Studying the Cosmic Microwave Background with

Spider’s First Flight

Edward Young

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance
by the Department of
Physics
Adviser: Professor William C. Jones

June 2018
Abstract

Gravitational waves are a prediction of many early Universe models. These waves generate a divergence-free polarization component, called $B$-modes, of the cosmic microwave background (CMB). Spider is a balloon-borne telescope designed to study the polarized emission from the CMB.

The Spider payload flew in January 2015, with detectors at 95 GHz and 150 GHz. This thesis focuses on the development and deployment of the SPIDER-1 payload and the analysis of the science data. In the hardware section, this thesis pays particular attention to the development of the flight cryostat and the implementation of an electrical gain monitor for characterizing time variable detector gain. In the analysis section, this thesis focuses on lowlevel data cleaning and calibrations through deprojection. In the final section, techniques for eliminating foreground contributions to CMB maps and the limitations of those techniques are discussed.

Our understanding of foregrounds currently limits our ability to measure $B$-modes from the CMB. SPIDER-2, with its 280 GHz channel, will be a powerful tool in constraining foregrounds.
Acknowledgments

I think everyone presupposes that Princeton professors are brilliant, and Bill certainly radiates the genius that we expect. But the truly remarkable thing about Bill is that he couples that brilliance with an equal kindness. Bill truly cares about his students and all those who work with him. And for that, more than anything else about my graduate school experience, I am grateful.

Perhaps Bill’s greatest achievement at Princeton was convincing Jón and Sasha to join his nascent lab. Jón and Sasha taught me how to do science, with more patience than I ever deserved. Working on SPIDER is always a trial by fire, but most days Jón and Sasha made it seem like a campfire rather than a conflagration. So much of my life as a grad student has been an attempt at imitating the two of them; this thesis is no exception.

Anne, somehow we got through the last seven years without hating each other. I would have already considered that a win. We managed to fail prelims, retake them, learn to do science, find a leak in the cryostat, realize we didn’t, stycast the main tank, deploy the instrument, survive Jon and Sasha forsaking us, analyze the data, and deal with all the life that happened between. I am fortunate to share all of that with someone I consider one of my best friends.

Aurelien has always been a voice of nervous calmness at Princeton for me. He always seems to have a proposal deadline due in two days and grades due the day after. Despite this, he always lets me intrude and answers my analysis questions.

I’m honored Ziggy followed my footsteps from Berkeley to Princeton. He and I are apparently in an unending struggle to speak faster and faster, something I never noticed until Sasha pointed it out. I think that’s because I knew he has so much knowledge that I wanted to learn. Ziggy has always been a wellspring of not only scientific knowledge, but also life advice.

I will miss early morning rowing with Steve. I will always be proud of our third place finish in the Thanksgiving row. Thank you for teaching me about mapmakers and pie making. Also, after all the times I’ve crashed on your extra bed, you have infinite dibs on staying at my place no matter where in the world I am.

The Princeton crew has never been stronger. Stevie, Sherry, Joseph, Steven, and Corwin have
been amazing fellow grad students and coworkers. And the crew will be all the better when Vy and Susan join. I have no doubt SPIDER-2 will be a resounding success with your leadership. And I hope you all have as much fun in the process as I did.

The Princeton crew is but a subset of what I consider the greatest scientific collaboration. I recently was at a conference with Jeff and Sean, and someone apropos of nothing said “You SPIDER folks spent three months in Antarctica together and still want to hang out. You must really like each other!” And it’s true. It’s always a joy to see each and every one of you, whether at a conference or just hanging out in a backyard.

Perhaps the most surprising outcome of graduate school is that there is now one dog in the world I like. I don’t know why she followed me that one day in Palestine, but I’m glad she did. Laika, you are the best thing to come out of Texas.

My time in Jadwin would not have been the same without Angela and Ted. Angela solved every one of my administrative tasks while simultaneously listening to me complain about my life. I already miss swinging by her office and chatting. For every bit of emotional support Angela has provided me, Ted has provided me with an equal amount of scientific support. First as the stock room manager and then as the purchasing manager, Ted has always gone out of his way to help me. I always looked for excuses to come by your office and hang out. Also, I am still honored that you two invited us to your wedding.

If I could live my life with half the joy that Darryl lives his with, I would be a happy man. Thank you for always brightening my day with your optimism. I hope you have the best one.

I have always loved Jim’s approach to teaching lift and helium safety. It is five parts experience, five parts knowledge, and ten parts mockery. Jim has saved SPIDER multiple times by conjuring helium from nothing. I think some of his magic helped the Eagles win the Super Bowl.

Julio is a younger, better looking version of Ted. And now the stock room is better than ever. Thank you for all your help. I hope your Raiders beat Jim’s Eagles in the next Super Bowl.

If there is one truth in the Universe, it is that machine shop guys hate sentimentality. So I hope Steve reads this. Thank you for teaching me how to safely and efficiently work in the machine shop. It’s a running joke in SPIDER that we will come to you with a problem, you’ll set up the
solution, show us how to do it, and by that point the project will be 95% done. SPIDER certainly would not have worked with out your help.

I am particularly grateful for everyone who read this text and found numerous typographical and mathematical errors. Jenn was the first to read this thesis, and caught every alignment issue, spelling mistake, and missing Oxford comma. Uncle Kenneth helped clarify the introduction. Sasha helped on the detector section. And Norm lent his incomparable experimental expertise to the entire thesis.

The British Antarctic Survey kept me alive for nearly a month in the harshest environment I have ever been in. I literally owe you my life. Also, more importantly, SPIDER owes you all our focal planes. In particular, Sam Burrell was crucial to the success of the recovery effort.

I am lucky to have shared my physics graduate school experience with so many wonderful people. Sara has always tortured me with her puns. And when that didn’t cause me enough pain, Shawn would jump in. It was the worst. Kenan knew more physics at age ten than I do now, and he has always enthusiastically explained basic concepts to me. Guangyong’s dedication to scoping out free food is unrivaled. I am proud that a department-wide email was sent because of us, chastising students for taking free food before explicit notification. Katie might as well have been a physicist. Thank you for always making Princeton a fun place with your themed events; I am more than happy to lose more rounds of Game of Thrones trivia. Vuk, Anthony, and Jack had the good sense to leave Princeton early. For that I am incredibly bitter. The three of them made Princeton physics a much better place.

I was lucky to live with Aaron, Jordan, Pawel, Omer, and Yasmin. Thank you Aaron for making me paranoid about Internet security before it was cool. Also thank you for dragging me into the moguls while I was on a snowboard. I think my tailbone still hurts. Jordan likes to believe he’s the best at lawn games. He’s not. But I’ve always found his confidence endearing. I’m excited to be on the West Coast with him. I hope Pawel also joins us. I still think you and Tanya are crazy for letting me officiate your wedding, but I will forever remember that moment. Thank you Omer and Yasmin for being huge improvements over Aaron and Pawel.

Our house was always filled with the best of friends. Alex, Julia, and Baxter were over almost
every weekend. I can’t wait to start FYI with you all. Tanya makes everyone around her a better person with her sheer force of good will. We need to perfect our smoked coconut recipe so we can recreate that BLT. Alison and Rainy brought more energy into our house than everyone else combined. I have learned so much about the digestive track from the both of you. Harvey and Daniel have always been my favorite couple. Thank you for teaching me that ball is life.

Thank you Ojas, Savannah, Eldon, and Lindsay for being crazy enough to fly out to celebrate my finishing graduate school. My brain shut down when you guys surprised me outside Jadwin. Your friendship and support from afar has helped me through this entire process.

One late August day in 2014, I was riding on the 7 train headed into Manhattan, talking to a girl I had met earlier in the day. I thought we were having a great conversation, when I noticed that she had fallen asleep. Now we talk every night before falling asleep together. Thank you Jenn for bringing so much love into my life. Neither poetry nor science could adequately describe the joy I find spending every day with you. Thank you for giving me the immeasurable. Wherever the next adventure lies, I am so lucky to have my best friend by my side.

Mom and Dad, I owe everything to you. “Thank you” doesn’t even begin to express the depths of my gratitude, but thank you nonetheless. Dad, I am a scientist today because of you. From my first day in lab, which I certainly do not remember because I was less than a year old, I have been inspired by your curiosity. Thank you for always encouraging me to find the answers on my own. Mom, I am the person I am today because of you. The older I get, the more I realize that your caring is unrivaled in this world. Thank you for all the love that a child could ever want. I am forever grateful for our family.
Collaboration

Figure 1: A picture of the SPIDER deployment team in 2014-2015. From left to right: Ivan Padilla, Jamil Shariff, Edward Young, Alexandra Rahlin, Jeffrey Filippini, Lorenzo Moncelsi, Anne Gambrel, Cynthia Chiang, Zigmund Kermish, Sean Bryan, Johanna Nagy, John Ruhl, Natalie Gandillo, William Jones, Don Wiebe, Barth Netterfield, Jón Emil Guðmundsson, Steven Benton, John Hartley.
Contents

Abstract ii

Acknowledgments iii

Collaboration vii

Contents viii

List of Figures xv

List of Tables xx

I Introduction 1

1 Modern Cosmology 2

1.1 The Fundamentals of Cosmology 3

1.1.1 The Friedman Equations and the Metric 3

1.2 History of the CMB 6

1.3 Inflation 10

1.3.1 Motivating Problems 10

1.3.2 Solving Problems with Inflation 12

1.3.3 The $\phi^2$ Potential 12

1.4 Alternatives to Inflation 14

1.5 Measurables 15
CONTENTS

5.3.2.2 Noise Correlation Matrix ..................................................... 81

6 Calibration .................................................................................. 97

6.1 Beam and Calibration ................................................................. 97

6.2 Deprojection ............................................................................ 99

6.2.1 Differential Gain ................................................................. 101

6.2.2 Differential Pointing ............................................................. 103

6.2.3 Differential Beam Width ...................................................... 105

6.2.4 Differential Ellipticity .......................................................... 107

6.2.5 Template Fitting ................................................................. 110

6.3 Deprojection Results ............................................................... 111

6.3.1 Detector Pointing ................................................................. 112

6.3.2 Boresite Pointing ............................................................... 112

6.3.3 Beam Parameters .............................................................. 114

6.4 MCE and Pointing Solution Timing .......................................... 120

7 Mapmaking ................................................................................. 124

7.1 Mapmaker Basics .................................................................... 124

7.1.1 Stokes Parameters ............................................................... 125

7.1.2 Polarimeter Model .............................................................. 126

7.1.3 Half-wave plate ................................................................ 128

7.1.4 Linear Mapmaking ............................................................. 129

7.1.5 Angular Coverage and Condition Numbers ....................... 132

7.2 SPIDER Mapmaker ................................................................. 136

7.2.1 Filtering ............................................................................. 136

7.2.2 Transfer Function ............................................................. 136

7.2.3 SPIDER Maps ................................................................. 138

7.3 Noise Simulation .................................................................... 138

7.3.1 Uncorrelated Noise .......................................................... 141
CONTENTS

7.3.2 Correlated Random Noise .......................................................... 141
7.3.3 SPIDER Noise Maps ................................................................. 143
7.4 Power Spectrum ................................................................. 144
  7.4.1 Power Spectrum Basics ......................................................... 144
  7.4.2 Cut Sky .................................................................... 150
7.5 Null Tests ................................................................... 152

IV SPIDER First Flight Results ...................................................... 157

8 Foregrounds ........................................................................ 158
  8.1 Foreground Components .......................................................... 158
    8.1.1 Dust ................................................................ 159
      8.1.1.1 Thermal Dust .................................................... 159
      8.1.1.2 Spinning Dust .................................................... 161
    8.1.2 Synchrotron .............................................................. 161
    8.1.3 Free-Free ................................................................. 162
  8.2 Component Separation ............................................................ 162
    8.2.1 Map Based Component Separation ..................................... 163
      8.2.1.1 Map Based Component Separation Theory .............. 163
      8.2.1.2 Generating Foreground Templates ....................... 165
      8.2.1.3 Choosing Templates ........................................... 169
      8.2.1.4 Joint Fitting IQU ............................................... 172
      8.2.1.5 Spatial Variations ............................................. 173
      8.2.1.6 Template Noise Variance and Bias ..................... 173
      8.2.1.7 CMB Map Noise Variance and Bias .................... 177
      8.2.1.8 CMB Signal Variance and Bias ......................... 178
      8.2.1.9 Map Based Foreground Variance and Bias Summary 179
      8.2.1.10 Relating Fit Parameters to $\beta_d$ and $T_d$ ............ 180
## CONTENTS

8.2.2 Spectral-Based Component Separation .......................... 181

8.2.2.1 Spectral-Based Component Separation Theory .................. 181

8.2.2.2 A Pedagogical Example ......................................... 183

8.2.2.3 Spectral-Based Component Separation with SPIDER ............... 184

9 Conclusion ............................................................. 188

V Appendix .............................................................. 190

A Amplified Noise Source .............................................. 191

A.1 Design Overview .................................................. 191

A.2 Components ....................................................... 192

B Carbon and Steel Stycast ............................................ 194

B.1 Power Attenuation ................................................ 194

B.2 Index of Refraction ................................................. 197

C Derivative Maps ....................................................... 199

D Plotting Beam Ellipticity ........................................... 203

E Explorer’s Notes ..................................................... 205

F IR Dewar ............................................................... 209

F.1 The Cryostat ....................................................... 209

F.2 Cooling Down ...................................................... 211

G Fluctuations in an Expanding Universe ......................... 214

G.1 Minkowski Metric ................................................ 214

G.2 deSitter Metric .................................................... 220

H Beam Parameters .................................................... 224
CONTENTS

Bibliography 243
# List of Figures

1. **Deployment of SPIDER in 2014-2015** ........................................ vii

1.1 The FIRAS CMB monopole spectrum ......................................... 7

1.2 CMB experiments ................................................................. 9

1.3 $BB$ spectra for various values of $r$ ...................................... 16

1.4 $\tau$ spectra for various values of $\tau$ .................................... 17

2.1 Cryostat cross-sectional view .................................................. 21

2.2 Vacuum window ................................................................. 22

2.3 Window bucket diagram ........................................................ 23

2.4 Hermetic conflat ................................................................. 24

2.5 MT, VCST, and VCS2 temperatures through flight ........................ 25

2.6 Stycast stripe on MT ............................................................ 26

2.7 SFT .................................................................................. 28

2.8 $^3$He fridge schematic ......................................................... 29

2.9 Fridge load curve ............................................................... 31

2.10 Pressure through successive pumpdowns .................................. 32

2.11 Cross-sectional view of SPIDER optical stack ........................... 35

3.1 Bolometer model ................................................................. 38

3.2 TES resistance as a function of $I$ and $T$ ................................. 40

3.3 The bias circuit ................................................................. 41

3.4 Bias step example ............................................................... 49
LIST OF FIGURES

3.5 Bias circuit diagram of TES coupled to a SQUID ........................................... 50
3.6 SQUID $V\Phi$ curve ......................................................................................... 51
3.7 A focal plane .................................................................................................... 52
3.8 Helium film noise on a detector ......................................................................... 53
3.9 Reaction wheel noise template ........................................................................... 54
3.10 Reaction wheel amplitude in physical coordinates ............................................. 55
3.11 An example of RF pickup .................................................................................. 56
3.12 An example of cross jumps ................................................................................ 57
3.13 An example of scan synchronous noise in the time domain ............................... 58
3.14 The scan synchronous noise amplitude through flight ...................................... 59

4.1 Flight path .......................................................................................................... 66
4.2 SPIDER on the launch pad .................................................................................. 66
4.3 Flight altitude ..................................................................................................... 67
4.4 SPIDER sky coverage ......................................................................................... 68
4.5 Payload landing site ............................................................................................ 69
4.6 The main tank being built and scrapped .............................................................. 69

5.1 PSD of a simulated signal timestream ................................................................. 83
5.2 Gaussian derivative kernel example .................................................................... 84
5.3 Transfer function of the triangle kernel ............................................................. 84
5.4 Gaussian derivative kernel convolved with SPIDER data ................................... 85
5.5 Cross jump matrix ............................................................................................. 86
5.6 Example of timestream cleaning ........................................................................ 87
5.7 Power spectral density of a SPIDER bolometer ................................................ 87
5.8 Correlated Noise FFT ....................................................................................... 88
5.9 Example noise correlation matrix ....................................................................... 89
5.10 Another example noise correlation matrix ......................................................... 90
5.11 Full X1 correlation matrix ................................................................................ 91
### LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.12</td>
<td>Full X2 correlation matrix</td>
<td>92</td>
</tr>
<tr>
<td>5.13</td>
<td>Full X3 correlation matrix</td>
<td>93</td>
</tr>
<tr>
<td>5.14</td>
<td>Full X4 correlation matrix</td>
<td>94</td>
</tr>
<tr>
<td>5.15</td>
<td>Full X5 correlation matrix</td>
<td>95</td>
</tr>
<tr>
<td>5.16</td>
<td>Full X6 correlation matrix</td>
<td>96</td>
</tr>
<tr>
<td>6.1</td>
<td>SPIDER beams at 95 GHz</td>
<td>98</td>
</tr>
<tr>
<td>6.2</td>
<td>SPIDER beams at 150 GHz</td>
<td>98</td>
</tr>
<tr>
<td>6.3</td>
<td>Detector pointing correction from deprojection</td>
<td>113</td>
</tr>
<tr>
<td>6.4</td>
<td>Boresite pointing correction from deprojection</td>
<td>114</td>
</tr>
<tr>
<td>6.5</td>
<td>Pointing offset map difference</td>
<td>115</td>
</tr>
<tr>
<td>6.6</td>
<td>Variable beam width in physical coordinates</td>
<td>116</td>
</tr>
<tr>
<td>6.7</td>
<td>Variable ellipticity in physical coordinates</td>
<td>117</td>
</tr>
<tr>
<td>6.8</td>
<td>Physical representation of elliptical beams</td>
<td>118</td>
</tr>
<tr>
<td>6.9</td>
<td>BB spectra from spurious beam parameters</td>
<td>119</td>
</tr>
<tr>
<td>6.10</td>
<td>Left-right null maps</td>
<td>121</td>
</tr>
<tr>
<td>6.11</td>
<td>Pointing offset optimization</td>
<td>123</td>
</tr>
<tr>
<td>7.1</td>
<td>Q and U Stokes vectors</td>
<td>125</td>
</tr>
<tr>
<td>7.2</td>
<td>SPIDER condition numbers</td>
<td>137</td>
</tr>
<tr>
<td>7.3</td>
<td>Filter transfer function and beam</td>
<td>138</td>
</tr>
<tr>
<td>7.4</td>
<td>The SPIDER 95 GHz map</td>
<td>139</td>
</tr>
<tr>
<td>7.5</td>
<td>The SPIDER 150 GHz map</td>
<td>140</td>
</tr>
<tr>
<td>7.6</td>
<td>Uncorrelated noise map</td>
<td>145</td>
</tr>
<tr>
<td>7.7</td>
<td>Correlated noise map</td>
<td>146</td>
</tr>
<tr>
<td>7.8</td>
<td>Correlated and uncorrelated noise map difference</td>
<td>147</td>
</tr>
<tr>
<td>7.9</td>
<td>Uncorrelated and correlated noise spectra</td>
<td>148</td>
</tr>
<tr>
<td>7.10</td>
<td>SPIDER masks</td>
<td>154</td>
</tr>
<tr>
<td>7.11</td>
<td>Interleaved chunks</td>
<td>155</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

7.12 Interleaved chunk null spectrum ........................................ 155
7.13 Interleaved chunk null maps ............................................. 156

8.1 Dust emissivity as a function of $\beta_d$ at $T_d$ ....................... 160
8.2 Example foreground templates ........................................ 167
8.3 Foreground template noise .............................................. 168
8.4 The re-smoothed foreground templates .................................. 170
8.5 Foreground template freq ................................................ 171
8.6 Spatial variation of fit parameters at 95 GHz .......................... 174
8.7 Spatial variation of fit parameters at 150 GHz ......................... 175
8.8 Foreground fit bias from template noise ................................ 176
8.9 Foreground fit uncertainty from SPIDER noise ....................... 177
8.10 Foreground fit uncertainty from CMB correlation .................... 178
8.11 SMICA frequency weights .............................................. 186
8.12 SMICA foreground parameters ........................................ 187

9.1 95 GHz and 150 GHz beam stacks .................................... 189

A.1 Amplified noise source schematic ......................................... 192
A.2 Tripler efficiency .......................................................... 193

B.1 Reflectivity measurement setup ........................................ 195
B.2 Reflectivity measurement samples ...................................... 195
B.3 Index of refraction of impregnated stycast ......................... 198

C.1 Planck T and first derivative maps ...................................... 201
C.2 Planck second derivative maps ........................................ 202

D.1 Schematic of beam ellipticity .......................................... 204

E.1 The recovery crew ......................................................... 208
LIST OF FIGURES

F.1 IR Dewar cross section ..................................................... 210

H.1 X1 beam widths in physical coordinates ........................................ 225
H.2 X2 beam widths in physical coordinates ........................................ 226
H.3 X3 beam widths in physical coordinates ........................................ 227
H.4 X4 beam widths in physical coordinates ........................................ 228
H.5 X5 beam widths in physical coordinates ........................................ 229
H.6 X6 beam widths in physical coordinates ........................................ 230
H.7 X1 ellipticities in physical coordinates .......................................... 231
H.8 X2 ellipticities in physical coordinates .......................................... 232
H.9 X3 ellipticities in physical coordinates .......................................... 233
H.10 X4 ellipticities in physical coordinates ........................................ 234
H.11 X5 ellipticities in physical coordinates ........................................ 235
H.12 X6 ellipticities in physical coordinates ........................................ 236
H.13 The X1 \( c \) parameter for the first and second half of flight .......... 237
H.14 The X2 \( c \) parameter for the first and second half of flight .......... 238
H.15 The X3 \( c \) parameter for the first and second half of flight .......... 239
H.16 The X4 \( c \) parameter for the first and second half of flight .......... 240
H.17 The X5 \( c \) parameter for the first and second half of flight .......... 241
H.18 The X6 \( c \) parameter for the first and second half of flight .......... 242
List of Tables

2.1 Approximate pumpdown timeline ........................................... 31
2.2 Focal plane list ..................................................................... 34
2.3 Filter list .............................................................................. 36

3.1 Detector parameters .............................................................. 49

4.1 The latitude and longitude of important locations ................. 64

6.1 Triple Gaussian beam parameters .......................................... 99
6.2 Deprojection templates .......................................................... 111

7.1 $f_{\text{sky}}$ for various masks .................................................. 151

8.1 Map based foreground fit parameters ..................................... 169
8.2 Map based foreground fit parameters with $I$ and $P$ separate .... 172
8.3 ......................................................................................... 179

B.1 Reflectivity measurements of impregnated stycast ............... 196
B.2 Reflectivity of impregnated stycast as a function of incident angle 197

xx
Part I

Introduction
Chapter 1

Modern Cosmology

Physics undergrads at Berkeley take 111 lab their Junior year. By this point, I had already worked for a year and half on a supernova satellite proposal that seemed to be going nowhere. I was ready to move on from cosmology.

My Junior year, 111 was taught by Bill Holzapfel. I was doing fairly well in the class, so I asked Holzapfel for some advice. I told him I was looking for a lab to work in, suggesting perhaps alternative energy research or quantum computing. His counterproposal was for me to work in his lab, which happened to be building a CMB telescope.

The next week, I walked into the lab for the first time. Liz and Tom were blasting Girl Talk. These were my people. For the next two years, I worked as much as possible in Holzapfel’s lab. I learned everything from soldering to SQUID tuning to wire bonding to cryogenics. I had such a blast. And I have been having a blast ever since.
1.1 The Fundamentals of Cosmology

From the earliest recorded history, and certainly before that, humans have considered the origins and evolution of the world around them. The earliest notions of cosmology were philosophical or religious in nature. Modern cosmology is the hubristic belief that we can apply the scientific method to these notions. That the entire history of the Universe can be described by simple, understandable physical laws is perhaps the most arrogant assumption a scientist can make. This thesis is an exercise in that arrogance.

1.1.1 The Friedman Equations and the Metric

To temper that arrogance, we remind ourselves that much of our understanding of the Universe comes from the mediocrity principle, the knowledge that our space in the Universe is in no way special. If in fact this is true, then the Universe must be homogeneous and isotropic on the largest scales. Then from Einstein’s theory of relativity [1], we can write the Friedman–Lemaître–Robertson–Walker (FLRW) metric [2–5] in spherical polar coordinates

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

(1.1)

where we have adopted unit convention with the speed of light as \( c = 1 \). The curvature of the Universe is described by \( k \), with values of \(-1\), \(0\), and \(1\) for open, flat and closed universes respectively. The expansion (or contraction) of the Universe is described by changes in the scale factor \( a(t) \). We can simplify the expression in eq. (1.1) by rewriting

\[ \frac{dr^2}{1 - kr^2} + r^2 \left( d\theta^2 + \sin^2 \theta d\phi^2 \right) = a^2 \gamma_{ij} dx^i dx^j \]  

(1.2)

where \( x^i \) is called comoving coordinate. The comoving distance describes the distance between two objects in a set of units that are invariant due to the expansion. The physical coordinate \( x_p \) is
1.1.1 The Friedman Equations and the Metric

related to the comoving coordinate simply by

$$x_p^i = a(t)x^i.$$  \hspace{1cm} (1.3)

The physical velocity of an object, a quantity we can actually measure, is

$$v_p = \dot{x}_p^i = a\dot{x}^i + \ddot{a}x^i = v_{\text{pec}} + Hx_p^i$$  \hspace{1cm} (1.4)

where we have defined the peculiar velocity $v_{\text{pec}}$ and the Hubble parameter $H = \dot{a}/a$. The second term on the right hand side of the equation is the Hubble flow, which describes the motion of objects due purely to expansion. We use the common notation where a dot above a parameter implies its time derivative, $\dot{f} \equiv df/dt$. Observational data shows that the second term on the right hand side eq. 1.4 dominates the first term at distances greater than approximately ten megaparsecs. Ten megaparsecs is where the local gravity of our galaxy is subdominant to the Hubble Flow. Thus, eq. 1.4 is simply the Hubble Law with corrections due to peculiar velocities. The peculiar velocities of galactic bodies are mean zero, so in aggregate do not contribute to measures of $H$.

Henrietta Leavitt’s discovery of the period luminosity relationship of Cepheid variable stars \[6\] enabled the first measurement of the Hubble constant by Edwin Hubble in 1929 \[7\]. The Hubble constant is not actually constant since $a$ is a function of time. Therefore, we denote the current value of $H$ as $H_0$. Hubble found the value $H_0 = 500$ km/s/Mpc, which differs by nearly an order of magnitude from the current best estimate of approximately $H_0 = 73$ km/s/Mpc \[8, 9\]. Hubble's measurement was the first direct evidence that the Universe itself is not static. And if the Universe is not static, then the Universe must have an age.

The FLRW metric describes the geometric properties of the Universe. The Einstein field equations connect them to the energy content of the Universe. This generates the Friedmann equations. The first Friedmann equation is

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{k}{a^2}$$  \hspace{1cm} (1.5)
and the second Friedmann equation is

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P) \quad (1.6)$$

where $G$ is the gravitational constant, $\rho$ is the energy density from all sources in the Universe, and $P$ is the pressure from all sources. The left side of eq. 1.5 is just $H^2$. For a spatially flat Universe $k = 0$, so from eq. 1.5 we can define the critical density as

$$\rho_c = \frac{3H^2}{8\pi G}. \quad (1.7)$$

Taking the derivative of eq. 1.5 and plugging the first and second Friedmann equations into it, we find the continuity equation

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + P) = 0. \quad (1.8)$$

This will help us derive a relationship between density and scale factor. Consider matter, which has pressure much less than energy density. This lets us set $P = 0$ in the continuity equation, and we get

$$\int \frac{d\rho}{\rho} = -3 \int \frac{da}{a} \quad (1.9)$$

which of course integrates to give the relationship

$$\rho_m \propto a^{-3}. \quad (1.10)$$

For radiation and all relativistic particles, $P = \rho/3$. Solving the exact same integral, we find

$$\rho_r \propto a^{-4}. \quad (1.11)$$

The recently discovered dark energy has the pressure/density relationship $P = -\rho$. This gives

$$\rho_\Lambda \propto a^0. \quad (1.12)$$
where $\Lambda$ denotes dark energy. So the energy density of dark energy is constant with scale factor.

For simplicity, we define the relative density

$$\Omega_i = \frac{\rho_i}{\rho_c} \quad \text{(1.13)}$$

where $i$ indexes the component. Now we can rewrite the Friedmann equation as a function of the constituent parts of the Universe

$$H^2(a) = H_0^2 \left[ \Omega_r \left( \frac{a_0}{a} \right)^4 + \Omega_m \left( \frac{a_0}{a} \right)^3 + \Omega_k \left( \frac{a_0}{a} \right)^2 + \Omega_\Lambda \right] \quad \text{(1.14)}$$

where we have also defined $\Omega_k \equiv -k/(a_0 H_0)^2$.

We have glossed over an incredibly important topic, dark energy. In fact, the relationship $P = -\rho$ is astounding and should stop one in one’s tracks. This relationship is in some sense not fundamental, although it is certainly true. The underlying nature of dark energy is that it is constant throughout space and time. As far as we can tell, at every moment in the history of the Universe, the energy density of dark energy is the same everywhere. From a theoretical standpoint, we derived eq. 1.12 backwards. We should have started with $\rho \propto a^0$ and then derived the corresponding pressure.

There is ample evidence for the existence of dark energy. The earliest direct measures of dark energy come from Type-1a supernova measurements which directly probe the Hubble parameter with standard candles. However, concerns were raised by the cosmic microwave background (CMB) community before that. And both CMB and baryon-acoustic oscillation measurements have put strong constraints on the expansion rate of the Universe.

### 1.2 History of the CMB

A continuously expanding Universe implies that the Universe was smaller and denser in the past. As we run the clock backwards from today, the temperature and density of the Universe rises. And if we run it back far enough, the Universe was too hot for even neutral hydrogen. In this stage in
1.2 History of the CMB

The Universe, the primordial plasma made it impossible for photons to free stream. Photons and electrons were locally in thermal equilibrium due to frequent collisions.

Now running the cosmic clock forward from the era where photons and electrons were coupled, the Universe cools, eventually allowing neutral hydrogen to form. Photons are no longer inhibited by free electrons and protons. The photons are allowed to free stream. This period is called recombination. These photons have been free streaming since decoupling, cooling as described by eq. 1.1. An observer sees photons from every direction because photons decoupled in all points in space. If we define the time from recombination to now as $t_{\text{recomb}}$, then all photons that were emitted $ct_{\text{recomb}}$ away arrive at the observer, forming a sphere. This sphere is called the surface of last scattering.

The first direct measurement of the CMB came from Penzias and Wilson in 1965 [12, 13], followed shortly by a measurement by Dicke, Peebles, Roll, and Wilkinson [14]. The CMB spectrum measured to great precision by the Cosmic Background Explorer (COBE) FIRAS mission. The spectrum is shown in Figure 1.1. This is the most perfect black body ever measured in nature and represents one of the strongest pieces of evidence for the Big Bang theory. The best measure of the black body temperature now comes from a joint analysis of WMAP and FIRAS, $2.72548 \pm 0.00057$ K [15].

Remarkably, the CMB has even more information to give us. Imprinted upon the black body
1.2 History of the CMB

is a set of fluctuations that are approximately 100 $\mu$K in amplitude. These fluctuations become visible after removing the CMB dipole generated by the Earth’s peculiar velocity relative to the CMB. The fluctuations are generated by slight overdensities and underdensities in the original plasma. The fluctuations have a characteristic size, determined by the fluid dynamical properties of the plasma. Since the fluid dynamical properties are consistent across the entire universe, this allows us to make a statistical statement about these fluctuations. Measuring these fluctuations, or anisotropies, has been one of the most fruitful endeavors in modern science.

The successor to COBE is the Wilkinson Microwave Anisotropy Probe (WMAP) [17], which had 45 times the sensitivity of COBE and 33 times the angular resolution. This allows WMAP to measure CMB anisotropies with far greater precision and fidelity, which helped cement the dark energy dominated, cold dark matter model of the universe as the leading cosmological model. This model is commonly referred to as $\Lambda$CDM, where $\Lambda$ is the symbol for dark energy and CDM is short for “cold dark matter”.

The successor to WMAP is Planck, the third generation space-based CMB telescope [18]. Planck has even greater sensitivity and angular resolution than WMAP, creating the best all-sky maps available to date. Additionally, Planck flew 9 frequency channels. This will play a crucial role in component separation, which is discussed in Section [8.2].

While the three satellite missions function as pillars in the CMB measurement community, that does not preclude important work happening on the ground and from balloons. Major discoveries were made from more terrestrial experiments. The Saskatoon experiment detected CMB anisotropies from the ground [19]. The QMAP and TOCO experiments were the first to measure the rise and fall of the CMB angular power spectrum [20]. BOOMERanG [21] and MAXIMA [22] were balloon payloads that measured the CMB temperature anisotropies. ACT and SPT have measured the CMB out to multipoles of several thousand, and their follow-up missions ACTPol and SptPol have added polarization sensitivity. There are numerous other experiments which have added valuable data on the CMB. An incomplete list starting from 1982 up until today is shown in Figure [1.2]
Figure 1.2: The approximate time frames of CMB experiments since 1982. Penzias and Wilson’s first measurements were made in 1965. The start date is broadly defined as first light. The end date is a vague combination of last data and end of analysis. All of the data is from [23].
1.3 Inflation

The success of modern cosmologies is belied by several theoretical issues. Inflation sets forth to alleviate some of them. The theory of inflation was first set forth by Alan Guth [24], and has since been developed by others [25–28]. It is fair to say that inflation is the most popular explanation for solving these problems, but it is not without its detractors. I will first go over the motivating problems, discuss how inflation attempts to solve them, then briefly discuss alternatives to inflation.

1.3.1 Motivating Problems

The Horizon Problem

The fact that the age of the Universe is finite sets a limit on the distance a photon could have traveled. In other words, the remarkably obvious statement is that a photon that travels at speed $c$ can only travel $c \times \Delta t_{\text{univ}}$ where $\Delta t_{\text{univ}}$ is the age of the Universe. To better motivate this, we introduce the notion of conformal time

$$\eta = \int_{t_i}^{t_f} dt \frac{1}{a(t)}.$$  \hspace{1cm} (1.15)

Then we can define the notion of the particle horizon, which is the maximum comoving distance that light can travel between two times

$$\chi_{\text{ph}}(\eta) = \eta - \eta_i = \int_{t_i}^{t_f} dt \frac{1}{a(t)}.$$  \hspace{1cm} (1.16)

Remember that we have chosen units where $c = 1$, so $\chi$ is a distance but the $c$ is dropped. So if the Universe is of age $\eta_{\text{univ}}$, then only objects that started $\chi(\eta_{\text{univ}})$ have had enough time to reach a location. Similarly we can define the event horizon as all future points that we can see. This has a nearly identical definition as the particle horizon

$$\chi_{\text{eh}}(\eta) = \eta_f - \eta = \int_{t}^{t_f} dt \frac{1}{a(t)}.$$  \hspace{1cm} (1.17)
1.3.1 Motivating Problems

The perhaps less obvious fact is that this does not necessarily go to infinity if $\eta_f = \infty$. Certainly if the Universe were static, $\chi_{eh} \to \infty$. However we live in a dynamical universe, and the metric can expand faster than photons can travel in it. If we want to evaluate eq. (1.18) we need to know the form of $a$, which is given by the Friedmann equations and the components of the Universe. To get a sense for its functional form, note that

$$\chi_{eh}(\eta) = \int_{t_i}^{t_f} \frac{dt}{a(t)} = \int_{a_i}^{a_f} \frac{da}{a} = \int_{\log a_i}^{\log a_f} \frac{d\log a}{aH}.$$  \hspace{1cm} (1.18)

So the comoving horizon scales with $1/(aH)$. If we assume the Universe is described by a single fluid with constant equation of state $P = w\rho$, we find that

$$\chi_{ph}(a) = \frac{2H_0^{-1}}{(1 + 3w)} \left[ a_i^{\frac{1}{2}(1+3w)} - a_i^{\frac{1}{2}(1+3w)} \right].$$  \hspace{1cm} (1.19)

Recombination occurred approximately 380,000 years after the Big Bang singularity. The limited distance that light can travel in that time period means the majority of the Universe has never been in causal contact. Despite this, the Universe is incredibly uniform in temperature. Without the possibility of interactions, there is no mechanism for particles to thermalize.

**Flatness Problem**

The Einstein field equations and the Friedmann equations allow any spatial curvature. In principle, we have no issue with any of the possibilities - open, flat, or closed. We can rewrite the first Friedmann equation as

$$\left( \Omega^{-1} - 1 \right) \rho a^2 = \frac{-3kc^2}{8\pi G}.$$  \hspace{1cm} (1.20)

where $\Omega = \rho/\rho_c$. We note that the right-hand side of this equation is constant. The $\rho a$ term decreases as we move forward in time from the Big Bang. The exact nature is dependent on the components of the Universe as described in Section 1.1.1. The best models have $\rho a$ changing by approximately $10^{60}$. This requires $(\Omega - 1)$ to change by the same amount. The Universe we measure today is very close to flat, or $\Omega \approx 1$. This implies that the initial universe must have had a value
of $\Omega$ that is $10^{60}$ times closer to 1. There is no a priori reason to believe the Universe should have arrived at this initial value. This is a fine tuning problem. In other words, to have the Universe that we observe today, the fundamental parameters describing the Universe must have very specific, but very unlikely values.

1.3.2 Solving Problems with Inflation

Inflation is the supposition that the Universe went through an initial phase of exponential expansion in the first instances after the Big Bang. In other words, the scale factor grows as $a \propto e^{\lambda t}$ where $\lambda$ is some constant. We will discuss later how to generate an exponential expansion, but for now let us explore what an exponential expansion does.

First we consider the horizon problem, where points in space that are not causally connected exhibit the same temperature. However, if we posit that the early universe was in fact exponentially smaller, then all patches have time to be in causal contact. Those points are then driven away from each other by inflation, causing them to appear causally disconnected.

Next we consider the flatness problem. From eq. 1.20 we can see that an exponential expansion drives $\rho a$ to an extremely large number. To keep the left hand side constant, $(\Omega - 1)$ must simultaneously be driven to an extremely small number. So a universe with an arbitrary initial density is driven to one that is very close to critical density.

The exact energy scale of inflation is not know, but when we plug in reasonable numbers, we find that inflation must have expanded the Universe by a factor of at least $10^{20}$ to solve both the horizon and flatness problems. Also the theory predicts that the inflationary epoch spanned $10^{-36} - 10^{-34}$ seconds after the Big Bang.

1.3.3 The $\phi^2$ Potential

The most popular explanation for sourcing the exponential expansion required for inflation is a scalar field that permeates all of space time. In a quantum field theory framework, fields permeate all of spacetime. In fact, every fundamental particle is a coupling to fundamental field. Inflation posits a new field. As the Universe rolls down this scalar field, the energy is deposited into spacetime
1.3.3 The $\phi^2$ Potential

itself, expanding it. The exact dynamics of this inflation are dependent on the model of the potential.

