Low-level Data Analysis, Calibration, and Ground Modeling with COMAP

Jonas Gahr Sturtzel Lunde

Thesis submitted for the degree of
Master of Science in Astronomy

Institute of Theoretical Astrophysics
University of Oslo

June 15, 2021

Copyright © 2021, Jonas Gahr Sturtzel Lunde

This work, entitled “Low-level Data Analysis, Calibration, and Ground Modeling with COMAP” is distributed under the terms of the Public Library of Science Open Access License, a copy of which can be found at http://www.publiclibraryofscience.org.

Abstract

The CO Mapping Array Project (COMAP) is a line intensity mapping experiment currently in its Pathfinder phase, where it is targeting CO emissions at redshifts $z = 2.4 - 3.4$. A main goal of the Pathfinder phase is to build a solid analysis and modeling framework and achieve detection of CO at the current redshifts. Central to this effort is the proper understanding and handling of different signal systematics.

In this thesis, we perform a comprehensive analysis of COMAP signal systematics and suggest improved methods for two of the current filters in the low-level data analysis pipeline, as well as the calibration. We claim that the currently employed frequency filter, primarily meant to target gain fluctuations and temperature continuum sources, is inadequate at handling the latter. We propose a new frequency filter, which performs a joint maximum likelihood fit of both quantities. One of the most critical systematics in the COMAP data is ground pickup by the far sidelobes of the telescope, currently handled by the pointing template filter. We propose a new way of constructing pointing templates, using ground pickup maps created from the COMAP data itself. Using a destriper mapmaking model, we create examples of such ground maps, and perform a preliminary analysis of the viability of introducing such a data-based pointing template model. We complement this analysis with ground pickup maps from a simulated beam profile.

We find that our new frequency filter outperforms the current filter at removing temperature systematics in simulated data. The filter also significantly improves the removal of continuum foregrounds. With Jupiter as a case-study, it reduces the mean squared signal residual from 51σ to 1.5σ. Our destriper demonstrates the feasibility of producing data-driven ground maps, which can be employed in a more refined pointing template than the one currently employed, but also highlights several challenges which must be overcome. Our simulated ground pickup qualitatively corresponds well to maps produced by our destriper, but there are discrepancies, especially in the observational range of CO2, which warrants further analysis. We find the simulated beam profile to have a complicated frequency structure, which results in a ground profile that depends non-trivially on both frequency and pointing. Finally, our calibration analysis concludes that a calibration vane angle of 69° or below produces acceptable results, and we use this analysis to implement a new and more robust scheme for hot load measurements.

Acknowledgments

I would like to thank my great supervisors Hans Kristian Eriksen and Ingunn Wehus, for making this thesis possible, and for having their hearts and brains in the right place. I would also like to thank the brilliant Dr. Håvard Ihle. I do not dare imagine what this thesis would look like without your insights and supervision. I want to thank the rest of the COMAP collaboration, whom I very much look forward to meeting, once the world has stopped going under. It goes without saying that all my accomplishments rest on the shoulders of my friends, who are there to lift me up when I'm exhausted and challenge me when I'm too cocky. This year would not have been the same without the greatest cubicle the Stellar Cellar has ever seen. Looking at you guys, Julie, Jakob and Daniel. This of course also extends to the rest of the amazing master students at the institute, whose way too long lunch breaks I would not have survived without. I would like to thank my brilliant fellow researchers, Nils-Ole Stutzer and Jowita Borowska. It has been a true pleasure working with you on COMAP. I would also like to thank my family for feeding me, supporting me, and providing me with fruitful discussions on non-astrophysics related topics. Finally, to the rest of my countless friends out there who make my world a better place, thank you. And remember, if you dare venture into a thesis on telescope noise characteristics and survive, you can do anything. Rock on.

Jonas Gahr Sturtzel Lunde

Preamble

Modern observational astronomy is pushing the dark frontiers of space ever farther back, bringing light and knowledge to eras never before explored. COMAP, the CO Mapping Array Project, is one new such experiment pursuing the distant past of our universe. It is a Line Intensity Mapping (LIM) experiment, which itself is a relative newcomer as far as observational paradigms go. LIM revolves around mapping the 3D density fluctuations of the universe using a chosen line emitter, and it excels both at covering vasts areas of space in an effective manner, and reliably doing so at high redshifts. COMAP uses carbon monoxide (CO) as a line emitter, the second most abundant molecule in the universe, emitting photons at multiples of 115.27 GHz when transitioning between quantized rotational states.

The ultimate goal of COMAP is to trace the distribution of star-forming galaxies at the Epoch of Reionization, at redshifts $z = 4.5 - 8$. This is a future stage of the experiment, which we call COMAP-EoR. In the current stage of the experiment, called COMAP-Pathfinder, we target the Era of Galaxy Assembly at redshifts $z = 2.4 - 3.4$. We currently employ a single 10.4 m telescope, located at the Owens Valley Radio Observatory in California. The purpose of the pathfinder stage is to prove the feasibility of line intensity mapping with CO as a line emitter, and to lay a solid foundation of relevant modeling and analysis efforts. These objectives are more easily obtained in the pathfinder region, which is both closer, and better explored by other experiments.

COMAP started observing during the summer of 2019. Being in such an early stage, we do not as of date have confirmed detection of CO luminosities. The CO signal is incredibly weak, and we do not expect to have a clean detection until several more years of observational time. In addition to increased observational time, a successful detection will be conditioned upon the proper removal of noise and systematics in our data. In COMAP, this is performed by the `l2gen` program, operated by the COMAP group here at the University of Oslo, with the purpose of transforming raw telescope data into data free of systematics, and suitable for mapmaking. This is done through a series of *filters*, each aimed at dealing with specific systematics.

The purpose of this thesis is the improved understanding of a selection of these systematics, and concrete propositions of new and improved ways of dealing with these in the `l2gen` program. We propose improved methods for two of the filters currently employed, namely the *pointing template subtraction filter*, and the *polynomial filter*. The pointing filter deals with signal systematics correlated in the pointing coordinates of the telescope, such as the atmosphere and the telescope sidelobe ground pickup. The current filter fits and subtracts a linear function in azimuth and a $1/\sin(\text{el})$ profile in elevation. We introduce a more sophisticated model for dealing with the ground pickup, based upon data-driven maps of the ground, created using a destriper mapmaking model. The current polynomial filter is aimed at dealing with so-called gain fluctuations, as well as any continuum temperature fluctuations in the data. While it is great at the former, we explore its current inadequacies in dealing with the latter, where it fails to take into account the system temperature frequency profile. We propose a new and improved frequency filter, which is tailored to remove both types of fluctuations.

While these two new methods constitute the main contributions of this thesis, our work also involves a broader analysis of COMAP data systematics. We explore the properties and impact of each individual filter in the pipeline, and the characteristics of the systematics they target. The most important systematics for our work are temperature fluctuations, especially in the form of ground pickup, which we perform a comprehensive analysis of, both from a simulation and a data-driven perspective. We also perform an analysis of the telescope hot load calibration and implement an alternative scheme for calculating the calibration parameters more directly from the signal data.

The thesis is structured as follows. In Part I, we give the required background knowledge to understand this thesis in the context of both COMAP and cosmology in general. Chapter 1 gives a general overview of the field of modern cosmology, while chapter 2 gives a brief introduction to the field of line intensity mapping. Chapter 3 outlines the COMAP experiment as a whole, with a special focus on the telescope and data collection, which helps understand the sources of the different systematics in our signal. In chapter 4, we introduce the current COMAP data analysis effort, with `l2gen` at the center. We explain each of the filters in the program, and other relevant concepts to COMAP data processing. In Part II, we present the contributions of this thesis, together with relevant theory. Chapter 5 introduces mapmaking, first in general, and then tailored towards our efforts of producing ground pickup maps with a destriper mapmaking model. Chapter 6 presents the first of our main goals, namely a comprehensive analysis of the telescope ground pickup, and a proposition for an improved pointing template filter. In chapter 7 we analyze the COMAP calibration procedure and introduce a simple new scheme for performing it. In chapter 8 we analyze the noise properties of the data at different stages of the `l2gen` pipeline. Finally, chapter 9 presents the second of our main goals, namely the introduction of a new and improved frequency filter. We implement and extensively test this filter on both simulated and real data.

Contents

Abstract iii

Acknowledgments v

Preamble vii

List of Figures xi

I Background 3

1 Modern Cosmology 5

1.1 What is cosmology? 5
1.2 History of the universe 6
  1.2.1 The Big Bang and inflation 6
  1.2.2 The primordial soup 7
  1.2.3 Recombination and the cosmic microwave background radiation . 7
  1.2.4 A universe of light 8
1.3 The theoretical framework of astronomy 8
  1.3.1 General relativity in a nutshell 8
1.4 The ΛCDM model 9
  1.4.1 Theory 10
  1.4.2 Dark matter 10
  1.4.3 Dark energy 11
  1.4.4 Observations and alternatives 11
1.5 Observational astronomy 12
  1.5.1 Time travel and redshifts 12
  1.5.2 The state of modern observational astronomy 13

2 Line Intensity Mapping **17**
2.1 Why line intensity mapping? 17
2.2 Science goals 19
2.3 Experimental landscape 20

3 COMAP **23**
3.1 CO as a line emitter . . . 23
3.2 Telescope details . . . 24
3.2.1 Signal processing . . . 25
3.2.2 Calibration vane . . . 26
3.3 Fields . . . 26
3.4 Scanning strategies . . . 27
3.5 Noise sources . . . 29

4 The Oslo COMAP Data Analysis Pipeline **31**
4.1 Data files . . . 31
4.2 Data model . . . 32
4.3 System temperature calibration . . . 33
4.3.1 T_{sys} derivation . . . 34
4.3.2 Calibration summary . . . 35
4.4 l2gen . . . 36
4.4.1 Normalization . . . 37
4.4.2 Pointing template removal . . . 38
4.4.3 Polynomial filter . . . 39
4.4.4 PCA filter . . . 41
4.4.5 Masking . . . 43
4.4.6 Calibration . . . 43
4.4.7 Decimation . . . 44
4.4.8 Pipeline summary . . . 44

II New developments **47**

5 Mapmaking **49**
5.1 Setup . . . 49
5.2 The binning scheme . . . 50
5.2.1 Noise weighted mapmaking . . . 50
5.3 Destriping . . . 51
5.3.1 Maximum likelihood solution . . . 52
5.3.2 Some intuition . . . 52
5.4 Ground maps . . . 53
5.4.1 Dataset . . . 53
5.4.2 Ground map destriper . . . 55

6 Ground Modeling **57**
6.1 Context and idea . . . 57
6.1.1 Current ground model . . . 57
6.1.2 Improved ground model . . . 58
6.2 Beam convolution map . . . 58

6.2.1 Gradient maps 60
6.3 Destriped maps overview 64
6.4 Data driven pointing filter 66
6.5 Frequency profile analysis 70
6.6 Ground pickup summary and discussion 72

7 Calibration 75
7.1 Calibration properties 75
7.2 A new calibration technique 77

8 Noise Characterization 79
8.1 Noise theory 79
8.1.1 White noise 79
8.1.2 1/f noise 80
8.2 Data power spectrum and noise level 82
8.3 Polyfilter coefficients 83

9 A New Frequency Filter 87
9.1 Theoretical foundation 87
9.1.1 Maximum likelihood solution 89
9.1.2 Comparison to polynomial filter 91
9.2 Proof of concept - Real scans 92
9.3 Proof of concept - Simulated data 94
9.3.1 Data generation 94
9.3.2 Correlation 95
9.3.3 Temperature correlation results 96
9.3.4 Noise level and gain fluctuation priors 99
9.3.5 χ^2 goodness of fit 101
9.4 Continuum foregrounds and applications to continuum science 102

10 Conclusion and Future Work 107
10.1 Conclusion 107
10.2 Future work 108
10.3 The outlook of COMAP 109

Appendices 111

A Frequency Filter Destriper Setup 111

List of Figures

1.1 History of the universe 8
1.2 Planck 2015 CMB map 15

2.1 Line intensity mapping versus galaxy surveys 18
2.2 Redshift scope of line intensity mapping 18
2.3 Line intensity mapping in the Epoch of Reionization 19
2.4 LIM overview 21

3.1 CO molecule angular momentum 24
3.2 COMAP redshift and science goals 25
3.3 COMAP observational fields and Planck LFI data 27
3.4 COMAP observational fields in Earth coordinates 28
3.5 Telescope scanning strategies 29

4.1 Gain and system temperature frequency profiles 35
4.2 Gain and system temperature over time 36
4.3 Pipeline normalization in time 38
4.4 Pipeline normalization in frequency 38
4.5 Pipeline pointing template subtraction 39
4.6 Pipeline polynomial filter in time 40
4.7 Pipeline polynomial filter in frequency 40
4.8 Pipeline PCA filter in time 42
4.9 Pipeline PCA filter in frequency 42
4.10 Hot load measurement illustration 44
4.11 l2gen transfer function 45

5.1 Destriper baselines demonstration 53

6.1 Simulated telescope beam profile 59
6.2 Beam convolution ground temperature map 60
6.3 Ground convolution horizontal gradient for CO6 and CO7 61
6.4 Ground convolution horizontal gradient for CO2 62
6.5 Ground convolution total gradient for CO6 and CO7 63
6.6 Ground convolution total gradient for CO2 63

6.7 Ground convolution overplotted by destriper maps 64
6.8 Binned versus destriped CO7 map 65
6.9 Destriped ground map for upper-right CO7 region 67
6.10 TOD with linear azimuth fit and ground map amplitude for CO7 67
6.11 Destriped ground map for middle-left CO2 region 69
6.12 TOD with linear azimuth fit and ground map amplitude for CO2 69
6.13 Vertically slices of simulated beam powers 71
6.14 Difference of 34 GHz versus 26 GHz simulated beam profiles 72
6.15 Frequency derivative maps from simulation and destriper 73

7.1 Physical arrangement of telescope feeds 76
7.2 Feed power angle relation during calibration 76
7.3 Calibration TOD power response 78
7.4 New calibration method 78

8.1 White noise and 1/f noise demonstration 81
8.2 l2gen $1/f$ power spectrum fits 82
8.3 Chi2 white noise consistency test 83
8.4 $1/f$ best fits for c0 84
8.5 $1/f$ best fits for c1 85

9.1 Frequency filter demonstration in good weather 92
9.2 Frequency filter demonstration in bad weather 93
9.3 Frequency filter correlation in good weather 93
9.4 Frequency filter correlation in bad weather 94
9.5 Correlation distribution of simulated frequency filter data 97
9.6 Correlation distribution of simulated frequency filter data, split by system temperature 97
9.7 Simulated power spectra 98
9.8 Excess correlations for new and old frequency filter 98
9.9 Best fit δg and δT for different priors 100
9.10 Chi2 white noise test on different filters and priors 102
9.11 COMAP observational fields and Planck LFI data, zoomed in 103
9.12 Jupiter temperature remainder after filtering 104
9.13 Jupiter map decomposition from frequency filtering 104
9.14 Jupiter temperature frequency profile 105

10.1 COMAP power spectrum outlook 110

Part I

Background

Chapter 1

Modern Cosmology

In this chapter, we give a very general introduction to the history and modern understanding of our universe, as well as modern cosmology as a scientific field. The purpose of such an overview is mainly to provide some context as to why what we do as astronomers is both important and interesting.

1.1 What is cosmology?

Cosmology is a branch of astronomy mostly concerned with the "big picture" of our universe, studying how the very largest structures, as well as the universe as a whole, change in space and time. Questions like how our universe was born, how it came to look the way it does today, and if and how it will die. A few centuries ago, these questions were left as much in the hands of philosophers as scientists. Today, the field of modern cosmology has come far enough to put forward quantifiable and scientific answers. Luckily for us, the answers have raised a lot of new and exciting questions, many of them of equally abstract and absurd nature as the ones we started out with, concerning dark energy, dark matter, black holes, and the like.[1]

Cosmology has had a golden age since the birth of what we call *modern cosmology*, dated about a century ago. This field owes its huge success to the interplay between two scientific advances:

- The emergence of a revolutionary theoretical framework for describing the universe as a whole, starting with Einstein's theory of General Relativity [1], and built upon by numerous astrophysicists since then.
- The exponential increase in astronomical observations and available astronomical data, together with the capability of storing and analyzing these vasts amounts of data, through modern computing.

---

[1] Cosmology is still going through its teenage years where everything needs to be a bit dark and edgy.

The interplay between these two success stories has made modern cosmology enormously potent at both presenting theories for how the universe works, and extensively and quantitatively test and challenge these theoretical models through enormous amounts of observational data.

1.2 History of the universe

The modern understanding of our universe is that it had some sort of beginning and that it will have some sort of end. In other words, the universe is *dynamic*, its size and properties changing over time. This view, although widely accepted today, was far from common knowledge not that long ago. In fact, the assumption that the universe was of a much more stationary nature was originally so strong that when Einstein discovered that his equations could only explain a dynamic universe, not a static one, he altered his equations by adding what we know as the *cosmological constant*. It was later discovered [2] that the universe was not at all stationary, but was instead rapidly expanding, strongly indicating that it had some sort of beginning. Einstein called the alteration of his equations and thus failure to predict a dynamic universe "his biggest blunder". Today, we are very confident that the universe (at least as we know it), began sometime around 13.7 billion years ago [3], and that it since then has undergone a series of wildly different epochs, finally evolving into the universe we know and love today. We will in this section give a brief summary of some of the more important epochs in the life of the universe.

1.2.1 The Big Bang and inflation

The earliest event in the modern understanding of the universe is the Big Bang, which is just what we call the fact that the universe suddenly stopped being incredibly dense, and started expanding. We do not really have any idea of why or how this happened, as our current theories break down in this regime.[2] The Big Bang was followed by a period called inflation [4], which is both quite descriptive and the understatement of the century, where the universe abruptly expanded by a factor of at least 100,000,000,000,000,000,000,000,000, over a period of less than 0.00000000000000000000000000000000001 seconds [5, 6]. Prior to inflation, the universe was unimaginably hot and dense, consisting of quantum fluctuations in thermal equilibrium. These quantum fluctuations were suddenly blown up to span large stretches of space. While the processes driving this extremely rapid expansion are not well understood, inflation is still a vital part of modern cosmology, as the rapid expansion of these local quantum fluctuations help explain several observations that would otherwise be in direct conflict with established theory, such as the *Horizon problem* [7].

---

[2] The two most advanced theories in physics are General Relativity, which explains how massive universe-sized systems works, and Quantum Mechanics, which explain how really small stuff works. Physicists have been unable to unify the two theories, meaning that our understanding of the early universe, where both theories are needed, is very limited.

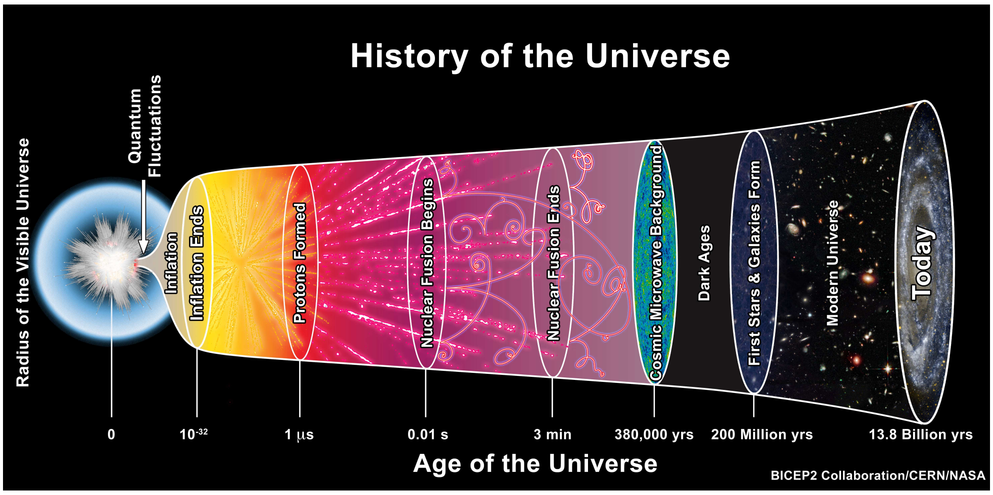
1.2.2 The primordial soup

After inflation dwindled out, the universe kept expanding, albeit at a much slower pace. It was still incredibly dense and hot, too hot for the everyday particles we see around us to exist, as elementary particles were coupled together in a dense plasma. During this time, the average distance light could travel before bumping into something and changing direction, the so-called *mean free path*, was short. The universe was, in other words, opaque. As the universe kept expanding, it eventually cooled down to the point where quarks could form protons and neutrons [8]. Not long after this, protons and neutrons combined to form some of the lighter atomic nuclei, like helium and deuterium in a process called Big Bang Nucleosynthesis. The abundance of these elements in the universe today is believed to stem from this process, only minutes after the Big Bang. The predicted abundances of these elements also match observed values to high accuracy, making it one of the strongest pieces of observational evidence we have in support of the Big Bang model [9].

In this early stage of the universe, and for some time still, the pressure in the universe was far too large for gravity to pull ordinary matter together to form any sort of structures, and the universe was very uniform. It is however theorized that much of the matter in our universe is so-called *dark* matter, which experiences little or no pressure. Dark matter is believed to have collapsed into galaxy-like structures called dark matter halos in the early stages of the universe [10]. These become important for structure formation later in the universe.

1.2.3 Recombination and the cosmic microwave background radiation

The early universe was in thermal equilibrium of uniform temperature and pressure. As the universe expanded, this pressure and temperature fell gradually, until the temperature more or less simultaneously across the universe fell below the threshold where electrons bind to protons to form hydrogen. This process is known as *recombination* [8], and happened some 380,000 years after The Big Bang. The universe was now suddenly transparent to photons, which had up until now been bouncing between the charged electrons through *Thomson scattering.* Hydrogen is neutral and therefore does not partake in the sport of Thomson scattering. The photons which had been bouncing around the plasma soup since the dawn of time suddenly found themselves traveling in lonely straight lines, without bouncing into anything at all. Most of them have today yet to meet a single other particle, over 13 billion years later. Some of them are lucky enough to hit Earth (or even better, a space telescope), where they are given an especially warm welcome. These photons are known to us today as the Cosmic Microwave Background (CMB) radiation, and because they have traveled mostly unobstructed through space since shortly after the Big Bang, they give us unique insight into the early workings of the universe. I take some relief in knowing that, even though they are among the loneliest particles in the universe, they are today thoroughly appreciated, having more or less revolutionized modern cosmology, and enabled several Nobel Prizes in physics.


History of the Universe
Radius of the Visible Universe
Quantum Fluctuations
Inflation
Inflation Ends
Protons Formed
Nuclear Fusion Begins
Nuclear Fusion Ends
Cosmic Microwave Background
Dark Ages
First Stars & Galaxies Form
Modern Universe
Today
0
10^-32
1 μs
0.01 s
3 min
380,000 yrs
200 Million yrs
13.8 Billion yrs
Age of the Universe

BICEP2 Collaboration/CERN/NASA

Figure 1.1: A simple outline of some of the more important events in the history of our universe. Figure credit: [14]

1.2.4 A universe of light

After recombination and the release of the CMB radiation, the universe went into a period of a few hundred million years known as the dark ages, named so due to the lack of light sources in the universe. The pressure in the universe had now fallen to a point where gravity became able to attract ordinary matter (which was mainly hydrogen and helium) into star-forming gas clouds. A few million years after the Big Bang, these clouds formed the first stars in our universe, known as Population III stars [11]. These stars usually formed in the gravitational wells of the dark matter halos created in the earlier universe and became part of the first galaxies.

The first stars and galaxies brought with them a period known as *reionization* [12], lasting from around 200 million to 1 billion years after the Big Bang. At this point, the universe mostly consisted of neutral hydrogen, formed when electrons and protons coupled during recombination. During the long period of reionization, the universe would again become ionized, as radiation from early galaxies ripped electrons from their atoms [13]. The end of the reionization era gives rise to what we call the modern universe, which bears much resemblance to the universe we see today, around 13 billion years later. Formation of stars and galaxies play an important role during this era.

1.3 The theoretical framework of astronomy

1.3.1 General relativity in a nutshell

The by far most important theoretical building block of cosmology is Einstein's theory of general relativity. General relativity is most importantly a theory about gravity, and how stuff in the universe gravitationally moves and interacts. It can be thought of as a natural extension of Newton's laws of motion, which, unlike general relativity, completely breaks down if looking at something with high density or velocity. Very simply, the central parts of general relativity can be summarized in two concepts.

Concept one: Space and time are actually part of a unified, dynamic spacetime, which can curve and bend. In general relativity, the three spatial dimensions and the single time dimension are intertwined into a single concept called spacetime. The most important property of this spacetime is that it can curve. The curvature of spacetime causes the force we know as gravity. In general relativity, gravity is actually not a force at all. Particles under the effect of gravity are moving along straight lines known as *geodesics*. They appear to be under influence of some force simply because spacetime itself bends, shaping the path of the particle. This does however not explain where this curvature comes from, which brings us to the second concept.

Concept two: The curvature of spacetime tells stuff how to move, and stuff tells spacetime how to curve. This interaction between spacetime and the content of the universe is described by Einstein's field equations,[3]

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}, \tag{1.1}$$

or, more intuitively,

$$[\text{spacetime}] = [\text{some constants}] \cdot [\text{content of the universe}]. \tag{1.2}$$

On the left-hand side of the equation, we have the Einstein tensor, $G_{\mu\nu}$, which represents the shape and behavior of the spacetime. The energy-momentum tensor on the right, $T_{\mu\nu}$, describes the content of the universe. The take-away message from this equation is that the spacetime and content of the universe are neatly intertwined, and to understand either we need to look at both.

1.4 The ΛCDM model

The ΛCDM model is sort of the poster child of modern cosmology, holding most of the ingredients of what has become the accepted version of components and events in the universe. Most efforts in cosmology today either directly or indirectly involve testing, extending, or, if you are bold, disproving the ΛCDM model.

The ΛCDM model aims to explain a series of disagreements between theory and observations which have emerged in the last few decades [8] by introducing two new ingredients to the universe. Without ΛCDM, the story goes something like this: We have a theory for how the universe works (mainly general relativity); we put into it the ingredients we believe our universe to consist of; and out pops a universe. Except it does not look anything like our universe. Galaxies should not exist at all, the distribution of matter is way off, and the expansion of the universe is all wrong. So what happened? Either our theory is wrong, and Einstein's general relativity is off, or we put in the wrong ingredients and were wrong about what our universe consists of. ΛCDM postulates the latter.

---

[3] If this looks like a single equation to you, and you wonder why I said "equations", this is because $G_{\mu\nu}$ and $T_{\mu\nu}$ are mathematical objects called tensors, and this single equation does, in fact, hold 16 equations. The tensor notation makes it much prettier to look at.

1.4.1 Theory

The ΛCDM model proposes that, in addition to the conventional components of the universe (baryonic matter, photons, neutrinos, etc.), there exist two additional components, known as *dark matter* and *dark energy*. The two are commonly confused but are in fact very different phenomena. We could in theory be right about the existence of one of them, but not the other. Both phenomena are introduced to explain a number of observations that would otherwise be at odds with established theory, and they are both therefore best understood in the context of why we need them to exists.

1.4.2 Dark matter

It is believed that about 85% of the matter in the universe is dark matter, which differs from "ordinary" matter in that it does not interact electromagnetically with other matter[4]. The lack of electromagnetic interaction means that dark matter does not absorb or emit electromagnetic radiation, such as light (thereby the name), and is therefore difficult to observe directly. Electromagnetism is also responsible for a lot of other ways in which ordinary matter can interact, making dark matter completely collisionless. It does, however, have mass, which is exactly the reason we believe it exists, as it interacts gravitationally with other objects. If the Sun was made of dark matter, Earth would keep its ordinary orbit, as if nothing was amiss, however, the Sun would be completely invisible. This hints at why need dark matter to exists in the first place: We observe that the amount of matter we can see in the universe is not sufficient to explain the observed behavior of our universe, nor how it came to be as it is. Galaxy formation in the early universe would not be possible, galaxies would have different structure and rotation properties than we observe them to have, and light coming from distant galaxies would not twist and bend as much due to gravitational lensing as we observe it to do.

For all of these reasons, the astrophysical community is rather confident that some sort of dark matter must exist. Fascinatingly enough, even though we are quickly narrowing down on the properties dark matter must have in order to explain our universe, we have rather little clue about what exactly it *is*. Potential candidates include sterile neutrinos, Weakly Interacting Massive Particles (WIMPs), or even tiny black holes, as well as a dozen others.

---

[4] Many theories also suggest that it *does* interact electromagnetically, but very weakly. In that case, everything I am about to say is only true to a large extent.

1.4.3 Dark energy

It was discovered in 1929 by Hubble [2] that the universe is not at all static, but instead expands. This in itself could be explained by the fact that the universe was once much smaller, inflation rapidly expanded it, and when inflation ended, everything was moving away from everything else at some speed. The expansion would then gradually slow down, as every object in the universe gravitationally pulled on everything else. However, it has in later years been discovered that the expansion is not at all slowing down. In fact, it is instead speeding up [15, 16]. This acceleration would require some sort of force driving it. This unknown force is what we call dark energy.

Since the discovery of the acceleration of the universe, several other observations also strongly support the existence of dark energy. As with dark matter, we do not really know what dark energy is, but we are today rather confident that it exists. We also know pretty well what it does. Dark energy acts sort of like a negative gravity, or a negative pressure, evenly spread throughout the universe. It is believed to be a property of vacuum itself. An important consequence of this is that, as the universe expands, there will be more dark energy, as there is more empty space between everything. The more dark energy, the faster the expansion will happen, and so on. In other words, any two objects that are not gravitationally bound to each other (like two galaxies far away from each other), will drift further and further away, faster and faster.

As an extra treat, it turns out that the modification you need to make to Einstein's equations to explain dark energy is to add a *cosmological constant*. You remember that thing Einstein added to his equations in order to explain a static universe, and then shamefully removing again? Turns out, he might have been right about that, just for all the wrong reasons. In the ΛCDM model, the cosmological constant is now back in his equations, as the term for dark energy.

1.4.4 Observations and alternatives

Since its gradual introduction throughout the 80s and 90s, the ΛCDM model has been enormously successful and goes almost hand-in-hand with the success of modern cosmology itself. It has predicted numerous observables with astonishing accuracy [17], and so far withstood all attempts at overthrowing the model.[5]

As hinted in the beginning of this section, one could go another way in explaining the disagreements between theory and observations. Instead of introducing dark matter and dark energy into the universe, some theories, like the theory of *modified Newtonian dynamics* [18] try to explain observed phenomena by instead modifying or completely replacing general relativity. These theories have generally been much less successful than ΛCDM at explaining the vast set of observational phenomena we have available today.

---

[5] Although, as most models in astronomy, it is of course not without its smaller problems, and is not a complete theory of the universe. It is, however, the closest we are to such a thing so far.

1.5 Observational astronomy

We now move from the realm of theory and into the field of observational astronomy, which is the field this thesis falls within. It is worth taking a step back and reflecting on the premise of observational astronomy, which differs somewhat from a lot of other scientific fields. First of all, an astronomer cannot really conduct experiments the way a physicist can. The universe is our laboratory, and we are left observing whatever experiments the universe feels like doing. Because of this, astronomers have to be clever in finding observable phenomena which can test their theories. Secondly, the objects we are interested in are, quite inconveniently, located very far away from Earth, often halfway across the universe. This poses numerous problems, both in actually finding phenomena that can test our theories, and then actually observing them properly. In other words, what you (can) see is what you get.

1.5.1 Time travel and redshifts

While that might not sound too promising, there is a silver lining. The fact that much of the universe is so far away means that light from distant objects takes millions or billions of years to reach Earth. When looking at far-away objects, we are looking into the distant past. In some cases, like with the CMB, almost all the way back to The Big Bang. This is so important for our understanding of the universe because of how dynamic the universe is. The universe of today is vastly different than it was long ago, and having access to the state of the universe over different timescales is a huge help in testing our theories.

Astronomers also got another ace up their sleeve, which helps us gauge how far away something is, and by extension, how long ago it happened. Because of the expansion of the universe (believed to be caused by dark energy), everything in the universe is moving away from everything else, and the velocity is proportional to the distance. If something is twice as far away from us as something else, it will move away from us twice as fast. This is known as Hubble's law: $v = H_0 D$. The law serves as a useful approximation, but has two limitations. Galaxies usually have some random velocity not related to the expansion of the universe, known as a peculiar velocity, which dominates the velocity profile of closer galaxies. Additionally, the expansion of the universe is believed to accelerate over time, and the Hubble constant H_0 is only an approximation of the time-dependent Hubble parameter $H(t)$. Hubble's law is therefore only accurate up to a certain distance. Within these limitations, Hubble's law allows us to easily approximate the distance to most of the galaxies we can observe in our universe.

Calculating how fast something is moving away or towards us is done using something called redshift. When an observer moves relative to a light source, that light will have its wavelength either contracted or expanded when observed by the observer. The redshift z of a photon is defined as

$$z = \frac{\lambda^{\text{obs}} - \lambda^{\text{source}}}{\lambda^{\text{source}}}, \tag{1.3}$$

where λ^{source} and λ^{obs} are the wavelengths of the photon as it is sent from the source, and observed by the observer, respectively.

In our case, every celestial object (except very close objects) is moving away from us, and we simply talk about how much longer the wavelengths of the light becomes, e.g. redshift.[6] We know a lot about what kind of colors and light spectra different types of celestial objects should have, by studying close-by objects. If we compare this to the spectra we observe distant objects to have, we can cross-match them, and calculate how much the light has been redshifted. This then allows us to calculate how fast the object is moving, how far away it is, and then, finally, how long ago the light was emitted.

As astronomers do not directly measure how long ago some observed event took place, redshift is often used as a measure of time in astronomy, as it is the quantity directly observed. We use the variable z for redshift, and the redshift today is $z = 0$, while the emission of the CMB is estimated to have happened at $z = 1100$. Ahead in the next chapter, figure 2.2 shows the redshifts at certain important events in the universe.

1.5.2 The state of modern observational astronomy

As with any scientific discipline, the progression of astronomy is dependent on both the development of new theories, as well as experimental data which can confirm or challenge these theories. Just as there have been great strides in the theoretical fields of astronomy, observational astronomy has also been entirely revolutionized in the past century. The Earth is now filled with telescopes, satellites, and detectors collecting data in amounts that were previously completely unfeasible. These data are analyzed and processed with the help of modern computing, often with huge computational requirements.

Until recently, photons were the only method of observing the universe (and until a bit less recently, only photons on the visible light spectrum). Photons are still by far the dominant source of information in astronomy, but they are today accompanied by a few other sources, such as gravitational waves, neutrinos, and protons (in the form of cosmic rays). The emergence of these new fields has been a huge success for astronomy, both in probing phenomena for which photons are inadequate, and for cross-referencing or improving existing results. While there exist a magnitude of different observational fields and paradigms in astronomy, we will give a brief summary of two of the most important ones; CMB experiments, and star and galaxy surveys. In the next chapter, we go into more detail on line intensity mapping, the paradigm under which COMAP falls.

---

[6] In our case, the redshift is not actually caused by the relative velocity of the object itself, but rather the expansion of space, which stretches the light as it travels towards us. It is important to remember that the objects are not moving away from us in a conventional manner, but rather because the space itself between us and the object expands. However, it results in the same redshift as it would if the object was moving away from us in a "conventional" manner.

CMB experiments

The cosmic microwave background radiation was emitted only 380,000 years after The Big Bang, and is one of our most important ways of understanding events in the early universe. It was released from everywhere in the universe at once, and can therefore be observed all the time by looking in any (relatively unobstructed) direction in space. The CMB had a temperature of approximately 3000 K at emission,[7] observed today at a redshifted temperature of 2.725 K. This places it in the microwave range, and its frequency distribution peaks at about 160 GHz. Several satellites have been dedicated to observing the features of the CMB, like COBE (1989) [19], WMAP (2001) [20], and Planck (2009) [21], each giving us access to a more accurate and sophisticated picture of the early universe. The most important analysis of the CMB radiation is the study of how different regions deviate from this temperature of 2.725 K, so-called anisotropies, as this tells us a lot about astrophysical properties in the early universe. The deviation from this average temperature, as observed by the Planck satellite, is shown in figure 1.2.

Star and galaxy surveys

Star and galaxy surveys are some of the most widely applied observational paradigms in astronomy. They work by targeting, observing, and cataloging a large number of stars or galaxies. Surveys can be classified as either photometric or spectroscopic. Spectroscopic surveys aim at resolving the emitted frequency spectrum of the observed object, which is useful for, e.g., resolving star abundances through absorption lines in the spectrum. Photometric surveys trade the frequency resolution of spectroscopic surveys for higher brightness sensitivity (or shorter integration time) by measuring only the aggregate emission over a range of frequencies. Such surveys might instead focus on time-correlated features, such as the periodic luminosity profile of Cepheids or RR Lyrae, which was used to produce the cosmic distance ladder, and measure a value of H_0 [15]. Surveys can also be classified by the frequency range they target, as there exist surveys targeting everything from radio waves to gamma rays. Some of the largest surveys as of date are the Sloan Digital Sky Survey [22], a photometric survey targeting galaxies and quasars, the Dark Energy Survey [23], a recent optical photometric survey, and the Gaia mission, a satellite survey mainly targeting stars in the Milky Way [24], just to mention a few.

---

[7] Most photon sources, like the CMB, emit photons at a range of different frequencies. Astronomers therefore often find it more convenient to talk about the temperature of a photon source instead of the frequency of the photons they emit. If the photon source is what we call a perfect *black body* (which the CMB is), its temperature alone is enough to know everything about the frequency of emitted photons, which follow a distribution known as a *Planck spectrum*.

