Title
Constraints on Short Lived Signals at 150 GHz

Author
Harrington, Nicholas Lee

Publication Date
2018

Peer reviewed|Thesis/dissertation

Constraints on Short Lived Signals at 150 GHz

by

Nicholas Harrington

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor William Holzapfel, Chair
Professor Adrian Lee
Associate Professor Aaron Parsons
Associate Professor Daniel Kasen

Fall 2018

Constraints on Short Lived Signals at 150 GHz

Abstract

Constraints on Short Lived Signals at 150 GHz

by

Nicholas Harrington

Doctor of Philosophy in Physics

University of California, Berkeley

Professor William Holzapfel, Chair

Fast radio bursts (FRBs) are bright, millisecond flashes of radio radiation detected at frequencies between 800 MHz and 8 GHz. We present the first rate constraints on millisecond duration celestial transients like FRBs at 150 GHz. Using a year of data from the South Pole Telescope we searched for signals with a similar phenomenology at 150 GHz and found no evidence for a celestial population. The 0.9 upper confidence limit is $1.1 \times 10^5 \mathrm{sky}^{-1}\ \mathrm{day}^{-1}$ for 1 ms signals with fluence above 10 Jy ms. The search sensitivity is expected to increase dramatically with future microwave telescopes due to increases in spectral information and a reduced background rate.

The first four chapters are dedicated to detailing how we convert a measured optical signal into a measurement of the power spectrum of the CMB. Particular emphasis is placed on the transition edge sensor bolometer readout, the map making and the power spectrum estimation from these maps. Chapter five focuses on how we can analyze the data to look for fast radio bursts.

For Lauren.

Contents

List of Figures

List of Tables

Acknowledgments

In alphabetical order thanks to Bill, Bryan, Jessica, Kyle, Lauren, Liz, Martin, Nathan, Neil, Nick, and Tom. Also thanks to the many people I forgot to include.

Chapter 1

Introduction and Software

SPTpol was a polarized, microwave receiver mounted on the South Pole Telescope. It used transition edge sensor (TES) bolometers to observe light in two bands centered at 90 GHz and 150 GHz. SPTpol's primary science goal was to measure the polarized power spectrum of the cosmic microwave background (CMB) radiation. It was the first experiment to detect B mode polarization through cross correlation with cosmic infrared background[30]. The experiment later went on to measure the E mode and B mode power spectra of the CMB without relying on cross correlation[14, 47]. While the experimental sensitivity is not high enough yet, B mode polarization in the future could be used to place constraints on the sum of the masses of neutrinos[1] and/or provide evidence for inflation[39]. In addition to the primary science goals, the telescope was used to look for orphan gamma ray burst (GRB) afterglows[103]. Currently there are papers in preparation that use SPTpol data to locate galaxy clusters and to place constraints on the rate of Fast Radio Bursts (FRBs) at 150 GHz.

1.1 Telescope

The telescope is an off-axis Gregorian Dragone telescope with a 10 meter primary mirror. After hitting the primary mirror, light is reflected off of a cooled secondary mirror maintained at 10 K with a pulse tube cooler (PTC). See figure 1.1. Lenses then focus the light onto a feedhorn array. For the 90 GHz detectors, the feedhorns couple the radiation to two dipole antennas. Power from the radiation is transported to the TES bolometers via waveguides and then dissipated as heat by a lossy section at the end of the waveguide. This lossy section is tightly thermally coupled to the TES. The TES acts as a sensitive thermistor that responds to this change in dissipated heat. The 150 GHz pixels are similar except orthomode transducers are used to couple the light in the feedhorn to the striplines.

The TES thermistors convert the deposited optical power into an electrical current measured by the readout system. The current is amplified with a DC SQUID and warm electronics chain and then digitized with an analog to digital converter. This digitized signal is fed

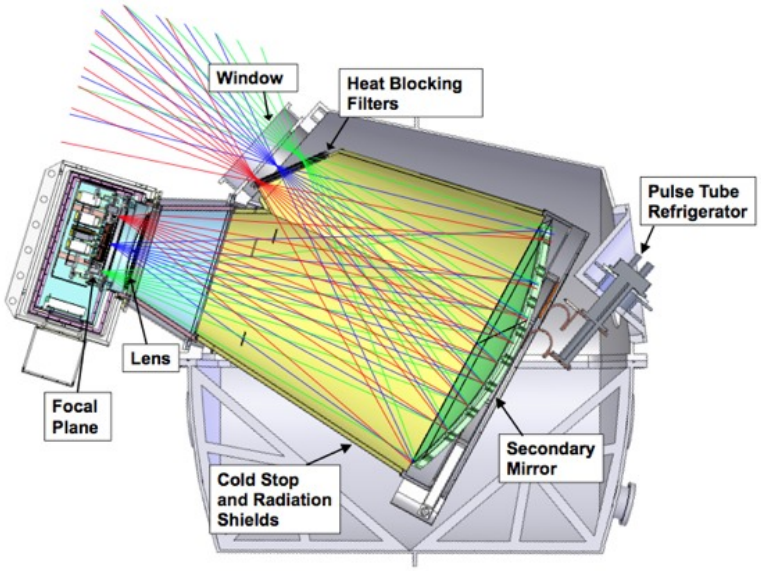


Figure 1.1: The incident rays come from the primary mirror which is not pictured but physically exists to the upper left of the image. The heat blocking filter is tilted to prevent reflections of the focal plane from being imaged by the focal plane

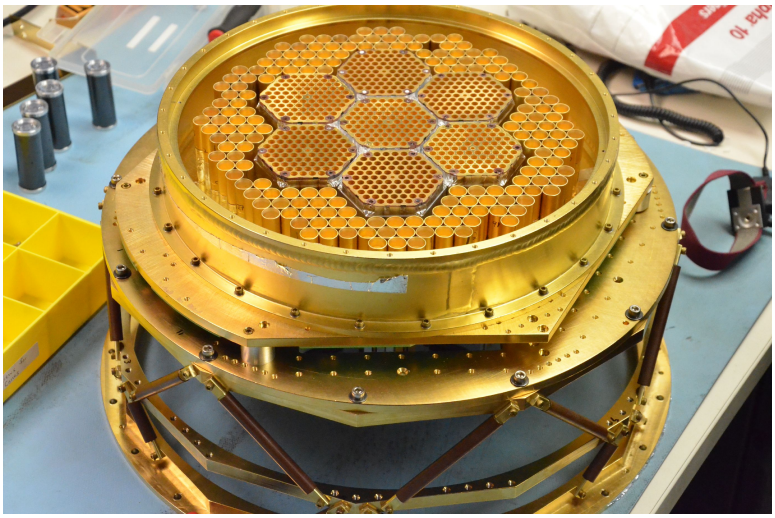


Figure 1.2: The interior 7 hexagons are the feedhorns for the 150 GHz detectors. The outside ring of feedhorns are attached to the 90 GHz detectors. The entire focal plane is mounted on the wedding cake structure. The brown struts are thermal insulators.

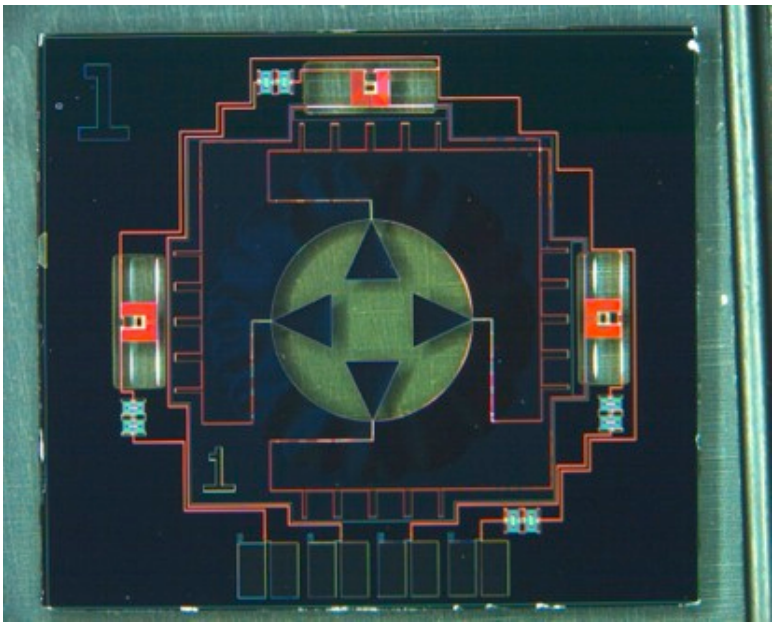


Figure 1.3: This is a close up image of a 150 GHz pixels. The transition edge sensors are at the center of the orange squares along the edge. The TESs at the top and right of the image are connected via a waveguide to the orthomode transducer, the triangle structures at the center. The TES on the left is not optically coupled. This dark detector is primarily used for debugging and characterization purposes. The feedhorn for this pixel directly mounts to the circle at the center of the pixel.

into an FPGA and then processed. The output of this processing is the our digitized signal proportional to the optical power deposited on the detector. We call this our time ordered data (TOD). The TOD is sent to a readout computer and passed along to the software processing.

The transition edge sensors were designed to operate with a base temperature around 250 mK. To cool them to this temperature a Simon Chase Helium sorption fridge is used. The sorption fridge is cooled to 4K with a PTC. The Chase fridge consists of one He_4 and two He_3 coolers. These coolers condense helium and then use the evaporative phase transition to extract heat from the system. The He_4 cooler is used to cool the two He_3 coolers below the He_3 condensation point. After condensation, these He_3 baths are pumped on to extract cooling power. This cooling power is used to cool the focal plane. The focal plane is divided into multiple stages that are thermally isolated from the previous stage. This structure is colloquially referred to as a "wedding cake". This wedding cake has two main layers, the intermediate cold (IC) stage and the ultra cold (UC) stage. The UC stage is strongly thermally linked to the TES bolometers and antennas, so it is maintained at 250mK. The UC stage is mechanically connected to the warmer IC stage with a thermal insulator. The IC stage is then connected to the 4K shell with a thermal insulator. One of the He_3 coolers

is used to cool the UC stage and the other He_3 cooler is attached to the IC stage. The IC stage is used to thermally isolate the UC stage from the 4K shell; any wiring attached between the 4K shell and the focal plane (UC stage) is thermally connected in the middle to the IC stage. The IC stage is used to pre-cool this wiring in order to reduce the thermal load on the UC stage.

The UC stage has a cooling power on the order of microwatts. Because any wiring connected between the 4K shell and the UC stage increases the heat load on the system, efforts are made to reduce the number of wires needed to operate the detectors. To accomplish this, we use a frequency based multiplexing scheme to read out the detectors. For SPTpol 12 TES detectors are operated with one pair of wires; with SPT3G, the next generation receiver, it is 64. We label the group of detectors operated with one pair of wires a module. Each TES is attached to an LC filter in series. Each LC filter in a module has a different resonant frequency ranging from 300kHz to 1.5 MHz for SPTpol. When operating the detectors, a sinusoidal voltage bias is applied at the resonant frequency of the LC filter for each TES. To zeroth order the TES acts as a sensitive thermistor, where changes in optical power will change the resistance of the TES. Changes in the optical power deposited on the TES will appear as an amplitude modulation of the sinusoidal current. Because we only care about low frequency (<50Hz) optical signals we only need $\approx$ 100 Hz of bandwidth to readout an individual detector. The spacing in resonant frequencies means that we can read the optical signal from multiple detectors with one pair of wires with minimum crosstalk. The entire summed signal for a module is digitized at the ADC. Quadrature amplitude demodulation is performed digitally on the FPGA to separate the timestreams for each detector. This demodulated signal is our TOD. After demodulation this signal is transmitted to a control computer that stores it and then all the work shifts to software processing.

1.2 Software

There are several software stacks used to operate the telescope, operate the readout system, collect the data, and finally analyze the data. During the transition from SPTpol to SPT-3G various software systems went through overhauls. We will only discuss the newest set of software.

The telescope is controlled with the software labeled GCP (Generic Control Program). This handles communicating with the telescope control computer, acquiring telemetry from the telescope and cryogenic system, displaying the state of the telescope, and running control programs we label "schedules". The schedules dictate what the telescope does when observing various celestial sources or fields. GCP is implemented primarily in C++ though there is a smattering of other languages and a unique scheduling language used.

The readout system is controlled with the PyDfMux software stack. PyDfMux handles communicating with and controlling the readout boards which handle operating the TES bolometers and DC SQUIDs used in the cryogenic readout system. PyDfMux is written primarily in Python, though it does use a SQL database.

Because of the unique needs of displaying the information from the 15,000 detectors used by SPT-3G, a custom display program is used to display the state of each detector. This program is written in C++ with some wrapper code written in Python. We call it lyrebird. Lyrebird gets the information it displays from the data acquisition process via sockets.

The data acquisition is handled by spt3g_software. The readout boards send the time ordered data and some telemetry over a local area network to the data acquisition computer. spt3g_software receives this data and then puts it into a useful format for future processing. After packing the data it compresses and stores the data. A collection of scripts are then used to transfer this data stateside. Because there is limited bandwidth from the South Pole we transmit a downsampled version of the data. The full rate data is stored to an array of hard drives. During the Austral summer these hard drives are physically shipped north.

Analysis tasks are also handled in spt3g_software. This includes making maps from the time ordered data, looking for FRBs and generating power spectra.

lyrebird

The main purpose of lyrebird is to quickly represent the state of the focal plane to speed up debugging and identifying issues with the operation of the TESs. Each detector is represented as a polygon that is rendered to the screen. The information that lyrebird is presenting is encoded in the color of these polygons. Because of this, it can only display one piece of information at a time. In practice, the most useful piece of information is operating resistance of the detector displayed in terms of the fractional resistance. The fractional resistance is the operating point divided by the normal resistance of the detector. Beyond the basic functionality, lyrebird can plot the data in a graph, perform real time PSD estimation on the data and display the information in text form.

Architecturally it is relatively simple. Data collecting functions operate in threads that write data to a buffer. Each buffer has an associated string ID. An equation parser takes equations written using those data IDs and applies basic arithmetic operations to them. These equations produce a number between 0 and 1. A color map converts the output of the equations into a displayed color. Each detector's rendered polygon has an associated equation and color map. When drawing these polygons to the screen this information is used to set the color.

spt3g_software

spt3g_software is written in a combination of C++ and Python. The low level processing routines are written in C++ in order to speed up processing, while the high level analysis choices all take place in Python code to speed up development time. Analysis software frameworks are not often discussed in the field of software engineering. They are very niche pieces of software and in most small scale analysis cases the architecture is unimportant. As the number of people contributing and the complexity of the analysis task increases, a

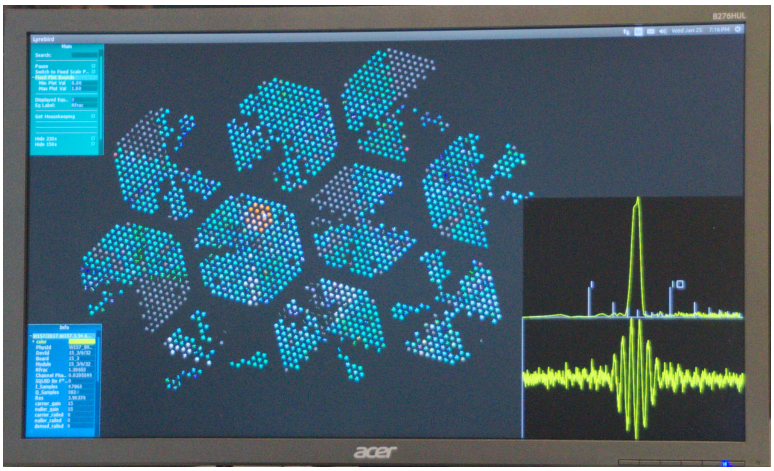


Figure 1.4: A picture of lyrebird displaying SPT-3G bolometer data. The frequency response of the detector is currently being characterized with a Fourier Transform Spectrometer.

poorly designed analysis framework can dramatically increase the time it takes to perform scientific analyses and ultimately limit the scientific results.

spt3g_software forces a structure for TOD processing with the goal of making all of the steps taken in the data processing explicit with no unintended side effects. This is to, hopefully, make it easier to share processing scripts and modify existing pipelines to do interesting things. There are three main types of objects used in the data processing. These are frames, modules, and pipelines. A frame is an associative array that maps strings to data. Frames store all of the data being processed. The data the strings map to in a frame include vectors, floating point numbers, and maps of the sky. Modules accept a frame as an input and returns zero, one or many frames. Modules represent discrete units of processing. A module that applies a high pass filter to the data and a module that calculates individual detector pointing from boresight pointing are both examples of modules used in the processing. A pipeline is just a list of modules. The first module generates the initial data; this can be a module that reads data from disk or a module that accepts data over the network from the readout boards. Any frames output from the initial module are passed to the next module. Frames returned by the second module are passed to the third module and so on. The pipeline forms what is effectively a bucket brigade of data processing. The main goal of this is so at a glance every step of data processing used in a script can be ascertained. The forced structure of the processing also enables people doing analysis to share the various modules they have written and easily incorporate them into existing scripts.

The frames have an enumerated type value that specify what their main content is. The three main frames used in data processing are scan frames, calibration frames, and map frames. Calibration frames store slow or unchanging information about the detectors. This includes information like the polarization angle of light a detector is sensitive to and the detector's pointing relative to boresight. Scan frames store all of the fast changing

information: the actual TOD, the boresight pointing, etc. Map frames store maps of the sky. During processing a module will often need to cache information in the calibration frames for future processing of scan frames. Because references to the data get stored for future use the requirement is placed that modules not modify any data stored in the frames. The modules can modify frames, but not the underlying data the frame stores. With the high pass filter module this means that instead of modifying the input TOD it copies the TOD, high pass filters it, and then stores it in the frame. If a module were to modify data a previous module was caching it could lead to modules later in the pipeline affecting earlier modules in the pipeline. While this can increase the memory used during processing, it has the added benefit that it makes it easy to inspect the changes in the data caused by individual modules. In the case of the high pass filter, because the input data remains, a quick comparison of the data before and after high pass filtering can be made.

For units that have a purely multiplicative relationship e.g. Kelvin and Rankines, meters and feet, a units system exists. When storing values with physical units, they are multiplied by a value that converts them to the stored units form. As an example, when storing a number in meters in a frame we multiply it by G3Units.meters:

```
frame['TheKey'] = number_in_meters * G3Units.meters
```

When using the data in some specific units we simply divide the desired units:

```
number_in_cm = frame['TheKey'] / G3Units.cm.
```

The trade off to all of this structure is that starting analysis can be a bit slow. Learning how to properly interact with the pipeline structure and the units system requires effort. The advantage is that once the knowledge has been gained the framework facilitates cooperation on the development of analysis techniques.

Chapter 2

Frequency Multiplexed Readout

2.1 Overview

SPTpol used a frequency multiplexed readout system to supply power and measure the impedance of the transition edge sensor (TES) bolometers. For legacy reasons, we call this the DfMux readout system. A TES is a superconductor held between the normal and superconducting state by electrical power. Because the resistance of this device changes very rapidly with changes in temperature, to first order it behaves like a very sensitive thermistor. With the frequency multiplexed readout system, each TES has an LC filter placed in series with it. In a given readout grouping, each LC filter is built to have a different resonant frequency. A voltage bias is applied to each of the TESs in the form of sine waves at the resonant frequency of the LC filters. We label the sum of these sine waves a comb after how they appear when plotted in frequency space. Each of the sine waves in a comb produces a current determined by the resistance of the TES that sine wave is supplying power to. The current is then amplified with a DC SQUID and op-amp amplifier chain. After amplification it is digitized and processed on an FPGA. On the FPGA the sine waves are demodulated. After demodulation the sine waves are low pass filtered and decimated in order to reduce the volume of data for transmission and analysis. This output signal is our time ordered data.

Transition edge sensors are held between the superconducting and normal state by electrical power. This is stabilized by what is labeled electro-thermal feedback. For any circuit supplying power to a TES resistor, the output impedance of the circuit needs to be much smaller than the TES resistance, R. In this limit we can write the electrical power as: $P_e = \frac{V^2}{R}$, which means that $\frac{dP_e}{dR} < 0$. Any increase in optical power deposited on the TES will raise the temperature and increase the resistance of the TES. This causes the electrical power supplied to the TES to drop. This acts as a negative feedback mechanism that keeps the TES at stable operating point. In the opposite limit, where the circuit supplying current to the TES has a large impedance, $P_e = I^2R$ and $\frac{dP_e}{dR} > 0$. Any increase in optical power is met with an increase in electrical power creating a positive feedback loop. The resistance will either drop entirely until the TES is superconducting or increase until it is in the normal

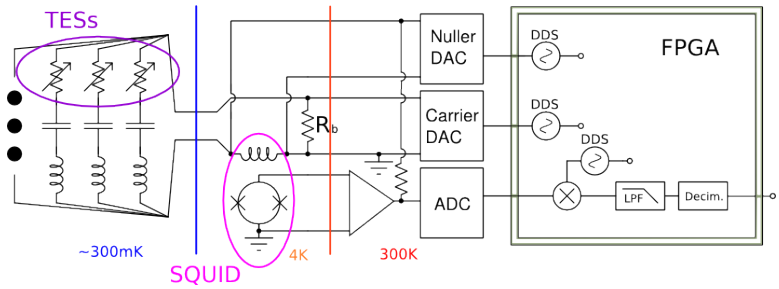


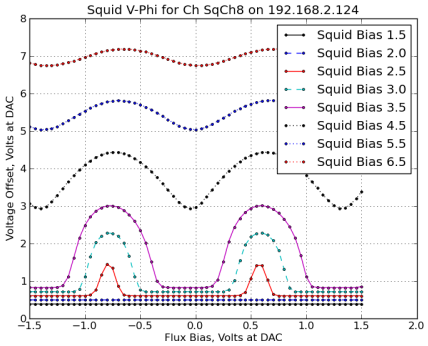
Figure 2.1: A simplified schematic of the DfMux readout system. The TESs on the left are attached to LC filters. The signal for the entire comb is amplified with a DC SQUID and then digitized. The voltage bias applied to the TES is generated by the carrier DAC. The nuller DAC generates a signal to help reduce the current through the non-linear SQUID.

state.

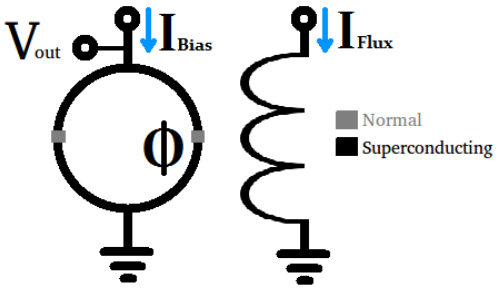
In order to measure the current flowing through the TES, we use a DC SQUID amplifier. The DC SQUID converts an input current into an output voltage that we further amplify before demodulating. DC SQUIDs intrinsically have a non-linear gain. In order to linearize the output and increase the dynamic range of the SQUID, we rely on two forms of negative feedback. We use an analog negative feedback loop called a flux locked loop(FLL). We also digitally synthesize sine waves that are 180 degrees out of phase with the current flowing through the TESs and feed this into the SQUID coil. We call the signal applying voltage to the TESs the carrier signal. We call this out of phase signal the nuller signal. Two schemes of nulling the carrier signal have been used on SPTpol. One where we set the nuller amplitudes and phases once after we have set the TES detectors operating point, which we call static nulling. In the other scheme we continuously update the nuller amplitudes and phases to cancel out the signal through the SQUID coil, called dynamic active nulling (DAN). DAN suppresses current through the SQUID coild at frequencies near the frequency of the nuller signal. The FLL is used to suppress low frequency current through the SQUID.

The goal of the frequency multiplex (FMux) system is to reduce the number of wires between the 4 Kelvin stage housing the SQUIDs and the millikelvin stage housing the TESs. This reduces the complexity of the wiring and the thermal load on the millikelvin stage. There are a number of complications that come from this multiplexing scheme, however. It needs to manage supplying power to the TESs with low impedance and reading out the current with low noise. This combination of requirements drives the various negative feedback mechanisms and added complications of the system. The goal of this section is to provide an overview of how TESs and DC SQUIDs operate and interact with the FMux readout system. There is quite extensive literature on the details of TESs and DC SQUIDs and the treatment in this document of the device physics is very minimalistic. The goal is to provide the simplest model of the device operation in order to understand how TESs and DC SQUIDs interact with the FMux readout system.

2.2 DC SQUIDs



(a) The x axis is the current through the SQUID coil. The y-axis is voltage out. The various lines correspond to current through the Josephson junctions. When the current through the Josephson junction is low, the entire current can become superconducting leading to the humped structure. Once it is above I_0 the voltage vs current at the SQUID coil becomes more sinusoidal.



(b) A DC SQUID consists of a coil tightly coupled to a pair of Josephson junctions. A Josephson junction consists of a thin insulator between two superconductors.

Figure 2.2: DC SQUID

Operating TES bolometers requires a circuit that measures the current with an input impedance much lower than the resistance of the TES. In order to do this, we use a DC SQUID[16]. DC SQUIDs consist of a superconducting loop with two insulating Josephson junctions tightly coupled to a coil that generates a magnetic field. See figure 2.2. The superconducting current through the Josephson junction pair has the form:

$$I_{super} = I_0 \cos \frac{q_e \Phi}{\hbar} \tag{2.1}$$

Where I_{super} is the superconducting current, I_0 depends on physics on more complicated physics. From an operational standpoint I_0 is somewhat dependent on the current through the Josephson junctions and the temperature of the SQUID. q_e is the elementary charge and Φ is the magnetic flux through the loop containing the two Josephson junctions[24]. For this section we will use the variables in figure 2.2.

We apply a total bias current to the Josephson junction pair. Only the normal current produces a voltage difference based on the impedance of the Josephson junction pair.

$$I_{bias} = I_{super} + I_{normal} \tag{2.2}$$

$$V_{out} = I_{normal} Z_{Joseph} = \left(I_{bias} - I_0 \cos \frac{q_e \Phi}{\hbar} \right) Z_{joseph} \tag{2.3}$$

A magnetic coil is tightly coupled to the Josephson junction loop. Current through the coil generates a flux in the Josephson junction loop based off of the mutual inductance M of:

$$\Phi = MI_{Flux} \tag{2.4}$$

The DC SQUID converts current through the input coil into an output voltage. The input impedance of this circuit is the impedance of the superconducting coil, which is purely inductive. The SQUIDs we use are actually many SQUIDs connected in series to increase their amplification[35], but they behave like a single SQUID.

There are of course a couple complications in actual use. The structure of the voltage out versus magnetic flux in is not sinusoidal due to more complicated device physics. The peak to peak voltage at the output when varying the input flux is also a function of the current through the Josephson junction pair. When selecting the operating point of the SQUID we vary the current through the Josephson junctions to maximize the peak to peak voltage. We then select a DC flux level that maximizes the dynamic range of the SQUID. During operation we set these values when the temperature of the SQUID is slightly higher than it is during normal use. To compensate for that we select a current bias through the Josephson junctions that is 10% higher than the optimal point at the time of measurement.

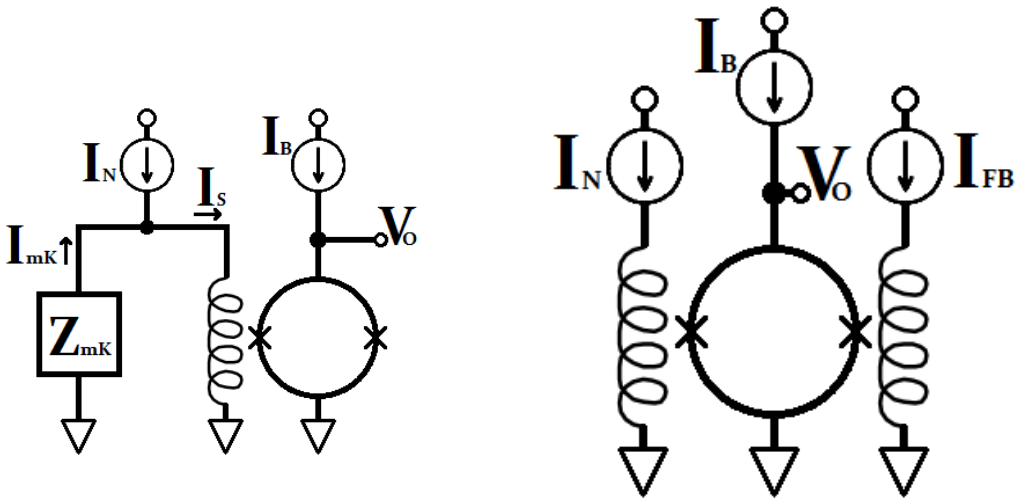
The DC SQUIDs serve as a low input impedance, high gain amplifier, but their output voltage is a highly non-linear function of the input current. Negative feedback is used to linearize the gain of our SQUIDs.

Feedback Topology

Some SQUIDs are manufactured with one coil. Any feedback is applied as current to that feedback coil. This is the form used by SPTpol and SPT3g's SQUIDs. Some manufactured SQUIDs also use a second coil that is specifically designed to be a coil for applying the negative feedback. Negative feedback in the single coil topology effectively reduces the input impedance of the SQUID coil. It does, however, also lead to significant complications with high frequency operation. Feedback current can flow through the millikelvin circuitry instead of the SQUID coil which makes to the feedback current through the SQUID have a very complicated gain and phase structure. With SQUIDs that have recently become available to the SPT collaboration the input coil inductance has been lowered to 20nH. Future work should consider the reduced high frequency complexity of using a second feedback coil.

Single Coil Topology

With a single coil, any feedback gets applied as current at the input coil. This topology has the benefit that the negative feedback lowers the effective impedance of the SQUID coil. It has the downside that not all of the feedback current flows through the SQUID coil; some current is diverted through the millikelvin circuitry. This downside can lead to instabilities with the FLL feedback system and an increased noise with the DAN feedback system.



(a) Diagram of single coil SQUID scheme employed by SPTpol. Z_{mK} is the impedance of the millikelvin circuitry. A low frequency, the impedance is set by the LC filters attached to the TESs. At high frequencies, the impedance is set by the capacitance of the transmission lines between the 4K and millikelvin stage, capacitive paths to ground, and any resonances in the system.

(b) Diagram of double coil feedback topology.

Figure 2.3: SQUID Operational Schematics

For this section we will discuss the circuit found in figure 2.3. With the single coil topology some of the current we apply to the system through the nuller wiring, I_N, is diverted through the millikelvin circuitry. We can write the ratio of the current through the SQUID to the current that is applied on the nuller line as a variable r_N:

$$r_N = \frac{I_S}{I_N} = \frac{Z_{mK}}{Z_m K + j\omega L_{SQUID}} \tag{2.5}$$

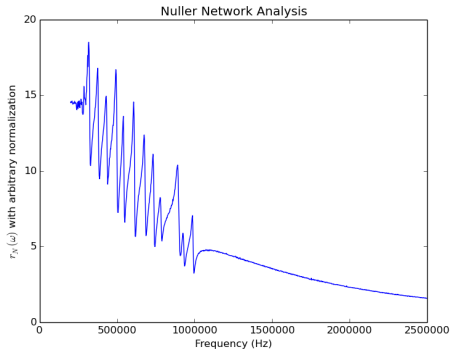
Where Z_{mK} is the series impedance of the millikelvin circuitry. In the case of the FMux multiplexing scheme this is many LCR series circuits placed in parallel.

If we write the nuller current as some gain factor A multiplied by the SQUID coil current:

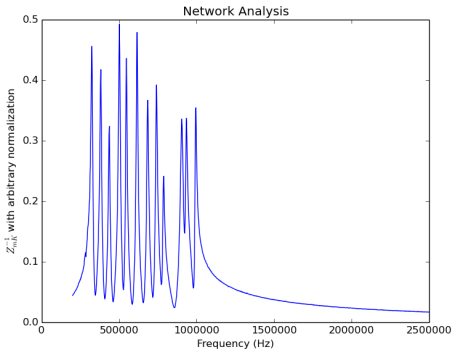
$$I_N = -AI_S \tag{2.6}$$

we can write the effective impedance of the SQUID coil in series with the TES as:

$$Z_{SQUID} = \frac{j\omega L_{SQUID}}{1 + A} \tag{2.7}$$



(a) This plot is what we call the nuller network analysis. It is a measurement of r_N. The low frequency structure comes from the rising impedance of the SQUID coil inductor. The low frequency spiky structure of r_N is dominated by the power diverting through the LC combs.



(b) This plot is the admittance of the millikelvin circuitry. Over our operating frequencies the structure of the admittance is dominated by the LC filters at millikelvin temperatures.

Figure 2.4: Circuit Network Analyses

In short, we reduce our SQUID coil impedance by the gain of our circuit, but any current delivered by the feedback system is modified by a factor of r_N.

Multiple Coil Topolgy

With a second feedback coil, the behavior of the system depends on the mutual inductance between the primary coil and the feedback coil. If we assume the impedance of the circuit generating the feedback current is much larger than the coil self and mutual inductance then the feedback current and SQUID coil effective impedance has a very simple form:

$$r_N = \frac{I_S}{I_N} = 1 \tag{2.8}$$

$$Z_{SQUID} = j\omega L_{SQUID} + j\omega MA \tag{2.9}$$

In the limit of small mutual inductance, the impedance is just the coil impedance.

Dynamic Active Nulling

With dynamic active nulling (DAN), the nuller signal is constantly updated to cancel the carrier current. The SPTpol implementation of DAN is an integral feedback loop, where the integral of the demodulated signal is used as the amplitude of a synthesized sine wave with the same frequency as the demodulation. For a low pass filter operator LPF[...], arbitrary gain G, and processing delay t_{dan}, the nuller current has the form:

$$I_N = -\sin(\omega t)\int_{-\infty}^{t-t_{dan}} G\text{LPF}[\sin(\omega t')I_S]dt' - \cos(\omega t)\int_{-\infty}^{t-t_{dan}} G\text{LPF}[\cos(\omega t')I_S]dt' \quad (2.10)$$

The measured SQUID current, I_S, is the residual signal that consists of any signal the nuller is not canceling out. This is the error that the integral feedback loop is correcting. In these equations, the units of G are $1/t$.

The integral form of the equation can easily be solved by picking a modulated basis for the current at the SQUID. This modulated basis has the form:

$$I_S = A_I(w_0)\sin(\omega t) + A_Q(w_0)\cos(\omega t) \quad (2.11)$$

Where $A_I(w_0)$ and $A_Q(w_0)$ are the Fourier transforms of the amplitude of the sine waves.

$$A_I(w_0) = \int_{-\infty}^{\infty} A_I(t)e^{2\pi j t\omega_0}d\omega_0 \quad (2.12)$$

With that form, we can rewrite the synthesized nuller current as:

$$I_N = -\frac{G}{2j\omega_0}(A_I(w_0)\sin(\omega t) + A_Q(w_0)\cos(\omega t))e^{j\omega_0 t_{dan}} = -\frac{G}{2j\omega_0}e^{j\omega_0 t_{dan}}I_S \quad (2.13)$$

With this feedback loop operating the feedback at the SQUID coil has the form:

$$I_S = I_{mK}\frac{1}{\left(1+\frac{G}{2j\omega_0}e^{j\omega_0 t_{dan}}\right)} \quad (2.14)$$

The loop gain of the feedback circuit is:

$$\text{Loopgain} = \frac{G}{2j\omega_0}e^{j\omega_0 t_{dan}}r_N \quad (2.15)$$

The DAN feedback loop drops the current through the SQUID coil to zero at the DAN operating frequency. This current suppression drops off at frequencies away from the DAN operating frequency. The bandwidth of the system is set by the time delay t_{dan}. The gain needs to be below unity before $e^{j\omega_0 t_{dan}} = j$ in order to prevent the negative feedback from becoming positive feedback.

Flux Locked Loop

The flux locked loop (FLL) used by SPTpol consists of an op-amp amplifier that amplifies and filters the output of the DC SQUID. This output signal is then fed back in the single coil configuration to the input of the SQUID coil through a resistor R_{FB} creating an analog feedback loop.