The task of inventing new inflationary models has employed theorists for a decade. There is a literal encyclopedia of inflationary models [29]. We will discuss briefly the $\phi^2$ potential, which is already excluded by data but is a useful illustrative example. Consider the potential

$$V = \frac{1}{2} m^2 \phi^2$$

(1.21)

where $\phi$ is scalar field often called an inflaton. For this potential, Einstein’s equations gives

$$\rho = \frac{1}{2} \dot{\phi} + V(\phi)$$

(1.22)

$$P = \frac{1}{2} \dot{\phi} - V(\phi).$$

(1.23)

This allows us to write the first Friedmann equation as

$$H^2 = \frac{8\pi G}{3} \left[ \frac{1}{2} \dot{\phi} + V(\phi) \right].$$

(1.24)

We can also write the continuity equation for the $\phi^2$ potential by first noting that

$$\dot{\rho} = \ddot{\phi} + \frac{\partial V}{\partial \phi} \dot{\phi}$$

(1.25)

which allows us to rewrite eq. 1.8 as

$$\ddot{\phi} + V'(\phi) + 3H \dot{\phi} = 0$$

(1.26)

where we use $'$ to denote a derivative with respect to the inflaton.

We now introduce the idea slow roll inflation [28, 30]. The concept of slow roll is expressed in the relations $\dot{\phi}^2/2 \ll V(\phi)$ and $\ddot{\phi} \ll 3H \dot{\phi}$. The first equation is the statement that the kinetic energy is much less than the potential energy. The second equation guarantees that the slow roll
1.4 Alternatives to Inflation

The equation is satisfied at all times. Now the Friedmann and fluid equation are even simpler

\[ H^2 = \frac{4\pi G}{3} m^2 \phi^2 \]  
\[ m^2 \phi + 3H \dot{\phi} = 0 \]

where we have also plugged in the form of \( V(\phi) \). These two equations can be combined by solving for \( H \) in the first equation and plugging it into the second, giving

\[ \dot{\phi} = \frac{m}{2\sqrt{3\pi}} \]

which integrates to give

\[ \phi(t) = \phi_i - \frac{m}{2\sqrt{3\pi}} t. \]

This is then plugged back into eq. 1.28 and integrated to give

\[ a(t) = a_i \exp \left[ 2m \sqrt{\frac{\pi}{3}} \left( \phi_i t - \frac{m}{4\sqrt{3\pi}} t^2 \right) \right]. \]

So the \( \phi^2 \) potential generates an exponential expansion with respect to \( t \), just as we desired for inflation. This particular field requires inflation to occur at an energy scale that is already excluded by current data.

1.4 Alternatives to Inflation

There is a growing movement that inflation is a problematic theory \[31,32\]. In some sense, inflation is simply to trying to explain away fine tuning problems. The universe could randomly have \( \rho = \rho_c \) and be isothermal on large scales. While this is unlikely, it is not provably false. The general criticism of inflation is that instead of solving these fine tuning parameters, it just moves them to new inflationary parameters. Or in other interpretations of inflation, the theory predicts infinite multiverses, each with unique parameter values. And any theory that predicts all possible outcomes is not prescriptive.
As these concerns about inflation come to the forefront, alternatives to inflation are being more strongly considered. One such theory is the ekpyrotic model \[33\]. This is a cyclic universe that undergoes a series of “crunches” and “bangs”. In this model, the expansion and contractions phases are slow, rather than fast as inflation predicts. These slow contraction phases create homogeneity and flatness.

Another alternative is the String Gas Cosmology \[34, 35\]. In this theory, we essentially forgo the quantum mechanical view of the Universe and assume that the early universe is described by string theory. It can be shown that thermal fluctuations of this string gas can lead to an almost scale-invariant spectrum of curvature fluctuations.

This thesis will focus on inflation. However, this does not preclude other theories from simultaneously being tested. As we mentioned, String Gas Cosmologies predict curvature fluctuations that are in theory measurable. But to limit the scope of this text, we will assume inflation as the model being tested.

1.5 Measurables

As far as an experimentalist is concerned, all the theory in the world is useless unless it is testable. Fortunately inflation makes several predictions which are right on the edge of what is detectable today\[7\]. Specifically, inflation predicts tensor perturbations. This is not surprising, since inflation is a dramatic change in the space-time metric, we expect it to generate oscillations. The recent direct measurement of gravitational waves by the Laser Interferometer Gravitational-Wave Observatory (LIGO) \[36\] demonstrates tensor perturbations propagate through space-time, adding to our confidence that if inflation truly describes the early universe then primordial tensor perturbations must exist.

Any local quadrupole anisotropy generates polarized signal through Thompson scattering. Photons flow from hot to cool regions, so the first order polarization is simply probing velocity in the primordial fluid. The regions with the greatest polarization are regions between hot and cold spots

* Theorists always make predictions right at the bounds of testability...
1.5 Measurables

![Graph showing the BB spectra for various values of r. The combined power of both SPIDER flights is forecasted to be able to put an upper limit on r of .03.](image)

Figure 1.3: The BB spectra for various values of $r$. The combined power of both SPIDER flights is forecasted to be able to put an upper limit on $r$ of .03.

in the CMB. This generates a $TE$ correlation that is out of phase with the temperature anisotropies. We see this effect in the SPIDER data shown in ??.

These scalar modes generate even parity patterns on the sky. These modes are curl free, so they are called $E$-modes. Tensor perturbations generate even and odd parity modes, or $E$-modes and $B$-modes. $B$-modes are divergence free. A measurement of low $\ell$ $B$ modes would be strong evidence for primordial gravitational waves. This motivates us to define the tensor to scalar ratio

$$r = \frac{A_t}{A_s}$$

where $A_t$ is the amplitude of the primordial tensor perturbations and $A_s$ is the amplitude scalar perturbation. Then the goal of any inflationary probe is to measure $r$. The BB spectrum is shown for various values of $r$ in Figure 1.3.

These measurements can be confounded by the epoch called reionization. This is the period in the history of the Universe that the first stars formed. Energy from collapsing hydrogen clouds provide enough energy to reionize hydrogen. This allows Thompson scattering, which introduces a
second source of polarized signal at low $\ell$. We defined the optical depth to reionization

$$\tau = \int d\ln n_e \sigma_T$$  \hspace{1cm} (1.33)

where $n_e$ is the electron density, $\sigma_T$ is the Thompson scattering cross section, and the integral is evaluated over the line of site. $\tau$ should be thought of as a measure of the opacity that suppresses the observed primordial signal while also inducing large scale polarization signal. The $BB$ spectrum for various values of $\tau$ is shown in Figure 1.4 Notice that it peaks at extremely low $\ell$s. SPIDER does not have sensitivity in these multipoles.

SPIDER is designed to either directly measure or put an upper limit on $r$ by flying multiple frequencies and observing a large patch of sky [37]. A large patch of sky is crucial in constraining foreground modes. The lack of atmosphere allows for sensitive detectors with high instantaneous signal to noise. SPIDER-2 will fly an additional 280 GHz band, which will help further constrain foregrounds.
Part II

Hardware
Chapter 2

Cryostat

In Fall of 2013, we packed the payload and shipped it to New Zealand, a stopping point on its journey to Antarctica. Then the government shut down. The Antarctic season was cancelled.

The SPIDER experiment was left in a shipping yard for weeks, waiting to find a way home. The cryostat was housed in a custom built crate, but it was not designed for large storms. But that is exactly what it experienced sitting in the shipping yard.

Eventually, all our cargo made it home. When we opened up the crate with the cryostat in it, we knew we were in trouble. There was water everywhere. The water managed to seep into vacuum seams, developing rust.

Jón and Sasha decided this was a good time to start writing a thesis. So Anne and I cleaned and leak checked the cryostat for months, making it all work again.

2.1 Vacuum Vessel

The vacuum vessel is the outer shell of the cryostat and maintains a low pressure environment for cryogenic operations. The vacuum vessel must be robust enough to hold a 1 atmosphere pressure
2.1.1 Top Dome

differential for ground operations and characterizations. It must allow CMB signal and detector and thermometer signals out. The SPIDER cryostat in shown in cross-section in Figure 2.1.

2.1.1 Top Dome

The top dome houses six windows (see Section 2.1.2) which hold vacuum while allowing light through. It is connected to the vacuum vessel via stainless steel bolts, with a polymeric o-ring to maintain vacuum.

In practice, there were two main problems with the top dome. First is that the top dome hole pattern was not well referenced to the corresponding thread pattern in the vacuum vessel body. This was due to premeasuring the hole pattern, then welding on the flange ring to the dome. The welding process caused slight deformations in the roundness of the ring. This required us to use a set of special tools which pushed or pulled the top dome into round, while installing the bolts.

The second problem is a large weld bead fouling with the recessed window buckets. This made it impossible to install the window buckets. The flange on the window bucket was milled down to provide extra clearance.

Both these problems were addressed in the second flight cryostat. The first problem was solved by first welding the ring to the dome, then milling it down and drilling the holes. This required that the manufacturer have a large mill since the top dome is about two meters in diameter. The second problem was solved trivially by having a smaller weld specification, and ensuring that the manufacturer meet it.

2.1.2 Windows

SPIDER’s windows are designed to maintain the cryostat’s vacuum and maximize the transmission of microwave radiation. The first condition requires a robust material, while the second favors thin components. To balance these, SPIDER uses molded Ultra-High Molecular Weight Polyethylene (UHMWPE) windows to maintain vacuum over 40 cm clear apertures with only 0.3 cm thick

* Material was purchased from McMaster Carr. Product number 8752K411.
2.1.2 Windows

Figure 2.1: The cross-sectional view of the SPIDER cryostat. Image from [38]. From outside to inside are the vacuum vessel (VV), VCS2, VCS1, and main tank (MT). The superfluid tank (SFT) is directly mounted to the MT via low thermal conductivity struts.
2.1.2 Windows

Figure 2.2: A window sliced in cross section after 3 months of use. The cut window measured in thickness .159” on the left, .165” at the center, and .172” on the right.

components. The THz community has a long history of successfully deploying UHMWPE windows which provided a starting point for the Spider development effort. UHMWPE has high impact strength allowing a very thin sheet to minimize absorption loss without sacrificing strength. UHMWPE has a loss tangent, $\tan \delta$, of less than $3 \times 10^{-4}$ at 150 GHz [39], and the mm-wave photon scattering within the material has been constrained to be less than 1%. Additional measurements have been made at 300 GHz[40]. UHMWPE’s relatively low refractive index allows simple anti-reflection coats to be effective over wide observing bands.

Because the windows are thin, they plastically deform when the cryostat is evacuated. This deformation occurs the first time vacuum is pulled on the windows. There is no measurable creep after 48 hours, even after repeated pressure cycles. Figure 2.2 shows a cross-sectional view of a window that successfully held vacuum for three months. It was then cut in half for visual inspection, but no signs of damage were found.

To check the mechanical robustness, a window was attached to a window tester that allowed us to pull vacuum behind it. We pulled vacuum behind the window and then overpressurized the volume behind the window until it bowed out. This cycle was repeated three times to mechanically work the UHMWPE. The deflection was then measured again, with no measurable change. One difficulty of ballooning is the cold flight temperatures. To ensure the window could withstand flight conditions, liquid nitrogen (LN2) was poured onto the window while under vacuum until the window equilibrated with the LN2. This caused no failures. To further test the robustness of the system, a 5 lb hammer was dropped from approximately 2 feet above the window onto the center of the window while the window was still 77K. Again the window was undamaged.

The Spider windows are mounted on the cryostat in a recessed structure, and a cross-sectional
2.1.3 Housekeeping Hermetic

![Diagram of window assembly with clamp and teeth](image)

Figure 2.3: On the left is a cross-section of the window bucket which holds the UHMWPE window. The window is held down using the clamp ring which is screwed into threaded holes at the bottom of the window bucket. On the right is a zoom in on the interface between the clamp and the window. Note the teeth milled into the clamp ring.

UHMWPE has a low coefficient of friction, making it difficult to hold the windows in place with clamping force alone. To provide additional gripping force, concentric teeth are milled into the clamp. These teeth push into the plastic material, shown on the right of Figure 2.3.

### 2.1.3 Housekeeping Hermetic

Each telescope, described in Section 2.4, has a suite of thermometers and heaters. These devices are powered and read out with external electronics. This necessitates a hermetic feedthrough, which passes signal between the inside and outside of the cryostat while simultaneously keeping vacuum. SPIDER used 6 ConFlat flanges with three DD-50 connectors per ConFlat. A drawing provided by the manufacturer is shown in Figure 2.4. Two of the DD-50 connectors are used for powering and reading out thermometers and heaters. The third connector is used to power the half-wave plate motors and encoders.

In addition to the telescope thermometers and heaters, SPIDER also employed thermometers and heaters across the cryostat itself. Some of the cryostat thermometers are plotted in Figure 2.5. These signals were fed through their own ConFlat flanges, which exited the cryostat near the bottom dome. Each ConFlat had 2 DD-25 connectors.
2.2 Cryogenics

The large scale cryogenic system consists of three primary components: the main tank (MT), the super fluid tank (SFT), and the vapor cooled shields (VCS).

2.2.1 Main Tank

The MT is the primary reservoir for liquid helium-4 for the SPIDER cryostat. The MT holds 1284 L of liquid helium, which in practice sets the limit for SPIDER’s operational time. The main tank is held at approximately one atmosphere of pressure using a 13.5 psi pressure regulator at the outlet of the vapor cooled shields. There is additionally a 17.5 psi pressure regulator at the main tank vent line as a safety outlet to prevent overpressurization. As a further safety, there is a burst disk on the main tank vent line. On the ground, the 13.5 psi regulator functions purely as a check valve to ensure that room air does not backflow into the MT. In flight, this regulator prevents the MT from losing too much pressure, which would reduce the flow through the vapor cooled shields. The
Figure 2.5: The temperature of MT (Section 2.2.1), VCS1, and VCS2 (Section 2.2.2). The periodic spikes in the MT temperature are due to $^3$He fridge cycles. There is a diurnal variation in temperature on all stages due to changing optical load. The overall temperature rises through the flight due to the diminishing liquid level. The pings in MT temperature are due to transient power from fridge cycles.
2.2.2 Vapor Cooled Shields

Figure 2.6: The stycast stripe across the front of the MT. This was used to mitigate a known helium leak. Also stycasted are visible weld beads.

MT is made of aluminum 5083 because of its high post-weld strength.

Figure 2.5 shows the MT bottom temperature through flight. The temperature is stable because it is constantly covered with liquid helium. The spikes are due to cycling the $^3$He fridges. When the heat switch is closed, the 40K pump is connected to the MT to cool it rapidly.

Through extensive testing, it became clear that the cryostat had a helium leak from the main tank. The leak was narrowed down to a band that circles around the main tank. A layer of Stycast 2850FT was applied to this band, as shown in Figure 2.6. Stycast is an epoxy that has a coefficient of thermal expansion similar to steel. Further testing pointed to concerns at the weld beads, so all accessible weld beads were also stycasted in an attempt to minimize the leak. These efforts appeared to help the problem, but did not solve it completely.

2.2.2 Vapor Cooled Shields

The gaseous helium leaves the MT at about 10 K. To take advantage of the additional cooling power available, the gas is passed through a set of heat exchanges connected to a pair of radiative shields referred to as vapor cooled shields 1 and 2 (VCS1 and VCS2). VCS1 is the internal shield
2.2.3 Superfluid Tank

directly surrounding the MT. VCS2 is the outer shield lying between VCS1 and the vacuum vessel. The VCSs are made of aluminum 1100, which has high thermal conductivity. In flight, VCS1 and VCS2 equilibrated at approximately 40 K and 160 K respectively.

Both VCSs are wrapped in multi-layer insulation (MLI). The MLI is 16 and 52 layers thick at VCS1 and VCS2 respectively. MLI consists of interleaved layers of reflective Mylar and polyester sheet. The Mylar is made of a thermoplastic polymer with a 35 nm thick aluminum layer. The aluminum reflects incident photons sourced by hotter stages. The polyester sheet prevents layers of Mylar from thermally shorting. The downside to MLI is that it is unwieldy and has a tendency to fray. Also, it limits pumpout speed of the cryostat.

VCS1 and VCS2 sit at around 30K and 115K respectively through flight, as shown in Figure 2.5. The diurnal variation in temperature is due to variable optical loading. The scan strategy scans up and down once a day. The optical loading from reflections off the Earth as well as clouds below the cryostat couple strongest when the cryostat is pointed to low elevations.

2.2.3 Superfluid Tank

The superfluid tank (SFT) is a small tank at the bottom of the cryostat which holds superfluid \(^4\text{He}\). When \(^4\text{He}\) is below its \(\lambda\) point, it becomes a two component fluid. One component is normal fluid while the other component is superfluid. The superfluid component has zero viscosity and zero entropy. The \(\lambda\) point for \(^4\text{He}\) at saturated vapor pressure is 2.172 K.

The SFT was designed to hold 16 L of liquid, but a post flight analysis showed that its volume is much closer to 12.5 L. In practice, the exact volume is not important as the SFT is continuously fed by a capillary system drawing liquid helium from the MT to the SFT. The capillary system is fully described in Jón Guðmundsson’s thesis [42]. The SFT is never more than half full of superfluid helium at equilibrium and sits at approximately 1.5 K during flight. Copper heat straps connect the SFT to the condensation point of the \(^3\text{He}\) fridges.

\[1\text{ The polyester sheet is sometimes called wedding veil.}\]
2.2.4 $^3$He Fridge

![Image of SFT during integration in Princeton](image)

Figure 2.7: A picture of the SFT during integration in Princeton. The 4K contact pads are covered in blue painter’s tape in this photo. The blue tape is not there during operations. The center port is used for cable routing.

2.2.4 $^3$He Fridge

The SPIDER detectors have a transition temperature around 500 mK. This necessitates a cooler that reaches well below that temperature. SPIDER employs 6 closed cycle $^3$He fridges, one in each telescope insert, to provide sub-Kelvin cooling. These fridges have a base temperature around 270 mK. The base temperature of the fridge is much lower than the transition temperature of the detectors to allow margin for optical loading. Because SPIDER is a balloon-borne telescope, the on-sky loading onto the detectors is unknown until flight, when it is too late to reduce optical loading onto the detectors.

Figure 2.8 shows a schematic of the $^3$He fridge used by SPIDER. It is incredibly important to be careful handling the fridges. $^3$He is expensive. Further, there is approximately 10 atm of gas in the fridge at room temperature, making the fridges dangerous. Moreover, the stainless steel tubing (the gray tubes in Figure 2.8) are fragile thin walled stainless steel. They must be thin to reduce parasitic loads from the post to the condensation point and from the condensation point to the
2.2.4 $^3$He Fridge

Figure 2.8: A schematic of the $^3$He fridge. When operating, the heat switch (HSW) is closed, shorting the pump to the 4K cold plate. When the pump is cold, it adsorbs free helium in the fridge. This drops the vapor pressure over the helium bath in the still, dropping its boiling point. This lowers the fridge temperature to 270 mK. Once the still is empty, the heat switch is opened and a heater in the pump is turned on. This causes helium on the pump to outgas. The gas flows to condensation point (CP) which is thermally sunk to the SFT so it is at 2K. Once the gas reaches the condensation point, it condenses and drips into the still. Once the still is full, the heater is turned off and the heat switch is closed, cooling down the fridge again.
2.3 Pumpdown and Cooldown

Starting from a fridge that is entirely at 4K, except the condensation point which is at 2K, the procedure for cycling is as follows:

1. Open the heat switch and turn on the heater. Heat the pump to 30 K. This will outgas all the $^3$He in the pump. The $^3$He convects to the 2K condensation point where it condenses and drops into the still.

2. Once the still is full, turn off the heater. In practice the heater is servoing to 30K because the pump has high heat capacity, so turning off the heater amounts to turning off the servo.

3. Close the heat switch. This rapidly cools the pump. Once the pump is cold, free $^3$He gas adsorbs to the pump, reducing the vapor pressure in the fridge. This drops the temperature of the $^3$He bath in the still. Now the still should be 250 mK.

Once the fridge is cold, it should hold for about 72 hours with no thermal load. The fridges were cycled much more often in flight due to leak in the cryostat which is discussed in Section 3.4.1. We can estimate the load onto the fridge just by looking at its equilibrium temperature. Figure 2.9 shows a load curve of one of the $^3$He fridges called Thelma 8. The load curve shows the temperature as a function of applied electrical power. This does not take into account the radiative load from the environment surrounding it. The data was taken in a small wet dewar which surrounds the fridge with a 2K shield, so the loading should be small.

2.3 Pumpdown and Cooldown

Due to the large volume of the cryostat, it takes a long time to evacuate and cooldown the cryostat. It requires approximately 11 days from the moment the cryostat is close to the first time $^3$He fridges can be cycled. Table 2.1 breaks down the various components of a pumpdown and cooldown.

---

1 All the fridges are called Thelma. The initial SPIDER proposal had two fridges per insert, the first providing the condensation point for the second, reducing the parasitic load on the cooler fridge. The pair of fridges were called Thelma and Louise.
Figure 2.9: A load curve of Thelma 8, one of the $^3$He fridges used by SPIDER. The y-axis is applied electrical power; not included is the radiative power from the cryostat around the fridge. This data was taken in a small wet dewar, so the radiative environment around the fridge was below 2K. The black line is a quadratic fit to the blue points. This wet dewar is described in Appendix F.

<table>
<thead>
<tr>
<th>Item</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nitrogen Purges</td>
<td>3 days</td>
</tr>
<tr>
<td>Turbo Pumpdown</td>
<td>2 days</td>
</tr>
<tr>
<td>Nitrogen Cooldown</td>
<td>4 days</td>
</tr>
<tr>
<td>Helium Cooldown</td>
<td>3 days</td>
</tr>
</tbody>
</table>

Table 2.1: Time required for pumping out and cooling down the cryostat.
2.3.1 Pumpdown

Empirically the quality of the cooldown is a strong function of final vacuum vessel pressure before the first cryogens are filled. In runs where cryogens were filled before a low pressure was achieved, the cryogenic hold times and detector performances were measurably impaired. In the worst instance, the SFT could only hold liquid for approximately 24 hours before all of its helium would boil off. In other runs, the detector performance would degrade due to helium condensation, requiring the fridges to be re-cycled.

We discovered that pumping down the vacuum space to a rough background, backfilling with nitrogen, and repeating several times, would result in both a faster pumpdown and lower final pressure. We theorize that this is due to nitrogen molecules physically dislodging water vapor and other contaminants which have adsorbed to the surfaces in the cryostat, particularly the multilayer insulation. Figure 2.10 shows vacuum vessel pressure as a function of pumpdown number.

In practice the pump-purge cycles stops reducing the final pressure after about five cycles. The speed in which we can pump and purge is set by optical filters installed to reduce incoming load onto cooler stages. These filters are thin and fragile. Because there is a significant flow impedance
through the VCSs, it is easy to set up a large pressure differential across the filters, threatening to break them.

2.3.2 Cooldown

Once the cryostat is at low pressure, the system is precooled with liquid nitrogen. Liquid nitrogen has the advantage of being affordable and having high enthalpy relative to liquid helium. The moment liquid nitrogen is introduced into the MT, the vacuum vessel pressure immediately drops. This is due to adsorption onto the cold surfaces. However the helium background does not change because helium does not condense out at these temperatures.

If telescopes are installed, it takes about four days for the system to cool to 77 K. The telescopes represent a large thermal mass and are designed to be well isolated from the rest of the system. Therefore the telescopes, and specifically the focal planes, are the slowest to cool down.

After the entire cryostat is at 77K, we empty out the liquid nitrogen. To ensure that there is no residual nitrogen in the system, we wait until the coolest thermometer is above 80K. Finally, the MT is filled with LHe. The initial fill effectively flash boils, but the cold vapor still cools the system. There is a large pressure transient associated with the flash boiling of helium, so monitoring the system is vital in this period. Successive fills are less eventful, and the system is at 4 K after 500 L of helium and three days.

2.4 The Telescopes

The entire cryostat is designed with the idea of cooling the telescopes themselves. The telescopes contain most optical elements, cryogenic baffling, and the focal plane itself. The focal planes are further explained in Section 3.3. A cross-sectional view of the SPIDER optical stack is shown in Figure 2.11. The individual telescopes are designed to be modular, allowing installation into any of the six telescope ports.

SPIDER flew 6 telescopes, 3 at 95 GHz and 3 at 150 GHz. The focal planes are all prepended with “X” because all they are the “Rev X” architecture from the Jet Propulsion Laboratory’s
2.4 The Telescopes

<table>
<thead>
<tr>
<th>Focal Plane</th>
<th>Freq [GHz]</th>
<th>Detectors</th>
<th>NET [$\mu$K$\sqrt{s}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>95</td>
<td>403</td>
<td>7.75</td>
</tr>
<tr>
<td>X2</td>
<td>150</td>
<td>232</td>
<td>11.37</td>
</tr>
<tr>
<td>X3</td>
<td>95</td>
<td>369</td>
<td>11.19</td>
</tr>
<tr>
<td>X4</td>
<td>150</td>
<td>211</td>
<td>14.55</td>
</tr>
<tr>
<td>X5</td>
<td>95</td>
<td>412</td>
<td>9.64</td>
</tr>
<tr>
<td>X6</td>
<td>150</td>
<td>232</td>
<td>12.03</td>
</tr>
</tbody>
</table>

Table 2.2: The focal planes and the number of functional detectors used in the SPIDER analysis pipeline. Also shown are the focal plane noise equivalent temperatures, which are further described in Section 5.3.1.2. Odd-numbered focal planes are 95 GHz and even-numbered focal planes are 150 GHz.

The telescope elements are supported by a lightweight carbon-fiber structure, which has minimal thermal contraction. The lenses are high-density polyethylene with expanded Teflon antireflection coatings. The lenses are cooled to minimize in-band loading. A 1.5 K blackened sleeve is placed between the two lenses to further reduce instrument loading. The sleeve is blackened with Eccosorb HR10 after experiments with carbon and steel impregnated Stycast were demonstrated to be insufficient. The measurements are shown in Appendix B. Substantial effort was expended to minimize HR10 dust collecting on the lenses.

The telescope is cooled to 4 K via the cold plate, which is thermally sunk to the MT. The carbon-fiber structure is designed to be relatively non-conductive. Copper straps are connected to the outside of the telescope to thermally couple the top of the telescope to the cold plate.

An auxiliary post presents the 2 K cooling power to the insert space. The post is connected to the 2 K stage via a long copper strap. The focal plane is stood off from the 2 K stage via carbon fiber legs, and is connected to the fridge via its own copper strap.

The optical filtering is done in multiple stages, only some of which are connected in the insert. The goal of filtering is to minimize loading onto the cold stages while allowing signal to pass through. Table 2.3 lists the full stack of SPIDER filters. The window, VCS1, and VCS2 filters reject out of band radiation from overwhelming the cooler stages. The 10 inverse centimeter (icm) filter at the
2.4 The Telescopes

Reflective forebaffle
Wide-angle baffling
UHMWPE window
Metal-mesh filters
(300 K, 150 K, 40 K)
HWP (4 K)
HWP encoder & motor mount
Dual-layer magnetic shield (4 K)
Optics sleeve with baffle rings (1.5 K)
Snout
Eyepiece lens (4 K)
Low-pass filter (1.5 K)
Focal plane (300 mK)
1.5 K standoff
4K cold plate
3He cooler (300 mK)
Detector tile
Magnetic shield (1.5 K)
SQUID modules
Carbon fiber truss
1.5 K standoff

Figure 2.11: A cross-sectional view of the SPIDER optical stack. The “telescope” refers to everything between and inclusive of the 4K filter and the cold eplate. Image from [43].
2.4 The Telescopes

<table>
<thead>
<tr>
<th>Location</th>
<th>Filter Type</th>
<th>Cutoff [GHz]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Window</td>
<td>C15</td>
<td>5200</td>
</tr>
<tr>
<td></td>
<td>C30</td>
<td>2300</td>
</tr>
<tr>
<td>VCS2</td>
<td>C15 (2)</td>
<td>5200</td>
</tr>
<tr>
<td></td>
<td>C30 (2)</td>
<td>2300</td>
</tr>
<tr>
<td>VCS1</td>
<td>C15</td>
<td>5200</td>
</tr>
<tr>
<td></td>
<td>C30 (2)</td>
<td>2300</td>
</tr>
<tr>
<td></td>
<td>12 icm</td>
<td>3600</td>
</tr>
<tr>
<td>4K stage</td>
<td>10 icm</td>
<td>290</td>
</tr>
<tr>
<td></td>
<td>Nylon</td>
<td>540</td>
</tr>
<tr>
<td>2K stage</td>
<td>4/6 icm (95 GHz/150 GHz)</td>
<td>115/180</td>
</tr>
</tbody>
</table>

Table 2.3: The filter stack from warmest to coolest stages. The best estimate for the cryostat temperature during flight is approximately -20°C [44]. The other temperatures can be read off from Figure 2.5. The multi-layer hot press filters are described by their cutoff wavelength in inverse centimeters (icm).

The top of the 4K stack rejects infrared radiation while the nylon filter is designed to catch any blue leaks above 1 THz. The upper band of the SPIDER focal planes is defined by the 4 and 6 icm filters for 95 GHz and 150 GHz respectively.
Chapter 3

Detectors

3.1 Bolometer Theory

Superconducting transition edge sensors (TESs) have been a workhorse for detecting astrophysical photons, including CMB photons, for over a decade [45–48]. TESs are conceptually simple. A superconductor is biased between its superconducting and normal state. When energy is deposited on the superconductor, the superconductor heats and increases its resistance. The change in resistance is read out using superconducting quantum interference devices (SQUIDs). The following section will describe basics of the TES as a bolometer. Almost all of the discussion will rely on Irwin and Hilton’s paper [49].

3.1.1 Bolometer Basics

A bolometer is a device consisting of an absorptive element coupled to a thermal bath. A diagram of a simple bolometer is shown in Figure 3.1. The absorptive element has heat capacity $C$ and is coupled to a thermal bath at temperature $T_{\text{bath}}$ via a thermal link with thermal conductance $G$. If the bolometer is made of a material with a temperature dependent resistance, $R(T)$, then incident power onto the bolometer changes the resistance. This also allows power to be applied to the bolometer via Joule heating. Thermal equilibrium is achieved when the incident radiative power and the electrical power match the cooling power supplied by the bath.
3.1.1 Bolometer Basics

Figure 3.1: A diagram of a simple bolometer. The bolometer is shown in orange and the thermal bath is shown in blue. The bias circuit is on the left. Incident power, \( Q \), is deposited on the bolometer, raising its temperature. This changes the resistance, \( R(T) \), of the TES bolometer, which alters voltage drop across the TES. The change in voltage can be measured by inductively coupling the bias circuit to a readout circuit.

The thermal circuit is described by

\[
C \frac{dT}{dt} = Q + P_J - P_G
\]  

(3.1)

where \( Q \) is the incident thermal power, \( P_J \) is the power from Joule heating, and \( P_G \) is the cooling power provided by the thermal bath. \( P_J \) is simply given by the power version of Ohm’s law. \( P_G \) can be calculated from the definition of \( G \), which is typically defined using a power law with index \( \beta \)

\[
G \equiv \frac{dP_G}{dT} = G_0 \left( \frac{T}{T_0} \right)^\beta
\]

(3.2)

By integrating eq. (3.2), we find

\[
P_G = \frac{G_0 T_0}{\beta + 1} \left[ \left( \frac{T}{T_0} \right)^{\beta + 1} - \left( \frac{T_{\text{bath}}}{T_0} \right)^{\beta + 1} \right].
\]

(3.3)

The exact nature of the physical thermal link determines \( \beta \), but typical values range between 1 and 3.

Now consider using a TES as the bolometer element. This introduces a crucial feature. TES
resistance is a function of both temperature and current. The physical nature of the temperature dependence is obvious. A TES below its transition temperature has 0 resistance and above its transition temperature has constant resistance $R_N$, the normal resistance. Varying between the superconducting and normal state is a continuous function.

The more subtle feature is the current dependence of $R$. It is important to note that this current dependence is separate from the Joule heating. The Ginzburg-Landau theory\[50\] for a two-fluid model description of a superconductor around the transition describes the current dependence of $R$ well. The two fluids are the supercurrent and ohmic current. The supercurrent has Cooper pairs as carriers. The ohmic current has unpaired electrons as carriers. Using this two fluid model, one can derive a first-order correction term in the transition\[51, 52\]. Now the TES resistance can be modeled as

$$R(T, I) = \frac{R_N}{2} \left\{ 1 + \tanh \left[ \alpha \left( \frac{T}{T_C} - 1 + \left( \frac{I}{I_C} \right)^{2/3} \right) \right] \right\}$$

where $T_C$ is the critical temperature, $I_C^0$ is the critical current at $T = 0$, and $\alpha$ is the temperature sensitivity which will be discussed further in this section. This model should be treated as a descriptive tool rather than a scientific description. For this purpose, taking the partial derivatives of the model as a function of $T$ and $I$ gives $\alpha$ and $\beta_I$ respectively. The latter is the current sensitivity of the TES, which is also described in more detail further in this section.

Figure 3.2 is an example of $R(T, I)$ for a TES with $\alpha \approx 100$ and $\beta_I \approx 1$ at the center of the transition. These are relatively good approximations for SPIDER TESs. As temperature and current change, whether through optical coupling or base temperature changes, the TES resistance moves along the surface in Figure 3.2.

Now consider the electrical circuit used to bias the TES on transition, shown on the left of Figure 3.3. The circuit consists of a shunt resistor ($R_s$) in parallel with an inductor ($L$), TES ($R$), and any parasitic load ($R_{\text{par}}$). There are actually two inductors, one to couple the TES circuit to the read out circuit and one to slow down the circuit to reduce aliased noise. The external electronics provide a current bias, $I_b$. This circuit can be restated in terms of a voltage bias, $V_b$,
3.1.1 Bolometer Basics

Figure 3.2: The TES resistance as a function of both $I$ and $T$ assuming the simple model in eq. 3.4. This model uses $\alpha \approx 100$ and $\beta I \approx 1$ at the center of the transition as inputs.

and load resistor, $R_L$, using the Thévenin equivalent circuit shown on the right of Figure 3.3.

Using Thévenin’s theorem, we find

$$V_b = I_b R_s \quad \text{and} \quad R_L = R_s + R_{\text{par}}. \quad (3.5)$$

Using the Thévenin equivalent circuit will simplify calculations.

Applying Ohm’s law to the circuit on the right of Figure 3.3 gives

$$V_b = IR_L + IR + L \frac{dI}{dt}. \quad (3.6)$$

Note that $I$ refers to the current through the TES, while $I_b$ is the current supplied by the external electronics. These are not the same.

Equations [3.1] and [3.6] are a set of coupled equations which describe how the TES bolometer responds to input power. Before solving these equations, it is helpful to expand them in the small

* The Thévenin voltage is given by the open circuit voltage. The Thévenin resistance is given by replacing all voltage sources with shorts and all current sources with opens.
3.1.1 Bolometer Basics

Figure 3.3: The physical bias circuit on the left. A current bias $I_b$ is provided by the external electronics. This can be restated as the Thévenin equivalent circuit shown on the right where $V_b = I_b R_s$ and $R_L = R_s + R_{\text{par}}$.

signal limit around an equilibrium value. The resistance of the TES can be rewritten as

$$R = R_0 + \frac{\partial R}{\partial T} \delta T + \frac{\partial R}{\partial I} \delta I = R_0 + \alpha R_0 \frac{\delta T}{T_0} + \beta I R_0 \frac{\delta I}{I_0}$$

(3.7)

where

$$\alpha \equiv \frac{\partial \log R}{\partial \log T} = \frac{T R}{R_0} \frac{\partial R}{\partial T} \quad \text{and} \quad \beta I \equiv \frac{\partial \log R}{\partial \log I} = \frac{I \frac{\partial R}{\partial I}}{R_0}.$$

(3.8)

The logarithmic sensitivities $\alpha$ and $\beta_I$ describe how the TES resistance changes with temperature and current respectively. These are intrinsic properties of the TES. Note that $\beta_I$ and $\beta$ are different variables.

Similarly, for small inputs of thermal power and electrical power, $\delta Q$ and $\delta V_b$ respectively, $P_G$ and $P_J$ can be expanded

$$P_G = P_{G0} + G_0 \delta T$$

(3.9)

and

$$P_J = P_{J0} + 2I_0 R_0 \delta I + I_0^2 \delta R.$$
Now eq. 3.1 can be rewritten as

$$C \frac{d}{dt}(\delta T) = \delta Q + 2I_0R_0\delta I + I_0^2\delta R - G_0\delta T.$$  \hspace{1cm} (3.11)

This is put in terms of $T$ and $I$ using eq. 3.8.

$$\delta Q = C \frac{d}{dt}(\delta T) - 2I_0R_0\delta I - I_0^2 \left( \frac{R_0}{T_0} \delta T + \frac{R_0}{I_0} \delta I \right) + G_0\delta T.$$  \hspace{1cm} (3.12)

Similarly, eq. 3.6 becomes

$$I_0\delta V_b = I_0R_L\delta I + I_0L \frac{d}{dt}(\delta I) + I_0R_0\delta I + I_0^2 \left( \frac{R_0}{T_0} \delta T + \frac{R_0}{I_0} \delta I \right).$$  \hspace{1cm} (3.13)

Everything up until this point has been written in the time domain. However, we can decompose these equations into individual Fourier modes. This gives

$$\delta Q = i\omega C\delta T - 2I_0R_0\delta I - I_0^2 \left( \frac{R_0}{T_0} \delta T + \frac{R_0}{I_0} \delta I \right) + G_0\delta T
= \left( G_0 - I_0^2 \frac{R_0}{T_0} + i\omega C \right) \delta T - I_0R_0(2 + \beta_I)\delta I$$  \hspace{1cm} (3.14)

and

$$I_0\delta V_b = I_0R_L\delta I + i\omega I_0L\delta I + I_0R_0\delta I + I_0^2 \left( \frac{R_0}{T_0} \delta T + \frac{R_0}{I_0} \delta I \right)
= \alpha I_0^2 \frac{R_0}{T_0} \delta T + \left[ I_0R_L + i\omega I_0L + I_0R_0(1 + \beta_I) \right] \delta I.$$  \hspace{1cm} (3.15)

where all the terms are a function of frequency, $\omega$, rather than time. These two equations can be

† This trivially sends $d/dt(X(t)) \rightarrow i\omega X(\omega)$. 