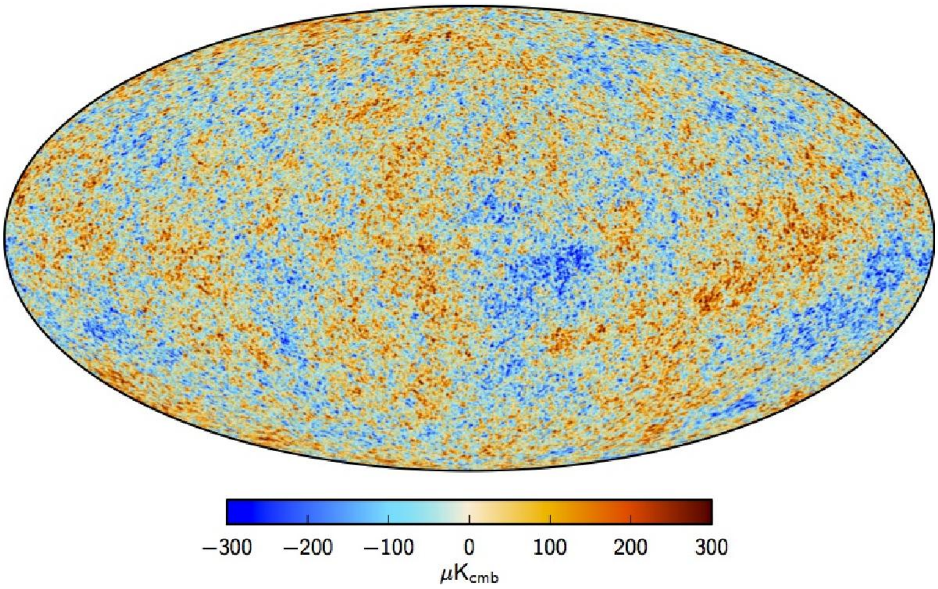

−300
−200
−100
0
100
200
300
μK_cmb

Figure 1.2: Temperature fluctuations in the CMB from the Planck 2015 results [21]. Figure credit: ESA and the Planck Collaboration

Chapter 2

Line Intensity Mapping

Line intensity mapping (LIM) is a relatively new but rapidly growing field in observational astronomy, aimed at mapping large volumes of space by measuring the integrated emission of spectral lines, instead of trying to resolve individual galaxies or stars. This is done using a chosen emission line—a specific photon frequency arising from some phenomena. The most common such lines are Lyα, emitted when the hydrogen electron falls from the $n = 2$ to the $n = 1$ orbital, and the 21cm line, emitted when the hydrogen electron flips from a parallel to anti-parallel state with the nucleus proton. Other common examples are the [CII], OII, and CO lines [25]. By knowing the specific frequency of the chosen emission line, the distance to the observed target can also be calculated through redshift, which allows the mapping of a three-dimensional volume.

2.1 Why line intensity mapping?

As a newcomer to the stage of observational astronomy with few detections to date, line intensity mapping needs convincing strengths to prove itself worthwhile the decades of research and testing it takes to mature an entirely new observational paradigm. As of today, the two largest observational paradigms in cosmology have been CMB and galaxy surveys. In terms of observational goals, LIM has perhaps most in common with galaxy surveys, as the main goal of LIM has long been fulfilled by galaxy surveys, namely, to map large 3D volumes of space at different redshifts [26]. Discussing the strengths of LIM as an observational paradigm is easiest done by comparing it to galaxy surveys. In this regard, LIM has a handful of obvious advantages.

At low redshifts, one of the main strengths of LIM is its ability to probe huge patches of the sky in a fast and cost-effective fashion [27]. Galaxy surveys are limited by their ability to catalog a large and representative set of galaxies from the relevant patch of sky. LIM can also combine emissions from other sources than galaxies and starts, such as interstellar or intergalactic medium.

At higher redshifts, the fact that galaxy surveys need to distinguish individual galaxies is a glaring weakness that LIM does not possess. There will be a redshift threshold where galaxy surveys are virtually incapable of gathering data, and even at moderate redshifts they will have a clear bias towards brighter galaxies [28]. LIM possesses no such weakness, as it merely integrates the luminosity contribution of all galaxies in the observational range, even if they can't be individually distinguished. Figure 2.1 illustrates this difference. LIM can therefore give new insight into the relatively unexplored region between the regions already well explored by the CMB (z = 1100), and galaxy surveys (z = 0-2), as illustrated by figure 2.2. One of the main scientific goals of line intensity mapping is to probe into the Epoch of Reionization (z = 5-27) [25], something galaxy surveys struggle to do.

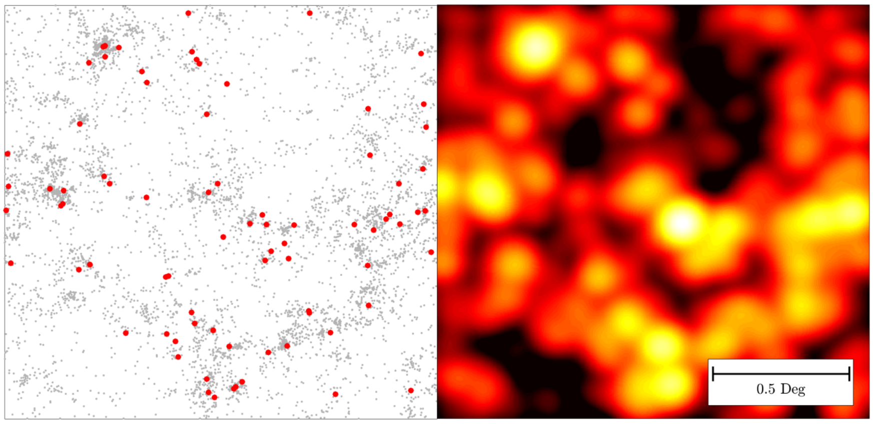


Figure 2.1: Figure illustrating the observational advantage of LIM as opposed to galaxy surveys. Left: Simulated galaxy assembly, with sources bright enough to be detected by low-threshold galaxy survey marked in red. Right: Intensity map from corresponding LIM survey. Figure credit: Patrick Breysse

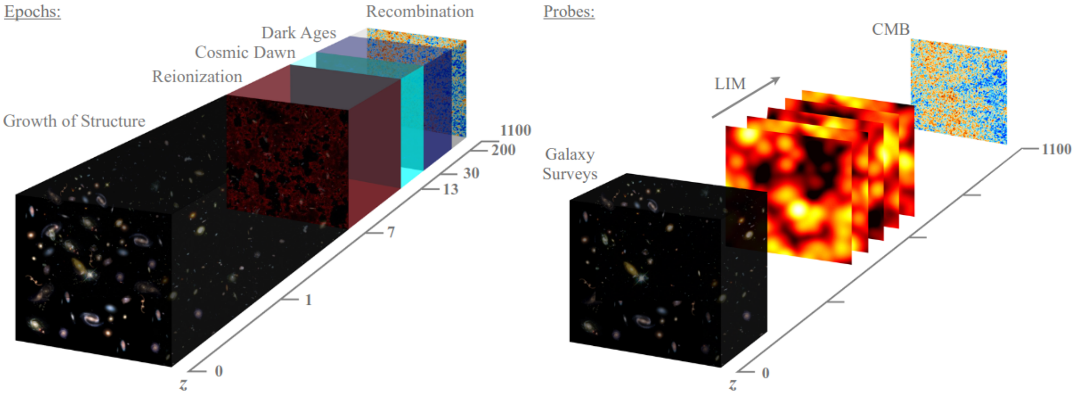


Figure 2.2: LIM aims to cover a large period of time in between the regions already well explored by galaxy surveys and the CMB. Figure credit: [27].

2.2 Science goals

Epoch of Reionization

Line intensity mapping looks set to drastically improve our understanding of the Epoch of Reionization (EoR) (z=5-27), a very important period in the formation of our universe, which so far has been challenging to probe. We expect to see a wide variety of lines suited for LIM originating in this epoch, as illustrated by figure 2.3. The ability to correlate different types of emission lines will be a considerable strength in future efforts to better understand the EoR.

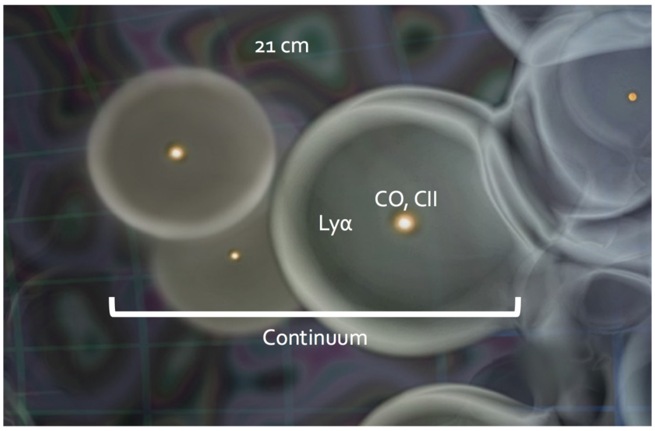


Figure 2.3: The 21cm line traces neutral intergalactic medium, while Lyα traces the ionized bubbles around galaxies. The CO and CII lines originate from the galaxies themselves. Figure credit: Patrick Breysse

Galaxy evolution and star formation

Many questions regarding galaxy evolution and star formation can be better answered with the help of LIM. These include questions about star formation rates, properties of active galactic nuclei, interstellar medium, and intergalactic medium. Non-LIM surveys targeting these phenomena are often limited in the redshift range they can target without bias, or the area they can span at lower redshift without suffering from cosmic variance [29]. LIM can also offer powerful cross-analysis data with other types of surveys targeting the same phenomena.

Large scale structure and dark energy

Many astrophysical models can be constrained through large-scale low redshift surveys, like ΛCDM parameters, dark energy models, and primordial non-Gaussianities. Although wide-field galaxy surveys will fill this need to an adequate level, LIM look poised to be a much cheaper and faster alternative. LIM can also probe baryonic acoustic oscillations over large continuous redshift ranges [27].

2.3 Experimental landscape

The number of line intensity mapping experiments is quickly growing, and we will here present some of the current experimental efforts in the field. An introduction to COMAP as a LIM experiment will be given in the next chapter. Figure 2.4 gives an overview of a large number of current and potential line intensity mapping efforts, illustrating the large span of both angular scales and redshift ranges which LIM is able to target.

TIME

The Tompgraphic Ionized-carbon Mapping Experiment (TIME) [30] is an observational array targeting the emission line of singly ionized carbon (the [CII] line) at redshifts of $z = 5.3 - 8.5$, at the Epoch of Reionization. Being a LIM experiment, TIME will among other things be sensitive to dwarf galaxies, which may have played a central role in producing ionizing UV photons. The survey will also target rotational CO emissions at redshifts of $z = 0 - 2$, which can be used to constrain the density of star-forming molecular gasses.

COPSS

The COPSS I and II surveys [31, 32] target the same CO(1-0) transition as COMAP does, at redshifts of $z = 2.3 - 3.3$, using an array of 8 3.5 m antennas. The experiment has a non-zero detection of CO of $3.1 \times 10^3 \pm 1.3\,\mu\mathrm{K}(\mathrm{h}^{-1}\mathrm{Mpc})^3$ at $z = 3$, which serves as a ground for comparison for future COMAP detections.

HIRAX

The Hydrogen Intensity and Real-time Analysis eXperiment (HIRAX) [33] is a planned radio telescope array, that will measure the Baryon Acoustic Oscillation (BAO) from the 21 cm line at redshifts $z = 0.8 - 2.5$. The array will consist of 1024 6 m telescopes at the SKA site in South Africa. The experiment aims to use accurate BAO measurements over a wide redshift range to constrain the dark energy equation of state.

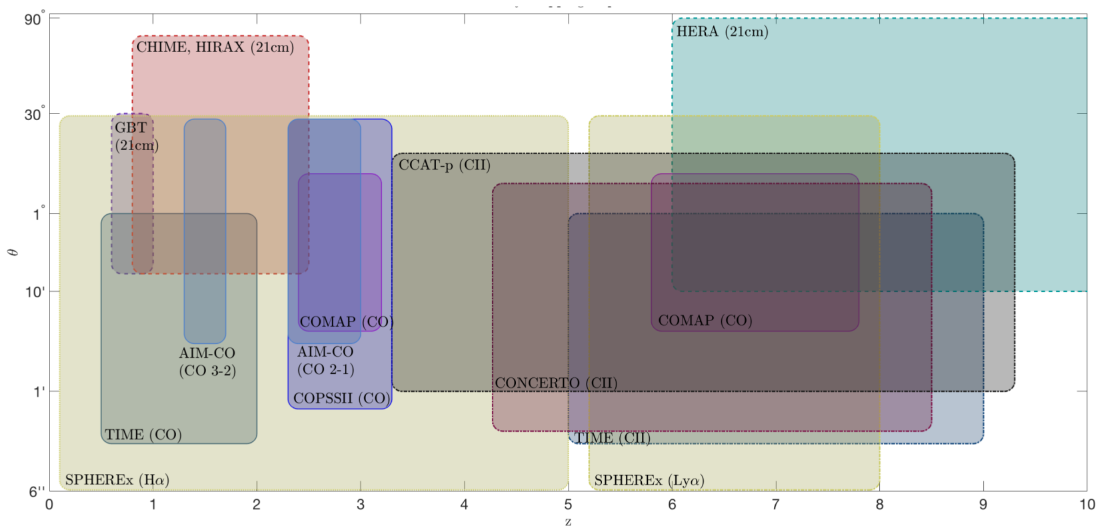


Figure 2.4: Overview of current and proposed LIM experiments by the redshift over which they operate at the x-axis, and the upper and lower angular resolution of the experiments on the y-axis. Figure credit: [25].

Chapter 3

COMAP

The CO Mapping Array project (COMAP) is a line intensity mapping project, using CO as a line emitter, ultimately aimed at tracing the distribution of star-forming galaxies at the Epoch of Reionization (EoR) (z=6-8) [34]. The project is currently at a pathfinder state, tracing CO at the Epoch of Galaxy Assembly (z=2.4-3.4), meant to provide both a proof-of-concept and valuable experience, for moving to the EoR stage of the project.

The project is a collaboration between a series of universities and institutions [35], including The California Institute of Technology, The University of Oslo, The University of Manchester, Stanford University, The University of Maryland, Princeton University, The University of Toronto, and UC Berkeley. The project currently employs a single telescope, operated by the Owens Valley Radio Observatory (OVRO) of the California Institute of Technology.

3.1 CO as a line emitter

As most of the previous section went into the aims and advantages of LIM in general, it is worth taking a look into how CO distinguishes itself from other lines. Carbon monoxide is the second most abundant molecule in the galaxy, after H_2. However, unlike H_2, which is symmetric, CO has an electric dipole moment, giving it a quantized ladder of allowed rotational energy levels [36], determined by the total angular momentum quantum number J, as

$$E_{\mathrm{rot}} = \frac{J(J+1)\hbar^2}{2I}, \quad J = 0, 1, 2, ... \tag{3.1}$$

where $I = \frac{m_C m_O}{m_C + m_O} r^2$ is the moment of inertia (see figure 3.1). The energy emitted from allowed transitions of $\Delta J \pm 1$ then becomes

$$\Delta E_{\mathrm{rot}} = \frac{\hbar^2 J}{I}, \tag{3.2}$$

which in the case of CO translates to line frequencies of

$$\nu_{J\to J-1} = \frac{\hbar J}{2\pi I} \approx 115\,\mathrm{GHz}\cdot J. \tag{3.3}$$

The frequency of the emitted lines form a ladder-like structure, as multiples of the lowest energy transition. This is an important argument for the use of CO. With a wide enough observed frequency range, cross-correlation between these ladder steps gives a powerful constraint on the CO brightness. This is an especially important property for line intensity mapping because there often is a certain level of degeneracy between CO and other so-called interloper lines. These are emission lines from other sources, happening at redshifts such that they overlap with the CO frequencies we are looking for. Interloper lines can be difficult to distinguish from CO signal, but an efficient way of breaking the degeneracy is cross-correlation between different frequencies on the CO emission ladder.

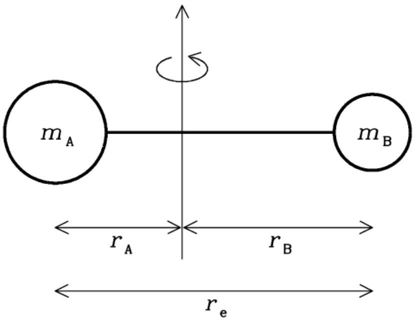


Figure 3.1: CO is a molecule consisting of a carbon and an oxygen atom. Due to being a molecule, it has an angular momentum, which is quantized into a series of allowed energies. Figure credit: [37].

This effect also works the other way around. CO lines from any narrow frequency band will in reality contain contributions from several vastly different redshifts, with different transitional lines on the ladder. This does of course introduce degeneracy which must be broken by other means, but in theory, enables the probing of several entirely different redshift ranges within a single narrow frequency band.

COMAP mainly considers the CO(1-0) transition at 115 GHz, with secondary considerations for the CO(2-1) transition at 230 GHz. As COMAP Pathfinder is observing in the 26-34 GHz range, it will pick up CO(1-0) signal at redshift $z = 2.4 - 3.4$ and C(2-1) at $z = 5.8 - 7.8$, as demonstrated by figure 3.2.

3.2 Telescope details

The COMAP collaboration currently employs a single retrofitted telescope, originally from the millimeter-wave array, and later used as a part of the CARMA project. It is a 10.4m Leighton telescope [38] situated in the Owens Valley Radio Observatory (OVRO)

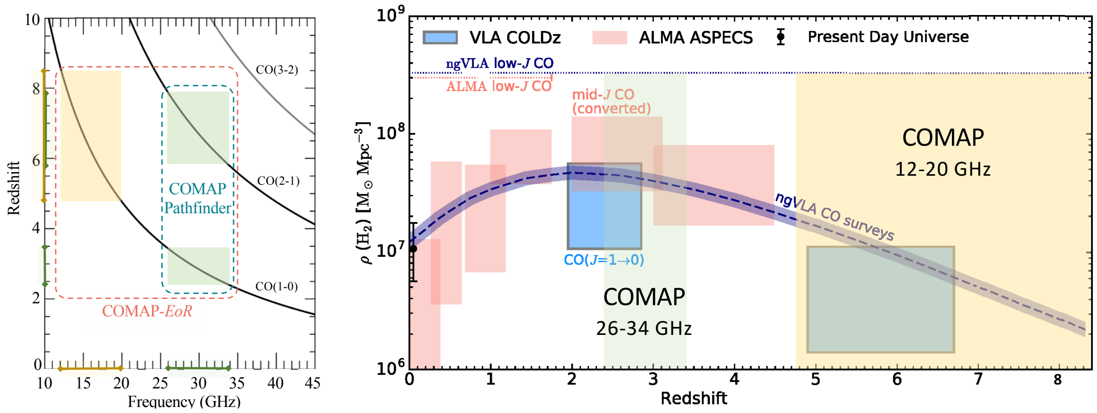


Figure 3.2: Left: Redshift-frequency ranges probed by the current COMAP Pathfinder experiment, and the future COMAP-EoR, with the three first rotational CO line emissions shown as solid black lines. Right: The H_2 density constraint as function of redshift for different surveys. Redshift ranges which COMAP has the potential to constrain are shown as shaded regions. Figure credit: Kieran Cleary.

in California, and operated by the California Institute of Technology. Later phases of the collaboration aim to add more telescopes of the same type to the array.

The telescope dish employs 84 honeycomb aluminum panels. Being a Cassegrain reflector telescope, the main dish which reflects signal onto the 1.1 m secondary reflector [39], mounted over the main dish upon four support legs. The secondary reflector again reflects the signal onto the 20 pixels, or "feeds", each mounted in a feed horn at the center of the main dish. The feeds, together with the rest of the signal processing equipment, is placed in a cryostat, a section of the telescope cooled down to $10-20$ K, for minimal thermal noise. The 20th feed is a "blind" feed, added to distinguish between internal and external signals in the receivers, and this feed is excluded from all our analysis.

3.2.1 Signal processing

The signal is passed through a series of modules connected to each feed. First is polarization, where, as of writing, 15 of the 19 feeds have two circular polarizers, 2 feeds have a single such polarizer, and 2 feeds have none. The polarizers separate the data into one or two orthogonal circular polarized signals. A discovered disadvantage of the double polarization has been the presence of standing waves in between the polarizers.

Each feed contains a total of five low-noise amplifiers (LNA) at different stages of the signal processing. LNAs serve the purpose of amplifying the power of a signal, with as small a degrading of the signal-to-noise ratio (SNR) as possible. Combined, the LNAs of each feed boost the signal power by over 140 dB [40]. Part of the signal processing is also two down-converter modules (DCMs), as well as an in-phase quadrature (IQ) mixer. The DCMs and IQ-mixer convert the signal from the observed signal range of $26-34$ GHz into four sidebands, which we refer to as A:LBS, A:USB, B:LSB, and B:USB ("LSB" and

"USB" referring to lower and upper sideband), with downconverted frequency ranges of $2 - 4\,\text{GHz}$, $4 - 6\,\text{GHz}$, $6 - 8\,\text{GHz}$ and $8 - 10\,\text{GHz}$, respectively. These frequency ranges correspond to $2\,\text{GHz}$ sections of the observed signal range of $26 - 34\,\text{GHz}$, in the same increasing order. Finally, the signal passes through a spectrometer, which splits each sideband into 1024 frequency channels, of frequency resolution $\approx 1.95\,\text{MHz}$. An important point to keep in mind is that all these signal processing modules are individual to each feed, and so are any individual biases or systematics introduced by them.

3.2.2 Calibration vane

The telescope has a calibration vane mounted next to the feed array. This vane can be rotated to fill the field of view of the entire feed array. The temperature of the vane is accurately measured with a temperature sensor and is used together with the observed power from the receivers to perform a gain calibration, which we will come back to in detail later.

3.3 Fields

The current COMAP-Pathfinder focuses on three main observational fields, which we have named CO2, CO6, and CO7. They have been chosen because of their high elevation from the Galactic disk, low amount of astrophysical foregrounds, and overlap with the HETDEX survey, which allows for useful cross-correlation. The location of the fields compared to the Galactic plane is shown in figure 3.3, against a $30\,\text{GHz}$ emission map from the Planck survey. Figure 3.4 shows the fields as we observe them drifting across the sky, from the perspective of the telescope.

The size of the fields in parallel direction (along the line of sight) can be approximated for small observational frequency bin $\delta\nu^{\text{obs}}$ as [41]

$$\delta l_{||} \approx \frac{c(z+1)^2}{H(z)} \frac{\delta\nu^{\text{obs}}}{\nu_0}, \tag{3.4}$$

where $H(z)$ is the local Hubble parameter, $\nu_0 = 115.27\,\text{GHz}$ is the frequency of the emitted CO$1 \rightarrow 0$ signal, and $\delta\nu^{\text{obs}}$ is some small frequency range we observe the signal. We observe redshifts around $z \approx 2.9$, giving a local Hubble parameter of [8]

$$\begin{aligned} H(z = 2.9) &\approx H_0 \sqrt{\Omega_{m,0} a^{-3} + \Omega_{\Lambda,0}} \\ &= 70\,(\text{km/s})/\text{Mpc} \times \sqrt{0.31(3.9)^3 + 0.69} \\ &\approx 306\,(\text{km/s})/\text{Mpc}. \end{aligned} \tag{3.5}$$

Our frequency bins are of size $1.95\,\text{MHz}$, which, together with the Hubble parameter, gives a co-moving parallel distance of

$$\delta l_{||} \approx 0.25\,\text{Mpc} \tag{3.6}$$

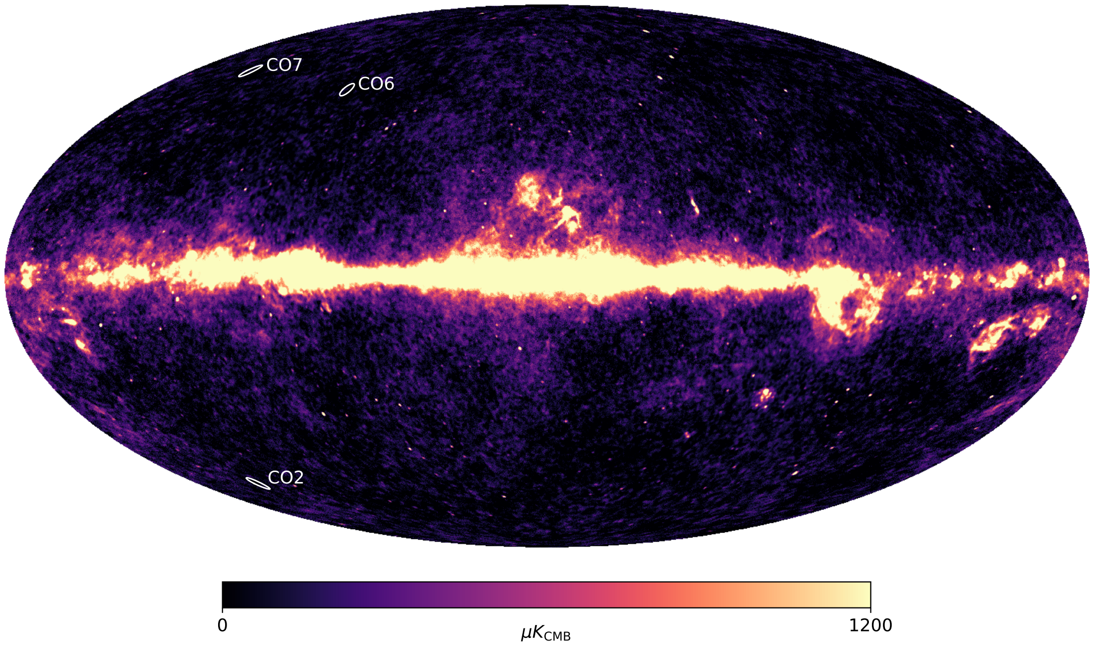


Figure 3.3: The three main CO fields observed by the telescope overplotted as circles of radius 2° on top of the *Planck* LFI 30 GHz full-mission map. Figure credit: Nils-Ole Stutzer.

per frequency bin. In our finished maps, we combine 16 frequency channels in 31.25 Mhz bins, instead giving a spatial resolution of 4 Mpc. Combining all 4096 high resolution frequency bins, we observe a field of around 1 Gpc in the parallel direction.

In the perpendicular direction, a stretch on the sky of angular size $\delta\theta$ will translate to a comoving length of [41]

$$\delta l_\perp = r(z)\delta\theta = \delta\theta \int_0^{z_1} \mathrm{d}z \, \frac{c}{H(z)}. \tag{3.7}$$

Our maps are binned to pixels of 2×2 arcmin^2. Inserting for this as well as $H(z)$ and numerically integrating up to the mean redshift of $z = 2.9$, we get

$$\delta l_\perp \approx 3.6 \, \mathrm{Mpc}. \tag{3.8}$$

An entire observational field is about 4×4 degrees, which corresponds to a size of around 433 Mpc, in other words, about half the size of the parallel field direction.

3.4 Scanning strategies

The choice of scanning strategy is a major design decision in ground-based astronomy. COMAP divides scanning periods into *observations* of typically 1 hour each. At the

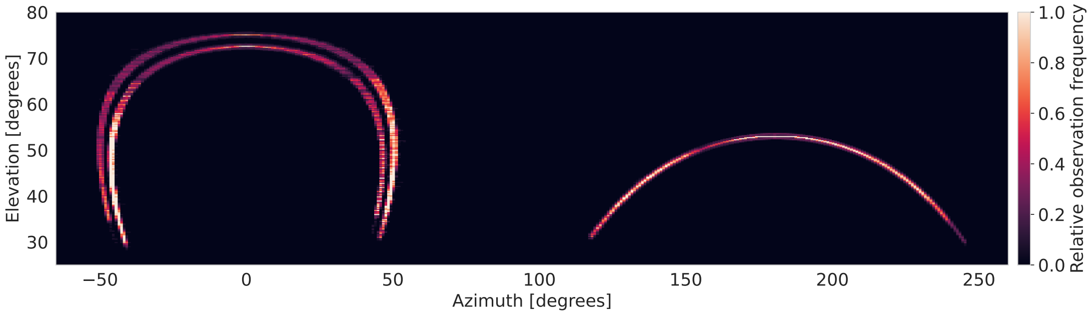


Figure 3.4: The three observational fields of COMAP, as we observe them from Earth. CO2 on the right, CO6 as the inner circle on the left, and CO7 as the outer circle. The relative frequency of observation is shown, to indicate where the fields are most often observed.

beginning and end of each observation, the calibration vane is rotated across the feed array, to perform a gain calibration. Each observation is again divided up into a series of *scans*. A scan is a period where the telescope observes around a single point on the sky for 5-10 minutes, until the observational field has drifted past the telescope field of view, due to the rotation of the earth. The telescope then performs a repointing to catch up with the field and starts a new scan.

COMAP has three main scanning strategies, namely circular scans, Lissajous scans, and constant elevation scans (CES) [41]. Each type of scanning strategy has its benefits and drawbacks. Circular scans were abandoned early in the experiment, and as of late 2020, CES scans are employed full-time in favor of Lissajous scans. This decision was made because the CES data was observed to give cleaner power spectra. It is, however, not entirely unthinkable that this decision is reconsidered if we achieve a better understanding of the Lissajous systematics, and become capable of effectively counteract them.

Circular scans simply circulates the central scan point at some constant radius for the duration of the scan. This type of scan is both easy to implement and perform, with little strain on the telescope mechanics, due to constant velocity and acceleration during the entire scan duration. It does, however, suffer from low coverage of the observational field, and bad cross-linking (meaning that we always observe the same observational patch from the same angle).

Constant elevation scans repoints the telescope back and forth in azimuth while keeping the elevation of the telescope constant. This means that the telescope needs to constantly accelerate and decelerate, and also gives much higher pointing velocities in the middle of the field. CES makes up for these drawbacks by holding the same elevation during an entire scan, which means a much more predictable atmosphere and ground pickup in the signal. CES does, however, suffer from bad cross-linking, as it is only motion is in the azimuth direction.

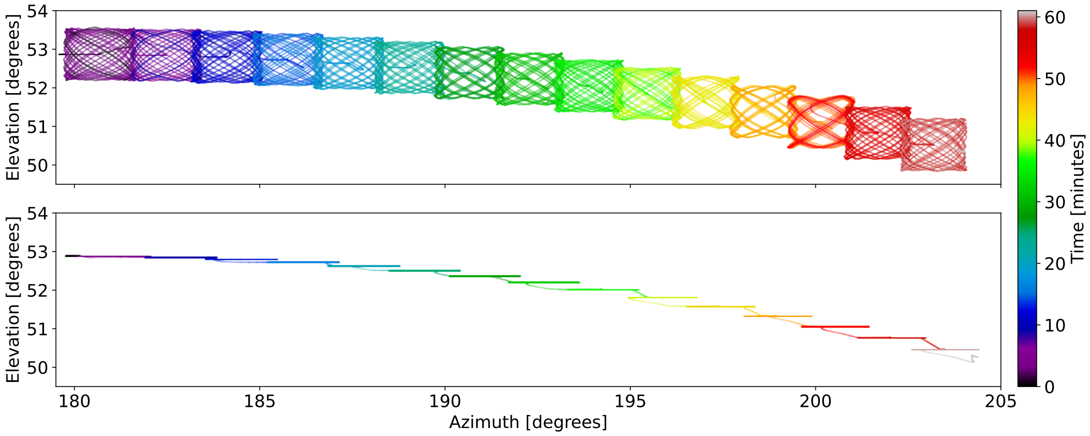


Figure 3.5: Telescope pointing pattern for an observation employing Lissajous scans (top) and another employing CES scans (bottom). For both observation, we see 15 individual scans, with a repointing in between, to follow the observational field across the sky.

Lissajous scans perform circle-like harmonic motions across the field, described by

$$\mathrm{az} = A\sin(at+\phi), \qquad \mathrm{el} = B\sin(bt), \tag{3.9}$$

where A and B are parameters decided by the size of the observational field, while a and b are randomly generated parameters that decide the shape of the harmonic motion. This harmonic motion of non-repeating patterns gives good cross-linking, often observing the same patch of sky from many different directions in a single scan. Lissajous does, however, have a more complicated scanning pattern, and spends more time at the edges of the observational field than the center.

The motion of the telescope for Lissajous and CES scans across the sky are illustrated in figure 3.5.

3.5 Noise sources

One of the most central goals of this thesis is the improved understanding and handling of different types of noise in the COMAP data. We will now summarize the most important sources of this noise, and give a brief summary of their properties. We can roughly divide the noise into three categories

- **White noise** - Completely random, uncorrelated noise. There is not really anything we can do with this type of noise, and it will integrate down as $\sim 1/\sqrt{t}$, as we increase our observational time.
- **Temporally correlated noise** - Noise which has a non-flat power spectrum, usually on a $1/f$ form, which we will get into more detail on later. This noise is correlated in time, but not in telescope pointing. We can usually remove a lot of

this noise due to its predictable frequency dependency, but it is also not especially dangerous, and will simply increase our required observational time.

- **Biased noise** - Noise which is in some way correlated with the telescope pointing. Since this type of noise is not independent between different observation, it will not integrate down with increased observational time. Biased noise could be interpreted as actual CO signal, and bias our results, making this type of noise dangerous and important to deal with.

There exist many sources of all these noise types, and we will now summarize the most important ones for the COMAP data.

Atmosphere and weather
Emission from the atmosphere is one of the largest contributions to the observed signal, and variations in the atmospheric thickness will result in varying signal strengths. Additionally, weather effects such as clouds, wind, humidity, and temperature all affect the signal behavior of the atmosphere. The atmospheric thickness follows a predictable $1/\sin(\mathrm{el})$ profile, which usually describes the signal strength well. However, weather effects will deviate from this profile but will be heavily correlated across both feeds and frequencies, which simplifies detection and removal.

Ground
The far sidelobes of the telescope will, for almost all elevations we observe, pick up emission from the ground around the telescope. While the far sidelobes contain little power compared to the main lobe, the ground has a brightness temperature of 300 K, far more than any other signal source. The ground pickup is also correlated both in telescope pointing and in time (assuming stationary mountains). This makes ground pickup a dangerous systematic, which it is important to remove properly.

Astrophysical Foregrounds
There are a number of other astrophysical sources in the same frequency range as we observe on. These are, obviously, completely correlated in both time and Galactic coordinates. Luckily for us, most of these are not line emissions, but rather continuum sources, which are well approximated as linear on short frequency ranges.

Standing waves
Wires and electronics in the telescope give rise to standing waves in the signal they output. These are strongly correlated across frequencies and time. They often have predictable frequency profiles and are correlated across feeds, which helps us to remove them.

Gain fluctuations
The LNAs of each feed give rise to fluctuations in the gain of the telescope. These fluctuations are observed to be constant across frequencies, and be correlated in time with a $1/f$ power spectrum [42]. We will come back to this in great detail.

Chapter 4

The Oslo COMAP Data Analysis Pipeline

Having given a general overview of the COMAP experiment, we will now go into more detail on how we analyze and process COMAP data. We will focus mostly on noise characterization and removal, which are performed by the `l2gen` program, and are very central to this thesis.

4.1 Data files

As discussed in section 3.4, the telescope performs a series of **observations**, which begin and end with a gain calibration measurement. These observations are divided into **scans** of typically 5-10 minutes, with repointing of the telescope happening in between the scans. Each observation is stored as a separate file in the hdf5 file format [43], and transferred from OVRO to Oslo. We refer to these as **level 1 files**, and they contain all information about the observation. The signal data itself we refer to as **TOD** (Time Ordered Data).

The signal, as stored in the level 1 files, is sampled at an interval of 20 ms, which over a typical observation of an hour gives a total of 180,000 samples. This is over 4 sidebands, each containing 1024 frequency bins, for each of the 20 feeds, giving each TOD array a shape of approximately $[20, 4, 1024, 180000]$, for a total of $\approx 1.5 \times 10^{10}$ data values. The TOD are stored as 32-bit floats, meaning that the total storage requirement per level 1 file is approximately 60 GB. (Everything else in the level 1 files is of negligible size.) This translates to about 500 TB per year of observation.

After passing through the low-level data processing code, known as `l2gen` and described in detail below, the level 1 files are converted into what we call **level 2 files**. These are files containing filtered and down-sampled TOD, which are deemed suitable for mapmaking. They are separated on a per-scan basis instead of a per-observation basis.

The total storage requirements of the level 2 files are much lower, as they have been decimated by a factor of 16 in the frequency domain, down from 1024 channels per sideband, to 64.

File type	Contains	Typical data shape	File size	Storage/ obs. year
Level 1	A single unprocessed observation	$[20, 4, 1024, 180000]$	60 GB	500 TB
Level 2	A single processed scan	$[20, 4, 64, 20000]$	0.4 GB	30 TB

4.2 Data model

We will in this section outline the different components involved in the signal picked up by our detectors. For a general introduction to radiometers, see [44]. The power response of the detectors of a radiometer can be modeled by integrating over all frequencies as

$$P = k_B G \int_{-\infty}^{\infty} \tau(\nu) T_{\mathrm{sys}}(\nu)\, \mathrm{d}\nu\,, \tag{4.1}$$

where k_B is the Boltzmann constant, $T_{\mathrm{sys}}(\nu)$ is the system temperature, which is the brightness temperature perceived by the detectors at a given frequency, and $\tau(t)$ is the bandpass response of the detectors, which represents the sensitivity of the instrument to different frequencies. The gain, G, is a unitless amplification of the signal by the instrument. COMAP uses a spectrometer, meaning that it detects signal over a large set of neighboring frequency bins. For each small frequency interval, $\nu_1 < \nu < \nu_2$, $\nu_2 - \nu_1 = \Delta\nu$, the bandpass response can be approximated as a top-hat function on the form

$$\tau(\nu) = \begin{cases} 1 & \text{if } \nu_1 < \nu < \nu_2 \\ 0 & \text{if else.} \end{cases} \tag{4.2}$$

or each such frequency interval, we then get the power response

$$P = k_{\mathrm{B}} \Delta\nu G(\nu) T_{\mathrm{sys}}(\nu), \tag{4.3}$$

where we have chosen to absorb the bandpass response into the gain, such that the gain is now a frequency dependent parameter.