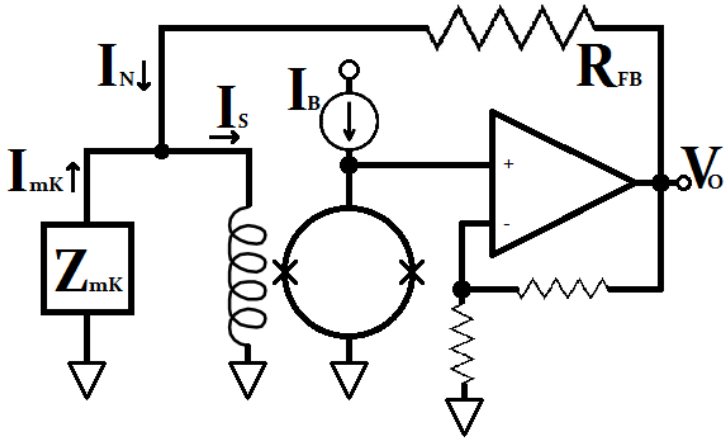


Figure 2.5: The output voltage, V_O generates a feedback current I_N that cancels the signal through the SQUID coil. While not explicit in this schematic the feedback is negative feedback to help linearize the SQUID.

The op-amp is operated with a large enough gain that over the frequencies of interest it behaves like a one pole filter. In the absence of feedback, $R_{FB} \to \infty$, the circuit has the gain:

$$V_O = \frac{A}{1 + j\omega\tau_{op-amp}} I_{mK} \tag{2.16}$$

For finite R_{FB} the gain of the flux locked loop circuit is:

$$V_O = \frac{R_{FB}A}{A + R_{FB}} \frac{1}{1 + j\omega\frac{\tau}{1+A/R_{FB}}} I_{mK} b \tag{2.17}$$

In the large gain limit, $A >> R_{FB}$, this has the simple form of:

$$V_O = R_{FB} I_{mK} \tag{2.18}$$

In practice, the total bandwidth of the FLL system ended up being limited. Two factors primarily limit the bandwidth of the flux locked loop. First, the amplifier system is located on the outside of the cryostat and the SQUID is located inside. This is separated by roughly 10 cm of wiring. This gives a time delay of roughly $t_{delay} = 1ns$. While this seems small, at high frequencies this can impart a phase delay in the feedback system that pushes the feedback loop to being unstable. The other complication comes from the fact that not all of the feedback current produced is delivered to the SQUID coil. The factor r_N estimated above has a complicated structure that depends on numerous factors with the cold circuitry. Capacitive paths to ground, stripline resonances from poorly terminated impedances and inductor self

resonances create a very rich structure in the high frequency sub-kelvin impedance. These resonances lead to instabilities with the FLL.

The equation for the loopgain of the circuit, now including the time delay, is:

$$\text{Loopgain} = r_N \frac{A}{1 + j\omega\tau_{op-amp}} \frac{1}{R_{FB}} e^{jwt_{delay}} \tag{2.19}$$

The various resonances abstracted away in r_N create large phase shifts in the loopgain. In order to maintain a stable feedback loop the loop gain needs to be below unity by the time these resonances appear. In SPTpol, one of these resonances appeared at $\approx 9MHz$ highly limiting the gain and bandwidth of the FLL.

To operate a FLL with an increased bandwidth, two changes would need to be made. First, the semiconductor based amplifier would need to be moved closer to the SQUID and operated at $4K$. Second, we would either need to switch to using the second feedback coil so the complicated impedance of the millikelvin circuitry would not interfere with the stability of the feedback loop, or we would need to carefully design the millikelvin circuitry to avoid unwanted resonances.

2.3 TES Properties

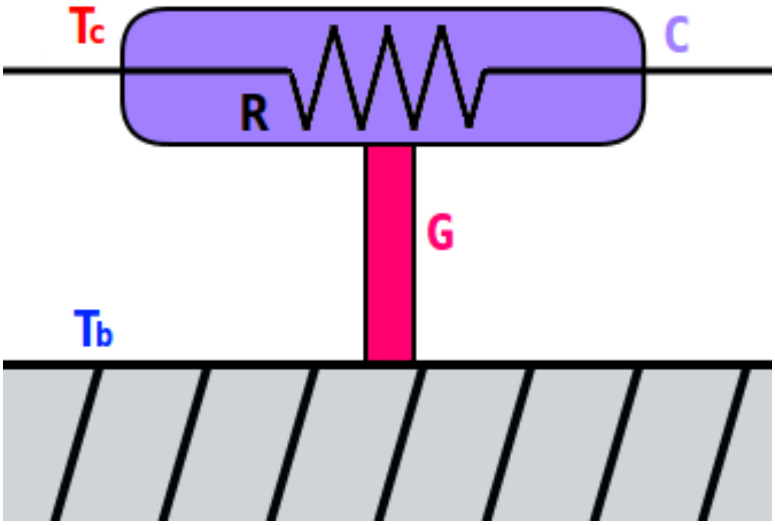


Figure 2.6: A TES consists of a transitioning superconductor attached to a thermal bath. Power is removed from the TES through a weak thermal link, G, to a bath at temperature, T_b. $T_b < T_c$. Electrical power is applied to keep the TES at its transition temperature T_c.

Transition edge sensors (TESs) are superconductors held between the superconducting and normal state with electrical power. The TESs are connected to a thermal bath that is at a temperature, T_b below their superconducting transition temperature, T_c. In our system

we use antenna coupled TESs. Optical power is transported to the TES and used to heat the TES. The TES converts the deposited heat into a current signal we can measure. The superconducting transition of a TES has a very steep resistance versus power curve. See figure 2.7. This sharp transition makes a TES a very sensitive thermistor and in our case, a very sensitive measurement of optical power.

During operation, we raise the temperature of the millikelvin stage, which as at a temperature T_b, to be above the TES transition temperature of T_c. We then supply sufficient electrical power to keep the TES in its normal state when the millikelvin stage is at its operating temperature, $T_b < T_c$. The SPT collaboration uses the term overbiased for TESs in this state. After the stage temperature is lowered, the bias voltage is dropped until the electrical power is low enough that the TES enters the superconducting transition. At this point the small signal impedance of the TES changes from positive and real $\frac{dV}{dI} = R_{normal}$ to being negative $\frac{dV}{dI} = -R$[37]. This transition from the normal to superconducting regime appears as the turn around point in the IV curve. See figure 2.7. This turn around is driven by the electro-thermal feedback of the TES.

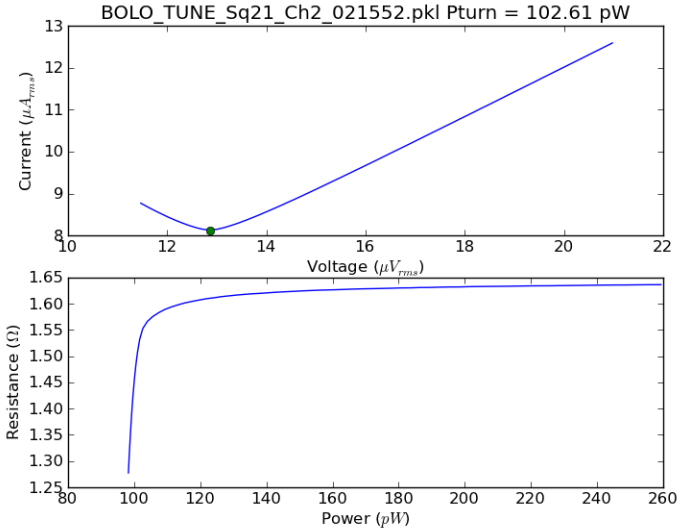


Figure 2.7: Transition Edge Sensor I-V Curve and R-P_e Curve. At high bias power, the TES is in the normal regime and behaves like a resistor. At lower bias power, the TES is in between the superconducting and normal state. At small enough bias power, the TES becomes purely superconducting. Once it is entirely superconducting, electrical power cannot be applied to the TES, $P_e = I^2 \times (R = 0) = 0$. It is stuck in the superconducting state. The SPT collaboration uses the slang "latched" to describe a TES in this state.

TES Thermal Link

The thermal link and heat capacity of the TES affect the operation. For our responsivity characterization in section 2.3 we use the small signal conductance and capacity:

$$G = \frac{dP}{dT_c} \tag{2.20}$$

$$C = \frac{dE}{dT} \tag{2.21}$$

For the thermal link we can model the thermal conductance of an infinitesimal section as a power law $k(T) = k_0 T^n$. We can the write the total thermal power conducted as:

$$P_T = \int_{T_b}^{T_c} \frac{A}{l} k(T) = \frac{A}{l} \frac{k_0}{n+1} \left(T_c^{n+1} - T_b^{n+1}\right) = G_0 \left(T_c^{n+1} - T_b^{n+1}\right) \tag{2.22}$$

The exponent, n depends on the carrier of the thermal energy and the defects in the conducting medium. n is typically in the range of 1-3[48]. With this form the small signal thermal conductance has the form:

$$G = G_0 \left((n+1)T_c^n - T_b^{n+1}\right) \tag{2.23}$$

TES Responsivity

For TES bolometers operated with a DC voltage bias, [37], is the standard resource for understanding the responsivity of the system and the stability requirements. With the frequency multiplexed readout system, the electrical circuit generating the voltage bias is significantly more complicated. Unfortunately, the behavior of a TES depends heavily on the circuit supplying power to the TES. Its stability and responsivity are both a function of this circuit. For the purposes of the responsivity derivation, we are going to make some simplifying assumptions about this electrical circuit. Mainly, we are going to neglect any phase changes with the current delivering power to the TES. In order to encode the effects of the bias circuit, we define the logarithmic derivative of the electrical power deposited on the TES with respect to the resistance as λ.

$$\lambda = \frac{R}{P_e} \frac{dP_e}{dR} \tag{2.24}$$

For a bias circuit with zero output impedance (voltage biased) $\lambda = -1$. For a bias circuit with infinite output impedance (current biased) $\lambda = 1$. This parametrization allows us to abstract away the details of the electrical circuit as long we can neglect any phase changes in the electrical circuit supplying power. This parametrization is not used in the literature about TESs because they focus on the much simpler DC biased system.

In the case of the frequency multiplexed system, frequency space is ambiguous. In the following derivation, frequency space refers to the frequency bolometer response, not the

frequency used to supply power and read out the TES. This analysis will hold true for bolometers biased at frequencies of hundreds of kilohertz to hundreds of megahertz. The biasing circuit is assumed to be a circuit that delivers a power that is some function of the TES resistance. Fortunately, when analyzing the bias circuit we can use traditional impedance analysis methods. By that I mean we do not need to worry about any ringing in the biasing LC circuit caused by the TES changing impedances. Any transient behavior in the LC resonant circuit dies off with a time constant of $2L/R_{TES}$ where L is our inductor and R_{TES} is the TES resistance. For our system, $L \approx 20\mu H$ and $R \approx 1\Omega$ giving a time constant of $\approx 5 \times 10^{-6}$ which is a couple orders of magnitude smaller than any time constant associated with the TESs used. In short, any ringing in the LC circuit dies off quickly enough that we will just ignore it.

To be as explicit as possible, here is what we are assuming:

- The signals we are measuring are sufficiently small that linear analysis is adequate
- The TES resistance is only a function of its temperature
- The time constant associated with the LC resonator $\frac{L}{R}$ is sufficiently small that we can neglect it
- The TES is stable.
- Any phase change in the circuit delivering power is negligible.

To begin, let us define a few terms we will be using for the rest of the derivation. P_{op} is the optical power. P_e is the electrical power. P_T is the thermal power. R is the TES resistance. T is the TES temperature. G is the small signal thermal conductance, $\frac{dP_T}{dT_c}$. C is the small signal heat capacity. I_{rms} is the RMS current through the TES. α is the logarithmic derivative of the TES resistance at the operating resistance, R.

$$\alpha = \frac{T}{R}\frac{dR}{dT} \tag{2.25}$$

It is convenient to include a couple derived quantities, such as the electrothermal loopgain, L, and the intrinsic bolometer time constant, τ_0.

$$L = -\frac{\alpha \lambda P_e}{GT} \tag{2.26}$$

$$\tau_0 = \frac{C}{G} \tag{2.27}$$

Note that this definition of the bolometer loopgain includes a factor of λ as compared to the traditional definition of loopgain. A circuit diverging from the pure voltage bias regime will have $\lambda > -1$ which results in an effective decrease in the loopgain of the system.

We will be solving everything in the linear small signal perturbation limit. Note $P_e = I_{rms}^2 R$. Here are two useful identities:

$$\frac{dI_{rms}}{dR} = \frac{I_{rms}}{2R}(\lambda - 1) \tag{2.28}$$

$$\delta T = \frac{dT}{dI_{rms}}\delta I_{rms} = \frac{dT}{dR}\frac{dR}{dI_{rms}}\delta I_{rms} = I_{rms}R\frac{T}{\alpha P_e}\frac{2}{\lambda - 1}\delta I_{rms} \tag{2.29}$$

Alright, now that we've gotten all of those definitions and identities out of the way, we start with conservation of energy:

$$P_T = P_{op} + P_e \tag{2.30}$$

We can divide this equation into a steady state equation and a small signal frequency dependent perturbation. For some sinusoidal optical power perturbation, we can write the conservation of energy in terms of the change in temperature of the TES.

$$\delta P_{op} e^{j\omega t} = (G + j\omega C)\delta T e^{j\omega t} - \frac{dP_e}{dR}\frac{dR}{dT}\delta T e^{j\omega t} = (G + j\omega C - \frac{P_e \alpha \lambda}{T})\delta T e^{j\omega t} \tag{2.31}$$

Using the two identities above, we rewrite this as:

$$\delta P_{op} = (G + j\omega C - \frac{P_e \alpha \lambda}{T}) I_{rms} R \frac{T}{\alpha P_e}\frac{2}{\lambda - 1}\delta I_{rms} = (1 + j\omega\tau_0 + L)\frac{I_{rms}R}{L}\frac{2}{1 - \frac{1}{\lambda}}\delta I_{rms} \tag{2.32}$$

Rearranging the terms gives the power to current responsivity:

$$s_i = \frac{\delta I_{rms}}{\delta P_{op}} = -\frac{1}{I_{rms}R}\frac{1 - \frac{1}{\lambda}}{2}\frac{L}{L+1}\frac{1}{1 + j\omega\frac{\tau_0}{1+L}} \tag{2.33}$$

In our system we apply a voltage bias, V_{rms} and the dominant source of loss behaves like a resistance that is not in thermal contact with the TES. We label this R_s. In this case:

$$\lambda = -\frac{R - R_s}{R + R_s} \tag{2.34}$$

$$s_i = \frac{\delta I_{rms}}{\delta P} = -\frac{1}{V_{rms}}\frac{R + R_s}{R - R_s}\frac{L}{L+1}\frac{1}{1 + j\omega\frac{\tau_0}{1+L}} \tag{2.35}$$

In the low loopgain limit we get no response. In the high loopgain limit $\frac{L}{L+1} \approx 1$. Increasing loopgain speeds up the detector which means it can respond to higher frequency signals. Off island resistance introduces a resistance dependent term in the responsivity. Because the operating resistance of the TES is a function of the optical loading, this introduces a loading dependent responsivity. This off island resistance also slows down the detector by reducing its loopgain.

Demodulation

In our system we have n many sine waves being amplitude modulated by the n many TESs on the comb. We would like to extract their amplitudes. We do this all digitally. The signal coming from the SQUID is amplified and filtered and sent to the ADC. This ADC is referred to as the demodulator (demod) ADC. To extract the amplitudes of each sine wave we multiply the signal by a sine wave at the carrier frequency and then low pass filter it. For a carrier at frequency ω we read out a modulated amplitude of $A(t)\sin(\omega t + \delta)$:

$$I(t) = A(t)\sin(\omega t + \delta) \times \sin(\omega t) = A(t)\frac{\cos(\delta) - \cos(2\omega t + \delta)}{2} \overset{lowpass}{\rightarrow} A(t)\frac{\cos(\delta)}{2} \tag{2.36}$$

$$Q(t) = A(t)\sin(\omega t + \delta) \times \cos(\omega t) = A(t)\frac{\sin(\delta) + \sin(2\omega t + \delta)}{2} \overset{lowpass}{\rightarrow} A(t)\frac{\sin(\delta)}{2} \tag{2.37}$$

We would like the phase of the demodulator to line up with the phase of the bolometer response. In that situation $\delta = 0$ so $I(t) = A(t)/2$ and $Q(t) = 0$. When the signal is adequately modeled as just a modulation of a sine wave, the demodulator phase δ can be set by maximizing the signal in the I phase. In practice, the phase of the response can be slightly different from the phase that maximizes the signal in I. As a toy example, let's say the bias circuit has some residual imaginary impedance, jZ at the operating frequency of the TES. The current through the circuit has the phase:

$$I = V/(R + jZ) \tag{2.38}$$

While any changes in resistance has a response phase:

$$\frac{dI}{dR} = -\frac{V}{(R + jZ)^2} \tag{2.39}$$

leading to a difference in phase between the response of the circuit and the maximum amplitude signal of $-1/(R + jZ)$.

This full I-Q basis describes all the information in the signal around the carrier frequency. If for some reason we would like to determine if a signal is amplitude modulated or if it is some signal being injected on either side of the carrier, one can take an DFT of $I(t) + jQ(t) = e^{j\omega t}A(t)$. The lower sideband and upper sideband are distinguished in the negative and positive frequency bins of the discrete fourier transformed data.

Anti-Aliasing Filter

The anti-aliasing filter consists of cascaded integrator-comb (CIC) filter and finite impulse response (FIR) filters. After low pass filtering the data, the data is decimated. The decimation process means that the data is not shift invariant. The decimation also means that there are multiple impulse responses, depending on where the impulse is in the input data stream relative to the output data stream.

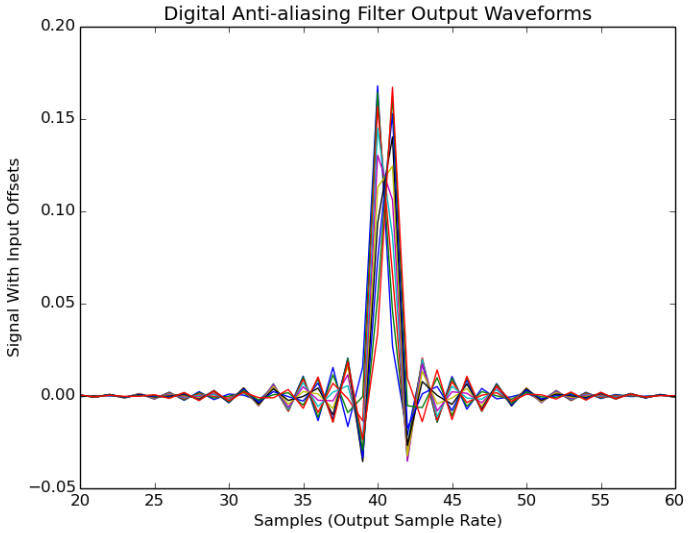


Figure 2.8: A Subset of the Decimation Filter Impulse Responses

2.4 Noise

Our time ordered data output is the demodulated and low pass filtered waveform. The respective sources of noise in system have different transfer functions to this demodulated waveform depending of if they are noise in the power applied to the TES or electronic noise. The coupling of the noise also depends on the location of the electronic noise. For convenience of calculation, we will refer to the noise in terms of the current at the input coil of the SQUID. This is normalized to be the signal that appears in the I phase of the demodulation.

Current Noise Demodulation

When demodulating the data we have freedom to choose the amplitude of the demodulation signal. Our goal is to have the demodulated spectrum be normalized to have the same amplitude as the input spectrum.

For some white noise spectrum of sine waves, we can write our time ordered data as:

$$T_n = \sum_{k=0}^{N-1} a_k \cos(\frac{2\pi kn}{N}) + b_k \sin(\frac{2\pi kn}{N}) \tag{2.40}$$

Where a_k and b_k are normally distributed with identical variance.

We demodulate with a signal $\chi \sin(\frac{2\pi k_{bias} n}{N})$ where we are solving for χ. Demodulation and low pass filtering causes a single Fourier term to appear as:

$$\chi \sin(\frac{2\pi k_{bias} n}{N}) a_k \cos(\frac{2\pi k n}{N}) = \frac{\chi}{2} \sin(\frac{2\pi k_{bias} n}{N} - \frac{2\pi k n}{N}) \tag{2.41}$$

$$\chi \sin(\frac{2\pi k_{bias} n}{N}) b_k \sin(\frac{2\pi k n}{N}) = \frac{\chi}{2} \cos(\frac{2\pi k_{bias} n}{N} - \frac{2\pi k n}{N}) \tag{2.42}$$

We can now write the demodulated timestream in its fourier basis where $\frac{2\pi k' n}{N} = \frac{2\pi k_{bias} n}{N} - \frac{2\pi k n}{N}$:

$$D_n = \frac{\chi}{2} \sum_{k'=0}^{N-1} (b_{k'} + b_{k'}) \cos(\frac{2\pi k' n}{N}) + (a_{k'} - a_{-k'}) \sin(\frac{2\pi k' n}{N}) \tag{2.43}$$

With the goal of having the spectrum have the same variance, we find $\chi = \sqrt{2}$, noting that $(< a_{k'} - a_{-k'})^2 >= \sqrt{2} < a_k^2 >$. To have a one to one mapping of white noise level between the input spectrum and the demodulated I spectrum, we demodulate with a signal $\sqrt{2}\sin(\omega t)$.

Power Noise Demodulation

The current, I_{rms}, in the TES calculation of the responsivity 2.33, $s_i = \frac{\delta I_{rms}}{\delta P}$, is the RMS current of the sine wave, meaning we need to scale that by $\sqrt{2}$ to get the amplitude. For some power signal $p(t)$, the response current has the form:

$$i(t) = \sqrt{2} s_i p(t) \sin(\omega t) \tag{2.44}$$

Using the normalization derived in 2.4, this demodulates to:

$$\sqrt{2}\sin(\omega t) i(t) = 2\sin(\omega t) s_i p(t) \sin(\omega t) \overset{lowpass}{\rightarrow} s_i p(t) \tag{2.45}$$

Noise from Sources After DC SQUID Amplification

Because it is convenient to refer to the noise in terms of the current at the input coil to the DC SQUID, noise that is generated in the amplifier chain after the DC SQUID gets divided by the amplification of the DC SQUID until it is the equivalent of current noise at the SQUID coil. When operating with static nulling this representation of the noise is valid. With DAN feedback this situation is somewhat more complicated. To first order, the current generated by the DAN feedback loop is set to cancel the measured current at the ADC. For noise generated before the DC SQUID in the amplifier chain this results in DAN noise being identical to the FLL operated noise. For noise that is generated after DC SQUID, the DAN feedback loop also creates current that cancels out that noise. This current, however, needs to go through the entire amplification chain to cancel it out. Because some of the current is diverted through the millikelvin wiring rather than flowing entirely in the SQUID, the current needed to cancel the post SQUID noise is larger by a factor of r_N^{-1}.

Sources of Noise

Power Noise

Noise in the power deposited on the TES comes from the quantization of power carriers. These power carriers are the optical photons, the phonons and electrons carrying thermal energy, and the electrons carrying electrical energy. In practice, the shot noise from electrons carrying electric power is minimal so we will neglect that.

The shot noise from the random arrival of photons produces a noise equivalent power (NEP) of:

$$NEP^2 = 2h\nu P_{op} + 2\frac{P_{op}^2}{\Delta\nu} \tag{2.46}$$

This equation comes from [109] with a conversion from occupation number to power.

The thermal carriers between the TES and the bath are electrons and phonons. These thermal carriers are also quantized. The expression for their noise can be found in [62].

$$NEP^2 = 4k_B T_c^2 G\gamma \tag{2.47}$$

The γ factor is an order unity factor that accounts for the thermal link being colder closer to the bath. k_B is Boltzmann's constant.

These NEPs manifest as demodulated noise with our coupling factor of s_i assuming the $\sqrt{2}\sin(\omega t)$ demodulation.

Current Noise

Every resistor in the system generates Johnson-Nyquist noise with a voltage variance per Hz of:

$$< v^2 >= 4k_B TR \tag{2.48}$$

The various amplifiers, SQUIDs, DACs and ADCs also create white noise. These manifest as a total readout noise of $\approx 12pA/\sqrt{Hz}$ at the input of the SQUID coil for SPTpol when properly accounting for amplifier noise.

2.5 Response To An External Voltage

With the frequeny multiplexed system we can measure the responsivity of a dark TES through purely electrical means. We do this by sending in a small voltage signal at a frequency other than the carrier frequency. This dissipates power on the TES which the TES then responds to.

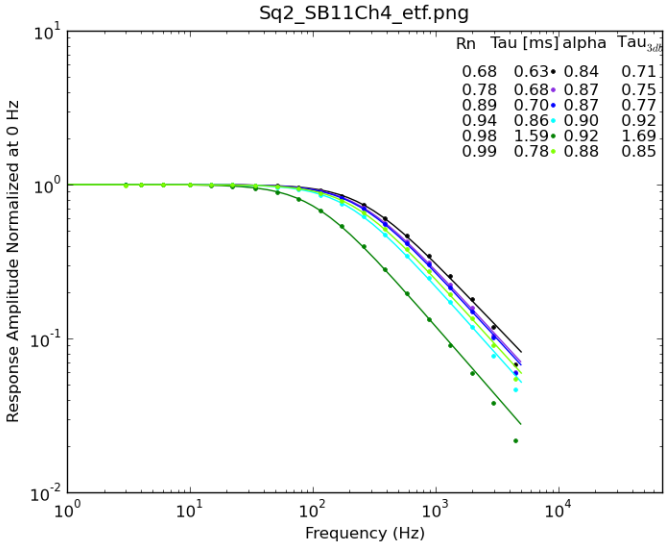


Figure 2.9: The frequency response of a TES measured with direct electrical stimulation. The frequency response has the one pole structure predicted by our model.

For a tickle signal, $v_t(t) = v_t \sin((\omega + \delta\omega)t)$ and a carrier signal, $v_c(t) = v_c \sin(\omega t)$ the power dissipated on the bolometer is:

$$P = \left(\frac{v_c \sin(\omega t) + v_t \sin((\omega + \delta\omega)t)}{R + R_s}\right)^2 R = \frac{(v_c R \sin(\omega t))^2}{(R+R_s)^2} + \frac{(v_t R \sin((\omega+\delta\omega)t))^2}{(R+R_s)^2} \quad (2.49)$$

$$+ \frac{2 v_c v_t R \sin(\omega t)\sin((\omega+\delta\omega)t)}{(R+R_s)^2} \quad (2.50)$$

Removing the DC components and the terms second order in v_t we find:

$$P = \frac{2 v_c v_t R \sin(\omega t)\sin((\omega + \delta\omega)t)}{(R+R_s)^2} = \frac{v_c v_t R}{(R+R_s)^2}(\cos(\delta\omega t) - \cos(2\omega t + \delta\omega t)) \overset{lowpass}{\rightarrow} \frac{v_c v_t R \cos(\delta\omega t)}{(R+R_s)^2} \quad (2.51)$$

Our response current is:

$$I(t) = \frac{2 v_t R \cos(\delta\omega t)}{(R+R_s)^2} \times \sin(\omega t) \times -\frac{R+R_s}{R-R_s}\frac{L}{L+1}\frac{1}{1 + j\omega\frac{\tau_0}{1+L}} \quad (2.52)$$

The factor of 2 appears in this equation because in our calculation of the responsivity of the TES, the current and voltage terms were all written in terms of the RMS values, while for this derivation we use the amplitude of the sine waves. One factor of $\sqrt{2}$ appears from the V_{rms} in the denominator of the equation and one factor of $\sqrt{2}$ appears from the fact that our responsivity is defined in terms of the change to RMS current, δI_{rms}.

Simplifying things a bit gives:

$$I(t) = -2v_t \cos(\delta\omega t)\sin(\omega t)\frac{R}{R^2 - R_s^2}\frac{L}{L+1}\frac{1}{1+j\omega\tau} \tag{2.53}$$

$$I(t) = -v_t \frac{R}{R^2 - R_s^2}\frac{L}{L+1}\frac{1}{1+j\omega\tau}(\sin(\delta\omega t + \omega t) + \sin(\omega t - \delta\omega t)) \tag{2.54}$$

In the $R_s << R$ and $L >> 1$ limit the signal we sent in is perfectly canceled by the response of the TES and a signal with identical magnitude appears reflected across the carrier frequency.

2.6 Response To An Internal Voltage Source

The TES response to a voltage source in good thermal contact with the TES is slightly different than the response to an external voltage source. For an external voltage source, power is deposited on the TES in the form of ohmic heating. For an internal voltage source, we also need to account for the work done on that voltage source by the bias current. The only internal voltage source we will concern ourselves with is the Johnson-Nyquist noise. The bias current does work on the voltage fluctuations caused by electron thermal fluctuations. The analysis of the TES Johnson noise in the DC bias case can be found in [37]. For the AC biased case, we will make the simplifying assumption that the circuit consists entirely of the TES, R, a voltage source $v_c(t)$, and some off-island resistance, R_s. The Johnson noise acts as a voltage source in good thermal contact with the TES, $v_N(t)$.

$$P_e = IV_{bolo} = \left(\frac{v_c + v_N}{R + R_s}\right)\left[v_c - \left(\frac{v_c + v_N}{R + R_s}\right)R_s\right] \tag{2.55}$$

Dropping terms that are second order in v_N we find:

$$P_e = v_c^2\frac{R}{(R+R_s)^2} + v_N v_c\frac{R - R_s}{(R+R_s)^2} \tag{2.56}$$

The left term is the steady state power supplied to the TES by the bias current. The right term is our applied power. We harmonically expand our Johnson noise:

$$v_N(\omega_0) = \int_{-\infty}^{\infty} v_N(t)e^{j\omega_0 t}d\omega_0 \tag{2.57}$$

And then solve for the measured current for an individual mode $\omega_0 = \omega + \Delta\omega$ for our carrier frequency, ω.

$$v_N(t) = N\sin((\omega + \Delta\omega)t) \tag{2.58}$$

$$v_C(t) = C\sin(\omega t) \tag{2.59}$$

$$P_1 = v_N v_c \frac{R - R_s}{(R + R_s)^2} = NC \sin((\omega + \Delta\omega)t) \sin(\omega t) \frac{R - R_s}{(R + R_s)^2} \overset{lowpass}{\rightarrow} \frac{NC}{2} \frac{R - R_s}{(R + R_s)^2} \cos(\Delta\omega t) \tag{2.60}$$

The power dissipated on an internal voltage source, in the low R_s limit, differs by a factors of 1/2 from the external voltage source. The two currents of interest are the response current of the TES to the Johnson noise:

$$I_r(t) = -\frac{N}{R + R_s} \frac{L}{L + 1} \frac{1}{1 + j\Delta\omega \frac{\tau_0}{1+L}} \left[\frac{1}{2} \sin(\omega t + \Delta\omega t) + \sin(\omega t - \Delta\omega t) \right] \tag{2.61}$$

And the regular current generated by the Johnson noise:

$$I_N(t) = \frac{v_N}{R + R_s} = \frac{N \sin(\omega t + \Delta\omega t)}{R + R_s} \tag{2.62}$$

In the high loopgain $L >> 1$, low frequency $\Delta\omega << \frac{1+L}{\tau_0}$ limit the sum of these currents is:

$$I_t = \frac{1}{2} \frac{N}{R + R_s} \left[\sin(\omega t + \Delta\omega t) - \sin(\omega t - \Delta\omega t) \right] \tag{2.63}$$

When we demodulate these currents with a sine wave at our carrier frequency we find the pleasant result:

$$\frac{1}{2} \frac{N}{R + R_s} \left[\sin(\omega t + \Delta\omega t) - \sin(\omega t - \Delta\omega t) \right] \sin(\omega t) \overset{lowpass}{\rightarrow} \frac{1}{4} \frac{N}{R + R_s} \left[\cos(\Delta\omega t) - \cos(-\Delta\omega t) \right] = 0 \tag{2.64}$$

In effect, the Johnson noise from the TES in the I phase of our demodulation is canceled out. It is important to note that the Johnson noise from the TES is not canceled in the Q phase demodulation. This result mirrors the canceling of Johnson noise in DC biased case.

2.7 Bolometer Biasing

Stable operation of a TES bolometer requires that $\lambda < 0$. For purely real impedances, this means that any resistance not in thermal contact with the TES be smaller than the resistance of the TES. For purely imaginary stray impedances, this constraint is loosened to be that the magnitude of the stray impedance is less the three times the TES resistance.

In practice, these strays are a combination of real and reactive impedances. With the FMux system, the circuit supplying power is relatively complicated. We can write down a relatively simple form for the thevenin equivalent impedance of the circuit supplying power:

$$Z_{eff} = j\omega L + \frac{1}{j\omega C} + \frac{Z_p Z_s}{Z_p + Z_s} \tag{2.65}$$

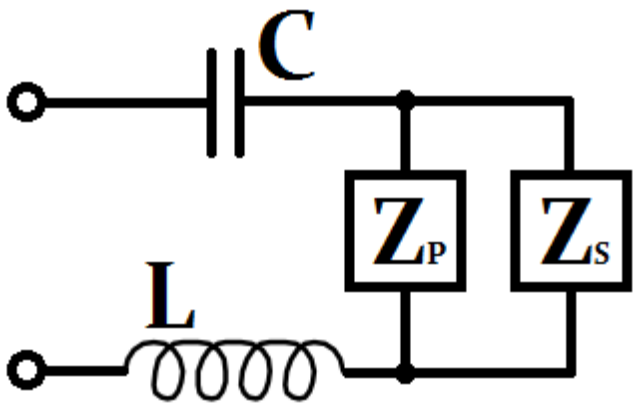


Figure 2.10: The circuit providing power to the TES, or what the readout looks like from the TES's perspective.

To first order λ is a function of Z_{eff}. For the FMux system Z_s is dominated by the inductance of the SQUID coil and the inductance of the transmission lines between the 4K SQUIDs and the millikelvin stage. We model $Z_s = j\omega L_{par}$. Z_p consists of all of the other LCR filters in parallel, capacitance of the transmission lines and any termination resistors placed at the end of the transmission lines.

In the limit of $Z_p \to \infty$, the impedance of the circuit supplying power has the form:

$$Z_{eff} = j\omega(L + L_{par}) + \frac{1}{j\omega C} \tag{2.66}$$

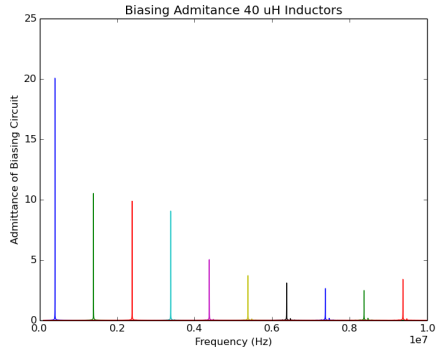
Which notably has a point where the circuit supplying power has zero impedance. The effective impedance of the bias circuit depends strongly on what the ratio of the impedance parallel to the stray series impedance. Even with an extremely large parasitic inductance it is possible to operate TES bolometers with the frequency multiplexed readout in the limit of $Z_p \to \infty$. Many discussions of supplying a voltage bias with the FMux system just discuss the magnitude of the parasitic inductance and claim that it must be smaller than the TES resistance. This is at best intentionally misleading, and at worst shows woeful misunderstanding of how circuits and electromagnetism works on a very basic level. Various admittance calculations for possible strays are plotted in figure 2.11.

Frequency Selection

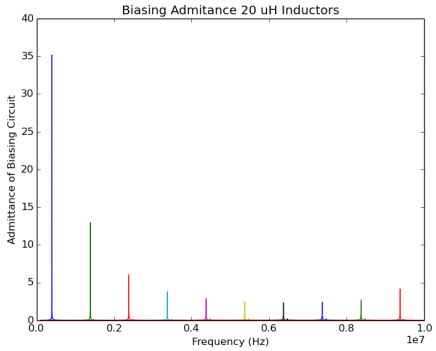
The LC resonator frequencies can shift as a function of temperature. Several factors contribute to this. The capacitor dielectric constants are a function of temperature. The geometry of the system is affect by thermal contraction. The kinetic inductance of the superconducting inductors is also a function of temperature[4]. Because we cannot know the

Figure 2.11: The admittance of the circuit providing the bias voltage. This was estimated with $Z_s = 400nH$, with 99 LCR filters separated by 100kHz, a TES resistance of 1Ω, a capacitance of 1nF included in Z_p, and $40uH$ LC inductors. Each plot includes the overplotted admittance for 10 different detectors. For stable operation, we require admittances in excess of 1 mho. At the frequency range of interest, the parasitic inductance has an impedance in excess of 10Ω. Each plot has slight variations on the circuit parameters. Including a termination resistor, a resistor is added in parallel to Z_p, significantly degrade our ability to deliver a voltage bias to the detectors. This resistor lowers the value Z_p which increases the amount of parasitic inductance that can be removed by the capacitor in the LC filter of the circuit. In short, using a small termination resistor degrades our ability to bias the detectors because it effectively increased the impedance of the bias circuit.

