

Data Analysis and Systematic Error Mitigation in the COMAP Experiment

Nils-Ole Stutzer



Thesis submitted for the degree of
Master of Science in Astronomy

Institute of Theoretical Astrophysics
University of Oslo

31st May 2021

Copyright © 2021, Nils-Ole Stutzer

This work, entitled “Data Analysis and Systematic Error Mitigation in the COMAP Experiment” 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>.

*"The only difference
between screwing
around and science is
writing it down!"*

- Adam Savage

Abstract

Although modern cosmology has been able to reveal a wealth of information about the Universe, there are still many questions to be answered, such as the details of how the first stars ionized the Universe in the Epoch of Reionization (EoR). The CO Mapping Array Pathfinder (COMAP) is a line intensity mapping experiment aiming to answer this question. Currently in the early stages, much of the present effort of COMAP goes into understanding and mitigating systematic effects. To do this we process the raw telescope data through a series of filters, cleaning it of correlated noise and systematics, before making sky maps. In this thesis, we contribute to this effort by presenting three improvements to the COMAP data analysis pipeline.

The first of these is to extend the COMAP pipeline by a simulation pipeline, with which we can simulate the COMAP telescope picking up additional signal. By estimating the factor by which the simulated signal has been attenuated by filtering the raw data, we estimate the pipeline transfer function. This additional step of the COMAP pipeline is critical as it gives us the possibility to adjust for any underestimated signal power and its error bars. In particular, when computing the ensemble-averaged transfer function from estimates of three different signal realizations as well as ~ 63 h of observational data, the peak transfer efficiency of the filters is found to be 85 – 90% at scales $k_{\parallel}, k_{\perp} \geq 0.1 \text{ Mpc}^{-1}$, while resulting in an almost complete attenuation at scales $k_{\parallel}, k_{\perp} \leq 0.05 \text{ Mpc}^{-1}$. Thus, on these scales, the signal estimates of the COMAP pipeline and its errors should be adjusted upwards by a factor inverse proportional to this transfer function estimate.

Secondly, we implement a baseline destriper in the COMAP mapmaker, which by fitting residual long-timescale modes such as ground pickup, standing waves or unknown systematics, can better resolve the large-scale modes of the astrophysical signal perpendicular to the line-of-sight. Of the tested baseline lengths, the 10 s baseline destriper is found to maximize the transfer function, yielding a peak transfer efficiency of $\sim 95\%$ and outperforming the currently implemented noise weighted binning of the highpass filtered time stream by up to 20 – 25% at the lower k_{\perp} region.

Finally, a principal component analysis of the dataset of feed-feed pseudo-cross spectra, computed from different data splits, is shown to aid in identifying spectra showing signs of systematic effects. Subtracting leading principal components of the data that could correspond to systematic effects is found to result in a modest signal loss of 20 – 50% in a few of the k -bins, introducing a potential new tool in the COMAP pipeline to clean otherwise discarded data of systematic effects.

Acknowledgments

Firstly I would like to extend a great thanks to my thesis supervisors Hans Kristian and Ingunn for their excellent supervision and advice throughout the previous years, and for giving me the chance to participate in their exciting field of research. I would also like to thank Håvard and Marie for all their support and advice, and not least for all their patience when repeatably answering all my stupid questions. In addition, I would like to thank my awesome fellow master students Jonas and Jowita, as well as Kieran, Bade, Duncan and all the rest of the COMAP collaboration, for all the great collaboration and discussions we have had throughout the previous year. It is truly an honor to be a part of COMAP!

I also want to thank all my fellow students and friends Mats Ola, Håkon, Jon, Johan, Markus, Jakob, Bernhard, Lise, Julie, Daniel, Jessie, Christina, Bruce and Ida for all the great company and fun, and not unimportantly for reminding me to be a little social once in a while.

Finally, I would like to thank my family; Thank you Alexa and Jens for being truly great parents and for giving me support whenever I needed it; A special thanks to Joachim who ignited the spark of astronomy that now burns bright within me; Thanks to my brothers Lars, Lennart and Adrian and my cats Ramses and Isis for just being awesome.

Preamble

Whenever humans throughout the millennia have looked into the night sky they have wondered what might lie behind this seemingly unreachable final frontier. However, through the ever-developing chase between technological development and scientific discovery, a great deal has been learned about the Universe. Especially in the last century or so the knowledge about the Universe has progressed at a seemingly exponential speed. That is true both on a theoretical level, but most importantly also on the observational side, through the ever-advancing technological development.

Nevertheless, there are still large parts of the Universe's history that have yet to be systematically researched. One of the eras we currently lack detailed knowledge about is the period when the Universe went from being dark and cold to being ionized by the birth of the first stars and galaxies; The Epoch of Reionization (EoR). As the sources of light from this epoch are extremely far away and correspondingly faint, few detections of individual resolved emitters have been made. This is where the relatively new field of line intensity mapping comes in. By measuring the aggregate line emission of all sources at a low angular resolution, line intensity mapping has the great potential to capture information about the large-scale structure and detailed processes that governed the time of first star formation in a fast and cheap way. One of the line intensity mapping experiments currently ongoing is the CO Mapping Array Pathfinder (COMAP) experiment. COMAP aims to trace star-forming galaxy clusters at the Epoch of Reionization through CO emission from molecular clouds.

However, being in its very early stages of observation, there is a multitude of systematic effects affecting the gathered data. The noise and systematic effects in the data are typically orders of magnitudes stronger than the astrophysical signal itself. Thus, to detect the signal, understanding and accounting for the systematic effects is imperative. To clean the raw time-ordered data from systematic effects it is therefore processed through a series of filters before maps are made. In this light, this thesis will consider some important improvements to the current COMAP pipeline.

Firstly, we will show how to extend the COMAP pipeline to compute the filter transfer function that quantifies how much the filtering of the signal will bias its estimate. This additional step of the pipeline is crucial as COMAP would otherwise underestimate the measured signal strength and its errors. This will be done by building a simulation pipeline that simulates the pickup of additional signal in the telescope, which can be used to feed simulated observations into the COMAP pipeline. By use of this simulation pipeline, the signal prior to and after filtering in the pipeline can be compared to

estimate the transfer function. This subsequently gives the COMAP pipeline the ability to account for the filter-induced signal bias.

Next, we will consider how making maps using a destriper algorithm can better handle residual slowly varying modes in the cleaned time stream and thus increase the signal recovery of large-scale structures, compared to currently used noise weighted binning of the highpass filtered time stream.

Finally, we will investigate the possibility of identifying feed-feed pseudo-cross-spectra (FPXS) with large residual systematic artifacts using a principal component analysis (PCA) on the dataset of cross-spectra that we produce in the COMAP pipeline. By doing so, we will investigate how we can potentially identify and clean data that would otherwise be neglected from the analysis, potentially including a powerful additional data selection and filtering tool in the COMAP pipeline.

This thesis will begin by presenting a short overview of the history of the Universe and cosmology as well as a summary of mathematical cosmology. Thereafter, the field of line intensity mapping and the COMAP experiment will be introduced before presenting our currently implemented COMAP analysis pipeline. Next, we will present the methods used in this thesis to improve the COMAP pipeline by estimating the amount by which we bias the signal by filtering it, by improving the mapmaking of the pipeline using a destriper as well as presenting a way forward to identify and save previously discarded FPXS. The results of this thesis will subsequently be presented and discussed before finally concluding and presenting some suggestions for natural future improvements of the methods and results.

Contents

Abstract	v
Acknowledgments	vii
Preamble	ix
I Background	1
1 A Brief History of Cosmology	3
1.1 The History of the Universe - From its Birth to the Current Age	4
1.1.1 The Very Early Times	5
1.1.2 Big Bang Nucleosynthesis	6
1.1.3 Recombination and the Universe Today	6
1.2 An Introduction to Cosmology in the Language of Mathematics	7
1.2.1 The Fundamental Principles of Cosmology	7
1.2.2 The Theory of General Relativity	7
2 The CO Mapping Array Pathfinder Experiment - COMAP	11
2.1 What is Line Intensity Mapping?	11
2.1.1 Science Goals	13
2.1.2 The Carbon Monoxide Emission Ladder	16
2.1.3 The Problems of Line Intensity Mapping	17
2.2 The COMAP Experiment	17
2.2.1 The COMAP - Pathfinder (Phase I) Instrument	18
2.2.2 Target Fields	20
2.2.3 Scanning Strategies	23
3 The COMAP Data Analysis Pipeline	27
3.1 The Data Modeling and Noise Properties	28
3.2 Level 1 Data	30
3.3 Scan Detect	31
3.4 Level 2 Data	32
3.4.1 The Ambient Load	32

3.4.2	Normalization	32
3.4.3	Removing the Pointing Template	34
3.4.4	Polynomial Filter	36
3.4.5	Principal Component Analysis	36
3.4.6	Frequency Masking	39
3.4.7	Calibration	40
3.4.8	Decimation	41
3.4.9	Highpass Filtering	42
3.5	Mapmaking	42
3.5.1	Maximum Likelihood Estimates	42
3.5.2	Noise Weighted Binning	44
3.6	High-Level Analysis	45
3.6.1	The Theoretical CO Power spectrum and Voxel Intensity Distribution	46
3.6.2	Bayesian Parameter Estimation: MCMC	47
II	Methods	49
4	Estimating Power Spectra	51
4.1	The Auto-Power Spectrum	51
4.1.1	The Pseudo-Power Spectrum	53
4.2	The Cross-Spectrum	55
4.2.1	Data Splits	56
4.2.2	Feed-Feed Pseudo-Cross-Spectrum	57
4.3	The Map Grid	58
5	Improvements to the COMAP Pipeline	61
5.1	The Simulation Pipeline	61
5.1.1	The Pipeline Transfer Function	62
5.1.2	The Upper Branch - Estimating the Noise	64
5.1.3	The Middle Branch - Simulating a Telescope Time Stream	64
5.1.4	The Lower Branch - The Unfiltered Signal Map	67
5.1.5	Computing the Transfer Function Estimate	67
5.2	The Destriper	67
5.2.1	The Destriper Algorithm	68
5.2.2	The Conjugate Gradient Method	70
5.2.3	Implementation of the Destriper	72
5.3	Principal Component Analysis of Feed-Feed Pseudo Cross-Spectra	73
5.3.1	The PCA Decomposition	74
5.3.2	Cleaning FPXS of Residual Systematics	76

III	Results	77
6	Results	79
6.1	The Simulation Pipeline	79
6.1.1	Simulation Pipeline Maps	79
6.1.2	From Auto-Power Spectra to Transfer Function	80
6.1.3	From Feed-Feed Pseudo-Cross Spectra to the Transfer Function	84
6.1.4	Transfer Function for Different Filters and Scanning Strategies	86
6.2	The Improved Mapmaker	91
6.3	The Eigen-Spectrum Analysis	97
6.3.1	The Eigen-Spectra	97
6.3.2	Identifying Cross-Spectrum Outliers	99
6.3.3	Signal Loss due to Principal Component Subtraction	102
7	Conclusion and Future Outlooks	107
7.1	Summary	107
7.2	Future Outlooks	108
	Bibliography	110

Part I

Background

Chapter 1

A Brief History of Cosmology

Since ancient times humankind has wondered about how the world came to be. Long before telescopes were invented people could see mysterious celestial objects such as stars, planets and faint nebulae on the night sky, but no one yet knew the sheer scale of things in space, let alone that there was such a thing as space. Since those ancient times a great many technological inventions have been made, often followed by great scientific discoveries. One of the best-known examples being the invention of the telescope by Dutch spectacle makers in the 17th century that led Galilei to discover that Jupiter had in fact moons orbiting the planet. This clearly showed that not everything in the Universe orbits Earth, an idea which supported Copernicus' heliocentric model of the Universe, in favor of the old Ptolemaic view that the Universe was geocentric [1]. This race between invention and scientific discovery, as we will see later, continues in the modern era of astronomy and cosmology.

In the centuries after Galilei astronomers uncovered more and more about the nature of the Cosmos. It was for instance soon realized and experimentally confirmed in the 19th century, that the stars were, in fact, distant "Suns" within our Galaxy which appeared so faint due to their enormous distance from our Solar System. Hence astronomers would heavily debate the question of how big the Universe actually was. In particular, the question of whether or not the Milky Way Galaxy was the entire Universe, or if there was more, was discussed for a long time. However, soon a strong indication was found that the Universe was far bigger than just the Milky Way when Hubble discovered that the Andromeda nebula in fact had to be its own galaxy as it was too far away to be contained within the Milky Way [2]. It was not until this point, discovering more and more distant galaxies, that cosmologists realized just how truly enormous the scales of the Universe were.

Nonetheless, progress was not limited to the observational side of cosmology. Theorists like Albert Einstein helped firmly establish the modern era of cosmology. In 1915 he published his discovery of the general theory of relativity, the modern theory of gravity, which could be used to mathematically describe the evolution of the Universe as a whole [3, 4]. However, the question of whether the Universe as a whole was a static thing or whether it was evolving dynamically, was still wildly debated. Einstein himself

thought at first the Universe was static, but in 1929 Hubble discovered that the spectra of distant galaxies all seemed to be redshifted, suggesting that they moved away from each other. The conclusion had to be that the very space in which the galaxies were situated was expanding. Soon after, in 1931, the Belgian cosmologist George Lemaître used this observation, together with equations that he and Alexander Friedmann had independently derived from Einstein's work, to claim that if space is expanding the content of the Universe must have been packed closely together at earlier times [2]. The idea that the Universe had sprung from some sort of infinitely dense primordial singularity was soon known as the Big Bang paradigm, a term initially coined by Fred Hoyle who was a supporter of the adversary steady-state theory [2].

However, the discovery of the Cosmic Microwave Background (CMB), the last remnants of the very first light that traveled through the Universe, by Penzias and Wilson in 1965 [5] firmly cemented the Big Bang paradigm as the currently accepted theory. This discovery was also the birth of modern cosmology as a precision science after which many new insights into the nature of the Universe have been made. In recent decades, observations of supernovae of type Ia by Riess et al. [6] and Perlmutter et al. [7] have shown that not only is the Universe expanding, but it is accelerating as well. This suggests that the expansion of the Universe is driven by some sort of negative pressure component often referred to as dark energy or vacuum energy. However, this dark sector of the Universe also seems to include some sort of non-baryonic matter only interacting gravitationally with regular baryonic matter (e.g. protons and neutrons). It is needed to consistently explain the observations of galactic rotation curves [8] and for primordial density fluctuations to collapse under gravity into the structures we observe today. This model of the Universe in which dark energy (Λ) and cold dark matter (CDM) are the dominating components became the standard model of cosmology called the Λ CDM model. Further CMB experiments such as COBE, WMAP and Planck have since delivered pristine insight into the nature of the Universe, tightly constraining the amount of each matter-energy component in the Universe and hence strengthening the Λ CMD model [9]. However as there are still some tensions left between the different cosmological experiments, especially between the supernova and CMB experiments, new methods such as intensity mapping could, as we will discuss later, be highly complementary and resolve some of the tension [9, 10].

Unless otherwise stated the following sections will be based on the theory presented in Carroll [11], Dodelson [12] and Schneider [13].

1.1 The History of the Universe - From its Birth to the Current Age

Now that we have taken a look at how Universe has been researched through history we will take a look at the current understanding of the history of the Universe itself. The history of the Universe can be roughly divided into three main parts; the very early times right after the Big Bang, the time between the very first fractions of a second up until the Epoch of Recombination when the CMB became to be, and lastly the era

after recombination.

1.1.1 The Very Early Times

Evidence that the Universe is expanding suggests that if one follows time backward, all of the Universe and its content may have been packed together infinitely close; a singularity. We do not know much about the very beginning of time, let alone if there was such a thing as the beginning of time, however it is still useful to consider such an extrapolation to $t = 0$ as a reference point. The details of what happened at the very first instance of space and time, in what is commonly known as the Big Bang, are therefore largely unknown as the conditions were so extreme that our current theories of physics break down.

Nonetheless, we know that the Universe in the pre-recombination era was very hot, dense and smooth, only having small temperature and density perturbations of the order 10^{-5} from the mean. This remarkable smoothness is even observed on the CMB across angular scales larger than 1° on the sky today. At the time of recombination this scale, known as the *Hubble sphere*, characterizes the maximum distance between two causally connected points. However, to establish thermal equilibrium between two such points photons must have been able to travel the distance between them. Yet seemingly this should not be possible as the two regions are causally disconnected. This issue, called the *horizon problem*, is not the only one. Another peculiarity observed is the so-called *flatness problem*. Today we observe that the Universe is spatially flat (i.e. parallelly emitted light rays forever remain parallel), only having a very small curvature. Due to the expansion of space, this curvature must be decreasing when looking back in time, until it is not even measurable at very early times. However, to obtain such remarkably flatness, the Universe must have had very special initial conditions, unless there is some process that can smooth out any initial curvature. A phenomenon that could potentially solve these problems is a short phase of exponential expansion in the early Universe; cosmic inflation.

The phase of cosmic inflation is believed to have taken place a very short time after the Big Bang when the Universe was only about $\sim 10^{-34}$ s old. By the following brief period of exponential expansion, the Universe increased in scale by at least 60 *e*-foldings (increasing by a factor e^{60}) and was finished by the time the Universe was $\sim 10^{-32}$ s old. Such an enormous expansion would both smooth out any initial spatial curvature and cause all of, what we today call, the observable universe to be initially in causal contact. Thereby it would solve both the flatness and horizon problems. What exactly caused inflation is largely unknown, but a likely candidate is a so-called inflation scalar (inflaton) field "rolling down" a potential. The initially large potential energy of the inflaton caused the exponential expansion of space, and this expansion was in the end halted by the inflaton decaying into standard model particles in the reheating process. Small quantum fluctuations in the primordial vacuum thus became blown up and transformed to the classical perturbations seen in the CMB. The fluctuations of the CMB in turn seeded the formations of structures like stars and galaxies we see today.

1.1.2 Big Bang Nucleosynthesis

At the end of inflation, the first particles and anti-particles were produced, which in turn would annihilate to create photons. At this early stage, the temperature of the Universe was so high that any nuclei or atoms forming would instantly be blown apart by the high-energy photons dominating the Universe. As the Universe expanded the temperature dropped, eventually reaching $k_B T \sim 1 \text{ MeV}$, at which point the cosmic neutrinos could decouple from the hot and dense plasma. Yet, it was not cold enough for nuclei or neutral atoms to form. It was not until the temperature fell below $k_B T \sim 0.1 \text{ MeV}$ that photons had sufficiently low energy that light atomic nuclei could form without instantly being disintegrated. In this epoch called the Big Bang Nucleosynthesis (BBN), the relative abundance of the light atoms like hydrogen, deuterium and helium we see today was laid down. The remarkable agreement between the predicted (by BBN) and observed light element abundance, as well as the mass fraction of baryons in the Universe, is one of the great successes of the Big Bang paradigm [9].

1.1.3 Recombination and the Universe Today

After having formed the first atomic nuclei the Universe was still much too hot for neutral atoms to form. In the hot photon-baryon plasma the photons were tightly coupled to the electrons through Compton scattering, while in turn the electrons and protons (and other nuclei) were tightly coupled through Coulomb scattering, effectively maintaining the equilibrium. Any electron binding to a nucleus was very efficiently ionized again, as there were many more photons than electrons. However, while the formation of baryonic structure was not yet possible, the earlier decoupled cold dark matter could begin to cluster and start forming gravitational potential wells. Once the temperature of the Universe dropped well below the ionization energy of hydrogen at 13.6 eV , electrons and protons could bond to form the first neutral atoms. At this point, about 380 000 years after the Big Bang, called recombination, the Universe became transparent and photons could now free-stream to form what we today can see as the Cosmic Microwave Background (CMB). The baryonic matter perturbations, still being very small, could now fall into the dark matter potential wells and grow into the structures we see today.

The following period is called the Cosmic Dark Ages, in which the Universe was cold and filled with neutral hydrogen and helium, the only radiation being CMB photons as well as 21 cm radiation from neutral hydrogen. It took a few 100 million years before the Dark Ages ended and the Cosmic Dawn begun with the first stars and galaxies [14, 15]. Soon after these first massive hot stars ionized more and more of the neutral Universe in what is commonly known as the *Epoch of Reionization*, until finally, the whole Universe was ionized. Nevertheless, the Universe was still matter-dominated. It was not until relatively recently that the Universe gradually became dominated by some dark energy causing space to expand at an exponential rate.

1.2 An Introduction to Cosmology in the Language of Mathematics

Now that we have seen the outlines of how the Universe became to be and has developed ever since, as well as an overview of the scientific research done over the ages, it is time to dive a little deeper into the laws and principles that govern our Universe in terms of mathematical language. Modern cosmology is based on only a few fundamental principles, which when applied to Einstein's theory of general relativity produces the equations needed to describe much of how the Universe evolves.

1.2.1 The Fundamental Principles of Cosmology

The foundation of modern cosmology rests on only a few principles; the Copernican principle, the principle of isotropy and the theory of general relativity. The first of these states that no position in the Universe is unique and that all observers are equivalent. The second principle states that the Universe as seen by any observer looks the same in any direction on large enough scales, i.e. it is isotropic. Together these two principles imply that the Universe must also be homogenous. This combined principle is called the cosmological principle. The third principle governs the evolution of space and time through the theory of general relativity as governed by its matter-energy content.

1.2.2 The Theory of General Relativity

The birth of modern mathematical cosmology can be seen as the discovery of the theory of general relativity (GR) by Albert Einstein [3, 4]. The essence of his theory states that the matter and energy contained in a spacetime manifold will curve space and time. In turn, the geometry of spacetime will then govern how matter and energy move under the influence of gravity. This geometrical interpretation of gravity can be formulated in terms of Einstein's field equations,

$$G_{\mu\nu} + \Lambda g_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}\mathcal{R} + \Lambda g_{\mu\nu} = 8\pi GT_{\mu\nu}, \quad (1.1)$$

where the Einstein tensor $G_{\mu\nu}$ on the left-hand-side (l.h.s.) describes spacetime's geometry and is given the Ricci tensor $R_{\mu\nu}$, the Ricci scalar \mathcal{R} being the contraction of the Ricci tensor $g^{\mu\nu}R_{\mu\nu}$, as well as the metric tensor $g_{\mu\nu}$. The metric tensor describes the physical distance between two events in spacetime and can hence be seen as the fundamental quantity describing spacetime geometry. The Ricci tensor and scalar are dependent on the metric and its derivatives only. The cosmological constant is denoted by Λ and is needed to produce for instance an accelerating expansion of the Universe. The right-hand-side (r.h.s.) of the equation is given by the energy-momentum tensor $T_{\mu\nu}$ describing all matter-energy content of the Universe and the gravitational constant G . The mutual dependence of spacetime geometry and matter-energy can now easily be noticed, as the l.h.s of the equation only depends on the metric and its derivatives and therefore the geometric properties of spacetime. Meanwhile the r.h.s. contains only matter-energy terms.

To obtain the solution to Einstein's equation most interesting for cosmology we simply apply the cosmological principle, i.e. a smooth homogenous and isotropic energy-momentum content of the Universe. Writing out the metric tensor as a line element one then finds it to be given by

$$ds^2 = dt^2 - a^2(t) \left(\frac{dr^2}{1 - kr^2} + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 \right) \quad (1.2)$$

in spherical coordinates. This particular line element is called the Friedmann-Lemaître-Robertson-Walker (FLRW) line element. One can see that as opposed to flat Euclidean space, the distance between two spacetime events ds is not only dependent on the temporal dt , radial dr and angular distances $d\theta$ and $d\phi$, but also on the scale factor $a(t)$ and curvature parameter k . The scale factor $a(t)$ relates the comoving coordinates (t, r, θ, ϕ) to the physical (or sometimes called proper) coordinates. The former of which make up a coordinate system that remains constant under the expansion of the Universe, while the latter characterizes the physical distances between different points. Thus $a(t)$ is a measure of how much space has expanded. Meanwhile, the curvature parameter, taking the values $k \in \{-1, 0, 1\}$ determines whether the Universe is open and negatively curved, open and flat, or closed and positively curved respectively. However, as has been measured to very high precision by Planck [9], the Universe is remarkably flat. We will hence only consider flat universes hereafter.

From the FLRW metric and Einstein's equations one can now find the two powerful Friedmann equations

$$H^2(t) \equiv \left(\frac{\dot{a}(t)}{a(t)} \right)^2 = \frac{8\pi G}{3} \rho(t) - \frac{k}{a^2(t)} \quad (1.3)$$

$$\frac{\ddot{a}(t)}{a(t)} = \frac{4\pi G}{3} (\rho + 3p), \quad (1.4)$$

the first of which relates the matter-energy content ρ of the Universe and its curvature k the expansion rate of the Universe $H(t)$, called the Hubble parameter. The second equation, meanwhile, characterizes the acceleration of the Universe expansion caused by the matter-energy content ρ exerting a pressure p . The pressure and matter-energy densities are in turn related through the equation of state $p = w\rho$, which when using the cosmological continuity equation,

$$\frac{\dot{\rho}}{\rho} = -3(1+w)\frac{\dot{a}}{a}, \quad (1.5)$$

gives the density as a function of the spatial expansion,

$$\rho \propto a^{-3(1+w)\frac{\dot{a}}{a}}. \quad (1.6)$$

The equation of state parameter w takes the values 0, 1/3 or -1 for a non-relativistic and relativistic fluid as well as vacuum energy, respectively.

Using this we can define the densities in terms of the critical density, i.e. the density needed to obtain a completely flat Universe; $\Omega_i \equiv \frac{\rho_i}{\rho_{\text{crit}}}$, where the critical density

$\rho_{\text{crit}} \equiv \frac{3H_0^2}{8\pi G}$. The Hubble constant $H_0 = 67.37 \pm 0.54 \text{ km s}^{-1} \text{ Mpc}^{-1}$ simply gives the present day expansion rate of the Universe [9]. The first Friedmann equation in terms of these density parameters becomes

$$\frac{H^2(t)}{H_0^2} = \sum_i \Omega_i(t) = \Omega_{B,0} a^{-3} + \Omega_{CDM,0} a^{-3} + \Omega_{\gamma,0} a^{-4} + \Omega_{\nu,0} a^{-4} + \Omega_{\Lambda,0}, \quad (1.7)$$

where the sum of the density parameters must sum to unity if the spatial curvature is zero, i.e. $\sum_i \Omega_{i,0} = 1$ at present (or any other) time. Here we have explicitly written the equation in terms of the five energy-matter constituents of the Universe in the Λ CDM model; baryonic and cold dark matter (CDM), photons, neutrinos as well as the cosmological constant (Λ) also commonly known as dark energy.

An often-used quantity that is both easier to grasp and measure than the scale factor $a(t)$ is the cosmological redshift z . When a photon is emitted and travels a given distance through space and time, it is subject to the spatial expansion of the Universe. The spatial expansion consequently stretches the wavelength λ of the photon, hence redshifting it. One can therefore express this redshift z in terms of the scale factor at the time of emission t_{em} and receiving t_{rec} as

$$1 + z \equiv \frac{\lambda_{\text{rec}}}{\lambda_{\text{em}}} = \frac{\lambda_c a(t_{\text{rec}})}{\lambda_c a(t_{\text{em}})} = \frac{a(t_{\text{rec}})}{a(t_{\text{em}})}, \quad (1.8)$$

since the wavelength of the photon at any time can be written as a comoving wavelength λ_c times the scale factor at that time. Usually the scale factor today $a_0 \equiv 1$, often used as $a(t_{\text{rec}})$ if we receive photons at present time.

As mentioned earlier, the scale factor, and consequently the redshift, are most commonly used to transform between comoving and physical distances. Consider for instance the comoving distance, which is a distance scale that expands with the Universe, simply found by setting $ds = 0$ in Eq. (1.2) for photons and subsequently integrate over time, scale or redshift;

$$r_c(t) = \int_{t_{\text{em}}}^t \frac{dt'}{a(t')} = \int_{a(t_{\text{em}})}^{a(t)} \frac{da'}{a'^2 H(a')} = \int_z^{z(t_{\text{em}})} \frac{dz'}{H(z')}. \quad (1.9)$$

The corresponding physical distance r_p is then simply found by multiplying r_c by a or $1/(1+z)$ at any given time t .

There are, however, two additional distance measures often used as they depend only on measurable quantities. The first of these is the angular diameter distance r_A which is the distance to an object, of proper size D , needed for it to extend an angular distance θ on the sky. In a flat universe this simply becomes related to the comoving and proper distances as

$$r_A = \frac{D}{\theta} = \frac{r_c}{1+z} = r_p, \quad (1.10)$$

and can be determined from observables only if the size D of the observed object can be assumed to be known.

The second important distance measure is called the luminosity distance and is defined to be the distance r_L to an object of luminosity L needed to observe the measured energy flux F , thus satisfying the equation

$$F = \frac{L}{4\pi r_L^2}. \quad (1.11)$$

If we draw a sphere of comoving radius r_c around the source with luminosity L , all emitted photons will at some time hit this sphere. However, as time goes on and physical space expands, a photon has to travel further to transverse the same comoving distance. As a result, photons will cross the spherical surface less frequently as time goes on. In addition, spatial expansion redshifts the photons, together resulting in a flux

$$F = \frac{L}{4\pi r_c^2(1+z)^2}, \quad (1.12)$$

which implies that the luminosity distance must be

$$r_L = (1+z)r_c. \quad (1.13)$$

Again we see that if we can assume to know the luminosity of an object and measure its flux F , we can directly infer its distance to us. Objects of known size or luminosity are therefore often called *standard rulers and candles* respectively, and are important for observational cosmology.

Chapter 2

The CO Mapping Array Pathfinder Experiment - COMAP

2.1 What is Line Intensity Mapping?

Modern cosmology has shed a lot of light onto the nature of the Universe we live in and most cosmological parameters are known to within the one percent level through the combined power of CMB [9], galaxy redshift [16], supernova Type Ia surveys [6, 7] and other observations. However, even though we know much about the Universe as a whole there is still a lack of detailed understanding of the astrophysics on smaller scales, like star and galaxy formation, within different epochs of the Universe. Gaining a deeper understanding of these is in and of itself important to obtain a deeper model of the whole Universe, but it is also important as the local astrophysics affects the global cosmology.

Thus, there are for instance still some degeneracies between cosmological and astrophysical parameters, such as the degeneracy between the optical depth of the Universe at the Epoch of Reionization τ and the scalar fluctuation amplitude A_s , which current experiments have not yet been able to constrain well due to the lack of detailed knowledge about the astrophysics of Reionization [10]. In addition to parameter degeneracies, there are still some tensions between CMB and lower redshift surveys. An example is the disagreement between the value of H_0 measured by the CMB and that inferred from Type Ia supernovae [17].

Hence we would like to compliment the existing astrophysical and cosmological surveys by probing directly the Epoch of Reionization and the large-scale structure of the Universe, to gain stronger leverage on parameter degeneracies, help alleviate tensions, and most importantly to develop more sophisticated models of the astrophysical processes within our Universe. This is where line intensity mapping comes in.

Line intensity mapping is a relatively new and promising field in observational astronomy that can provide valuable information, both cosmological and astrophysical, from a wide range of redshifts and scales. Intensity mapping aims to integrate all photons emitted from molecular and atomic line emission, like 21 cm, Ly α , [CII] as well as CO

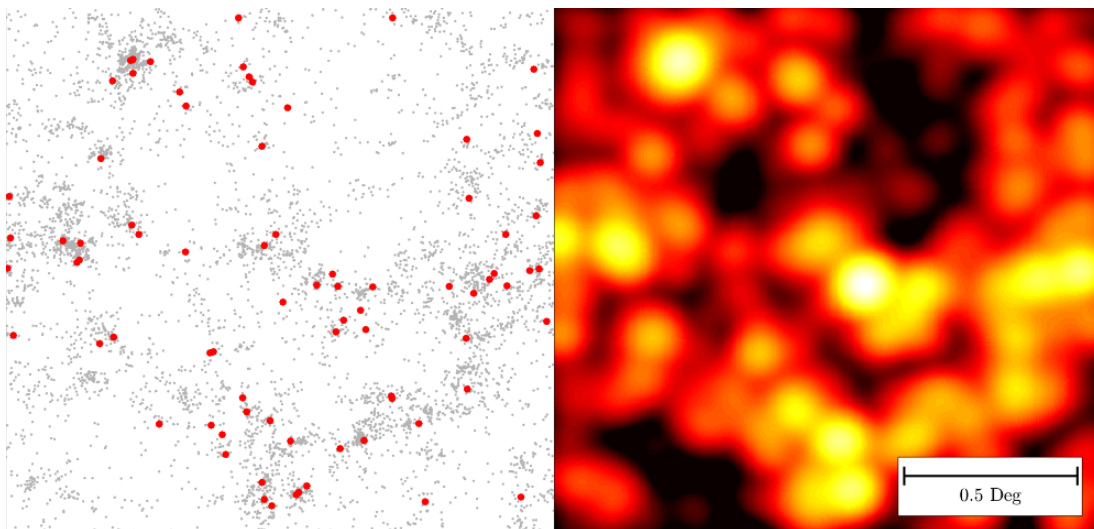


Figure 2.1: Illustration of the difference between galaxy redshift surveys and line intensity mapping. The left panel shows a distribution of galaxies with luminosities drawn from a Schechter distribution. The brightest of the galaxies, being most accessible to galaxy surveys, are marked in red. The right panel shows the corresponding CO line intensity map, which follows the distribution of galaxies in the left panel, yet includes faint emitters in between the galaxies as well. Figure from Kovetz et al. [10]

rotational lines, from both bright and faint sources across a larger field of the sky [10]. This way not only a 2D map of spatial line intensity fluctuations can be produced, but also the fluctuations along the line-of sight through the redshift of the lines. The main advantage of intensity mapping over galaxy surveys is that an intensity mapping survey is significantly cheaper to perform as it is faster in covering the sky, while not being limited to bright sources only, making its intensity maps cheap and unbiased [10].

As the scales probed are generally quite large, there is a trade-off between the statistical information needed for cosmology and losing the details of the small-scale physics of the environment of the emitters themselves [18]. Because a galaxy survey is limited to a high angular resolution and needs to spend a longer time integrating on each source to claim a high signal-to-noise ratio (SNR) detection. Meanwhile, the lower angular resolution and SNR requirements of intensity mapping make it a much faster and more economical survey technique [10].

Another problem with galaxy surveys is that only bright galaxies are probed, hence being biased towards the bright end of the galaxy distribution. Line intensity mapping, however, having an unbiased nature as it integrates both faint and bright sources, can thus capture details lost in traditional galaxy surveys [10]. The difference between a galaxy and intensity mapping survey is illustrated in Fig. 2.1. This figure nicely illustrates how the intensity map of a simulated galaxy field follows the distribution of discrete emitters and also includes all diffuse emission sources between the galaxies. One can, however, see that the individual galaxies, even the brightest ones marked as red

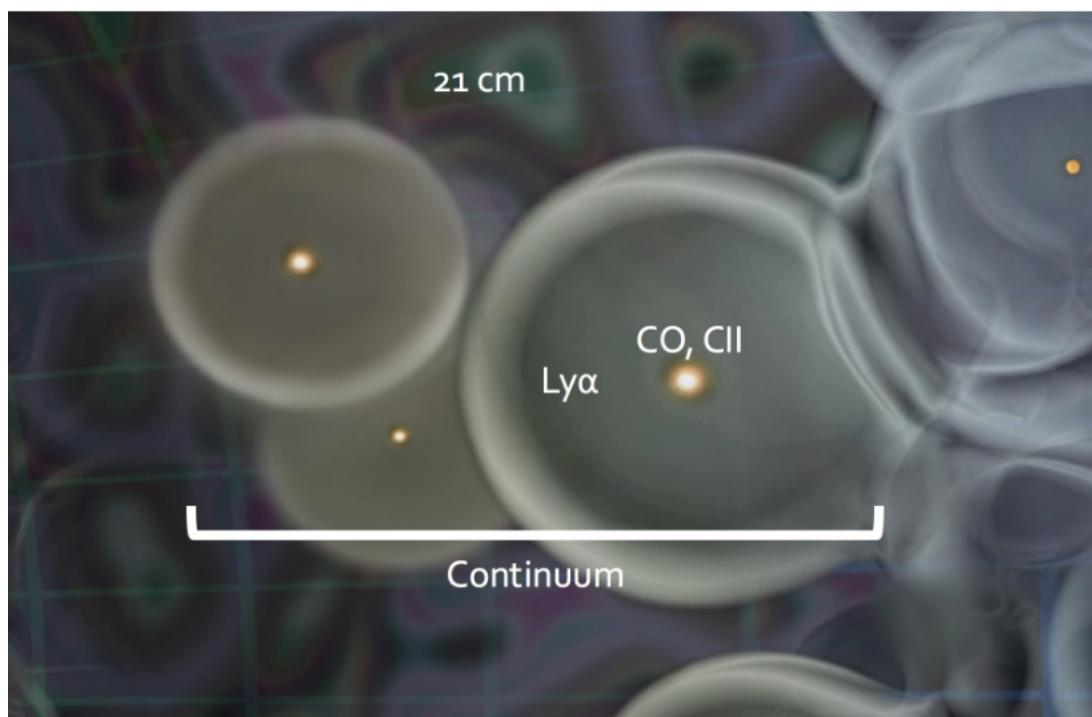


Figure 2.2: Illustration of the interplay between the different sources of line emission and how they are correlated across different scales and parts of the local environment. Inside ionizing halos, one can see how star-forming galaxies glow in CO and [CII], ionizing the surrounding neutral intergalactic medium. The ionized regions within the halo are thus traced by Ly α radiation, while the neutral surroundings emit 21 cm radiation. Figure from Kovetz et al. [10]

dots, are themselves not resolved in the intensity map. Nevertheless, mapping out the faint emission is especially important since it for instance is believed that much of the interesting physics of baryonic matter, star formation and the processes of reionization happened in the diffuse intergalactic medium (IGM) and low-mass galaxies at high redshifts [17, 19, 20].

2.1.1 Science Goals

Now that we have discussed what line intensity mapping is, we will venture a little deeper into the scientific goals intensity mapping aims to achieve. Much of the initial efforts in line intensity mapping have been to observe the Epoch of Reionization and Cosmic Dawn out to redshift $z \sim 10$ [10]. However, because intensity mapping surveys generally have a large field of view and are sensitive to even faint emission, there are several other interesting goals line intensity mapping aims to achieve. These scientific goals can roughly be divided into three sectors; the Epoch of Reionization; star and galaxy assembly; as well as large-scale structure and cosmology [10].

The Epoch of Reionization

As mentioned earlier, the Universe after recombination was largely cold and neutral. However, we see evidence, in the form of a characteristic Gunn-Peterson trough in the spectra of distant quasars as well as a smoothing of the CMB structures, that the Universe changed to an ionized state at some point [21, 9].

Even though we know the Universe must at some point have been reionized, with some constraints from CMB and quasar surveys, the detailed processes and history causing the reionization are still largely unknown. This is where line intensity mapping comes in.

It is believed that the Universe was progressively reionized by massive low-metallicity, Pop III stars residing in faint dwarf galaxies [19]. Around these galaxies, the neutral atomic hydrogen gas would be ionized to form ionization bubbles, as seen in Fig. 2.2. The neutral hydrogen gas surrounding the ionized bubbles is mainly traced by 21 cm radiation, while the ionized bubble and the galaxies inside it will be traced by the Ly α line. Meanwhile, the galaxies in the center of the ionization region can be traced by CO rotational lines and the [CII] fine-structure line of ionized carbon, lines which both trace star-forming regions within galaxies.

By using intensity mapping surveys targeting each of these lines, a wide range of scales and processes can be effectively probed. However, the real power of intensity mapping is to combine these individual surveys and cross-correlate them to reveal a detailed interplay of ionizing star-forming galaxies and their surrounding Universe. Thereby the details of how local astrophysical processes affect cosmological parameters can be better understood [10].

Some examples of current and future intensity mapping surveys include HERA and SKA-LOW, which are interferometric arrays that aim to trace 21 cm radiation at redshifts $z = 5 - 27$ and $z = 3 - 7$ respectively, with among others the goal of directly imaging the reionization bubbles [22, 23, 10]. Other probes that survey [CII] and/or CO lines at redshifts between $z = 3.3 - 9.5$ are for instance CCAT-prime ([CII]) [24, 25], TIME ([CII] and CO) [26, 27] and CONCERTO ([CII] and CO) [28], which aim to constrain the processes of how star formation affected reionization [10].

Star Formation and Galaxy Assembly

Another promising area of research is to constrain the star formation rate inside galaxies around the time of peak star formation, also known as the Epoch of Galaxy Assembly at redshift $z \sim 2 - 3$ [10].

Most of the knowledge of star formation at high redshift is based on the emission of hot ionized gas in the interstellar medium (ISM) or direct starlight from bright galaxies. Hence, the results from such galaxy surveys will be highly biased towards bright emission sources. However, much of the star formation at high redshift is happening inside the cold molecular clouds within relatively small galaxies. These are traced by CO and [CII] emission too faint to be observed by galaxy surveys, but well within the capabilities of an intensity mapping survey [10].

By using a CO or [CII] intensity mapping experiment, the cumulative effect of a large number of faint emitters on their surrounding can be used to constrain the star formation rate density (SFRD) across different epochs. To connect the observed CO intensity to some SFRD one can use empirical relations between CO luminosity and the far-infrared (FIR) emission of a galaxy, which in turn is related to the SFRD [29].

Some noteworthy experiments targeting star formation and galaxy assembly observing CO lines are COPSS [30, 31], mmIME (CO and [CII]) [32, 10] and COMAP [33, 34] at respective redshift ranges $z = 2.3 - 3.3$, $z = 1 - 5$ and $z = 2.4 - 3.4$. There is also a mission on its way to launch a space telescope called SPHEREx targeting among others, $H\alpha$, $H\beta$, [OII] and [OIII] as well as $Ly\alpha$ at redshifts $z = 0.1 - 5$ and $z = 5.2 - 8$ respectively [35, 36]. SPHEREx aims to investigate the galaxy formation history since the Epoch of Reionization, star formation rates (SFR), large-scale structure and the inflationary history of the Universe. A secondary goal of SPHEREx is to research the origin of water in planetary systems [35, 10].

