher, the FRMs in BICEP were limited by the bolometer time constants, which had a median of 21 ms across the focal plane. A 1 Hz bias was found to be the best compromise between maximizing integration time per cycle, which lowers the frequency but decreases the noise in demodulation (Section 2.4.2 in the sub-section “Demodulation”), and minimizing the $1/f$ atmospheric fluctuations.

### 2.3.2 Instrumental Polarization

Instrumental polarization (IP) is spurious polarization detected when only unpolarized light is observed. IP can result from the use of FRMs if the FRM toothpick assembly is tilted relative to the optical axis of light to the detector. The IP values of the FRMs were measured by placing unpolarized aperture-filling 300 K and 77 K sources at the cryostat window and biasing the FRMs through their full range of rotation angles. The resultant fractional IP is defined as

$$IP = \frac{1}{2} \left( \frac{V_{AC}(300K)}{V_{DC}(300K)} - \frac{V_{DC}(77K)}{300K} \right) \left( \frac{300K - 77K}{300K} \right)$$

where $V_{DC}(300K)$ is the mean bolometer voltage at 300 K, $V_{DC}(77K)$ is the mean at 77 K, and $V_{AC}(300K)$ is the peak-to-peak bolometer voltage at 300 K.

Individual FRM pixels were found to have a repeatable and time-invariant IP signal, with a standard deviation of approximately 6% (and $\pm$ 10%) over the testing period of several months. The scatter of the mean IP between FRM assemblies, however, varied considerably. The average value of all FRM assemblies was found to be 0.59% with a standard deviation of 0.43%. The maximum value among all the FRM pixels was found to be less than 1.2%, for FRM150A.

### 2.3.3 Rotation Angle Calibration

Two rotation angle calibration runs were performed with the BICEP FRMs. A dielectric sheet calibrator (DSC) was used based on the design from POLAR [75, 76]. The DSC consisted of an eccosorb-lined cylinder with a polypropylene sheet mounted at 45° that served as a beam splitter, creating a small polarized signal
Figure 2.4: Dielectric sheet calibration for FRM150A. The top panel shows the 1 Hz FRM modulated signal. The bottom panel shows the split demodulated timestreams at ±45 degrees.

from input ambient temperature and sky loads. The direction of the polarized signal was determined by the angle of the DSC relative to the detector. The DSC was placed above BICEP’s optical window, and as the boresight was rotated each PSB exhibited a sinusoid-like response that varied as a function of radial position in the focal plane. Model fits using the known properties of the DSC yielded detector polarization angles for each PSB in the focal plane.

Figure 2.4 shows the timestream response of a BICEP PSB to DSC calibration with FRM modulation. To extract polarization angle, the FRM output was demodulated into two separate timestreams as described in Section 2.4.2 in the sub-section “Demodulation”. The result is shown in the bottom panel of Figure 4. Each timestream was fit to the DSC model for detector angle. Given a nominal PSB orientation angle $\psi$, derived from DSC tests in the absence of modulation, the two demodulated timestreams with proper bias should yield detector angles $\psi \pm 45^\circ$.

DSC calibration showed proper bias rotation for four of the six FRMs. One
Table 2.1: DSC calibration results for all six PSB pairs. The table lists the total rotation modulation of the polarization angle. For perfect $\pm45^\circ$ polarization modulation, $90^\circ$ is expected. Polarization modulation in the plus and minus direction is one-half of the listed value. The standard deviation is calculated across each individual run and averaged over the two different rotation angle calibrations.

<table>
<thead>
<tr>
<th>FRM PSB</th>
<th>Avg Rotation Range ± Standard Deviation (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100A</td>
<td>85.5 ± 7.3</td>
</tr>
<tr>
<td>100B</td>
<td>41.4 ± 12.5</td>
</tr>
<tr>
<td>100C</td>
<td>79.7 ± 7.4</td>
</tr>
<tr>
<td>150A</td>
<td>90.2 ± 0.1</td>
</tr>
<tr>
<td>150B</td>
<td>88.2 ± 4.5</td>
</tr>
<tr>
<td>150C</td>
<td>73.8 ± 18.8</td>
</tr>
</tbody>
</table>

of the rotator biases was incapable of supplying the currents necessary for full $\pm45^\circ$ rotation (FRM 100B) and another was deemed unreliable based on the high standard deviation in its rotation angle (FRM 150C). In addition, the PSB pair associated with FRM 100C had excess noise in both the modulated and demodulated timestreams, rendering it unusable for celestial polarization studies.

The results of the DSC calibration are summarized in Table 2.1 for all six PSB pairs. The standard deviations listed in the table reflect a true deviation of the rotation range during bias periods and are not limited by measurement readout accuracy.

2.4 Galactic Observations

2.4.1 Observing Strategy

In 2006, five observations of the Galactic plane were made with the FRMs biased. The target fields are shown in Figure 2.5. The first of these observations, referred to as the “shallow” observation, is defined by the celestial coordinate system $(\alpha, \delta) = ($Right Ascension, Declination$)$ as $180^\circ < \alpha < 290^\circ$ and $-70^\circ < \delta < -45^\circ$ (Figure 2.5). The shallow FRM observation was used mainly for diagnostic
Table 2.2: Details of each FRM observation, ordered sequentially. Each column describes a single observation with date specified in the format MM/DD/YY. Each observation was segmented into 9-hr sections described by the No. Sections row. For each of these sections, the rotation angle of the boresight is described. The next two rows describe the range of elevation and step size at the boresight center. Azimuth parameters include the length of all azimuthal half-scans, the scan speed, and the number of uni-directional half-scans performed at each step in elevation.

<table>
<thead>
<tr>
<th>Date(s)</th>
<th>4/21/06</th>
<th>10/14/06</th>
<th>10/18/06</th>
<th>10/28/06</th>
<th>10/30/06</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sections</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Boresight (°)</td>
<td>{315, 315, 180, 180, 0}</td>
<td>{154, 164, -152, -142}</td>
<td>{154, 164, -152, -142}</td>
<td>{120, -150, -150, 120}</td>
<td>{120, -150, -150, 120}</td>
</tr>
<tr>
<td>El (°)</td>
<td>55.5–60.5</td>
<td>55.5–61.2</td>
<td>55.5–61.2</td>
<td>55.5–61.2</td>
<td>55.5–61.2</td>
</tr>
<tr>
<td>El Step (°)</td>
<td>0.25</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
<tr>
<td>Az Range (°)</td>
<td>≈77</td>
<td>≈12</td>
<td>≈12</td>
<td>≈12</td>
<td>≈12</td>
</tr>
<tr>
<td>Speed (°/s)</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>Half-Scans</td>
<td>4</td>
<td>22</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
</tbody>
</table>

purposes; this was the first test of FRM modulation on a celestial source.

All deep observations focused on a small portion of the Galactic plane 238° < α < 248° and −53° < δ < −50°. The small area was necessary to maximize integration time, as the goal of the deep observations was a detection of polarization from a celestial source. Due to their respective positions in the BICEP focal plane, no two FRMs could scan this same region of sky simultaneously, so integration time was split evenly between two FRMs: FRM100A and FRM150A. FRM100A was chosen because it had the highest optical efficiency, lowest IP, and a reliable current bias signal. FRM150A was chosen for its extremely consistent bias signal during DSC calibration and its close proximity to FRM100A in the focal plane. An integration time of approximately 72 hours on this deep patch was achieved for each FRM used for analysis.

For all FRM observations, azimuth-elevation raster scans were used. At each step in elevation, the telescope performed a number of unidirectional scans in azimuth (known as “half-scans”) back and forth across the target patch with a scan
Figure 2.5: 100 GHz FDS Model 8 Galactic dust emission prediction [1] plotted in celestial coordinates on a logarithmic scale. The outlined regions show the areas of integration for the shallow and deep FRM observations.

speed 0.25°/s. The elevation step size and number of half-scans at each step were altered between the shallow and deep observations. These values are summarized in Table 2.2 along with the other pertinent features of each observation. During the deep observations, a single boresight rotation angle corresponds to a single FRM pixel focused on the Galactic region of interest.

2.4.2 Timestream Processing

When characterized by Stokes $T$, $Q$, and $U$ parameters, the time domain response of an unmodulated BICEP PSB as a function of frequency $\nu$ and direction $\Omega$ is modeled by

$$
d(t) = K_t \otimes \left\{ n(t) + g \int d\nu A_e F(\nu) \int d\Omega P(\Omega) \right\} \left[ T + \frac{1 - \epsilon}{1 + \epsilon} \left( Q(\cos(2\psi)) + U \sin(2\psi) \right) \right],
$$

where $\psi$ is the polarization orientation angle of the PSB, the parameter $\epsilon$ quantifies the amount of cross polar leakage of the PSB, $P(\Omega)$ is the beam response function,
$F(\nu)$ is the spectral response, $A_e$ is the effective antenna area, $g$ is the gain factor that converts voltage to temperature, $n(t)$ is a term describing total atmospheric and detector noise, and $K_t$ is the time domain impulse response of the detector. A complete description of the measurements and experimental procedures used to obtain these terms is given in [73].

**Preliminary Processing**

Data processing begins by deconvolving the complete, measured transfer functions for all FRM timestreams. Deconvolution plays a necessary and important role in FRM analysis. Without deconvolution, the non-instantaneous response of the bolometers results in the perfect square-wave bias signal transforming into a rounded square wave. This rounding affects between forty to fifty percent of the useable timestream depending on the bolometer time constant. This effect can be seen in the left half of Figure 2.6. The bolometer bias signal is plotted as a reference at the top, and the non-instantaneous switching in response is shown at below.

For BICEP science data, both CMB and Galactic, low pass filtering is applied simultaneously to the deconvolution by enforcing a frequency cut-off in Fourier space as the deconvolution is performed. However, for the FRMs, a low-pass cutoff eliminates the higher order harmonics necessary to reconstruct a perfect square wave signal, and thus would defeat the purpose of deconvolving. As such, the deconvolution is performed with a low frequency cut-off of 1 Hz, the frequency of the square wave bias, and a high frequency cut-off of 25 Hz, which is half the sample frequency and used to avoid ringing. The right half of Figure 2.6 shows the effects of deconvolution. Many points are recovered along each bias switching and the square-wave behavior has been restored.

After deconvolution, the beginning and end of each half-scan is identified and periods of non-zero telescope mount acceleration are trimmed from the end points. Horizon and celestial boresight coordinates are calculated by applying a pointing model to the raw boresight pointing using the focal plane coordinates of the pixel of interest.
Figure 2.6: Figure showing the effects of deconvolution on the modulated timestreams. The bottom panel on the left shows the raw timestream and the bottom panel on the right shows the timestream after being deconvolved by the measured transfer function of the corresponding bolometer. The rounding of the bolometer response can be clearly seen in the pre-deconvolution timestream, and has been removed post-deconvolution. On each half, the FRM bias signal is shown as a switching reference.
Relative Gains

Relative detector gains are applied by fitting to the elevation nods, a period at the beginning of each fixed elevation scan when the telescope performs a rounded triangle wave motion with an elevation amplitude of 1.2 degrees and a duration of approximately 27 seconds. During this time, for consistency, the FRMs are modulating the sky signal. The responsivity factor for each PSB is found by fitting the detector voltage versus the changing line of atmospheric sight, given by the cosecant of elevation multiplied by the scale height. For the shallow observation, the relative gains were derived in two ways: via the modulated and the demodulated timestreams (as described in the next section). The two values were found to agree to within 0.1%. This makes intuitive sense, as the relative gains are derived through a linear least-squares fit which would split the different between any modulation of the sky signal. Therefore, demodulation became unnecessary for the relative gain calculation, and the relative gains for the deep integration observations were derived from the raw, modulated timestreams. Before applying the relative gains, each PSB gain factor was weighted by the average of the PSB pair during the scan set.

Demodulation

Although many demodulation techniques were considered, time domain demodulation was used due to its simple and robust properties and accurate polarization reconstruction. Figure 2.8 is a graphical example of the timestream demodulation using FRM 150A.

The first step of demodulation was to clean the bias current signals of software spikes and glitches via nearest-neighbor interpolation. Bias points that were more than three standard deviations from the total bias signal were identified and a weighted average of the four surrounding points, two on either side, were used to define the value of the bias signal at this point. After removing the software glitches, the current signal displayed fluctuations on the < 1% level, and so perfect ±45° rotation was assumed for positive and negative values of the bias, respectively. Using this assumption, and to ease computation, positive values of the bias were
reassigned to the value one, and negative values of the bias were assigned to the value negative one. This process is plotted in Figure 2.7.

The current biases were then used as the lock-in phase reference signals for demodulation of the corresponding bolometer timestreams. All points at which the current bias transitioned across zero were identified and the median number of points between transitions was computed. Any region where the switching period differed from the median by more than 10%, along with one plateau region on either side, was identified and the corresponding region in the bolometer timestream was excluded from analysis. This resulted in a loss of < 1% of the usable timestream data as the bias signal exhibited very few deviations from normal periodicity.

The midway point between each two adjacent transitions in the bias signal was identified, dividing each “bias plateau” into two halves. The cleaned bias signal and its segmentation are shown in the top panel of Figure 2.8. The average value of the corresponding bolometer timestream, along with other pertinent variables (pointing, time, etc.), for each half was computed, less the two to four points closest to the transition to account for the non-instantaneous bolometer response. Although deconvolution using the measured time constant does recover a small number of points along each transition – without deconvolution, a minimum of six points were cut on each side – the bolometer response to bias switching is still imperfect and a fraction of the data surrounding each transition must still be removed from analysis. The number of points cut from each side of the transition was derived independently for each bolometer using iterative fits to the elevation nods. To maximize the signal-to-noise, the minimum number of data points was cut that yielded a convergent fit to the elevation nod signal.

The bolometer timestream averaged values were then separated based on the sign of the corresponding bias signal. Essentially, the timestream of a single bolometer at polarization angle $\psi_{ref}$ has been demodulated and downsampled into two demodulated bolometer timestreams at $\psi_{ref}+45^\circ$ and $\psi_{ref}-45^\circ$. This is shown in the last panel of Figure 2.8. The sum and difference of the two demodulated timestreams from each individual PSB was then calculated.

The demodulation scheme was put through many simulated timestream
Figure 2.7: Plots of the bias cleaning process. The top left panel shows the raw bias timestream, which still exhibits software spikes and glitches. The bottom left panel has the glitches removed via nearest neighbor interpolation. The top right panel shows a zoom-in of the first 1000 bias points that are now deglitched but still exhibit small, less than 1% fluctuations in bias magnitude. The bottom right panel shows the transformed perfect bias signal, where the fluctuations have been removed and all values of the bias have been transformed to one or zero. To finish processing, one-half is subtracted from the entire bias timestream and all values are multiplied by two, giving the bias a switching magnitude of one and negative one.
Figure 2.8: Illustration of time domain demodulation. The top plot shows the bias current as a function of time. Points in pink are positively biased, corresponding to a $+45^\circ$ rotation. Points in blue are negatively biased, corresponding to $-45^\circ$ rotation. The black crosses are the midpoints of each bias plateau. The middle plot shows the corresponding points mapped onto the bolometer timestream. The red crosses are points of gradual switching resultant from bolometer time constants, and are cut from analysis. The remaining points are averaged together for each half-plateau and are stored as two new demodulated timestreams as shown in the lowest panel. The pink filled circles comprise the demodulated $+45^\circ$ timestream, whereas the filled blue circles correspond to $-45^\circ$ timestreams.
tests to verify the accuracy. The first and simplest test was to confirm that a square wave bias could be demodulated in the presence of a white noise signal. The bias was designed arbitrarily for this test such that each plateau had 34 points and the noise signal was generated randomly using a Gaussian distribution with a mean of zero and a standard deviation of 0.2. The demodulation scheme was tested for accuracy by verifying that each demodulated point decreased the noise by a factor of the square root of the number of averaged samples. The standard deviation among demodulated points was found to be 0.035, which is approximately equal to the $0.2/\sqrt{34}$. To check that this was not just coincidental, three points from either side of the transition were cut from the demodulation analysis (also confirming that this could be done accurately) and the means were recomputed. As expected, the standard deviation among the demodulated points was found to be 0.045, approximately equal to $0.2/\sqrt{27}$.

Next, the demodulation scheme was tested for glitch finding and gap filling. Figure 2.9(a) shows an example of a white noise timestream with a modulation reference signal using 37 points along each plateau, but with a break from the normal periodicity of the signal in the middle. The top subplot is the fabricated data input and the bottom panel is the resultant demodulated timestream. The demodulation scheme effectively locates the regions where the periodicity of the bias is different from nominal behavior and these regions are blanked from analysis. The sum of the demodulated timestreams is labeled in the demodulated figure as “temperature” and the difference as “polarization”. Notice that both exhibit white noise of the same magnitude and that both effectively screen the improper bias section.

Finally, a more realistic signal input was generated which had three independent terms: white noise generated from a Gaussian distribution with standard deviation 0.2, a Lorenzian temperature signal with an amplitude of one, and a polarized sine wave with period of five times the bias and an amplitude of one. All of these are shown separately in the top subplot of Figure 2.9(b) and combined into one input timestream as the sum of the Lorenzian temperature signal, white noise, and the polarized signal multiplied by the bias amplitude. This process
mirrors true sky signal - unpolarized white noise and temperature are not modulated by the FRMs, whereas true polarized signal, in this case the sine wave, is modulated. The demodulation scheme showed an accurate reconstruction of the individual input signals, which can be seen in the bottom subplot of Figure 2.9(b). The temperature and white noise both contribute to the demodulated temperature timestream, which mirrors the input Lorenzian signal with an additive noise fluctuation term. The polarized sine wave was reconstructed as the demodulated polarization signal, showing that the demodulation scheme worked effectively for both temperature and polarization.

**Masking and Filtering**

Post-demodulation, a DC-offset was removed from both the sum and difference timestreams via mean-subtraction of each azimuthal half-scan. An obvious distortion arises when fits for the mean include the strong Galactic signal, so the Galaxy was masked from the fits. The top two panels of Figure 2.11 show an example of such distortion and the lack of distortion in the masked fit.

BICEP CMB half-scans are filtered using a third order polynomial [19] and BICEP Galactic half-scans utilize a second order polynomial [77] in order to accurately remove atmospheric $1/f$ noise. For the deep FRM observations, the length of the half-scans across the galaxy prohibited such a filtering scheme. A mask large enough to remove all Galactic signal from the fits resulted in the majority of half-scans beginning or terminating on the masked region. Interpolated fits that are not constrained by data on both sides of the mask tend to diverge and no longer reasonably approximate noise within the mask. Conversely, smaller masks avoid this problem but leave residual Galactic signal for the polynomial fit, causing higher order polynomial fits to remove true Galactic signal instead of only noise. These results are shown in the last two panels of Figure 2.11. While mean subtraction does not remove large-scale noise from the data, FRM modulation mitigates some of the $1/f$ atmospheric fluctuations that would otherwise need to be filtered.

The optimal mask cuts maximal Galactic signal from the fits while retaining
Figure 2.9: FRM demodulation test schemes. Panel (a) shows the demodulation test with a white noise signal and a break in the periodicity of the bias signal. Panel (b) shows a demodulation test with an input temperature and polarization signal combined with white noise. In both cases, the demodulation scheme accurately reconstructs temperature and polarization and screens for flagged areas.
Figure 2.10: Four panels showing masks of the Galactic plane region. The temperature is predicted by the 100 GHz FDS Model 8 Galactic dust emission prediction [1] plotted in celestial coordinates on a logarithmic scale. From left to right, top to bottom the panels show a 1°, 1.5°, 2°, and 2.5° radial Galactic mask. None of the masks completely screen intensity from the Galactic plane. As the mask size is increased the number of scans beginning or terminating on a masked region increases accordingly. At 2.5°, the majority of half-scans begin or terminate on the Galactic mask.

sufficient off-Galactic data to yield statistically significant fits for the mean. Given the Galactic coordinates defined by \((b, l) = (\text{latitude}, \text{longitude})\), the radius of the masked region was reduced from \(|b| < 3°\) (the maximal possible mask for the deep integration scans) until the DC offset fits converged for the majority of half-scans. The optimal mask was found to be \(|b| < 1.5°\).

Following DC offset subtraction, residual high-frequency power from demodulation was removed via a Butterworth low pass filter at 0.6 Hz. A data quality check was then performed where any half-scan with a signal spike more than 7 times the standard deviation was removed from analysis. These regions comprised < 1% of the total usable data.
Figure 2.11: Summed detector timestreams for FRM deep integration Galactic half-scans with different filtering methods. The top plot shows the timestream (blue), DC offset fit (red), and the resultant filtered timestream (black). Note the severe distortion from fitting the Galactic region. The second plot shows the same data with Galactic intensity removed from the fit. The boundaries of the mask are shown by black dotted vertical lines. The DC offset is subtracted from the masked region, leading to minimal artifacts. The third plot shows data that terminates on the masked Galactic region, fit with a higher-order polynomial; the interpolation diverges through the Galactic plane. The final plot illustrates the effects of subtracting a higher order polynomial fit from data with an insufficient mask to completely remove Galactic signal.
Instrumental Polarization Removal

Temperature to polarization leakage from a bright source such as the Galaxy can cause significant artificial structure in polarization. Laboratory testing revealed that FRM150A showed approximately 1% IP, the highest of any of the devices; a value that could distort Galactic polarization maps. To remove this temperature to polarization leakage, the IP value for each half-scan was calculated using a fit to the preceding elevation nod. The derived fractional IP value was then multiplied by the half-scan sum data and subtracted from the difference data.

Figure 2.12 shows an example of IP leakage derived from elevation nods. Because the atmosphere is unpolarized, the elevation nods should cause a change in the summed (but not differenced) data for each demodulated PSB. The data from each elevation nod for each PSB was demodulated, summed/differenced, mean subtracted, and low-pass filtered. The summed and differenced timestreams are both fit to the cosecant of elevation to derive average responses to the changing line of sight. Average IP values were computed by taking the mean of the absolute value of the difference divided by the sum across all scans and all runs. The average values at both frequencies across all observations was found to be IP = 0.21% and IP = 0.88% at 100 GHz and 150 GHz respectively.

To confirm that the elevation nods revealed IP and not some other source of spurious polarization, the fit values for both the sum and difference were plotted as a function of elevation. The summed data increased with elevation due to greater atmospheric loading on the PSBs, and the differenced data was shown to scale proportionally to the sum. The correlation between summed and differenced data indicate true IP. This correlation for each of the elevation nods in the final deep observation is shown in Figure 2.13.

Figure 2.14 summarizes the IP found for each of the PSBs across all the FRM observations. The x-axis is labeled in terms of the number of overall testing runs performed on the FRMs. Runs 5–8 are the FRM deep observations, while runs 9–11 are FRM observations pointing near the weak arm of the galaxy. Due to pointing issues, these scans were never used but provide a good blank sky calibration. If the strength of the observed signal in the field of view were somehow
Figure 2.12: From top to bottom: Elevation angle, FRM150A demodulated sum, and FRM150A demodulated difference data for an elevation nod versus time for FRM 150A. The mean has been subtracted from both the sum and the difference data so that small changes can be seen. The difference data is approximately 1% of the sum data, indicative of instrumental polarization.
Figure 2.13: A plot showing the results of the summed (left half) and differenced (right half) demodulated timestreams for each of the individual FRM bolometers. Each blue circle represents the gains derived from a single elevation nod and the red line is the fit to the data. Each individual detector is well fit by a linearly increasing or decreasing slope as a function of elevation angle, and the summed data correlates heavily with the differenced data for the same bolometer.
the cause of the bolometer IP, a difference would be observed between runs 5–8 and runs 9–11. Because no difference was seen, the possibility of the observing patch yielding IP was ruled out.

The points shown represent the mean IP found across each full observation and the error bars represent the standard deviation among those values. The choice to use a dynamic IP subtraction, where the IP value derived from the preceding elevation nod was used for subsequent half-scans, was made due to the relatively high standard deviation of IP throughout a single observation. It is evident from Figure 2.14 that, taking into account the standard deviation, the IP for each individual PSB remains relatively stable over time, though some bolometers experience a much higher IP value than others. Many of the PSBs are consistent with zero instrumental polarization, though even these pixels were corrected dynamically during processing for IP leakage. Again, the elevation nods confirm what the laboratory testing revealed, FRM150A (RTC29 and RTC30) experience the highest IP at approximately 1%.

Map Making

After timestream data processing was complete, Stokes $T$, $Q$, and $U$ were derived following the formalism in [78]. Temperature data were obtained by binning the filtered detector timestreams into map pixels, $p$, as

$$ T(p) = \frac{\left( \sum_{i} \sum_{j \in p} w_{ij}^+ d_{ij}^+ \right)}{\left( \sum_{i} w_{ij}^+ \right)}, $$

where the indices $i$ and $j$ denote the PSB channel number and timestream sample number respectively. The variable $d^+$ is the sum demodulated timestream, $n$ is the number of PSBs, and $w^+$ is the weight assigned to each demodulated sum. Stokes $Q$ and $U$ were calculated by

$$ \sum_{i} \sum_{j \in p} w_{ij}^- \begin{pmatrix} d_{ij}^- \alpha_{ij} \\ d_{ij}^- \beta_{ij} \end{pmatrix} = \frac{1}{2} \sum_{i} \sum_{j \in p} w_{ij}^- \begin{pmatrix} \alpha_{ij}^2 & \alpha_{ij} \beta_{ij} \\ \alpha_{ij} \beta_{ij} & \beta_{ij}^2 \end{pmatrix} \begin{pmatrix} Q(p) \\ U(p) \end{pmatrix} $$

(2.5)
Figure 2.14: A plot summarizing the instrumental polarization values found for each bolometer during each observation. The FRM deep observations correspond in the plots to runs 5–8. Runs 9–11 were pointing at blank sky in the weak arm of the Galaxy, but serve as a good blank sky reference for IP values.
where $w^{-}$ is the weight assigned to the demodulated timestream difference and $\alpha$ and $\beta$ are the differenced orientation factors

\[
\alpha_{ij} = \gamma_i (\cos[2(\psi_{ij} + 45^\circ)] - \cos[2(\psi_{ij} - 45^\circ)]) \quad (2.6)
\]

\[
\beta_{ij} = \gamma_i (\sin[2(\psi_{ij} + 45^\circ)] - \sin[2(\psi_{ij} - 45^\circ)]) \quad (2.7)
\]

The factors of ±45 degrees come from the Faraday rotation of the nominal polarization angle of the $i^{th}$ PSB. The same correction factor describing cross polar leakage, $\gamma_i = \frac{1 - \epsilon_i}{1 + \epsilon_i}$, is applied to both terms $\alpha_{ij}$ and $\beta_{ij}$. Equation 2.5 is degenerate for a single FRM PSB at time $j$. To break the degeneracy, the same sky pixel $p$ is observed by the same modulated detector at a different polarization angle $\psi_{ij}$. With more than one observing angle the off-diagonal elements of Equation 5, $\alpha_{ij}\beta_{ij}$, average to zero and the matrix can be inverted to solve for $Q$ and $U$.