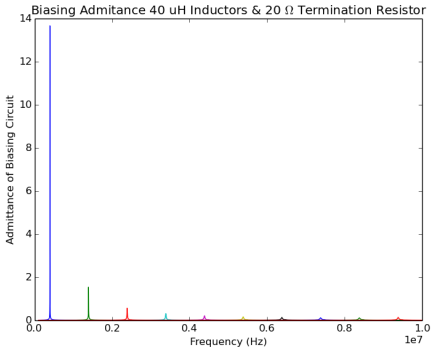
(a) Baseline circuit



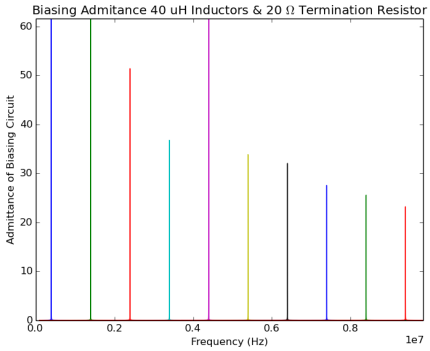
(b) LC filter inductance lowered to $20nH$.



(c) 20Ω termination resistor included.



(d) Parasitic inductance reduced to $100nH$.



impedance of the millikelvin circuitry prior to cooling the LC resonators, we measure the impedance of the circuit cold. A circuit model is fit to the admittance measurement and frequencies are selected from the fit. In order to reduce the effects of harmonic and intermodulation distortion, the selected frequencies are all multiples of 117 Hz. The only meaning of the 117 Hz is that this frequency is larger than the science bandwidth. Any harmon-

ics or intermodulation products will show up at the bias frequencies or shifted by 117 Hz. When selecting the frequencies it is also important that they are not representable as simple fractions of the clock frequency of the digital synthesizer. Simple fractions means fractions where the denominator is a small integer. Sine waves constructed at those frequencies are sampled very coarsely and end up having a lot of harmonic distortion.

Chapter 3

Map Making

3.1 The Map Making Problem

Making maps entails converting time ordered detector signals into a measurement of the intensity and polarization on a specific patch of sky. This subject has been discussed quite extensively in the literature [45, 46, 21, 77, 10, 93, 78, 105, 91, 19, 31].

SPT uses a biased map maker that is based on the MASTER method [31] mapmaker that has been extended to work for polarized data. Our map maker produces maps that are convolved with the beams. Whenever we produce scientific data products we need to correct for this bias and our beams. To motivate this choice of mapmaker, we first need to work through the math of a maximum likelihood mapmaker. This derivation is almost a direct replication of the math in the polarized MADAM mapmaker paper [46] though it has been simplified somewhat. For all of this we assume that the noise for each detectors sample is drawn from a Gaussian distribution. This derivation works if the noise for different detectors and samples is correlated. This assumption of Gaussianity is not true, but for now let us pretend it is and we will justify it later.

Pixel Response Parametrization

For a polarized CMB map, at each pixel there are three values of interest, $\boldsymbol{I}$, $\boldsymbol{Q}$ and $\boldsymbol{U}$. These correspond to their respective stokes parameters. Since our telescope is primarily imaging the CMB we usually write $\boldsymbol{I}$ as $\boldsymbol{T}$. That is to say we label the intensity with $\boldsymbol{T}$ since the intensity corresponds to the temperature of the CMB at that patch of sky. Because this is the way it is written in the code implementing this, we will use the $\boldsymbol{T}$ convention. We have neglected the $\boldsymbol{V}$ stokes parameter because our detectors are not sensitive to circularly polarized light and the polarized radiation from the CMB has no (or at least very little) $\boldsymbol{V}$ component[75]. On the celestial sphere, $Q = 1$ when the linearly polarized light is oriented towards the pole of our coordinate system. Unfortunately, two different conventions for the the definition of U differing by a sign exist. The convention used in most CMB experiments for U is flipped

in sign from the IAU prescribed format. In spt3g_software and sptpol_software, we have used the IAU definition of U. See figure 3.1.

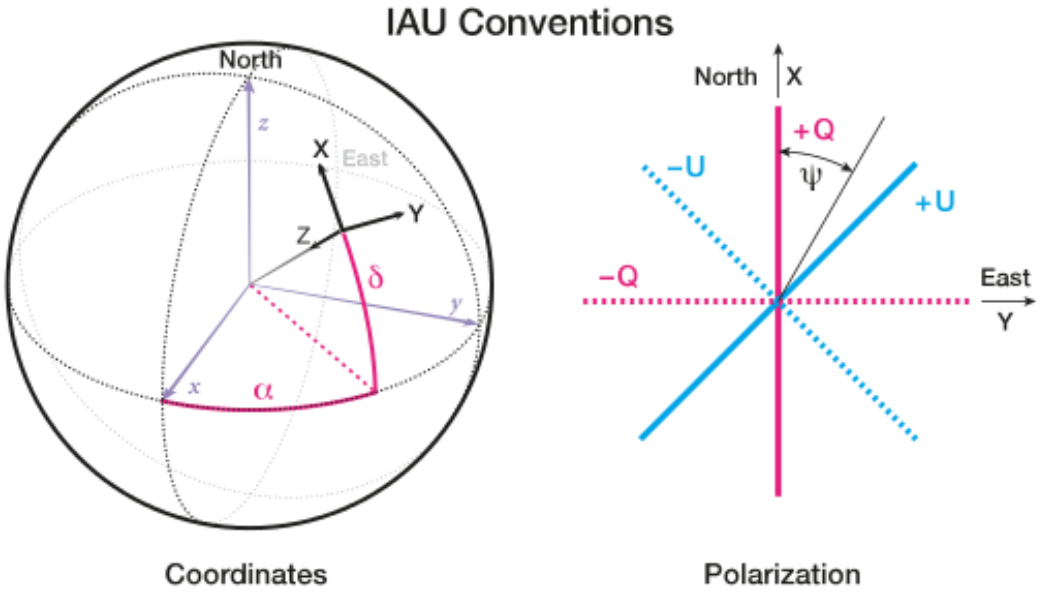


Figure 3.1: The polarization convention used defines Q and U relative to the $\hat{\alpha}$ direction. This image was taken from [36].

For a noiseless detector looking at a patch of sky with stokes parameters (T_t, Q_t, U_t) we measure a signal y_t of:

$$y_t = T_t + \frac{\gamma}{2-\gamma}(Q_t \cos(2\psi_t) + U_t \sin(\psi_t)) \tag{3.1}$$

Where we have parametrized the detectors polarization response with the polarization angle ψ and the polarization efficiency γ. In antenna literature this polarization efficiency is labelled the cross-polarization. It characterizes the detectors response to a polarized signal orthogonal to its polarization angle. For a detector with no cross-polarization, $\gamma = 1$. For a detector with no polarized response, $\gamma = 0$. In this definition of detector response we have normalized it so the signal $y_t = T_t$ when viewing an unpolarized source.

Map Making

The general map making procedure is to write down an equation for our data in terms of the map, convert that into a likelihood of our data in terms of our map, $P(\boldsymbol{y}|\boldsymbol{m})$, and then solve for the map that maximizes this likelihood. All of this math is just a replication of the MADAM paper[46]. To make the math easier, we use vectors to represent maps and

timestreams. We combine all of the time ordered data into one vector $\boldsymbol{y}$. We combine all of the pixels of the T,Q and U maps into one vector $\boldsymbol{m}$. This allows us to write where our detectors are pointing as a matrix, $\boldsymbol{P}$, and the noise covariance as a matrix, $\boldsymbol{C_n}$. In the actual processing, the data is separated by timestream and by scan. We also store the maps as separate maps of T, Q and U rather than one vector.

We start with a simple form of our time ordered data $\boldsymbol{y}$:

$$\boldsymbol{y} = \boldsymbol{P}\boldsymbol{m} + \boldsymbol{n} \tag{3.2}$$

$\boldsymbol{y}$ is a vector containing every detectors data. It has length $N_{detsamps}$. $\boldsymbol{m}$ is our map of the sky, a vector of $3N_{pixel}$ length. $\boldsymbol{P}$ is the pointing matrix that maps between detector TOD space and map space. It has dimension $(N_{detsamps}, 3N_{pixels})$. This matrix encodes where each detector is looking on the sky and its polarized response. $\boldsymbol{n}$ is a vector of the added noise for each sample.

With our assumption of Guassian noise, we can write down the likelihood of our data:

$$P(\boldsymbol{y}|\boldsymbol{m}) = (2\pi\boldsymbol{C}_n)^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{n}^T\boldsymbol{C}_n^{-1}\boldsymbol{n}\right) \tag{3.3}$$

$\boldsymbol{C}_n = \left\langle \boldsymbol{n}\boldsymbol{n}^T \right\rangle$. To maximize the likelihood we can minimize its inverse logarithm. Neglecting an added constant we write:

$$\chi^2 = -2\ln P(\boldsymbol{y}|\boldsymbol{m}) = \boldsymbol{n}^T\boldsymbol{C}_n^{-1}\boldsymbol{n} = (\boldsymbol{y} - \boldsymbol{P}\boldsymbol{m})^T\boldsymbol{C}_n^{-1}(\boldsymbol{y} - \boldsymbol{P}\boldsymbol{m}) \tag{3.4}$$

We then minimize this with respect to $\boldsymbol{m}$ and find:

$$\boldsymbol{m} = (\boldsymbol{P}^T\boldsymbol{C}_n^{-1}\boldsymbol{P})^{-1}\boldsymbol{P}^T\boldsymbol{C}_n^{-1}\boldsymbol{y} \tag{3.5}$$

In the code we've labelled $(\boldsymbol{P}^T\boldsymbol{C}_n^{-1}\boldsymbol{P})$ the weight matrix. We use the moniker weighted map for $\boldsymbol{P}^T\boldsymbol{C}_n^{-1}\boldsymbol{y}$.

Now that we have gotten that relatively simple equation out of the way, it is time to mention that in practice, even with our assumptions, this form will never work. For a year of SPTpol data $N_{detsamps} \approx 10^{12}$. This means that a covariance matrix that encompasses every samples correlation to every other sample would be one million exabytes of data. For context, if one were to purchase $20 sticks of 8 gigabytes of RAM, it would cost a quadrillion dollars to purchase that much RAM. Inverting that would be nigh impossible.

There are two issues that must be addressed by every CMB map maker. How does it handle non-Gaussian noise? How does it make the noise covariance inversion tractable? SPT employs the solution presented in the MASTER paper[31]. We filter the data until the noise is white and the covariance matrix is diagonal. In all of the above equations we replace $\boldsymbol{y}$ with our now filtered $\boldsymbol{y_{filtered}}$ and then solve for the map assuming a diagonal covariance matrix. Because our data has been filtered the map produced is by construction not an accurate representation of that patch of sky. For any science result we produce we need to account for this bias. In the case of our N-point estimators, we usually do this via

Monte-Carlo simulations. We produce a large number of simulations of our data, apply the filters to said simulations and then repeat the map making procedure. This gives us a set of simulated maps to estimate the effects of our filtering.

There are a number of improvements we can make. We can incorporate some restricted set of correlations into the map making procedure. If we assume the noise covariance terms are only non-zero for some small time scale, we can make the noise covariance matrix block diagonal. This would make the inversion process tractable. There are a number of improvements we can also make to the noise modelling. Instead of assuming Gaussianity we can incorporate a low frequency component like MADAM[46] or we could try to incorporate some atmospheric model. A lot of work was done to make the map making code modifiable so that improvements like this are as easy as possible to incorporate.

Each part of the map making equation and the assumptions embedded in this equation needs to be discussed.

3.2 One Pixel Map

Before we get to the discussion of all the individual pieces of the map making equation, let's work through a highly simplified version of the map making problem. Let's make a map with one pixel where we have data from one sample for N detectors. Let's assume that every detector has the same noise so $\boldsymbol{C}_n = \nu \boldsymbol{I}$ where I is the identity matrix. This pixel encodes the T, Q and U information. With these assumptions we can write down our inputs into the map making equation.

$$\boldsymbol{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} \qquad \boldsymbol{P} = \begin{bmatrix} 1 & \frac{\gamma_1}{2-\gamma_1}\cos(2\psi_1) & \frac{\gamma_1}{2-\gamma_1}\sin(2\psi_1) \\ & \vdots & \\ 1 & \frac{\gamma_N}{2-\gamma_N}\cos(2\psi_N) & \frac{\gamma_N}{2-\gamma_N}\sin(2\psi_N) \end{bmatrix} \tag{3.6}$$

With that input our weighted map has the form:

$$\boldsymbol{P}^T\boldsymbol{C}_n^{-1}\boldsymbol{y} = \frac{1}{\nu}\begin{bmatrix} \sum y_i \\ \sum y_i \frac{\gamma_i}{2-\gamma_i}\cos(2\psi_i) \\ \sum y_i \frac{\gamma_i}{2-\gamma_i}\sin(2\psi_i) \end{bmatrix} = \frac{1}{\nu}\begin{bmatrix} \sum (T + \frac{\gamma_i}{2-\gamma_i}(Q\cos(2\psi_i) + U\sin(2\psi_i)) + n_i) \\ \sum (T + \frac{\gamma_i}{2-\gamma_i}(Q\cos(2\psi_i) + U\sin(2\psi_i)) + n_i)\frac{\gamma_i}{2-\gamma_i}\cos(2\psi_i) \\ \sum (T + \frac{\gamma_i}{2-\gamma_i}(Q\cos(2\psi_i) + U\sin(2\psi_i)) + n_i)\frac{\gamma_i}{2-\gamma_i}\sin(2\psi_i) \end{bmatrix} \tag{3.7}$$

This equation becomes a lot cleaner if we make a few assumptions. Let's imagine we have a bunch of detectors with random polarization angles and which are 100% polarized, $\gamma = 1$. So $\langle \sin(2\psi) \rangle = \langle \cos(2\psi) \rangle = \langle \sin(2\psi)\cos(2\psi) \rangle = 0$ and $\langle \sin^2(2\psi) \rangle = \langle \cos^2(2\psi) \rangle = 1/2$. In that case our weighted map is:

$$\boldsymbol{P}^T\boldsymbol{C}_n^{-1}\boldsymbol{y} = \frac{1}{\nu}\begin{bmatrix} \sum T \\ \sum \frac{1}{2}Q \\ \sum \frac{1}{2}U \end{bmatrix} \tag{3.8}$$

Applying the same math for our weight matrix we find:

$$(\boldsymbol{P}^T\boldsymbol{C}_n^{-1}\boldsymbol{P}) = \frac{1}{\nu}\begin{bmatrix} 1 & \sum \frac{\gamma_i}{2-\gamma_i}\cos(2\psi_i) & \sum \frac{\gamma_i}{2-\gamma_i}\sin(2\psi_i) \\ \sum \frac{\gamma_i}{2-\gamma_i}\cos(2\psi_i) & \sum(\frac{\gamma_i}{2-\gamma_i})^2\cos^2(2\psi_i) & \sum(\frac{\gamma_i}{2-\gamma_i})^2\cos(2\psi_i)\sin(2\psi_i) \\ \sum \frac{\gamma_i}{2-\gamma_i}\sin(2\psi_i) & \sum(\frac{\gamma_i}{2-\gamma_i})^2\sin(2\psi_i)\cos(2\psi_i) & \sum(\frac{\gamma_i}{2-\gamma_i})^2\sin^2(2\psi_i) \end{bmatrix} \tag{3.9}$$

Using the same assumptions we made above we can write this weight matrix in a really simple form:

$$(\boldsymbol{P}^T\boldsymbol{C}_n^{-1}\boldsymbol{P}) = \frac{1}{\nu}\begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix} \tag{3.10}$$

With a quick inversion and multiplication we can see that we are able to replicate our (T, Q, U) map. Let's now consider the weight matrix in a different limit. Let's say we only have one detector with polarization angle $\psi = 0$.

$$(\boldsymbol{P}^T\boldsymbol{C}_n^{-1}\boldsymbol{P}) = \frac{1}{\nu}\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{3.11}$$

This matrix is not invertible. This is just a statement that if we only measure the sky with a detector with only one polarization angle we cannot accurately measure the T, Q and U maps since we have absolutely no information about any polarization information orthogonal to the direction of our detector. This is just a restatement of the fact that if we are trying to measure three numbers we need at least three independent measurements of them.

In practice, some pixels on the edge of our maps are only observed by a very small subset of the detectors. This results in a weight matrix with very poor conditioning. These pixels in the output map often have extremely large values because of numerical accuracy issues with the matrix inversion. Any pixels with a poorly conditioned weight matrix should not be included in any analysis.

3.3 Pointing Matrix

Each row of the pointing matrix encodes how much response each detector is getting from each map pixel. This row of information is a combination of where the detector is pointing and its sensitivity to polarization.

Beams

Our detectors response to a point on the sky has a finite extent. If an unpolarized detector is looking at a point on the sky $\boldsymbol{p}$ with our unpolarized beam $B_T(\boldsymbol{x})$ we can write its response as:

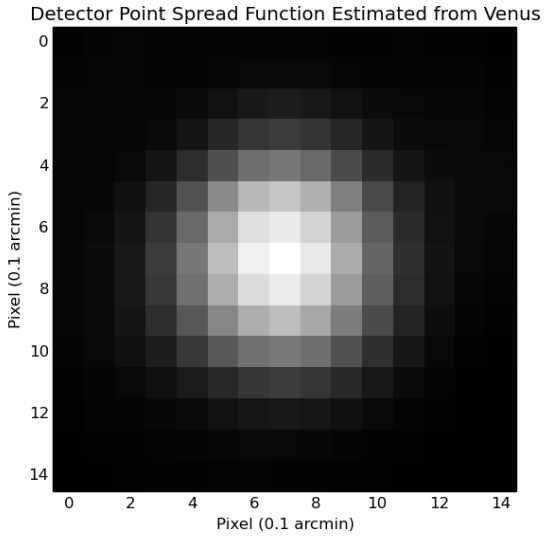


Figure 3.2: This is an individual detector's response to a point source as a funciton of the detector pointing.

$$y_t = \int\int B_T(\boldsymbol{x}) T(\boldsymbol{p} - \boldsymbol{x}) d^2\boldsymbol{x} \tag{3.12}$$

If we blindly mapped this to our pointing matrix, each row in the pointing matrix would contain a mapping of the detectors response to each pixel on the map. This would result in many non-zero entries in our pointing matrix. We can reduce this number, however, if we make the assumption that every detector has the same beam. We can redefine our map to be the beam convolved map. That is to say:

$$\boldsymbol{m}(\boldsymbol{p}) = \int\int B(\boldsymbol{x}) T(\boldsymbol{p} - \boldsymbol{x}) d^2\boldsymbol{x} \tag{3.13}$$

In our map making equation we are now solving for this beam convolved map and our pointing matrix maps where each detector is pointing to a unique pixel in the map.

Extending this to polarization requires a few extra assumptions. Writing down the response again and assuming our polarized beam is the same as the intensity beam:

$$\begin{aligned} y_t &= \int\int B_T(\boldsymbol{x}) T(\boldsymbol{p} - \boldsymbol{x}) + \frac{\gamma}{2-\gamma} \cos(2\psi) B_Q(\boldsymbol{x}) Q(\boldsymbol{p} - \boldsymbol{x}) + \frac{\gamma}{2-\gamma} \sin(2\psi) B_U(\boldsymbol{x}) U(\boldsymbol{p} - \boldsymbol{x}) d\boldsymbol{x} \\ &= \int\int B(\boldsymbol{x}) \left(T(\boldsymbol{p} - \boldsymbol{x}) + \frac{\gamma}{2-\gamma} \cos(2\psi) Q(\boldsymbol{p} - \boldsymbol{x}) + \frac{\gamma}{2-\gamma} \sin(2\psi) U(\boldsymbol{p} - \boldsymbol{x}) \right) d\boldsymbol{x} \end{aligned}$$

A couple sins were embedded these assumptions/equations.

- The polarized beams are the same as the intensity beams.

- The cross polarization is independent of location in the beam.

For SPTpol analyses we have mostly assumed these to be true. Some limitted constraints were placed on the leakage. Characterizing these beam effects is still an open problem and should be done for SPT-3G. I cannot emphasize enough how important properly characterizing our detectors response to a point source is. Every drop of science we produce depends on doing this accurately.

Once we have committed to these assumptions, we can just assume our Q and U maps are the beam convolved maps allowing us to reduce the number of non-zero entries in each row of the pointing matrix to three (T,Q,U). In the map making software, we store this pointing matrix as the pixel each detector is pointing at at each sample. We store the polarization angle and polarization efficiency for each detector.

The upside to committing to making beam convolved maps is that every pixel in the map making equation is independent. Any knowledge we glean from studying the one-pixel map can be extended to a full map.

Calculating the Pointing Matrix

We can divide estimation of the pointing matrix into several steps:

- Constructing boresight pointing model.
- Estimating detector pointing offsets from boresight.
- Measuring detector polarization angle and efficiency.
- Applying all these numbers to create the pointing matrix.

Boresight Pointing Model Construction

To zeroth order, we use various sensors/encoders on the telescope to create a best guess of where the telescope is pointing. This model is tested by observing point sources with known celestial coordinates. The lion's share of the work was done by Jason Henning. This is just a brief summary of his work.

For SPTpol this model includes a variety of effects. The corrections to the azimuth and elevations, δaz and δel, have the form:

$$\delta az = -(a_2 \cos[az] + a_3 \sin[az]) \tan[el] + (a_4 - det) \tan[el] - a_5 / \cos[el] - az_0 \quad (3.14)$$

$$\delta el = a_0 \sin[el] + a_1 \cos[el] + (a_2 \sin[az] - a_3 \cos[az]) - del - a_6 - \theta_{refr} \quad (3.15)$$

Where the various variables in this model are:

$\boldsymbol{a_0}$, $\boldsymbol{a_1}$ Boom flexure. This is how much the boom arm flexes during operation. The correction is obtained from point source fits.

a_2, a_3 Azimuth axis tilt. The bearing the telescope is on slowly tilts. Encoder measurements are used to correct for this.

a_4 Elevation axis tilt.

az_0 Azimuth encoder offsets. The encoders on the telescope used to measure the azimuth and elevation slowly drift. Direct measurements of this drift are done to correct this.

dET, dEL Metrology and thermometry corrections. Weather and temperature can influence the geometry of the telescope. This is modelled as a simple offset of the pointing.

θ_{refr} Refraction. Light refracts as it travels through the atmosphere. A simple atmospheric model is used to correct this.

a_5, a_6 Elevation collimation. This is the difference in the optical axis from the physical elevation axis. This is measured with observations of HII regions.

For SPTpol, we ended up with errors on the order of 10 arcseconds. See figure 3.3.

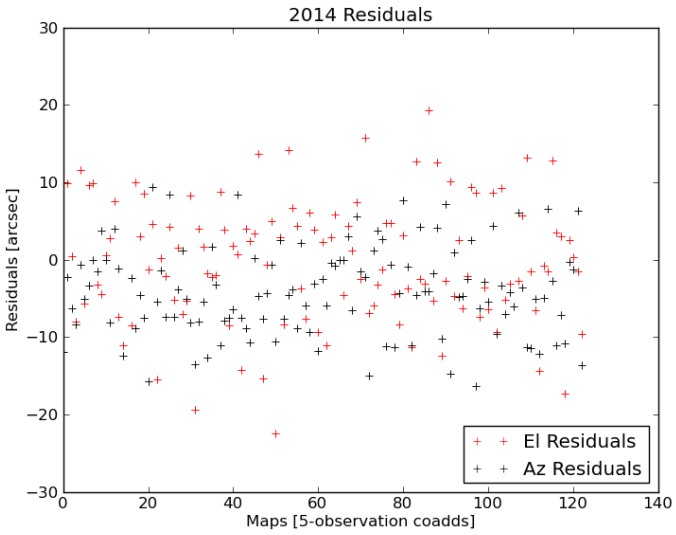


Figure 3.3: The error in the boresight pointing. This plot was generated by Jason Henning.

Detector Pointing Offset

The detector pointing offsets from boresight pointing are estimated by observing a celestial source and then fitting for the offsets. We use RCW 38 for this measurement. RCW 38 is used because its flux is stable in time and its intensity is in a Goldilocks zone. It is bright enough that we can observe it quickly with good signal to noise. It also is not so bright as to create worrying non-linear response effects in our transition edge sensors.

With our telescopes resolution, RCW 38 is an extended source. See figure 3.4. In order to measure the detector pointing offsets, we first generate a template of RCW 38 and then fit this template to maps made from individual detector data. For SPTpol this template was generated from multiple observations of RCW 38. For SPT-3G, this template is constructed from the observation we are fitting to. This process is done iteratively. Our best guess of pointing offsets is used to construct the initial map. We then fit individual detector maps to this sum map and extract pointing offsets. These pointing offsets are then used to make a better estimate of the sum map and we repeat the procedure. We iterate until the detector offsets converge. These detector offsets are stored in telescope local Azimuth and Elevation coordinates.

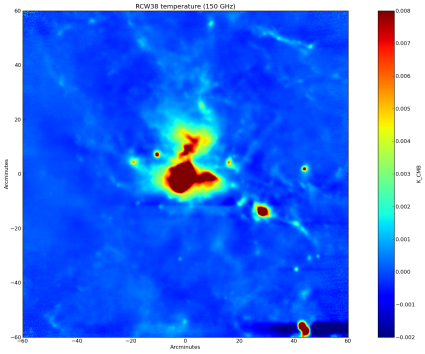


Figure 3.4: RCW 38 at 150GHz. This plot was generated by Stephen Hoover.

Polarization Angle and Efficiency

For SPTpol, a calibration source was used to measure the polarization angle and efficiency. JT Sayre, Tyler Natoli and Ryan Keisler did most of the work on this. The general idea was to place a modulated polarized calibration source 3 km away from the telescope, look at it and measure the detector response as a function of the polarization angle. From these measurements we can then fit for the polarization angle and efficiency. This is a terribly brief description of a lot of work done to get this measurement.

For SPT-3G, there are no plans to repeat this polarization angle measurement. There are a couple driving factors for this choice. First, making the measurement for SPTpol required a herculean effort. Extending this measurement for the increased detector count is not entirely feasible. The other reason relates to the new antenna design. For the sinuous antenna technology employed by SPT-3G, the polarization angle axis varies as a function of frequency[94]. This means the polarization angle and polarization efficiency we use in our map making is a function of the spectrum of the source we are measuring.

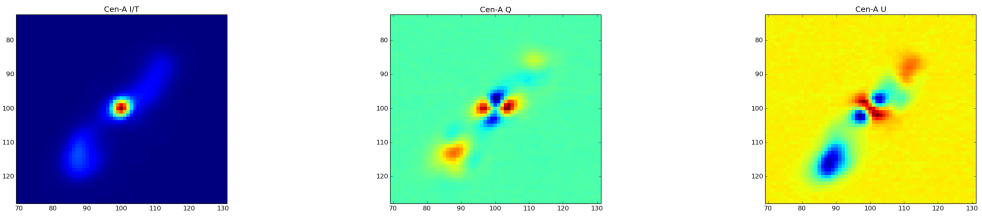
Some initial work has been done trying to estimate polarization angles for individual detectors from celestial sources. The general idea is to make a template of a celestial source,

either the CMB or Centaurus A, and then fit individual detector maps to this sum map. The fit varies the polarization efficiency and angle of the detector. This potentially will work up to a global rotation angle error and a global polarization efficiency error. For the polarization efficiency, we will need to use another CMB telescope to measure it. For the polarization angle error we can exploit the statistical properties of the CMB. The EB power spectrum of the CMB is consistent with zero[75]. This power spectrum is sensitive to this global rotation error. This allows us to fit from our own data this mean error.

We already employ the assumption of the EB power spectrum being consistent with zero to constrain a global rotation error with SPTpol's analysis. Some work has been done fitting polarization angles from our individual detector maps. This analysis so far has not created usable data products. A quick test of this method was done using Centaurus A's radio lobes as a polarized calibration source.

Figure 3.5: Centaurus A Template

(a) T (b) Q (c) U



Actually making this technique work is crucial for making SPT-3G work, and still needs to be done. The backup plan if self calibration doesn't work is to fit individual detector angles to Planck maps.

Converting Estimated Parameters to Map Pixel

Calculating Detector Pointing

The full calculation converting azimuth and elevation to a celestial coordinate system like equatorial coordinates is actually rather involved and inefficient. It is not computationally feasible to make this calculation for every detector at every sample. The reason this calculation, when done properly, is inefficient is because the appropriate transform also accounts for effects beyond the coordinate rotation. These effects include the relativistic aberration of light and the gravitational deflection of the light by bodies in our solar system.

Since we can not make this full calculation many times, we just make the calculation once. We store this as the FK5 (Ra, Dec) pointing of boresight and also as a rotation. The rotation we store is a rotation that maps the coordinate (1,0,0) on the unit sphere in local coordinates to the boresight pointing of the telescope in our desired coordinate system.

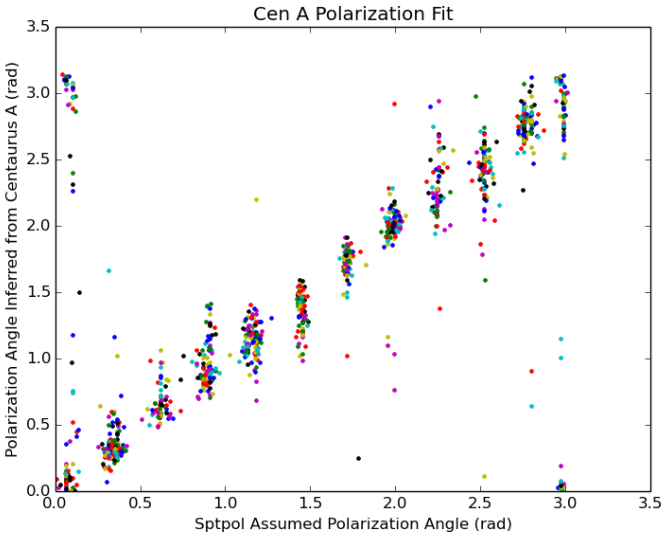


Figure 3.6: The x axis is the polarization angle fit from observations of the polarization calibrator. The y axis is the polarization angle fit from observations of the polarization structure of Centaurus A's radio lobes. The correlation between the polarization angle constructed from the two sources indicates that fitting polarization angle from the radio lobes of Centaurus A is a potential way of constraining the detector polarization angle.

Using the offsets we estimated from RCW 38, we compute the location every detector is looking on the celestial sphere if the boresight were pointing at (1,0,0) and then apply our stored rotation to get the approximate pointing of each detector in our desired celestial coordinate system. This way of calculating the detector pointing neglects any variation in aberration or refraction over the extent of the focal plane. In practice these errors are on the order of 1 arcsecond, which is well below our pointing jitter error.

An Aside About Map Projections

The mapping of right ascension and declination to a specific pixel in the map is a map projection. Map projections have quite a long and storied history[85]. For our use, this is just a function that maps (Ra, Dec) coordinates to a pixel on the map.

For most of our analyses, we require that the map projection we use be equal area. That is to say, each pixel stores an equal area on the sky. In practice there are three map projections used in analysis. We use equal area cylindrical projections, the Lambert azimuthal equal-area projection, and healpix. The first two project the sky coordinates onto a flat euclidean plane. Equal area cylindrical projections are nice because given our unique location on the Earth, the x and y axes of the map correspond to Ra and Dec. The Lambert azimuthal equal-area projection is employed, because for small maps of the celestial sphere, this projection limits

the distortion of the map. Healpix [27] is a pixelization of a sphere that is equal area. It also has a couple other nice properties for speeding up spherical harmonic transforms.

Whenever flat sky projections are used, one needs to be careful with the definition of the polarization angle. The map maker produces maps where the Q/U values fit with the IAU definition of polarization angle. That is to say, the polarization angle is defined as based off of the direction of the celestial pole. For some analyses, like a flat sky power spectrum estimation, having polarization angle defined relative to the vertical direction in the map is important. In the software, the routine that applies this correction is called flatten_pol.

Errors

There are several numbers estimated from data used in construction of the pointing matrix. Let's figure out how errors in those measurements will affect our maps.

Errors in Pointing

Errors with the boresight pointing and the detector offsets manifest as an error for where each detector is looking. With our scan strategy we observe the same patch of sky many, many times. We can divide the pointing errors into two categories. Errors that are the same every time we observe a patch of sky and errors that are different every time we observe the same patch of sky. The former manifests as a distortion of the map. How this distortion affects any analysis is a function of this distortion and the analysis. This is a terrifying source of error.

The random errors are more innocuous. With our scan strategy any individual detector will image the same patch of sky multiple times. Let's write down the average signal it sees. For a unpolarized detector looking at $\boldsymbol{p}$ with an error $\boldsymbol{e}$:

$$\begin{aligned} y &= \frac{1}{N}\sum T(\boldsymbol{p}-\boldsymbol{e}) \\ &\approx \int p(\boldsymbol{e})T(\boldsymbol{p}-\boldsymbol{e}) \end{aligned}$$

Where in the many measurement limit, we have changed our average sum into an expectation value where $p(\boldsymbol{e})$ is the probability of having an error $\boldsymbol{e}$. This form looks exactly like our equation for the beam convolved map, except in this case, the beam is our probability density. On our final summed map, any random error in pointing manifests as a convolution of the error in our pointing with the map. The jargon used for this error is the "pointing jitter". Any scientific result needs to include our pointing jitter in the definition of the beam.

Errors in Detector Polarization Sensitivity

Any errors in our estimate of a detectors polarization sensitivity will manifest as errors with our weighted map.

Since errors with the polarization efficiency is easiest to evaluate, lets once again solve for the weighted map in the large detector limit. For some error δ_i in our polarization efficiency γ_i:

$$\boldsymbol{P}^T\boldsymbol{C}_n^{-1}\boldsymbol{y} = \frac{1}{\nu}\begin{bmatrix} T \\ Q\sum \frac{\gamma_i+\delta_i}{2-\gamma_i-\delta_i}\frac{\gamma_i}{2-\gamma_i}\cos^2(2\psi_i) \\ U\sum \frac{\gamma_i+\delta_i}{2-\gamma_i-\delta_i}\frac{\gamma_i}{2-\gamma_i}\sin^2(2\psi_i)\end{bmatrix} \approx \frac{1}{\nu}\begin{bmatrix} T \\ Q\sum(\frac{\gamma_i}{2-\gamma_i}+\frac{\delta_i}{(2-\gamma_i)^2})\frac{\gamma_i}{2-\gamma_i}\cos^2(2\psi_i) \\ U\sum(\frac{\gamma_i}{2-\gamma_i}+\frac{\delta_i}{(2-\gamma_i)^2})\frac{\gamma_i}{2-\gamma_i}\sin^2(2\psi_i)\end{bmatrix} \tag{3.16}$$

The last section is in the small δ_i limit. Errors with the constraint $\langle\delta_i\rangle = 0$ will result in extra noise. While errors where $\langle\delta_i\rangle \neq 0$ manifest as a global error in the amplitude of Q and U.