Large Scale Structure and Cosmology

As most of the baryonic matter in the Universe resides in faintly glowing gas tracing the underlying dark matter field on large linear scales, an intensity mapping survey could produce unbiased estimates of the large-scale dark matter distribution fast and economically. This is mainly done by observing the 21 cm spin-flip transition of neutral hydrogen in the IGM. Subsequently, the statistical information of the observed large-scale structure can be used to constrain cosmological parameters [10].

In particular one can use the baryonic acoustic oscillations (BAO). The BAO are remnants of sound waves that traveled through the primordial plasma at relativistic speeds. Consider for instance a primordial density perturbation in the matter-energy field, containing both photons, baryons and dark matter. As photons and baryons are affected by pressure, while dark matter is not, the photons and baryons will be pushed radially outwards from the central perturbation. Meanwhile, as the photon-baryon plasma travels away from the perturbation, the dark matter part of the perturbation stays behind as it is not affected by pressure [16, 13].

At recombination, when the photons decoupled from the baryons, the baryons froze in as they rapidly became non-relativistic. The distance this wave traveled from the Big Bang until decoupling is called the sound horizon. Due to the baryonic and dark matter perturbation remnants coalescing after recombination, the enhanced baryon density at the sound horizon scale led to a peak in the correlation function of matter [16].

Therefore, by mapping the large-scale matter field with intensity mapping, one could potentially recover the BAO peak at redshifts and scales not yet probed by galaxy surveys [10]. Using this, one can construct a powerful standard ruler calibrated at the CMB to form an inverse distance ladder, as the BAO clustering will always remain at a constant comoving scale of $\sim 100h^{-1}\text{Mpc}$ [16, 37]. The result would be very powerful in researching for instance the nature of dark energy and break degeneracies of existing cosmological parameters at low and intermediate redshifts.

Telescopes targeting this epoch include for instance CHIME and HIRAX both tar-

getting the 21 cm at redshifts $z = 0.8 - 2.5$, where they both aim to map out large HI structures to investigate the expansion rate of the Universe and dark energy through BAO [38, 39, 40, 41]. These two telescopes are also examples of experiments with unexpected secondary science goals as it turns out that they are ideally built for real-time analysis of the sky. Through their interferometric nature, transient events such as pulsars and fast radio bursts may be detected effectively [42, 43, 39].

2.1.2 The Carbon Monoxide Emission Ladder

As the work presented in this thesis is a contribution to the COMAP experiment, we will now take a closer look at why the CO emission lines, which COMAP targets, are convenient to use. If one wants to gain more insight into star formation by using intensity mapping, it is important to find a tracer for cold molecular gas in the ISM of a galaxy or cluster where stars are believed to form [44]. The most common molecular gas inside molecular clouds is molecular hydrogen, due to the large abundance of hydrogen in the Universe. However, since molecular hydrogen is a symmetric molecule and therefore does not have an electric dipole, it cannot emit rotational dipole radiation in the radio or infrared bands. As such it has no low-energy excitations suitable for tracing large-scale structure distributions [45]. Therefore one can instead use carbon monoxide (CO), which is asymmetric and requires only a few kelvin to become excited in its rotational states. Importantly, it is also the second most abundant molecule after H_2 [45]. Since H_2 gas is, however, the most important for star formation due to its high abundance, it is essential to relate the CO density to that of H_2 when using a CO line as a tracer. This can be done through the so-called X -factor which relates the CO luminosity to the H_2 column density [44].

The rotational energy of a two-atom molecule like CO is discretized and determined by its angular momentum, which is given by the total angular momentum quantum number J as

$$E_{\text{rot}} = \frac{1}{2} \frac{\mathbf{J}^2}{I} = \frac{J(J+1)\hbar}{2MR_e^2}, \quad (2.1)$$

since the absolute square of the angular momentum $|\mathbf{J}|^2 = J(J+1)\hbar$ and the moment of inertia of the molecule $I = MR^2$ with the reduced mass of the two-atom system M and the distance R between the atoms [46]. Thus, using the energy-frequency relation of a photon, one can easily find that the energy difference between excited states corresponds to an emitted photon of frequency

$$\nu = \frac{\Delta E(J)}{h} = \frac{E_{\text{rot}}(J+1) - E_{\text{rot}}(J)}{h} = \frac{(J+1)\hbar}{MR_e^2}. \quad (2.2)$$

The resulting emission lines hence have a frequency that grows linearly with J and is hence sometimes called the CO emission ladder. The first two transitions $J = 1 \rightarrow 0$ and $J = 2 \rightarrow 1$, for instance, correspond to emission of $\nu_0 = 115.27$ GHz and $\nu_0 = 230.54$ GHz photons in the CO molecules rest frame [47]. Therefore it is possible to design a telescope that can at the same time observe several different CO molecular lines at different redshifts with a relatively narrow bandwidth.

In addition to the regular CO with a ^{12}C isotope in it, there is another quite common isotopologue of CO with a ^{13}C isotope instead. Due to ^{13}CO being slightly heavier than the regular ^{12}CO its emission line has a slightly lower frequency at about $\nu = 110$ GHz [48]. Thus, if we suspect a signal to come from ^{12}CO gas, which usually has some ^{13}CO mixed into it, will see a correlation peak between the two lines at a $\Delta\nu_0 = 5$ GHz separation (in the emitters rest frame) [10]. Such a clear correlation signature can be a powerful tool when determining whether the observed signal is indeed from CO or from some other line that is redshifted into the telescope’s bandwidth.

2.1.3 The Problems of Line Intensity Mapping

Until now we have mainly talked about the advantages of intensity mapping, however as in all surveys, some problems need to be handled to achieve the scientific goals. For instance, in intensity mapping the signal observed is generally weak and the data are therefore highly noise-dominated. Other problems are strong Galactic continuum foreground emissions contaminating the signal. These two problems we are, as we will see later when describing the COMAP analysis pipeline, able to handle to some degree. However, there are still challenges like emission from unwanted lines being redshifted into the observed frequency bandwidth. For example if we aim to observe [CII] at redshift $z = 7$ at an observed frequency of $\nu = 238$ GHz with some survey, we can get some signal from for instance $\text{CO}(J = 3 \rightarrow 2)$ and $\text{CO}(J = 4 \rightarrow 3)$ at respectively redshifts $z = 0.45$ and 0.88 mixed in [10]. These interloper lines are hard to separate from the target line and the most promising technique to separate them is to cross-correlate the intensity mapping survey at hand with some external galaxy redshift survey. Finally, it is worth mentioning that the complex, and often, non-linear processes in the environment of the emitters at high redshift can lead to large modeling uncertainties.

2.2 The COMAP Experiment

The CO Mapping Array Pathfinder (COMAP)¹ experiment is a line intensity mapping experiment aiming to map the large-scale distributions of star-forming galaxies in the Epoch of Reionization and the Epoch of Galaxy Assembly. COMAP is a pathfinder mission to demonstrate the feasibility and potential of wide-field intensity mapping and will be used to test observational strategies and technological development [33].

The COMAP experiment is proposed to take place in three phases. COMAP Phase I (Pathfinder) will consist of a single 26 – 34 GHz (the K_a -band) telescope targeting the $\text{CO}(J = 1 \rightarrow 0)$ line at $z = 2.4 - 3.4$ and $\text{CO}(J = 2 \rightarrow 1)$ lines at $z = 6 - 8$. Later, COMAP - Pathfinder will expand its operations to include an array of several more K_a -band telescopes. The goal of COMAP - Pathfinder will be to constrain the

¹The acronym’s name was changed to CO Mapping Array Project at the time of writing to prevent naming redundancies in COMAP - Pathfinder (Phase I).

star formation at the Epoch of Galaxy Assembly when the star formation rate peaked [49].

The third phase, COMAP - EoR, will add another receiver sensitive to 12 – 20 GHz radiation in the K_u -band which will map out $\text{CO}(J = 1 \rightarrow 0)$ at $z = 6 - 8$ and will cross-correlate its signal with that of the K_a -band receiver from the previous phases of COMAP to obtain constraints on the power spectrum of star-forming galaxies at the end of the Epoch of Reionization [34].

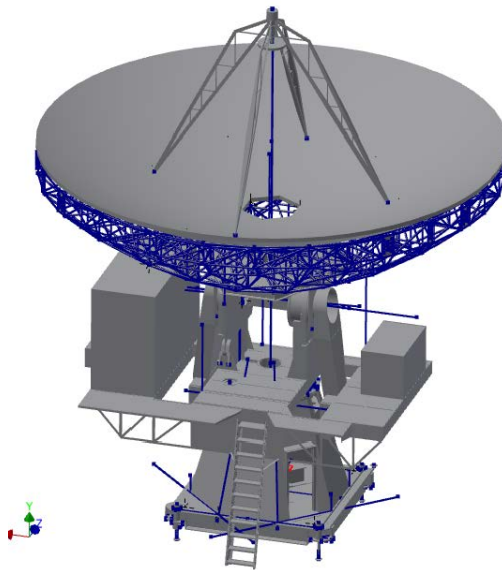


Figure 2.3: A drawing of the COMAP Phase I telescope. Courtesy of David Woody.

2.2.1 The COMAP - Pathfinder (Phase I) Instrument

The Telescope

The Phase I telescope of COMAP (see Fig. 2.3) is a modified 10.4 m Leighton telescope. It was originally built for the Millimeter Array at Owens Valley Radio Observatory (OVRO) in California. Due to high humidity in Owens Valley in summertime millimeter observations become difficult. Therefore, in 2005 the telescope was moved to a higher elevation where the air is dryer and became a part of the Combined Array for Research in Millimeter-wave Astronomy (CARMA). In 2015 the telescope was moved back to OVRO due to CARMA lacking funding, after which the telescope was retrofitted for COMAP [50, 34].

The primary mirror is made up of a honeycomb structure of 84 hexagonal aluminum panels with a surface skin [51], while the 1.1 m secondary mirror is held up by four secondary support legs. The backing structure of the primary mirror, preventing deformations of the reflective surface under the influence of gravity or other external

Table 2.1: The frequency coverage of the four sidebands of the COMAP - phase I instrument.

Band	Sideband	Frequency (GHz)
A	LSB	26 – 28
	USB	28 – 30
B	LSB	30 – 32
	USB	32 – 34

factors, is made by steel tubing arranged in a lattice of equilateral triangles resulting in a relatively lightweight telescope. Leighton designed the backing structure of the dish in such a way that was highly homogenous and isotropic and deformed homologously under its own weight [52, 53]. The telescope dish itself rests on an alt-azimuth mount.

The Receiver Chain

The receiver of the telescope is made up of an array of 19 feed-horns (sometimes referred to as pixels), in addition to a blind feed that does not observe the sky, but can be used for systematic error mitigation. Each of the 19 feed-horns has its own independent detector chain, which we often just call "feed". A detector chain starts at a feed-horn which gathers up incoming radiation, which is subsequently split into two orthogonal circular polarizations by a polarizer. The current setup of the telescope uses a two-stage polarizer for 15 of the feeds, while there are two feeds with a single-stage polarizer and two without polarizers. As the incoming signal is very weak, the subsequent induced power in the receivers has to be boosted in power by several orders of magnitude. This is done in a low-noise amplifier (LNA) which is cooled to 15 – 18 K to minimize thermal noise contamination [34].

The subsequent amplified signal is then down-converted by two down-converter modules (DCM1 and DCM2), the first of which down-shifts the 26 – 34 GHz frequency bandwidth to a 2 – 10 GHz bandwidth. The second DCM further divides the signal into two 4 GHz bands, A at 2 – 6 GHz and B at 6 – 10 GHz. These two bands are then subsequently each quadrature down-converted by an IQ mixer into an "in-phase" (I) and "quadrature" (Q) component. To split the signal into four sidebands (SB) with 2 GHz bandwidths and a resolution of about 2 MHz, the I and Q components of band A and B are sent through two CASPER ROACH2 spectrometers. Each of the bands A and B are then associated with each their "lower" (LSB) and "upper" (USB) sidebands (see Table 2.1). Thereafter the data are stored to disk for further processing [34].

The Telescope Beam

When observing any source of electromagnetic radiation the received radiation will be affected by the very receiver itself due to the wave nature of the radiation. Due to the diffraction of the incoming waves on a telescope, the instrument will have a distinct

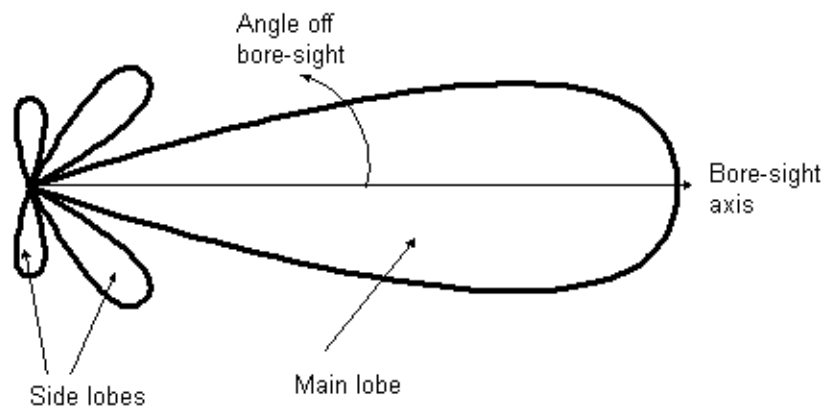


Figure 2.4: A side-view example of a radiation pattern of some antenna with its main-lobe along the optical axis or bore-sight (x-axis in the figure), as well as the sidelobes and far-sidelobes caused by the diffraction of the incident radiation. Figure from ESA [54].

pattern of sensitivity. For a parabolic antenna, like the one used in COMAP, the bulk part of the received radiation comes from the region close to the optical axis (sometimes referred to as the bore-sight); the so-called main beam or lobe [55]. The width of the main beam is often given in terms of the full width at half maximum (FWHM), which in the case of the COMAP instrument is $\sim 4.5'$ [34].

However, the inevitable secondary maxima of the radiation pattern, called sidelobes, result in a pickup of radiation at some angle from the main beam [55, 56]. This is illustrated in Fig. 2.4. In the case of the COMAP telescope, there is an additional pickup of radiation from four far-sidelobes. These four far-sidelobes are a result of radiation entering the telescope sideways, hitting one or more of the four secondary support legs and reflecting into the detector array. Since reflective angles stay constant for different wavelengths the far-sidelobe intensity maxima will always point in the same direction, while the received intensity changes for different wavelengths due to the diffraction of the sidelobe radiation in the telescope [57].

An example of a calculated beam pattern of the COMAP telescope for feed 1 can be seen in Fig. 2.5. Because the COMAP telescope is located at OVRO, where the terrain is mountainous, the far-sidelobes will sweep across different parts of the terrain and result in a ground pickup that can change in azimuth and elevation. Hence the telescope's data will, depending on the pointing and frequency observed, be contaminated by ground pickup to some degree. The ground contamination of the signal is, therefore, one of the most important systematic sources of error that are currently investigated.

2.2.2 Target Fields

There are several areas of the sky that are covered by the COMAP telescope. The three main targeted fields are called CO2, CO6 and CO7 and are chosen to avoid bright 30

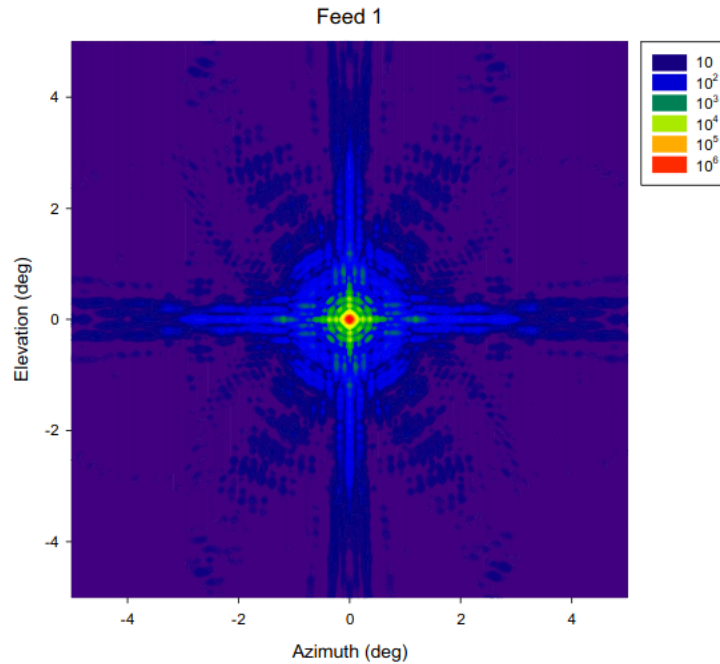


Figure 2.5: The head-on view of the radiation pattern of feed 1 of the COMAP telescope at 30 GHz. One can clearly see the four extended far-sidelobes due to reflection on the secondary support legs at 90° separation. Figure by Lamb [58].

GHz foreground point sources, as seen in Fig. 2.6. They are also chosen to overlap with the Hobby-Eberly Telescope Dark Energy eXperiment (HETDEX) Ly α galaxy survey [34]. The HETDEX survey will map large-scale structures to constrain the nature of dark energy [59]. Specifically, the CO7 field lies within the HETDEX spring field and the CO2 field is located within the HETDEX fall field. The latter is also encompassing the Spitzer/HETDEX Exploratory Large-Area(SHELA) survey area [34, 60]. The SHELA survey will for a wide range of redshifts measure the stellar-mass-halo-mass relation for galaxies [60]. The CO6 field was originally chosen to lie within the HETDEX spring field, but due to changes of the HETDEX field boundaries CO6 no longer has any overlap with the HETDEX spring field [34]. By cross-correlating the HETDEX galaxy catalog with the CO intensity map of the COMAP survey the significance of detection can be increased by a factor of two [18].

However, COMAP has some secondary science goals as well such as constraining the spectrum of Galactic anomalous microwave emission (AME) at around 30 GHz. Other fields observed are star-forming regions in the Andromeda Galaxy (M31), the Westerhout 43 (W43) region of star formation and the Perseus molecular cloud as well as a Galactic plane survey between longitudes $22.5^\circ < l < 50^\circ$. Lastly, it is worth mentioning that sources emitting in radio such as Jupiter, supernova remnants like Taurus A (Tau A), i.e. the Crab Nebula, and Cassiopeia A (Cas A) as well as the radio

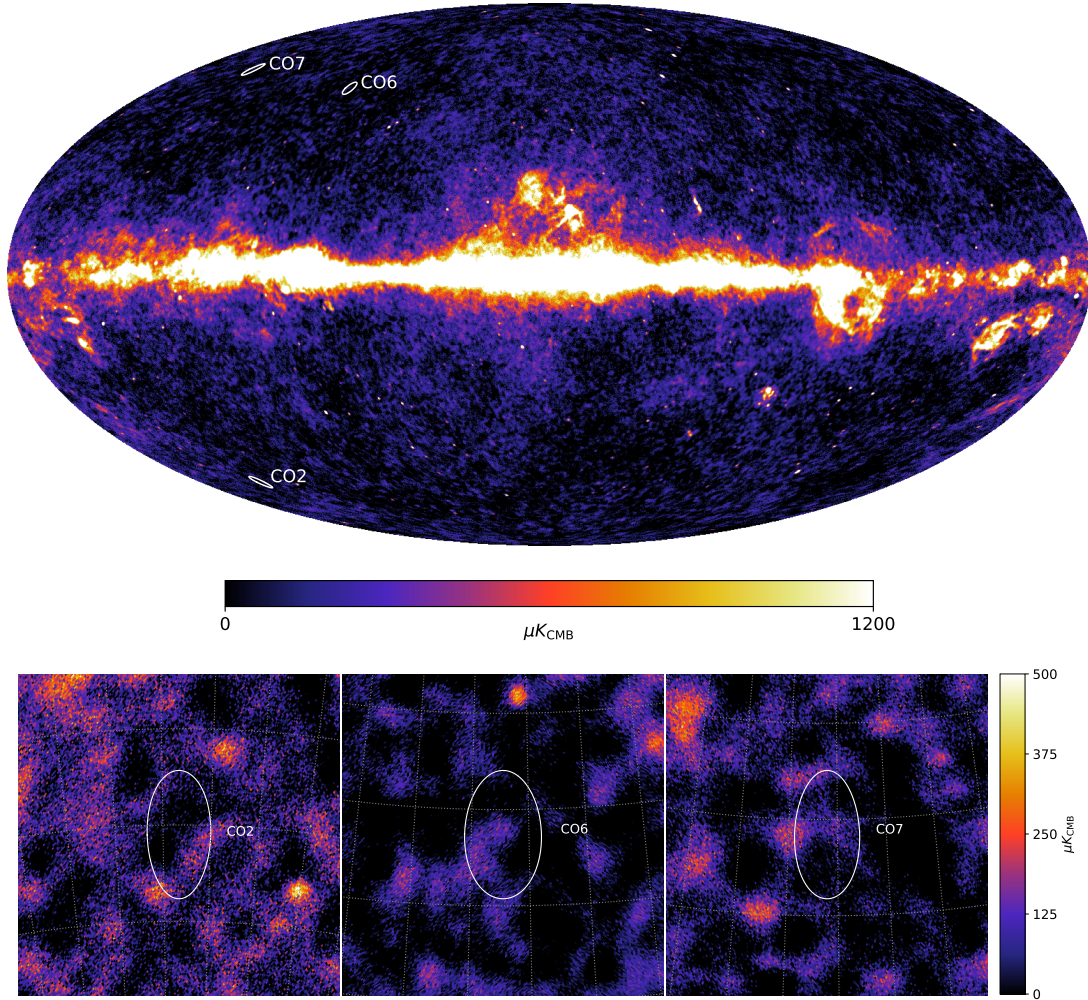


Figure 2.6: **(Upper panel)** The *Planck* LFI 30 GHz full-mission map (downloaded from the *Planck Legacy Archive*, <https://pla.esac.esa.int>, [9]) with overlaid circles of radius 2° that are centered on the three main CO fields COMAP is currently targeting. **(Lower panel)** A $10^\circ \times 10^\circ$ zoom-in on each of the three main CO science fields of COMAP illustrated in the full-sky map. The fields CO2, CO6 and CO7 are centered at Galactic coordinates (lon, lat) = $(149.0^\circ, -60.3^\circ)$, $(91.35^\circ, 53.22^\circ)$ and $(150.64^\circ, 59.53^\circ)$ respectively. These circles serve to illustrate the position of the fields with respect to the Galactic plane, as well as the approximate coverage of the fields on the sky.

galaxy Cygnus A (Cyg A) are used for calibrations [34]. As these sources exhibit a known behavior they can be used as extra-atmospheric control sources when calibrating the telescope. We will come back to the calibration in more detail in Sec. 3.4.7.

2.2.3 Scanning Strategies

Next, we will consider how the COMAP telescope observes a target field while it drifts across the sky. A typical scan starts with the COMAP telescope starting to scan at the edge of a sky patch. Due to the rotation of the Earth, the target patch will drift across the sky while the telescope scans across it. After the patch has completely drifted through, the telescope re-points itself back to the leading edge of the patch to start a new equivalent scan. Each of these scans take 5-10 min, while the collection of individual scans (typically 3-11) are referred to as an observation. Each observation is assigned with an observational identification (obsID) number [34].

We will here take a look at the three main scanning strategies that are or have been used by the COMAP telescope at some point; Constant elevation scans (CES), circular scans and Lissajous scans. The scanning path on the sky can in all three cases be described in sidereal time by the harmonic motion in local telescope coordinates

$$az(t) = A \sin(at + \phi) \quad (2.3)$$

$$el(t) = B \sin(bt), \quad (2.4)$$

where the specific value of the parameters A and B describe the extent of the observed field, while the ratio between a and b gives the shape of the scanning pattern and ϕ determines the relative phase between azimuth az and elevation el [34].

In particular, there are three main scanning strategies used by COMAP. If $b = 0$ then the telescope will remain at constant elevation and only scan back and forth in azimuth. Meanwhile, if one lets $a = b$ and $A = B$, with a phase shift $\phi = 0$, the telescope will sweep across the sky in a circular pattern around a fixed point. The circular scanning pattern may be altered to a Lissajous pattern by using different a and b . Then the ratio a/b determines how many sweeps in azimuth are performed for every elevation cycle [34].

Two main factors determine how good a given scanning strategy is; coverage efficiency and cross-linking. The former of the two is important because a higher total integration time per map pixel will result in a lower noise level and hence a higher signal-to-noise ratio. The latter factor, cross-linking, refers to whether a pixel is repeatedly approached from different angles throughout the scan. Using the spatial and temporal information of two different sweeps across a pixel, approaching from different directions, makes it possible to separate the signal to observe from correlated $1/f$ -noise (more on $1/f$ -noise in Sec. 3.1) [61]. Hence a poorly cross-linked map can look "stripy". We will now take a look at the three main scanning strategies implemented by COMAP and discuss their pros and cons.

Constant Elevation Scans

In a CES scan, the telescope will scan back and forth in azimuth while remaining stationary in elevation. The path across the sky in azimuth and elevation as a function of time can be seen in Fig. 2.7.

The advantage of a CES scan is that it remains at the same elevation over the duration of each scan. This makes the resulting measurement less prone to systematic errors like correlations in elevation (e.g. varying amount of air along the line-of-sight or ground pickup). As COMAP aims for very weak CO signatures the observed signal is easily dominated by such systematics, which is a strong argument for using the CES scan. A typical CES scan has better cross-linking than a circular scan, but worse cross-linking than a Lissajous scan. However, the observed field on the sky will rotate with respect to the alt-azimuth telescope mount, hence increasing the cross-linking somewhat for longer observations. Other drawbacks are that a CES scan is generally less efficient in covering the sky than for instance a Lissajous scan. Moreover, in a CES scan the telescope must repeatedly accelerate back and forth in azimuth. This increases the strain on the telescope's motors and wears them down quicker [62].

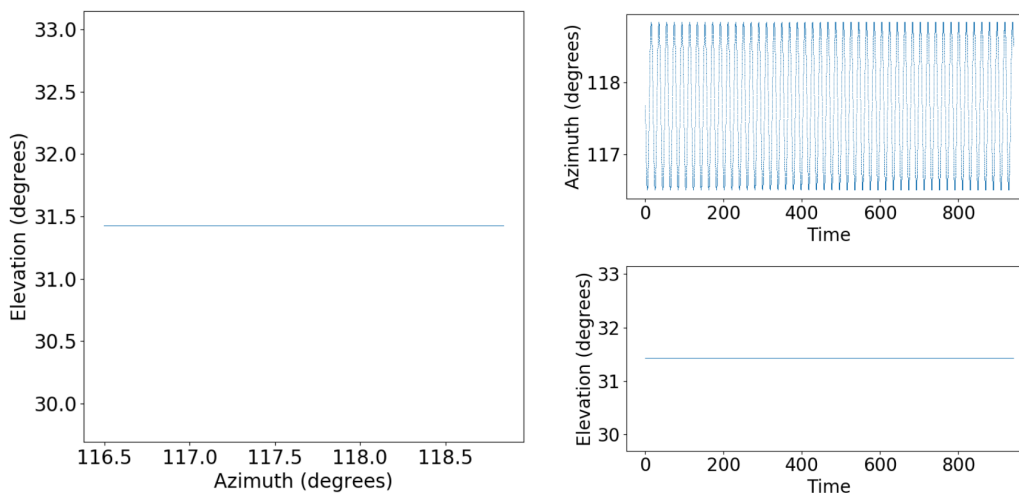


Figure 2.7: (**Right panel**) Example of a constant elevation scanning path of the telescope on the sky and (**left panels**) the corresponding azimuth and elevation coordinates as a function of time. Courtesy of Stuart Harper.

Circular Scans

In a circular scan both the azimuth and elevation of the telescope will, as can be seen in Fig. 2.8, move in a sinusoidal motion and trace out circles on the sky in local telescope coordinates.

An advantage of such a scanning pattern is that the telescopes scanning speed can remain at higher speeds, and hence trace out more of the sky. This also reduces the wear of the telescope's mechanical parts as there is less back and forth acceleration. However, as the telescope now changes in azimuth and elevation it will be more prone to correlations in these coordinates such as ground pickup by the far-sidelobes. Also, the angle of approach on a given pixel tends to be more predictable than in a CES or Lissajous scan, resulting in a less effective cross-linking [62].

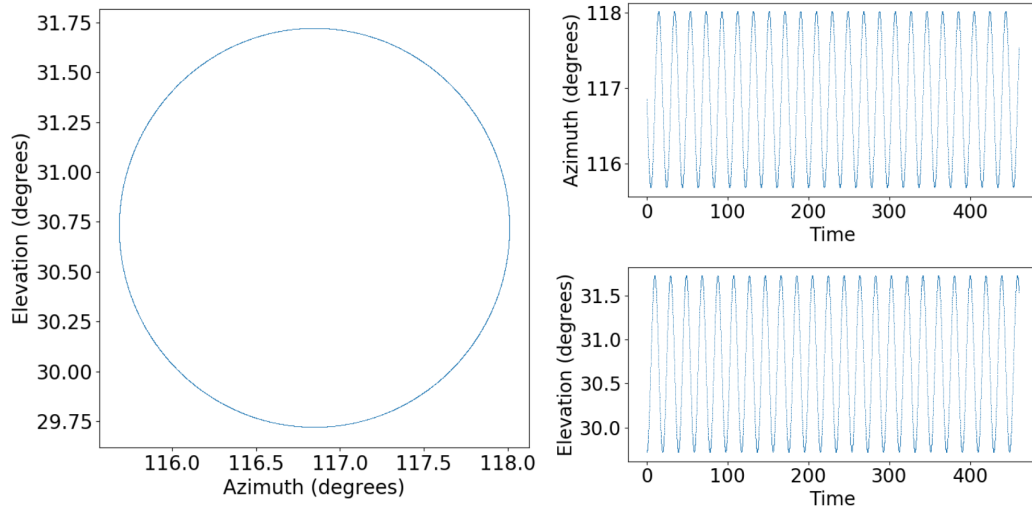


Figure 2.8: (**Right panel**) Example of a circular scanning path of the telescope on the sky and (**left panels**) the corresponding azimuth and elevation coordinates as a function of time. Courtesy of Stuart Harper.

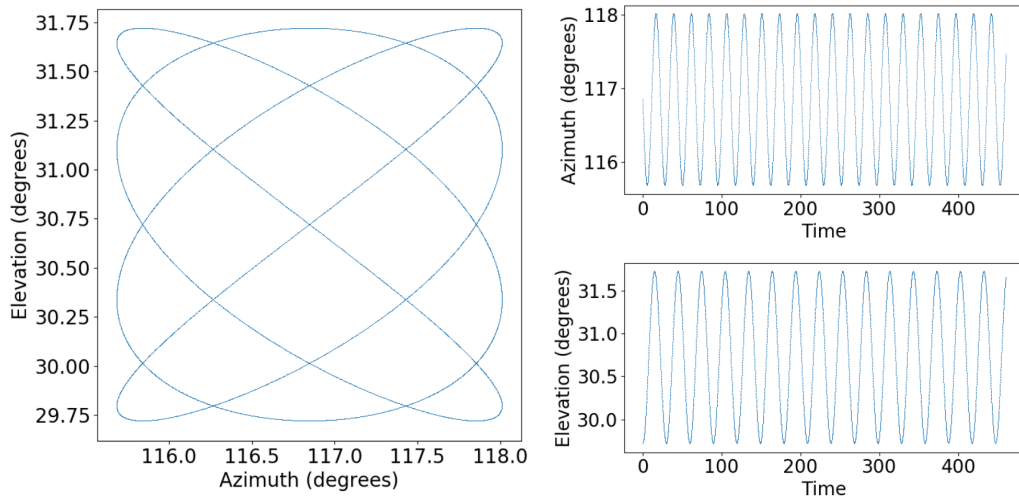


Figure 2.9: (**Right panel**) Example of a 3/4 Lissajous scanning path of the telescope on the sky and (**left panels**) the corresponding azimuth and elevation coordinates as a function of time. Courtesy of Stuart Harper.

Lissajous Scans

A Lissajous scan will trace out a pattern on the sky in which the harmonic motion in azimuth and elevation have a different period. The resulting pattern across the sky for a 3/4 ($a = 3$ and $b = 4$) Lissajous scan as well as the corresponding coordinate changes can be seen in Fig. 2.9. Here it sweeps up and down three times in elevation for every

four oscillations in azimuth.

The obvious advantage of such a Lissajous scan is that it has very efficient sky coverage and cross-linking. That is because the telescope can spend a lot of time at full speed, while each map pixel is approached from many different directions across the span of an observation. The full-speed scanning will also reduce the forces on the mount compared to CES scans. However, as the telescope changes in both azimuth and elevation, it will also be affected easier by systematic effects correlated in telescope coordinates, e.g. ground pickup [62]. This is a potential argument against using Lissajous scans in COMAP, as a ground-based CO intensity mapping experiment is very vulnerable to such systematics.

Chapter 3

The COMAP Data Analysis Pipeline

When performing an astrophysical survey one often thinks about the remarkable efforts of engineering that go into building and operating a high-precision instrument to gain such remarkable insight into the nature of reality. However, a part of such a survey that is of equal importance is the data analysis. For COMAP, the data gathered by the telescope is highly contaminated by continuum foreground, noise and systematic errors that dominate the wanted CO signal by orders of magnitude. Therefore, to make the data usable for science we must remove those contaminants.

In the following section, we will therefore dive into some of the notable features that contaminate and characterize the data. Thereafter we present an overview of the filters and methods used in the analysis pipeline we are developing in the Oslo group of COMAP to clean the data. A schematic structure of the pipeline can be seen in Fig. 3.1 to illustrate the order of steps from raw to cleaned data.

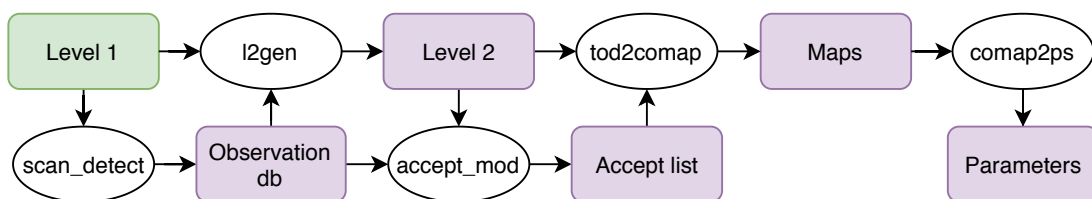


Figure 3.1: The basic structure of the analysis pipeline from the raw level 1 (green box, representing data produced outside of the Oslo group) and how it is processed by the various codes (white ellipses) producing intermediate and final products (purple boxes). Courtesy of Marie Foss.

3.1 The Data Modeling and Noise Properties

When the telescope is gathering data by scanning across the sky, it receives the observed electromagnetic radiation in the detector. Once the radiation hits the feeds it induces an electrical impulse which is amplified in the telescope's electronics. The output power per frequency bandwidth is then given by

$$P = Gk_{\text{B}}\Delta\nu T_{\text{sys}}, \quad (3.1)$$

where k_{B} is the Boltzmann constant, G is the so-called gain, $\Delta\nu$ is the frequency bandwidth used, and T_{sys} is the system temperature [63]. The gain G quantifies how much the electrical signal, induced in the detector, is amplified before it is registered, while the system temperature T_{sys} quantifies how much thermal noise is in the data. The thermal noise in the data is distributed according to a Gaussian distribution $\mathcal{N}(0, \sigma_0)$, where the standard deviation is given by the radiometer equation [63]

$$\sigma_0 = \frac{T_{\text{sys}}}{\sqrt{\tau\Delta\nu}}. \quad (3.2)$$

Here the resolution of the time and frequency samples are given respectively as τ and $\Delta\nu$. Furthermore, the system temperature consists of the temperature of all sources that contribute to white noise in the system [63] and is in our case written as

$$T_{\text{sys}} = \sum_i T_i = T_{\text{rec}} + T_{\text{atm}} + T_{\text{cmb}}e^{-\tau} + T_{\text{forgr}}e^{-\tau} + T_{\text{ground}} + T_{\text{src}}e^{-\tau}, \quad (3.3)$$

i.e. the sum of the receiver, atmosphere, CMB, foreground, ground and CO line emission source temperatures respectively [34]. The optical depth of the atmosphere is given by τ , which quantifies the amount of air that attenuates the received radiation between entering the atmosphere and hitting the detector. Therefore all contributions to the system temperature that originate outside the atmosphere will be damped by a factor of $e^{-\tau}$. However, as the atmosphere is also a thermal emitter itself, the atmosphere will contribute with T_{atm} to the system temperature, in addition to its attenuating effect on the extra-atmospheric sources.

We know from CMB measurements that the CMB has a nearly perfect black-body spectrum with temperature $T_{\text{CMB}} \approx 2.7255$ K [64] with only small fluctuations $\Delta T/T \sim 10^{-5}$. The foreground term T_{forgr} consists mainly of the signal from continuum emission within our own galaxy. Meanwhile, the T_{src} term represents the brightness temperature of some CO emitter, for example, a faint galaxy cluster at high redshift. As described earlier, the telescope beam also has a far-sidelobe pickup, that when hitting the ground will contribute to the signal with its thermal radiation. Even though the ground pickup may be modeled by simple black body radiation, the frequency dependence of the far-sidelobes will cause a different degree of ground contamination at different frequencies. Due to the complexity of the ground profile and the frequency dependence of the sidelobes, the ground pickup might therefore be hard to model.

The receiver temperature $T_{\text{rec}} \sim 35 - 45$ K [34] cannot directly be thought of as the "thermometer" temperature of the receiver, but rather as the temperature equivalent

of the white noise added by the detector alone. Nevertheless, because the receiver temperature is determined by the number of thermal electrons in the system it can be inferred by the thermal temperature of the detector. To minimize this contamination the electronics of the telescope are therefore cooled down in the cryostat.

The uncorrelated thermal noise has the nice property of averaging out in the large sample limit. Hence the way to reduce the white noise is to simply observe longer. Because we not only have white noise in our data, but also correlated noise, we cannot estimate the level of white noise σ_0 by simply taking the standard deviation of the data. Instead, we need to use an approach in which the standard deviation of the white noise is estimated locally between two neighboring data points, giving [34]

$$\sigma_0 = \sqrt{\frac{\text{Var}(d_i - d_{i-1})}{2}}. \quad (3.4)$$

In the COMAP data there are primarily three known sources of correlated noise; gain fluctuations within the low noise amplifier, temperature fluctuations in the atmosphere and standing waves between the primary and secondary reflector of the telescope as well as inside the telescope electronics. These correlated noise sources all exhibit roughly a characteristic $1/f$ spectrum in the temporal Fourier domain, i.e. where f is the frequency of for example the scanning frequency of order Hz (as opposed to the frequency domain characterizing the electromagnetic radiation received, being in the GHz regime). Figure 3.2 shows this characteristic $1/f$ spectrum at various stages of the pipeline. It is, therefore, a good approximation to model the correlated noise as

$$N(f) = \sigma_0^2 \left(1 + \left(\frac{f}{f_{\text{knee}}} \right)^\alpha \right), \quad (3.5)$$

where σ_0^2 is the aforementioned white noise level, while f_{knee} quantifies the frequency f at which the correlated noise level equals the white noise level. The parameter α quantifies the slope of the power spectral density in the low- f region where the $1/f$ noise is dominant [34]. Sources of noise like atmospheric temperature fluctuations or standing waves are not only correlated in time, but also in frequency and across different feeds. The correlated noise will therefore not simply integrate down in the limit large sample sizes, and leave artifacts in the signal that can affect the final scientific analysis negatively. However, while correlated noise and unwanted contaminants are highly correlated over frequencies and feeds, the wanted CO signal is very weak and is not expected to span a large volume in pixel-frequency space. This is the key to separate the two, as we will see in Sec. 3.4, to obtain a cleaned dataset usable for scientific parameter estimates.

Now that we have introduced the contributions that go in to form the signal received by the COMAP telescope we can look at how to model the time stream of data. As described by [34] the time-ordered data (TOD) can be modelled as

$$d_\nu^i(t) = \langle d_\nu^i \rangle (1 + \delta_G^i) [1 + P_{\text{cel}}^i (s_{\text{cont}} + s_{\text{CO}}^\nu) + P_{\text{tel}}^i s_{\text{ground}} + n_{\text{corr}} + n_{\text{w}}^{\nu,i}], \quad (3.6)$$

where the raw TOD is given by $d_\nu^i(t)$ at time t , feed i and frequency channel ν . The running mean of the time stream is here denoted as $\langle d_\nu^i \rangle$ containing all the modes of the

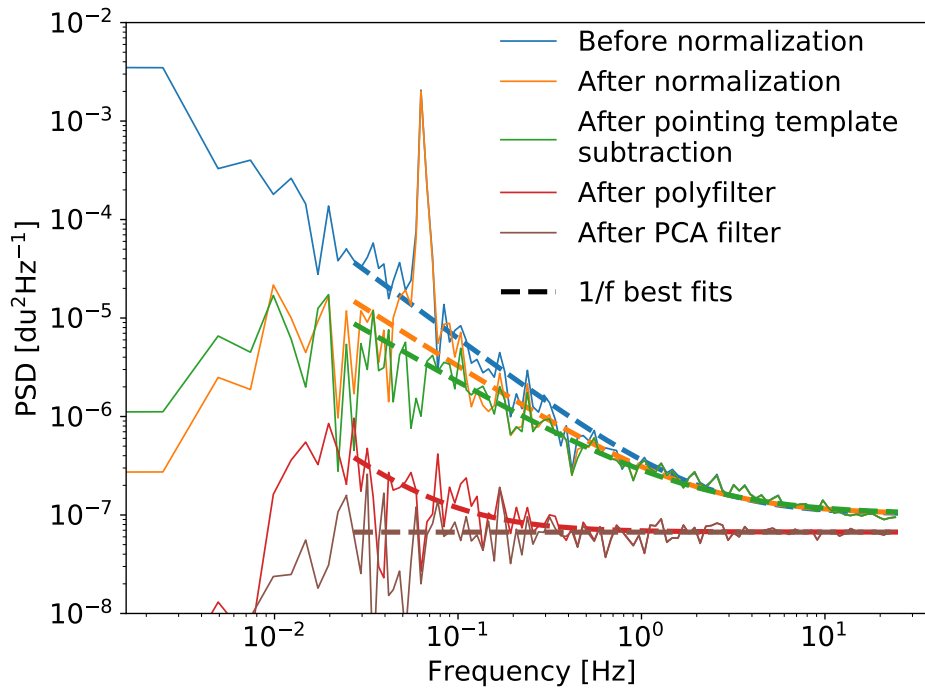


Figure 3.2: The power spectral density of the TOD at different stages of the pipeline, as well as the best fit $1/f$ spectra corresponding to the correlated noise components, while white noise only would yield flat power spectra. Note the spike at ~ 0.1 Hz in the data after normalization, but before eliminating the pointing template (orange), corresponding to the correlations in the TOD induced by the pointing of the telescope. Courtesy of Jonas Lunde.