42
combined into a matrix form

\[
\begin{bmatrix}
\delta Q \\
I_0 \delta V_b
\end{bmatrix} =
\begin{bmatrix}
G_0(1 - \mathcal{L} + i\omega\tau_0) & -I_0R_0(2 + \beta I) \\
G_0\mathcal{L} & I_0(R_L + R_0(1 + \beta I))(1 + i\omega\tau_{el})
\end{bmatrix}
\begin{bmatrix}
\delta T \\
\delta I
\end{bmatrix}
\]

(3.16)

\[
\begin{bmatrix}
G_0\tau_0(1/\tau_I + i\omega) & -I_0R_0(2 + \beta I) \\
G_0\mathcal{L} & I_0L(1/\tau_{el} + i\omega)
\end{bmatrix}
\begin{bmatrix}
\delta T \\
\delta I
\end{bmatrix}
\]

(3.17)

where the thermal time constant

\[
\tau_0 = \frac{C}{G_0},
\]

(3.18)

the biased thermal time constant

\[
\tau_I = \frac{\tau_0}{1 - \mathcal{L}},
\]

(3.19)

the electrical time constant

\[
\tau_{el} = \frac{L}{R_0(1 + \beta I) + R_L},
\]

(3.20)

and the low frequency loop gain

\[
\mathcal{L} = \frac{\alpha P J_0}{G_0 T_0}.
\]

(3.21)

It is worth pausing here and considering what all these newly defined terms actually mean. To aid in the discussion, write eq. (3.16) as

\[
\bar{Q} = M\bar{\Delta}
\]

(3.22)
and consider the matrix

\[
S \equiv M^{-1} = \left\{ \begin{array}{c}
\frac{1}{G_0} \left( \frac{1}{\tau_I} + i\omega \right) \left( \frac{1}{\tau_{el}} + i\omega \right) + \mathcal{L} R_0 (2 + \beta_I) \\
-\mathcal{L}/I_0 \\
\end{array} \right\}^{-1} \\
\times \left[ \begin{array}{cc}
I_0 G_0 \tau_0 L & I_0 R_0 (2 + \beta_I) \\
-\mathcal{L} & G_0 \tau_0 (1/\tau_I + i\omega) \\
\end{array} \right]
\]  
(3.23)

This new matrix is called the responsivity matrix. Equations 3.23 and 3.24 are equivalent matrices expressed differently and will be useful for exploring how the detector reacts to various inputs. With the responsivity matrix in hand, we trivially invert eq. 3.22 so that the state variables \((\delta I, \delta T)\) are a function of the stimuli \((\delta Q, \delta V_b)\)

\[
\delta = S \delta Q.
\]  
(3.25)

This matrix equation gives the full response of a bolometer as a function of input. So while the equation is complicated, it cannot be overstated how useful it is.

Now let us briefly pause to consider the significance of all the parameters. First note that \(\mathcal{L}\) is the ratio of the coefficients to the \(\delta T\) terms in eqs. 3.9 and 3.10, so \(\mathcal{L}\) must describe something about how input energy changes the Joule heating versus the bath cooling. In fact, just by inspection we can see \(\delta P_J/\delta P_G \propto \mathcal{L}\), so if \(\mathcal{L}\) is large then input energy changes the Joule heating more than the thermal cooling. To state this more formally, it is apparent from eq. 3.23 that \(\delta T\) goes to zero as \(\mathcal{L} \to \infty\). The current term as \(\mathcal{L} \to \infty\) is

\[
\delta I = \frac{-1}{R_0 - R_L} \left( \frac{\delta Q}{I_0} + \delta V_b \right).
\]  
(3.26)
3.1.1 Bolometer Basics

The frequency term in the denominator is dropped because SPIDER’s signal band is approximately DC relative to $\tau_{el}$. In the limit of high loop gain, when input power is applied to the TES the temperature of the TES does not change, but the current does. A subtle note is that while the current changes, the Joule power does not. Note from eq. (3.26) that $\delta I/\delta V_b = -1/(R_0 - R_L)$ if $\delta Q$ is zero. Using this fact, we write

$$
\delta P_J = \delta \left[ I V_b \left( \frac{R}{R + R_L} \right) \right]
= V_b \left( \frac{R_0}{R_0 + R_L} \right) \delta I + I_0 \left( \frac{R_0}{R_0 + R_L} \right) \delta V_b + I_0 V_b \left[ \frac{1}{R_0 + R_L} - \frac{R_0}{(R_0 + R_L)^2} \right] \delta R
= -I_0 \beta I \frac{R_0 + R_L}{R_0 - R_L} \left[ \left( \frac{R_0}{R_0 + R_L} \right) - \left( \frac{R_0}{R_0 + R_L} \right)^2 \right] \delta V_b.
$$

(3.27)

So if $R_L \ll R_0$, then $\delta P_J \rightarrow 0$.\footnote{This is a good assumption for SPIDER. Typical $R$ and $R_S$ are 30 and 3 m$\Omega$ respectively.}

Now consider $\tau_0$. From its definition it is obvious that $\tau_0$ describes something about the thermal system. If a bolometer has no electrical bias, in other words $I_0 = P_{f0} = 0$, then the only nonzero term in eq. (3.16) is

$$
\delta Q = G_0 (1 + i\omega\tau_0) \delta T.
$$

(3.28)

This makes it obvious that $\tau_0$ is the natural time constant of a damped harmonic oscillator of a bolometer with no electrical bias. In other words, $\tau_0$ describes the timescale in which the bolometer thermalizes to changes in input power.

Similarly, $\tau_{el}$ describes the timescale in which the electrical bias circuit adapts to changes in input power. If we consider the high loop gain limit so $\delta T = 0$

$$
\delta V_b = [R_L + R_0(1 + \beta_I)] (1 + i\omega\tau_{el}) \delta I
$$

(3.29)

then we see that $\tau_{el}$ is analogous to $\tau_0$ in the electrical equation.
3.1.2 Gain Monitoring with Bias Steps

Random gain drifts can couple into the map as spurious signal, so monitoring variations in gain is important. In the proposal phase of SPIDER, it was estimated that SPIDER required a 0.5% determination of gain [37]. In post-flight analysis, it appears that the requirement is weaker than the proposal value. Nonetheless, having an accurate gain measurement is valuable in constraining systematics, even if just to provide confidence in our final result.

Past experiments have implemented a calibration lamp which injects a known optical signal into the detector path [53]. However, this was deemed impractical for SPIDER. As an alternative, Jeff Filippini proposed using known steps in electrical bias current as a proxy for changes in optical responsivity [54]. The fundamental principle underlying this process is that the electrical gain is related to the optical gain. This entire subsection will derive and explore how to relate the electrical response, $S_{IB} \equiv \partial I / \partial I_b$, to the power response $S_{IQ} \equiv \partial I / \partial Q$. With that relation in hand, we will be able to monitor input optical power response by injecting a known electrical power input.

To derive $S_{IB}$, consider eq. 3.24 with $dQ = 0$.\[ \frac{\delta I}{\tau_0 (1/\tau_I + i \omega) + \mathcal{L} R_0 (2 + \beta_I) \delta V_b} \frac{\tau_0}{\tau_0 L (1/\tau_I + i \omega) (1/\tau_{el} + i \omega)} = \frac{L R_0 (2 + \beta_I) \tau_0}{1 + \beta_I + L (2 + \beta_I) \tau_0} \]

which can be rearranged to

\[
\frac{\partial V_b}{\partial I} = \left[ R_L + i \omega L + R_0 (1 + \beta_I) \right. + \left. \frac{R_0 \mathcal{L} (2 + \beta_I)}{1 - \mathcal{L} + i \omega \tau_0} \right] \]

which is just the complex impedance of the Thévenin circuit. So rewrite eq. 3.31 as

\[
\frac{\partial V_b}{\partial I} = Z = R_L + i \omega L + Z_{TES} \]

where

\[
Z_{TES} = R_0 \left[ 1 + \beta_I + \frac{\mathcal{L} (2 + \beta_I)}{1 - \mathcal{L} + i \omega \tau_0} \right].
\]

As a brief aside, note that this agrees with our previous observation that as $\mathcal{L} \to \infty$, $Z_{TES} = -R_0$. 

46
3.1.2 Gain Monitoring with Bias Steps

Continuing on, we already established that \( V_b = I_b R_S \), so we get

\[
S_{IB} = \frac{\partial I}{\partial I_b} = \frac{R_S}{R_L + i\omega L + Z_{TES}}
= \frac{R_S}{Z}.
\]  

(3.34)

In a similar manner, we can solve for \( S_{IQ} \equiv \partial I/\partial Q \) by consider eq. 3.24 with \( \delta V_b = 0 \). This gives

\[
\delta I = -\frac{\mathcal{L}}{I_0(1/\tau + i\omega)(1/\tau_0 + i\omega) + \mathcal{L}R_0(2 + \beta_I)} \delta Q.
\]  

(3.35)

This is rearranged and simplified to give

\[
S_{IQ} = \frac{\partial I}{\partial Q} = -\frac{\mathcal{L}}{I_0(1 - \mathcal{L} + i\omega\tau_0)(R_L + i\omega L + Z_{TES})}
= -\frac{\mathcal{L}}{I_0(1 - \mathcal{L} + i\omega\tau_0)} Z.
\]  

(3.36)

Now we want to find a way to relate \( S_{IQ} \) and \( S_{IB} \). In other words, we want to know how small changes in optical input power relate to small change in input electrical power. Both \( S_{IQ} \) and \( S_{IB} \) change with \( R_0 \), but that is a quantity that is empirically difficult to know while taking science data. \( R_0 \) drifts slowly with variable loading and helium-3 fridge base temperature. Therefore, we will construct a relationship that cancels \( R_0 \). eq. 3.36 has a hidden factor of \( R_0 \) in the \( I_0 \) and \( Z \) term. From the bias circuit, we know

\[
I_0 = I_b \frac{R_S}{R_S + R_0}
\]  

(3.37)

so rewrite eqs. 3.34 and 3.36 as

\[
S_{IB} = \frac{R_S}{R_S + i\omega L + R_0\lambda}
\]  

(3.38)

\[
S_{IQ} = \frac{-\mathcal{L}(R_0 + R_S)}{I_0 R_S(1 - \mathcal{L} + i\omega\tau_0)(R_S + i\omega L + R_0\lambda)}
\]  

(3.39)
where we have assumed that $R_S = R_L$ ($R_{par} = 0$) and defined the convenience variable

$$\lambda = 1 + \beta_I + \frac{\mathcal{L}(2 + \beta_I)}{1 - \mathcal{L} + i \omega \tau_0}.$$  \hfill (3.40)

Rewriting eq. 3.38 in terms of $R_0$ and plugging it into eq. 3.39, we find

$$S_{IQ} = \left\{ S_{IB} [R_S (1 - \lambda) + i \omega L] - R_S \right\} \frac{\mathcal{L}}{\lambda R_S R_S^2 (1 - \mathcal{L})}. \hfill (3.41)$$

This is the exact relation between $S_{IQ}$ and $S_{IB}$, and in principle we are done here. If we know all the parameters in this equation well, we can reconstruct any $S_{IQ}$ from a known $S_{IB}$. In other words, we can get our relative gain on optical signals, $\partial I / \partial P$ by injecting a known electrical signal.

In practice, we can and do make some simplifying assumptions. First, given the scan strategy of SPIDER, the input signal of interest is approximately DC. In other words, we can evaluate eq. 3.41 at $\omega = 0$. This simplifies the expression to give

$$S_{IQ} = \left\{ S_{IB} R_S [1 - \lambda (\omega = 0)] - R_S \right\} \frac{\mathcal{L}}{\lambda (\omega = 0) I_b R_S^2 (1 - \mathcal{L})}. \hfill (3.42)$$

The detectors are built to have steep $\alpha$, making $\mathcal{L} \gg 1$. This allows yet further simplification,

$$S_{IQ} = \frac{2 S_{IB} - 1}{I_b R_S}. \hfill (3.43)$$

This is the equation SPIDER uses to measure the time variable gain of the detectors through flight. Using the flight electronics, a 2 Hz square wave of known amplitude is injected onto the TES with amplitude $\delta I_b$. The change in TES current, $\delta I$, is simply the change in the DC level of the detector. This allows a direct measure of $S_{IB}$. Figure 3.4 shows an example of a bias step.

### 3.2 Multichannel Electronics

Up until this point in the discussion, we have not specified how to read out the detector current. However, we had the foresight to draw an inductor in our bias circuit in Figure 3.3. We can
Figure 3.4: An example bias step from flight data. The y-axis is the detector amplitude. The ringing is due to the finite response time of the detectors. To resolve the amplitude of the step, the bias step algorithm skips the first 3 samples after $I_b$ changes.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T_c$</td>
<td>500 mK</td>
</tr>
<tr>
<td>$R_n$</td>
<td>30 mΩ</td>
</tr>
<tr>
<td>$G_0$</td>
<td>19 pW/K</td>
</tr>
<tr>
<td>$P_{\text{sat}}$</td>
<td>2.5 pW</td>
</tr>
<tr>
<td>$dP/dT_{RJ}$</td>
<td>.1/.15 pW/K\text{RJ}</td>
</tr>
<tr>
<td>$P_{\text{int}}$</td>
<td>.25/.35 pW</td>
</tr>
</tbody>
</table>

Table 3.1: Typical parameters for the SPIDER bolometers. For parameters with frequency dependence, the 95 GHz value is on the left and 150 GHz value is on the right.
3.2 Multichannel Electronics

![Bias Circuit Diagram]

Figure 3.5: The bias circuit of the TES coupled to a SQUID (specifically the SQ1). Note that the symbol for a SQUID is the circle with the Josephson junctions represented by the Xs. A feedback current is applied to the SQUID to keep it on its lock point on the $V\Phi$ curve.

couple the inductor to a superconducting quantum interference device (SQUID), which is a sensitive magnetometer. We will call this SQUID $SQ1$ for reasons that will be obvious soon. The circuit is shown in Figure 3.5.

A SQUID is a pair of Josephson junctions in parallel. If a magnetic field $\Phi$ is applied to the superconducting loop, a screening current begins to circulate, canceling the input applied magnetic field. As soon as the current surpasses the critical current, the voltage across the SQUID changes. The voltage response is periodic with period $\Phi_0 = h/2e$. An externally controlled source applies a feedback flux to the SQUID to keep it on its lock point on the $V\Phi$ curve. So this feedback current is directly proportional to the magnetic flux through the SQUID, which itself is proportional to the current through the TES.

In principle, we can just read out the feedback current applied to keep SQ1 at its lock point and be done. However this would necessitate $\mathcal{O}(N_{det})$ cryogenic wires per FPU. Specifically, you would need to be able to bias each TES, bias each SQ1, and supply feedback current. This proves to be intractable.
3.3 Focal Planes

The multichannel electronics (MCEs) [55–57] set out to solve this problem. The MCEs use time-domain multiplexing § to readout multiple detectors using fewer wires. They achieve this goal by connecting a group of SQ1s to a second SQUID, SQ2. The external electronics power on each SQ1 in series, so the SQ2 receives power from only one SQ1 at a time. The signal from the SQ2 is then sent to an amplifier, called the SQUID series array (SSA), and then is read out. The SSA is critical because it provides impedance matching for the cold and warm electronics.

3.3 Focal Planes

The SPIDER 1 focal planes (FPUs) share heritage with BICEP/Keck Array [48, 58]. A picture of a partially assembled FPU is shown in Figure 3.7. Each focal plane consists of four tiles, each of which are made from 100 mm silicon wafers. Each tile has a 6×6 and 8×8 grid of orthogonally polarized detector pairs at 95 GHz and 150 GHz respectively.

Each pixel consists of two orthogonally polarized detectors. Incoming radiation is absorbed by polarized planar antenna arrays. Each element of the antenna array is a slot sub-radiator patterned in a superconducting niobium ground plane. All of these slots are coherently summed with no phase lag through a microstrip summing network ¶. The antenna signal is then passed through an on-chip

---

§ The primary alternative to time-domain multiplexing is frequency domain multiplexing. Detectors are biased with a known bias frequency and then demodulated in the readout system. Also currently in development is a microwave multiplexer.

¶ In principle one could change the beam or pointing by changing the phase lag of the individual antennas.
3.4 Time Domain Noise

Figure 3.7: The 150 GHz SPIDER focal plane. Note that this photo was taken during assembly, so the upper right tile is missing. The individual antenna arrays are also visible in each tile. Photo courtesy of Marc Runyan and Tony Turner.

band-defining filter and then deposited onto a TES.

3.4 Time Domain Noise

An ideal detector is perfectly described by the equations in Section 3.1, but we inhabit a world that is far from ideal. Various pathologies of noise exist in real detector data of every experiment, and SPIDER is no exception. In this section, I describe the five biggest sources of noise in terms of effect on SPIDER’s total sensitivity — helium film noise (Section 3.4.1), reaction wheel noise (Figure 3.9), radio frequency pickup (Section 3.4.3), flux jumps and cross jumps (Section 3.4.4), and scan synchronous pickups (Section 3.4.5). The helium film noise and radio frequency pickup limit the total duty cycle of the detectors. The remaining three primarily add to the timestream noise. We will defer the discussion of mitigating the noise to Chapter 5.
3.4.1 Helium Film Noise

The helium film noise is pathologically the least well understood of all the sources of noise. The belief is that small amounts of $^4$He in the vacuum space condense onto the detectors. This thermally shorts the detectors to the focal plane body, dramatically increasing the $G$ and correspondingly increasingly the detector noise. The only repeatable way to solve this problem was to cycle the fridges, as described in Section 4.1.4. This would boil off the condensed helium-4.

The helium film would start on the outer edge of the focal plane and creep in to the center of the focal plane over about an hour. The effect first presents in the time domain as a set of spikes with a quick decay, shown in Figure 3.8. The features look similar to the dorsal fin of a shark, so we named them “sharks”. These are identified both algorithmically and by hand and flagged.

3.4.2 Reaction Wheel Noise

A signal correlated with the reaction wheel position appeared on a small fraction of detectors. The reaction wheel is discussed briefly in Section 4.1.3 and described in full detail in Jamil Shariff’s thesis [59]. It is important to note that this reaction wheel locked signal is not believed to be sourced by the reaction wheel or its motor; rather the metal in the reaction wheel modulates the existing RF environment. In ground testing, inserting metal in the space between the reaction wheel and
3.4.3 RF Transmitter Pickup

Figure 3.9: The reaction wheel noise signal as a function of reaction wheel angle. This plot is generated by binning a single 10 minute detector time stream as a function of reaction wheel angle. The six-fold symmetry is due to the six spokes of the reaction wheel. The amplitude is in Analog to Digital Units [ADUs], and 1 ADU is approximately 1 mK.

In flight, the reaction wheel is the only fast-moving metallic element. The reaction wheel is 6 spoked, with brass bricks at the end of each spoke to increase its moment of inertia. This creates the six-fold symmetry in the reaction wheel noise. Figure 3.9 shows a binned template of the reaction wheel noise signal as a function of reaction wheel angle. Figure 3.10 shows the reaction wheel amplitude on X3 in physical coordinates. The response is dramatic in some detectors, and nonexistent in others. There is some evidence that the pickup comes from the MCE crates; however, this has not been confirmed. Empirically, re-seating the cooling fins would change the reaction wheel noise in a nondeterministic way.

3.4.3 RF Transmitter Pickup

SPIDER communicates with the ground via several radio frequency antennas. One dominant source of problems for SPIDER was the Iridium antenna, which would pulse for approximately 2 seconds every minute. An example of the effect on a detector is shown in Figure 3.11. The signal to noise on the Iridium signal is large, making it easy to flag. However this led to complete data loss in this
3.4.3 RF Transmitter Pickup

Figure 3.10: The amplitude of the reaction wheel response on X3 in physical coordinates. Note that color scale is $\log_{10}$. Each pixel has two orthogonal polarizations, denoted A and B. The values for A and B are in the lower left and upper right respectively.
3.4.4 Flux Jumps and Cross Jumps

Among the more pernicious time-domain problems in SPIDER are flux jumps and their associated cross jumps. The readout architecture is discussed in detail in Section 3.2. Because the SQUID is periodic, large signals can drive the response onto the next period of the $V\Phi$ curve. The feedback circuit will drive the SQUID response to exactly one $\Phi_0$ away. This in itself is not a large problem since the amplitude of a flux quanta is fixed per detector, so it is easy to find discontinuities of a given amplitude and stitch them together. In practice, the vast majority of flux jumps happen when the cryostat is pointed high in boresite elevation and is entirely coincident with Iridium pings. The high energy deposition from the Iridium antenna drives the SQUID to the next period of the $V\Phi$ curve.

The bigger problem is that the flux jumps cause a correlated DC step in other detectors on the same multiplexer column. These channels share a feedback current supply loop. An example is shown in Figure 3.12. In this example, the detector generating the timeline shown is not flux jumping.

In principle, reconstructing DC level discontinuities is not difficult, particularly with the knowledge of when it happens. The problem is made more complicated by Iridium and scan synchronous period.

Figure 3.11: An example of an Iridium ping on a detector in flight. Each Iridium pulse is approximately 2 seconds. All of this data is thrown out in the analysis. Positive power deposited on the bolometer results in a negative response in units of ADU.
3.4.5 Scan Synchronous Noise

Figure 3.12: An example of cross jumps. Note that this detector is not flux jumping in this time period. Other detectors in column are flux jumping which is leading to a small DC step in this detector. This is a particularly bad channel shown for illustrative purposes and should not be thought of as typical.

noise. Almost all flux jumps are caused by Iridium pulses, which cause approximately 2 seconds of data loss. So all cross jumps also occur during a time period with no detector data. Again, this is not a problem if the signal is dominated by the dipole and white noise. We would simply subtract the known dipole and then average both sides of the discontinuity to reconstruct a DC level. Unfortunately the scan synchronous noise, which is approximately the amplitude of the dipole, injects an additional offset that is difficult to measure to the required precision.

3.4.5 Scan Synchronous Noise

In terms of effect on SPIDER’s total instrumental sensitivity on foreground parameters and \( r \), scan synchronous pickup is currently the worst source of noise. Scan synchronous noise is defined as any non-sky signal that is at the scan frequency of the payload or any of its harmonics. The problem with scan synchronous noise is that it is by definition phase locked with sky signal, meaning filtering out scan synchronous noise modes will simultaneously attenuate signal. Currently SPIDER uses an aggressive poly-5 filter, which is discussed in detail in Section 7.2. There are alternative mapmakers that potentially filter signal less aggressively or recover science modes without amplifying noise.
3.4.5 Scan Synchronous Noise

Figure 3.13: An example of the scan synchronous signal in the time domain on X3. These timelines are generated by binning detector timestreams into boresite coordinate maps, then reobserving those maps with the boresite pointing. In the upper panel, the data is generated by splitting all detectors in the focal plane into top and bot, where top/bot refers to detector pointing in telescope coordinates. In other words, top detectors point further away from the horizon. The lower panel is the lowpassed difference between A and B polarized detectors in each of the top and bot splits.

These mapmakers are currently being explored, but have none have converged to a functional state yet. The scan synchronous noise can be generated by various sources, including magnetic pickup, sloshing of liquid cryogens, or sidelobes. Given the current analysis, the prevailing theory is that scan synchronous noise is sourced from sidelobes below the payload, either clouds or the ground.

Scan synchronous templates were generated by binning detector timestreams into boresite binned maps after first removing a line on 10 minute chunks. Note that a single scan is approximately 1 minute, so subtracting a line on 10 minute chunks only eliminates long timescale drifts. These maps were then reobserved with the boresite pointing. This effectively averages the detector data between elevation steps, which occur approximately every 3 minutes depending on
3.4.5 Scan Synchronous Noise

Figure 3.14: The amplitude of the scan synchronous signal on X3 as a function of time and boresite elevation on the left and right respectively. The plot is color coded by local sidereal day, while the x-axis is the standard 24 hour day. The scan strategy was changed for the last 4 days to a narrow scan. This by construction reduces the amplitude of the scan synchronous component. The scan synchronous pickup has a strong boresite elevation dependence; high elevation observations have almost no scan synchronous pickup.

The scan synchronous noise has a strong boresite dependence. At high elevations, the signal is effectively nonexistent. At an elevation of approximately 35 degrees, the scan synchronous noise turns on and increases as the elevation drops, as shown in Figure 3.14. An elevation of 35 degrees is broadly consistent with the time at which the top of the windows have a direct line of sight to the horizon. This argues that the scan synchronous pickup is due to illumination of the window from sources at or below the horizon, possibly sourced by either clouds below the payload or Earth
albedo. The rays hit the window and Lambertian scatter, illuminating the detectors.

3.4.6 Cosmic Rays

A search for cosmic rays was done by template fitting to unflagged time streams. SPIDER detected approximately 1 cosmic ray per 3 minutes during flight across all detectors and focal planes. Cosmic rays can deposit energy onto the wafer as a whole or onto the TES island itself. Wafer hits should create correlated signals across many detectors, while TES hits should create single detector responses. It is believed that the majority of cosmic ray detections were due to cosmic rays hitting the TES island.
The British Antarctic Survey flies most of their twin otters with a single pilot. This leaves the co-pilot seat open for passengers. I was lucky enough to sit co-pilot for the flight from Fossil Bluff to SkyBlu. The BAS pilots are the model of professionalism, and my pilot Mark is no exception.

Once we reached cruising altitude, Mark let me take control of the plane. I banked us left and right. I took us from 9,000 ft to 11,000 feet and back down again. It was amazing. At your feet in the cockpit, there are two foot pedals. I asked Mark what they did. He replied, “Push down the left one.” And I did.

“See how we are turning left?”

“Yes. But why don’t we just bank left?”

The mischievous look that came across his face belied his cool demeanor. This was exactly the question he had hoped I would ask. “Keep pushing the pedal, hard.” He reached above our heads and pulled a lever. And to my terror, the left engine had shut off. “Notice how we’re going straight, even with one engine.” I nodded in nervous agreement, hoping he would turn the engine back on. But we flew that way for minutes, Mark clearly delighted by the frightened look on my face.
4.1 2015 Flight

SPIDER launched from the Long Duration Balloon Facility at McMurdo Station on January 1, 2015. The experiment took scientific data for 14 days, and floated for 3 more days with no cryogens. The payload was dropped in western Antarctica on January 17, 2015.

4.1.1 Launch

Launching a payload is perhaps the most frightening moment of ballooning, and certainly the most awe inspiring. Setting SPIDER into flight is a moment I will never forget.

The day of launch, the instrument is suspended via a hook directly above the pivot to the launch vehicle. The launch vehicle is a large truck that carries the payload from the high bay, where it is integrated, onto the launch pad. The payload on the launch vehicle is shown in Figure 4.2.

The SPIDER payload was launched on New Year’s Day 2015 from McMurdo Station, Antarctica. NASA ballooning provides various platforms of balloons to fly payloads, each with a specified weight and altitude range. SPIDER flew on a 34.43H MCF, where MCF is short for “million cubic feet”. This balloon is colloquially known as a “34-Heavy”. It is estimated that SPIDER saw a 3G shock at launch, which is the maximum specified shock a payload should see. There was no known damage due to the shock.

4.1.2 Flight Path

The circumpolar winds set up above Antarctica in late October to early November, and typically begin to break down in late January. These winds tend to keep balloons above continent which is vital because water recoveries are impossible. The circumpolar winds are also helpful because balloons tend to make laps around the perimeter of the continent, meaning the payload will hopefully return near McMurdo Station. This makes recovery easy.

Unfortunately the circumpolar winds started to break down in mid-January, and the SPIDER payload started drifting towards the ocean on the opposite side of Antarctica. This necessitated an immediate termination. The payload was dropped on January 17, 2015, about 5,000 miles from
4.1.3 Scan Strategy

McMurdo base. Fortunately there was a British Antarctic Survey (BAS) base about 1,500 miles from the drop point, and BAS was able to aid in the recovery effort, which is discussed more in Section 4.2.

The full flight path is shown in Figure 4.1. The payload floated at an altitude of 36 km with diurnal variations due to variable sun loading of about 1 km, as shown in Figure 4.3

4.1.3 Scan Strategy

I will only briefly cover the scan strategy here. It is described fully in Jamil Shariff’s thesis [59]. The scan strategy must balance sky coverage, half-wave plate coverage, total flight time available, avoiding the sun on one extreme of right ascension, and avoiding the galaxy on the other extreme of right ascension.

The scan strategy consisted of slewing up and down in elevation once every local sidereal day. The telescope scans in azimuth sinusoidally with a period of about 1 minute. A sinusoidal scan balances the torque limitations of the pivot/reaction wheel and the desire for a smooth scan strategy.

The SPIDER observation region is shown in Figure 4.4 with other experiments’ regions shown for comparison. SPIDER has a relatively large patch of sky compared to its ground based analogue, the BICEP/Keck Array set of experiments. SPIDER does not use the full observation region for its science result. The specific subregion of our observed sky used for science results is currently being explored.

4.1.4 Fridge Cycling Strategy

The $^3$He fridge cycles are designed to hold for 72 hours under normal flight loading conditions. Unfortunately the cryostat had a leak in the main tank which caused helium film noise as described in Section 3.4.1. The hypothesis is that $^4$He would condense on the focal plane, eventually thermally shorting the TES to the focal plane body, dramatically increasing $G$ and correspondingly increase the noise. The only way to repeatably solve this problem was to cycle the fridge, which would bring the focal plane up to 2 K.
Table 4.1: The latitude and longitude of various locations in Antarctica, including the launch location in McMurdo Station and the landing site of Spider. These are expressed in decimal degrees.

<table>
<thead>
<tr>
<th>Location</th>
<th>Lat</th>
<th>Lon</th>
</tr>
</thead>
<tbody>
<tr>
<td>McMurdo Station</td>
<td>-77.850</td>
<td>166.666</td>
</tr>
<tr>
<td>Spider Landing Site</td>
<td>-76.366</td>
<td>-87.450</td>
</tr>
<tr>
<td>Rothera Station</td>
<td>-67.566</td>
<td>-68.133</td>
</tr>
<tr>
<td>Fossil Bluff</td>
<td>-71.333</td>
<td>-68.283</td>
</tr>
<tr>
<td>Sky-Blu</td>
<td>-74.856</td>
<td>-71.586</td>
</tr>
<tr>
<td>South Pole Station</td>
<td>-89.997</td>
<td>139.272</td>
</tr>
<tr>
<td>Concordia Station</td>
<td>-75.103</td>
<td>123.358</td>
</tr>
</tbody>
</table>

Each focal plane had a different functional time until onset of helium film noise. Ground testing in Antarctica made it clear that the 6 focal planes fell into three distinct groups — 8-hour (X1 and X6), 12-hour (X4 and X5), and 24-hour (X2 and X3). To minimize data loss subject to these constraints, all 6 fridges were cycled at the end of every LST day. X4/X5 were cycled at noon every LST day. X1/X6 were cycled every 8 hours, which notched out sections of the Spider maps.

4.2 Recovery

The payload was dropped on January 17, 2015 and landed at the GPS coordinates S76°21.9122′ W87°27.1499′. There are no United States bases near the drop site. The nearest station is the Rothera Research Station run by the British Antarctic Survey (BAS). All recovery efforts were lead by BAS using BAS resources. The Spider payload was recovered in two trips, one shortly after termination in February 2015 and one at the beginning of the next Austral summer in November 2015.

4.2.1 February 2015 Recovery

The first recovery effort was a quick trip targeting only small, high value items, the most important of which were the hard drives holding the flight data. This allowed the Spider collaboration to start the analysis process immediately, instead of waiting for the next Austral summer. Also recov-
4.2.2 November 2015 Recovery

were the master control computers, 3 star cameras, the flight computer, the communications package, and the GoPros which recorded the launch.

The initial recovery was conducted entirely by BAS, lead by Sam Burrell. The team landed on February 2, 2015 at 1500 UTC and was on the ground for 6.5 hours.

4.2.2 November 2015 Recovery

The remaining payload was recovered in November 2015. After a winter in Antarctica, the majority of the payload was encased in snow. Figure 4.5 shows a photo of the payload in February 2015 on the left and in November 2015 on the right. I flew out with BAS to the landing site and led the recovery. See Appendix E for some notes about things I wish I knew before leaving for recovery.

The primary goal of the second recovery was to remove all high value items from the payload before the heavy components were dragged to the Antarctic coast for removal by the Polarstern, a German icebreaker and research vessel. The most important recovered items were the telescope inserts. All 6 were removed at the landing site, and flown out on the Twin Otters. In addition to the inserts, optical filters, half-wave plates, detector readout crates, and the temperature readout computer were also removed and flown back. The cryostat body was loaded onto a large polyethylene sled and dragged to the Antarctic coast by the PistenBully.

It was obvious at recovery that the main tank was no longer mounted to the vacuum vessel. The flexures holding the main tank were likely broken during chute shock. This made the cryostat unusable and expensive to repair. Upon arrival back at Princeton, the cryostat was stripped of valuable items and recycled, shown in Figure 4.6

---

* The sled is affectionately called the magic carpet.
Figure 4.1: The SPIDER flight path, shown in black. The flight started at McMurdo Station, and the line ends at the end of science data. The payload floated for three additional days before dropping at the orange x. McMurdo Station (orange dot), Rothera (pink), and Fossil Bluff (yellow) are also shown on the map. South Pole is shown as a black cross in the middle of the map. This plot was made using the Basemap package for Python [60].

Figure 4.2: SPIDER on the launch pad. The payload is suspended from the launch vehicle on the left. The red strip in the center left is the parachute. The balloon at nearly full inflation is on the right.
Figure 4.3: The altitude of the payload through the full flight. The plot begins at launch and ends at termination. The grayed region is the science period of the flight.
Figure 4.4: The Spider sky coverage. This is adapted from Sasha Rahlin’s thesis [61]. The background is the 150 GHz polarized dust component of the Commander foreground maps [62]. The bright region is the galactic plane. Included are also the regions of BICEP2 [63]; SptPol, and Spt3G [64, 65]; ACTPol and AdvACT [66, 67]; QUIET [68]; POLARBEAR [69]; and the proposed observation region of SIMONS ARRAY [70]. BICEP3’s observation overlaps with the BICEP2 region, but is larger.
4.2.2 November 2015 Recovery

Figure 4.5: Left: The payload at the landing site 2 weeks after termination. Right: The payload after a season in Antarctica. The snow built up several feet and was heavily impacted. Removal from the snow required the use of a PistenBully; removing by hand would have been impossible in the 1 week recovery window. The left photo is courtesy of Sam Burrell.

Figure 4.6: The main tank being built at Redstone in Colorado on the left. The main tank being scrapped in a New Jersey scrap yard on the right. Left photo by Jón Guðmundsson. Right photo provided by the scrapping company.
Part III

Analysis
Drive any direction away from Princeton, and the New England manors guarded by white picket fences are replaced by comfortable houses protected by rows of New Jersey sweet corn. The roads meander more freely. The grass really is a little greener and a little more unkempt.

About twenty minutes away from campus is a farm that sells whole chicken, lamb, and pigs. It’s off a small road and uses a small dirt patch as its parking lot. The smell of hay and manure permanently hangs in the air, but it’s not entirely unpleasant. I’ve hosted two whole pig roasts, and I’ve gotten both my pigs from here.

To order a pig, you enter the slaughterhouse which doubles as their office. The farmer, between gutting various animals, takes you out to the pen where three or four dozen pigs live. If you are anything like me, you have no idea how to choose a pig. The pigs mill about, ignoring the farmer walking between them. When you finally choose one, the farmer pulls out a sharpie and marks the pig for slaughter. You pay and leave.

Early the next morning, you arrive to pick up the pig and the farmer uncere-
moniously drops the now gutted pig into your waiting arms. It is still warm. You
put the pig into the trunk of your car, wrapped in plastic, and drive away.

For the first pig roast, we trussed the pig to the spit as soon as we got home. We lit the coals and tended the roast for 10 hours. Sipping beers and watching a pig roast for nearly half a day works up quite an appetite. And when the cooking was finally done, everyone was ready to devour the pig. The first few bites were glorious - smoky, crackling pig skin enveloping unctuous meat. But as we dug deeper into the pig, we discovered the unmistakable taste of innards, which has the aroma of shit. We did not know that even though the animal had been gutted, the residual smell of the innards naturally permeate the meat.

Armed with this knowledge at the second pig roast, we scrubbed the inside of the pig by hand using salt as an abrasive. By the end of this process, our hands were raw, salt was in our nails, our entire bodies were covered in pig extract, and we smelled awful. But all this effort was well worth it. The roasted pig was perfect. So remember, the next time you are biting into a delicious pork chop, someone had to stick their arms elbow deep into a dead pig’s chest cavity to make it taste right.

5.1 Lowlevel Theory

Section 3.4 describes in detail various pathologies of non-Gaussian detector noise in the SPIDER dataset. Here we will discuss tools to mitigate these effects in the time domain. The goal will be to remove as much noise as possible while preserving as much signal. Since the filtering in this section works in the time domain, let us first look at the expected sky signal in the time domain given SPIDER’s scan strategy. Figure 5.1 show the power spectral density (PSD) of a detector timestream that is generated by reobserving a Planck 143 GHz map that has been smoothed by the SPIDER beam. The SPIDER beam suppresses high-$\ell$ power. Note that the power in this plot is dominated by temperature anisotropies.

The goal of lowlevel processing is to find and flag irrecoverably noisy data before it gets to the
mapmaker, all while preserving as much data as possible. The processing of the timestream data relies almost entirely on running specific kernels through the detector data and making cuts based on the output. The type and specific parameters of the kernel depend on the features of the noise which we want eliminated. In this subsection, we will discuss some of the kernels used in SPIDER.

5.1.1 Kernel to Transfer Function

Before exploring the kernels themselves, it is helpful to consider the transfer function of a kernel. Every kernel functions as a filter; the question is what features it picks out and which features it suppresses. The transfer function describes this in the Fourier domain.