Errors with the polarization angle require a bit more writing. For the sake of brevity, let's set the polarization efficiency to one. Let's assign $\delta_i/2$ to our error and take the large detector limit. Let's make the assumption that δ_i is uncorrelated with our polarization angle.

$$\begin{aligned}
\boldsymbol{P}^T\boldsymbol{C}_n^{-1}\boldsymbol{y} &= \frac{1}{\nu}\begin{bmatrix} \sum(T + (Q\cos(2\psi_i+\delta_i) + U_t\sin(2\psi_i+\delta_i))) \\ \sum(T + (Q\cos(2\psi_i+\delta_i) + U\sin(2\psi_i+\delta_i)) + n_i)\cos(2\psi_i) \\ \sum(T + (Q\cos(2\psi_i+\delta_i) + U\sin(2\psi_i+\delta_i)) + n_i)\sin(2\psi_i)\end{bmatrix} \\
&= \frac{1}{\nu}\begin{bmatrix} T \\ \sum(Q[\cos(2\psi_i)\cos(\delta_i) - \sin(2\psi_i)\sin(\delta_i)] + U[\sin(2\psi_i)\cos(\delta_i) + \cos(2\psi_i)\sin(\delta_i)])\cos(2\psi_i) \\ \sum(Q[\cos(2\psi_i)\cos(\delta_i) - \sin(2\psi_i)\sin(\delta_i)] + U[\sin(2\psi_i)\cos(\delta_i) + \cos(2\psi_i)\sin(\delta_i)])\sin(2\psi_i)\end{bmatrix} \\
&= \frac{1}{\nu}\begin{bmatrix} T \\ \sum(Q\cos(2\psi_i)\cos(\delta_i) + U\cos(2\psi_i)\sin(\delta_i))\cos(2\psi_i) \\ \sum(-Q\sin(2\psi_i)\sin(\delta_i) + U\sin(2\psi_i)\cos(\delta_i))\sin(2\psi_i)\end{bmatrix} \\
&\approx \frac{1}{\nu}\begin{bmatrix} T \\ \sum(Q\cos(2\psi_i) + U\cos(2\psi_i)\delta_i)\cos(2\psi_i) \\ \sum(-Q\sin(2\psi_i)\delta_i + U\sin(2\psi_i))\sin(2\psi_i)\end{bmatrix} \\
&\approx \frac{1}{\nu}\begin{bmatrix} T \\ \frac{1}{2}Q + \sum U\cos^2(2\psi_i)\delta_i \\ \frac{1}{2}U + \sum -Q\sin^2(2\psi_i)\delta_i\end{bmatrix}
\end{aligned}$$

In short, polarization angle errors will mix Q and U together.

3.4 Noise Estimates

In our map making algorithms we call the elements of the diagonal inverse covariance matrix the weights in analogy to weighted averages. Were our post filtering noise truly white the optimal estimate of the weight would be the inverse variance of the noise of the timestream.

Unfortunately, even post filtering, our detector noise is often not white. This turns estimation of the weight into more of an art than a science. The general method that's been employed for SPT's science results is to assign the weight to be the inverse of the power spectral density averaged over some frequency range. Since we know the scan speed of the telescope we can convert angular scales on the sky into time ordered data frequency. We pick the set of frequencies that correspond to angular scale of the science we are trying to produce. This weight assignment is only valid in the low signal to noise limit.

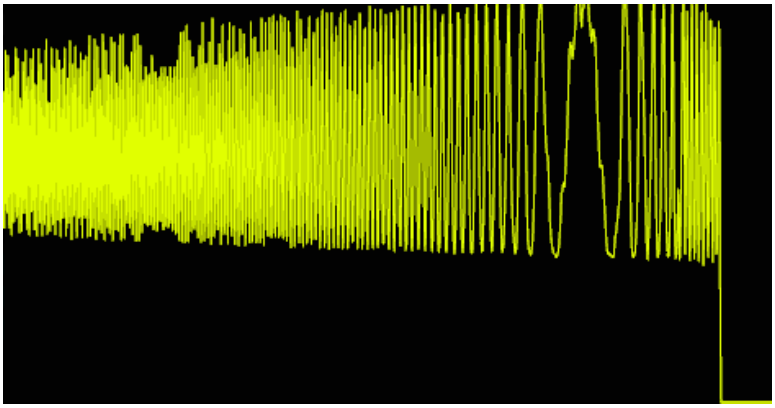


Figure 3.7: The time ordered data for a detector becoming unstable.

We also have the issue that sometimes detector time streams vary so far from a Gaussian noise distribution that the inclusion of them will corrupt the output maps. An example of this would be a transition edge sensor becoming unstable and then latching. We remove this data when making maps. This step is labelled as timestream flagging. We generally flag detectors that are not responsive or have extremely large, non-gaussian noise. For an example of this see figure 3.7.

3.5 Detector Calibration

In order to make maps of the sky in meaningful units, in our case K_{CMB}, we must convert our time ordered data into these units. The K_{CMB} units are normalized to be the change in power from a change in temperature in Kelvin of the CMB. Maps constructed with TOD in these units produce a direct map of the primary anisotropy temperature fluctuations. When observing sources with a non-CMB spectrum corrections need to be made to convert the maps into physically meaningful units.

The transition edge sensor current response to optical power has the form:

$$s_i = \frac{\delta I}{\delta P} = -\frac{1}{V_0}\frac{R+R_s}{R-R_s}\frac{\mathcal{L}}{\mathcal{L}+1}\frac{1}{1+j\omega\tau} \tag{3.17}$$

[37]

In this equation, V_0 is our voltage bias. R is the operating resistance of the TES. R_s is any off island series impedance. $\mathcal{L}$ is the loopgain. τ is the time constant of the detector. This response is frequency dependent and since R, τ and $\mathcal{L}$ are dependent on sky loading, a function of the DC optical power on the detector.

If τ is sufficiently small, we can neglect the frequency dependent response. In the case of SPTpol this was true. For SPTsz, the opposite of the detector response, $1 + j\omega\tau$, was applied to the TOD using a DFT based filter. In the case of SPT-3G, our detector response is also affected by the dynamic active nulling's transfer function. An appropriate transfer function needs to be constructed.

Our time ordered data starts its life in arbitrary counts units. To convert it to physical units we use our calibrator and a celestial source, RCW 38. The calibrator is a chopped thermal source that illuminates the focal plane. At least once every fridge cycle, we measure the integral response of each detector to RCW 38. Because our detector response is a function of sky loading we measure our detectors response to the calibrator as a function of elevation. For a given telescope elevation we scale the response to RCW 38 by the ratio of the calibrator response at our elevation to the calibrator response at RCW 38's elevation. Because the sky opacity can also change, throughout the course of observations we repeatedly measure the average response to RCW 38. We scale the response of the TES by this average change.

Conversion from RCW 38 response to our preferred unit, Kelvin, is a function of the exact structure of a detectors band. This is because the spectrum of RCW 38 differs from the CMB's spectrum by more than a multiplicative constant. The conversion between RCW38 flux and K_{cmb} is derived by cross correlating our maps or power spectra with whatever satellite experiment is most popular at the time. For SPTpol this is Planck. The important thing to remember is that this scaling factor must be different for detectors with different bands.

The last thing that necessitates mentioning is that our readout system has cross talk. The measured signal for a detector is a function of what neighbouring detectors are measuring. This can result in a map with many shadows of compact objects we are observing. See figure 3.8. We estimate this crosstalk from point source observations and then remove it before doing any science processing.

3.6 Pixel Effects

Our polarized map making procedure necessitates grouping together multiple observations of some finite patch of sky. In our case, this is an individual pixel of the map. There are a couple effects one should be aware of relating to pixelization.

Pixel Window Function

Our recorded T,Q,U values for a pixel are the average of the values over the extent of the pixel. We can write this as:

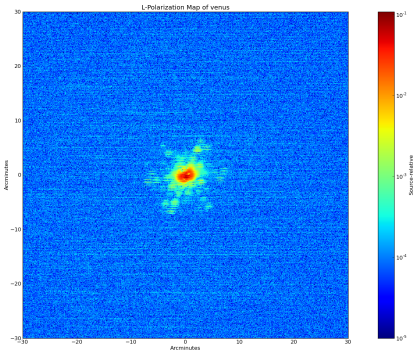


Figure 3.8: Observation of a compact object with crosstalk. This plot was generated by Stephen Hoover.

$$T_i(\boldsymbol{p}) = \int T(\boldsymbol{p} - \boldsymbol{x}) w_p(\boldsymbol{x}) d\boldsymbol{x} \tag{3.18}$$

Where $w_p(\boldsymbol{p})$ is 1 inside the pixel and 0 everywhere else. In effect, we convolve our map with something the shape of our pixel and then sample this convolved surface. This acts as a modification to our beam. For most of our N-point analyses this effect is handled through Monte Carlo simulations by using a pixel size in the input map simulation that is much smaller than the output maps. The simulations will naturally measure this pixel window function for us.

Gradients

The intensity and polarization naturally vary over the pixel. Because the intensity signal is so much larger than the polarized signal let's write down how that intensity gradient appears in our maps. Returning yet again to the single pixel model let's write down what the effect of an intensity gradient is. Without loss of generality, let's just solve this for a gradient in the declination or y direction. Once again returning to our single pixel weighted map equation and taking the large number of detectors limit:

$$\begin{aligned}
\boldsymbol{P}^T \boldsymbol{C}_n^{-1} \boldsymbol{y} &= \frac{1}{\nu} \begin{bmatrix} \sum y_i \\ \sum y_i \cos(2\psi_i) \\ \sum y_i \sin(2\psi_i) \end{bmatrix} = \frac{1}{\nu} \begin{bmatrix} \sum (T + (Q\cos(2\psi_i) + U\sin(2\psi_i) + \Delta y_i \nabla T) + n_i) \\ \sum (T + (Q\cos(2\psi_i) + U\sin(2\psi_i) + \Delta y_i \nabla T) + n_i)\cos(2\psi_i) \\ \sum (T + (Q\cos(2\psi_i) + U\sin(2\psi_i) + \Delta y_i \nabla T) + n_i)\sin(2\psi_i) \end{bmatrix} \\
&= \frac{1}{\nu} \begin{bmatrix} \sum T + \Delta y_i \nabla T \\ \sum Q\cos^2(2\psi_i) + \Delta y_i \nabla T \cos(2\psi_i) \\ \sum U\sin^2(2\psi_i) + \Delta y_i \nabla T \sin(2\psi_i) \end{bmatrix}
\end{aligned}$$

This gradient term will manifest as temperature to polarization leakage if there are any correlations between the locations on a pixel imaged by the detectors of a given polarization

angle and the gradient. As a toy example, imagine that we have an unpolarized sky with a gradient. We scan across the pixel with two detectors. One with $\psi = 0$ and one with $\psi = \pi/4$. Let's say their pointing is offset from each other by half a pixel. Because these detectors would see signals with different intensity, the solved map maker equation would find that this pixel has a non-zero polarization because two detectors with different polarization angles measured a different signal when observing the same pixel. In practice, this effect can be pretty large. Jason Gallicchio discovered this effect when generating maps with interpolated sims. Interpolated sims are where values not at pixel centers are set as the interpolated value of neighbouring pixels rather than the closest pixel value.

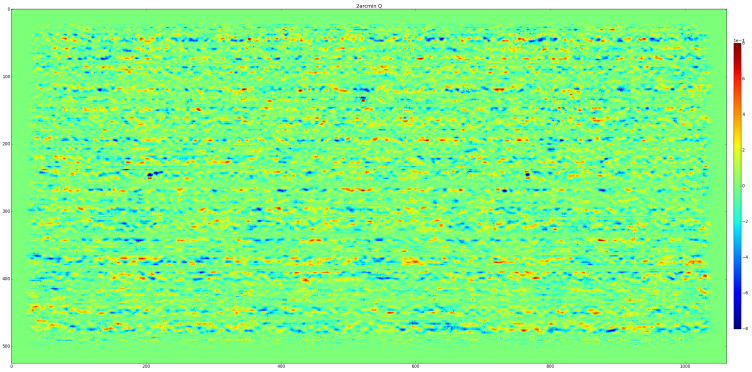


Figure 3.9: This is a map of the Q stokes parameters of the sky. The horizontal striping are because of uneven coverage coupling intensity gradients into polarization. This map was generated from a small subset of observations. The stripes are horizontal because of our constant elevation scan strategy. This plot was generated by Jason Gallicchio.

We can simulate the size of this effect as long as we use interpolated sims. For power spectra this extra leakage shows up as an increased variance in our maps. We can limit the size of this effect by only including data when both pixels in a pair are live. That means for detectors with roughly the same pointing we will be generating the map with a detector at ψ and $\psi + \pi/2$ which can help cancel out this effect.

3.7 Filtering

Our map maker necessitates filtering the data. The largest signal we need to filter out is the detector's response to the atmosphere. Fortunately, on the time scales of our scans, the atmosphere is slowly varying, so we can use a high pass filter the data to remove it. The impact of filtering on analysis is usually understood through Monte Carlo simulations. Because of this, it is crucially important that every filter be linear. Otherwise, Monte Carlo simulations would need a completely accurate noise model to properly replicate the effects of the filtering.

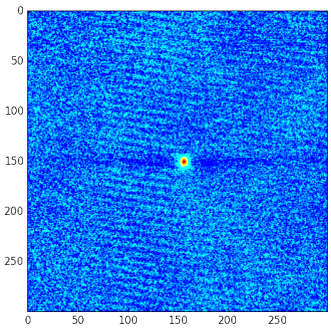
For SPTpol we used two types of filters on the data. Linear least squares based filters and discrete fourier transform based filters. We used DFT based filters to low pass filter the data. This step is to prevent high frequency noise from aliasing into our map. We also used a DFT based filter to remove bright lines in the data. The most prominent line usually corresponds to the frequency the pulse tube cooler is operated at. The linear least squares filters were used as high pass filters. The fit functions consisted of Legendre polynomials, sines and cosines.

High pass filters can create something colloquially called "filtering wings" near point sources. Fitting a slowly varying function to a delta function results in fit being inaccurate near the delta function. In practice, this shows up as dark patches along the scan direction near point sources. See figure 3.10. In order to remove these filtering wings, we exclude the region in the data where a timestream is observing a bright point source when we fit our polynomials and trig functions.

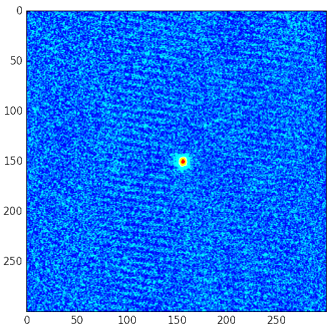
While all of these filters are linear, they are all non-local. Each sample in the filtered timestream is a function of all the samples in the filtered timestream. When filtering we are acting on the detector signal that is a combination of its response to the T, Q and U components of the radiation. This means our filtering can mix T, Q and U together. Studying this effect ends up being rather complicated, so it is usually left to Monte Carlo simulations.

Figure 3.10: Filtering wings.

(a) Polynomial filtered map with no filter masking.



(b) Polynomial filtered map with filter masking.



3.8 Room for Improvement

Our map maker assumes uncorrelated noise. Even after filtering, our noise is correlated. The most direct improvement to the map maker would be to modify it to include these correlations in the map making process. The other avenue of approaching this problem is to develop better filters such that the filtered data is actually uncorrelated.

Beyond that, there are a lot of open questions for SPT-3G.

- How well can we characterize the polarization angle and efficiency of the detectors?
- Does the cross polarization of the detector vary across the beam?
- With our new understanding of polarized gradients, how can we actually study the polarization of compact objects since we know any asymetries in our coverage will create a fake polarization signal?
- Relative to our noise, how large of an effect is our filtering T to P leakage?

Pixel-differenced Map Making

Pixel-differenced map making is something a lot of CMB experiments do that SPT does not. The general idea is that for two time streams for detectors with the same band on the same pixel, t_a and t_b, we replace their timestreams with $t_a + t_b$ and $t_a - t_b$. Almost all of the intensity information lives in $t_a + t_b$, and almost all of the polarization information lives in $t_a - t_b$. The basic idea is that any noise that is correlated in the two detector signals will be suppressed in the subtracted timestreams. The downside to this technique is that it implicitly assumes the location each detector is looking in a pixel is identical.

If by some small miracle, we actually have uncorrelated noise between detectors in a pixel, pixel-differenced map making is equivalent to our original map maker. Let's work through the math. First, let's assume that both detectors in a pixel are looking at the same pixel for every sample and that our samples are sorted such that samples from timestreams on a pixel are always next to each other. We can modify our map making equation by including a pixel difference operator, which we will define in matrix form.

$$\boldsymbol{D} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & -1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 1 & -1 \end{pmatrix} \tag{3.19}$$

This difference operator has a couple nice properties, mainly, $\boldsymbol{D} = \boldsymbol{D}^T$ and $\boldsymbol{D}\boldsymbol{D} = \boldsymbol{1}$. In our map making equation, we need to replace $\boldsymbol{y'} = \boldsymbol{D}\boldsymbol{y}$ and $\boldsymbol{P'} = \boldsymbol{D}\boldsymbol{P}$. If we go into the map maker assuming uncorrelated noise, that is to say C_n is diagonal and the noise is mean 0 we can write $\boldsymbol{C'}_n = \boldsymbol{D}\boldsymbol{C}_n\boldsymbol{D}$. And since $\boldsymbol{D}$ is its own inverse, $\boldsymbol{C'}_n^{-1} = (\boldsymbol{D}\boldsymbol{C}_n\boldsymbol{D})^{-1} = \boldsymbol{D}\boldsymbol{C}_n^{-1}\boldsymbol{D}$.

$$\boldsymbol{m} = (\boldsymbol{P'}^T\boldsymbol{C'}_n^{-1}\boldsymbol{P'})^{-1}\boldsymbol{P'}^T\boldsymbol{C'}_n^{-1}\boldsymbol{y'} = (\boldsymbol{P}^T\boldsymbol{D}^T\boldsymbol{C'}_n^{-1}\boldsymbol{D}\boldsymbol{P})^{-1}\boldsymbol{P}^T\boldsymbol{D}^T\boldsymbol{C'}_n^{-1}\boldsymbol{D}\boldsymbol{y}$$
$$\boldsymbol{m} = (\boldsymbol{P}^T\boldsymbol{C}_n^{-1}\boldsymbol{P})^{-1}\boldsymbol{P}^T\boldsymbol{C}_n^{-1}\boldsymbol{y}$$

Under our assumptions of uncorrelated noise, the map making equation results in the same map using pixel differencing. In the case of correlated noise, however, the reprojected data could result in a noise covariance matrix that is more diagonal. The differential detector pointing information is lost in the pixel differencing as well. This effect will need to be corrected in any maps made with a pixel differenced map. We can reap some of the benefits of the pixel differenced map maker by placing some requirements in our data selection and our inverse noise covariance assignment. If we only use timestreams in map making that also have usable data from their pixel partner and if we make the weights of each timestream in a pixel equal, we can roughly approximate the pixel differenced map maker without losing the differential pointing information. The setting the weight of each timestream in a pixel to the same value is equivalent to the assumption that the noise covariance matrix in in the differenced/summed projection is diagonal.

Point Sources

During most of SPTpol, we did not know about the large effect temperature gradients could have on our produced polarized maps. This has hamstrung our understanding of the polarization of compact objects. Any study of the polarization of point sources will be affected by our map maker's coupling of intensity gradients into polarized signals. Around bright point sources this gradient is large. The gradient is large because for a point source and SPT's compact beams a large change in intensity happens over a very small patch of sky.

For any study of compact objects this coupling will need to either be carefully modelled in our map maker or new methods of understanding the polarization of compact objects will need to be developed. The above discussion of the equivalence of pixel differencing was included to highlight the fact that the coupling of intensity gradients into polarization will be unaffected by a pixel differenced map maker.

3.9 Notes About Scan Strategy

When observing a CMB field we use a constant elevation, constant velocity scan. Given our unique geographic position, this maps to roughly constant declination scans. We are just varying the right ascension of our telescope. For other locations on the Earth, the celestial sphere is rotating relative to the local coordinates.

This creates a few unique problems for SPT. Our time ordered data filtering is non-local. With the lack of sky rotation, each scan at the same elevation maps to to the same patch of sky. Because of this, any filtering T,Q,U leakage adds coherently between scans. This lack of sky rotation also affects the coverage of every map pixel. This enhances our sensitivity to intensity gradients across the pixel.

Our scans are at a fixed angular velocity. For the filtering this allows us to map our time ordered data frequencies to angular scales on the sky. These angular scales are mapped to

the angular scales of the spherical harmonics in terms of l.

For science observations we repeatedly measure the same patch of sky. In order to smooth the telescope's coverage of various parts of the sky we vary the elevations we perform scans at. These elevation variations are colloquially called "dither steps".

Chapter 4

Power Spectra Estimation

4.1 Overview

The cosmic microwave background (CMB) radiation was first measured in 1964[67]. Thirty years later small temperature anisotropies were measured in the CMB. These temperature variations are on the order of 10^{-5} times smaller than the average temperature[84]. The temperature fluctuations of the photons from the CMB provide a snapshot of the universe when these photons by and large stopped interacting with matter. This slice in time forms a spherical shell around us we label the surface of last scattering. The small scale of the temperature fluctuations means that linear perturbation theory can be used to solve the kinetic equations for particles in the early universe. The physics is simple enough that if we are given a starting distribution of fluctuations and a small set of cosmological parameters, we can to high accuracy[81] predict what these fluctuations will look like at the surface of last scattering. While analytic solutions do not exist for the full set of equations, many software packages exist for numerically simulating the early universe[108, 11, 53]. Let's take a moment to appreciate how stunning this fact is. We live in a world where we cannot adequately predict the behavior of a double pendulum moments after releasing it, and yet we are able to model the entire universe for hundreds of thousands of years.

We as experimentalists can also invert the process. Instead of taking cosmological parameters and predicting the properties of the surface of last scattering, we can measure the surface of last scattering and estimate the likelihood of various cosmological parameters. The first step in this process is characterizing the distribution of the temperatures of the photons we measure. Up to experimental errors, the primary anisotropies in the CMB are Gaussian distributed[74]. If we use a spherical harmonic basis to describe the anisotropies of the CMB, this Gaussianity property means that the multipole moments are fully characterized by their power spectrum, $C_{T,l}$. For a band limited temperature field we expand in terms of spherical harmonics:

$$T(\boldsymbol{n}) = \sum_{lm} a_{T,lm} Y_{lm}(\boldsymbol{n}) \tag{4.1}$$

We will be discussing multiple type of a_{lm} values. These can be T, E or B a_{lm} values. When what we are discussing applies to any of these sets we will use the notation $a_{X,lm}$ to denote this where $X \in [T, E, B]$.

The power spectrum $C_{T,l}$ is defined as the the variance of the $a_{T,lm}$ values:

$$\langle a_{T,lm} a_{T,l'm'} \rangle = \delta_{ll'} \delta_{mm'} C_{T,l} \tag{4.2}$$

This power spectrum has an incredibly rich structure [18, 106, 55]. Experimental measurements of it provide constraints on many parameters of cosmological interest[92, 25, 73]. For a good introduction of the structure of the CMB power spectrum please read [33]. The numerical simulations of the early universe primarily focus on predicting this power spectrum.

We, however, will not be focusing on the theoretical predictions for the CMB power spectrum. This chapter will focus entirely on the experimental measurement of the CMB power spectrum. We will be touching on some spherical harmonic math, the flat sky approximation, and then how various analysis choices affect the power spectrum estimation. This is meant to be a primer on how SPT estimates power spectra.

With that said, it is important to understand what the power spectrum feeds into. The power spectrum estimation code generates band powers, band power covariances, and band power window functions. Band powers are $C_{X,l}$ values averaged over an l range. Band power window functions describe the response of our binned measurement to a change in the underlying theory power spectrum:

$$C_{X,b} = \sum_l w_{X,b,l} C_{X,l} \tag{4.3}$$

The simulation software converts a set of cosmological parameters, Θ, into a theory power spectrum. The likelihood of the data is then estimated from this theory spectrum and our measured band powers, $p(C_{X,l}|\Theta)$. Bayes theorem is the used to generate a posterior on the cosmological parameters of interest:

$$p(\Theta|C_{X,l}) = \frac{p(C_{X,l}|\Theta)p(\Theta)}{p(C_{X,l})} \tag{4.4}$$

$p(\Theta)$ is the prior which is set from other measurements of the cosmological parameters. Because no meaningful way to estimate $p(C_{X,l})$ from the data exists, it is treated as an overall normalization constant. Markov chain Monte Carlo (MCMC) methods are used to integrate $p(C_{X,l}|\Theta)p(\Theta)$ to measure the normalization $p(C_{X,l})$. For SPT, the CosmoMC software library is used to do this[54].

While the power spectrum fully characterizes the primordial anisotropies, many mechanisms exist after initial CMB emission that introduce higher order correlations between the $a_{X,lm}$ values. Gravitational lensing, the KSZ and TSZ effect, point sources, and many other physical phenomena make the power spectrum an incomplete characterization of the microwave sky. Many of these phenomena impart a measurable signal in the power spectrum.

Characterizing the power spectrum can also provide some understanding of these secondary anisotropies.

4.2 Power Spectrum Definitions

As mentioned before, we characterize the intensity fluctuations in terms of the variance of the spherical harmonics:

$$\langle a_{T,lm} a_{T,l'm'} \rangle = \delta_{ll'} \delta_{mm'} C_{T,l} \tag{4.5}$$

For a fully measured sky we can invert this relationship:

$$C_{T,l} = \frac{1}{2l+1} \sum_m |a_{T,lm}|^2 \tag{4.6}$$

The factor of $2l+1$ comes from the number of available m values for a given l.

For polarized fluctuations the situation is a bit more complicated. When measuring the polarized sky we can parameterize each point of the sky in terms of the Stokes parameters, I, Q, U and V. Because when we measure the intensity we are actually measuring the temperature of the field, we often use T for I. On the celestial sphere the angle used to define the Q and U values is the direction to the pole of the coordinate system $\hat{\theta}$. Theoretical predictions and experimental measurements suggest circular polarization of the CMB as being consistent with 0, so we will neglect it. So in practice at every point on the sky we have three values characterizing the intensity and polarization, T, Q, and U.

In analogy to the intensity field, we may consider estimating the polarized power spectrum in terms of the spherical harmonic coefficients of the Q and U fields. This has some theoretical issues, however. At each point of the sphere the Q and U values are not uniquely defined. Since Q and U are defined relative to $\hat{\theta}$, if at a point on the sphere we were to rotate our $\hat{\theta}$ and $\hat{\phi}$ vectors by an angle ψ we would need to transform our Q and U values like:

$$\begin{bmatrix} Q' \\ U' \end{bmatrix} = \begin{bmatrix} \cos(2\psi) & \sin(2\psi) \\ -\sin(2\psi) & \cos(2\psi) \end{bmatrix} \begin{bmatrix} Q \\ U \end{bmatrix} \tag{4.7}$$

This is in stark contrast to the intensity signal that is rotationally invariant.

Zaldarriaga and Seljak identified that $Q \pm iU$ behaved like spin ± 2 quantities on the sphere and that spin 2 weighted spherical harmonics should be used to expand these quantities[107].

$$(Q + iU)(\boldsymbol{n}) = \sum_{lm} a_{2,lm} \, {}_2Y_{lm}(\boldsymbol{n}) \tag{4.8}$$

$$(Q - iU)(\boldsymbol{n}) = \sum_{lm} a_{-2,lm} \, {}_{-2}Y_{lm}(\boldsymbol{n}) \tag{4.9}$$

The paper had more insight, but first we should discuss spin weighted spherical harmonics. Spin weighted spherical harmonics are not particularly common in physics and, to be frank, somewhat confusing objects. The clearest derivations of the properties I have found are in [20, 65]. We will be following the work of these two papers.

First, we need a representation of linear polarization on the sphere that encompasses our transformation rules. First we construct an orthonormal basis $(\hat{\boldsymbol{e_1}}, \hat{\boldsymbol{e_2}})$ at each point on the sphere. In practice $(\hat{\boldsymbol{e_1}}, \hat{\boldsymbol{e_2}})$ often corresponds to $(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\phi}})$, but this choice is arbitrary. Much of the following math is to create a way of describing polarized fields on the sphere in a way that is independent of our constructed orthogonal basis. Under a local rotation of the angle ψ these vectors transform as:

$$\hat{\boldsymbol{e_1'}} = \cos(\psi)\hat{\boldsymbol{e_1}} - \sin(\psi)\hat{\boldsymbol{e_2}} \tag{4.10}$$

$$\hat{\boldsymbol{e_2'}} = \sin(\psi)\hat{\boldsymbol{e_1}} + \cos(\psi)\hat{\boldsymbol{e_2}} \tag{4.11}$$

We say a function ${}_sf(\hat{\boldsymbol{n}})$ has spin weight s if under this type of rotation the function transforms as ${}_sf(\hat{\boldsymbol{n}}) \rightarrow e^{-is\psi}\,{}_sf(\hat{\boldsymbol{n}})$. We are following the sign convention of [65]. We define the vectors $\boldsymbol{m}$ and $\overline{\boldsymbol{m}}$ in this basis:

$$\boldsymbol{m} = \frac{1}{\sqrt{2}}\left[\hat{\boldsymbol{e_1}} + i\hat{\boldsymbol{e_2}}\right] = \frac{e^{-i\gamma}}{\sqrt{2}}\left[\hat{\boldsymbol{\theta}} + i\hat{\boldsymbol{\phi}}\right] \tag{4.12}$$

$$\overline{\boldsymbol{m}} = \frac{1}{\sqrt{2}}\left[\hat{\boldsymbol{e_1}} - i\hat{\boldsymbol{e_2}}\right] = \frac{e^{i\gamma}}{\sqrt{2}}\left[\hat{\boldsymbol{\theta}} - i\hat{\boldsymbol{\phi}}\right] \tag{4.13}$$

Where we have represented the arbitrary nature of our basis through the γ variable. The choice of γ we will call the spin gauge. Local rotations of our coordinate basis correspond to a transformation of the form:

$$\boldsymbol{m} \rightarrow e^{-i\Lambda}\boldsymbol{m} \tag{4.14}$$

Given this transformation for some vector v_α the quantities $v_\alpha m^\alpha$ and $v_\alpha \overline{m}^\alpha$ transform as spin 1 and spin -1 quantities. At a point on the sphere this quantity, $v_\alpha m^\alpha$, is defined in term of the basis vectors at that point on the sphere. The local rotations we have been discussing are a rotation of the local coordinate basis. It is just a rotation of the basis vectors at that one point on the sphere; it does not correspond to a global change in variables or a global rotation of the coordinate system. This type of transform corresponds to a spin gauge change from equation 4.13 or the rotation of Q and U in equation 4.7. In the following discussion of spin weighted functions when we say "transforms as" we explicitly mean this type of local basis change.

This extends to a rank-s tensor. For a tensor $T_{i_i \ldots i_s}$ the quantity $T_{i_i \ldots i_s} m^{i_1} \ldots m^{i_s}$ transforms as a spin-s object since every m constributes a factor of $e^{-is\psi}$. We can extend this definition somewhat. A quantity with t contractions with m and u contractions with $\overline{m}$,

$$T_{i_i \ldots i_{t+u}} m^{i_1} \ldots m^{i_t} \overline{m}^{i_1} \ldots \overline{m}^{i_u} \tag{4.15}$$

transforms like a spin $t-u$ quantity.
These vectors have the property:

$$\boldsymbol{m} \cdot \boldsymbol{m} = \overline{\boldsymbol{m}} \cdot \overline{\boldsymbol{m}} = 0 \tag{4.16}$$
$$\boldsymbol{m} \cdot \overline{\boldsymbol{m}} = 1 \tag{4.17}$$

We can represent linear polarization on the sphere as a trace free symmetric tensor and then express it in terms of m and $\overline{m}$:

$$\mathcal{P}_{ij} = \begin{bmatrix} Q & iU \\ iU & -Q \end{bmatrix} = (Q+iU)\overline{m}_i \overline{m}_j + (Q-iU) m_i m_j \tag{4.18}$$

This means that we have a representation of the polarization on the sphere that transforms like we want. This representation generalizes to higher rank tensors. m and $\overline{m}$ can provide a complete basis for the totally symmetric trace-free portion of a rank s tensor.

$$T_{i_1 \ldots i_s} = {}_s f(\hat{\boldsymbol{n}}) \overline{m_{i_s}} \ldots \overline{m_{i_1}} + {}_{-s} f(\hat{\boldsymbol{n}}) m_{i_1} \ldots m_{i_s} \tag{4.19}$$

The chief insight [107] had was that if we have a transform that changes a spin ± 2 quantity into a spin 0 quantity we can make theoretical predictions about that spin 0 quantity without worrying about the spin gauge freedom we have when picking out polarization angles. This motivates us to define a raising and lowering operators, $\eth$ and $\bar{\eth}$, that have the property that under a gauge transform, a spin s function transforms like: $\eth \; {}_s f(\hat{\boldsymbol{n}}) \rightarrow \eth e^{-i(s+1)\Lambda} \; {}_s f(\hat{\boldsymbol{n}})$. That is to say the raising and lowering operators change the spin by plus or minus one. For our spin s function we represent as:

$${}_s f(\hat{\boldsymbol{n}}) = m^{i_1} \ldots m^{i_s} T_{i_1 \ldots i_s} \tag{4.20}$$

We can define a raising and lowering operators as:

$$\eth \; {}_s f(\hat{\boldsymbol{n}}) = -\sqrt{2} m^j m^{i_1} \ldots m^{i_s} \nabla_j T_{i_1 \ldots i_s} \tag{4.21}$$
$$\bar{\eth} \; {}_s f(\hat{\boldsymbol{n}}) = -\sqrt{2} \overline{m}^j m^{i_1} \ldots m^{i_s} \nabla_j T_{i_1 \ldots i_s} \tag{4.22}$$

Where we have used the normalization from [65]. Since $\nabla_j T_{i_1 \ldots i_s}$ is a tensor and we have added the appropriate contraction with m^j or $\overline{m}^j$ this construction has the appropriate transform under a spin gauge change. This construction is a bit awkward, though, since the covariant derivative has already commuted through the m vectors. We would like to convert the operator to something that we apply to the function ${}_s f(\hat{\boldsymbol{n}})$. Defining the vector K_j as $m^j m^i K_j = -m^j \nabla_j m^i$ we have the relationship:

$$m^j m^{i_1} ... m^{i_s} \nabla_j T_{i_1 ... i_s} = m^j (\nabla_j + s K_j) m^{i_1} ... m^{i_s} T_{i_1 ... i_s} \tag{4.23}$$

Where we have summed s copies of K_j since each m^i contributes one copy of K_j when commuting the covariant derivative. This gives:

$$\eth = -\sqrt{2} m^j (\nabla_j + s K_j) \tag{4.24}$$

We can use a similar construction for the lowering operator. Since $\nabla_\theta m_\theta = \nabla_\theta m_\phi = 0$ the quantity $m^j \nabla_j m^i = -\overline{m}^j \nabla_j m^i$ because only the ϕ index contributes and $m^\phi = -\overline{m}^\phi$.

$$\bar{\eth} = -\sqrt{2} \overline{m}^j (\nabla_j - s K_j) \tag{4.25}$$

Evaluating these operators in (θ, ϕ) coordinates gives[20]:

$$\eth \, {}_s f(\theta, \phi) = -e^{-i\gamma} \left[\partial_\theta + \frac{i}{\sin(\theta)} \partial_\phi + s(-\cot(\theta) - i\gamma_{,\theta} + \frac{1}{\sin\theta} \gamma_{,\phi}) \right] \tag{4.26}$$

$$\bar{\eth} \, {}_s f(\theta, \phi) = -e^{i\gamma} \left[\partial_\theta + \frac{i}{\sin(\theta)} \partial_\phi - s(-\cot(\theta) - i\gamma_{,\theta} + \frac{1}{\sin\theta} \gamma_{,\phi}) \right] \tag{4.27}$$

$$\tag{4.28}$$

Fixing the spin gauge at $\gamma = 0$ we can simplify this to:

$$\eth \, {}_s f(\theta, \phi) = -\left[\partial_\theta + \frac{i}{\sin(\theta)} \partial_\phi - s \cot(\theta) \right] \tag{4.29}$$

$$\bar{\eth} \, {}_s f(\theta, \phi) = -\left[\partial_\theta + \frac{i}{\sin(\theta)} \partial_\phi + s \cot(\theta) \right] \tag{4.30}$$

$$\tag{4.31}$$

Since we now have a representation of the raising and lowering operators we can apply these to the spherical harmonics to generate what are called the spin weighted spherical harmonics. In the case of a spin-s weighted spherical harmonic we have the form:

$${}_s Y_{lm} = \left[\frac{(l-s)!}{(l+s)!} \right]^{1/2} \eth^s Y_{lm}, (0 \leq s \leq l) \tag{4.32}$$

$${}_s Y_{lm} = \left[\frac{(l+s)!}{(l-s)!} \right]^{1/2} (-1)^s \bar{\eth}^s Y_{lm}, (-l \leq s \leq 0) \tag{4.33}$$

$${}_s Y_{lm} = 0, (l < |s| \leq l) \tag{4.34}$$

With equations 4.29 and 4.34 the spin weighted spherical harmonics are presented as scalar functions. At least for the author, this lead to some confusion since scalar functions

have no meaningful notion of a change in spin. It is important to remember, however, that the spin weighted spherical harmonics are effectively short hand for applying the operator in equation 4.24 to the spherical harmonics. When we apply a spin gauge transform to the spin weighted spherical harmonics the $\eth$ and $\bar{\eth}$ operators will change, not the Y_{lm} values since that is a scalar field.