The system temperature is the temperature measured by the detectors, both from the brightness temperature entering the telescope and from internal contributions, such as the physical temperature of the detectors. The gain of the telescope represents the conversion from temperature into the arbitrary power units recorded by the telescope, decided by the inner workings of the telescope, like the LNAs. Both the gain and system temperature fluctuate over time, and both are complicated functions of frequency. The CO signal will enter the system temperature term, T_{sys}, which typically takes values of $40 - 50\,\mathrm{K}$. This is many orders of magnitude larger than the CO signal itself, which is estimated to be around a few μK. The system temperature is mainly decided by the physical temperature of the detectors, atmospheric thickness, weather, and the CMB temperature.

4.3 System temperature calibration

An important aspect of our data analysis is the capability of separating the system temperature and the gain, as the signal we wish to study is solely contained in the system temperature. This separation is performed by a calibration, where the telescope points at some source with a known brightness temperature, preferably much higher than that of the ordinarily observed source. In the case of COMAP, we mainly perform this calibration by placing a calibration vane of known temperature in front of the telescope. This is referred to as a hot load, and we refer to its temperature as T_{hot}, and the observed power as P_{hot}. We refer to a normal scan of the sky as a cold load, with temperature T_{cold} and observed power as P_{cold}.

The observed signal of the telescope is modeled as[1]

$$P = GT_{\text{sys}}. \tag{4.4}$$

When observing the sky, i.e. a "cold load", we model the signal as

$$P_{\text{cold}} = GT_{\text{cold}} = G[T_{\text{rc}} + (1 - e^{-\tau})\eta T_{\text{atm}} + (1 - \eta)T_{\text{grd}} + e^{-\tau}\eta T_{\text{cmb}}]. \tag{4.5}$$

where T_{rc}, T_{atm}, T_{grd} and T_{cmb} are the temperatures of the receiver and electronics, the atmosphere, the ground, and the CMB, respectively. Additionally, τ is the optical depth of the atmosphere, such that any signal moving through the atmosphere gets reduced by a factor $e^{-\tau}$, as is the case for the CMB. Finally, η represents the fraction of signal lost due to ground spillover, mostly due to signal hitting the support legs holding the secondary reflector.

The first term in the brackets is simply the receiver temperature. The second term is the atmosphere, which emits a signal of brightness temperature $(1 - e^{-\tau})T_{\text{atm}}$, and gets multiplied by a factor η by the ground spill. The third term is the ground, which emits a signal of system temperature T_{grd}, but only the signal reflected into the telescope gets observed, which is a factor of $(1 - \eta)$. The last term is the CMB, which gets reduced both by the atmospheric absorption, and the ground spill. The CO signal itself, as well as other astrophysical foregrounds, can be included in the T_{cmb} term, although we often neglect them in the calibration calculations, as they are small and/or unknown.

During the calibration measurement, a calibration vane is placed in front of the receiver, blocking all sky signals. This is known as a "hot load", and we model its signal as

$$P_{\text{hot}} = GT_{\text{hot}} = G[T_{\text{rc}} + T_{\text{cal}}], \tag{4.6}$$

where T_{cal} is the temperature of the calibration vane.

---

[1] We have, for simplicity's sake, omitted k_{B} and $\Delta\nu$. They can be included in the gain if desired.

4.3.1 T_{sys} derivation

The purpose of the calibration is to separate the measured data into gain and temperature. By default, the gain is an unknown quantity, and we cannot directly tell its value from a cold load. However, using a calibration source of known temperature as a reference, the decomposition becomes possible.

Combining equations 4.6 and 4.5, we get

$$P_{\text{hot}} - P_{\text{cold}} = G[T_{\text{hot}} - T_{\text{cold}}], \tag{4.7}$$

$$G = \frac{P_{\text{hot}} - P_{\text{cold}}}{T_{\text{hot}} - T_{\text{cold}}} = \frac{P_{\text{hot}} - P_{\text{cold}}}{T_{\text{cal}} - (1 - e^{-\tau})\eta T_{\text{atm}} - (1 - \eta)T_{\text{grd}} - e^{-\tau}\eta T_{\text{cmb}}}. \tag{4.8}$$

We introduce the ambient temperature T_{amb} to be the approximate temperature at and around the telescope, such that we can assume $T_{\text{atm}} = T_{\text{grd}} = T_{\text{cal}} = T_{\text{amb}}$. Inserting for this we get

$$G = \frac{P_{\text{hot}} - P_{\text{cold}}}{e^{-\tau}\eta T_{\text{amb}} - e^{-\tau}\eta T_{\text{cmb}}} = e^{\tau}\eta^{-1}\frac{P_{\text{hot}} - P_{\text{cold}}}{T_{\text{hot}} - T_{\text{cmb}}}. \tag{4.9}$$

While P_{hot}, P_{cold}, T_{amb} and T_{cmb} are all known quantities, we do not accurately know what values η and $e^{-\tau}$ take. This means that, while we do not have all required quantities to calculate G, we **do** however have everything we need to calculate the quantity $e^{-\tau}\eta G$, which takes the form

$$e^{-\tau}\eta G = \frac{P_{\text{hot}} - P_{\text{cold}}}{T_{\text{amb}} - T_{\text{cmb}}}. \tag{4.10}$$

This again allows us to calculate $e^{\tau}\eta^{-1}T_{\text{sys}}$, as

$$e^{\tau}\eta^{-1}T_{\text{sys}} = \frac{P_{\text{cold}}}{e^{-\tau}\eta G} = \frac{T_{\text{amb}} - T_{\text{cmb}}}{P_{\text{hot}}/P_{\text{cold}} - 1}. \tag{4.11}$$

In summary, while we do not have a solution for T_{sys} and G without knowing τ and η, we do have solutions for the related quantities

$$T'_{\text{sys}} = e^{\tau}\eta^{-1}T_{\text{sys}} \qquad \text{and} \qquad G' = e^{-\tau}\eta G. \tag{4.12}$$

This new definition of the system temperature takes into account signal lost to both the ground spill and the atmosphere. We have moved these contributions from the temperature to the gain. While T_{sys} represents the observed brightness temperature of the telescope, T'_{sys} represents the brightness temperature observed if we imagine that the telescope was moved to the top of the atmosphere (under the assumption that all observed brightness temperature originates from atop the atmosphere), and the ground spill was removed.

Apart from actually being able to separate gain and system temperature, one advantage of this rescaling is that a 1 K increase in the CO source temperature will directly

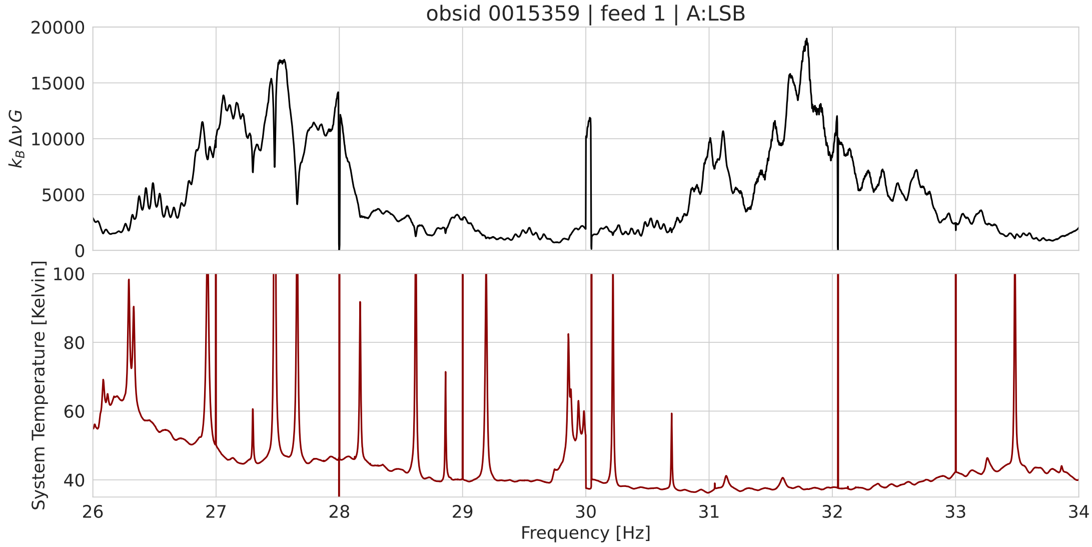


Figure 4.1: Gain and system temperature of feed 1 as function of frequency, for a randomly chosen observation.

translate to a 1 K increase in the system temperature T'_{sys}. This is not the case for T_{sys}, where it would be scaled by both a factor of $e^{-\tau}$ from the signal loss in the atmosphere and a factor of η for the signal lost to ground spill. From now on, when speaking of the system temperature T_{sys}, we will be referring to the rescaled "effective" system temperature, $T'_{\text{sys}} = e^{\tau}\eta^{-1}T_{\text{sys}}$.

4.3.2 Calibration summary

The hot-load calibration is performed before and after each observation, where we can precisely separate the telescope gain and system temperature, using equations 4.10 and 4.11. An example of a gain and system temperature profile in frequency is shown in figure 4.1. Each feed has a unique (but similar) gain and system temperature profile. As we see from the figure, some frequencies are prone to larger system temperature than others. Typical values range from 40 K to 60 K. Another noticeable feature are what we call the T_{sys} spikes, which are specific frequencies with much higher T_{sys} values than neighboring frequencies. It is unknown exactly what causes these spikes, but they are found at specific, unchanging frequency channels, and are found to be more susceptible to some types of correlated noise, and therefore usually masked away in practice.

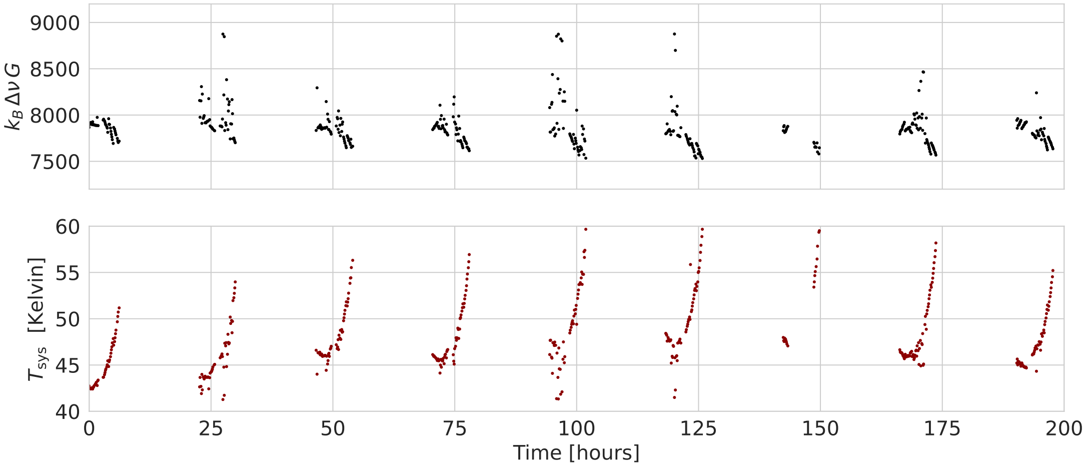


Figure 4.2: Frequency averaged gain and system temperature of feed 1 as a function of time, for nine days of observations around the 1st of November 2019.

Both the system temperature and gain fluctuate in overall amplitude over time, but they mostly retain their overall frequency profile per feed. Figure 4.2 shows the mean gain and system temperature for a series of observations over 9 days. The system temperature varies substantially over the course of a day but follows a repeating pattern from day to day. This is both because it is impacted by the outdoor temperature at the telescope and because we perform the same observational patterns each day, and T_{sys} varies with atmospheric thickness, which depends on elevation.

4.4 l2gen

The perhaps most central part of the Oslo data analysis pipeline is the program known as `l2gen`, which translates level 1 files into level 2 files. The program filters and processes the data with the intention of removing as much noise and systematics as possible. The underlying structure of a level 2 file remains unchanged from that of a level 1 file, with the exceptions of a decimation in the frequency domain and the addition of some housekeeping data.

As mentioned earlier, each level 1 file contains a single observation, which consists of a series of scans, with a calibration measurement at the beginning and end of the observation. `l2gen` is performed on every scan individually, completely agnostic to the existence of other scans, except the values obtained from the previous and next calibration scans. Table 4.1 gives an overview of the four filters in the pipeline, summarizing their most important properties.

Filter name	Fits over	Works by	Removes
Normalization	time	Fitting and dividing by a running mean	Slow-running gain and temperature fluctuations.
Pointing template	time	Fitting an subtracting an az/el template	Atmosphere, ground, and other pointing systematics.
Polyfilter	frequencies (single sideband)	Fitting and subtracting a polynomial in frequency space for each timestep.	Correlated noise, leftover atmosphere and ground, continuum foregrounds.
PCA filter	feeds, sidebands, frequencies	Removing the first 4 principal components of a PCA fit on the data.	Standing waves, leftover weather and atmosphere fluctuations.

Table 4.1: Overview of the filters in the `l2gen` pipeline, together with what dimension they fit over, a summary of how they work, and what specific systematics they target.

4.4.1 Normalization

The first step in the pipeline is a **normalization** of the data. The purpose of this is to remove the bandpass response, which we talked about in 4.2, and which we have incorporated into the gain. Without this, it would be impossible to compare values across frequencies, as they would be multiplied by some arbitrary factor. We also want to remove slow-running fluctuations in the gain and system temperature. The normalization applies a highpass filter to the data, which in practice is done by dividing the data by their running mean, and subtracting one,

$$d_{\text{norm}} = \frac{d}{\bar{d}} - 1. \tag{4.13}$$

The running mean $\bar{d}$ is obtained by applying a lowpass filter to the data. The filter is applied by multiplying by a weight function in Fourier space, as

$$\bar{d} = \mathcal{F}^{-1}\{W \cdot \mathcal{F}\{d\}\}. \tag{4.14}$$

where W is defined as

$$W = \left[1 + \left(\frac{f}{f_{\text{knee}}}\right)^{\alpha}\right]^{-1}. \tag{4.15}$$

The filter has two parameters, f_{knee} and α; f_{knee} is the frequency above which signal gets suppressed, while α is the slope of the suppression effect beyond this frequency, defining how harsh the filter is. We use values of $\alpha = -4.0$ and $f_{\text{knee}} = 0.01\,\text{Hz}$, meaning that all modes below approximately 1.5 minutes get suppressed. We then divide the data upon this running mean, such that, for the actual data, all modes *longer* than 1.5 minutes are suppressed. We are left with short-term fluctuations in the data. We expect very little actual CO signal to be correlated on timescales anything close to this, as it takes seconds, not minutes, to cross an entire observational field once. The purpose of the normalization is to make the data comparable for different frequencies by bringing them to the same scale, as well as remove slow-running fluctuations in both system temperature and gain.

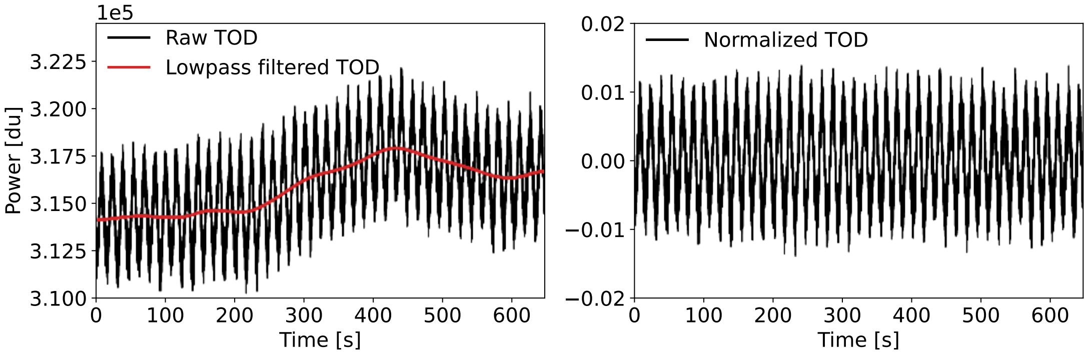


Figure 4.3: Visualization of normalization filter on the TOD for the 29.14 GHz frequency channel of feed 1, as function of time. Left: TOD of a scan before normalization, with the lowpass filtered version of the TOD overplotted in red. Right: TOD after normalization.

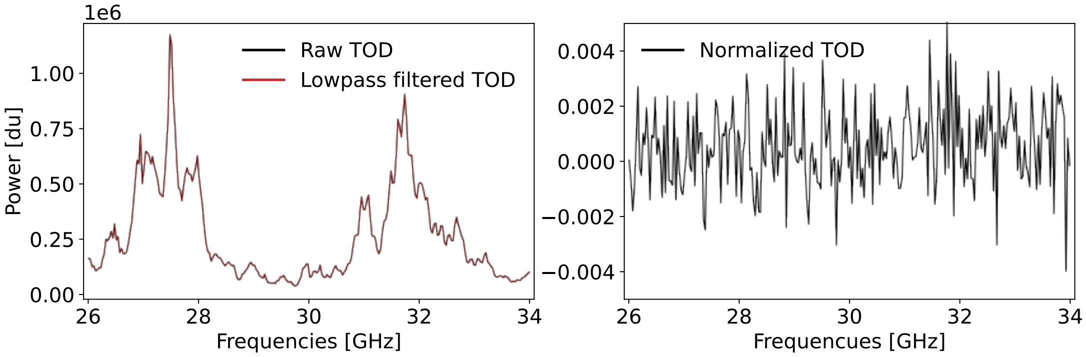


Figure 4.4: Visualization of of normalization filter on the TOD for a randomly selected timepoint of feed 1, as function of frequency. Left: TOD of a scan before normalization, with the lowpass filtered version overplotted in red. The lowpass filtered TOD is virtually indistinguishable from the TOD itself. Right: TOD after normalization.

4.4.2 Pointing template removal

The second part of the pipeline is the removal of pointing templates. The intention behind this filter is to remove any systematic which is correlated to the pointing direction of the telescope in Earth coordinates (elevation and azimuth). This will typically be any signal originating from Earth, most notably from the atmosphere, and the ground around the telescope.

For each frequency, a three-parameter pointing template d_{point} is fitted to and subtracted from the scan, on the form

$$d_{\text{out}} = d - d_{\text{point}}, \qquad d_{\text{point}}(t) = \frac{\tau_0}{\sin \text{el}(t)} + a\,\text{az}(t) + c, \tag{4.16}$$

where az(t) and el(t) is the pointing direction of the telescope, and τ_0, a and c are parameters found with a maximum likelihood fit of the model to the data.

The first term in the pointing template accounts for the atmospheric thickness at different elevations, which approximately follows a $1/\sin(\text{el})$ profile. τ_0 is the atmospheric optical depth at zenith. Apart from the atmosphere, the data might contain an unknown number of other pointing-related systematics. Principal among these is sidelobe ground contamination, where a telescope sidelobe picks up a signal from the nearby mountain range. As the ground has a temperature that is orders of magnitude larger than the sky signal, even small sidelobe contaminations might have a large impact. In our current pointing model, we fit any remaining pointing systematic as a simple linear function in azimuth, fitting a as a slope parameter. The reasons for not choosing a more complicated model are mainly the unknown nature of the remaining systematics (most importantly the ground pickup) and the fear of removing actual signal.

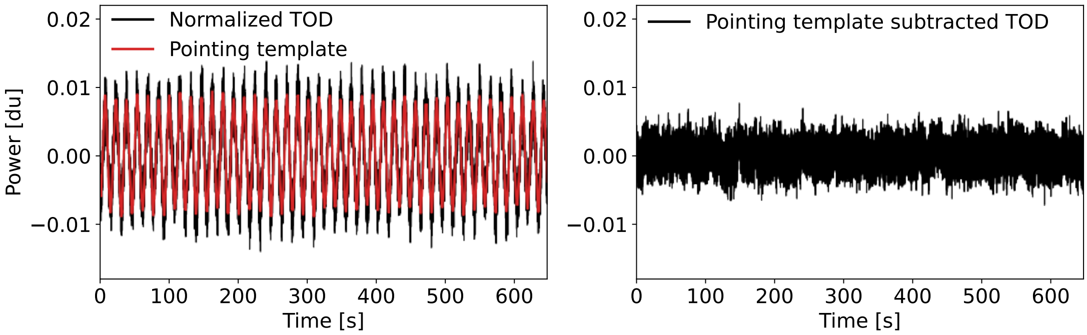


Figure 4.5: Left: TOD of a scan before pointing template subtraction, with the pointing template fit overplotted in red. Right: TOD after pointing template subtraction.

4.4.3 Polynomial filter

The polynomial filter is our strongest general purpose noise filter, aimed at removing correlated noise and continuum sources. The filter works by fitting and subtracting a polynomial in frequency across each sideband for every single timestep of the data. Our pipeline currently uses a first order polynomial, on the form

$$d_{\text{out}} = d - d_{\text{poly}}, \qquad d_{\text{poly}} = c_0 + c_1\nu, \tag{4.17}$$

with c_0 and c_1 fitted to each sideband and timestep in the TOD.

The polyfilter aims at removing a number of systematics correlated across frequency. Most importantly, the filter removes correlated gain fluctuations, often referred to as

$1/f$ noise, as it has a power spectrum that is inversely proportional to the frequency to some power. These gain fluctuations are mostly introduced by the low noise amplifiers in the telescope. The polyfilter also removes temperature continuum sources, such as astronomical continuum foregrounds (thermal dust, synchrotron radiation, free-free emission, etc), as well as any leftovers of atmosphere and ground from the pointing template removal.

Where actual CO signal is concerned, the polyfilter will only remove the very largest structures, as it is fit to each sideband, which corresponds to physical sizes of $\approx 250\,\mathrm{Mpc}$. Any CO structure substantially smaller than this will be affected little by the filter. We will come back to this when we look at the pipeline transfer function at the end of the chapter.

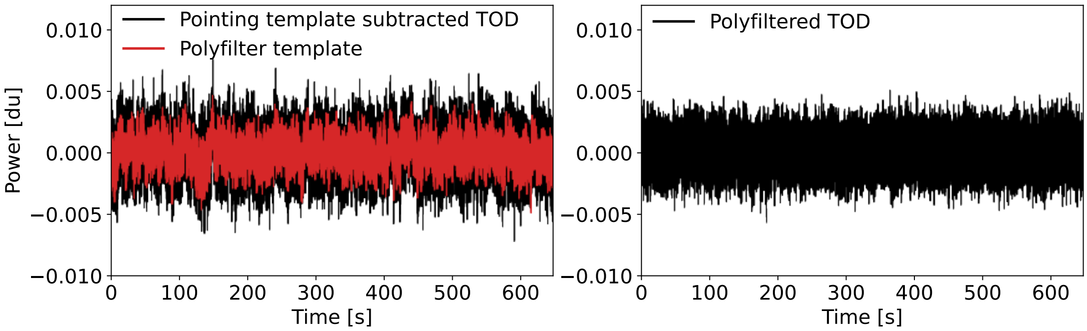


Figure 4.6: Left: TOD of a scan before the polynomial filter, with the polynomial fit overplotted in red. Right: Filtered TOD.

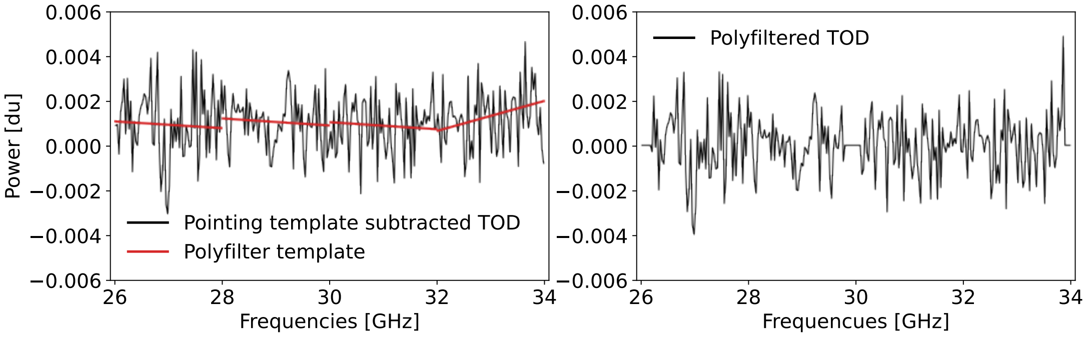


Figure 4.7: Visualization of polynomial filter on the TOD for a randomly chosen timepoint of feed 1, as function of frequency. Left: TOD of a scan before the polynomial filter, with the four polynomial fits on each sideband overplotted in red. Right: Filtered TOD.

4.4.4 PCA filter

Principal component analysis (PCA) is a decomposition technique which, from a set of n vectors of dimension p, finds an ordered basis spanning the vector space, known as the *principal components* of the vectors. These principal components come in an ordered fashion, with each component, in descending order, containing as much as possible of the common features in the vectors. Another way of putting this is that the first component points in the direction of the largest variance in the vector set, the second component in the orthogonal direction of second-most variance, and so forth. PCA is often employed as a dimensionality reduction technique, as a reduced vector set of size m where $m < n$ will still well explain much of the data, as long as m is not too small and the data not too high in entropy. In practice, the principal components are found as the eigenvectors of the covariance matrix of the vector set. It is the eigenvalues of these eigenvectors which define how much of the vectors are contained in each principal component. If the eigenvalues of subsequent components fall quickly, the set of vectors are easy to reduce and low in entropy.

In the case of COMAP, the set of vectors are the TOD of each feed and frequency, meaning we typically have a set of $n = 19 \cdot 4096$ vectors, with dimension equal to the number of timesteps in the scan, p. Contrary to typical usage of PCA, in l2gen we find and *subtract* the first 4 principal components of the data. In other words, we remove the largest cross-vector features in the TOD. This is because these are features correlated across feeds and frequencies, corresponding predominantly to systematics.

Let D be a data matrix where the rows are the TOD of each feed and frequency, such that

$$\mathbf{D} = \begin{pmatrix} \mathbf{d}_{\nu_1} \\ \mathbf{d}_{\nu_2} \\ \vdots \\ \mathbf{d}_{\nu_n} \end{pmatrix} = \begin{pmatrix} d_{\nu_1,t_1} & d_{\nu_1,t_2} & \cdots & d_{\nu_1,t_p} \\ d_{\nu_2,t_1} & \ddots & & \\ \vdots & & & \\ d_{\nu_n,t_1} & \cdots & & d_{\nu_n,t_p} \end{pmatrix}. \tag{4.18}$$

The covariance matrix of this vector set is

$$\mathbf{C} = \mathbf{D}^T\mathbf{D} = \begin{pmatrix} C_{1,1} & C_{1,2} & \cdots & C_{1,n} \\ C_{2,1} & \ddots & & \\ \vdots & & & \\ C_{n,1} & \cdots & & C_{n,n} \end{pmatrix}. \tag{4.19}$$

Let $\mathbf{w}_{(i)}$ be the set of eigenvectors of $\mathbf{C}$, with corresponding eigenvalues λ_i. The l2gen PCA filter works as follows

$$\mathbf{d}_{\mathrm{out}} = \mathbf{d} - \sum_{i=1}^{4} a_i \mathbf{w}_{(i)}, \tag{4.20}$$

where a_i is the inner product of the data with the ith principal component (which represents "how much" of that principal component is within that TOD),

$$a_i = \mathbf{d} \cdot \mathbf{w}_{(i)}. \tag{4.21}$$

The PCA filter is great at removing weather contaminations that slips past the polyfilter, as these are strongly correlated across feeds. It also removes standing waves very well. We do not expect CO signal to be correlated across frequencies or feeds to any large extent, leaving the signal mostly intact. The PCA is in some ways a more conservative filter than the polyfilter, as it looks for correlated features across both all feeds *and* frequencies.

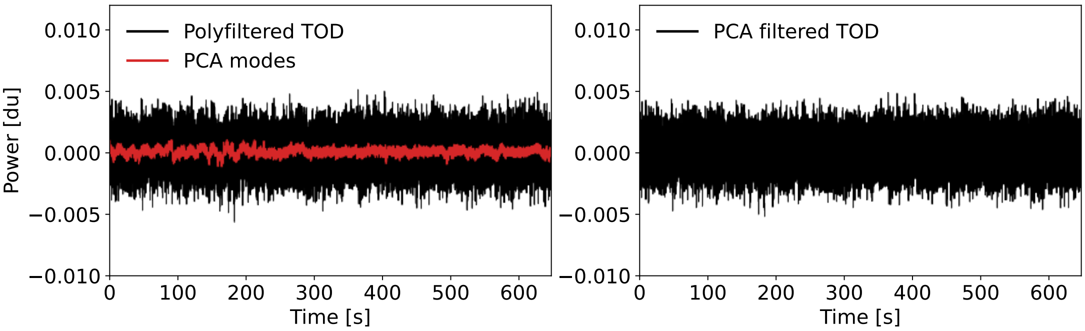


Figure 4.8: Visualization of PCA filter on the TOD for 29.14 GHz frequency channel of feed 1, as function of time. Left: The TOD before the PCA filter, with the sum of the four leading PCA components overplotted in red. Right: The TOD after PCA filtering.

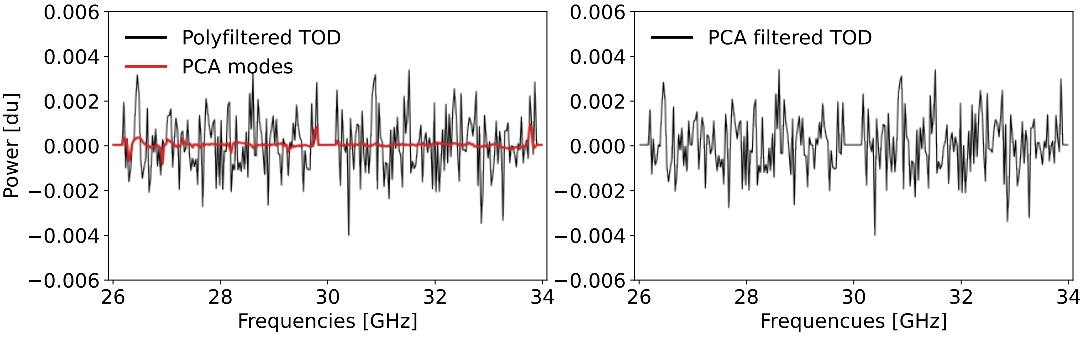


Figure 4.9: Visualization of PCA filter on the TOD for a randomly chosen timepoint of feed 1, as function of frequency. Left: The TOD before the PCA filter, with the sum of the four leading PCA components overplotted in red. Right: The TOD after PCA filtering.

4.4.5 Masking

After all filters have been applied to a scan, `l2gen` masks out badly behaving feeds and frequency channels in the data. Masking simply means that we flag these feeds or channels to be ignored in future analysis, such as decimation, maps, or power spectra. The usual way of doing this is by setting the variance of the masked data to infinity $\sigma^2 = \infty$, effectively weighting that data by 0 when e.g. creating a map.

A number of factors are considered when masking a frequency channel, but they mostly have to do with cross-frequency correlation. Filtered TOD are expected to be almost consistent with white noise, and uncorrelated across frequencies. The expected correlation between two uncorrelated white noise timestreams of length N is $1/\sqrt{N}$. A large correlation excess above this is considered a sign of bad data. Entire sidebands can also be masked out if the majority of frequency channels in them are masked.

The masks are calculated from the TOD after they have passed through the entire pipeline. However, once the masks are decided, we apply them to the data *before* the polyfilter and PCA filter, and then redo those filters with the masking. This is because, once we have figured out which channels behave badly even after filtering, we do not want these channels to impact the effectiveness of the filters. Since the polyfilter and PCA filter take all frequencies into consideration, leaving out badly behaving frequencies can improve the performance of the filters.

4.4.6 Calibration

After the scan has gone through all filtering and masking, we calibrate the TOD from normalized units to temperature units. We do this by multiplying the TOD by the system temperature, derived in section 4.3. The vane calibration, and thereby the acquisition of P_{hot} and T_{hot}, is only performed at the beginning and end of each observation, while T_{sys} is calculated on a per-scan basis. This is done by linearly interpolating P_{hot} and T_{hot} to the center of each scan, as

$$P_{\text{hot}}(t) = \frac{P_{\text{hot}}^1(t_2 - t) + P_{\text{hot}}^2(t - t_1)}{t_2 - t_1} \tag{4.22}$$

$$T_{\text{hot}}(t) = \frac{T_{\text{hot}}^1(t_2 - t) + T_{\text{hot}}^2(t - t_1)}{t_2 - t_1}, \tag{4.23}$$

where the 1 and 2 superscripts refer to the calibrations at the beginning and end of the observation, respectively. T_{sys} is then calculated for each scan as

$$T_{\text{sys}} = \frac{T_{\text{hot}} - T_{\text{cmb}}}{P_{\text{hot}}/P_{\text{cold}} - 1}, \tag{4.24}$$

where we for P_{cold} use the mean power of each scan. An illustration of the calculation and interpolation of P_{hot} between two calibrations is shown in figure 4.10. We get our calibrated data by multiplying the TOD of each scan by the scan system temperature,

$$d_{\text{calib}} = d \cdot T_{\text{sys}}. \tag{4.25}$$

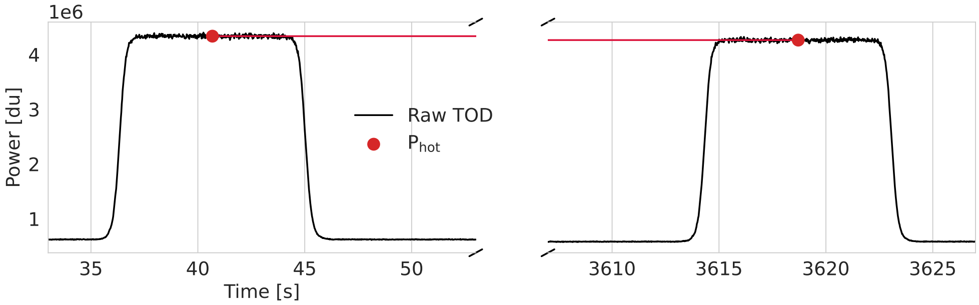


Figure 4.10: Power profile of the TOD during two calibrations, at the beginning and end of an hour long observation. The P_{hot} values are calculated from the power peak, and interpolated to the center of each scan within the observation, as shown by the red line.

4.4.7 Decimation

After calibration, the data are decimated down from 1024 frequency channels per sideband to 64, weighted by the inverse variance of the TOD (or $\sigma^2 = \infty$, in the case of masked channels),

$$d^{\text{dec}} = \frac{1}{\sum_j \sigma_j^{-2}} \sum_j \frac{d_j}{\sigma_j^2}. \tag{4.26}$$

This is done to prepare the signal for mapmaking, for which we do not need as high resolution, and instead prefer the higher signal to noise ratio we get by decimating the frequency channels.

4.4.8 Pipeline summary

The decimated data are finally written to disk as a level 2 file. The ultimate goal of the pipeline is the removal of as much noise as possible while retaining as much of the actual CO signal as possible. It is not trivial to predict the impact each filter has on any potential CO signal, but we do have the means to simulate it. By adding a simulated signal into a series of level 1 files, we can estimate the transfer function imposed by the pipeline on the CO signal, defined as [45]

$$T^{\text{pipeline}}(\mathbf{k}) \approx \left\langle \frac{P_{\mathbf{k}}^{\text{full}} - P_{\mathbf{k}}^{\text{noise}}}{P_{\mathbf{k}}^{\text{signal}}} \right\rangle, \tag{4.27}$$

where $P_{\mathbf{k}}^{\text{full}}$, $P_{\mathbf{k}}^{\text{noise}}$ and $P_{\mathbf{k}}^{\text{signal}}$ are the power spectra of the full simulated data with noise, the noise only, and the signal only, respectively.

Figure 4.11 shows this simulated transfer function for both the default pipeline setup and a series of variations of the pipeline. Generally, all the transfer functions follow the same trend, with the most power retained at scales of $k = 2 - 10\,\text{Mpc}^{-1}$. We also see that none of the largest scales survive the pipeline, because of filters such as the normalization and polynomial filter, which removes large-scale correlations in the signal (in time and frequency, respectively).