TOD of a long timescale such as atmospheric temperature fluctuations or slowly varying gain fluctuations. The small scale gain fluctuations characteristic to each feed are given by δ_G^i , while the actual signal contribution from astrophysical sources is given by s_{cont} and s_{CO}^ν . These represent the continuum foreground and extragalactic CO signal's map respectively, which are spread from the pixel domain to the temporal domain by the pointing matrix P_{cel}^i (more on that in the mapmaking section) in celestial coordinates. Similarly the signal contribution from ground pick-up is given by the map s_{ground} which is projected into the time domain by the telescope coordinate pointing matrix P_{tel}^i . Lastly, the raw TOD is also dependent on the correlated and uncorrelated noise sources, n_{corr} and $n_w^{\nu,i}$, as introduced earlier.

3.2 Level 1 Data

The data that is gathered by the COMAP telescope consists of a time stream of data often referred to as time-ordered data or TOD. Hence a raw TOD consists of an array of

time samples taken by each detector, in each sideband and each frequency channel, and are thus best represented by a $n_{\text{samp}} \times n_{\text{feed}} \times n_{\text{sb}} \times n_{\text{freq}}$ matrix. Here n_{samp} , n_{feed} , n_{sb} and n_{freq} are the number of time samples, feeds, frequency sidebands and channels per sideband. This TOD dataset for a single obsID, along with its pointing information for each feed (as the feeds do not overlap, they have their separate pointing information), diagnostic information and housekeeping data, make up a so-called level 1 file. The level 1 files are of the HDF5 format, made for efficient storage of large multidimensional hierarchical datasets. Hence all the necessary information needed for further analysis is neatly stored in one place.

An example of how the raw data of all feeds from several scans look like can be seen in Fig. 3.3. The data from the telescope are thereafter sent to Caltech, where level 1 files are assembled for compact storage. Subsequently, the level 1 file is transferred to Oslo where further analysis is done to obtain cleaned and calibrated level 2 data.

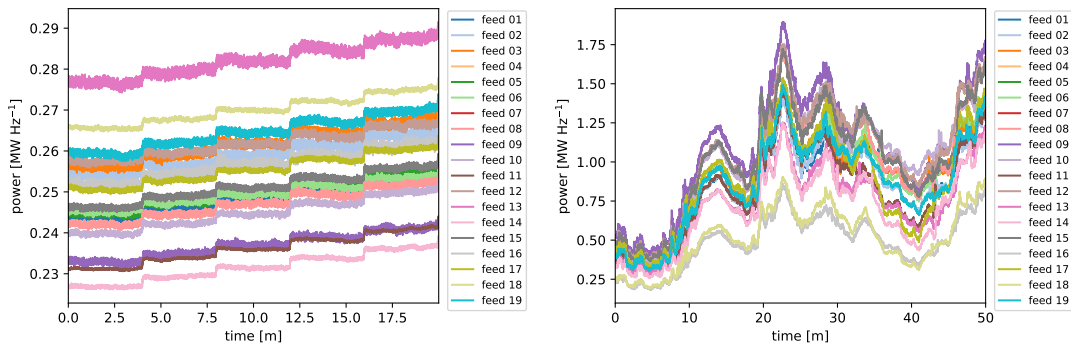


Figure 3.3: The TOD corresponding to all 19 feeds for several scans of two different obsIDs; (**left**) one where the data are gathered under good conditions, and (**right**) one that is contaminated by bad weather. The data gathered under good weather conditions is dominated by instrumental noise and systematics, while the data from the bad weather observation clearly shows strong common-mode contamination at all feeds. Courtesy of Håvard Tveit Ihle.

3.3 Scan Detect

The first step we perform in the Oslo analysis pipeline is to read in the raw level 1 data sent from Caltech, which is then fed into the `scan_detect` code. The `scan_detect` script's task is to go through the raw TOD and divide it into individual scans between each re-pointing of the telescope. As mentioned in Sec. 2.2.3 each obsID consists of several 3-10 minutes long scans. Each time the observed sky field has drifted by, due to the rotation of the Earth, the telescope must re-point itself. The data gathered during re-pointing is usually not used as it does not contain any information valuable for the scientific goal.

After dividing the raw level 1 data into a series of individual scans, a runlist is

generated. The runlist is a separate file listing all obsID and their constituent scans. For each obsID, the individual scans making it up are provided, along with the scanning strategy used, the time as the Modified Julian Date (MJD) of when the scans start and end as well as some pointing information. The runlists are usually generated according to each observed field, i.e. only one source per runlist [34].

3.4 Level 2 Data

After we have identified each scan from the generated runlist we can process the data through the level 2 file generator in the `12gen` code. The main purpose of `12gen` is to clean the raw data of correlated noise and continuum foreground emission and perform the signal calibration of the data.

The very first step in generating level 2 data is to read in all the data provided by a runlist. Each level 1 file listed in the runlist, corresponding to a single obsID, is assigned to an MPI process and the subsequent analysis of the data is then performed separately on each obsID. The scans in each obsID are then in turn assigned to each their parallel process, and the following analysis of the data in `12gen` is performed separately on each scan.

In the following we describe each of the steps we perform in `12gen` to go from the raw level 1 data to obtain the cleaned level 2 data that is ready for mapmaking.

3.4.1 The Ambient Load

Because the raw data from the telescope is in digital units (DU), but we want to do the final scientific analysis in units of brightness temperature, we need to do a calibration. However, as the cleaning of the data is done before the calibration step (see Sec. 3.4.7) we need to extract the information needed for calibration before introducing the dimensionless units used in the filtering of the TOD. For calibration, we need the ambient temperature of the telescope which is easily estimated using thermometers on the instrument. In addition, we need the corresponding ambient load measurement of the telescope, being the value of the TOD (in DU) when the telescope observes something of ambient temperature. To do this a disk-shaped absorber, called the calibration vane, is rotated in front of the telescope feeds. As the vane has the same temperature as the surrounding environment its thermal radiation registered by the telescope serves as our estimate of the ambient load measurement. For each obsID two measurements of the calibration vane are performed, respectively at the very beginning and end of the obsID. Two such ambient load measurements are illustrated in Fig. 3.4. As we will see in the later calibration section 3.20, we can estimate the ambient load measurements and temperatures at any given time by interpolating between the two measurements.

3.4.2 Normalization

The very first method of cleaning we apply to the raw data is the normalization step. Many of the systematics and noise artifacts in the raw data show a strong correla-

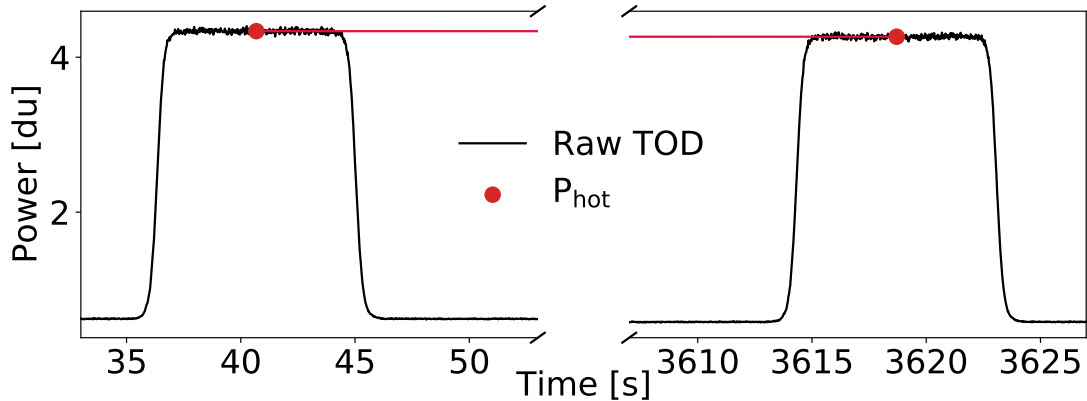


Figure 3.4: The two ambient load measurements at the beginning and end of an obsID when the calibration vane is rotated in front of the detector feeds. Courtesy of Jonas Lunde.

tion across frequencies. To remove these correlated artifacts we thus ideally want the frequency channels to be directly comparable. However, as the telescope electronics boost the signal somewhat differently for each frequency the TOD power level of each frequency needs to be normalized to the same level for them to be comparable. By flattening out the instrumental bandpass we also eliminate unwanted long-timescale structures like atmospheric temperature fluctuations, slowly varying gain variations or large astrophysical foreground structures [34].

To normalize the TOD it is divided by its running mean \bar{d} , which corresponds to the lowpass filtered TOD, found by applying a weight function in Fourier space. The weight function, in this case, is given as

$$W = \left[1 + \left(\frac{f}{f_{\text{knee}}} \right)^\alpha \right]^{-1}, \quad (3.7)$$

where f is the temporal frequency of the TOD and the knee frequency is usually $f_{\text{knee}} = 0.01$ Hz. Thus, f_{knee} corresponds to a slightly lower temporal frequency than that of the scanning of the telescope at roughly 0.1 Hz, below which we want to remove the corresponding large-scale structures. The slope of the cut-off is given by the parameter α and is in our case equal to $\alpha = 4$ [34]. The lowpass filtered TOD is then simply represented as

$$\bar{d} = \mathcal{F}^{-1}[\mathcal{F}[d] \cdot W], \quad (3.8)$$

in which \mathcal{F} represents the Fourier transform from the time domain to temporal frequency space. Subsequently, we divide the data d by its lowpass filtered equivalent and subtract unity [34];

$$d_{\text{norm}} = \frac{d}{\bar{d}} - 1. \quad (3.9)$$

This step effectively removes all running modes of a large temporal scale $\langle d_t^i \rangle$ from

Eq. (3.6), leaving the small scale gain fluctuations, pointing correlations and noise contributions for later filters to deal with.

In the first row of Fig. 3.5 it is illustrated how the normalization removes the large running modes, leaving oscillations caused by telescope pointing as well as noise.

3.4.3 Removing the Pointing Template

Because COMAP is a ground-based experiment the signal we receive will always to some extent be affected by the environment of the telescope. Most noticeable in that regard is the alteration of astrophysical signals by the atmospheric attenuation, as well as ground pickup by the far-sidelobes of the telescope.

When pointing to low elevations the astrophysical signal has to travel through more air to reach the telescope than when observing at high elevations, resulting in a higher attenuation of the signal. Thus if the telescope uses a scanning strategy like Lissajous, where it changes its elevation over time, the TOD will show an oscillatory pattern. The atmospheric attenuation can, if the atmosphere is assumed to be roughly homogenous, be described by the optical depth τ at elevation el by [65]

$$\tau = \frac{\tau_0}{\sin el}. \quad (3.10)$$

The optical depth of the atmosphere when the telescope points directly upwards to the zenith, i.e. $el = 90^\circ$, is defined to be τ_0 .

In addition to the oscillations in the time stream stemming from the change in elevation pointing, there are also signs that there are some correlations in the TOD with changes in azimuth [34]. The origins from these are not fully understood but could originate from for instance ground pick-up by the far-sidelobes of the telescope, as described earlier in Sec. 2.2.1.

Using Eq. (3.10) to model the elevation dependent extinction and a low order polynomial in azimuth we get the elevation template

$$d = \frac{g}{\sin el(t)} + \alpha az(t) + \beta + n \quad (3.11)$$

where g , α and β are constants and n represents Gaussian noise. The time dependent pointing information of the telescopes is given by the elevation $el(t)$ and the azimuth $az(t)$ measured in local telescope coordinates [34]. Subsequently, the model with best-fit parameters is subtracted from the time stream so that

$$d_{\text{new}} = d_{\text{old}} - \frac{g}{\sin el(t)} - \alpha az(t) - \left\langle \frac{g}{\sin el(t)} + \alpha az(t) \right\rangle, \quad (3.12)$$

where the average of the template is also subtracted to remove the contribution of the mean atmospheric effects over the time of the scan [34]. The fitting and subsequent removal of the pointing template is done separately on each frequency channel.

The removal of the pointing template from the TOD is illustrated in the second row of Fig. 3.5, showing a clear elimination in the periodic motion stemming from elevation

changes. Correspondingly in Fig. 3.2, showing the TOD's power spectral density, the spike at ~ 0.1 Hz corresponding to the telescopes scanning period of ~ 10 s, is effectively removed after subtracting the pointing template.

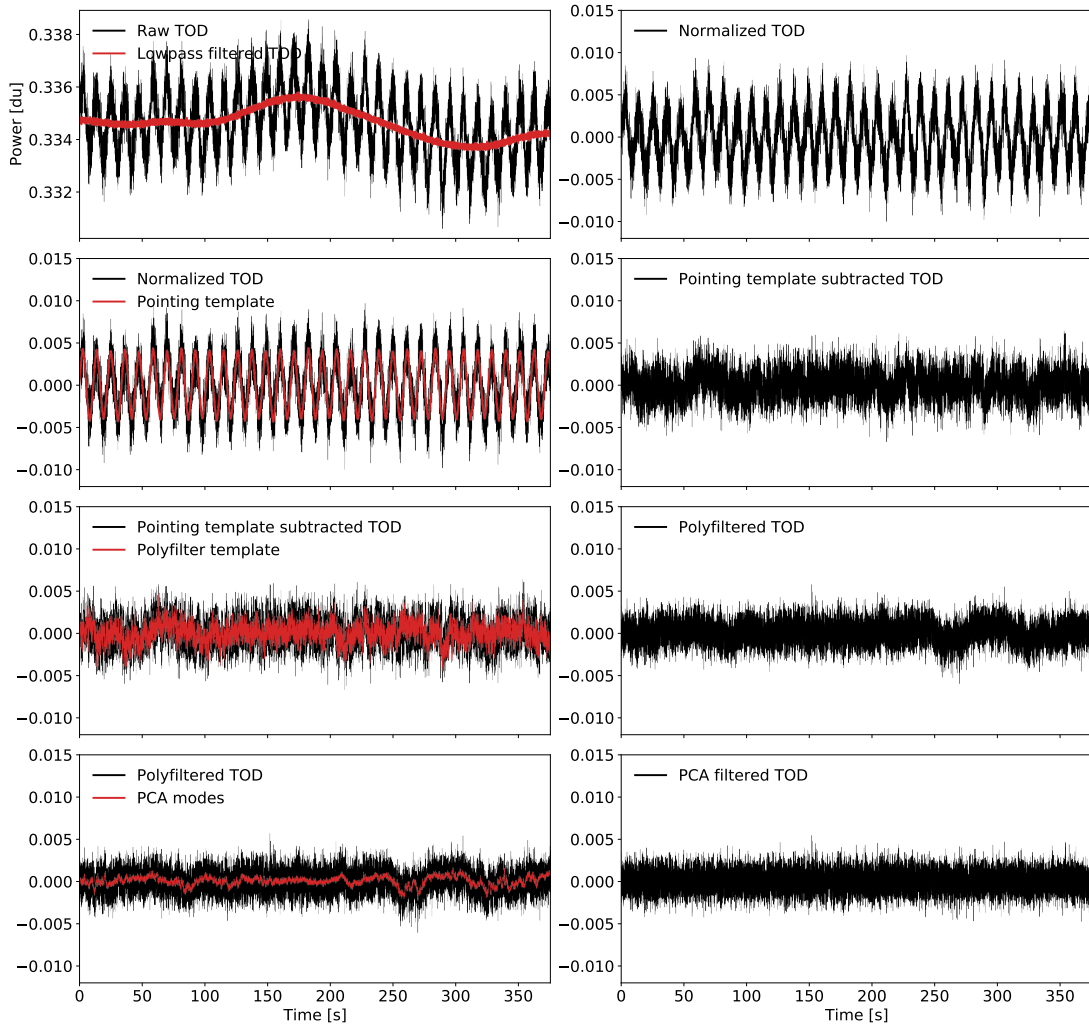


Figure 3.5: Before and after (respectively in the left and right columns of panels) view examples of the TOD at different stages in the analysis pipeline. In the normalization step (**first row**), we see how the division by the running mean (in red) mostly flattens the TOD, leaving only pointing oscillations and noise. After fitting and removing the pointing template, (**second row**) the data are even closer to pure noise. However, there are still some correlated noise artifacts left, which are quite effectively removed by the polyfilter (**third row**). The little that is left of correlated residuals is then taken out by the PCA filter (**fourth row**), leaving mostly white noise. Courtesy of Jonas Lunde.

3.4.4 Polynomial Filter

After having normalized and removed the pointing template from the data it already looks a lot cleaner. Nevertheless, the TOD is still contaminated by a lot of correlated noise such as $1/f$ noise from electronic gain fluctuations and atmospheric turbulence, as we see by the TOD's power spectral density in Fig. 3.2. In addition continuum foregrounds like the CMB, synchrotron, free-free and spinning dust components (see Fig. 2.6 for illustration of the 30 GHz foreground level) will be contained in the signal and will if not removed, totally obscuring the wanted CO signal. As mentioned, these unwanted frequency correlated modes do not average out over time and must hence be removed.

The key idea of removing these contaminants is to use that they are both highly correlated across frequencies within a sideband and slowly varying in frequency space. Thus we can fit a low order polynomial

$$d_\nu = c_0 + c_1\nu + c_2\nu^2 \dots \quad (3.13)$$

with constants c_i , to the data d_ν in frequency space ν at each individual time step to model the common modes. Usually the polynomial used is of first degree to allow a good fit of for instance the spectral distribution of continuum foreground [34, 9]. To get rid of the common frequency modes across each sideband we remove the best-fit polynomial over a sideband from the TOD giving

$$d_\nu^{\text{new}} = d_\nu^{\text{old}} - c_0 - c_1\nu - c_2\nu^2 - \dots \quad (3.14)$$

Although the idea of this so-called polynomial filter (polyfilter) is very simple, it is powerful when removing the common-mode noise [34]. The effect of the polyfilter is illustrated in the third row of Fig. 3.5 as a before and after view, where the output time stream looks close to uncorrelated noise with little structure. Similarly, we see in Fig. 3.2, that the polyfilter removes a significant portion of the $1/f$ noise from the TOD's power spectral density.

3.4.5 Principal Component Analysis

Although the TOD at this stage should be fairly well cleaned some signal contaminants like bad weather, or artifacts we were unable to remove with the previous filters, may still be left. These often span the whole focal plane of the telescope, i.e. over multiple feeds, and extend over several frequency channels. Meanwhile, the CO signal we want to measure is expected to only weakly correlate between feeds, as well as spanning only small frequency intervals. This difference in structure is the key to separating the remaining contaminants from the CO signal.

A principal component analysis (PCA) is exactly what we need to separate the correlated common modes that have survived the previous filters. The idea behind a PCA is to decompose the data in some parameter space into its components of most spread, i.e. most variance. These components called the principal components (PCs) or PCA modes, corresponding to the eigenvectors of the covariance matrix of the data,

with the eigenvector of highest variance pointing in the direction of most spread. This is illustrated on a simple 2D Gaussian mock dataset in some arbitrary parameter space in Fig. 3.6.

In the case of COMAP our data can be similarly decomposed in the parameter space of feeds and frequencies, and subsequently cleaned by eliminating the leading PCA components that explain the most information in the signal. Because the CO signal will cause little spread in the data in frequency-feed space, compared to for instance a cloud causing weather spikes in all feeds and frequencies, it will be left unaffected when removing the leading PCA components. A somewhat extreme example of such weather spikes in the TOD of all feeds is seen in the right plot in Fig. 3.3.

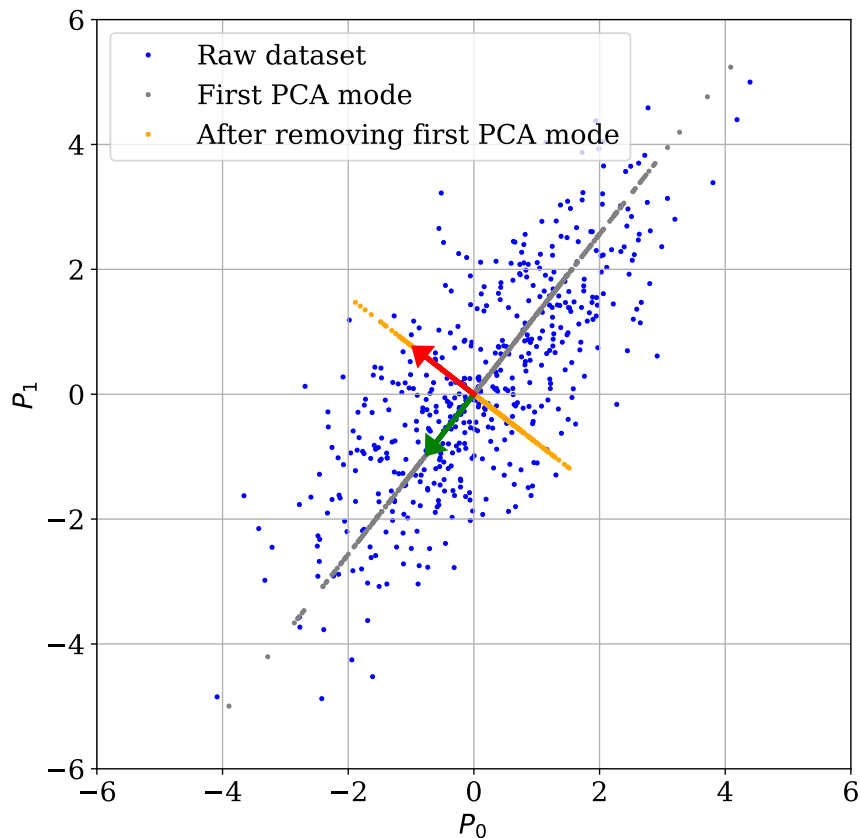


Figure 3.6: A mock example of a simple PCA analysis on a dataset that follows a two-variate Gaussian distribution (seen as blue dots) in an arbitrary $P_1 - P_2$ parameter space. The dataset has two directions of variance, as marked by the two unit vectors. The two corresponding PCA modes are seen as gray and orange dots respectively.

Mathematically we need to quantify the spread in the data in frequency-feed space. This can be done by first organizing the data in a data matrix

$$\mathbf{D} = \begin{bmatrix} D_{11} & \cdots & D_{1n_{\text{samp}}} \\ \vdots & \ddots & \vdots \\ D_{n_{\text{freq}}1} & \cdots & D_{n_{\text{freq}}n_{\text{samp}}} \end{bmatrix} \quad (3.15)$$

in which the rows correspond to frequencies and feeds, while the columns contain time samples. More precisely we sort all channels of all sidebands for all detectors along the first axis of the matrix, making \mathbf{D} a $n_{\text{freq}} \times n_{\text{samp}}$ matrix, where $n_{\text{freq}} = n_{\text{feeds}} \cdot n_{\text{sidebands}} \cdot n_{\text{channels}} = 19 \cdot 4 \cdot 1024$ [34].

Next, we find the covariance matrix \mathbf{C} of this data matrix by taking the outer product between \mathbf{D} with itself [34];

$$\mathbf{C} \propto \mathbf{D}^T \mathbf{D}. \quad (3.16)$$

The eigenvectors \mathbf{v}_k of this covariance matrix with the largest associated eigenvalues λ_k will correspond to the direction in frequency-feed space with the largest variance. To find out how much of the TOD points in the direction of some eigenvector \mathbf{v}_k we simply project the TOD at some frequency channel ν into the frequency-feed sub-space spanned by the eigenvector to obtain the PCA amplitude

$$a_k^\nu = \mathbf{d}_\nu \cdot \mathbf{v}_k = \sum_{i=1}^{n_{\text{samp}}} d_\nu^i v_k^i. \quad (3.17)$$

The TOD's PCA mode pointing in the direction of \mathbf{v}_k is subsequently given by

$$\mathbf{d}_\nu^k = a_k^\nu \mathbf{v}_k. \quad (3.18)$$

Finally, to eliminate the component that explains the most correlation in the data we can simply subtract the leading PCA components (normally $n_{\text{comp}} = 4$) from the TOD frequency-wise in each feed i , giving [34]

$$\mathbf{d}_{\text{new}}^{\nu,i} = \mathbf{d}_{\text{old}}^{\nu,i} - \sum_{k=1}^{n_{\text{comp}}} a_k^{\nu,i} \mathbf{v}_k. \quad (3.19)$$

As shown in Fig. 3.2, removing the leading PCA modes from the TOD even further flattens its power spectral density towards that of white noise. The corresponding change of the data in the time domain is illustrated in the last row of Fig. 3.5. As this, however, is also an example of relatively clean data the power of the PCA filter does not show that clearly. If we instead look at Fig. 3.7 we see that even data that are contaminated to a large degree by bad weather and/or standing waves becomes remarkably clean after removing the leading PCA components.

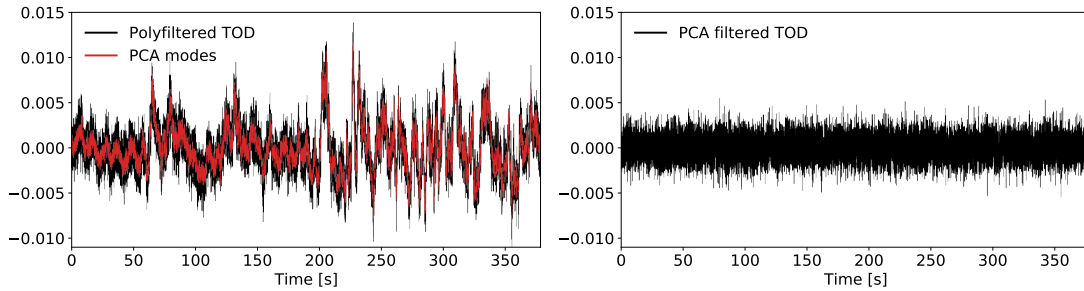


Figure 3.7: Before and after view of the effect of the PCA filter on a TOD that is highly contaminated by bad weather and/or standing waves. Courtesy of Jonas Lunde.

3.4.6 Frequency Masking

Up until this stage of the pipeline, the raw data have been through four different filters that target each of their artifact that must be removed to obtain scientifically viable results. Even still, at this stage, there can still be some frequency channels that do not behave as well as we would like. These channels we would like to mask out, to not contaminate the final cleaned data with unwanted systematic effects.

We do this by first applying the poly- and PCA filters on a copy of the data, and subsequently identify bad channels based on their correlation structure. We then apply the mask of bad channels to the original data, and thereafter poly- and PCA filter the data again. The idea behind this is that deriving a mask of bad channels of somewhat cleaned data is more effective on the raw data than if the mask was made based on the uncleaned data [34].

When identifying bad channels we use two main approaches. In the first approach, we utilize the fact that we can quantify the expected correlation structure of good data. Specifically, since we know that the correlation between two independent Gaussian variables is proportional to $\frac{1}{\sqrt{n_{\text{samp}}}}$, for a (large) number of samples n_{samp} from which the correlations are computed. However, we may also see some known correlations between frequencies induced by the polyfilter. Channels that deviate more than 5σ from the expected white noise and polyfilter correlations are then masked out. This approach is best if we consider longer scans where n_{samp} is large and deviations from the expected correlation are easier to estimate. This in turn makes it difficult to state a threshold for when to mask away a channel, as it depends on the length of the time stream [34].

In the second approach, we compute several diagnostics for each frequency channel, such as the average correlation of a channel to all the others on the same sideband. Thereafter the diagnostics from different frequencies are compared. If any significant outliers are found, we simply mask these out. However, this also brings a disadvantage, since a set of too many bad channels might not show any significant relative deviations that are to be masked out [34].

Another thing that we look for, and often masked out, are channels on the edges

of each frequency sidebands. These frequencies behave especially badly due to strong correlations between channels at the edges of different sidebands, caused by an aliasing effect. Hence we mask out these "edge correlations" of the sidebands [34].

3.4.7 Calibration

Ever since we normalized the raw level 1 data by its running mean, as described in Sec. 3.4.2, the data are on a dimensionless form. However, we want the units of the data to be in units of brightness temperature. To obtain the right units we can simply multiply the data by the system temperature T_{sys} . Meanwhile, as the system temperature is generally different for each frequency, we must do this separately for each channel so that the TOD in temperature units becomes [34]

$$d(\nu, t)|_K = d(\nu, t)\langle T_{\text{sys}}(\nu, t) \rangle \quad (3.20)$$

Therefore we must find the system temperature T_{sys} to perform the calibration.

Fortunately, we have the measurements of the ambient load from Sec. 3.4.1 which we can use to estimate the unknown system temperature. From Eq. (3.1) we see that the power output of the telescope P is proportional to the system temperature T_{sys} , with a constant of proportionality given by $\tilde{G} = Gk_B\Delta\nu$. Hence, if we have measurements of P at two different temperatures, where one of the temperatures is known, we can estimate the unknown one quite easily. The observed object of known temperature should ideally be outside the atmosphere to make it as comparable to the actual CO signal as possible, being subject to the altering effects of the atmosphere in the same way.

However, putting something of known temperature into space is, for obvious reason, highly impractical. Nevertheless, one could in theory use a calibration source with known properties such as Jupiter or Tau A. These sources are, however, not always visible and can hence not always be used. Luckily we can use a few assumptions to circumvent this problem. It turns out that we can use the calibration vane measurements as an approximation for having a measurement of something with ambient temperature in space. That is because both the atmosphere and the calibration vane can be assumed to have the same ambient temperature T_{amb} , thus emitting the same thermal radiation. Because the vane and the air together are an optically thick system, the ordering of the vane and the air becomes irrelevant for the measured telescope output. Of course, the atmosphere having a temperature gradient induced a small error, which however is assumed to be quite small.

Now knowing this we can perform the T_{sys} measurement by comparing the power output of the telescope, P_{cold} and P_{amb} , when observing the cold sky and the ambient load respectively. These two power measurements then simply correspond to Eq. (3.1) respectively evaluated at the corresponding cold sky and ambient system temperatures

$$T_{\text{sys}} = T_{\text{rec}} + T_{\text{atm}} + e^{-\tau}T_{\text{CMB}} \quad (3.21)$$

$$T_{\text{sys}}^{\text{amb}} = T_{\text{rec}} + T_{\text{amb}}. \quad (3.22)$$

The atmosphere, being both an emitter and absorber, removes and adds equally much of the calibration vane's signal if it were in space. This can be reformulated

mathematically as $T_{\text{atm}} = (1 - e^{-\tau})T_{\text{amb}}$. Using this knowledge we can find T_{sys} by dividing P_{cold} by P_{amb} , eliminating the proportionality constant \tilde{G} , and subsequently solve the resulting equation to give the effective system temperature of the telescope

$$T_{\text{sys}}^{\text{eff}} \equiv e^{\tau}T_{\text{sys}} = \frac{T_{\text{amb}} - T_{\text{CMB}}}{\frac{P_{\text{amb}}}{P_{\text{cold}}} - 1}. \quad (3.23)$$

This relation being a function of time, however, requires us to know both P_{cold} and P_{amb} at the same time. But because we cannot observe the sky and the calibration vane at the same time we must use the interpolated value for the ambient load measurements P_{amb} between the measurements at the beginning and end of each obsID. We therefore interpolate the ambient load measurement and corresponding ambient temperature to the time t of each scan so that

$$P_{\text{amb}}(t) = \frac{(t - t_0)P_{\text{amb}}^1 + (t_1 - t)P_{\text{amb}}^0}{t_1 - t_0} \quad (3.24)$$

$$T_{\text{amb}}(t) = \frac{(t - t_0)T_{\text{amb}}^1 + (t_1 - t)T_{\text{amb}}^0}{t_1 - t_0}, \quad (3.25)$$

where P_{amb}^i , T_{amb}^i and t_i are the power output, temperature and time of each ambient load measurement. Now that we have estimates for both the cold and ambient load measurements, P_{cold} and P_{amb} at the time t of each scan, as well as the corresponding temperatures T_{cold} and T_{amb} , we can finally perform the calibration in Eq. (3.20).

3.4.8 Decimation

The last step in generating the level 2 data is to reduce the number of frequency channels in each sideband in the cleaned data. We do this because we simply do not need such a high resolution along the line-of-sight and because the storage requirements and further analysis time is greatly reduced.

The down-sampling of the frequency grid is simply done by coadding a number of adjacent TOD points in frequency. The process of coadding refers to the noise weighted arithmetic average of a number of elements. In the analysis pipeline we normally coadd $n_{\text{dec}} = 16$ neighboring TOD values of the high-resolution channels m to a single low-resolution TOD value [34]

$$d_i^{\text{low}} = \frac{1}{\sum_m w_m} \sum_{m=(i-1)n_{\text{dec}}+1}^{in_{\text{dec}}} w_m d_m^{\text{high}} \quad (3.26)$$

at the downgraded frequency channel i . The noise weighting is done through the inverse variance weights $w_m = \frac{1}{\sigma_n^2}$ of the high-resolution grid. To ensure that masked out frequency channels in the high-resolution grid do not contaminate the resulting low-resolution grid, we set their weights to zero [34].

3.4.9 Highpass Filtering

Sometimes the filtered level 2 data have some residual long-timescale modes that can be removed by highpass filtering the TOD before making maps. In the current setup of the analysis pipeline, we apply the highpass filter by transforming the TOD to frequency space and then cut off the data at a given cut-off frequency before inverse transforming back to the time domain. This step will remove some of the larger temporal modes from the TOD and may hence save potentially bad data at the cost of some large-scale signal loss. Currently, we use a cut-off frequency of 0.02 Hz as the default highpass cut.

3.5 Mapmaking

Now that we have seen how the raw level 1 data are cleaned of correlations it is time for the next step in our analysis. Because our ultimate goal in the COMAP experiment is to use the gathered data to estimate astrophysical and cosmological parameters from its statistical properties, we want to perform a likelihood (or posterior probability) analysis on the data in parameter space. Assuming Gaussian parameter distributions we, therefore, have to compute matrix determinants of $n_{\text{samp}} \times n_{\text{samp}}$ -covariance matrices at many different points in parameter space if the analysis is directly done on the time samples [66]. Now, in our case, a typical obsID has of the order $n_{\text{samp}} \sim 10^5$ samples in time per frequency per detector alone. Thus, computing a determinant of the resulting covariance matrix, scaling as $\mathcal{O}(n_{\text{samp}}^3)$, quickly becomes completely unpractical [66].

We must therefore perform a data reduction, without losing any statistically valuable information, to ease the later scientific parameter estimations. To do this we want to project the TOD of each frequency and feed into an image of the observed sky patch. As we in COMAP typically have maps of $n_{\text{pix}} = 120 \times 120$ samples per frequency per detector, regardless of the length of the time stream, the data volume is drastically reduced.

The conversion of the data from the time to the pixel domain is done via a linear transformation;

$$\mathbf{m} \approx \hat{\mathbf{m}} = \mathbf{W}\mathbf{d}, \quad (3.27)$$

in which the map \mathbf{m} is estimated by the vector of pixels $\hat{\mathbf{m}}$ found by applying a $n_{\text{pix}} \times n_{\text{samp}}$ matrix \mathbf{W} to the TOD vector \mathbf{d} [66]. The exact form of the projection matrix \mathbf{W} will be dependent on the mapmaking scheme used, which we in the following will see is dependent on the noise properties of the TOD. We will in the following subsection consider the methods we use for making maps in the `tod2comap` script in our current COMAP pipeline and will discuss a possible extension to the current framework in the methods (see Part II).

3.5.1 Maximum Likelihood Estimates

When choosing a mapmaking scheme we want to use a method that can compress the information in the TOD to a much smaller data volume and at the same time lose as little as possible of the cosmologically and astrophysically valuable information. A method

often used for this is a maximum likelihood (ML) estimate. Using an ML-estimate will create unbiased minimum-variance maps from a TOD with both correlated and uncorrelated noise, and will hence form the basis for all mapmaking schemes considered in this thesis. Maximum likelihood estimates are very popular because in the limit of large amounts of data the ML-estimate will saturate the Cramér-Rao bound, thus becoming asymptotically the best-unbiased-estimator of the map [67].

We will therefore have a look at how to arrive at the ML-estimate of a map. We first assume the TOD can be written as

$$\mathbf{d} = \mathbf{P}\mathbf{m} + \mathbf{n}, \quad (3.28)$$

consisting of the astrophysical signal map \mathbf{m} , with length n_{pix} which is then projected from the pixel domain to the time domain by the pointing matrix \mathbf{P} [68]. The pointing matrix is a $n_{\text{samp}} \times n_{\text{px}}$ matrix containing ones for each pixel i that is hit by the telescope beam at a time t , and zero otherwise. The random noise n that is added to the signal can in general contain both white uncorrelated noise and correlated components.

As often is the case we can assume that the noise \mathbf{n} is distributed according to a Gaussian likelihood

$$L \equiv P(\mathbf{d}|\mathbf{m}) = \frac{1}{\sqrt{2\pi \det \mathbf{N}}} \exp\left(-\frac{1}{2}\mathbf{n}^T \mathbf{N}^{-1} \mathbf{n}\right), \quad (3.29)$$

where the time-domain covariance matrix $\mathbf{N} = \langle \mathbf{n}\mathbf{n}^T \rangle$, as the expectation value of the noise, $\langle n \rangle = 0$, vanishes [66]. Assuming white noise only, the covariance matrix will be diagonal, however, we can generalize this case to also include correlated noise by either allowing \mathbf{N} to have off-diagonal elements, or by including a prior to the likelihood. In the latter case the ML-estimate becomes a maximum posterior (MP) estimate.

If the data model in Eq. (3.28) is then solved for the noise \mathbf{n} and subsequently inserted into the likelihood function of Eq. (3.29) we obtain the log-likelihood

$$\mathcal{L} \equiv \log L \propto -\frac{1}{2}(\mathbf{d} - \mathbf{P}\mathbf{m})^T \mathbf{N}^{-1}(\mathbf{d} - \mathbf{P}\mathbf{m}). \quad (3.30)$$

Maximizing the likelihood (3.29) is then equivalent to maximizing the log-likelihood Eq. (3.30) as a logarithm is a monotonically increasing function.

We can then find the ML-estimate of the map $\hat{\mathbf{m}}$ that maximizes the log-likelihood by letting the gradient of \mathcal{L} with respect to the map \mathbf{m} vanish, i.e.

$$\nabla_{\mathbf{m}} \mathcal{L} = \frac{d\mathcal{L}}{d\mathbf{m}} = \mathbf{P}^T \mathbf{N}^{-1}(\mathbf{d} - \mathbf{P}\mathbf{m}) + (\mathbf{d} - \mathbf{P}\mathbf{m})^T \mathbf{N}^{-1} \mathbf{P} = 0. \quad (3.31)$$

Because \mathbf{N}^{-1} is symmetric we get $(\mathbf{d} - \mathbf{P}\mathbf{m})^T \mathbf{N}^{-1} \mathbf{P} = (\mathbf{P}^T \mathbf{N}^{-1}(\mathbf{d} - \mathbf{P}\mathbf{m}))^T$. Therefore Eq. (3.31) is fulfilled only if

$$\mathbf{P}^T \mathbf{N}^{-1}(\mathbf{d} - \mathbf{P}\hat{\mathbf{m}}) = 0. \quad (3.32)$$

Solving for the ML-estimate $\hat{\mathbf{m}}$ yields the so-called mapmaker equation

$$\hat{\mathbf{m}} = (\mathbf{P}^T \mathbf{N}^{-1} \mathbf{P})^{-1} \mathbf{P}^T \mathbf{N}^{-1} \mathbf{d}, \quad (3.33)$$

which is on the form of Eq. (3.27) with $\mathbf{W} = (\mathbf{P}^T \mathbf{N}^{-1} \mathbf{P})^{-1} \mathbf{P}^T \mathbf{N}^{-1}$.

The map found by this general form of an ML-estimate is applicable for a general case with correlated noise as this is taken into account by the covariance matrix \mathbf{N} . Nevertheless, this method trades generality for cost, as it is expensive to compute inverses of the covariance matrix when \mathbf{N} has off-diagonal elements. Luckily things get simpler in the case of uncorrelated white noise, in which case the covariance matrix is simply represented by a diagonal matrix that is trivial to invert.

3.5.2 Noise Weighted Binning

Because we manage to remove most correlated artifacts from the raw data using `12gen` we can assume that the level 2 TOD is close to white noise [34]. However, even though most correlated noise is removed by the filtering described earlier so will some of the signal in the data. Nevertheless, we can take this bias into account, as we will discuss in more detail in Ch. 5, by using simulated signals.

As our cleaned data by now are mainly dominated by white noise we can assume the covariance matrix \mathbf{N} of the remaining noise to be diagonal. Then the mapmaker equation (3.33) has the nice property of reducing to a simple noise weighted binning of the TOD into pixel bins. Specifically, the matrix product from Eq. (3.33)

$$\mathbf{C}_m = (\mathbf{P}^T \mathbf{N}^{-1} \mathbf{P})^{-1} \quad (3.34)$$

is easily identified as the covariance matrix of the final map estimator $\hat{\mathbf{m}}$ [68]. Having only white noise the covariance matrix of the final map simply reduces to a diagonal matrix. The n_{pix} diagonal elements of \mathbf{C}_m will then equal the sum of all inverse variances of the time samples that correspond to a pixel i

$$C_m^{ii} = \left(\sum_j \frac{1}{\sigma_j^2} \right)^{-1}. \quad (3.35)$$

Correspondingly $\mathbf{P}^T \mathbf{N}^{-1} \mathbf{d}$ reduces to sum of the noise weighted TOD data points that hit each pixel; $\sum_i d_i / \sigma_i^2$. Hence the full mapmaker equation in case of white noise only becomes the simple noise weighted binning algorithm [34]

$$m = \frac{\sum_i d_i / \sigma_i^2}{\sum_i \frac{1}{\sigma_i^2}}. \quad (3.36)$$

We see that this simply amounts to coadding up all the TOD samples that hit each pixel bin.