For a detector timeline \( x(t) \) and kernel \( h(t) \), the convolution is defined

\[
y(t) = \int_{-\infty}^{\infty} x(\tau) h(t - \tau) d\tau.
\] (5.1)

If the Fourier transform of \( x(t) \) is \( \mathcal{F}[x(t)] = \tilde{x}(k) \) and \( \mathcal{F}[y(t)] \) is \( \tilde{y}(k) \), then the transfer function is simply given by

\[
H(k) = \frac{\tilde{y}(k)}{\tilde{x}(k)}.
\] (5.2)

Typically we work with discrete values rather than continuous data. So consider the timeseries \( x \) with \( N \) elements and the kernel \( h \) with \( N_k \) elements. Then the \( n \)th element of the convolved timeline \( y \) is

\[
y_n = \sum_{j=-N_k/2}^{N_k/2} x_{j+n} h_j.
\] (5.3)

Now we must consider the discrete Fourier transform, \( Y_k \), of \( y \)

\[
Y_k = \sum_{n=0}^{N-1} \sum_{j=-N_k/2}^{N_k/2} x_{j+n} h_j \exp\left(-2\pi i k \frac{n}{N}\right)
\] (5.4)
5.1.2 Gaussian Derivative Kernel

We can explicitly write out the sum on \( j \)

\[
Y_k = \sum_{n=0}^{N-1} \left[ \left( x_{n-N_k/2} h_{-N_k/2} + x_{n+1-N_k/2} h_{1-N_k/2} + \cdots + x_{n+N_k/2} h_{N_k/2} \right) \times \exp \left( -2\pi j k \frac{n}{N} \right) \right]
\]  

We know from the shift theorem that the Fourier transform of a timestream that is shifted by \( m \) samples

\[
\sum_{n=0}^{N-1} x_{n-m} \exp \left( -2\pi i k \frac{n}{N} \right) = \sum_{n=0}^{N-1} x_n \exp \left( -2\pi j k \frac{n+m}{N} \right)
\]  

Now consider the first term in the summation in eq. 5.5

\[
\sum_{n=0}^{N-1} x_{n-N_k/2} h_{-N_k/2} \exp \left( -2\pi j k \frac{n}{N} \right) = \sum_{n=0}^{N-1} x_n h_{-N_k/2} \exp \left( -2\pi i k \frac{n}{N} \right) \exp \left( 2\pi i k \frac{N_k/2}{N} \right)
\]

\[
= X_k h_{-I} \exp \left( 2\pi i k \frac{N_k/2}{N} \right)
\]  

If we consider all the terms in the summation in eq. 5.5 we find

\[
Y_k = X_k \sum_{j=-N_k/2}^{N_k/2} h_j \exp \left( -2\pi i k \frac{j}{N} \right).
\]  

The right side of this equation is just the discrete Fourier transform of \( h_i \). So the transfer function of the kernel is

\[
H_k = \frac{Y_k}{X_k} = \sum_{j=-N_k/2}^{N_k/2} h_j \exp \left( -2\pi i k \frac{j}{N} \right).
\]  

5.1.2 Gaussian Derivative Kernel

We will begin the discussion of kernels with the Gaussian derivative kernel, which is an implementation of a Canny edge detector \[71\]. As the name implies, start with the Gaussian function

\[
g(x|\sigma) = \frac{1}{\sigma \sqrt{2\pi}} \exp \left( -\frac{x^2}{2\sigma^2} \right)
\]  

\[74\]
5.1.2 Gaussian Derivative Kernel

and take its derivative,

\[ g'(x|\sigma) = \frac{-x}{\sigma^3\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma^2}\right). \]  \hspace{1cm} (5.11)

For now treat the kernel as a continuous function, so using eq. 5.2 we find

\[ H(k) = \mathcal{F}[g'(x)] = \frac{ik}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}\sigma^2k^2\right). \]  \hspace{1cm} (5.12)

The easiest way to derive this is to leverage the fact that \( \mathcal{F}[z'] = ik\mathcal{F}[z] \). To see how this is useful, consider the convolution of the Gaussian derivative kernel with a Heaviside (or step) function, \( \mathcal{H} \). The Fourier transform of a Heaviside function is

\[ \mathcal{F}[\mathcal{H}] = \frac{1}{2} \left[ \delta(k) - \frac{i}{\pi k} \right]. \]  \hspace{1cm} (5.13)

From the transfer function in eq. 5.12 we can see that the Gaussian derivative kernel allows the imaginary component through. This component dominates eq. 5.13 at \( k \neq 0 \). White noise by definition is flat in \( k \)-space, so it is heavily down weighted by this kernel. Thus the Gaussian derivative kernel is ideal for finding step functions.

To see what this means in a practical application, we will switch to the discrete Fourier basis, as the SPIDER time domain data is discrete. Figure 5.2 shows a worked example. The top left shows the kernel, the top right shows its transfer function, and the bottom is the kernel convolved with a Heaviside function. It is obvious that the maximum of the convolved timestream coincides with the location of the step in the original timestream.

Tuning the width, \( \sigma \), of the Gaussian derivative kernel is dependent on the nuances of the data. For SPIDER, most steps occur during Iridium pings, so there is no information in the samples near the step function. This forces SPIDER to use a wide \( \sigma \) value.

For a kernel that is much wider than \( \sigma \), we can normalize the kernel

\[ \kappa(x) = \frac{1}{\sigma^2} x \exp\left(-\frac{x^2}{2\sigma^2}\right). \]  \hspace{1cm} (5.14)
This has the feature that \( \int_{-\infty}^{\infty} \kappa(x)dx = 1 \) and \( \int_{-\infty}^{0} \kappa(x)dx = -1 \), making the maximum value of the kernel convolved timestream equal to the amplitude of the DC step.

### 5.1.3 Triangle Kernel

The triangle kernel is perhaps the simplest kernel one could write. This specific kernel is applied to discrete data. The kernel is simply \([-1/2, 1, -1/2]\), which has transfer function \( H(k) = [3/2, 0, 3/2] \). The transfer function has no imaginary component. Since the kernel is so small, it is fast to convolve with the data. Obviously this is a very simple, but perhaps peculiar, highpass filter.

The transfer function, shown in Figure 5.3, shows that this kernel suppresses even \( k \)-modes while leaving odd \( k \)-modes largely unchanged. We can extend the triangle filter by adding zeros between the \(-1/2\)s and the center 1. For example, with 2 spaces the kernel becomes \([-1/2, 0, 0, 1, 0, 0, 1/2]\). The transfer function for various spacings is also shown in Figure 5.3.

### 5.1.4 Tophat Kernel

The tophat kernel is another trivial but useful kernel. The kernel can be of any width and every element has the same value. If the kernel has width \( w \) and amplitude \( 1/w \), then the kernel simply takes the mean of the samples in the kernel width. The transfer function of this kernel is zero everywhere except for \( k = 0 \), where it is unity. In other words, this is a strong lowpass filter.

### 5.2 Timestream Cleaning

With the knowledge of various kernels in hand, we are now able to clean actual SPIDER data. Life is never as easy as the simple models described in Section 5.1, however those models will provide a guide for the actual cleaning.

#### 5.2.1 Iridium

Transmitter pickup was discussed in Section 3.4.3. The primary source of radio-frequency transmitter noise is the Iridium antenna. An example of an Iridium pulse is shown in Figure 3.11.
Iridium pulses are extremely large relative to the white noise and are easy to detect. There are a handful of detectors that are particularly sensitive to Iridium for unknown reasons. We use these channels to quickly generate an instrument wide flag based on deviations from a threshold value. In practice this threshold is not very important because of the extremely large amplitude of the Iridium pulses.

### 5.2.2 Cross Jumps

The first step to solving cross jumps is to attempt to find them. We will employ the Gaussian derivative kernel outlined in Section 5.1.2. Figure 5.4 shows an example timestream, shown in blue, convolved with the kernel, shown in orange. The white noise is lower in convolved timestream, as we expected. The kernel is normalized so the maximum of the convolved timestream corresponds to the best estimate of the step amplitude. We then define a threshold based on the noise level of the convolved timestream, and every spike with amplitude greater than this threshold is considered a step.

In theory, this should fully solve the DC step problem. Simply subtract the amplitude of the DC step from the raw data and we have fully corrected the DC discontinuity. Unfortunately the existence of the known dipole and unknown scan synchronous noise inject an unknown offset over these time periods which is significant relative to our required sensitivity.

Instead, we leverage the knowledge that cross jumps are caused by flux jumps. The MCE catches most flux jumps on its own and records them. Missed flux jumps are easy to find because they are large discontinuities, with amplitude of many thousands of ADUs. We then correlate every cross jump with a known flux jump. This too is made complicated by the fact that flux jumps are correlated because they are caused by Iridium. The only way to disambiguate this is to construct a full flux jump/cross jump matrix and solve for a correlated amplitude. This matrix for 1 MCE column is shown in Figure 5.5.

We can see that for this MCE column, if row 2 flux jumps, almost every detector in the column cross jumps. A similar feature is in row 23. The cause of this effect is unknown, but is probably related to the SQ2. There is also a feature in the diagonal of the matrix, suggesting that when any
5.3 Noise Power Spectrum

detector flux jumps, its neighboring detector cross jumps. The time domain multiplexing reads out
detectors in sequential row order, so the diagonal feature may be due to the sequential readout.

5.3 Noise Power Spectrum

With the entire suite of flags, we hope to have data which is composed of only signal, white noise,
and $1/f$ noise. The $1/f$ noise can be sourced by drifting fridges, readout issues in the MCE, or
external loads from sidelobes. To minimize this $1/f$ component, we subtract a 5th order polynomial
per half-scan before binning the data into a map. The processing is shown in Figure 5.6. The top
panel is the untouched detector timestream. The middle panel shows the data with the flag applied.
The bottom panel shows the flagged data with the 5th order polynomial subtracted and then gap
filled. The gap filling does not go into the mapmaker and is only there to allow Fourier operations.

5.3.1 Naive Detector Noise

First we will consider the noise of a single detector. For a timestream $x(t)$ with elements $x_n =
x(n\Delta t)$, measured of the time period $T = N\Delta t$, we define the power spectral density (PSD)

$$P(f) = \frac{(\Delta t)^2}{T} \left| \sum_{n=1}^{N} x_n e^{-2\pi ifn\Delta t} \right|^2 = \frac{(\Delta t)^2}{T} \tilde{x}(f)\tilde{x}(f)^*.$$  \hspace{1cm} (5.15)

Figure 5.7 shows the PSD of one of the SPIDER detectors in flight. The CMB, dipole, and
scan synchronous components are subtracted from the data. A first order polynomial is subtracted
from all unflagged 1 minute chunks, then their PSD is calculated. These values are then averaged
to generate the curve shown. The first order polynomial subtraction per 1 minute chunk heavily
suppresses the first bin.
5.3.1 Naive Detector Noise

5.3.1.1 Noise Equivalent Power

The PSD allows for quick calculation of the noise equivalent power (NEP) \[\text{NEP}\]. The NEP is the signal power required to generate signal-to-noise ratio of one in a one hertz output bandwidth. NEP is convenient because many of the noise terms are intrinsic to the detector and readout system, which are easily expressible in power units.

As a brief aside, many implementations of power spectral estimators will return the PSD with the time unit as \(1/\sqrt{\text{Hz}}\). The CMB community often prefers NEPs in units of \(\sqrt{s}\). This allows quick calculations of detector or instrumental sensitivity as a function of integration time. To change from units of \(1/\sqrt{\text{Hz}}\) to \(\sqrt{s}\),

\[
P[\sqrt{s}] = \frac{P[1/\text{Hz}]}{\sqrt{2}}. \tag{5.16}
\]

The factor of \(\sqrt{2}\) comes from the fact that an output bandwidth of one hertz is equivalent to half a second of integration time due to the Nyquist-Shannon theorem.

5.3.1.2 Noise Equivalent Temperature

Alternatively, we can define the noise equivalent temperature (NET). NET is the noise in \(\mu K_{\text{CMB}}\) units. This requires knowledge of the efficiency of the optical system, \(\eta\). In other words, this is what fraction of total photons (or total power) incident on the instrument are actually measured. Typical efficiencies are 30%. Starting from the Planck blackbody equation, the spectral radiance is given by

\[
B_\nu(\nu, T) = \frac{2h\nu^3}{c^2} \frac{1}{\exp(x) - 1} \tag{5.17}
\]

where

\[
x = \frac{h\nu}{k_B T} \tag{5.18}
\]

\(\nu\) is the frequency, \(c\) is the speed of light, \(h\) is the Planck constant, \(k_B\) is the Boltzmann constant, and \(T\) is temperature. We can calculate

\[
\frac{dP}{dT} = \eta \frac{d}{dT} \left[ \int_{\nu_0}^{\nu_f} B_\nu(\nu, T) d\nu \right] \tag{5.19}
\]
where $P$ is the power incident on the detector, and $\nu_0$ and $\nu_f$ define the bandpass of the detector. Then we can write

$$\text{NET} = \frac{dT_{\text{CMB}}}{dP} \text{NEP}. \quad (5.20)$$

SPIDER’s NETs are $7.1 \mu K \sqrt{s}$, $5.3 \mu K \sqrt{s}$ at 95 and 150 GHz with 675 detectors and 1,188 detectors at 95 and 150 GHz.

5.3.2 Correlated Noise

The mapmaker used by SPIDER, outlined in Chapter 7, assumes that the detector noise is uncorrelated. At some level, that statement must be false. For example, all the detectors in the focal plane are coupled to the same thermal bath, so they must suffer from the same thermal drifts. Some detectors share a bias line. Others share an SQ2. All of these can couple noise to the detectors in a correlated manner.

5.3.2.1 Cross Spectral Density

One way to explore this is to take the cross correlation of two detector timestreams in the Fourier domain, otherwise known as the cross spectral density (CSD). For two detector timestreams $x(t)$ and $y(t)$, and the Fourier transforms $\tilde{x}(f)$ and $\tilde{y}(f)$, the cross spectral density is

$$C(f) = \frac{(\Delta t)^2}{T} \tilde{x}(f)\tilde{y}(f)^* \quad (5.21)$$

This quantity describes the degree to which the detectors vary together by frequency. This is particularly powerful because we expect things like bath temperature drifts to be very slow, so it should appear as a $1/f$ term in the correlation. Noise sourced by readout issues can present anywhere in frequency space.

Figure 5.8 shows the Fourier correlation between two pairs of detectors. On the left is a pair that is positively correlated, and on the right is a pair that is anti-correlated. This correlation is stable across a day, and the plots are generated by taking cross correlations of approximately 1 minute chunks of data and averaging them.
5.3.2 Correlated Noise

5.3.2.2 Noise Correlation Matrix

To explore the correlations further, we construct a noise correlation matrix for an MCE column. This allows us to view the correlations across many detectors. We average the CSDs over frequency bands, specifically 0.5-2 and 5-10 Hz. We then normalize the CSDs with their auto spectra. Then

$$C_{ij}^{\text{norm}} = \frac{C_{ij}}{\sqrt{C_{ii}C_{jj}}} \quad (5.22)$$

where $C$ indicates an average over the frequency band. Note that $P_i = C_{ii}$.

The real component of the correlation matrix for a single SPIDER column is shown in Figure 5.9. We also calculate the imaginary component of the correlation matrix; however, it is effectively 0 for SPIDER so it is not shown here. The diagonal element is omitted because that is simply the PSD, which would be unity with this normalization. The left is the 0.5-2 Hz bin, and the right is the 5-10 Hz bin. These correspond to the gray bands in Figure 5.8. The detectors in Figure 5.8 are MCE rows $8 \times 22$ on the left and $8 \times 23$ on the right.

In this specific MCE row, the values in the 0.5-2 Hz bins are largely the same as the values in the 5-10 Hz bin. Thus the correlated noise is flat in power spectrum across frequencies. This agrees with the power spectrum shown in Figure 5.8.

Figure 5.9 says the correlated noise is 15% of the total noise, even out to 10 Hz. This is very surprising. Long drifts in the focal plane temperatures are well below 10 Hz, and should appear as a $1/f$ component. Even more surprising is the large number of anti-correlated detectors and the periodic nature of the correlations as a function of MCE row. This implies that the problem almost certainly is in the readout architecture.

Fortunately, the column in Figure 5.9 is not typical. Figure 5.10 shows another correlation matrix which is more typical for SPIDER. The correlation goes approximately as $1/f$. Neighboring readout rows, the off diagonal elements of the matrix, have a slightly higher noise correlation, consistent with known crosstalk. There are no anti-correlations.

The full FPU noise correlation matrices are shown in Figures 5.11 to 5.16. The strong correlation in X2 column 1 is due to known popcorn noise. X3 is strongly correlated due to reaction
5.3.2 Correlated Noise

wheel noise. This is minimized, although not fully removed, by subtracting the reaction wheel noise template.
Figure 5.1: The PSD of a detector timestream generated by reobserving a Planck 143 GHz map using the SPIDER scan strategy. The blue line is CMB and dipole. This is the best estimate of the actual cosmological sky signal onto the detectors. The orange line is CMB only. The yellow line is the CMB only signal filtered with a 5th order polynomial every turnaround. The vast majority of this power is from temperature anisotropies, which are much brighter than the signals of interest for SPIDER, the polarization anisotropies.
5.3.2 Correlated Noise

Figure 5.2: A worked example of the Gaussian derivative kernel. All data is unitless in this example but can be thought of as detector amplitudes in ADU. The top left is the Gaussian derivative kernel. The top right is its Fourier transform. The bottom is the kernel convolved with a Heaviside function. The maximum of the convolved function locates the sample where the step happens.

Figure 5.3: Transfer function of the triangle kernel for different spacings. The kernel suppresses even modes and nulls DC. This acts as a highpass filter.
Figure 5.4: A channel with cross jumps convolved with the Gaussian derivative kernel. Blue is the raw data and orange is the kernel convolved data. The kernel is normalized so that the maximum amplitude of the kernel convolved timestream is the amplitude of the DC step.
Figure 5.5: The matrix showing the best flux jump to cross jump response. The x-axis is the MCE row number of the detector that is flux jumping. The y-axis is the MCE row number of the detector that is responding to the flux jump. The color indicates the sign and amplitude of the cross jump. This is generated by averaging DC level shifts on response detectors every time a source detector flux jumps. Not every detector flux jumps, so some columns in the matrix are empty. Not every detector cross jumps, so some rows in the matrix are empty.
5.3.2 Correlated Noise

Figure 5.6: An example of timestream cleaning. The top panel shows the raw detector timestream. The orange vertical bars show where data has been flagged. The middle panel shows the flag applied. The bottom panel shows the data that has been poly-5 filtered per half-turn and gap filled.

Figure 5.7: The power spectral density of a SPIDER bolometer in flight. The data is CMB, dipole, scan synchronous subtracted. Then contiguous, unflagged 1 minute chunks are Fourier transformed after removing a first order polynomial. The PSD is then estimated from all the 1 minute chunks in an LST day.
5.3.2 Correlated Noise

Figure 5.8: The cross FFT of two correlated detectors on the left and anti-correlated detectors on the right. The top panel shows the real and imaginary components of the cross FFT, and the bottom panel is the phase. These plots are generated by taking the FFT of all continuous 1 minute chunks in a day and averaging them. The majority of detector pairs are uncorrelated, the primary exception being detectors on X1. The gray bands indicate the 0.5-2 and 5-10 Hz bins which will be used for discussing correlated noise matrices in Section 5.3.2.2.
5.3.2 Correlated Noise

Figure 5.9: The real component of the correlation matrix for one SPIDER column. The left is the 0.5-2 Hz bin; the right is the 5-10 Hz bin. The values are normalized with their PSDs; thus, diagonal elements are unity by construction so they are omitted. The detector on MCE row 15 is dead. The cross spectral density shown in the left plot of Figure 5.8 is MCE rows 8×22. Similarly the right plot of Figure 5.8 is MCE rows 8×23 on the right.
5.3.2 Correlated Noise

Figure 5.10: A more typical noise correlation matrix. The low frequency bin on the left has greater correlations than the higher frequency bin on the right. This is broadly consistent with the expected $1/f$ nature of correlations. The elevated off diagonal elements are due to known crosstalk between neighboring readout rows.
5.3.2 Correlated Noise

Figure 5.11: The full FPU correlation matrix for X1. All elements are normalized by their PSDs and the diagonal elements are not shown. The correlations are binned between 1-3 Hz and 3-10 Hz on the top and bottom respectively.
5.3.2 Correlated Noise

Figure 5.12: The full FPU correlation matrix for X2. All elements are normalized by their PSDs and the diagonal elements are not shown. The correlations are binned between 1-3 Hz and 3-10 Hz on the top and bottom respectively.
Figure 5.13: The full FPU correlation matrix for X3. All elements are normalized by their PSDs and the diagonal elements are not shown. The correlations are binned between 1-3 Hz and 3-10 Hz on the top and bottom respectively.
5.3.2 Correlated Noise

Figure 5.14: The full FPU correlation matrix for X. All elements are normalized by their PSDs and the diagonal elements are not shown. The correlations are binned between 1-3 Hz and 3-10 Hz on the top and bottom respectively.
5.3.2 Correlated Noise

Figure 5.15: The full FPU correlation matrix for X5. All elements are normalized by their PSDs and the diagonal elements are not shown. The correlations are binned between 1-3 Hz and 3-10 Hz on the top and bottom respectively.

Figure 5.15: The full FPU correlation matrix for X5. All elements are normalized by their PSDs and the diagonal elements are not shown. The correlations are binned between 1-3 Hz and 3-10 Hz on the top and bottom respectively.
5.3.2 Correlated Noise

Figure 5.16: The full FPU correlation matrix for X6. All elements are normalized by their PSDs and the diagonal elements are not shown. The correlations are binned between 1-3 Hz and 3-10 Hz on the top and bottom respectively.

96
Chapter 6

Calibration

6.1 Beam and Calibration

The instrumental beam determines how sky signal couples into the detectors. The calibration is effectively the value that connects power from the sky into detector readout units. We need both of these values to construct sensible SPIDER maps in units of $\mu$K$_{\text{CMB}}$.

SPIDER determines its beam and calibration from Planck maps. In broad strokes, Planck maps are smoothed with a theoretical SPIDER beam and compared to the SPIDER data maps. The input theoretical beam is altered slightly, and the residual with Planck is recomputed. The beam that minimizes the residual is considered the beam which is used in simulations. The focal plane averaged beams are shown in Figures 6.1 and 6.2.

The best fit model is a sum of 3 Gaussians. The dominant component is the full width half max of the main beam. We believe we have a sidelobe, which is currently modeled as a high width Gaussian. This is likely unphysical, in the sense that the sidelobe is probably dominated by a small beam at wide angle. The beam parameter values are shown in Table 6.1. Note that the beam and calibration are degenerate up to a single scalar value.

One way to check this model is to take cross spectra with Planck maps. In addition to the full mission maps, Planck released half-mission maps. These maps have independent noise. This allows
Figure 6.1: The per focal plane SPIDER beams at 95 GHz. The beams are modeled as the sum of 3 Gaussians beams. The top panel is the beam model. The bottom panel is the difference of the model and data divided by the model.

Figure 6.2: The per focal plane SPIDER beams at 150 GHz. The beams are modeled as the sum of 3 Gaussians beams. The top panel is the beam model. The bottom panel is the difference of the model and data divided by the model.
6.2 Deprojection

<table>
<thead>
<tr>
<th></th>
<th>width 1</th>
<th>Amp 1</th>
<th>width 2</th>
<th>Amp 2</th>
<th>width 3</th>
<th>Amp 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>X1</td>
<td>28.39</td>
<td>.99</td>
<td>42.17</td>
<td>.83</td>
<td>3.95</td>
<td>.05</td>
</tr>
<tr>
<td>X3</td>
<td>28.21</td>
<td>.98</td>
<td>42.37</td>
<td>1.10</td>
<td>3.83</td>
<td>.05</td>
</tr>
<tr>
<td>X5</td>
<td>28.88</td>
<td>.99</td>
<td>52.08</td>
<td>2.80</td>
<td>5.30</td>
<td>.06</td>
</tr>
<tr>
<td>X2</td>
<td>41.62</td>
<td>1.00</td>
<td>30.61</td>
<td>.83</td>
<td>4.74</td>
<td>.08</td>
</tr>
<tr>
<td>X4</td>
<td>41.23</td>
<td>.98</td>
<td>34.50</td>
<td>2.50</td>
<td>3.37</td>
<td>.07</td>
</tr>
<tr>
<td>X6</td>
<td>41.12</td>
<td>1.00</td>
<td>34.77</td>
<td>2.97</td>
<td>4.80</td>
<td>.09</td>
</tr>
</tbody>
</table>

Table 6.1: The triple Gaussian beam parameters. The width is in units of arcminutes. The first parameter can be thought of as the width of the main beam. Note that the amplitudes are degenerate with calibration.

This allows us to take cross spectra with no noise bias. We consider

\[
\frac{S \times P_1}{P_1 \times P_2} = \frac{C_\ell F_\ell B_\ell^S B_\ell^P}{C_\ell F_\ell B_\ell^P B_\ell^P} = \frac{B_\ell^S}{B_\ell^P}
\]

(6.1)

where \( S \) is a SPIDER map, \( P_1 \) is the resampled Planck half-mission 1 map, \( P_2 \) is the resampled Planck half-mission 2 map, \( F_\ell \) is the filter transfer function, \( B_\ell^P \) is the Planck beam window function, and \( B_\ell^S \) is the SPIDER beam window function. Since Planck released its beam window function, we can solve for \( B_\ell^S \). The bottom panels in Figures 6.1 and 6.2 are the difference between the model and the solution for \( B_\ell^S \) described here divided by the model.

The noise begins to dominate the solution at high multipoles. This is because the wide SPIDER beams have very little sensitivity on small angular scales. Thus the SPIDER map, \( S \), is effectively a noise only map at those scales. The SPIDER 150 GHz beam is narrower than the SPIDER 95 GHz beam, so the noise dominates earlier at 95 GHz.

We generate calibrations per detector, but we determine a beam per focal plane. This averages down known beam effects like variable beam widths and beam ellipticity. We will explore those with deprojection.

6.2 Deprojection

SPIDER’s analysis pipeline implements a novel use case for deprojection. Deprojection is an analysis
6.2 Deprojection

technique developed by the BICEP/Keck Array experiment \cite{73,75}. Deprojection is a process used to characterize and remove spurious polarization signal due to instrumental uncertainties. Uncertainties in properties like gain and differential pointing couple temperature anisotropies, $\Theta$, into polarization, $P$. SPIDER’s implementation considers 4 sources of spurious polarization parameterized by 6 variables: differential gain ($\delta g$), differential pointing ($\delta x, \delta y$), differential beam width ($\delta \sigma$), and differential ellipticity ($\delta p, \delta c$). The analysis relies on derivatives of temperature maps. Generating these maps is described in Appendix C.

The BICEP/Keck Array mapmaker pair differences co-pointed, crosspolar detectors before binning into a map. This eliminates the temperature anisotropies as well as any common systematics, particularly the atmospheric emission. The assumption is any remaining signal in the differenced timeline is a combination of polarized signal and noise. However, miscalibrations or asymmetries between A and B detectors would result in imperfect cancellation of $\Theta$ as well as systematics, creating a leakage signal. The mapmaker treats the leaked signal as polarized signal. Even worse, many of the asymmetries are constant, so the leaked signal does not average down with more observations. For example, beam asymmetries are typically due to the geometry of the physical optics, which are invariant across seasons. The time stability of the beam parameters is explored in Figures H.13 to H.18. Deprojection attempts to remove these signals, and is crucial for the success of BICEP/Keck Array’s mapmaker.

The SPIDER mapmaker does not explicitly pair difference its own detectors, so the implementation and use of deprojection is slightly different. Instead SPIDER differences its timestream with a reobserved Planck timestream. Another key difference between the SPIDER and BICEP/Keck Array implementations is that the BICEP/Keck Array deprojection pipeline works in the map domain, while the SPIDER one is done in the time domain. In the SPIDER implementation, deprojection is an powerful tool for constraining pointing, beam parameters, and gain drifts. In fact, SPIDER’s best boresite pointing solution comes from deprojection rather than the dedicated pointing sensors. Moreover, it provides an efficient and flexible method to check beam systematics in a computationally efficient manner.

Sections 6.2.1 to 6.2.4 derive the templates and attempt to give some intuition. Section 6.2.5
6.2.1 Differential Gain

describes how to fit the templates.

6.2.1 Differential Gain

The simplest coupling of temperature into polarization is from differential gain. Even though the SPIDER mapmaker does not explicitly use difference timelines to solve for maps, gain miscalibrations still induce spurious polarization signals. As a contrived example, imagine two detectors with identical pointing and beams, but with mismatched gains

\[ \delta g = \frac{g_A - g_B}{2}. \]  

(6.2)

We will employ the mapmaker formalism which is discussed in Chapter \[ \text{7} \]. To ensure that the mapmaker equation is well conditioned, imagine each detector observes the same point in the sky twice, once with half-wave plate angle 0 degrees and once with half-wave plate angle 22.5 degrees. Also for simplicity make this example noiseless. Then the pointing vector is

\[
\mathbf{A} = \begin{bmatrix}
1 & 1 & 0 \\
1 & 0 & 1 \\
1 & 1/\sqrt{2} & 1/\sqrt{2} \\
1 & 1/\sqrt{2} & -1/\sqrt{2}
\end{bmatrix}
\]  

(6.3)

where the observations are ordered \([A_0, B_0, A_{22.5}, B_{22.5}]\). Consider detector timelines with temperature anisotropies as the only input. Solving the noiseless mapmaker equation

\[
\hat{s} = \left(\mathbf{A}^T \mathbf{A}\right)^{-1} \mathbf{A}^T \mathbf{d}
\]

\[
= \begin{bmatrix}
4 & 1 + \sqrt{2} & 1 \\
1 + \sqrt{2} & 2 & 0 \\
1 & 0 & 2
\end{bmatrix}^{-1} \begin{bmatrix}
1 & 1 & 1 & 1 \\
1 & 0 & 1/\sqrt{2} & 1/\sqrt{2} \\
0 & 1 & 1/\sqrt{2} & -1/\sqrt{2}
\end{bmatrix} \begin{bmatrix}
\Theta + \delta g \Theta \\
\Theta - \delta g \Theta \\
\Theta + \delta g \Theta \\
\Theta - \delta g \Theta
\end{bmatrix}
\]  

(6.4)
we find that
\[
\begin{bmatrix}
I \\
Q \\
U
\end{bmatrix} = \begin{bmatrix}
\Theta \\
(2 + \sqrt{2}) \delta g \Theta \\
\sqrt{2} \delta g \Theta
\end{bmatrix}.
\] (6.5)

So despite inputting temperature only data, a gain mismatch produces spurious Q and U signal. Obviously as $\delta g$ goes to 0, so does the spurious polarization.

While SPIDER does not explicitly difference, I will continue from here talking about difference timelines for simplicity. It is immediately obvious from the example above that the Q and U spurious signals scale as $\Theta$. To better motivate this, consider a timeline for a single detector
\[
d(t) = g [B \cdot (\Theta + P)] (\vec{p}(t)) + n(t)
\] (6.6)

where B is the beam, $n(t)$ is the noise in the detector timestream, and $\vec{p}(t)$ is the detector pointing. This equation simply says that the detector timeline is the sum of two terms. The first term is given by the beam convolved sky signal (with both temperature and polarization) evaluated at a given detector pointing, which is then multiplied by the detector gain. The second term is noise. Then the difference in the timelines between A and B detectors is
\[
\Delta d(t) = d_A(t) - d_B(t) = 2 \delta g (B \cdot \Theta) (\vec{p}(t)) + \delta n(t)
\] (6.7)

where the polarization signal is assumed to be much smaller than the temperature signal and we have defined the term $\delta n(t) = n_A(t) - n_B(t)$. Therefore a template timeline $T_g$ can be constructed based on the nominal beam smoothed temperature map
\[
T_g(t) = g (B \cdot \Theta) (\vec{p}(t)).
\] (6.8)

$B$ is the focal plane average beam width, although results are largely independent of the beam width chosen as long as it is close to the actual beam width. The template timeline is then fit to
6.2.2 Differential Pointing

the difference timeline, and the fit coefficient $\alpha_g$ gives the relative gain mismatch

$$\alpha_g = \frac{2\delta g}{g}.$$  \hfill (6.9)

6.2.2 Differential Pointing

We can now consider a pair of detectors that are pointed a small angle away from their co-pointed nominal beam center

$$\delta \bar{p}(t) = \frac{\bar{p}_A - \bar{p}_B}{2}.$$  \hfill (6.10)

In this section we will assume that the beam shapes and gain are perfectly matched. Then we can calculate the beam function for small displacements using a Taylor expansion

$$B(\bar{p}(t) + \delta \bar{p}) = B(\bar{p}(t)) + \delta \bar{p} \cdot \nabla B(\bar{p})|_{\bar{p}=\bar{p}_0} + \mathcal{O}(\delta \bar{p}^2).$$  \hfill (6.11)

The pair-differenced beam is

$$B_A - B_B = B(\bar{p}(t) + \delta \bar{p}) - B(\bar{p}(t) - \delta \bar{p}) = 2\delta \bar{p} \cdot \nabla B(\bar{p})|_{\bar{p}=\bar{p}_0} + \mathcal{O}(\delta \bar{p}^3).$$  \hfill (6.12)

Eventually the derivative in eq. 6.12 will be operated on a map, which is in celestial coordinates. The detector offsets are invariant in focal plane coordinates. We will denote the celestial coordinates in units of $\theta$ and $\phi$ using the HEALPix conventions. We will denote the focal plane coordinates in $x$ and $y$ which is referenced to the gondola. The two coordinate systems are related by

$$\hat{x} = -\cos \gamma \hat{\phi} + \sin \gamma \hat{\theta}$$  \hfill (6.13)

$$\hat{y} = -\sin \gamma \hat{\phi} - \cos \gamma \hat{\theta}$$  \hfill (6.14)

where $\gamma$ is the sky rotation (in BICEP it is the deck rotation because they have no sky rotation due to being at the South Pole). Then derivatives in the $\hat{x}$ and $\hat{y}$ directions are given by
6.2.2 Differential Pointing

\[
\hat{x} \cdot \nabla = \left( -\cos \gamma \hat{\phi} + \sin \gamma \hat{\theta} \right) \cdot \left( \frac{1}{\sin \theta} \frac{\partial}{\partial \phi} \hat{\phi} + \frac{\partial}{\partial \theta} \hat{\theta} \right) \\
= -\frac{\cos \gamma}{\sin \theta} \frac{\partial}{\partial \phi} + \sin \gamma \frac{\partial}{\partial \theta} 
\]

(6.15)

\[
\hat{y} \cdot \nabla = \left( -\sin \gamma \hat{\phi} - \cos \gamma \hat{\theta} \right) \cdot \left( \frac{1}{\sin \theta} \frac{\partial}{\partial \phi} \hat{\phi} + \frac{\partial}{\partial \theta} \hat{\theta} \right) \\
= -\frac{\sin \gamma}{\sin \theta} \frac{\partial}{\partial \phi} - \cos \gamma \frac{\partial}{\partial \theta} 
\]

(6.16)

Now we can consider two detectors that displaced only in the \( \hat{x} \) direction by angle \( \delta x \). The difference beam is then

\[
B_A - B_B \approx 2\delta x \left( -\frac{\cos \gamma}{\sin \theta} \frac{\partial}{\partial \phi} + \sin \gamma \frac{\partial}{\partial \theta} \right) \cdot B(\vec{p})|_{\vec{p}=\vec{p}_0(t)}. 
\]

(6.17)

The spurious signal from a pointing offset is then

\[
\Delta d(t) = d_A - d_B = g(B_A - B_B) \cdot \Theta(\vec{p}(t)) + n(t) \\
= g \left( -\frac{\cos \gamma}{\sin \theta} \frac{\partial}{\partial \phi} + \sin \gamma \frac{\partial}{\partial \theta} \right) \cdot (B \cdot \Theta)(\vec{p}(t))|_{\vec{p}=\vec{p}_0(t)} 
\]

(6.18)

where we’ve again assumed that the power in polarization is much smaller than the power in temperature. We also take advantage of the fact that the derivative of a beam convolved map is the same the derivative of the beam which is then convolved with the map, ie \( \partial (B \cdot \Theta) = \partial B \cdot \Theta \).

Then the fit template is

\[
T_x(t) = g\sigma \left( -\frac{\cos \gamma}{\sin \theta} \frac{\partial}{\partial \phi} + \sin \gamma \frac{\partial}{\partial \theta} \right) \cdot (B \cdot \Theta)(\vec{p}(t))|_{\vec{p}=\vec{p}_0(t)}. 
\]

(6.19)

In the same vein

\[
T_y(t) = g\sigma \left( -\sin \gamma \frac{\partial}{\sin \theta \partial \phi} - \cos \gamma \frac{\partial}{\partial \theta} \right) (B \cdot \Theta)(\vec{p}(t))|_{\vec{p}=\vec{p}_0(t)}. 
\]

(6.20)
6.2.3 Differential Beam Width

The fit parameters are

$$\alpha_x = \sigma_0 \delta x \quad \alpha_y = \sigma_0 \delta y.$$  \hfill (6.21)

Unsurprisingly, the spurious signal due to differential pointing scales directly with pointing offset.

6.2.3 Differential Beam Width

An elliptical Gaussian beam can be described by its covariance matrix

$$\Sigma = \begin{bmatrix} (1 - p)\sigma^2 & c\sigma^2 \\ c\sigma^2 & (1 + p)\sigma^2 \end{bmatrix}$$  \hfill (6.22)

where

$$B(\bar{x}) = \frac{1}{2\pi\sigma^2} \exp \left[ \bar{x}^T \Sigma^{-1} \bar{x} \right].$$  \hfill (6.23)

We have defined \(\sigma\) as the beam width, \(p\) as the “plus” ellipticity, and \(c\) as the “cross” ellipticity.