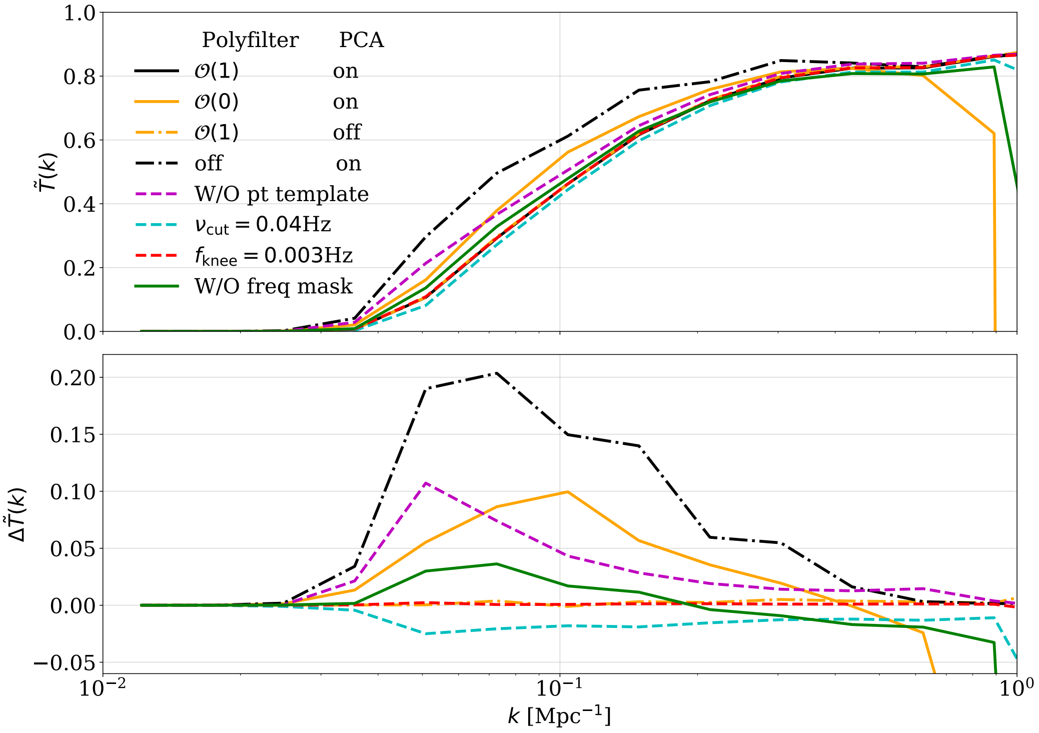


Figure 4.11: Transfer function of the pipeline for different configurations of the pipeline parameters. The top plot shows the total transfer function, while the bottom plot shows the absolute difference from the default parameter set, shown as a solid black line in the top figure. Figure credit: Nils-Ole Stutzer.

Part II

New developments

Chapter 5

Mapmaking

A fundamental challenge in observational astronomy is how to transform a time-ordered signal into some sort of map, consisting of magnitude values over a series of pixels. Many receivers, including the one used by COMAP, employ a very limited set of observational feeds, and instead rely on the telescope movement across the sky to cover the observational field. This means that each feed often covers a vast range of coordinates on the sky in the span of a few minutes or seconds. This is in strong contrast to stationary observers, like a camera, where each detector focuses on the same observational point for the duration of the exposure.

The most trivial problem arising from this difference is figuring out which datapoint on the TOD corresponds to which pixel on the map. We will start off by introducing the easiest and most intuitive mapmaking scheme, which solves this problem only, namely the naive binning scheme. We will then improve upon this by taking noise weighting into consideration. The last scheme we will look into is destriping, a mapmaking technique capable of dealing with temporally correlated noise, and is the method we aim to employ throughout the COMAP experiment. We end the chapter by looking at how we can apply the destriping mapmaking scheme to produce maps of the telescope ground pickup. This will be a central part of our efforts to produce an improved pointing template model for our pipeline.

5.1 Setup

The premise of mapmaking is that we have TOD consisting of N_s datapoints, which we want to transform into a map of N_{px} pixels. We will assume that the dimensionality of the map does not matter and that the spatial arrangement of the pixels does not impact the solution to any of our mapmaking schemes. We can then consider all maps to be one-dimensional, where 2D or 3D maps are simply flattened. The spatial arrangement of the pixels would matter if we imposed a power spectrum prior on the map itself, but we will not cover map priors in this thesis.

5.2 The binning scheme

Our simplest data model makes the assumption that the TOD are a combination of the "true" map values, and uncorrelated white noise. In this model, we write the TOD $\mathbf{y}$ as a function of the map $\mathbf{m}$ as

$$\mathbf{y} = \mathbf{P}\mathbf{m} + \mathbf{n}, \tag{5.1}$$

where $\mathbf{n}$ is the white noise, and $\mathbf{P}$ is the so-called *pointing matrix*. In general, the pointing matrix contain all information about how the TOD picks up data from the map, including the scanning pattern and telescope beam profile. In our model, we will assume a simple beam profile where the telescope picks up only a single pixel of the map in each TOD sample. The pointing matrix then takes a very simple form. The matrix element P_{tp} contains a value of 1 if TOD sample t points at pixel number p, and 0 otherwise. Equation 5.2 shows an example pointing matrix for a TOD of 4 datapoints and a 3 pixel map. Here, the telescope points first once at the first map pixel, then twice at the second map pixel, and then once at the third map pixel, giving the pointing matrix

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{5.2}$$

Under the assumption that the white noise level does not change throughout the TOD, the maximum likelihood solution to the map m in equation 5.1 is

$$\mathbf{P}^T\mathbf{P}\mathbf{m} = \mathbf{P}^T\mathbf{y}. \tag{5.3}$$

Note that $\mathbf{P}^T\mathbf{P}$ is a diagonal matrix, with the diagonal elements representing the number of hits on each pixel. Diagonal matrices are trivial to invert, making the solution of $\mathbf{m}$ explicit, as

$$\mathbf{m} = (\mathbf{P}^T\mathbf{P})^{-1}\mathbf{P}^T\mathbf{y}. \tag{5.4}$$

As the solution of $\mathbf{m}$ is explicit, we might as well write it out in component form. The map pixel m_p takes a value

$$m_p = \frac{1}{N_p}\sum_{j\in p} y_j, \tag{5.5}$$

where $j \in p$ are all the indices of y_j which hits pixel p, and N_p are the number of hits in this pixel.

5.2.1 Noise weighted mapmaking

If we have information about the noise level in different parts of the TOD, we can incorporate this information into the solution by downweighing the contribution of data with a higher noise level. The maximum likelihood solution to equation 5.1 then takes the form [46]

$$\mathbf{P}^T\mathbf{C}_n^{-1}\mathbf{P}\mathbf{m} = \mathbf{P}^T\mathbf{C}_n^{-1}\mathbf{y}, \tag{5.6}$$

where $\mathbf{C}_n = \langle \mathbf{n}\mathbf{n}^T \rangle$ is the noise covariance matrix. We can solve for $\mathbf{m}$ as

$$\mathbf{m} = (\mathbf{P}^T\mathbf{C}_n^{-1}\mathbf{P})^{-1}\mathbf{P}^T\mathbf{C}_n^{-1}\mathbf{y}. \tag{5.7}$$

If the noise is uncorrelated, i.e. white noise, the noise covariance matrix is diagonal, and $\mathbf{P}^T\mathbf{C}_n^{-1}\mathbf{P}$ becomes trivial to invert. We can then write out the map solution in component form as

$$m_p = \frac{1}{\sum_{j \in p} \frac{1}{\sigma_j^2}} \sum_{j \in p} \frac{y_j}{\sigma_j^2} \tag{5.8}$$

where σ_j^2 is the estimated variance of y_j, or $C_n(i,j) = \sigma_j^2 \delta_{i,j}$

It is usually impossible to estimate the white noise level of each individual TOD point. A reasonable approach is to estimate the white noise level on intervals. This will make $\mathbf{C}_n$ piecewise stationary. We will in section 8.1.1 outline how the white noise level of the TOD can be estimated. In COMAP, we usually estimate the white noise level on scan-length sections, separately for each frequency channel.

5.3 Destriping

The destriping mapmaking scheme is an attempt to model not only uncorrelated noise but also temporally correlated noise, which is common in telescope detectors. This is the mapmaking scheme COMAP aims to employ in general and is also what we will use to use to produce ground pickup maps later in this thesis. Let $\mathbf{y}$ be some TOD with both correlated and uncorrelated noise, which we model as

$$\mathbf{y} = \mathbf{P}\mathbf{m} + \mathbf{n}_{\mathrm{corr}} + \mathbf{n}, \tag{5.9}$$

where $\mathbf{n}_{\mathrm{corr}}$ is the correlated noise. The destriper works by dividing the TOD into sections of b consecutive datapoints. For each of these sections, the correlated noise is modeled by one or more *base functions.* The goal of the destriper is to find amplitudes for these base functions which minimizes the variance of the datapoints hitting each pixel. We can write this model as

$$\mathbf{y} = \mathbf{P}\mathbf{m} + \mathbf{F}\mathbf{a} + \mathbf{n}, \tag{5.10}$$

where $\mathbf{F}$ is a known (chosen) matrix defining the shape of the baselines, and $\mathbf{a}$ is a vector containing their unknown amplitudes. $\mathbf{F}$ has shape $(N_s \times N_b)$ where N_s is the number of TOD points, and N_b is the number of base functions. $\mathbf{F}_{ij}$ represents the value of base function number i at the TOD point j (which will be zero except for on the interval of y where that specific base function belongs). The base functions may take any shape, and come in any number per interval. We will here only consider a single base function per interval, in the form of uniform baselines, which take the same value across the entire interval. Such base functions have, despite their simplicity, been found to perform well compared to more complicated base functions, such as Fourier components [46]. Each column of $\mathbf{F}$ will then contain b consecutive 1's, each column starting on the row after the end of the former column.

5.3.1 Maximum likelihood solution

Our signal model presented in equation 5.10 has two unknown quantities, namely the map, $\mathbf{m}$, and the amplitudes of the base functions, $\mathbf{a}$. The idea is to solve first for the base functions and subtract those from the data, hopefully removing as much correlated noise as possible. We then solve for the map of the baseline-subtracted data using the noise weighted binning scheme, as

$$\mathbf{m} = (\mathbf{P}^T \mathbf{C}_n^{-1} \mathbf{P})^{-1} \mathbf{P}^T \mathbf{C}_n^{-1} (\mathbf{y} - \mathbf{F}\mathbf{a}). \tag{5.11}$$

First, however, we need to solve for the baseline amplitudes. The maximum likelihood solution for $\mathbf{a}$ is [46]

$$(\mathbf{F}^T \mathbf{C}_n^{-1} \mathbf{Z}\mathbf{F})\mathbf{a} = \mathbf{F}^T \mathbf{C}_n^{-1} \mathbf{Z}\mathbf{y}, \tag{5.12}$$

where

$$\mathbf{Z} = \hat{\mathbf{I}} - \mathbf{P}(\mathbf{P}^T \mathbf{P})^{-1} \mathbf{P}^T, \tag{5.13}$$

and $\hat{\mathbf{I}}$ is the identity matrix.

Unlike for the map, the left-hand side matrix is not diagonal, and we cannot easily invert it to solve for $\mathbf{a}$. We will instead have to solve it as a matrix equation $\mathbf{A}\mathbf{x} = \mathbf{b}$ by other means, which we will get back to at the end of this chapter.

The solution presented assumed no prior knowledge of the behavior of $\mathbf{a}$. If we wish to put a prior on the power spectrum of $\mathbf{a}$, we can add the covariance matrix of $\mathbf{a}$ to our problem. In this case, the maximum likelihood solution of $\mathbf{a}$ takes the form [47]

$$(\mathbf{F}^T \mathbf{C}_n^{-1} \mathbf{Z}\mathbf{F} + \mathbf{C}_a^{-1})\mathbf{a} = \mathbf{F}^T \mathbf{C}_n^{-1} \mathbf{Z}\mathbf{y}, \tag{5.14}$$

where $\mathbf{C}_a = \langle \mathbf{a}\mathbf{a}^T \rangle$ is the covariance matrix of a. If the power spectrum properties of the correlated noise can be estimated to high accuracy, the addition of a prior has been shown to improve the correlated noise removal, especially for shorter baseline lengths [47].

5.3.2 Some intuition

The destriper can be hard to get ones head around, so let us finish with some intuition. The point of the destriper is to, for each interval of b datapoints (say a hundred) in the TOD, fit some unknown constant a_i to all 100 points. Now, what value should this constant take? The hundred points will hit a number of pixels, and these pixels will also be hit by a large number of other datapoints elsewhere in the TOD. The idea is to look at all the other data hitting the same pixel, and ask if our interval of 100 points would, on average, be more in line with the other data if we added or subtracted some constant from it. If there is correlated noise in our data, this will manifest as raising or lowering the data on stretches. Now we ask this same question, simply on a larger scale: What set of values $\mathbf{a}$, added to different stretches of the of the TOD, will produce minimal variance in the datapoints hitting each pixel? These are the values $\mathbf{a}$ given by our maximum likelihood solution.

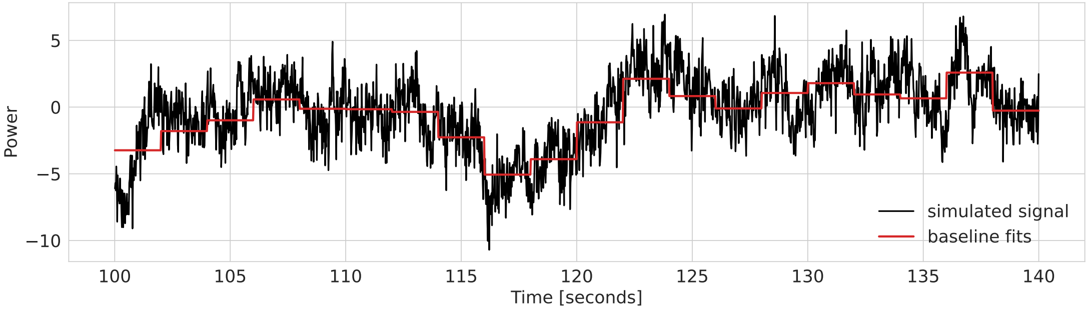


Figure 5.1: A cutout of a signal, and the baseline fits from performing a destriper mapmaking on it. The signal is simulated correlated noise, together with pointing information from a real scan. The destriper baselines are of length 100 points (2 seconds).

A demonstration of a simulated signal, overplotted with the fitted baselines is shown in figure 5.1. We can see how the baselines trace the correlated noise in the TOD to a large extent. Shorter baselines can also be employed to try and better fit the noise, and will usually improve the correlated noise removal, down to a certain point. However, decreasing the baseline size too far is both computationally expensive, and will at some point make the basis functions more degenerate with the white noise, and we will lose more of the actual signal, effectively getting a worse fit. This effect can be mitigated by including an accurate prior, as we discussed at the end of the previous section.

5.4 Ground maps

One of the primary goals of this thesis is the creation of data-driven ground maps, to accurately portray the telescope ground pickup as a function of pointing coordinates. To achieve this, we will employ the destriper to create azimuth-elevation maps with as much of our data as possible.

5.4.1 Dataset

To create a suitable dataset for mapmaking, we create a slightly altered version of the `l2gen` pipeline. Our dataset on hand contains a total of 50,000 scans, distributed over the three observational field of COMAP. To avoid our data being contaminated by other sources, especially pointing-correlated sources, we perform two preliminary cuts on the data. We keep only nighttime data, as to avoid any problems with the Sun hitting our sidelobes, and we remove any scan suspected of being weather contaminated. This leaves us with 17,000 scans. We then perform the following filtering steps on these scans.

Normalization

We perform per-frequency normalization of all scans, identical to the one performed by `l2gen`, using a lowpass filter, as described in 4.4.1. This will get rid of a lot of correlated noise, such as slow running gain and system temperature fluctuations.

Sideband average

To increase our signal-to-noise ratio, and reduce the computational requirements associated with our large dataset, we average our data over each sideband. This still leaves us with some resolution in frequency (4 sidebands). We do a weighted average, to account for some frequencies being noisier than others. From the radiometer equation (which we will talk about in section 8.1.1), we know that the white noise variance of a frequency is proportional to T_{sys}^2, such that the sideband averaged signal becomes

$$d^{\text{sbavg}} = \frac{1}{\sum_{i=1}^{1024} {T_{\text{sys}}}_i^2} \sum_{i=1}^{1024} \frac{d_i}{{T_{\text{sys}}}_i^2}. \tag{5.15}$$

Elevation template removal

We perform an elevation fit and subtraction to the sideband averaged data. This is similar to the pointing template subtraction done in `l2gen`, except that we exclude the linear azimuth term. The filtered TOD of each scan are now

$$d^{\text{out}} = d - d^{\text{el}}, \qquad d^{\text{el}} = \frac{\tau_0}{\sin \text{el}(t)} + c, \tag{5.16}$$

where τ_0 and c are found by a maximum likelihood fit to the scan. The purpose of this filter is to remove the atmosphere from the scan. However, because the atmosphere term is virtually linear on a 1-degree elevation section, any linear elevation structure in the ground will also be fitted and removed. There is not really anything we can do about this, as linear ground structure and the atmosphere is virtually entirely degenerate. A potential solution would be constraining τ_0 with a strong prior around some reasonable atmospheric optical depth. This could introduce other complications and biases, and we have chosen to leave it be for the moment, and settle for losing linear elevation structures in the ground.

Calibration

Finally, we calibrate the scans, to get them in temperature units. The calibration process was described in section 4.4.6, and we use a new database of scan hot load measurements we have created, which we will describe in chapter 7. The system temperature is sideband averaged the same way as the data, before calculating the calibrated from the normalized data as

$$d^{\text{calib}} = T_{\text{sys}} d. \tag{5.17}$$

We then write the filtered scans to file, ready to be made into ground maps. As we have sideband averaged them, the storage requirements of the data are very reasonable ($\sim 1\,\mathrm{TB}$). The pipeline itself spends around 10 minutes per scan per cpu core on an AMD EPYC 7H12, most of it on file read and decompression, and on the normalization, as it is performed before the sideband average.

5.4.2 Ground map destriper

We employ a standard destriper mapmaking method to create azimuth-elevation binned maps from our scan dataset. After specifying an azimuth and elevation range for the map, we select all scans which are entirely within that range. We prefer to avoid having to cut up scans that partially cover a region, as it makes things complicated with the baselines. The white noise correlation matrix $\mathbf{C}_n$ is created as a piecewise stationary diagonal matrix, where we estimate σ_0^2 per scan using equation 8.3. The pointing matrix $\mathbf{P}$ is constructed from binning up azimuth and elevation pointing information from the scan. We employ baselines of length 100 in our destriper setup. The matrix $\mathbf{F}$ has columns of 100 consecutive 1s, each starting the row after the last one ended. We also stop the baselines short upon a scan end.

The primary goal of the destriper is solving the equation

$$(\mathbf{F}^T\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{F})\mathbf{a} = \mathbf{F}^T\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{y} \tag{5.18}$$

for $\mathbf{a}$. As we mentioned earlier, the left hand side of this equation is not trivial to invert, and not really even feasible to calculate or store as a matrix. It is, however, very sparse, and we solve it using a *conjugate gradient method*, an iterative solver which is excellent for sparse $\mathbf{Ax} = \mathbf{b}$ systems [48]. One of the advantages of such a method is that we do not ever need to calculate or store the matrix $\mathbf{A}$, which (while sparse) is large and expensive to compute. We only need to be able to calculate the resulting vector from applying $\mathbf{A}$ to some vector $\mathbf{x}$, which is often, and in our case, much simpler to do. For our conjugate gradient method, we additionally employ a *preconditioner*, which is a matrix $\mathbf{M}$ which approximates $\mathbf{A}$, and for which we can easily compute $\mathbf{M}^{-1}\mathbf{A}$ and $\mathbf{M}^{-1}\mathbf{x}$. Multiplying each side of the equation with a preconditioner is a common way of helping the conjugate gradient converge faster. The equation we actually solve is then $\mathbf{M}^{-1}\mathbf{Ax} = \mathbf{M}^{-1}\mathbf{b}$. We use the perhaps simplest imaginable preconditioner, which is simply the diagonal of $\mathbf{A}$. This is much easier to compute than the entirety of $\mathbf{A}$, and is trivial to invert. If $\mathbf{A}$ is diagonally dominant, which ours is, such a preconditioner is a very cost-effective way of obtaining faster convergence.

Joint elevation template fit

We previously mentioned how performing the elevation template subtraction on our data costs us potential linear elevation ground structures. It is possible to include the elevation fit into the destriper model itself, instead of performing it ahead of time. Remember that the baselines laid out in $\mathbf{F}$ are simply chosen sets of functions spread throughout the TOD. We can introduce a set of new baselines, each the length of one scan, and with amplitudes $\frac{1}{\sin \mathrm{el}(t)}$, were $\mathrm{el}(t)$ are the elevations of each individual datapoint. These new baselines represent the elevation fit. They will overlap with the original ones, and will not be constant. The vector $\mathbf{a}$ will then also contain the τ_0 amplitudes of each scan, which will be fitted the same way as the other baseline amplitudes. This joint fit has a lot of theoretical appeal and was how we originally designed our destriper model. The problem is, however, that the amplitude profile is still entirely degenerate with any linear ground structure, and the conjugate gradient method had huge problems with converging on a solution. This can be somewhat mitigated by placing a strong prior on τ_0, but we are then basically predetermining the atmosphere subtraction, which defeats much of the purpose of a joint fit. We therefore redesigned the destriper model to not include an elevation term, and instead presubtract it.

Chapter 6

Ground Modeling

One of the dominating sources of systematics in the COMAP data is the ground around the telescope, picked up by the telescope sidelobes. These ground systematics are correlated with pointing, and dependent on both the complicated beam profile and the mountain structure around the telescope, making them difficult to properly model and remove. The goal of this chapter is two-fold. First, we perform an analysis of the effect of the ground on the telescope signal, both relative to pointing and frequency. Secondly, we lay the groundwork for a new and improved pointing template subtraction filter, built on a data-driven model of the ground, to replace the existing filter. These objectives are achieved by looking at both convolutions with a simulated telescope beam profile, and by creating ground pickup maps from actual COMAP data using a destriper mapmaking model.

6.1 Context and idea

6.1.1 Current ground model

Our current pipeline includes a filter for fitting and removing pointing correlated systematics. As we explored in section 4.4.2, this filter fits and subtracts a function on the form

$$d_{\text{point}} = \frac{\tau_0}{\sin \text{el}(t)} + a\,\text{az}(t) + c, \tag{6.1}$$

where $\text{el}(t)$ and $\text{az}(t)$ are the telescope pointing coordinates, and τ_0, a, and c are values we fit for. This model has a few obvious flaws when it comes to dealing with ground pickup. First of all, the ground is assumed to have a very simple model, with a linear dependence on both elevation and azimuth.[1] Exactly how poor this model is will be explored in more detail, but it is natural to assume that the mountain profile around the telescope does not perfectly follow a linear model. Secondly, the current model is employed entirely independently for each frequency channel of the data. While this, in theory, allows for the removal of complicated frequency profiles, it throws away significant constraining power, as it is likely that the ground systematics have a predictable frequency profile, that we can use to constrain the fit. We will therefore also look into the frequency dependence of the ground pickup in our analysis.

---

[1] If you are wondering where the linear elevation term is, it is a part of the first term. The fields have such a small elevation span that the $1/\sin \text{el}$ term is virtually linear, and the first term can be used to fit both the atmosphere and linear ground pickup.

6.1.2 Improved ground model

The ultimate goal of this analysis is the implementation of an improved ground model, based upon the az/el structure observed in actual COMAP data. Not only can this improve the quality of our current data, it might also allow us to loosen the elevation cut criteria of $35^\circ < \text{el} < 65^\circ$, giving us greater observational efficiency. As a replacement for the model outlined in equation 6.1, we propose a more general model, on the form

$$d_{\text{point}} = \frac{\tau_0}{\sin \text{el}(t)} + a\, A(\text{az}(t), \text{el}(t), \nu) + c, \tag{6.2}$$

where $A(\text{az}(t), \text{el}(t), \nu)$ is now some unknown function of pointing and frequency, representing the pickup from the ground (or other az/el correlated systematics), which we aim to create a data-driven model for. The ground is well approximated as a continuum source, but may still have a non-trivial frequency profile which depends on the relevant signal beam shape, which we will also look into.

6.2 Beam convolution map

We start our analysis by simulating the ground pickup. This is done by first considering the beam profile of the telescope, which we have plotted in figure 6.1. The beam profile consists of a *main beam*, at the very center of the pointing pattern. About 5 degrees from the center are the *near sidelobes*. These can pose a problem if they e.g. hit the sun, but are too close to the main beam to hit the ground at our observation elevations. The ground pickup relevant to us is caused by the *far sidelobes*, which are the four circles that peaks about 65° from the center. Note that the plot is logarithmic, so a lot of the structure contributes very little to the signal pickup (although it is very pretty). The far sidelobes are a problem because they both have significant power, and hit the ground, due to being so far from the pointing center. As the ground has a brightness temperature of $\approx 300\,\text{K}$, far more than e.g. the atmosphere, the signal contribution can become significant.

We create a map of the mountain range around the telescope from [49], where we assume that the mountains have a temperature of $300\,\text{K}$, and the sky of $0\,\text{K}$. We convolve the beam profile with the mountain map to simulate the ground pickup of the telescope. This is done as a 2D convolution of the sphere G, representing the mountain profile, and the sphere B, representing the beam. We can formulate this convolution as [50]

$$G'(\theta, \phi) = \int \mathrm{d}\Omega\, \hat{R}(\theta, \phi) B \cdot G, \tag{6.3}$$

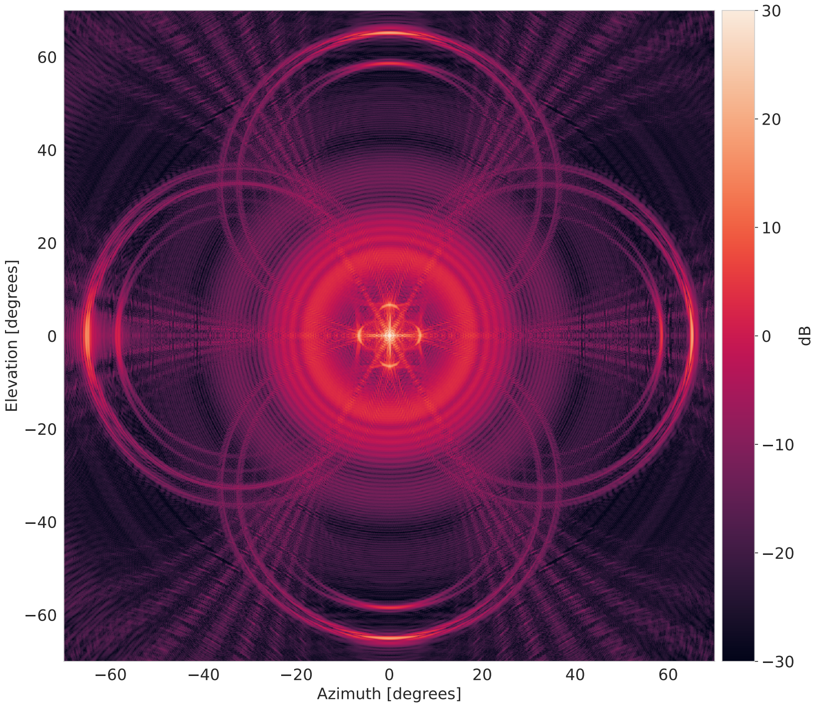


Figure 6.1: Simulated beam profile of feed 1 of the telescope at 30 GHz. The values represent the amount of signal the telescope will pick up at a given angle away from the central pointing location. The beam simulations were performed by James Lamb.

where $\hat{R}(\theta, \phi)$ is a rotation operator, with the effect of rotating B to be centered at (θ, ϕ). In practice, we represent both spheres as vectors in the 1D HEALPix [51] pixelization scheme, and performing rotations of the beam using a healpy [52] function.

Figure 6.2 shows the result of such a convolution, for the entire sky. COMAP currently only uses data from between 35° and 65° elevation, the reason for which is very clear from the figure, as it is the elevation span with the least ground contamination. We see that the regions close to the ground have very heavy pickup, as a result of the near sidelobes. We also see an imprint of the mountain profile at $70° - 80°$ elevation, as a result of the downward far sidelobe. A less prominent such imprint a few degrees lower is also present, due to the slightly closer far sidelobe, although is virtually invisible in this figure. Another interesting feature is the beams stretching up at around 340° $(-20°)$ and 130° azimuth. These are the result of the left and right far sidelobes hitting the largest mountain, and are the most prominent features to cross the 35° elevation line.

We will now look more closely at the regions covered by the COMAP fields. While figure 6.2 illustrates the large-scale effects of the ground, there does not seem to be much going on in the region relevant for COMAP. This is not the case and is because the color scale covers a range of several Kelvin, which is a lot in our context. It is

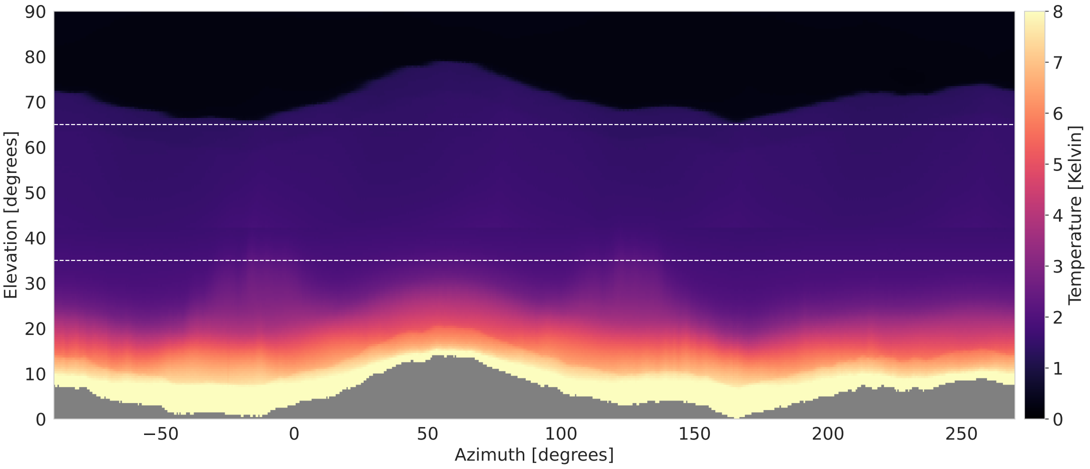


Figure 6.2: Predicted map of the ground pickup by pointing location, assuming a ground temperature of 300 K. The ground profile itself is shown in grey. The currently used scanning region, of 35° to 65° elevation, is shown as striped lines. The convolutions are based on code written by Duncan Watts.

also important to keep in mind that it is the change in the ground pickup in a scan that poses a problem. A constant addition to the signal will simply be removed by the normalization, while a gradient in the ground will not. We will therefore now take a look at the gradient of this map.

6.2.1 Gradient maps

We solve for the gradient of the map by applying a 5x5 Sobel kernel separately in the elevation and azimuth direction. This means that we convolve the image with a derivative kernel on the form

$$\mathbf{D}_x = \begin{pmatrix} -0.25 & -0.2 & 0 & 0.2 & 0.25 \\ -0.4 & -0.5 & 0 & 0.5 & 0.4 \\ -0.5 & -1. & 0 & 1. & 0.5 \\ -0.4 & -0.5 & 0 & 0.5 & 0.4 \\ -0.25 & -0.2 & 0 & 0.2 & 0.25 \end{pmatrix} \tag{6.4}$$

for the azimuth direction, and the corresponding kernel $\mathbf{D}_y = \mathbf{D}_x^T$ for the elevation direction. The kernel size of 5x5 pixels corresponds roughly to half the size of an observational field on the sky.

We will limit our analysis to the patches of the sky which we actually observe, which we illustrated all the way back in figure 3.4. Figures 6.3 and 6.4 shows the ground map derivative in the azimuth direction. To give some context to the maps, we have overplotted the azimuth amplitude parameter from the pointing filter, averaged over all available scans. As a reminder, this is the amplitude of the azimuth slope fit, such that

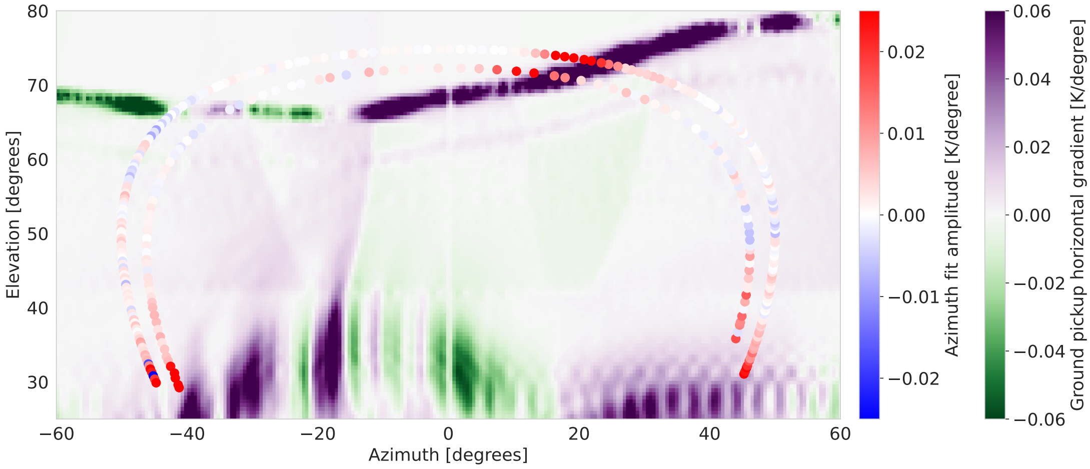


Figure 6.3: Azimuth derivative (background) of the ground pickup convolution for the area around the CO6 and CO7 fields. Overplotted are the mean amplitudes of the linear azimuth fit performed by the pointing template subtraction filter in the pipeline.

a larger amplitude means that we consistently see an azimuth dependency in this area of the sky. We have chosen to plot only the azimuth derivative of the ground convolution here, exactly because we happen to have a corresponding pipeline parameter to plot it against. If our ground convolution model is accurate, we expect to see a correlation between the parameter and the map. Such a correlation is very visible in the CO6/CO7 map, where we see a large positive amplitude at both at 70° elevation, and towards both lower ends of the field. An interesting disagreement between the map and the parameter is that the parameter has high values further up than the ground convolution gradient map does, both at 30° and 70° elevation. The values would have correlated better if the map was shifted by a few degrees, perhaps indicating that some beam features might be slightly off. A final interesting observation is that but parameter and map agree to the lack of a horizontal ground derivative in the upper left region of the map. We point this out because there is still a very noticeable transition in the vertical direction (see figure 6.2), but both the simulations and the pipeline data agree that there is no azimuth structure, which is a nice sanity check.

The CO2 map mainly shows azimuthal structure in the lower-left region. The azimuth amplitude fits are also larger in this region, but we do not see a great small-scale correlation between the two quantities. There are large negative azimuth amplitudes at about 140° azimuth, which does not immediately correlate with the map, but would so if the beam-line feature right below it was shifted up by a few degrees.

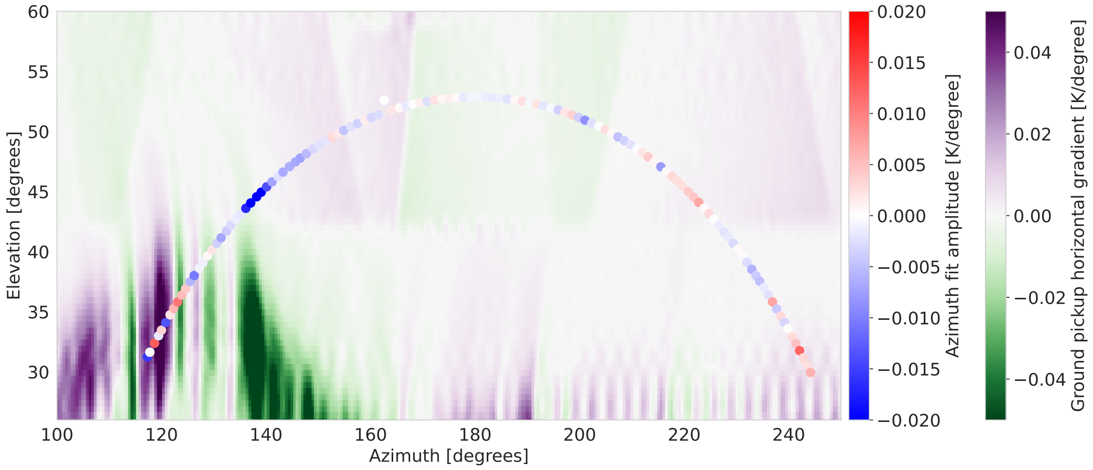


Figure 6.4: Same as figure 6.3 for the CO2 field.

The second comparison we want to make is with another metric from the pipeline, called `ps_chi2`. This value represents the χ^2 statistic from performing a white noise goodness of fit on the 3D power spectrum of the filtered data. If there are leftover systematics, such as badly modeled ground pickup in our scans after all filtering has been performed, this should show up as an excess in `ps_chi2`. Figure 6.5 shows this value overplotted on the total gradient amplitude[2] in both azimuth and elevation. As in the previous plots, we see the same excess in the upper-right corner of CO7, where the far sidelobe crosses the mountain horizon. CO6 meanwhile, has a more constant excess across its whole high-elevation region, which is not present in CO7. This lends even further support to our theory that the sidelobe mountain gradient, in reality, lies slightly higher than the beam convolution predicts and that CO6, therefore, keeps crossing this boundary during its high-elevation scans. While this explains some phenomena we observe, there are others still unexplained. There is a `ps_chi2` excess around 35° azimuth in CO6, with no corresponding excess in close proximity for CO7. There are also larger excesses at the bottom-left than bottom-right of both fields, while the beam convolution definitively claims there are larger ground gradients on the right. The lack of a large `ps_chi2` excess in some of these areas could also just mean that the ground in these areas are well modeled as linear. There is virtually no `ps_chi2` excess in the top-left part of CO7, as it crosses the horizon. As we have already established, this ground pickup transition is almost entirely in the elevation direction, with virtually no azimuth structure, which could help explain this. One of the shortcomings of the current pointing template subtraction model is that it fits for linear azimuth and elevation structures entirely independently, which is an especially bad model in places with ground gradients in both, such as the upper-right CO7 corner.