Figure 3.8 illustrates three maps, one per CO field, made by the noise weighted binning of all our currently available data on each field. As the telescope sweeps across the field center more often than the field edges, the noise in the center gets integrated down the most in the mapmaking process.

Although the noise weighted binning method is quite simple, it is only really applicable in case the TOD is dominated by white noise. If this is not the case and correlated components are included the situation becomes more complex as the inversion of \mathbf{N} becomes very time-consuming. In such a case one can consider a more sophisticated method that is better suited to handle residual artifacts. We will thus later in Part II discuss how we can improve our current COMAP pipeline by using a destriper instead of a noise weighted binning scheme in the mapmaker.

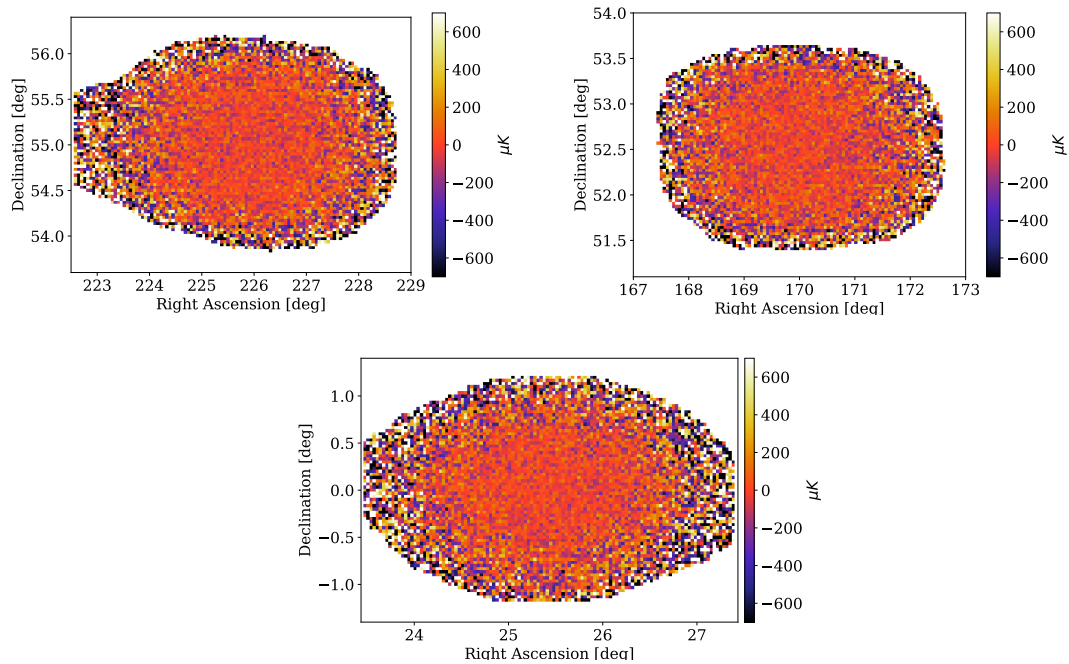


Figure 3.8: Maps of the three fields CO2 (**bottom**), CO6 (**top left**) and CO7 (**top right**), each made from all the current data (~ 1800 obsIDs) per field.

3.6 High-Level Analysis

Now that we have seen how the raw telescope data are cleaned from correlated noise and systematics as well as how we project the cleaned TOD into a map, it is time to consider how we go from the statistical properties of a map to a final estimate of astrophysical parameters. However, as modeling the CO signal, the theoretical statistical properties and estimating parameters from these is not the scope of this thesis we will here only briefly outline how we can estimate parameters from a map.

3.6.1 The Theoretical CO Power spectrum and Voxel Intensity Distribution

As we ultimately want to estimate physical parameters that affect the structure of the astrophysical signal we hope to see in the COMAP data, we need to form some quantity that captures the statistical properties of the data. From these statistical properties, we can subsequently do a parameter estimation, comparing the data to some theoretical model, to gain an understanding of the physics that governs the Universe. The two statistics that are most frequently used in intensity mapping are the power spectrum and the voxel intensity distribution (VID) [10, 69, 29, 48]. From these we can constrain for instance the CO luminosity function, SFR, etc. at the Epoch of Galaxy Assembly.

The CO Power Spectrum

The power spectrum, to which we will come back to in more detail in Ch. 4, is a two-point statistic that roughly speaking measures the two-point correlation in map space. If the CO field we observe with the COMAP telescope was a perfect Gaussian random field the power spectrum would incorporate all the statistical information in the field. However, because the small-scale CO emitters within a dark matter halo have been affected by non-linear processes the field of emitters is not entirely Gaussian, hence the power spectrum only captures part of the statistical information of the map. Nevertheless, the power spectrum still contains a lot of interesting information [48].

The scale and redshift dependent theoretical line emission power spectrum, for instance for CO($J = 1 \rightarrow 0$) emission, is commonly written as

$$P(k, z) = \bar{T}^2(z) \bar{b}^2(z) P_m(k, z) + P_{\text{shot}}(z). \quad (3.37)$$

The first term, commonly referred to as the clustering component of the power spectrum is made up of the dark matter power spectrum $P_m(k, z)$ which the CO emitters, of average intensity (in units brightness temperature) $\bar{T}(z)$, trace up to the average luminosity weighted bias $\bar{b}(z)$. The second term, referred to as the shot noise term P_{shot} , is caused by the fact that the CO signal is sourced by discrete emitters, which results in a scale invariant power spectrum contribution [48].

Given some luminosity function $\Phi(L)$, quantifying the number of emitters at a given luminosity L , we can write the luminosity weighted bias as

$$\bar{b} = \frac{\int L b(L) \Phi(L) dL}{\int L \Phi(L) dL}. \quad (3.38)$$

Here $b(L)$ is the bias of the emitters of luminosity L , which determines how well the galaxies at L follow the underlying dark matter field. However, because the galaxies, with power spectrum $P_{\text{gal}}(k) = \bar{b}^2 P_m(k)$, are observed by their CO line emission the clustering term must be weighted by the averaged CO intensity, which is written as [48]

$$\bar{T} = \frac{c^3(1+z)^2}{8\pi k_B \nu_{\text{em}}^3 H(z)} \int_0^\infty L \Phi(L) dL. \quad (3.39)$$

Similarly, the shot noise term caused by the discrete nature of the emitters can be related to the luminosity function by [48]

$$P_{\text{shot}} = \left(\frac{c^3(1+z)^2}{8\pi k_B \nu_{\text{em}}^3 H(z)} \right)^2 \int_0^\infty L^2 \Phi(L) dL, \quad (3.40)$$

which hence quantifies the intrinsic "graininess" of the source field.

The Voxel Intensity Distribution

Meanwhile, as mentioned before, using solely the power spectrum when describing a non-Gaussian field will cause a significant loss of statistical information. Hence a complementary statistic is needed; for instance the VID. The VID is, simply said, the histogram of observed intensities, and is as such more closely related to the luminosity function itself [48, 69]. In particular, the VID can be written as

$$\mathcal{P} = \sum_{N=0}^{\infty} \mathcal{P}_N(T) \mathcal{P}(N), \quad (3.41)$$

where the probability of observing N sources is denoted by $\mathcal{P}(N)$ and the corresponding probability of these N emitters having a brightness temperature of T is given by $\mathcal{P}_N(T)$ [48]. From the VID we can then subsequently get the expected number of emitters in each voxel

$$\langle B_i \rangle = N_{\text{vox}} \int_{T_i}^{T_{i+1}} \mathcal{P}(T) dT, \quad (3.42)$$

within each temperature bin $T \in [T_i, T_{i+1}]$. This histogram of temperatures then forms the basic observable related to the VID [69].

3.6.2 Bayesian Parameter Estimation: MCMC

When estimating parameters from a dataset, such as the one we have in COMAP, and some parameterized model one typically constructs a posterior distribution. The posterior probability is given by

$$P(\theta|\mathbf{d}) \propto P(\mathbf{d}|\theta)P(\theta), \quad (3.43)$$

for the likelihood $P(\mathbf{d}|\theta)$ quantifying the probability of observing the data \mathbf{d} given some model parameters θ (e.g. SFR at the Epoch of Galaxy Assembly) and the prior probability on the model parameters $P(\theta)$ that quantifies our prior knowledge of the parameters. The posterior probability is then typically sampled using Markov Chain Monte Carlo (MCMC) methods like the Metropolis or Gibbs algorithms. The sampling is beyond the scope of this work and we will hence not consider it in great depth. The idea behind these MCMC methods is to make a "jump" in parameter space and either reject or accept the move according to some acceptance criterion (which differs between MCMC methods) [70]. The resulting points corresponding to each step will

then generate a point cloud in parameter space that after a burn-in period sample the posterior distribution. Thus the point cloud density at any point in parameter space will reflect the posterior probability. From the latter we can then easily read the parameter constraints [67].

Note that to do the sampling we need a model, in addition to the data. A model we often use in COMAP is the Li et al. [29] model. It uses the relation between dark matter halo mass and SFR from Behroozi et al. [71, 72], to relate the dark matter halo mass to the CO luminosity by a series of parametrized relations. Specifically, the SFR is coupled to the infrared (IR) luminosity L_{IR} by

$$\text{SFR} = \delta_{\text{MF}} \times 10^{-10} L_{\text{IR}}. \quad (3.44)$$

The IR luminosity is then subsequently related to the CO luminosity through the relation

$$\log L_{\text{IR}} = \alpha \log L'_{\text{CO}} + \beta. \quad (3.45)$$

Additionally log-normal scatters are determined by σ_{SFR} and $\sigma_{L_{\text{CO}}}$, and added to respectively SFR and L_{CO} . Hence the parameters of the model that the MCMC samples over are $\theta = \{\sigma_{\text{SFR}}, \log \delta_{\text{MF}}, \alpha, \beta, \sigma_{L_{\text{CO}}}\}$ [29].

When it comes to the observed data, Ihle et al. [69] showed that one can use the combined leverage of the power spectrum and VID to constrain the model parameters. In doing so, the combined parameter constraints from the power spectrum and the VID are superior to the ones obtained from either statistic alone, as more of the available information is utilized.

Part II

Methods

Chapter 4

Estimating Power Spectra

In the previous chapters of this thesis, we have seen how we clean the raw telescope data from correlated noise and systematics, project it into maps as well as showing a rough outline of how the map's statistical properties (e.g. the power spectrum) can be used to estimate astrophysical parameters. We have, however, not yet considered how we compute power spectra from a map numerically. We will in this chapter thus present how we compute estimates of the power spectra which we will use later, in Ch. 5, when presenting the improvements to the COMAP pipeline that we implement in this thesis. Specifically, we will consider both auto- and cross-power spectra, respectively quantifying the variance and covariance of structures in maps.

In the following chapter, the theory on estimating auto- and cross-spectrum analysis is unless otherwise stated, based on the methods presented by Ihle et al. [73].

4.1 The Auto-Power Spectrum

The auto-power spectrum (sometimes just referred to as "the power spectrum") of a three-dimensional map $x_{m_1 m_2 m_3}$ is closely related to the Fourier transformation of the map. More precisely the power spectrum equals the variance of the absolute square of the Fourier coefficients of the map $x_{m_1 m_2 m_3}$. If the transverse direction of the 3D map would cover a large sky fraction we would have to use a spherical harmonic decomposition instead of a Fourier transform. However, because we in COMAP until now only have observed relatively small sky patches we will only consider Fourier transformations. The Fourier coefficients of a map, and its inverse, are given as

$$f(\mathbf{k}) = \int x(\mathbf{r}) e^{-2\pi i \mathbf{r} \cdot \mathbf{k}} d\mathbf{r} \quad (4.1)$$

$$x(\mathbf{r}) = \int f(\mathbf{k}) e^{2\pi i \mathbf{r} \cdot \mathbf{k}} d\mathbf{k} \quad (4.2)$$

if we would consider a continuous case [74]. Our maps in COMAP, however, are discretized into 3D voxels and thus we must use a discrete Fourier transform. In 1D the

Fourier transform is given as

$$f_l = \sum_{m=0}^{n-1} x_m \exp\left(-2\pi i \frac{ml}{n}\right), \quad l = 0, \dots, n-1 \quad (4.3)$$

where f_l are the Fourier coefficients of the 1D map x_m , both in this example in 1D. The inverse of this relation is simply given as

$$x_m = \frac{1}{n} \sum_{l=0}^{n-1} f_l \exp\left(2\pi i \frac{ml}{n}\right). \quad (4.4)$$

The corresponding wave number will then be given as

$$k = \frac{2\pi j}{\Delta x n}, \quad \text{for } j \in \left\{-\frac{n}{2}, \dots, -1, 0, 1, \dots, \frac{n}{2}\right\}, \quad (4.5)$$

which will be used to quantify the size of a given scale.

However, the maps we consider in COMAP are given along two perpendicular angular directions, x and y , as well as the dimension parallel to the line-of-sight. Hence the wave number will become a vector $\mathbf{k} = (k_x, k_y, k_z)$, with each k_i defined analogously to the 1D case. The resulting Fourier transform of the map will be along three axis. More precisely, we get the three-dimensional Fourier transformation and its inverse given as

$$f_{l_1 l_2 l_3} = \sum_{m_1, m_2, m_3=0}^{n-1} x_{m_1 m_2 m_3} \exp\left(-2\pi i \frac{m_1 l_1 + m_2 l_2 + m_3 l_3}{n}\right) \quad (4.6)$$

$$x_{m_1 m_2 m_3} = \frac{1}{n^3} \sum_{l_1, l_2, l_3=0}^{n-1} f_{l_1 l_2 l_3} \exp\left(2\pi i \frac{m_1 l_1 + m_2 l_2 + m_3 l_3}{n}\right). \quad (4.7)$$

With the 3D Fourier transformation in hand we can define the power spectrum by the variance of the absolute square of the Fourier coefficients;

$$P(\mathbf{k}) = \frac{V_{\text{vox}}}{N_{\text{vox}}} \langle |f_{\mathbf{k}}|^2 \rangle. \quad (4.8)$$

Here the number of voxels $N_{\text{vox}} = n_x n_y n_z$, where n_i is the number of grid points along each of the three axis. The comoving volume of a voxel V_{vox} is given by the comoving grid resolution $V_{\text{vox}} = \Delta x \Delta y \Delta z$. We will come back to how the maps grid is defined in terms of cosmological distances in Sec. 4.3.

Because the two perpendicular directions, k_x and k_y , behave statistically the same we can merge the perpendicular modes by binning the three-dimensional power spectrum in perpendicular bins of size $k_{\perp} = \sqrt{k_x^2 + k_y^2}$. The parallel modes will then be represented in $k_{\parallel} = |k_z|$ bins. This way we can reduce the dimensionality of the power spectrum, while keeping the same statistical information. To obtain an estimate of the 3D power

spectrum in terms of the two variables k_{\parallel} and k_{\perp} we can bin up the 3D power spectrum (4.8) in bins $\{\mathbf{k}_i\}$ to get

$$P(\mathbf{k}_i) \approx \frac{V_{\text{vox}}}{N_{\text{vox}}N_{\text{modes}}} \sum_{j=0}^{N_{\text{modes}}} |f_{\mathbf{k}_j}|^2 \equiv P_{\mathbf{k}_i}. \quad (4.9)$$

Here N_{modes} is the number of modes binned up in bin i .

4.1.1 The Pseudo-Power Spectrum

Because we can further split up the power spectrum of some map into its signal and noise part

$$P(\mathbf{k}) = P_{\text{signal}}(\mathbf{k}) + P_{\text{noise}}(\mathbf{k}), \quad (4.10)$$

the power spectrum will be highly dependent on the properties of the noise in the map. For instance assuming that a map dominated by uniform white noise its power spectrum will be estimated to be

$$P_{\text{noise}} = V_{\text{vox}}\sigma_0^2, \quad (4.11)$$

being scale invariant. The white noise level in each voxel is then given by the radiometer equation defined in Eq. (3.2) and depends on the frequency resolution $\Delta\nu$ of a voxel, the total integration time for a pixel τ as well as the system temperature T_{sys} of the telescope.

Yet, as we saw in Fig. 3.8, the maps we consider seldom have uniform noise properties, due to the higher total integration time per pixel in the map center compared to the edge. The resulting noise power spectrum will therefore receive a high contribution from the high-noise regions, which in turn will bias our signal power spectrum estimate. For each map region to contribute equally we can therefore weight the map by its noise before power spectra are computed. We then get the so-called pseudo-power spectrum

$$\tilde{P}(\mathbf{k}) = \frac{V_{\text{vox}}}{N_{\text{vox}}} \langle |f_{\mathbf{k}}|^2 \rangle. \quad (4.12)$$

instead of the regular power spectrum defined earlier in Eq. (4.8). The Fourier coefficients used in (4.12) are this time defined analogously, but with the map x_m in (4.7) exchanged with the noise weighted map

$$\tilde{x}_m \equiv w_m x_m. \quad (4.13)$$

The noise weights w_m are defined as

$$w_m = \frac{\mathcal{N}}{\sigma_m^2}, \quad (4.14)$$

with a noise map σ_m that quantifies the white noise level in each voxel and an overall normalization constant \mathcal{N} .

Nevertheless, solving the problem of non-uniform noise in the map by using the noise weighted maps instead introduces another problem. When applying a noise weighing

or masking (i.e. infinite noise) to the map, the different Fourier modes of the map get mixed up when computing power spectra [74, 73].

To see how this happens we can, without loss of generality, consider the discrete Fourier transform of a 2D noise weighted map

$$\tilde{f}_{l_1 l_2} = \sum_{m_1=0}^{n-1} \sum_{m_2=0}^{n-1} w_{m_1 m_2} x_{m_1 m_2} \exp\left(-2\pi i \frac{m_1 l_1 + m_2 l_2}{n}\right). \quad (4.15)$$

Inserting the inverse coefficients, i.e. the map, back into this equation yields

$$\tilde{f}_{l_1 l_2} = \frac{1}{n^2} \sum_{\substack{m_1, m_2=0 \\ l'_1, l'_2=0}}^{n-1} w_{m_1 m_2} f_{l'_1 l'_2} \exp\left(-2\pi i \frac{m_1(l_1 - l'_1) + m_2(l_2 - l'_2)}{n}\right).$$

Subsequently, by inserting the weights in terms of their Fourier transform we can expand our expression for the weighted Fourier coefficients to

$$\begin{aligned} \tilde{f}_{l_1 l_2} &= \frac{1}{n^4} \sum_{m_1, m_2=0}^{n-1} \sum_{l'_1, l'_2=0}^{n-1} \sum_{l''_1, l''_2=0}^{n-1} \tilde{w}_{l'_1 l'_2} f_{l'_1 l'_2} \\ &\cdot \exp\left(-2\pi i \frac{m_1(l_1 - l'_1 - l''_1) + m_2(l_2 - l'_2 - l''_2)}{n}\right). \end{aligned} \quad (4.16)$$

As the exponential functions for large enough n are well approximated by a Kronecker delta, and due to their complex nature behave periodically, we can write

$$\begin{aligned} \tilde{f}_{l_1 l_2} &= \frac{1}{n^2} \sum_{l'_1, l'_2=0}^{n-1} \sum_{l''_1, l''_2=0}^{n-1} \tilde{w}_{l'_1 l'_2} f_{l'_1 l'_2} \delta_{l'_1(l_1 - l'_1) \% n, l''_1(l_1 - l'_1) \% n} \delta_{l'_2(l_2 - l'_2) \% n, l''_2(l_2 - l'_2) \% n} \\ &= \sum_{l'_1, l'_2=0}^{n-1} f_{l'_1 l'_2} K_{l_1 l_2 l'_1 l'_2}, \end{aligned} \quad (4.17)$$

where the matrix $K_{l_1 l_2 l'_1 l'_2} \equiv \frac{1}{n^2} \tilde{w}_{(l_1 - l'_1) \% n, (l_2 - l'_2) \% n}$. The modulo operation is here defined by $\%$.

To see how these pseudo-Fourier coefficients affect the power spectrum we insert them into Eq. 4.12 and go back to the general N -dimensional case to get

$$\tilde{P}(\mathbf{k}') = \frac{V_{\text{vox}}}{N_{\text{vox}}} \langle \tilde{f}_{\mathbf{k}} \tilde{f}_{\mathbf{k}'}^* \rangle = \frac{V_{\text{vox}}}{N_{\text{vox}}} \frac{1}{n^{2N}} \sum_{\mathbf{k}'} \sum_{\mathbf{k}''} \langle \tilde{f}_{\mathbf{k}'} \tilde{f}_{\mathbf{k}''}^* \rangle K_{\mathbf{k}, \mathbf{k}'} K_{\mathbf{k}, \mathbf{k}''}^*. \quad (4.18)$$

Finally we can define the mode mixing matrix $M_{\mathbf{k}, \mathbf{k}'} \equiv \frac{1}{n^{2N}} |K_{\mathbf{k}, \mathbf{k}'}|^2$ and write

$$\tilde{P}(\mathbf{k}) = \sum_{\mathbf{k}'} P(\mathbf{k}') M_{\mathbf{k}, \mathbf{k}'}. \quad (4.19)$$

Thus noise weighing of the maps will to some extent bias the power spectrum by some mode coupling matrix $M_{\mathbf{k},\mathbf{k}'}$. By doing so, the pseudo-power spectrum will be distorted with respect to its unbiased equivalent. Because we know exactly how maps were weighted before computing the power spectrum, we can in theory invert Eq. (4.19) to find the unbiased power spectrum estimator $P(\mathbf{k})$. An algorithm doing just that already exists; the MASTER algorithm, which scales as $\mathcal{O}(n^3)$ for a 2D map [74, 75]. However, for our 3D maps in COMAP the scaling will be even worse, making it computationally expensive to invert the mode coupling. We will hence, for now, simply stick to the pseudo-power spectra in our analysis and keep in mind that the power spectra we estimate here are biased by a mode coupling kernel.

4.2 The Cross-Spectrum

The second statistic often used to quantify structures in maps is the so-called cross-spectrum. While the auto-power spectrum measures the variance of the Fourier modes of a given map, the cross-spectrum measures the covariance between the Fourier modes of two distinct maps. Hence it can be defined as

$$C(\mathbf{k}) = \frac{V_{\text{vox}}}{N_{\text{vox}}} \langle \text{Re}\{f_{1\mathbf{k}}^* f_{2\mathbf{k}}\} \rangle \approx \frac{V_{\text{vox}}}{N_{\text{vox}} N_{\text{modes}}} \sum_{j=0}^{N_{\text{modes}}} \text{Re}\{f_{1\mathbf{k}_j}^* f_{2\mathbf{k}_j}\} \equiv C_{\mathbf{k}_i}, \quad (4.20)$$

where $f_{i\mathbf{k}}$, for $i = 1, 2$, are the Fourier coefficients of two different sky maps. We thus see that if we were to compute the cross-spectrum between a map and itself, we would get the auto-power spectrum back.

Because a cross-spectrum by nature measures correlations between two maps we can, if the maps are independent, eliminate contributions to the maps that are not common among both. This is in fact quite powerful as it means that the noise, which can be assumed to be independent for two different maps, will not bias the signal estimate from the cross-spectrum. Hence we can write that

$$\langle C_{\mathbf{k}_i} \rangle = P_{\text{signal}}(\mathbf{k}_i), \quad (4.21)$$

where P_{signal} is the power spectrum of the signal that is common in both maps. It is, however, also important to note that P_{signal} will pick up everything common between the maps, both astrophysical signal and common systematic effects. Hence a systematic error that affects the two maps differently will be effectively canceled in a cross-spectrum analysis.

Meanwhile, to obtain an estimate of the signal power P_{signal} from the auto-spectrum analysis we needed to estimate the noise extremely well or else it could bias the signal estimate. This is, nevertheless, not as important in the case of a cross-spectrum analysis since the noise only affects the uncertainties, and not the estimate itself. Hence having a thorough noise estimate is still important, but only to get a good error estimation.

As previously presented when presenting the auto-power spectrum analysis, we found that the maps from which auto-spectra were computed needed to be noise weighed

to compensate for their uneven noise properties and sky cuts. In the cross-spectrum case, we therefore define the pseudo-cross spectrum in a completely analogous fashion. However, as we combine two separate maps in a cross-spectrum, each with its own noise properties, the maps must be weighted by a combined noise weight. To do this we let the weights

$$w_{1,2} \propto \frac{1}{\sigma_{m_1} \sigma_{m_2}}, \quad (4.22)$$

for the noise σ_{m_1} and σ_{m_2} from respectively map m_1 and m_2 . This weight is applied, the same way as in the auto-spectrum case, to the maps when computing the pseudo-Fourier coefficients of each map.

4.2.1 Data Splits

To compute cross-spectra from maps we need to have two maps to cross-correlate. These two maps should ideally be affected by different sources of systematics, canceling them when determining the cross-spectrum.

Hence we can split the data into N_{split} parts under different split criteria. We could for instance split the data by high/low elevation, high/low ambient temperature or by separating the Lissajous and CES scans from one another. Any such split can be performed on its own, or we can successively split the data under several different split criteria.

In particular, we define two distinct split variables; the test and control variables. They are defined so that the data are always split into N_{ctrl} parts by the control variables for each of the N_{test} test variables. For each test variable there will then be $2^{N_{\text{ctrl}}+1}$ parts of the data. That way we can always split the data according to, for instance, both the scanning strategy used and the elevation, in addition to N_{test} test criteria. With all these data split in hand, we obtain a rich dataset from which to compute cross-spectra between portions of the data with different systematic effects.

Because the current mapmaker does only support single data splits and no successive ones, we modified `tod2comap` to perform the wanted successive splits. Specifically, the `accept_mod` routine takes in a split definition file, which defines which split is a control or test variable, and a runlist to define a list of binary numbers. Each binary number in the list corresponds to a given scan. The bits in the binary number are associated with one of the splits in the split definition file. If we for instance consider a scan with the assigned number 101, where the bits correspond respectively to an elevation, scanning and temperature split, then this particular scan will be at high elevation (i.e. above the median elevation of the scans in the runlist), Lissajous scanning and high ambient temperature.

Then when making maps from the data, the mapmaker reads in the list of binary numbers and sorts the data according to the split they belong to. In doing so we obtain feed maps for each frequency channel and sideband from each of the splits.

There is however a small caveat in splitting the data. Since we when computing average cross-spectra do not want to include the auto-spectrum information from all the cross-combinations, we lose some of the sensitivity. In particular, if we consider

a split into two halves of the data, we lose at least a factor of $\sqrt{2}$ in sensitivity. To compensate for that, we can split the data into a large number of parts to subsequently average over the possible cross-spectrum combinations. In such a case the sensitivity can be estimated by

$$\sigma_C^{N_{\text{split}}} \geq \sqrt{\frac{1}{1 - 1/N_{\text{split}}}} \sigma_P. \quad (4.23)$$

Here the upper limit on the cross-spectrum sensitivity is given by the number of splits N_{split} as well as the optimal sensitivity σ_P from the auto-spectrum analysis found from the full data set used. Hence we see that the limit in sensitivity asymptotically approaches that of the optimal auto-spectrum, in the limit of a large number of splits N_{split} . Hence the loss in sensitivity will be somewhat compensated by computing an average cross-spectrum of a large number of split combinations.

4.2.2 Feed-Feed Pseudo-Cross-Spectrum

The main power spectrum method we currently implemented in the COMAP pipeline is the so-called Feed-Feed Pseudo-Cross-Spectrum (FPXS) analysis. Although the method of computing FPXS is based upon the data splits we implemented in the mapmaker as part of this work, the development of this particular power spectrum method itself, lead by Jowita Borowska and Håvard Tveit Ihle, is outside the scope of this work. However, as we will later suggest how to improve the current FPXS methods by investigating a PCA algorithm to identify and clean FPXS with strong residual systematics, we will in the following only give a brief overview of the FPXS methods.

Because the data from each feed in COMAP is largely unique, as each feed has its independent feed-chain with amplifiers, cables, bandpasses, etc., it is possible to circumvent many of its associated systematics by computing cross-spectra between each combination of feeds. In addition to this, many systematic effects like for instance ground pickup are correlated in time. Therefore, by cross-correlating the data of the high and low elevation split we can eliminate for instance much of the ground pickup and other systematic effects that correlate with the elevation of the telescope pointing.

Therefore, the averaged FPXS is a statistic that is highly robust against systematic effects and at the same time has a high sensitivity. The following shows the main four steps in computing the average FPXS under a single (elevation) split;

1. The data are split as described in Sec. 4.2.1. One of the split variables is then defined as the primary split variable used in computing the cross-spectrum. The upper and lower part of the primary split are denoted as part A and B .
2. Subsequently, 19 feed maps are generated from both part A and B . Thus we denote m_{A_5} as the map from part A and feed 5.

3. From all possible combinations of maps m_{A_j} and m_{B_k} cross-spectra $\tilde{C}_{\mathbf{k}_i}^{jk}$ are computed. The resulting dataset of FPXS can be used to compute an averaged spectrum, or it can function as the basis for a systematic error classification algorithm, as we will see later.
4. Lastly an average FPXS is computed by the noise weighted average of all cross-spectra from the previous step for which $j \neq k$

$$\tilde{C}_{\mathbf{k}_i}^{\text{FPXS}} = \frac{1}{\sum_{j \neq k} 1/\sigma_{\tilde{C}_{\mathbf{k}_i}^{jk}}^2} \sum_{j \neq k} \frac{\tilde{C}_{\mathbf{k}_i}^{jk}}{\sigma_{\tilde{C}_{\mathbf{k}_i}^{jk}}^2}. \quad (4.24)$$

The standard deviation of the i th \mathbf{k} bin of the pseudo-cross-spectrum between maps m_{A_j} and m_{B_k} is then given by $\sigma_{\tilde{C}_{\mathbf{k}_i}^{jk}}$.

4.3 The Map Grid

As we have seen we need to know the comoving volume of the voxels in our map to estimate the power spectra. To do this we can utilize our earlier discussion of cosmological distances in Sec. 1.2.2. In our case, a voxel consists of two distance components perpendicular to the line-of-sight as well as one parallel to it.

Because we want to know the grid resolution of the map in terms of cosmological redshift we must translate the map's native frequency dimension into redshift. The redshift can in turn be turned into a comoving distance. The former can easily be done using the definition of cosmological redshift showed earlier in Eq. (1.8) and using that the frequency of light $\nu = c/\lambda$. Then we get

$$z = \frac{\nu_{\text{em}}}{\nu_{\text{rec}}} - 1, \quad (4.25)$$

where ν_{em} and ν_{rec} are the emitted rest-frame and received frequencies respectively. As we target the CO($J = 1 \rightarrow 0$) line at $\nu_{\text{em}} = 115.27$ GHz and the central frequency channel of the maps is at about 30 GHz, the redshift of the middle of the map is $z_{\text{cent}} \approx 2.9$. Furthermore, we assume that all voxels have approximately the same shape as the one in the middle of the map grid at z_{cent} .

The redshift resolution of the map can then easily be found by differentiating Eq. (4.25) with respect to ν_{rec} and combine it with the original equation. We then get that

$$\frac{dz}{d\nu_{\text{rec}}} \approx \frac{\Delta z}{\Delta \nu_{\text{rec}}} = -\frac{(1+z)^2}{\nu_{\text{em}}}, \quad (4.26)$$

which at $z = z_{\text{cent}} = 2.9$ and $\Delta \nu_{\text{em}} \approx 32.3 \cdot 10^{-3}$ GHz gives $\Delta z \approx 4.3 \cdot 10^{-3}$.

Because the redshift interval corresponding to a voxel depth is a fairly small interval, we can get the comoving depth of the voxel by assuming that the integrand of Eq. (1.9)

stays approximately constant over the short interval. Thus we have

$$\Delta r_c^{\parallel}(z) = \int_{z_{\text{rec}}}^{z_{\text{rec}} + \Delta z_{\text{rec}}} \frac{cdz'}{H(z')} \approx \frac{c\Delta z_{\text{rec}}}{H(z)} \approx -\frac{c(1+z)^2}{H(z)} \frac{\Delta \nu_{\text{rec}}}{\nu_{\text{em}}} \quad (4.27)$$

from expression (4.26) for Δz_{rec} . Using this relation, we can easily translate the map's native frequency grid to its comoving line-of-sight distance equivalent.

Similarly, the map's perpendicular dimension must be translated from angular distance on the sky to comoving distance space at redshift $z_{\text{cent}} = 2.9$ in order to estimate the correct voxel volume. From Eq. (1.10) the comoving perpendicular size of a voxel at z_{cent} will be

$$\Delta r_c^{\perp}(z_{\text{cent}}) = \Delta r_p^{\perp}(z_{\text{cent}})(1 + z_{\text{cent}}) = \Delta \theta r_c(z_{\text{cent}}) = \Delta \theta \int_0^{z_{\text{cent}}} \frac{cdz'}{H(z')}. \quad (4.28)$$

Inserting numbers and assuming every voxel has the same shape as the central one we get $\Delta r_c^{\perp}(z_{\text{cent}}) \approx 2.55 \text{Mpc}/h$ for an angular resolution of $\Delta \theta = (2/60)^\circ$.

However, the map's grid is not completely cartesian, but part of a spherical sky. Therefore we must take into account that the angular distances in direction of right ascension (in celestial coordinates) on the map are modified by a factor $\cos \bar{\theta}$, where $\bar{\theta}$ is the mean declination of the map. This is due to two north-south geodesics on a sphere getting closer when approaching the poles, decreasing their angular separation.

Having computed the comoving grid resolution of an average voxel we can now simply set

$$\Delta x = \Delta r_c^{\perp} \cos \bar{\theta} \quad (4.29)$$

$$\Delta y = \Delta r_c^{\perp} \quad (4.30)$$

$$\Delta z = \Delta r_c^{\parallel}, \quad (4.31)$$

which can be multiplied to get the comoving volume of a voxel needed to define the power spectrum.

Chapter 5

Improvements to the COMAP Pipeline

We have now had a thorough look at how we clean the raw telescope data and make maps from it. There are however several points on which we can improve our current design of the COMAP pipeline.

Therefore, this thesis primarily considers improvements to the following main aspects of the COMAP pipeline; how to use a simulated signal to estimate the factor by which we bias the signal estimate when filtering the raw telescope data, using a destriper in the mapmaking to better resolve large-scale structures of the signal and lastly how to use a principal component analysis approach to aid in identifying and cleaning feed-feed pseudo-cross-spectra with strong signs of residual systematics.

5.1 The Simulation Pipeline

By filtering the raw telescope data from correlated noise and systematic effects in `12gen` we will not only remove the unwanted contaminants in the data, but also some of the astrophysical signal. Knowing the extent to which we bias the signal estimate by the filtering is imperative, as we would otherwise underestimate both the detected signal power and the associated uncertainties. This would in turn reflect negatively in the final astrophysical parameters we want to estimate. Hence we will in the following consider how we can use an extension to the COMAP pipeline to estimate the factor of signal attenuation; the pipeline transfer function. We will refer to this pipeline improvement as the *simulation pipeline*.

The main idea behind the simulation pipeline is to enable us to simulate the pickup of additional artificial signal in the COMAP telescope. Knowing the structure of the generated signal a priori, we can compare the "before" and "after" views of the signal to estimate the pipeline transfer function, which in turn can be used to produce unbiased signal estimates. As seen in Fig. 5.1 the simulation pipeline can be divided into three main branches; the upper branch in which the raw telescope data functioning as our noise analog is processed, the middle branch simulating and processing an artificial

telescope TOD as well as the lower branch formatting the unfiltered simulation.

In the following, we will first discuss how to estimate the filter transfer function and subsequently show how this is in practice done using our simulation pipeline. Unless otherwise stated, the theory on transfer functions presented here is based on Ihle et al. [73].

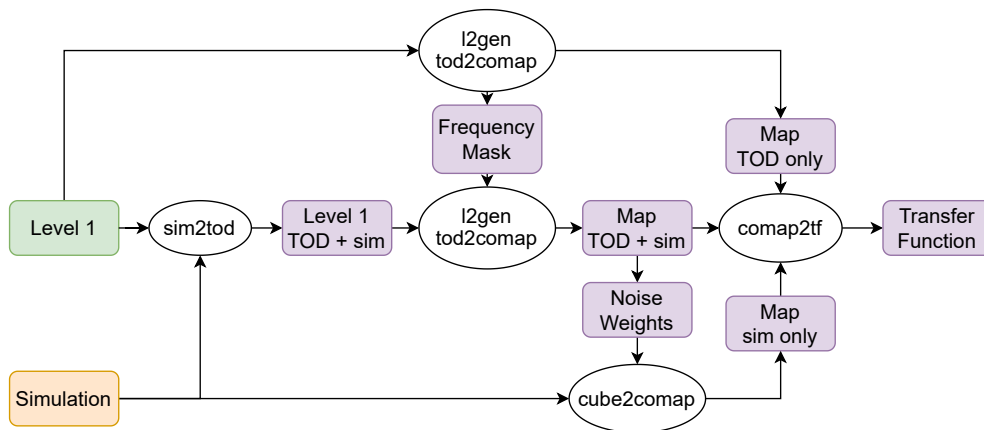


Figure 5.1: Flow chart of the simulation pipeline extension to the COMAP pipeline, with which the filter transfer function, used to produce unbiased signal power spectra, can be estimated. The simulation pipeline takes in raw level 1 data (green box, produced at Caltech) and a simulation cube (yellow box, produced by the COMAP modeling group). These are processed by several scripts (white ellipses) producing intermediate and final results (purple boxes).

5.1.1 The Pipeline Transfer Function

As we have seen in Eq. (4.10), we can decompose the (pseudo-) auto-power spectrum of a given map in terms of its signal and noise components. Similarly, a cross-spectrum estimator $C_{\mathbf{k}}$ can be divided into its signal and noise contribution. However, what we have not taken into account so far is that there is an inevitable attenuation of the astrophysical signal when filtering the raw data.

To model the combined effect of the pipeline filters on the received signal we can modify the decomposition of the auto- and cross-power spectrum estimators to respectively

$$P_{\mathbf{k}} = T(\mathbf{k})P_{\mathbf{k}}^{\text{signal}} + P_{\mathbf{k}}^{\text{noise}}, \quad (5.1)$$

$$C_{\mathbf{k}} = T(\mathbf{k})P_{\mathbf{k}}^{\text{signal}} + C_{\mathbf{k}}^{\text{noise}} \quad (5.2)$$

where $T(\mathbf{k})$ is the so-called pipeline transfer function and $C_{\text{noise}}(\mathbf{k})$ is the cross-spectrum of the noise (which from Eq. (4.21) has a vanishing expectation value). The transfer function will thus quantify how much the pipeline attenuated the signal compared to when it was received by the telescope. Ideally speaking we would therefore like the

transfer function to be as close to unity, as this would mean that no signal has been lost at any given scale \mathbf{k} . It is important to also note that there are other transfer functions, such as for the beam smoothing and pixel resolution of the telescope, in addition to the one of the analysis pipeline. We will, however, only consider the pipeline transfer function in the following as we in this work are interested in how the signal is attenuated by the filtering in the pipeline.

To obtain a transfer function estimate we can simply solve Eq. (5.1) to obtain

$$T(\mathbf{k}) \approx \frac{P_{\mathbf{k}} - P_{\mathbf{k}}^{\text{noise}}}{P_{\mathbf{k}}^{\text{signal}}}, \quad (5.3)$$

in terms of the auto-power spectrum estimate of a map $P_{\mathbf{k}}$, the noise power spectrum $P_{\mathbf{k}}^{\text{noise}}$ as well as the signal power spectrum $P_{\mathbf{k}}^{\text{signal}}$. Alternatively, it should be equivalently possible to estimate the transfer function using the cross-spectrum $C_{\mathbf{k}}$ from Eq. (5.2) as

$$T(\mathbf{k}) \approx \frac{C_{\mathbf{k}} - C_{\mathbf{k}}^{\text{noise}}}{P_{\mathbf{k}}^{\text{signal}}}, \quad (5.4)$$

which ideally should be equal to the transfer function found using the auto-spectrum approach in Eq. (5.3). Nevertheless, because the computational effort needed to estimate the pipeline transfer function is considerably less when choosing the auto-spectrum approach we will mainly consider the the transfer function as found by using auto-spectra. That said, we will still compute the transfer function using cross-spectra as a consistency check, to confirm that the two approaches are in fact largely equivalent.

However, even though these expressions look fairly simple there are a few problems we need to handle to estimate $T(\mathbf{k})$. Firstly, the CO signal the telescope measures is extremely weak and we do not know a priori how the power spectrum of the signal looks before it is affected by the pipeline. Secondly, in order to obtain an estimate of $T(\mathbf{k})$ we need an estimate of the noise $P_{\mathbf{k}}^{\text{noise}}$ or $C_{\mathbf{k}}^{\text{noise}}$.

To solve the first of these issues we must therefore generate some artificial signal and add it to the TOD. That way we can simulate the telescope picking up additional signal which is fully known. The resulting auto- or cross-spectrum, $P_{\mathbf{k}}$ or $C_{\mathbf{k}}$, will then be of a map made from a raw TOD with additional simulated signal.

When it comes to the second issue of modeling the noise power spectrum we could of course simply generate white noise TODs from the radiometer equation (3.2) to subsequently compute maps and power spectra from these. However, even though a white noise TOD would have the same local white noise properties of the TOD with additional signal, it would miss out on the possible residual correlated noise. Therefore we instead choose to use the auto- and cross-power spectrum of a raw TOD without any added signal. That is because the raw TOD is highly noise-dominated and besides has the same noise properties as the TOD with additional signal.