First consider a beam with no ellipticity, setting both \(p\) and \(c\) to zero. The general equation for the beam then is

$$B(\bar{x}, \sigma) = \frac{1}{2\pi\sigma^2} \exp \left( -\frac{x^2 + y^2}{2\sigma^2} \right).$$  \hfill (6.24)

Then for a pair of detectors, the pair average beam width is

$$\sigma_0 = \frac{\sigma_A + \sigma_B}{2}$$  \hfill (6.25)

and the pair difference beam width is

$$\delta\sigma = \frac{\sigma_A - \sigma_B}{2}.$$  \hfill (6.26)

Then plugging this definition of differential beam width into eq. \ref{6.24} we find

$$B(\bar{x}, \sigma) = \frac{1}{2\pi(\sigma + \delta\sigma)^2} \exp \left( -\frac{x^2 + y^2}{2(\sigma + \delta\sigma)^2} \right).$$  \hfill (6.27)
6.2.3 Differential Beam Width

Assuming that the differential beam width is small, we Taylor expand \( B(\vec{x}, \sigma_A) \) about the pair averaged beam center.

\[
B(\vec{x}, \sigma_A) \approx B(\vec{x}, \sigma_0) + \delta \sigma \frac{\partial}{\partial \sigma} B(\vec{x}, \sigma)|_{\sigma = \sigma_0} + \mathcal{O}(\delta \sigma^2) \tag{6.28}
\]

Now consider the differential beam width \( B(\vec{x}, \sigma_A) - B(\vec{x}, \sigma_B) \), or equivalently and perhaps more instructively \( B(\vec{x}, \sigma_0 + \delta \sigma) - B(\vec{x}, \sigma_0 - \delta \sigma) \). The even order terms cancel, leaving us with

\[
B(\vec{x}, \sigma_A) - B(\vec{x}, \sigma_B) = 2 \delta \sigma \frac{\partial}{\partial \sigma} B(\vec{x}, \sigma)|_{\sigma = \sigma_0} + \mathcal{O}(\delta \sigma^3). \tag{6.29}
\]

This tells us the polarization leakage in terms of derivatives of the beam width \( \sigma \); we want it in physical focal plane coordinates. To calculate this, we go back to the definition of the beam width in eq. 6.24

\[
\frac{\partial B}{\partial \sigma} = \frac{x^2 + y^2 - 2 \sigma^2}{2 \pi \sigma^5} \exp\left( -\frac{x^2 + y^2}{2 \sigma^2} \right) \tag{6.30}
\]

This can be related to the second derivative of the beam with respect to \( x \) and \( y \).

\[
\frac{\partial^2 B}{\partial x^2} = \frac{x^2 - \sigma^2}{2 \pi \sigma^6} \exp\left( -\frac{x^2 + y^2}{2 \sigma^2} \right) \tag{6.31}
\]

\[
\frac{\partial^2 B}{\partial y^2} = \frac{y^2 - \sigma^2}{2 \pi \sigma^6} \exp\left( -\frac{x^2 + y^2}{2 \sigma^2} \right) \tag{6.32}
\]

Comparing eq. 6.30 with eqs. 6.31 and 6.32 we find

\[
\frac{1}{\sigma} \frac{\partial B}{\partial \sigma} = \frac{\partial^2 B}{\partial x^2} + \frac{\partial^2 B}{\partial y^2}. \tag{6.33}
\]

Now we have the effect of beam asymmetries as a function of derivatives in the \( x \) and \( y \) basis. However the \texttt{healpy} gives the derivatives in spatial coordinates. As before, we must relate the derivatives in focal plan coordinate with those on the sky. Note that eqs. 6.15 and 6.16 are the derivatives of the map in the \( \hat{x} \) and \( \hat{y} \) directions respectively. So to get the second derivative in sky
6.2.4 Differential Ellipticity

coordinates, we just apply the equations to themselves

\[
\frac{\partial^2}{\partial x^2} = \left( -\cos \gamma \frac{\partial}{\sin \theta \partial \phi} + \sin \gamma \frac{\partial}{\partial \theta} \right) \cdot \left( -\cos \gamma \frac{\partial}{\sin \theta \partial \phi} + \sin \gamma \frac{\partial}{\partial \theta} \right) = -\cos \gamma \frac{\partial}{\sin \theta \partial \phi} \left( -\cos \gamma \frac{\partial}{\sin \theta \partial \phi} + \sin \gamma \frac{\partial}{\partial \theta} \right) + \sin \gamma \frac{\partial}{\partial \theta} \left( -\cos \gamma \frac{\partial}{\sin \theta \partial \phi} + \sin \gamma \frac{\partial}{\partial \theta} \right) \\
= \frac{\cos^2 \gamma}{\sin^2 \theta \partial \phi^2} + \sin^2 \gamma \frac{\partial^2}{\partial \theta^2} + \cos \gamma \sin \gamma \left( \frac{\cos \theta}{\sin^2 \theta \partial \phi} - \frac{2}{\sin \theta \partial \phi \partial \theta} \right) \tag{6.34}
\]

The 2 in the last term comes from the commutativity of partial derivatives. Similarly for \( y \)

\[
\frac{\partial^2}{\partial y^2} = \left( -\sin \gamma \frac{\partial}{\sin \theta \partial \phi} - \cos \gamma \frac{\partial}{\partial \theta} \right) \cdot \left( -\sin \gamma \frac{\partial}{\sin \theta \partial \phi} - \cos \gamma \frac{\partial}{\partial \theta} \right) = -\sin \gamma \frac{\partial}{\sin \theta \partial \phi} \left( -\sin \gamma \frac{\partial}{\sin \theta \partial \phi} - \cos \gamma \frac{\partial}{\partial \theta} \right) - \cos \gamma \frac{\partial}{\partial \theta} \left( -\sin \gamma \frac{\partial}{\sin \theta \partial \phi} - \cos \gamma \frac{\partial}{\partial \theta} \right) \\
= \frac{\sin^2 \gamma}{\sin^2 \theta \partial \phi^2} + \cos^2 \gamma \frac{\partial^2}{\partial \theta^2} + \cos \gamma \sin \gamma \left( -\frac{\cos \theta}{\sin^2 \theta \partial \phi} + \frac{2}{\sin \theta \partial \phi \partial \theta} \right) \tag{6.35}
\]

We finally arrive at

\[
\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} = \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{\partial^2}{\partial \theta^2} \tag{6.36}
\]

giving us the template

\[
T_\sigma(t) = g \left( \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{\partial^2}{\partial \theta^2} \right) (B \cdot \Theta) (\tilde{p}(t)) \bigg|_{p=p_0(t)} \tag{6.37}
\]

and the corresponding fit parameter

\[
\alpha_\sigma = \sigma_0 \delta \sigma. \tag{6.38}
\]

6.2.4 Differential Ellipticity

So far we’ve assumed that the detector beams are well described by circular Gaussians. This analysis can be extended to detectors with elliptical Gaussians. If a pair of detectors have different ellipticities, the difference timeline will introduce a polarization signal. The differential beam ellipticity can be fully described as the sum of the differential ellipticity in the \( x \) and \( y \) axis, and

*Note that this solves for the standard deviation, \( \sigma \). Often beams are defined by full width at half max. FWHM = \( 2\sqrt{2 \ln 2} \sigma \).*
6.2.4 Differential Ellipticity

its cross direction, rotated 45 degrees from the $x$ and $y$ direction. We call these “plus” and “cross”
ellipticities respectively. First consider the differential ellipticity aligned with $x$ and $y$.

$$B(\bar{x}, p) = \frac{1}{2\pi\sigma^2(1 - p^2)} \exp \left[ -\frac{1}{2} \left( \frac{x^2}{2\sigma^2(1 + p)} + \frac{y^2}{2\sigma^2(1 - p)} \right) \right] \quad (6.39)$$

We can again Taylor expand the beam and take difference between $A$ and $B$ detectors, finding

$$B(\bar{x}, p_A) - B(\bar{x}, p_B) = 2\delta p \frac{\partial}{\partial p} B(\bar{x}, p) + O(p^3). \quad (6.40)$$

Just as we did in Section 6.2.3 with differential beam width, we can express derivatives of $p$ in
terms of derivatives of $x$ and $y$. In fact, it’s identical except for a minus sign from the $1 - p$ in the
exponent.

$$\left( \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} \right) B(\bar{x}, p = 0) = \frac{1}{\sigma^2} \frac{\partial}{\partial p} B(\bar{x}, p) \bigg|_{p=0} \quad (6.41)$$

We have already solved for $\partial^2/\partial x^2$ and $\partial^2/\partial y^2$ in eqs. 6.34 and 6.35. We find

$$\left( \frac{\partial^2}{\partial x^2} - \frac{\partial^2}{\partial y^2} \right) = \left( \cos^2 \gamma - \sin^2 \gamma \right) \left( \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{\partial^2}{\partial \theta^2} \right)$$

$$+ 2 \cos \gamma \sin \gamma \left( \frac{\cos \theta}{\sin^2 \theta} \frac{\partial}{\partial \phi} - \frac{2}{\sin \theta} \frac{\partial^2}{\partial \phi \partial \theta} \right)$$

$$= \cos 2\gamma \left( \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{\partial^2}{\partial \theta^2} \right) + \sin 2\gamma \left( \frac{\cos \theta}{\sin^2 \theta} \frac{\partial}{\partial \phi} - \frac{2}{\sin \theta} \frac{\partial^2}{\partial \phi \partial \theta} \right) \quad (6.42)$$

where we’ve used the double angle formulas $\sin 2\gamma = 2 \cos \gamma \sin \gamma$ and $\cos 2\gamma = \cos^2 \gamma - \sin^2 \gamma$. To
get a template timeline to fit the data, we act these derivatives on the beam smoothed temperature
map

$$T_p(t) = g \left[ \cos 2\gamma \left( \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{\partial^2}{\partial \theta^2} \right) \right. \left.+ \sin 2\gamma \left( \frac{\cos \theta}{\sin^2 \theta} \frac{\partial}{\partial \phi} - \frac{2}{\sin \theta} \frac{\partial^2}{\partial \phi \partial \theta} \right) \right] \left( \mathbf{B} \cdot \mathbf{\Theta} \right) \left( \mathbf{\bar{p}}(t) \right) \bigg|_{\mathbf{p} = \mathbf{\bar{p}}(t)} \quad (6.43)$$
6.2.4 Differential Ellipticity

with a fit parameter

\[ \alpha_p = \sigma_0^2 \delta p \]  

(6.44)

A crosspolar beam is described as

\[ B(\bar{x}, p) = \frac{1}{2\pi\sigma^2 (1-c^2)} \exp \left[ - \frac{1}{2} \frac{x^2 - 2cxy + y^2}{\sigma^2 (1-c^2)} \right]. \]  

(6.45)

Taylor expanding and differencing like before, we find

\[ B(\bar{x}, c_A) - B(\bar{x}, c_B) = 2\delta p \frac{\partial}{\partial c} B(\bar{x}, c) + \mathcal{O}(c^3) \]  

(6.46)

which we can rewrite as

\[ \left( \frac{\partial^2}{\partial x \partial y} + \frac{\partial^2}{\partial y \partial x} \right) B(\bar{x}, c) = -\frac{2}{\sigma^2} \frac{\partial}{\partial c} B(\bar{x}, c). \]  

(6.47)

Obviously \( \partial^2/\partial x \partial y = \partial^2/\partial y \partial x \), but I leave it here because it’s the convention in Randol’s thesis and he makes an error there. I will be more explicit here.

\[
\frac{\partial}{\partial x} \left( \frac{\partial}{\partial y} \right) = \left( -\frac{\cos \gamma}{\sin \theta} \frac{\partial}{\partial \phi} + \frac{\sin \gamma}{\sin \theta} \frac{\partial}{\partial \theta} \right) \left( -\frac{\sin \gamma}{\sin \theta} \frac{\partial}{\partial \phi} - \cos \gamma \frac{\partial}{\partial \theta} \right)
\]

\[
= \frac{\cos \gamma \sin \gamma}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{\cos^2 \gamma}{\sin \theta} \frac{\partial^2}{\partial \phi \partial \theta} + \frac{\sin^2 \gamma \cos \theta}{\sin^2 \theta} \frac{\partial}{\partial \phi} + \frac{\sin^2 \gamma \cos \theta}{\sin^2 \theta} \frac{\partial}{\partial \phi} - \frac{\sin^2 \gamma}{\sin \theta} \frac{\partial^2}{\partial \theta^2} - \frac{\sin \gamma \cos \gamma}{\sin^2 \theta} \frac{\partial^2}{\partial \phi \partial \theta}
\]

\[
= \frac{1}{2} \sin 2\gamma \left( \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} - \frac{\partial^2}{\partial \phi \partial \theta} \right) + \cos 2\gamma \frac{\partial^2}{\sin \theta \partial \phi \partial \theta} + \frac{\sin^2 \gamma \cos \theta}{\sin^2 \theta} \frac{\partial}{\partial \phi}
\]
6.2.5 Template Fitting

\[
\frac{\partial}{\partial y} \left( \frac{\partial}{\partial x} \right) = \left( -\sin \gamma \frac{\partial}{\sin \theta \partial \phi} - \cos \gamma \frac{\partial}{\partial \theta} \right) \left( -\cos \gamma \frac{\partial}{\sin \theta \partial \phi} + \sin \gamma \frac{\partial}{\partial \theta} \right) \\
= \cos \gamma \sin \gamma \frac{\partial^2}{\sin^2 \theta \partial \phi^2} - \sin^2 \gamma \frac{\partial^2}{\sin \theta \partial \phi \partial \theta} - \cos^2 \gamma \cos \theta \frac{\partial^2}{\sin^2 \theta \partial \phi} \\
+ \cos^2 \gamma \frac{\partial^2}{\sin \theta \partial \phi \partial \theta} - \cos \gamma \sin \gamma \frac{\partial^2}{\partial \phi \partial \theta} \\
= \frac{1}{2} \sin 2\gamma \left( \frac{1}{\sin^2 \theta \partial \phi^2} - \frac{\partial^2}{\partial \theta^2} \right) + \cos 2\gamma \left( -\frac{\cos \theta \partial}{\sin^2 \theta \partial \phi} + \frac{2}{\sin \theta \partial \phi \partial \theta} \right) 
\]

So then the sum is

\[
\left( \frac{\partial^2}{\partial x \partial y} + \frac{\partial^2}{\partial y \partial x} \right) = \sin 2\gamma \left( \frac{1}{\sin^2 \theta \partial \phi^2} - \frac{\partial^2}{\partial \theta^2} \right) + \cos 2\gamma \left( -\frac{\cos \theta \partial}{\sin^2 \theta \partial \phi} + \frac{2}{\sin \theta \partial \phi \partial \theta} \right) 
\]

(6.48)

and we get the template

\[
T_c(t) = g \left[ \sin 2\gamma \left( \frac{1}{\sin^2 \theta \partial \phi^2} - \frac{\partial^2}{\partial \theta^2} \right) \\
+ \cos 2\gamma \left( -\frac{\cos \theta \partial}{\sin^2 \theta \partial \phi} + \frac{2}{\sin \theta \partial \phi \partial \theta} \right) \right] (B \cdot \Theta) \left( \tilde{p}(t) \right) \big|_{\tilde{p} = \tilde{p}_0(t)} 
\]

(6.49)

with the fit parameter

\[
\alpha_c = \sigma_0^2 \delta_c 
\]

(6.50)

6.2.5 Template Fitting

Fitting the data is a simple linear least squares problem, analogous to a naive mapmaker. If we define \( T \) as a matrix with rows populated by the templates outlined above and \( N^{-1} \) as the weight matrix, and \( d \) as the detector data vector, then the best fit parameters are given by

\[
\hat{\alpha} = \left( T^T N^{-1} T \right)^{-1} T^T N^{-1} d. 
\]

(6.51)

\(^{1}\) Note the factor of \( 1/\sin^2 \theta \) in the \( \partial/\partial \phi \) term. The derivative map output by \texttt{healpy} only has 1 factor of \( 1/\sin \theta \) in it, so we need to add an extra factor when constructing our template timelines.
### 6.3 Deprojection Results

<table>
<thead>
<tr>
<th>Template</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\delta g$</td>
<td>$T_g(t) = g \Theta$</td>
</tr>
<tr>
<td>$\delta x$</td>
<td>$T_x(t) = g \sigma \left( -\cos \gamma \frac{\partial}{\partial \phi} + \sin \gamma \frac{\partial}{\partial \theta} \right) \cdot \Theta$</td>
</tr>
<tr>
<td>$\delta y$</td>
<td>$T_y(t) = g \sigma \left( -\sin \gamma \frac{\partial}{\partial \phi} - \cos \gamma \frac{\partial}{\partial \theta} \right) \cdot \Theta$</td>
</tr>
<tr>
<td>$\delta \sigma$</td>
<td>$T_\sigma(t) = g \left( \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{\partial^2}{\partial \theta^2} \right) \cdot \Theta$</td>
</tr>
<tr>
<td>$\delta p$</td>
<td>$T_p(t) = \cos^2 (\gamma) \left( \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{\partial^2}{\partial \theta^2} \right) + \sin 2 \gamma \left( \frac{\cos \theta \frac{\partial}{\partial \phi} - \frac{2}{\sin \theta} \frac{\partial^2}{\partial \phi \partial \theta}} \right) \cdot \Theta$</td>
</tr>
<tr>
<td>$\delta c$</td>
<td>$T_c(t) = \sin^2 (\gamma) \left( \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} - \frac{\partial^2}{\partial \theta^2} \right) + \cos 2 \gamma \left( \frac{\cos \theta \frac{\partial}{\partial \phi} + \frac{2}{\sin \theta} \frac{\partial^2}{\partial \phi \partial \theta}} \right) \cdot \Theta$</td>
</tr>
</tbody>
</table>

Table 6.2: The 6 deprojection templates. The beam parameter and pointing vector are omitted in this notation.

We can then calculate the covariance matrix, $M^\alpha$, where the diagonal elements contain the variance in the parameter estimates. First, calculate the objective function of the best fit parameters

$$S(\hat{\alpha}) = \sum_{i=1}^{m} \left| d_i - \sum_{j=1}^{n} T_{ij} \hat{\alpha}_j \right|^2 \tag{6.52}$$

where the sum on $i$ is over samples and the sum on $j$ is over templates. Then

$$M^\alpha = \frac{S(\hat{\alpha})}{m-n} \left( T^T T \right)^{-1} \tag{6.53}$$

### 6.3 Deprojection Results

As mentioned before, deprojection is used as a calibration tool and a check on systematics in the SPIDER analysis pipeline. Instead of differencing A/B detectors, SPIDER differences detectors with a reobserved Planck timeline which has been beam smoothed by the best FPU averaged beam as described in Section 6.1. The initial boresite pointing comes from the dedicated star cameras, and each detector’s boresite offset comes from the designed plate scale. Both detector and boresite
pointings are improved by deprojection.

6.3.1 Detector Pointing

The detector pointing is a function of two things, the boresite pointing which is time variant and the detector boresite offset which is constant. The boresite pointing describes SPIDER’s scan strategy, while the boresite offset describes the intrinsic optics of the payload. In this section, deprojection is used to constrain the boresite offset of each detector.

Each detector pair is nominally co-pointed, and each tile is designed to project a square pattern of detector beams onto the sky. Four tiles are assembled into an array, also nominally designed to create a 2x2 square pattern of tiles onto the sky. However, small lithographic impurities can slightly alter the pointing of individual detectors on a tile. And small assembly uncertainties can misalign tiles in relation to each other. Thus it is important to determine the detector boresite offsets empirically rather than relying on designed values.

An example of the detector boresite offset solution is shown in Figure 6.3. We can see that the top two tiles are shifted, implying a small offset during assembly. The bottom right tile is very close to the design value. The bottom left tile has a rotation, probably due to assembly as well. In fact, we can express the offsets in physical units of the tile by noting that a 1 degree offset it on sky pointing corresponds to a shift of about 0.65” of a tile at 95 GHz. The typical offset is less than 0.1 degrees, so the physical offset of the tile is less than 0.065”. These assembly offsets are not a problem, as long as we measure them accurately.

6.3.2 Boresite Pointing

SPIDER solves for its boresite pointing using 3 star cameras. Two of the cameras sit on the deck and are called the rotating star cameras (RSC). The other camera is mounted directly to the cryostat and is called the boresite star camera (BSC). These cameras image the sky and reconstruct the boresite pointing by comparing the pictures with star catalogs. See Jamil Shariff and Natalie Gandilo’s theses for more details [59, 77].
6.3.2 Boresite Pointing

Figure 6.3: The change in detector pointing relative to the designed plate scale as solved for by deprojection scaled by 10. Note that A/B detectors are all shown, most of which largely overlap. The large outliers are thrown out.

One disadvantage of this system is that the star cameras only act as a proxy for the true pointing of the detectors. Deprojection allows SPIDER to measure its boresite pointing with the detectors themselves. Where in Section 6.3.1 the deprojection fit parameters are solved for a single detector for the full length of the flight, all the detectors are averaged over a 10 minute chunk. This generates accurate boresite pointing per 10 minute chunk; in fact the uncertainty is lower than the dedicated pointing sensors in elevation.

Figure 6.4 shows the difference between the boresite pointing solution from star cameras and deprojection. Boresite elevation changes cause minor flexes in the outer frame, which coupled with
6.3.3 Beam Parameters

Figure 6.4: The difference in boresite pointing from the dedicated pointing system versus the pointing as resolved by the detectors using deprojection. The diurnal variations are due to the scan strategy. The rolloff in the later days is due to the cryostat running out of liquid helium, changing the center of mass and generating a thermal gradient across the main tank.

the daily up-down scan strategy create the diurnal variation in Figure 6.4. Similarly, the slow evaporation of liquid helium shifts the center of mass towards the back of the cryostat and additionally creates thermal gradients across the tank, causing the flight-long drift which is most apparent towards the end of flight. This boresite correction has been incorporated into the mapmaking pipeline.

A new pointing solution is generated using the data in Figure 6.4. Figure 6.5 shows the difference of SPIDER temperature maps made with the old (incorrect) and new (fixed) pointing solution at 150 GHz. Note that the features are greatest at the top and bottom of the map, precisely because that is where the flex is the greatest.

6.3.3 Beam Parameters

The beam width and ellipticity parameters are invariant through flight, so the fit parameters are averaged per detector for all the data. There is measurable variation in both beam width and ellipticity in a focal plane. The pattern is radially symmetric and is in agreement with physical
6.3.3 Beam Parameters

Figure 6.5: The difference between the new and old SPIDER T maps at 150 GHz. The map is smoothed with a half degree beam. The difference is largest at high and low elevations because that is where the correction is the greatest.

optics simulations. Figures 6.6 and 6.7 show the beam width and ellipticity respectively for X1, where ellipticity is defined in Appendix D. All the FPU beam parameters are shown in Appendix H. Figure 6.8 shows the ellipticity in a more physical representation. Note that the elliptical shape represents the shape of the 1-sigma contour of the beam, but the size is much smaller than the actual beam.

Knowing the beam parameters is the first half of the problem. The more interesting question is whether one should care about the beam differences. The SPIDER mapmaker naively bins the detector data with the assumption that all the beams in a focal plane are the same circular Gaussian. Figures 6.6 and 6.7 show that this assumption is measurably incorrect.

Simulation is the most obvious way to quantify the effect of the variable beam parameters on science results. However, smoothing maps is a slow operation. Moreover the ellipticities of the beams are different per detector and project onto the sky differently depending on the scan strategy and sky rotation. In other words, to fully simulate beam ellipticities by reobserving raw smoothed
Figure 6.6: The beam width as solved by deprojection of the X1 in physical coordinates. The radial pattern is consistent with physical optics simulations.
Figure 6.7: The ellipticity as solved by deprojection of the X1 in physical coordinates. The radial pattern and beam shapes are consistent with physical optics simulations.
Figure 6.8: A physical representation of the beam ellipticities. Each ellipse is centered on that detector’s solved beam center. The ellipse shape represents the actual shape of the beam. The size (radius) of the ellipse is much smaller than the actual beam size to aid in visual clarity.
6.3.3 Beam Parameters

Figure 6.9: The BB spectra due to beam non-idealities. The left is for 95 GHz and the right is for 150 GHz. Blue points are from the difference in beam width; orange points are from differences in ellipticity.

Planck maps would require having unique maps per detector per 10 minutes\(^1\). Deprojection offers a much more efficient method of simulation. Detector timestreams of differential beam widths and ellipticities can be constructed using the last 3 equations in Table 6.2. This means a full simulation of every detector for the full flight only requires holding the 4 maps in memory. Note that a full polarization map effectively has 3 maps in it (I, Q, and U), so this method is only slightly more memory intensive than reobserving a polarized map. These 4 maps can be presmoothed with a nominal beam and stored to disk.

To measure the effect of beam non-idealities in SPIDER, detector timestreams were generated using the solutions from deprojection. These detector timestreams were binned into maps with the nominal detector pointing. The spectra of the maps were calculated. The BB spectra is shown in Figure 6.9 for 95 GHz and 150 GHz. The amplitudes of the spectra are very small relative to SPIDER’s instrumental noise. Note that the shape of the spectra are similar between beam width and ellipticity parameters. This is because the maps they are sampled from are effectively the same

\(^1\) The 10 minute number is an approximate value given SPIDER’s scan strategy and physical location. The exact number depends on the accuracy desired and the amount of sky rotation.
6.4 MCE and Pointing Solution Timing

The detector data read out via the MCE is synced to the pointing data using a sync box provided by the UBC experimental cosmology group. The sync box provides a 25 MHz clock, much faster than the sample rate of either the MCE or the pointing solution. This system guarantees that the data stays locked relative to each other, up to a phase shift. The phase shift comes from the uncertainty of time delay between acquiring data by a subsystem and writing that data to disk.

With the initial pointing solution, Spider failed the left-right null test. The left-right null test is generated by making a map with all left going scans, then making a map with all right going scans, and then differencing the two:

\[ m_{\text{null}} = m_{\text{left}} - m_{\text{left}}. \] (6.54)

The expectation is that \( m_{\text{null}} \) is simply white noise. The center panel of Figure 6.10 shows the null map using the original pointing solution, which is clearly not null.

The key to solving this null test failure was to note that the null maps look very similar to spatial derivative maps generated for deprojection which are shown in Figure C.1. Since the scan strategy of Spider is dominated by left-right scans (as opposed to vertical scans), a left-right null test with phase delay in the pointing relative to detector data is identical to taking a spatial derivative in the left-right direction. To test this theory, I generated maps that phase shifted the pointing solution forwards and backwards by 1 sample relative to the detector data. One sample is approximately 8 ms. These shifted maps are shown on the top and bottom panel of Figure 6.10. The maps are smoothed with a half degree beam to highlight degree scale features.

It is obvious that the top panel has much less power than the middle or bottom panel. In fact the bottom panel is even worse than the original, which is consistent with the timing being even worse. The phase offset between the detector and pointing data is not required to be exactly 1 sample.
Figure 6.10: The temperature map of the left-right null tests. The maps are all smoothed with a half degree beam. The middle panel is generated using the original pointing solution. The top and bottom panels are generated by shifting the pointing solution backwards and forwards by sample respectively. One sample is approximately 8 ms.
6.4 MCE and Pointing Solution Timing

The offset is due to a combination of two effects. This first effect is the unknown timing between acquiring and writing data, which is set by complex electronics in the system. The second effect is the time lag in the detector data due to the MCE filtering. We can instead solve for this offset by interpolating the pointing solution to subsample the offset to values. With interpolated pointing solutions, we generate new left-right null maps and calculate the standard deviation of values in the temperature map. The residuals as a function of offset are shown in Figure 6.11. The offsets are solved per focal plane per quarter flight. A quadratic is fit to the points, and the minima are noted with a star and are written on the legend.

The solutions per FPU are very consistent. The solutions across FPUs are close, but not identical. We first attempt a correction with an instrument wide offset between detectors and pointing that is 0.787 samples. This is run through the left-right null test and there is no measurable non-null signal. Thus an instrument wide offset was deemed sufficient for SPIDER.
There are solutions per FPU per quarter flight map, indicated by the "N of 4". The stars indicate the minimum of the quadratic fit to the data. The minimum is also noted on the right of the legend.
Chapter 7

Mapmaking

Before we would go on road trips, my dad would always drag me to AAA to pick up maps. We would bring them home, lay them out on the dinner table, and chart our entire trip – every highway, every exit, and every stop along the way. I would sit in the back of the car as we drove towards Los Angeles or the Grand Canyon or the Hoover Dam, tracking our progress.

7.1 Mapmaker Basics

With thousands of polarized detector timestreams and a pointing solution in hand, the task is now to generate a map of the sky. The unpolarized case is conceptually easy to understand; simply bin the detector amplitudes for every pixel on the sky. The polarized case is largely the same, but the mapmaker must also have knowledge of the detector polarization angle. The description of mapmaker basics largely flows from Sasha Rahlin and Steve Benton’s theses\cite{61,78}. 

124
7.1.1 Stokes Parameters

Figure 7.1: The Stokes parameters. For all of these plots, \( V = 0 \). For the non-zero \( Q \), \( V = 0 \). For the non-zero \( U \), \( Q = 0 \).

7.1.1 Stokes Parameters

For polarized mapmaking, start with the description of a plane wave

\[
E(x, t) = (\epsilon_1 A_1 + \epsilon_2 A_2) \exp(i k \cdot x - i \omega t) \tag{7.1}
\]

from Jackson\cite{79}, where \( k \) is the direction of propagation, \( \epsilon_1 \) and \( \epsilon_2 \) are orthogonal basis vectors, and \( A_1 = E_1 e^{i \delta_1} \) and \( A_2 = E_2 e^{i \delta_2} \) are complex amplitudes.\footnote{The pair \((A_1, A_2)\) is sometimes called the Jones vectors.} The terms \( \delta_1 \) and \( \delta_2 \) encode the polarization state of the plane wave. If \( \delta_1 = \delta_2 \), the wave is linearly polarized. If \( \delta_1 - \delta_2 \), the wave is circularly polarized.

We want to construct scalar quantities that are functions only of the wave intensity. To do this, we use a 4-parameter model called the *Stokes parameters*. The parameters are defined as

\[
I = |\epsilon_1 \cdot E|^2 + |\epsilon_2 \cdot E|^2 = E_1^2 + E_2^2 \tag{7.2}
\]

\[
Q = |\epsilon_1 \cdot E|^2 - |\epsilon_2 \cdot E|^2 = E_1^2 - E_2^2 \tag{7.3}
\]

\[
U = 2 \text{Re}[(\epsilon_1 \cdot E)(\epsilon_2 \cdot E)] = 2 E_1 E_2 \cos(\delta_2 - \delta_1) \tag{7.4}
\]

\[
V = 2 \text{Im}[(\epsilon_1 \cdot E)(\epsilon_2 \cdot E)] = 2 E_1 E_2 \sin(\delta_2 - \delta_1) \tag{7.5}
\]

These values are convenient because they are simple sums and differences of measurable intensities for quasi-monochromatic light. One property of quasi-monochromatic light is the complex amplitudes vary slowly relative to the frequency. This allows us to take the time-averaged measurement of the Stokes parameters, which we will denote by \( \langle \rangle \).
To gain some physical intuition for what the Stokes parameters are, consider 3 different coordinate axes. The first and most trivial one is $(x, y)$ which is aligned with $\epsilon_1$ and $\epsilon_2$. The second is a simple $45^\circ$ rotation of $(x, y)$, which we will denote $(a, b)$. The third coordinate axis includes a $\pm 90^\circ$ phase shift, which we will denote $(+, -)$. This can be expressed as

\[
\begin{align*}
\epsilon_x &= \epsilon_1 \quad \epsilon_y = \epsilon_2 \\
\epsilon_a &= \frac{1}{\sqrt{2}}(\epsilon_1 + \epsilon_2) \quad \epsilon_b = \frac{1}{\sqrt{2}}(\epsilon_1 - \epsilon_2) \\
\epsilon_+ &= \frac{1}{\sqrt{2}}(\epsilon_1 + i\epsilon_2) \quad \epsilon_- = \frac{1}{\sqrt{2}}(\epsilon_1 - i\epsilon_2).
\end{align*}
\]

Now using the time averaged parameters, we find

\[
\begin{align*}
I &= \langle E_1 \rangle^2 + \langle E_2 \rangle^2 = \langle E_a \rangle^2 + \langle E_b \rangle^2 = \langle E_+ \rangle^2 + \langle E_- \rangle^2 \\
Q &= \langle E_1 \rangle^2 - \langle E_2 \rangle^2 \\
U &= \langle E_a \rangle^2 - \langle E_b \rangle^2 \\
V &= \langle E_+ \rangle^2 - \langle E_- \rangle^2.
\end{align*}
\]

The $I$ parameter is the total intensity. $Q$, $U$, and $V$ are the differences between the two linear polarizations, rotated linear polarizations, and circular polarizations respectively. Figure 7.1 shows a schematic of $Q$ and $U$ parameters.

### 7.1.2 Polarimeter Model

We can write the Stokes parameters as a Stokes vector

\[
s = \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix}.
\]
7.1.2 Polarimeter Model

Any optical element can be represented by a Mueller matrix, $M$, which acts on $s$. For example, we can rotate $s$ using a Mueller matrix for rotation

$$M_{\text{rot}}(\theta) \cdot s = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\theta & \sin 2\theta & 0 \\ 0 & -\sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = \begin{bmatrix} I \\ Q \cos 2\theta + U \sin 2\theta \\ -Q \sin 2\theta Q + U \cos 2\theta \\ V \end{bmatrix}.$$  \(7.14\)

First note that $I$ and $V$ do not change under rotation. Linearly polarization rotates by $2\theta$ because polarization is spin-2.

Before constructing the basic bolometer model, we will introduce one more Mueller matrix. For a linear polarizer for horizontal transmission,

$$M_{\text{pol}} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$  \(7.15\)

To see this, act $M_{\text{pol}}$ on $s$

$$M_{\text{pol}} \cdot s = \frac{1}{2} \begin{bmatrix} I + Q \\ I + Q \\ 0 \\ 0 \end{bmatrix}.$$  \(7.16\)

where only the $I$ and $Q$ terms pass through. It is worth explicitly stating the fact that the linear polarizer eliminates $V$. If Spider were an ideal instrument, it would have no sensitivity on circular polarization.

Now we are armed with all the tools we need to describe a polarimeter. Considering elements in the time reverse sense, the bolometer is simply an intensity measurement device, so it can be
modeled as

\[
a = \begin{bmatrix}
1 \\
0 \\
0 \\
0 \\
\end{bmatrix}.
\] (7.17)

The instrument has multiple polarization angles, so we need to rotate the horizontal polarizer into arbitrary angles, \( \phi \). To do this, note that any arbitrary Mueller matrix can be rotated by \( M' = M_{\text{rot}}(\theta)MM_{\text{rot}}(-\theta) \). So we construct the rotated polarizer matrix

\[
M_{\text{pol}}(\phi) = \frac{1}{2} \begin{bmatrix}
1 & \cos 2\phi & \sin 2\phi & 0 \\
\cos 2\phi & \cos^2 2\phi & \sin 2\phi \cos 2\phi & 0 \\
\sin 2\phi & \sin 2\phi \cos 2\phi & \sin^2 2\phi & 0 \\
0 & 0 & 0 & 0 \\
\end{bmatrix}.
\] (7.18)

So for an arbitrary input of energy described by \( s \), the signal on the bolometer is

\[
d = a^T M_{\text{pol}}(\phi)s = I + Q \cos 2\phi + U \sin 2\phi.
\] (7.19)

### 7.1.3 Half-wave plate

SPIDER modulates incident polarization using a half-wave plate (HWP)\[80–82\]. The SPIDER half-wave plates are made from sapphire, a birefringent crystal. The Mueller matrix for a half-wave plate with the fast axis aligned vertically is

\[
M_{\text{pol}} = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & -1 & 0 \\
0 & 0 & 0 & -1 \\
\end{bmatrix}.
\] (7.20)
which can be rotated to an arbitrary angle

\[
M_{\text{HWP}}(\psi) = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & \cos 4\psi & \sin 4\psi & 0 \\
0 & \sin 4\psi & -\cos 4\psi & 0 \\
0 & 0 & 0 & 1 \
\end{bmatrix}.
\] (7.21)

Now we can rewrite eq. 7.19 accounting for the half-wave plate matrix as

\[
d = a^T M(\phi) M_{\text{HWP}} s = I + \cos(2\phi - 4\psi) + U \sin(2\phi - 4\psi)
\] (7.22)

where we have used the trigonometric sum and difference identities to simplify the expression. To further simplify the expression, we define the total polarization angle as the sum of grid and HWP angles

\[
\phi = \phi - 2\psi.
\] (7.23)

### 7.1.4 Linear Mapmaking

Assuming a noiseless detector, eq. 7.22 gives the detector amplitude for any input sky signal. Unfortunately, SPIDER is not noiseless, so the full detector signal is

\[
d = I + Q \cos(2\varphi) + U \sin(2\varphi) + n.
\] (7.24)

This is for one measurement by one pixel. Of course we have multiple detectors with multiple samples and multiple pixels on the sky. Let us index the \(N_{\text{det}}\) detectors with \(i\), the \(N_{\text{samp}}\) samples for a given detector with \(t\), and the \(N_{\text{pix}}\) pixels with \(p\), giving

\[
d_{it} = I_p + Q_p \cos(2\varphi_{it}) + U_p \sin(2\varphi_{it}) - n_{it}.
\] (7.25)
7.1.4 Linear Mapmaking

Clearly the raw \( I, Q, \) and \( U \) from the sky are unchanged across observations, so they are only indexed by pixel. The observations of the polarization terms do change with time due to sky rotation and half-wave plate rotation, giving the \( \sin \) and \( \cos \) terms with detector and time indexes. The noise is assumed to be uncorrelated with sky power, and therefore has no pixel dependence but does have a detector and time index.

Equation 7.25 can be rewritten to in matrix notation as

\[
d = Am + n. \tag{7.26}
\]

Both \( d \) and \( n \) are length \( N_{\text{samp}}N_{\text{det}} \). The map vector \( m \) is length \( 3N_{\text{pix}} \). Explicitly written out

\[
d = \begin{bmatrix}
d_{00} \\
d_{01} \\
\vdots \\
d_{0N_{\text{samp}}} \\
d_{10} \\
\vdots \\
d_{N_{\text{det}}N_{\text{samp}}}
\end{bmatrix}, \quad n = \begin{bmatrix}
n_{00} \\
n_{01} \\
\vdots \\
n_{0N_{\text{samp}}} \\
n_{10} \\
\vdots \\
n_{N_{\text{det}}N_{\text{samp}}}
\end{bmatrix}, \quad m = \begin{bmatrix}
I_0 \\
Q_0 \\
U_0 \\
\vdots \\
I_{N_{\text{pix}}} \\
Q_{N_{\text{pix}}} \\
U_{N_{\text{pix}}}
\end{bmatrix}. \tag{7.27}
\]

The pointing matrix \( A \) is a sparse matrix of size \( N_{\text{det}}N_{\text{samp}} \times N_{\text{pix}} \). The only nonzero elements of \( A \) are the prefactors to \( I, Q, \) and \( U \) in eq. 7.25 corresponding to the observed pixel. Explicitly, the rows of \( A \) are

\[
A_{it} = \begin{bmatrix}
0 & \ldots & 1 & \cos(2\varphi_{it}) & \sin(2\varphi_{it}) & \ldots & 0
\end{bmatrix}. \tag{7.28}
\]

The ultimate goal is find the best estimate of the sky signal \( m \), given the raw data \( d \) and pointing \( A \), subject to the instrumental noise \( n \). The most general linear solution \( \hat{m} \) is

\[
\hat{m} = (A^TN^{-1}A)^{-1}A^TN^{-1}d. \tag{7.29}
\]
where the noise covariance matrix in the time domain is

\[ N = \langle \mathbf{n}\mathbf{n}^T \rangle. \]  

(7.30)

Equation 7.29 is unbiased and minimizes the noise covariance in the pixel domain, \((\mathbf{A}^T \mathbf{N}^{-1} \mathbf{A})^{-1}\), also called the \textit{pixel-pixel noise covariance}. Explicitly, this procedure minimizes

\[ \chi^2(\mathbf{m}) = (\mathbf{d} - \mathbf{A}\mathbf{m})^T \mathbf{N}^{-1} (\mathbf{d} - \mathbf{A}\mathbf{m}). \]  

(7.31)

The true nature of \( \mathbf{N} \) is complicated. However, the filtering of \textsc{Spider} data attempts to remove correlations, the off diagonal elements of the noise matrix. The majority of the correlated noise is at low frequencies and is removed by a polynomial filter described in Section 7.2.1. Moreover, data flagging attempts to guarantee that the data is Gaussian. We implemented a skewness check in the flagging routine and found that it rejected less than 0.1% of the data. This allows us to assume in the mapmaking stage that \( \mathbf{N} \) is diagonal with values \( 1/\sigma^2 \) where \( \sigma^2 \) is the noise variance of the detector. The disadvantage of filtering the data is that it makes \( \hat{\mathbf{m}} \) biased. This necessitates simulations to understand the effects of filtering.