---

[2] There is no longer any reason to look only at the gradient in one direction, as badly fit ground in both elevation and azimuth should show up in `ps_chi2`.

Figure 6.6 shows a similar plot for the CO2 field. Interestingly, the 140° azimuth region in which we saw huge azimuth fit values does not contain the largest `ps_chi2` excess in CO2 (although there is definitively something there). This could be explained if the ground in this area, while substantial, is well modeled by our current filter. The most noise-contaminated regions of CO2 seem to be the lower left and upper right regions. The former makes a lot of sense, as it is the region of the largest predicted ground gradients in the whole area. The latter, however, does not seem to be explainable from the beam convolution.

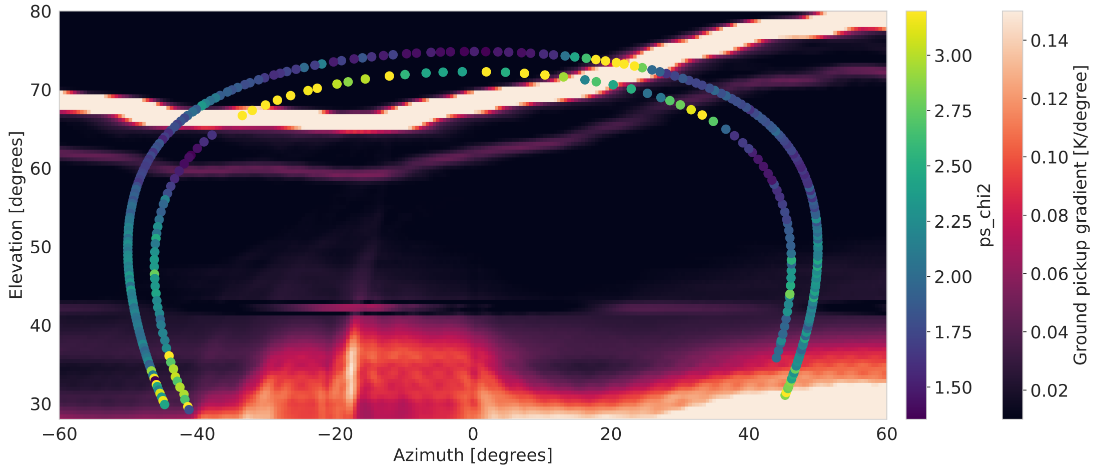


Figure 6.5: Gradient amplitude (background) of the ground pickup convolution for the area around the CO6 and CO7 fields. Overplotted are the mean amplitudes of `ps_chi2`, a noise residual metric from the pipeline.

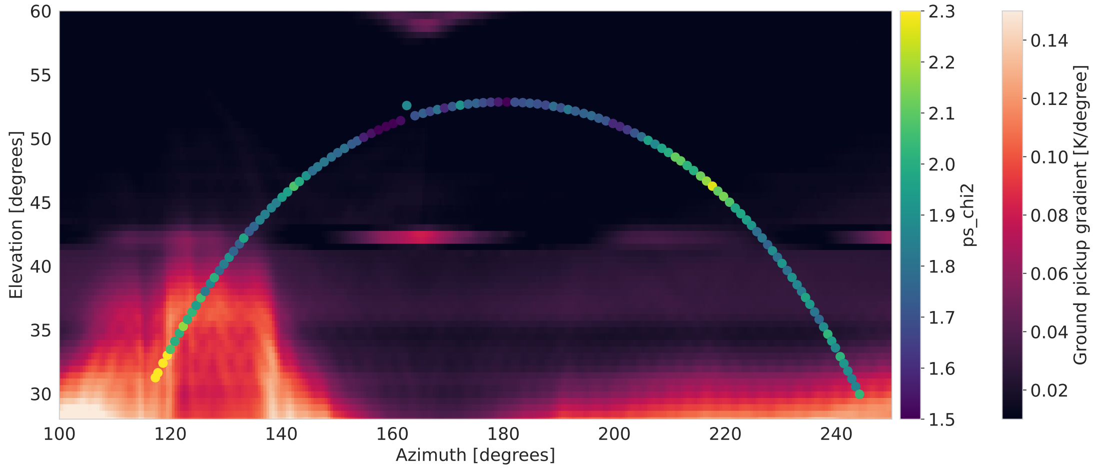


Figure 6.6: Same as figure 6.6 for the CO2 field.

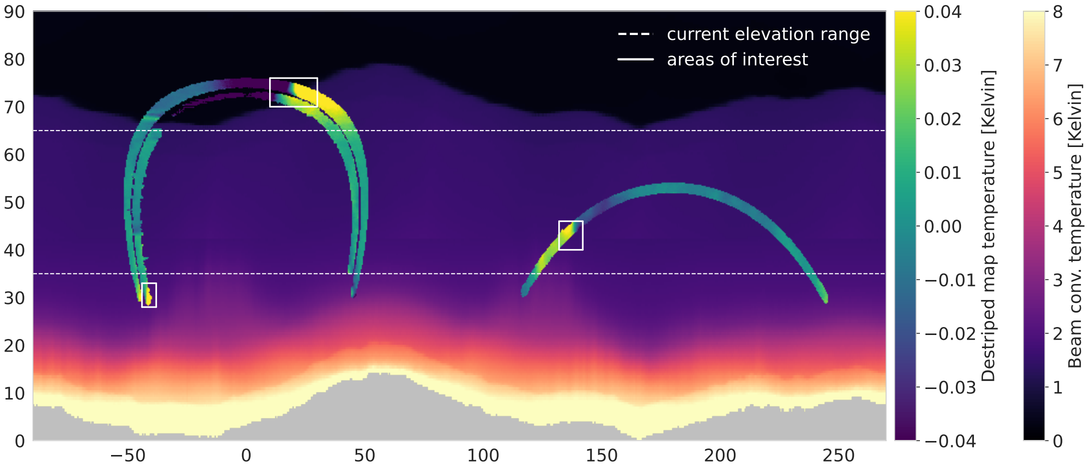


Figure 6.7: The same ground convolution simulation as seen in figure 6.2 (background). Overplotted is the temperature profile of the destriper maps generated from actual COMAP scans.

6.3 Destriped maps overview

We will now transition over to looking at maps created by applying the destriper on the scan dataset we outlined in section 5.4. These maps are intended to form the basis of a new and improved pointing template subtraction filter, and we will outline some of the ideas and challenges involved in such a method.

Figure 6.7 shows a destriper-made map of all three observational fields across the entire sky, overplotted on the beam convolution from the previous section. We see some clear temperature gradients in places we recognize from the previous section, such as the upper and lower regions of CO6 and CO7, and around 140° azimuth on CO2. We have marked three regions of especially large gradients, which we will study in more detail later. We show this full-field map mostly for illustrational purposes, as we do not entirely trust the map to be well constrained on such large scales. Remember that each scan is only $\sim 1^\circ \times 1^\circ$ across the sky, meaning we need several hundred such overlapping scans to reach all the way across an entire observational field. The destriper might struggle properly constraining the normalization of the scans on such large scales, and we will from now only make destriped maps of considerably smaller regions. This is fine, as it is the small-scale ground pickup gradients that impact our scans either way.

Before considering the implications of our data-driven maps, we want to take a short look at them in the context of the destriper mapmaking model we are using. As it turns out, for our use, the destriper does more than just remove correlated noise. Figure 6.8 shows two maps of the same upper-right region of CO7, created using the noise-weighted binning mapmaker, and the destriper mapmaker. The difference between the two is quite significant. First of all, we see that the striped features in the binned map is entirely gone in the destriped map, which looks much cleaner. However, the perhaps

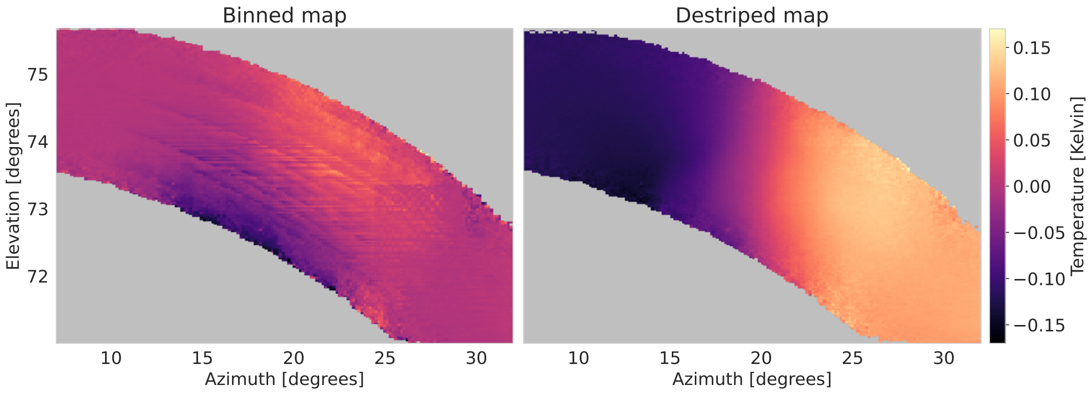


Figure 6.8: Left: Noise weighted binned maps from all available scans, as outlined in section 5.4.1, for the upper-right section of CO7. Right: Same map, created instead using a destriper mapmaker.

most surprising difference is the much larger overall gradient in the destriped map, The reason we do not see this gradient in the binned map is because we normalize each scan (which we have to, as we cannot trust absolute signal temperature, only relative temperature). This means that a scan in a lower signal region, e.g. the upper-left corner of the figure, and a scan in a higher signal region, e.g. the lower right part of the figure, will both be normalized to fluctuate around zero. We would need a scan to cover the entire map range to accurately extract the gradient of the ground. Since our scans are of size $\sim 1^\circ \times 1^\circ$, only gradients on this scale will show up in a binned map, and because we have many scans stacked beside each other, they will overlap and cancel each other out. Looking more closely at the binned map, we see that the upper-right arch is brighter, and the lower-right is dimmer. This is because the scans that cross the heavy gradient in the middle of the map will be much brighter towards higher azimuth, and dimmer towards lower azimuth. However, the part of the scan that overlaps with lots of other scans, who themselves see no gradient, will be washed out. We therefore only see hints of the gradient at the very edges of the binned map.

So how does the destriper retrieve the large-scale structure, which the binned scheme lost due to normalization? The problem is that each scan should contain some overall normalization factor, which is unknown to us. The whole point of the destriper is to add a constant offset to intervals of the TOD, to minimize the total variance in data hitting each pixel. This offset can very well be raised or lowered on entire scans. The scan crossing back and forth between the bright and the dim regions in the middle of the figure can now constrain the pixels on each side, insisting to all the scans on the right that they are in fact very bright, and the scans to its left that they are dimmer. This shows that, for producing maps from normalized scans, the destriper not only helps us remove short and medium timescale correlated noise, in the form of the stripes we see in the binned map, but also helps constrain the overall normalization of each scan. The destriper of course relies on good cross-linking between scans, with scans covering different patches of the sky, and overlapping to varying degrees. This is very much the case for our scans, which are centered around a range of different coordinates, and contain both Lissajous and CES scans.

6.4 Data driven pointing filter

Let us now look at some destriped maps in the context of actual scans. We wish to assess whether a ground map modeled pointing template might outperform the linear ones we already employ. For this purpose, we will look at regions with strong, preferably non-linear ground features. Let us start off with the same region we just explored, at the top-right of the CO7 field. Figure 6.9 shows the destriped map for this region, together with the average temperature as a function of azimuth only. The azimuth profile of the map shows fairly constant ground profiles at $\mathrm{az} < 13^\circ$ and $\mathrm{az} > 28^\circ$ with an almost linear region in between. However, in the transition between the constant and linear regions, the profile would be badly approximated by a linear fit. We have overplotted the pointing of two scans, specifically selected to fall within these regions.

The ultimate purpose of these ground maps is to serve as the function $A(\mathrm{el}, \mathrm{az})$ in a better pointing template, as outlined by equation 6.2. In figure 6.10 we have plotted the TOD for both scans from the map figure. The linear azimuth fit of the TOD is also plotted, as well as the destriped map amplitude (normalized to 0) at the scan pointing coordinates. The fact that the map amplitude is so similar to the azimuth fit in both overall amplitude and shape is very encouraging, although the map amplitude seems somewhat lower in scan 2. Note that the red fit in these plots is just the raw map temperature at the pointing coordinates of the scan, not in any way fitted to the relevant TOD. Apart from being similar in amplitude, the map fit shows clear deviations from the linear azimuth fit, and is also not constrained to a repeating pattern, the way the linear fit is. This is because the map can have an arbitrarily complicated azimuth and elevation structure. It is important to remember, however, that it is still a one-parameter model, just like the linear azimuth fit. This is a crucial advantage of the new method, as increasing the number of free parameters in the model can introduce other problems, like a larger loss of signal.

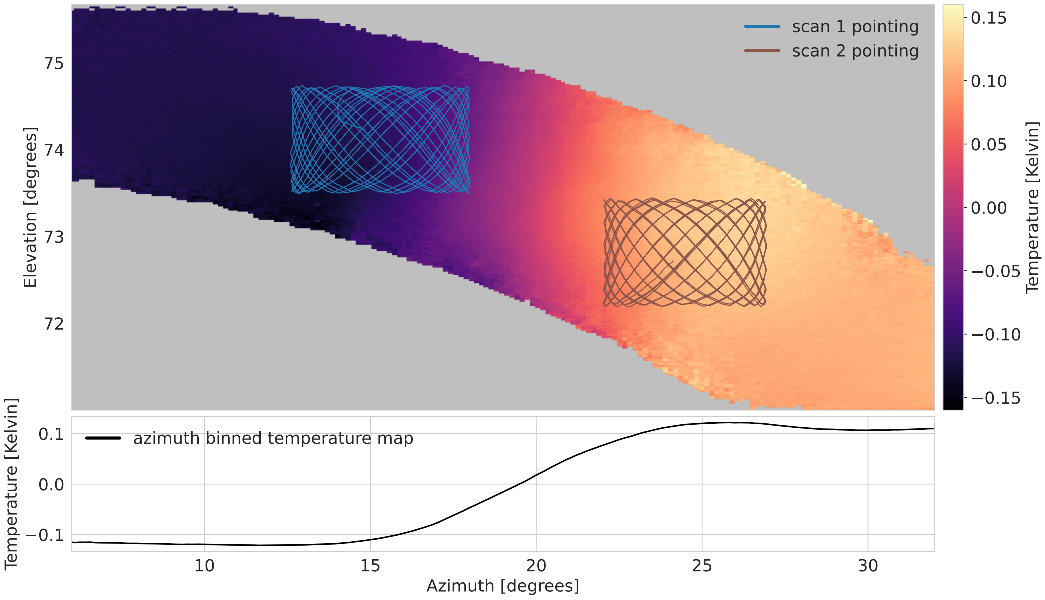


Figure 6.9: Top: Destriped map of the upper-right region of CO7, with the pointing pattern of two selected scans overplotted. Bottom: The azimuth temperature structure of the destriped map, averaged over elevations.

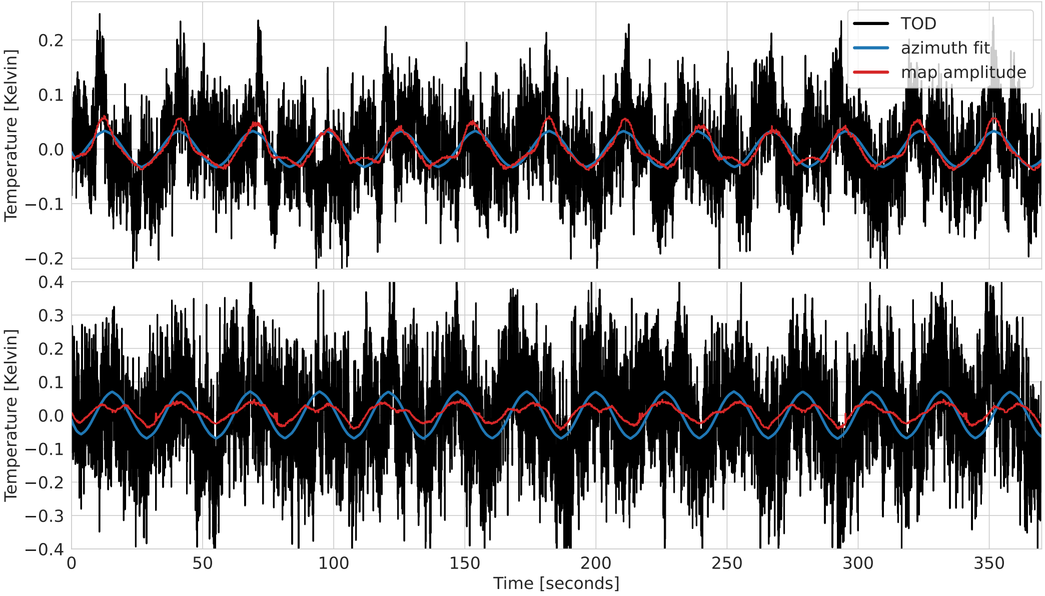


Figure 6.10: Sideband averaged TOD for scan 1 (top) and scan 2 (bottom) from figure 6.9. The temperature profile from the destriped ground map at the scan pointing pattern is shown in red, and the best fit linear azimuth profile is shown in blue.

Moving onto an area of even more interest, the 140° azimuth region on CO2. This is the region within our observational cut of 35° to 65° elevation with the most significant ground gradients. Figure 6.11 shows the destriped map of this region, also with two selected scans. We see a strong gradient in the map, with a 0.1 K change in ground pickup over a few degrees. This is in good agreement with what we saw from the mean linear azimuth amplitudes from figure 6.4, and in less agreement with the ground convolution map, in which this feature seemed lacking or displaced. This region of CO2 seems to be more linear in frequency than the region of CO7 we looked at did, as we see from the bottom of figure 6.11. This could help explain why we saw no noticeable `ps_chi2` excess when studying the area back in figure 6.6, since the linear azimuth fit would leave less excess systematics if the ground profile was well approximated as linear.

Figure 6.12 shows the two scans we overplotted on the CO2 field, together with the best fit linear azimuth template, and the map amplitude. In scan 1, there is good agreement between the linear azimuth fit and the destriped map amplitude. Scan 2 shows a potential pitfall of a data-driven pointing template model if the model were to be wrong. Here, we see that the destriper map actually entirely disagrees with the scan about which direction the ground varies in azimuth. The two profiles look, interestingly enough, rather similar, but they are anti-symmetric. If we were to fit this as a function $A(\mathrm{az}, \mathrm{el})$ to the scan, a negative amplitude would obviously tell us that something was amiss. We would expect to see amplitudes around a value of 1, which could be used as a sanity check to how effective the fit was.

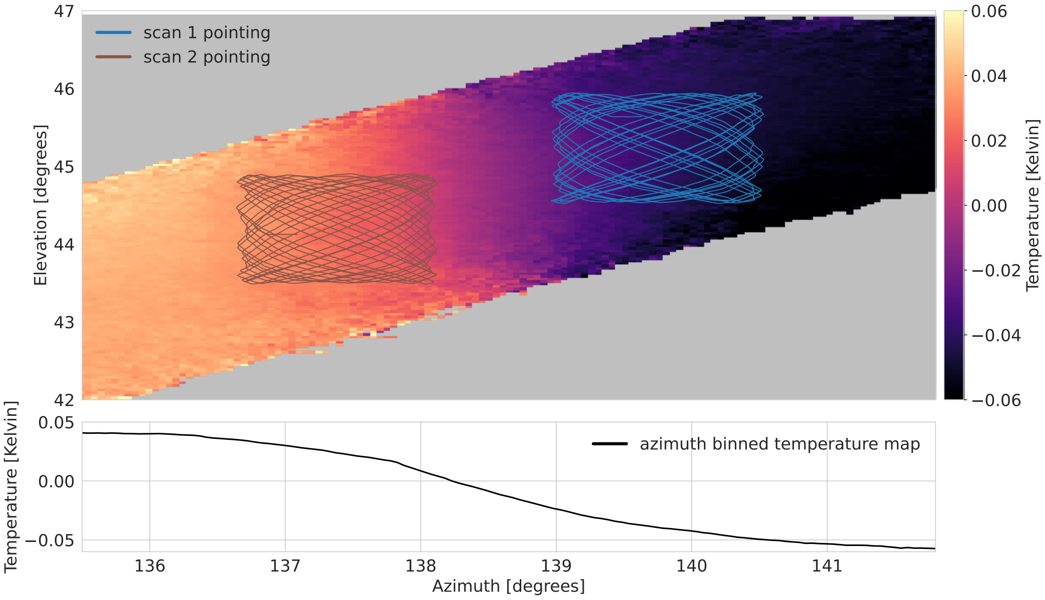


Figure 6.11: Top: Destriped map of the middle-left region of CO2, with the pointing pattern of two selected scans overplotted. Bottom: The azimuth temperature structure of the destriped map, averaged over elevations.

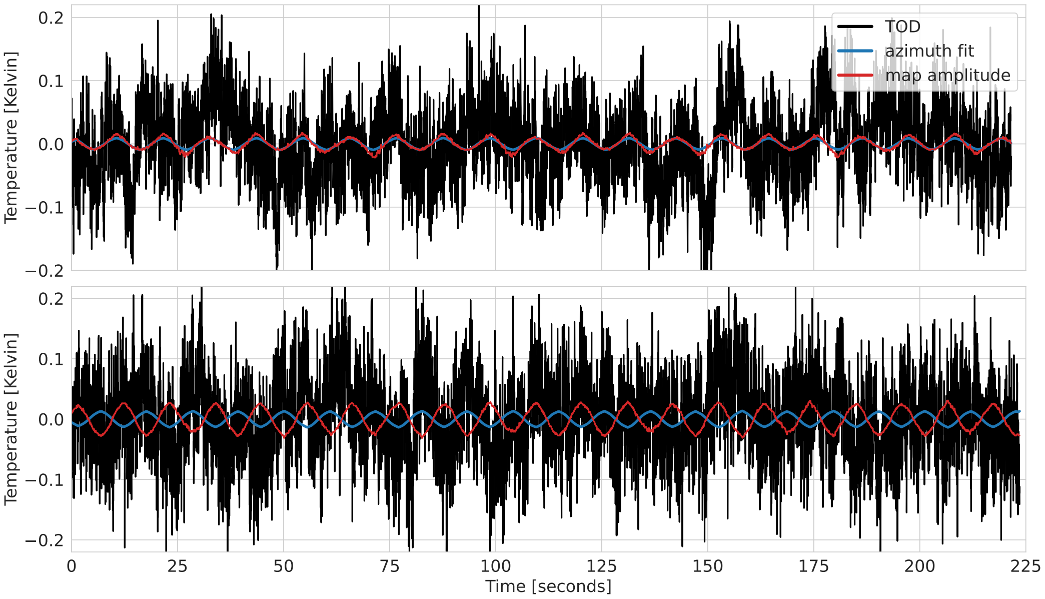


Figure 6.12: Sideband averaged TOD for scan 1 (top) and scan 2 (bottom) from figure 6.9. The temperature profile from the destriped ground map at the scan pointing pattern is shown in red, and the best fit linear azimuth profile is shown in blue.

6.5 Frequency profile analysis

We will finish off our ground analysis by taking a look at the frequency dependence of the ground pickup. The ground itself is more or less a blackbody, and has an approximately constant brightness temperature in frequency. We do, however, not observe the ground directly, but rather from the far sidelobes, and these do have a frequency structure. In the current version of the pipeline, the frequency dependence of the ground pickup is not very important, as we perform the pointing template fit independently for each frequency. If we, however, want to leverage frequency information in a new pointing template model, we need an understanding of how the far sidelobes behave in frequency. We have the opportunity to compare both simulated and data-based maps across frequencies. The beam profile has been simulated for 26, 30, and 34 GHz, which represent the mean and outer edges of our frequency range. Our data-based destriper maps use sideband averaged signal, meaning we have four maps in frequencies, at mean frequencies of 27, 29, 31, and 33 GHz.

In the case of the main beam, higher frequencies produce a sharper beam profile, which is higher in amplitude at the center and falls off quicker from the beam center. This relation is described by the diffraction limit, which (among other things) states that the size of the main beam is inversely proportional to the frequency of the signal. For the rest of the beam profile, things get a bit more complicated. Figure 6.13 shows the simulated beam power, sliced vertically down from the main beam, for three different frequencies. There are some important both differences and similarities between these profiles. First of all, the sidelobe peaks are all in the exact same locations for all frequencies. They are, however, varying in amplitude. Especially the most important sidelobe for ground pickup at $-65°$ elevation, is about 5 dB higher in the 34 GHz band than the 26 GHz, while the 30 GHz band falls in the middle. While this peak itself seems like it might follow some simple frequency trend, the rest of the beam power shows an incredibly complex and unpredictable profile. In some regions, it is even the middle 30 GHz frequency band that holds the most power. There is little reason to suspect that the resulting frequency profile from convolving this sort of profile over a mountain range will be simple. This graph alone also gives far from the whole picture, as we are only looking at a vertical slice of the beam profile.

Figure 6.14 shows the difference between the power in the simulated 34 and 26 GHz beams for the entire beam profile. We see an extremely complicated frequency dependence in the power, which could be very complicated to accurately model, with different portions of this beam structure passing in and out of the mountain profile. The center of the far sidelobes has a positive frequency derivative, but there are several nearby points where it is negative. This is also just the average derivative between the two outermost points of our beam, and as we saw from the inclusion of the 30 GHz profile in figure 6.13, the structure within those two outermost points does not always appear to be simple or linear.

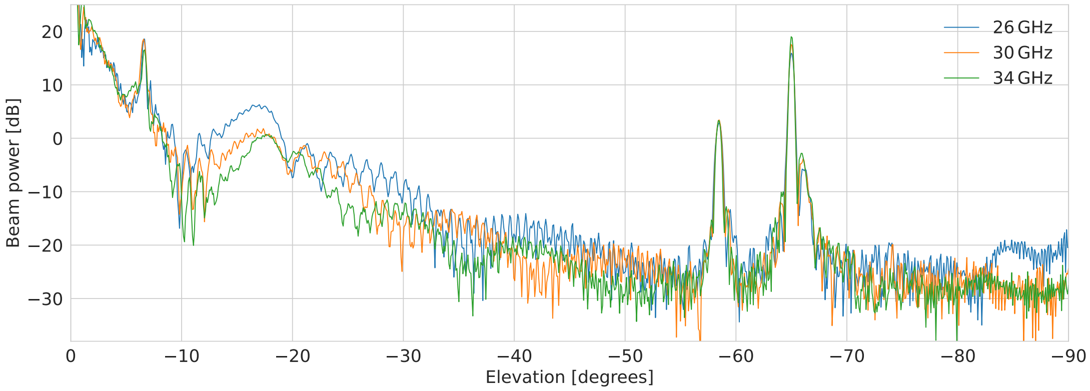


Figure 6.13: Relative power of the telescope beam as function of elevation, vertically down from the beam center, simulated at frequencies of 26 GHz, 30 GHz, and 34 GHz.

Figure 6.15 shows the simulated frequency map derivative (in the background) and the frequency derivative of destriped maps, for two of the regions we previously looked at. Let us start off by looking at the simulated ground convolution. At the very top, we have an area with a negative frequency derivative of $-0.05\,\mathrm{K/GHz}$, where only the outermost parts of the beam contribute. As the far sidelobes hit the mountain, the derivative sharply shifts to being approximately zero. This confirms what we saw in figure 6.13, namely that the far sidelobe is stronger for higher frequencies. In the intermediate elevation region, where most of our observations happen, there are much more modest features, but some small fluctuations can be seen. Around 20° above the mountain range, the inner, low-frequency dominated features of the beam (see figure 6.14) move into the mountains, and see a strong negative derivative. As we hit the mountain range itself, this shifts entirely to be dominated by higher frequencies, as more of the power is concentrated around the main beam in higher frequencies.

We have also overplotted two frequency derivative destriper maps, one for the upper-right region of CO7, and one for the center-left part of CO2. In the CO7 plot, we see the same transition from negative to positive gradient as the far sidelobe moves into the mountain range. This is another useful sanity check for both our simulations and data-based maps, although there are some differences. The destriped map shows a change in gradient by about $0.01\,\mathrm{K/GHz}$ across the transition, while this value is around $0.05\,\mathrm{K/GHz}$ for the simulated map, quite a bit stronger. As a reminder, there is no absolute temperature information in the destriped map, and all maps have been normalized to zero-mean, so it only makes sense to compare relative values on the map. We, therefore, compare only how much the temperature frequency gradient changes over certain areas. Secondly, the transition happens at a couple of degrees higher elevation and lower azimuth than the convolution map predicts. This actually conforms well with the discrepancies we previously saw from these maps, where the data also seemed to suggest that the far sidelobe transition should have happened at higher elevation (figure 6.5).

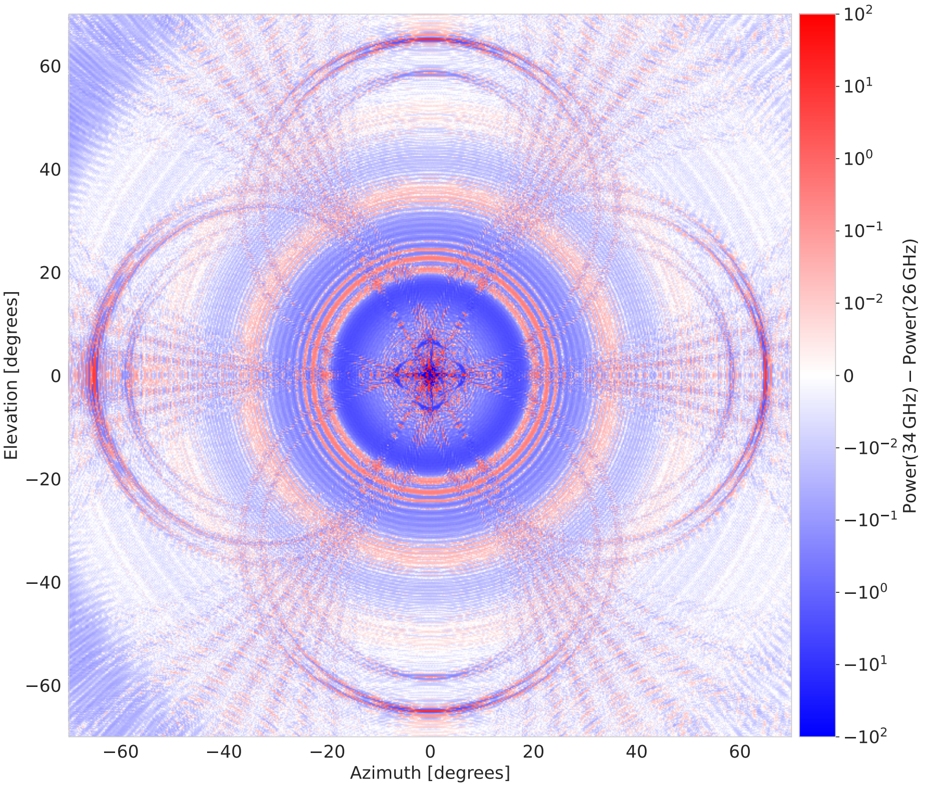


Figure 6.14: Power difference between the simulated 34 GHz and 26 GHz beam profiles.

The CO2 plot also predicts a sharp change in the temperature frequency derivative across the simulated map. This change is in no way predicted by the simulated convolution map, which has no noticeable features at that exact location. There is, however, a feature 5-10 degrees below it, which seems to show a similarly sharp transition. We know from our previous analysis that the observed features of this region of CO2 were hard to explain with the convolution map, and it, therefore, comes as no huge surprise that the frequency profile also looks different in this area. Note that we have only looked at the signal frequency derivative over a substantial frequency range, and there are probably more complicated small-scale fluctuations, as figure 6.13 hints to.

6.6 Ground pickup summary and discussion

In this chapter, we have laid the groundwork for a new and improved ground template model, based on a data-driven destriped map. We will now quickly summarize the findings, and foreseen challenges in such a model.

The ground maps we have produced employ all available feeds, and use sideband averaged maps, for maximum signal to noise. In the future, we want to perform an analysis of to what degree the ground pickup depends on feed position, and whether it is feasible

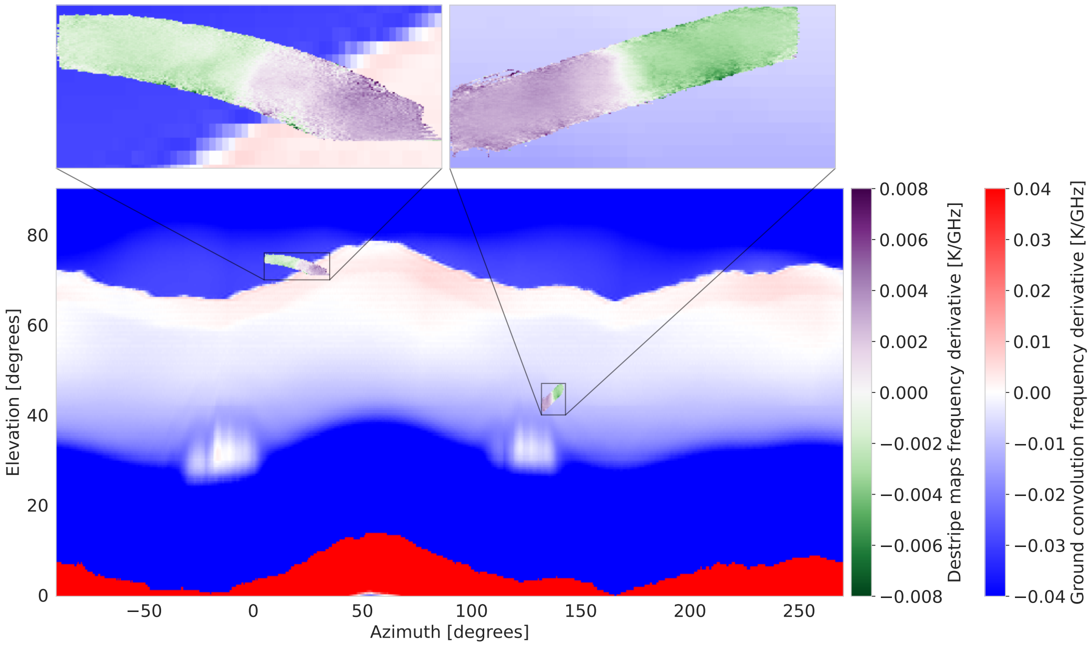


Figure 6.15: Frequency derivative, as calculated by the difference between the simulated 34 GHz and 26 GHz convolution maps (background). Overplotted are frequency derivative maps calculated from the difference between the last and first sideband of the data-driven destriper maps, for two selected regions of CO2 and CO7. Zoomed versions of the regions of the relevant maps are shown above.

to use the same ground map for all feeds. In the ground convolution analysis at the beginning of the chapter, we used the simulated beam profile specifically of feed 1, which is the center feed of the feed array.[3] The far sidelobe features of the beam depend on the physical placement of each feed. It is unknown if this effect is large enough to constitute an entirely independent ground pickup map for each individual feed, and to what degree this discrepancy will need to be taken into account when producing the ground maps. Producing individual ground maps for each feed would reduce the available number of scans by a factor of 19, in theory reducing the signal to noise by a factor of $\sqrt{19} \approx 4.4$. It could also impact the level of cross-linking in the maps.

Producing individual frequency maps would reduce the signal to noise by a factor of $\sqrt{1024} = 32$. As we saw from the analysis of section 6.5, the complicated frequency structure of the beam profile again produces an even more complicated frequency structure in the actual maps. Our findings suggest that the ground pickup may have structures in both frequency and pointing within a single scan-sized patch of the sky. This poses a problem, as a single-parameter sideband averaged template would not be sufficient to model variances in both pointing and frequency within the same scan, without intimate knowledge of the frequency structure.

---

[3] Figure 7.1 in the next chapter shows the physical arrangement of the feeds in the feed array.

Chapter 7

Calibration

In this chapter, we present an analysis of the calibration phase of an observation, where a calibration vane covers the feeds of the telescope to measure a reference hot load power and temperature. We study how the angle of this vane relates to the power of the different feeds, and look at what angles provide acceptable calibration results. Additionally, we outline a new method for extracting the hot load power and temperature from a calibration, as the current method sometimes fails.

7.1 Calibration properties

During a calibration observation, a vane of known temperature is placed in the field of view of all telescope feeds, blocking other incoming signals. The purpose of this calibration is for each feed to measure a hot load power P_{hot} and temperature T_{hot}, as we outlined in section 4.3. Being of approximate outdoor temperature, the brightness temperature of the vane is usually an order of magnitude brighter than the sky, meaning that we see a sharp power increase in the feeds as the vane covers them. The position of the calibration vane is denoted by an angle, which is around 215 degrees when fully disengaged, and lowers to around 65 degrees when fully covering the feeds. The feeds physical arrangement in the cryostat is shown in figure 7.1, and there is a 26 cm spacing between completely opposite feeds. This means that the vane will cover some feeds before others when moving. Figure 7.2 illustrates this relation, where we see the power response of the feeds as a function of the vane angle. Some feeds are fully covered by the vane already at 120 degrees, while feed 8 does not reach 99% power until 68.9 degrees. This is very close to the lowest angle state of the vane, which we have observed to vary anywhere from 65.5 to 66.8 degrees from calibration to calibration. With this, we can conclude that hot load measurements can safely be taken at an angle of 69 degrees, with less than 1% signal loss.