Finally, a third important point is that the estimates of Eq. (5.3) and (5.4) are only estimated from one realization of the simulated signal. However, different realizations of the signal will vary from one another, especially in their largest scales. Therefore, the transfer function computed from any given realization of the signal is merely an

estimate. To get the most robust estimate of the transfer function we must compute several estimates according to Eq. (5.3) and subsequently compute its ensemble average

$$T(\mathbf{k}) = \left\langle \frac{P_{\mathbf{k}} - P_{\mathbf{k}}^{\text{noise}}}{P_{\mathbf{k}}^{\text{signal}}} \right\rangle. \quad (5.5)$$

Note that the transfer function shown here can be estimated using both power and pseudo-power spectra. In the latter case however, the mode coupling of the pseudo spectra may lead to the transfer function looking distorted in a similar way to the power spectra.

Now that we know what is needed to estimate the transfer function by which our signal estimates are biased, we will in the next subsections consider each step in the simulation pipeline by which we estimate the filter-induced bias in practice.

5.1.2 The Upper Branch - Estimating the Noise

As we discussed earlier in Sec. 5.1.1, to find the transfer function of the pipeline filters we need to know the noise power spectrum of our simulated observation. For this, we simply process the raw level 1 TOD (without added simulation) to produce a map using `l2gen` and `tod2comap`. This step is shown in the upper-most branch of Fig. 5.1.

Because we must ensure that the unmodified and simulated level 1 time streams are processed the same way, we need to use the same frequency mask on both time streams in case the masking step, described in Sec. 3.4.6, is performed. However, as we will see later, we will add a simulated signal that may be boosted in strength to make it easier to detect, to the raw telescope data. Therefore, the frequency masking may perceive the amplified signal as unwanted correlated artifacts and subsequently mask it out completely. Thus we implement an option in `l2gen` to import a frequency mask from another level 2 file. Thereby we can first process the unmodified raw data with frequency masking applied, and subsequently import the resulting mask into `l2gen` when processing the modified TOD. Doing this will ensure equal treatment of the two datasets.

5.1.3 The Middle Branch - Simulating a Telescope Time Stream

To estimate the transfer function we need to simulate the telescope measuring some known signal. This can simply be done by taking a cube of simulated astrophysical structure, place it at some celestial coordinate and sweep across it using the telescope's pointing information. The signal from the simulation is then subsequently added to the already existing TOD of a raw level 1 file. This step is performed in `sim2tod` as shown in the flow chart of Fig. 5.1.

The Simulation Cubes

To produce the simulation cubes used in our analysis we used the `make_cube` code that was kindly provided by Håvard Tveit Ihle. The code takes a $9^\circ \times 9^\circ$ dark matter halo

mock catalog produced by an N-body simulation and separates out some region that matches the COMAP survey patches of $4^\circ \times 4^\circ$. Subsequently, the dark matter halos in the mock catalog are populated with CO signal using a double power law model relating the halo masses to CO luminosity by Chung et al. [76]. The CO luminosities are then easily translated to units of brightness temperature by the Reighly-Jeans approximation

$$T = \frac{c^2 I_{\nu, \text{obs}}}{2k_B \nu_{\text{obs}}^2}. \quad (5.6)$$

The intensity $I_{\nu, \text{obs}} = L_{\text{CO}}/4\pi D_L^2 \Delta\nu$ observed at frequency ν_{obs} is given in terms of the cube's CO luminosity L_{CO} , the luminosity distance to the source D_L (as defined earlier in Eq. (1.13)) and the frequency resolution $\Delta\nu$ [29].

To simulate the later pickup of the signal by the telescope the cube is convolved with the $4.5'$ FWHM instrumental beam. The box dimensions will correspond to those of an actual sky map, i.e. 120×120 pixels of $2' \times 2'$ resolution by 4096 (or 256, depending on whether the high- or low-resolution grid is needed) frequency channels. An example of a frequency slice of a simulation cube in the down-sampled frequency grid can be seen in Fig. 5.2.

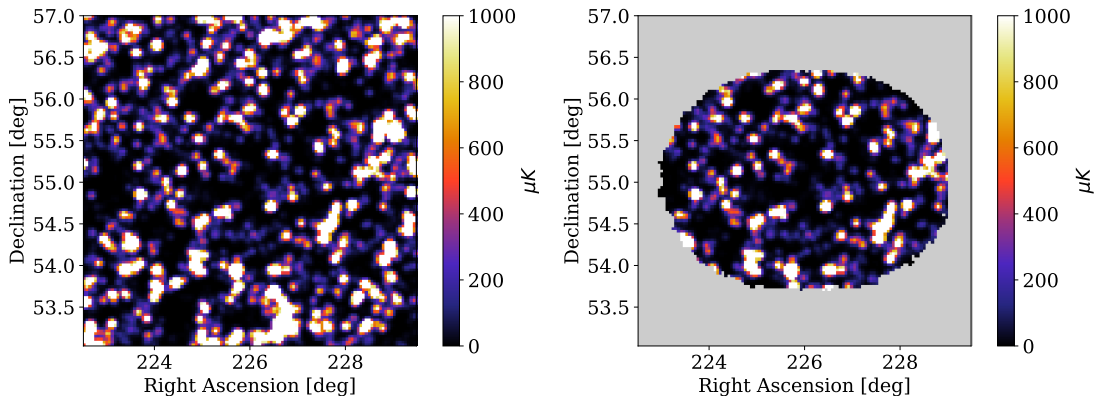


Figure 5.2: A frequency slice of a simulation cube used, at channel 30 of the second sideband at $\nu = 28.922$ GHz. The cube was normalized to have a maximum brightness of about $0.135 \mu\text{K}$ and has been centered at the coordinates of the CO6 field. The slice has 120×120 pixels. The right and left plot show the simulation cube, respectively, before and after masking the area with zero hits from the telescope pointing in `cube2map`.

These so-called semi-analytic models, combining numerical and physically motivated models, are generally most suited for intensity mapping modeling. That is, because they cover a large variety of scales without losing the physical context while also being comparatively cheap to produce [10]. In our case the model parameters are also constrained by previous experiments, making it a mix of a semi-analytic and semi-empirical model [76].

Adding Signal to the TOD

With the signal simulation in hand, we need to somehow simulate the telescope observing it at some celestial coordinates. To do this we first center the simulation cube at the celestial coordinates of one of the CO science fields.

Subsequently, we want to use the telescope’s pointing information to draw a path across the simulation cube’s perpendicular face. The brightness temperature of the cube T_{signal} at any given part of the path at some frequency ν is then used to add in additional signal to the raw TOD. The raw TOD, P_{TOD} , will then simulate the noise part of the data as it is noise-dominated and has the same noise properties as an actual observation.

The power of the TOD with added signal $P_{\text{TOD+signal}}$ is then given by the sum of Eq. (3.1) when evaluated with respectively the system temperature of the raw observation T_{sys} (estimated prior to adding signal, see Sec. 3.4.7) and the signal brightness temperature T_{signal} from Eq. (5.6). We then get

$$\begin{aligned} P_{\text{TOD+signal}}(\phi, \theta) &= P_{\text{TOD}}(\phi, \theta) + P_{\text{signal}}(\phi, \theta) \\ &= P_{\text{TOD}}(\phi, \theta) \left(1 + \frac{T_{\text{signal}}(\phi, \theta)}{T_{\text{sys}}(\phi, \theta)} \right), \end{aligned} \quad (5.7)$$

where the telescope pointing is given in terms of the right ascension (RA) and declination (Dec) coordinates (ϕ, θ) . When adding in the simulated signal, it may be amplified in strength, to make it more easily detectable.

Meanwhile, to add signal from the simulation to the TOD we must first translate the telescope’s pointing information to the pixel coordinates of the cube in order to know which cube pixel is hit by the telescope’s line-of-sight at a given time. However, because the cube used by default has a flat cartesian pixel face, we must use a flat-sky approximation when projecting the curved coordinates of the celestial sphere onto the box. For this we use a gnomonic projection in which the pixel coordinates (i, j) become

$$i = \frac{180^\circ}{\pi} \sin \phi \cot \theta \quad (5.8)$$

$$j = -\frac{180^\circ}{\pi} \cos \phi \cot \theta \quad (5.9)$$

as a function of RA ϕ and Dec θ [77].

Fortunately, we can use the WCS Python module, which was kindly provided by the Manchester group of COMAP, for the transformation. We can thus easily identify the pixel the telescope’s line-of-sight crosses at a given time and plug the corresponding signal T_{signal} into Eq. (5.7).

Because the COMAP telescope has 19 feeds, each with a slightly different pointing, we need to do the coordinate mapping separately for each feed. However, since each pixel is observed in all frequencies simultaneously, the coordinate transformation done in one frequency channel is valid for all the others. The resulting new TOD is saved to a new level 1 file for further processing, which differs from a regular obsID only in its modified TOD.

Processing the Simulated Data

Having produced the TOD with additional simulated signal inside `sim2tod`, we can process the modified TOD by `l2gen` and `tod2comap` to produce a map. As mentioned, to ensure an equal treatment of the unmodified and simulated time streams we import the frequency mask from the unmodified data into `l2gen` when processing the simulated data. The resulting map of the modified TOD will then serve as the "after" view of the pipeline filters.

5.1.4 The Lower Branch - The Unfiltered Signal Map

Finally, to make the simulation cube comparable to the maps of the raw and modified data we must first down-sample its frequency grid by averaging (as the simulation has no uncertainties) neighboring channels. Next, since we perform our analysis on pseudo-power spectra we assign the noise weights and sky cut of the map of the modified data to the simulation cube. An example of this is shown in Fig. 5.2, where the simulation is seen before and after masking the regions of the sky not observed by the telescope. That way the simulation will be affected by the same mode-mixing matrix as the simulated observation, making the pseudo-power spectra of the two as comparable as possible.

The down-sampling and weighting of the simulation cube is done in the `cube2map`-routine seen in Fig- 5.1. The resulting map of the unfiltered simulation will then serve as our "before" view of the pipeline filters.

5.1.5 Computing the Transfer Function Estimate

Now having obtained the "before" and "after" maps of the simulated data, as well as that of the noise, we can compute their power spectra as described in Ch. 4 and compare these according to Eq. (5.3) (or (5.4)) to get a transfer function estimate. This is done in the `comap2tf`-routine in Fig. 5.1.

Most importantly, the subsequent estimate of the transfer function can then be used to reverse the filter-induced bias of the signal. This is essential as we want the parameters to be estimated from our data to be unbiased. However, we can also use the developed simulation pipeline to investigate the impact of each pipeline filter on the signal, as well as how well possible improvements to the COMAP pipeline can recover additional signal at different scales.

5.2 The Destriper

As described in Sec. 3.5 on mapmaking, in the current COMAP analysis pipeline we implement a noise weighted binning with optional high pass filtering of the TOD. The underlying assumption that justifies the use of a simple binning algorithm such as this is that the processed level 2 data are fairly close to white noise.

We do, however, see that the level 2 TOD sometimes still contains some residual long-timescale modes after we have filtered it, such as residual ground pickup, standing

waves or other unknown artifacts. As these residuals are not random they will not integrate down properly. That is because the regular noise weighted binning can have trouble correctly comparing the TOD along different scan directions, as the cross-linking between the scans is poor. The improper handling of these long-timescale modes may then introduce stipes in the final map, that can obscure the large-scale structure of the astrophysical signal. Currently we handle this by highpass filtering the TOD prior to binning, removing all long-timescale modes, both that of the residuals and the actual signal. By doing so there is a significant loss in the large-scale structure of the signal. Therefore, a better-suited mapmaking algorithm is the destriper, which can distinguish between the dynamic long-timescale residuals and the static signal to produce maps without stripes.

We will in the following describe how the algorithm works, and some advantages and challenges of implementing it as part of our mapmaker in `tod2comap`.

5.2.1 The Destriper Algorithm

We begin by explicitly writing out the noise in terms of its uncorrelated and correlated components, so that the time stream of data becomes

$$\mathbf{d} = \mathbf{P}\mathbf{m} + \mathbf{n}_{\text{corr}} + \mathbf{n}_{\text{w}}, \quad (5.10)$$

where the white noise \mathbf{n}_{w} is simple Gaussian thermal noise [78]. Even though the term \mathbf{n}_{corr} traditionally refers to correlated noise, long-timescale residuals in the time stream can be modelled the same way. We thus stick to the traditional terminology hereafter, even though we are strictly speaking not fitting the correlated noise but residual long-timescale systematics. The residuals \mathbf{n}_{corr} can be modelled by a set of basis functions that are contained as columns in a matrix \mathbf{F} with corresponding amplitudes contained in a vector \mathbf{a} . The contribution from the long-timescale modes can then be written as [78]

$$\mathbf{n}_{\text{corr}} = \mathbf{F}\mathbf{a} \quad (5.11)$$

In the simplest case the basis functions used to model the slowly varying modes will be simple flat baselines. The $n_{\text{samp}} \times n_{\text{basis}}$ template matrix then will become something like

$$\mathbf{F} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \cdots & \vdots \\ 0 & 1 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \cdots & \vdots \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 1 \end{bmatrix}. \quad (5.12)$$

In this case each of the n_{basis} columns has n_1 subsequent elements that are equal to 1, while the rest of the n_{samp} elements are set to zero. Hence all elements equal to one

together represent a baseline. The number of ones n_1 times the temporal resolution Δt will then equal the baseline's length in time. In general, all baselines should have the same length. However, we do not want the baseline fit of several level 2 TODs to overlap. Therefore, if a given level 2 TOD is slightly too long or short to fit an integer number of baselines to it, the excess may be assigned a shorter baseline length n_1 .

In fact, this method is quite powerful because we can also fit more complicated basis functions or templates to the TOD, potentially even replacing or complementing some of the filters in `l2gen`. We could for instance fit for the pointing template of the data as well as ground pickup, to remove the corresponding correlated artifacts. In doing so we would add additional columns to \mathbf{F} that are for instance filled with values of the elevation template $1/\sin(\text{el}(t_i))$ at each time step t_i . However, we will in our analysis only consider the simple flat baseline case because we want to see how well we can recover large-scale modes previously removed by highpass filtering of the TOD.

Again using the assumption of gaussianity we can write the log-likelihood as

$$\mathcal{L} \propto (\mathbf{d} - \mathbf{F}\mathbf{a} - \mathbf{P}\mathbf{m})^T \mathbf{N}_w^{-1} (\mathbf{d} - \mathbf{F}\mathbf{a} - \mathbf{P}\mathbf{m}) + \mathbf{a}^T \mathbf{N}_a^{-1} \mathbf{a}, \quad (5.13)$$

instead of Eq. (3.30) [78]. The total covariance matrix is now $\mathbf{N} = \mathbf{F}\mathbf{N}_a\mathbf{F}^T + \mathbf{N}_w$, respectively made up of the correlated and white noise covariances matrices $\mathbf{N}_{\text{corr}} = \mathbf{F}\mathbf{N}_a\mathbf{F}^T$ and \mathbf{N}_w . The covariance of the baselines is given in a similar fashion to that of the white noise; $\mathbf{N}_a = \langle \mathbf{a}\mathbf{a}^T \rangle$ [68].

Even though there might be some long-timescale artifacts surviving the pipeline filters we can still assume their contribution to be fairly small. Hence we can fairly safely neglect the prior term $\mathbf{a}^T \mathbf{N}_a^{-1} \mathbf{a}$ of Eq. (5.13) and let the total covariance matrix equal that of white noise.

To find $\hat{\mathbf{m}}$ we this time cannot only maximize \mathcal{L} with respect to \mathbf{m} , but need to find the best-fit basis amplitudes \mathbf{a} as well. We therefore start by maximizing \mathcal{L} with respect to the map \mathbf{m} , i.e. we find $\hat{\mathbf{m}}$ for which

$$\nabla_{\mathbf{m}} \mathcal{L} = 0. \quad (5.14)$$

This equation is easily solved to obtain the slightly modified mapmaker equation

$$\hat{\mathbf{m}} = (\mathbf{P}^T \mathbf{N}_w^{-1} \mathbf{P})^{-1} \mathbf{P}^T \mathbf{N}_w^{-1} (\mathbf{d} - \mathbf{F}\mathbf{a}). \quad (5.15)$$

Interestingly, we see that this mapmaking equation simply is a noise weighted binning of the TOD, corrected by some baselines $\mathbf{F}\mathbf{a}$, prior to binning.

To evaluate the above mapmaking equation we must find the baselines $\mathbf{F}\mathbf{a}$ that best fit the TOD. To find this we first substitute the map estimate (5.15) back into \mathcal{L} to obtain

$$\mathcal{L} \propto (\mathbf{d} - \mathbf{F}\mathbf{a})^T \mathbf{Z}^T \mathbf{N}_w^{-1} \mathbf{Z} (\mathbf{d} - \mathbf{F}\mathbf{a}), \quad (5.16)$$

where the matrix $\mathbf{Z} = \mathbf{I} - \mathbf{P}(\mathbf{P}^T \mathbf{N}_w^{-1} \mathbf{P})^{-1} \mathbf{P}^T \mathbf{N}_w^{-1}$, for the identity matrix \mathbf{I} . The effect of \mathbf{Z} on a time domain vector is to subtract the noise weighted average of all samples that hit each pixel from that time domain vector [78]. By doing so it separates the

parts of the TOD that correspond to stationary signals on the sky from the dynamic long-timescale systematic residuals.

Subsequently, the new \mathcal{L} is maximized with respect to \mathbf{a} so that

$$\nabla_{\mathbf{a}}\mathcal{L} = 0. \quad (5.17)$$

This happens when the equation

$$\mathbf{F}^T\mathbf{N}_w^{-1}\mathbf{Z}\mathbf{F}\mathbf{a} = \mathbf{F}^T\mathbf{N}_w^{-1}\mathbf{Z}\mathbf{d}, \quad (5.18)$$

is fulfilled [78]. Solving this equation for \mathbf{a} then yields the best-fit baseline amplitudes which can be inserted into the destriper equation (5.15) to solve for the unbiased ML map estimate $\hat{\mathbf{m}}$.

5.2.2 The Conjugate Gradient Method

The combined matrices $\mathbf{F}^T\mathbf{N}_w^{-1}\mathbf{Z}\mathbf{F}$ and $\mathbf{F}^T\mathbf{N}_w^{-1}\mathbf{Z}$ on respectively the left- and right-hand-side of Eq. (5.18) can be very large and are in general not sparse. We therefore cannot represent them on a computer, let alone compute their inverses to solve the equation. It is therefore way more efficient to apply each constituent matrix on each side of Eq. (5.18) to respectively \mathbf{a} and \mathbf{d} individually as a linear operator. Subsequently, we can use an iterative numerical method like the conjugate gradient (CG) method to solve for the best fit baseline amplitudes \mathbf{a} . In the following, we will summarize the CG method used to solve Eq. (5.18) for the amplitudes \mathbf{a} . Unless otherwise stated the summary is based on Shewchuk's introduction to CG methods [79].

The idea behind the method of conjugate gradients is that we want to minimize a function on the form

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} - \mathbf{b}^T\mathbf{x} + \mathbf{c}, \quad (5.19)$$

by moving a given number of steps $\{\mathbf{x}_k\}$ along a set of A -orthogonal directions $\{\mathbf{p}_k\}$ until we have found the function's minimum. The index $k = 1, 2 \dots n$, where n is the dimension of the parameter space in which we want to minimize f . Similar to regular orthogonality, two vectors \mathbf{v}_i and \mathbf{v}_j are said to be \mathbf{A} -orthogonal, or conjugate, if $\mathbf{v}_i^T\mathbf{A}\mathbf{v}_j = \delta_{ij}$. Intuitively we can think about it as the orthogonality criterion extended to a cartesian space that is "warped" by some matrix \mathbf{A} . That way we only need to move along each linearly independent direction one time, by choosing the next step so that f is minimized for each new step [80]. Compared to the gradient method of steepest descent, which often moves along the same direction several times, the CG descent can thus more efficiently converge.

Going further, it is easy to show that Eq. (5.19) is minimized when

$$\nabla f(x) = \mathbf{A}\mathbf{x} - \mathbf{b} = 0, \quad (5.20)$$

if the matrix \mathbf{A} is positive-definite. This problem thus reduces to a $N \times N$ linear system on the same form as the one in Eq. (5.18) that we want to solve to get the best-fit baseline amplitudes for the destriper.

Next, we want to move from the current position \mathbf{x}_k along the search direction \mathbf{p}_k to get closer to the minimum of f . We can write the next step after our move as $\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k$, for some step size α_k along the vector \mathbf{p}_k .

In order to find the optimal step size α_k we must first define the error $\mathbf{e}_k = \mathbf{x} - \mathbf{x}_k$ and residual $\mathbf{r}_k = \mathbf{A}\mathbf{e}_k = \mathbf{b} - \mathbf{A}\mathbf{x}_k$. These are used to quantify how close the position \mathbf{x}_k is to the minimum of f at \mathbf{x} . Since we only want to move along \mathbf{p}_k a single time we can require that

$$\mathbf{p}_k^T \mathbf{A}\mathbf{e}_{k+1} = \mathbf{p}_k^T \mathbf{r}_{k+1} = 0. \quad (5.21)$$

From the definition of the residual and the next step we can write the residual of the next step as

$$\mathbf{r}_{k+1} = \mathbf{b} - \mathbf{A}\mathbf{x}_{k+1} = \mathbf{b} - \mathbf{A}(\mathbf{x}_k + \alpha_k \mathbf{p}_k) = \mathbf{r}_k - \alpha_k \mathbf{A}\mathbf{p}_k. \quad (5.22)$$

Note here that the residual decreases for each step taken, which is exactly what we want when approaching the minimum of f .

Inserting back into Eq. (5.21) we get that

$$\mathbf{p}_k^T \mathbf{r}_k - \alpha_k \mathbf{p}_k^T \mathbf{A}\mathbf{p}_k = 0, \quad (5.23)$$

from which we get that the step size is given by

$$\alpha_k = \frac{\mathbf{r}_k^T \mathbf{r}_k}{\mathbf{p}_k^T \mathbf{A}\mathbf{p}_k}. \quad (5.24)$$

However, in order to actually perform the next move in the first place we must find an expression for the step direction \mathbf{p}_k . It turns out that we can utilize the conjugate properties of search directions $\{\mathbf{p}_k\}$ for this by first choosing an arbitrary basis $\{\mathbf{q}_k\}$ and then make it \mathbf{A} -orthogonal by the Conjugate Gram-Schmidt process. We do this by first letting $\mathbf{p}_0 = \mathbf{q}_0$ and subsequently iteratively subtract all non-conjugate components of \mathbf{q}_k to give

$$\mathbf{p}_{k+1} = \mathbf{q}_k - \sum_{j=0}^{k-1} \frac{\mathbf{q}_k^T \mathbf{A}\mathbf{p}_j}{\mathbf{p}_j^T \mathbf{A}\mathbf{p}_j} \mathbf{p}_j. \quad (5.25)$$

As it turns out a convenient choice for the basis $\{\mathbf{q}_k\}$ is to let $\mathbf{q}_k = \mathbf{r}_k$. That is because the residuals are orthogonal to the previous search direction \mathbf{p}_k so that we can write the conjugate Gram-Schmidt coefficients as

$$\beta_{ij} = \frac{\mathbf{q}_i^T \mathbf{A}\mathbf{p}_j}{\mathbf{p}_j^T \mathbf{A}\mathbf{p}_j} = \begin{cases} \frac{\mathbf{r}_i^T \mathbf{r}_i}{\mathbf{r}_{i-1}^T \mathbf{r}_{i-1}}, & \text{for } i = j + 1 \\ 0, & \text{for } i > j + 1. \end{cases} \quad (5.26)$$

Due to this simple trick we can in fact always find the next search vector \mathbf{p}_k by knowing only the previous one. Thus we only need to keep track of two search vectors at the same time. This reduces the order of computations per iteration from $\mathcal{O}(n^3)$ to the order of non-zero elements in the matrix \mathbf{A} . The next search direction will then simply be given as

$$\mathbf{p}_{k+1} = \mathbf{r}_k + \beta_{k+1,k+1} \mathbf{p}_k. \quad (5.27)$$

Connecting the loose threads, we can summarize the algorithm by the following steps

1. Chose some initial position \mathbf{x}_0 and compute the corresponding $\mathbf{p}_0 = \mathbf{r}_0$.
2. Loop over $k = 1; k = n$

$$\alpha_k = \frac{\mathbf{r}_k^T \mathbf{r}_k}{\mathbf{p}_k^T \mathbf{A} \mathbf{p}_k} \quad (\text{by inserting Eq. (5.27) into Eq. (5.24).)} \quad (5.28)$$

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \alpha_k \mathbf{p}_k \quad (5.29)$$

$$\mathbf{r}_{k+1} = \mathbf{r}_k - \alpha_k \mathbf{A} \mathbf{p}_k \quad (5.30)$$

$$\beta_{k+1,k+1} = \frac{\mathbf{r}_{i+1}^T \mathbf{r}_{i+1}}{\mathbf{r}_i^T \mathbf{r}_i} \quad (5.31)$$

$$\mathbf{p}_{k+1} = \mathbf{r}_{k+1} + \beta_{k+1,k+1} \mathbf{p}_k \quad (5.32)$$

3. Ideally speaking the CG method should converge at the minimum of f after at most n steps, but due to round-off errors etc. on a computer this might not be the case. In that case we can stop the loop if the residual \mathbf{r}_{k+1} is smaller or equal to some chosen tolerance, or if we have performed some chosen maximum number of iterations.

5.2.3 Implementation of the Destriper

When implementing the destriper we compute the baseline fit for the TOD per frequency. Because computing the baselines at a given frequency is independent, we can fit baselines to multiple frequencies in parallel at the same time. However, there are two main approaches on how to fit baselines to the COMAP data at a given frequency, both of which we will consider. As the COMAP telescope has 19 feeds we could sequentially row up all the TOD's from each feed into a large times stream and subsequently fit the baselines to it. We will denote this approach as the all-feeds approach. In theory, as we fit the baselines to all feeds simultaneously we should have better cross-linking. However, destripping a larger amount of data at the same time also requires more memory. For COMAP, if we treat all the data per CO field at the same time we can exceed the memory capabilities of the computers at hand.

Another approach would be to fit the baselines to the data contained in each data split (described in 4.2.1) separately. Thus, the TOD to which baselines are fit will be a lot shorter, decreasing the overall cross-linking somewhat. The advantage of this per-split destriper, however, is that we destripe the data within each split separately, keeping the splits independent from each other. That way, we can make sure that the systematic effects from one split do not leak over to another through some common baseline fit. Additionally, fitting baselines to smaller batches of the data also requires much less memory, which makes it easier to process the whole data from a given CO field.

As we saw in Sec. 5.2.1, the process of fitting the baselines and projecting the time stream into map pixels are independent of each other. We can thus use the same noise weighted binning algorithm previously used in `tod2comap` to perform the

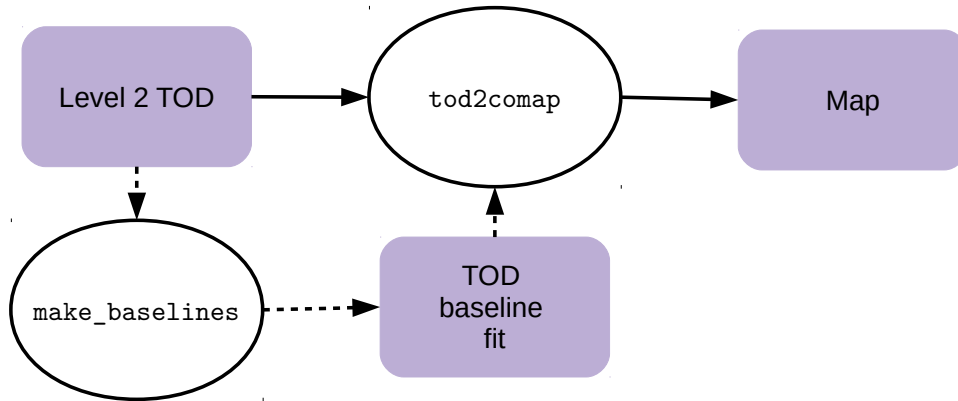


Figure 5.3: Flow chart of the improved mapmaker. The dotted lines show optional steps of producing a baseline fit to the TOD using `make_baselines`, which can be performed between `l2gen` and `tod2comap`.

projection of the TOD. Meanwhile, the baseline fitting is performed in the Python script `make_baselines` that can be used as an optional step between `l2gen` and `tod2comap` as illustrated in Fig. 5.3. The baseline amplitudes \mathbf{a} and the number of time samples per baseline are then saved to file by `make_baselines`, and can be imported to `tod2comap` in order to perform the noise weighted binning of the corrected TOD $\mathbf{d} - \mathbf{F}\mathbf{a}$.

Note also that in some cases the CG solver can have trouble converging, especially if it has to handle a lot of correlations in the TOD. In that case we can use a precondition matrix $\tilde{\mathbf{A}}$, for which $\tilde{\mathbf{A}}^{-1}\mathbf{A} \approx \mathbf{I}$, where \mathbf{A} represents the matrix on the left-hand-side of Eq. 5.18, in order to speed up convergence. Finding a suited preconditioner can be somewhat of a challenge, but a common point to start is to use the inverse of the diagonal of \mathbf{A} [80]. Thus, when testing the destriper we can compare whether using a preconditioner brings any run-time advantages.

Finally, we need to find the optimal baseline lengths, which should produce clean maps without new additional structure and at the same time resolve additional large-scale structures of the signal. In the default pipeline, we perform TOD highpass cuts at 0.02 Hz. Thus a good starting point should be baselines of length ~ 50 s.

5.3 Principal Component Analysis of Feed-Feed Pseudo Cross-Spectra

As we discussed in Sec. 4.2 we can compute cross-spectra between different parts of the COMAP data and thus eliminate systematic effects that are not common among the split halves. Nevertheless, when looking at individual FPXS of our data we see that the spectra still contain some surviving systematic effects, as the FPXS are not completely consistent with white noise. Although the sources of several systematic effects are known to some extent, e.g. the standing wave artifacts seen in the middle

panel of Fig. 5.4, there are certainly still a significant number of systematic artifacts seen in spectra that have a yet unknown origin. Examples of such artifacts are seen in the left and right panels of Fig. 5.4.

If spectra with such artifacts are not discarded or cleaned in some way, the averaged FPXS can be contaminated. Therefore we will in the following consider a principal component analysis (PCA) of the FPXS dataset found in Step 3. presented in Sec. 4.2.2. The subsequent PCA will then potentially aid in identify and clean bad spectra, before the average FPXS is computed, and could in the future be a valuable data selection and filtering tool.

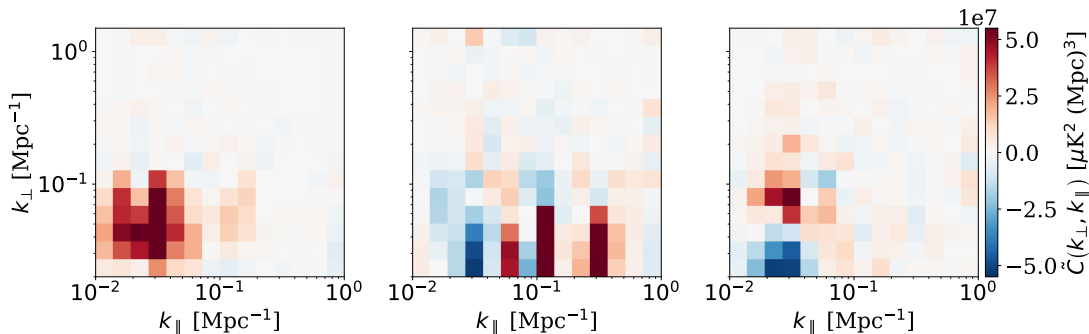


Figure 5.4: Example of individual 2D FPXS which contain significant systematic artifacts; (**left panel**) a large excess structures at large scales, (**middle panel**) bar-like features at constant k_{\parallel} caused by standing waves, as well as (**right panel**) a dipole-shaped structure at low k .

5.3.1 The PCA Decomposition

The idea behind doing a PCA decomposition of the FPXS dataset is that it consists of cross-spectra from multiple different combinations of data, each with different contributions from systematic effects. Thus some of the cross-spectra will be affected by a given systematic effect and others will not. The corresponding variation in the cross-spectrum dataset will hence introduce a given amount of spread, i.e. variance in the data. Meanwhile, the astrophysical signal which should be common to all maps will not introduce any variance in the data, up to some sample variance. Luckily, this is exactly what a PCA is made for; to separate out the components of a dataset that describe most of the variance, while leaving the mean (i.e. the common signal) alone.

The goal of the PCA of the cross-spectra is thus twofold; 1) identification of systematic effects and 2) possibility cleaning spectra from these. One approach we will consider to the former is to look at the multidimensional dataset in the low dimensional space spanned by the first principal components that explain most of the variance in the data. Doing so could uncover interesting patterns in which the data are clustered or arranged in low dimensional space, which in turn could help identify spectra with systematic artifacts. In the second point, we utilize that the common astrophysical

signal should not be picked up by the PCA. Hence if a principal component of the data likely explains some systematic effect, it can in theory be removed from contaminated spectra without too much loss of signal.

To make it clearer what is meant by the above let us consider an FPXS \tilde{C}_i from our dataset. Each spectrum consists of 14×14 bins, $\mathbf{k}_i = (k_{\parallel}^i, k_{\perp}^i)$, which we can simply treat as a 196 pixel image. All of these images are subsequently mean centered;

$$\tilde{C}_i^m = \tilde{C}_i - \frac{1}{n} \sum_{j=1}^n \tilde{C}_j, \quad (5.33)$$

where we consider an FPXS dataset of n (usually 3000-8000) spectra and denote any mean centered spectrum by a superscript "m". Mean centering the dataset is important when doing a PCA via the SVD method, as shown here, or else the first principal component may reflect the mean of the dataset [81], which we want to leave out as it contains the signal. Each mean centered image is then flattened into a 1D vector of 196 entries and collected in the data matrix

$$\mathbf{D} = [\tilde{\mathbf{C}}_1^m \quad \tilde{\mathbf{C}}_2^m \quad \tilde{\mathbf{C}}_3^m \quad \cdots \quad \tilde{\mathbf{C}}_n^m]. \quad (5.34)$$

The covariance matrix $\mathbf{N} = \mathbf{D}^T \mathbf{D}$ of the data will then contain all the information needed to separate out all the directions, and the corresponding amount, of most spread. In particular, we need to find the eigenvectors and corresponding eigenvalues of the covariance matrix \mathbf{N} . This can be done by the singular value decomposition (SVD) of the data,

$$\mathbf{D} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T, \quad (5.35)$$

where \mathbf{U} and \mathbf{V} are orthogonal matrices, while the diagonal elements of $\mathbf{\Sigma}$ are the singular values of the data σ_i . Now, the columns of \mathbf{V} and the square of the singular values in $\mathbf{\Sigma}$ correspond, respectively, to the eigenvectors and eigenvalues of the covariance matrix \mathbf{N} . The eigenvectors \mathbf{v}_i can then be sorted according to size of their corresponding eigenvalues $\lambda_i = \sigma_i^2$. Thus any eigenvector (or eigen-spectrum), and its corresponding eigenvalue, give the direction and amount of variance along the given direction in k -bin space [82].

Next, consider a given cross-spectrum from the dataset, $\tilde{\mathbf{C}}_j$. To quantify the contribution to $\tilde{\mathbf{C}}_j$ around the mean from a given PCA eigen-spectrum \mathbf{v}_i we can simply project the mean-centered cross-spectrum onto the eigen-spectrum by performing the dot product

$$a_{ij} = \mathbf{v}_i \cdot \tilde{\mathbf{C}}_j^m = \sum_{l=1}^{196} v_i(\mathbf{k}_l) \tilde{C}_j^m(\mathbf{k}_l). \quad (5.36)$$

This PCA amplitude a_{ij} can therefore be used to quantify the amount of contamination of $\tilde{\mathbf{C}}_j$ from the systematic effect described by \mathbf{v}_i . Therefore we may use a_{ij} to classify the cross-spectra according to their systematic effect, given that the PCA eigen-spectra indeed describes the latter. Furthermore, if we find that the first two or so eigen-spectra of the data explain a majority of the total variance in the data, it may indicate that

these eigen-spectra represent some dominant systematic effects. We can thus compute the mean-centered PCA amplitudes of each FPXS in our data, for each of the two primary eigen-spectra. Subsequently, using the two amplitudes for each FPXS as x and y coordinates in a 2D scatter, we can project the whole 196D FPXS dataset into the plane spanned by the two primary eigen-spectra. After this dimensionality reduction, the dataset becomes much easier to interpret, at the cost of only a relatively modest loss of information. Subsequently, we can look at the outliers and clustering patterns observed in the 2D scatter, thus potentially revealing some interesting patterns that hint to which splits or feeds are affected the most by the systematic effects described by the two primary (or higher) eigen-spectra.

5.3.2 Cleaning FPXS of Residual Systematics

We can also use the mean-centered PCA amplitude and the corresponding eigen-spectrum to possibly remove the contamination of the systematic effect from a cross-spectrum, not dissimilar to how we can remove bad weather modes from the TOD in the PCA filter (see Sec. 3.4.5). To clean a given contaminated FPXS $\tilde{\mathbf{C}}_j^{\text{contamin}}$ from its systematic excess we simply subtract out the PCA mode $a_{i,j}\mathbf{v}_i$ that describes the systematic effect;

$$\tilde{\mathbf{C}}_j^{\text{clean}} = \tilde{\mathbf{C}}_j^{\text{contamin}} - a_{i,j}\mathbf{v}_i. \quad (5.37)$$

Because we would rather not remove any signal from the cross-spectra, we need to check how much the signal is affected by the subtraction in Eq. (5.37). In fact, doing so is essentially the same as finding an effective transfer function, which quantifies the combined signal loss due to the filtering of the data and the PC subtraction of the FPXS. Luckily, for this we can use our newly developed simulation pipeline, i.e. produce a dataset with a known signal and compare it to the filtered and PC subtracted equivalent. The resulting effective transfer function is then found in a similar fashion as seen in Eq. (5.4), only this time the $C_{\mathbf{k}}$ represents the FPXS after PC subtraction. Finally, we can estimate the loss of signal due to the eigen-spectrum subtraction by comparing the effective transfer function of cleaned the FPXS to the transfer function of the un-subtracted FPXS.

Part III

Results

Chapter 6

Results

In the following chapter, we will present the main results of the improvements of the COMAP pipeline as previously described in Ch. 5. We will present the estimates of the pipeline transfer function produced by our simulation pipeline as well as how different filtering of the raw TOD biases our signal estimates. Subsequently, we will present the results of the mapmaking with a destriper and how it improves the recovery of large-scale modes in the final maps. Finally, we will scratch the surface on how well we can potentially classify and remove artifacts from systematic errors from FPXS using the PCA approach.

6.1 The Simulation Pipeline

We begin our analysis of the simulation pipeline results by looking at some maps produced by its three branches. Subsequently, we will show the power spectra and transfer function we can obtain from these, before investigating the ensemble-averaged transfer function used to reverse the underestimation of the signal we measure in COMAP. Lastly, we will consider how the different filters in `12gen` and different scanning strategies attenuate the signal, and thus quantify the resulting bias.

6.1.1 Simulation Pipeline Maps

In Fig. 6.1 we see an example of a 26.92 GHz slice of the three maps that are produced by the three branches of the simulation pipeline. These were produced using a single realization of the signal and 63 obsIDs worth of data on the CO6 field. The artificial signal added to the telescope time stream in `sim2tod` was in this case boosted by a factor of 1000, as we found this to ultimately yield the cleanest looking transfer functions. The maximum temperature of the cube after boosting was thus about 2 K.

The map of the raw simulation seen in Fig. 6.1a is an illustration of how the simulated signal looks before it is affected by the pipeline filtering. Meanwhile, the map made from a TOD with added signal can be seen in Fig. 6.1b. We can clearly recognize some of the simulated structures in this map that are also present in the pure simulation

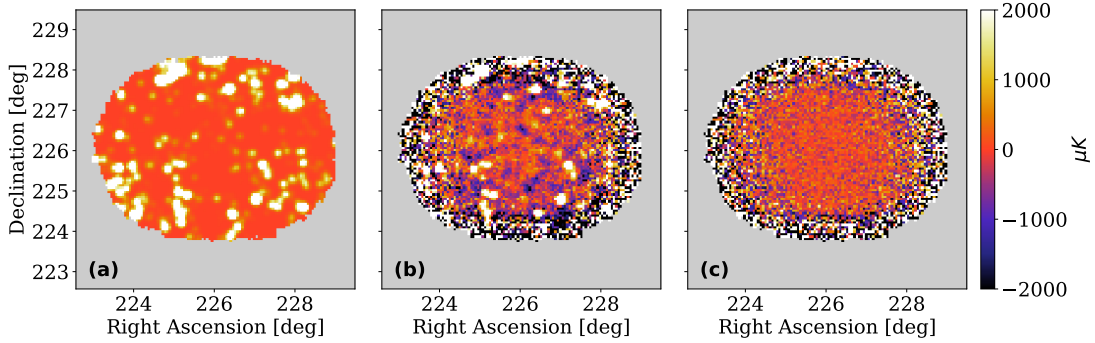


Figure 6.1: Three 26.92 GHz (i.e. A:LSB channel 30) maps from each of the three branches of the simulation pipeline; (a) the raw simulation, (b) the map made from a TOD of 63 CO6 obsIDs with added signal as well as (c) the map made from the corresponding TOD without additional signal. Each map is coadded over the feeds. The signal cube was boosted in temperature by a factor of 1000 to make it detectable, resulting in a max temperature of about 2 K.

map. In addition, it has similar noise properties as the map in Fig. 6.1c made from only the raw TOD, as the noise level is the highest at the edges of the map where the telescope sweeps by less frequently.

6.1.2 From Auto-Power Spectra to Transfer Function

From each of the maps seen in Fig. 6.1 we computed pseudo-auto-spectra. These are seen in Fig. 6.2. In the upper left panel of the figure, we see the 2D binned pseudo-auto-spectrum of the unfiltered simulation from Fig. 6.1a. As we can see the simulation has the largest contribution to its power spectrum by the largest scales, gradually fading off in both the parallel and perpendicular directions. The reason why the power spectrum is not completely isotropic in k in this plot is due to the beam smoothing, which causes the power spectrum not to fade off to $\sim 4 \mu\text{K}^2 \text{Mpc}^3$ the same way in the perpendicular direction as in the parallel one.