The spin weighted spherical harmonics satisfy the same orthogonality and completeness relationships the standard spherical harmonics satisfy[107]. When integrating over a sphere Ω:

$$\int_{\Omega} dn \; {}_sY_{lm} \, \overline{{}_sY_{l'm'}} = \delta_{ll'}\delta_{mm'} \tag{4.35}$$

$$\sum_{lm} \; {}_sY_{lm}(\theta,\phi) \, \overline{{}_sY_{lm}}(\theta',\phi') = \delta(\theta-\theta')\delta(\phi-\phi') \tag{4.36}$$

Which makes them a useful basis for band limited spin-s quantities on the sphere. In our case, the quantities $Q \pm iU$ can be represented as a tensor contraction which we can then describe with spin ± 2 spherical harmonics since these are also tensor contractions with our m vectors.

$$Q + iU = m^i m^j \mathcal{P}_{ij} = \sum_{lm} a_{2,lm} \; {}_2Y_{lm} \propto \sum_{lm} a_{2,lm} \eth^2 Y_{lm} \tag{4.37}$$

$$Q - iU = \overline{m}^i \overline{m}^j \mathcal{P}_{ij} = \sum_{lm} a_{-2,lm} \; {}_{-2}Y_{lm} \propto \sum_{lm} a_{-2,lm} \bar{\eth}^2 Y_{lm} \tag{4.38}$$

Inverting this relationship we can write:

$$a_{\pm 2,lm} = \int dn (Q \pm iU) \, \overline{{}_{\pm 2}Y_{lm}} \tag{4.39}$$

The spin ± 2 basis is useful because our function expanded in terms of spin 2 spherical harmonics can easily be converted to a spin 0 quantity. The quantities $\bar{\eth}^2(Q+iU)$ and $\eth^2(Q-iU)$ are both spin gauge invariant which means any theoretical predictions for the statistical properties of these quantities can ignore the complications brought about from the spin gauge freedom.

Rather than use just the spin 2 basis for spin 2 functions, in CMB physics we use a basis that's a linear combination of the spin ± 2 harmonics. This is called the E/B basis and is defined by:

$$a_{E,lm} = -\left(a_{2,lm} + a_{-2,lm}\right)/2 \tag{4.40}$$

$$a_{B,lm} = -\left(a_{2,lm} - a_{-2,lm}\right)/(2i) \tag{4.41}$$

The modes on the sky composed entirely of E modes are unchanged under a parity transformation. The B modes flip sign. This comes from the fact that:

$$ {}_sY_{lm}(\pi - \theta, \phi + \phi) = (-1)^l \ {}_{-s}Y_{lm} \tag{4.42} $$

This basis is used because scalar perturbations only contribute to the E modes on the sky while tensor modes contribute to both E and B modes[107].

The power spectrum that we are trying to estimate for the polarized sky is defined in a similar way to the intensity power spectrum:

$$ C_{T,l} = \frac{1}{2l+1} \sum_m |a_{T,lm}|^2 \tag{4.43} $$

$$ C_{E,l} = \frac{1}{2l+1} \sum_m |a_{E,lm}|^2 \tag{4.44} $$

$$ C_{B,l} = \frac{1}{2l+1} \sum_m |a_{B,lm}|^2 \tag{4.45} $$

$$ C_{TE,l} = \frac{1}{2l+1} \sum_m < \overline{a_{T,lm}} a_{E,lm} > \tag{4.46} $$

4.3 The Flat Sky Approximation

Our observable power spectrum is defined in terms of spin 0 and spin ± 2 spherical harmonics. This would suggest that when we are estimating the power spectrum we would need to do spin spherical harmonic transforms on the quantities T, $Q + iU$ and $Q - iU$. For full sky CMB surveys this is what is done. The HEALPIX pixelization of the sphere is used [26] and the power spectrum is estimated with spherical harmonics. When analyzing data we are more often limited by our understanding of systematics and how the data processing affects the results. Our ability to understand the processing is inversely proportional to the complexity of the processing. We can make our data processing simpler by using the flat sky approximation. The flat sky approximation is built on the idea that for large enough l values over a small patch of sky, the spherical harmonics look like sine waves on the surface of the sphere. So to estimate the power spectrum of the CMB using a small patch of sky we do not need to use spherical harmonic transforms and can instead use discrete Fourier transforms (DFT).

Terrestrial CMB experiments often create maps of relatively small patches of sky explaining the widespread use of the flat sky approximation. The observation of a small patch of sky is driven by a number of factors. The observing time per sky area increases when only imaging a small patch of sky lowering the noise of the maps made. Optical loading on the bolometers from the atmosphere changes as a function of telescope elevation. Because the optical loading can change the responsivity of the TES bolometers there is a driver for

limiting the changes in elevation when observing a patch of sky. Since they are focused on extra-galactic sources of radiation terrestrial CMB experiments tend to observe patches of sky at high galactic latitude limiting the available observation points for terrestrial experiments, especially those at the South Pole.

Spin-0, Intensity

For the temperature power spectrum our goal is to relate spherical harmonics to plane waves over a small patch of sky. We will be following the derivation in [34].

Starting with a band limited scalar field described on the sphere by:

$$X(\hat{\boldsymbol{n}}) = \sum_{lm} a_{lm} Y_{lm}(\hat{\boldsymbol{n}}) \tag{4.47}$$

We define a quantity $X(\boldsymbol{l})$ as being the plane wave expansion of the a_{lm} coefficients:

$$X(\boldsymbol{l}) = \sqrt{\frac{4\pi}{2l+1}} \sum_{m} i^{-m} a_{lm} e^{im\phi_l} \tag{4.48}$$

This expansion has the inverse relationship:

$$a_{lm} = \sqrt{\frac{2l+1}{4\pi}} i^{m} \int \frac{d\phi_l}{2\pi} e^{-im\phi_l} X(\boldsymbol{l}) \tag{4.49}$$

Our goal is to show that the Fourier transform of our plane wave expansion is a good approximation to the map. To be explicit, we would like to show:

$$X(\hat{\boldsymbol{n}}) \approx \int \frac{d^2 l}{(2\pi)^2} X(\boldsymbol{l}) e^{i\boldsymbol{l}\cdot\hat{\boldsymbol{n}}} \tag{4.50}$$

This is done by first relating the associated Legendre polynomials to Bessel functions in the large l limit. The Bessel functions are then converted into plane waves using the Jacobi-Anger expansion.

To relate the Bessel functions to the associated Legendre polynomials we use a relationship in [2]:

$$\lim_{\nu\to\infty} \left[\nu^{\mu} P_{\nu}^{-\mu} \left(\cos \frac{x}{\nu} \right) \right] = J_{\mu}(x) \tag{4.51}$$

The Ps are associated Legendre polynomials. The J are Bessel functions. With a quick change of variables $\mu \to m$, $x \to l\theta$, and $\nu \to l$:

$$P_l^{-m}(\cos\theta) \approx \frac{1}{l^m} J_m(l\theta) \tag{4.52}$$

We can now plug in our approximation for the associated Legendre polynomials into the definition of the spherical harmonics:

$$Y_l^m(\theta,\phi) = \sqrt{\frac{(2l+1)}{4\pi}\frac{(l-m)!}{(l+m)!}}P_l^m(\cos\theta)e^{im\phi} \approx \sqrt{\frac{(l)}{2\pi}\frac{(l-m)!}{(l+m)!}}\frac{1}{l^{-m}}J_m(l\theta)e^{im\phi} \tag{4.53}$$

Noting that:

$$\lim_{l\to\infty}\sqrt{\frac{(l-m)!}{(l+m)!}l^{2m}} = 1 \tag{4.54}$$

We can now write the spherical harmonics at large l in terms of Bessel functions:

$$Y_l^m \approx J_m(l\theta)\sqrt{\frac{l}{2\pi}}e^{im\phi} \tag{4.55}$$

We use Jacobi-Anger expansion and the above relationship we can write:

$$e^{i\boldsymbol{l}\cdot\hat{\boldsymbol{n}}} = \sum_m^\infty i^m J_m(l\theta)e^{im(\phi-\phi_l)} \tag{4.56}$$

$$\approx \sqrt{\frac{2\pi}{l}}\sum_m i^m Y_l^m(\hat{\boldsymbol{n}})e^{im\phi_l} \tag{4.57}$$

Now with all those identities out of the way, we can simply plug in our definition of a_{lm} and use our approximations for Y_{lm} to show the desired relationship:

$$X(\hat{\boldsymbol{n}}) = \sum_{l,m} a_{lm}Y_{lm}(\hat{\boldsymbol{n}}) \tag{4.58}$$

$$\approx \sum_l \frac{l}{2\pi}\int\frac{d\phi_l}{\pi}X(\boldsymbol{l})\sum_m J_m(l\theta)i^m e^{im(\phi-\phi_l)} \tag{4.59}$$

$$\approx \int\frac{d^2l}{(2\pi)^2}X(\boldsymbol{l})e^{i\boldsymbol{l}\cdot\hat{\boldsymbol{n}}} \tag{4.60}$$

Using the definition of $X(\boldsymbol{l})$ if we integrate that around a loop we find that:

$$\int_0^{2\pi} d\phi X^*(\boldsymbol{l})X(\boldsymbol{l}) = 8\pi^2\frac{1}{2l+1}\sum_m |a_{lm}|^2 \tag{4.61}$$

So up to a constant the integral of $X(\boldsymbol{l})$ around a loop of constant l should be roughly equal to the angular averaged power at that l value.

All this math means that it is acceptable for us to treat a small patch of sky when dealing with large l values as a flat surface and use the Fourier transform of the map in order to estimate the power spectrum. Since the data is quantized, we use the DFT to do this transform.

Spin-2 Polarization

For the $\gamma = 0$ spin gauge in the flat sky limit, $\sin(\theta) \to 1$, we can write the spin raising and lowering operators in a very simple form:

$$\eth \approx -(\partial_x + i\partial_y) \tag{4.62}$$
$$\bar{\eth} \approx -(\partial_x - i\partial_y) \tag{4.63}$$

For the small scale, flat sky limit we apply the raising and lowering operators to our flat sky derived Y_{lm}[106].

$${}_2Y_{lm} = \left[\frac{(l-2)!}{(l+2!)}\right]^{1/2} \eth^2 Y_{lm} \to (2\pi)^{-2}\frac{1}{l^2}\eth^2 e^{i\boldsymbol{l}\cdot\hat{\boldsymbol{n}}} \tag{4.64}$$
$${}_{-2}Y_{lm} = \left[\frac{(l-2)!}{(l+2!)}\right]^{1/2} \bar{\eth}^2 Y_{lm} \to (2\pi)^{-2}\frac{1}{l^2}\bar{\eth}^2 e^{i\boldsymbol{l}\cdot\hat{\boldsymbol{n}}} \tag{4.65}$$

$$\frac{1}{l^2}\eth e^{i\boldsymbol{l}\cdot\hat{\boldsymbol{n}}} = -e^{2i\phi_l} e^{i\boldsymbol{l}\cdot\hat{\boldsymbol{n}}} \tag{4.66}$$
$$\frac{1}{l^2}\bar{\eth} e^{i\boldsymbol{l}\cdot\hat{\boldsymbol{n}}} = -e^{-2i\phi_l} e^{i\boldsymbol{l}\cdot\hat{\boldsymbol{n}}} \tag{4.67}$$

This form of the equation differs from the literature[106, 34] by a factor of $e^{\pm 2i\phi}$. This term rotates the curved sky Q/U angle, defined relative to the direction of the pole, to the flat sky angle, defined relative to the y-axis.

Combining these approximations with equation 4.8, we find:

$$Q = \frac{1}{(2\pi)^2}\int d^2\boldsymbol{l}\,[E(\boldsymbol{l})\cos(2\phi_l) - B(\boldsymbol{l})\sin(2\phi_l)]\,e^{i\boldsymbol{l}\cdot\hat{\boldsymbol{n}}} \tag{4.68}$$
$$U = \frac{1}{(2\pi)^2}\int d^2\boldsymbol{l}\,[E(\boldsymbol{l})\sin(2\phi_l) + B(\boldsymbol{l})\cos(2\phi_l)]\,e^{i\boldsymbol{l}\cdot\hat{\boldsymbol{n}}} \tag{4.69}$$

4.4 Power Spectrum Estimators

Our goal is to take a map of the sky and convert it into an estimate of the power spectrum of the full sky. To be explicit we would like to convert a data vector of polarized pixels $\boldsymbol{x}$ into an estimate of $C_{X,l}$ values that we parameterize as a set of values q_i. While many power spectrum estimators exist, three classes of estimators are the most common. These estimators are the maximum likelihood, quadratic minimum variance and pseudo C_l. When discussing the estimators we will use $\boldsymbol{C}$ for the pixel covariance and assume that the data is mean zero, $\langle x_i \rangle = 0$. The number of elements in $\boldsymbol{x}$ we will label N.

Maximum likelihood estimators rely on finding the power spectrum that maximizes the likelihood of the data. For normally distributed noise we can write this as[101]:

$$p(\boldsymbol{x}|\boldsymbol{q}) \propto \exp(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{C}^{-1}\boldsymbol{x}) \tag{4.70}$$

Solving the maximization problem requires $O(N^3)$ operations. SPTpol maps are on the order of $10^6 - 10^8$ pixels making this sort of maximum likelihood calculation computationally infeasible for full resolution data. For a map with a greatly increased pixel size (reducing the total number of pixels) with a known pixel covariance matrix we could feasibly perform this operation. These operations are nonlinear which can make estimating the statistical properties of the inferred power spectrum difficult.

Another class of estimator is the quadratic minimum variance estimator discussed in [96, 97]. The general idea behind these estimators is that we assume our estimate of the power spectrum has a quadratic form:

$$q_i = x^T\boldsymbol{Q_i}x - b_i \tag{4.71}$$

And then we solve for the $\boldsymbol{Q_i}$ that minimizes the variance of our estimate of the $C_{X,l}$ parameters. While this quadratic ansatz is better justified in the original paper, let's discuss why it has this form. The power spectrum of some map of the sky is a quadratic function of the sky data. We can see that this is the case since the $C_l \propto |a_{lm}|^2$ and the a_{lm} values are just a linear transform of the map. It's not unreasonable to assume that since the thing we are measuring is proportional to the square of the data that our best estimate could also be constructed from the square of the data.

The solution that minimizes the variance of the estimate has the form:

$$\boldsymbol{Q_i} = \frac{1}{2}\sum_j (\boldsymbol{B})_{ij}\boldsymbol{C}^{-1}\boldsymbol{P_j}\boldsymbol{C}^{-1} \tag{4.72}$$

Where $\boldsymbol{B}$ is an arbitrary invertible matrix, $\mathcal{N}_i$ are a normalization constants and $\boldsymbol{P_j}$ is defined by:

$$\boldsymbol{P_i} = \frac{\partial \boldsymbol{C}}{\partial p_i} \tag{4.73}$$

Where p_i represents one power spectrum l mode for one type of spectrum. When measuring l_{max} modes and the 7 spectra (the signal T, E, B, TE, TB, EB modes and the noise) there are $7l_{max}$ components. Tegmark argues that this approach has more easily understood statistical properties since the estimate is quadratic in the data and that the computational complexity scales as N^2[97]. For a reasonable estimate of the pixel covariance, this method of estimating the power spectrum does not destroy any cosmological power spectrum information.

The third class of estimators are the pseudo-C_l estimators. The general idea behind these estimators is that we estimate a set of pseudo $\widetilde{a_{X,lm}}$s from our noisy, biased data by just

applying the appropriate spherical harmonic transforms. We then relate these pseudo $\widetilde{a_{X,lm}}$ values or the summed pseudo $\widetilde{C_{X,l}}$, to an estimate of the full sky $C_{X,l}$ values via analytic expressions or Monte Carlo simultion[32]. The pseudo-C_l estimators tend to under perform the quadratic minimum variance estimators at low l values[22]. They do have the advantage that they are computationally cheap, can handle non-Gaussian noise, and can easily work with biased map makers.

In practice, SPT has used pseudo-C_l techniques for all of our power spectrum estimation. We have generally focused on large enough l values that the benefit of the quadratic estimator is outweighed by the extra computation and person time it would cost to implement a quadratic estimator. An exploration of the quadratic estimator could potentially reduce the variance of our low l power spectrum. Because it is what we have used, we will be discussing the issues with Pseudo-C_l estimation of the power spectrum.

Intensity Pseudo C_ls

Survey limits and foregrounds restrict any map constructed of the CMB to only have partial coverage of the sky. The direct calculation of C_l values via a spherical harmonic transform is complicated by this incomplete sky coverage. Rather than estimating the actual a_{lm} values, we estimate the windowed, pseudo $\widetilde{a_{lm}}$ values:

$$\widetilde{a_{lm}} = \int d\boldsymbol{n} T'(\boldsymbol{n}) W(\boldsymbol{n}) \overline{Y_{lm}} \tag{4.74}$$

Where $W(\boldsymbol{n})$ is our smoothed coverage map of the sky and $T'(\boldsymbol{n})$ is our constructed map of the sky. We say constructed because our map maker produces a biased map of the sky. With the pseudo-C_l estimators the coverage map is apodized on the edges. We label this the apodization mask.

Throughout our discussion of these pseudo $\widetilde{a_{lm}}$ values we will be using the curved sky notation. All of these derivations apply equally well in the flat sky limit.

During the map making procedure, we filter the data with the intent of making the detector time ordered data have a diagonal covariance matrix. Because of this linear filtering, the $T'(\boldsymbol{n})$ is a biased map of the sky. With the restriction we have placed that the filters used are linear we can write down a mapping of the true a_{lm} values to our pseudo $\widetilde{a_{lm}}$ in terms of a simple mapping which we arbitrarly label as $J_{lml'm'}$:

$$\widetilde{a_{lm}} = \sum_{l'm'} a_{lm} J_{lml'm'} \tag{4.75}$$

While analytic forms for this mapping exist for an unbiased map of the sky, no such luxury exists with the biased map maker. Given the complicated nature of the interactions of the filtering with the pointing of the telescope we must rely on Monte Carlo simulations to construct this mapping. Unfortunately, the number of unique elements in this construction

scales like l^4 making its construction computationally infeasible for an experiment like SPT that makes meaningful measurements to angular scales of $l > 10^4$.

We rely on the ansatz constructed in [32], typically called the MASTER method, to reduce the computational requirements of this estimation. Rather than rely on the construction for individual a_{lm} values we construct this mapping for the ensemble average. With the inclusion of a map noise term $\left\langle \widetilde{N_l} \right\rangle$, we can write our average pseudo $\left\langle \widetilde{N_l} \right\rangle$ with the coupling term $K_{ll'}$.

$$\left\langle \widetilde{C_l} \right\rangle = \sum_{l}^{'} K_{ll'} \left\langle C_l \right\rangle + \left\langle \widetilde{N_l} \right\rangle \tag{4.76}$$

We have included the noise because every real map of the sky is noisy. We further divide $K_{ll'}$ into several terms that contribute to it:

$$K_{ll'} = M_{ll'} F_{l'} B_{l'}^2 \tag{4.77}$$

$M_{ll'}$ is the mode coupling kernel resulting from the incomplete sky coverage of the map. The term mode mixing is often used for the mode coupling kernel. $F_{l'}$ is the transfer function that models the effects of filtering the data when producing the map. $B_{l'}^2$ is a smoothing function that combines the pixel window function and detector's response to a point source. $B_{l'}$ is estimated in terms of its effect on a_{lm} values which is why it is squared in this equation. This simplification of the coupling neglects the fact that the filtering transfer function is not truly diagonal in l space. For some time domain filters, like a polynomial filter, various l modes can be mixed together. For these constructions we have assumed that the filters applied to the data are linear. Nonlinear filtering causes the mapping of a_{lm} to $\widetilde{a_{lm}}$ to be sensitive to the realization of the noise in the time ordered data.

The mode coupling kernel described the l mode mixing from performing spherical harmonic transforms, or DFTs in the flat sky limit, on a finite patch of sky. To first order it is the 2D equivalent of an DFT window function. The mode coupling kernel can be analytically calculated or estimated with Monte Carlo simulations. The beam function is composed of our detectors response to a point source and the pixel window function. The detectors response to a point source is estimated from observations of celestial sources. The pixel window function, which encodes the fact that a pixel is not sample of an infinitesimal point on the sky but an average of the sky over the size of pixel, is either estimated using Monte Carlo simulations or an approximate ansatz.

The estimation of the transfer function, $F_{l'}$, is done with Monte Carlo simulations. A set of fake microwave skies with a power spectrum matching some fiducial cosmology is generated. We then mock observe these skies in software to generate detector time ordered data. The map making and power spectrum estimation code is then repeated on this simulated data. Because we know the input power spectrum we can easily compute the ratio of the estimated power spectrum to the input power spectrum. This Monte Carlo simulation frame-

work also provides a way for understanding how various systematics affect our measurement of the power spectrum.

When estimating the transfer function we remove the effects of mode leakage and the beam so for some estimated pseudo spectra, $\widetilde{C_{l'}}$ with an input spectra $B_l^2 C_l$:

$$F_l = \frac{\sum_l' M_{ll'}^{-1} \widetilde{C_{l'}}}{B_l^2 C_l} \tag{4.78}$$

Because the patch of sky is relatively small, neighboring l values of the C_l are indistinguishable. To get around this when estimating the power spectra, we combine multiple C_l measurements to create our bin averaged power spectrum, $\hat{C}_b$, by averaging over an l range. We describe the binning with the operation P_{bl} and the reciprocal interpolation operation with Q_{lb}. Using these we can write a binned version of $K_{ll'}$:

$$K_{bb'} = P_{bl} M_{ll'} F_{l'} B_{l'}^2 Q_{l'b'} \tag{4.79}$$

With this binned version we are now able to write our band power window function defined in equation 4.3:

$$w_{bl} = K_{bb'}^{-1} P_{b'l'} M_{l'l} (F_l B_l^2) \tag{4.80}$$

Bundles

SPTpol made on the order of 10^4 observations of the same patch of sky. These 10^4 observations are grouped together and added to form 10^2 bundles. When estimating the power spectrum of the sky we estimate the cross spectrum between bundles in order to remove the noise bias from our estimate.

To see why this works, let's assume the bundles have the same noise. We can write our full map as a sum of all of the bundles:

$$M = \frac{1}{N_{bundles}} \sum_{i=1}^{N_{bundles}} S + N_i = \frac{1}{N_{bundles}} \sum_i (a_{lm}^S + a_{lm}^{N_i}) Y_{lm} \tag{4.81}$$

Where S is the sky signal and N_i is the noise in that bundle. If we were to sum all the bundles together and estimate the pseudo spectra we would have:

$$\frac{1}{N_{bundles}^2} \int \sum_m (\sum_i (a_{lm}^S + a_{lm}^{N_i}) Y_{lm}) * \overline{(\sum_j (a_{lm}^S + a_{lm}^{N_j}) Y_{lm})} = < |a_{lm}^s|^2 > + \sum_i < |a_{lm}^{N_i}|^2 > \tag{4.82}$$

With this form the estimated variance of the a_{lm}s has a noise bias. We can remove this, however. We expect the noise between different bundles to be uncorrelated so when estimating the pseudo spectra we never estimate the spectra of a bundle with itself. That is to say we only estimate the spectra using separate bundles, we can remove this noise bias.

$$\frac{1}{N_{bundles**2}} \int \sum_m (\sum_i (a^S_{lm} + a^{N_i}_{lm}) Y_{lm}) * \overline{(\sum_{j \neq i} (a^S_{lm} + a^{N_j}_{lm}) Y_{lm})} =< |a^s_{lm}|^2 > \tag{4.83}$$

We have also assumed that the noise is uncorrelated with the signal.

The bundled maps are also useful for jackknife systematics checks. Jackknives are tests performed with the data that check for various systematics. They are done by dividing the data into halves that would be sensitive to systematics and then comparing the power spectrum estimated form the two groups. As an example, we could compare the estimated power spectrum from data taken when the sun is above the horizon and when the sun is below the horizon. If these two spectra are not statistically consistent then there is evidence for sun contamination.

Bundles are also useful for estimating the band power covariances. We can divide the full covariance into the sum signal and noise covariance, $\boldsymbol{C} = \boldsymbol{S} + \boldsymbol{N}$. We can use simulations to estimate the signal covariance. To do this we generate many fake realizations of the microwave sky, mock observe them, and then estimate the band powers. The covariance can be directly estimated from these measurements. Estimating the noise covariance requires multiple realizations of the noise. To generate these we randomly divide the maps into two halves. One half is multiplied by -1 and these are added. This map should have the signal subtracted leaving only the noise with roughly the statistical properties of the full map noise. We can repeat this process with other random split halves to generate more noise realizations.

Polarization Pseudo C_ls

With the inclusion of polarized data we now have three sets of $a_{X,lm}$ values to keep track of, where $X \in (T, E, B)$. Our projection operator now has the form:

$$\widetilde{a_{X,lm}} = \sum_{l'm'X'} a_{X,lm} J_{lml'm'XX'} \tag{4.84}$$

When mapped to the ansatz of [32] we need to promote the mode coupling kernel and transfer function to include the mixing from T, E and B. $M_{ll'} \rightarrow M_{ll'XX'}$. $F_{l'} \rightarrow F_{l'X'}$. At a given l', $F_{l'X'}$ is a dense 3x3 matrix. For some fixed l and l', $M_{ll'XX'}$ has the form:

$$M_{ll'XX'} = \begin{bmatrix} M_{ll'TT'} & 0 & 0 \\ 0 & M_{ll'EE'} & M_{ll'EB'} \\ 0 & M_{ll'BE'} & M_{ll'BB'} \end{bmatrix} \tag{4.85}$$

While filtering can potentially mix temperature and polarization signals together, the effect of a finite patch of sky only mixes E and B together. It can not leak intensity signals into polarized signals. Because the T and E modes of the sky are so much larger than the B modes of the sky we will need to worry about the leakage from T and E into B. If these

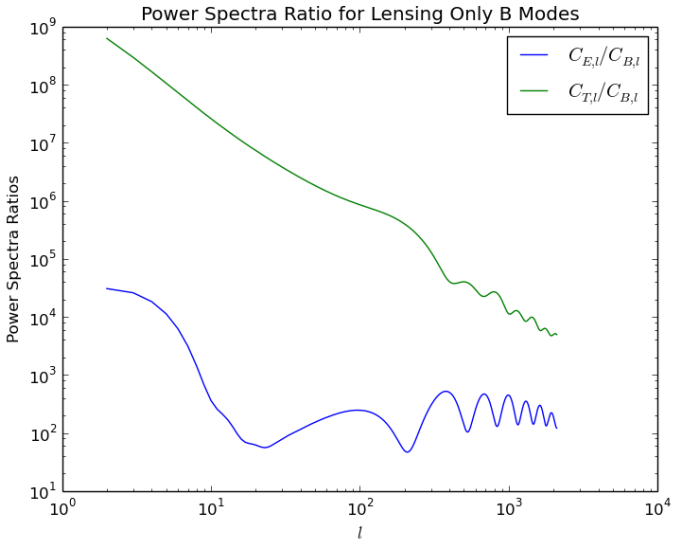


Figure 4.1: Much of the discussion of BB power spectrum estimation focuses on limitting the leakage terms. This is primarily because the power in the BB spectra is so much smaller than the other terms. Acceptable levels of E to B leakage are those below 10^{-3}. Acceptable levels of T to B leakage are much more scale dependent but on the order of $10^{-7} - 10^{-6}$.

leakage terms are subtracted off in $C_{X,l}$ space, $C_{X,l}^{True} = C_{X,l}^{Est} - C_{X,l}^{Leak}$, they contribute to the variance since the variance of two normally distributed variables subtracted is the sum of their respective variances.

Estimating this leakage can be easily done via Monte Carlo simulations. We can mock observe fake skies with only T and E content and directly measure the B modes our power spectrum estimator produces. If the variance of the leakage terms is large relative to our estimated spectrum, however, this leakage will degrade our measurement. This has motivated some work on constructing pseudo-C_l BB estimators that limit this leakage.

4.5 E/B Mixing from a Finite Patch of Sky

The orthogonality relationship of the spin-s spherical harmonics, $\int_\Omega {}_sY_{lm}\ \overline{{}_sY_{l'm'}} dn = \delta_{ll'}\delta_{mm'}$, is only valid when integrated over the entire sphere. This lack of orthogonality when observing a finite patch of sky leads to a mixture of E modes and B modes with the naive pseudo-C_l estimator.

From equations 4.40 and 4.8, we can write down Q and U values for a sky consisting entirely of E modes. For our $B = 0$ sky, we define $b_{lm} = {}_2a_{lm} = {}_{-2}a_{lm}$:

$$Q = \sum_{lm} b_{lm} \frac{{}_2Y_{lm} + {}_{-2}Y_{lm}}{-2} \tag{4.86}$$

$$U = \sum_{lm} b_{lm} \frac{{}_2Y_{lm} - {}_{-2}Y_{lm}}{-2i} \tag{4.87}$$

We then solve for the $a_{B,lm}$ coefficients in the presence of some window of the sky $W(s)$.

$$a_{B,lm} = \frac{-1}{2i} \int_\Omega dn W(s) \left[(Q + iU) \; {}_2Y_{lm} - (Q - iU) \; {}_{-2}Y_{lm} \right] \tag{4.88}$$

$$= \frac{-1}{2i} \int_\Omega ds \sum_{l'm'} b_{l'm'} \left[W(s) \; {}_2Y_{l'}^{m'} \overline{{}_2Y_{lm}} - W(s) \; {}_{-2}Y_{l'}^{m'} \overline{{}_{-2}Y_{lm}} \right] \tag{4.89}$$

For a window $W(s) = 1$ over the sphere we can use our orthogonality relationship, $\int_\Omega dn \; {}_sY_{lm} \; \overline{{}_sY_{l'm'}} = \delta_{ll'}\delta_{mm'}$ to find that $a_{B,lm} = 0 \forall l, m$. If we set $W(s) \neq 1$ for some portions of the sky this orthogonality relationship can no longer be used and we find that $a_{B,lm} \neq 0$. A sky composed entirely of E modes when only measured over a finite portion appears to have B modes. B modes will also leak into E modes from measuring a finite patch of sky, but that term is usually negligible because the magnitude of E modes is significantly larger than the B modes[14, 47]. For a measured E mode spectrum we can estimate the leaked B mode contribution to $C_{l,BB}$ and then subtract it off. This subtraction contributes extra variance to our estimate of $C_{l,BB}$, however. Because of this, various pseudo-C_l estimators for the BB spectrum that do not leak E to B have been developed. We will touch on a couple of them.

χ_B Estimator

The χ_B relies on converting the Q and U maps into maps that are proportional to E and B and then estimating the power spectrum from those maps. The constructed χ_B is only a function of the B modes of the sky. The χ_E is only a function of the E modes of the sky. Because the windowing happens in this χ_B data space, it does not mix E into B modes. This method was developed in [83].

The general idea is that with these raising and lowering identities:

$$\eth \; {}_sY_{lm} = \left[(l - s)(l + s + 1) \right]^{1/2} \; {}_{s+1}Y_{lm} \tag{4.90}$$

$$\bar{\eth} \; {}_sY_{lm} = -\left[(l + s)(l - s + 1) \right]^{1/2} \; {}_{s-1}Y_{lm} \tag{4.91}$$

$$\tag{4.92}$$

We can apply the raising and lowering operators to Q and U maps to generate a map entirely composed of E or B modes:

$$\chi_E = \frac{1}{2}\left[\bar{\eth}\bar{\eth}(Q+iU) + \eth\eth(Q-iU)\right] \tag{4.93}$$

$$= -\sum_{lm}\left[(l+2)!/(l-2)!\right]^{1/2} a_{E,lm}Y_{lm} \tag{4.94}$$

$$\chi_B = \frac{i}{2}\left[\eth\eth(Q+iU) - \bar{\eth}\bar{\eth}(Q-iU)\right] \tag{4.95}$$

$$= -\sum_{lm}\left[(l+2)!/(l-2)!\right]^{1/2} a_{B,lm}Y_{lm} \tag{4.96}$$

Once we have these χ maps we can then use our intensity power spectrum estimation code to get the E and B modes on the sky.

The produced maps cleverly sidestep the mixing issue because the windowing done in the χ_X map space. The windowing only mixes $B \to B$ and $E \to E$. This method does come with its share of drawbacks, however. The χ_x construction relies on taking numerical derivatives of a pixelized map which can make the operation imperfect. Also, the power spectrum of the χ_X maps is proportional to l^4 which is a very blue spectrum. This blue spectrum complicates the construction of the apodization mask. See figure 4.3. Any mode mixing from the apodization mask will be multiplied by an l^{-4} factor creating a very red estimated spectrum.

Pure Estimators

Instead of thinking of the our window operation as something scaling our Q and U maps, one can think of it as a modification of our E and B conversion operators. Instead of integrating our Q and U maps with respect to ${}_sY_{lm}$ we integrate with respect to $W(s)\ {}_sY_{lm}$. The pure estimator modifies this operation so that the produced pseudo $\widetilde{a^{pure}_{X,lm}}s$ are only proportional to the full sky $a_{X,lm}$ values and does not mix the values. This method was developed in [82].

The pure estimator takes our naive pseudo-C_l estimator, which we write as:

$$\widetilde{a_{B,lm}} = \frac{-1}{2i}\int dn(Q+iU)W(s)N_{lm}\overline{\left[\eth\eth Y_{lm}\right]} - (Q-iU)W(s)N_{lm}\overline{\left[\bar{\eth}\bar{\eth}Y_{lm}\right]} \tag{4.97}$$

And moves the window, $W(s)$, inside the raising and lowering operators.

$$\widetilde{a^{pure}_{B,lm}} = \frac{-1}{2i}\int dn(Q+iU)N_{lm}\overline{\left[\eth\eth(W(s)Y_{lm})\right]} - (Q-iU)N_{lm}\overline{\left[\bar{\eth}\bar{\eth}(W(s)Y_{lm})\right]} \tag{4.98}$$

To see why this has the desired effect, first note that since the spherical harmonics form a complete basis we can represent $W(s)Y_{lm} = \sum_{l'm'}\alpha_{l'm'}Y_{l'm'}$. If we plug that into our pure equations we find:

$$\widetilde{a^{pure}_{B,lm}} = \frac{-1}{2i}\int dn(Q+iU)N_{lm}\overline{\left[\eth\eth(\sum_{l'm'}\alpha_{l'm'}Y_{l'm'})\right]} - (Q-iU)N_{lm}\overline{\left[\bar{\eth}\bar{\eth}(\sum_{l'm'}\alpha_{l'm'}Y_{l'm'})\right]} \tag{4.99}$$

$$= \frac{-1}{2i}\sum_{l'm'}\frac{N_{lm}}{N_{l'm'}}\alpha_{l'm'}\int dn(Q+iU)\overline{{}_2Y_{l'm'}} - (Q-iU)\overline{{}_{-2}Y_{l'm'}} \tag{4.100}$$

$$= \sum_{l'm'}\frac{N_{lm}}{N_{l'm'}}\alpha_{l'm'} \times a_{B,lm} \tag{4.101}$$

The estimated $\widetilde{a^{pure}_{B,lm}}$ are only a function of the full sky $a_{B,lm}$ coefficients. For the above equations, $N_{lm} = \left[\frac{(l-2)!}{(l+2)!}\right]^{1/2}$. We have assumed that the raising and lowering operators are well behaved when applied to $W(s)Y_{lm}$. This requires that $W(s)Y_{lm}$ is continuous and has a continuous first and second derivative. This method relies on taking numerical derivatives of the window function $W(s)$ applied to the sky. Using integration by parts, it can be shown that the χ_B estimator and pure estimator are mathematically equivalent[83].