For each half-scan the inverse of the variance of the pair sum and difference timestreams is used as the weight $w^\pm$. Total integration time is also computed for each pixel.

**Absolute Calibration**

To relate detector units to CMB units, BICEP measurements are cross-correlated with WMAP temperature maps to derive an absolute gain calibration. A complete description is given in [19].

This same comparison cannot be made for the FRMs, as the FRMs were never biased and observing during CMB observations. Instead, FRM detectors were calibrated via comparison with the calibrated 3-year BICEP temperature maps of the bright arm of the Galaxy. The absolute calibration factor was computed as the slope of a linear fit to the pixel-pixel scatter plot of the two maps. This is shown in Figure 2.15. The variance of the difference between the fit data and the BICEP data is used to compute the uncertainty. This method yields values of $1.1288 \pm 0.0023$ and $1.0011 \pm 0.0017$ times the nominal absolute calibration factors for BICEP at 100 and 150 GHz respectively.
Figure 2.15: Pixel-pixel scatter plots of the Galactic FRM maps versus the nominal BICEP maps with identical masking and filtering. The slope of the linear fit to the pixel-pixel scatter plot was used to derive the absolute gain calibration factors for the 100 GHz (top) and 150 GHz (bottom) FRMs. The sample variance from the fitted line was used to determine an uncertainty on this absolute calibration.
2.5 Map Results

Figure 2.16 shows both the integration time and Galactic temperature maps derived from the shallow FRM observation at 100 GHz and 150 GHz. Because the shallow observation featured long sweeps of the Galaxy, half-scans were filtered using a second order polynomial rather than mean subtraction. The maps were binned with Healpix [79] pixelization $0.25^\circ$. A $\sigma = 0.5^\circ$ Gaussian smoothing function has been applied to all maps.

Due to glitches in the rotator bias signals, only the first and last 9 hour sections resulted in usable data, a total of approximately seventeen hours of integration time including data cuts. The deepest integration time was $\approx 15$ s per pixel, with an average of 4.1 and 5.4 s per pixel at 100 and 150 GHz, respectively. Data from 5 out of 6 of the FRM pixels (two at 100 and three at 150) were used for these maps; one was omitted due to biasing problems.

The temperature maps from this observation reveal bright Galactic emission confined mainly to $|b| < 1^\circ$, with both large scale features and compact sources. The intensity is greater at 150 GHz than at 100 GHz, as predicted by models of dust in the interstellar medium [1], and consistent with other recent mm-wave experiments [80, 77]. Although the off-Galactic portions of the map show relatively large noise fluctuations due to the low integration time, the Galactic signal dominates the noise for both bands. This is, however, not true for the much fainter polarization signal (not pictured); a detection of statistically-significant polarization was not seen in this observation.

The left two columns of Figure 10 show integration time as well as the Galactic $T$, $Q$, $U$, polarization vector, and $Q$ and $U$ difference maps at both 100 GHz and 150 GHz from the FRM deep observations. Only two FRMs, 100A and 150A, were used to accumulate the maps. Maps are binned into Healpix [79] $0.1^\circ$ pixels and high-pass filtered via DC offset subtraction. A $\sigma = 0.2^\circ$ Gaussian smoothing kernel has been applied to the maps at both frequencies.

The integration time in this region averages 77.7 and 67.2 seconds per map pixel at 100 and 150 GHz respectively. The $T$, $Q$, $U$ maps all show an increase in signal from 100 to 150 GHz. At both frequencies, large-scale positive $Q$ and nega-
Figure 2.16: Integration time and temperature maps from the shallow FRM observation on April 21, 2006. The temperature of the Galactic plane can clearly be seen above the residual noise. The large scale noise features at right ascension \(180^\circ - 200^\circ\) and declination of \(\approx -63^\circ\) are an artifact of masking and filtering. Although the non-FRM PSBs are observing during this shallow observation, the slow scan speed places their observing band well below the \(1/f\) knee, leading to excessive noise. For this reason, these maps are intentionally omitted from further analysis.
Figure 2.17: FRM and BICEP maps of the deep integration region. From top to bottom, each column displays a single plot of temperature with polarization vectors (T & P), integration time ($T_{\text{int}}$), Stokes $Q$, Stokes $U$, and $Q$ and $U$ scan direction difference maps (Diff $Q$ and Diff $U$). All maps are displayed in $\mu$K except for integration time, which is displayed in seconds. The first two columns are the FRM deep observation results at 100 and 150 GHz. The last two columns show maps made with the nominal BICEP observations using non-FRM PSBs at 100 and 150 GHz.
Figure 2.18: A schematic of the difference between filtering the nominal BICEP Galactic half-scans (left) and the FRM observation half-scans (right). The mean of the FRM observations is weighted more heavily by the strong signal of the Galaxy, which comprises a much larger fraction of the overall scan data.

tive $U$ polarization can be seen along the plane of the Galaxy. These polarization features are generally confined to within $|b| < 1.5^\circ$ of the center of the Galactic plane. The polarized portions of the maps are coincident with the largest magnitude intensity in Declination and slightly offset towards lower Right Ascension. At both frequencies, the magnitude of the signal in Stokes $U$ is approximately five times larger than the signal in Stokes $Q$. The average signal in Stokes $U$ confined to $|b| < 1.5^\circ$ is given by $-16.2 \, \mu K$ and $-48.4 \, \mu K$ at 100 and 150 GHz respectively. Given the error bars defined by difference maps (Section 2.5.1 in the subsection “Map Noise”), this corresponds to an average S/N ratio of 1 and 3 in each smoothed map pixel for the average galactic polarization signal seen at 100 and 150 GHz respectively.

BICEP observations of the same region were made during the 2006 - 2008 BICEP observing seasons and compared to the FRM data. The BICEP Galactic scan strategy uses long sweeping scans across the entire Galactic plane at a step size of $0.25^\circ$, prohibiting a direct comparison of a single non-FRM PSB pair to a single FRM PSB pair; for a single non-FRM PSB pair, the integration time is too low and the number of boresight angles is insufficient to accumulate polarization. As such, maps were accumulated using all non-flagged PSB pairs in the focal plane. PSB pairs containing FRMs (among others) were flagged and excluded from
analysis. The data were taken over all three BICEP observing seasons, totaling 763 observing hours for the entire gal bright region. Non-FRM data were processed using the same analysis pipeline as the FRM data with a few exceptions. The initial deconvolution was applied with a low-pass filter at 5 Hz. Following deconvolution, the data were downsampled to 10 Hz before preliminary processing and relative gain calibration. The demodulation step was omitted; the sum and difference data were taken between the two PSBs within each pair. The data were then subjected to the same masking and filtering strategy as the FRM pixels, with the additional complication that all data outside the region of deep integration were also masked from the fits. This is to keep large portions of off-Galactic data from weighting the DC offset subtraction, which was not possible with the FRM deep integration scan strategy. The overall idea is shown pictorially in Figure 2.18. The relative amount of data taken off the Galactic plane during the BICEP scans makes the mean value subtracted during filtering much lower than it would be for the FRM scans. To remove such an effect, a masking scheme was applied to the nominal BICEP data as shown in Figure 2.19. The endpoints of the mask were chosen to mimic the FRM scan strategy such that the high pass filtering between the two data sets was identical.

Maps were accumulated using the same formalism as Section 2.4.2, except Equations 2.6 and 2.7 are modified to

$$\alpha_{i,j} = \gamma_i \left( \cos[2(\psi_{i+1,j})] - \cos[2(\psi_{ij})] \right)$$  (2.8)

$$\beta_{ij} = \gamma_i \left( \sin[2(\psi_{i+1,j})] - \sin[2(\psi_{ij})] \right)$$  (2.9)

where $i$ and $i + 1$ are the indices of the two PSBs within the pair that have been subtracted. The binning and smoothing of the maps is identical to the FRM deep observations.

The right two columns of Figure 2.17 show the integration time as well as Galactic $T$, $Q$, $U$, polarization vector, and $Q$ and $U$ difference maps at both 100 GHz and 150 GHz from the BICEP observations accumulated using the non-flagged focal plane PSBs. A total integration time of approximately 324 hours and 194 hours was taken at 100 and 150 GHz respectively, with a mean integration time of 412 seconds per pixel at 100 GHz and 246 seconds per pixel at 150 GHz. The
Figure 2.19: Masks used for processing BICEP data to mimic the FRM observations. The left column shows the unmasked plots: FDS fiducial model (top) and BICEP (bottom). The right column shows the remaining data after the mask has been applied.
maps exhibit very good agreement with the FRM maps in both temperature and polarization at both frequencies. As evidenced by the polarization vector maps in Figure 2.17, both data sets show the same strong Galactic polarization, roughly perpendicular to the Galactic plane, with the highest polarized signal in the upper left quadrant. Both the FRM and non-FRM BICEP maps exhibit a diminution of polarization across the Galaxy with increasing RA. At each frequency and for all polarization maps, the contribution of Stokes U dominates the polarized signal. Both the FRM and non-FRM BICEP maps are consistent with the polarized emission expectations based on Galactic plane maps published previously by BICEP and other recent mm-wave experiments [81, 82, 83, 80, 77].

The polarization fraction is found by calculating the median value of the quotient of polarization and temperature across all map pixels where $|b| < 1.5^\circ$. The variance in this same region is used to compute the uncertainty on these values. The FRM maps yield a polarization fraction of $1.32\% \pm 2.17\%$ and $2.36\% \pm 0.21\%$ at 100 and 150 GHz respectively. The results derived from BICEP’s non-FRM maps show strong agreement: $1.53\% \pm 0.61\%$ at 100 GHz and $2.31\% \pm 0.02\%$ at 150 GHz. For both the FRM and non-FRM BICEP maps, 150 GHz displays little variance in polarization fraction across the galactic plane, whereas the 100 GHz variance is much higher.

### 2.5.1 Noise and Systematics

The deep FRM observations were analyzed for noise and other transient issues that may have resulted from the use of the devices. The results of the FRM observations were also subjected to many self-consistency checks to in order to verify the accuracy of the data presented here. Due to the low signal-to-noise achieved for the shallow observations, only the deep observations are subjected to the difference map and noise analysis.

**Difference Maps**

Difference maps, in which the map data is split in half and differenced, were used to check the self-consistency of the Galactic maps. Although some residual
Table 2.3: The rms map noise for each of the three difference maps analyzed for the deep FRM observations. The rms is quoted for an effective smoothed pixel size of $0.24 \text{ deg}^2$. For comparison, the Galactic U signal for the FRM maps was found to be $-16.2 \mu K$ and $-48.4 \mu K$ at 100 and 150 GHz respectively.

<table>
<thead>
<tr>
<th>Difference Map Type</th>
<th>100 GHz rms ($\mu K$/pixel)</th>
<th>150 GHz rms ($\mu K$/pixel)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Scan Direction</td>
<td>12.3</td>
<td>20.5</td>
</tr>
<tr>
<td>PSB Pair</td>
<td>13.4</td>
<td>12.8</td>
</tr>
<tr>
<td>Observation</td>
<td>22.6</td>
<td>45.0</td>
</tr>
</tbody>
</table>

The signal may remain due to timestream filtering effects, the expected signal of the maps is nearly zero and all large scale Galactic temperature and polarization signals should vanish. Statistical polarization errors are quantified by taking the standard deviation for all $Q$ and $U$ map pixels at both 100 and 150 GHz. These results are summarized in Table 2.3. For consistency with the map results, a $\sigma = 0.2^\circ$ Gaussian smoothing function is applied to all maps. As such, the rms map noise for each pixel in this Table 2.3 is quoted for an effective pixel size of $0.24 \text{ deg}^2$.

Three separate divisions of the data were analyzed: scan-direction, PSB pair, and observation. The scan-direction split, where the two data sets are separated based on half-scan direction, can generate noise based on thermal instability at half-scan endpoints. The scan-direction difference maps are shown in the last two rows of Figure 2.17. PSB pair difference maps subtract the maps accumulated from individual detectors within a pair. This difference map is unique to FRM analysis, as a FRM turns a single BICEP PSB into a polarimeter, allowing $T$, $Q$, and $U$ maps to be accumulated for individual detectors within a PSB pair. This difference map is perhaps the most robust data quality test for the FRMs as it is sensitive to many factors including relative gain mismatches, demodulation errors between detectors, and thermal stability. The PSB pair difference maps are shown in Figure 2.20. Finally, observation difference maps subtract maps accumulated from odd and even numbered observing runs. This jackknife probes for sensitivity to weather changes. The FRM maps show the highest sensitivity to changes in weather, especially in the 150 GHz PSB. The main source of the elevated signal
Figure 2.20: $T$, $Q$, and $U$ differenced data from individual detectors within a single PSB pair. The $Q$ and $U$ scales are identical to Figure 7 for comparison, whereas the temperature scale is reduced to $\pm 500 \mu K$. Large scale features in both polarization and temperature are no longer apparent in the maps.

at 150 GHz is a very high noise contribution at the lowest declination, where the least integration time occurs.

Difference maps of the two data sets, the FRM and non-FRM BICEP, are displayed in Figure 2.21. The rms map noise per pixel is given by 20.6 $\mu K$ rms/pixel at 100 GHz and 35.6 $\mu K$ rms/pixel at 150 GHz. This differencing should remove all true sky signal and leave only uncorrelated noise, which should combine as the quadrature sum of the noise in the two individual maps. To estimate the map noise in the non-FRM BICEP map, the scan direction difference map was utilized, yielding 12.5 and 12.0 $\mu K$ per effective pixel at 100 and 150 GHz respectively. Given the map noise values in Table 2.3 for the FRM map, the difference map noise is consistent with the quadrature sum of the map noise of the FRM and non-FRM BICEP difference maps.

Although many properties of the FRMs were characterized during initial
Figure 2.21: Maps of the difference between the FRM and non-FRM BICEP observations. Galactic signal, in both temperature and polarization, has been removed. This is particularly evident in celestial U, where both the FRM and non-FRM BICEP maps originally exhibited significant polarized structure. No similar structure is evident in the Stokes U difference map.
testing and deployment, the lack of time for field testing prevented a complete characterization for the FRMs and is beyond the scope of this paper. Specifically, properties such as cross-polar leakage and polarization inefficiency were not studied in-depth, and it was assumed throughout this analysis that the FRMs behaved as ideal devices in these respects. Further, the DSC calibration revealed an imperfect bias signal for most of the FRMs, but for the deep observations it was assumed that the FRMs accomplished an ideal rotation of $\pm 45^\circ$ to the nominal angle of the coupled PSBs. Deviations from this perfect modulation would result in miscalibrated detector angles and reduced polarization efficiency, which would alter the polarization signal in the $Q$ and $U$ maps.

Although the FRM maps are absolutely calibrated via non-FRM BICEP maps, the calibration is done using temperature alone. Deviations from the (assumed perfect) polarization properties of the FRMs (described above) would result in marked differences between the FRM and non-FRM $Q$ and $U$ difference maps. The fact that these maps are statistically consistent with each other limits the presence of FRM non-idealities to the percent level.

**Map Noise**

Map noise in FRM observations 2–5 was quantified via the scan-direction difference map. If the map was not sensitive to systematic effects (e.g. systematics are not present), then the noise should be Gaussian white and integrate down with the square-root of the integration time and the square-root of the number of detectors. To validate this assumption for these maps, the data were reprocessed several times, each time removing a different number of points from the transitions during demodulation. Maps were accumulated using a range of cuts from four points (two on either side of the transition) to twelve points (six on either side of the transition). The results are shown in Figure 2.22. The data exhibit a decrease in rms map noise with a $\sqrt{t}$ dependence based on the data cuts. Additionally, at 100 GHz, the data show a $\sqrt{2}$ increase in the map noise when accumulating the difference maps with a single PSB as opposed to the pair. At 150 GHz, however, the $\sqrt{2}$ dependence on PSB number is not observed.
Figure 2.22: Plot of rms map noise for the scan-direction difference maps as a function of the portion of total integration time used. At each frequency, two lines are displayed representing the rms difference map noise from individual PSB and the PSB pair. The thinner dashed lines are a guide to the eye exhibiting square-root of integration time dependence; they are plotted in the same color as their corresponding solid lines. At 100 GHz, the square-root of integration time scaling is obvious, whereas for a single PSB at 150 GHz there is no such dependence. In addition, the 100 GHz PSB pair scales as the square-root of the number of detectors ($\sqrt{2}$) from the individual PSBs to the pair. At 150 GHz, the $\sqrt{2}$ reduction in noise is not seen.
Histograms of the noise at 100 GHz (left) and 150 GHz (right) in the unsmoothed scan direction jackknife maps for the FRM deep observations. Gaussian fits using the derived mean and standard deviation at each frequency are shown for comparison.

Histograms of the noise distribution were also computed using the unsmoothed difference maps at both frequencies and are shown in Figure 2.23. Only map pixels with an integration time of 50 seconds or more are used in this analysis to avoid pixels with low integration time weighting the result. The histograms show that the amplitude distribution of the noise is roughly Gaussian and distributed about zero within one standard deviation of the mean. The mean $\pm 1$ standard deviation are given by $-4.1 \pm 75.3 \mu K$ and $-11.1 \pm 99.4 \mu K$ at 100 GHz and 150 GHz respectively.

Assuming these maps are Gaussian white noise-dominated, detector noise was quantified by multiplying the $Q$ and $U$ pixel values by the square-root of the integration time per pixel and taking the standard deviation across the maps. This yielded an average instantaneous (i.e., single Stokes parameter) “NEQ per feed” of $420 \mu K\sqrt{s}$ and $501 \mu K\sqrt{s}$ for 100 GHz and 150 GHz respectively. The value at 100 GHz is comparable to the values in [73, 77], though the value at 150 GHz is elevated by a factor of 1.6. The source of the excess noise at 150 GHz
is unknown, though both the shape of the histogram, which is skewed towards positive amplitude, and the lack of $\sqrt{2}$ dependence when doubling the number of PSBs indicate that systematic effects can not be ruled out.

### 2.5.2 Non-FRM Pixels

During the FRM observations, the non-FRM pixels were observing similar patches of sky based on their relative positions in the focal plane. Because of the FRM observing strategy, each of the nominal BICEP pixels observed several small area patches near the center of the Galactic plane. Therefore, using all the FRM pixels, a fuller image of the Galactic plane can be created with a deeper integration time. Maps were accumulated using the non-flagged BICEP pixels with the same process described in Section 2.5 for the nominal BICEP scans. The result is shown in Figure 2.24.

The Galactic temperature maps again show bright Galactic emission confined mainly to $|b|<1^\circ$, with both large scale features and compact sources. A greater intensity, increased polarized dust intensity, and a greater number of small features are observed at 150 GHz, consistent with theoretical predictions and BICEP’s decreasing beam size with increasing frequency. Due to the increased integration time from using the pixels across the entire focal plane, the signal to noise ratio is greatly increased in temperature. However, in polarization, large scale noise can still be observed in both $Q$ and $U$ and at both frequencies. Despite the noise, the dominant $U$ polarization can be seen in both frequency bands along the Galactic plane, with an obvious increase in magnitude from 100 GHz to 150 GHz. A slightly lower $Q$ polarization can also be seen along the plane of the Galaxy at both bands.

During the FRM scans, however, the scan velocity is much slower than nominal BICEP scans, as explained in Section 2.4.1. Since this velocity is below the $1/f$ knee for unmodulated detectors, excess noise in the polarization can be seen in the maps in Figure 2.24. Further, because of the short raster scans, the data must still be processed with the small galactic mask ($1.5^\circ$) and mean subtraction as our filtering, which does not eliminate large scale atmospheric noise. For an
Figure 2.24: Maps accumulated from the timestreams of the non-FRM pixels during FRM observations. The bright arms of the Galaxy can be seen above the residual noise in temperature and the negative $U$ polarization is apparent at both 100 GHz and 150 GHz.
Figure 2.25: Maps accumulated from the timestreams of the non-FRM pixels during FRM observations. The bright arms of the Galaxy can be seen above the residual noise in temperature and the negative $U$ polarization is apparent at both 100 GHz and 150 GHz.
in-depth analysis of the non-FRM noise properties, the scan-direction jackknives were generated and are shown in Figure 2.25. Given the integration time, the level of noise is excessively high. Performing the same noise analysis as was done for the FRM pixels, where the average of the $Q$ and $U$ map pixel signal multiplied by the square-root of the integration time per pixel is used to determine an estimate for the noise, noise values of approximately 1100 and 1400 $\mu K \sqrt{s}$ is observed for 100 and 150 GHz respectively. Obviously, there is much residual noise left in these maps; the associated noise quantities are factors of three and four higher than the values from the FRM maps. In light of this analysis, the use of the FRMs opens the parameter space for scan strategy, including, but not limited to, scan speed and angular distance in azimuth.

2.6 Results from Unbiased FRMs

For the majority of the 2006 observing season, the FRMs remained unbiased in the BICEP focal plane. Without modulation, the PSB pairs coupled to the FRMs functioned as nominal BICEP PSBs, where the sum and difference of the pair were used to calculate temperature and polarization respectively. Using 2006 Galactic scans, noise was computed for all FRM feeds to analyze whether the presence of the FRMs in the optical path introduced additional noise within the detectors.

To compute the noise, the spectral power distributions for detector sums and differences for each calibrated half-scan is calculated and averaged for all FRM pixels at each band. This is done after masking, filtering, and sum-differencing each PSB pair. Figure 2.26 shows the results of this analysis. The pair-sum data exhibits increasing $1/f$ contamination from 100 to 150 GHz. Pair differencing removes this contamination, resulting in nearly white noise above 0.1 Hz. Averaging the pair-difference periodogram between 0.1 and 1.0 Hz gives a Noise Equivalent Temperature (NET) per detector of 620 and 430 $\mu K \sqrt{s}$ for FRM feeds at 100 and 150 GHz respectively. Previous studies [73, 77] characterizing the pixels in BICEP (which do not include the FRMs) give BICEP NET values of 525 $\mu K \sqrt{s}$ at 100
Figure 2.26: Average power spectral distributions calculated for the pair-sum and pair-differenced FRM PSBs at 100 and 150 GHz during Galactic observations in 2006 while the FRMs remained unbiased. Data for the plots were taken while the FRMs were unbiased in the BICEP focal plane. The second order polynomial filtering can be seen at 0.05 Hz, where the pair-difference noise is lower than the white noise floor.

GHz and 450 $\mu K \sqrt{s}$ at 150 GHz. The elevated noise at 100 GHz was found to stem from the PSB pair corresponding to a single FRM (100C). When removed, the FRM NET at 100 GHz is 520 $\mu K \sqrt{s}$. Both bands show excellent agreement with the quantities derived from observation quality pixels, demonstrating that unbiased FRMs do not introduce any excess noise into the detectors.

2.7 Conclusions

In this chapter, I report on the first detection of Galactic polarization using fast, active optical polarization modulation. The observed polarized signal is consistent with both BICEP’s and other recent experiments’ maps of Galactic polarization. Although observation and testing time limited the scope of the comparison that could be made between the nominal BICEP instrument and the BICEP instrument modulated by FRMs, the FRMs have been shown to be effective solid-state polarization modulators for both laboratory and celestial polarimetry.
applications. In particular, the 100 GHz FRM was found to exhibit equivalent noise to similar BICEP PSB pairs without FRMs, but expands the parameter space for the instrument’s scan strategy, relaxing constraints on scan speed and potentially increasing the observed sky fraction. Ultimately, the choice of whether to use a fast active modulator such as a FRM depends upon the details and constraints of the experiment. Devices similar to the FRMs are currently being explored for use in several experiments [84, 85]. FRM technology could be applied wherever fast solid-state polarization modulation is called for.

FRMs are a promising technology due to their design flexibility, low systematic polarization, large bandwidth, and ability to be used over a large frequency range. Further, FRMs have been shown to be functional in both modulating and non-modulating modes without adding noise, making them a flexible option for a mm-wave polarization modulation. FRMs also have potential application as modulators for cm and sub-mm wavelength polarimeters.

2.8 Acknowledgements

Chapter 3

POLARBEAR Pointing

“The polar bear is classified as a vulnerable species, with eight of the nineteen polar bear subpopulations in decline. Being both curious and scavengers, polar bears investigate and consume garbage where they come into contact with humans. Polar bears may attempt to consume almost anything they can find, including hazardous substances.” (Wikipedia)

This chapter highlights my work on the POLARBEAR project and, more specifically, telescope boresight pointing.

3.1 An overview of POLARBEAR

Figure 3.1 shows the projection of the POLARBEAR CMB angular power spectrum constraints for both $E$-mode and $B$-mode polarization. The projection is for a total observation time of 8 hours per day for 9 months. The blue error bars are modeled with no foreground consideration, and red error bars accounts for foregrounds following [86, 87, 88]. Two $B$-mode gravitational wave spectra are shown for comparison: $r = 0.1$ and $0.025$. POLARBEAR is projected to achieve a $2\sigma$ detection of $r = 0.025$ and will also characterize the gravitational lensing signals in an attempt to measure properties of cosmological neutrinos [89, 90].
Figure 3.1: Projection of POLARBEAR limits on the B-mode power spectra after one observing season. The inflation model plots with $r=0.1$ and $r = 0.025$ are also shown for comparison. Figure courtesy of Nathan Miller.

### 3.1.1 POLARBEAR Design

POLARBEAR consists of the POLARBEAR receiver and the Huan Tran Telescope (HTT) and is located at the James Ax Observatory in the Atacama Desert in Chile. POLARBEAR is designed to have unprecedented sensitivity over an extremely large range of angular scales, surveying four $15^\circ \times 15^\circ$ regions (to detect the large angular scale $B$-modes from the GWB) with 3.5 arcminute resolution with precise control and mitigation of systematic effects. The following section describes the instrument’s detectors and optics. For more details on the design of POLARBEAR we refer the reader to [91].