Comparing the pure-signal spectrum to the pseudo-auto-spectrum resulting from the processed TOD with signal, seen in the upper right of the same figure, we can see that there are some changes from that of the pure simulation. Firstly, we see that the processing of the signal has led to a general reduction in power. This effect is most significantly seen on small k_{\parallel} and k_{\perp} and is caused by the removal of the largest modes in the pipeline. Thus we would underestimate the signal, most notably on large scales, if we do not take the filtering into account. If we look at the high k_{\perp} we can, however, see that the power spectrum shows a higher value than that of the raw simulation. This effect is caused by the fact that the map made from a TOD with added signal is dominated by the TOD's noise properties on small scales (i.e. high k), as opposed to the pure simulation which is noise-free. Therefore the power spectrum of the simulated observation approaches the power spectrum of the map made from a raw TOD (i.e. the

noise-spectrum) seen in the lower left of Fig. 6.2. Now, the noise-spectrum approaches

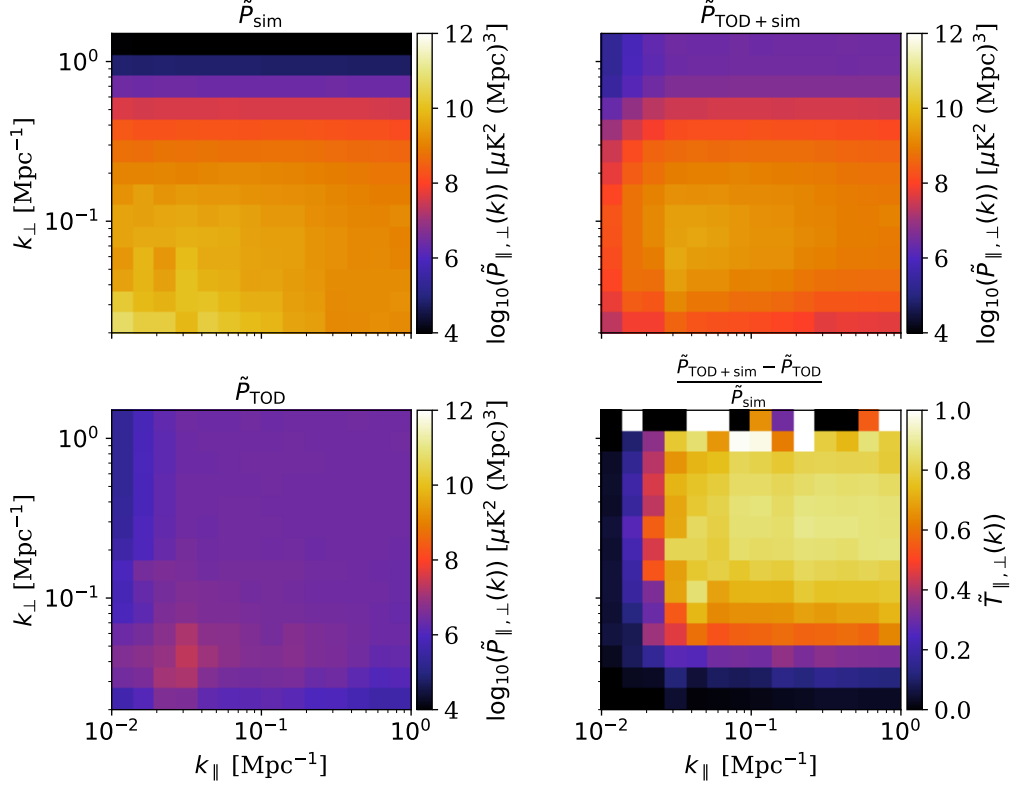


Figure 6.2: An example of the pseudo-auto-spectra from the (**top left**) raw simulation cube, noise weighted in the same fashion as the map of the signal with noise. The corresponding (**top right**) pseudo-auto-spectrum of the map resulting from adding signal to the observed TOD. The (**bottom left**) pseudo-auto-spectrum of the observed TOD (i.e. the noise-only spectrum) is subsequently subtracted from the spectrum of the TOD with signal, and divided by the pure signal spectrum, to yield an estimate of the (**bottom right**) pipeline pseudo transfer function. The spectra which contain the TOD were all produced by 63 obsIDs of CO6, the same as used for the maps in Fig. 6.1.

a constant power of $\sim 10^7 \mu\text{K}^2\text{Mpc}^3$ at most scales, which corresponds to the white noise power. The exception being the excess structures seen at small k , which could be caused by systematic errors not yet handled. We also see that the noise spectrum has a lower value than that of the white noise at small k_{\parallel} and k_{\perp} due to the pipeline filters removing the largest modes in the TOD. Thus we should, if being completely rigorous, mitigate the pipeline's effect on even the white noise. However, as the effect on the noise-spectrum seems to not be dramatically large, we simply neglect the corresponding white noise transfer function explaining this effect. However, it could be worth checking

the white noise transfer function in the future when we close in on a detection of CO signal and finding all systematic effects becomes much more important.

Combining all the power spectra seen in Fig. 6.2 according to Eq. (5.3) we get the transfer efficiency seen in the lower right of Fig. 6.2. This plot should hence quantify approximately the fraction of signal that survives the combined filtering of the pipeline on different scales. We see that the peak efficiency at about $k_{\parallel}, k_{\perp} \geq 0.1 \text{ Mpc}^{-1}$ is about 85 – 90% transfer of signal. However, at small k the transfer function approaches zero, due to the pipeline filtering out the largest modes of the simulation.

Another noteworthy point to make about the transfer function in Fig. 6.2 is its behavior at high k_{\perp} . As we can see, this region looks considerably noisier than all the others, even showing some k bins that are over-saturated to above unity. Considering this happens on scales in which the map is noise dominated, it may be that our noise approximation (i.e. the TOD without any signal) breaks down. In investigating this effect we for instance found that a stronger signal than the one used would worsen this issue. Hence it could also be a small round-off error in the computations that gets worse when the signal is too strong compared to the noise level. As we will see in the remainder of this section, this is a problem that seems to recur in most (if not all) estimated 2D transfer functions. Thus as our noise approximation in this region is probably somewhat off, the corresponding transfer function values should not be taken too seriously.

As mentioned, the transfer function seen in Fig. 6.2 was computed using a single realization of the simulation cube, added to 63 obsIDs of CO6. However, as discussed in Sec. 5.1.1 the true transfer function is found by ensemble averaging over all transfer function estimates from several realizations to account for sample variance in the simulations. Because estimating these transfer functions for 63 obsIDs (or more) is expensive in terms of run time and needed storage space, we only computed the transfer function from three realizations to compute an ensemble average. Each of the simulations used was boosted in temperature the same way, i.e. by a factor of 1000, to make the signals more easily detectable. Although this is far from a perfect estimate of a transfer function it will even out some of the sample variance and help quantify the approximate variation of the transfer functions from different realizations of the signal.

The 2D ensemble-averaged transfer function, as well as its k -averaged 1D equivalent, can be seen in Fig. 6.3. The individual transfer functions from which we computed the average can be seen in the upper panels (a - c) of Fig. 6.4 as well as the differences to the mean transfer function in the lower panels (d - f). The first thing to note in the ensemble-averaged transfer function is how much less "noisy" it seems overall compared to the individual estimates. Note also that the ensemble-averaged transfer function looks considerably cleaner at the noise-dominated high k_{\perp} region. This suggests that computing the ensemble-averaged transfer function in the limit of even more realizations could largely benefit the transfer function in this region and counteract the break-down of the noise-spectrum approximation.

However, when considering the ensemble-averaged 1D transfer function we see that the over-saturated bins of the 2D transfer function are effectively reduced to a locally

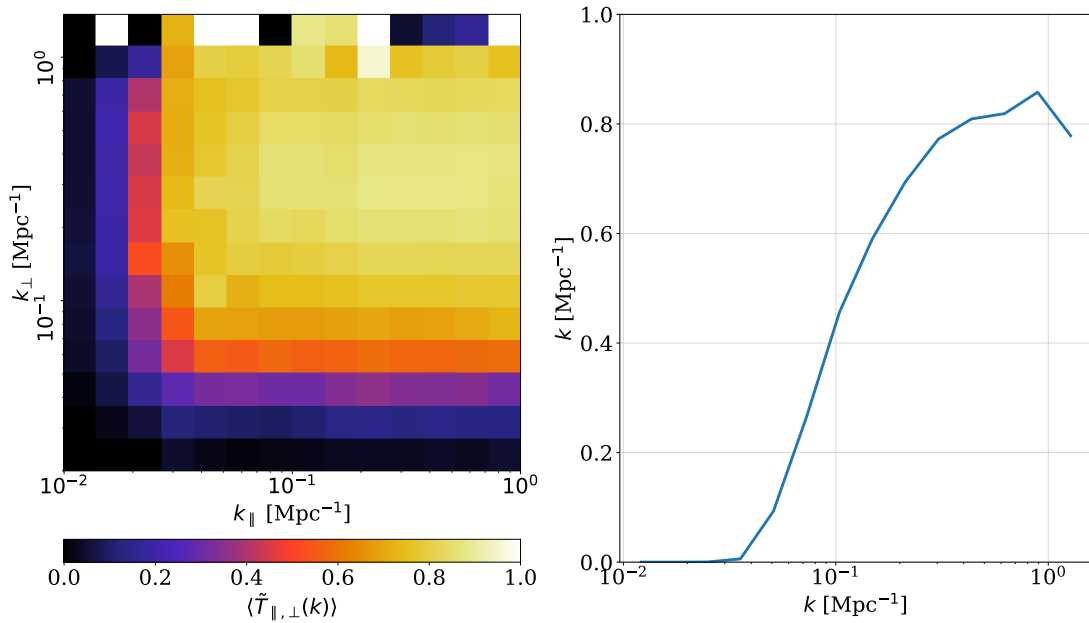


Figure 6.3: The ensemble-averaged 2D and 1D transfer function from three different realizations of the signal cube, each of which were added to our dataset of 63 obsIDs of CO6.

enhanced peak. Because we are averaging over shells of constant $k = \sqrt{k_{\parallel}^2 + k_{\perp}^2}$ when computing the 1D transfer function the overall transfer function has a lower peak than the 2D equivalent. Thus the problems of the noise-dominated region get considerably less notable.

Furthermore, the fluctuations in the transfer function estimates at low k_{\parallel} are nicely evened out in the ensemble average. We can see this by looking at the difference plots in Fig. 6.4. These show both negatively and positively varying bins of at most $\leq \pm 0.2$ from the mean. Hence the individual realizations of the transfer function are not biased compared to the average one. This suggests that the fluctuating bins are indeed the result of sample variance in the simulation at large scales and that these will average out nicely in a large number of realizations.

Even though our ensemble-averaged transfer function is far from perfect, when considering the relatively modest amount of data that went into it, the averaged transfer function in Fig. 6.3 is the so far best estimate of the transfer efficiency of the COMAP pipeline filters. Having a transfer efficiency of $\sim 85 - 90\%$ at $k_{\parallel} \geq 0.1 \text{ Mpc}^{-1}$ and $k_{\perp} \geq 0.2 \text{ Mpc}^{-1}$ and almost zero at low k_{\parallel} and k_{\perp} , we have the greatest accuracy at intermediate to small scales while significantly underestimating the signal at large scales. Nevertheless, we can now correct for the filter-induced bias of our signal estimates, as well as the underestimated error bars, by applying the inverse of the transfer function to the signal spectrum estimate. By doing so we can prevent the filter's attenuation to bias our ultimate astrophysical parameter estimates.

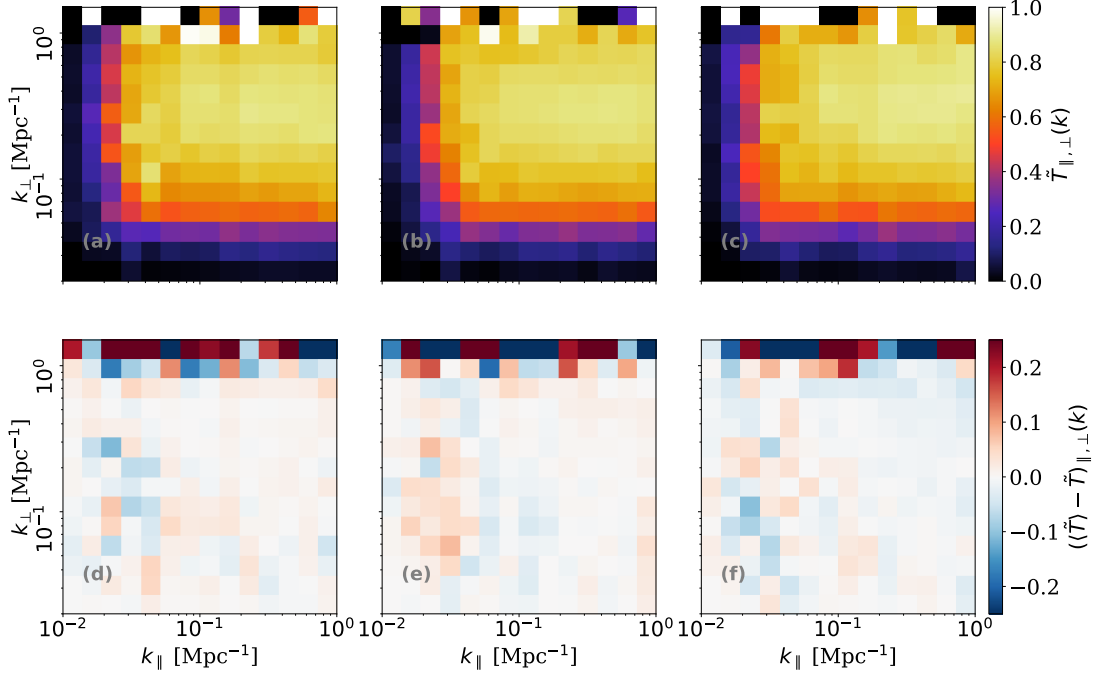


Figure 6.4: The transfer function estimate of each of the three different realizations (**a** - **c**) used to compute the ensemble-averaged transfer function in Fig. 6.3, as well as the corresponding difference to the mean (**d** - **f**). Note how the transfer function estimate of each realization and the ensemble-averaged one differ by at most about 10-20% at low k_{\parallel} .

Some natural improvements to our average transfer function estimate would as discussed, be to compute a larger ensemble of realizations. This could be done by sampling at random from both obsID- and simulation-space, to take variations in both the obsIDs and simulations into account. In addition, the fact that we used pseudo-spectra to compute the transfer functions will to some extent distort them, and thus undoing this effect could provide a more unbiased transfer function.

6.1.3 From Feed-Feed Pseudo-Cross Spectra to the Transfer Function

Similarly to computing the transfer function from pseudo-auto-spectra we can, as discussed in Sec. 5.1.1, compute the transfer function from feed-feed pseudo-cross spectra (FPXS). An example of this can be seen in Fig. 6.5. The averaged FPXS of the TOD with and without added signal, respectively seen in the upper right and lower left, were computed according to the methods in Sec. 4.2.2 using the 63 obsIDs of CO6 previously also used. In particular, the cross-spectra are computed by cross-correlating the high and low elevation splits of the Lissajous-only data. Hence this is only one of the possible averaged FPXS resulting from the splits of Split Tree 2. from Table 6.1. The signal pseudo-auto-spectrum seen in the upper left of Fig. 6.5, is the same in Fig. 6.2, the

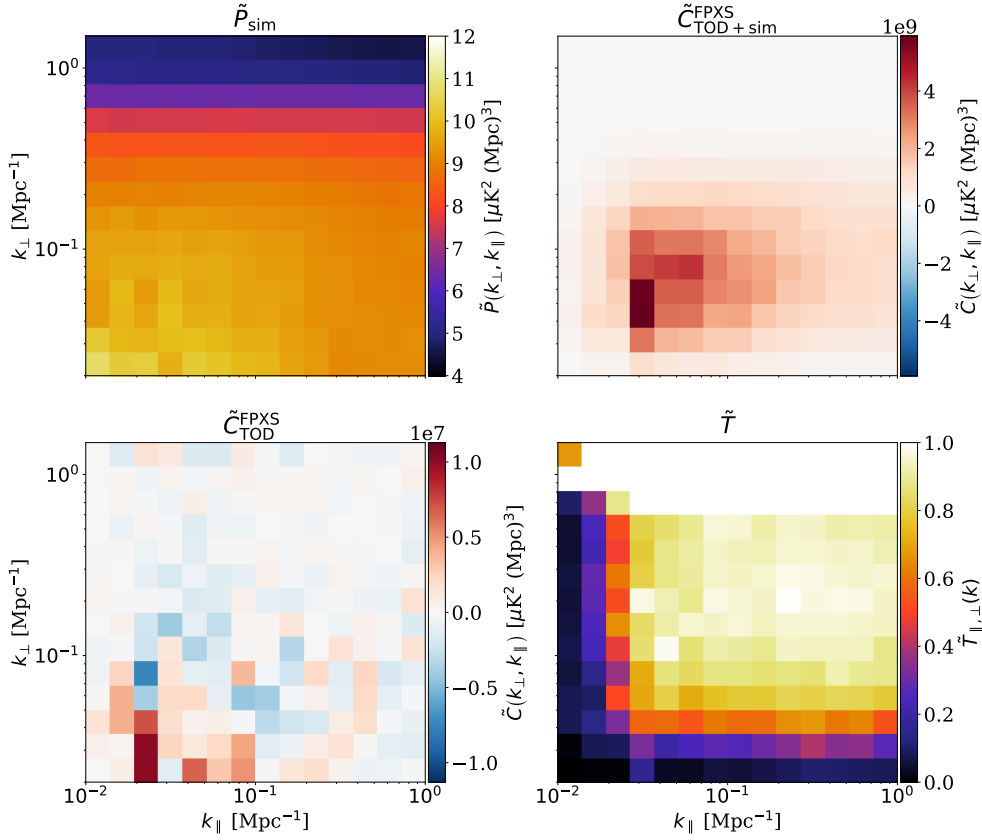


Figure 6.5: The (**lower right**) pipeline transfer function derived from a (**upper right**) FPXS made from the maps of the TOD with additional signal, (**lower left**) the FPXS of the raw time stream as well as (**upper left**) the signal pseudo-auto-spectrum. The two FPXS are averaged Lissajous cross-elevation spectra containing the same 63 obsIDs of CO6 data used to compute the spectra and transfer function in Fig. 6.2. Thus the average FPXS seen here are only one of several possible ones resulting from Split Tree 2. of Table 6.1.

only difference being that the noise weights used were different.

We see that the transfer function resulting from this analysis, seen in the lower right of Fig. 6.5, looks qualitatively equal to the one estimated from auto-spectra seen in Fig. 6.2. Both transfer functions show the same canceling of large-scale modes, as well as a roughly equally shaped peak region. Nevertheless, we can see that the cross-spectrum derived transfer function is highly over-saturated in the noise-dominated high- k_{\perp} region. This suggests that the subtracted noise approximation or the noise weighting applied in computing the signal pseudo-auto-spectrum were for some reason not good enough. We were, however, unable to find any better functioning solutions to this problem, and will have to investigate this issue further in the future.

There are, however, also some subtleties important to note when computing transfer

functions from FPXS. In particular, as an FPXS from Eq. (4.22) is always computed from two maps its noise weights are affected by both input maps, as well as a constant of proportionality. Thus a wrongly estimated constant of proportionality, or a lack of overlap of the high-SNR regions of the two cross-correlated maps may lead to the combined noise weights $w = \mathcal{N}/\sigma_1\sigma_2$ making cross-spectra more susceptible to mode mixing than auto-spectra. Thus in theory, since different FPXS could be affected by mode coupling to a different degree, so could the derived transfer functions, potentially making it necessary to compensate for the signal bias differently for each FPXS.

Nevertheless, as the transfer function derived from an average FPXS (for now) only serves as a sanity check, we are satisfied by the qualitative agreement between the transfer function estimated from the two species of power spectra.

6.1.4 Transfer Function for Different Filters and Scanning Strategies

We have now had a thorough look at the transfer function quantifying the loss of signal power due the default filtering we apply to the raw data (see Ch. 3 for default filter parameters). However, we want to see the effect of the different filters on the signal that is received by the telescope, to determine which filters bias the signal the most and if we can potentially even use a stronger filter without attenuating the signal too much.

As we are here mostly interested in the qualitative effect of the individual filters we compute the transfer function using only six obsIDs of CO6, three Lissajous and three CES scans, for different runs of the simulation pipeline. In each run we use the same signal realization, to make the estimated transfer functions as comparable as possible, and change only a single filter setting at a time.

Before going on to present the results we note that the frequency mask applied in each run, except the one where the masking was turned off, was derived from the `12gen`-run with default parameter settings and without added signal. The reason is twofold; firstly the signal added to the TOD is so strong that the corresponding correlation patterns will be perceived as bad data resulting in almost complete masking of the data. The second reason is that when generating the mask in the first place, both the poly- and PCA filters are involved in making the mask (as described in Sec. 3.4.6), and hence changing their settings would inevitably also change the mask. By always using the same mask we therefore make sure the observed signal bias is caused by a single filter.

The 1D transfer function estimates from runs with different filtering can be seen in Fig. 6.6 along with the difference between each transfer function estimate to the one estimated from the default settings (solid black line). Correspondingly, the difference between the 2D transfer functions from each filter change to the default one, shown to spot the different treatment of scales along the parallel and perpendicular directions of the line-of-sight, can be seen in Fig. 6.7. The top-most left plot in Fig. 6.7 shows the default 2D transfer function.

Looking at Fig. 6.6 we see that the largest difference in the transfer function is observed when we turn off the polyfilter, and to a similar but lesser extent when using a $\mathcal{O}(0)$ polyfilter. In particular, we observe that most of the additional signal that

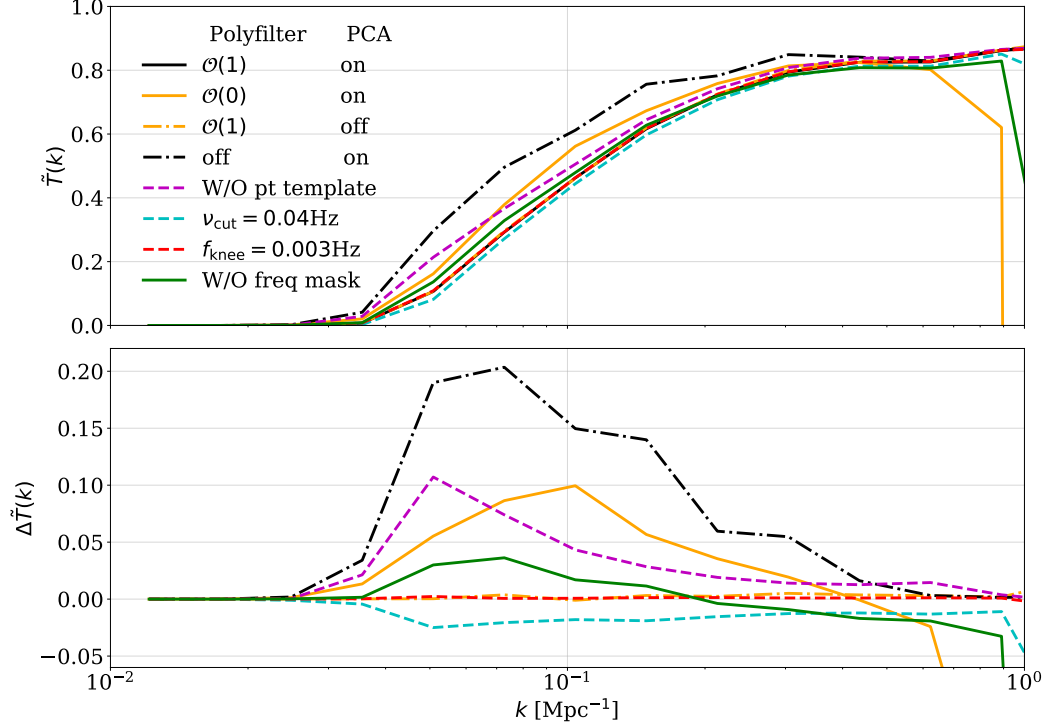


Figure 6.6: (**Upper panel**) the 1D binned transfer function from different filter combinations. The black line shows the transfer function with the current default filter settings, i.e. a polyfilter of first order, a removal of four PCA components, normalization knee frequency 0.01 Hz, with frequency masking turned on, removal of the elevation template as well as a TOD highpass cut 0.02 Hz. (**Lower panel**) the differences between the transfer function from each changed filter to the default transfer function. The color coding and line styles correspond to those of the upper panel. The transfer functions are produced using a single simulation realization as well as six obsIDs of CO6 (mixed Lissajous and CES type scans).

survives the filtering is at large to intermediate scales. However, when looking at the 2D equivalent in Fig. 6.7 (i.e. the second row from the top) we see that the changes in transfer efficiency are mainly limited to the large-scale parallel modes. This makes sense as the frequency filter's main task is to subtract the frequency dependence of each time sample, which is most effectively done by a $\mathcal{O}(1)$ polyfilter. Thus the largest modes of the signal that are correlated along the frequency dimension of the map will be subtracted out the same way as continuum foreground contamination. This is most effectively done if the subtracted polynomial at each time step is allowed to vary in both its constant and slope. Therefore when we are using no polyfilter or a polynomial of order zero, fewer of these larger line-of-sight modes are affected and can survive the

pipeline.

Next, consider the transfer function when turning off the pointing template removal. When we remove the pointing template, the modes that are correlated with the telescopes pointing are removed from the TOD. The pointing correlated modes that are canceled are therefore not only the atmospheric effects and potential ground pickup, but also the part of the signal that causes a larger pointing correlation. Thus the removed structures of the signal should predominantly be on large perpendicular scales. This is confirmed by the difference between the transfer function without and with the removal of the pointing correlations, which we can see in Fig. 6.7 shows a significant excess at constant low k_{\perp} . Similarly, in the 1D equivalent of this transfer function difference, this effect on average amounts to almost as much additional signal transmission as when using a $\mathcal{O}(0)$ polyfilter. However, the effect is primarily limited to the modes at $k \leq 0.1 \text{ Mpc}^{-1}$.

The remaining filter combinations we consider here appear to have a relatively modest effect on the transfer function. For instance, turning off the masking of frequencies completely will, as can be seen in Fig. 6.6, have a subtle effect on the transfer function at $k \leq 0.1 \text{ Mpc}^{-1}$. If neglecting the unrealistically low transfer function at the highest k bin in Fig. 6.6, where the result is somewhat unpredictable due to a poor noise approximation, the peak difference in transfer efficiency between not masking and masking is at below 5% at scales below 0.1 Mpc^{-1} , and even less at the high k modes. When looking at the corresponding 2D difference plot in Fig. 6.7 we can see that turning off the mask ever so slightly increases the transfer of signal by $\leq 10\%$ at the low k_{\parallel} scales similarly, although to a significantly lesser extent, as turning off the polyfilter. That is because masking some frequencies will decrease the amount of data along the large-scale parallel modes, resulting in a somewhat decreased transfer function.

When it comes to changing the default 0.02 Hz highpass cut frequency to 0.04 Hz we see a relatively modest effect on the transfer function of $\sim 2.5\%$ peak difference in the 1D difference plot of Fig. 6.6. However, when considering the 2D difference plot in Fig. 6.7 we see that the peak absolute difference is $\sim 5\%$ at low k_{\perp} . As a 0.04 Hz highpass cut is a somewhat harsher filter than the default 0.02 Hz it makes sense that somewhat less signal survives the pipeline. It also makes sense that the effect of the harsher filter is limited to low k_{\perp} modes as the highpass cut will affect the slowly varying temporal that correspond to large-scale structures along the perpendicular direction. However, the result is also encouraging as the effect seen in the harsher filter only has a relatively subtle effect on the bias of the signal, meaning that we could save slightly contaminated data by a less conservative highpass cut at the price of only $\sim 5\%$ peak loss in signal.

Lastly, we have two remaining filter settings to consider, both of which have negligible effects on the transfer function according to what is seen in Fig. 6.6 and 6.7. The first of these is the PCA filter, which when turned off changes the transfer function by at $\leq 5\%$ in a few bins at low k_{\parallel} , as seen in Fig. 6.7. This is indeed what is expected from the PCA filters removal of the first 4 principal components since these components correspond to the modes that are correlated the strongest across the focal plane and frequencies. Hence the structures of the signal, which are expected to be contained within a small part of the survey volume, will only correlate weakly over the focal plane

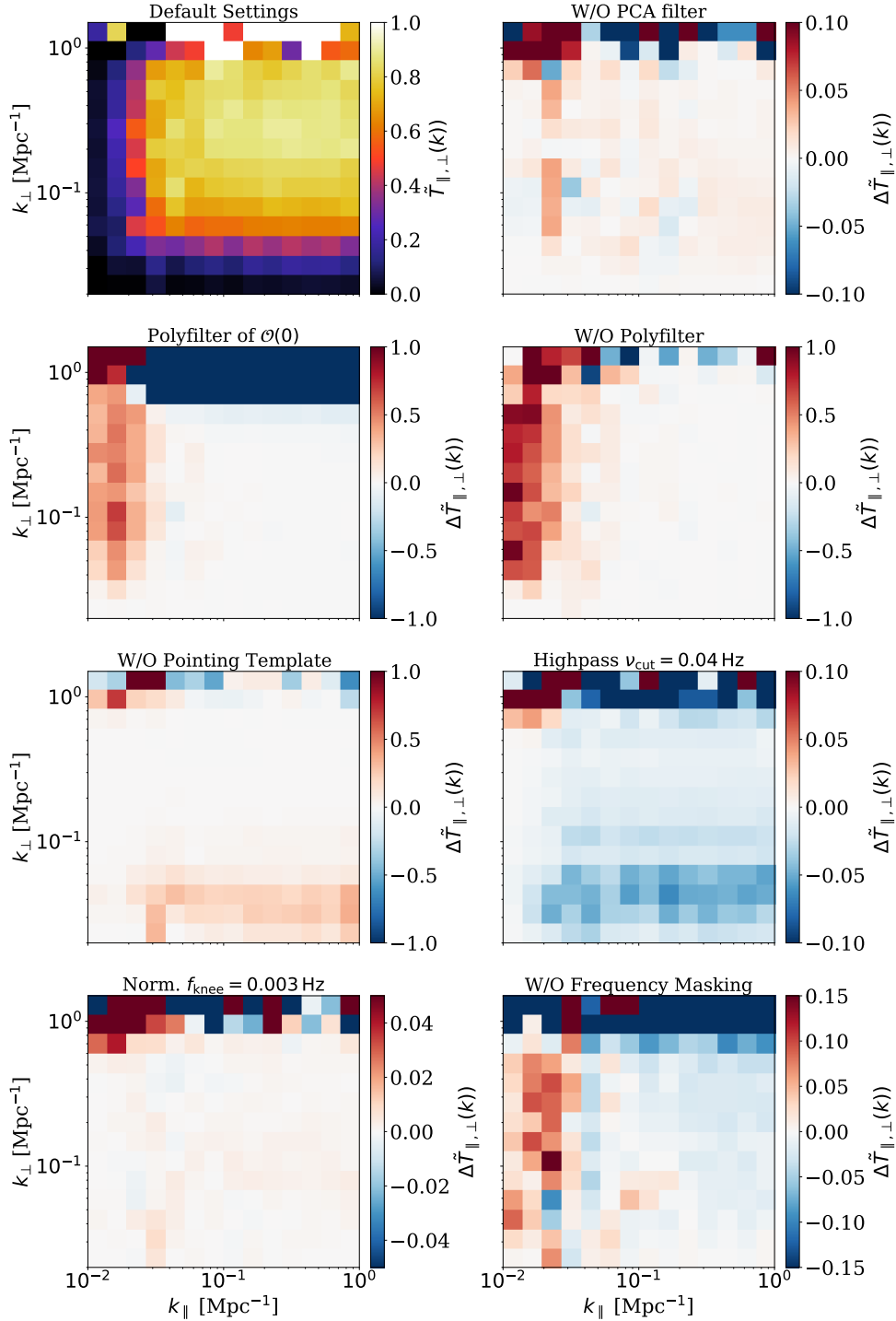


Figure 6.7: The 2D transfer function corresponding to the default filter settings (**uppermost left**) as well as the difference between the transfer function corresponding to each filter change and that of the default settings (see figure titles for which filters were considered). The 2D transfer function differences shown here correspond to the 1D differences of Fig. 6.6. As in Fig. 6.6 the transfer functions are produced using a single simulation realization as well as six obsIDs of CO6 (mixed Lissajous and CES type scans).

and frequencies. It thus makes sense that we only observe a minor change in transfer efficiency of the largest modes. Thus if removing four principal components hardly changes the recovered amount of signal we could potentially even subtract more modes from the TOD in the future, without biasing the signal estimate in any significant way.

The very last filter we consider is the normalization by the running mean where we changed the knee frequency from the default 0.01 Hz to 0.003 Hz. By using a lower knee frequency the running mean by which the time stream is normalized will be even more slowly varying. Thus only the very slowest temporal modes of the time stream should be normalized out by dividing by the running mean. Nevertheless, as we see in Fig. 6.6, and most notably in Fig. 6.7, the lower knee frequency only lets through a few percent additional signal at the lowest k bins. This suggests that since the signal has very few modes that are of a long enough timescale to be affected by the normalization in any significant way, using a knee frequency $f_{\text{knee}} = 0.003$ Hz or the slightly harsher $f_{\text{knee}} = 0.01$ Hz makes almost no difference. We can therefore use the default 0.01 Hz knee frequency to take out slightly more long-timescale modes from the TOD without any significant signal bias.

Different Scanning Strategies

When we estimate the astrophysical signal from FPXS we often split the data according to the scanning strategy, which could bias our COMAP signal estimates differently. Therefore, we also computed the transfer function for a dataset of six mixed CO6 obsIDs of Lissajous and CES, as well as separately for the dataset of the three obsIDs from each scanning strategy. The resulting transfer functions can respectively be seen in Fig. 6.8a - c. Each of the three transfer functions were computed from a single realization of the signal, in addition to a small obsID dataset, because we again are only interested in qualitative tendencies.

When comparing the recovery of the signal of the Lissajous and CES scans we can see that the Lissajous transfer function has a higher peak efficiency. The reason for this is most probably that the higher scanning speed of Lissajous scans results in covering more sky, hence collecting more data, in a comparable amount of time than a CES scan.

Another point to note is that the peak of the Lissajous transfer function seems to be shifted towards larger scales than that of the CES scans. Also, the transfer functions of both scanning strategies have a similar tendency to decrease somewhat when going from the transfer function's peak towards high k_{\perp} , although this tendency seems to be somewhat emphasized in the Lissajous case. The reason for this shift in the peak transfer function to lower k is not completely known, but it could perhaps be caused by Lissajous scans having a better cross-linking than the CES scans, which would result in a somewhat better recovery of the larger scales. However, this is debatable and needs some more investigation in the future.

Finally, we can see that the mixed transfer function does not quite reach the same peak efficiency as the Lissajous transfer function, but peaks a little higher than that of the CES scans. Also, the mixed transfer function's peak is located approximately between the peaks of the two other transfer functions.

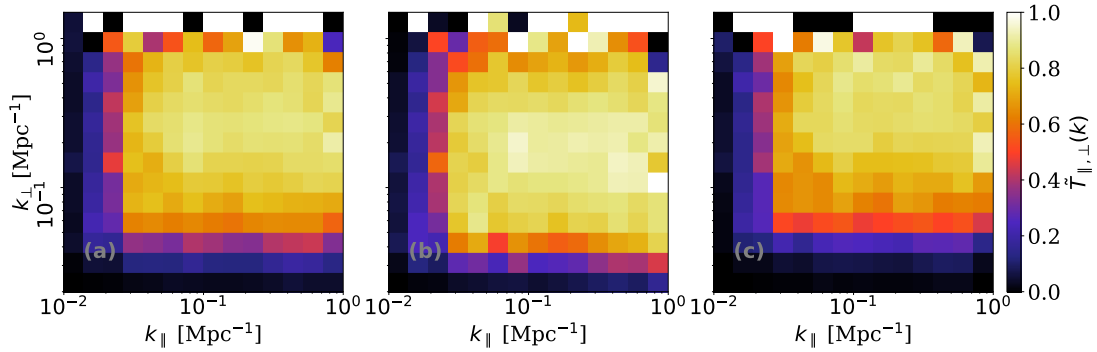


Figure 6.8: The transfer function of a single realization of the simulation cube using (a) three obsIDs of type Lissajous and three CES scans, (b) the transfer function of the three Lissajous obsIDs as well as (c) the three CES type observations.

Because we observe quite different transfer functions for each scanning strategy, they will not upwards-adjust our estimated signal power spectra and their error bars equally. Hence when we compute spectra from Lissajous- or CES-only data (as for example in FPXS) the respective transfer function should be used when taking the pipeline filtering into account. Meanwhile, when using a dataset of an approximately equal amount of each scan type, the transfer function of a mixed dataset should be used when accounting for the bias of the signal.

6.2 The Improved Mapmaker

Having seen how we can use our simulation pipeline to account for the filter-induced signal attenuation, it is time to have a look at the destriper and how well we can improve upon the current mapmaker design using a noise weighted binning algorithm of the highpass filtered TOD.

In Fig. 6.9 the TOD of an arbitrary frequency and feed is shown in relation to some 10s baseline fits, amplified by a factor of 10 to make them visible. When comparing the uncorrected level 2 TOD to its baseline fit, we can see that the latter is far higher in amplitude than the former. Only looking at the time domain data hence gives the impression that the correction of the TOD by the baseline fit would be quite subtle.

Furthermore, we see that the baseline fit of the inverse diagonal preconditioned and un-precondition CG solver are essentially the same. Also, we found that the number of iterations needed to converge in the CG solver was quite similar in the two compared cases and usually close to 10-30. Nevertheless, the preconditioned solver was somewhat more stable in the number of iterations, while the un-preconditioned solver occasionally spiked to 100-200 iterations. We also noted a tendency for the number of iterations to increase somewhat when fitting baselines to increasingly larger datasets. Thus, as we in this section only consider relatively small datasets, the regular CG solver should be sufficient. That said, in the future when testing the destriper on a full dataset of a

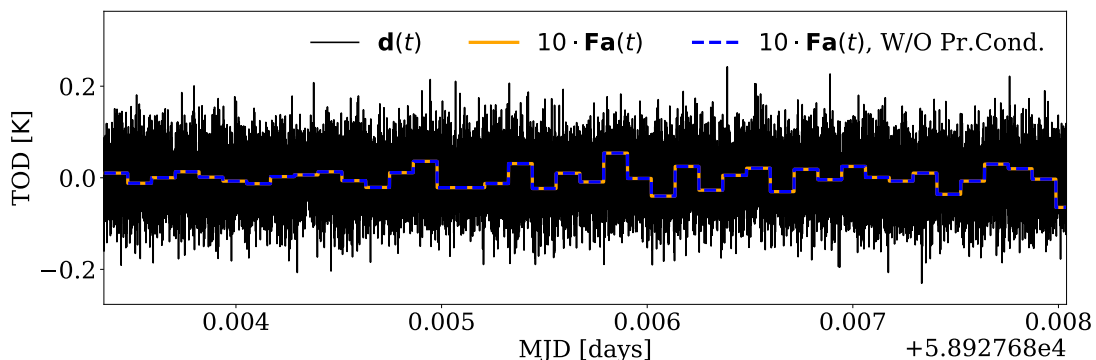


Figure 6.9: An example of a level 2 TOD $d(t)$ with some corresponding (10 times) amplified baseline fits overplotted. The baselines of both the preconditioned fit (dashed blue) and the un-preconditioned fit (orange) are shown.

given CO field, using the preconditioner could be beneficial if the number of iterations becomes large and the danger of numerical errors increases.

To gain a better idea of which baseline lengths maximize the signal recovery we computed the 1D binned transfer functions using different baseline lengths. Because we are only interested in the relative performance between different baseline lengths we used a small dataset of a single signal realization added to six obsIDs (of mixed scan types). Figure 6.10 shows the resulting 1D transfer functions in relation to the one made by the binned mapmaker using a 0.02 GHz highpass cut of the TOD prior to mapmaking.

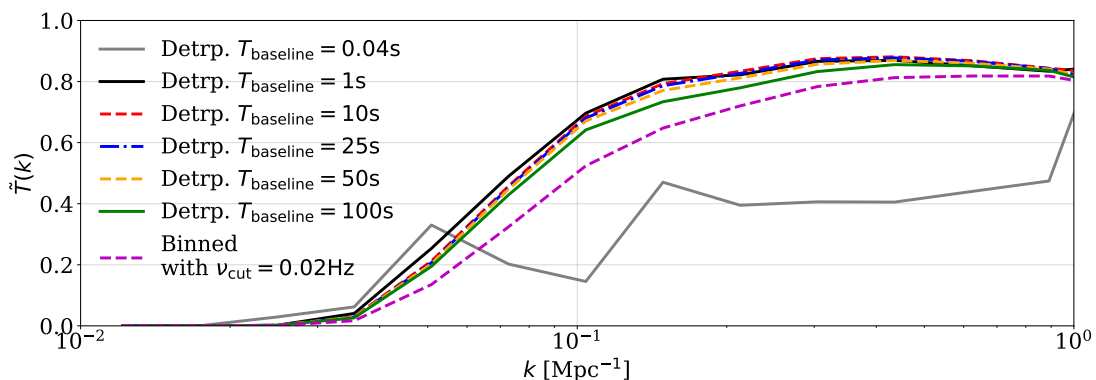


Figure 6.10: The 1D binned transfer function corresponding to the destriper, fitting baselines of different lengths simultaneously to the time stream of all feeds, compared to the transfer function from a highpass filtered and noise weighted binned mapmaker.

What we can see is that all baseline lengths, except the very shortest one at 0.04s, provides a better transfer efficiency than the binned mapmaker with highpass cuts. Specifically, we see that when starting at 100s baselines and decrease the baseline

lengths stepwise to 10 s we get an ever higher transfer efficiency. However, we also see that if the baselines get too short, i.e. 1 s and below, that the transfer function starts to behave unpredictably. Although the 1 s baselines have the highest transfer function at scales $k \leq 0.1 \text{ Mpc}^{-1}$, the transfer function of the 10 s baselines is slightly better at higher k and behaves more predictable overall. Going down to baselines of length 0.04 s we see that the transfer function starts to break down and oscillate wildly below even that of the binned map.