Equation 7.29 is often simplified to

\[ \hat{\mathbf{m}} = \mathbf{P}^{-1}\mathbf{v} \]  

(7.32)

where

\[ \mathbf{P} = (\mathbf{A}^T \mathbf{N}^{-1} \mathbf{A}) \]  

(7.33)

and

\[ \mathbf{v} = \mathbf{A}^T \mathbf{N}^{-1} \mathbf{F}\mathbf{d}. \]  

(7.34)

We have introduced the matrix \( \mathbf{F} \) to represent the the time domain filtering applied to the data. This can be any arbitrary filter. Some of the filters used in \textsc{Spider} are discussed in Chapter 5. With the simplifying assumption that the noise is Gaussian and uncorrelated, we can write for a
7.1.5 Angular Coverage and Condition Numbers

single pixel in the map

\[
\begin{bmatrix}
I_p \\
Q_p \\
U_p
\end{bmatrix} = \left( \sum_{it} \frac{1}{\sigma_{it}^2} \begin{bmatrix}
1 & \cos 2\varphi_{it} & \sin 2\varphi_{it} \\
\cos 2\varphi_{it} & \cos^2 2\varphi_{it} & \cos 2\varphi_{it} \sin 2\varphi_{it} \\
\sin 2\varphi_{it} & \cos 2\varphi_{it} \sin 2\varphi_{it} & \sin^2 2\varphi_{it}
\end{bmatrix} \right)^{-1}
\times \left( \sum_{it} \frac{1}{\sigma_{it}^2} \begin{bmatrix}
d_{it}' \\
\cos 2\varphi_{it} d_{it}' \\
\sin 2\varphi_{it} d_{it}'
\end{bmatrix} \right)
\]

(7.35)

where we have defined \(d' = Fd\). Remember that \(i\) indexes detector and \(t\) indexes sample. In words, for a single pixel in the map, take every sample from every detector that falls into the given pixel. Calculate \(\varphi_{ij}\) for every observation, which is a function of sky rotation, HWP rotation, and detector polarization angle. With those values, calculate \(P\) and \(v\). Inverting \(P\) gives \(\mathbf{m}_p\).

This generates a naively binned map. The process requires one iteration over all the data to generate \(P\) and \(v\). Naive mapmaking is a relatively fast process.

7.1.5 Angular Coverage and Condition Numbers

Simply by multiplying both sides of eq. 7.35 with \(P\) and treating \(I_p\), \(Q_p\), and \(U_p\) as actual sky signal rather than our estimate of the sky

\[
\sum_{it} \begin{bmatrix}
1 & \cos 2\varphi_{it} & \sin 2\varphi_{it} \\
\cos 2\varphi_{it} & \cos^2 2\varphi_{it} & \cos 2\varphi_{it} \sin 2\varphi_{it} \\
\sin 2\varphi_{it} & \cos 2\varphi_{it} \sin 2\varphi_{it} & \sin^2 2\varphi_{it}
\end{bmatrix} \begin{bmatrix}
I_p \\
Q_p \\
U_p
\end{bmatrix} = \sum_{it} \begin{bmatrix}
d_{it}' \\
\cos 2\varphi_{it} d_{it}' \\
\sin 2\varphi_{it} d_{it}'
\end{bmatrix}
\]

(7.36)

we can see that \(P\) describes how true signal couples into the timestreams. The top row of \(P\) multiplied with the sky signal returns eq. 7.24 The second and third rows generate the \(Q\) and \(U\) equivalents. The right side of the equation contains all the detector data, and \(P\) describes the angular coverage of the experiment.

To quantify the angular coverage of the experiment, we defined the condition number \(\kappa\) as
the ratio of the largest eigenvalue to the smallest eigenvalue of $P$. The condition number is used to quantify how errors in $v$ couple to errors in $\hat{m}$. Or in terms relevant to us, how do small uncertainties in detector values couple to uncertainties in $I$, $Q$, and $U$? A derivation of $\kappa$ is given in Sasha Rahlin’s thesis[61]. I will instead attempt to motivate the condition number and provide some intuition with examples.

It is obvious from $P$ that our knowledge of $I$, $Q$, and $U$ are codependent. The easiest way to see this is that to reconstruct the $I$ parameter, we must subtract both polarization parameters $Q$ and $U$. So large uncertainties in one parameter couple to others. To minimize this coupling uncertainty, we want equal $Q$ and $U$ coverage. To gain intuition on how condition number depends on $Q$ and $U$ coverage, we will go over some single pixel examples. None of the examples are groundbreaking, but they are instructive. For simplicity, we will divide all the examples by the number of measurements.

**One Polarized Detector**

Assume we have one polarized detector sensitive to one polarization angle with no sky rotation or HWP rotation. For simplicity let’s have $\varphi = 0$. Then $P$ is simply

$$
\begin{bmatrix}
1 & 1 & 0 \\
1 & 1 & 0 \\
0 & 0 & 0
\end{bmatrix}
$$

which has an infinite condition number. It is obvious that this setup provides no measurement on $U$, since $U$ is aligned $45^\circ$ away from $0^\circ$. Additionally there is no measurement on $Q$ because this setup does not measure the horizontal component of the polarization, as seen in Figure 7.1.

**Two Orthogonal Detectors**

Now let’s assume we have 2 orthogonal, polarized detectors. Let’s again assume that there is no sky rotation or HWP rotation. So we have detectors with $\varphi = 0$ and $\varphi = 90^\circ$. This is analogous to
7.1.5 Angular Coverage and Condition Numbers

an instantaneous measurement by a single detector pair in SPIDER. This makes $P$

$$
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 0
\end{bmatrix}
$$

This also has an infinite condition number. Again we have no constraint on $\pm U$. In some sense, we have a measurement of $Q$ since our detectors now measure both the vertical and horizontal component. However we cannot decouple $Q$ from $U$.

One Polarized Detector with HWP rotation

Assume we have one polarized detector sensitive to one polarization angle. We make 2 measurements of the same map pixel, one at $\psi = 0$ and one at $\psi = 22.5^\circ$. This gives a $P$ of

$$
\begin{bmatrix}
1 & .5 & .5 \\
.5 & .5 & 0 \\
.5 & 0 & .5
\end{bmatrix}
$$

This yet again has an infinite condition number, as we can not measure either $Q$ or $U$.

Two Orthogonal Detectors with HWP rotation

Now let’s consider 2 orthogonal, polarized detectors with HWP rotation. For simplicity, let’s assume that each detector makes 2 measurements, one at $\psi = 0$ and $\psi = 22.5^\circ$. This is equivalent to four measurements at $\varphi = [0, 45^\circ, 90^\circ, 135^\circ]$. This produces a diagonal $P$

$$
\begin{bmatrix}
1 & 0 & 0 \\
0 & .5 & 0 \\
0 & 0 & .5
\end{bmatrix}
$$
which has a condition number of 2. By plugging this into equation 7.35 we can see that we have perfect reconstruction of $I$, $Q$, and $U$. Of course this makes sense, as we have measured $Q$ with the $\varphi = 0^\circ, 90^\circ$ and $U$ with $\varphi = 45^\circ, 135^\circ$. And now we can fully disambiguate the components. In fact, a condition number of 2 is the best possible value since it represents equal $Q$ and $U$ coverage.

A low condition number is not sufficient to indicate that a pixel is useful for science. In fact, the contrived example we just came up with is terrible for science since each Stokes polarization parameter is measured exactly once. We also need to take into account a high number of measurements. SPIDER uses a combination of hits count and condition number to select pixels to analyze.

**Two Orthogonal Detectors with HWP and Sky rotation**

Now consider 2 orthogonal, polarized detectors with HWP and sky rotation. The SPIDER scan strategy measured the middle of the map 12 hours apart. This maximized the sky rotation between observations. The largest sky rotation SPIDER saw in 12 hours was about 20 degrees. In between the measurements, the HWPs were rotated 22.5 degrees. We can model this as four measurements, $\varphi = [0, 90^\circ, 45^\circ + 20^\circ, 135^\circ + 20^\circ]$. The third and fourth terms represent orthogonal detectors that have been rotated by the HWP then rotated by the sky. The $P$ for these four observations is

\[
\begin{bmatrix}
1 & 0 & 0 \\
0 & .707 & -.246 \\
0 & -.246 & .292
\end{bmatrix}
\]

which has a condition number of 5.62.

**Condition Number in Spider**

Now armed with some intuition, we can look at SPIDER’s condition number, shown in Figure 7.2. SPIDER has excellent condition number over a large region in the sky. This is primarily due to the HWP rotations, with some additional rotation due to sky coverage. There is no sky rotation at the poles. However, the SPIDER flight stayed away from the South Pole, so we saw about 20° sky
rotation each day. In the analysis, we reject all points with condition number greater than 3. In practice this does not matter because systematics force us to eliminate more data.

7.2 Spider Mapmaker

This section will discuss SPIDER’s mapmaker implementation. The data is assumed to be composed of only sky signal, white noise, and scan synchronous signal. In other words, it is assumed that the data cleaning described in Chapter 5 has removed all spurious signals.

7.2.1 Filtering

As described in Section 3.4.5, the SPIDER data has a large scan synchronous component. The SPIDER collaboration has attempted multiple techniques to remove spurious signal including destriping borrowed from Planck [83], a singular value decomposition solution, and template building. None of these approaches successfully recover a systematics free map. The only demonstrably successful filtering scheme is through simple polynomial filtering of the time domain data. In this scheme, we fit a fifth order polynomial to the azimuth binned detector data per half scan. Other orders were attempted as well. A third order polynomial still had measurable residuals. The choice of binning in azimuth versus time was made because the SPIDER scan strategy is sinusoidal rather than continuous, and the scan synchronous pickup is believed to live in physical space, which is better represented in azimuth.

7.2.2 Transfer Function

Any filtering suppresses signal in addition to noise. In the simplest case, the SPIDER flagging described in Chapter 5 throws away samples. In the $\ell$ domain, this down weights power in the modes represented by that time domain data. Similarly, the filtering described in Section 7.2.1 suppresses modes that are coupled with the polynomial filter and scan strategy.

The filter transfer function, $F_\ell$, is defined as the observed signal divided by the actual sky signal. SPIDER derives its $F_\ell$ by generating ensembles of $\Lambda$CDM maps and running them through
Figure 7.2: The condition number for 95 and 150 on top and bottom respectively. Most pixels are well conditioned, with a value well below 3. The right edges of the maps are poorly conditioned due to a reduction in our observation region during flight.
7.2.3 Spider Maps

![Graph showing filter transfer function and Spider beam.](image)

Figure 7.3: The filter transfer function and the Spider beam. The low transfer function at low $\ell$ is due to the aggressive polynomial filtering. The beam is not normalized to one due to degeneracy with calibration.

The simulation pipelines. This pipeline flags and filters the simulated data identically to the real data. The simulated power is then compared to the known input signal. Note that $F_\ell$ is defined in spectral space, while the beam is typically thought of in spatial or map domain. The beam is re-expressed in spectral coordinates, denoted $B_\ell$. So $B^2_\ell$ is directly comparable to $F_\ell$.

7.2.3 Spider Maps

After flagging and polynomial filtering, the data is ready to be binned into a map. We implement the naively binned map described in Section 7.1.4. Spider 95 and 150 maps are shown in Figure 7.4 and Figure 7.5 respectively. The maps are smoothed with a 20 arcminute beam to highlight the degree scale anisotropies.

7.3 Noise Simulation

To understand the significance of Spider's scientific results, we must understand how the instrumental noise propagates into the maps and spectra. The simplest way to achieve this is to generate noise timestreams and run them through the mapmaker. In this section, we will discuss
Figure 7.4: The SPIDER 95 GHz maps. The map is smoothed with a 20 arcminute beam to highlight degree scale anisotropies.
Figure 7.5: The SPIDER 150 GHz maps. The map is smoothed with a 20 arcminute beam to highlight degree scale anisotropies.
7.3.1 Uncorrelated Noise

the uncorrelated noise simulations in Section 7.3.1 then extend the analysis to correlated noise in Section 7.3.2.

7.3.1 Uncorrelated Noise

To generate uncorrelated noise simulations, we generate timelines based on the power spectral densities (PSD) of the detectors. An example PSD is shown in Figure 5.7. Each detector has its own noise PSD. In principle the noise has time variance. However, the SPIDER simulation pipeline makes the simplifying assumptions that the noise is time invariant and described by a simple Gaussian. These two assumptions allow us to generate noise simulations in the time domain by simply using a pseudo-random number generator with a given variance.

7.3.2 Correlated Random Noise

We want to generate a set of noise timestreams that are consistent with a correlation matrix \( C \). The correlation matrix is constructed by taking the Fourier transform, \( \tilde{d}_n \), of a timeseries, \( d_n \), where the index \( n \) indicates an individual detector. So if there are \( N \) detectors, we construct a vector

\[
Y(f) = [\tilde{d}_0(f), \tilde{d}_1(f), \ldots, \tilde{d}_N(f)]
\]

(7.37)

then the correlation matrix is the outer product of \( Y \) and \( Y^\dagger \)

\[
C(f) = Y^\dagger(f) \otimes Y(f).
\]

(7.38)

† This is also written often as \((Y^\dagger, Y)\), where the parentheses represent the outer product.
This creates a matrix that has diagonal elements that are the power spectral densities (PSD) and off diagonal elements that are cross spectral densities (CSD).

\[
C(f) = \begin{bmatrix}
\mathcal{P}_0(f) & C_{01}(f) & \ldots & C_{0N}(f) \\
C_{10}(f) & \mathcal{P}_1(f) & & \vdots \\
\vdots & \ddots & \ddots & \vdots \\
C_{N0}(f) & \ldots & \ldots & \mathcal{P}_N(f)
\end{bmatrix}
\]

We have actually already generated these matrices in Section 5.3.2.2. Examples of binned correlation matrices are shown in Figures 5.9 and 5.10.

To generate random noise with the properties described by the correlation function, we need to generate random realizations of the correlation function then inverse Fourier transform them. Note that the PSDs and CSDs are effectively the square of the Fourier transforms, so before we inverse Fourier transform them, we need to take the square-root. To do this, we calculate the eigenvectors \( u_i \) and eigenvalues \( \lambda_i \) of the correlation matrix. Then we can represent the correlation matrix as

\[
C = U\Lambda U^\dagger
\]

where \( U \) is a matrix where the eigenvectors are the columns, \([u_0, u_1, \ldots, u_N]\) and \( \Lambda \) is a diagonal matrix with the corresponding eigenvalues. This can be rewritten as

\[
C = U\Lambda^{1/2}I\Lambda^{1/2}U^\dagger
\]

where \( \Lambda^{1/2} \) is a diagonal matrix filled with \( \sqrt{\lambda_i} \). To create random realization of the correlation matrix, we need to create realizations of \( I \) that on average are the identity matrix. If we create a random, complex vector \( \mathbf{z} = [z_0, z_1, \ldots, z_N] \), where the elements \( z \) are mean 0, variance 1, and
7.3.3 Spider Noise Maps

random in phase\(^\dagger\) we can generate a matrix that on average is the identity matrix.

\[
I = \langle z^* \otimes z \rangle
\]  
(7.42)

For clarity, we can look at an \(N = 2\) example. \(z = [z_0, z_1]\). So the outer product of \(z^*\) and \(z\) is

\[
\begin{bmatrix}
z_0^* z_0 & z_0^* z_1 \\
z_1^* z_0 & z_1^* z_1
\end{bmatrix}.
\]  
(7.43)

The expectation value of \(z_i^* z_i\) is 1 because we generated numbers of variance 1. The expectation value of \(z_i^* z_j\), where \(i \neq j\), is 0 because \(z_i\) and \(z_j\) are uncorrelated.

Equation 7.41 can be rewritten as

\[
C = U \Lambda^{1/2} Z^* \otimes Z \Lambda^{1/2} U^\dagger.
\]  
(7.44)

By comparing eq. 7.44 with eq. 7.38, we immediately see that

\[
Y = Z \Lambda^{1/2} U^\dagger.
\]  
(7.45)

To get random noise in the time domain, simply inverse Fourier transform \(Y\).

7.3.3 Spider Noise Maps

In this section, we will generate uncorrelated and correlated noise and compare the two. For simplicity, we look at 8 detectors, each detector consists of two orthogonally polarized bolometers, for a total of 16 bolometers. Combined with the half-wave plate rotation and sky rotation, this will give excellent Q and U coverage through the flight. The input correlation structure is from the first 16 detectors on X1. The real correlation matrix that generated this structure is shown in Figure 5.9. Specifically, the correlation between the two detectors is given by Figure 5.8. The

\(^\dagger\)This can be easily generated by drawing from a Gaussian distribution. If \(r_1\) and \(r_2\) are two random, independent draws from a Gaussian distribution of mean 0 and variance 1, then we can construct a random, complex number \(z = (r_1 + i r_2)/\sqrt{2}\).
uncorrelated map has the same noise amplitude as the correlated simulation. The difference of the
two maps is shown in Figure 7.8. In other words, the off diagonal elements of the correlation matrix
are set to zero while the diagonal elements are the same as the correlated simulation. The noise
timestreams are then fed into the mapmaker with the real SPIDER scan strategy and filtering.

Figures 7.6 and 7.7 show the uncorrelated and correlated maps generated for this discussion.
By eye, there is very little difference between the two cases. The edges of the map are extremely
noisy because of poor matrix conditioning, as discussed in Section 7.1.5.

We can further explore the noise by looking at its power spectrum, shown in Figure 7.9. These
are the cross spectra from interleaved maps described in Figure 7.11. We expect the noise to be
zero mean, which is borne out in the simulation.

7.4 Power Spectrum

The WMAP beach balls floating around the Princeton physics building are ample evidence that
the CMB can be projected onto a sphere. This is fairly obvious; the surface of last scattering is a
sphere around an observer with radius equal to the distance light can travel since decoupling. Any
pattern on the surface of a sphere can be decomposed using spherical harmonics, $Y_{\ell m}$.

7.4.1 Power Spectrum Basics

Consider a temperature anisotropy field $\Theta(\hat{n}) = \delta T(\hat{n})/T_0$. This scalar field can be decomposed into

$$\Theta(\hat{n}) = \sum_{\ell=1}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m}^T Y_{\ell m}(\theta, \phi)$$

(7.46)

where

$$a_{\ell m}^T = \int d\Omega \Theta(\theta, \phi) Y_{\ell m}(\theta, \phi).$$

(7.47)
Figure 7.6: The uncorrelated noise simulation map generated from two orthogonal, co-pointed detectors. The maps are I, Q, and U from top to bottom.
Figure 7.7: The correlated noise simulation map generated from two orthogonal, co-pointed detectors. The detectors have the same pointing as the maps in Figure 7.6. The maps are $I$, $Q$, and $U$ from top to bottom.
Figure 7.8: The difference of uncorrelated noise simulation shown in Figure 7.6 and correlated noise simulation shown in Figure 7.7. The maps are $I$, $Q$, and $U$ from top to bottom.
Figure 7.9: The noise spectra of correlated and uncorrelated noise simulations. This is simulated for 8 detectors, each with two orthogonally polarized bolometers. The noise is generated from the correlation matrix for X1. The data is binned with bin width 25. The center of the bins are offset by ±1 for visual clarity.
The superscript \( T \) notates that this describes only the temperature anisotropies. From this, we can construct the power spectrum by averaging over all \( m \) modes

\[
C^{TT}_\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} a^T_{\ell m} a^T_{\ell m}.
\]  

(7.48)

Of course, we are only allowed to average over \( m \) modes if the field is rotationally invariant. The \( C^{TT}_\ell \)s represent the variance in \( \Theta(\hat{n}) \) for each \( \ell \). We can see this by reexpressing eq. 7.48 as

\[
\langle \Theta(\hat{n})\Theta(\hat{n}') \rangle = \frac{1}{4\pi} \sum_{\ell=1}^{\infty} (2\ell + 1) C^{TT}_\ell P_\ell (\hat{n} \cdot \hat{n}')
\]

(7.49)

where \( P_\ell (\hat{n} \cdot \hat{n}') \) is a Legendre polynomial of order \( \ell \). This is exactly why we represent CMB maps as power spectra. It contains all the statistical information of the map, assuming the underlying process is Gaussian.

To extend this analysis to polarized data, we note that we made use of the rotation invariance of \( \Theta(\hat{n}) \) to generate \( C^{TT}_\ell \) in eq. 7.48. The Stokes \( Q \) and \( U \) maps are not invariant under rotation, so we must construct a new quantity that is invariant under rotation of angle \( \phi \)

\[
(Q \pm iU)(\hat{n}) = e^{\mp 2i\phi} (Q \pm iU)(\hat{n}).
\]

(7.50)

This is a spin-2 function, so we must use the spin weighted spherical harmonics for \( s = 2 \), typically denoted \( \pm 2Y_{\ell m} \) [85, 86]. This allows us to write the \( E \) and \( B \) mode coefficients as

\[
a^E_{\ell m} = -\left( 2a_{\ell m} + -2 a_{\ell m} \right) / 2
\]

(7.51)

\[
a^B_{\ell m} = -i \left( 2a_{\ell m} + -2 a_{\ell m} \right) / 2.
\]

(7.52)

These values can be combined in the same way as eq. 7.48 to generate \( C^{EE}_\ell \) and \( C^{TT}_\ell \). We can also calculate the cross terms, \( C^{TE}_\ell, C^{TB}_\ell \), and \( C^{EB}_\ell \). These are generated by multiplying \( a_{\ell m} \)s with
different terms. In other words, we rewrite eq. 7.48 as

\[ C_{XY}^\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^\ell a_{lm}^{X*} a_{lm}^Y \]  

(7.53)

where \( X \) and \( Y \) are all elements of \( \{T, E, B\} \).

We can also take the cross spectra between two different maps. To do this, we generate \( a_{\ell m} \) from two different maps, \( a_{\ell m1} \) and \( a_{\ell m2} \). Then we can generate all the terms described in eq. 7.53 but with two different input maps

\[ C_{XY}^{\ell_1\ell_2} = \frac{1}{2\ell + 1} \sum_{m=-\ell}^\ell a_{\ell m1}^{X*} a_{\ell m2}^Y. \]  

(7.54)

The advantage of using cross spectra is that the uncorrelated white noise in the maps does not square. In eq. 7.53 both \( a_{\ell m1}^{X*} \) and \( a_{\ell m} \) contain the same noise. If the noise is uncorrelated, then \( a_{\ell m1} \) and \( a_{\ell m2} \) contain different noise components, and these components when multiplied in eq. 7.54 are mean zero.

7.4.2 Cut Sky

SPIDER, like many experiments, does not image the entire sky. The spherical harmonics are, as the name implies, defined on the full sphere. Thus we must find a way to express the partial sky in terms of the full sky harmonics. SPIDER’s baseline power spectral estimator is POLSPICE [87, 88].

Again starting with the temperature field, we defined the pseudo-coefficients

\[ \tilde{a}_{\ell m}^T = \int d\Omega(\theta, \phi) W(\theta, \phi) Y_{\ell m}(\theta, \phi). \]  

(7.55)

where \( W(\theta, \phi) \) is a weight or apodization mask. Obviously \( W(\theta, \phi) \) is set to zero where there are no observations. We can use uniform weights in the observation region or we can weight by the square root of the integration time. Now we recover the pseudo-\( C_{\ell s} \)

\[ \tilde{C}_{\ell}^{TT} = \frac{1}{2\ell + 1} \sum_{m=-\ell}^\ell \tilde{a}_{\ell m}^{T*} \tilde{a}_{\ell m}^T. \]  

(7.56)
7.4.2 Cut Sky

<table>
<thead>
<tr>
<th>Mask</th>
<th>$f_{\text{sky}}$ [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>hitmask</td>
<td>6.45</td>
</tr>
<tr>
<td>latlon</td>
<td>4.83</td>
</tr>
<tr>
<td>vpol</td>
<td>3.50</td>
</tr>
</tbody>
</table>

Table 7.1: $f_{\text{sky}}$ for various masks. The masking is still being explored for SPIDER.

An analogous value is calculated for $\tilde{a}_{\ell m}^E$ and $\tilde{a}_{\ell m}^B$, generating analogous $\tilde{C}_{\ell}^{EE}$ and $\tilde{C}_{\ell}^{BB}$. The pseudo-$C_{\ell}$s are related to the true sky $C_{\ell}$s by the mode coupling kernel

$$\tilde{C}_{\ell} = \sum_{\ell'} K_{\ell\ell'} C_{\ell'}.$$  (7.57)

The values of $K_{\ell\ell'}$ are due to the windowing and its specific values are dependent on the window and weighting chosen. POLSPICE actually can return $K_{\ell\ell'}$.

As a practical notion, want to keep as large a mask as possible to include as much data as possible. Particularly important is keeping as many modes as possible, and a larger map contains more modes. We are pushed to smaller masks from noise and poor matrix conditioning, discussed in Figure 7.2.

SPIDER attempted multiple masks; the three most commonly used in our internal analysis are shown in Figure 7.10. The hitmask is generated by finding all pixels in the SPIDER map that are above a threshold for hits count and condition number. This is the largest mask, but it does not pass null tests. The latlon mask is a simple by-eye mask with edges based on latitude and longitude. The edges are (-55, -15) in latitude and (18, 80) in longitude. The vpol mask is another mask based on latitude and longitude cuts that was used for the circular polarization SPIDER paper [89]. The edges are (-55, -15) in latitude and (30, 75) in longitude. It was chosen because it minimized systematics.

The masking scheme is still being explored in the SPIDER analysis pipeline. The $f_{\text{sky}}$ of the various masks are shown in Table 7.1.
null tests

The goal of a null test is to generate noise-only maps by differencing maps that ideally have the same signal. Consider a map

\[ m_i = s + n_i \]  

(7.58)

where \( s \) is the sky signal, \( n_i \) is the noise in the map, and \( i \) indexes the map. The sky signal has no index because it is the same in all maps. The difference between two maps is

\[ \Delta m_{ij} = m_i - m_j = n_i - n_j \]  

(7.59)

which is a noise-only map. If the noise components are systematic free, then these maps should be purely Gaussian noise. To test this, we take the cross product with another null map. Consider

\[ \Delta m_{ij} \times m_{kl} = (n_i - n_j) \times (n_k - n_l) \]

\[ = n_i \times n_k - n_i \times n_l - n_j \times n_k + n_j \times n_l \]  

(7.60)

where \( \times \) denotes a cross spectrum between two maps. If the noise is white, the values of all the cross spectra will be null in the mean. However if there are systematics that correlate maps, then some of the terms will be non-null. For example, if the maps \( i \) and \( k \) have a shared systematic, the \( n_i \times n_k \) term will be non zero. All the other crosses may be zero (or have their own correlations). This will create a residual signal in the spectra.

The results of null tests can be extremely sensitive to the choice of data splits. We can easily hide a systematic by choosing a data split which cancels that systematic. For example, if \( i \) and \( j \) share a systematic, that systematic is canceled in this null test. We can permute \( i, j, k, \) and \( l \) to see the systematic.

A huge suite of null tests is currently being performed on the SPIDER data, and a comprehensive review is presented in Anne Gambrel’s thesis\[90\]. Here, we will review the most basic null test, called the *interleaved chunk* null test. We divide the entire flight into approximately 10 minute
7.5 Null Tests

chunks where the boundaries of the chunk are set by turn arounds. We then collect every fourth chunk and generate a map. We do this for every set of four. A schematic to better explain the setup is shown in Figure 7.11.

With four maps in hand, we difference them and then calculate their cross spectra. An example null map is shown in Figure 7.13. The null map has been smoothed with a 20 arcminute beam. Six null maps are generated by differencing all permutations of the four chunks. These can be crossed to generate 3 cross spectra. It is important to note that these are not independent null tests because the same underlying maps are used in multiple crosses. These are shown in Figure 7.12.

These cross spectra do not have error bars. To assess the uncertainty in the null maps, we have two options. We can either generate noise ensembles and assume that their distribution represents the true instrumental uncertainty. Or we can look at the statistics of a greater number of cross spectra. SPIDER is currently employing both strategies.
Figure 7.10: Three of the masks that SPIDER uses in its analysis pipeline. From top to bottom are the hitmask, latlon, and vpol.
7.5 Null Tests

Figure 7.11: A schematic of the interleaved chunks over an 80 minute period. 10 minute chunks represented by the same color are binned into a single map.

Figure 7.12: The three null cross spectra generated using four interleaved chunks. Note that these null spectra are not entirely independent because some of the cross spectra repeat. $EE$ and $BB$ are shown on the left and right respectively.
Figure 7.13: The difference map of chunk 1 and chunk 2. The maps are smoothed with a 20 arcminute beam. There is visible large scale structure in the temperature map. It is impossible to distinguish systematics versus noise in the $Q$ and $U$ maps.
Part IV

SPIDER First Flight Results
Chapter 8

Foregrounds

8.1 Foreground Components

The CMB community has been considering foreground emissions for decades \[91, 92\]. The ultimate goal of a CMB experiment is to measure only photons from the CMB. However, sources between the telescope and the surface of last scattering confound the data.

Before beginning the discussion on foregrounds, it is important to clarify that we will be discussing two different types of spectra. The first is the frequency spectrum of the source, which lives in $\nu$ space. The second is the multipole spectrum, which lives in $\ell$ space.

There are several difficulties associated with separating the CMB data from foreground sources. Proper component separation requires high signal to noise measurements of the foreground components. High signal to noise regions are by definition where the foreground component is bright. In
these regions, we have detailed maps of the foreground emission. In fact, the galactic astrophysics community is able to create high resolution 3D maps of dust in the galactic plane \cite{93}. However the CMB community looks away from the galactic plane precisely to avoid the high intensity dust component. The dust is difficult to measure because we have deliberately chosen a region with a low dust amplitude.

Also of concern is decorrelation between frequencies. Measurements of the spatial shape of the foreground components at one frequency must scale to another frequency according to some model. If this statement is not true, and in fact there is no correlation between measurements at two different frequencies, then it is impossible to use measurements at one frequency to clean another.

In this section we will discuss various sources of foreground contamination, including dust (Section 8.1.1), synchrotron (Section 8.1.2), and free-free (Section 8.1.3).

### 8.1.1 Dust

It has been demonstrated that foreground dust contaminates polarized measurements of the CMB \cite{92}. Characterizing and removing dust contributions from the CMB is a critical task for making precision measurements of the polarized CMB. The dust is classified into several different categories\cite{94,95}. In this subsection, we will discuss thermal dust and spinning dust.

#### 8.1.1.1 Thermal Dust

Interstellar dust consists mainly of graphite, silicates, and polycyclic aromatic hydrocarbons. The intensity of dust is often described as a modified black-body

\[
I_d(\nu) \propto \nu^{\beta_d} B_\nu(T_d)
\]

(8.1)

where \(\nu\) is the frequency of the emission, \(B_\nu\) is the black body spectrum (or Planck spectrum), \(T_d\) is the dust temperature, and \(\beta_d\) is the dust spectral index. The emissivity as a function of \(\beta_d\) and \(T_d\) is shown in Figure 8.1. The values are normalized to the dust intensity at 353 GHz. The SPIDER bands are shown in gray.
Figure 8.1: The dust emissivity for various values of $\beta_d$ and $T_d$ shown in the top and bottom respectively. Values are normalized to the amplitude at 353 GHz. The SPIDER bands are shown in gray. The 285 GHz band will fly in SPIDER-2. The 2015 Planck results suggest the best fit parameters are $\beta_d = 1.53 \pm .05$ and $T_d = 21 \pm 2$. 
8.1.2 Synchrotron

As $\beta_d$ increases, the relative dust power at 95/150 GHz versus 353 GHz decreases. The same is true for increasing $T_d$. Given this model, the dust is almost 8 times brighter at 280 GHz than 150 GHz, making the SPIDER-2 280 GHz channel ideal for measuring dust.

The Planck intermediate results found that $\beta^I_d$, the value of $\beta_d$ for intensity only maps was $1.52 \pm .01$. It also found the polarized value $\beta^P_d = 1.59 \pm .02$ [96]. However a more recent analysis by Planck suggests that the detection of $\beta^I_d \neq \beta^P_d$ may be purely due large scale systematics in the Planck dataset [62]. They reestimate the value of dust parameters, only reporting the joint temperature and polarization value, as $\beta_d = 1.53 \pm .05$ and $T_d = 21 \pm 2$.

Typically the sky is modeled with a single modified black body, as described above. However the galactic astrophysics community believes there is more than one type of dust in the interstellar medium, with different emission properties and temperatures [97].

8.1.2.2 Spinning Dust

Early measurements of the sky at 10-60 GHz discovered excess power over previously known sources. This excess power was interpreted as rotational emission from rapidly rotating nanoparticles [98,99]. For polarization measurements, it is natural to be concerned with spinning dust as the rotation creates an angular momentum which could align with local magnetic fields. Fortunately for CMB experiments, a detailed study of the quantum mechanical nature of the dust particles shows that dust alignment is heavily suppressed [100]. Grains which are small enough to have GHz rotations have quantized vibrational states that do not allow for rotation to vibration energy transfer. This greatly limits the ability for a dust particle to align. It is believed that the polarized fraction of spinning dust can not be greater than 0.5%. For the purposes of SPIDER, we will ignore spinning dust.

8.1.2 Synchrotron

Synchrotron radiation is sourced by cosmic rays interacting with magnetic fields in the Galaxy. Electrons are accelerated by the magnetic field, generating photons. The synchrotron emission is difficult to model because it varies spatially and in the electromagnetic domain. In other words, it
8.1.3 Free-Free

is not well described by a single power law

\[ I_s(\nu) \propto \nu^{\beta_s} \]  \hspace{1cm} (8.2)

where \( \beta_s \) is the synchrotron spectral index. Although the frequency spectrum flattens at lower frequencies, the spectrum above 30 GHz is well described by a power law with \( \beta_s = -3 \) \cite{62, 101}. This varies spatially by \( \pm 2 \).

Since synchrotron radiation is generated by electrons in a magnetic field, emission is polarized with the magnetic field lines. WMAP estimates that the polarized fraction is 2-4% near the galactic plane and 20% at high latitude \cite{102}.

Higher signal to noise maps at low frequencies where synchrotron is brighter are needed to clean CMB maps. Low frequency maps are more difficult because the pixels must be larger due to the longer wavelengths. Despite these difficulties, efforts are underway to measure the sky at these low frequencies \cite{103}.

8.1.3 Free-Free

In a plasma, free electrons scatter off ions, generating bremsstrahlung radiation \cite{104}. This process is seen in the interstellar plasma, and is tracked with high significance using H\( \alpha \) line emission \cite{105}.

The spectral index for free-free emission is \( \beta_b \approx -2.14 \) for \( T_e \approx 8000K \) \cite{105}.

While every individual scattering event is polarized, the sum of all events from a source is unpolarized because the scattering directions are isotropic and random. Interstellar magnetic fields are too weak to polarize ions to generate a polarized emission. WMAP puts an upper limit of about 1% at high galactic latitudes (away from the galactic plane).

8.2 Component Separation

The literature is replete with different methods of component separation, but all the methods fall into three broad categories based on the domain they work in – map based, spectral based, and needlet based. SPIDER implement ma and spectral based methods which will be described
in Section 8.2.1 and Section 8.2.2 respectively. A needlet-based method is also being tentatively explored. The trade off between the two methods is a choice of which uncertainties are believed to be dominant. Foregrounds are known to be anisotropic across the sky, making spectral based methods inherently incomplete descriptions of foregrounds. On the other hand, purely map based approaches provide no freedom for different weights at different Fourier components. This is particularly useful for maps with non-negligible noise. The hope is that both approaches, with their different strengths and shortcomings, recover the same result.

Every method leverages several assumptions, the most important being that foregrounds have a different frequency spectrum than the CMB. The CMB is a nearly perfect black body, while foreground sources tend to be modeled by modified power laws. If foregrounds and the CMB had the same spectral shape, it would be nearly impossible to separate them.

### 8.2.1 Map Based Component Separation

The most conceptually simple method of component separating is the map based approach. In the most simple case, one constructs a foreground template in the map domain and simply fits it to the CMB map.[106] Any signal that correlates with the foreground map is assumed to be foreground signal, and the subtracted map is considered a clean CMB map with some bias and noise.

#### 8.2.1.1 Map Based Component Separation Theory

Every map is a linear combination of CMB signal, $s$, foregrounds, $f$, and noise, $n$,

$$m_{\nu_i}(p) = s_{\nu_i}(p) + f_{\nu_i}(p) + n_{\nu_i}(p)$$  \hspace{1cm} (8.3)

where $p$ indexes a pixel, and $\nu_i$ indexes observation frequency. Typically, maps are made in units of $K_{\text{CMB}}$, so the frequency dependence of $s(p)$ is normalized away. We will drop the $\nu_i$ subscripts from $s(p)$. 

163
8.2.1 Map Based Component Separation

This normalization makes generating foreground templates easy

\[ t_{\nu_i\nu_j}(p) = m_{\nu_i}(p) - m_{\nu_j}(p) \]

\[ = s(p) + f_{\nu_i}(p) + n_{\nu_i}(p) - s(p) - f_{\nu_j}(p) - n_{\nu_j}(p) \]

\[ = f_{\nu_i}(p) - f_{\nu_j}(p) + n_{\nu_i}(p) - n_{\nu_j}(p). \] (8.4)

If the foreground had the same frequency dependence as the CMB, then \( f_{\nu_i}(p) - f_{\nu_j}(p) \) would be zero, and the template would be purely noise. Fortunately, all the sources we discussed in Section 8.1 are not black bodies. In fact, if there were no spatial variation in foreground spectra, \( t_{\nu_i\nu_j}(p) \) would completely describe the specific foreground component up to a scale factor.

We can generate template maps with any pair of frequency maps. SPIDER only observed in two frequencies, so it is impossible to generate many templates with SPIDER data only. Fortunately Planck observed at many frequencies which allow for many different combinations. Using Planck also has the advantage of having systematics that are uncorrelated with SPIDER.

With templates in hand, we can now attempt subtracting foreground from our CMB maps. For every CMB map, we attempt to find the parameters \( \beta_{\nu_i\nu_j} \)

\[ \hat{s}_{\nu_k}(p) = m_{\nu_k}(p) - \sum_{\nu_i\nu_j} \beta_{\nu_i\nu_j} t_{\nu_i\nu_j}(p). \] (8.5)

that minimizes

\[ \chi^2 = \sum_{pp'} \left[ m_{\nu_k}(p) - \sum_{\nu_i\nu_j} \beta_{\nu_i\nu_j} t_{\nu_i\nu_j}(p) \right] C_{pp'} \left[ m_{\nu_k}(p') - \sum_{\nu_i\nu_j} \beta_{\nu_i\nu_j} t_{\nu_i\nu_j}(p') \right] \] (8.6)

where \( C_{pp'} \) is the pixel-pixel covariance matrix. This covariance matrix is defined as

\[ C_{pp'} = \langle m(p)m(p') \rangle - \langle m(p) \rangle \langle m(p') \rangle \] (8.7)

where \( \langle \rangle \) indicates an expectation value. In principle this matrix can be estimated with simulations. However, in practice this matrix is intractably large. For example, a HEALPix map with nside
512 has approximately 150,000 pixels in the SPIDER region. So the covariance matrix is 150,000^2 elements large. We make the simplifying assumption that the covariance matrix is diagonal, making eq. (8.6)

\[ \chi^2 = \sum_p \left[ m_{\nu_k} (p) - \sum_{\nu_i, \nu_j} \beta_{\nu_i, \nu_j} t_{\nu_i, \nu_j} (p) \right]^2. \]  

This fitting routine has very few free parameters. If there are \( N_t \) templates, then there are \( N_t \) free parameters, specifically the values of \( \beta_{\nu_i, \nu_j} \).