**Huan Tran Telescope**

The Huan Tran Telescope (Figure 3.2) is an off-axis Gregorian telescope with a 3.5 meter primary aperture that provides the 3.5 arcminute angular resolution necessary to characterize the gravitational lensing signal. The primary is composed of a 2.5 meter high-precision monolithic reflector surrounded by a 0.5 meter radius guard ring, which is used to prevent beam spillover and minimize sidelobes. Incident radiation is directed from the primary to the baffled secondary
Figure 3.2: Image of the Huan Tran Telescope at the James Ax Observatory in 2012 with important components labeled. Incident light is reflected off the primary mirror, onto the secondary mirror (not seen in the image), into the cryogenic receiver, and onto the focal plane. Figure courtesy of Zigmund Kermish.

mirror that re-images the primary image onto the detector array. The telescope was designed for large optical throughput while simultaneously mitigating systematic effects such as temperature-to-polarization leakage and cross-polarization. The co-moving ground shield and secondary baffle blocks stray light from entering the optical path [92].

Receiver Optics

The fundamental noise limit for any CMB experiment is known as photon noise, which is set by the quantum fluctuations in the arrival rate of photons. CMB polarimeters seek to minimize all other noise sources (e.g., phonon noise, readout noise, etc), such that photon noise dominates. To achieve a high level of sensitivity, the POLARBEAR focal plane is cooled to 250 mK so that thermal carrier noise in the detectors is smaller than the photon noise. POLARBEAR achieves this cooling with closed-cycle refrigeration: a pulse tube cooler and a three-stage helium sorption refrigerator.
Figure 3.3: Cross section of the POLARBEAR receiver with optical ray trace paths shown.

Figure 3.3 shows a cross-section through the receiver. Before incidence on the focal plane, millimeter-wave radiation entering the receiver encounters, sequentially, a Zotefoam vacuum window, a rotating sapphire half wave plate (HWP), single- and multi-layer metal mesh filters, porous Teflon IR blocking filters, and anti-reflection coated ultra-high molecular weight polyethylene re-imaging lenses. The re-imaging lenses and multi-layer metal mesh filters are emissive in the spectral band of the detectors and are therefore cooled to prevent an increase in optical loading.

The rotating half-wave plate (HWP) is a 3.1 mm thick single crystal disk of A-plane sapphire. Because of its birefringent properties, rotation of the HWP modulates the polarization of the signal (only), and thus allows for mitigation of instrumental systematic effects which do not have the requisite symmetry of true CMB polarization. The HWP is cooled to 70K to reduce thermal emission. POLARBEAR’s focal plane features a planar array of bolometers, requiring a flat, telecentric image from the primary reflector. This is achieved using three re-imaging lenses coupled to the telescope optics. A cold aperture stop with an absorbing edge is imaged by the inner high-precision monolithic section of the
primary mirror, suppressing sidelobe response.

### 3.1.2 Detectors and Focal Plane

While there are a number of competing detector technologies for CMB experiments, only bolometers operated from sub-Kelvin platforms and at frequencies \( \geq 90 \) GHz are sensitive enough to be photon noise limited. A detector in this regime cannot make significant gains in sensitivity by improvements to the detector itself; instead, gains are made by increasing the number of detectors and the throughput of the telescope. POLARBEAR is designed to meet this goal, using entirely lithographed superconducting transition edge sensor (TES) bolometers in scalable arrays. Figure 3.4a shows an image of the entire POLARBEAR focal plane. The focal plane is composed of seven separate hexagonal sub-arrays. One such array, along with readout electronics, is shown in Figure 3.4b.

![Figure 3.4: (a) An image of the full POLARBEAR focal plane as seen from the top (b) A single hex wafer plus vertically integrated readout (c) Microscope image of a single POLARBEAR pixel with important components labeled. Image courtesy of Zigmund Kermish.](image)

The POLARBEAR focal plane features 637 pixels (1274 TES bolometers) with a spectral band between 120 and 180 GHz to avoid atmospheric contamination. Each pixel (Figure 3.4c) features several lithographed components: a double-slot dipole antenna that separates millimeter-wave signals into orthogonally polarized components, distributed bandpass filters that reject frequencies outside of the desired observing band, and superconducting transition edge sensors that are thermally isolated to provide photon noise limited detection of the
Each antenna is coupled to an anti-reflection coated silicon lenslet\textsuperscript{1} that serves to increase forward gain, decrease radiation lost to substrate modes, and magnify the effective size of the antenna (thus increasing the area available for the rest of the pixel). The lenslets are quarter-wavelength anti-reflection coated to minimize reflection loss at the surface. The coupling of the extended hemispherical lenslet above the planar antenna produces a diffraction-limited beam with Gaussianity and directivity similar to that of a conical feedhorn \cite{95}. Both TES bolometer theory and the design of the POLARBEAR pixels are discussed in much greater depth in Sections 4.2 and 4.3.

Another benefit of superconducting TES bolometers is that they can be read-out and multiplexed by low noise SQUID amplifiers, allowing for simultaneous readout of thousands of detectors. POLARBEAR detector signals are read-out with a frequency-domain multiplexed readout using cryogenic SQUID ammeters. A group of eight transition edge sensor bolometers is AC voltage-biased, each with a different frequency, and its current response is summed using a single SQUID. The “frequency comb” SQUID signals exit the cryostat and are demodulated at 300 K. This multiplexed readout reduces the otherwise prohibitive amount of wiring, and is designed as a vertical stack (Figure 4b) allowing POLARBEAR to take full advantage of (precious) cold focal plane area.

3.1.3 Inyo Mountains Engineering Run

An engineering run of POLARBEAR was performed in 2010 at the CARMA site in the Inyo Mountains of Eastern California. The focal plane of this engineering run contained three of seven hexagonal sub-arrays and a 50% attenuating filter at 4 K was placed in the optical path to reduce the high atmospheric power present in California (which is absent in the James Ax Observatory in the Atacama Desert). Bright astrophysical point sources were observed in order to characterize POLARBEAR’s beam parameters. Data was accumulated for each source by scanning the

\textsuperscript{1}Except for the antennas on a single wafer that are coupled to alumina, and not silicon, lenslets.
telescope in azimuth and stepping in elevation. In this way, each pixel crosses the source many times. Co-added maps of the brightest celestial source, Jupiter, are shown in Figure 3.5a. These maps give a best fit Gaussian beam with full width at half maximum of 3.8 arcminutes.

Observations of the bright, polarized mm-wave source Tau A were performed during the Cedar Flat engineering run. A row of 12 observing pixels were chosen and rastered back and forth across Tau A for approximately two hours while the HWP was stationary. Maps were subsequently produced by rotating the HWP and reobserving Tau A. Figure 3.5b shows the co-added temperature and polarization maps of Tau A. The polarization magnitude and angle show good agreement with published results [96].

![Figure 3.5](image_url)

Figure 3.5: (a) Co-added map of Jupiter used to characterize beam Gaussicity, ellipticity, and beam-width. b) Temperature and polarization map of Tau-A from 2 hours of observing data with 12 pixels. Data for both plots were obtained during the 2010 Inyo Mountains, California engineering run.

Other tests (such as polarization response) were also performed during the Cedar Flat engineering run to characterize beam parameters. POLARBEAR is more than a factor of two better than the most stringent requirements for differential pointing, differential ellipticity, and differential beam size in order to detect CMB B-Modes from inflationary gravitational waves with $r = 0.025$. 
3.1.4 POLARBEAR Current Status

After the successful engineering run, POLARBEAR was moved to the James Ax observatory in late September 2011. The observatory is located on Cerro Toco in the Atacama desert in Chile at an altitude of 5250 m. The location and telescope are shown in Figure 3.6. The telescope was fully assembled by the start of 2012 and first light was achieved on January 10 with an observation of Jupiter.

Figure 3.6: Geographical location of POLARBEAR and an image of the assembled telescope in the Atacama Desert. Figure courtesy of Josquin Errard.

Since April 2012 the telescope has been performing routine observations on a 36 hour cycle: 20 hours are dedicated to CMB observations, 11.5 hours are used for calibration and instrumental characterization, and 4.5 hours are necessary to cycle the $^3$He refrigerator. The following sections are meant to highlight the status of POLARBEAR results.

Beam Properties

Beam properties are investigated by accumulating temperature maps of sources that are significantly smaller than the beam size, such as planets and quasars. Planets are desirable for the magnitude of their signal, which allows
for relatively quick characterization of individual detectors with high signal to noise and yield. Measurements of planets similar to those performed during the engineering run are used to investigate the beam shape and size, asymmetries, offsets, and detector gains.

Figure 3.7(a) shows the pixel locations found by fitting the timestreams of all the active detectors to a Gaussian beam model during a single Saturn observation in April 2012. Figure 3.7(b) shows the co-added beams for each device wafer. Properties such as beam ellipticity and differential pointing can be quantified using these maps and have been found to be comparable to the values obtained in Cedar Flat, well within the requirements set for $r = 0.025$. By accumulating the fits for the beam maps over many months of observations, POLARBEAR has obtained beam fits and locations for over 90% of the detectors in the focal plane.

**Figure 3.7:** (a) Pixel locations in the POLARBEAR focal plane found by fitting a Gaussian to the detector timestreams for a single Saturn observation. (b) Coadded beams using the known pixel locations for each device wafer. Figure courtesy of Nathan Miller.

**Throughput and Noise**

The fractional throughput, $\eta$, is a measure of the detected power from a source at the input of the receiver compared to the expected theoretical power. For
POLARBEAR, the quantity $\eta \Delta \nu$ has been found to be consistent with the expectation of $\eta \approx 35\%$ using the beam maps and elevations nods and a measured band of $\Delta \nu \approx 34$ GHz. Figure 3.8 shows the array NET obtained from power spectral densities (PSDs) of the elevation nods. A substantial fraction of the POLARBEAR detectors operate at a NET of less than $550 \mu K\sqrt{s}$, giving a total array NET of approximately $21 \mu K\sqrt{s}$ [91]. The PSD also allows us to quantify how well the unpolarized atmosphere is suppressed at low frequencies in the differenced data. The $1/f$ knee for POLARBEAR is found from the PSD to occur at approximately $100$ mHz.

![Noise Power Spectrum](image)

Figure 3.8: Noise power spectrum of the POLARBEAR array summed and differenced timestreams. POLARBEAR’s science range is labeled and is above the measured $1/f$ knee.

CMB Observations

POLARBEAR has been observing three CMB patches since April 2012 and, to date, over 2500 hours of CMB data has been taken. Figure 3.9 shows the three observing patches, a preliminary temperature map from a single patch, and the
total observing time accumulated since instrumental first light.

Figure 3.9: POLARBEAR’s (a) current CMB observing patches, (b) preliminary CMB temperature map for a single patch, and (c) total observation and CMB observation time since first light in April 2012.

Each CMB observation utilizes Constant Elevation Scans (CESs) as the patch drifts through a fixed elevation via sky rotation. In order to accurately reconstruct and co-add data accumulated at a range of elevations and over several months time, a pointing model is required that accurately reconstructs the nominal sky locations versus the recorded telescope position. The remainder of this chapter describes my work on POLARBEAR pointing, both in Cedar Flat and Cerro Toco.

3.2 Pointing Model

A reliable pointing model of the telescope is necessary to reconstruct the true boresight pointing of the telescope as a function of the commanded values.
Table 3.1: A description of the physical meaning of the individual parameters that contribute to the modeled pointing error.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IA</td>
<td>Azimuth encoder zero</td>
</tr>
<tr>
<td>CA</td>
<td>Telescope boom not perpendicular to elevation axis</td>
</tr>
<tr>
<td>NPAE</td>
<td>Elevation axis not perpendicular to azimuth axis</td>
</tr>
<tr>
<td>AN</td>
<td>Azimuth axis tilt north of vertical</td>
</tr>
<tr>
<td>AW</td>
<td>Azimuth axis tilt west of vertical</td>
</tr>
<tr>
<td>AN2</td>
<td>Quadrupole of the azimuth bearing north of vertical</td>
</tr>
<tr>
<td>AW2</td>
<td>Quadrupole of the azimuth bearing west of vertical</td>
</tr>
<tr>
<td>DT</td>
<td>Offset in time between measured and nominal data</td>
</tr>
<tr>
<td>IE</td>
<td>Elevation encoder zero + telescope beam not perpendicular to azimuth axis</td>
</tr>
<tr>
<td>TF</td>
<td>Cosine flexure under telescope weight</td>
</tr>
<tr>
<td>TFS</td>
<td>Sine flexure under telescope weight</td>
</tr>
</tbody>
</table>

The full pointing model developed for POLARBEAR is given by:

\[
\Delta az(az_s, el_s) = IA \cos(el_s) - CA + NPAE \sin(el_s) - AN \sin(az_s) \sin(el_s) \\
- AW \cos(az_s) \sin(el_s) - AN2 \sin(2az_s) \sin(el_s) \\
- AW2 \cos(2az_s) \sin(el_s) \\
+ DT(- \sin(lat) + \cos(az_s) \cos(lat) \tan(el_s)) \tag{3.1}
\]

\[
\Delta el(az_s, el_s) = - IE + AN \cos(az_s) - AW \sin(az_s) \\
- AN2 \cos(2az_s) - AW2 \sin(2az_s) + TF \cos(el_s) \\
+ TFS \sin(el_s) - DT \cos(lat) \sin(az_s) \tag{3.2}
\]

where \(\Delta az\) and \(\Delta el\) are the measured offsets in azimuth (az) and elevation (el) from the nominal source az \((az_s)\) and el \((el_s)\) positions and lat is the geographic latitude of the telescope. Table 3.1 describes the physical meaning of each parameter in the model above. The parameters exhibit degeneracies for data taken in a narrow range of azimuth and/or elevation or for small data sets. Therefore, to fit for the pointing parameters that most accurately reconstruct the true telescope boresight location, it is desirable to have a large dataset (in excess of 100 points) that spans a continuous and wide range of azimuth \((> 145^\circ)\) and elevation \((> 40^\circ)\).
Offsets in telescope boresight pointing have several effects on polarization experiments. Random pointing errors smear POLARBEAR’s 3.5 arcminute beam, giving rise to a slight suppression of the recovered $B$-mode lensing signal at high multipoles [97, 98]. In principle, random pointing errors can also lead to $E$ to $B$ leakage [99, 98], but models and simulations have shown that the expected leakage from random pointing jitter is minimal in the case of a scan/modulation strategy that uniformly covers the full range of polarization sensitivity angles in each sky pixel. Specifically, the pointing error results in a white-noise spectrum of $B$-modes and, for jitter constrained to a small fraction of the beam size, the leakage is very small, less than 1% of the $B$-mode power from gravitational lensing on large angular scales [97]. Asymmetric or systematic pointing errors, however, can distort the instrument beam from a perfectly circular Gaussian profile to an ellipse (or other shape), increasing differential ellipticity that also results in $E$ to $B$ leakage. For these reasons, it is in the best interest of POLARBEAR to constrain the pointing rms to the smallest possible value and to have a complete understanding of the pointing residuals.

A simulation was performed to investigate the magnitude of the change in beam size caused by random pointing errors, the result of which is shown in Figure 3.10. The nominal POLARBEAR beam is Gaussian with a 3.5’ FWHM as shown in (a). The simulated nominal beam is shifted 1000 times by generating random, independent errors in az and el from Gaussian distributions with an exaggerated 2’ root-mean-squared (rms) pointing error in each direction. The shifted beams are co-added and fit for a Gaussian, the width of which confirms that the rms pointing error of 2’ adds in quadrature with the known beam width to form the observed beam.

### 3.3 Cedar Flat Results

During the engineering run in 2010, the rms pointing error of the HTT was characterized via optical and telescope measurements.
3.3.1 Optical Pointing

Star Camera Design

Optical pointing was used to help establish the pointing model for the telescope and as a complement to the pointing corrections derived from the receiver itself. Using an optical camera, data can be collected much more quickly and in a larger range of elevation. The design of the optical pointing apparatus, referred to as the star camera, is based on the design from BICEP. Optical signals from the sky pass through a focusing lens and an aperture in the star camera box, the optical axis of which is aligned along the main axis of telescope. Light is reflected off of a 45° angled mirror and continues onward until it is recorded by the StellaCam EX CCD. Figure 3.11 shows the main components of the star camera design.

The only focusing element of the star camera is the lens, so the resolution and field of view are determined solely by this component. The resolution provided by the lens places the fundamental limit on the pointing accuracy of our model.

Figure 3.10: The nominal POLARBEAR beam, a 3.5 arcminute FWHM Gaussian, (a) before and (b) after the consideration of pointing errors. Offsets are generated in azimuth and elevation drawing from random distributions with a FWHM of 2’. The center coordinate offsets shown in (b) are random errors resultant from the finite number of shifts that were performed.
Figure 3.11: The main components of the star camera used for optical pointing of POLARBEAR: (a) the outward facing part of the star camera enclosure, (b) the StellaCam EX, (b) the focusing lens, and (d) the side of the enclosure that is mounted to the telescope boom.

reconstruction. Two available lenses provide two modes of operation: one lens has a short focal length lens that gives a large field of view comparable to POLARBEAR and the other has a longer focal length that gives a small field of view and high resolution. The former is used as a coarse evaluation of the pointing offsets such as azimuth encoder zero; it is used mainly to find the base offsets necessary to locate a source of interest. The majority of interest is on the larger field of view lens, as it has the appropriate resolution for pointing reconstruction. Both lenses have a diameter of 50 mm.

The StellaCamEX CCD camera is sensitive through 1100 nm and is capable of daytime observation with the use of a near IR filter (700 nm – 900 nm). The higher resolution lens has a focal length of 1018.6 mm and an F-number of 20.4, giving a point source resolution of 4.53 arcseconds. Each pixel in the CCD has a sky-side angular size of 1.7 arcseconds by 2.0 arcseconds giving a total field of view of 23.6 arcminutes by 19.6 arcminutes.
Figure 3.12: The star camera mounted to the telescope during assembly in Chile, with me in the foreground. The star camera is attached to the forward most component of the telescope boom through two extension pieces of aluminum. The primary and secondary mirrors can be seen in the background.

The star camera mounts to the boom of the telescope through two extended pieces of aluminum as shown in Figure 3.12 (this picture was taken in Chile during telescope assembly and not in Cedar Flat). The star camera is mounted at an inclination that matches the POLARBEAR optics through two machined aluminum braces on the back as shown in Figure 3.11(d).

Observations

Preliminary optical pointing observations using the short focal length lens were performed on June 22nd, 2010. First, the star camera was used to find the azimuth encoder zero, the dominant source of pointing error in Cedar Flat. Multiple stars were tracked at different elevations and the offsets in azimuth were
manually adjusted until the stars were all centered. A fit for the measured offsets versus elevation yielded the az encoder zero (IA) term. Using the same lens, the star camera was run on a looped schedule observing approximately eight bright stars repeatedly for several hours of the night. The telescope was slewed from star to star at approximate one minute intervals, and at each new location data was acquired while the telescope remained motionless. Approximately ten seconds of data were taken for each source at one image per second. A total of 127 scans were accumulated in a 70° range of elevation.

Data was taken on two more nights, June 27, 2010 and August 28, 2010, using the same data acquisition strategy on but with the long focal length lens for increased resolution and a more accurate pointing reconstruction. On July 12, in between these two scans sets, the star camera mount was improved by bolting the CCD camera in-place. This fixed the optics of the camera much more rigidly in place; less shaking was observed as the telescope moved.

### 3.3.2 Data Analysis

Star camera data is analyzed by first fetching the relevant scans from the observation table and then downloading and unpacking the corresponding archive files. Because the star camera operates independently of the focal plane detectors, only the az and el timestream registers need to be unpacked. The location of the star is then found in the image by fitting a two-dimensional Gaussian about the maximum signal location. Images with a signal-to-noise ratio of less than ten are cut from analysis. It is important that a later image of the ten to twelve images taken for each source be used for analysis or that a number of images are co-added together. This reduces any jitter induced by the motion of the telescope that would otherwise be mistaken for pointing error. The motion induced jitter is shown in Figure 3.13 and is on the order of a few arcseconds.

The RA and Dec of the source are loaded from a published source catalog and converted to az and el coordinates based on the time of observation. The az and el telescope encoder times differ slightly from the recorded star camera times by a constant offset, so the measured location of the star is found using the star
camera observation time with the known offset programmed in. The difference in the measured location of each star from the nominal source catalog position is recorded. Once all data has been accumulated, the azimuth and elevation offsets are fit to the pointing model using a linear least-squares regression that minimizes the total rms error. Outlying points that deviate from the pointing fit by more than three sigma are cut from analysis. Both pointing parameters and estimated error bars are stored for the fit to the data points that remain. The difference between the pointing model and the measured offsets is used to quantify the rms az and rms el residuals.

**Figure 3.13**: Star camera data taken for a single source. (a) Shows an optical image taken on Jupiter by the long focal length CCD. The black dot is the measured signal maximum and the Gaussian interpolation box is drawn-in. (b) The azimuth and elevation offsets (blue circles) for twelve sequential scans of the north star. The motion of the telescope can still be observed on the first and last few points. The points in between exhibit jitter on the order of a few arcseconds. The black +’s are an applied pointing model that can be used to quantify the error resultant from the jitter.

Figure 3.14(a) shows the pointing model fits to the star camera data taken on June 27, 2010. The azimuth encoder zero term was found to be more than two degrees and is the dominant source of the offsets. It has been subtracted from the fits for clarity. The az and el offsets are each plotted against azimuth and elevation.
to illustrate the model dependence on the sky coordinates. Observed data points are represented by blue circles and the pointing model fits are represented by black +’s. The full eleven parameter model was used to fit this data. The tilt parameters (AN and AW) of the HTT were constrained to two-tenths of an arcminute, though the parameters NPAE, TF, and TFS were found to be higher than expected from telescope construction at several arcminutes each.

The rms offsets were found to be 8.8 arcseconds in azimuth and 13.9 arcseconds in elevation across the entire data set, showing the pointing error of the HTT can be constrained to a fraction of an arcminute across the observable sky. The residuals of the measured offsets and pointing model are shown in Figure 3.14(b). The residuals show no obvious systematics that the model fails to take into account.

The error on the star positions can be quantified by looking at Polaris, which is observed many times throughout the night at the same fixed location of azimuth = 0° and elevation = 37°. The “walking” that the north star exhibits has an rms of approximately four arcseconds in elevation and 5 arcseconds in azimuth. This is the lowest possible limit on the pointing rms, the repeatability at which a single source can be observed at the same position but at different times.

Figure 3.14: Fits and residuals for the star camera pointing model derived from the data taken on June 27th, 2010. The model uses all 11 pointing parameters and gives an rms error of 8.8 and 13.9 arcseconds in azimuth and elevation respectively. The residuals show no obvious systematics that the model fails to take into account.

After the camera mount was secured in July 2010, the August 2010 data
exhibited a seven arcsecond decrease in elevation rms offset with a consistent rms offset in azimuth. Both elevation flexure terms (TF and TFS) and NPAE were found to decrease by over an arcminute, closer to the expected parameters for the HTT. Polaris exhibited the same amount of walking as it had in the June data.

### 3.3.3 Radio Pointing

**Observations**

Radio pointing scans were taken using a single pixel in the POLARBEAR receiver for five separate observations in mid-June 2010. Each observation tracked a bright planet – Jupiter, Saturn, or Venus – across the sky from the time it rose above elevation 25° to the time that it dropped back below this value. Only a single row of bolometers was active during these scans. Multiple raster scans of the planet were taken during its course across the sky, each of which took approximately 20 minutes. From the time the planet rises to the time it sets, 20 – 30 data points were typically accumulated. The telescope tracks the planet’s motion as it performs individual raster scans. Without planet tracking, the source drifts through the field too quickly to map it with the required arcsecond precision.

**Data Analysis**

Like optical pointing, radio pointing also begins by fetching the relevant scans from the archive table and downloading and unpacking the corresponding archive files. The azimuth, elevation, and active bolometer timestreams are all unpacked. The raw data is shown in Figure 3.15 for a single Saturn observation. The telescope (a) tracks in elevation while (b) making several back and forth rasters (half-scans) in azimuth that are also tracking planet location. The (c) bolometer response exhibits high atmospheric noise and spikes in the negative direction when the bolometer “sees” the planet.

The RA and Dec coordinates of each planet are stored at ten minute intervals in a file that is generated by JPL Horizons. RA/Dec is used instead of az/el because the planet moves much more slowly in RA/Dec space than in az/el space;
the time interval required for an accurate reconstruction using az/el coordinates is fractions of a second and generating/parsing such a file is time consuming and unnecessary. The RA and Dec coordinates of the planet are transformed to azimuth and elevation based on the telescope observing location and time. For each ten minute interval, the planet is considered a fixed source in the sky and the location is determined strictly by sky rotation and other astronomical aberrations. In practice, this ten minute approximation makes very little difference in the planet position; the measured change differs by an arcsecond at most. The map coordinates are converted from azimuth and elevation to azimuth and elevation offset space by subtracting the known planet location at each step and accounting for the focal plane position of each bolometer. The conversion allows the offsets from the nominal source positions to be found directly from the peak values in the map.

Figure 3.15: Raw (a) azimuth, (b) elevation, and (c) bolometer timestreams for a single Saturn scan. The atmospheric noise and planet signal can both be seen in the bolometer timestream.

After converting to offset space, the azimuth offset, elevation offset, and timestream registers are trimmed to constant elevation portions of the scans and
each half-scan is polynomial filtered to remove the atmospheric signal. The planets are not masked during filtration, as the planet position is not yet known and the signal is sufficiently bright to remain without masking. A time domain fit for the 3.8 arcminute Gaussian beam is performed and the output azimuth, azimuth offset, elevation, and elevation offset are determined from the mean of the Gaussian fit. If the Gaussian fit deviates from the nominal POLARBEAR beam size by more than 25% or the signal-to-noise ratio is less than five, the data point is cut from analysis. Finally, the atmosphere cleaned map is binned and the centroid is marked for evaluation by eye. Figure 3.16 shows the end product of the timestream analysis. The horizontal dark blue streak at the planet elevation is a filtering effect; for the half-scan where the planet is present, the unmasked filtering over subtracts from the background because the polynomial fits include the strong planet signal. Note that a signal-to-noise of greater than 100 has been achieved for this scan, which is a typical result for planet sources.

![Radio Pointing Fit: saturn](image)

**Figure 3.16**: Radio pointing beam map and pointing center fit for a single Saturn observation. The map is trimmed from the full scan size to better illustrate the pointing offset and beam size.