The reason this happens is most probably due to the short baselines starting to become degenerate with the white noise in the TOD. Take for instance a case with baselines of length one time sample. Since we can freely vary each baseline amplitude in the fit, we will inevitably fit the whole data, both the white noise and the signal. The result is a residual of almost zero. In the case of 0.04 s baselines, i.e. two time samples in length, the effect is similar and way too much signal is taken out by the baseline fit. What we want is hence a baseline length with which the long-timescale residuals are fitted well, but the white noise and the signal are left out. Hence, for now, it seems as if the 10 s baselines seem to behave the best in terms of transfer efficiency, increasing the transfer function at intermediate scales by up to $\sim 15\%$ with respect to the default binning and highpass filtering.

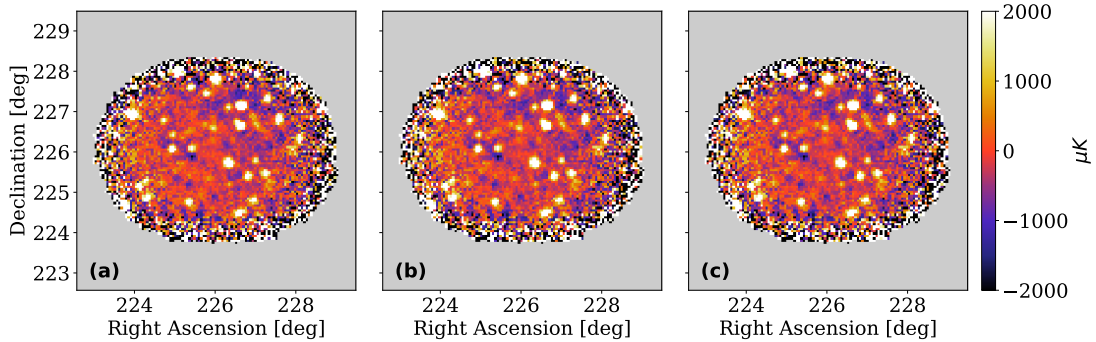


Figure 6.11: Three 32.33 GHz (i.e. B:USB channel 11.) maps; (a) made using the highpass filtered and noise weighted binned mapmaker, (b) the all-feeds destriper map as well as (c) the map from the per-split destriper.

However, the transfer function alone is not enough to determine whether the 10 s baseline destriper yields an acceptable result, the maps and power spectra must also look sufficiently good. To compare the 10 s destriper to the binned mapmaker, we thus made some maps of the dataset of 63 obsIDs of CO6 previously used, with additional signal. In particular, both the all-feeds and the per-split baseline fit were tested. In Fig. 6.11 one can see the resulting maps of a 32 GHz frequency channel, made by (Fig. 6.11a) the binned, (Fig. 6.11b) the destriper with a simultaneous baseline fit on all feeds as well as (Fig. 6.11c) the map made by destripping the TOD on a per-split basis. The additional signal was added to the TOD to test whether it would properly cancel out in the difference between each of the maps, ensuring that the signal was left untouched by the destriper. The corresponding relative difference between the maps can be seen

in Fig. 6.12.

When comparing between the maps in Fig. 6.11 it is hard to spot large qualitative differences by eye. However, looking at the difference maps in Fig. 6.12 between the binned map and respectively the all-feeds and the per-split destriper maps (respectively in the left and middle panels) we see a clear "stripy" pattern appearing. This large-scale difference pattern is of the order of several $100 \mu\text{K}$ and seems like a significant improvement in case the structure removed by the destriper is indeed that of the residual systematics. Nevertheless, the difference is not too large to wake any larger suspicions of an un-physical result.

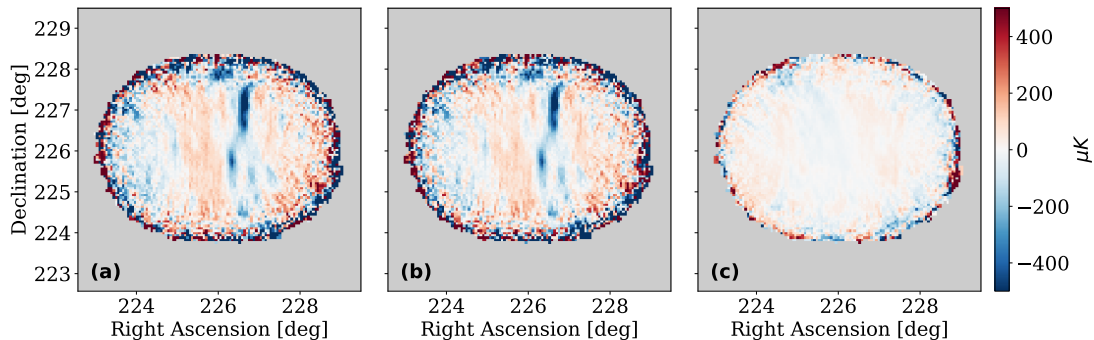


Figure 6.12: Three 32.33 GHz (i.e. B:USB channel 11.) difference maps; (a) between the highpass filtered and binned map (Fig. (6.11a)) and the all-feeds destriper map (Fig. (6.11b)), (b) between the highpass filtered and binned map and the data split-wise destriper map (Fig. (6.11c)) as well as (c) between the maps of the two destriper approaches.

What can also be noted is that the difference between the two destriper maps, seen in the right panel, is much lower than between any of the destriper maps and the binned one. This suggests that in fact, the destripping of all feeds and the destripping on a per-split basis are roughly equivalent. Note also that the boundary regions of the difference maps show the largest deviations, which most probably is due to slightly different noise properties at the high-noise boundary. But as the most interesting and trustworthy regions are the central ones, where the SNR is the highest, we can fairly safely neglect the edge excess observed.

A common feature observed in all of the difference maps is that the added signal is nicely canceled. This suggests that the signal itself was left untouched by the improved mapmaker schemes. This is good to see as it is further evidence that the destriper only attacks the dynamic components of the TOD and leaves the static signal alone.

Meanwhile, the maps seen in Fig. 6.11 and 6.12 merely show the result of the destriper from a single frequency slice. However, a conclusion in favor of the destriper must be justified by the whole 3D map. In particular, we want to ensure that the destriper also yields an acceptable result on the same 63 obsIDs but without added simulation. To illustrate the structures in both the parallel and perpendicular dimensions of the maps from each mapmaker scheme we computed the 2D (pseudo) auto-spectra.

The resulting power spectra can be seen in Fig. 6.13 and correspond, respectively from (a)-(c), to the binned, all-feeds destriper and split-wise destriper maps.

As we can see, the two power spectra from the destriper look virtually identical. Compared to the spectrum of the binned map there are, however, some clear differences seen in the large-scale excess structure. We see that this large-scale excess looks comparable in structure in all three spectra, but is considerably amplified in strength in the spectra corresponding to the destriper. The similarities in the shape of the excess suggest that we are not inducing any large-scale structures when destripping the maps. Hence the amplified excess seen is probably some already existing structure that is simply resolved better by the destriper algorithm than when highpass cutting the TOD and binning the maps.

However, we recall from our earlier discussion of Eq. (5.13) that we neglected the prior term $\mathbf{a}^T \mathbf{N}_a^{-1} \mathbf{a}$. This could introduce some artificial constraints on the large-scale structures, preventing the noise on large scales to integrate down as effectively as on small scales. Potentially, this can lead to a higher power spectrum value at low k , especially if we in the future test the destriper on large datasets where the small-scale noise can be integrated down significantly. Thus including the prior term on the baseline amplitudes could in the future potentially improve the destriper on large scales.

As the 10s baseline fit maximizes the transfer function and seems to produce relatively clean-looking maps we want to know in more detail how well it recovers additional signal on different scales. Thus we computed the 2D transfer functions corresponding to the 63 obsID dataset of CO6. This is shown in Fig. 6.14, with the panel order following the same mapmaking schemes as in the auto-spectra in Fig. 6.13. As we can see immediately when comparing the transfer functions side-by-side is that the one resulting from the binned mapmaker scheme is inferior to either of the two transfer functions of the destrippers. This can also be seen in the difference plots in Fig. 6.15. Compared to the highpass filtering and binning, either of the destrippers appear to recover $\sim 20 - 25\%$

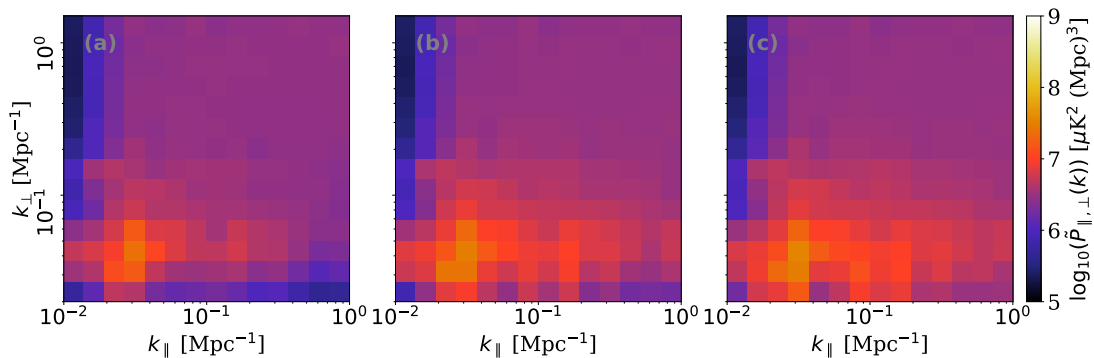


Figure 6.13: The pseudo-auto-spectra of 63 obsIDs of the CO6 field resulting from a mapmaking with (a) highpass filtering and noise weighted binning algorithm, (b) a destriper with a baseline fit on all feeds simultaneously as well as (c) from a destriper using a baseline fit on data split batches.

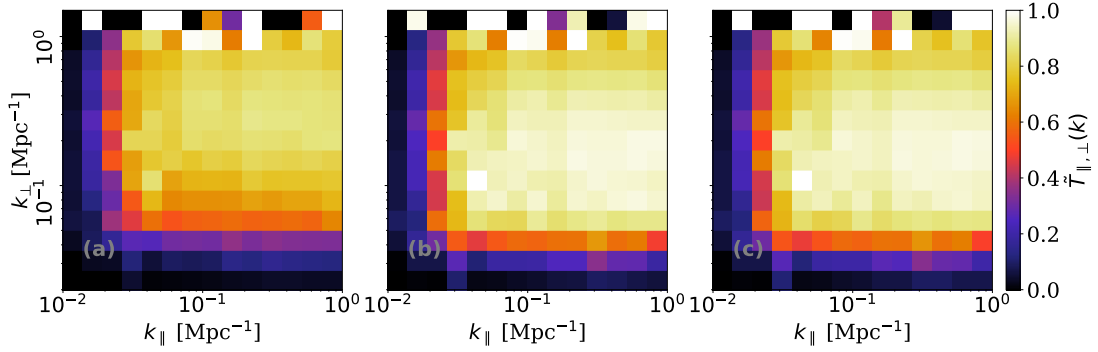


Figure 6.14: The transfer function resulting from a mapmaking with a (a) highpass filtering and noise weighted binning algorithm, (b) an all-feeds destriper (c) from a per-split destriper. The transfer functions are all computed based on a single realization of the simulated signal as well as 63 obsIDs of CO6.

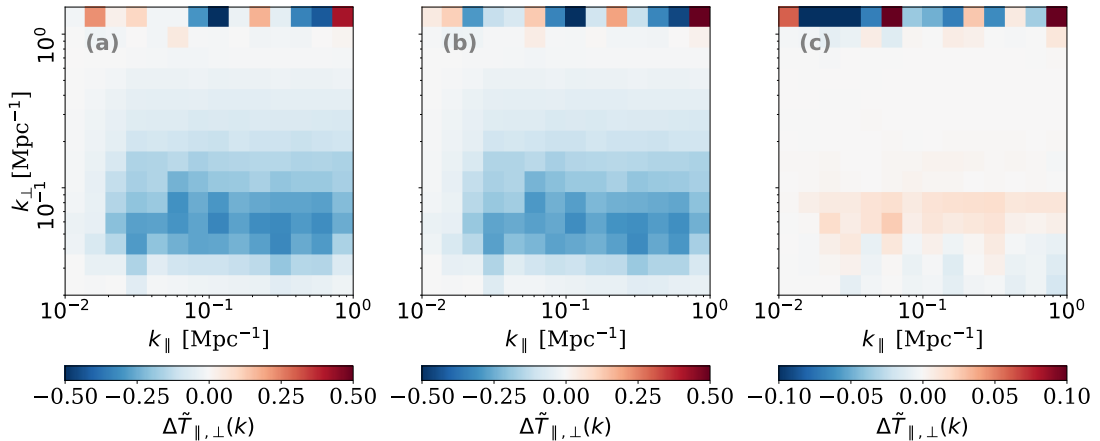


Figure 6.15: The difference between the transfer function computed by a mapmaker using; (a) highpass filtering and noise weighted binning (Fig. (6.14a)) and the all-feeds destriper (Fig. (6.14b)), (b) highpass filtering and noise weighted binning and a per-split destriper (Fig. (6.14c)) as well as (c) between the two destripping approaches.

more signal at large perpendicular scales. Before, when using only the highpass filtered and binned mapmaker the maximum transfer efficiency was at $\sim 85 - 90\%$, while the destrippers manage to reach a peak transfer function of $\sim 95\%$. Meanwhile, we see only a small improvement of $\sim 5\%$ in transfer efficiency in the large perpendicular modes in the all-feeds destripping over the destripping per-splits, as seen in panel (c) of Fig. 6.15.

All in all, it seems as if the destriper with 10s baselines nicely recovers additional large-scale structures in the signal that were previously canceled by the 0.02 Hz highpass cuts prior to the binned mapmaker. Moreover, the two approaches of destripping perform almost equally well. Having better cross-linking when fitting baselines to the combined

TOD of all feeds only improves the transfer efficiency by a small amount.

However, when it comes to run time of the mapmaker the binned mapmaker is the fastest. While the binned mapmaking takes under an hour when parallelizing sufficiently, the destriper need between two and five hours for the 63 obsID dataset. In particular, the all-feeds destriper is slightly faster than the per-split one, as there are fewer I/O bound read and write operations needed. Also, the pointing matrix, etc. can be precomputed for all frequencies in the all-feeds destriper, while these need to be computed for each data split batch in the other destriper approach.

Meanwhile, the far more limiting factor of memory usage is effectively held at a fairly low level when destripping the TOD in smaller split batches. Another important advantage of destripping the TOD in data split batches is that the data in each batch is completely independent which enables us to maintain the same flexibility of computing cross-spectra as before. Hence due to all the aforementioned advantages, the destriper fitting baselines on a per-split basis proves to be the superior scheme. Nonetheless, the binned mapmaker with highpass cuts is still useful when in need of a quick run-through of the data.

6.3 The Eigen-Spectrum Analysis

In this section, we present and discuss the results from the PCA of the cross-spectrum dataset resulting from a multitude of successive data splits and, as described in Sec. 5.3. In particular, we will consider the PCA eigen-spectra from each of the CO fields and discuss commonalities. Subsequently, we will discuss how we may use the PCA amplitudes to identifying spectra with contamination by systematic effects and investigate the loss of signal when subtracting the systematic artifacts.

6.3.1 The Eigen-Spectra

When computing feed-feed cross-spectra from the full dataset of each CO field (i.e. ~ 1800 obsIDs per field), according to the methods described in Sec. 4.2.2 using the data splits of Split Tree 1., presented in Table 6.1, we get in total 8192 different combinations of the data from which to compute FPXS. As one of these successive splits is between Lissajous and CES data we get 4096 spectra per scanning strategy. This gives sizable data space over which to compute the PCA eigen-spectra that capture the directions of most spread in k -bin space.

Because of the movement in both azimuth and elevation in a Lissajous scan, as opposed to motion in azimuth only in CES scans, their corresponding spectra should in theory show different amounts of contamination from possible residual systematic effects like ground pickup. We, therefore, chose to treat the data from each scanning strategy separately, to not pick up the variance in the data induced by the differences in the scanning strategy when computing PCA eigen-spectra. In Fig. 6.16 the resulting first five eigen-spectra (i.e. unit eigen-vectors) for each scanning strategy and CO field can be seen. The amount of variance in the direction of k -bin space corresponding to

each eigen-spectrum is given above each plot in the percentage of the total variance of the data.

What we immediately see in Fig. 6.16 is how unevenly the explained variance per eigen-spectrum is distributed. For instance, the first CO6 and CO7 eigen-spectra for each scanning strategy describe about 50 – 60% of the variance in their respective dataset, while the remaining modes contain below 10% of the variance each. In the case of CO2, the percentage of variance is somewhat more evenly distributed over the first five eigen-spectra. However, the first couple of eigen-spectra showing 10 – 25% of the variance each is still a considerable amount for a single eigen-spectrum.

Table 6.1: **Left table:** The abbreviations for the different data splits considered as well as a short description of the abbreviation's meaning. **Right table:** The two split "trees" considered, with the test and control variables according to which the splits are performed. After the test variable split the data are subsequently split (successively) according to the control variables. The variables marked with "s" are the so-called split variable across which the cross spectra are computed for each feed-feed combination.

Abbreviation:	Description:	Split Tree 1.	
		Test variables:	Control variables:
cesc	Lissajous/ CES	ambt	elev (s)
elev	High/ low elevation	dayn	cesc
ambt	High/ low ambient temperature	fpol	
dayn	Day/night	half	
fpol	High/low value of knee frequency of the first polyfilter constant (c_0 from Eq. 3.13)	odde	
		rise	
		wind	
		wint	
		Split Tree 2.	
		Test variables:	Control variables:
half	Half mission	elev (s)	cesc
odde	Odd/even obsID	dayn (s)	
rise	Field rising/setting	sidr (s)	
wind	High/low wind speed	ambt (s)	
wint	Winter/summer	wind (s)	
sidr	High/low sidereal time	wint (s)	
		rise (s)	

If the data we perform the PCA on were containing only noise-spectra we would expect each eigen-spectrum to only resemble random fluctuations. Also, each eigen-spectrum would be explained by a small percentage of the variance. Therefore, since we both see some clear structure in the leading eigen-spectra, and these contain most of the explained variance in the data, it indicates that systematic effects are picked up by the eigen-spectra. That is, since a given systematic effect is only showing up in some of the splits and or feed-feed combinations, the corresponding variance in the dataset

of different splits will be correspondingly large.

Furthermore, we can clearly see that the first couple of PCA modes of CO6 and CO7 look very similar in structure and are described by similar amounts of variance. This is especially notable in the Lissajous eigen-spectra, but also to some extent in the CES spectra. As both CO6 and CO7 are quite similar in their path across the sky, the similarities in the structures of the first eigen-spectra of the two fields seem to further support the hypothesis that the eigen-spectra could describe common systematic features of the two fields.

When comparing the Lissajous and CES eigen-spectra within each field we also see some similarities. This is, in particular, the case for the first eigen-spectra of CO6 and CO7, but also some of the other spectra. For instance, the dipole structure observed in the second Lissajous components of CO6 looks similar in structure to the fourth component of CES, as well as the fourth Lissajous component of CO6 resembling the third CES eigen-spectrum. In the former of these examples, it is important to note that the overall sign of the spectra is irrelevant when comparing them as these are only unit vectors in k -bin space. Hence two overall anti-correlated eigen-spectra will simply point in opposite directions along the same principal axis of the data. Therefore, it is likely that some systematic effects are common among CES and Lissajous scans in the CO6 and CO7 fields, although contributing to the variance in the cross-spectra to different levels.

Interestingly enough, the same tendencies for structures seen in Lissajous eigen-spectra to resemble those seen in the CES equivalent do not seem to be observed in the CO2 eigen-spectra. The exception, being perhaps the last of the spectra of Lissajous and CES. Thus the systematic effects in Lissajous and CES seem to look different in the CO2 data. In fact, when comparing the CO2 spectra to either of the other two field's spectra there are no obvious similarities. This suggests that CO2 has quite different systematic effects than the other two fields, which makes sense considering that CO2 moves quite differently across the sky than the other fields.

6.3.2 Identifying Cross-Spectrum Outliers

Apart from directly comparing the structures observed in the eigen-spectra to those caused by known effects, there are some other approaches we can use to identify contaminated FPXS. To be clear, this type of analysis is the first of its kind performed on this particular dataset and we are hence only scratching the surface of potential ways of using the data from the PCA. Nevertheless, even our tentative results we present could have great potential in classifying the spectra in our dataset by their systematic contamination.

As mentioned in Sec. 5.3 we can use the mean-centered PCA amplitudes of the first two eigen-spectra of an FPXS dataset to reduce its dimensionality from 196 down to the plane spanned by the two eigen-spectra. In the case of CO6 and CO7, these two components explain about 60 – 70% of the information in the data, which means we can reduce the dimensionality of the data by a huge amount and still only lose 30 – 40% of the information. As an example, we performed this dimensionality reduction on the

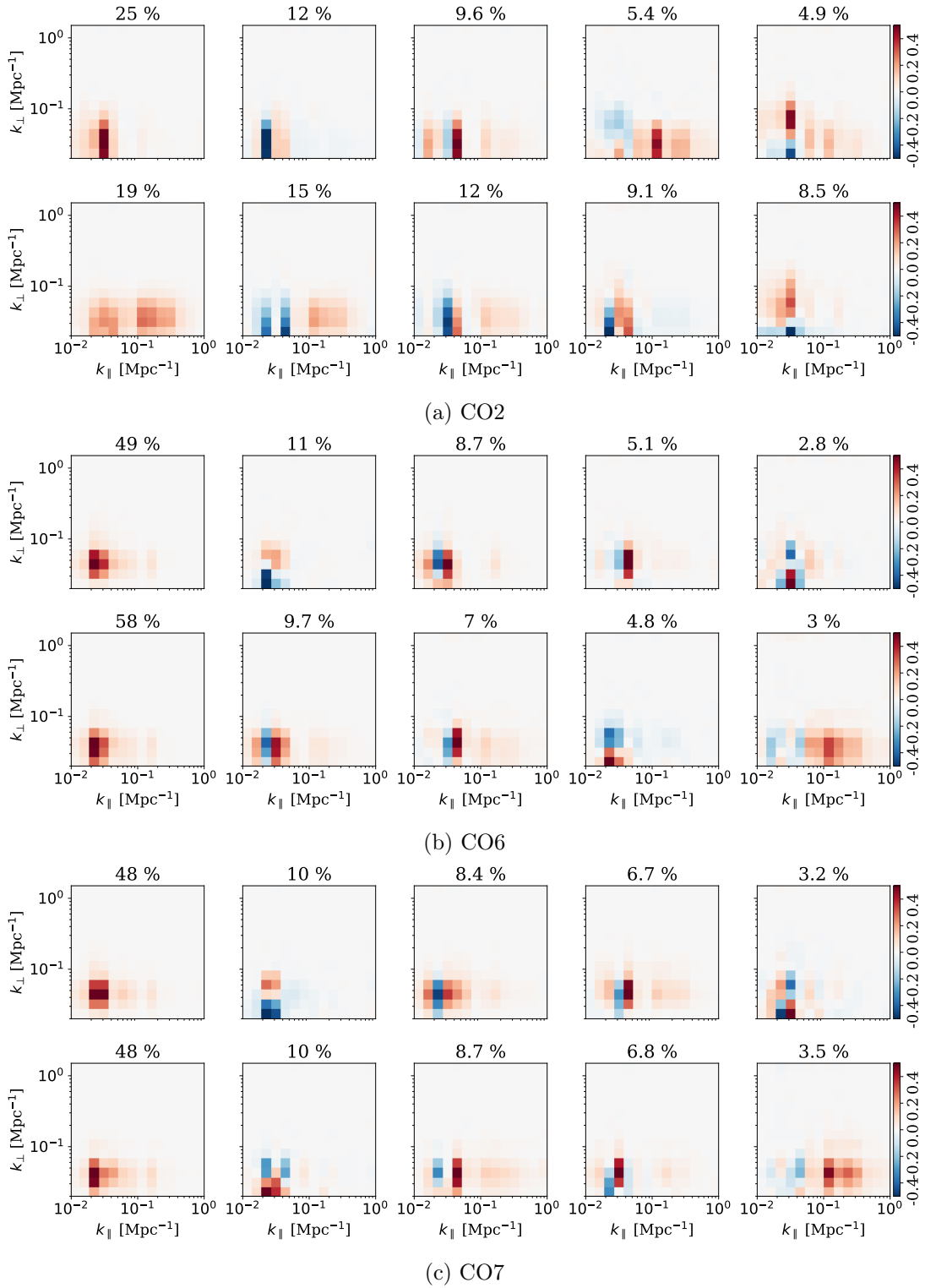


Figure 6.16: The first five PCA eigen-spectra for each field (respectively CO2, CO6 and CO7 from top to bottom). The upper and lower row of eigen-spectra for each field respectively are derived from Lissajous and CES data only. The eigen-spectra corresponding to each field and scanning strategy are derived from FPXS that are computed from all the 4096 possible, split and feed, combinations of the data described in Split Tree 1. of Table 6.1. The percentage of the total variance explained by each eigen-spectrum is seen above each plot.

dataset of CO7 Lissajous FPXS.

In Fig. 6.17 this is shown as a scatter plot. Each dot corresponds to a given FPXS in the dataset, where the planar x and y coordinates are respectively given by the mean-centered PCA amplitudes of the first two CO7 Lissajous eigen-spectra from Fig. 6.16. What becomes evident when displaying the data in this way is the clustering patterns of the dots. It looks like there are about three or so distinct clusters in the data, a central one with the bulk of all dots, and two outlier clusters to the top and right of the central cluster. As we see in the left panel of Fig. 6.17, each of the three clusters has a relatively even distribution of spectra from each data split. This indicates that the outliers in the dataset are not primarily associated with a given split. That being said, this could very well look different when using other splits, and should be investigated more in the future.

However, if we look at the right panel of Fig. 6.17, which indicates which feeds were crossed in a given FPXS, we can easily see that the outlier clusters belong to distinct feed-feed combinations. In particular, we see that the right-most outlier cluster is given by all the spectra which cross feed 8 with itself. Because these spectra are crossing the same feeds the common systematic effects in feed 8 are amplified, resulting in a far outlier cluster. This is one of the reasons we leave out these auto-feed combinations of FPXS when computing the final averaged FPXS. Thus in the future, it could be interesting to repeat the PCA with the auto-feed spectra excluded.

Meanwhile, we also see in Fig. 6.17 that the upper-most outlier cluster is defined by feed 17×8 , or vice versa. In general, it seems as if almost all outliers are associated with feed 8 in some way. This could mean that feed 8 is affected by some systematic effect associated with the two first principal axis of the dataset. Alternatively feed 8 could be somewhat more sensitive than the other feeds, which could lead to it picking up slightly more of a systematic effect than the other feeds. Nevertheless, looking into feed 8, in particular, could be worthwhile in the future when trying to rule out all the sources of systematic error in the data.

To determine if large absolute PCA amplitudes correspond to spectra with large systematic effects, and vice versa, we plotted some arbitrary FPXS with high and low absolute amplitudes in the first principal axis of the data. These spectra are displayed in Fig. 6.18, where the high and low absolute PCA amplitudes correspond to spectra in the upper and lower rows respectively. As can be seen, the high PCA amplitude spectra all show clear signs of systematic effects, while the low amplitude spectra look much cleaner. The first of the contaminated spectra for instance has a large excess structure at large scales reminiscent of the first eigen-spectrum of CO7 Lissajous. Meanwhile, the third of the bad spectra has several bar-like structures at constant k_{\parallel} which are reminiscent of standing waves in our data. Lastly, we find the same dipole-like excess structure at large scales in the second of the contaminated spectra that was previously seen in the second eigen-spectrum of CO7 Lissajous.

Of course, the examples shown here are only a few, but in general, there seems to be a correlation between the high absolute values of the PCA amplitudes and spectra showing signs of larger systematic effects. To conclude, we can thus use this type of PCA to make reasonably good estimates on whether a given cross-spectrum is contaminated

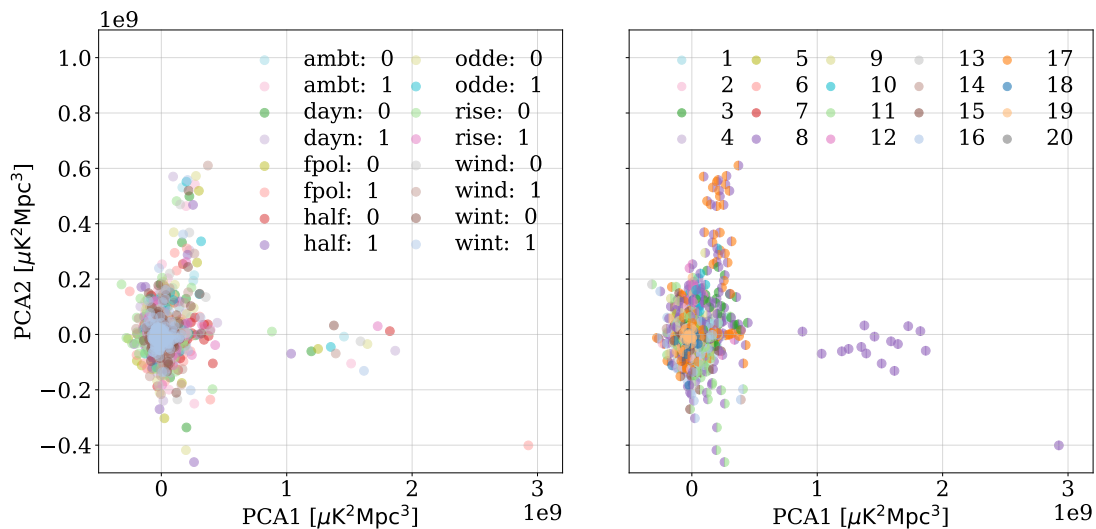


Figure 6.17: Scatter of the CO7 Lissajous cross-elevation FPXS dataset in the 2D subspace spanned by its two first eigen-spectra (seen in Fig. 6.16). In the **left panel** the dots are colored according to the which split the corresponding FPXS belongs to (see Table 6.1 for description of split abbreviation), while the **right panel** shows the same scatter colored according to the crossed feeds. The left and right semicircles of each dot respectively indicate the two feeds that were crossed.

by systematic artifacts. However, an exact threshold for when a given amplitude is considered to be high needs to be established in the future. The identified outliers could be masked in the further analysis, or as we will see in the next subsection we could potentially clean the outliers. Hence there is still a wealth of interesting data that can be mined from this kind of PCA of FPXS and we are here merely scratching the surface in terms of identifying systematic effects.

6.3.3 Signal Loss due to Principal Component Subtraction

Finally, we want to see what happens when we subtract out some of the principal components of a spectrum believed to be associated with a systematic effect. To do this we simply use the same dataset of FPXS from which we derived the transfer function seen in Fig. 6.5. From the resulting dataset of 3586 FPXS without additional signal, we computed the PCA eigen-spectra, seen in the top row of Fig. 6.19. The middle two rows of the figure respectively show the principal components of the averaged Lissajous cross-elevation spectra from Fig. 6.5 with and without additional signal.

Now, the reason the FPXS dataset without added signal was chosen when computing the eigen-spectrum fit, is that the signal is orders of magnitude stronger than the TOD it was added to. Thus small variations in the observed sky portion or different mode mixing in the different splits could result in each FPXS picking up slightly different Fourier modes of the signal when computing FPXS. As the signal is very strong this

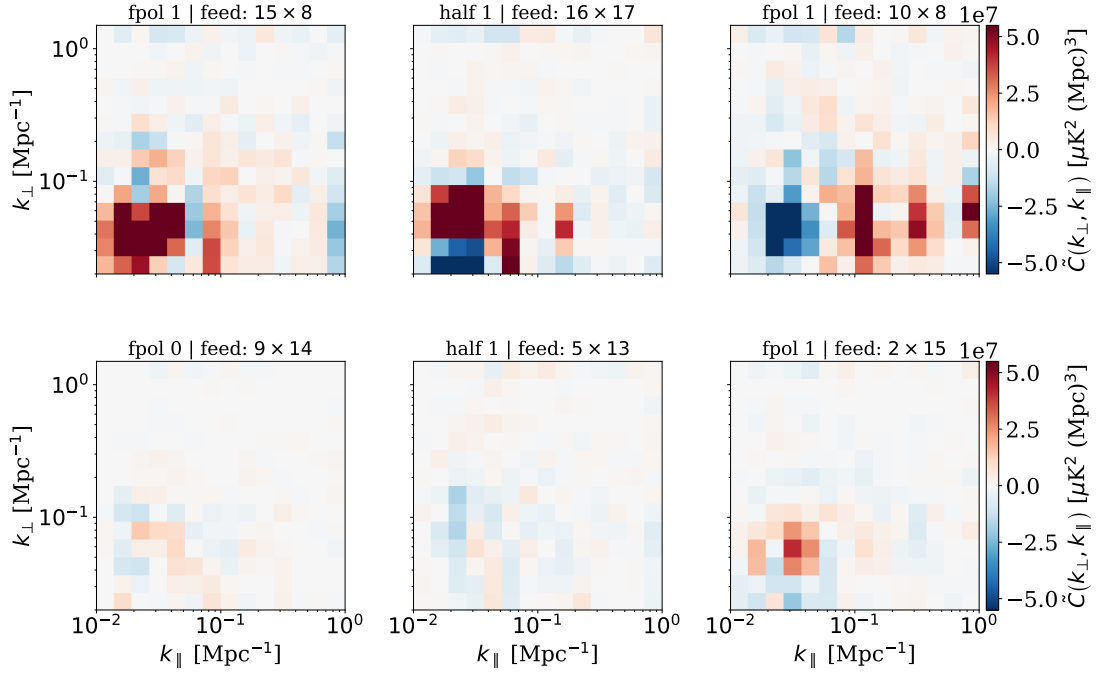


Figure 6.18: Some selected spectra from the CO7 Lissajous cross-elevation FPXS dataset. The **upper panel** of spectra were selected according to a relatively high absolute value of the mean-centered PCA amplitudes of the first CO7 Lissajous eigen-spectrum. Meanwhile, the **lower panels** show the opposite case, being selected from the lower PCA amplitudes.

could lead to a correspondingly sizable variance in the FPXS dataset, which in turn would be picked up by the first PCA eigen-spectrum. Therefore fitting, and subsequently removing, the PCA modes found by an eigen-spectrum fit on the FPXS data with signal may result in unrealistically much signal loss.

To estimate how much signal is lost when subtracting the first five principal components from the FPXS with added signal, we computed the corresponding effective transfer functions. In the bottom row of Fig. 6.19 the relative difference between the resulting effective transfer functions of the cumulatively subtracted principal components and the default FPXS derived transfer function are shown. What we can see in the difference of the transfer functions is that we at most alter the signal by 20 – 50% in removing up to five principal components. As we previously discussed, the noise dominated regime of the FPXS-derived transfer functions was found to be somewhat unpredictable. Hence the large (absolute) value of the transfer function differences in the upper three to four rows of k_{\perp} -bins should not be taken very seriously.

Furthermore, the additional signal loss due to subtraction of the principal components is mainly contained in a region of a few of the large-scale bins. Thus subtracting the principal components of the data does not overall bias the transfer function.

It is also important to mention that the transfer function differences could look some-

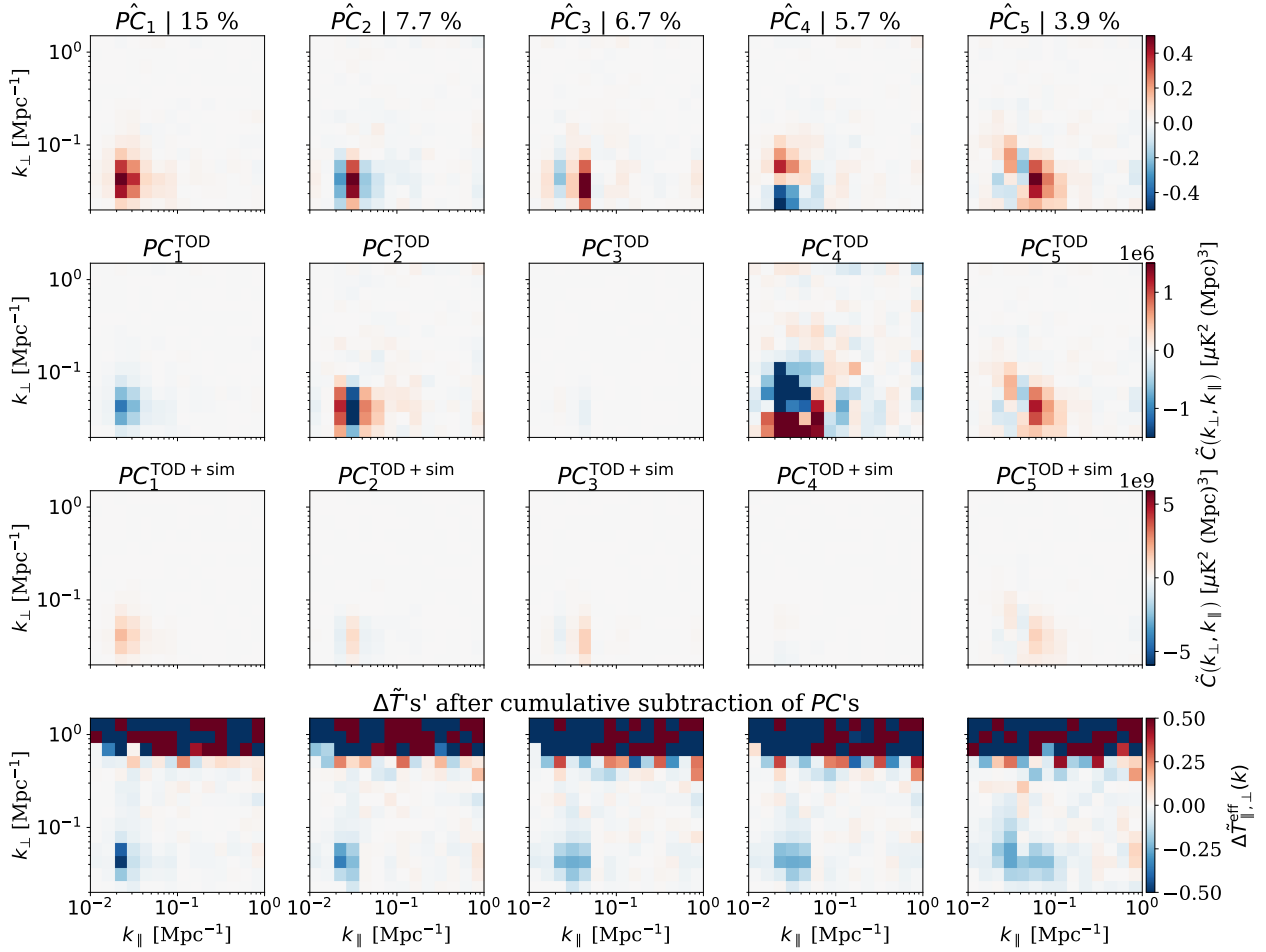


Figure 6.19: **Upper row:** the five first eigen-spectra (with the corresponding fractional explained variance above each plot) from the fit to the FPXS dataset without added signal, resulting from our 63 obsID subset of CO6 split according to Split Tree 2. of Table 6.1. **Upper middle row:** first five principal components of the average FPXS of the TOD without added signal seen in Fig. 6.5. **Lower middle row:** the corresponding principal components of the average FPXS of the TOD with additional signal from Fig. 6.5. **Lower panel:** the difference between the effective transfer function after a cumulative removal (from left two right) of the first five principal components (from the two middle rows) and the transfer function without any removal of principal components (see Fig. 6.5).

what different when removing the principal components from another set of averaged FPXS. That is, if another FPXS has a higher or lower contribution (i.e. corresponding PCA amplitude) to the systematic effects represented by the subtracted eigen-spectra, the subsequent loss in signal may be correspondingly higher or lower. Nevertheless, the shown example serves as a reasonably realistic illustration of the amount of signal loss

due to the PCA filtering of FPXS.

Although losing up to 50% of the signal in some of the bins may seem like a significant loss, it may be weighed up by the fact that we could potentially clean the data of some of the systematic effects at the price of slightly upwards adjusted uncertainties in affected bins. In the current cross-spectrum methods we implement in COMAP any FPXS of $\chi^2 > 5$ above the white noise expectations are discarded when computing an averaged FPXS. Thus, if the subtraction of systematic effects through PCA modes proves to be effective in cleaning bad spectra in the future we could potentially include previously discarded data in the FPXS averages without suffering too much additional signal loss. This could be an especially powerful tool in the COMAP pipeline if eliminating the remaining systematic effects seen in the data proves to be challenging by traditional means. However, as we are only scratching the surface on the potential use of PCA on FPXS, it remains to be tested systematically in future work how well we can mitigate systematic errors using this method.

Chapter 7

Conclusion and Future Outlooks

Through this thesis, we have worked on three main improvements to the analysis pipeline of the COMAP line intensity mapping experiment. We extended the COMAP pipeline to compute the pipeline transfer function, which is essential to account for the filter-induced bias of the signal estimate. In addition, we have implemented a destriper in the mapmaker and shown that we can significantly improve the signal recovery on large scales. Finally, we used a principal component analysis on a COMAP feed-feed pseudo-cross-spectrum dataset to introduce a potential way forward on identifying spectra according to how contaminated they are by systematic artifacts and potentially clean them of these.