4.6 Apodization Mask Construction and Inpainting

When imaging the sky, SPT also measures various compact objects. These foregrounds effectively act as extra sources of noise in our constructed power spectra. For the sources bright enough to be individually resolved, we can limit their contribution to power spectra by excluding them from the data when estimating the power spectrum.

This can be done by modifying our apodization mask to exclude those regions or by inpainting. Inpainting involves replacing the portions of the map corrupted by a bright point source with what is effectively a guess of the data in that region. The problem of inpainting maps to solving a set of under-determined equations. This can be done by setting some form of regularization condition to make the problem tractable[90] or assuming the solution follows some simple form[29]. For our power spectrum estimation we require that the inpainting operation be linear. Much to my chagrin, the inpainting operation used in [47] was a nonlinear operation. For sources with unpolarized fluxes in excess of 50mJy at 150GHz, we replaced all the pixels within 6' of the source with the median value of the pixels in the annulus of distance r from the source where $6' < r < 10'$. This operation is non-linear because the median of a set is a non-linear operation on the set. The inpainting operation for compact foregrounds used in [72] is, however, linear. We will focus on that method.

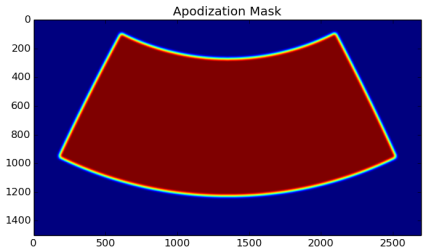
For the pixels in the region we wish to inpaint we first set the pixels to 0. We then replace each pixel with the average of its four neighboring pixels. This process is repeated until the values converge. This operation is equivalent to solving Laplace's equation in the inpainted region with Dirichlet boundary conditions.

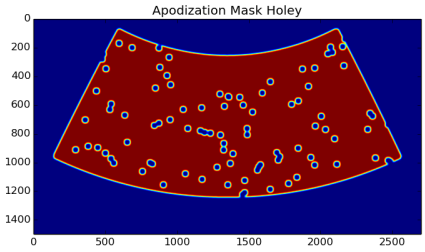
Inpainting can provide some benefits over masking the sources, namely it reduces the off diagonal elements of the mode coupling kernel. In the flat sky limit the apodization mask is just a 2 dimensional DFT window function. The mode coupling kernel is the angular average of the DFT of the apodization mask. Adding narrow holes in map space results in the spectrum spreading in frequency space. Inpainting allows us to avoid this issue somewhat. See figure 4.2.

Figure 4.2: An input sky consisting entirely of E modes around $l = 200$ is generated. This is then either windowed with a holey apodization mask or windowed with the smooth apodization mask and inpainted over the point sources. The BB and EE spectra is estimated with the naive pseudo C_l estimator, the χ estimator, and the pure estimator. A comparison is made between the estimated power spectra for the various estimators with inpainting or apodizing point sources. The inpainting estimators perform the best at l values close to 200, particularly when inpainting the χ_B estimate.

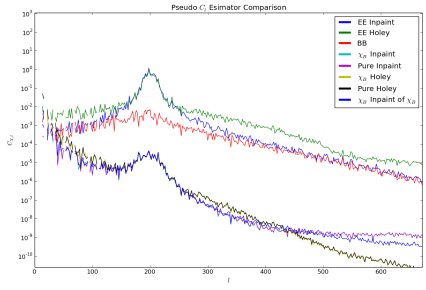
(a) Apodization mask used with the inpainted BB power spectrum estimate.

(b) Apodization mask used when not inpainting.





(c)



(d)

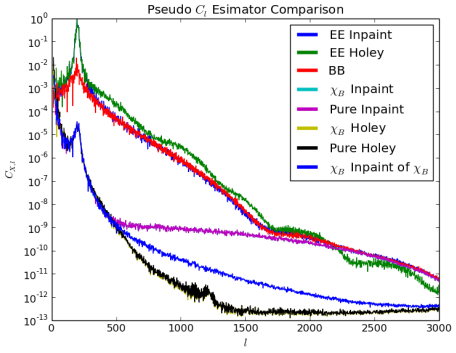
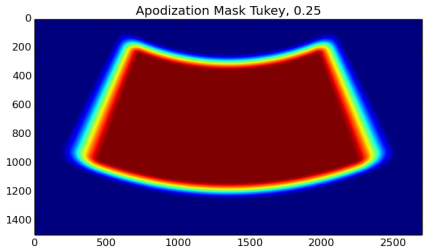
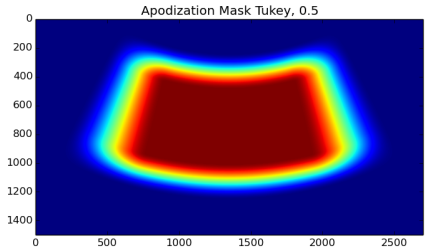


Figure 4.3: A fake microwave sky is realized with the best fit Planck cosmological parameters. The BB power spectrum is then estimated with various apodization masks using the χ_B estimator. The l^4 of the χ_b estimator creates the sharp upturn at low-l values since that l^4 dependence needs to be divided out. Any spreading of the BB spectra to low l or residual EE leakage manifests a low frequency upturn in our estimated power spectrum because it is multiplied by this l^{-4} factor. Using a smoother apodization mask reduces these two leakage terms and pushes this upturn lower in l.

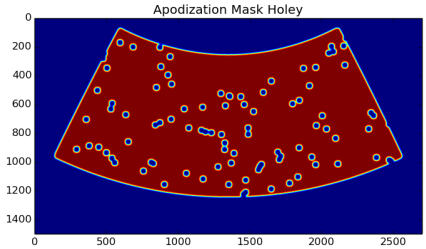
(a)



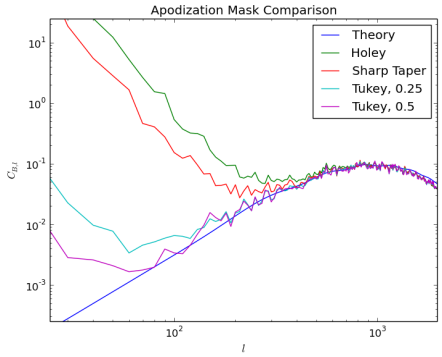
(b)



(c)



(d)



4.7 FRANK BB

Overview

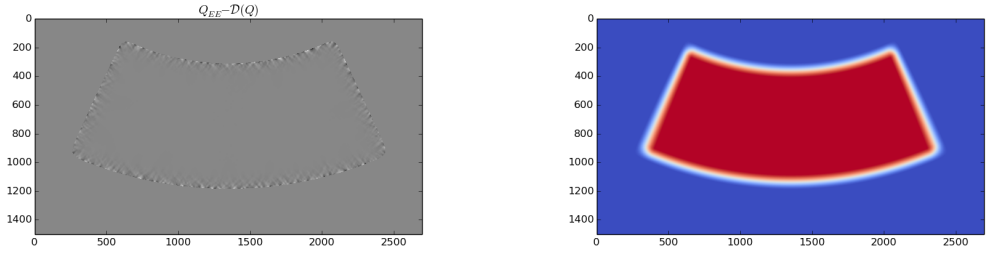
Our biased map maker filters the time ordered data before binning it to make a map. This filtering operation can leak E modes into B modes. The FRANK (First Remove Any Nearly Known) BB estimator is designed to remove this filtering leakage. The general idea is that we first make a map that is a best estimate of the T and E modes on the sky with no B mode content. We then mock observe that TE sky and repeat the map making operation. Any B mode content in that second map is an artifact from filtering and power spectrum

estimation.

While this method seems fairly circular, to get a better estimate of the B modes we remove the B modes and then subtract that map, there are a couple properties of the fields we can exploit. The B modes on the sky are much smaller than the E modes on the sky. This means that the fractional error in our estimate of the E mode map from filtering leakage will be much smaller than the error from filtering leakage with the B mode map. The other property we can use is the semi-locality of the E-B basis. As seen in equation 4.93 the E and B maps can be found in terms of second derivatives of the Q and U maps. Knowing Q and U at a single point on the sky cannot meaningfully provide information about the E/B content. Knowing the values at a point and at a locus of points around it can, however, provide information about the E/B content of that patch. Given that we observe a patch of sky this means that our estimate of the E mode only polarization map will have a small error near the center of our patch of sky and a larger error near the edge. For a picture of this see figure 4.4. Because we know there is a region of sky where our estimate of the E mode map is corrupt we can just exclude this outside ring of data when estimating the power spectrum.

Figure 4.4: B Mode Removal Error

(a) (b)



The map making, mock observation, and removal of B modes are all linear operations that can be written as matrices. We write our map maker as $\mathcal{M}$. The filtering operator as $\mathcal{F}_\alpha$. The mock observation operator is $\mathcal{O}$. The mock observation is just the pointing matrix from the map making equation $\boldsymbol{P}$. The B mode removal operator we will write as $\mathcal{D}$. This means we can replace our map making operation with our leakage removed map making operation:

$$\boldsymbol{m} = \mathcal{M}\mathcal{F}_\alpha \boldsymbol{x} \rightarrow \boldsymbol{m} = \mathcal{M}\mathcal{F}_\alpha \boldsymbol{x} - \mathcal{M}\mathcal{F}_\alpha \mathcal{O}\mathcal{D}\mathcal{M}\mathcal{F}_\beta \boldsymbol{x} \tag{4.102}$$

$$\boldsymbol{m} = \mathcal{M}\mathcal{F}_\alpha (\boldsymbol{I} - \mathcal{O}\mathcal{D}\mathcal{M}\mathcal{F}_\beta)\boldsymbol{x} \tag{4.103}$$

The map produced has the E mode signal removed and the B modes created by the filtering of the E mode data removed. It consists of the uncorrupted B modes. The two nuances of this equation are how the B mode removal operator, $\mathcal{D}$, is constructed and how the filtering options, $\mathcal{F}_\alpha$ and $\mathcal{F}_\beta$, interact.

Filtering Concerns

Two primary types of time domain filters are used by SPT. A discrete fourier transform based filter is used to low pass filter the data. This filter removes high frequency noise that can appear as excess map noise. Atmospheric noise is removed using high pass filters. These high pass filters use linear least squares fitting to fit a small set of basis functions to the time ordered data. The most commonly used set of these functions in SPT's analysis are the first few Legendre polynomials. These remove large scale atmospheric changes and any slow thermal drifts in the detector focal plane temperature. These filters operate on an individual constant elevation scan.

When constructing our estimate of the B mode free sky we can neglect the low pass filter, but we still need to high pass filter the data in some way. The time drift in focal plane temperature would dramatically increase the noise in the map if this filtering were not done. This filtering biases our constructed B mode free map because we lose information on the map fourier modes that we filter. This means with FRANK we will not estimate the leakage from the E modes we filter when generating our best estimate of the E mode only sky. This results in our final map having some modes with an unknown E to B leakage contribution. If we design the second filtering step, $\mathcal{F}_\alpha$, to remove the modes with unknown E to B leakage, then we can guarantee that the output map has properly removed the filtering leakage.

To be a bit more explicit, first note that the linear least squares filtering operators act as projection matrices, $\mathcal{F}_\beta\mathcal{F}_\beta\boldsymbol{x} = \mathcal{F}_\beta\boldsymbol{x}$. We place the requirement that the final filtering, $\mathcal{F}_\alpha$, has the form:

$$\mathcal{F}_\alpha = \mathcal{F}_\phi\mathcal{F}_\beta \tag{4.104}$$

Where $\mathcal{F}_\beta$ is the filtering we use to estimate our B mode free map and $\mathcal{F}_\phi$ is the extra filtering we do on top of that when producing our final map.

We then can rewrite the map making equation as:

$$\begin{aligned}
\boldsymbol{m} &= \mathcal{M}\mathcal{F}_\alpha(\boldsymbol{I} - \mathcal{O}\mathcal{D}\mathcal{M}\mathcal{F}_\beta)\boldsymbol{x} = \mathcal{M}\mathcal{F}_\phi\mathcal{F}_\beta(\boldsymbol{I} - \mathcal{O}\mathcal{D}\mathcal{M} - [\mathcal{O}\mathcal{D}\mathcal{M}, \mathcal{F}_\beta])\boldsymbol{x} && (4.105)\\
&= \mathcal{M}\mathcal{F}_\phi\mathcal{F}_\beta(\boldsymbol{I} - \mathcal{O}\mathcal{D}\mathcal{M})\boldsymbol{x} - \mathcal{M}\mathcal{F}_\phi\mathcal{F}_\beta[\mathcal{O}\mathcal{D}\mathcal{M}, \mathcal{F}_\beta])\boldsymbol{x} && (4.106)\\
&\approx \mathcal{M}\mathcal{F}_\alpha(\boldsymbol{I} - \mathcal{O}\mathcal{D}\mathcal{M})\boldsymbol{x} && (4.107)
\end{aligned}$$

The last line in the equation is the ideal form of this type of filtering. We have a perfectly constructed B mode free map in the form $\mathcal{O}\mathcal{D}\mathcal{M}\boldsymbol{x}$ that we apply our filtering to to estimate the E to B leakage. This assumes that we can safely ignore $\mathcal{M}\mathcal{F}_\alpha[\mathcal{O}\mathcal{D}\mathcal{M}, \mathcal{F}_\beta]\boldsymbol{x}$.

The commutator describes our error in the construction of the B mode free timestreams from polynomial filtering before constructing the data. Fortunately we can construct our filter $\mathcal{F}_\alpha$ to remove this error in estimate. For data filtered with a low order polynomial we expect this error signal to be approximately described by a low order polynomial. By applying a higher order polynomial filter as our second pass filter the corrupted modes will be removed from our estimated B free sky during the second map making process. To summarize, it doesn't matter if we do a bad job estimating part of the B free sky if we filter that part away.

One last thing worth mentioning about filtering is purely a bookkeeping concern. When estimating our B free map of the sky, part of our estimate is corrupted near the edge of the map. When constructing our linear least squares filter, $\mathcal{F}_\alpha$, it must only operate on the data that we have a reasonable B free estimate of. When polynomial filtering the data, the data on the edge of the map that we do not have a B free estimate for will need to be excluded. This is to ensure that the filtering operation is identical between the data constructed from the B free map and the actual data is identical.

$\mathcal{D}$, De-B Operator Construction

Our method for B mode removal is fairly naive. We multiply our Q and U maps by some apodization mask that smooths the transition to zero. We use then convert our apodized Q/U maps into E/B in Fourier space using equation 4.68. We zero all of the B modes below some l cut-off and apply the inverse transform back to Q and U space. We then divide out the apodization mask. This does a fairly reasonable job of estimating the B mode free sky except around the edges of the map. See figure 4.4. Some other options, like using a pure estimator for the E map have been explored, but they tend to perform worse than this estimate.

The main complication with the construction of the B mode free map comes from our desire to not have a noise bias in our estimate of the BB spectrum. As mentioned in section 4.4 we use bundle cross spectra to remove this noise bias. If we were to only produce one estimate of the B free sky we would be biased by the noise. This bias would come from B mode leakage of the noise signal of the E mode map.

Performance

This method of map making has the downside that some data on the edge of the map gets thrown away. In practice this ends up being $\approx 10\%$. Due to uneven coverage of the map this data is often the noisiest so the loss in sensitivity is less than expected from geometry. It provides a factor of 10 reduction in filtering leakage at low l. See figure 4.5. This technique also acts to suppress E to B leakage from the apodization mask. The efficacy of FRANK depends on if the variance in the estimate of the BB power is improved by reducing filtering leakage by sacrificing some map area.

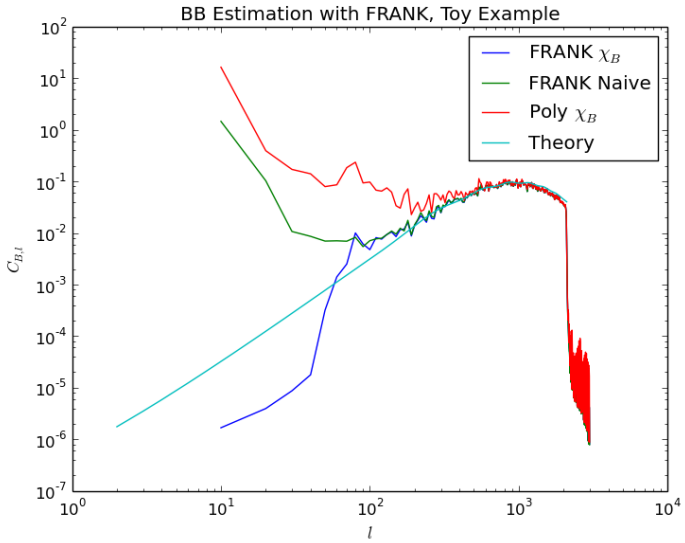
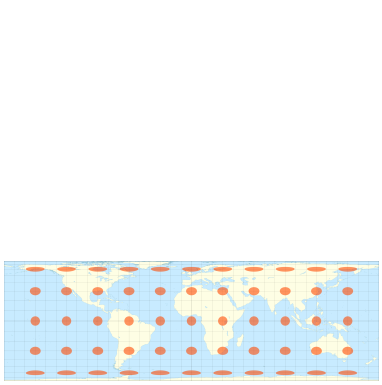


Figure 4.5: A microwave fake sky is generated with best fit Planck cosmological parameters. We then compare the BB power spectrum estimate with a fourth order polynomial filter applied to the data with the FRANK estimated power spectrum where $\mathcal{F}_\alpha$ is a fourth order polynomial filter and $\mathcal{F}_\beta$ is a first order polynomial filter. When estimating the FRANK power spectrum, because the mode mixing has been removed, one can use a leakage free Pseudo-C_l estimator or just the naive Pseudo-C_l estimator.

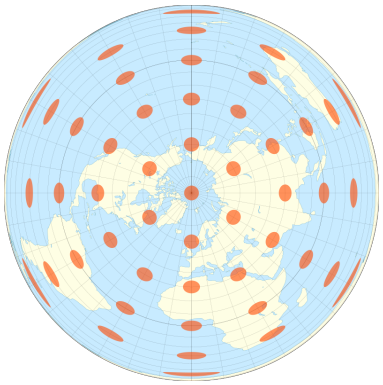
4.8 Flat Sky Map Projections

The flat sky approximation requires mapping points on the sky to a Euclidean, two dimensional coordinate system. The projection of the sphere onto a 2D plane necessarily distorts the map. Many subclasses of map projections that preserve various qualities of the underlying structure such as area, shape, or distance exist[86]. For the purposes of power spectrum estimation, we restrict the set of projections used to be equal area projections in order to minimize the effects of the inhomogeneous pixel size creating an uneven weighting over the map. These equal area map projections, unfortunately, distort the shape of the underlying structures. With our power spectrum estimation we correct for the effects of the detector beam in l space. The l space correction assumes that the produced map is the true map of the sky convolved with the detector beam. This assumption breaks down if the shape of the beam changes over the map. For an illustration of these distortions see figure 4.6.

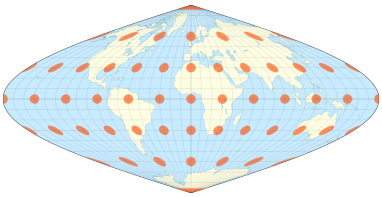
In order limit the distortions in our map it is beneficial to use an oblique projection. Oblique projections redefine the equatorial axis or the center point of a projection to limit the map distortions near a point on the celestial sphere. This rotation of the equatorial axis complicates the mapping of the x and y directions in the projection to θ and ϕ values of the celestial sphere. While a necessary evil for limiting distortion of our beam it can complicate our understanding of various non-idealities in the system.



(a) Lambert cylindrical equal-area projection



(b) Lambert azimuthal equal-area projection



(c) Sinusoidal projections

Figure 4.6: The Tissot's indicatrices for three equal area map projections constructed by [49]. Each ellipse shows the local distortion of the map projection. The projections tend to limit distortions around some local point in the map.

Table 4.1: DFT Library Normalizations

Lib	Forward	Inverse
IDL	$\frac{1}{N}\sum_{x=0}^{N-1} f(x)\exp[-i2\pi xk/N]$	$\sum_{k=0}^{N-1} F(k)\exp[i2\pi xk/N]$
Numpy	$\sum_{x=0}^{N-1} f(x)\exp[-i2\pi xk/N]$	$\frac{1}{N}\sum_{k=0}^{N-1} F(k)\exp[i2\pi xk/N]$
Scipy	$\sum_{x=0}^{N-1} f(x)\exp[-i2\pi xk/N]$	$\frac{1}{N}\sum_{k=0}^{N-1} F(k)\exp[i2\pi xk/N]$
GSL	$\sum_{x=0}^{N-1} f(x)\exp[-i2\pi xk/N]$	$\frac{1}{N}\sum_{k=0}^{N-1} F(k)\exp[i2\pi xk/N]$
FFTW	$\sum_{x=0}^{N-1} f(x)\exp[-i2\pi xk/N]$	$\sum_{k=0}^{N-1} F(k)\exp[i2\pi xk/N]$

Our scan strategy consists of constant declination scans with a varying right ascension, see figure 4.7. Time varying sources of noise will couple along that axis. Any effects of filtering on the map will also be aligned with our scan axis. With a map projection that maps the x and y axis simply to the θ/ϕ directions these distortions are aligned with the axes of the map. The scan aligned distortions will then be localized in the 2D DFT of the map allowing for the construction of a simple filter. An oblique map projection does not have this property, unfortunately.

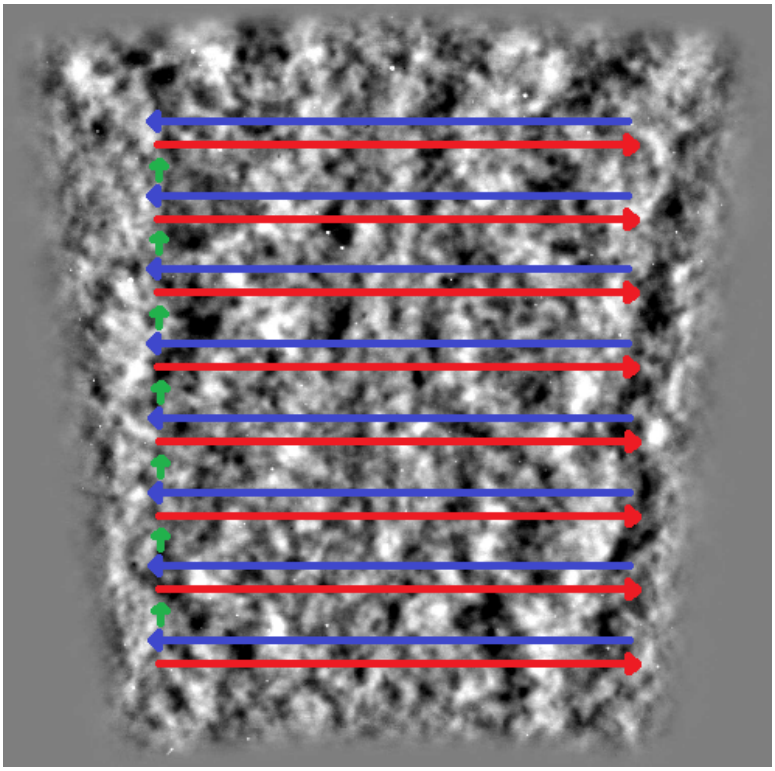


Figure 4.7: SPTpol data was taken with constant elevation scans. Only the azimuth was varied. For our unique geographic location these azimuthal scans are (almost) equivalent to constant declination scans in the equatorial coordinate system.

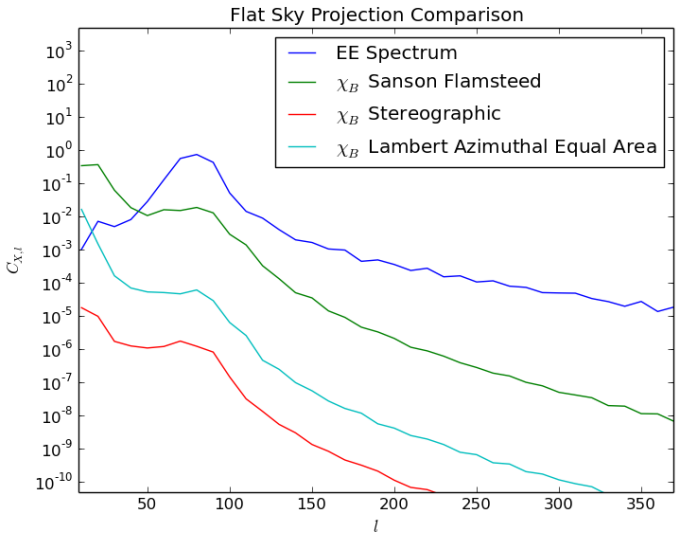


Figure 4.8: A full sky consisting entirely of E modes in the range $70 < l < 90$ is generated with the healpix pixelization scheme. This full sky is projected into various flat sky map projections. The BB power spectra is then estimated with the flat sky χ_B estimator. Using the flat sky power spectrum estimators mixes power between different l bins and between E and B. Because $E >> B$ we are most sensitive to that form of the leakage. For the stereographic and oblique Lambert azimuthal equal area projection the leakage of E to B is small enough to not meaningfully contribute to the variance of our estimate of the BB spectra for all $l > 50$.

Chapter 5

Fast Transient Search

5.1 Introduction

Fast radio bursts (FRBs) are millisecond-long bright ($\approx$ 2 Jy ms) bursts of radio emission. [58, 44, 98]. FRBs are characterized by a dispersion measure larger than can be reasonably sourced in our galaxy, $\approx 300 - 1800 cm^{-3} pc$[70, 41].

They were initially reported in 2007 after having been discovered in legacy pulsar survey data at the Parkes telescope. [58] Further bursts were found in the Parkes High Time Resolution Universe survey [98]. Shortly after, a fast radio burst was found in Arecibo data[89] helping to alleviate worries that the Parkes FRBs were being sourced by terrestrial interference. They have been detected at 800 MHz[61, 9] 1.4 GHz [98, 89], 3 GHz[13], 4 GHz, and 8 GHz[63].

FRB 121102 was originally detected in Arecibo data. Follow up observations of FRB 121102 found additional bursts at the same patch of sky. [88, 80] The repeating nature of the burst allowed for localization of the FRB to its host galaxy [13] and to a persistent radio source[60]. The repeating source has been found to be linearly polarized with an extremely large rotation measure[60], $10^5\ rad\ m^{-2}$. The repeating pulses of FRB 121102 at low frequencies appear to be over a longer time scale (2-9ms) than the pulses measured by Parkes[80, 12]. The repeating pulses measured at 4 and 8 GHz have much shorter time scales, $< 1ms$[63]. This difference in pulse width is roughly consistent with multi-path broadening from the propagation through plasma.[52]

FRB 121102 has a highly variable spectral structure over the measured frequency range of 1.4-8 GHz[88, 80, 51]. This spectra is poorly matched by a power law and can best be described as an envelope with a width of roughly 500 MHz[51]. This envelope starts at zero, peaks, and returns to zero flux. The spectral structure of other FRBs has been complicated by the frequency dependence of diffraction limited beams leading to a large variation in constraints placed on the spectral index[70].

Higher frequency (8 GHz) measurements at Arecibo and the Green Bank Telescope have provided rotation measure measurements of $10^5\ rad\ m^{-2}$[63]. These bursts were found to

be highly linearly polarized, with polarization fractions statistically consistent with 100%. Previous measurements that found smaller [61] or no measurable linear polarization [68] polarization could have been caused by the bandwidth used to measure the signal if other FRBs have as large of rotation measure as the repeater[63].

Follow up observations of other FRBs have yielded no measurements of repeated bursts [69, 61]. It is possible that the non-detections are a result of the lowered fluence of the repeated bursts or that the FRB progenitor was not active during follow-up observations[88]. It is also possible that there are multiple progenitors of fast radio bursts. Many cataclysmic models involving magnetic field shedding predict a relatively flat spectrum of fast radio bursts, though this is complicated by the theoretical difficulty of modelling cataclysmic events[23]. Previous to this paper, no search has been done to see if the phenomenology extends out to frequencies of hundreds of GHz.

Many source models for fast radio bursts have been proposed. Prior to the detection of the repeated fast radio burst many models involving cataclysmic events such as core collapses or compact object mergers were proposed[57, 7, 104, 64, 40, 99, 23]. If FRB 121102 is characteristic of the population, these models are no longer adequate to explain the phenomenology of FRBs. Non-cataclysmic models centered around young neutron stars, [43, 41, 102, 43, 59] black holes[42] and active galactic nuclei[100] are consistent with observations so far. Many more exotic models for FRBs have also been proposed[38, 79][56].

5.2 Relevant SPTpol Information

SPTpol observed the sky in two frequency bands centered at 90 GHz and 150 GHz using transition edge sensor (TES) bolometers. The 150 GHz antennas consist of a feed horn coupled orthomode transducer. The two orthogonal linear polarization modes are separated in an orthomode transducer and coupled to the individual TES bolometers via low-loss striplines. The 150 GHz frequency band is set by the waveguide cut off of the feed horn and a metal mesh filter[3]. The bandwidth of the system is 45 GHz centered at 150 GHz. See figure 5.1. The beams for the two polarizations overlap on the sky. We label this group a pixel. The full width at half maximum of the 150 GHz beams is 1.3 arcminutes. Neighboring pixels have a pointing separation of 3.4 arcminutes.

The TES bolometers are operated with a frequency multiplexed readout system[17] in which 16 detectors are connected in parallel. Each TES is connected in series with a LC filter that defines a resonant frequency. A sum of sine waves at the resonant frequencies is used to supply a constant amplitude voltage biases to each of the 16 detectors in the comb. The total power on the TES is roughly fixed, so changes in optical power result is the opposite change in the electrical power applied to the TES. Increasing optical power decreases the current through the TES. The summed current through the comb is amplified with a DC SQUID and op amp amplifier chain. The amplified signal is digitized and sent to an FPGA. The FPGA demodulates the sine waves and then decimates the demodulated signal. The output time ordered data (TOD) has a sample rate of 190.7 Hz.

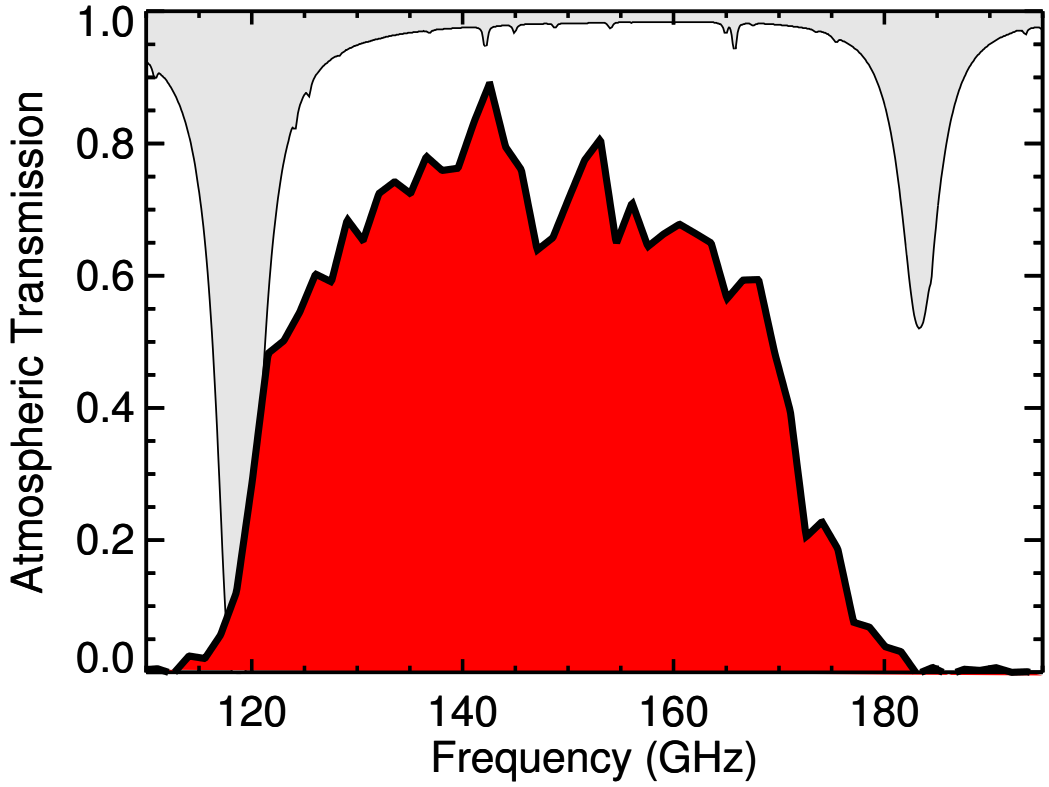


Figure 5.1: SPTpol band response measured with a Fourier-transform spectrometer.

We analysed a year of data from the 2015 and 2016 austral winter. During that time, SPTpol observed a patch of sky centered at right ascension 0 and declination $-57.5°$ with an extent of 4 hours in RA and 15° in declination. While observing this patch the telescope scanned across the sky with an angular speed of $2°s^{-1}$.

5.3 Search Overview

This work presents a search for fast transient events (<10 ms) using the 150 GHz detectors. For most cosmic microwave background analyses, transients with this phenomenology are expected from cosmic rays interacting with the detectors and are actively removed to avoid contaminating the data[50, 47, 21, 76, 71, 8].

The search strategy consists of looking for impulses of flux, filtering known sources of contamination, estimating the background rate, and using simulations to map our residual count rates to a constraints on the rate, flux, and spectra of sources. When estimating the background rate, we rely on the fact that the 150 GHz detectors were monolithically

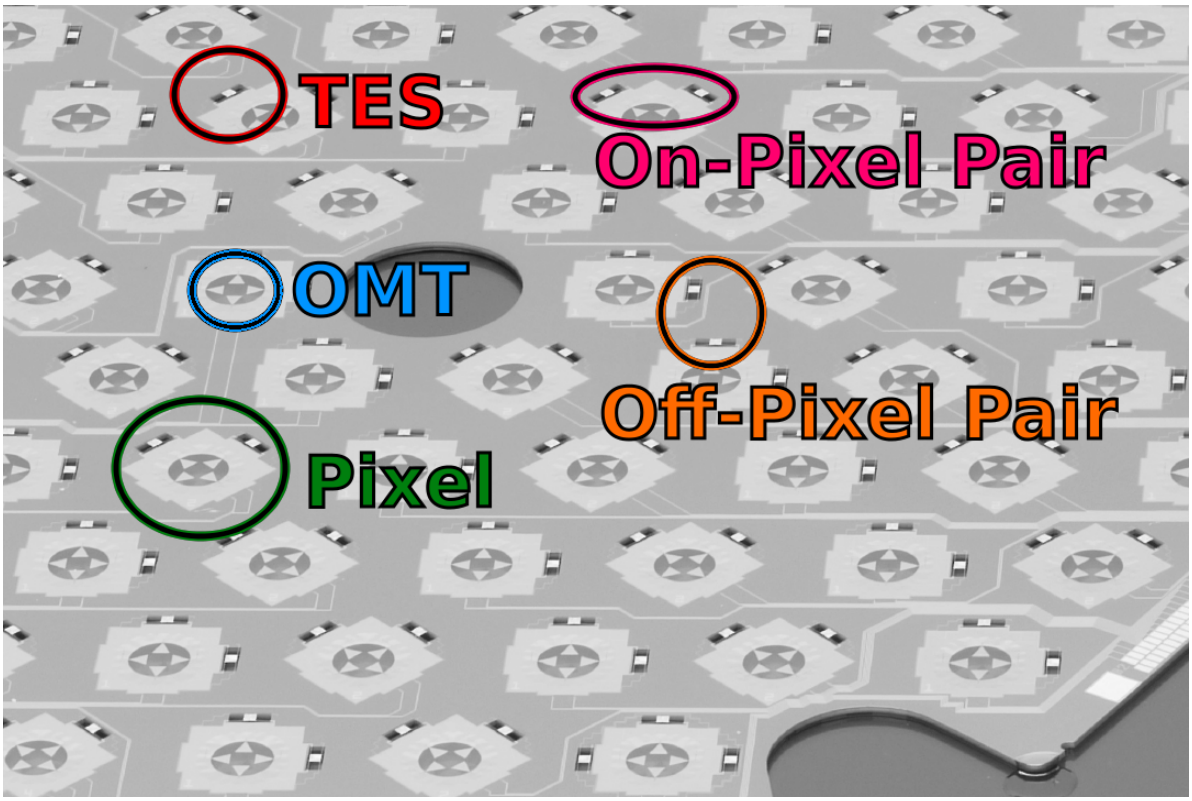


Figure 5.2: An image of SPTpol's 150 GHz wafers. A pixel is comprised of two TESs that measure the optical power from two orthogonal polarization modes split by and OMT. When doing the background rate estimation, we use the rate of off-pixel pair events to estimate the rate of on-pixel pair events. Because the FWHM of the beam is much smaller than the pointing separation of different pixels, TESs that do not share a pixel have negligible overlapping optical response.

fabricated on wafers allowing us to use other TESs on the wafer to estimate the background rate. We use pairs of neighboring detectors that do not share optical response to the same area of sky to estimate the background rate for detectors with overlapping optical response. The 90 GHz detectors were individually manufactured making this method of background rate estimation impossible for them. Each detector beam is measured independently by two different TES bolometers which are measuring the incident power in the two orthogonal polarization modes. In order to reduce our background rate to a manageable level, we require that any impulse be coincidentally detected in both detectors in a pixel.