After all the planet scans have been analyzed, the radio pointing fit is performed utilizing the same code developed for the star camera. Like star camera
analysis, radio pointing eliminates data points that are more than three standard deviations from the measured fits before computing the final pointing parameters. Figure 3.17(a) shows the radio pointing fit for a single Saturn scan set from rise until set on June 15, 2010. Figure 3.17(b) displays the residuals of the same observation. The motion of the planet across the sky can be seen in the az and el positions in the plots. The fit utilizes six free parameters: IA, IE, CA, AN, AW, and TF. The rms azimuth for this scan is 23.4 arcseconds and the rms elevation is 4.8 arcseconds. Obviously, it is extremely undesirable to have such a high and unbalanced offset in az and el because of their combined effects on beam shape. However, the residuals exhibit interesting behavior in the azimuth offset; points that are sequential in time show the planet “jumping” back and forth across the best fit pointing model. Although this could be explained by random telescope jitter, the magnitude of the change between jumps is approximately 40 arcseconds, much higher than any random motion seen in elevation, and so the jitter explanation was highly suspect. Further, the star camera displayed no such jitter, and so the motion would have to be confined locally to the focal plane or receiver as opposed to the telescope itself. The elevation residuals also exhibited systematic effects to a much lesser degree. The same jumping phenomenon can be seen in the elevation offset versus elevation plot and the elevation offset versus azimuth residuals exhibit a sine wave like pattern. The pattern exhibited in these azimuth and elevation residuals was consistent across every planet observation, independent of source and location on the sky.

The residual problem was found to stem from the uncertainty in the Gaussian fitting routing used to locate the beam centers. The fitting routine locates the beam center by minimizing the chi-squared of a Gaussian distribution fit to the timestream with five free parameters: center az, center el, beam width in az, beam width in el, and amplitude. The chi-squared values form a complicated distributed in phase space that is strongly dependent upon both azimuth and elevation centers. If the fitting routine finds a local minima that meets the termination criteria it stops, but there is no way of determining if this is the absolute minimum in the parameter space. The results of the fitting routine were found to be strongly
Figure 3.17: (a) Radio pointing fit for a single Saturn scan set from rising till setting and (b) the corresponding residuals. The residuals exhibit systematic behavior where the planet “jumps” back and forth across the pointing model fit.

dependent on the initial hypothesis for the az and el position of the center of the Gaussian. Due to the scan strategy and binning, two local minima in azimuth were found to exist approximately 40 arcseconds apart that both yielded convergent results for the Gaussian fitting routine. This result explains the jumping phenomenon seen in the residuals.

To combat this problem, iterative pointing was employed. The entire radio pointing data set was fit to derive a set of initial pointing parameters, then and these parameters were used to calculate expected locations for the beam centers. Because systematic offsets are known to exist in the data, the pointing model is computed initially using only five parameters – IA, IE, CA, AN, and AW – to avoid overfitting the incorrect data. The expected locations were then used to seed the Gaussian fits, and the new fits were used to find the final set of pointing parameters. The results are shown in Figure 3.18. Eight parameters were used in the final fit – IA, IE, AN, AW, CA, NPAE, TF, and TFS – to fully characterize the telescope pointing. The large and obvious trends in the azimuth and elevation residuals have been removed. The systematic residual error that appears to remain in elevation offset is thought to stem from the difference in pointing data taken during the day and at night, when differential thermal contraction and expansion of the telescope might play a role. Unfortunately, the data set was too small to confirm or deny this hypothesis with any certainty; only one of the five scan sets taken was at night, so repeatability could not be verified even if the individual data sets had
been larger. Using iterative pointing, the rms residual in azimuth was reduced to 11.1 arcseconds and the rms residual in elevation was 14.3 arcseconds. Note that the elevation rms quoted previously was for a single set of Saturn observations, whereas the data used for this fit accumulates all the radio pointing scans taken during the Cedar Flat engineering run. The parameters CA, AN, AW, and NPAE were found to agree with the optical pointing to within the quantified $1\sigma$ error bars on the fit parameters generated by the least squares algorithm.

![Figure 3.18](image)

**Figure 3.18**: (a) Radio pointing fit using the second iteration of the beam fits for all the pointing data taken in Cedar Flat and (b) the resultant residuals. The residuals are vastly improved from those shown in Figure 3.17. The rms error in azimuth and elevation are 11.1 and 14.3 arcseconds respectively.

### 3.4 Chile Results

#### 3.4.1 Observations

Pointing performed in Chile was fundamentally different from the pointing that was performed in Cedar Flat. Due to the heavy optical loading of the Cedar Flat environment and the goals of the engineering run, no CMB data were taken. This left all available telescope observation time for calibrators and instrument characterization; taking pointing data was one of the highest priorities. In Chile, however, the highest priority is obviously given to CMB observations. The long
radio pointing scans that tracked planets across the sky require hours of continuous, dedicated telescope time. Further, planet sources were often available only when POLARBEAR’s CMB patches were in an observable position. To combat this issue, radio pointing scans were modified to match the star camera strategy from Cedar Flat. A list of approximately 40 bright mm-wave sources were identified in a range of RA and Dec from existing source catalogs so that proper sky coverage could be achieved. During times when no CMB patches were observable (eg - all CMB patches were below el of 30°), the telescope controls would parse the list of pointing sources, prioritized in order of flux, and raster scan sequentially across each that was available. This continued until either the list was exhausted, other calibration data was necessary, or the CMB patches rose back into an observable position. During pointing scans in Chile, all bolometers are biased and operational to within normal yield constraints. In addition to these dedicated radio pointing scans, polarization, gain, and pixel offset calibration data taken on fixed sources (in RA/Dec), such as RCW38 and TauA, were included in the list of usable pointing data. Although the star camera was mounted on the telescope and operational, it has not been used reliably to quantify the pointing error during the Chilean deployment.

3.4.2 Analysis and Results

The pointing scans designed for Chile raster across fixed sources that have much lower signals than those of the planets. In order to improve the yield and signal-to-noise on these measurements, the 35 bolometers closest to the array center were co-added in the map domain using the previously measured pixel offsets and then converted back into a single timestream for Gaussian fitting. It is important that only pixels close to the boresight center are used to co-add the maps. A pointing model is required to accurately shift the sky side positions of the bolometers but, since it is the goal of pointing to find these parameters, no pointing model can be applied when the data is unpacked. For pixels close to the center of the focal plane, however, these pointing induced shifts have a minimal effect. To completely decouple any errors from boresight pointing and beam offsets, both are
computed relative to the center pixel of the array. For further improvements on
signal-to-noise and yield, each bolometer signal is given an eight arcminute radial
mask about the predicted beam location before atmospheric polynomial subtrac-
tion is performed. The change in the result for a single pixel scanning over RCW38
in October 2012 can be seen in Figure 3.19.

Figure 3.19: Maps illustrating the effects of masking and co-adding on radio
pointing observations using RCW38 data from October 2012. The color scale has
units of volts. (a) Unmasked data from a single pixel exhibits residual striping
at the source elevation and a scarcity of binned data. (b) Signal-to-noise is in-
creased and the striping is removed by masking the data with an eight arcminute
radial mask about the expected source position. (c) The residual striping can be
seen clearly in the unmasked map created by co-adding all pointing bolometer
timestreams. This effect is not present in (d) when the mask has been employed.

POLARBEAR has been taking science quality CMB data since late April
2012, and so it is important that a reliable pointing model with low pointing
rms is found for these dates. During this time, only one change to the telescope
that would affect the pointing parameters took place. On June 20, 2012 the field
team re-zeroed the azimuth encoder. It is worth noting that in the Cedar Flat deployment the azimuth encoder zero was set very imprecisely, but in Chile the zero was set with a precision compass and a sky source. This allowed for a much more highly calibrated zero point, on the order of half a degree or less. Although any arbitrary azimuth encoder offset can be fit in the pointing model (as was shown in Cedar Flat), a lower value helps to break pointing model parameter degeneracies and can be quantified far more accurately.

One would expect that if the telescope was mechanically stable over the time period of science observations, then the pointing model would not change except for the azimuth encoder zero parameter on June 20. To assess this mechanical stability and derive the pointing model parameters and their uncertainty, pointing and/or calibration scans on bright, fixed sources were taken every two to three days during this time. All scans from June 20th until the present were subjected to a first iteration of radio pointing analysis. Figure 3.20 shows the results of the preliminary fits. The pointing model fits five free parameters—AN, AW, IA, IE, and CA—using 235 scan sets on 19 fixed sources. The rms in azimuth was 15.8 arcseconds and the rms in elevation was 19.4 arcseconds.

Figure 3.20: (a) The pointing model fits and (b) residuals for the first iteration of radio pointing data from June 20th until the present in the Chilean deployment of POLARBEAR. Five parameters—AN, AW, IA, IE, and CA—have been used in this pointing model.

The residuals show a surprising amount of movement for RCW38, a fixed
source in RA/Dec that was repeatedly scanned at elevation of 65° and azimuths of 165° and 190°. This magnitude and direction of this movement were found to be random over time, but the mean offset of the jitter was found to vary as a function of azimuth. To a certain degree, this is modeled by the tilt terms in telescope. Repeated scans of other fixed sources, such as TauA, at lower elevations (el ≈ 42° and az ≈ 20°) exhibit far less jitter. It appears that the precision to which the telescope can map a single source is an elevation dependent phenomenon.

The parameters from the first iteration of pointing were then used to seed the fits for the second iteration. Seeding the fits actually removed many improper beam fits that occurred in the first iteration; with better initial guesses for the source’s position, many false detections of sources could be ruled out. The second iteration fit a pointing model with six free parameters – AN, AW, IA, IE, CA, and TFS – utilizing 212 scan sets on 13 fixed sources. The derived parameters were found to deviate from the first iteration only incrementally, with three of the five previously fit parameters varying by less than 1%, and the other two by less than 3%. The added flexure term fit a small value of 0.15 arcminutes. These incremental changes however, coupled to the removal of bad beam fits, reduced the pointing error to 10.5 arcseconds in azimuth and 16.8 arcseconds in elevation, a gain of approximately five and three arcseconds respectively. Figure 3.21 shows (a) the pointing fits for the second iteration, (b) the residuals as a function of azimuth and elevation, and (c) the sky coverage plot with vector residuals at each source location. From the vector plot shown in Figure 3.21(c), it is obvious that there are no azimuth or elevation dependent systematics that the pointing model fails to take into account. The scatter exhibited at a single point in both direction and magnitude indicates random error that cannot be further reduced. If systematic effects are hiding, they obviously do not contribute much to the total pointing rms.

To quantify the true random error of the pointing model, histograms and time dependent offsets were plotted in both azimuth and elevation. The results are shown in Figure 3.22. The histograms follow a Gaussian distribution defined by the mean and standard deviation (equal to the rms) of the pointing data. The expected Gaussian distributions are shown in red in Figure 3.22(a). In both
Figure 3.21: (a) The pointing model fits, (b) residuals, and (c) vector residuals in az/el space for the second iteration of radio pointing data from June 20th until the present in the Chilean deployment of POLARBEAR. Six parameters – AN, AW, IA, IE, TFS, and CA – have been used in this pointing model.

azimuth and elevation, the mean is consistent with zero to within 0.001% of the standard deviation. The residuals plotted as a function of time (in Modified Julian Date) are shown in Figure 3.22(b). The residuals show random scatter in both azimuth and elevation offsets as a function of time. The offsets were fit for a constant drift in time using a linear equation whose parameters were found via a least-squares algorithm. The quantified drift, given by the slope of the line, was found to be less than 0.01 arcseconds per day in both coordinates. These values were far less than the measured uncertainties as quantified from the scatter in the data. This clearly illustrates the stability of the HTT over time. Over 300 days, the HTT is found not to drift in pointing offsets in either azimuth or elevation, and has a stable global pointing model constrained to less than 17 arcseconds in
both coordinates.

![Figure 3.22](image)

**Figure 3.22**: Pointing error systematics in azimuth and elevation from June 20 to present date investigated via (a) histograms of the residuals and (b) scatter plots of the residuals as a function of time. The histograms exhibit a Gaussian like distribution with mean zero and the scatter plots display no obvious systematics that would indicate changes in the telescope parameters over time.

A separate pointing model is required for the time period before the azimuth encoder was re-zeroed, but this period has a data set with only eleven scan sets as shown in Figure 3.23(a). These scan sets have acceptable sky coverage to constrain the azimuth encoder zero, which is a function of elevation, but do not have enough points to constrain the other parameters of the model simultaneously. Fortunately, since the parameters describing the telescope appear to be stable over time, the eleven scan sets could be fit for *only* the azimuth encoder zero while holding the other parameters constant at the values obtained from June 20, 2012 onward. The results are shown in Figure 3.23(b). The fits give an rms azimuth offset of 21 arcseconds and an rms elevation offset of 29 arcseconds. It should be noted, however, that these rms values are not statistically significant because of the small size of the dataset. The azimuth encoder zero value was found to differ from the re-aligned value by twelve arcminutes.

Using both pointing models, science data from late April onward can be unpacked with the proper pointing parameters and co-added across multiple data sets. To evaluate the accuracy of the model, planet scans are co-added over this
Figure 3.23: (a) Source coverage for scans taken from late April until June 20th before the azimuth encoder was re-zeroed. (b) The pointing model fits to the data using the model parameters from June 20th onward but allowing the azimuth encoder zero to vary.

entire time period using the derived pointing parameters and over all the pixels in the focal plane using the known beam offsets. The results are shown in Figure 3.24. An analysis of the beam size gives a Gaussian beam with 3.5 arcminute FWHM. This value confirms that the pointing model is accurate to a fraction of an arcminute error across the sky and during this entire time period.

A new student on the collaboration from the University of Colorado at Boulder will also validate the pointing model by locating resolved point sources in POLARBEARs CMB patches and correlating these against the Planck source catalog. This test will confirm that the pointing model does not induce any constant offset shifts or rotations to our CMB fields. After this is completed, the known locations of the point sources will be used to analyze the beam size on resolved sources within our CMB patches as was done for the planets. This will be the final confirmation of the pointing error uncertainties before publication of POLARBEAR’s initial dataset, anticipated for August 2013.
Figure 3.24: Coadded map of Jupiter for all pixels and all observations taken from late April 2012 until April 1, 2013. The fits to the beam profile give a 3.5 arcminute Gaussian beam. Figure courtesy of Zigmund Kermish.

3.5 Acknowledgements

The introduction to Chapter 3 is a reprint of material as it appears in B. Keating, S. Moyerman, et. al., Ultra High Energy Cosmology with POLAR-BEAR, DPF 2011 conference proceedings, arXiv 1110.2101v1 (2011). The dissertation/thesis author was the primary author of this paper.
Chapter 4

Detector Technology

“It’s hotter than heck outside. That’s why we got the refrigerator open, dummy!” -My Name Is Earl

As discussed in previous sections, the current and near future science targets - particularly CMB polarization anisotropies - have signal amplitudes so small that large arrays of background limited detectors are required for a chance of detection. This chapter describes my work in detector design, development, and fabrication at UCSD. Two main technologies are discussed: Normal metal- Insulator- Superconductor (NIS) tunnel junctions and superconducting transition edge sensor (TES) bolometers. The first two sections describes the theory of each and motivates future design possibilities, the third section describes my fabrication work on specific devices, and the fourth section describes recent work in cryogenics done as a collaboration with NIST Boulder.

4.1 NIS Junction Background

4.1.1 Superconductivity

In 1911, Heike Kamerlingh-Onnes first observed superconductivity after dipping mercury in liquid helium and watching the resistance drop to an immeasurably small value [100]. It wasn’t until 46 years later when Bardeen, Cooper and Schrieffer were able to pin down a successful microscopic theory explaining
low temperature superconductivity, coined BCS theory in their honor [101].

In BCS theory, superconductivity occurs when two electrons bind together in boson-like “Cooper” pairs, acting as one particle. The energy that binds the Cooper pairs is separated from the conduction band by a small energy gap ($\Delta$). At sufficiently low temperatures ($k_B T \ll \Delta$), no excited states exist with energy less than $\Delta$, preventing the lattice phonons from scattering the Cooper pairs. Simply put, the binding energy associated with Cooper pairing inhibits electron scattering, allowing current to flow with zero resistance through a material.

BCS theory predicts that, at absolute zero, all superconducting electrons exist as Cooper pairs. At finite temperature, however, electrons exist have a density of states given by

$$ N(E) = N(0)\nu_0(E) $$  \hspace{1cm} (4.1)

where $N(0)$ is the density of states at the Fermi level in the same material above the superconducting transition temperature, $T_c$, and $\nu_0(E)$ is a scale factor describing how the normal density of states scales with energy. The function $\nu_0(E)$ is given by

$$ \nu_0(E) = \frac{E}{\sqrt{E^2 - \Delta^2}} $$  \hspace{1cm} (4.2)

which takes on a value of zero when the energy is less than the gap edge. Electrons are prohibited from occupying these energy states. This function is plotted in Figure 4.1.

The superconducting gap energy is not constant but rather varies with temperature as shown in Figure 4.2. At 0 K, the gap $\Delta_0$, is described by $\Delta_0 \approx 1.76k_B T_c$ where $k_b$ is the Boltzmann constant [102]. The temperature dependence is minimal at $T < T_c/2$. For example, the value of $\Delta$ at $T = T_c/3$ is given by $0.994\Delta_0$. As temperature approaches the superconducting transition temperature, however, the energy gap decreases asymptotically as

$$ \Delta(T) = \Delta_0 \sqrt{1 - \frac{T}{T_c}} $$  \hspace{1cm} (4.3)
Figure 4.1: Normalized density of states function, \( \nu(E) \), plotted as a function of energy. When the magnitude of the energy is less than the superconducting band gap the density of states goes to zero, when the energy equals the band gap a singularity occurs, and when the magnitude of the energy is higher than the gap the values asymptote back to zero.

4.1.2 Corrections to BCS Theory

In practice, measured superconducting materials often deviate from standard BCS theory. These deviations can be modeled by adding what is known as the Dynes parameter, \( \gamma \), to the density of states as \[ \nu_0(E) = \frac{E/\Delta - i\gamma}{\sqrt{(E/\Delta - i\gamma)^2 - 1}}. \] The addition of this parameter allows for non-zero states within the gap region and effectively smears the gap edge singularity without changing the total number of occupied states. For this reason, the change to the density of states is known as Dynes Smearing. The modified density of states is shown in Figure 4.3 for several different values of \( \gamma \). Typical values of \( \gamma \) in sputtered superconducting Aluminum films are 1/5000 \[105\]. Note that in the case of \( \gamma = 0 \), Equation 4.2 is recovered.

In practice, there are many factors that can cause non-ideal behavior that results in finite \( \gamma \), such as the proximity effect, environmentally assisted tunneling,
Figure 4.2: Data for several different superconductors fit to the model for the superconducting band gap as a function of temperature. For every superconductor, the gap drops dramatically to zero as you approach the critical temperature. Image taken from [103].

and Andreev currents. Although most of these effects do not actually alter the superconducting density of states (in fact, of these examples, the only one that does is the proximity effect), they alter the tunneling currents that flow through superconducting materials in a way that is conveniently modeled by Dynes Smearing.

4.1.3 NIS Tunneling Theory

In this section we derive the fundamental equations governing current flow through NIS tunnel junctions.

Historical Context

In 1960, E. Giaever, a research scientist at GE laboratory, first demonstrated superconductive tunneling through an insulating barrier [2]. After having just learned about superconductivity and the superconducting energy gap,
Figure 4.3: BCS density of states with the addition of Dynes Smearing for several values of Dynes parameter. Panel (a) has identical axes to Figure 4.1, whereas the y-axis of (b) is plotted in logarithmic units to highlight the effects of smearing.
Figure 4.4: Current-voltage curves for an Al-AlOx-Pb tunnel junction, as measured in [2]. The different lines correspond to IV curves taken at different combinations of temperature and applied magnetic field.
Gaiever’s intent was to measure the gap using superconducting tunnel junction properties. When Gaiever presented this idea to his colleagues, he was met with skepticism for his oversimplified ideas of superconductivity. However, his colleagues encouraged him to pursue his experiments nonetheless, and within a matter of days Gaiever had measured the current-voltage (IV) curves of a lead-aluminum tunnel junction, shown in Figure 4.4. This measurement was the most accurate measurement of the superconducting energy gap and density of states to date - from idea to conception in less than a week’s time. So inexperienced was Gaiever that he became distressed that his measurements did not agree with other contemporary experiments. In his Nobel Prize lecture in 1973, Gaiever stated that:

I was, of course, not the first person to measure the energy gap in a superconductor, and I soon became aware of the nice experiments done by M. Tinkham and his students using infrared transmission. I can remember that I was worried that the size of the gap that I measured did not quite agree with those previous measurements. Bean set me straight with words to the effect that from then on other people would have to agree with me; my experiment would set the standard, and I felt pleased and like a physicist for the first time.

Shortly after his discovery of superconductive tunneling, Gaiever and his colleague Fisher also measured the first metal-aluminum oxide-metal tunnel junctions [106] and superconducting-insulating-superconducting junctions [107]. To date, tunneling is still the best method for measuring the density of states in a superconductor.

**NIS Current Flow**

A simple schematic of a NIS junction is given in Figure 4.5. Two metals, labeled 1 and 2, are separated by a thin, insulating barrier such as Al$_2$O$_3$ or vacuum gap. As a voltage bias is applied to the junction, current flow occurs via quantum mechanical tunneling. The system can be modeled effectively as electrons tunneling across a potential barrier as in elementary quantum mechanics. The scale of the energy barrier is set by the properties of the insulator, such as shape, though varying these parameters has essentially negligible effect for the small voltage biases used in NIS junctions [108].
Figure 4.5: Schematic of a tunnel junction; two metals separated by an insulating barrier. When a voltage, $V$, is applied to the closed loop containing the junction, a current, $I$, results.

To understand the tunneling current that results from the application of a voltage bias, we begin our treatment by considering the electron energy level occupation within the two metals. For any metal, the electron energy level population is determined by the Fermi function

$$f_x(E) = \frac{1}{1 + e^{E/(k_B T_x)}},$$  \hspace{1cm} (4.5)$$

where $T_x$ is the temperature of metal $x$ and $E$ is the excitation energy relative to the Fermi level [109]. A voltage bias $V_b$ applied to an NIS tunnel junction increases the Fermi level of the normal metal by a factor of the electron charge, $e$, times the bias voltage relative to the Fermi level of the superconductor. Tunneling across the barrier occurs when the wave function of an excited electron moves between the normal metal and the superconductor with no change in energy. Therefore, defining the electron excitation in the superconductor as energy $E$, the equivalent normal metal electron has an intrinsic energy $E - eV_b$.

We can now determine the tunneling rate of electrons/holes across the tunnel barrier. The tunneling rate from the normal metal to the superconductor is proportional to the number of occupied states at energy $E - eV_b$ in the normal metal with corresponding unoccupied states at energy $E$ in the superconductor. The number of occupied states in the normal metal is determined by multiplying
the density of states in the normal metal, \( N(E - eV_b) \), by the Fermi function, \( f_N(E - eV_b) \). For a normal metal the density of states is slowly varying and the function \( N(E - eV_b) \) can be well approximated as a constant, \( N_N(0) \). The number of unoccupied states in the superconductor is given similarly by the product of the density of states in the superconductor, \( N(E) \), and the complement to the Fermi function, \( 1 - f_S(E) \). Here, the subscripts \( N \) and \( S \) on the Fermi functions represent the normal metal at temperature \( T_N \) and superconductor at temperature \( T_S \), respectively. For simplicity, we define \( \Gamma_{NS} \) as the tunneling rate from normal metal to superconductor

\[
\Gamma_{NS} = P N_N(0)f_N(E - eV_b)N(E)[1 - f_S(E)],
\]

(4.6)

where \( P \) is the normalization factor.

In order to calculate the total current, we must also calculate the rate of tunneling from the superconductor to the normal metal. Following the formalism above, this is given by

\[
\Gamma_{SN} = P N_N(0)[1 - f_N(E - eV_b)]N(E)f_S(E).
\]

(4.7)

The net tunneling rate is then found by integrating the difference of Equation 4.6 and Equation 4.7 over all energies. Multiplying this value by the charge of the electron gives the tunneling current, \( I_{NIS} \) as

\[
I_{NIS} = \int_{-\infty}^{\infty} e(\Gamma_{NS} - \Gamma_{SN})dE
\]

\[
= ePN_N(0)\int_{-\infty}^{\infty} N(E)[f_N(E - eV_b) - f_S(E)]dE.
\]

(4.8)

Substituting for \( N(E) \) using Equation 4.1 gives

\[
I_{NIS} = ePN_N(0)N_S(0)\int_{-\infty}^{\infty} \nu_0(E)[f_N(E - eV_b) - f_S(E)]dE.
\]

(4.9)

(4.10)

The prefactor in Equation 4.10 can be determined by imposing the condition that above the superconducting temperature, the NIS junction behaves as a normal-insulating-normal (NIN) junction, which is a standard resistor. Using Ohm’s law, we apply the constraint that the product of the current and the normal state resistance, \( R_n \), should equal the voltage bias. Solving the integral in this case gives

\[
P N_N(0)N_S(0) = 1/(e^2 R_n).\]
Figure 4.6: Body of IV curves for a NIS junction with a normal state resistance of 10 Ω and a gap energy of $\Delta = 180 \mu$eV. Panel (a) shows the full range of voltage biased behaviors of the junction, from the nonlinear (low bias) out into the linear (high bias) regime. Panel (b) is a zoom in of the diode-like part of the curve to highlight the dependence on temperature and the nonlinear behavior.
It is now constructive to convert Equation 4.10 into an integral over only positive energies, as the dependence on the superconducting temperature will be removed from the equation. Making the appropriate conversions to only positive energies gives

\[ I_{NIS} = \frac{1}{e^2 R_n} \int_{0}^{\infty} \nu_0(E)[f_N(E - eV_b) - f_S(E)] \\
+ \nu_0(-E)[f_N(-E - eV_b) - f_S(-E)]dE \]

and substituting the properties \( \nu_0(E) = \nu_0(-E) \) and \( f(-E) = 1 - f(E) \) yields

\[ I_{NIS} = \frac{1}{e^2 R_n} \int_{0}^{\infty} \nu_0(E)f_N(E - eV_b) - f_S(E)] \\
+ \nu_0(E)[1 - f_N(E + eV_b) - (1 - f_S(E))]dE. \]

Finally, we expand the terms and simplify to arrive at

\[ I_{NIS} = \frac{1}{e^2 R_n} \int_{0}^{\infty} \nu_0(E)[f_N(E - eV_b) - f_N(E + eV_b)]dE, \quad (4.11) \]

which clearly illustrates that the current response of an NIS junction to an applied voltage is a unique function of the normal metal temperature alone. It is this property that makes NIS junctions very simple thermometers; one achieves a different current-voltage (IV) curve at each temperature below the superconducting band gap. Equation 4.11 must be integrated numerically to derive full IV curves. A body of such IV curves is given in Figure 4.6.