### 8.2.1.2 Generating Foreground Templates

SPIDER uses a single template derived from the difference of Planck 353 and 100 GHz maps. The discussion of that choice is deferred to Section 8.2.1.3. Since dust grows as a modified black body with a positive spectral index, the higher the frequency the higher the dust amplitude. Thus, the 353 GHz map is perfect for dust template generation. The 100 GHz map has low dust amplitude, so subtracting it out removes primarily CMB only, creating a high signal to noise template. There are some details of making a template map that are worth going over briefly, and are described in more detail in Ivan Padilla’s thesis [44].

- **Point Sources**: Any survey of the sky will have bright point sources in it, and the Planck maps are no exception. Planck publishes a compact source catalog [107] which we use to mask the brightest point sources. The masked points are then inpainted using a nearest neighbor average. It is important to do the inpainting on the raw Planck maps, rather than the smoothed ones we are about to create. A raw inpainted map is shown in the top panel of Figure 8.2.

- **Beam Smoothing**: Planck [108] and SPIDER have very different beams. Since we are going to use these templates to fit the sky as seen by SPIDER, we must deconvolve the Planck beam and then apply the SPIDER beam. A beam convolved map is shown in the middle panel of Figure 8.2.
8.2.1 Map Based Component Separation

- **Downgrading**: To prevent aliasing effects, the template maps are downgraded from nside 2048 to nside 512.

- **Reobservation**: The SPIDER mapmaker, described in Chapter 7, flags the data and removes a fifth order polynomial per half scan. The polynomial filter in particular removes low $\ell$ modes. To make the template maps compatible with the SPIDER maps, we must remove those modes in the templates as well. To do this, we take the template maps and reobserve them with the SPIDER mapmaker. This is shown in the bottom panel of Figure 8.2.

In principle, we can immediately fit the template we just created to the SPIDER 150 and 95 GHz maps. However eq. 8.4 tells us that the templates we generated are not only foreground signal, but also Planck noise. This noise will bias the fit parameter, $\beta_{\nu_i\nu_j}$.

There are multiple methods to handle the Planck noise. The first method is to attempt to down-weight $\ell$ modes that have low signal to noise. We know that dust has high power at low multipoles, while the noise is approximately flat in $\ell$ space. So we want to down-weight high-$\ell$ modes in the template before fitting. A simple way to do this is to smooth the template map.

We can use the Planck data splits to characterize the noise in the templates generated by differencing Planck 353 and 100 GHz maps. In addition to the full mission map, Planck also released half-ring and half-mission maps. If we take the cross spectra of the half-ring (or half-mission) maps, and divide by the full mission map, we get an estimate of the signal to noise ratio. To see this, calculate the auto-spectra of the full mission map

$$m_{\text{full}} \times m_{\text{full}} = s_{\text{full}} \times s_{\text{full}} + n_{\text{full}} \times n_{\text{full}} + 2s_{\text{full}} \times n_{\text{full}}$$

$$= s_{\text{full}} \times s_{\text{full}} + n_{\text{full}} \times n_{\text{full}}$$

(8.9)

and the cross spectra of the half maps

$$m_{H1} \times m_{H2} = s_{H1} \times s_{H2} + n_{H1} \times n_{H2} + s_{H1} \times n_{H2} + s_{H2} \times n_{H1}$$

$$= s_{H1} \times s_{H2}.$$  

(8.10)
8.2.1 Map Based Component Separation

Figure 8.2: An example of the foreground template generated from differencing Planck 353 and 100 GHz maps. Shown are the $I$ maps. The top panel is the raw difference map after inpainting. The middle panel is the template after correcting for the Planck beam. The bottom panel is the reobserved map. The top two maps have been median subtracted.
Figure 8.3: The signal divided by signal plus noise for the foreground templates generated by differencing Planck 353 and 100 GHz maps. The black line is a best fit beam model to the data. The beam model is not physically motivated; it is chosen because it fits the data well. The rise in power at $\ell > 250$ in the 94 GHz map is due to low total signal in the maps from the beam smoothing step in the template generation, so the denominator is effectively zero.

We use the symbol $s$, which is the sky signal (foregrounds and CMB), to distinguish from $s$ introduced in eq. 8.3, which is only CMB signal. The $\times$ symbol implies a cross spectra between two maps as described by eq. 7.54. We assume the signal and noise are uncorrelated, so the expectation value of the $s \times n$ term goes to zero in eq. 8.9. This is also true in eq. 8.10, with the additional feature that the expectation value of the noise goes to zero as well because the maps have uncorrelated noise. The expectation value of the $s$ is the same independent of input map, since the underlying sky is the same, so we can drop the subscripts for the $s$ term. So the ratio of the two spectra is

$$\frac{m_{H1} \times m_{H2}}{m_{\text{full}} \times m_{\text{full}}} = \frac{s \times s}{s \times s + n_{\text{full}} \times n_{\text{full}}},$$

or in other words just the signal divided by the signal plus noise. Remarkably, this is the form of a
8.2.1 Map Based Component Separation

<table>
<thead>
<tr>
<th>Freq</th>
<th>Raw Fit</th>
<th>Smooth Fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>95 GHz</td>
<td>.030</td>
<td>.042</td>
</tr>
<tr>
<td>150 GHz</td>
<td>.050</td>
<td>.066</td>
</tr>
</tbody>
</table>

Table 8.1: The fit parameters for map based foreground removal. In other words, they are the values of $\beta_{\nu_i\nu_j}$ in eq. 8.6. The left column is using the reobserved, but not resmoothed Planck templates. The right column is with the resmoothed Planck templates based on Figure 8.3.

Simple Wiener filter. If we can convert the ratio in the spectral domain into a map domain kernel, we can optimally filter our data for fitting the template to the foregrounds.

Figure 8.3 shows the ratio described in eq. 8.11 for the Planck 353-100 GHz template. I show both the half-mission and half-ring crosses. We could simply smooth the maps with the values shown in Figure 8.3; however, fitting a functional model to the data will reduce the noise coupling in via the kernel. I fit a Gaussian beam model to the ratio at each observation frequency. It should be made clear that there is no physical motivation for fitting a beam model (as opposed to some other functional form), except that it fits the data well. The fit is shown in black. This best fit model is applied to the template and the Spider maps, then the template is fit to the Spider map. The fit parameters are shown in Table 8.1. These values are still changing as we are still in the process of understanding Spider systematics.

The values of the fit parameter are higher in the smoothed case. Uncorrelated noise in the template tends to reduce the fit parameter, since the fit wants to drive $\beta_{\nu_i\nu_j}$ to zero for uncorrelated components. So reducing the noise in the template should result in higher fit values. The resmoothed templates are shown in Figure 8.4.

8.2.1.3 Choosing Templates

Planck observed at 30, 44, 70, 100, 143, 217, 353, 545, and 857 GHz. In theory, we could generate a template using eq. 8.4 for every frequency pair. Unfortunately the Planck low frequency instrument (LFI) has systematics that make it difficult to use to constrain foregrounds [109]. These channels would be ideal for constraining synchrotron because synchrotron has a negative spectral index.
Figure 8.4: The foreground templates after smoothing to reduce high-\(\ell\) noise. From top to bottom are the Stokes \(I\), \(Q\), and \(U\). The edges of the map are noisy due to low Stokes coverage from the scan strategy and are masked in the foreground fit and the analysis.
Figure 8.5: The 150 GHz EE spectrum after foreground template subtracting using different Planck foreground templates. The spectra are generated using cross-spectra of quarter chunk maps, with Spider providing the 150 GHz map. Note that the y-axis is plotted in raw $C_\ell$, not $D_\ell$.

The two highest frequency channels do not have polarization data, so those are omitted from consideration as well.

The remaining channels are 100, 143, 217, and 353 GHz. Before computing, consider the possible trade-offs. Dust gets brighter with rising frequency, so the highest dust signal should be at 353 GHz. However 353 GHz is also the noisiest Planck channel, with a $\Delta T/T$ sensitivity of 29.8 in polarization. In contrast 100, 143, and 217 have sensitivities of 4.0, 4.2, and 9.8 respectively. We may also be concerned that the dust map at 353 GHz begins to decorrelate from the maps at our CMB frequencies. If that is the case, then 217 GHz maps may be the best foreground templates.

Figure 8.5 shows the 150 GHz EE spectra after foreground subtraction done with three different templates, 353-217, 353-100, and 217-100. The 150 GHz map comes from Spider, while the templates come from Planck differences. The spectra are generated using cross-spectra of quarter chunk maps. An important note is that the subtraction is done with full mission Planck maps, so the Planck noise is being squared. This exercise was repeated with different half mission maps on both sides of the cross spectra, and we arrived at the same result.

All templates reduce EE power at low multipoles, which is the expected performance of fore-
8.2.1 Map Based Component Separation

<table>
<thead>
<tr>
<th>Freq</th>
<th>Joint</th>
<th>I Only</th>
<th>P Only</th>
</tr>
</thead>
<tbody>
<tr>
<td>95 GHz</td>
<td>.042</td>
<td>.043</td>
<td>.015</td>
</tr>
<tr>
<td>150 GHz</td>
<td>.066</td>
<td>.067</td>
<td>.036</td>
</tr>
</tbody>
</table>

Table 8.2: The map based foreground fit parameters for a joint solution as well as $I$ and $P$ ($Q$ and $U$) separately. The joint fit is driven almost entirely by the $I$ solution, as evidenced by the $I$ only value being nearly identical to the joint solution. The $P$ solution more noise biased than the $I$ solution.

ground removal. The 217-100 template clearly suffers from excess noise. This is because the 217-100 template requires a much higher fit parameter, so the Planck noise couples into the map with high amplitude. The 353-100 and 353-217 templates are largely the same. 353-100 is slightly better because it has slightly higher foreground amplitude, so the fit parameter is smaller. Also the 100 GHz Planck map has lower noise than the 217 GHz map.

8.2.1.4 Joint Fitting IQU

Up until this point, we have ignored an important detail. All the maps in question are polarized, so it is an open question how to use the polarization maps ($Q$ and $U$) and the intensity map. We could choose to fit $I$, $Q$, and $U$ maps simultaneously, or we could fit $I$ independent of $Q$ and $U$.

Table 8.2 shows the fit parameters for $IQU$ jointly fit on the left, and the $I$ and $P$ values separately on the middle and right. The joint values are the same as those in Table 8.1.

The prior from Planck is that $\beta_d$ is the same between $I$ and $P$ [62]. We should be cautious interpreting the data in Table 8.2 as evidence that we see different spectral indexes for $I$ and $P$. We already showed in Section 8.2.1.2 that the noise in the templates bias the fit parameters low, and we will continue to explore the noise bias in Sections 8.2.1.6 and 8.2.1.7. For now, it is enough to note that for polarization templates the signal amplitude is lower the noise amplitude is higher. Thus we expect the amplitude of the noise bias to be greater in polarization than dust.
8.2.1 Map Based Component Separation

8.2.1.5 Spatial Variations

By fitting a single parameter to the entire SPIDER region, we have assumed that the dust spectral properties are constant across the sky. This is an assumption worth checking. We will leverage our large sky coverage and break up the full SPIDER observation region into subregions and then fit the foreground templates.

Figures 8.6 and 8.7 show the fit parameters in twelve subregions at 95 GHz and 150 GHz respectively. The background image is the Planck 353-100 GHz foreground template $I$ for the top two plots and $P = \sqrt{Q^2 + U^2}$ for the bottom plot. There is large variation in foreground fit parameters. The features at 95 GHz track the features in 150 GHz. In other words, subregions with high (or low) values in 95 GHz are also high (or low) in 150 GHz. The absolute amplitudes are different because of the rising amplitude of dust with frequency.

Particularly concerning are the negative values in some of the subregions; we do not expect any anti-correlations between the template and SPIDER. This suggests that chance correlations between the CMB maps and template are driving the fit negative in those subregions. We can see that there is much greater spatial variation in $I$ than $P$. This is due to the fact that the $I_{\text{CMB}}/I_{\text{fg}} > P_{\text{CMB}}/P_{\text{fg}}$. So small chance correlations between the template and the CMB have large impacts on the fit parameters.

8.2.1.6 Template Noise Variance and Bias

There are several ways the foreground fit parameters can be biased. Even if we had a perfect, noiseless measurement of the sky and a perfect, noiseless foreground template, chance correlations or anti-correlations between the template and CMB component of the sky will introduce an error in the fit parameter. And when we consider the real measurements that are neither perfect nor noiseless, we can have correlations between noise and systematics in both the CMB maps and the templates. In fact, we have already seen how uncorrelated noise in the template will tend to drive the fit parameter low in Section 8.2.1.2

We further study the template noise bias by simulating Planck noise using the Planck noise
Figure 8.6: The foreground fit parameters in different subregions at 95 GHz. Plotted in the background is the foreground $I$ and $P$ template to provide context for dust amplitude. The plots are joint, $I$ only, and $P$ only from top to bottom.
Figure 8.7: The foreground fit parameters in different subregions at 150 GHz. Plotted in the background is the foreground I and P template to provide context for dust amplitude. The plots are joint, I only, and P only from top to bottom.
8.2.1 Map Based Component Separation

![Figure 8.8](image-url)

Figure 8.8: The fit parameters for an input foreground of amplitude .05 as a function of noise bias. 

$I$ and $P$ are fit separately in blue and orange respectively. All $I$, $Q$, and $U$ maps fit simultaneously in yellow. The error bars are $5 \times$ the standard deviation.

covariance matrix. In this simulation, we assume that the resmoothed 353-100 GHz template is the true foreground sky with no noise. We then treat the sky signal as a $\Lambda$CDM realization plus the resmoothed 353-100 GHz templates scaled by .05. In other words the true value of the fit parameter $\beta_{353-100}$ is .05 by construction. The sky is assumed to have no noise in this simulation. We then create the template by summing the resmoothed 353-100 GHz map and the noise simulation.

Before looking at how template noise affects the fit results, there is an important nuance. The *Planck* noise covariance does not generate noise that is consistent with the data. We know an approximate value of the noise in the spectral domain from Figure 8.3. We do not reproduce these results with the noise simulations. In fact, the noise simulations need to be multiplied by a factor of 10 in the map domain before the results match.

With that in mind, the results are shown in Figure 8.8. It is obvious that increasing noise decreases the fit parameters. The fit bias is small for $I$ only, but dramatic for $P$. This is due to the dramatically higher signal in $I$. The joint fit is an intermediate case, as expected. Another interesting note is that at a noise scaling of 10, the fit $P$ amplitude is about 1/3 of the actual input value. That is approximately the offset of the $P$ only value in Table 8.2 relative to the $I$ only value at 95 GHz. If we believe that the $I$ value in Table 8.2 is not significantly noise biased, then
8.2.1 Map Based Component Separation

![Map Based Component Separation](image)

Figure 8.9: The standard deviation of fit parameters per subregion due to Spider noise. The background is the foreground template in $I$. The fit was done with $I$ and $P$ jointly.

it represents the true value of the fit parameter.

8.2.1.7 CMB Map Noise Variance and Bias

In Section 8.2.1.6, we studied how noise in the templates bias the fit parameters. In this section, we study how the noise in the Spider CMB maps bias the fit parameters. The Spider mapmaker can also make Spider noise simulations. We add realizations of the Spider noise to a $\Lambda$CDM map and treat that as the observed sky. We then calculate the variance in the fit parameter due to the noise.

This procedure is repeated for subregions and shown in Figure 8.9. The variance on fit parameter due to Spider noise is insignificant relative to expected foreground amplitude in individual subregions. The standard deviation for the entire lat-lon region is .00013.
8.2.1 Map Based Component Separation

8.2.1.8 CMB Signal Variance and Bias

While the CMB and foregrounds are generated by entirely separate sources, there are still chance correlations. In fact, the CMB is much brighter than the foregrounds in $I$ that even small chance correlations with the foreground template can significantly bias the fit result. To quantify this, I generate an ensemble of reobserved $\Lambda$CDM sims and fit the foreground template to it. I do not inject any foreground signal into the $\Lambda$CDM map. The variance in fit parameters gives an estimate of the uncertainty introduced by the CMB sky. Over the latlon region, the standard deviation of fit parameters is .0166. This is a significant variance relative to the expected fit parameter.

We can repeat this exercise for subregions, shown in Figure 8.10. It is no surprise that the variance per subregion is much higher. In fact, this implies that we should not be able to resolve the foreground parameters in any of the subregions by themselves, with perhaps two exceptions. The variance due to chance correlations is lowest in the brightest regions.
8.2.1 Map Based Component Separation

<table>
<thead>
<tr>
<th></th>
<th>Joint Bias</th>
<th>Joint $\sigma$</th>
<th>$I$ Only Bias</th>
<th>$I$ Only $\sigma$</th>
<th>$P$ Only Bias</th>
<th>$P$ Only $\sigma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temp. noise naive</td>
<td>-3.8E-5</td>
<td>7.7E-5</td>
<td>-2.8E-5</td>
<td>8.0E-5</td>
<td>-1.4E-4</td>
<td>2.5E-4</td>
</tr>
<tr>
<td>Temp. noise 10x</td>
<td>-0.062</td>
<td>7.9E-4</td>
<td>-1.3E-4</td>
<td>8.0E-4</td>
<td>-0.036</td>
<td>5.8E-4</td>
</tr>
<tr>
<td>CMB map noise</td>
<td>-0.0011</td>
<td>1.3E-4</td>
<td>-2.0E-5</td>
<td>1.3E-4</td>
<td>-0.0050</td>
<td>0.0010</td>
</tr>
<tr>
<td>CMB corr.</td>
<td>-0.0080</td>
<td>0.017</td>
<td>-0.0081</td>
<td>0.017</td>
<td>-2.5E-4</td>
<td>0.0012</td>
</tr>
</tbody>
</table>

Table 8.3: The bias and uncertainty, $\sigma$, for an input foreground amplitude of .05 in the latlon region. The bias and uncertainty are the mean and standard deviation of an ensemble of sims. The top two lines are described in Section 8.2.1.6. The third line is described in Section 8.2.1.7. The fourth line is described in Section 8.2.1.8.

8.2.1.9 Map Based Foreground Variance and Bias Summary

There are many separate but related points in Sections 8.2.1.6 to 8.2.1.8. The take away points are:

- The fit parameter is driven primarily by $I$.

- Noise in the SPIDER CMB maps do not affect the foreground fit in any appreciable way when joint fitting $I$ and $P$.

- Noise in the Planck foreground templates can and do affect the fit parameters, depending on the underlying assumption of the Planck noise. In particular, if the Planck noise is high, the fit parameter is biased low.

- The underlying CMB sky causes large chance correlations particularly in $I$.

The values bias and uncertainty for the latlon region is shown in Table 8.3 for an input foreground amplitude of 0.05. Both joint and separate fits shown. The CMB both biases and adds variance to the fit parameter in $I$. This suggests that it may be worth attempting looking at $P$ only results, not because we believe that the true fit parameters are different in $I$ and $P$, but because the variance is less in $P$. 

179
8.2.1 Map Based Component Separation

8.2.1.10 Relating Fit Parameters to $\beta_d$ and $T_d$

We want to derive a relation between the foreground fit parameter and the foreground model parameters, $\beta_d$ and $T_d$. The following derivation follows from an internal note by Jamil Shariff [110]. The subtle detail that makes the extrapolation nontrivial is that the maps are all in units of $\mu K_{CMB}$, which scales as a black-body, while the dust component scales as a modified black-body as described by eq. [8.1]. Typically the dust amplitude is normalized to the amplitude at some reference frequency, $\nu_0$.

$$I_d^d = I_{\nu_0} \left( \frac{\nu}{\nu_0} \right)^{\beta_d} \frac{B_{\nu}(T_d)}{B_{\nu_0}(T_d)}. \quad (8.12)$$

For SPIDER, we set $\nu_0$ to 353 GHz because it is the highest amplitude dust channel.

We want to relate the dust emission, which we have defined in units of power, to a change in CMB temperature. To first order, this is given by

$$\Delta T_d(\nu) = I_\nu \left( \frac{\partial B_\nu(T)}{\partial T} \bigg|_{T=T_{CMB}} \right)^{-1}. \quad (8.13)$$

The derivative of $B_\nu$ with respect to $T$ is

$$\frac{\partial B_\nu(T)}{\partial T} = \frac{c^2}{2\nu^2T^2} \exp \left( \frac{h\nu}{k_B T} \right) B_\nu(T)^2. \quad (8.14)$$

To be explicit about the notation here, $\nu_i$ is the CMB band used to generate the template, $\nu_0$ is the frequency used to trace dust, and $\nu_k$ is the frequency of the map that the template is fit to. So the fit parameter $\beta_{\nu_i\nu_0}$ for a map at frequency $\nu_k$ is

$$\beta_{\nu_i\nu_0} = \frac{I_{\nu_k}}{I_{\nu_0}} \left( \frac{\nu_k}{\nu_0} \right)^{2+\beta_d} \exp \left[ \frac{h}{k_B T_{CMB}} (\nu_0 - \nu_k) \right] \frac{B_{\nu_k}(T_d)}{B_{\nu_0}(T_d)} \frac{B_{\nu_k}(T_{CMB})}{B_{\nu_0}(T_{CMB})^2}. \quad (8.15)$$
8.2.2 Spectral-Based Component Separation

Spectral Matching Independent Component Analysis (SMICA) is harmonic space component separation technique [111, 112]. Fundamentally, the technique is attempting to solve for $s_{lm}$, the $a_{lm}$ associated with the CMB component of the input maps.

8.2.2.1 Spectral-Based Component Separation Theory

I will describe the polarized implementation here. Given $N_{\text{chan}}$ input maps at different observation frequencies, we write

$$\begin{bmatrix} s_{lm}^{E} \\ s_{lm}^{B} \end{bmatrix} = W_{\ell}^{T} \begin{bmatrix} x_{lm}^{E} \\ x_{lm}^{B} \end{bmatrix}. \quad (8.16)$$

$W_{\ell}$ is a $2N_{\text{chan}} \times 2$ matrix of weights that describes how each channel contributes to the CMB at a given multipole. The frequency spectrum of the CMB is given by $a$, an $N_{\text{chan}} \times 1$ matrix, which should be thought of as a recalibration term. If the input channels were perfectly calibrated to each other, $a$ would be uniformly 1. The weights are given by

$$W_{\ell} = \left( A^{T} R_{\ell}^{-1} A \right)^{-1} A^{T} R_{\ell}^{-1} \quad (8.17)$$

where $R_{\ell}$ is the $2N_{\text{chan}} \times 2N_{\text{chan}}$ spectral covariance matrix and

$$A = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}. \quad (8.18)$$

We could use the natural sample estimate

$$\hat{R}_{\ell} = \frac{1}{2l + 1} \sum_{m} x_{lm} x_{lm}^{\dagger}. \quad (8.19)$$

However, this estimate is likely noise biased at large scales due to chance correlations. SMICA attempts to alleviate this problem by fitting a semi-parametric model to the data. The CMB,
8.2.2 Spectral-Based Component Separation

foregrounds, and noise are independent, so we can linearly decompose the correlation matrix

$$R_\ell = R_\ell^{\text{CMB}} + R_\ell^F + R_\ell^\text{noise}$$  \hspace{1cm} (8.20)$$

where $R_\ell^{\text{CMB}}$ is the component from the CMB, $R_\ell^F$ is from foregrounds, and $R_\ell^\text{noise}$ is from noise. The noise is assumed to be uncorrelated between different channels, so $R_\ell^\text{noise}$ is diagonal. The CMB and foregrounds are modeled as

$$R_\ell^{\text{CMB}} = A \begin{bmatrix} C^{EE}_\ell & C^{EB}_\ell \\ C^{EB}_\ell & C^{BB}_\ell \end{bmatrix} A^T$$  \hspace{1cm} (8.21)$$

$$R_\ell^F = \begin{bmatrix} F_E & 0 \\ 0 & F_B \end{bmatrix} P_\ell \begin{bmatrix} F_E & 0 \\ 0 & F_B \end{bmatrix}^T$$  \hspace{1cm} (8.22)$$

where $C_\ell$ is the CMB power spectrum, $F_E$ and $F_B$ are $N_{\text{chan}} \times d$ matrices and $P_\ell$ is a $2d \times 2d$ matrix, where $d$ is the number of foreground templates. $P_\ell$ describes how the foregrounds vary as a function of multipole, and $F$ describes the frequency spectra. This means that the frequency dependence of the foreground is constant across multipoles.

SMICA fits eq. 8.20 to eq. 8.19 by minimizing the spectral matching criterion

$$\phi(\theta) = \sum_\ell (2\ell + 1) \left[ \text{Tr} \left( \hat{R}_\ell R_\ell^{-1} \right) + \log \det R_\ell \right]$$  \hspace{1cm} (8.23)$$

where $\theta$ is a vector with all the fit parameters, or in other words all the elements of $R_\ell$. The fit can be performed with binned or unbinned $a_{lm}$s.

The fitting algorithm works in three steps.

1. Recalibrate the spectra by solving for $a$ in a clean patch of the sky. After this step, $a$ will be kept constant.

2. Fit the foreground emission spectrum, $F$ over a large sky patch using $l$ bins less than 150.

3. Fit for the CMB spectra, $C_\ell$s, and foreground spectra, $P_\ell$s
8.2.2 Spectral-Based Component Separation

8.2.2.2 A Pedagogical Example

It is instructive to write down the elements of the model covariance matrix, $R_\ell$. Consider an experiment with three frequencies ($N_{\text{chan}}=3$) and the SMICA algorithm solving for one foreground template ($d=1$). Then

$$A = \begin{bmatrix}
  a^{\nu_1} & 0 \\
  a^{\nu_2} & 0 \\
  a^{\nu_3} & 0 \\
  0 & a^{\nu_1} \\
  0 & a^{\nu_2} \\
  0 & a^{\nu_3}
\end{bmatrix} \quad (8.24)$$

where $\nu_i$ indexes the observation frequency. Plugging this into eq. 8.21, we get the matrix

$$R_{CMB}^\ell = \begin{bmatrix}
  a^{\nu_1}a^{\nu_1}C_{\ell}^{EE} & a^{\nu_1}a^{\nu_2}C_{\ell}^{EE} & a^{\nu_1}a^{\nu_3}C_{\ell}^{EE} & a^{\nu_1}a^{\nu_1}C_{\ell}^{EB} & a^{\nu_1}a^{\nu_2}C_{\ell}^{EB} & a^{\nu_1}a^{\nu_3}C_{\ell}^{EB} \\
  a^{\nu_2}a^{\nu_1}C_{\ell}^{EE} & a^{\nu_2}a^{\nu_2}C_{\ell}^{EE} & a^{\nu_2}a^{\nu_3}C_{\ell}^{EE} & a^{\nu_2}a^{\nu_1}C_{\ell}^{EB} & a^{\nu_2}a^{\nu_2}C_{\ell}^{EB} & a^{\nu_2}a^{\nu_3}C_{\ell}^{EB} \\
  a^{\nu_3}a^{\nu_1}C_{\ell}^{EE} & a^{\nu_3}a^{\nu_2}C_{\ell}^{EE} & a^{\nu_3}a^{\nu_3}C_{\ell}^{EE} & a^{\nu_3}a^{\nu_1}C_{\ell}^{EB} & a^{\nu_3}a^{\nu_2}C_{\ell}^{EB} & a^{\nu_3}a^{\nu_3}C_{\ell}^{EB} \\
  a^{\nu_1}a^{\nu_1}C_{\ell}^{EB} & a^{\nu_1}a^{\nu_2}C_{\ell}^{EB} & a^{\nu_1}a^{\nu_3}C_{\ell}^{EB} & a^{\nu_1}a^{\nu_1}C_{\ell}^{BB} & a^{\nu_1}a^{\nu_2}C_{\ell}^{BB} & a^{\nu_1}a^{\nu_3}C_{\ell}^{BB} \\
  a^{\nu_2}a^{\nu_1}C_{\ell}^{EB} & a^{\nu_2}a^{\nu_2}C_{\ell}^{EB} & a^{\nu_2}a^{\nu_3}C_{\ell}^{EB} & a^{\nu_2}a^{\nu_1}C_{\ell}^{BB} & a^{\nu_2}a^{\nu_2}C_{\ell}^{BB} & a^{\nu_2}a^{\nu_3}C_{\ell}^{BB} \\
  a^{\nu_3}a^{\nu_1}C_{\ell}^{EB} & a^{\nu_3}a^{\nu_2}C_{\ell}^{EB} & a^{\nu_3}a^{\nu_3}C_{\ell}^{EB} & a^{\nu_3}a^{\nu_1}C_{\ell}^{BB} & a^{\nu_3}a^{\nu_2}C_{\ell}^{BB} & a^{\nu_3}a^{\nu_3}C_{\ell}^{BB}
\end{bmatrix} \quad (8.25)$$

The upper left block is the EE component, the lower right block is the BB component, and the off diagonal blocks are EB. Within each block, there is the recalibration factor between two frequencies on off diagonal elements. If $a$ is uniformly one, meaning the channels are perfectly calibrated, then $R_{CMB}^\ell$ is simply the $C_\ell$s from the CMB.
The foreground covariance model is similar in form to the CMB covariance model

\[
F = \begin{bmatrix}
  f^{\nu_1} & 0 \\
  f^{\nu_2} & 0 \\
  f^{\nu_3} & 0 \\
  0 & f^{\nu_1} \\
  0 & f^{\nu_2} \\
  0 & f^{\nu_3}
\end{bmatrix},
\]

Again this is plugged into eq. [8.22], giving us

\[
R_F^\ell = \begin{bmatrix}
  f^{\nu_1}P_{EE}^\ell & f^{\nu_2}P_{EE}^\ell & f^{\nu_3}P_{EE}^\ell & f^{\nu_1}f^{\nu_2}P_{EB}^\ell & f^{\nu_1}f^{\nu_3}P_{EB}^\ell & f^{\nu_2}f^{\nu_3}P_{EB}^\ell \\
  f^{\nu_2}f^{\nu_1}P_{EE}^\ell & f^{\nu_2}f^{\nu_2}P_{EE}^\ell & f^{\nu_2}f^{\nu_3}P_{EE}^\ell & f^{\nu_2}f^{\nu_1}P_{EB}^\ell & f^{\nu_2}f^{\nu_2}P_{EB}^\ell & f^{\nu_2}f^{\nu_3}P_{EB}^\ell \\
  f^{\nu_3}f^{\nu_1}P_{EE}^\ell & f^{\nu_3}f^{\nu_2}P_{EE}^\ell & f^{\nu_3}f^{\nu_3}P_{EE}^\ell & f^{\nu_3}f^{\nu_2}P_{EB}^\ell & f^{\nu_3}f^{\nu_2}P_{EB}^\ell & f^{\nu_3}f^{\nu_3}P_{EB}^\ell \\
  f^{\nu_1}f^{\nu_1}P_{BB}^\ell & f^{\nu_1}f^{\nu_2}P_{BB}^\ell & f^{\nu_1}f^{\nu_3}P_{BB}^\ell & f^{\nu_1}f^{\nu_2}P_{EB}^\ell & f^{\nu_1}f^{\nu_2}P_{EB}^\ell & f^{\nu_1}f^{\nu_3}P_{EB}^\ell \\
  f^{\nu_2}f^{\nu_1}P_{BB}^\ell & f^{\nu_2}f^{\nu_2}P_{BB}^\ell & f^{\nu_2}f^{\nu_3}P_{BB}^\ell & f^{\nu_2}f^{\nu_1}P_{EB}^\ell & f^{\nu_2}f^{\nu_2}P_{EB}^\ell & f^{\nu_2}f^{\nu_3}P_{EB}^\ell \\
  f^{\nu_3}f^{\nu_1}P_{BB}^\ell & f^{\nu_3}f^{\nu_2}P_{BB}^\ell & f^{\nu_3}f^{\nu_3}P_{BB}^\ell & f^{\nu_3}f^{\nu_2}P_{EB}^\ell & f^{\nu_3}f^{\nu_2}P_{EB}^\ell & f^{\nu_3}f^{\nu_3}P_{EB}^\ell
\end{bmatrix}.
\]

Here it is easy to see that \( R_F^\ell \) for every \( \ell \) will have the same \( f^{\nu_i}f^{\nu_j} \) structure. The \( P_\ell \) terms allow the amplitude of the foreground contribution to vary by \( \ell \) and spectrum type (EE, BB, and EB).

\( R_{CMB}^\ell \) and \( R_F^\ell \) have identical forms, but the implementation of the fitting algorithm makes them non-degenerate in practice. The recalibration factors, \( a \), are solved for in the first step and then held constant, presumably at a value close to one. The remaining differences in the spectra are now accounted for by \( R_F^\ell \) only.

### 8.2.2.3 Spectral-Based Component Separation with Spider

With the theory in hand, we apply SMICA to the Spider and Planck maps. I use the Spider 95 and 150 GHz maps and the Planck 100, 143, 217, and 353 GHz maps [113]. The Spider noise model comes from the difference between the auto and cross spectra of four interleaved chunks.
The Planck noise model comes from the difference between the auto spectrum of the full mission map and cross spectra of half mission maps. I fit for one foreground template. So in the language of Section 8.2.2, \( N_{\text{chan}} = 6 \) and \( d = 1 \).

For this implementation, I skip the recalibration step because I find that it does not change the result. Moreover, SPIDER is already well calibrated to Planck by construction, as discussed in Chapter 6.

The spectra is first binned into bins of width 25 with the lowest edge of the lowest bin at \( \ell \) of 8. The frequency calibration is done by solving for \( F \) over all bins with \( \ell < 150 \). We omit the EB component because SPIDER maps currently have unaccounted for signal in EB that ruins the fit.

The fitting routine described by eq. 8.23 requires an initial guess to converge. For the CMB, I assume \( \Lambda \)CDM cosmology with \( r=0 \) and lensing. For the single foreground template, I assume a dust model with \( \beta_d = 1.59 \) and \( T_d = 19.6 \) K. These are merely initial guess parameters; the converged values do not depend strongly on them.

After we solve for \( F \), we then solve for \( P_\ell \) bin by bin. Now we can construct the weights according to eq. 8.17. The results are shown in Figure 8.11. Remember that these weights represent how much each map contributes to a measurement of the CMB sky. The CMB peaks at \( \sim 150 \) GHz, and Planck and SPIDER have the highest signal to noise maps at 143 and 150 GHz respectively. Thus the CMB spectrum is expected to be dominated by those two maps. Moreover, SPIDER has higher sensitivity than Planck at low \( \ell \) so it contributes more to the CMB spectrum. As \( \ell \) rises, the SPIDER beam cuts in, reducing SPIDER’s signal to noise, making the Planck contribution rise.

The Planck 353 GHz map’s contribution to the CMB spectrum is negative at \( \ell \lesssim 150 \) and zero above that. This is because the primary component of the 353 GHz map is dust at low multipoles. Therefore to construct the optimal CMB only map, we want to subtract the 353 GHz low multipole signal.

We can also look at the foreground component, described by \( R^F_\ell \). As mentioned before, the frequency dependence is fully described by the matrix \( F \). So we can derive the full frequency dependence just by looking at \( R^F_\ell \) at one \( \ell \). The EE and BB values are shown in Figure 8.12. The foreground model described by eq. 8.1 is plotted in a dashed black line with \( \beta_d = 1.53 \) and \( T_d = 21 \)
Figure 8.11: The SMICA frequency weights. The CMB measurement is dominated in lower $\ell$ by SPIDER maps because SPIDER has higher signal to noise than Planck in this region. At higher $\ell$, SPIDER’s beam reduces SPIDER’s signal to noise, so Planck’s weight increases. The Planck 353 GHz channel is negative at low $\ell$ and essentially zero at $\ell > 150$. This is consistent with dust being dominant at low multipoles.
8.2.2 Spectral-Based Component Separation

Figure 8.12: The SMICA foreground parameters, normalized by the amplitude of $R^F_{\ell}$ at 353 GHz. A foreground model with $\beta_d = 1.53$ is shown in the black line. It is renormalized with a free amplitude to fit the data. These should be seen as preliminary results, and will not be the final science result by SPIDER.

K. The model is fit to the data with one free amplitude parameter. These plots are preliminary, and will not be the final science result of SPIDER.
Chapter 9

Conclusion

As of the writing of this thesis, the SPIDER collaboration is still finalizing its analysis. The low-level analysis described in Chapter 5 is finalized. The calibrations are largely done as well, with the exception of possible small adjustments to the fiducial beam model. This will likely affect the per detector calibration as well.

The maps are passing null tests which are briefly described in Chapter 7 and fully described in [90]. The bulk of the work now is going towards understanding foregrounds and developing a robust power spectrum estimator. The difficulty in foregrounds is driven by chance correlations between CMB and foreground models and noise in the template maps. All the development is ongoing with low-ℓ BB spectra blinded.

SPIDER is successfully measuring temperature and polarization fluctuations, which is obvious from the foreground spectra in Figures 8.5 and 8.12. We can also generate beam stacks by combining hot and cold spots, shown in Figure 9.1. The blue ring around the center hot spot corresponds to a circular polarization pattern. This is the TE correlation.

SPIDER-2 is nearly ready to fly, with three telescopes at 280 GHz. The instrument is scheduled to launch in December 2019. With the high frequency channel, the combined SPIDER program will further constrain dust in its observation region.
Figure 9.1: Beam stacks for 95 GHz (top) and 150 GHz (bottom). The hot and cold spots in the CMB temperature map are stacked. The black lines show polarization.
Part V

Appendix
Appendix A

Amplified Noise Source

A.1 Design Overview

Using a thermal source to do beam mapping creates several difficulties, the greatest of which is the low power in the SPIDER observation band. If we consider the thermal source as a Rayleigh-Jeans source, then the power per unit area per frequency is

\[ I = \frac{2k_B T \nu^2}{c^2} \]  \hspace{1cm} (A.1)

where \( c \) is the speed of light, \( k_B \) is Boltzmann’s constant, \( \nu \) is the frequency, and \( T \) is the temperature. Then the power is just

\[ P = A \int_{\nu_0}^{\nu_f} I d\nu = A \frac{6k_B T}{c^2} (\nu_f^3 - \nu_0^3) \] \hspace{1cm} (A.2)

where \( A \) is the area of the radiative element. If we consider the frequency band between 80GHz and 180GHz for a hand-sized thermal source, we find the total power emitted by the thermal source at 500K is about 3\( \mu \)W. The amplified noise source is designed to provide about two orders of magnitude more power in a small, portable device.

The amplified noise source consists of a 50 \( \Omega \) resistor, four amplifiers, a waveguide filter, and a frequency tripler. A schematic drawing is shown in Figure A.1. The Johnson noise from the
A.2 Components

Figure A.1: A schematic of the amplified noise source. From the left to right: the 50 Ω resistor, three amplifiers, a waveguide cutoff, another amplifier, and finally the tripler. The source outlined here produces a total output power of approximately 150 µW.

resistor is fed through three stages of amplifiers. This nearly saturates the third amplifier. The signal then goes through a waveguide filter which attenuates out-of-band signal, lowering the total power. The filtered signal then goes into the final amplifier which should saturate. Finally the power is sent into a tripler, which as the name suggests triples the frequency of the signal and broadcasts it.