With the 1.3 arcminute FWHM of the beams and the 3.4 arcminute pointing separation between neighboring pixels, we expect any short transient to only be measurable by the detectors in one pixel. For our 5 ms sample time, and $2^{\circ}s^{-1}$ scan speed, we scan across 0.6 arcminutes per sample. With our beam size, for sources longer than 5 ms, this has an appreciable affect on the power deposited on the antenna that is taken account in the

simulations. For sources shorter than that, we are effectively just sampling the detector beam at a point.

5.4 Details on the Expected Event Structure

Power on the Transition Edge Sensor From A Celestial Source

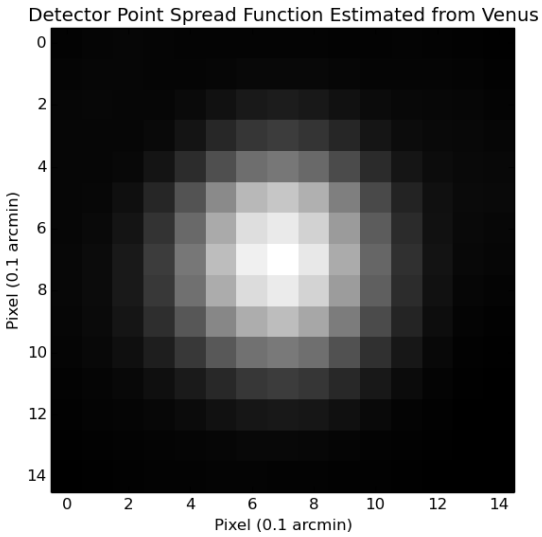


Figure 5.3: SPTpol Beam/Point Spread Function for 150GHz Detector

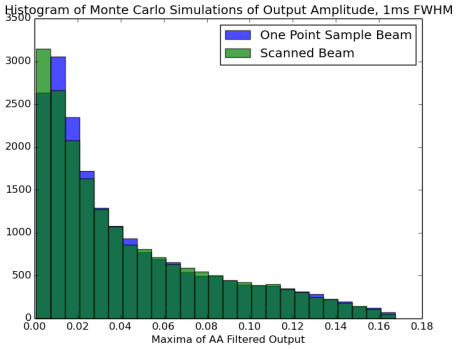
For any fast event, we only observing it in some random section of our detector's beam. This means that for any population of celestial transient signals we will be randomly sampling our beam. Our beam estimated from Venus is seen in figure 5.3.

Typical event durations of FRBs are $\approx$ 4 ms. The mean duration of known FRBs is 6.1 ms, median duration is 4.3 ms[70]. This means we sweep $\approx$ 0.6 arcminutes of our beam across a transient source in 5 ms. This angular scale is smaller than the pointing separation of different pixels. Having multiple pixels respond to an event would require a very large flux or a very long duration event. In the event that there is a sizeable number of these multi-pixel detections we would also see a much larger number of individual pixel events. Because of this we can focus solely on events where two detectors see a noticeable impulse.

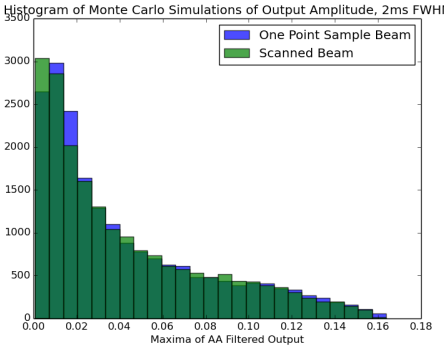
The 2 degree/second scan speed of the telescope means that even for a brief event we are scanning across a fairly large portion of the beam. Monte Carlo simulations were used to understand how different the output time ordered data is when one includes the effects of scanning across the beam as compared to just sampling the beam. See figure 5.4.

Figure 5.4: Input Gaussian signals with random time offsets were fed through the decimation filter. The maximum of the output was histogrammed. For the blue points, the input data stream was scaled by a constant multiplicative factor that is the value of the beam at one point. This is the equivalent to the telescope not moving or scanning very slowly. For the green data points, the effects of the telescope scanning over the source are included. After low-pass filtering the data, the effect of scanning is to push some of the extremes of the response to be closer to the median response. On the low end of the response, we are more likely to have a portion of the signal be when the telescope is looking at the source. For the high end of the response, we will never just be looking at the maximum of the telescope response.

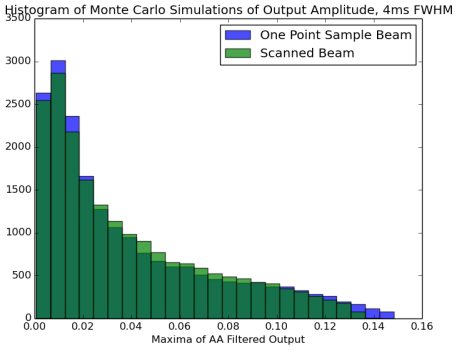
(a)



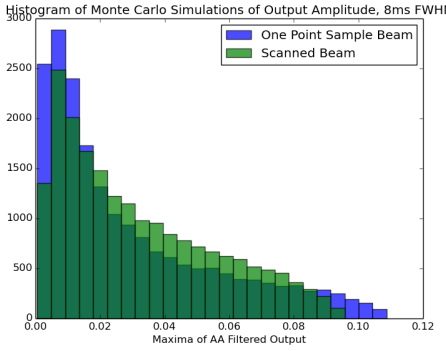
(b)



(c)



(d)



TOD Unit Calibration

Our detector timestream data is normalized to have units of K_{cmb}. For simulations of on sky events we'd like a map of each detector's response to a 1 Jy source in units of K/Sr. We first scale the amplitude of the beam map such that the integral is 1 Jy. For a pixel size s and pixel in Jy/sr units p_i:

$$\sum_{i \in pixels} s \times p_i = 1Jy \tag{5.1}$$

We then apply a conversion factor of Jansky/steradian to Kelvin, J, as estimated from our bands measured with the Fourier transform spectrometer. The number used is $J = 1/(396.3 \times 10^6)K/(Jr/sr)$. This conversion assumes a flat source spectrum. All of the science results are presented in terms of Jy ms for this flat source spectrum.

$$Jp_i = k_i \tag{5.2}$$

Now we have a map of a 1 Jy source in units of Kelvin for the 150 GHz detectors with our beam. When simulating the events, we imitate the scanning of the telescope over this 1 Jy source and then scale our model of the input events.

Filter Response to a Transient

For simulation and constructing the event searching code we need to understand the impulse response of the bolometer and readout system. The frequency response of the TES and the digital anti-aliasing filter could potentially both contribute.

Detector Response

To first order, the transition edge sensor current response to power (s_i) looks like a one pole filter[37]:

$$s_i \propto \frac{1}{1 + j\omega\tau} \tag{5.3}$$

The detector response has typical time constants on the millisecond to sub millisecond scale, $\tau < 1ms$ 5.5. For a 1 ms time constant the response frequency knee is at $\approx$ 160 Hz.

Digital Readout Filters

After the digitization of the detector response signal, we demodulate the data, low-pass filter the data, and finally decimate the data. This digital low-pass filter primarily acts as an anti-aliasing filter. For our output sample rate of 190.7 Hz, this filter's cut off frequency is by construction below the Nyquist frequency of 95 Hz. Because the digital anti-aliasing filter is a multi-pole filter at a frequency roughly an octave below any other filter, this filters response will dominate how power on the TES will be converted into time ordered data.

Anti-aliasing Filter and Decimation

The decimation filter consists of a low-pass anti-aliasing filter and a decimator. The low-pass filter is a CIC filter followed by an FIR filter. The low-pass filter is linear, acausal, and shift

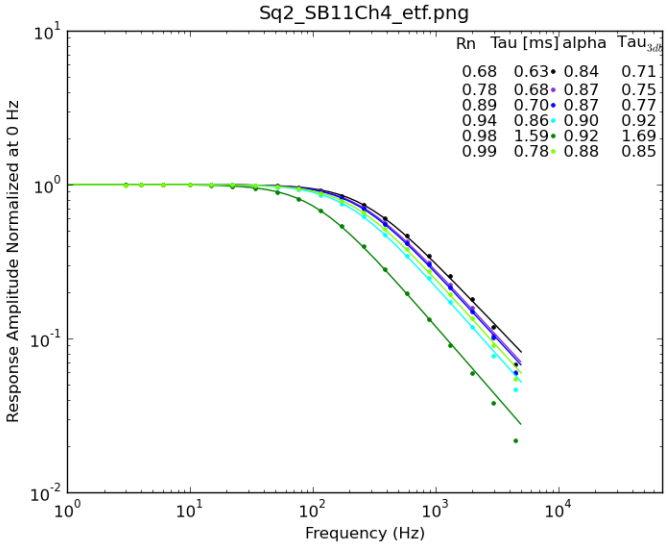


Figure 5.5: TES Frequency Dependent Response

invariant. Once we decimate the output of the anti-aliasing filter, we cause the filter to *not* be shift invariant. The decimator just takes every Nth sample of the output data. If we were to shift the input data by one sample, it would change the form of the output signal. For a cartoon of this see figure 5.6. Because the data is not shift invariant, we cannot construct a matched filter for identifying transient signals. One could hope that it would be possible construct an approximate matched filter by averaging the multitude of impulse responses we have. This ends up being somewhat problematic, due to the acausal nature of the digital anti-aliasing filter. For the acausal filter there is significant ringing in the impulse response. This ringing varies wildly depending on the phase of the input impulse relative to the output sample. See figure 5.7. Any average impulse response one could construct bears little to no resemblance to each individual impulse response.

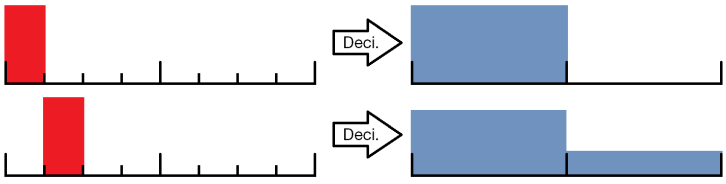


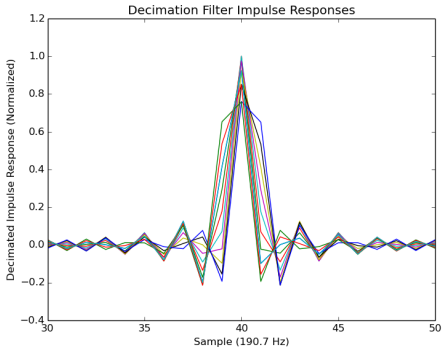
Figure 5.6: Shifting an impulse by one input sample causes the form of the output signal to change. In effect, the decimation filter has multiple impulse responses depending on where the input impulse is relative to the output signal.

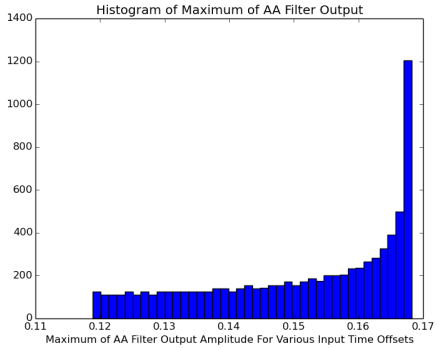
The sample rate of the output data is 190.7 Hz which gives a sample time of $\approx$ 5ms.

Figure 5.7: Decimation Filter Response

(a) A subset of the impulse responses of the decimation filter. The acausal low-pass filter output rings. The exact shape of the impulse depends on the relative phase of the input impulse to the output sample.

(b) The maximums of the various decimation filter outputs for a 1 ms FWHM Gaussian input signal. Only the relative phase of the input signal to the output sample is varied.





For any celestial signal shorter than 5ms our sensitivity is degraded, but not removed. The anti-aliasing filter and detector response act as integrators. The output signal is very appoximately the integral of the signal over a 5 ms sample. One unfortunate aspect of looking through legacy data is we can't increase that sample rate. Because of our sample rate, we cannot resolve the peak flux of an FRB in our data. We are only sensitive to the integral of the flux, fluence, of an event. This does have some benefit in simplifying the analysis. For signals shorter than 5 ms we need not concern ourselves with the actual source profile. The output response is only a function of the source fluence. See figure 5.8.

The decimation filter directs many of the choices in this analysis. It means we cannot use a matched filter. It also means when simulating events, we care a lot about the fluence and the relative phase of the source in the input data stream, but not the actual source profile.

5.5 Event Finding

Short impulses in the detector TOD are located, and every detector with an above threshold significance in a three sample (15 ms) window is grouped into an event. To locate these impulses, we fit a null and alternative model to every sample of the detector TOD and assign a significance defined by the difference of log likelihoods. The null model consists of a linear polynomial and a Heaviside step function. The alternative model includes a delta function along with the linear polynomial and Heaviside step function. The Heaviside function is included to account for rapid changes in the operating resistance of our TES bolometers that can be caused by a number of environmental effects. When estimating the likelihood the noise is assumed to be normally distributed with a variance estimated from

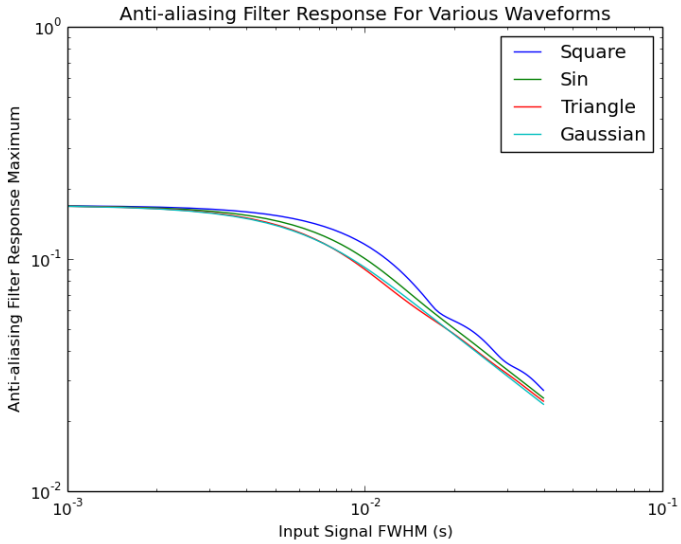


Figure 5.8: The maxima of the decimation filter output TOD as a function of source profile. For each input, the total fluence is held constant. For short signals, the detector response is only a function of the fluence.

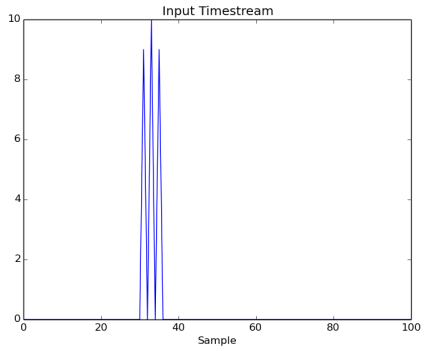
the data. An initial search threshold of a significance of 6 is used when locating events.

The impulse response of the TES bolometers and the readout system is dominated by the digital anti-aliasing (AA) and decimation filter used to reduce the data volume. The AA and decimation filter is not shift invariant, making the construction of a matched filter for use in the event finding impossible. The impulse response of the AA decimation filter depends on the location of the impulse in the input data stream relative to the sample boundaries in the output data stream. Input impulses close to a sample edge of the output data stream have a filter response with the signal spread over two output samples. When multiple TESs observe an identical input impulse noise in the TOD can cause the apparent peak maximum to shift and spread in time. When one includes a noise bias, an identical input impulse in two different detectors with different noise can appear as a maximum signal in neighboring, but different output samples. To find these events we collect above significance threshold events over a three sample window, rather than an individual sample.

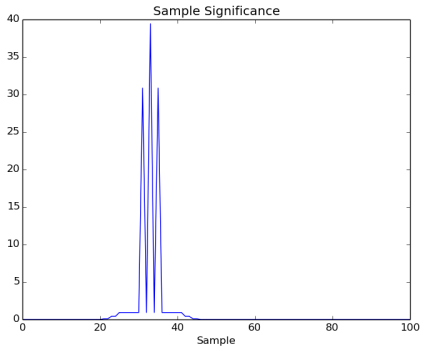
There is an edge case where many consecutive samples have a detector with an above threshold significance. As an example, sample 3, 4, 5, 6, 7, ..., 20 all have above threshold significances. This leads to an ambiguity as to how we would group the event, so we just group all of those samples into an event. This event is removed later when filtering in the events. A structure like that probably indicates an issue with the telescope.

Figure 5.9: A simulation of events and the assigned significance.

(a) Simulated timestream



(b) Timestream Significance



5.6 Event Filtering

We first remove any events without exactly two detectors with an above threshold significance. Filtering to two detectors reduces the number of events by a factor of 25, primarily removing single detector events. It is also a useful veto of radio frequency interference that manifests as more than two detectors with an above threshold significance.

Along with filtering to two detectors in an event, we filter sources of known backgrounds. We filter events if:

- The Heaviside amplitude fit when doing the significance assignment is too large indicating a persistent change in detector resistance.
- The detectors are pointing at a bright point source in the field.
- The assigned significance of these other detectors sharing a DC SQUID and amplifier chain at the event sample are above a threshold indicating an amplifier glitch.
- The weighted sum of the TOD for all the detectors in the same wafer as the event detectors sees a significant event. Cosmic rays interacting with the wafer can heat it leading to a low significance spike for every detector in the wafer.

For the filters with some cut-off value, the threshold was selected by repeating the filtering procedure on simulated transient events while varying the cut-off parameter. The threshold values were set to remove $\approx 2\%$ of simulated events.

While assigning significance we assumed normally distributed noise. This assumption is invalid for some of the time ordered data, however. Various physical mechanisms operate to create these non-Gaussianities. During operation TESs can occasionally become unstable and oscillate. These oscillations manifest as a series of events in the data. This non-Gaussian noise results in a large population of spurious background events. In order to remove these

Filter	Sim Evs Cut	Background Evs Cut
Heaviside Amp.	1.5%	1.4%
Point Source	0.2%	0.7%
Amplifier Glitch	4.0%	2.4%
Wafer Heating	2.0%	0.5%
Non-Normal Noise	40.0%	87.9 %

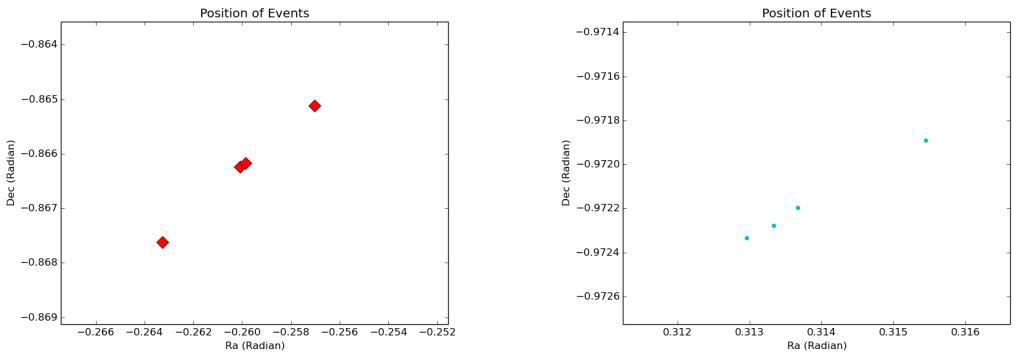
Table 5.1: Simulated events filtered and real background events filtered by the different filters. The non-normal noise filter is the only filter that cuts a large percentage of the data.

events, we record the number of samples over minute and hour-long time scales above a significance threshold for each detector. If the number of super-threshold samples for the detectors in an event is above some cut-off parameter, we remove the event. The cut-off scale was set by maximizing our signal to noise. The signal to noise ratio was estimated as the number of simulated transient events surviving filtering divided by the square root of background events surviving filtering. This resulted in a large number of simulated events being cut by the filter.

Satellite Filtering

Figure 5.10: For the events that were very closely spaced in time, the celestial location of the event was moving at an approximately constant angular velocity.

(a) (b)



From the population of remaining events, a statistically significant population of celestial on-sky signals were detected. The details of the on-sky and background rate estimation are found in section 5.7. These on-sky events were found to be highly clustered in time. Three or four detections occured over the timespan of seconds. The location of the progenitor of these events were moving across the sky with a constant angular velocity. This angular velocity was

consistent with a satellite in low Earth orbit. Through the use of the Joint Space Operations Center (JSpOC) catalog of satellite ephemerides[87], these events could be unambiguously associated (chance coincidence $< 10^{-5}$) with the GRACE and Swarm satellites being in the telescope's field of view.[95, 66] Harmonic distortion of a K-band transmitter on the GRACE satellites appears to create a measurable signal at 150 GHz. [15] The emission mechanism from the Swarm satellites is unknown at this time, though flux estimates indicate that the emission is non-thermal.

For our detectors bandwidth and sample rate, thermal emission from large objects in low Earth orbit can appear as low significance events in this search. In order to filter this background, we remove any events with an object in a patch of sky centered on the detector's field of view. The ephemerides of satellites are not perfectly accurate. In order to correct for that error we scaled the patch of sky used to exclude satellite events until we recovered every Swarm and GRACE detection. The area was then increased by 50% to approximately 2 square degrees. This assumes that the error in the trajectory of the GRACE and Swarm satellites is characteristic of all of the tracked satellites. That coupled with the fact that the JSpOC catalogue excludes classified satellites means that there are possibly an excess of low significance events in the output data. We do not see this low significance excess indicating the satellite filtering is working. When estimating the background rate, we apply a count correction factor for the chance coincidence of events with satellites being in the field of view. We expect to lose roughly 9% of events to being falsely attributed to a satellite signal.

5.7 Background Rate Estimation

After filtering we divide the remaining events into four groups: positive amplitude events with overlapping optical response, positive events without overlapping optical response, negative amplitude events with overlapping optical response and negative amplitude events without overlapping optical response. Only detectors sharing a pixel have overlapping optical response, which we label as on-pixel. For all other TES pairs we use the moniker off-pixel. Cosmic rays interacting with the TES bolometers appear as an impulse in the time ordered data mimicking the signals we are searching for. This creates an excess of positive events as compared to negative events motivating us to use the off-pixel positive events to estimate the background rate of the on-pixel events. This procedure is repeated on the negative amplitude events as a cross check.

In a cosmic ray shower, the density of particles is a function of the lateral distance from the center of the shower[28]. Because of this, we parameterize our background rate estimate in terms of the lateral separation of our TESs, d. We assume that the rate of events between two detectors, i and j is Poisson distributed with mean $l_i l_j \textsc{Ldf}(d_{ij})$, where $\textsc{Ldf}(d_{ij})$ is the rate for two detectors separated by distance d_{ij}. l_i and l_j are order unity liveness correction factors for each detector. We can extend this individual detector rate into a total rate for every detector in the camera at separation d, $R(d)$, by the inclusion of a geometric correction factor, $X(d)$.

$$R(d) = \sum_i l_i \sum_j l_j \textsc{Ldf}(d)\delta(d - d_{ij})b = \textsc{Ldf}(d)X(d) \tag{5.4}$$

We label $\textsc{Ldf}(d)$ our lateral distribution function in analogy to cosmic ray physics. The geometric correction factor $X(d)$ is estimated from the physical layout of the TES bolometers and their yield. To first order, it is the average number of detectors at the physical separation d from a TES in the array.

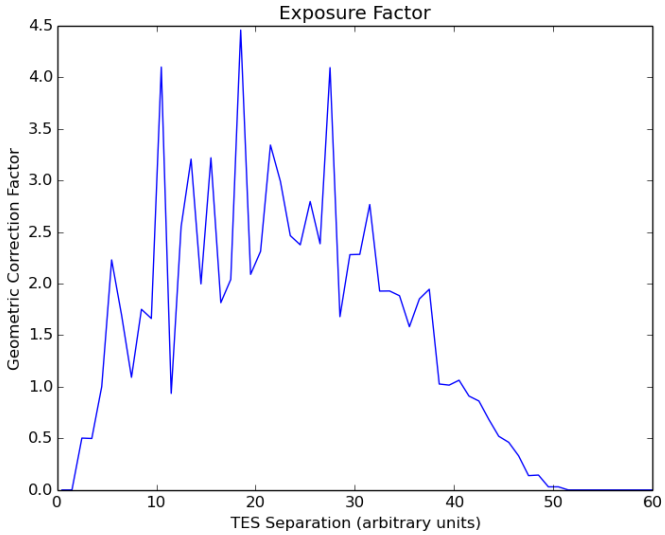


Figure 5.11: The geometric correction factor. This is the average number of live detectors at a distance away from an individual detector. It grows roughly linearly from the flat geometry and then tapers off from arriving at the edge of the wafer.

We use the off-pixel events to estimate $r(d)$ and then use that as a predictor of the on-pixel background rate. With the monolithic array fabrication of the 150 GHz detectors, the intra-pixel TES separation is constant for all pixels. The inter-pixel TES separations cover distances that are smaller and larger than the intra-pixel separation. This allows us to constrain $r(d)$ on both sides of the intra-pixel separation from off-pixel information.

With the frequency-multiplexed readout system, there are other possible ways of grouping detectors. These include the readout frequency, the physical location of the inductor used in the LC filter attached to the TES, and whether the two detectors share an amplifier chain. Physically close inductors and close frequencies would manifest as an increased sensitivity to radio frequency interference. The physical closeness of the inductors is highly correlated with sharing a readout amplifier chain. We compared the background rate estimate using only detectors sharing an amplifier chain to the background rate estimate from the entire data set and found it to be statistically consistent.

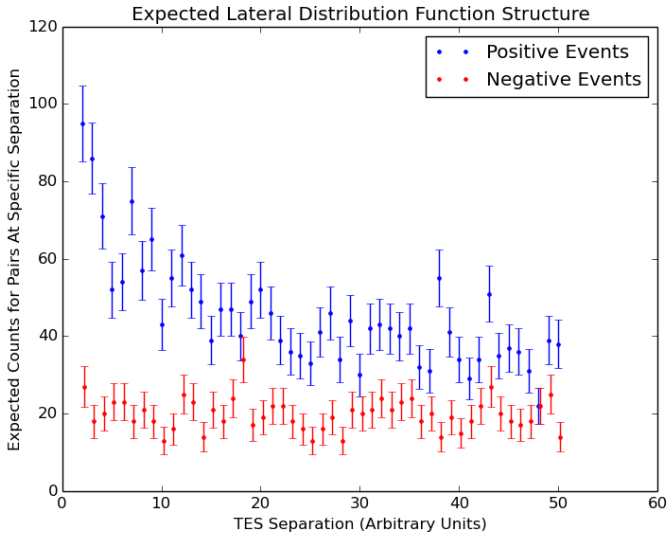


Figure 5.12: This is a simulated lateral distribution function structure. This plot is included to give an idea of what the expected structure of a lateral distribution function is. The plotted negative events were generated from uncorrelated noise. The positive events included a an additional rate component that is correlated with separation.

5.8 Event Simulation

When estimating on-sky counts and the background rate, we divide the events into groups based off of their significance. To convert the significance thresholds into physically meaningful parameters, we rely on Monte Carlo simulations. Sources with a known fluence are added to the time ordered data and the search process is repeated on the simulated events. This simulation process gives us the percent of events found as a function of significance range and fluence.

To simulate events we pick an input source profile, scale the flux over time by the detector beam scanning over the source, apply the decimation filter, and then add it to the time ordered data. When adding the simulated signal to the time ordered data, we assume an unpolarized source signal, and add the same simulated signal to both polarization sensitive detector's time ordered data.

For sources shorter than our sample rate, the input source profile has little effect on the output signal. It is only a function of the signal fluence. Because the exact structure of the input signal does not matter we just use a Gaussian input source profile. For the flux of the input events we assumed a flat source spectrum. All results are quoted assuming a flat spectrum.

We modelled the on-sky event rate as the power law of the form:

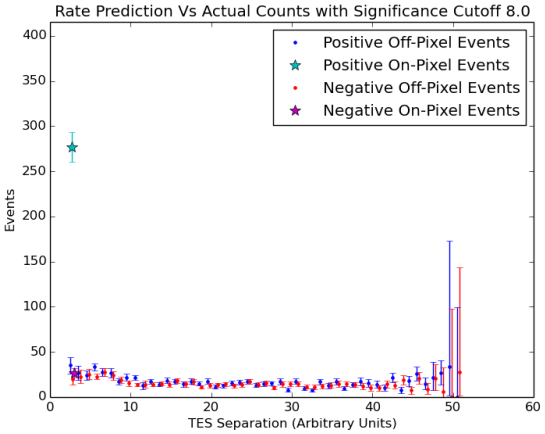


Figure 5.13: Lateral distribution function normalized to units of expected on pixel count rate. The negative on-pixel events are statistically consistent with the estimated background rate from off pixel events, while the positive on-pixel rate shows a significant excess. This excess is caused by satellites being in view of the detectors. The satellite signals appear as transients in the data due to the scanning of the telescope. The telescope scan speed is $2^{\circ}s^{-1}$ while low Earth orbit satellites move with an apparent angular velocity less than $1^{\circ}s^{-1}$.

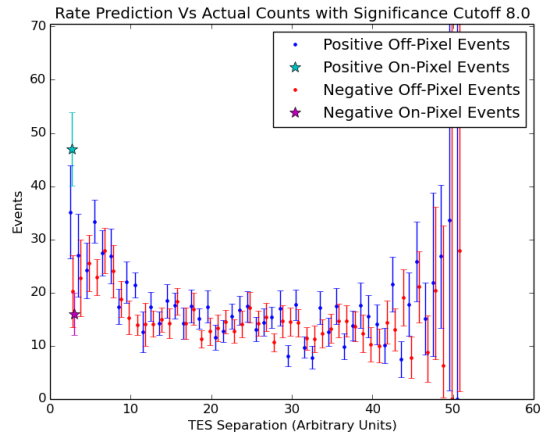


Figure 5.14: This is exactly the same plot as above, but now with the satellites filtered from the data. With our satellite filtering procedure we expect to remove only 9% of events if this is a chance coincidence.

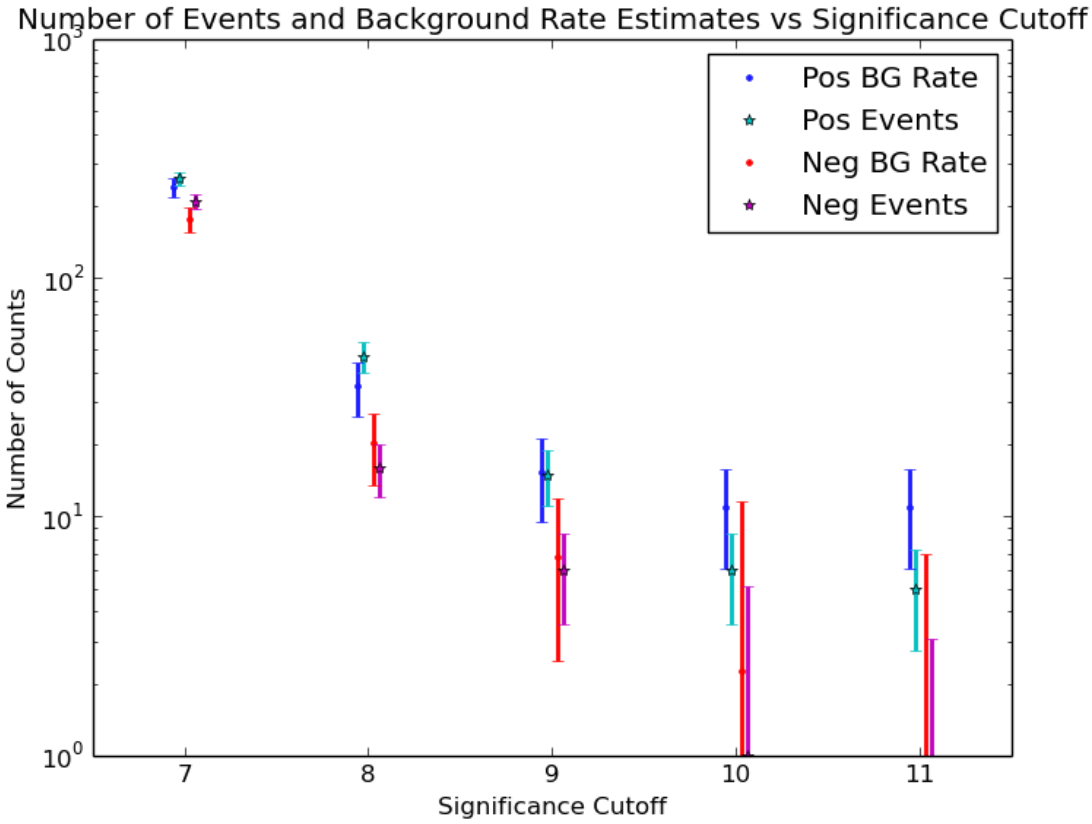


Figure 5.15: The background rate prediction and on pixel rate are consistent after satellite filtering for both positive and negative amplitude events. There are slightly more positive events than negative events.

$$R_{sky} = R_0 \left(\frac{F}{F_0} \right)^{\alpha} \tag{5.5}$$

Using the detector live time and our simulations we are able to convert and on-sky rate into our expected event rate with the equation:

$$R_{counts} = \phi \Omega T \int_0^{\infty} P_{sig}(F) R_0 \frac{F^{(\alpha-1)}}{F_0^{\alpha}} dF \tag{5.6}$$

Where Ω is the instantaneous sky area observed. T is the time of observation. ϕ is the correction factor for the chance coincidence of an event with a satellite. $P_{sig}(F)$ is the percent of events found after filtering in the significance range.

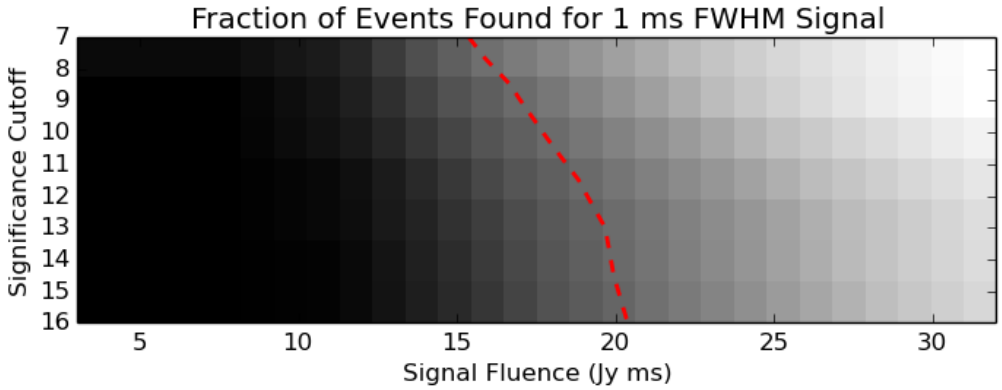


Figure 5.16: Percent found as a function of significance cutoff and source fluence. The contour plots where 33% of events are recovered compared to the 32 Jy ms, significance cutoff 7 value. The kink in the contour is just noise from the Monte Carlo based estimation of the percent found.