Although Figure 4.6 was derived through numerical integral of Equation 4.11, it can be physically motivated by looking at the energy band diagrams of an NIS junction. In fact, energy band diagrams are such a vital tool for understanding physical phenomenon that Herbert Kroemer, in his Normal Prize lecture \[110\] stated his Lemma of Proven Ignorance:

If, in discussing a semiconductor problem, you cannot draw an Energy Band Diagram, this shows that you don’t know what you are talking about, with the corollary - If you don’t draw one, your audience won’t know what you are talking about.

Figure 4.7 shows the energy band diagrams on an NIS junction at zero temperature with an increase in applied voltage from left to right. The left-most
Figure 4.7: Energy band diagrams of an NIS junction at $T = 0$ with an increase in applied bias voltage from left to right. The blue arrow represents current flow.

Panel shows the energy configurations of the normal metal and superconductor with no bias voltage applied. Although tunneling can occur at energies below the Fermi energy, $E_f$, the tunneling proceeds in both directions and no net current flows. The second panel shows the altered energy diagram when the normal metal is voltage biased with a magnitude equal to $V = \Delta/e$. The increase in energy is given by $E = Ve = \Delta$, which shifts the Fermi level in the normal metal to coincide with the lowest energy states of the superconductor. The electrons are equal to the highest prohibited states, but do not exceed them, and so are still prohibited from tunneling. It is at this point, however, that any increase in the voltage bias will result in a tunneling current. This is shown in the third panel. Finally, as the voltage bias is increased far beyond $\Delta/e$, that tunneling from the normal metal to the superconductor is no longer limited by the number of available states; current flows through the junction like a normal resistor. This is shown in the right-most panel of Figure 4.7.

Putting this together, one can explain the full nature of the IV curves at $T = 0$. As the voltage bias is increased from 0 to the $\Delta/e$, the behavior of the junction remains constant; no current flows ($I = 0$) and the IV curve response is flat. As the voltage bias is increased past $\Delta/e$ the tunneling current begins to flow sharply through the junction. Finally, as the bias is increased far beyond $\Delta/e$, the tunneling rate is sufficiently high that the junction approximates a normal resistor with linear behavior. The curve exhibits symmetric behavior for negative values of the bias. The normal state resistance of the junction, $R_n$, can be determined
by computing the inverse slope of the IV curve at high values of the bias voltage when the junction is in the linear regime.

Figure 4.6 also shows the NIS curve dependence on temperature. At zero K, the current response remains perfectly flat until exactly $V_b = \Delta/e$, at which the transition in behaviors is very sharp. As the temperature is increased, less bias is required for the tunneling current to begin and the transitional behavior in the IV curve becomes rounded. The increase in temperature effectively smears the density of states about the Fermi energy, resulting in some electrons with energy above $E_f + \Delta$ even when $V_b < \Delta/e$. This effect is shown in Figure 4.8. This effect becomes more dramatic with increasing temperature and, additionally, the energy band gap of the superconductor decreases as temperature approaches $T_c$. At 0.8 K $\approx 0.66T_c$ (for Al) the curve appears almost linear, with only slight curvature near the gap energy.

4.1.4 Cooling Power

The energy band diagrams of NIS tunneling show another interesting consequence of their non-linear behavior; because of the quantum mechanical nature of tunneling, only the highest energy, or hottest, electrons from the normal metal are able to tunnel into the superconductor. This process selectively removes energy from the normal metal and displaces it into the superconductor. The amount of energy removed from the normal metal can be quantified by taking the differ-
ence in the tunneling rates multiplied by the energy associated with each electron tunneling event and integrating it over all energies. This is written as

\[
P_N = \int_{-\infty}^{\infty} (E - eV_b)(\Gamma_{NS} - \Gamma_{SN})dE
\]

\[
= \frac{1}{e^2 R_n} \int_{-\infty}^{\infty} (E - eV_b)[f_S(E) - f_N(E - eV_b)]\nu_0(E)dE. \tag{4.12}
\]

Numerical integration of the above expression gives the cooling power at each value of the bias. The result is shown in Figure 4.9(a). When \(P_N < 0\), power is being removed from the normal metal, yielding a remaining electron bath with an effectively colder temperature. Notice that the cooling power of the normal metal increases to positive values above a voltage bias of \(\Delta/e\). For maximum cooling power, the junctions should be biased at \(\approx 0.8\Delta/e\) with small variations based on operational temperature. The optimal operating temperature is one that gives maximum cooling power over all voltage biases and all superconducting temperatures. This is shown in Figure 4.9(b). The maximum cooling power of an NIS junction is achieved when the operational temperature is between \(0.5T_c\) and \(0.6T_c\).

The tunneling power deposited into the superconductor can be computed by

\[
P_S = \int_{-\infty}^{\infty} (E)(\Gamma_{SN} - \Gamma_{NS})dE
\]

\[
= \frac{1}{e^2 R_n} \int_{-\infty}^{\infty} (E)[f_N(E - eV_b) - f_S(E)]\nu_0(E)dE, \tag{4.13}
\]

and conservation of energy dictates that

\[
I_{NIS}V_b = P_N + P_S. \tag{4.14}
\]

The inverse relation between \(R_n\) and \(P_N\) is obvious in Equation 4.12. Since, the normal state resistance varies exponentially with the insulating barrier thickness and scales linearly with the area, one might naively expect that larger area junctions with thinner barriers would supply the most cooling power. However, these modifications experience practical limitations; larger area and thinner barriers have a higher risk of developing pinholes through which non-tunneling current
Figure 4.9: (a) Numerical integration of Equation 4.12 as a function of voltage bias evaluated at several different bath temperatures using $R_n = 10 \, \Omega$ and $\Delta = 180 \, \mu eV$. (b) Optimization of cooling power as a function of bath temperature.
Figure 4.10: NIS junction IV curves for different values of the Dynes parameter. The Dynes parameter is exaggerated beyond typical values to illustrate the effects of subgap leakage. In practice, the Dynes parameter of a working junction is often orders of magnitude lower than those shown here.

can flow. Atomic layer deposition of the insulating barrier has the potential to avoid these pitfalls, and chapter 5 describes the development of a novel fabrication process using atomic layer deposition that has been developed at UCSD. Such a process could significantly reduce nontunneling current flow and allow for much larger, thinner insulating barriers. Future junctions fabricated using this methodology could result in orders of magnitude more cooling power.

Current flow through pinholes in the insulating barrier is known as subgap leakage, because current flows at values of the bias below the gap. Although this is a different physical phenomenon from Dynes Smearing, it is well modeled by this modification to BCS theory. Figure 4.10 shows the IV curves of a junction with a high leakage factor modeled by increases in Dynes parameter.

The quality factor, $Q$, of an NIS junction is generally characterized by the ratio of the dynamic resistance, the inverse slope of the IV curve at $-0.8\Delta/e < V_b < 0.8\Delta/e$, to the normal state resistance; a higher $Q$ value indicates a higher quality NIS tunnel junction. As illustrated in Figure 4.10, if the subgap leakage, modeled by Dynes parameter, is high, the dynamic resistance decreases appreciably.
Figure 4.11: Image of a SINIS structure - two junctions in series sharing the same normal metal. The normal metal is shown in gold with cross-hatching. The insulator is shown in red and the superconducting metal is shown atop the structure in light gray. The material shown in white is an insulator that defines the area of current flow through the tunnel junction; no current flows through these white regions. When a voltage bias is applied as shown, the resulting current must flow through both junction interfaces to complete the loop.

Because Equation 4.12 is symmetric for positive and negative voltage biases, NIS junctions can be stacked in symmetric structures where two junction barriers are coupled to the same normal metal. This forms a SINIS structure as shown in Figure 4.11. The cooling power of such a device is double that of a single NIS junction, as both junctions are now tunneling the highest energy electrons from the same normal metal. The modifications to the theory are minimal; the tunneling behavior in the IV curves now begins at $2\Delta/e$ and the optimal bias occurs at $\approx 1.6\Delta/e$.

Refrigeration

Refrigeration by NIS junctions will only occur if the cooling power of the normal metal exceeds the external power loads. There are four main contributions to the external power load:

- electron-phonon coupling
- subgap heating
normal metal heating

quasiparticle backflow

Electron-phonon coupling is the dominant contribution to the external power load and describes the power exchange between electrons and phonons in the normal metal. Electrons can be heated by either absorbing or scattering phonons from the lattice. The power load is given by

\[ P_{ep} = \Sigma \Omega (T_n^p - T_n^e) \]  

where \( \Sigma = 2.3 \text{ nW/(K}^6\text{m}^3) \) is the electron-phonon coupling constant describing the strength of the electron-phonon interaction, \( T_p \) is the phonon temperature, \( T_e \) is the electron bath temperature, and \( \Omega \) is the volume of the normal metal. The exponent \( n \) is found empirically to vary between 5 and 6 for metals depending upon their purity [111].

Subgap leakage is power lost to normal currents passing through pinholes in the junction barrier. A parallel resistor, \( R_{lk} \) can be used to model the leakage across the interface, dissipating power as normal Joule heating

\[ P_{lk} = I_{NIS}^2 R_{lk}. \]  

Similarly, normal metal heating is power dissipated by joule heating in the normal metal with resistance \( R_{pad} \). It is important to remember that the total current flowing through the normal metal, \( I_{tot} \), is defined by both the leakage current and the tunneling current, giving a power dissipation of

\[ P_J = I_{tot}^2 R_{pad} = (I_{NIS} + \frac{|V_b|}{R_{lk}})^2 R_{pad}. \]

Finally, quasiparticle backflow occurs when quasiparticles in the superconductor recombine and produce phonons that heat the normal metal. The power load can be approximated as a small percentage, \( \beta \), of the superconductor power, Equation 4.13, flowing back into the normal metal [112]

\[ P_{qp} = \beta P_S. \]
A reasonable value of the parameter $\beta$ is 0.05, though this parameter can be lowered with the addition of quasiparticle traps. Traps are normal metal layers outside the current path of the junction that are separated from the superconductor by a thin, insulating barrier. Quasiparticles in the superconductor are transported by random scattering mechanisms and, based on size, shape, etc., some tunnel back into the normal metal where they can recombine. However, if another normal metal layer and tunnel barrier are added in coincidence with the superconductor, the quasiparticles have an equal probability of tunneling and recombining in the trap direction (again, this depends on the size, shape, etc. of the traps). In this way, traps helps to minimize the fraction of the total power that recombines in the normal metal, effectively lowering the $\beta$. Much research has been done to optimize trap design [105] and increase cooling power.

It is often reasonable to approximate the temperature of the phonons in the SINIS structure as a constant equal to the bath temperature. With this approximation in place, one need only solve the power balance equation for the electron bath to arrive at the cooled temperature. This is given by

$$0 = 2P_N + P_{ep} + P_{lk} + P_J + 2P_{qp},$$

which states that the sum of all power loads must equal zero, i.e. - the junction is in equilibrium. The factors of 2 on $P_N$ and $P_{qp}$ arise for a SINIS junction structure, with two NIS junctions both of these power loads are doubled. The above equation can be solved for $T_N$ for a fixed bias and bath temperature, which gives the equilibrium self-cooled temperature of the junction (if $2P_N$ exceeds the loading terms). Experimental device parameters such as $R_{lk}$ and $\beta$ are required to compute the expected base temperature of NIS operation. A reduction in temperature of 200 mK is expected from a well designed junction at a bath temperature of 300 mK and a bias $0.9\Delta$. It is worth noting that the optimal operational bias point is increased slightly when the power loads are considered [113].

Figure 4.12 shows the numerically computed power loads for each contributing factor and their influence on SINIS junction performance. The solid blue line is the computed SINIS junction cooling power at a range of electron bath temperatures and in the absence of applied loads; the solid green line is the net cooling
Figure 4.12: SINIS junction cooling power and the different contributions to the power loading. The parameters used for this model are $R_N = 5.6\,\Omega$, $\beta = 0.02$, $\Sigma = 2.5\,\text{nW}/(\mu\text{m}^3\text{K}^6)$, $\Omega = 6.8\mu\text{m}^3$, $R_{lk} = 15000\,\Omega$, and $R_{pad} = 0.56\,\Omega$. The phonon temperature is assumed to be equal to the bath temperature at 300 mK.

Power when loading is considered. Each of the dashed lines represents a different contribution to the external power load. At different electron bath temperatures, it is obvious that different power loads become the dominant contribution to the total loading. The equilibrium temperature is achieved when the solid green line crosses the x-axis, at approximately 150 mK. Because the volume of the normal metal in a SINIS junction is very small, the heat capacity is sufficiently low at cryogenic temperatures that reduction in temperature occurs almost simultaneously to the application of the bias.

NIS tunneling cools the electron bath in the normal metal, but for NIS refrigeration to be practical it must be capable of cooling both electrons and phonons.
Fortunately, NIS junctions are capable of cooling membranes, suspended structures with only a weak thermal link to the substrate, generally through long skinny legs. The normal metal of the NIS junction is extended out onto the legs of the membrane, surrounding the thermally isolated region with a lower effective temperature. Cooling is possible because the phonons of the suspended membrane are more tightly coupled to the electrons in the NIS junction normal metal than to the phonons of the substrate [114]. Optimizing the shape of the NIS junctions as well as the shape and thermal conductivity of the cold fingers (by plating with Au or another high conductivity material) maximizes the ratio of the electron-phonon coupling between the NIS junctions and the membrane to the phonon-phonon coupling between the membrane and the rest of the substrate. Because the relative coupling is maximized, the loading on the membrane from the substrate is minimal, allowing for more effective cooling of the membrane and any associate power loads, such as those from attached or lithographed devices [114].

Figure 4.13 shows a schematic of NIS junctions coupled to two different types of membranes: (a) a traditional membrane and (b) a suspended brick or channel-etch membrane. The traditional suspended membrane floats like a bridge over the removed substrate below. Individual CMB detectors are often thermally isolated from their environment in this manner, making it seemingly (and deceptively) easy to couple them to NIS junction coolers. Suspended brick membranes are isolated from their substrate by removing channels of material through the entire substrate and leaving the material directly underneath the suspended portion. Because they are very thick, these membranes are stable enough for mounting external samples and offer a viable option for NIS junctions to serve as bulk coolers instead of cooling individual detectors. This will be discussed in greater detail in Section 4.4.

4.2 TES Bolometers

The overarching goal of the UCSD NIS junction project is to cool POLAR-BEAR detectors, increasing their sensitivity by lowering their intrinsic detector
4.2.1 Introduction

As mentioned in Section 3.1.2, many current and future CMB experiments utilize large arrays of superconducting transition edge sensor (TES) bolometers for maximum sensitivity to sky signal. Bolometers are detectors of thermal power that absorb electromagnetic radiation over a wide range of wavelengths. Figure 4.14 shows the simplest model of thermal operation. An absorber is connected by a weak thermal link to a heat sink held at constant temperature $T_b$. As the absorber collects incident photons, their optical power, $P_{\text{opt}}$, is dissipated and raises the temperature, $T$, of the absorber above $T_b$. The change in resistance of a coupled thermistor is used to measure these changes in temperature, corresponding to changes in optical power. If the absorber has heat capacity $C$, the time constant in response to changes in optical power is given by $\tau = C/G$. The higher the heat capacity, the longer it takes the absorber to change in temperature. The higher the thermal conductivity between the absorber and the heat sink, the faster the absorber settles back to temperature $T_b$.

Voltage-biased superconducting TES bolometers utilize a voltage biased superconducting material as the thermistor, operating on the transition from normal metal to superconductor. In the region near the transition temperature, $T_c$, small
Figure 4.14: Simple cartoon schematic of a bolometer. Photons are absorbed by an absorber connected to a heat sink through a weak thermal link. The thermistor measures changes in the temperature of the isolated material. Image taken from http://www3.mpifr-bonn.mpg.de/.

Changes in temperature correspond to marked changes in resistance, making TES bolometers extremely sensitive to small changes in incident power. In order to take full advantage of this steep temperature change, one must be able to operate the bolometer stably at the same point along the superconducting transition. For every small change in incident power, the thermistor must be quickly restored to its base operational point. This process is shown in Figure 4.15.

Such a condition is met by the voltage bias on the TES, which provides negative Electro-Thermal Feedback (ETF) that stabilizes the device. Above the transition temperature, a voltage bias $V_b$ is applied to the superconductor, which dissipates a power $P_{bias} = V_b^2/R$, where $R$ is the resistance of the thermistor. $T_b$ is then lowered below $T_c$, but the joule heating from the voltage bias prevents the thermistor from superconducting. The voltage is then slowly lowered until the thermistor is operating on its transition edge. An increase in optical power increases the total power, $P_{opt} + P_{bias}$, raising the temperature of the thermistor. A rise in temperature results in a steep increase in resistance, causing $P_{bias} = V_b^2/R$ to decrease in response. As a result, the total power and, correspondingly, the temperature are decreased. Thus, for a voltage-biased TES, the increases in to-
Figure 4.15: The superconducting transition of niobium illustrating thermistor operation and ETF. The point of operation is near the center of the transition. As optical power is absorbed the bolometer temperature rises, causing an increase in resistance (y-axis), represented by arrow 1 on the graph. The increase in resistance causes a decrease in the bias power, lowering total power and thus the temperature.

Total power are smaller than the increases in $P_{\text{opt}}$ from incident radiation, allowing these thermistors to operate in the narrow range of their most sensitive temperature responsivity. This property also makes voltage-biased TESs very stable; the electro-thermal feedback fights to restore them to their nominal operating point.

The level of ETF can be quantified through the loop gain, which is a function of the steepness of the superconducting transition. This loop gain is given by

$$L = \frac{P_{\text{bias}} \alpha}{gT},$$

(4.20)

where $g$ is the derivative of the total power with respect to temperature

$$g = \frac{dP_{\text{total}}}{dT},$$

(4.21)

and $\alpha$ describes the steepness of the superconducting transition

$$\alpha = \frac{T}{R} \frac{dR}{dT}.$$

(4.22)

Loop gain effectively sharpens the response time of a TES to incident power. When optical power is deposited the ETF reduces the excess power needing to
be transported to the heat sink. The effective thermal time constant is reduced by a factor of $L + 1$ to $\tau = \frac{C}{G(L+1)}$, making bolometers with a high $\alpha$ into very fast optical power detectors. Strong ETF also linearizes the current responsivity to incident power, making it relatively simple to analyze the output signal of the TES.

### 4.2.2 Bolometer Noise

In order to compare the detector noise properties among various devices, it is necessary to quantify sensitivity that is inherent only to the detector and is not dependent on any specific experimental details (such as the signal strength or detector location). The figure of merit often used is the Noise Equivalent Power (NEP) of the detector. NEP is the optical power required to achieve a signal to noise ratio of one in a one Hertz output bandwidth, which is equivalent to a one-half second of integration time. In terms of the NEP then, the signal to noise is defined as

$$\frac{S}{N} = \frac{P_{signal}}{P_{noise}} = \frac{P_{signal}}{NEP} \sqrt{2t},$$

(4.23)

where $t$ is the observing time of the detector. The NEP is most commonly quoted in spectral density units of Watts per square-root of bandwidth (W/√Hz) or Watts times square-root seconds (W√s). From Equation 4.23, it is obvious that small NEPs give higher signal-to-noise on equivalent timescales, and so correspond to more sensitive detectors.

Many individual sources of detector noise can contribute to total detector NEP. If these noise sources are completely uncorrelated, then the total NEP is calculated by the quadrature sum of each individual component. Typically, for bolometers, there are several sources of noise that need to be considered

$$NEP^2 = NEP_g^2 + NEP_\gamma^2 + NEP_J^2 + NEP_r^2,$$

(4.24)

where the subscripts $g$, $\gamma$, $J$, and $r$ correspond to thermal carrier noise, photon noise, Johnson noise, and readout noise, respectively. Minimizing NET requires understanding the contribution by each of these noise terms.
Thermal Carrier Noise

Any bolometer is ultimately limited in sensitivity by the random fundamental energy fluctuations across the weak thermal link to the heat sink. Such noise is often referred to as thermal carrier noise, because it is the thermal carriers themselves that cause these fluctuations. In the case that the operating temperature of the bolometer, $T$, is equal to the bath temperature, the thermal carrier noise is given by $NEP_g = \sqrt{4kT^2g}$, which is derived by calculating the statistical variance in the energy of the absorber.

However, operational bolometers experience loading from optical sources and bias power, forcing $T > T_b$. In this case, the gradient in temperature across the thermal link gives a modified $NEP_g$ of

$$NEP_g = \sqrt{4\gamma k_B T^2 g},$$

(4.25)

where the additional factor $\gamma$ is given by

$$\gamma = \sqrt{\left(\frac{n+1}{2n+3}\right)\left(\frac{1-(T_b/T)^{2n+3}}{1-(T_b/T)^{n+1}}\right)}$$

(4.26)

and $n$ is the exponential dependence of the thermal conductivity, generally assumed at low temperatures to behave as $g \propto T^n$. For dielectric materials $n = 3$ and for metals $n = 1$. Assuming dielectric suspension (such as the legs in POLARBEAR bolometers), one can minimize the $NET_g$ with respect to the quantity $T/T_b$. For a thermal bath temperature $T_b$, the operating temperature of a TES bolometer with minimal thermal carrier contribution is given by $T = 1.705 T_b$. The operational bath temperature is set by the limitations of the cryogenic system that houses the receiver, so the superconducting transition temperature of the TES thermistor must be carefully tuned to meet these requirements. For example, the base cryogenic temperature of POLARBEAR is approximately 270 mK, so the TES bolometers are tuned for a transition temperature of $T \approx (270)(1.705) \approx 480$ mK to minimize the noise.

From Equation 4.25, there are two options for further reducing the thermal carrier noise in any experiment. The first is by reducing the factor $g$, but since this factor contributes to other bolometer properties such as loop gain and time
constant, it cannot be set arbitrarily. The less restrictive way to reduce the noise is to reduce the operational temperature of the bolometer, which lowers the noise without consequence. This is the motivating factor for lower temperature operation of bolometers, and will be revisited later in this chapter.

**Photon Noise**

Photon noise is noise contributed by variations in the observed source over time. Because photons are quantized, the arrival rate of photons at the detector varies even for a perfect (or almost perfect) blackbody such as the CMB. The \( \text{NEP}_\gamma \) for a single bolometer is given by \[116\]

\[
\text{NEP}_\gamma = \sqrt{2hfP_{\text{opt}}} + \frac{2P_{\text{opt}}^2}{\Delta f},
\]

(4.27)

where \( P_{\text{opt}} = \eta hf \Delta f n_0 \) is the total optical power received at the bolometer, \( n_0 = 1/(e^{hf/(k_B T)} - 1) \) is the Boltzmann occupation number for bosons as a function of frequency, \( \eta \) is the optical efficiency of the system, \( f \) is the center frequency of the detector, and \( \Delta f \) is the bandwidth.

The first term in Equation 4.27 is proportional to \( n_0 \), so when \( n_0 \) is small (low temperature or high frequency) this term dominates the \( \text{NEP}_\gamma \). It describes statistical variances in the arrival rate of photons following a Poisson distribution. The second term goes as \( n_0^2 \) and consequently dominates the expression when \( n_0 \) is high (high temperature or low frequency). At high occupation, photons tend to bunch together as non-classically interacting particles, which correlates their arrival times, increasing this noise term.

**Johnson Noise**

Johnson noise is due to the thermal motion of electrons through a resistor at finite temperature. Any resistor can be modeled as a perfect resistor in parallel with a Johnson current noise source \( I_J = \sqrt{4k_B T/R} \) where \( R \) is the resistance of the resistor. This gives an \( \text{NEP}_J \) of

\[
\text{NEP}_J = \sqrt{4k_B T P_{\text{bias}}},
\]

(4.28)
which is on the order of the thermal carrier noise. Fortunately, ETF suppresses
the Johnson noise by a factor of the loop-gain

\[ NEP_J = \frac{\sqrt{4k_BTP_{bias}}}{L + 1} . \] (4.29)

Thus, for TES detectors with a sharp superconducting transition and correspond-
ingly high loop gain, \( NEP_J \) is suppressed below \( NEP_g \) by several orders of mag-
nitude. For this reason, this factor can be ignored when considering the total
\( NEP \).

**Readout Noise**

Superconducting quantum interference devices (SQUIDs) are often used as
amplifiers to readout the signal from TES bolometers. Commercial SQUIDs have
a noise floor of 0.1 pA/\( \sqrt{Hz} \), which is several orders of magnitude lower than the
thermal noise carrier contribution to \( NEP \). For this reason, the readout term
contribution to the total \( NEP \) can often be ignored as well.

### 4.2.3 Noise Limitations

The two dominant terms remaining in the total \( NEP \) are the thermal carrier
noise and the photon noise. An ideal experiment would be limited by only photon
noise, a condition that is referred to as Background LImited Photometry (BLIP).
Once this condition is met, detector sensitivity can be improved only by increasing
system throughput, i.e. - collecting more photons. This higher throughput is
generally achieved by increasing the number of individual detectors, though other
means, such as improving detector efficiency, would also work.

In terms of an individual detector, what are the parameters required to
meet the BLIP limit? Consider the POLARBEAR detectors as a toy model, with
a center bandwidth of 150 GHz, \( \Delta f =40 \) GHz, and an optical efficiency of 40%.
Due to other design constraints, such as avoiding observing through atmospheric
water lines, these parameters are somewhat fixed and cannot be used freely to
optimize \( NEP \). Additionally, the detector efficiency is set not only by the detectors
themselves but also by the telescope as a whole. Once again, the only truly free parameter left to optimize is the detector temperature.

Assuming the CMB is the single source of optical power for POLARBEAR, one can compute $NEP_\gamma$ and $NEP_g$ at each value of operational temperature. In doing so, it is assumed that at each temperature the $NEP_g$ is reoptimized and the lowest possible value is taken. For BLIP performance, $NEP_g \leq NEP_\gamma$, the condition is found that $T \leq 0.2$ K. For a dielectric suspension, this gives a required bath temperature of 117 mK or lower. It should be noted that actual detectors will also experience emission from components inside the receiver and from the atmosphere if they are ground based. This increases the temperature at which the BLIP limit is achieved since the level of photon noise is higher.