A.2 Components

The amplifiers are Marki A2050. The gain is approximately 23 dB across the 20 to 50 GHz band. The saturation power is 21 dBm. The waveguide is a standard WR-19. It has a cutoff frequency of 31 to 54 GHz. It has a small bandpass loss of .8 dB and a large stop band rejection of 45 dB. The tripler is a Virginia Diodes WR6.5X3. Figure A.2 shows the efficiency of the tripler as a function of output frequency. It is crucial to power the tripler before the amplifiers. If signal is input into the tripler before it is powered, the tripler components may fry.

The power from Johnson noise of a resistor is

\[ P = 4k_BT\Delta f \]  (A.3)

where \( \Delta f \) is the frequency band. For this setup, the resistor is at room temperature, 300 K. The relevant \( \Delta f \) is 30 GHz. So the total power in band is 497 pW. The signal goes through three gain stages.
A.2 Components

Figure A.2: The efficiency of the tripler as a function of output frequency [114]. Note that the efficiency is not flat across the band. This may be important in analysis depending on the application.

stages, each increasing the power by 23 dB, so the total power is 3.9 mW (5.9 dBm). The third amplifier is nearly saturated. The signal then is filtered, dropping the power by approximately 3/5. More importantly, it attenuates the out-of-band signal so that when it is fed into the final amplifier stage, there is only signal in the desired frequency bands. The final stage amplifier is saturated primarily by power between 31 and 54 GHz (set by the waveguide filter) and outputs a signal at 21 dBm. The signal is then fed into the tripler which is approximately 5% efficient in band. Then the final power output should be 6.3 mW. The actual measured output is 150 µW, suggesting loss in the system that is not accounted for in the above calculation.
Appendix B

Carbon and Steel Stycast

B.1 Power Attenuation

The optical performance of the 2K baffles is a strong function of its reflectivity. One suggestion to reduce reflectivity is to coat the 2K baffles with carbon and steel impregnated stycast. Zemax simulations indicate acceptable sidelobes for SPIDER require relatively high attenuation in the stycast. However if the coating is too thin (less than one wavelength), it may not act as an absorber.

We built new optics sleeves in February 2014 and coated them with a stycast mixture with 5% lampblack and 5% steel by weight. This layer is measured to be about 0.02” (0.5 mm) thick. The old iteration of optics sleeves is measured to be about 0.04” (1 mm) thick with a 7% carbon and 6.5% steel mix. Both samples are also textured, the new ones using a paintbrush and the old ones using a comb. We were able to create larger ridges with the paintbrush method.

To test the reflectivity of impregnated stycast, I used a 90 GHz IMPATT source and 90 GHz diode detector. I built a V bracket to hold them at fixed angles. The full setup, shown in Figure B.1 provided repeatable results.

I then made plates of stycast with 5% lampblack and 5% steel by weight of varying thickness by first applying an overly thick layer and then milling it down to the desired thickness. This created a very uniform, flat surface. Then on two of the samples, I added texture using a paintbrush in the
B.1 Power Attenuation

Figure B.1: On the left is the IMPATT source. On the right is the detector diode, which is read out with a digital volt meter. The sample being measured is in the center well. There is eccosorb lining the sides of the center well to minimize stray reflections.

same manner as was done on the sleeves. I used Stycast 2650FT with catalyst 23LV. The sample was baked in a 80 C oven for 2 hours to set. The samples are shown in Figure B.2

The results below are reported in mV. The detector diode has a linear response between measured signal and voltage. In this part of the experiment, I took care to make sure the aluminum plates were always at the same height. This means that as the layer of stycast got thicker, the term that was reflecting off the stycast would be more and more out of alignment. However this test will always do a good job of measuring the reflection off the aluminum.

Figure B.2: The stycast plates are shown from a side view and top view. From left to right are 0.02”, .04”, .09”, .19”, .02”+bumps, and .04”+bumps.
B.1 Power Attenuation

<table>
<thead>
<tr>
<th>Sample</th>
<th>Voltage (mV)</th>
<th>Std (mV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bare Al.</td>
<td>29.5</td>
<td>0.7</td>
</tr>
<tr>
<td>0.02”</td>
<td>30.7</td>
<td>2.0</td>
</tr>
<tr>
<td>0.04”</td>
<td>27.1</td>
<td>1.5</td>
</tr>
<tr>
<td>0.09”</td>
<td>8.6</td>
<td>2.4</td>
</tr>
<tr>
<td>0.19”</td>
<td>4.6</td>
<td>0.7</td>
</tr>
<tr>
<td>0.02” + bumps</td>
<td>23.5</td>
<td>3.3</td>
</tr>
<tr>
<td>0.04” + bumps</td>
<td>14.5</td>
<td>2.5</td>
</tr>
<tr>
<td>Eccosorb</td>
<td>0.005</td>
<td></td>
</tr>
</tbody>
</table>

Table B.1: The reflectivity measurements for various samples of impregnated stycast. The attenuation is minimal from thin layers of stycast, which is consistent with expectations. The bumps make reflections more specular, also minimizing measured reflections.

The first test was done at an incident angle of 34 degrees from normal. I took 5 measurements of each sample, rotating through all of them. The samples labeled “0.02 + bumps” and “0.04 + bumps” were samples that were first machined down to 0.02” and .04” respectively and then texture was added. Thus, the minimum thickness is probably close to 0.01” greater than the listed value.

The results are shown in Table B.1 and remarkably physics seems to work. Thin layers of stycast mix seem to do absolutely nothing. The thickness must approach the wavelength before any attenuation occurs. When attenuation starts happening, it turns on dramatically.

The bumps seem to help. The comparison between 0.02” and 0.02” + bumps suggests that adding bumps scatters about 25% of the power. More interestingly, the 0.04” + bumps seems to do quite well. I propose that is because in addition to scattering, you also have regions that are thick enough to absorb.

The 34° incident angle is much higher than the expected value in SPIDER. Side-lobe photons will come in at a shallower angle, meaning the photons will travel through more stycast mix before it reflects off the underlying aluminum. This setup allows us to vary the angle of incidence too. The results are shown in Table B.2. The “ratio” column is the measured voltage divided by the voltage measured for bare aluminum at a given angle. In other words, it is the fraction of power that is absorbed or scattered by the stycast for a given incidence angle.
Table B.2: The reflectivity as a function of incident angle for all the samples. “Ratio” refers to the measured voltage divided by the voltage for bare aluminum. It should be thought of as a measure of how much incident power is attenuated by the stycast.

### B.2 Index of Refraction

In this section I measured the reflection off the stycast surface. I took the ratio of power observed with an aluminum plate versus a thick stycast mix. I use the 0.19” sample because it is the thickest, meaning most of the signal is attenuated before it hits the aluminum, so the only detected signal is from reflections off the stycast itself. This time I made sure the top surface of the aluminum plate was at the same height as the top surface of the stycast sample. This ensures that we are properly aligned to measure the reflection off stycast. I then took measurements at various angles of incidence.

Note that both the source and detector are polarized. In this setup, the polarization is parallel to the plane defined by the incident and reflected rays. From Fresnel’s equations we know the reflectivity for the “p-polarization” goes as

$$R_p = \left| \frac{n_1 \cos \theta_t - n_2 \cos \theta_i}{n_1 \cos \theta_t + n_2 \cos \theta_i} \right|^2$$

where $\theta_i$ and $\theta_t$ are the incident and transmission angles respectively and $n_1$ and $n_2$ are the index of refraction for the two media. The conventional wisdom said that steelcast has an index of about 3; however, an index of 2.3 matches much better.
B.2 Index of Refraction

Figure B.3: The transmission and reflection amplitude between \( n_1 = 1 \) and \( n_2 = 2.3 \) media. The data matches well an index of 2.3.
Appendix C

Derivative Maps

This section describes how to generate derivative maps used for deprojection described in Section 6.2. Below is a code snippet briefly detailing how to make derivative maps. I use the custom package `spider_analysis` to load a temperature-only Planck map. The `spider_analysis` package is simply a wrapper around `healpy`. Note that converting maps to \( a_{lm} \)s and calculating their derivatives are slow operations.

```python
import healpy as hp
import spider_analysis as sa

# Load raw Planck map
raw_map = sa.map.read_map('/path/to/map/planck_map.fits')

# Convert to alms and take derivative
alm = hp.map2alm(raw_map)
raw_map, dt, dp = hp.alm2map_der1(alm, 2048)
alm_theta = hp.map2alm(dt)
alm_phi = hp.map2alm(dp)
dt, dt_dt, dt_dp = hp.alm2map_der1(alm_theta, 2048)
dp, dp_dt, dp_dp = hp.alm2map_der1(alm_phi, 2048)
```
For the spider_analysis pipeline, the derivative maps are precomputed and stored to disk. They are then read using the nominal detector timeline to get time-domain templates. Remember that the relations between physics angles and RA and DEC are

$$\theta = 90 - DEC, \quad \phi = RA.$$  

The derivative maps for X1 are shown in Figures C.1 and C.2. They are generated from Planck 143 GHz intensity maps, smoothed with the Spider beam. The derivatives are taken in galactic coordinates but shown in celestial coordinates, making the poles obvious.
Figure C.1: The beam smoothed *Planck* map in the SPIDER region is shown in the top panel. The middle and bottom panels are the derivatives with respect to $\theta$ and $\phi$ respectively. Note that the derivatives are taken in galactic coordinates and shown in celestial coordinates for ease of understanding.
Figure C.2: The second derivatives of the beam smoothed Planck in the SPIDER region. Note that the derivatives are taken in galactic coordinates and shown in celestial coordinates for ease of understanding.
Appendix D

Plotting Beam Ellipticity

The method outlined in Section 6.2 solves for the ellipticity as a function of $p$ and $c$, the “plus” and “cross” polar terms. Most plotting algorithms want values of the semi-major and semi-minor axis, $\sigma_a$ and $\sigma_b$ respectively. For the generic beam profile in eq. 6.23, we can express the equation as $\theta$, $\sigma_a$, and $\sigma_b$ given by

$$
\theta = \frac{1}{2} \arctan \left( \frac{c}{p} \right)
$$

$$
\sigma_a^2 = \frac{\sigma^2 (1 - p^2 - c^2)}{2(1 - \sqrt{p^2 + c^2})}
$$

$$
\sigma_b^2 = \frac{\sigma^2 (1 - p^2 - c^2)}{2(1 + \sqrt{p^2 + c^2})}
$$

where $\theta$ is defined as a counter-clockwise rotation. Figure D.1 shows a schematic drawing. Also note that implementation of this requires the use of the two parameter arctan function.*

It is helpful to then define eccentricity

$$
e = \sqrt{1 - \left( \frac{\sigma_b}{\sigma_a} \right)^2}
$$

$$
= \sqrt{1 - \left( \frac{1 - \sqrt{p^2 + c^2}}{1 + \sqrt{p^2 + c^2}} \right)}
$$

* In python, this is np.atan2.
Figure D.1: A drawing of an elliptical beam with semi-major, semi-minor axis, and rotation.

or alternatively ellipticity

\[
\epsilon = \frac{\sigma_a - \sigma_b}{\sigma_a + \sigma_b} = \frac{(1 + \sqrt{p^2 + c^2})^{1/2} - (1 - \sqrt{p^2 + c^2})^{1/2}}{(1 + \sqrt{p^2 + c^2})^{1/2} + (1 - \sqrt{p^2 + c^2})^{1/2}} = \frac{1 - \sqrt{1 - p^2 - c^2}}{\sqrt{p^2 + c^2}}.
\]
Appendix E

Explorer’s Notes

These are some things that either made my life better or would have if I knew/had them. This is largely based on my experience recovering SPIDER with British Antarctic Survey (BAS), but at least some of it applies to deployment as well.

Things I wish I knew

- Don’t be afraid to pack heavy. BAS in particular travels with tons of things. If you pack an extra pair of shoes or a couple extra layers, no one will ever notice. It’ll be a minor inconvenience traveling to Antarctica, but your life will be better when you get there.

- Be prepared to get very bored. You’re going to be sitting around a lot. Books are great, Kindles are better. Both Rothera and McMurdo have ebooks of questionable legality on hard drives which can be loaded onto Kindles.

- Get strong. When you’re out in the field, you’ll spend far more time than you’d like shoveling snow. It’s hard work and you’ll get tired fast. Walking across loose snow is tiring.

- You’re going to get cold. To warm up, shovel snow. Good news is that there’s lots of snow to shovel.

- Only bring things with a twist cap. My toothpaste had a snap-on cap, which immediately
shattered in the cold. I had to deal with a leaky toothpaste tube the entire trip. Everything now smells like peppermint.

- Sleep with your contacts in your sleeping bag or they will be frozen in the morning.
- Learn to tie knots. At the very least, know how to tie a trucker’s hitch.

**Things I am glad or wish I had**

- Multiple pairs of gloves – At least two pairs for working (you may lose or break one) and a pair of gauntlets.
- Multiple pairs of socks – I had five pairs of long SmartWool socks.
- Warm hat – They say you lose 90% of your heat from your head. That’s a lie. But your head will still get cold.
- Neck protection – There are million different types that you can find. It makes a huge difference, especially when the wind is blowing.
- Sunglasses – Sun is bright.
- Multiple layers – Things tend to go from very hot to very cold very quickly. Heaters in huts are incredibly powerful, so you spend an inordinate amount of time trying to cool rooms in Antarctica. I wish I had an extra base layer and a thin fleece of my own. USAP and BAS both provide base layers if you want them, but remember these are designed to fit everyone, so they’re generally oversized. I found the fleeces they provided to be way too baggy. Wind would blow underneath them all the time.
- Wet wipes – There are basically no showers out in the field. This is the closest you’ll get.
- Water bottle – Nalgenes are popular because they don’t tend to break if you accidentally freeze your water, which you will.
• Ethernet jack adapter – New Macs don’t have ethernet jacks. Make sure to bring the right adapter.

• Power adapter – Remember that UK and Chile have different plugs. Also some adapters don’t support grounding pins, so make sure you have one that works with the equipment you’re bringing.

• Large-ish dry bag – Things outside get very cold. Huts tend to be very humid because people are cooking/boiling water. This means water condenses all over things until they warm up. Put your electronics in a dry bag so they warm up in dry air. If you don’t want to get a dry bag, large ziplock bags work pretty well too.

• Noise canceling headphones – The planes are very loud. All the pilots wear them and so should you.

• Multiple copies of important tools – Things get lost in the snow easily. We lost a two foot long adjustable wrench in the snow after 20 minutes of use. Allen keys just disappear. If I redid our recovery, I would bring 6-8 sets of Allen keys.

• As much fresh fruit, alcohol, and chocolate as you can carry into Antarctica. It’s the best way to make friends.
Figure E.1: From left to right: Chucky, Cheese, Tim, and Sam. Steve is unfortunately not in this photo. All are members of the British Antarctic Survey. The SPIDER recovery would have been impossible without them. The SPIDER payload is open to remove the inserts before shipping the cryostat back on the land train, driven by Tim.
Appendix F

IR Dewar

Wilbur is an IR Labs open cycle, wet dewar that can be cooled to 2K with liquid helium, and can support a millikelvin helium-3 fridge. Because of its small size, it has a quick turnaround time, allowing for rapid testing. In the past, Wilbur has been used to calibrate diodes, measure transition temperatures of transition edge sensors, and measure thermal conductivity of cryogenic materials. This is a brief guide to using the Princeton IR dewar. There is no substitute for having an experienced hand teach you in person.

F.1 The Cryostat

Figure F.1 shows Wilbur in cross-section. There are three key parts, the vacuum vessel: the 77 K stage, and the 4 K stage. Note that the configuration when the cryostat is on a table is upside-down. Figure F.1 shows Wilbur in its running configuration; we will discuss it in reference to the orientation shown.

Vacuum vessel

The stock vacuum vessel has an aluminum extension to allow for a larger work space. This was originally necessary to mount the relatively large helium-3 fridges for SPIDER. The vacuum vessel
Figure F.1: A cross-section of the IR Dewar. This is in the running configuration. All test devices are mounted to the baseplate. The 77 K and 4 K stages are filled manually with a funnel.
also has a port, which can be used to optically couple to a test device inside the cryostat if desired.

77 K stage

The 77 K stage sits above the 4 K stage and has a 4.2 L capacity. The 77 K stage has a shield which extends around the 4 K stage and shield, which is conductively cooled. When closing up, it is important to screw down and tape the shields well, since this the only conductive path for the shields. Additionally it is important to cover up light leaks. Unfortunately the shields are somewhat poorly built, so only a subset of the screws can go in, but try your best.

4 K stage

The 4 K stage has a 4.4 L capacity. The term “4 K” is a bit of a misnomer, since it will first go to 77 K, then 4 K, then finally 2 K. This will be discussed more in Appendix F.2. The fill line for the 4 K stage runs through the center of the 77 K stage. The 4 K stage also has a shield which is conductively cooled. It is attached in the same manner as the 77 K shield. The spacing between the 4 K and 77 K stage shields is incredibly small, making it very easy to have a touch that creates a thermal short. Even a small touch between these two shields will make it impossible to cool down the cryostat.

The baseplate of the 4 K stage is where everything will mount. The baseplate presents a square hole pattern of #4 screws. No matter what test you are doing, you must have a diode on the 4 K baseplate.

F.2 Cooling Down

Pumping

Due to its small volume, pumping out Wilbur takes only a few hours, but it is best to leave it on a turbo pump overnight. Typically I use the Tri-Scroll to rough out the cryostat for a few hours then switch over to the small turbo overnight.

* As far as I know, this has never been done with this dewar.
F.2 Cooling Down

Nitrogen

Fill the 77 K stage and 4 K stage with liquid nitrogen. This is done with a funnel and an external liquid nitrogen bucket. It will be obvious when the tank is full because nitrogen will be spurting out of the fill port. It doesn’t seem to matter which stage is filled first. Now wait until the stages equilibrate to 77 K. How long this takes depends on what is mounted to the 4 K stage.

Helium

There will undoubtedly be LN2 in the 4 K tank when it is time to fill LHe. To get rid of it, invert the cryostat so the excess nitrogen pours out onto the floor (or a bucket). This also empties out the 77 K tank. Refill the 77 K tank with LN2.

Now it is time to fill the 4 K tank with LHe. Drop the stinger into the LHe dewar. Once there is liquid coming out of the transfer line, drop the transfer line into the fill port. Make sure to use the center fill port that connects to the 4 K tank and not the 77 K tank! They are very close to each other and look very similar. If you are recovering helium during the transfer, be careful to listen for the clicking of the recovery counter. There will be a sudden change in the frequency of the clicking, indicating that the fill is complete. If you are not recovering helium, watch the plume coming out of the fill port. Liquid helium will be coming out in a large plume. Stop the transfer and remove the stinger.

The 4 K stage will take a few hours to equilibrate, again depending on how much thermal mass is mounted to it. Once it is equilibrated, refill the 77 K tank with nitrogen and the 4 K tank with helium. Now the system is ready to cool down.

Attach the pump-out manifold to the 4 K fill port. Attach the other end of the manifold to a scroll pump. Begin to pump on the 4 K bath, starting very slowly. Often, oscillations can set up, which can dump power into your 4 K bath, making it impossible to cool down. Try backing off the pumping power. If this does not work, sometimes pumping through a large volume can damp oscillations. If all goes well, the stage should be at 2 K in less than two hours, and should hold for about 48 hours.
Warming Up

To warm up, back fill the 4 K stage with gaseous helium at 1 atmosphere pressure. When it’s at atmospheric pressure, disconnect all the fittings. Dump all the cryogens out of the cryostat by inverting the cryostat. Without cryogens, the 4 K baseplate should begin warming immediately.
Appendix G

Fluctuations in an Expanding Universe

This is a set of notes I wrote with Herman Verlinde for my prethesis about how fluctuations evolve in an expanding universe.

G.1 Minkowski Metric

We want to know the spectrum of fluctuations in a deSitter universe (one that grows exponentially). But before we do that, we will derive the spectrum of fluctuations in a Friedmann universe and use the results we get there to help understand the deSitter solution. We start with the spatially flat Friedmann universe, where we write the metric as

\[ ds^2 = dt^2 - a(t)^2 \delta_{ik}dx^i dx^k. \]  

(G.1)

It is convenient to work in conformal time, which we defined in eq. [1.15]. This gives

\[ ds^2 = a(\eta)^2 [d\eta^2 - \delta_{ik}dx^i dx^k] \]  

(G.2)
which is conformally equivalent to the Minkowski metric. Now we can write down the minimally coupled massive scalar field with action

\[ S = \frac{1}{2} \int \sqrt{-g} d^4x \left( g^{\alpha \beta} \phi_{,\alpha} \phi_{,\beta} - m^2 \phi^2 \right). \quad (G.3) \]

We can make the substitution \( g^{\alpha \beta} = a^{-2} \eta^{\alpha \beta} \), where \( \eta^{\alpha \beta} \) is the Minkowski metric, and \( \sqrt{-g} = a^4 \). Then the action becomes

\[ S = \frac{1}{2} \int d^3x d\eta a^2 \left[ \phi'^2 - (\nabla \phi)^2 - m^2 a^2 \phi^2 \right] \quad (G.4) \]

where we have defined terms with ‘ to mean derivatives with respect to \( \eta \). For even greater simplification, we can introduce the auxiliary field

\[ \chi \equiv a(\eta)\phi \quad (G.5) \]

which gives

\[ \frac{\partial \chi}{\partial \eta} = \frac{\partial a(\eta)}{\partial \eta} \phi + a(\eta) \frac{\partial \phi}{\partial \eta} \quad (G.6) \]

or reorganizing terms

\[ a(\eta)\phi' = \frac{\partial \chi}{\partial \eta} - \frac{\partial a(\eta)}{\partial \eta} \phi \]

\[ = \frac{\partial \chi}{\partial \eta} - \frac{1}{a(\eta)} \frac{\partial a(\eta)}{\partial \eta} \chi. \quad (G.7) \]

Of course \( a(\eta) \) is independent of spatial coordinates, so

\[ \nabla \chi = a(\eta) \nabla \phi. \quad (G.8) \]
From here on out, we will assume that $a$ is always a function of $\eta$ unless otherwise specified. We can plug the expression of $\chi$ into eq. (G.4) allowing us to simplify our equation for the action.

\[
S = \frac{1}{2} \int d^3x d\eta \left[ \left( \frac{\partial \chi}{\partial \eta} - \frac{1}{a} \frac{\partial a}{\partial \eta} \chi \right)^2 - (\nabla \chi)^2 - m^2 a^2 \chi^2 \right]
\]

\[
= \frac{1}{2} \int d^3x d\eta \left[ \chi'^2 - (\nabla \chi)^2 - m^2 a^2 \chi^2 + \left( \frac{a'}{a} \chi \right)^2 - 2 \frac{a'}{a} \chi' \chi \right] \quad (G.9)
\]

We can simplify this even further by noting that

\[
\frac{\partial}{\partial \eta} \left( \frac{a'}{a} \chi^2 \right) = 2 \frac{a'}{a} \chi' \chi + \frac{a''}{a} \chi^2 - \frac{a'^2}{a^2} \chi^2.
\]

\[
(G.10)
\]

This allows us to write

\[
S = \frac{1}{2} \int d^3x d\eta \left[ \chi'^2 - (\nabla \chi)^2 - m^2 a^2 \chi^2 \right] - \frac{a''}{a} \chi^2 \quad (G.11)
\]

where we have omitted the total derivative term. From direct comparison, we can see that the effective mass is

\[
m^2_{\text{eff}} \equiv m^2 a^2 - \frac{a''}{a}.
\]

\[
(G.12)
\]

The action can then be written as

\[
S = \frac{1}{2} \int d^3x d\eta \left[ \chi'^2 - (\nabla \chi)^2 - m^2_{\text{eff}} \chi^2 \right] \quad (G.13)
\]

or alternatively the Lagrangian, $\mathcal{L}$, is

\[
\mathcal{L} = \chi'^2 - (\nabla \chi)^2 - m^2_{\text{eff}} \chi^2.
\]

\[
(G.14)
\]

Now the field $\chi$ functions as a massive scalar field in Minkowski spacetime with the addition of a time dependent effective mass. Use the variational principle to get

\[
\chi'' - \nabla^2 \chi + m^2_{\text{eff}} \chi = 0.
\]

\[
(G.15)
\]
Now we want to expand $\chi$ into its Fourier modes,

$$\chi(x, \eta) = \int \frac{d^3k}{(2\pi)^{3/2}} \chi_k(\eta) e^{ik \cdot x}. \quad (G.16)$$

Plug the Fourier expanded $\chi$ into eq. \textbf{G.15}. The time derivative of $\chi(x, \eta)$ in the Fourier domain only acts on $\chi_k$ and the spatial derivatives only act on the exponential.

$$\chi''(x, t) = \int \frac{d^3k}{(2\pi)^{3/2}} \chi''_k(\eta) e^{ik \cdot x} \quad (G.17)$$

$$\nabla^2 \chi(x, t) = -k^2 \int \frac{d^3k}{(2\pi)^{3/2}} \chi_k(\eta) e^{ik \cdot x} \quad (G.18)$$

Immediately, we can see that

$$\chi'' + \left( k^2 + m_{\text{eff}}^2 \right) \chi_k = 0. \quad (G.19)$$

In complete analogy to the Klein-Gordon problem, we can write

$$\omega_k^2(\eta) = k^2 + m_{\text{eff}}^2 = k^2 + m^2 a^2 - \frac{a''}{a} \quad (G.20)$$

simplifying our equation of motion in the Fourier expansion to

$$\chi'' + \omega_k^2 \chi_k = 0. \quad (G.21)$$

The general solution to this differential equation is

$$\chi_k(\eta) = \frac{1}{\sqrt{2}} \left[ a_k^- v_k^*(\eta) + a_k^+ v_k(\eta) \right] \quad (G.22)$$

where $a_k^+$ and $a_k^-$ are the creation and annihilation operators associated with the field $\chi$ with frequency $k$. We plug this equation into eq. \textbf{G.16} finding

$$\chi(x, \eta) = \frac{1}{\sqrt{2}} \int \frac{d^3k}{(2\pi)^{3/2}} \left[ a_k^- v_k^*(\eta) e^{ik \cdot x} + a_k^+ v_k(\eta) e^{-ik \cdot x} \right] \quad (G.23)$$
Now we have the tools to solve for the amplitude of quantum fluctuations. Consider the equal time correlation function for a vacuum state.

\[
\langle 0 | \chi(x, \eta) \chi(y, \eta) | 0 \rangle = \frac{1}{2} \int \frac{d^3k d^3k'}{(2\pi)^3} \langle 0 | \left[ a_k^- a_{k'}^+ v_k^* v_{k'}^* e^{i k \cdot x + i k' \cdot y} + a_k^- a_{k'}^+ v_k^* v_{k'}^* e^{i k \cdot x - i k' \cdot y} + a_k^+ a_{k'}^- v_k^* v_{k'}^* e^{i k \cdot x - i k' \cdot y} + a_k^+ a_{k'}^- v_k^* v_{k'}^* e^{i k \cdot x + i k' \cdot y} \right] | 0 \rangle
\]  
(G.24)

We eliminate the first, third, and fourth terms in the bracket because applying the lowering operators to the vacuum state returns 0 and applying the raising operator twice gives \( \langle 0 | 2 \rangle \), which is also 0. Additionally, we know that \( k \) must equal \( k' \) from the commutation relation

\[
\left[ a_k^-, a_{k'}^+ \right] = \delta(k - k')
\]  
(G.25)

where \( \delta \) is the Dirac delta function. This dramatically simplifies the correlation function to

\[
\langle 0 | \chi(x, \eta) \chi(y, \eta) | 0 \rangle = \frac{1}{2} \int \frac{d^3k}{(2\pi)^3} | v_k^* |^2 e^{i k \cdot (x - y)}
= \frac{1}{2} \int_0^{2\pi} d\theta \int_0^\infty \frac{dk}{(2\pi)^3} (4\pi) k^2 | v_k |^2 \sin(kL)
\]  
(G.26)

where \( L = |x - y| \). Finally integrating of \( \theta \) by substitution, we get the result

\[
\langle 0 | \chi(x, \eta) \chi(y, \eta) | 0 \rangle = \frac{1}{4\pi^2} \int_0^\infty dk k^2 | v_k(\eta) |^2 \frac{\sin(kL)}{kL}
\]  
(G.27)

The main contribution to the integral comes from wavenumbers \( k \approx L^{-1} \), so the correlation function is estimated as

\[
\langle 0 | \chi(x, \eta) \chi(y, \eta) | 0 \rangle \sim k^3 | v_k |^2.
\]  
(G.28)

Another way to characterize typical fluctuations of scale \( L \) is to calculate the field operator
averaged over some region. The fluctuation is defined as

$$\delta \chi^2_L(\eta) \equiv \langle \psi | [\chi_L(\eta)]^2 | \psi \rangle$$

where

$$\chi_L(\eta) \equiv \int d^3x \chi(x, \eta) W_L(x).$$

$W_L(x)$ is a window function that is nearly 1 for $x < L$ and rapidly drops off for $|x| \gg L$. The most common example of a window function is a Gaussian

$$W_L(x) = \frac{1}{(2\pi)^{3/2}L^3} \exp \left( -\frac{|x|^2}{2L^2} \right).$$

As always, we want to work in the Fourier domain, so define the Fourier transform of the window function

$$w(kL) = \int W_L(x) e^{-ikx} d^3x.$$  

Doing the exact same algebra as we did before, we find that the fluctuation is

$$\delta \chi^2_L(\eta) = \frac{1}{2} \int \frac{d^3k}{(2\pi)^3} |v_k(\eta)|^2 |w(kL)|^2.$$  

By construction, this function decays quickly as $k < L^{-1}$. Thus the integral can be approximated as

$$\delta \chi^2_L(\eta) \sim \int_0^{L^{-1}} dk k^2 |v_k(\eta)|^2$$

$$\sim \frac{1}{L^3} |v_k(\eta)|^2$$

$$\sim k^3 |v_k(\eta)|^2.$$  

This result scales exactly the same as the solution to the correlation function. Therefore we define

$$\delta(k) \equiv \frac{1}{2\pi} k^{3/2} |v_k|$$

$$\text{(G.35)}$$
as the spectrum of quantum fluctuations in a static Minkowsky metric.

### G.2 deSitter Metric

Now we want to consider a scalar field, \( \phi \), in a deSitter universe. We demonstrated that a simple slow roll inflation creates an exponentially growing Universe. In this Universe, the scale factor evolves as

\[
a(t) = a_0 e^{H \Lambda t}. \tag{G.36}
\]

If the Universe is homogeneous and isotropic with a positive cosmological constant \( \Lambda \), then it is called a deSitter universe. We know that energy density is constant with time; thus, the Hubble parameter for a \( \Lambda \) dominated universe is

\[
H_\Lambda = \sqrt{\frac{8\pi G}{3} \rho_\Lambda}. \tag{G.37}
\]

This allows us to write the deSitter metric

\[
ds^2 = dt^2 - H_\Lambda^{-2} \exp (2H_\Lambda t) \delta_{ij} dx^i dx^j \tag{G.38}
\]

where we set \( a_0 = H_\Lambda^{-1} \). As in the Minkowsky case, we find the conformal time

\[
\eta = - \int_{-t_a}^{\infty} \frac{dt}{a(t)} = - \exp (-H_\Lambda t) \tag{G.39}
\]

and rewrite the scale factor in terms of \( \eta \)

\[
a(\eta) = - \frac{1}{H_\Lambda \eta}. \tag{G.40}
\]

This allows us to write

\[
ds^2 = a(\eta)^2 \left( d\eta^2 - \delta_{ij} dx^i dx^j \right) \tag{G.41}
\]
G.2 deSitter Metric

which is identical to the form we derived in the Appendix G.1. We define an auxiliary field in the exact same manner. But now we must plug in the deSitter scale factor instead of the Minkowsky scale factor, giving us

$$\omega_k^2(\eta) = k^2 + \left(\frac{m^2}{H_\Lambda^2} - 2\right) \frac{1}{\eta^2}.$$  \hspace{1cm} (G.42)

and the mode function

$$v''_k + \omega_k^2 v_k = v''_k + \left[k^2 + \left(\frac{m^2}{H_\Lambda^2} - 2\right) \frac{1}{\eta^2}\right] v_k = 0.$$ \hspace{1cm} (G.43)

The solution to this differential equation involves the Bessel functions $J_n(x)$ and $Y_n(x)$

$$v_k(\eta) = \sqrt{k|\eta|} [A_k J_n(k|\eta|) + B_k Y_n(k|\eta|)]$$ \hspace{1cm} (G.44)

with

$$n = \sqrt{\frac{9}{4} - \frac{m^2}{H_\Lambda^2}}.$$ \hspace{1cm} (G.45)

While the mode equations are not trivial to understand, the asymptotic behavior can be found directly from eq. (G.43). For early time, $k|\eta| \gg 1$, meaning the wavelength of the associated mode

$$L_p \sim a(\eta) k^{-1} \approx \frac{H_\Lambda}{k|\eta|}$$ \hspace{1cm} (G.46)

is much smaller than the curvature scale, $H_\Lambda^{-1}$. In other words, the waves function as if they were in Minkowski space. Using the fact that $k|\eta| \gg 1$, eq. (G.43) simplifies to

$$v''_k + k^2 v_k = 0$$ \hspace{1cm} (G.47)

which has the solution

$$v_k = e^{ik\eta}.$$ \hspace{1cm} (G.48)

As the universe expands, $|\eta|$ decreases. So for a given $k$ with $k|\eta|$ initially greater than one, it eventually becomes smaller than unity. This process is called event horizon crossing. This
corresponds to a mode of length $L$ (or equivalently $k$) crossing from subhorizon to superhorizon scales. Superhorizon modes are of those that have $k|\eta| \ll 1$. After horizon crossing, eq. (G.43) can be simplified to

$$v''_k - \left(2 - \frac{m^2}{H^2_\Lambda}\right) \frac{1}{\eta^2} v_k = 0 \quad (G.49)$$

that has the general solution

$$v_k(\eta) = A_k|\eta|^{n_1} + B_k|\eta|^{n_2} \quad (G.50)$$

where

$$n_{1,2} = \frac{1}{2} \pm \sqrt{\frac{9}{4} - \frac{m^2}{H^2_\Lambda}}. \quad (G.51)$$

Ultimately we would like to solve for the evolution of fluctuations that start subhorizon and that are then stretched to be superhorizon. To guarantee that the mode starts subhorizon, we must require that

$$v_k(\eta) \to \frac{1}{\sqrt{\omega_k}} e^{i\omega_k \eta} \quad (G.52)$$

as $\eta \to -\infty$. The specific form of eq. (G.50) that satisfies this condition and has the proper normalization is

$$v_k(\eta) = \sqrt{\frac{\pi |\eta|}{2}} \left[ J_n(k|\eta|) - i Y_n(k|\eta|) \right]. \quad (G.53)$$

This can be directly plugged into eq. (G.35) to get

$$\delta(k) = \frac{1}{\sqrt{8\pi}} k^{3/2} \sqrt{\eta} \left[ J_n(k|\eta|)^2 + Y_n(k|\eta|)^2 \right]^{1/2}. \quad (G.54)$$

It is useful to transition back into physical coordinates, $k_p = k/a$. At horizon crossing $k|\eta| = k_p H^{-1}_\Lambda$, so we can rewrite the amplitude of fluctuations as

$$\delta(k_{ph}) = \frac{H_\Lambda}{\sqrt{8\pi}} \left( \frac{k_{ph}}{H_\Lambda} \right)^{3/2} \left[ J_n^2 \left( \frac{k_{ph}}{H_\Lambda} \right) + Y_n^2 \left( \frac{k_{ph}}{H_\Lambda} \right) \right]^{1/2}. \quad (G.55)$$
We can evaluate the Bessel function asymptotically to find that

\[ \delta \approx \begin{cases} 
  k_p & k_p \gg H_\Lambda \\
  \frac{2^n \Gamma(n)}{\sqrt{8\pi^3}} H_\Lambda \left( \frac{k_p}{H_\Lambda} \right)^{3/2-n} & k_p \ll H_\Lambda 
\end{cases} \]  

(G.56)
Appendix H

Beam Parameters

The beam widths and beam ellipticities are derived from deprojection, described in Section 6.3.3. The beams are more elliptical for the 95 GHz FPUs. There is also greater beam width variance for 95 GHz FPUs.
Figure H.1: The beam width as solved by deprojection of X1 in physical coordinates.
Figure H.2: The beam width as solved by deprojection of X2 in physical coordinates.
Figure H.3: The beam width as solved by deprojection of X3 in physical coordinates.
Figure H.4: The beam width as solved by deprojection of X4 in physical coordinates.
Figure H.5: The beam width as solved by deprojection of X5 in physical coordinates.
Figure H.6: The beam width as solved by deprojection of X6 in physical coordinates.
Figure H.7: The ellipticity as solved by deprojection of X1 in physical coordinates.
Figure H.8: The ellipticity as solved by deprojection of X2 in physical coordinates.
Figure H.9: The ellipticity as solved by deprojection of X3 in physical coordinates.
Figure H.10: The ellipticity as solved by deprojection of X4 in physical coordinates.
Figure H.11: The ellipticity as solved by deprojection of X5 in physical coordinates.
Figure H.12: The ellipticity as solved by deprojection of X6 in physical coordinates.
Figure H.13: The X1 cross polar parameter $c$ for the first half (top) and second half (bottom) of flight. Not all detectors are in both halves due flagging.
Figure H.14: The x2 cross polar parameter c for the first half (top) and second half (bottom) of flight. Not all detectors are in both halves due flagging.
Figure H.15: The x3 cross polar parameter $c$ for the first half (top) and second half (bottom) of flight. Not all detectors are in both halves due flagging.
Figure H.16: The x4 cross polar parameter $c$ for the first half (top) and second half (bottom) of flight. Not all detectors are in both halves due flagging.
Figure H.17: The x5 cross polar parameter $c$ for the first half (top) and second half (bottom) of flight. Not all detectors are in both halves due flagging.
Figure H.18: The x6 cross polar parameter $c$ for the first half (top) and second half (bottom) of flight. Not all detectors are in both halves due flagging.
Bibliography


A. Planck early results. i. the planck mission. Astronomy and Astrophysics, 536:A1, 2011. doi: 10.1051/0004-6361/201116464. URL https://doi.org/10.1051/0004-6361/201116464


251


258


[106] F. K. Hansen, A. J. Banday, H. K. Eriksen, K. M. Górski, and P. B. Lilje. Foreground Subtraction of Cosmic Microwave Background Maps Using WI-FIT (Wavelet-Based High-


BIBLIOGRAPHY


[100] Jamil Shariff. Dust scaling between bands, in cmb units, as a function of dust model parameters. Internal Publication.