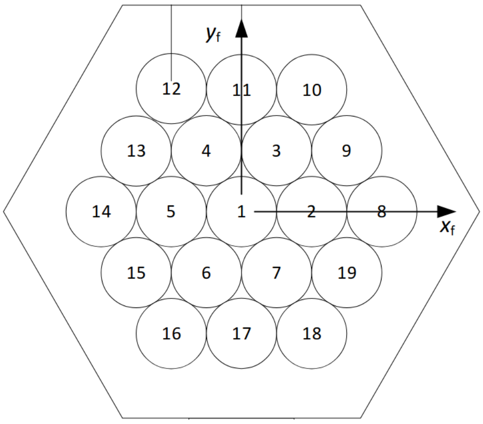


Figure 7.1: Physical arrangement of the telescope feeds. The calibration vane enters the feeds field of view from the lower-left corner of this image. Figure credit: James Lamb.

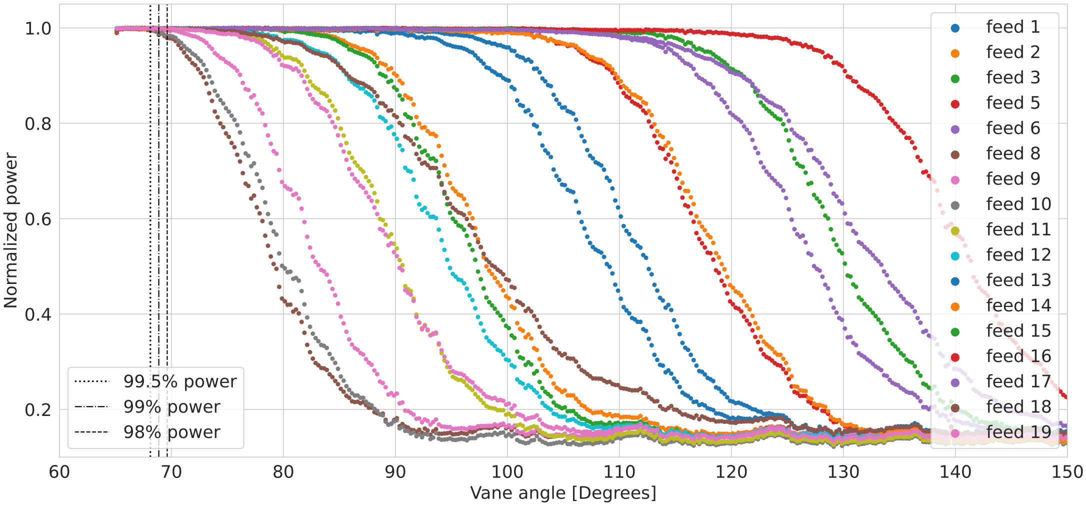


Figure 7.2: Relation between the observed power of each feed and the angle of the calibration vane. The values are an average of over 7000 calibrations. The angle at which the first feed loses 0.5%, 1%, and 2% power is shown as striped lines, and happen at 68.1, 68.9 and 69.6 degrees, respectively.

7.2 A new calibration technique

We have traditionally calculated values for P_{hot} and T_{hot} averaging the TOD over a section marked out by a flag in the level 1 files. The flag indicates that the calibration vane fully covers the observational array, and the hot load measurements can safely take place. This flag has turned out to be somewhat unreliable, often not triggering, probably because the calibration vane did not move to an angle deemed acceptable. However, in many such cases, the calibration data is still fully usable. For this reason, we have implemented an alternative technique for calculating P_{hot} and T_{hot}, ignoring the flag and instead looking at the data itself to figure out when a calibration is underway.

Figure 7.3 shows an example of the power response for a handful of feeds over a single calibration, together with the vane angle. As we have already seen, the different feeds increase in power at very different angles, with some requiring the vane to be almost completely descended. We have also marked the interval flagged for hot load measurements in the level 1 file. As we can see, it is a relatively short interval of 0.6 seconds. This is not much of a problem, as the hot load power is very stable in time, but it does not hurt to use more datapoints in the calibration if we can. As mentioned, sometimes the flag does not trigger, even if the calibration was otherwise successful.

We, therefore, propose a simple alternative scheme for performing the hot load measurements in the absence of this flag, which can be summarized as follows.

- Identify the part of the TOD marked for calibration. The level 1 files have a separate flag for indicating that a calibration is underway, which we use to approximate the location of the calibration.
- Find the point of highest power within this regime.
- From this point of highest power, move outwards in both directions until the power has dropped by 5%. Add a 1-second safety margin to this, and mark those points as the beginning and end of the calibration measurements.
- Calculate P_{hot} and T_{hot} by averaging over this range.

We then perform a couple of sanity checks to confirm that the calibration was indeed successful, as all of these steps could very well have been performed on nonsensical data. First of all, we check that the vane reached an angle of less than 69 degrees, which we found to be a natural limit in the previous section, and that this happened during the region we marked for calibration. Secondly, we check that the calculated P_{hot} value is more than twice the power of the scan mean power, and the interval used for calculating these values exceed 1 second. This ensures that we actually observe a consistent power increase in the data, regardless of what the vane angle and flag claims. If these tests are not all successful, the calibration is discarded. Note that each feed will now have individual calibration durations, as the power falls off differently in each feed. In theory, the uncertainty in both P_{hot} and T_{hot} will be lower, as it is calculated from the maximum available number of points. The uncertainties were, however, already very low.

Figure 7.4 demonstrates this new scheme, where we show both the measured power, the vane temperature, and their averages. We have also marked the interval over which we average the measurements and their resulting values.

This new calibration scheme was run on all available level 1 files and collected in a database for later usage. `l2gen` now features an option to employ these calibration values instead of the original calculation. We have also employed this database in the calibration of the scans we prepared for the destriper ground pickup maps, as we described in section 5.4.1.

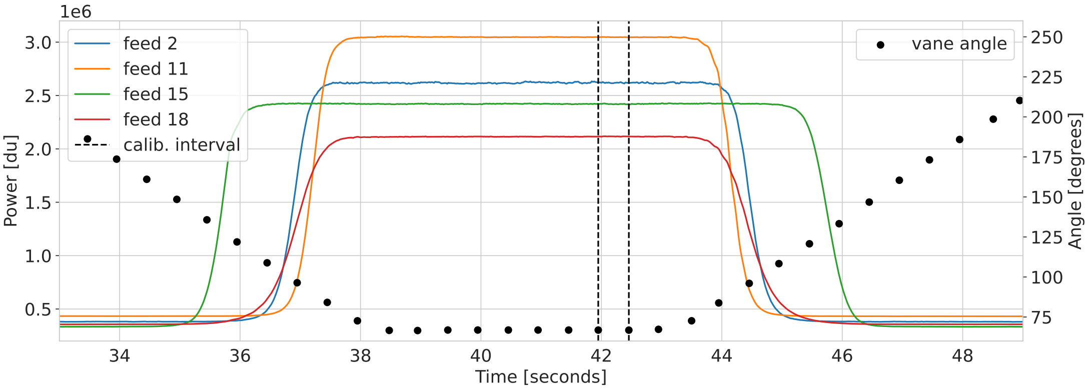


Figure 7.3: Power curves of the TOD from four selected feeds during a calibration. The angle of the calibration vane is shown as black dots, while the striped black lines shows the interval in which we calculate the hot load power.

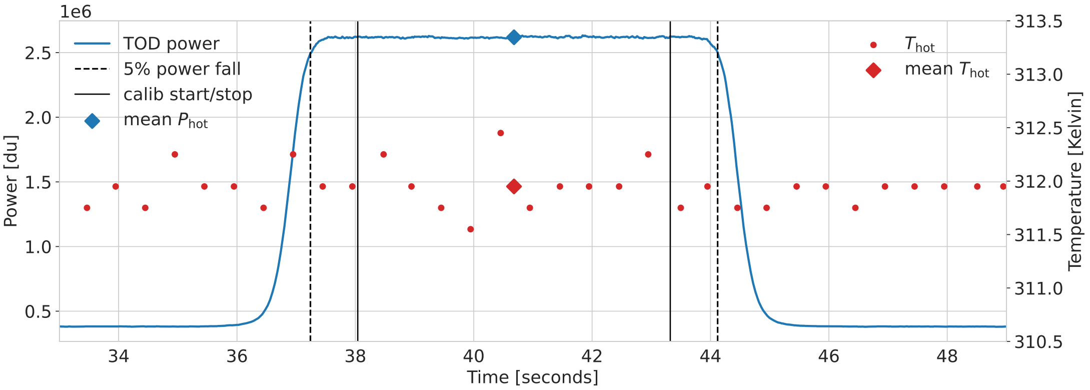


Figure 7.4: Demonstration of new hot load calculations on feed 2. The vertical black lines show the points of 5% power fall, and the added 1-second margin, which constitutes the beginning and end of a hot load measurement. The red dots show the vane temperature. The temperature and hot load power are averaged over the interval marked by solid black lines, and the average values of both quantities are shown as diamonds.

Chapter 8

Noise Characterization

The effective removal of systematics in our data requires a thorough understanding of how they behave and effective ways of measuring their presence or absence. The distinction between different types of systematics, and their respective risk to our experiment is also important. In this chapter, we present an analysis of the COMAP data as they move through the pipeline, with a special focus on noise characterization. The CO signal itself is too weak to have any meaningful impact on the characteristics of the TOD on a single scan. The filtered scans will also not be meaningfully correlated across feeds, or on larger frequency or timescales, because these modes are suppressed by the filtering process. This means that we desire our filtered signal to be as consistent with white noise as possible, and as uncorrelated across time, feeds, and frequencies as possible. Before we look at the COMAP data, we present a short analysis of common noise characteristics.

8.1 Noise theory

We gave a brief summary of the different noise sources in the COMAP data back in section 3.5, and will now quickly revisit this topic, after having been more thoroughly introduced to the COMAP data and analysis pipeline.

8.1.1 White noise

The simplest and most fundamental noise in any experiment is white noise, which is uncorrelated noise sampled from a Gaussian distribution. An example of white noise, and its temporal power spectrum (which is flat) is shown in the top panels of figure 8.1. Even under perfect observational conditions, white noise poses a theoretical limit on the sensitivity of a radio receiver. This limitation is given by the Radiometer Equation, which states that the standard deviation in the observed signal due to temperature white noise is [53]

$$\sigma_0 = \frac{T_{\rm sys}}{\sqrt{\Delta\nu\tau}}, \tag{8.1}$$

where $T_{\rm sys}$ is the system temperature of the receiver, $\Delta\nu$ is the frequency range over which we observe, and τ is the observation time. The COMAP data is currently sampled at intervals of $\tau = 20\,\mathrm{ms}$, with frequency bins of size $\nu = 1.953\,\mathrm{MHz}$. The radiometer equation predicts a white noise level in our raw data of $\sigma_0 = 0.00506 T_{\rm sys}$, or $\sigma_0 = 0.228\,\mathrm{K}$ for $T_{\rm sys} = 45\,\mathrm{K}$.

We can also turn this equation on its head, and ask what sort of observational time we would require to detect a CO signal. In our analysis data, we use decimated frequency bins of $\Delta\nu = 31.25\,\mathrm{MHz}$. If we require a white noise level on the same order of magnitude as the expected CO signal brightness [34] $\sigma_0 = 1\,\mu\mathrm{K}$, we get an observational time of

$$\tau = \left(\frac{T_{\rm sys}}{\sigma_0}\right)^2 \frac{1}{\Delta\nu} = 6.48 \times 10^7\,\mathrm{s} \approx 2\,\mathrm{yrs}. \tag{8.2}$$

The radiometer equation provides a nice theoretical basis for our expected white noise level, but we also desire a way of calculating this direction from a TOD. We cannot simply look at the standard deviation of the data itself, since that will contain a contribution from temporally correlated noise as well. However, if we only look at the standard deviation between neighboring points in the TOD, there will be no contribution from the correlated noise. We calculate the white noise level in our data as [34]

$$\sigma_0 = \sqrt{\frac{\mathrm{Var}(d_i - d_{i-1})}{2}}. \tag{8.3}$$

8.1.2 1/f noise

One of the most common types of temporally correlated noise is so-called 1/f noise. This is noise which has a power spectrum proportional to $1/f^\alpha$, where f is the temporal frequency, and α is some positive value. 1/f noise has been shown to occur naturally in everything from music [54] to sea tides [55], but most importantly for us, it occurs in semiconductor electronics [42] such as the low noise amplifiers employed by the telescope. This means a substantial amount of our temporally correlated noise can be modeled as 1/f noise. The power spectrum of 1/f noise can be written as

$$P(f) = \sigma_0^2\left[1 + \left(\frac{f}{f_{\rm knee}}\right)^\alpha\right] \tag{8.4}$$

where σ_0 is the white noise level, and $f_{\rm knee}$ is the knee frequency, which represents the transition point between correlated noise and uncorrelated noise dominated regimes, and α is the correlated noise slope. Figure 8.1 shows an example of simulated 1/f noise, with a white noise term.

Some signals are completely dominated by the correlated noise, such that the power spectrum never really flattens out. In such cases, we neglect the white noise contribution

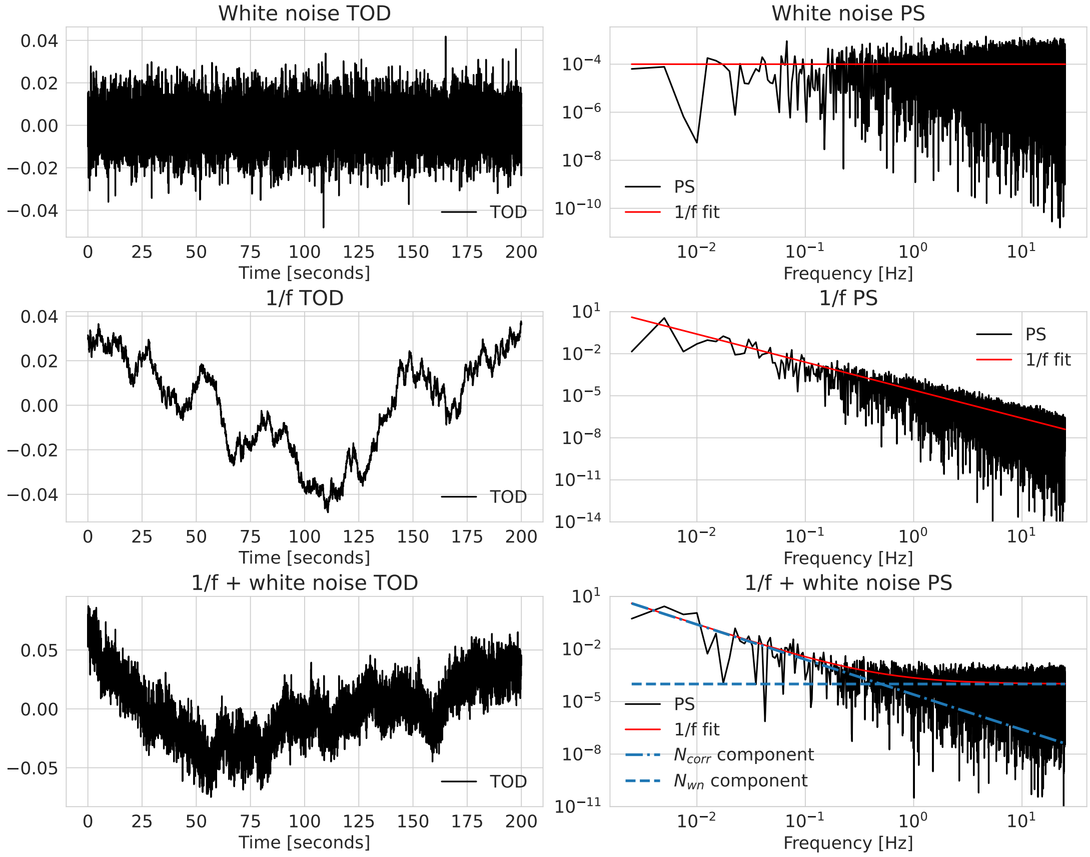


Figure 8.1: Figure showing the TOD (left) and power spectrum (right) of a generated 1/f noise dataset, with $\sigma_0 = 0.01$, $f_{\text{knee}} = 0.5\,\text{Hz}$ and $\alpha = -2.0$. The top plot shows the white noise part of the 1/f spectra only, while the middle plot shows the sloped part only. The bottom plot shows the combined signal.

and write the $1/f$ power spectrum simply as

$$P(f) = \sigma_0^2 \left(\frac{f}{f_{\text{knee}}} \right)^{\alpha} \tag{8.5}$$

It might not always be obvious which of these two models is being applied when simply referring to $1/f$ noise. We will here try to refer to these as $1/f$ noise with and without a white noise term, respectively.

When doing a best-fit of the 1/f PS parameters, we can either fit for the white noise level σ_0 together with the two other parameters, or fix it beforehand using either the radiometer equation 8.1 or equation 8.3. We usually prefer to fix σ_0, to reduce the number of free parameters. α and f_{knee} are then found by doing a least squared fit to the power spectrum. In the case of a 1/f spectrum on the form 8.5, the white noise can not be calculated this way, but it is also completely degenerate with f_{knee}, so we can simply fix σ_0 to something reasonable, e.g. using the radiometer equation.

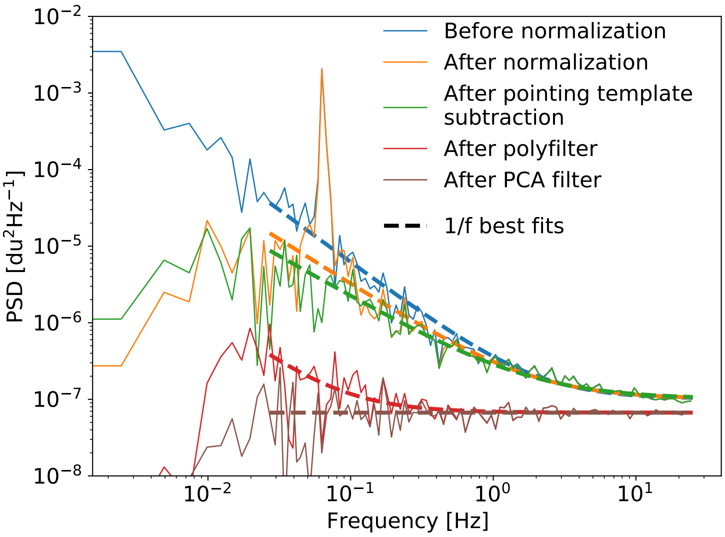


Figure 8.2: Power spectrum of a TOD at different stages of the `l2gen` pipeline. The respective $1/f$ best fit power spectra are overplotted in striped lines.

8.2 Data power spectrum and noise level

For ideal white noise-dominated data, the output power spectrum should be flat, as this is the shape of a white noise power spectrum. Figure 8.2 shows the power spectrum of the TOD for a typical Lissajous scan after each filter in the `l2gen` pipeline, together with best-fit $1/f$ spectra to each. Each filter has a clear effect on the power spectrum. The normalization heavily suppresses low-frequency modes. The pointing template subtraction knocks out the excess around $\sim 0.1\,\mathrm{Hz}$, due to the periodic pointing in elevation. The polynomial filter heavily suppresses the $1/f$ spectrum over the whole frequency spectrum. Finally, the PCA filter further suppresses it. Two important takeaways from this figure are that 1: Apart from specific, known deviations, the power spectrum is well approximated by a $1/f$ fit at every step of the pipeline, and 2: By the end of the pipeline, the power spectrum is very flat, meaning it is close to white noise.

A quantitative way of testing the removal of correlated noise is through a χ^2 goodness-of-fit test. For each scan, we define a χ^2 statistic on the form

$$\chi^2 = \frac{\sum_{i=0}^{N} \left(\frac{d_i}{\sigma_0}\right)^2 - N}{\sqrt{2N}}, \tag{8.6}$$

where d_i are the N datapoints of the filtered TOD of a the scan, and σ_0 is the white noise level, set by the radiometer equation 8.1. Under the hypothesis that the data are pure white noise, $\sum_{i=0}^{N} \left(\frac{d_i}{\sigma_0}\right)^2$ is the sum of squared standard Gaussian random variables, meaning it follows a χ^2_N distribution. For large N, a χ^2_N distribution becomes

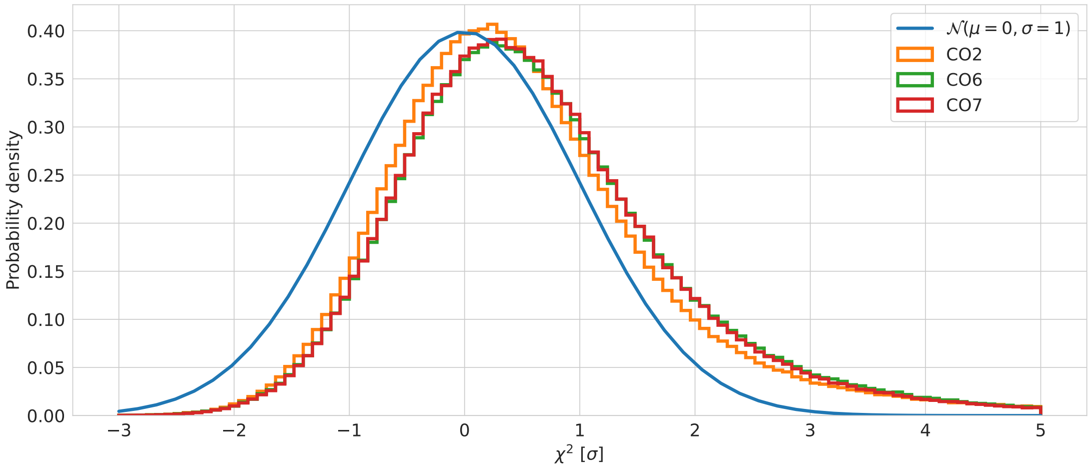


Figure 8.3: Distribution of χ^2 metrics as defined in equation 8.6 for level 2 data of our three observational fields. A standard normal distribution is plotted for comparison, which corresponds to pure white noise in line with the radiometer equation.

a Gaussian distribution, with $\mu = N$ and $\sigma = \sqrt{2N}$. Therefore, the χ^2 metric described in equation 8.6 is expected to be a standard Gaussian $\chi^2 \sim \mathcal{N}(\mu = 0, \sigma = 1)$.

We combine all frequency channels from each sideband in calculating a single χ^2 value, such that $N = n_{\text{samp}} \cdot n_{\text{freq}}$, where $n_{\text{freq}} = 64$, and n_{samp} is typically $\sim 20,000$. Figure 8.3 shows the distribution of these χ^2 values for each available scan. We have divided the metrics up by observational field, where each field contains $\sim 15,000$ scans. Although we see a clear positive bias in all three fields, it is well within 1σ, which for a white noise χ^2 test with $N \approx 10^6$ datapoints puts us very close to white noise. The CO6 and CO7 fields have near-identical distributions, while the CO2 field outperforms them somewhat. This is likely due to CO6 and CO7 being observed at very similar pointing coordinates, while the CO2 field is somewhere else entirely on the sky. Perhaps most importantly, CO2 is observed at a much lower (and more consistent) elevation than CO6 and CO7, as we saw all the way back in figure 3.4.

8.3 Polyfilter coefficients

The dominating correlated noise component in our data are the gain fluctuations coming from the LNAs attached to each feed in the COMAP telescope. These gain fluctuations are believed to have a constant frequency profile (for normalized data), and they are individual to each feed, as each feed has its separate LNA. The main purpose of the polynomial filter, where we fit and subtract a linear function in frequency in each sideband, is to suppress these gain fluctuations. The polynomial coefficients of this filter, therefore, provide valuable insight into the behavior of these fluctuations. They also provide a good starting point for proposing a new frequency filter.

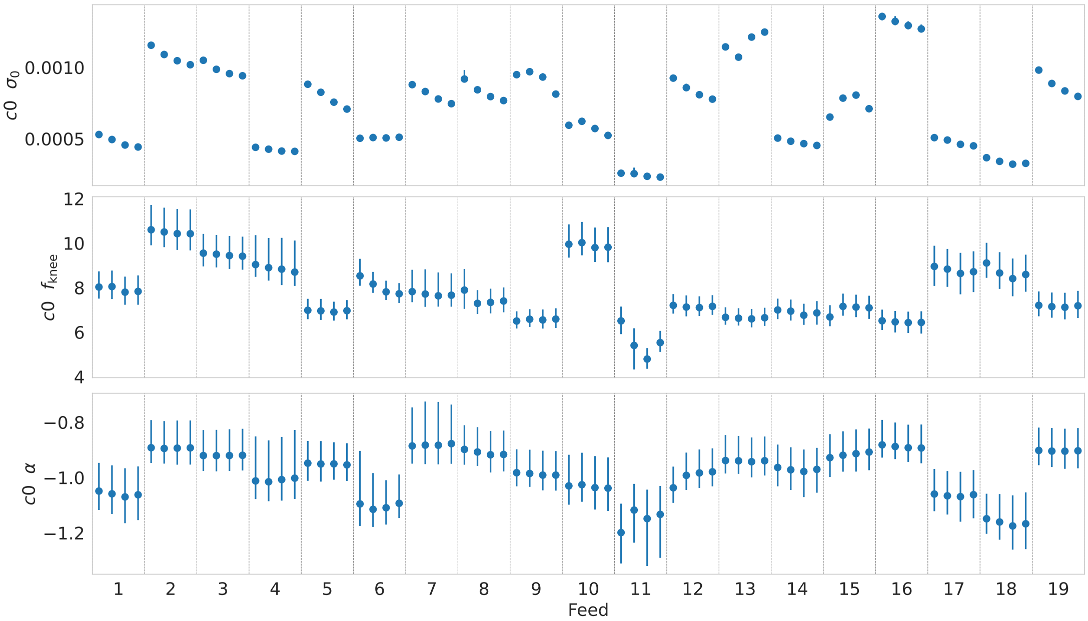


Figure 8.4: The $1/f$ best fit parameters of the constant term in the polyfilter, for the four sidebands of each of the 19 feeds. The figure shows the median, and 1σ upper and lower deviation of the $c0$ parameter, obtained from best-fit of a $1/f$ power spectrum on each available scan.

The gain fluctuations are well modeled by a $1/f$ power spectrum. This is evident from looking at the power spectrum of the data before the polynomial filter, which clearly follows a $1/f$ spectrum, as we saw in figure 8.2. While we cannot prove from this figure that what we are removing are gain fluctuations, it is at least evident that whatever we are removing follows a $1/f$ spectrum. Under the assumption that the gain fluctuations do indeed follow a $1/f$ spectrum, a convenient way of studying their behavior is simply looking at the $1/f$ best-fit parameters.

For each scan, `l2gen` performs a power spectrum best fit of the c0 and c1 polyfilter parameters (see equation 4.17), and stores the values of this fit. In figure 8.4, we show the mean and standard deviation of these across feeds and sidebands. The reason for separating the values by feed is, as already discussed, the fact that each feed has its own LNA, and the noise properties induced by the LNA are therefore expected to differ by feed. We have also chosen to show the parameters for each sideband separately, to see if the gain behaves notably different between sidebands.

We see a large spread in the white noise levels σ_0 across feeds, with some at above twice the value of others. The values are, however, very consistent within each feed, with the error bars barely visible, meaning that each LNA has a very consistent white noise profile. Across the sidebands of each feed, we also see some variation, although much smaller than between different feeds. Among most of the feeds, there seems

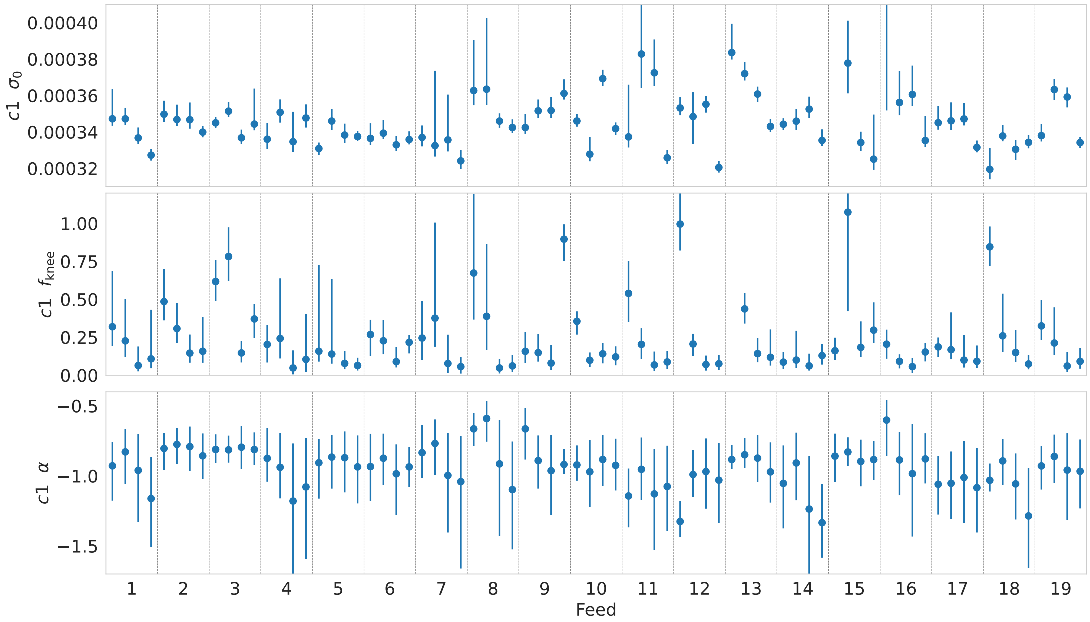


Figure 8.5: The $1/f$ best fit parameters of the linear slope term in the polyfilter, for the four sidebands of each of the 19 feeds. The figure shows the median, and 1σ upper and lower deviation of the $c1$ parameter, obtained from best-fit of a $1/f$ power spectrum on each available scan.

to be a slight linear decrease in the white noise level towards the higher frequency sidebands, although there are some exceptions to this, like feed 13 and 15. The presence of structure across the sidebands of each feed hints at some frequency dependence in these fluctuations (although the sideband-sideband structure does not necessarily imply inter-sideband structures). The gain fluctuations are primarily believed to be constant in frequency, and these structures could perhaps be attributed to some temperature source, like the atmosphere. (We will come back to this in great detail when discussing the new frequency filter.) We are already using a linear polyfilter model in our polynomial filter, and the linear structure of most of these noise levels seems to lend some support to this choice. For the knee frequency and slope, f_{knee} and α, the story is similar but less pronounced. The differences between different feeds still noticeable, but are smaller, as are the differences between different sidebands of each feed. However, the variations between scans are much larger, as seen from the error bars. Note that there is a strong correlation between the parameters within each feed (or across all feeds), which the error bars do not convey.

Figure 8.5 shows corresponding plots for the linear term in the polynomial filter. This term is generally much smaller, and as we can see, behaves more erratic. There is much less consistent structure across sidebands and feeds in these parameters. However, the errorbars, especially on σ_0, are on many sidebands very small, meaning that the $1/f$ profile is consistent within the sideband.

Chapter 9

A New Frequency Filter

In this chapter, we propose a new frequency filter for the COMAP pipeline, to replace the current polynomial filter. The polynomial filter is built primarily to handle fluctuations in the telescope gain, which is known to be the largest correlated noise term in our data. However, we believe it may be inadequate at handling systematic fluctuations in the system temperature, which, while being of much smaller amplitude than the gain, can build up as a systematic bias in the data over time. The new frequency filter is meant to better handle such temperature contaminations in the data, by performing a joint maximum likelihood fit of the signal into system temperature and gain, on short timescales. While the separation between gain and system temperature on observation-length timescales are resolved to high accuracy by the telescope calibration every hour, these values are simply interpolated to the center of each scan, offering no real separation between the two on shorter timescales.

We first set up a new and more realistic data model for our signal, and find a maximum likelihood solution to the joint fit of gain and temperature. We then present our findings from testing the new filter on both real and simulated data. We first apply the filter to a couple of chosen scans in the COMAP dataset, and then move on to working on simulated data, as this allows us to better study the amount of temperature fluctuations left in the output signal. The impact of different priors on the gain fluctuations will also be considered. We finish the chapter by taking a look at how well the filter removes astronomical continuum foregrounds, with Jupiter as a case study, and discuss whether the new filter could have applications in continuum science as well.

9.1 Theoretical foundation

The observed power of the telescope can be modeled as[1]

$$d_{\nu t} = T_{\nu t} g_{\nu t} \tag{9.1}$$

---

[1] Again, we have left out the terms of $\Delta\nu k_B$, which can be included in the gain.

where $T_{\nu t}$ is the system temperature, and $g_{\nu t}$ is the gain, both of which are unknown functions of frequency and time. We chose to model these as $T_{\nu t} = \bar{T}_\nu + \delta T_{\nu t}$ and $g_{\nu t} = \bar{g}_{\nu t}(1 + \delta g_{\nu t})$, where $\bar{T}_\nu$ and $\bar{g}_\nu$ are the constant (or very slow-running) system temperature and gain, and $\delta T_{\nu t}$ and $\delta g_{\nu t}$ are the fluctuations around these means. The reason for the gain and temperature fluctuation taking different shapes is that the gain fluctuations are observed to follow the frequency profile of the mean gain. However, a temperature increase of 1 K from a continuum source would raise the observed temperature by 1 K for all frequencies, and it would not increase more for frequencies with higher average system temperatures. Inserting these quantities in equation 9.1, we get the very general data model

$$d_{\nu t} = \bar{T}_\nu \bar{g}_\nu (1 + \delta g_{\nu t}) \left[1 + \frac{\delta T_{\nu t}}{\bar{T}_\nu} \right] (1 + n_{\nu t}), \tag{9.2}$$

where we have also included a white noise term $n_{\nu t}$.

In the most general case, $\delta T_{\nu t}$ and $\delta g_{\nu t}$ could be any function of time and frequency. This would, however, make them completely degenerate with both each other and the white noise, and there are physical justifications for assuming they are simple functions of frequency. In the simplest case, we assume that both are constant across frequency, but make no assumptions about their time-dependence, such that $\delta T_{\nu t} = \delta T_t$ and $\delta g_{\nu t} = \delta g_t$, giving the signal model

$$d_{\nu t} = \bar{T}_\nu \bar{g}_\nu (1 + \delta g_t) \left[1 + \frac{\delta T_t}{\bar{T}_\nu} \right] (1 + n_{\nu t}). \tag{9.3}$$

The next natural step in model complexity is to allow the fluctuations to depend linearly on frequency. Most temperature contaminations in our signal are continuum sources, like ground, atmosphere, weather, and astronomical foregrounds. Continuum sources are well approximated as linear on smaller frequency ranges, like the 2 GHz sidebands of the COMAP data. We, therefore, have a lot of physical justification for wanting to expand our temperature model to include a linear term. With this added, our model now reads

$$d_{\nu t} = \bar{T}_\nu \bar{g}_\nu (1 + \delta g_t) \left[1 + \frac{\delta T_t}{\bar{T}_\nu} + \alpha \frac{\delta T_t}{\bar{T}_\nu}(\nu - \bar{\nu}) \right] (1 + n_{\nu t}), \tag{9.4}$$

where $\bar{\nu}$ is the frequency mean, and we have introduced the temperature frequency slope α. We could expand the gain with a similar linear term, but in our experience, the gain is well modeled without it. The reason our current polynomial filter model includes a linear term is exactly to account for continuum temperature systematics, which we are now dealing with in a more adequate way. We, therefore, propose this as our final model.

The frequency filter is performed after the normalization and pointing template removal, where the signal is left with only the small and short-term fluctuations. We model the normalized data as

$$y_{\nu t} = d_{\nu t}/(\bar{T}_\nu \bar{g}_\nu) - 1, \tag{9.5}$$

$$= (1 + \delta g_t)\left[1 + \frac{\delta T_t}{\bar{T}_\nu} + \alpha \frac{\delta T_t}{\bar{T}_\nu}(\nu - \bar{\nu})\right](1 + n_{\nu t}) - 1, \tag{9.6}$$

$$\approx \delta g_t + \frac{\delta T_t}{\bar{T}_\nu} + \alpha \frac{\delta T_t}{\bar{T}_\nu}(\nu - \bar{\nu}) + n_{\nu t}, \tag{9.7}$$

where we have assumed that all perturbation terms are small, such that any product between them is negligible. To clean this up a bit, we introduce $\alpha' = \alpha \delta T_t$ and $\nu' = (\nu - \bar{\nu})/\bar{T}_\nu$, such that our model can be written as

$$y_{\nu t} = \delta g_t + \frac{\delta T_t}{\bar{T}_\nu} + \alpha'_t \nu'_\nu + n_{\nu t}. \tag{9.8}$$

9.1.1 Maximum likelihood solution

This section introduces the maximum likelihood solutions of δg, δT and α'_t from equation 9.8. We can write this equation as a matrix equation, on the form

$$\mathbf{Y} = \mathbf{Pm} + \mathbf{Fa} + \mathbf{N} \tag{9.9}$$

written out as

$$\begin{pmatrix} y_{1,1} & y_{1,2} & \cdots & y_{1,n} \\ y_{2,1} & \ddots & & \\ \vdots & & & \\ y_{m,1} & & & y_{m,n} \end{pmatrix} = \begin{pmatrix} \bar{T}_1^{-1} & \nu'_1 \\ \bar{T}_2^{-1} & \nu'_2 \\ \vdots & \\ \bar{T}_m^{-1} & \nu'_m \end{pmatrix} \begin{pmatrix} \delta T_1 & \delta T_2 & \cdots & \delta T_n \\ \alpha'_1 & \alpha'_2 & \cdots & \alpha'_n \end{pmatrix}$$
$$+ \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \begin{pmatrix} \delta g_1 & \delta g_2 & \cdots & \delta g_n \end{pmatrix} + \begin{pmatrix} n_{1,1} & n_{1,2} & \cdots & n_{1,n} \\ n_{2,1} & \ddots & & \\ \vdots & & & \\ n_{m,1} & & & n_{m,n} \end{pmatrix}. \tag{9.10}$$

The equation is now on a similar form to that of the destriper signal equation, except with the dimensionality of some values a bit off. If you want some intuition on how this equation relates to the destriper setup, and where the pixels are, we have written out the equation on a less compact and more destriper-like form in appendix A. What is most important is that we are familiar with the maximum likelihood solutions to such an equation.