Obviously, POLARBEAR (and a range of other ground based experiments using similar cryogenic systems) are still operating above the BLIP condition, and a reduction in noise can still be achieved by further cooling the detectors. The current upper limit on the tensor-to-scalar ratio, $r$, constrains the $B$-mode signal to an amplitude that has not yet been measured by CMB polarization studies; the closest experiment is still a factor of approximately six above this value [19]. Obviously, increasing detector sensitivity is a must for future studies. Furthermore, since the amplitude of the $B$-mode signal is as-yet unknown (if it exists at all), every push for an increase in instrumental sensitivity is warranted.

Recomputing the NET for POLARBEAR with a base operating temperature of 150 mK shows an increase in array sensitivity of 17%. This is equivalent to increasing the number of detectors by approximately 40%, a change that would increase focal plane area so appreciably that a complete reconfiguration of the telescope optics would be required. Obviously, a lower detector bath temperature provides an easier solution, but the question of how to lower the temperature still remains. While commercial refrigeration systems, such as adiabatic demagnetization refrigerators (ADRs) and dilution refrigerators, exist that are capable of reaching 100 mK base operational temperatures they often present monetary or operational issues for small, ground-based telescopes and TES bolometers. The dilution refrigerator for Planck cost approximately nine million dollars, a higher
value than the entire budget for the first generation POLARBEAR project. ADRs cool by exchanging magnetic field energy for thermal energy, requiring magnetic fields on the order of a few Teslas to achieve base temperature\(^1\). Large magnetic fields can have unintended effects on TES bolometers and SQUIDs. Magnetic fields are known to quench superconductivity and alter \(T_c\), which would greatly affect bolometer operation. SQUIDs amplify the TES signal due to their sensitivity to changing magnetic flux; they can detect one single Bohr magneton of flux difference. The danger of placing arrays of such sensitive magnetometers next to several Tesla magnetic fields explains why no CMB experiment to date has used an ADR to cool its detectors.

NIS junctions present a viable alternative for cooling TES bolometers to below their BLIP limit. NIS junctions can be lithographed on-chip to cool individual detectors (traditional membrane coolers) or as external coolers used to cool the entire focal plane (suspended brick membrane coolers). With regards to the latter, NIS junctions as external coolers are not limited to operation with TES bolometers; they present an interesting alternative to commercial cryogenic coolers (discussed in more detail in Section 4.4). The remainder of this chapter describes the work I performed at UCSD in order to couple TES bolometers and NIS junctions at lower operational temperatures.

### 4.3 UCSD Fabrication

The coupling of NIS junctions and TES bolometers requires many separate, intermediary steps. This section discusses the methods for fabricating bolometers with lower \(T_c\) in the POLARBEAR architecture and highlights my contribution to the fabrication of future POLARBEAR focal plane arrays. The fabrication and testing of thermally oxidized NIS junctions are discussed in Appendix A and B respectively.

\(^1\)Typical commercial ADRs can achieve base temperatures of 50 – 100 mK.
Figure 4.16: POLARBEAR pixel made by the author with labeled components: (1) polarization sensitive dual slot-dipole antenna, (2) superconducting microstrip line, (3) microstrip crossover/crossunder structure, (4) distributed bandpass filters, and (5) thermally released TES bolometers.

4.3.1 POLARBEAR Pixel Fabrication

The original idea for coupling NIS junctions to POLARBEAR bolometers was on-chip, integrated fabrication, known as distributed cooling. As the first step in this process, POLARBEAR detectors were fabricated in UCSD’s nano3 facility for in-house research and development. Figure 4.16 shows a POLARBEAR pixel fabricated at UCSD with the main components labeled.

Each POLARBEAR pixel integrates several different technologies. The dual slot-dipole antenna couples free-space radiation to a superconducting microstrip line, separating the incident signal into its linearly polarized components.
Spectral bandpass filters along the microstrip line define the center frequency and bandwidth of the transmitted radiation, and the microstrip line terminates on a TES bolometer that measures the sky signal. In the following paragraphs, we very briefly describe the individual POLARBEAR pixel components. For a complete description, please see [117, 118, 119].

The antenna chosen for use in POLARBEAR is a superconducting, planar structure which couples optically to a dielectric lenslet through the backside of the Si substrate. This optical coupling has several benefits. For the dual-slot dipole, the beam is formed more strongly in the direction of high-dielectric constant. The figure of merit for an antenna is its gain, which quantifies how well an antenna converts radio waves arriving from a specified direction into electrical power. Typically, it is given by the ratio of the power received by the antenna when a far-field source is incident on the beam axis to the power received by a lossless isotropic antenna, which is uniformly sensitive to all directions. Coupling the antenna to the relatively high dielectric constant Si substrate and Si lenslet gives a gain that is 10 times higher in the forward direction (direction of incident sky radiation) than in the backward direction [120]. Additionally, coupling the antenna to a dielectric lenslet limits the total internal reflection within the substrate, reducing unwanted power from substrate modes. The lenslet also effectively magnifies the active size of the antenna, leaving additional room for other pixel elements without increasing overall pixel size. This maximizes precious focal plane real estate that must be cooled to a temperature of 250 mK as described earlier. A depiction of a POLARBEAR antenna coupled to a Si lenslet is given in Figure 4.17(a).

Any antenna chosen for POLARBEAR must effectively couple to the telescope optics, lenslets, and other pixel elements such as band-defining filters. The dual slot-dipole was chosen for its low cross-polarization, relatively low impedance, and relatively circular beams [121, 18]. The antenna dimensions are shown in Figure 4.17(b), and are tuned for a resonant frequency of 148 GHz such that Gaussicity of the beam is maximized. The antenna is manufactured as the negative cut-out in a 300 nm thick layer of superconducting Nb.

Radiation from the slot-dipole antenna is coupled to the TES bolometers
Figure 4.17: (a) A dual slot dipole antenna fabricated at UCSD with planar dimensions labeled. (b) A depiction of a dual slot coupled antenna and Si lenslet. Radiation couples to the antenna from the bottom in this depiction, through the backside of the device wafer. The size of the antenna has been exaggerated for clarity. Image taken from [117].

through a microstrip line transformer. The antenna has an impedance of 30 Ω and thus, to avoid reflection loss, the microstrip must be well matched. A line width of 4 µm and thickness of 600 nm provides this desired impedance, but issues with optical lithography, such as resolution and repeatability, present a problem for overall pixel yield at these values. Thus, the line width is chosen as 4 µm where the microstrip couples to the antenna, but is tapered out gradually along the length of the microstrip line until it encounters the filters. The effect of this tapering on the impedance of the microstrip transformer is well-modeled [122] and can be used to determine the efficacy of transmission. For the POLARBEAR pixel architecture, the microstrip line effectively transmits 97% of the total radiation from the antenna [118]. The four transformer lines couple to radiation from the antenna through four small vias into the ground plane, and the remainder of the line is separated from the plane of the antenna by a 300 nm thick dielectric layer of SiO₂.

Each slot on the antenna is coupled to an individual microstrip transformer, yielding four lines in total. For the two lines of the same polarization to couple to a single bolometer, two microstrip lines transmitting orthogonal polarizations
Figure 4.18: A depiction of the sequential steps in the crossover fabrication process as viewed from above. Within the images, blue represents SiO$_2$ and gray represents Nb. (a) A SiO$_2$ layer is deposited and patterned, providing insulation from the ground plane. (b) A small Nb transmission line is patterned atop the SiO$_2$. (c) A layer of SiO$_2$ is deposited and circular vias are etched, exposing the Nb but leaving the rest covered. (d) The remainder of the transmission lines are patterned: a continuous transmission line over the upper SiO$_2$ layer, shown in the vertical direction, and a “broken” transmission line connected through the vias, shown in the horizontal direction.

must cross. A crossover structure is required to prevent the signals from mixing. The design of this structure is shown in Figure 4.18. A small insulating barrier of SiO$_2$ is first deposited between the ground plane and the would-be crossover structure. A small transmission line of Nb is then patterned on top of the barrier with enlarged endpoints for later contact to the remainder of the stripline. A conformal layer of SiO$_2$ is deposited and holes are etched in the dielectric to expose the endpoints of the Nb line below. Finally, Nb is deposited and patterned to form two striplines: one that connects to the exposed Nb link below and one that crosses continuously over in the transverse direction. The simulated crosstalk of this structure is less than 0.01% [118]. To keep all microstrip lines identical, matching crossover structures are added to the other two lines that do not cross and are terminated with an Al/Ti bilayer on the non-transmitting ends.

Radiation passes through a set of distributed filters when the impedance of the line has decreased to 10 $\Omega$. Distributed filters, as opposed to lumped element
filters, were chosen for their size and corresponding ease of fabrication. The filters are standard shorted quarter-wavelength stub filters [123, 124]. The operational parameters, such as frequency, bandwidth, and steepness of cutoff, are optimized for POLARBEAR’s CMB observations through the atmosphere.

The lines terminate on the load resistor of a thermally released bolometer structure as seen in Figure 4.19. The weak thermal link to the bath is provided by ≈1 µm thick silicon nitride legs. The Si substrate has been removed from below the bolometer structure by reactive ion etching with XeF₂, leaving the bolometer suspended by only these legs (Figure 4.19(b)). To minimize reflected power, the load resistor, fabricated from Ti, is tuned for an impedance of 20 Ω to match the two identical 10 Ω microstrip lines that terminate there. Incident power on the load resistor heats the TES thermistor. POLARBEAR employs a dual transition thermistor of two series resistors: a higher operational temperature mode made of Al (T_c ≈ 1.2 K) and a lower operational temperature mode composed of a bilayer of Al and Ti (T_c ≈ 500 mK). CMB observations are taken using the lower temperature mode, which is tuned for an operational resistance of approximately 1 Ω for responsivity and readout requirements. During observations using the lower T_c thermistor, the series Al thermistor is superconducting and does not contribute to the total resistance.

The time constant of the bolometer must be tuned for the read out system. The tuning can be achieved by manipulating either the heat capacity or the thermal conductivity, though the latter must be carefully chosen to optimize the operational power of the bolometer in relation to the sky signal. Thus, the heat capacity of the bolometer is altered to achieve our desired device parameters (Section 4.2.2 and 4.2.3). This is done by depositing approximately 1 µm of Au, called “bling”, onto the surface of the bolometer in good thermal contact with the thermistor. The decrease in overall heat capacity effectively lowers the bolometer time constant. A scanning electron microscope (SEM) image of the Au “bling” can be seen in Figure 4.19(b).

The original idea for on-chip NIS cooling of bolometers was motivated by the POLARBEAR pixel architecture. Due to the lenslet/antenna coupling for in-
Two microscope images of the POLARBEAR TES bolometer made by the author: (a) before Au deposition and dry etch release from the substrate and (b) after Au deposition and dry etch release. In (a) the dual transition thermistor is easily observed.

Coming radiation, the POLARBEAR pixels feature sufficient free-space to integrate NIS junction coolers. Additionally, the released bolometers used in POLARBEAR make immediately possible the traditional membrane cooling scheme.

Unfortunately, as one readily observes, the structure of POLARBEAR pixels is tremendously complicated. The observation quality pixels used in POLARBEAR were fabricated in the Berkeley MicroLab, which houses a 20:1 reduction stepper with a 1/6 µm alignment tolerance that was used for the photolithography of these devices. The 9 layer POLARBEAR pixel stack, as shown in Figure 4.20, has an approximate 1 µm alignment tolerance for the final fabrication step, the release of the bolometers, to work properly. At UCSD, the tightest alignment tolerance for photolithography is still achieved through contact printing, which has an individual layer alignment tolerance of 0.5 µm. Even with alignment close to these lower limits, random directional error would compound through the layers as \((0.5)(\sqrt{10}) \approx 1.5 \mu m\). This makes the reliable fabrication of POLARBEAR pixels at UCSD difficult if not impossible.
Additionally, the on-chip fabrication of NIS junctions requires, at the very least, three more layers on the total process: the normal metal, the insulator that defines the junction area, and the superconducting metal. Each of these must be deposited and patterned, adding increased complexity to the already complicated pixel. However, simple NIS coolers have been shown to reduce bath temperatures by only 20-30 mK [114], and the payoff for such a tremendous effort becomes very small. For optimized cooling down to a base temperature of 100-125 mK, an additional layer for trapping recombined quasiparticles and a layer for increasing the thermal conductivity of the normal metal electron bath are required, further increasing complexity.

However, the most important reason for abandoning integrated, distributed cooling was the bias and read-out lines that would be required. First, the NIS junctions must be voltage biased to cool the TES detectors, requiring on-chip lithographed bias lines. Due to the pixel layout, it would be impossible for bias lines to reach the junctions without crossing the microstrip lines. Obviously, the signal from a voltage bias cannot be mixed with the signal that is transmitted down the microstrip lines, and so another series of crossover structures would be necessary. This adds four additional layers of complexity and, more importantly, an additional
risk for the focal plane. Even relatively small power leakage from the bias line into the microstrip would result in a potentially fatal situation for POLARBEAR, as fluctuations in the bias power would falsely manifest as a changing optical signal. And, although the POLARBEAR pixels themselves feature ample space for incorporating junctions, the perimeter of the focal plane is completely filled with wirebond pads for TES bolometer readout. With the planned six NIS junctions per POLARBEAR pixel (one on each suspension leg of every bolometer), the number of bondpads increases six-fold, reducing the available focal plane area. Simply put, introducing on-chip NIS junctions would not affect pixel size, but would reduce the total size available for all POLARBEAR pixels. Of course, one can reduce the number of required bond pads by biasing sets of NIS junctions in series, but this further complicates the routing issues and with even higher values of voltage bias.

For all of these reasons, the distributed cooling idea was abandoned in favor of bulk cooling, an attempt to cool the entire POLARBEAR focal plane using NIS junctions. This is discussed in much greater detail in Section 4.4.

4.3.2 Seating Wafers for POLARBEAR-2

Although distributed cooling was abandoned, the fabrication of POLARBEAR pixels brought a working fabrication lab to the cosmology group at UCSD. For POLARBEAR, all focal plane components were fabricated at UC Berkeley, but for the extension project, POLARBEAR-2, the UCSD cosmology group will fabricate the lenslet seating wafers that couple each lenslet to its corresponding pixel.

POLARBEAR features a flat, telecentric focal plane, where parallel rays are incident in the perpendicular direction to the orientation of the focal plane. To maximize observing efficiency, these rays must be converged onto a focus at each antenna. Parallel rays incident on a dielectric elliptical surface of the right eccentricity (based on dielectric material) can be focused onto a single point, as shown Figure 4.21(a). For materials with high dielectric constant, the eccentricity required of the ellipse is sufficiently small that it can be approximated by an extended hemisphere, as shown in Figure 4.21(b). And, whereas elliptical lenses
Figure 4.21: Parallel rays incident upon (a) an elliptical surface of the proper eccentricity for focus on a single point and (b) an extended hemisphere where the rays have some offset transmitted angle dependent upon the length L. Figure taken from [118].

are very difficult to produce, hemispherical lenslets are not.

A range of acceptable extension lengths, \(0.576\lambda \leq L \leq 0.624\lambda\), where \(\lambda\) is the wavelength of the center of the detector band, produce beams with high Gaussicity and directivity [120] and are acceptable for coupling to POLARBEAR pixels. Rather than couple each detector individually to the hemispherical lenslet plus extension, each wafer in the focal plane array is coupled to a seating wafer that houses all the lenslets simultaneously with the proper extension lengths. A finished seating wafer is shown in Figure 4.22(d). The depressed regions are the seats that house the lenslets, and the center of each set corresponds to the center location of an antenna on the device wafer. For POLARBEAR, the fabrication of seating wafers became a roadblock to detector testing; all fabrication, both detectors and seating wafers, was done at UC Berkeley, placing a tremendous stress on one institution in the collaboration. To alleviate the stress and increase productivity for POLARBEAR-2 and future generations, I designed a lenslet seating wafer process at UCSD so that fabrication responsibilities could be divided among institutions.

Seating wafer production begins by cleaning the Si substrate with a three minute ultrasonic bath in each of acetone, isopropyl alcohol, and deionized water.
Figure 4.22: Several steps highlighting the fabrication process for POLARBEAR-2 seating wafers: (a) Patterned photoresist, (b-c) the post-etch wafer, and (d) the finished seating wafer. The depth of the etch (120 µm) can be easily observed in (c).
The wafer is then blown dry with \( \text{N}_2 \) and descummed with an oxygen plasma for two minutes. The negative photoresist NR9-8000 is spun onto the wafer at 3500 RPM for 40 seconds, yielding a thickness of approximately 2.5 \( \mu \text{m} \). The photoresist must be at least this thick to conformally coat the wafer during the entire deep etch process. Wafers are proximity exposed in the EVG mask aligner and developed for two minutes in resist developer RD6, rinsed in deionized water, and blown dry with \( \text{N}_2 \). The resulting pattern can be seen in Figure 4.22(a). The wafers are then oxygen plasma etched for one minute to remove any residual photoresist that many be clinging to the surface of a seat. A Bosch process is used for the etching; the etch rate is approximately 2 \( \mu \text{m} \) per minute for the full POLARBEAR-2 pattern, but varies as a function of the exposed area of Si on the wafer. Post-etching the seating wafers appear as in Figure 4.22(b)-(c). The photoresist is then removed in three steps: an ultrasonic bath in acetone for three minutes, a soak in resist remover RR4 at 50° C for one hour, and an oxygen plasma etch for five minutes. The final product is shown in Figure 4.22(d).

The use of a negative photoresist is important in the seating wafer process and was pioneered at UCSD. Residual photoresist in a seat will cause a local spot to remain selectively unetched, resulting in a pillar of material in the area where the lenslet must sit. Such a profile prevents the lenslets from sitting flush against the seat, resulting in vacuum gaps between the lenslet and spacer that degrade device performance. A positive photoresist develops, or is selectively removed, where it is exposed to UV light. Any particulate matter on the surface of the photoresist within the seat results in an unexposed region that forms a pillar during the etch process. A negative photoresist, however, develops where it remains unexposed. Particulate matter on the seats has no possible effect, because the entire seats must remain unexposed to develop. With this change in process, it is expected that seating wafer fabrication at UCSD should have very high yield.

The detector array in POLARBEAR-2 will be similar to that of POLARBEAR except it will house approximately six times as many bolometers, will have two operational frequencies, and will be fabricated on six inch wafers. The exact extension length is as-yet unknown, but the process will be similar to the one that
has been developed. The developmental seating wafers utilized substrates with a thickness of 500 µm and an etch depth of 150 µm. Etch depths less than 100 µm do not make good seats for the lenslets, so typically the substrate thickness is varied to give the proper extension lengths for the lenslets and the etch depth remains set at ≈ 150µm.

The quality of the prototype seating wafers was characterized using two parameters: edge profile and seat depth. For the lenslets to sit properly centered above the antennas, the edge of the seat must not taper more than 10% of the seat depth. Optical microscopy confirmed that seats in the center of the array had an average taper of 8.5 µm increasing to 14 µm at the perimeter of the array. Similarly, the depth was found to be 146 ± 2 µm for all seats from the center out to the second farthest ring, well within specification and with less variability than the seating wafers for POLARBEAR. The two outermost rings of seats, however, were found to increase precipitously in depth to 160 µm and 180 µm; the lack of exposed Si area around the perimeter of the wafer causes the etch to penetrate more quickly due to excess gas in this region. Alterations to the photomask exposing more area outside the perimeter of the seats have been made, and with this change seating wafers should be easily fabricated to specification with high yield. Preliminary seating wafers for testing full POLARBEAR-2 device wafers are currently being fabricated at UCSD. This process will continue until enough quality wafers for POLARBEAR-2 have been fabricated.

4.3.3 AlMn Bolometers

Whether using distributed or bulk cooling, a colder bath temperature requires bolometers that operate at lower $T_c$ to minimize the NET. The Al/Ti bolometers used in POLARBEAR are tuned for operation at approximately 500 mK. By changing the relative thicknesses of the layer one can manipulate $T_c$ in the range $0.4 \text{ K} \leq T_c \leq 1.2 \text{ K}$, but lower temperatures cannot be achieved using these materials. One promising alternative is to use the lightly doped alloy AlMn. The dilute addition of Mn lowers the superconducting transition temperature of Al dramatically while retaining most of the bulk properties of Al such as resistivity
and heat capacity. Mn is a magnetic impurity that retains its magnetic character even in dilution, scattering Cooper pairs at impurity sites, altering the density of states, $\nu$, and greatly suppressing $T_c$ [125, 126, 127]. By 3000 parts-per-million (ppm) atomic doping, the superconducting temperature drops from 1.2 K to below 50 mK [125].

An AlMn sputtering target was acquired from ACI Alloys with doping concentration of Mn at 3000 ppm. The superconducting critical temperature is extremely sensitive to the doping concentration and also to other impurities within the metals [105], and so the first step in fabricating TES bolometers using AlMn was to confirm the composition and purity of the sputtering target.

X-Ray Fluorescence (XRF) measurements of the sputtering target were taken to analyze its composition. XRF is a technique where high energy X-rays cause characteristic secondary (fluorescence) emission of an element/material by ionization, ejecting inner shell electrons from the atom. When outer shell electrons cascade down to fill these holes, photons of specific energy are emitted and their energies are recorded. The energies are characteristic of the elemental material that was ionized, so the energy of the emitted radiation is unique to each element. Recording count rates of the materials allows for extrapolation of the sample composition.

Figure 4.23 shows an example spectrum acquired at several different locations on our sputtering target. The sample shows strong Al and Mn peaks in all three spectra. The Rh peak is expected because Rh is the source of the incident X-ray excitation, but low levels of magnetic impurities can be observed that vary based on target location. Fortunately, compositional analysis confirmed that these did not contribute significantly to the overall composition of the target; the contribution was less than the lowest threshold set by the software, 0.01% by mass. The software also confirmed an atomic doping of Mn of 3130 ppm ± 190 ppm where the uncertainty comes from target location and not from instrumental precision.

Using this target, different doping concentrations of Mn could be achieved.

---

2Originally, we received a heavily contaminated and improperly doped sputtering target from Kurt Lesker that we did not check for quality. The repeated use of this target in both bolometers and NIS junctions set us back several months.
Figure 4.23: Three spectra taken on the 3000 ppm atomic composition AlMn sputtering target from ACI Alloys. Important peaks are labeled.
by co-sputtering the alloy target with a pure Al target as

$$X = \frac{r_{\text{AlMn}}}{r_{\text{AlMn}} + r_{\text{Al}}} (3000) \ [\text{ppm}]$$

(4.30)

where \( r_{\text{AlMn}} \) and \( r_{\text{Al}} \) are the deposition rates of AlMn and Al respectively. Figure 4.24(a) shows the resistance as a function of temperature for AlMn samples deposited in identical patterns and thicknesses but with different dilute alloy compositions. The superconducting transition temperatures and normal state resistances as a function of doping concentration are given in Figure 4.24(b). Samples prepared with identical sputtering parameters months later yielded similar superconducting transition temperatures to within 5%, indicating that this sputtering process was fairly stable over time. The same targets were used for both sets of samples.

To test the feasibility of using these AlMn bolometers in POLARBEAR and future generations of POLARBEAR, dark bolometers, those unconnected to an antenna and thus unexposed to optical power, were fabricated with the nominal POLARBEAR \( T_c \) of 500 mK. The top panel of Figure 4.25 shows the current-voltage (IV) curves of one of the AlMn dark bolometers. Data were taken at 270 mK to match the base temperature of POLARBEAR. At high values of the voltage bias, above approximately 13 \( \mu \text{V} \), the bias power is sufficient to Joule heat the TES above its superconducting transition temperature and the IV curve exhibits linear behavior. The inverse slope of this linear portion gives the normal state resistance of the TES. As the bias voltage is lowered the self-heating drops such that the TES enters the superconducting transition. At this point, the IV turns upwards and the resistance begins to drop as shown in the bottom panel. In this same region the power (middle panel) asymptotes to an approximately constant value, indicating that high loop gain has been achieved. The saturation power, \( P_{\text{sat}} \), is given by the value of the power at the turnaround point. This is the maximum amount of power that the TES can absorb before transitioning back to a normal metal. POLARBEAR’s design specification for \( P_{\text{sat}} \) is 16 pW. The fabricated AlMn bolometer had a saturation power of 19 pW and a \( T_c \) of 520 mK, demonstrating that AlMn is a viable alternative for POLARBEAR TES bolometers and POLARBEAR style bolometers that require a lower \( T_c \).
Figure 4.24: (a) Resistance versus temperature measurements for the superconducting transition of doped AlMn alloys. All samples are identical in size, shape, and thickness. (b) Transition temperature and normal state resistance as a function of Mn doping extracted from the data shown in (a).
Figure 4.25: Electronic transport tests taken on AlMn dark bolometers fabricated at UCSD. The top panel is a plot of the IV curve of the bolometer, the middle panel is a plot of the power versus voltage, and the bottom panel shows the resistance versus voltage.
4.4 Payload Cooling

As mentioned in previous sections, Normal metal-Insulator-Superconductor tunnel junctions can be used to cool the electron bath in the normal metal by selectively tunneling the hottest electrons into the superconductor. Although NIS junctions have been used to cool external payloads on a single lithographed sample [128, 129], this process is often overly complicated. As in the case of POLARBEAR detectors, two entirely different, complex technologies must be integrated during detector fabrication for on-chip cooling.

The flexibility of NIS refrigerators can be increased by designing an external cooling platform that can cool arbitrary payloads. This section describes the design and construction of a fully implementable NIS junction cooling platform, which can be backed by a Helium-3 refrigerator for cooling from 300 mK to 100 mK or by an ADR for cooling from 100 mK to 20 mK. The stage is lightweight, extremely mechanically robust, and easily implementable thermally, mechanically, and electronically. This work was performed as a collaboration between UCSD and the National Institute of Standards and Technology (NIST); the stage was designed and built at UCSD while NIST provided the NIS junction coolers.

4.4.1 NIST Junctions

The NIS coolers used for platform cooling were designed and systematically optimized for cooling power by Galen O’Neil [105] and fabricated by Peter Lowell\textsuperscript{3} at NIST. The junctions function as suspended brick membrane coolers. A SEM micrograph image of one such cooler is given in Figure 4.26(a). Figure 4.26(b) shows a schematic of a cross sectional slice of the junction coolers.