5.9 On-Sky Constraints

We generate Feldman Cousins confidence intervals with Monte Carlo simulations using the method of confidence belts. For the background rate estimate, we use profile likelihoods to set the background rate. The equation has the form:

$$\begin{aligned}\underset{B_{sig}}{\operatorname{argmax}}\ \mathcal{L}(x, \textsc{Ldf}|S_{sig}, B_{sig}) =& \mathcal{P}(x, S_{sig} + B_{sig}) + \\ & \mathcal{P}(\textsc{Ldf}, XB_{sig})\end{aligned} \tag{5.7}$$

$\mathcal{P}(x, \mu)$ is the Poisson likelihood of x counts with mean μ. S_{sig} is the on-sky rate for our significance band. B_{sig} is the estimated background rate in our significance band. When constructing the confidence belt for a S_{sig}, we use the B_{sig} that maximizes the likelihood

of the on-sky rate and the likelihood of the lateral distribution function. This reduces the dimensionality of the parameters needed for confidence interval construction. For estimating the likelihood of the lateral distribution function for a given background rate, we model the LDF rate in the region of the detector spacing as a low order polynomial that intersects our background rate. This is implicitly assuming the background rate is smoothly varying with separation. This allows us to estimate the likelihood of the lateral distribution function for a given background rate.

For the rate constraint Monte Carlo we divide the data into three significance regions: the events with both detector significances above 7, 9 and 13, above 9 but not above 13, and the events above significance of 7. Most of the constraining power comes from the low-background, high-significance region since the background rate drops exponentially with significance, while the expected on-sky signal drops at a polynomial rate with significance. Changing the lowest significance cut-off to 8 changes the on-sky constraints by less than 3% indicating that lowering the significance threshold further would have little effect on the analysis.

We found no statistically significant excess of on-sky events over our background rate once the satellite signals had been removed. The 90% upper confidence limit is 1.1×10^5 sky^{-1} day^{-1} for 1 ms signals with fluence above 10 Jy ms. With our analysis techniques, the constraints degrade for signals longer than our sample size. Our sensitivity roughly tracks the fluence integrated over one 5 ms sample.

5.10 Discussion

Our 150 GHz on-sky rate constraint is significantly worse than the 1.4 GHz rate constraint of roughly $1.7^{+1.5}_{-0.9} \times 10^3$ FRBs sky^{-1} day^{-1} above 2 Jy ms[6]. If we model the event rate as a function of frequency along with fluence in the form:

$$R = R_0 \left(\frac{F}{F_0}\right)^{\alpha} \nu^{\gamma} \tag{5.8}$$

we place a 0.95 confidence limit on the spectral index of FRBs of $\gamma < 1.6$ when using the 1.4 GHz rate stated above. The sensitivity of the survey to these transients is primarily limited by the large background event rate in the SPTpol instrument.

Some future CMB telescopes, like SPT3G and POLARBEAR 2[94, 5], have a multi-chroic pixel design. In contrast to the pixels in SPTpol, where only one frequency band is measured, these telescopes will have two or three frequency bands measured per pixel. This increases the number of detectors with overlapping beams from two to four or six. Requiring that 4 detectors detectors coincidentally see an event dramatically reduces the background rate of this type of search. When factoring in the increased pixel number, SPT3g will be 50 to 100 times more sensitive than SPTpol. This estimate was constructed by measuring the 4 detector event rate in SPTpol and extrapolating that event rate to SPT3g. Predicting

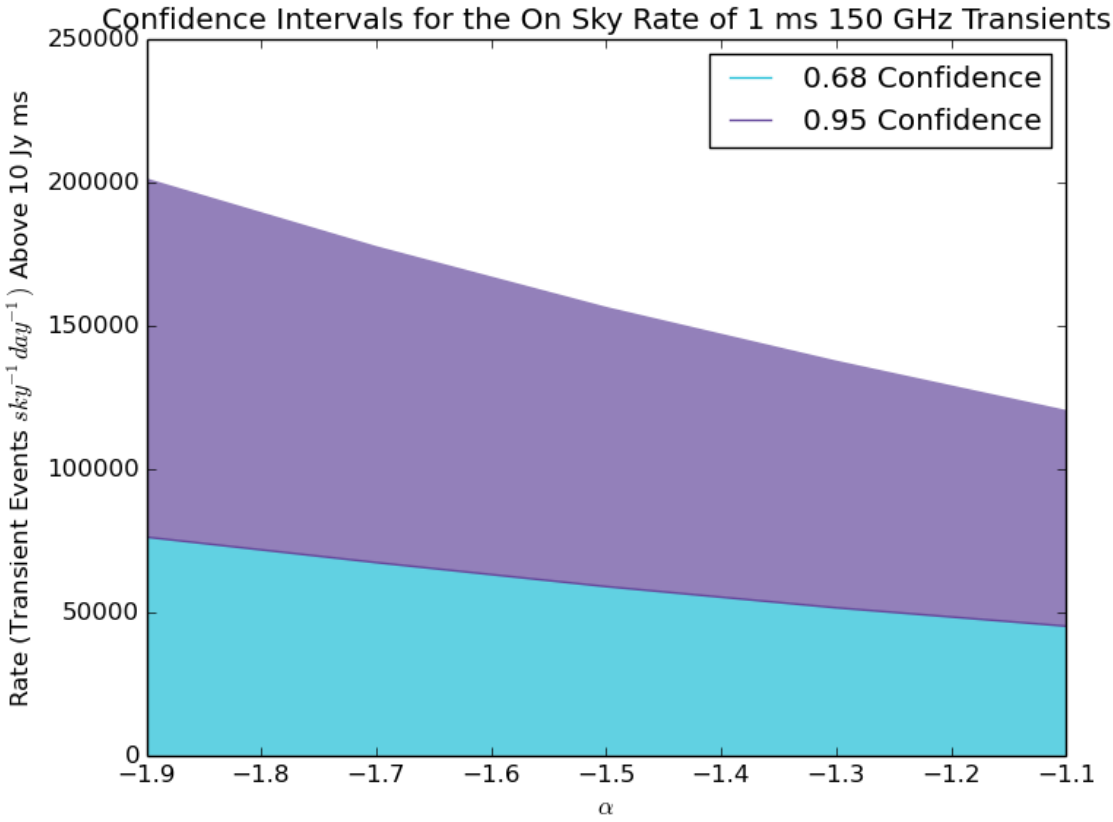


Figure 5.17: This is the on-sky rate Feldman-Cousins confidence intervals estimated with Monte-Carlo simulations using the method of confidence belts for a fluence distribution following a power law. α is the power law index.

the sensitivity of other telescopes to these sorts of transients is extremely difficult because it relies on a characterization of non-normal noise in the system which is rarely published.

For all of our rate constraints we assumed a flat spectrum for the transient source. For the only detected repeating burst, FRB 121102, the spectrum is best described by a peak with a width of 500 MHz[51]. The bandwidth does not appear to increase for FRBs detected at higher frequencies[63]. If the sources we are searching for have a similar narrow bandwidth at 150 GHz rather than flat spectrum, the constraints on spectral index will be degraded. This degredation is roughly equivalent to the source bandwidth divided by the bolometric bandwidth. For SPT-3G this is a ratio of 1/100. Improvements to the source constraints for SPT3g rely on measuring a larger bandwidth with multiple TES bolometers and would not be fully realized for narrow band signals.

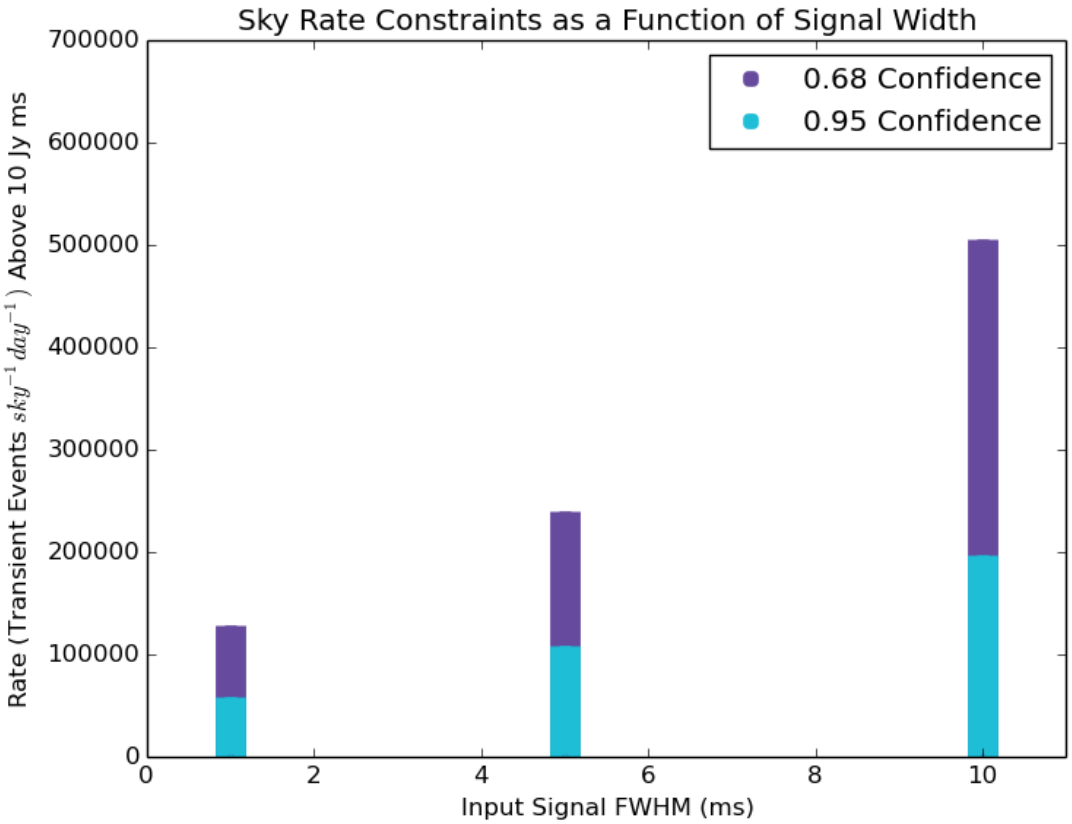


Figure 5.18: For $\alpha = -1.5$ this is the on-sky rate confidence intervals as a function of signal width for sources with a fixed fluence. With our search strategy we are sensitive to the integral of the fluence over our 5 ms sample window. This degrades our sensitivity to longer sources because the flux is spread over multiple samples.

Bibliography

[1] K. N. Abazajian et al. "Neutrino physics from the cosmic microwave background and large scale structure". In: *Astroparticle Physics* 63 (Mar. 2015), pp. 66–80. DOI: 10.1016/j.astropartphys.2014.05.014. arXiv: 1309.5383.

[2] M. Abramowitz and I.A. Stegun. *Handbook of Mathematical Functions*. Fifth. New York: Dover, 1964.

[3] P. A. R. Ade et al. "A review of metal mesh filters". In: *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. Vol. 6275. Proc. SPIE. June 2006, 62750U. DOI: 10.1117/12.673162.

[4] A. J. Annunziata et al. "Tunable superconducting nanoinductors". In: *Nanotechnology* 21, 445202 (Nov. 2010), p. 445202. DOI: 10.1088/0957-4484/21/44/445202. arXiv: 1007.4187 [cond-mat.supr-con].

[5] B. A. Benson et al. "SPT-3G: a next-generation cosmic microwave background polarization experiment on the South Pole telescope". In: *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII*. Vol. 9153. Proc. SPIE. July 2014, 91531P. DOI: 10.1117/12.2057305. arXiv: 1407.2973 [astro-ph.IM].

[6] S. Bhandari et al. "The SUrvey for Pulsars and Extragalactic Radio Bursts II: New FRB discoveries and their follow-up". In: *ArXiv e-prints* (Nov. 2017). arXiv: 1711.08110 [astro-ph.HE].

[7] S. Bhattacharyya. "Fast Radio Bursts from neutron stars plunging into black holes". In: *ArXiv e-prints* (Nov. 2017). arXiv: 1711.09083 [astro-ph.HE].

[8] BICEP2 Collaboration et al. "BICEP2. II. Experiment and three-year Data Set". In: *ApJ* 792, 62 (Sept. 2014), p. 62. DOI: 10.1088/0004-637X/792/1/62. arXiv: 1403.4302.

[9] M. Caleb et al. "The first interferometric detections of Fast Radio Bursts". In: *ArXiv e-prints* (Mar. 2017). arXiv: 1703.10173 [astro-ph.HE].

[10] C. M. Cantalupo et al. "MADmap: A Massively Parallel Maximum Likelihood Cosmic Microwave Background Map-maker". In: *ApJS* 187 (Mar. 2010), pp. 212–227. DOI: 10.1088/0067-0049/187/1/212. arXiv: 0906.1775.

[11] A. Challinor and A. Lewis. "Lensed CMB power spectra from all-sky correlation functions". In: *PhRvD* 71.10, 103010 (May 2005), p. 103010. DOI: 10.1103/PhysRevD.71.103010. eprint: astro-ph/0502425.

[12] D. J. Champion et al. "Five new fast radio bursts from the HTRU high-latitude survey at Parkes: first evidence for two-component bursts". In: *MNRAS* 460 (July 2016), pp. L30–L34. DOI: 10.1093/mnrasl/slw069. arXiv: 1511.07746 [astro-ph.HE].

[13] S. Chatterjee et al. "A direct localization of a fast radio burst and its host". In: *Nature* 541 (Jan. 2017), pp. 58–61. DOI: 10.1038/nature20797. arXiv: 1701.01098 [astro-ph.HE].

[14] A. T. Crites et al. "Measurements of E-Mode Polarization and Temperature-E-Mode Correlation in the Cosmic Microwave Background from 100 Square Degrees of SPTpol Data". In: *ApJ* 805, 36 (May 2015), p. 36. DOI: 10.1088/0004-637X/805/1/36. arXiv: 1411.1042.

[15] A. Davis. private correspondence. 2017.

[16] A. T. A. M. de Waele and R. de Bruyn Ouboter. "Quantum-interference phenomena in point contacts between two superconductors". In: *Physica* 41 (Feb. 1969), pp. 225–254. DOI: 10.1016/0031-8914(69)90116-5.

[17] M. A. Dobbs et al. "Frequency multiplexed superconducting quantum interference device readout of large bolometer arrays for cosmic microwave background measurements". In: *Review of Scientific Instruments* 83.7, 073113-073113-24 (July 2012), pages. DOI: 10.1063/1.4737629. arXiv: 1112.4215 [astro-ph.IM].

[18] Scott Dodelson. *Modern cosmology*. San Diego, CA: Academic Press, 2003. URL: https://cds.cern.ch/record/1282338.

[19] O. Doré et al. "MAPCUMBA: A fast iterative multi-grid map-making algorithm for CMB experiments". In: *AAP* 374 (July 2001), pp. 358–370. DOI: 10.1051/0004-6361:20010692. eprint: astro-ph/0101112.

[20] T. Dray. "The relationship between monopole harmonics and spin-weighted spherical harmonics". In: *Journal of Mathematical Physics* 26 (May 1985), pp. 1030–1033. DOI: 10.1063/1.526533.

[21] R. Dünner et al. "The Atacama Cosmology Telescope: Data Characterization and Mapmaking". In: *ApJ* 762, 10 (Jan. 2013), p. 10. DOI: 10.1088/0004-637X/762/1/10. arXiv: 1208.0050 [astro-ph.IM].

[22] G. Efstathiou. "Myths and truths concerning estimation of power spectra: the case for a hybrid estimator". In: *MNRAS* 349 (Apr. 2004), pp. 603–626. DOI: 10.1111/j.1365-2966.2004.07530.x. eprint: astro-ph/0307515.

[23] H. Falcke and L. Rezzolla. "Fast radio bursts: the last sign of supramassive neutron stars". In: *AAP* 562, A137 (Feb. 2014), A137. DOI: 10.1051/0004-6361/201321996. arXiv: 1307.1409 [astro-ph.HE].

[24] Richard P. Feynman, Robert B. Leighton, and Matthew Sands. *The Feynman lectures on physics. Vol. 3: Quantum mechanics*. Addison-Wesley Publishing Co., Inc., Reading, Mass.-London, 1965.

[25] E. M. George et al. "A Measurement of Secondary Cosmic Microwave Background Anisotropies from the 2500 Square-degree SPT-SZ Survey". In: *ApJ* 799, 177 (Feb. 2015), p. 177. DOI: 10.1088/0004-637X/799/2/177. arXiv: 1408.3161.

[26] K. M. Gorski et al. "The HEALPix Primer". In: *ArXiv Astrophysics e-prints* (May 1999). eprint: astro-ph/9905275.

[27] K. M. Górski et al. "HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere". In: *ApJ* 622 (Apr. 2005), pp. 759–771. DOI: 10.1086/427976. eprint: arXiv:astro-ph/0409513.

[28] K. Greisen. "Cosmic Ray Showers". In: *Annual Review of Nuclear and Particle Science* 10 (1960), pp. 63–108. DOI: 10.1146/annurev.ns.10.120160.000431.

[29] H. F. Gruetjen et al. "Using inpainting to construct accurate cut-sky CMB estimators". In: *ArXiv e-prints* (Oct. 2015). arXiv: 1510.03103.

[30] D. Hanson et al. "Detection of B-Mode Polarization in the Cosmic Microwave Background with Data from the South Pole Telescope". In: *Physical Review Letters* 111.14, 141301 (Oct. 2013), p. 141301. DOI: 10.1103/PhysRevLett.111.141301. arXiv: 1307.5830 [astro-ph.CO].

[31] E. Hivon et al. "MASTER of the Cosmic Microwave Background Anisotropy Power Spectrum: A Fast Method for Statistical Analysis of Large and Complex Cosmic Microwave Background Data Sets". In: *ApJ* 567 (Mar. 2002), pp. 2–17. arXiv: astro-ph/0105302.

[32] E. Hivon et al. "MASTER of the Cosmic Microwave Background Anisotropy Power Spectrum: A Fast Method for Statistical Analysis of Large and Complex Cosmic Microwave Background Data Sets". In: *ApJ* 567 (Mar. 2002), pp. 2–17. DOI: 10.1086/338126. eprint: astro-ph/0105302.

[33] W. Hu. "CMB temperature and polarization anisotropy fundamentals". In: *Annals of Physics* 303 (Jan. 2003), pp. 203–225. DOI: 10.1016/S0003-4916(02)00022-2. eprint: astro-ph/0210696.

[34] W. Hu. "Weak lensing of the CMB: A harmonic approach". In: *PhRvD* 62.4, 043007 (Aug. 2000), p. 043007. DOI: 10.1103/PhysRevD.62.043007. eprint: astro-ph/0001303.

[35] Martin Huber et al. "DC SQUID series array amplifiers with 120 MHz bandwidth". In: 11 (Apr. 2001), pp. 1251–1256.

[36] *IAU Polarization Convention*. https://lambda.gsfc.nasa.gov/product/about/IAU_Conventions_v2.svg. Accessed: 2018-06-02.

[37] K. D. Irwin and G. C. Hilton. "Transition-Edge Sensors". In: *Cryogenic Particle Detection*. Ed. by C. Enss. 2005, p. 63. DOI: 10.1007/10933596_3.

[38] A. Iwazaki. "Axion Stars and Repeating Fast Radio Bursts with Finite Bandwidths". In: *ArXiv e-prints* (July 2017). arXiv: 1707.04827 [astro-ph.HE].

[39] M. Kamionkowski and E. D. Kovetz. "The Quest for B Modes from Inflationary Gravitational Waves". In: *ARA&A* 54 (Sept. 2016), pp. 227–269. DOI: 10.1146/annurev-astro-081915-023433. arXiv: 1510.06042.

[40] K. Kashiyama, K. Ioka, and P. Mészáros. "Cosmological Fast Radio Bursts from Binary White Dwarf Mergers". In: *ApJL* 776, L39 (Oct. 2013), p. L39. DOI: 10.1088/2041-8205/776/2/L39. arXiv: 1307.7708 [astro-ph.HE].

[41] J. I. Katz. "Fast radio bursts - A brief review: Some questions, fewer answers". In: *Modern Physics Letters A* 31, 1630013 (Apr. 2016), p. 1630013. DOI: 10.1142/S0217732316300135. arXiv: 1604.01799 [astro-ph.HE].

[42] J. I. Katz. "FRB as products of accretion disc funnels". In: *MNRAS* 471 (Oct. 2017), pp. L92–L95. DOI: 10.1093/mnrasl/slx113. arXiv: 1704.08301 [astro-ph.HE].

[43] J. I. Katz. "How Soft Gamma Repeaters Might Make Fast Radio Bursts". In: *ApJ* 826, 226 (Aug. 2016), p. 226. DOI: 10.3847/0004-637X/826/2/226. arXiv: 1512.04503 [astro-ph.HE].

[44] E. F. Keane et al. "Rotating Radio Transients: new discoveries, timing solutions and musings". In: *MNRAS* 415 (Aug. 2011), pp. 3065–3080. DOI: 10.1111/j.1365-2966.2011.18917.x. arXiv: 1104.2727 [astro-ph.SR].

[45] E. Keihänen, H. Kurki-Suonio, and T. Poutanen. "MADAM- a map-making method for CMB experiments". In: *MNRAS* 360 (June 2005), pp. 390–400. DOI: 10.1111/j.1365-2966.2005.09055.x. eprint: astro-ph/0412517.

[46] E. Keihänen et al. "Making cosmic microwave background temperature and polarization maps with MADAM". In: *AAP* 510, A57 (Feb. 2010), A57. DOI: 10.1051/0004-6361/200912813. arXiv: 0907.0367.

[47] R. Keisler et al. "Measurements of Sub-degree B-mode Polarization in the Cosmic Microwave Background from 100 Square Degrees of SPTpol Data". In: *ApJ* 807, 151 (July 2015), p. 151. DOI: 10.1088/0004-637X/807/2/151. arXiv: 1503.02315.

[48] P. G. Klemens. "The Lattice Component of the Thermal Conductivity of Metals and Alloys". In: *Australian Journal of Physics* 7 (Mar. 1954), p. 57. DOI: 10.1071/PH540057.

[49] Justin Kunimune. *Map-Projections*. https://github.com/jkunimune15/Map-Projections. 2018.

[50] A. Kusaka et al. "Results from the Atacama B-mode Search (ABS) Experiment". In: *ArXiv e-prints* (Jan. 2018). arXiv: 1801.01218.

[51] C. J. Law et al. "A Multi-telescope Campaign on FRB 121102: Implications for the FRB Population". In: *ApJ* 850, 76 (Nov. 2017), p. 76. DOI: 10.3847/1538-4357/aa9700. arXiv: 1705.07553 [astro-ph.HE].

[52] L. C. Lee and J. R. Jokipii. "Strong scintillations in astrophysics. II - A theory of temporal broadening of pulses". In: *ApJ* 201 (Oct. 1975), pp. 532–543. DOI: 10.1086/153916.

[53] J. Lesgourgues. "The Cosmic Linear Anisotropy Solving System (CLASS) III: Comparision with CAMB for LambdaCDM". In: *ArXiv e-prints* (Apr. 2011). arXiv: 1104.2934.

[54] A. Lewis and S. Bridle. "Cosmological parameters from CMB and other data: A Monte Carlo approach". In: *PhRvD* 66.10, 103511 (Nov. 2002), p. 103511. DOI: 10.1103/PhysRevD.66.103511. eprint: astro-ph/0205436.

[55] A. Lewis and A. Challinor. "Weak gravitational lensing of the CMB". In: *PhR* 429 (June 2006), pp. 1–65. DOI: 10.1016/j.physrep.2006.03.002. eprint: astro-ph/0601594.

[56] M. Lingam and A. Loeb. "Fast Radio Bursts from Extragalactic Light Sails". In: *ApJL* 837, L23 (Mar. 2017), p. L23. DOI: 10.3847/2041-8213/aa633e. arXiv: 1701.01109 [astro-ph.HE].

[57] X. Liu. "A model of neutron-star–white-dwarf collision for fast radio bursts". In: *ArXiv e-prints* (Dec. 2017). arXiv: 1712.03509 [astro-ph.HE].

[58] D. R. Lorimer et al. "A Bright Millisecond Radio Burst of Extragalactic Origin". In: *Science* 318 (Nov. 2007), p. 777. DOI: 10.1126/science.1147532. arXiv: 0709.4301.

[59] M. Lyutikov, L. Burzawa, and S. B. Popov. "Fast radio bursts as giant pulses from young rapidly rotating pulsars". In: *MNRAS* 462 (Oct. 2016), pp. 941–950. DOI: 10.1093/mnras/stw1669. arXiv: 1603.02891 [astro-ph.HE].

[60] B. Marcote et al. "The Repeating Fast Radio Burst FRB 121102 as Seen on Milliarcsecond Angular Scales". In: *ApJL* 834, L8 (Jan. 2017), p. L8. DOI: 10.3847/2041-8213/834/2/L8. arXiv: 1701.01099 [astro-ph.HE].

[61] K. Masui et al. "Dense magnetized plasma associated with a fast radio burst". In: *Nature* 528 (Dec. 2015), pp. 523–525. DOI: 10.1038/nature15769. arXiv: 1512.00529 [astro-ph.HE].

[62] J. C. Mather. "Bolometer noise: nonequilibrium thoery". In: *ApOpt* 21 (Mar. 1982), pp. 1125–1129. DOI: 10.1364/AO.21.001125.

[63] D. Michilli et al. "An extreme magneto-ionic environment associated with the fast radio burst source FRB 121102". In: *ArXiv e-prints* (Jan. 2018). arXiv: 1801.03965 [astro-ph.HE].

[64] T. J. Moriya. "Radio Transients from Accretion-induced Collapse of White Dwarfs". In: *ApJL* 830, L38 (Oct. 2016), p. L38. DOI: 10.3847/2041-8205/830/2/L38. arXiv: 1610.00416 [astro-ph.HE].

[65] T. Okamoto and W. Hu. "Cosmic microwave background lensing reconstruction on the full sky". In: *PhRvD* 67.8, 083002 (Apr. 2003), p. 083002. DOI: 10.1103/PhysRevD.67.083002. eprint: astro-ph/0301031.

[66] N. Olsen et al. "The Swarm Satellite Constellation Application and Research Facility (SCARF) and Swarm data products". In: *Earth, Planets, and Space* 65 (Nov. 2013), pp. 1189–1200. DOI: 10.5047/eps.2013.07.001.

[67] A. A. Penzias and R. W. Wilson. "A Measurement of Excess Antenna Temperature at 4080 Mc/s." In: *ApJ* 142 (July 1965), pp. 419–421. DOI: 10.1086/148307.

[68] E. Petroff et al. "A real-time fast radio burst: polarization detection and multiwavelength follow-up". In: *MNRAS* 447 (Feb. 2015), pp. 246–255. DOI: 10.1093/mnras/stu2419. arXiv: 1412.0342 [astro-ph.HE].

[69] E. Petroff et al. "A survey of FRB fields: limits on repeatability". In: *MNRAS* 454 (Nov. 2015), pp. 457–462. DOI: 10.1093/mnras/stv1953. arXiv: 1508.04884 [astro-ph.HE].

[70] E. Petroff et al. "FRBCAT: The Fast Radio Burst Catalogue". In: *PASA* 33, e045 (Sept. 2016), e045. DOI: 10.1017/pasa.2016.35. arXiv: 1601.03547 [astro-ph.HE].

[71] Planck Collaboration et al. "Planck 2013 results. VI. High Frequency Instrument data processing". In: *AAP* 571, A6 (Nov. 2014), A6. DOI: 10.1051/0004-6361/201321570. arXiv: 1303.5067.

[72] Planck Collaboration et al. "Planck 2013 results. XXIV. Constraints on primordial non-Gaussianity". In: *AAP* 571, A24 (Nov. 2014), A24. DOI: 10.1051/0004-6361/201321554. arXiv: 1303.5084.

[73] Planck Collaboration et al. "Planck 2015 results. XIII. Cosmological parameters". In: *AAP* 594, A13 (Sept. 2016), A13. DOI: 10.1051/0004-6361/201525830. arXiv: 1502.01589.

[74] Planck Collaboration et al. "Planck 2015 results. XVII. Constraints on primordial non-Gaussianity". In: *AAP* 594, A17 (Sept. 2016), A17. DOI: 10.1051/0004-6361/201525836. arXiv: 1502.01592.

[75] Planck Collaboration et al. "Planck intermediate results. XLIX. Parity-violation constraints from polarization data". In: *AAP* 596, A110 (Dec. 2016), A110. DOI: 10.1051/0004-6361/201629018. arXiv: 1605.08633.

[76] Polarbear Collaboration et al. "A Measurement of the Cosmic Microwave Background B-mode Polarization Power Spectrum at Sub-degree Scales with POLARBEAR". In: *ApJ* 794, 171 (Oct. 2014), p. 171. DOI: 10.1088/0004-637X/794/2/171. arXiv: 1403.2369.

[77] D. Poletti et al. "Making maps of cosmic microwave background polarization for B-mode studies: the POLARBEAR example". In: *AAP* 600, A60 (Apr. 2017), A60. DOI: 10.1051/0004-6361/201629467. arXiv: 1608.01624 [astro-ph.IM].

[78] T. Poutanen et al. "Comparison of map-making algorithms for CMB experiments". In: *AAP* 449 (Apr. 2006), pp. 1311–1322. DOI: 10.1051/0004-6361:20052845. eprint: astro-ph/0501504.

[79] J. G. Rosa and T. W. Kephart. "Black hole lasers powered by axion superradiant instabilities". In: *ArXiv e-prints* (Sept. 2017). arXiv: 1709.06581 [gr-qc].

[80] P. Scholz et al. "The Repeating Fast Radio Burst FRB 121102: Multi-wavelength Observations and Additional Bursts". In: *ApJ* 833, 177 (Dec. 2016), p. 177. DOI: 10.3847/1538-4357/833/2/177. arXiv: 1603.08880 [astro-ph.HE].

[81] U. Seljak et al. "Comparison of cosmological Boltzmann codes: Are we ready for high precision cosmology?" In: *PhRvD* 68.8, 083507 (Oct. 2003), p. 083507. DOI: 10.1103/PhysRevD.68.083507. eprint: astro-ph/0306052.

[82] K. M. Smith. "Pseudo-C_l estimators which do not mix E and B modes". In: *PhRvD* 74.8, 083002 (Oct. 2006), p. 083002. DOI: 10.1103/PhysRevD.74.083002. eprint: astro-ph/0511629.

[83] K. M. Smith and M. Zaldarriaga. "General solution to the E-B mixing problem". In: *PhRvD* 76.4, 043001 (Aug. 2007), p. 043001. DOI: 10.1103/PhysRevD.76.043001. eprint: astro-ph/0610059.

[84] G. F. Smoot et al. "Structure in the COBE differential microwave radiometer first-year maps". In: *ApJL* 396 (Sept. 1992), pp. L1–L5. DOI: 10.1086/186504.

[85] John P. Snyder. *Map projections–a working manual*. English. Washington: U.S. Geological Survey, 1987.

[86] John Parr Snyder. *Map projections used by the U.S. Geological Survey*. 2nd ed. United States Government Printing Office, 1984.

[87] *Space-Track*. https://www.space-track.org/. Accessed: 2017-12.

[88] L. G. Spitler et al. "A repeating fast radio burst". In: *Nature* 531 (Mar. 2016), pp. 202–205. DOI: 10.1038/nature17168. arXiv: 1603.00581 [astro-ph.HE].

[89] L. G. Spitler et al. "Fast Radio Burst Discovered in the Arecibo Pulsar ALFA Survey". In: *ApJ* 790, 101 (Aug. 2014), p. 101. DOI: 10.1088/0004-637X/790/2/101. arXiv: 1404.2934 [astro-ph.HE].

[90] J.-L. Starck, M. J. Fadili, and A. Rassat. "Low-l CMB analysis and inpainting". In: *AAP* 550, A15 (Feb. 2013), A15. DOI: 10.1051/0004-6361/201220332. arXiv: 1210.6587.

[91] R. Stompor et al. "Making maps of the cosmic microwave background: The MAXIMA example". In: *PhRvD* 65.2, 022003 (Jan. 2002), p. 022003. DOI: 10.1103/PhysRevD.65.022003. eprint: astro-ph/0106451.

[92] K. T. Story et al. "A Measurement of the Cosmic Microwave Background Damping Tail from the 2500-Square-Degree SPT-SZ Survey". In: *ApJ* 779, 86 (Dec. 2013), p. 86. DOI: 10.1088/0004-637X/779/1/86. arXiv: 1210.7231.

[93] D. Sutton et al. "Map making in small field modulated CMB polarization experiments: approximating the maximum likelihood method". In: *MNRAS* 393 (Mar. 2009), pp. 894–910. DOI: 10.1111/j.1365-2966.2008.14195.x. arXiv: 0807.3658.

[94] A. Suzuki et al. "Multi-chroic Dual-Polarization Bolometric Focal Plane for Studies of the Cosmic Microwave Background". In: *Journal of Low Temperature Physics* 167 (June 2012), pp. 852–858. DOI: 10.1007/s10909-012-0602-y. arXiv: 1210.8256 [astro-ph.IM].

[95] B. D. Tapley et al. "The gravity recovery and climate experiment: Mission overview and early results". In: *Geophys. Res. Lett.* 31, L09607 (May 2004), p. L09607. DOI: 10.1029/2004GL019920.

[96] M. Tegmark. "How to measure CMB power spectra without losing information". In: *PhRvD* 55 (May 1997), pp. 5895–5907.

[97] M. Tegmark and A. de Oliveira-Costa. "How to measure CMB polarization power spectra without losing information". In: *PhRvD* 64 (2001), 063001 (15 pages).

[98] D. Thornton et al. "A Population of Fast Radio Bursts at Cosmological Distances". In: *Science* 341 (July 2013), pp. 53–56. DOI: 10.1126/science.1236789. arXiv: 1307.1628 [astro-ph.HE].

[99] T. Totani. "Cosmological Fast Radio Bursts from Binary Neutron Star Mergers". In: *PASJ* 65, L12 (Oct. 2013), p. L12. DOI: 10.1093/pasj/65.5.L12. arXiv: 1307.4985 [astro-ph.HE].

[100] F. L. Vieyro et al. "A model for the repeating FRB 121102 in the AGN scenario". In: *AAP* 602, A64 (June 2017), A64. DOI: 10.1051/0004-6361/201730556. arXiv: 1704.08097 [astro-ph.HE].

[101] B. D. Wandelt, E. Hivon, and K. M. Górski. "Cosmic microwave background anisotropy power spectrum statistics for high precision cosmology". In: *PhRvD* 64.8, 083003 (Oct. 2001), p. 083003. DOI: 10.1103/PhysRevD.64.083003. eprint: astro-ph/0008111.

[102] W. Wang et al. "FRB 121102: a star quake-induced repeater?" In: *ArXiv e-prints* (Oct. 2017). arXiv: 1710.00541 [astro-ph.HE].

[103] N. Whitehorn et al. "Millimeter Transient Point Sources in the SPTpol 100 Square Degree Survey". In: *ApJ* 830, 143 (Oct. 2016), p. 143. DOI: 10.3847/0004-637X/830/2/143. arXiv: 1604.03507 [astro-ph.HE].

[104] S. Yamasaki, T. Totani, and K. Kiuchi. "Repeating and Non-repeating Fast Radio Bursts from Binary Neutron Star Mergers". In: *ArXiv e-prints* (Oct. 2017). arXiv: 1710.02302 [astro-ph.HE].

[105] D. Yvon and F. Mayet. "Mirage: A new iterative map-making code for CMB experiments". In: *AAP* 436 (June 2005), pp. 729–739. DOI: 10.1051/0004-6361:20035920. eprint: astro-ph/0401505.

[106] M. Zaldarriaga and U. Seljak. "All-sky analysis of polarization in the microwave background". In: *PhRvD* 55 (Feb. 1997), pp. 1830–1840. DOI: 10.1103/PhysRevD.55.1830. eprint: astro-ph/9609170.

[107] M. Zaldarriaga and U. Seljak. "All-sky analysis of polarization in the microwave background". In: *PhRvD* 55 (Feb. 1997), pp. 1830–1840. DOI: 10.1103/PhysRevD.55.1830. eprint: astro-ph/9609170.

[108] M. Zaldarriaga and U. Seljak. "CMBFAST for Spatially Closed Universes". In: *ApJS* 129 (Aug. 2000), pp. 431–434. DOI: 10.1086/313423. eprint: astro-ph/9911219.

[109] J. Zmuidzinas. "Thermal noise and correlations in photon detection". In: *ApOpt* 42 (Sept. 2003), pp. 4989–5008. DOI: 10.1364/AO.42.004989.