The maximum likelihood solutions for **a** and **m**, assuming a diagonal correlation matrix $\mathbf{C}_n$, we know from section 5.3.1 to be

$$(\mathbf{F}^T \mathbf{C}_n^{-1} \mathbf{Z} \mathbf{F})\mathbf{a} = \mathbf{F}^T \mathbf{C}_n^{-1} \mathbf{Z} \mathbf{y}, \tag{9.11}$$

and

$$\mathbf{m} = (\mathbf{P}^T\mathbf{C}_n^{-1}\mathbf{P})^{-1}\mathbf{P}^T\mathbf{C}_n^{-1}(\mathbf{y} - \mathbf{F}\mathbf{a}), \tag{9.12}$$

where $\mathbf{Z} = \mathbf{I} - \mathbf{P}(\mathbf{P}^T\mathbf{C}_n^{-1}\mathbf{P})^{-1}\mathbf{P}^T\mathbf{C}_n^{-1}$.

As we explored in section 5.3.1, this solution can easily be extended with some prior on **a**, by including its correlation matrix. Since we have a good understanding of the power spectrum characteristics of the gain fluctuations, represented by **a**, this is a natural extension. The maximum likelihood solution for **a** now takes the form

$$(\mathbf{F}^T\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{F} + \mathbf{C}_a^{-1})\mathbf{a} = \mathbf{F}^T\mathbf{C}_n^{-1}\mathbf{Z}\mathbf{y}. \tag{9.13}$$

On shorter timescales, for a single detector, a common assumption is that the white noise level is stationary in time. Under this assumption, the white noise covariance is a uniform diagonal matrix, $\mathbf{C}_n = \sigma_0^2\hat{\mathbf{I}}$. Additionally, for our specific setup, the quantity $z = \mathbf{F}^T\mathbf{Z}\mathbf{F}$ is actually a scalar. This means that the only non-diagonal quantity on the left hand side of our equation is the covariance matrix $\mathbf{C}_a$, without which the solution for **a** would be explicit. This matrix is, however, diagonal in the Fourier domain. We therefore rewrite the solution for **a** as

$$\frac{z}{\sigma_0^2}\mathbf{a} + \mathcal{F}^{-1}[\mathcal{F}[\mathbf{a}]/\mathbf{C}_f] = \frac{1}{\sigma_0^2}\mathbf{F}^T\mathbf{Z}\mathbf{y}, \tag{9.14}$$

where $\mathbf{C}_f = \mathcal{F}[\mathbf{C}_a]$ is the Fourier representation of $\mathbf{C}_a$.

Doing a Fourier transformation of each side of the equation, the solution for **a** becomes explicit, as

$$\frac{z}{\sigma_0^2}\mathcal{F}[\mathbf{a}] + \mathcal{F}[\mathbf{a}]/\mathbf{C}_f = \frac{1}{\sigma_0^2}\mathbf{F}\left[\mathbf{F}^T\mathbf{Z}\mathbf{y}\right], \tag{9.15}$$

$$\mathcal{F}[\mathbf{a}] = \frac{\mathcal{F}\left[\mathbf{F}^T\mathbf{Z}\mathbf{y}\right]}{z + \sigma_0^2/\mathbf{C}_f}. \tag{9.16}$$

We know that the power spectrum of the gain fluctuations are very well approximated by a $1/f$ function, a claim we will explore in more detail later. The prior we wish to place on **a** therefore takes the form

$$\mathbf{C}_f = \mathcal{F}[\mathbf{C}_a] = \sigma_0^2\left(\frac{f}{f_\text{knee}}\right)^\alpha, \tag{9.17}$$

where f_knee and α has to be decided, and σ_0 is the white noise level.

Masking

In the COMAP data, some frequencies are always masked out due to their poor performance and large systematics. This is simple to implement in our model, by simply setting the corresponding rows in **y**, **P**, and **F** to zero. An example of this, with the second frequency channel masked, can be written as

$$\begin{pmatrix} y_{1,1} & y_{1,2} & \cdots & y_{1,n} \\ 0 & 0 & \ddots & 0 \\ \vdots & & & \\ y_{m,1} & & & y_{m,n} \end{pmatrix} = \begin{pmatrix} \bar{T}_1^{-1} & \nu_1' \\ 0 & 0 \\ \vdots & \\ \bar{T}_m^{-1} & \nu_m' \end{pmatrix} \begin{pmatrix} \delta T_1 & \delta T_2 & \cdots & \delta T_n \\ \alpha_1' & \alpha_2' & \cdots & \alpha_n' \end{pmatrix}$$
$$+ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \begin{pmatrix} \delta g_1 & \delta g_2 & \cdots & \delta g_n \end{pmatrix} + \begin{pmatrix} n_{1,1} & n_{1,2} & \cdots & n_{1,n} \\ 0 & 0 & \ddots & 0 \\ \vdots & & & \\ n_{m,1} & & & n_{m,n} \end{pmatrix}. \tag{9.18}$$

9.1.2 Comparison to polynomial filter

Before moving on to our results, we try building some intuition on why we believe that the new frequency filter is an improvement on the existing polynomial filter, and what implications this will have on our data. The currently used polynomial filter works by fitting and subtracting a linear function in frequency from every timestep, across each sideband of each feed. This essentially assumes that the normalized data can be well approximated as

$$y_{\nu t} = c_0 + c_1(\nu - \bar{\nu}) + n_{\nu t}. \tag{9.19}$$

This fits our profile of the gain fluctuations, which are assumed to be constant. However, the temperature fluctuations become modulated by the mean system temperature frequency profile, which we know is not linear (see figure 4.1). The reason the current polynomial filter works well enough is twofold. First, the gain fluctuations completely dominate the correlated noise profile. We do not know exactly how much temperature contaminations an ordinary scan contains, but it is far less than the gain. Secondly, although the system temperature frequency profile is not linear, it is not too bad of an approximation. The most obvious problem is the system temperature spikes. These are often masked away at some point in the pipeline, and therefore not considered too great a problem.

Even though the temperature fluctuations are assumed to be small, they pose a larger problem to the COMAP data than immediately obvious. Most sources of these contaminations are, unlike the gain, correlated in galactic coordinates or telescope pointing. This is a huge problem, as such systematics do not integrate down in time, and could be interpreted as actual CO signal. Any astrophysical foreground is directly correlated in galactic coordinates. Additionally, ground, atmosphere, and weather are all correlated with telescope pointing (azimuth/elevation). While this is not completely degenerate with galactic coordinates, as the fields drift across the sky, it is still a lot worse than the completely uncorrelated gain fluctuations. In summary, we have very good reasons to worry more about the temperature fluctuations compared to the gain fluctuations, even if their amplitudes are small within a single scan.

9.2 Proof of concept - Real scans

This section presents a qualitative analysis of the results obtained from applying the new frequency filter to a few real scans. Figure 9.1 demonstrates the difference between the old polynomial filter and the new frequency filter for an ordinary scan. The more complicated, non-linear structure of the new filter can be seen, especially around the system temperature spikes. Since the temperature fluctuations are very small compared to the overall noise level, we do not expect to clearly see which fit is better by eye, but the new fit should in theory be able to better model the systematics. The frequency correlation across the first sideband is shown in figure 9.3, where, again, no real difference between the two methods can be observed.

Figure 9.2 shows the same fit on another scan, which was picked specifically because of the presence of heavy weather contaminations, which is a temperature systematic. Here, the merit of the new fit is very clear, with the old fit being very inadequate at modeling the frequency profile of the data. A scan such as this is contaminated beyond recovery, and would never be used in actual analysis, but serves as a good proof-of-concept for our new filter. Figure 9.4, showing the frequency correlation of the first sideband, tells the same story. The old polynomial filter results in heavy correlation across all frequencies, while the new filter reduces the correlation almost down to the level of the well-behaved scan.

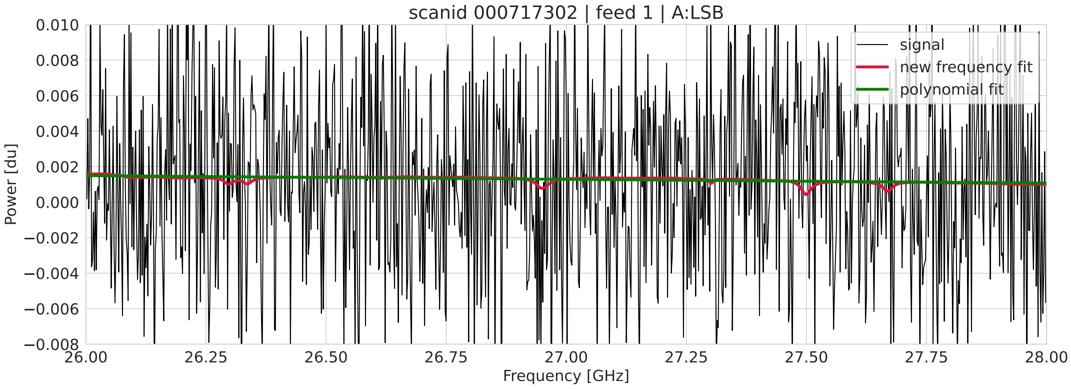


Figure 9.1: Comparison between the previous linear polynomial fit (green), to the new frequency filter fit (red) on an "ordinary" scan without obvious weather contaminations. The new fit has a non-linear profile containing an imprint of the system temperature profile, but it is not immediately clear if the new fit better models the data.

A problem in testing new filtering methods with real data is that we have no universal metric for determining the quality of output data. While we could test how consistent with white noise our signal is, tests like these fail to distinguish between systematics

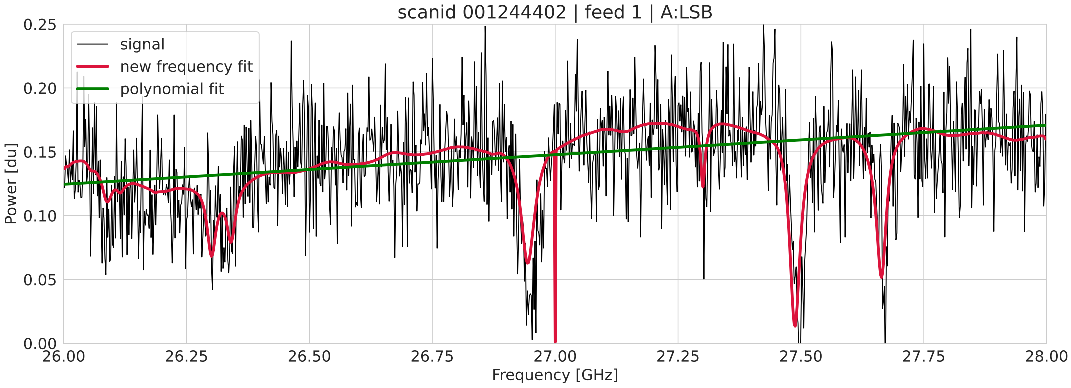


Figure 9.2: Comparison between the previous linear polynomial fit (green), to the new frequency filter fit (red) on a "bad" scan with obvious weather contaminations. It is very obvious that the data contains structures which are not adequately modeled by the polynomial fit, while the new frequency fit fares much better.

which will bias our data and not, which is a very important distinction. In general, we wish to see as little correlated noise as possible in our output TOD, but some correlated noise is far worse than others. Noise that is correlated in either Galactic coordinates or signal frequency is of most serious concern, as it could be interpreted as actual signal. Any noise not correlated in such a way will simply be integrated down when we produce maps. Temperature contaminations are often of the former variant, such as those from ground, atmosphere, or astrophysical foregrounds. Gain fluctuations on the other hand, while being our dominant source of correlated noise, are only temporally correlated. It therefore mainly increases our integration time, and do not bias our data the way temperature contaminations may. We will therefore now transition to looking at simulated signal, where we have better control over the different types of systematics.

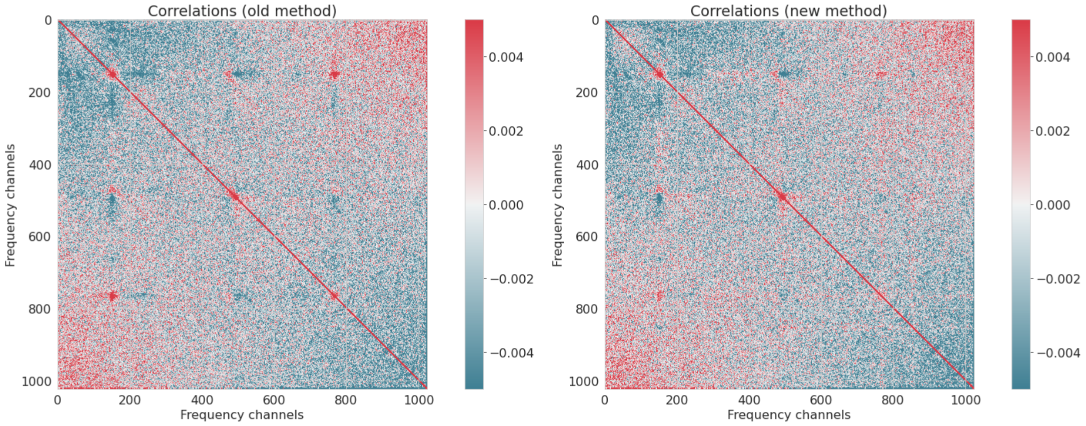


Figure 9.3: Frequency channel correlations across A:LSB of feed 1 for scan 000717302. This scan contains no obvious weather contaminations, and the correlation plots are similar.

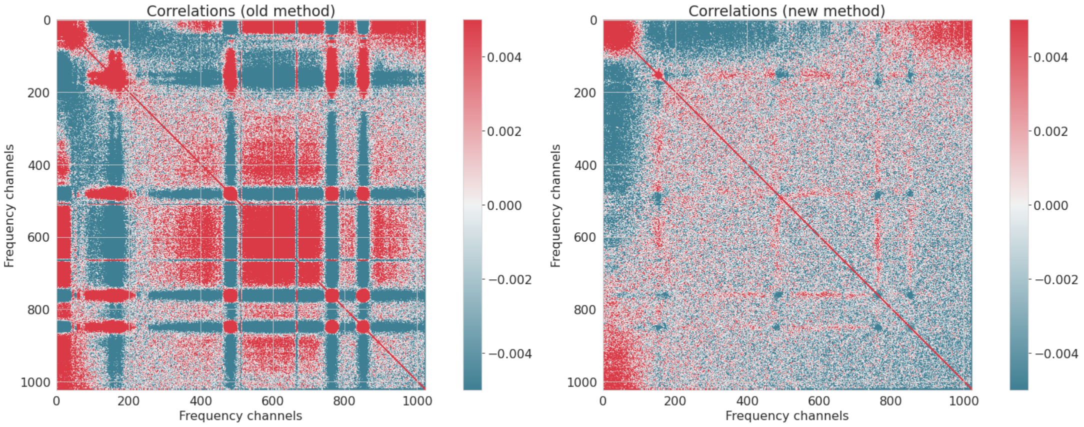


Figure 9.4: Frequency channel correlations across A:LSB of feed 1 for scan 001244402. This scan contains heavy weather contaminations, and a lot of the correlated noise passes through the old polynomial filter. The new filter also shows some large excess correlation at certain frequencies.

9.3 Proof of concept - Simulated data

In this section, we study the effect of the new frequency filter on simulated data, with similar noise properties as a real scan. As we explored in the last section, the new frequency filter is a significant improvement in scans with large excess temperature, and while this demonstrates the theoretical potency of the filter, these scans are usually unusable anyway. When applied to a scan with cleaner data, the temperature fluctuations were too small to see an obvious improvement from the single scan alone. This poses a problem, as we are worried that the temperature fluctuation might introduce a bias in our data over time, yet we are unable to confirm that the new filter is adequately removing these contaminations from a single scan. With simulated data, we know the solution to how the signal decomposes into gain and system temperature, and can more accurately explore how well the filter removes them.

9.3.1 Data generation

We generate our data using the data model we introduced in section 9.1, where the TOD takes the form

$$d_{\nu t} = \bar{T}_\nu \bar{g}_\nu \left[\delta g_t + \frac{\delta T_t}{\bar{T}_\nu} + \alpha \frac{\delta T_t}{\bar{T}_\nu}(\nu - \bar{\nu}) + n_{\nu t} \right]. \tag{9.20}$$

We use a setup close to that of a typical scan, with 20,000 timesteps, over 19 feeds, with 4 sidebands of 1024 frequency bins each. We adopt the mean gain and system temperature, $\bar{g}_\nu$ and $\bar{T}_\nu$, from a randomly selected scan. The white noise $n_{\nu t}$ is generated as an uncorrelated Gaussian signal with standard deviation taken from the radiometer equation, which we have found to conform well with our data. The gain and system temperature fluctuations, δg_t and δT_t, are generated as correlated signals with a $1/f$

power spectrum without the white noise term. For the gain, we use $1/f$ parameters based upon those of the polynomial parameter c0 we analyzed in section 8.2. We average these parameters across feeds and sidebands, and use the values $\sigma_0^{\delta g} = 0.00073$, $f_{\text{knee}}^{\delta g} = 7.78\,\text{Hz}$, and $\alpha^{\delta g} = -1$. For the system temperature fluctuations, we simply do not have accurate measurements of its $1/f$ behavior. We will experiment with a range of values, but as a baseline, we give it a white noise amplitude a factor 16 lower than that of the gain (a factor we will simply refer to as "the σ_0 factor", for simplicity), at $\sigma_0^{\delta T} = 0.000024$, and a slope of $\alpha^{\delta T} = -1.4$. We always keep $f_{\text{knee}}^{\delta T} = f_{\text{knee}}^{\delta g}$.

Now that we have several meanings to certain quantities, we will refer to the maximum likelihood fits as δg^{out} and δT^{out}, to clearly distinguish them from the "true" simulated data, δg and δT. Similarly, the filtered data will be referred to as d^{out}, while d refers to the simulated input TOD.

9.3.2 Correlation

To quantify the presence (or absence) of temperature fluctuations, we will look at the correlation between the temperature fluctuations δT_t, which we randomly generate, and the filtered output TOD. The correlation between two timeseries a and b is defined as

$$\text{Corr}(a,b) = \frac{E[(a-\bar{a})(b-\bar{b})]}{\sigma_a \sigma_b} = \frac{\sum_{i=1}^{N}(a_i-\bar{a})(b_i-\bar{b})}{\sigma_a \sigma_b} \tag{9.21}$$

where $\bar{a}$ and $\bar{b}$ are the mean values of the timeseries, and σ_a and σ_b are their standard deviations. The correlation is always a number between -1 and 1; it is 1 if the two are identical (up to a factor), and -1 if they are anti-symmetric (up to a factor).

Unless the timeseries are of infinite length, the correlation will always be non-zero even for uncorrelated data. The correlation between two timeseries of N independent gaussian random variables follow a normal distribution[2] $\mathcal{N}(0, 1/\sqrt{N})$. However, the quantities we will be comparing, like δT and d^{out}, are not independent Gaussian random variables, but instead correlated in time. Two timeseries with temporal correlation have larger expected correlation between each other than two independent Gaussian time series do. The correlation distribution is also non-trivial. To establish the expected correlation distribution between two completely uncorrelated timeseries of, for example, δT and d^{out}, we simulate a large number of δT timeseries, which are independent of the relevant TOD, and calculate the correlation distribution between these and the TOD. This serves as the uncorrelated baseline we compare our results to.

To further simplify the analysis of these correlations, we introduce a metric to quantify this correlation. Let d_j^{out} be the jth of M filtered TODs, and δT_j be the corresponding simulated temperature fluctuations. Additionally, let δT_j^{rnd} be an independent realization of δT_j, with the same statistical properties. We then introduce the mean squared correlation excess between the two timeseries, as

---

[2] This is an approximation which holds for $N \gg 1$, which will always be the case for us, as we are looking at thousands or tens of thousands of datapoints.

$$\mathrm{MSCE}(d^{\mathrm{out}}, \delta T) = \frac{\sum_{j=1}^{M} \Big[\mathrm{Corr}(d_j^{\mathrm{out}}, \delta T_j)\Big]^2}{\sum_{j=1}^{M} \Big[\mathrm{Corr}(d_j^{\mathrm{out}}, \delta T_j^{\mathrm{rnd}})\Big]^2}. \tag{9.22}$$

This quantity can be interpreted as the mean squared error above the expected correlation of d^{out} and δT, if they were completely uncorrelated. It will take a value of 1 if they are completely uncorrelated, and higher values the more correlated they are.

9.3.3 Temperature correlation results

Figure 9.5 shows the distribution of correlations between d^{out} and δT for a simulated scan with the default parameters we outlined in section 9.3.1. The new frequency filter shows a distribution very much in line with the randomly sampled correlations, while the old polynomial filter produces a slightly broader distribution. In figure 9.6, we have split this distribution up into frequencies with system temperatures above and below 60 K. Here we clearly see that the polynomial filter has a negative correlation bias for high temperatures and a slight positive bias for low temperatures. This is because, as figure 9.2 demonstrates, when there are temperature systematics in the data, the polynomial filter underestimates the amplitude of the signal for most frequencies, while overestimating it for the system temperature spikes. In an ordinary scan, this effect will be small but will build up a bias in the filtered signal over time.

As the amplitude and $1/f$ behavior of the temperature fluctuations in an ordinary scan are not known to us, we have simulated data with a range of parameters and calculated their correlations. For some intuition on the magnitudes of the different signals involved, we have plotted the power spectra of the gain, white noise, and temperature fluctuations in figure 9.7. We also show the effect of halving $\sigma_0^{\delta T}$, as well as changing $\alpha^{\delta T}$ from -1.4 to -1.6, to give some intuition for what these changes represent.

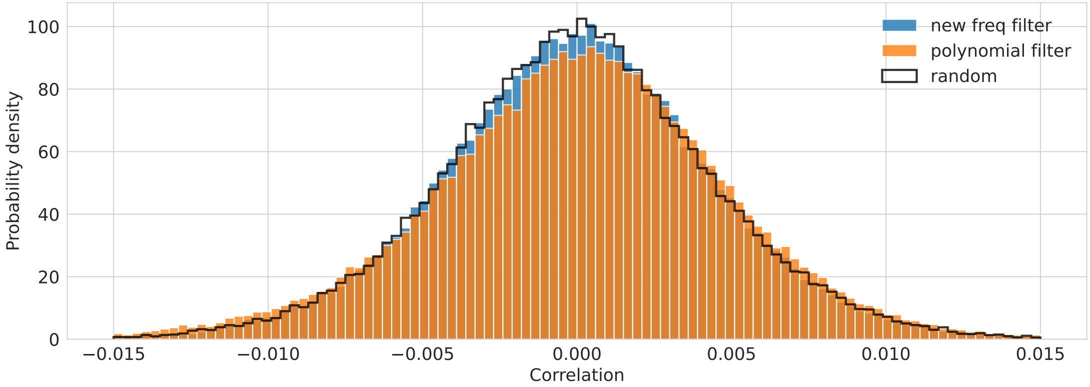


Figure 9.5: Distribution of correlations between filtered signal and temperature fluctuations for simulations of a scan with 4096 frequency channels and 19 feeds. Both the old polynomial filter and the new frequency filter, are shown. As a reference, the correlation distribution between two independent realizations of signal and temperature is shown in black, as described in section 9.3.2. The default $1/f$ parameters outlined in section 9.3.1 was used.

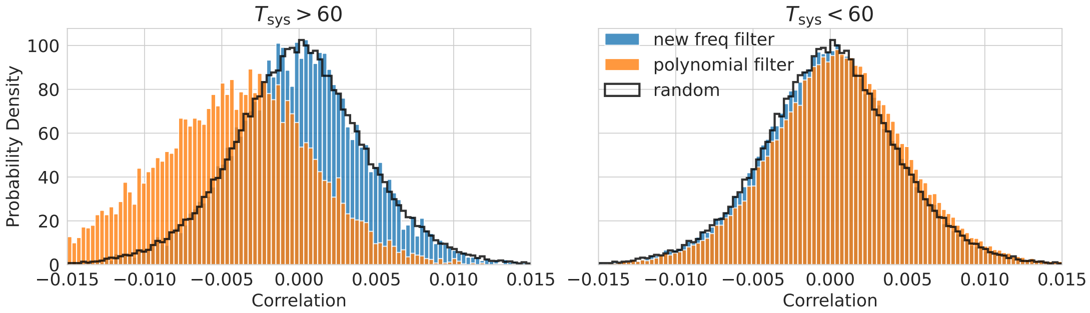


Figure 9.6: The same correlation distribution as shown in figure 9.5, split into frequencies above and below 60 K in system temperature.

In figure 9.8 we have plotted the MSCE (as defined in equation 9.22) between d^{out} and δT for both the old polynomial filter and the new frequency filter, with $\alpha^{\delta T}$ values ranging from -1 to -1.8, and σ_0 factors from 4 to 64. For some reference on the MSCE values, remember that a histogram of correlations for the default configuration of σ_0 factor $= 16$, $\alpha = -1.4$, with a MSCE value of 1.111, is plotted in figure 9.5. For every single configuration, the frequency filter shows less correlation with temperature than the polynomial filter. The polynomial filter has a large excess of correlation for all but the left-uppermost values, which represent the smallest temperature amplitudes. There is no noticeable excess in the frequency filter for any configuration. With this, we can conclude that the new filter will always be better at removing temperature systematics in our data, but how much better comes down to how much such systematics are present in the first place.

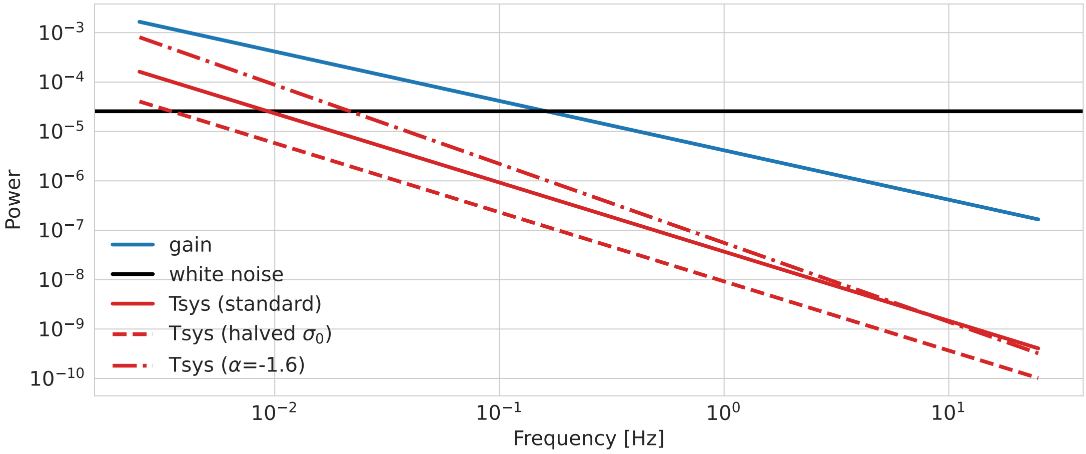


Figure 9.7: Power spectra used for the simulated data. We fix the gain and white noise power spectra from values from analysis of actual data, as outlined in section 9.3.1. The temperature fluctuations is by default set to $\alpha^{\delta T} = -1.4$, with $f_{\mathrm{knee}}^{\delta T} = f_{\mathrm{knee}}^{\delta g}$, and a $\sigma_0^{\delta T} = \frac{1}{16}\sigma_0^{\delta g}$, as shown by the solid line. The striped and dotted lines show the effect of halving $\sigma_0^{\delta T}$, and changing $\alpha^{\delta T}$ from -1.4 to -1.6 in the temperature power spectrum.

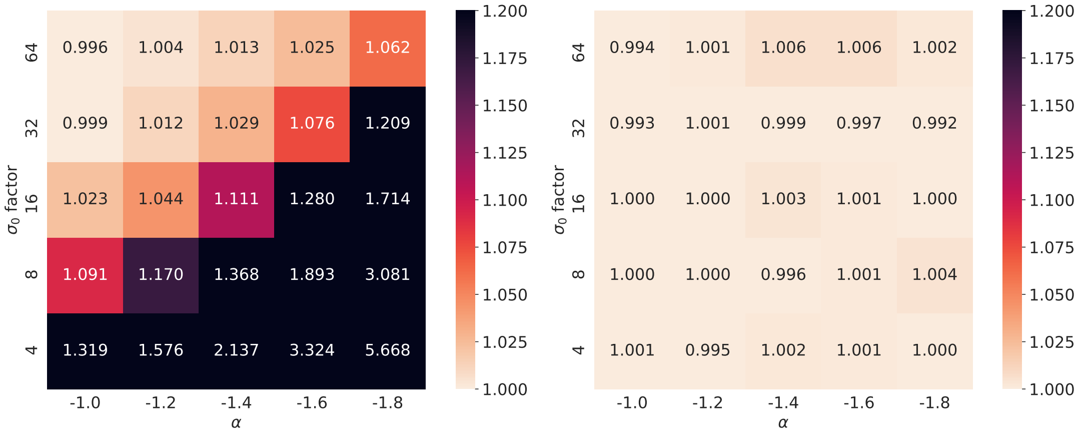


Figure 9.8: MSCE values between the simulated δT, and the filtered signal d^{out} for the old polynomial filter (left) and the new frequency filter (right). The values as plotted as function of the σ_0 factor (the ratio $\sigma_0^{\delta g}/\sigma_0^{\delta T}$), as well as the power spectrum slope $\alpha^{\delta T}$.

Filter type	Corr($\delta g^{\text{out}}, \delta T^{\text{out}}$)	MSCE($d^{\text{out}}, \delta g$)	MSCE($d^{\text{out}}, \delta T$)
Polyfilter	n/a	1.004	1.107
Freqfilter $f_{\text{knee}}^{\text{prior}} = 0.1 f_{\text{knee}}^{\delta g}$	0.3002	2.543	1.005
Freqfilter $f_{\text{knee}}^{\text{prior}} = f_{\text{knee}}^{\delta g}$	0.0005	1.583	1.005
Freqfilter $f_{\text{knee}}^{\text{prior}} = \infty$	-0.7522	1.003	1.005

Table 9.1: Correlation and mean squared correlation excess between different parameters for the polynomial filter, and three prior configurations of the frequency filter.

9.3.4 Noise level and gain fluctuation priors

We have now established that the new frequency filter succeeds in its primary purpose of removing temperature fluctuations from a range of realistic data configurations. With that settled, we will now explore differences between the old and new models in regard to the gain fluctuations and the overall noise level in the filtered data. It should be stressed that this is a secondary concern, and we would gladly offer a slight increase in the overall noise profile (within reasonable limits) in order to get rid of potential temperature fluctuations in our data. This section will also explore the impact of the prior we have placed on the gain fluctuations, which was left out of the previous analysis because it did not meaningfully influence the temperature correlation results.

We return to looking at simulated data with the default parameters outlined in section 9.3.1. Figure 9.9 together with table 9.1 summarizes the results of this simulation with different prior configurations. Looking at the figure we see that the choice of prior configuration has, unsurprisingly, the biggest impact on δg^{out}, but δT^{out} is also somewhat influenced. Reducing the $f_{\text{knee}}^{\text{prior}}$ by a factor of 10 gives δg^{out} both a lower amplitude and less short timescale fluctuations, as the high-frequency modes of the power spectrum is suppressed. In this configuration, we see that δT^{out} takes on a slightly higher amplitude than the default prior, probably in an attempt to model the gain. Additionally, the two quantities looks to be somewhat correlated. This is backed up by the table, which shows a positive correlation of 0.3. Again, this is probably because δT^{out} attempts to model the gain fluctuations, as δg^{out} itself is suppressed by the prior. The temperature also takes on a larger amplitude than the gain, which we know is a rather unrealistic fit.

In the default prior configuration, the two quantities show a correlation of only 0.0005, which is rather negligible. Here we also see from the power spectrum that the gain manages to fit the prior structure very well. The fact that this correlation vanishes with the use of a correct prior is very useful information to bring into a real data setup. Remember that both δg^{out} and δT^{out} are known quantities in a real data scenario, and we can use this fact to pinpoint the correct prior on δg.

In the configuration without a prior altogether, there is a much larger amplitude of both gain and temperature. This is because the two quantities are almost completely anti-correlated, with a correlation of -0.750, such that their total signal contribution is much smaller than their respective amplitudes. This is very non-physical, as we do not

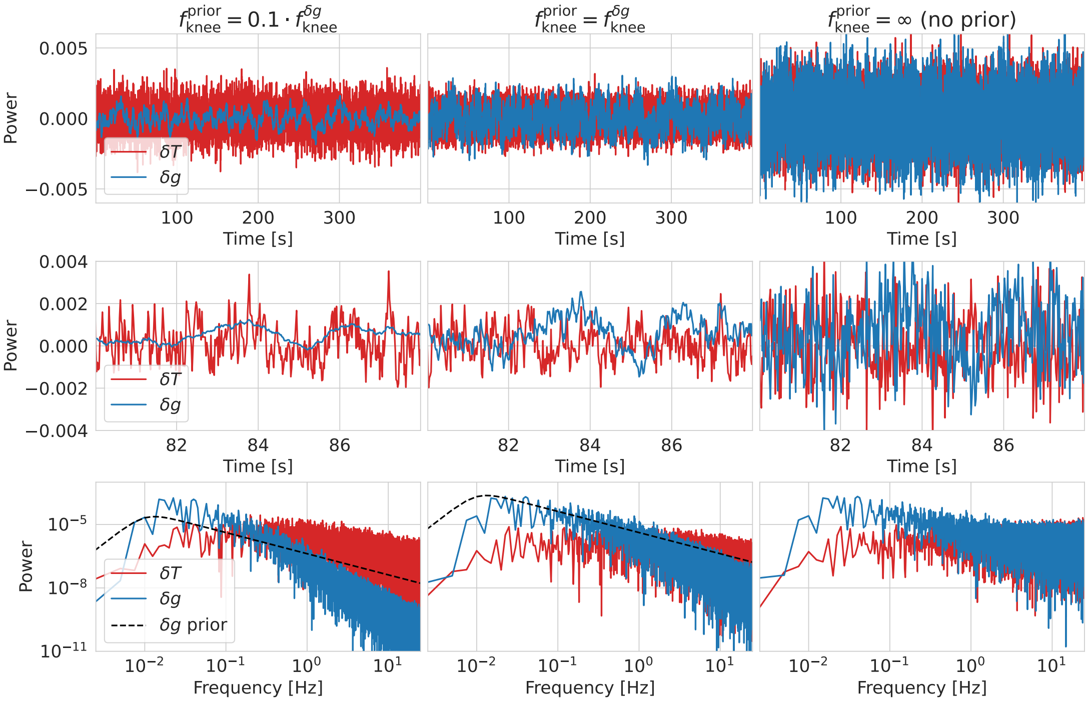


Figure 9.9: Best fit δg^{out} and δT^{out} on simulated data. The figures show three fits with different priors on δg. The middle column has a prior with the correct knee frequency, the left column has the knee frequency of the prior reduced by a factor of 10, and the right column has no prior, which is equivalent to an infinite knee frequency. The top row shows the entire TOD fits, the middle row shows a short section of the TOD, and the bottom row shows the power spectra of the fits, together with the priors on the gain.

expect any correlation between δg and δT at all. The power spectrum is similar to that of the priors, with the exception that both quantities, but especially the gain, take on a much larger amplitude in the high-frequency modes. This is not behavior we expect from the gain, and neither did we model the gain this way in the input simulation. This shows that putting the correct prior on the gain helps to break degeneracy with the temperature, by using the fact that δg is correlated in time. The anti-correlation between δg and δT is exactly what happens when we run a destriper mapmaking model on badly constrained data with little crosslinking, which results in anti-correlation between a and m. If we look at our model as a destriper, as we do in appendix A, we will see that we have absolutely no crosslinking. We therefore rely on the prior to constrain the model when the signal-to-noise ratio is low. The prior also makes the filter less harsh, since we no longer fit and subtract δg independently for each timestep. This means we could potentially lose less data, which we will come back to later.

Note also that the last column of the table supports our claim that the choice of prior does not impact the removal of the temperature fluctuations[3]. This means that, even for the very unrealistic case of the large amplitude anti-correlated fits we got from removing the prior on the gain, we do not bias our data with our fit, which is the absolutely most important point.

---

[3] These values are from a different simulation than the one which the data in figure 9.8 was based

The middle column of the table shows the correlation level between the output signal and the input gain fluctuations. We see that the polynomial filter, as well as the priorless frequency filter, removes the gain very well. The stronger the prior we place on the gain, the more residual gain is left in the filtered signal. This is because a strong prior will make us under-fit the gain fluctuations. A moderate residual in the gain is not necessarily in itself a large problem, as it is an unbiased noise. The correlation levels we see in the table are also relatively small, as the MSCE metric is extremely sensitive.

9.3.5 χ^2 goodness of fit

We have yet to assess how good the new model is at lowering the overall noise level in our data. Back in section 8.2 we introduced a χ^2 test of data consistency with white noise. We perform a similar test on the simulated data after having been filtered by both the old polynomial filter and the new frequency filter with different priors. Figure 9.10 shows the χ^2 distribution of the white noise goodness of fit, against a standard normal distribution. The distributions appear almost indistinguishable, meaning there is little difference in the overall noise level after any of the filtering variations.