The cooled membrane is suspended by eight 20 \( \mu \text{m} \) long and 60 \( \mu \text{m} \) wide SiNx/SiO\textsubscript{2} legs of thickness 370 nm and coated in the same material. A pair of NIS junctions, with lateral size 7 \( \mu \text{m} \) by 32 \( \mu \text{m} \) each, is used to cool the suspended membrane and located on the bulk substrate at the base of each suspension leg. The normal metal of the junctions is a 30 nm thick layer of AlMn with an Mn

\textsuperscript{3}A description of the complete fabrication process can be found in the supplemental files of [130].
Figure 4.26: (a) SEM micrograph of the membrane coolers fabricated at NIST. A pair of NIS junctions couples to each of the eight suspension legs. (b) Cross sectional schematic of the composition and architecture of a single NIS cooler on one suspension leg. Figure taken from [130].
doping concentration of 4000 ppm by atomic composition. The superconducting 
Al layer is 300 nm thick and the insulating barrier between the two is approximately 
1 nm thick, yielding a resistance area product of $2000 \Omega \mu m^2$. The AlMn normal 
metals are extended onto the membrane as a cold finger as shown in Figure 4.26(b). The extended portion on the membrane is coated with 150 nm of Cu and 100 nm 
of Au to improve thermal conductivity. Portions of AlMn on the bulk substrate 
are not coated to minimize the volume-dependent electron-phonon coupling that 
lowers cooling power.

The eight cold fingers extend to form a ring around a central Cu/Au film 
that can be connected to a payload. This central component is electrically isolated 
from the NIS junctions and cold fingers. The 380 $\mu$m Si substrate is left underneath 
the central film and the ends of the cold fingers to provide a thermal link and 
structural support for wirebonding to the payload.

4.4.2 Design Motivation

This NIS junction coolers described above were designed for a cooling power 
of hundreds of picowatts at 300 mK [113]. As such, to enable cooling, any stage 
designed to couple to the junctions much have a lower parasitic load. Addition-
ally, the stage must be a standalone unit that can be easily transferred between 
cryosystems and structurally sound for repeated use. The coupling to NIS junc-
tions provided many other practical limitations. The following design criteria were 
set for the design of a practical NIS junction refrigeration stage:

- **Thermal Loading** - A maximum of 1 nanowatt of thermal loading from 300 
mK to 100 mK. This was a practical limitation set by the maximum number 
of junctions that could be reasonably coupled to the stage.

- **Integrated Heat Switch** - The thermal isolation required by the stage 
results in a cooling time of several *months* from room temperature to 300 mK 
in any standard cryostat. Thus, to be used as a viable refrigerator, the stage 
must incorporate a heat switch. A mechanical heat switch, which contributes 
zero thermal loading in the “off” position, was required to meet the thermal
loading constraints. The thermal conductivity in the “on” position was set at a threshold value of at least 500 nW/K at 300 mK in order to cool the stage within one hour of an ADR or $^3$He cycle. Additionally, the heat switch must also be incorporated into the stage design, electronically controllable at room temperature, and dissipate less power per use than the cooling power supplied to the coldest stage in the cryostat.

- **Structural Integrity** - The stage must be able to support at least 2.3 kgs, the heaviest payloads used at NIST, with deflection under 250 µm. It must be rigid enough to resist vibrations that will dissipate heat into the stage and to withstand repeated use and user manipulation.

- **Modular Junction Platform** - Wirebonding the fragile membrane coolers to any external payload introduces many complications. Direct bonding to the stage is not possible due to the assembly procedure and the limitations of laboratory wirebonders. To overcome these constraints, a modular platform must be designed for external coupling that can then be attached to the stage. This platform must also shield the junctions from user damage.

- **Size** - The main stage must be reasonably sized to accommodate typical payloads, but small enough to fit inside a wide range of cryostats.

- **Heat Capacity** - The stage should have a sufficiently low heat capacity to cool in nine hours or less using several hundred picowatts of excess cooling power. The time was chosen to match the cycle time of the $^3$He refrigerator at UCSD.

- **Electronic Integration** - The system must be “plug-and-play” for users. It must be easy to connect both the junctions and any user supplied payloads to an external cryostat for temperature control and payload measurements.

- **Physical Integration** - The stage must be easy to integrate into multiple cryosystems. For practical purposes, the design criteria was set that it should be usable at both NIST and UCSD without any configuration changes.
4.4.3 Platform Design and Construction

The stage design is based on the design ideas presented in [131]. Figure 4.27(a) shows a CAD rendering of the main stage. The stage consists of two U-shaped brackets composed of three pieces of Copper each that are held together with 0-80 screws. Each U has a 1/4” diameter brass crossbeam for mechanical support that fastens into an Aluminum endcap on each side. The two brackets are held together by a Kevlar suspension that winds around the Aluminum endcaps as shown. Figure 4.27(b) shows a CAD rendering of a single Aluminum endcap. The endcap has machined grooves that seat the Kevlar and thread it to the next position at a 45° angle from vertical. The endcaps are beadblasted to increase friction when tensioning the Kevlar and to help the epoxy adhere to the surface after tensioning. The edges of the endcaps have been rounded to keep the Kevlar from meeting any sharp corners under tension. In Figure 4.27(a), each color represents a unique strand of Kevlar; two strands are required to keep an even tension on the stage. The bottom bracket has 4-40 through holes to attach to the UCSD cryogenic stage and the top bracket has tapped 4-40 holes so that external payloads can be fastened down. The NIS cooled stage is 63.75 mm long by 44.45 mm wide and 3.30 mm thick. It is gold plated to increase thermal conductivity and lower the heat capacity. The entire platform assembly weighs approximately 250 grams.

The completed cooling platform is shown in Figure 4.28(a) with the Kevlar supports holding the U-brackets in place. A closed Aluminum bracket was constructed to hold the U-brackets immobile while tensioning the Kevlar as shown in Figure 4.28(b). The tensioning bracket has a cylindrical extension that can be clamped by a collet block as shown in Figure 4.28(c). Each of the two strands of Kevlar is woven alternatingly from one endcap to the next until the strand meets back with its starting point. One strand on each of the four sides will resist stress under compression, and the other will resist stress under expansion. For maximal resistance under applied mechanical load, the Kevlar should be tensioned to one-half of the breaking strength. This even tension also resists movement under lateral torques and keeps the Kevlar from “creeping” with repeated thermal cycles.

In order to tension at one-half the maximum strength, breaking tests were...
Figure 4.27: CAD rendering of the cooling platform assembly. The full assembly is shown in (a); the Kevlar is drawn in two colors to show the winding of each cord. A single Aluminum endcap is shown in (b). The slots for the Kevlar are cut at a $45^\circ$ angle and the edges have been rounded.

Figure 4.28: (a) The completed cooling platform, (b) the platform held in place by the Aluminum tensioning bracket, and (c) the tensioning of the Kevlar during stage assembly.
performed on the Kevlar by incrementally increasing a hanging weight on the strand until it snapped. Special care was taken not to bend or kink the Kevlar, as this introduced a weakness where the Kevlar was always found to break. Instead, the Kevlar was wrapped around a smooth bend in a stainless steel rod from which the weights were hung. The breaking tension was found to be 68 N, approximately 7 kgs under the force of gravity.

To tension the Kevlar as shown in Figure 4.28(a), the platform cooler was fastened into the tensioning bracket and loaded into the collet block at a 45° angle. This angle was necessary to match the angle at which the Kevlar ran off of the endcaps. This is shown in Figure 4.28(c). Each Kevlar strand was started at the top of an arbitrary endcap and superglue was applied and given five minutes to dry. The collet block was then flipped 180° with the stage held in place and the Kevlar was threaded through the tensioning bracket and looped over the top of the next endcap. It was then tensioned by hanging a 3.5 kg mass from the end as described in the breaking tests. Superglue was applied again and given five minutes to dry while the Kevlar was still under tension. This process was repeated until both strands were complete. After all the Kevlar bonds were set the tensioning bracket was removed. Superglue was used for its convenience; it is extremely strong and dries quickly, but it does not hold at cryogenic temperatures. To combat this, each point of contact between the Kevlar and the stage was covered with the epoxy Lucsol, which is known to form a strong and durable bond through repeated cryogenic cycling.

4.4.4 Thermal Loading

The expected thermal load on the NIS junction cooled stage comes from three sources: the Kevlar suspension, any wires running to external payloads on the mainstage, and radiative loading from the environment. The sum of the expected contribution from each gives us the total thermal loading on the stage.

Assuming that the Kevlar is a homogeneous material with endpoints at constant temperature, the thermal power load contributed from the Kevlar can be
calculated using Fourier’s Law:

\[
P = \frac{A}{L} \int_{T_i}^{T_f} \kappa(T) dT
\]

(4.31)

where \(P\) is the power load, \(A\) is the cross-sectional area of the Kevlar strand, \(L\) is the length between points of thermal contact, \(\kappa(T)\) is the thermal conductivity as a function of temperature, and \(T_i\) and \(T_f\) are the temperatures at the endpoints. Because the thermal conductivity of the Kevlar is fixed, the only way to decrease the thermal loading is to decrease the cross-sectional area or to increase the length. To minimize the loading, we used the thinnest commercially available braided Kevlar\(^4\) with a diameter of 0.02 cm. Although thinner Kevlar exists as unbraided single strands, it is very easily broken under tension and could not be used to keep the stage sufficiently rigid. There are eight Kevlar connections from the lower stage to the upper stage, two at each endcap as shown in Figure 4.27(a). The distance between points of contact for the Kevlar is 4.52 cm due to the size constraints of both cryostats (at NIST and UCSD). The thermal conductivity of Kevlar has been found empirically to be \(3.9 \times 10^{-5} \text{ W/(cm K)} \times T^{1.71}\) at temperatures from 100 mK to 2.5 K [132].

Electrical connections to the upper U-platform are made with 12.7 cm long and 0.002 cm diameter NbTi wires with no Cu matrix. NbTi wire is commonly supplemented with a Cu matrix to ease electrical connections and lower room temperature resistivity, but the presence of Cu in the wires dramatically increases the thermal conductivity at millikelvin temperatures. Cu is a resistive normal metal down to absolute zero, and conducts heat at millikelvin temperature through electron scattering. Pure NbTi, however, superconducts around 10 K. At temperatures far below this value, almost all electrons are bound in Cooper pairs and do not scatter and conduct heat. At such temperatures only phonons, which are much less tightly coupled than electrons, transfer the heat, dramatically lowering the thermal conductivity. The thermal conductivity of pure NbTi has been measured to be \(0.15 \times 10^{-3} \text{ W/(cm K)} \times T^{2.0}\) below 2 K [133]. Preliminary testing of the NIS cooled stage requires seven wires: four to perform IV measurements on a ther-

\(^4\)Purchased from Edmund Scientific, Item number: 3034863.
Figure 4.29: Thermal loading from all parasitic heat leaks on the NIS junction cooled platform. Each point represents the thermal load when cooling from 300 mK to the given temperature. The dominant source of loading is the Kevlar, which is several orders of magnitude higher than the contributions from the wiring and radiation. The Kevlar thermal loading is approximately equal to the total power load; the lines overlap almost identically on the logarithmic scale.

The radiative power, $P_{rad}$ on the stage can be calculated using the Stefan-Boltzmann law

$$P_{rad} = \epsilon \sigma A T^4$$

where $\epsilon$ is the emissivity, $\sigma$ is the Stefan-Boltzmann constant, $A$ is the surface area, and $T$ is the temperature of the blackbody seen by the stage. Assuming a conservative estimate of the emissivity of Cu as 0.1 [134], the radiative power loading on the stage from the surrounding 4 K heat shield at both UCSD and NIST is calculated to be 9.2 nanowatts. Obviously, this large parasitic load would swamp the cooling provided by the junctions and so a heat shield was designed to surround the structure at 300 mK. Although the emissivity of Cu used here is just a rough estimate, the radiative power load with a 300 mK heat shield is orders of
magnitude less than the loads caused by the Kevlar and the NbTi wires.

Figure 4.29 shows the total thermal loading from 300 mK to 100 mK as well as each of the individual contributions to the total thermal load. Each individual Kevlar connection contributes approximately 36 pW of parasitic loading, each NbTi wire contributes approximately 0.3 pW of parasitic loading, and the radiative load from the heat shield is approximately 0.3 pW of parasitic loading. The total combined load is 294 pW, which is 3-4 times below the original specification. Additionally, the extremely low loading given by each NbTi wire suggests the easy usage of the junction cooled platform. Payloads, such as arrays of TES detectors, often require hundreds of readout wires to measure all detectors simultaneously. Using this design, over 2000 wires can be added before the total thermal loading exceeds the design specification.

### 4.4.5 Cooling Time

The stage is made entirely of Cu, brass, and Al in order to minimize the heat capacity and, consequently, the cooling time. The temperature of the stage can be obtained numerically by solving

$$C \frac{dT}{dt} = P_{NIS} + P_{load}$$  \hspace{1cm} (4.33)

where $C$ is the heat capacity of the stage, $T$ is the stage temperature, $t$ is the time, $P_{NIS}$ is the net power coming from the SINIS coolers (given by the right-hand side of Equation 4.19), and $P_{load}$ is the total parasitic loading of the stage itself. The term $P_{NIS}$ is computed using the parameters in the caption of Figure 4.12 and this value is multiplied by 96, since the stage design accommodates twelve junction coolers with eight SINIS junctions each. A total heat capacity of $6.6 \times 10^{-4}$ J/K was computed knowing the heat capacities of the specific composite materials [135] – Aluminum, Copper, and Brass – and the masses of each.

Figure 4.30 shows the numerical solution for the stage temperature as a function of time. For the first three hours, the NIS junctions cool the stage quite rapidly to 200 mK, a drop from the bath temperature of 100 mK. Between six and seven hours the stage cools to 150 mK, halving the original bath temperature.
Figure 4.30: Stage temperature as a function of time calculated by considering the net cooling power of the junction coolers and all parasitic loads present on the stage.

The stage cools asymptotically to 115 mK after 11 hours. This is the minimum temperature that the junctions can provide.

Most cryogenic cooling systems, ADRs and $^3$He refrigerators, have a several hour cooling cycle to reach their lowest temperatures, and so this cooling time was deemed acceptable for the NIS junction stage. It is important to note that unlike ADRs and closed-cycle helium systems, which have to be cycled to maintain their base operating temperatures, NIS junctions are continuous coolers; base operating temperatures, once achieved, can be held indefinitely.

4.4.6 Additional Stage Components

Aside from the platform cooler itself, the completed stage has several more important integrated components. The next few sub-sections highlight the main points of each.
Figure 4.31: An image of the completed stage mounted in the heat shield. The heat shield has two thick Cu walls for thermal conduction, one thick Al wall for housing electronics and the heat switch, and three thin walls made of bent Aluminum sheet metal.

Heat Shield

Figure 4.31 shows an image of the cooling platform within the radiative heat shield. Two walls of the heat shield are made of Cu for high thermal conductivity. The wall shown on the bottom in Figure 4.31 is used to mount to the cryosystem at UCSD while the other Cu wall is used to mount to the cryostat at NIST. The thicker Al wall is used to house the electronics and the heat switch; the cutouts are for microD-25 connectors. The remaining three sides are bent Aluminum sheet metal with drilled holes to connect to the other half of the shield. The connecting holes can be seen clearly on the thicker components of the shield. The heat shield is designed to be completely light-leak proof; only 300 mK radiation should be incident on the platform.

Modular Platform

Figure 4.32(a) shows an image of the assembled stage with the modular platform in place and Figure 4.32(b) shows the complete CAD rendering. Due to the difficulty in wirebonding the fragile membranes and the constraints imposed by the wirebonding process, a flat modular platform was designed so that the
Figure 4.32: (a) An image of the assembled stage with the modular platform and (b) a CAD rendering of the same thing with the Aluminum end braces in place.

wirebonding could take place external to the rest of the stage. The platform is 3.5” long, with enough space to house twelve NIS junction cooler chips and the Printed Circuit Board (PCB) used for electronic bias and readout. The modular platform sits atop two Cu pieces that stand vertically and mount to the exterior profile of the stage as shown.

The NIS junction chips are mounted atop a PCB on the half of the modular platform that does not contact the upper stage. This half of the platform is labeled “NIS Junctions” in Figure 4.32(a). The junctions are held in place with rubber cement and the membranes are fixed with low melting point wax so that they remain rigid during wirebonding. During the wirebonding process, the two ends of the modular platform are held together by Aluminum end braces as shown in Figure 4.32(b). Wirebonds are made from the NIS junctions across to the other half of the modular platform, which has been Au plated (with no magnetic undercoat so as not to harm junction performance) to help the wirebonds adhere to the surface and increase the thermal conductivity to the NIS cooled platform when reattached. Once the bonds have been made, the entire modular platform is soaked in acetone to remove the low temperature wax from below the membranes. This step is crucial; any physical touch (such as left-over wax) between the membranes and the PCB will act as a parasitic load and prohibit the junctions from cooling the suspension. This is yet another motivation for a removable platform. If the
platform had been fixed to the stage then the entire assembly would have to be soaked in acetone, which is extremely undesirable as it weakens the Kevlar. After the wax has been removed, the platform is attached to the stage by fastening the left half to the two vertical Cu pieces and the right, gold-plated half, to the cooling platform. The Aluminum end braces are removed so that there is no contact between the two halves except for the wirebonds. The half of the modular platform that houses the junctions is then covered with an Aluminum hood to prevent accidental damage, as shown in Figure 4.33(a).

Heat Switch

Figure 4.33(a) shows a CAD rendering of the heat switch design in the final assembly and Figure 4.33(b) shows a close-up photograph of the actual switch. The entire switch is seamlessly integrated into the stage and shield assembly. It is driven by a 2.5 cm diameter stepper motor\(^5\) whose stock Cu coils have been replaced with 1000 winds of 0.23 mm Cu cladded NbTi wire insulated in formvar\(^6\). The stepper was chosen for its small size and symmetric shape; the size allowed it to fit easily in the assembly and the shape lends to symmetric differential contraction so that moving parts still mesh at low temperatures. Operation of the stepper motor requires a minimum current of 0.3 A. The inspiration for using this particular motor came from [136].

The shaft of the stepper motor turns an anchored screw that was attached to the motor via differential heating and applied force. The screw head was sawed off and a hole was drilled into one end with a diameter 0.002” less than the diameter of the shaft of the stepper motor. The drilled screw was then heated with a torch for thermal expansion and immediately forced onto the motor shaft using a press. The screw mates with a tapped Cu block that is kept from rotating by a fixed guide. Stepper motor rotation advances (or retracts) the Cu block along the guide and towards (or away from) a hanging Cu tongue attached to the upper stage as shown in Figure 4.33(a). The movable Cu block is attached to a thick Cu foil that

\(^5\)Nippon Pulse product PF25
\(^6\)Supercon Wires, product number SC-T48B-M-0.08mm
Figure 4.33: (a) A CAD rendering of the heat switch in heat shield assembly integrated with the cooling platform. The stepper motor turns an attached screw that drives a Cu block forward and backward along a guide. The Cu block contacts a Cu tongue that hangs down from the upper, thermally isolated stage. (b) A close-up photograph of the heat switch and Cu foil that connects the switch thermally to the 300 mK shield. The Cu foil is bolted down to the shield wall for good thermal contact.
is well thermally sunk to the heat shield at 300 mK, as shown in Figure 4.33(b). This provides the complete thermal connection from the upper stage to the rest of the assembly. When the block is advanced and makes contact with the tongue on the upper stage, the switch is closed and heat flows between the upper stage and the shield. When the Cu block is retracted the contact is broken and no heat is transferred across the switch. Touch wires were fastened to the heat switch so that contact (i.e. - opening and closing) could be monitored from outside the closed cryostat.

**Electronics**

The electronics for the system were designed for easy integration with any cryostat. Two microD-25s were integrated into the thicker Aluminum sidewall to form a tight radiation seal with the heat shield and an electronic connection to the outside world, as shown in Figure 4.34. Once the internal wiring is assembled and connected on the stage and modular platform, user need only plug into the microD-25s to control the junction bias and heat switch motor and readout devices attached to the upper stage.

Three PCBs were designed and constructed to facilitate the electronics for the whole assembly, which is shown in Figure 4.34(c-d). Figure 4.34(a) shows the two coupled PCBs that integrate the junction bias and readout lines. The larger of the two PCBs is attached to the sidewall as shown in Figure 4.34(b-c) and a male microD-25 is ported out through the sidewall for connection to external wiring as shown in Figure 4.34(d). This PCB has twelve lines, each populated with 80 and 15 MHz low pass filters that eliminate RF noise that can degrade cooling performance. Schottky diodes are also present on each line to protect the junctions from damage by electrostatic discharge. The two junction PCBs are connected via Millmax connectors and Cu wires to meet space constraints on top of the stage. The PCB on the modular platform has four Cu placeholder pads for NIS cooler chips. Each chip houses nine suspended membranes, and three membranes per chip can be biased simultaneously, allowing for a total of twelve NIS coolers. The eight NIS junction pairs that are coupled to each suspended membrane are biased
Figure 4.34: Images showing the integration of the electronic circuitry into the cooling platform package: (a) the NIS junction cooler PCBs with important components labeled, (b) the larger NIS junction cooler PCB loaded into the heat shield side wall, (c) all electronic components loaded into the assembly with junctions bonded to the modular platform, and (d) the completed assembly showing the seamless integration that has been designed for an external user. Once the electronics are assembled they need not be altered; an external user can simply plug-in to the external microDs that are shown.

in series. Wirebonds from the samples to the PCB bondpads route the electronics for both NIS biasing and readout. It is important that junction readout is present; current feedback is necessary to optimally bias the junctions for cooling power.

The final PCB is used to measure payloads on the ultracold stage. NbTi wires are used to connect the upper stage electronics to the standalone microD-25 in the heat shield. The NbTi wires are very delicate and difficult to solder\(^7\), so

\(^7\)“Soldering” the NbTi wires requires first pressing Indium into the soldering port, then heating the Indium above \(\approx 300^\circ\) F, and then inserting the fine NbTi wire into the hot metal so that the heat of the Indium removes the formvar insulation. Stripping the NbTi wires of their insulation or bonding with conventional solder is difficult if not impossible. For reference, flux must **NOT**
once connected it is desirable that they are not manipulated by users. By coupling them to an external and robust microD through the heat shield, a user need only plug into the port on the outside of the shield to control the electronics on the cold stage. Four Cu placeholders are designed into the PCB on the NIS cooled stage and are routed out to the electronics on the shield. Four spaces in this microD are left intentionally open for the stepper motor wiring, which is threaded back through a small hole in the shield and soldered in using Cu wire. By incorporating the wires for the stepper into this electronics packet, all elements of the system can be controlled for simple “plug-and-play” operation.

### 4.4.7 Testing

After construction, the stage was subjected to many tests to confirm that it met the design criteria.

**Deflection Tests**

The platform cooler was evaluated for deflection under load using a dial indicator as shown in Figure 4.35(a). Testing was done after the Kevlar joints be used.

**Figure 4.35**: Photographs of the stage testing performed after assembly of the stage: (a) deflection tests using a dial indicator and an applied load of 2.5 kgs and (b) liquid Nitrogen dunks using a Kevlar suspension.
were epoxied with Lucsol to ensure that these joints had set. This ensured that all load would be on the Kevlar and not on the connections. One-half kg weights were added incrementally to the stage until the design tolerance of five pounds was met. Under an applied load of 2.5 kgs, the stage was shown to deflect 0.001 inches, well beyond the specification that had been set.

It is important that the deflection gives a resonant frequency that is much higher than the pulse tube cooler in the cryogenic testbed or vibrations will couple in from the surrounding environment and heat the stage. With the bottom of U of the platform assembly clamped in place, the upper stage can be modeled simply as a mass on a spring. The spring constant of the Kevlar was calculated using the stage deflection test as $k = \text{(applied load)} / \text{(deflection)} = 9.7 \times 10^5 \text{ kg/s}^2$. The resonant frequency was determined by

$$\nu = \frac{1}{2\pi} \sqrt{\frac{k}{m}}, \quad (4.34)$$

where $m = 0.1 \text{ kg}$ is the mass of the upper stage. The resonant frequency was found to be 495 Hz, orders of magnitude higher than the pulse tube frequency.

**Thermal Cycling**

For use as a practical millikelvin refrigeration system, the stage must be able to withstand repeated thermal cycles and thermal shock. Preliminary thermal cycling tests are shown in Figure 4.35(b). The stage was lowered into liquid Nitrogen by a Kevlar cord and held until the two reached equilibrium temperature, after which it was removed and exposed to room temperature for several minutes. This process was repeated in excess of 20 times to prove that repeated thermal contraction and expansion of the platform would not break the Kevlar under tension, that the Kevlar would not creep or loosen, and that the stage would not warp. After over 20 cycles, the stage was found to be as structurally rigid and well tensioned as it has been before dunking.
Stepper Motor Testing

The stepper motor assembly was tested using a liquid helium dunk to ensure that the motor coils would go superconducting, that the wiring and controls would work at cryogenic temperatures, that the assembly would remain intact under thermal contraction, and that a proper biasing scheme could be used to drive the stepper in its superconducting state. Before the dunk, a flag was attached to the stepper motor to mark its original position. Upon dunking, the stepper motor transitioned quickly to superconducting from a room temperature resistance of over 300 Ω. The only residual resistance was contributed by the bias wires, and was sufficiently small that it became difficult to apply the required 0.3 A of current with any reasonable precision. To combat this problem, 10 Ω load resistors were added at room temperature to each of the four bias lines driving the stepper motor. The motor was stepped in the same direction several hundred times (7.5° per step) and removed from the liquid He. Proper operation was confirmed.

Preliminary Cooldown and Testing

After testing the preliminary components, the stage was taken to NIST for cryogenic testing with all integrated components. Figure 4.36 shows the assembly loaded into the NIST cryostat. The entire assembly was kept below a total mass of two pounds and fit inside of a 5-inch cube.

To test the electronics, twelve NIS junction coolers were wirebonded to the PCB on the modular platform as shown in Figure 4.34(c) and a ruthenium oxide (RuOx) surface mount thermometer and bias resistor were attached to the upper stage. During this cooldown the junction coolers were not bonded to the upper stage for payload cooling, but instead were screened for performance using the new electronics. It is important that the twelve chosen coolers coupled to the stage have the highest yield and maximal cooling power and, since they are biased in series, that all of them have comparable device parameters for optimal biasing. Before the stage can be cooled, several rounds of junction screening must be performed, though the preliminary test showed that all electronic connections were good and that the junctions could be reliably biased and read out.
Figure 4.36: The platform assembly inside of the heat shield loaded into the cryostat at NIST.