It might seem surprising that the distributions are shifted to the left of the standard normal distribution. This does in fact mean that the noise level in the output data is lower than the white noise level of the simulated data. This happens because some of the white noise is fitted and removed by the models. This means that we would also lose actual signal to these filters. We already know this to be the case for the polynomial filter, which we saw negatively impact the signal transfer function back in figure 4.11. The different models look almost entirely overlapping in figure 9.10, but there are subtle differences. The mean values of the distribution, in the order they are shown in the label, are -0.248, -0.221, -0.271, and -0.346. It is, unsurprisingly, the priorless frequency filter that removes most of the white noise, and therefore probably also most data in an actual scan. We also see that the stronger prior lies closer to zero than the correct prior, even closer than the old polynomial filter does. This brings us to one of the strongest arguments for using a prior in the first place. We saw that the models constrained by a prior on the gain were worse at removing gain fluctuations, but they also *remove less CO signal.*

In the future, we will perform a proper transfer function analysis, where we include CO data in the simulations. We would also like to experiment with a prior on δT. We will then be presented with the same dilemma, where a prior will probably somewhat reduce the amount of temperature fluctuations we are capable of removing. It will, however, also remove less of the CO signal.

---

upon, explaining why these values do not exactly correspond to the values of 1.003 and 1.111 shown in that figure. The values are pretty close though, which is a nice sanity check.

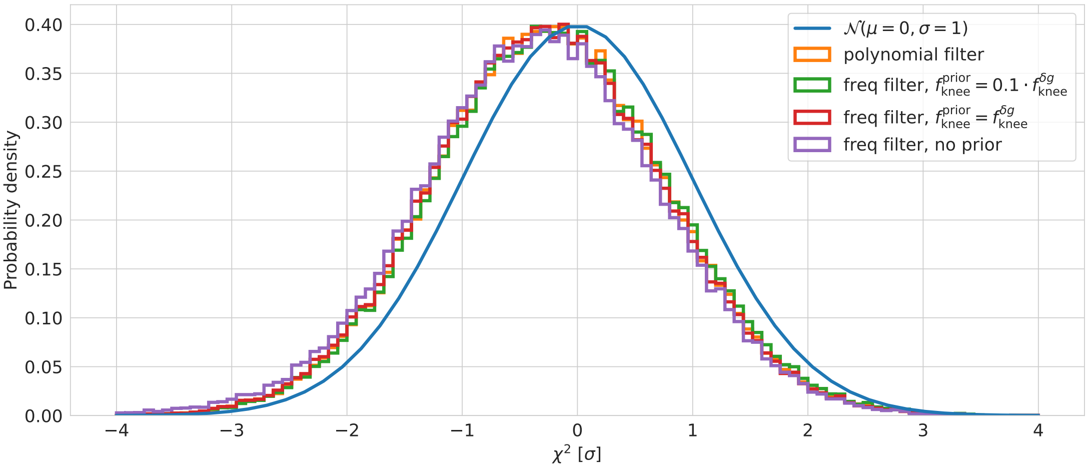


Figure 9.10: Distribution of χ^2 metrics defined in equation 8.6 for the filtered data of both the polynomial filter, and several prior configurations of the frequency filter. A normal distribution is plotted for comparison, which corresponds to a pure white noise distribution in line with the radiometer equation.

9.4 Continuum foregrounds and applications to continuum science

We will finish up our frequency filter chapter with a quick look at the removal of continuum foregrounds. Astrophysical foregrounds are one of the most important systematics to remove properly, as they are completely correlated with galactic coordinates, and will therefore in no way integrate down in time. Figure 9.11 shows the Planck LFI 30 GHz foregrounds at and around the three observational fields. We see that the foregrounds take values of the order of 100 μK or more. Since the CO signal is expected to have amplitudes of only a few μK, we need to suppress the foreground by at least two orders of magnitude for clean detections.

For this analysis, we use Jupiter as a case study, as it is a bright continuum source, and COMAP frequently does Jupiter scans for the purpose of calibration. The new frequency filter should in theory be very capable of removing such continuum sources. We will also take a quick detour and talk about how our frequency filter could have potential applications in continuum science. As COMAP is a line emission survey, we treat temperature continuum sources as contaminations and aim to remove them. For our purposes, the motivation for an accurate fit of continuum temperature sources is simply to remove them more accurately. However, in the field of continuum science, the aim is to map such continuum sources, while treating individual line emissions as contaminations.

For the purpose of studying to what degree we can remove Jupiter from our data, we put a single Jupiter scan through the new frequency filter and identified all points on

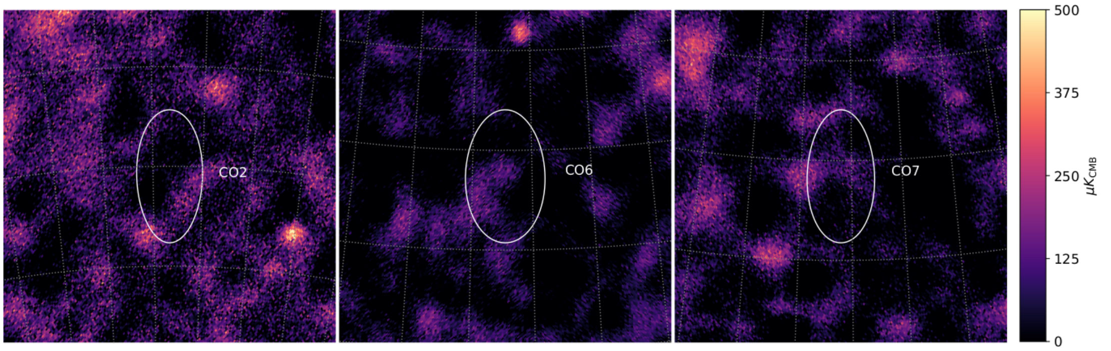


Figure 9.11: The approximate area of each of the three observational fields, overplotted on the Planck LFI 30 GHz full-mission map [56], showing the foreground levels at each field. Figure credit: Nils-Ole Stutzer.

the TOD which hit Jupiter. The temperature distribution of these points before and after the frequency filtering averaged over each sideband, is shown in figure 9.12. While it is not entirely evident from these plots how much of the remaining temperature is from Jupiter, and how much is white noise, we see that we have suppressed the power of Jupiter by *at least* 2 or 3 orders of magnitude. Figure 9.13 shows the binned maps of the different components of the frequency filter fit. Here, we also see that it is impossible to identify Jupiter by eye in the best-fit noise map, which corresponds to the TOD with gain and temperature fluctuations subtracted. Figure 9.14 shows the frequency profile of the central Jupiter pixels, before and after the frequency filter, where we clearly see that it has suppressed to fluctuations around zero, with no signs of bias. We have also included the same frequency profile when the same subtraction is done using the old polynomial filter. This filter is obviously inadequate at removing continuum foregrounds. Calculating the mean squared error of both methods from zero temperature, across frequencies, we get errors of 0.21 K and 0.03 K for the new and old filter, respectively. This corresponds to excesses of 56σ and 1.5σ above the expected value from the radiometer equation, a very significant reduction.

If we look at figure 9.13 from the perspective of a continuum scientist, the temperature fluctuation map is what we would actually be interested in. We could use the frequency filter to separate the gain from the temperature, and treat the upper-right plot in the figure as our result. We see that this plot looks cleaner than the merely binned map does.[4] In a situation such as this, with the desire to remove gain fluctuations from a continuum science map, we could also have employed a destriper mapmaker, which could also serve the purpose of removing gain fluctuations. The two techniques could perhaps also both be employed, in succession. A detailed study of this is outside the scope of this thesis.

---

[4] The gain plot shows a slight anti-correlation with the temperature plot in the center of Jupiter. This suggests that we might have placed a slightly too weak prior on the gain, as we would in such a case expect anti-correlation between δg and δT, like we saw from table 9.1.

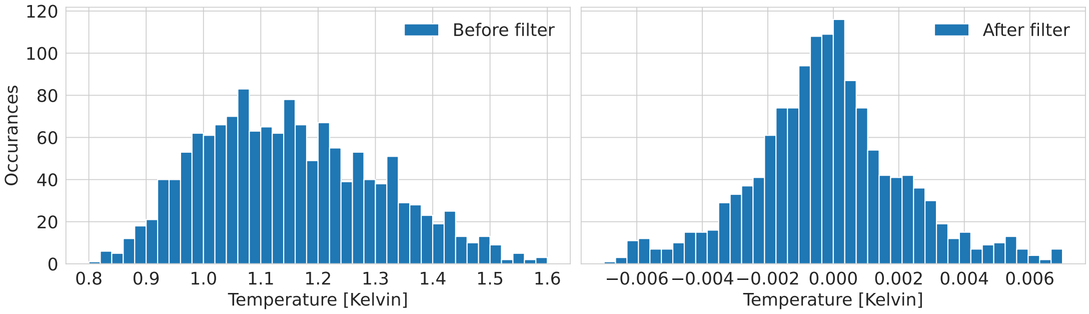


Figure 9.12: Distributions of the average sideband temperature before (left) and after (right) frequency filtering, for TOD points hitting Jupiter, over all 18 feeds and 4 sidebands. The TODs hitting Jupiter have an average temperature of 1.01 K across feeds and sidebands. After filtering, the temperature distribution has a mean of -1×10^{-4} K, and a standard deviation of 0.003 K.

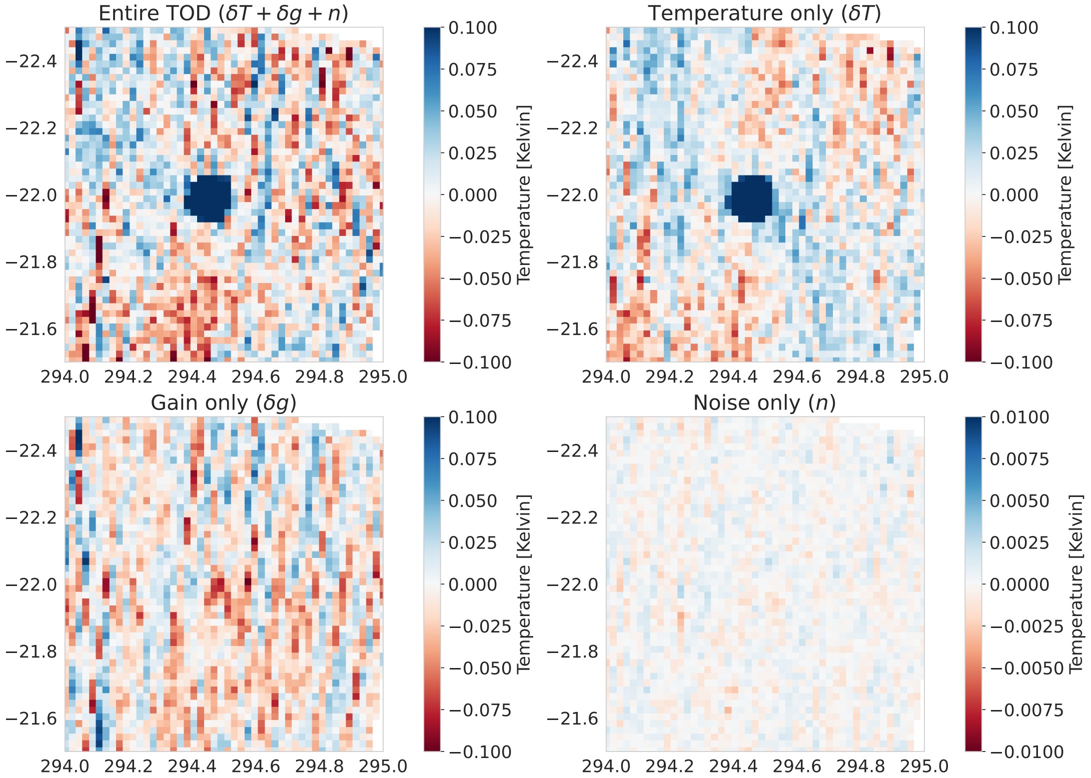


Figure 9.13: Frequency averaged binned maps of the Jupiter scan, as well as the individual parameters in the frequency filter fit.

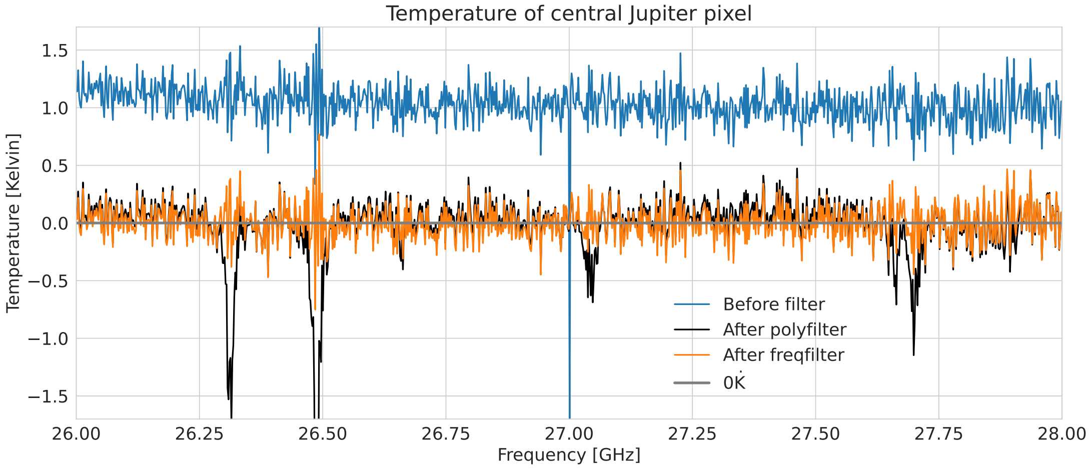


Figure 9.14: Temperature as function of frequency for the central Jupiter pixel of the different maps. The blue line shows the unfiltered TOD, corresponding to the upper-left corner of the previous figure. The orange line shows the frequency filtered map, corresponding to the lower-right corner of the previous figure. The black line shows the result of instead applying the old polynomial filter to the data. $y = 0\,\mathrm{K}$ is overplotted for clarity.

Chapter 10

Conclusion and Future Work

10.1 Conclusion

In this thesis, we have performed an analysis of the COMAP noise characteristics and put forward improved models for two of the currently employed pipeline filters. By analyzing the data as they move through the pipeline, we are able to confirm that our filtered data are almost entirely consistent with white noise, meaning that we have successfully suppressed all signal systematics to below white noise. We find that the correlated noise in our data are well approximated as $1/f$ power spectra and that they are strongly feed-dependent. We also study the power response of the feeds as functions of the angle of the calibration vane, and find that all feeds have reached acceptable power as the vane reaches 69°. Based on this, we developed a new and more robust scheme for calculating the calibration measurement values, and store them in a database.

We have implemented a more sophisticated frequency filter to replace the existing polynomial filter, as the current filter was found to be inadequate at dealing with continuum temperature fluctuations in our data. Most temperature systematics are a big threat to our data modeling, as they are often correlated in telescope pointing or Galactic coordinates, and will therefore bias our results. However, since the gain fluctuations dominate the correlated noise structure, we struggle to measure the level of temperature fluctuations present in our data. We therefore tested the new frequency filter primarily on simulated data, which allowed us total control over the data and noise composition. For all simulated configurations, the new filter removed all temperature fluctuations, whereas the polynomial filter left a large excess correlation with the input temperature in almost all simulation configurations. We further explored the impact of putting a prior on the gain fluctuation fit. We found indications that using a correct prior on the gain removes less CO signal from the data than using no prior does and that an even stricter prior further amplified this, at the cost of a worse gain fit. The new filter was also found to be excellent at removing astrophysical continuum foregrounds. Using Jupiter as a case study, the filter was able to suppress the Jupiter signal by at least 2 or

3 orders of magnitude and produced a mean squared signal residual only 1.5σ from the level expected by the radiometer equation, compared to 56σ for the polynomial filter.

This thesis also lays the groundwork for a new and improved pointing template, based upon data-driven ground maps produced by a destriper mapmaking model. As the current pointing template fits only linear functions separately in azimuth and elevation, it will not accurately be able to fit and subtract more complicated sidelobe ground pickup structures. We produced ground maps both from convolving the simulated telescope beam over the mountain profile around the telescope and from employing a destriper on a large set of scans from the COMAP dataset. We found that the current observational region of $35° < \mathrm{el} < 65°$ avoids most, but not all, of the of the larger ground pickup features. Most notably, we discovered a region around 140° azimuth of the CO2 observational field, with significant ground pickup, that was not entirely consistent with the simulated ground maps. The simulated maps overall proved moderately consistent with different metrics from the `l2gen` pipeline, but we suspect that certain of the simulated beam features might not be entirely accurate. The data-driven destriper maps were also largely in agreement with previous findings, picking up large ground excess at high and low elevations, as well as the noticeable excess in the CO2 field. Comparing the destriper maps to a few real scans, we demonstrate the potential advantages of a data-based pointing template, which may take on a lot more complicated structures in both azimuth and elevation than the linear azimuth fit may. It does this while still being a one-parameter model, and should therefore not increase the CO signal loss of the filter. We also explored some of the challenges of producing an accurate new pointing template model, including degeneracy between atmosphere and ground elevation profiles, and the feed- and frequency-dependency of the maps. We find that the beam has a complicated, non-linear frequency structure, that varies strongly with the distance to the main beam. This produces ground pickup maps with unpredictable frequency structures. We supplement this analysis with frequency derivative maps from the destriper model, where we confirm that the frequency dependence of the ground has a pointing-dependent structure.

10.2 Future work

An overarching goal of the low-level comap analysis is the tighter integration of the filtering steps, into larger filters, which perform join fits on several systematics at once. We have attempted to take a small step in this direction, by including a joint fit of temperature fluctuations in the frequency filter. In the future, we would like to merge the pointing template filter and frequency filter into a joint maximum likelihood fit and keep moving in this direction with the rest of the filters.

We will soon study the impact the new frequency filter has on our data more qualitatively. We have already implemented the filter into the `l2gen` program, but have not yet had the opportunity to test it on the entire COMAP dataset. We will also experiment with the effects of placing a prior on δT, just like we did with δg. This could further reduce the amount of signal we remove, at the expense of some ability to remove δT, which will hopefully be small. On a related note, we plan on sending simulated CO signal through the pipeline with the new frequency filter, to assess whether the new filter negatively impacts the CO signal transfer function.

Our ground pickup work simply laid the groundwork for future efforts, and there is a lot of future work to be done, hopefully ultimately leading to the implementation of a new pointing template filter. While we were able to demonstrate the potential advantages of a data-based approach, we also saw a case of the map predicting a gradient in the opposite direction of a chosen scan. We somehow need an improved mapmaking model, to produce even more accurate ground maps. One approach is to include a prior on the correlated noise, as done in [47], which can substantially improve the map quality. We have a good understanding of the correlated noise power spectrum properties from our work with the frequency filter and can use this to produce an accurate prior. Further, we need to more quantitatively assess the impact of feed- and frequency-dependence of the ground pickup. If our destriper model is good enough, we might be able to produce ground maps for each individual feed. Another important future goal of the destriper is finding a way of better breaking the degeneracy between elevation structures in the ground and the atmospheric elevation profile. In the future, COMAP aims to employ a sensor at the telescope with the sole purpose of measuring τ_0 accurately at every instance. It will then become possible to properly break this degeneracy.

10.3 The outlook of COMAP

COMAP is still in an early stage of its Pathfinder mission, yet to achieve a detection of CO at the targeted redshifts of $z = 2.4 - 3.4$. The successful detection of CO will depend not only on increased observational time but also on proper handling of systematics in our data, as we have demonstrated throughout this thesis. COMAP is already quickly integrating down the noise power spectrum, taking us closer to a future detection of cosmological CO. Figure 10.1 shows the CO power spectrum upper bounds for the first observational year of COMAP data, as well as the optimistic and pessimistic five-year predictions. We have already ruled out several of the more optimistic theoretical models [57, 58], and are in some tension with the COPSS results [59].

In the coming decade, we hope to enter the EoR phase of the COMAP experiment, including an array expansion to include more of the telescopes currently stationed at OVRO. In this phase, COMAP will actively target the C(1-0) emissions at redshifts of $z = 4.5 - 8$, which can be cross-correlated with C(2-1) emissions from the currently observed redshift range (see figure 3.2). With the Epoch of Reionization being a relatively unexplored epoch of our universe, COMAP-EoR is set to bring new and exciting observations into the field of cosmology. A long-term goal of the COMAP collaboration is a dedicated CO mapping satellite, which can probe emissions even further back into the depths of our universe. It truly is an exciting time to be a cosmologist.

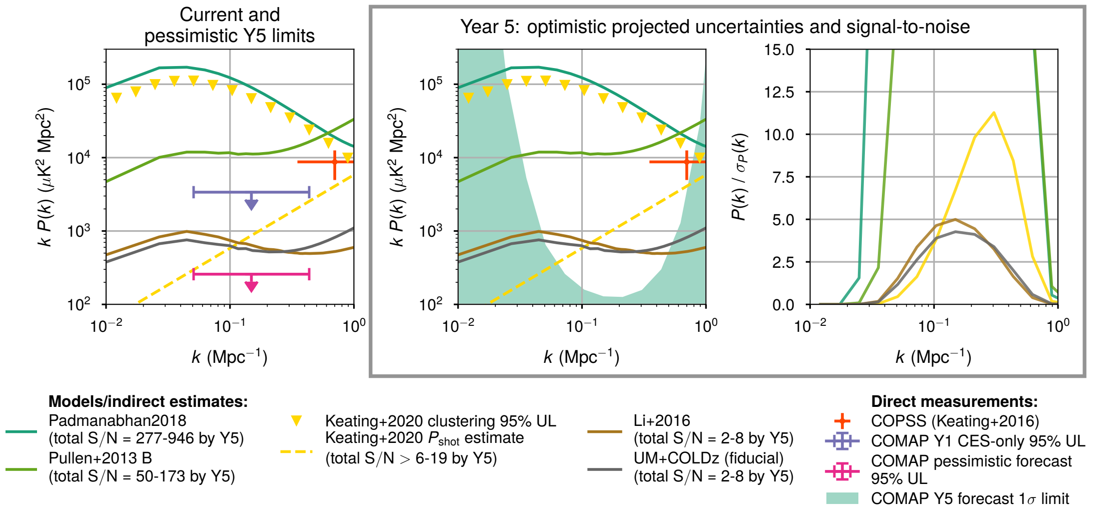


Figure 10.1: Left: The current (purple) and five-year pessimistic forecast (pink) upper bounds on the CO power spectrum by COMAP. Several theoretical CO models are also shown, as well as the COPSS estimate. Middle: Five-year optimistic forecast of the COMAP noise level. Right: The five-year signal-to-noise estimate of the COMAP data, given different theoretical models. Figure credit: Dongwoo Chung

Appendix A

Frequency Filter Destriper Setup

Assume a signal model on the form

$$y_{\nu t} = \delta g_t + \frac{\delta T_t}{\bar{T}_\nu} + n_{\nu t}. \tag{A.1}$$

We have neglected the linear slope term, as to not make the equations too overwhelming, but the equations can easily be extended to account for it. If we flatten the frequency and time dimension, such that $\mathbf{y}$ becomes a one dimensional vector, we get a matrix equation on the destriper form, as

$$\mathbf{y} = \mathbf{Pm} + \mathbf{Fa} + \mathbf{n}. \tag{A.2}$$

Written out, this looks like

$$\begin{aligned} \begin{pmatrix} y_{1,1} \\ y_{2,1} \\ \vdots \\ y_{1,2} \\ y_{2,2} \\ \vdots \\ y_{m-1,n} \\ y_{m,n} \end{pmatrix} &= \begin{pmatrix} \bar{T}_1^{-1} & 0 & \cdots & 0 \\ \bar{T}_2^{-1} & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & \bar{T}_1^{-1} & \cdots & 0 \\ 0 & \bar{T}_2^{-1} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \bar{T}_{m-1}^{-1} \\ 0 & 0 & \cdots & \bar{T}_m^{-1} \end{pmatrix} \begin{pmatrix} \delta T_1 \\ \delta T_2 \\ \vdots \\ \delta T_{n-1} \\ \delta T_n \end{pmatrix} \\ &+ \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \\ 0 & 1 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \delta g_1 \\ \delta g_2 \\ \vdots \\ \delta g_{n-1} \\ \delta g_n \end{pmatrix} + \begin{pmatrix} n_{1,1} \\ n_{2,1} \\ \vdots \\ n_{1,2} \\ n_{2,2} \\ \vdots \\ n_{m-1,n} \\ n_{m,n}. \end{pmatrix} \end{aligned} \tag{A.3}$$

It is now much clearer how this translates into the intuition of the destriper setup. The map we are trying to solve for is the temperature fluctuations δT, while the baselines are the correlated noise δg. Each separate point in time is represented by one pixel, and we hit each pixel 1024 times in a row, for each frequency channel in a sideband. We also only visit each pixel once, over the duration of those 1024 points, which is also the exact length of the baselines. This means that each pixel is resolved completely separately, meaning that the solutions to the destriper equations become explicit. The mechanism we have for breaking the degeneracy between the correlated noise δg and the map δT is instead that the pointing matrix contains a sort of "beam profile", in the form of the system temperature frequency profile. Placing a prior on the correlated noise, which we do, actually brings back some dependency between the pixels.

Bibliography

[1] Albert Einstein. “Die Feldgleichungen der Gravitation. (German) [The Field Equations of Gravitation]”. German. In: (1915), pp. 844–847.

[2] Edwin Hubble. “A relation between distance and radial velocity among extra-galactic nebulae”. In: *Proceedings of the National Academy of Sciences* 15.3 (1929), pp. 168–173. URL: `https://www.pnas.org/content/15/3/168`.

[3] N. Aghanim et al. “Planck 2018 results”. In: *Astronomy & Astrophysics* 641 (Sept. 2020), A6. URL: `http://dx.doi.org/10.1051/0004-6361/201833910`.

[4] Andrew R. Liddle and David H. Lyth. *Cosmological Inflation and Large-Scale Structure*. 2000.

[5] Michael Paul Hobson, George P Efstathiou, and Anthony N Lasenby. *General relativity: an introduction for physicists*. Cambridge University Press, 2006.

[6] *AST1100 lecture notes, lecture 25*. URL: `https://www.uio.no/studier/emner/matnat/astro/nedlagte-emner/AST1100/h09/undervisningsmateriale/lecture25.pdf`.

[7] Alan H. Guth. “Inflationary universe: A possible solution to the horizon and flatness problems”. In: *Phys. Rev. D* 23 (2 Jan. 1981), pp. 347–356. URL: `https://link.aps.org/doi/10.1103/PhysRevD.23.347`.

[8] Scott Dodelson. *Modern Cosmology*. Academic Press, Elsevier Science, 2003.

[9] Ann Merchant Boesgaard and Gary Steigman. “Big Bang Nucleosynthesis: Theories and Observations”. In: *Annual Review of Astronomy and Astrophysics* 23.1 (1985), pp. 319–378. URL: `https://doi.org/10.1146/annurev.aa.23.090185.001535`.

[10] Julio F Navarro. “The structure of cold dark matter halos”. In: *Symposium-international astronomical union*. Vol. 171. Cambridge University Press. 1996, pp. 255–258.

[11] BJ Carr, JR Bond, and WD Arnett. “Cosmological consequences of population III stars”. In: *The Astrophysical Journal* 277 (1984), pp. 445–469.

[12] Xiaohui Fan, CL Carilli, and B Keating. “Observational constraints on cosmic reionization”. In: *Annu. Rev. Astron. Astrophys.* 44 (2006), pp. 415–462.

[13] John H. Wise. “Cosmic reionisation”. In: *Contemporary Physics* 60.2 (Apr. 2019), pp. 145–163. URL: `http://dx.doi.org/10.1080/00107514.2019.1631548`.

[14] URL: `https://commons.wikimedia.org/wiki/File:The_History_of_the_Universe.jpg`.

[15] Adam G. Riess et al. "Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant". In: *The Astronomical Journal* 116.3 (Sept. 1998), pp. 1009–1038. URL: `http://dx.doi.org/10.1086/300499`.

[16] G. Paál, I. Horváth, and B. Lukács. "Inflation and compactification from Galaxy redshifts?" In: *Astrophysics and Space Science* 191.1 (May 1992), pp. 107–124. URL: `https://doi.org/10.1007/BF00644200`.

[17] N. Aghanim et al. "Planck2018 results". In: *Astronomy & Astrophysics* 641 (Sept. 2020), A1. URL: `http://dx.doi.org/10.1051/0004-6361/201833880`.

[18] Jacob D. Bekenstein. "Relativistic gravitation theory for the modified Newtonian dynamics paradigm". In: *Phys. Rev. D* 70 (8 Oct. 2004), p. 083509. URL: `https://link.aps.org/doi/10.1103/PhysRevD.70.083509`.

[19] C. L. Bennett et al. "Four-Year [ITAL]COBE[/ITAL] DMR Cosmic Microwave Background Observations: Maps and Basic Results". In: *The Astrophysical Journal* 464.1 (June 1996), pp. L1–L4. URL: `https://doi.org/10.1086/310075`.

[20] C. L. Bennett et al. "NINE-YEAR WILKINSON MICROWAVE ANISOTROPY PROBE (WMAP) OBSERVATIONS: FINAL MAPS AND RESULTS". In: *The Astrophysical Journal Supplement Series* 208.2 (Sept. 2013), p. 20. URL: `http://dx.doi.org/10.1088/0067-0049/208/2/20`.

[21] R. Adam et al. "Planck2015 results". In: *Astronomy & Astrophysics* 594 (Sept. 2016), A1. URL: `http://dx.doi.org/10.1051/0004-6361/201527101`.

[22] Donald G. York et al. "The Sloan Digital Sky Survey: Technical Summary". In: *The Astronomical Journal* 120.3 (Sept. 2000), pp. 1579–1587. URL: `https://doi.org/10.1086/301513`.

[23] DES Collaboration et al. *The Dark Energy Survey Data Release 2*. 2021. arXiv: `2101.05765 [astro-ph.IM]`.

[24] Collaboration Gaia et al. "Gaia Data Release 2 Summary of the contents and survey properties". In: *Astronomy & Astrophysics* 616.1 (2018).

[25] Ely D. Kovetz et al. *Line-Intensity Mapping: 2017 Status Report*. 2017. arXiv: `1709.09066 [astro-ph.CO]`.

[26] Jason D. Rhodes et al. "The End of Galaxy Surveys". In: *The Astronomical Journal* 160.6 (Nov. 2020), p. 261. URL: `http://dx.doi.org/10.3847/1538-3881/abbe86`.

[27] Ely D. Kovetz et al. *Astrophysics and Cosmology with Line-Intensity Mapping*. 2019. arXiv: `1903.04496 [astro-ph.CO]`.

[28] Fabian Schmidt et al. "Size Bias in Galaxy Surveys". In: *Phys. Rev. Lett.* 103 (5 July 2009), p. 051301. URL: `https://link.aps.org/doi/10.1103/PhysRevLett.103.051301`.

[29] Yuting Wang and Gong-Bo Zhao. "A brief review on cosmological analysis of galaxy surveys with multiple tracers". In: *Research in Astronomy and Astrophysics* 20.10 (Oct. 2020), p. 158. URL: `http://dx.doi.org/10.1088/1674-4527/20/10/158`.

[30] A.T. Crites et al. "The TIME-Pilot intensity mapping experiment". In: vol. 9153. July 2014.

[31] Garrett K. Keating et al. “FIRST RESULTS FROM COPSS: THE CO POWER SPECTRUM SURVEY”. In: *The Astrophysical Journal* 814.2 (Nov. 2015), p. 140. URL: `http://dx.doi.org/10.1088/0004-637X/814/2/140`.

[32] Garrett K. Keating et al. “COPSS: The CO Power Spectrum Survey”. In: *American Astronomical Society Meeting Abstracts #227*. Vol. 227. American Astronomical Society Meeting Abstracts. Jan. 2016, p. 108.05.

[33] Benjamin R. B. Saliwanchik et al. *Mechanical and Optical Design of the HIRAX Radio Telescope*. 2021. arXiv: `2101.06338 [astro-ph.IM]`.

[34] Marie K. Foss et al. *First Season COMAP Results: CO Data Processing*. To be submitted. 2020.

[35] URL: `http://comap.caltech.edu/team.html`.

[36] B. Judd. *Angular Momentum Theory for Diatomic Molecules*. Elsevier Science, 2012. URL: `https://books.google.no/books?id=0yprSc8ivl8C`.

[37] URL: `http://cosmo.fisica.unimi.it/assets/RadioAstro/2017-2018/Radioastro2/Lezione-P02L06-misure-di-idrogeno-molecolare-CO-expanded.pdf`.

[38] Robert B. Leighton. *A 10-Meter Telescope for Millimeter and Sub-Millimeter Astronomy*. Final Technical Report for National Science Foundation Research Grant AST 73-04908. Aug. 1977.

[39] D. Woody, D. Vail, and W. Schaal. “Design, construction, and performance of the Leighton 10.4-m-diameter radio telescopes”. In: *Proceedings of the IEEE* 82.5 (1994), pp. 673–686.

[40] Leven E. “Line-Intensity Mapping with COMAP”. In: (2019).

[41] Håvard Tveit Ihle. “Bayesian Data Analysis for Intensity Mapping and CMB Experiments”. PhD thesis. University of Oslo, 2021. URL: `https://www.duo.uio.no/handle/10852/85984`.

[42] M. B. Weissman. “$\frac{1}{f}$ noise and other slow, nonexponential kinetics in condensed matter”. In: *Rev. Mod. Phys.* 60 (2 Apr. 1988), pp. 537–571. URL: `https://link.aps.org/doi/10.1103/RevModPhys.60.537`.

[43] Mike Folk et al. “An Overview of the HDF5 Technology Suite and Its Applications”. In: *Proceedings of the EDBT/ICDT 2011 Workshop on Array Databases*. AD ’11. Uppsala, Sweden: Association for Computing Machinery, 2011, pp. 36–47. URL: `https://doi.org/10.1145/1966895.1966900`.

[44] James J. Condon and Scott M. Ransom. *Essential Radio Astronomy*. 2016.

[45] Håvard T. Ihle et al. *First Season COMAP Results: Power spectrum methodology and preliminary data quality assessment*. To be submitted. 2021.

[46] E. Keihänen et al. “A maximum likelihood approach to the destriping technique”. In: *Astronomy & Astrophysics* 428.1 (Nov. 2004), pp. 287–298. URL: `http://dx.doi.org/10.1051/0004-6361:200400060`.

[47] E. Keihänen, H. Kurki-Suonio, and T. Poutanen. “Madam- a map-making method for CMB experiments”. In: *Monthly Notices of the Royal Astronomical Society* 360.1 (June 2005), pp. 390–400. URL: `http://dx.doi.org/10.1111/j.1365-2966.2005.09055.x`.

[48] J. Shewchuk. "An Introduction to the Conjugate Gradient Method Without the Agonizing Pain". In: 1994.

[49] URL: `https://www.heywhatsthat.com/?view=BTV9VXUH`.

[50] Benjamin D. Wandelt and Krzysztof M. Górski. "Fast convolution on the sphere". In: *Physical Review D* 63.12 (May 2001). URL: `http://dx.doi.org/10.1103/PhysRevD.63.123002`.

[51] K. M. Gorski et al. "HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere". In: *The Astrophysical Journal* 622.2 (Apr. 2005), pp. 759–771. URL: `http://dx.doi.org/10.1086/427976`.

[52] Andrea Zonca et al. "healpy: equal area pixelization and spherical harmonics transforms for data on the sphere in Python". In: *Journal of Open Source Software* 4.35 (Mar. 2019), p. 1298. URL: `https://doi.org/10.21105/joss.01298`.

[53] URL: `https://casper.ssl.berkeley.edu/astrobaki/index.php/Radiometer_Equation_Applied_to_Telescopes`.

[54] RICHARD F. VOSS and JOHN CLARKE. "'1/fnoise' in music and speech". In: *Nature* 258.5533 (Nov. 1975), pp. 317–318. URL: `https://doi.org/10.1038/258317a0`.

[55] M. S. Bos et al. "The effect of temporal correlated noise on the sea level rate and acceleration uncertainty". In: *Geophysical Journal International* 196.3 (Dec. 2013), pp. 1423–1430. eprint: `https://academic.oup.com/gji/article-pdf/196/3/1423/1569563/ggt481.pdf`. URL: `https://doi.org/10.1093/gji/ggt481`.

[56] N. Aghanim et al. "Planck2018 results". In: *Astronomy & Astrophysics* 641 (Sept. 2020), A1. URL: `http://dx.doi.org/10.1051/0004-6361/201833880`.

[57] Hamsa Padmanabhan. "Constraining the CO intensity mapping power spectrum at intermediate redshifts". In: *Monthly Notices of the Royal Astronomical Society* 475.2 (Dec. 2017), pp. 1477–1484. eprint: `https://academic.oup.com/mnras/article-pdf/475/2/1477/23602609/stx3250.pdf`. URL: `https://doi.org/10.1093/mnras/stx3250`.

[58] Anthony R. Pullen et al. "CROSS-CORRELATIONS AS A COSMOLOGICAL CARBON MONOXIDE DETECTOR". In: *The Astrophysical Journal* 768.1 (Apr. 2013), p. 15. URL: `http://dx.doi.org/10.1088/0004-637X/768/1/15`.

[59] Garrett K. Keating et al. "An Intensity Mapping Detection of Aggregate CO Line Emission at 3 mm". In: *The Astrophysical Journal* 901.2 (Oct. 2020), p. 141. URL: `http://dx.doi.org/10.3847/1538-4357/abb08e`.