For the preliminary cooldown from room temperature, the heat switch was placed in the closed position such that the stage would cool with the rest of the cryostat. The heat switch was proven to be effective at cooling the stage; the temperature of the Kevlar suspended U-platform agreed with the temperature of the coldest portion of the cryostat within a few percent during the cooldown. The base operational temperature was achieved for the stage within standard cooling time. To test the efficacy of the heat switch at millikelvin temperature, a continuous current ($I = 4.6 \times 10^{-4}$ A) was run through the resistor on the stage (89.9 mΩ) and the temperature was measured using the RuOx thermometer. After several hours, the temperature rose asymptotically from a base temperature of 284.6 mK to 311.2 mK. The conductance of the heat switch was then found as $G = I^2 R/\Delta T = 7.2 \times 10^{-7} \text{ W/K}$. This value is competitive, to within an order of magnitude, with other recently developed cryogenic heat switches [137, 138], but, unlike these, has the added advantage of an infinite on/off ratio.

The NIST system is cooled using liquid cryogens, Nitrogen and Helium, and an adiabatic demagnetization refrigerator (ADR). This system is entirely different from the one used at UCSD, which is composed of a pulse tube cooler and
Simon-Chase $^3$He refrigerator. A demonstration of platform use in both cryostats would confirm that the design not only works and is robust, but that it is free of environmental constraints aside from size. Although the NIS platform cooling is still a work in progress, the next few months will be an exciting time for NIS refrigeration. It is expected that before September 2013 the stage will be demonstrated as a viable commercial cooler in both cryogenic systems, that a Cryogenics paper will be submitted for publication, and that the stage design will be presented at the upcoming Low Temperature Detectors Meeting in June 2013.
Chapter 5

ALD Junctions

“Nanotechnology will let us build computers that are incredibly powerful. We’ll have more power in the volume of a sugar cube than exists in the entire world today.” (Ralph Merkle)

The following section is an expanded reprint of a paper submitted to Superconductor Science and Technology in 2013. Over two years at UCSD were spent fabricating NIS tunnel junctions refrigerators without much repeatable success. The original fabrication process mirrored the junction refrigerators developed at NIST, with structure AlMn/AlOx/Al. The fabrication process and some highlights are provided in Appendix A. The setup developed for cryogenic testing is described in Appendix B.

In late 2012, I developed an entirely new process for superconducting tunnel junctions at UCSD that has the potential to produce higher yield and vastly superior junctions than thermal oxidation. The implications for this development extend far beyond refrigeration junctions, improving superconducting junction quality is an active area of interest for many groups worldwide. The design and fabrication process that I developed at UCSD has been submitted through the university as a provisional patent application (2013-242) named “ALD Method for Superconducting Tunnel Junction Fabrication”. The patent application was filed on April 19, 2013 with application number 61/814029.
5.1 Introduction

In recent years, superconducting tunnel junctions have become a viable technology for a wide range of cryogenic applications. There is presently intense interest in solid-state quantum computing based on a range of systems that include superconducting-insulating-superconducting (SIS) trilayers as the building blocks for quantum bits (qubits) in quantum computers [139]. Normal-insulating-superconducting (NIS) junctions have been used as on-chip solid-state refrigerators [128, 129] and more recently as bulk cryogenic coolers [140, 141].

Both SIS and NIS technologies require pristine dielectric barriers a few nanometers thick. These barriers are typically fabricated using thermal oxidation of Al or Al alloys using a controlled combination of temperature, partial pressure of oxygen, and time. However, the diffusive nature of the thermal oxidation process leads to point defects in the tunnel barrier that affect junction quality [142, 143, 144]. These defects result in decoherence in SIS junctions that limit the performance of qubits and subgap leakage in NIS junctions that limits their cooling power. Point defects also place an upper limit on the size and yield of superconducting tunnel junctions. For a fixed specific resistance, the cooling power of an NIS junction refrigerator scales directly with junction area; the ability to fabricate larger junctions would vastly improve refrigeration capabilities. Thus, to improve superconducting tunnel junction technology, it is imperative to eliminate, or at least significantly reduce, point defects in tunnel barriers.

Atomic layer deposition (ALD) has found widespread use in creating high permittivity coatings on the nanometer scale [145, 146] and provides a promising alternative to thermal oxidation. ALD is a chemical vapor deposition process that sequentially deposits atomic monolayers through a series of self-limiting reactions. ALD of the Al oxide, Al$_2$O$_3$, occurs by exposing a heated substrate to alternating pulses of trimethylaluminum (TMA) and H$_2$O in vapor form, separated by a purge of Nitrogen to ensure that the chemicals only react on the sample surface and not in their gaseous states [147]. Each full cycle typically deposits approximately 1 Å of Al$_2$O$_3$, giving sub-nanometer control of film thickness. Further, the sequential reaction used in the ALD process results in complete oxidation [147], potentially...
reducing defects in the barrier and thus allowing for larger area superconducting tunnel junctions.

Utilizing ALD, high quality tunnel barriers can be deposited on materials that do not natively oxidize, such as Au, or metals that form incomplete or conductive oxide layers, such as Cu, Ag, and Ti [148]. Historically, NIS junction refrigerators have used Al or Al alloys, such as AlMn, that are lightly doped with magnetic materials to suppress superconductivity but are still able to form a conformal native oxide tunnel barrier [149]. The use of Au as a normal metal has many benefits: Au remains resistive down to absolute zero and avoids dilute alloy defects such as migration and uneven stoichiometry. ALD also produces a finely tuned barrier thickness that is not strictly limited by the penetration depth of native oxide growth. Although thermal oxide thickness can be increased through mechanisms such as heating, these often result in unwanted effects that limit device performance, such as the clumping and migration of the dilute elements within the alloys.

In this chapter I discuss the demonstration of a large-area (2500 $\mu$m$^2$) NIS tunnel junction with a high quality barrier deposited via ALD on Au with an ultrathin (approximately 1 nm) wetting layer of Al. Direct deposition of ALD $\text{Al}_2\text{O}_3$ on Au is challenging as the ALD process requires hydroxyl bonding on the sample surface and Au lacks surface hydroxyl groups. Although specific growth parameters [148] and hydrous plasmas [150] have been found to deposit ALD on otherwise hydroxyl-free surfaces, another solution is the introduction of a metallic wetting layer, such as Al, that is readily oxidized by a similar process of hydroxylation. This process facilitates hydroxylation via attachment of hydroxyl groups to surface oxygen. For Al wetting layers thinner than the native oxide, the entire wetting layer oxidizes and contributes to the conformal tunnel barrier formed by the ALD $\text{Al}_2\text{O}_3$, forming a true NIS junction with no residual SIS tunneling. To the best of my knowledge, this is the first demonstration of a deposited superconducting tunnel barrier.
Figure 5.1: Schematic of NIS tunnel junction fabrication process. (a) The Ti/Au layer is deposited on a Si wafer coated in SiO$_2$ and lift-off patterned. (b) A SiO$_2$ layer is deposited and wet chemically etched to form the vias. (c) An insulating barrier is formed from an ultrathin layer of Al and Al$_2$O$_3$. In practice, this covers the entire sample surface, but for clarity we show it solely as the junction barrier. (d) A superconducting counterelectrode of Al is deposited and patterned, forming an NIS tunnel junction.

5.2 Sample Fabrication

Figure 5.1 shows a schematic of the junction fabrication process and topology. Devices were fabricated by DC magnetron sputtering a 10 nm adhesion layer of Ti followed (in situ) by a 50 nm Au film on a 2.25 cm$^2$ Si chip coated with 300 nm of SiO$_2$. The base electrode was then patterned via a lift-off process and covered with a 100 nm layer of SiO$_2$ via PECVD deposition at 250° C. The SiO$_2$ was wet etched with a buffered oxide to open vias to the base electrode. The vias were surface cleaned with an Ar sputter etch for 5 minutes at 25 mTorr and 30 Watts RF bias and then approximately 1 nm of Al was deposited without breaking vacuum. The devices were then loaded into the Beneq TFS200 commercial ALD machine after a one minute exposure to atmosphere. A 10 cycle TMA pre-clean, which helps to limit the depth of the interfacial layer between the native Al oxide and Al$_2$O$_3$ [151], was followed by an 18 cycle deposition of Al$_2$O$_3$, both performed at 250° C. Each Al$_2$O$_3$ cycle consisted of a 45 ms pulse of TMA, an 850 ms purge with Nitrogen, a 50 ms pulse of H$_2$O, and another 850 ms purge of Nitrogen. Each cycle of the TMA pre-clean repeated only the first two steps of the process. Post
ALD deposition, vacuum was broken and the sample was loaded into the DC sputtering system. Before depositing a 0.8 µm layer of Al, the sample was heated in vacuum to 105°C for 3 minutes to prevent any residual water vapor from potentially clinging to the junction surface. After deposition, the Al was chemically etched to form the counterelectrode. Devices are formed as shown in Figure 5.1(d).

5.3 Sample Analysis

To demonstrate the viability of the fabrication process for superconducting tunnel junctions, large area tunnel junctions were fabricated with lateral size 50 µm × 50 µm. The tunnel junctions were analyzed via spectroscopy and cryogenic electronic transport measurements.

5.3.1 Spectroscopy

Energy-dispersive X-ray (EDX) spectroscopy was performed locally at the junction interface during each barrier fabrication step. Before and after deposition of the Al wetting layer, the Al peak is not readily identifiable in the resulting spectra; the minute amount of Al provided by the wetting layer does not manifest in the scattering data. After the tunnel barrier has been deposited, however, the Al peak is easily observed, indicating that the Al₂O₃ deposition has adhered to the surface. The change in the scattering data is shown in Figure 5.2. No other new peaks appear in the spectrum post-ALD deposition, showing that the barrier is formed by the Al₂O₃ and not by another mechanism.

5.3.2 Electronic Transport

Cryogenic measurements were used to characterize the tunneling properties of the NIS junctions with ALD deposited insulating barriers. Current-voltage (IV) curves were taken at a range of bath temperatures from 230 mK to 1100 mK. The measured IV curves are shown in Figure 5.3 for four values of temperature, along with theoretical predictions for the curves. Control samples created with an
Figure 5.2: Energy dispersive X-ray spectra taken immediately before and after the Al$_2$O$_3$ barrier was deposited using ALD. Elemental peaks are labeled in the larger plot. The addition of the Al peak between the pre- and post-spectra is easily observed and is the only elemental change to the spectrum. The inset highlights the statistical significance of this change.

identical process, but without the ALD step, exhibited sub-Ohm resistances and linear behavior at superconducting temperatures, confirming that ALD Al$_2$O$_3$ was necessary for barrier formation.

For a superconducting tunnel junction, the predicted tunneling current $I$ in response to a voltage bias, $V_b$, is given by Equation 4.11. The current-voltage curves yield a value of $\Delta = 180 \mu$eV for Al at 230 mK and a normal state junction resistance of 105 kΩ, equivalent to a specific resistance of $R_N A = 2.6 \, \text{M}\Omega \mu \text{m}^2$.

The agreement exhibited between the theoretical model and the measured data indicate that a high quality, conformal tunnel barrier is present. The junctions have a quality factor of $Q = R(V_b = 0)/R_N = 245$ at 230 mK, the lowest base temperature achievable by the $^3$He sorption refrigerator used for junction barrier performance verification. While thermal oxidation has produced higher quality factors at lower specific resistances [152, 153], this is a promising start for tunnel barriers fabricated using ALD.
Figure 5.3: NIS junction IV curves as a function of temperature along with theoretical curves based on the measured stage temperature. The vertical line at 145 μV illustrates a point at which the curves exhibit maximally divergent behavior as a function of temperature. This is the most sensitive thermometry bias point.

5.4 Thermometry

Tunneling in NIS junctions transfers the highest energy electrons from the normal metal to the superconductor, cooling the normal metal. Although normal metal cooling can be utilized for cryogenic applications, the self-cooling of the electron bath in the normal metal limits the potential for thermometry. A self-cooled junction has an equilibrium temperature that is achieved when the junction cooling power is equal to the external power loads. Solving for the base temperature of operation requires a detailed thermal model.

Because the cooling power of a NIS junction is inversely proportional to the normal state resistance, NIS junctions fabricated with high resistance ALD barriers experience no self-cooling, reflect the true bath temperature, and therefore serve as self-calibrated thermometers. Knowing the normal state resistance and the superconducting band gap energy of the junction, one can fit for the only free parameter in Equation 4.10, the electron bath temperature in the normal metal. Current-voltage curves are not required; for a suitable value of the bias where the
curves in Figure 5.3 diverge maximally as a function of temperature (145 µV), one need only measure the resultant current to fit for temperature. This point is highlighted by a vertical line in Figure 5.3.

Using a bias value of 145 µV, the NIS junction electron bath temperature was calculated at a range of temperatures and a calibrated Cernox thermometer was used for comparison. The results are shown in Figure 5.4. The NIS junction data has a standard deviation of 23 mK from the Cernox temperature, which reflects uncertainties in both the Cernox and the NIS junction as well as any errors resultant from the difference in position of the two devices. Because the tunnel barrier is deposited and not natively oxidized, the superconducting material can be chosen almost without constraint, allowing NIS junction thermometers to be tuned for a range of operational temperatures. For instance, a superconducting Nb cathode and Au base electrode forms a self-calibrated thermometer usable in the temperature range of 0 to 9.2 K.

5.5 Future Considerations

For these devices, the film quality at the junction interface was measured to motivate current and future design parameters. At each step in the barrier
deposition process, phase sensitive interferometry (PSI) was used to quantify the roughness of the entire sample surface and scanning electron microscopy (SEM) was used to quantify grain size. The Au grain size was found to be invariant at approximately 30 nm. The roughness of the Au surface did vary; it was found to have a root-mean-square (rms) roughness of 1.73 nm before treatment for the barrier deposition. After the Ar sputter etch, the rms surface roughness decreased to 1.54 nm, indicating that the sputter clean helps to improve surface uniformity. The Al wetting layer was found to have no measurable effect on surface roughness with a rms value of 1.54 nm and, post ALD Al₂O₃, the roughness increased to 1.73 nm. The increase in roughness with ALD Al₂O₃ is attributed to the uneven adherence to Au and Al on the sample surface. Since the Al is non-uniform in its coverage, Al₂O₃ would be deposited more readily in some sites than others, leading to an overall increase in surface roughness. High quality tunnel junction barriers require a thickness that is greater than or equal to the roughness to avoid non-tunneling currents flowing through pinhole gaps. For these samples, the value of 18 cycles (corresponding to approximately 1.8 nm) of Al₂O₃ was chosen such that the thickness of the barrier was the minimum required to exceed the roughness of the Au surface. Because the barrier resistance is an exponential function of the insulator thickness, improvements in the fabrication process that decrease surface roughness would allow for much thinner and far less resistive barriers.

Future endeavors will focus on fabricating NIS refrigerators via modification to the current process. The net cooling power in the normal metal can be optimized by maximizing the junction cooling power and minimizing the external power loads. Decreasing surface roughness will allow for junctions with lower resistance and, as a result, an increased cooling power, \( P \), which scales as \( P \propto R_N^{-1} \). For example, large area junctions (2500 \( \mu \)m\(^2\) ) with low resistance barriers (1 \( k\Omega \mu \)m\(^2\) ) would yield nanowatts of continuous cooling power, an order of magnitude improvement over current NIS refrigerators. Unfortunately, Joule heating sets an upper limit on potential cooling power; a normal metal resistance \( R_p \) gives an external power load of \( I^2R_p \propto R_N^{-2} \). Since the Au used in this process is much less resistive than the standard Al or Al alloys, the upper limit on the cooling power of these junctions is
Figure 5.5: Surface roughness plots of the Au film before any treatment for the tunnel barrier. Two techniques are used: (a) scanning electron microscopy and (b) phase sensitive interferometry. The SEM fits give an Au grain size of 30 nm and the phase sensitive interferometry gives a rms roughness of 1.73 nm.
potentially much higher than has been shown previously. Additionally, (other than substrate size) there is no upper limit on the lateral dimensions of junctions of this design. Room temperature resistance measurements confirmed that all of the 8 fabricated 2500 $\mu m^2$ junctions had formed complete tunnel barriers; the size was chosen arbitrarily for the preliminary mask set. An increase in size would decrease both $R_N$ and $R_p$, and the high yield allows for many junctions to be fabricated in series for increased cooling.

5.6 Conclusions

We have demonstrated a large area superconducting tunnel junction using atomic layer deposition to form a high quality insulating tunnel barrier. The insulating layer has been deposited on Au, showing the viability of using non-oxidizing metals as the electrodes in the future fabrication of NIS refrigerators and SIS qubits. We have also demonstrated the viability of using these devices as precise, self-calibrated thermometers and have highlighted the potential for NIS refrigeration with device modifications. This work presents a pioneering example of superconducting tunnel junction technology with ALD deposited tunnel barriers.

5.7 Acknowledgements

Chapter 5 is a reprint of material as it appears in S. Moyerman, et. al., Atomic layer deposition of superconducting tunnel barriers, Submitted for Publication to Superconductor Science and Technology (2013). The dissertation/thesis author was the primary investigator and author of this paper.
Appendix A

Thermally Oxidized Junctions

Over two years of effort were put into fabricating thermally oxidized NIS tunnel junctions using the NIST material composition: AlMn/AIOx/Al. The following is the final fabrication recipe that was developed, which had an approximate 50% yield on samples, though cooling power was never achieved. It should be noted that this recipe was developed such that no photoresist or unnecessary material ever touched the junction surface, as this caused significant fabrication problems. It is also important that once the AlMn has been deposited, the chip is never heated above 150° and never exposed to heat greater than 100° C for longer than one minute.

1. A four inch wafer is cleaned using a five minute ultrasonic bath in acetone, a three minute ultrasonic bath in isopropyl alcohol (IPA), and a 3 minute ultrasonic bath in deionized water (DI). The wafer is blown dry with N₂. The wafer is surface cleaned with an oxygen plasma at 200 mTorr and 200 Watts for two minutes. From now on, step 1 will be referred to as the “standard cleaning procedure”.

2. The cleaned wafer is coated with 300 nm SiO₂ using the pre-supplied recipe on the Oxford Plasmalab Plasma Enhanced Chemical Vapor Deposition (PECVD) Machine. This dielectric SiO₂ coating allows us to evaluate the samples at room temperature once completed.

3. The wafer is diced by the nano3 staff to 1.5 × 1.5 cm squares.
4. The standard cleaning procedure is performed on a single $1.5 \times 1.5$ cm square chip.

5. The photoresist NR-9 1500PY is patterned onto the chip. Before spinning the resist, the chip is heated to $150^\circ$ C for one minute to help the resist adhere to the surface. The resist is then spun for 40 seconds at 4000 rpm, yielding a thickness of $1.3 \, \mu m$. A pre-exposure bake for one minute $150^\circ$ C was performed before the resist was exposed for 8.5 seconds at a dose of 11 mW/(s cm$^2$). The resist was post-exposure baked for 1 minute at $100^\circ$ C and then developed in resist developer RC6 for 8 seconds, rinsed in DI, and blown dry with nitrogen gas.

6. 50 nm of AlMn is deposited onto the chip using the AJA DC Sputter Deposition Tool at 300 Watts, 2.5 mTorr, and stage height of 30 mm.

7. Lift-off is performed using a five minute ultrasonic bath in acetone. This is followed by the standard cleaning procedure.

8. 100 nm of SiO$_2$ is sputter deposited using the AJA RF Sputter Deposition Tool at 400 Watts, 5 mTorr, and a stage height of 30.

9. Step 5 is repeated except with an exposure time of 4.2 seconds. The resist NR-9 is very sensitive to back-reflected light, and since the AlMn is much more reflective than the Si coated in SiO$_2$, the exposure time needs to be trimmed significantly.

10. The wafer is surface cleaned with an oxygen plasma at 200 mTorr and 200 Watts for one minutes. This is to remove any residual resist from the junction surface that will cause insulating material to remain after the etch.

11. The sample is etched using the standard SiO$_2$ etch recipe programmed into the Oxford P80 Reactive Ion Etcher (RIE). The etch is performed for ten minutes. In practice, it only takes between two and three minutes to etch through the SiO$_2$, but hundreds of failed samples made us paranoid about left-over insulator in the vias, and so the samples were vastly over-etched.
This had no effect on the AlMn below, as the reactive gas (CHF$_3$) used to perform the etch does not affect Al.

12. The standard cleaning procedure is used to remove the photoresist post-etch.

13. The Al cathode material is deposited. Before deposition, the sample is argon etched at 25 mTorr and 30 Watts for ten minutes in the AJA DC Sputter Deposition Machine to remove the native oxide from the sample surface. The sample is then exposed to atmosphere for 1 minute for controlled oxidation with a (time)(pressure) product of 30,000 Torr·s$^1$. The sample is then loaded back into the sputtering chamber and 600 nm of Al is deposited at 300 Watts, 2.5 mTorr, and a sputtering height of 30 mm.

14. The photoresist Shipley 1813 is patterned onto the chip. The resist is spun for 40 seconds at 4500 rpm for a resist thickness of 1.3 µm. The chip is pre-exposure baked for one minute at 115° C, exposed for 11 seconds at a dose of 11 mW/(s cm$^2$), post-exposure baked for 90 seconds at 90° C, developed in MF321 for 30 seconds, rinsed in DI, and blown dry with nitrogen gas.

15. The Al is wet etched using Transene Aluminum Etchant Type-A for 45 seconds at 50° C. It is important to let the Transene heat at 50° C for several hours before the etch so that the temperature can come to equilibrium. Non-ideal temperature causes dramatic changes in the etch rate.

16. The photoresist is removed using the standard cleaning procedure

17. Samples can be tested for performance at room temperature using the Agilent B1500 Semiconductor Device Analyzer. Using the IV curve feature, the room temperature resistance of the junctions should be found.

Figure A.1 shows microscopy images of the sequential steps in the fabrication process. Figure A.2 shows an IV curve taken on a more recent set of junctions.

\footnote{It is worth stating that this process is not ideal – pure, dry O$_2$ gas should be used instead of a mixture and especially one that contains water vapor. Many, many attempts were made to oxidize the junctions in chamber, all of which failed}
Figure A.1: Microscopy images of the sequential steps in the fabrication process for AlMn NIS tunnel junctions: (a) The AlMn deposited and patterned onto a single chip, (b) the chip after the SiO$_2$ has been deposited and etched, and (c) the finished junction after the Al cathode has been deposited and patterned. Two SINIS junctions are formed; the larger area SINIS junction is the refrigerator and the smaller area SINIS junction is the thermometer.
Figure A.2: IV curve of an AlMn SINIS refrigerator at 220 mK. The gap is at $2\Delta = 360 \mu eV$ as expected. The junction has a normal state resistance of 250 Ω.
Appendix B

NIS Junction Testbed

The cryogenic testbed at UCSD is composed of a commercial Cryomech CP1000 Pulse Tube Cooler (PTC) capable of achieving a base temperature of 1.5 K. This is coupled to a commercially available Simon Chase \(^0\)Helium refrigerator known as the “Berkeley Helium 10”. It is a closed cycle Helium 3, Helium 3, Helium 4 refrigerator that is capable of a base temperature of approximately 210 mK with a hold time of 48 hours under normal use.

For testing, NIS junctions are wirebonded into a sample box as shown in Figure B.1. The sample box is RF tight and bolts down to the coldest stage of the cryostat. The sample box accommodates eight sets of SINIS junctions with four wires each. Wirebonds are connected to a PCB which ports them to the outside world via a \(\mu\)D-37 connector. Each line of the PCB is populated with a 15 MHz low-pass filter and a Schottky diode.

The NIS junctions are constant current biased in series with a bias resistor at room temperature. The series resistors range from 20 k\(\Omega\) to 50 M\(\Omega\) and can be switched during testing. One half of the resistance value is added to each side of the junction circuitry to balance the lines, which minimizes the noise. The junctions must then be read differentially to extract the drop in voltage.

The four-wire measurements for NIS junction testing can be shunted to ground so that ESD damage can be prevented when connecting and disconnecting the main cable. The power supply is supplied by either the DAQ board used for acquisition or a manually controlled battery box with switchable polarity. The DAQ
Figure B.1: Sample box that houses the junctions on the cryogenic stage. The lid makes an RF tight seal with the bottom and contains a shield between the wires to the $\mu$D and the samples for radiative protection. The PCB is equipped with low pass filters and diodes on the underside.
board is capable of supplying a triangle wave signal with a peak-to-peak amplitude ranging from 0.02 V to 20 V. The output signal is amplified in series by operational amplifiers with a gain of two and a gain of 500 and low pass filtered via an LC circuit. The entire circuit is housed inside of an RF tight electronics box that was specially machined for this purpose. The room temperature electronics are shown in Figure B.2. The setup was structured such that two SINIS junctions could be measured simultaneously.
Bibliography


Roger Christopher O’Brient. *A Log-Periodic Focal-Plane Architecture for...*


[39] A. van Engelen, R. Keisler, O. Zahn, K.A. Aird, B.A. Benson, et al. van Engelen, a. and Keisler, r. and Zahn, o. and Aird, k. a. and Benson, b. a. and Bleem, l. e. and Carlstrom, j. e. and Chang, c. l. and Cho, h. m. and Crawford, t. m. and Crites, a. t. and de Haan, t. and Dobbs, m. a. and Dudley, j. and George, e. m. and Halverson, n. w. and Holder, g. p. and Holzapfel, w. l. and Hoover, s. and Hou, z. and Hrubes, j. d. and Joy, m. and Knox, l. and Lee, a. t. and Leitch, e. m. and Lueker, m. and Luong-Van, d. and McMahon, j. j. and Meh, j. and Meyer, s. s. and Millea, m. and Mohr, j. j. and Montroy, t. e. and Natoli, t. and Padin, s. and Plagge, t. and Pryke, c. and Reichardt, c. l. and Ruhl, j. e. and Sayre, j. t. and Schaffer, k. k. and Shaw, l. and Shirokoff, e. and Spieler, h. g. and Staniszewski, z. and Stark, a. a. and Story, k. and Vanderlinde, k. and Vieira, j. d. and Williamson, r. *Astrophys.J.*, 756:142, 2012.


[97] Michael L. Brown, Anthony Challinor, Chris E. North, Bradley R. Johnson,


[100] H. Kamerlingh-Onnes. The superconductivity of mercury, 1911.