7.1 Summary

To summarize, we built an extension to the COMAP pipeline with which simulated telescope data can be fed into the analysis pipeline filters. Using this simulation pipeline we were able to compute the pipeline transfer function, which is essential to account for the filter-induced bias of the signal estimate as well as accurately estimating its errors. In particular, we found that when using the default filter parameters the ensemble-averaged estimate of the transfer function yielded a peak transfer efficiency of 85 – 90% at scales $k_{\parallel}, k_{\perp} \geq 0.1 \text{ Mpc}^{-1}$, while on the largest scales, $k_{\parallel}, k_{\perp} \leq 0.05 \text{ Mpc}^{-1}$, the signal was almost completely attenuated. Furthermore, the filters that were found to bias the signal estimates the most were the polyfilter and pointing template removal, but several other filters were found to only attenuate the signal to a modest degree. In addition, we found that the transfer functions for Lissajous scans peaked higher than in the CES equivalent, while the mixed case peaked somewhere in between the two.

Secondly, we implemented a destriper algorithm in the mapmaker by fitting the long-timescale residuals in the TOD by straight baselines. Of the tested baseline lengths, we found that the 10s baselines maximized the transfer function. Furthermore, we found the 10s baselines to produce satisfactory maps and power spectra, while at the same time resolving $\sim 20 - 25\%$ more signal at large scales compared to the highpass filtered and binned mapmaker. The peak transfer efficiency of the destriper was found to be

$\sim 95\%$. Specifically, we found that fitting baselines to all feeds simultaneously and on a per-data-split basis essentially yielded the same result. However, due to its higher memory efficiency and flexibility, the per-split destriper was found to be superior to the all-feeds destriper.

Finally, we investigated how we could use a PCA approach to identify feed-feed pseudo-cross-spectra with contamination from systematic effects. In particular, using our full COMAP data per CO field and scanning strategy to generate an FPXS dataset, we derived their corresponding eigen-spectra. We found that the structures observed in each of the first five eigen-spectra as well as their explained variance of 10–50%, per field and scanning strategy, were consistent with systematic effects. Furthermore, projecting the whole CO7 Lissajous FPXS dataset into 2D principal component space showed that the outlier clusters of spectra were not associated with any data split in particular, but rather with spectra involving feed 8 in some way. Subsequently, we found that it was possible to estimate whether an FPXS was contaminated by systematic errors by the size of the corresponding absolute PCA amplitude. Lastly, we estimated that we lose at most 20–50% signal in a few of the k -bins when subtracting out up to five principal components from an average FPXS, introducing the possibility to use this method in the COMAP pipeline to clean spectra of systematic effects at the price of some signal loss.

7.2 Future Outlooks

Even though we were able to produce interesting results in this thesis there are several points on which to improve. Firstly, our analysis on transfer functions was only done on pseudo-spectra. Hence a natural next step in the analysis would be to undo the mode coupling to get an unbiased estimate of the power spectra and transfer functions.

Next, since we at most used three realizations of the simulated signal, the ensemble-averaged transfer function could be improved by averaging over a larger dataset of simulations. This could be done by for example adding 100 different realizations of the signal to a few random obsIDs at a time. This way we would sample both simulation- and obsID-space and in doing so increase our robustness against sample variance in the simulations and different looking TODs.

Although the simulation pipeline we built is primarily designed to compute transfer functions in order to estimate the filter-induced bias of the pipeline, it is trivial to extend the pipeline to produce any kind of simulation. For instance, the simulation pipeline could be used to produce data with generated systematic effects like ground pickup or standing waves. This could in turn help testing further improvements to the COMAP pipeline filters that target each of these systematic effects.

When it comes to the destripping mapmaker we have only tested it on a comparably small dataset of 63 obsIDs. Thus, the next step from our proof-of-concept would be to test the destriper on a full dataset of obsIDs. In particular, the per-split destriper could be further tested by computing cross-spectra from the destripped maps and see how these are affected. As our current destriper implementation only fits baselines to

the TOD per frequency a future improved version of the destriper could be generalized to the frequency domain also. That way several of the filters in `12gen` could potentially be performed in the mapmaking process itself.

Lastly, when it comes to the PCA of the COMAP feed-feed pseudo-cross-spectra, we clearly saw some interesting patterns appearing in the data. However, we merely scratched the surface on how to identify and clean bad FPXS. Thus for the future, a more concrete framework needs to be build to sort the FPXS according to their systematic errors. This could be achieved by looking at the clustering patterns in the low dimensional principal component space, or by identifying and labeling the structures seen in each cross-spectrum according to their eigen-spectra, by using a machine learning algorithm. Subsequently, it would be interesting to systematically test if including previously discarded spectra, after cleaning them using their principal components, can clear some of the remaining systematic residuals from the averaged FPXS.

Bibliography

- [1] Wilson Wall. *A History of Optical Telescopes in Astronomy*. Historical & cultural astronomy. Springer International Publishing AG, Cham, 2018.
- [2] Harm J Habing. *The Birth of Modern Astronomy*. Historical & cultural astronomy. Springer International Publishing AG, Cham, 2019.
- [3] Albert Einstein. Zur allgemeinen Relativitätstheorie. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin)*, pages 778–786, January 1915.
- [4] Albert Einstein. Die Feldgleichungen der Gravitation. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin)*, pages 844–847, January 1915.
- [5] A. A. Penzias and R. W. Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *Astrophysical Journal*, 142:419–421, July 1965.
- [6] Adam G. Riess, Alexei V. Filippenko, Peter Challis, Alejandro Clocchiatti, Alan Diercks, Peter M. Garnavich, Ron L. Gilliland, Craig J. Hogan, Saurabh Jha, Robert P. Kirshner, B. Leibundgut, M. M. Phillips, David Reiss, Brian P. Schmidt, Robert A. Schommer, R. Chris Smith, J. Spyromilio, Christopher Stubbs, Nicholas B. Suntzeff, and John Tonry. Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. *Astronomical Journal*, 116(3):1009–1038, September 1998.
- [7] S. Perlmutter, G. Aldering, G. Goldhaber, R. A. Knop, P. Nugent, P. G. Castro, S. Deustua, S. Fabbro, A. Goobar, D. E. Groom, I. M. Hook, A. G. Kim, M. Y. Kim, J. C. Lee, N. J. Nunes, R. Pain, C. R. Pennypacker, R. Quimby, C. Lidman, R. S. Ellis, M. Irwin, R. G. McMahon, P. Ruiz-Lapuente, N. Walton, B. Schaefer, B. J. Boyle, A. V. Filippenko, T. Matheson, A. S. Fruchter, N. Panagia, H. J. M. Newberg, W. J. Couch, and The Supernova Cosmology Project. Measurements of Ω and Λ from 42 High-Redshift Supernovae. *Astrophysical Journal*, 517(2):565–586, June 1999.
- [8] Vera C. Rubin and Jr. Ford, W. Kent. Rotation of the Andromeda Nebula from a Spectroscopic Survey of Emission Regions. *Astrophysical Journal*, 159:379, February 1970.

- [9] Planck Collaboration, Aghanim, N., Akrami, Y., Arroja, F., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R. B., Bartolo, N., Basak, S., Battye, R., Benabed, K., Bernard, J.-P., Bersanelli, M., Bielewicz, P., Bock, J. J., Bond, J. R., Borrill, J., Bouchet, F. R., Boulanger, F., Bucher, M., Burigana, C., Butler, R. C., Calabrese, E., Cardoso, J.-F., Carron, J., Casaponsa, B., Challinor, A., Chiang, H. C., Colombo, L. P. L., Combet, C., Contreras, D., Crill, B. P., Cuttaia, F., de Bernardis, P., de Zotti, G., Delabrouille, J., Delouis, J.-M., Désert, F.-X., Di Valentino, E., Dickinson, C., Diego, J. M., Donzelli, S., Doré, O., Douspis, M., Ducout, A., Dupac, X., Efstathiou, G., Elsner, F., Enßlin, T. A., Eriksen, H. K., Falgarone, E., Fantaye, Y., Fergusson, J., Fernandez-Cobos, R., Finelli, F., Forastieri, F., Frailis, M., Franceschi, E., Frolov, A., Galeotta, S., Galli, S., Ganga, K., Génova-Santos, R. T., Gerbino, M., Ghosh, T., González-Nuevo, J., Górski, K. M., Gratton, S., Gruppuso, A., Gudmundsson, J. E., Hamann, J., Handley, W., Hansen, F. K., Helou, G., Herranz, D., Hildebrandt, S. R., Hivon, E., Huang, Z., Jaffe, A. H., Jones, W. C., Karakci, A., Keihänen, E., Keskitalo, R., Kiiveri, K., Kim, J., Kisner, T. S., Knox, L., Krachmalnicoff, N., Kunz, M., Kurki-Suonio, H., Lagache, G., Lamarre, J.-M., Langer, M., Lasenby, A., Lattanzi, M., Lawrence, C. R., Le Jeune, M., Leahy, J. P., Lesgourgues, J., Levrier, F., Lewis, A., Liguori, M., Lilje, P. B., Lilley, M., Lindholm, V., López-Caniego, M., Lubin, P. M., Ma, Y.-Z., Macías-Pérez, J. F., Maggio, G., Maino, D., Mandolesi, N., Mangilli, A., Marcos-Caballero, A., Maris, M., Martin, P. G., Martinelli, M., Martínez-González, E., Matarrese, S., Mauri, N., McEwen, J. D., Meerburg, P. D., Meinhold, P. R., Melchiorri, A., Mennella, A., Migliaccio, M., Millea, M., Mitra, S., Miville-Deschênes, M.-A., Molinari, D., Moneti, A., Montier, L., Morgante, G., Moss, A., Mottet, S., Münchmeyer, M., Natoli, P., Nørgaard-Nielsen, H. U., Oxborrow, C. A., Pagano, L., Paoletti, D., Partridge, B., Patanchon, G., Pearson, T. J., Peel, M., Peiris, H. V., Perrotta, F., Pettorino, V., Piacentini, F., Polastri, L., Polenta, G., Puget, J.-L., Rachen, J. P., Reinecke, M., Remazeilles, M., Renault, C., Renzi, A., Rocha, G., Rosset, C., Roudier, G., Rubiño-Martín, J. A., Ruiz-Granados, B., Salvati, L., Sandri, M., Savelainen, M., Scott, D., Shellard, E. P. S., Shiraishi, M., Sirignano, C., Sirri, G., Spencer, L. D., Sunyaev, R., Suur-Uski, A.-S., Tauber, J. A., Tavagnacco, D., Tenti, M., Terenzi, L., Toffolatti, L., Tomasi, M., Trombetti, T., Valiviita, J., Van Tent, B., Vibert, L., Vielva, P., Villa, F., Vittorio, N., Wandelt, B. D., Wehus, I. K., White, M., White, S. D. M., Zacchei, A., and Zonca, A. Planck 2018 results - I. Overview and the cosmological legacy of Planck. *Astronomy and Astrophysics*, 641:A1, 2020.
- [10] Ely D. Kovetz, Marco P. Viero, Adam Lidz, Laura Newburgh, Mubdi Rahman, Eric Switzer, Marc Kamionkowski, James Aguirre, Marcelo Alvarez, James Bock, J. Richard Bond, Geoffroy Bower, C. Matt Bradford, Patrick C. Breyse, Philip Bull, Tzu-Ching Chang, Yun-Ting Cheng, Dongwoo Chung, Kieran Cleary, Asantha Corray, Abigail Crites, Rupert Croft, Olivier Doré, Michael Eastwood, Andrea Ferrara, José Fonseca, Daniel Jacobs, Garrett K. Keating, Guilaine Lagache, Gunjan Lakhani, Adrian Liu, Kavilan Moodley, Norm Murray, Aurélie Pénin, Gergö Popping,

- Anthony Pullen, Dominik Reichers, Shun Saito, Ben Saliwanchik, Mario Santos, Rachel Somerville, Gordon Stacey, George Stein, Francesco Villaescusa-Navarro, Eli Visbal, Amanda Weltman, Laura Wolz, and Micheal Zemcov. Line-Intensity Mapping: 2017 Status Report. *arXiv e-prints*, page arXiv:1709.09066, September 2017.
- [11] Sean Carroll. *Spacetime and geometry : An Introduction to General Relativity*. Pearson, Harlow, Pearson New international edition. edition, 2014.
- [12] Scott Dodelson. *Modern cosmology*. Academic Press, Amsterdam, 2003.
- [13] Peter Schneider. *Extragalactic Astronomy and Cosmology : An Introduction*. Springer Berlin Heidelberg : Imprint: Springer, Berlin, Heidelberg, 2nd ed. 2015. edition, 2015.
- [14] J. P Willis, R. E. A Canning, E. S Noordeh, S. W Allen, A. L King, A Mantz, R. G Morris, S. A Stanford, and G Brammer. Spectroscopic confirmation of a mature galaxy cluster at a redshift of 2. *Nature (London)*, 577(7788):39–41, 2020.
- [15] Takuya Hashimoto, Nicolas Laporte, Ken Mawatari, Richard S. Ellis, Akio K. Inoue, Erik Zackrisson, Guido Roberts-Borsani, Wei Zheng, Yoichi Tamura, Franz E. Bauer, Thomas Fletcher, Yuichi Harikane, Bunyo Hatsukade, Natsuki H. Hayatsu, Yuichi Matsuda, Hiroshi Matsuo, Takashi Okamoto, Masami Ouchi, Roser Pelló, Claes-Erik Rydberg, Ikkoh Shimizu, Yoshiaki Taniguchi, Hideki Umehata, and Naoki Yoshida. The onset of star formation 250 million years after the Big Bang. *Nature*, 557(7705):392–395, May 2018.
- [16] Daniel J. Eisenstein, Idit Zehavi, David W. Hogg, Roman Scoccimarro, Michael R. Blanton, Robert C. Nichol, Ryan Scranton, Hee-Jong Seo, Max Tegmark, Zheng Zheng, Scott F. Anderson, Jim Annis, Neta Bahcall, Jon Brinkmann, Scott Burles, Francisco J. Castander, Andrew Connolly, Istvan Csabai, Mamoru Doi, Masataka Fukugita, Joshua A. Frieman, Karl Glazebrook, James E. Gunn, John S. Hendry, Gregory Hennessy, Zeljko Ivezić, Stephen Kent, Gillian R. Knapp, Huan Lin, Yeong-Shang Loh, Robert H. Lupton, Bruce Margon, Timothy A. McKay, Avery Meiksin, Jeffery A. Munn, Adrian Pope, Michael W. Richmond, David Schlegel, Donald P. Schneider, Kazuhiro Shimasaku, Christopher Stoughton, Michael A. Strauss, Mark SubbaRao, Alexander S. Szalay, István Szapudi, Douglas L. Tucker, Brian Yanny, and Donald G. York. Detection of the Baryon Acoustic Peak in the Large-Scale Correlation Function of SDSS Luminous Red Galaxies. *Astrophysical Journal*, 633(2):560–574, November 2005.
- [17] Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak, R. Battye, K. Benabed, J. P. Bernard, M. Bersanelli, P. Bielewicz, J. J. Bock, J. R. Bond, J. Borrill, F. R. Bouchet, F. Boulanger, M. Bucher, C. Burigana, R. C. Butler, E. Calabrese, J. F. Cardoso, J. Carron, A. Challinor, H. C. Chiang, J. Chluba,

- L. P. L. Colombo, C. Combet, D. Contreras, B. P. Crill, F. Cuttaia, P. de Bernardis, G. de Zotti, J. Delabrouille, J. M. Delouis, E. Di Valentino, J. M. Diego, O. Doré, M. Douspis, A. Ducout, X. Dupac, S. Dusini, G. Efstathiou, F. Elsner, T. A. Enßlin, H. K. Eriksen, Y. Fantaye, M. Farhang, J. Fergusson, R. Fernandez-Cobos, F. Finelli, F. Forastieri, M. Frailis, A. A. Fraisse, E. Franceschi, A. Frolov, S. Galeotta, S. Galli, K. Ganga, R. T. Génova-Santos, M. Gerbino, T. Ghosh, J. González-Nuevo, K. M. Górski, S. Gratton, A. Gruppuso, J. E. Gudmundsson, J. Hamann, W. Handley, F. K. Hansen, D. Herranz, S. R. Hildebrandt, E. Hivon, Z. Huang, A. H. Jaffe, W. C. Jones, A. Karakci, E. Keihänen, R. Kesitalo, K. Kiiveri, J. Kim, T. S. Kisner, L. Knox, N. Krachmalnicoff, M. Kunz, H. Kurki-Suonio, G. Lagache, J. M. Lamarre, A. Lasenby, M. Lattanzi, C. R. Lawrence, M. Le Jeune, P. Lemos, J. Lesgourgues, F. Levrier, A. Lewis, M. Liguori, P. B. Lilje, M. Lilley, V. Lindholm, M. López-Cañiego, P. M. Lubin, Y. Z. Ma, J. F. Macías-Pérez, G. Maggio, D. Maino, N. Mandolesi, A. Mangilli, A. Marcos-Caballero, M. Maris, P. G. Martin, M. Martinelli, E. Martínez-González, S. Matarrese, N. Mauri, J. D. McEwen, P. R. Meinhold, A. Melchiorri, A. Mennella, M. Migliaccio, M. Millea, S. Mitra, M. A. Miville-Deschênes, D. Molinari, L. Montier, G. Morgante, A. Moss, P. Natoli, H. U. Nørgaard-Nielsen, L. Pagano, D. Paoletti, B. Partridge, G. Patanchon, H. V. Peiris, F. Perrotta, V. Pettorino, F. Piacentini, L. Polastri, G. Polenta, J. L. Puget, J. P. Rachen, M. Reinecke, M. Remazeilles, A. Renzi, G. Rocha, C. Rosset, G. Roudier, J. A. Rubiño-Martín, B. Ruiz-Granados, L. Salvati, M. Sandri, M. Savelainen, D. Scott, E. P. S. Shellard, C. Sirignano, G. Sirri, L. D. Spencer, R. Sunyaev, A. S. Suur-Uski, J. A. Tauber, D. Tavagnacco, M. Tenti, L. Toffolatti, M. Tomasi, T. Trombetti, L. Valenziano, J. Valiviita, B. Van Tent, L. Vibert, P. Vielva, F. Villa, N. Vittorio, B. D. Wandelt, I. K. Wehus, M. White, S. D. M. White, A. Zacchei, and A. Zonca. Planck 2018 results. VI. Cosmological parameters. *Astronomy and Astrophysics*, 641:A6, September 2020.
- [18] Dongwoo T Chung, Marco P Viero, Sarah E Church, Risa H Wechsler, Marcelo A Alvarez, J. Richard Bond, Patrick C Breyse, Kieran A Cleary, Hans K Eriksen, Marie K Foss, Joshua O Gundersen, Stuart E Harper, Håvard T. Ihle, Laura C Keating, Norman Murray, Hamsa Padmanabhan, George F Stein, and Ingunn K Wehus. Cross-correlating Carbon Monoxide Line-intensity Maps with Spectroscopic and Photometric Galaxy Surveys. *Astrophysical Journal*, 872(2):186, 2019.
- [19] Brant E. Robertson, Richard S. Ellis, Steven R. Furlanetto, and James S. Dunlop. Cosmic Reionization and Early Star-forming Galaxies: A Joint Analysis of New Constraints from Planck and the Hubble Space Telescope. *Astrophysical Journal Letters*, 802(2):L19, April 2015.
- [20] Shaghayegh Parsa, James S. Dunlop, and Ross J. McLure. No evidence for a significant AGN contribution to cosmic hydrogen reionization. *Monthly Notices of the Royal Astronomical Society*, 474(3):2904–2923, March 2018.
- [21] R. J. Bouwens, G. D. Illingworth, P. A. Oesch, J. Caruana, B. Holwerda, R. Smit,

- and S. Wilkins. Reionization After Planck: The Derived Growth of the Cosmic Ionizing Emissivity Now Matches the Growth of the Galaxy UV Luminosity Density. *Astrophysical Journal*, 811(2):140, October 2015.
- [22] David R. DeBoer, Aaron R. Parsons, James E. Aguirre, Paul Alexander, Zaki S. Ali, Adam P. Beardsley, Gianni Bernardi, Judd D. Bowman, Richard F. Bradley, Chris L. Carilli, Carina Cheng, Eloy de Lera Acedo, Joshua S. Dillon, Aaron Ewall-Wice, Gcobisa Fadana, Nicolas Fagnoni, Randall Fritz, Steve R. Furlanetto, Brian Glendenning, Bradley Greig, Jasper Grobbelaar, Bryna J. Hazelton, Jacqueline N. Hewitt, Jack Hickish, Daniel C. Jacobs, Austin Julius, MacCalvin Kariseb, Saul A. Kohn, Telalo Lekalake, Adrian Liu, Anita Loots, David MacMahon, Lourence Malan, Cresshim Malgas, Matthys Maree, Zachary Martinot, Nathan Mathison, Eunice Matsetela, Andrei Mesinger, Miguel F. Morales, Abraham R. Neben, Nipanjana Patra, Samantha Pieterse, Jonathan C. Pober, Nima Razavi-Ghods, Jon Ringuette, James Robnett, Kathryn Rosie, Raddwine Sell, Craig Smith, Angelo Syce, Max Tegmark, Nithyanandan Thyagarajan, Peter K. G. Williams, and Haoxuan Zheng. Hydrogen Epoch of Reionization Array (HERA). *Publications of the Astronomical Society of the Pacific*, 129(974):045001, March 2017.
- [23] L. Koopmans, J. Pritchard, G. Mellema, J. Aguirre, K. Ahn, R. Barkana, I. van Bemmell, G. Bernardi, A. Bonaldi, F. Briggs, A. G. de Bruyn, T. C. Chang, E. Chapman, X. Chen, B. Ciardi, P. Dayal, A. Ferrara, A. Fialkov, F. Fiore, K. Ichiki, I. T. Illiev, S. Inoue, V. Jelic, M. Jones, J. Lazio, U. Maio, S. Majumdar, K. J. Mack, A. Mesinger, M. F. Morales, A. Parsons, U. L. Pen, M. Santos, R. Schneider, B. Semelin, R. S. de Souza, R. Subrahmanyam, T. Takeuchi, H. Vedantham, J. Wagg, R. Webster, S. Wyithe, K. K. Datta, and C. Trott. The Cosmic Dawn and Epoch of Reionisation with SKA. In *Advancing Astrophysics with the Square Kilometre Array (AASKA14)*, page 1, 2015.
- [24] N. F. Cothard, S. K. Choi, C. J. Duell, T. Herter, J. Hubmayr, J. McMahon, M. D. Niemack, T. Nikola, C. Sierra, G. J. Stacey, E. M. Vavagiakis, E. J. Wollack, and B. Zou. The Design of the CCAT-prime Epoch of Reionization Spectrometer Instrument. *Journal of low temperature physics*, 199(3-4):898–907, 2020.
- [25] S. K. Choi, J. Austermann, K. Basu, N. Battaglia, F. Bertoldi, D. T. Chung, N. F. Cothard, S. Duff, C. J. Duell, P. A. Gallardo, J. Gao, T. Herter, J. Hubmayr, M. D. Niemack, T. Nikola, D. Riechers, K. Rossi, G. J. Stacey, J. R. Stevens, E. M. Vavagiakis, M. Vissers, and S. Walker. Sensitivity of the Prime-Cam Instrument on the CCAT-Prime Telescope. *Journal of Low Temperature Physics*, 199(3-4):1089–1097, 2020.
- [26] Z. Staniszewski, J. J. Bock, C. M. Bradford, J. Brevik, A. Cooray, Y. Gong, S. Hailey-Dunsheath, R. O’Brien, M. Santos, E. Shirokoff, M. Silva, and M. Zemcov. The Tomographic Ionized-Carbon Mapping Experiment (TIME) CII Imaging Spectrometer. *Journal of low temperature physics*, 176(5):767–772, 2014.

- [27] A. T. Crites, J. J. Bock, C. M. Bradford, T. C. Chang, A. R. Cooray, L. Duband, Y. Gong, S. Hailey-Dunsheath, J. Hunacek, P. M. Koch, C. T. Li, R. C. O'Brient, T. Prouve, E. Shirokoff, M. B. Silva, Z. Staniszewski, B. Uzgil, and M. Zemcov. The TIME-Pilot intensity mapping experiment. In Wayne S. Holland and Jonas Zmuidzinas, editors, *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII*, volume 9153 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 91531W, 2014.
- [28] Concerto Collaboration, P. Ade, M. Aravena, E. Barria, A. Beelen, A. Benoit, M. Béthermin, J. Bounmy, O. Bourrion, G. Bres, C. De Breuck, M. Calvo, Y. Cao, A. Catalano, F. X. Désert, C. A. Durán, A. Fasano, T. Fenouillet, J. Garcia, G. Garde, J. Goupy, C. Groppi, C. Hoarau, G. Lagache, J. C. Lambert, J. P. Leggeri, F. Levy-Bertrand, J. Macías-Pérez, H. Mani, J. Marpaud, P. Mauskopf, A. Monfardini, G. Pisano, N. Ponthieu, L. Prieur, S. Roni, S. Roudier, D. Tourres, and C. Tucker. A wide field-of-view low-resolution spectrometer at APEX: Instrument design and scientific forecast. *Astronomy and Astrophysics*, 642:A60, 2020.
- [29] Tony Y. Li, Risa H. Wechsler, Kiruthika Devaraj, and Sarah E. Church. Connecting CO Intensity Mapping to Molecular Gas and Star Formation in the Epoch of Galaxy Assembly. *Astrophysical Journal*, 817(2):169, 2016.
- [30] Garrett K. Keating, Geoffrey C. Bower, Daniel P. Marrone, David R. DeBoer, Carl Heiles, Tzu-Ching Chang, John E. Carlstrom, Christopher H. Greer, David Hawkins, James W. Lamb, Erik Leitch, Amber D. Miller, Stephen Muchovej, and David P. Woody. First Results from COPSS: The CO Power Spectrum Survey. *Astrophysical Journal*, 814(2):140, 2015.
- [31] Garrett K. Keating, Daniel P. Marrone, Geoffrey C. Bower, Erik Leitch, John E. Carlstrom, and David R. DeBoer. COPSS II: The Molecular Gas Content of Ten Million Cubic Megaparsecs at Redshift $z \sim 3$. *Astrophysical Journal*, 830(1):34, 2016.
- [32] Garrett K Keating, Daniel P Marrone, Geoffrey C Bower, and Ryan P Keenan. An Intensity Mapping Detection of Aggregate CO Line Emission at 3 mm. *The Astrophysical Journal*, 901(2), 2020.
- [33] Kieran Cleary, Marie-Anne Bigot-Sazy, Dongwoo Chung, Sarah E. Church, Clive Dickinson, Hans Eriksen, todd gaier, Paul Goldsmith, Joshua O. Gundersen, Stuart Harper, Andrew I. Harris, James Lamb, Tony Li, Ryan Munroe, Timothy J. Pearson, Anthony C. S. Readhead, Risa H. Wechsler, Ingunn Kathrine Wehus, and David Woody. The CO Mapping Array Pathfinder (COMAP). In *American Astronomical Society Meeting Abstracts #227*, volume 227 of *American Astronomical Society Meeting Abstracts*, page 426.06, January 2016.
- [34] Marie K. Foss, Håvard T. Ihle, Jowita Borowska, Morgan Catha, Kieran A. Cleary, Hans Kristian Eriksen, Stuard Harper, James W. Lamb, Jonas G. S. Lunde, Nils-Ole Stutzer, Duncan J. Watts, Ingunn K. Wehus, David P. Woody, and The

- COMAP Collaboration. First Season COMAP Results: CO Data Processing. To be submitted, 2020.
- [35] Olivier Doré, Jamie Bock, Matthew Ashby, Peter Capak, Asantha Cooray, Roland de Putter, Tim Eifler, Nicolas Flagey, Yan Gong, Salman Habib, Katrin Heitmann, Chris Hirata, Woong-Seob Jeong, Raj Katti, Phil Korngut, Elisabeth Krause, Dae-Hee Lee, Daniel Masters, Phil Mauskopf, Gary Melnick, Bertrand Mennesson, Hien Nguyen, Karin Öberg, Anthony Pullen, Alvise Raccanelli, Roger Smith, Yong-Seon Song, Volker Tolls, Steve Unwin, Tejaswi Venumadhav, Marco Viero, Mike Werner, and Mike Zemcov. Cosmology with the SPHEREX All-Sky Spectral Survey. *arXiv e-prints*, page arXiv:1412.4872, 2014.
- [36] Yan Gong, Asantha Cooray, Marta B. Silva, Michael Zemcov, Chang Feng, Mario G. Santos, Olivier Dore, and Xuelei Chen. Intensity Mapping of $H\alpha$, $H\beta$, [OII], and [OIII] Lines at $z < 5$. *Astrophysical Journal*, 835(2):273, 2017.
- [37] Ashley J. Ross, Florian Beutler, Chia-Hsun Chuang, Marcos Pellejero-Ibanez, Hee-Jong Seo, Mariana Vargas-Magaña, Antonio J. Cuesta, Will J. Percival, Angela Burden, Ariel G. Sánchez, Jan Niklas Grieb, Beth Reid, Joel R. Brownstein, Kyle S. Dawson, Daniel J. Eisenstein, Shirley Ho, Francisco-Shu Kitaura, Robert C. Nichol, Matthew D. Olmstead, Francisco Prada, Sergio A. Rodríguez-Torres, Shun Saito, Salvador Salazar-Albornoz, Donald P. Schneider, Daniel Thomas, Jeremy Tinker, Rita Tojeiro, Yuting Wang, Martin White, and Gong-bo Zhao. The clustering of galaxies in the completed SDSS-III Baryon Oscillation Spectroscopic Survey: observational systematics and baryon acoustic oscillations in the correlation function. *Monthly Notices of the Royal Astronomical Society*, 464(1):1168–1191, January 2017.
- [38] Philippe Berger, Laura B. Newburgh, Mandana Amiri, Kevin Bandura, Jean-François Cliche, Liam Connor, Meiling Deng, Nolan Denman, Matt Dobbs, Mateus Fandino, Adam J. Gilbert, Deborah Good, Mark Halpern, David Hanna, Adam D. Hincks, Gary Hinshaw, Carolin Höfer, Andre M. Johnson, Tom L. Landecker, Kiyoshi W. Masui, Juan Mena Parra, Niels Oppermann, Ue-Li Pen, Jeffrey B. Peterson, Andre Recnik, Timothy Robishaw, J. Richard Shaw, Seth Siegel, Kris Sigurdson, Kendrick Smith, Emilie Storer, Ian Tretyakov, Kwinten Van Gassen, Keith Vanderlinde, and Donald Wiebe. Holographic beam mapping of the CHIME pathfinder array. In Helen J. Hall, Roberto Gilmozzi, and Heather K. Marshall, editors, *Ground-based and Airborne Telescopes VI*, volume 9906 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 99060D, 2016.
- [39] L. B. Newburgh, K. Bandura, M. A. Bucher, T. C. Chang, H. C. Chiang, J. F. Cliche, R. Davé, M. Dobbs, C. Clarkson, K. M. Ganga, T. Gogo, A. Gumba, N. Gupta, M. Hilton, B. Johnstone, A. Karastergiou, M. Kunz, D. Lokhorst, R. Maartens, S. Macpherson, M. Mdlalose, K. Moodley, L. Ngwenya, J. M. Parra, J. Peterson, O. Recnik, B. Saliwanchik, M. G. Santos, J. L. Sievers, O. Smirnov, P. Stronkhorst, R. Taylor, K. Vanderlinde, G. Van Vuuren, A. Weltman, and

- A. Witzemann. HIRAX: a probe of dark energy and radio transients. In Helen J. Hall, Roberto Gilmozzi, and Heather K. Marshall, editors, *Ground-based and Airborne Telescopes VI*, volume 9906 of *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, page 99065X, 2016.
- [40] Andrej Obuljen, Emanuele Castorina, Francisco Villaescusa-Navarro, and Matteo Viel. High-redshift post-reionization cosmology with 21cm intensity mapping. *Journal of Cosmology and Astroparticle Physics*, 2018(5):004, May 2018.
- [41] Emanuele Castorina and Martin White. Measuring the growth of structure with intensity mapping surveys. *Journal of Cosmology and Astroparticle Physics*, 2019(6):025, 2019.
- [42] CHIME/FRB Collaboration, M. Amiri, K. Bandura, P. Berger, M. Bhardwaj, M. M. Boyce, P. J. Boyle, C. Brar, M. Burhanpurkar, P. Chawla, J. Chowdhury, J. F. Cliche, M. D. Cranmer, D. Cubranic, M. Deng, N. Denman, M. Dobbs, M. Fandino, E. Fonseca, B. M. Gaensler, U. Giri, A. J. Gilbert, D. C. Good, S. Guliani, M. Halpern, G. Hinshaw, C. Höfer, A. Josephy, V. M. Kaspi, T. L. Landecker, D. Lang, H. Liao, K. W. Masui, J. Mena-Parra, A. Naidu, L. B. Newburgh, C. Ng, C. Patel, U. L. Pen, T. Pinsonneault-Marotte, Z. Pleunis, M. Rafiei Ravandi, S. M. Ransom, A. Renard, P. Scholz, K. Sigurdson, S. R. Siegel, K. M. Smith, I. H. Stairs, S. P. Tendulkar, K. Vanderlinde, and D. V. Wiebe. The CHIME Fast Radio Burst Project: System Overview. *Astrophysical Journal*, 863(1):48, 2018.
- [43] CHIME/Pulsar Collaboration, M. Amiri, K. M. Bandura, P. J. Boyle, C. Brar, J. F. Cliche, K. Crowter, D. Cubranic, P. B. Demorest, N. T. Denman, M. Dobbs, F. Q. Dong, M. Fandino, E. Fonseca, D. C. Good, M. Halpern, C. Höfer, V. M. Kaspi, T. L. Landecker, H. H. Lin, K. W. Masui, J. W. McKee, J. Mena-Parra, B. W. Meyers, D. Michilli, A. Naidu, L. Newburgh, C. Ng, C. Patel, T. Pinsonneault-Marotte, S. M. Ransom, A. Renard, P. Scholz, J. R. Shaw, A. E. Sikora, I. H. Stairs, C. M. Tan, S. P. Tendulkar, I. Tretyakov, K. Vanderlinde, H. Wang, and X. Wang. The CHIME Pulsar Project: System Overview. *arXiv e-prints*, page arXiv:2008.05681, 2020.
- [44] Yasuo Fukui and Akiko Kawamura. Molecular Clouds in Nearby Galaxies. *Annual review of astronomy and astrophysics*, 48(1):547–580, 2010.
- [45] Norbert S Schulz. *The Formation and Early Evolution of Stars: From Dust to Stars and Planets*. Astronomy and Astrophysics Library. Springer-Verlag, Berlin, Heidelberg, 2. aufl. edition, 2012.
- [46] Wolfgang Demtröder. *Atoms, Molecules and Photons : An Introduction to Atomic-, Molecular- and Quantum Physics*. Graduate Texts in Physics. Springer Berlin Heidelberg : Imprint: Springer, Berlin, Heidelberg, 3rd ed. 2018. edition, 2018.

- [47] Il-Gyo Jeong, Bon-Chul Koo, Wan-Kee Cho, Carsten Kramer, Jürgen Stutzki, and Do-Young Byun. CO J=1-0 and J=2-1 Line Observations of the Molecular Cloud - Blocked Supernova Remnant 3C434.1. *Astrophysical Journal*, 770(2), 2013.
- [48] Patrick C. Breyse and Mubdi Rahman. Feeding cosmic star formation: exploring high-redshift molecular gas with CO intensity mapping. *Monthly notices of the Royal Astronomical Society*, 468(1):741–750, June 2017.
- [49] Kieran Cleary. The CO Mapping Array Pathfinder (COMAP). https://www.cfa.harvard.edu/~afialkov/talk_Cleary.pdf, 2018. Accessed: 12.16.2020.
- [50] The History of OVRO. <https://www.ovro.caltech.edu/index.php?page=history>. Accessed: 12.14.2020.
- [51] Jacob W. M. Baars and Hans J. Kärcher. *Radio Telescope Reflectors*, volume 447. Springer, 2018.
- [52] Robert B. Leighton. A 10-Meter Telescope for Millimeter and Sub-Millimeter Astronomy. Technical report, California Institute of Technology, 1977.
- [53] D Woody, D Vail, and W Schaal. Design, construction, and performance of the Leighton 10.4-m-diameter radio telescopes. *Proceedings of the IEEE*, 82(5):673–686, 1994.
- [54] ESA earthnet online - RADAR and SAR Glossary. <https://earth.esa.int/handbooks/asar/CNTR5-2.html>, 24. July 2014. Accessed: 05.07.2021.
- [55] Ian S. McLean, editor. *Planets, Stars and Stellar Systems : Volume 1: Telescopes and Instrumentation*. Springer Netherlands : Imprint: Springer, Dordrecht, 1st ed. 2013. edition, 2013.
- [56] Jingquan Cheng. *The Principles of Astronomical Telescope Design*. Springer, New York, NY, 2009.
- [57] Andrew Hunt and Alan Wright. Complex, off-axis sidelobes of a radio telescope caused by feed-support legs. *Monthly notices of the Royal Astronomical Society*, 258(1):217–224, 1992.
- [58] James W. Lamb. Chromaticity of COMAP beams. https://sites.astro.caltech.edu/CRAL1-wiki/images/b/b8/Chromaticity_of_COMAP_beams%2C_v5.pdf, 2018. Accessed: 12.16.2020.
- [59] G. J. Hill, K. Gebhardt, E. Komatsu, N. Drory, P. J. MacQueen, J. Adams, G. A. Blanc, R. Koehler, M. Rafal, M. M. Roth, A. Kelz, C. Gronwall, R. Ciardullo, and D. P. Schneider. The Hobby-Eberly Telescope Dark Energy Experiment (HET-DEX): Description and Early Pilot Survey Results. In T. Kodama, T. Yamada, and K. Aoki, editors, *Panoramic Views of Galaxy Formation and Evolution*, volume 399 of *Astronomical Society of the Pacific Conference Series*, page 115, October 2008.

- [60] C. Papovich, H. V. Shipley, N. Mehrrens, C. Lanham, M. Lacy, R. Ciardullo, S. L. Finkelstein, R. Bassett, P. Behroozi, G. A. Blanc, R. S. de Jong, D. L. DePoy, N. Drory, E. Gawiser, K. Gebhardt, C. Gronwall, G. J. Hill, U. Hopp, S. Jogee, L. Kawinwanichakij, J. L. Marshall, E. McLinden, E. Mentuch Cooper, R. S. Somerville, M. Steinmetz, K. V. Tran, S. Tuttle, M. Viero, R. Wechsler, and G. Zeimann. The Spitzer-HETDEX Exploratory Large-area Survey. *Astrophysical Journal Supplement Series*, 224(2):28, June 2016.
- [61] T. J. Waskett, B. Sibthorpe, M. J. Griffin, and P. F. Chaniai. Determining the optimum scan map strategy for Herschel-SPIRE using the SPIRE photometer simulator. *Monthly Notices of the Royal Astronomical Society*, 381(4):1583–1590, November 2007.
- [62] Stuard Harper. COMAP Memo XX:Scan Strategies II. https://sites.astro.caltech.edu/CRAL1-wiki/images/5/58/Scan_strategies_II.pdf, 2018. Accessed: 12.17.2020.
- [63] Thomas Wilson, Kristen Rohlf, and Susanne Hüttemeister. *Tools of Radio Astronomy*. Astronomy and Astrophysics Library. Springer, Berlin, Heidelberg, 6th ed. 2013 edition, 2013.
- [64] D. J Fixsen. The Temperature of the Cosmic Microwave Background. *Astrophysical Journal*, 707(2):916–920, 2009.
- [65] Erika Böhm-Vitense. *Introduction to stellar astrophysics : vol.2 : Stellar atmospheres*, volume vol.2. Cambridge University Press, Cambridge, 1989.
- [66] Max Tegmark. How to Make Maps from Cosmic Microwave Background Data without Losing Information. *Astrophysical Journal*, 480(2):L87–L90, 1997.
- [67] A. Heavens. *Statistical Techniques in Cosmology*, 2016.
- [68] E Keihänen, H Kurki-Suonio, and T Poutanen. Madam - a map-making method for CMB experiments. *Monthly notices of the Royal Astronomical Society*, 360(1):390–400, 2005.
- [69] H. T. Ihle, D. Chung, G. Stein, M. Alvarez, J. R. Bond, P. C. Breysse, K. A. Cleary, H. K. Eriksen, M. K. Foss, J. O. Gundersen, S. Harper, N. Murray, H. Padmanabhan, M. P. Viero, I. K. Wehus, and COMAP Collaboration. Joint Power Spectrum and Voxel Intensity Distribution Forecast on the CO Luminosity Function with COMAP. *Astrophysical Journal*, 871(1):75, 2019.
- [70] Željko Ivezić, Andrew J. Connolly, Jacob T. VanderPlas, and Alexander Gray. Statistics, data mining, and machine learning in astronomy: a practical Python guide for the analysis of survey data, by Željko Ivezić, Andrew J. Connolly, Jacob T. VanderPlas and Alexander Gray. *Contemporary Physics*, 58(1):99–99, 2017.

- [71] Peter S. Behroozi, Risa H. Wechsler, and Charlie Conroy. On the Lack of Evolution in Galaxy Star Formation Efficiency. *Astrophysical Journal Letters*, 762(2):L31, 2013.
- [72] Peter S. Behroozi, Risa H. Wechsler, and Charlie Conroy. The Average Star Formation Histories of Galaxies in Dark Matter Halos from $z = 0-8$. *Astrophysical Journal*, 770(1):57, 2013.
- [73] Håvard T. Ihle, Jowita Borowska, Nils-Ole Stutzer, Jonas G. S. Lunde, and The COMAP Collaboration. First Season COMAP Results: Power spectrum methodology and preliminary data quality assessment. To be submitted, 2021.
- [74] Eric Hivon, Krzysztof M. Górski, C. Barth Netterfield, Brendan P. Crill, Simon Prunet, and Frode Hansen. MASTER of the Cosmic Microwave Background Anisotropy Power Spectrum: A Fast Method for Statistical Analysis of Large and Complex Cosmic Microwave Background Data Sets. *Astrophysical Journal*, 567(1):2–17, 2002.
- [75] Thibaut Louis, Sigurd Naess, Xavier Garrido, and Anthony Challinor. Fast computation of angular power spectra and covariances of high-resolution cosmic microwave background maps using the toeplitz approximation. *Physical review. D, Particles, fields, gravitation, and cosmology*, 102:123538, 2020.
- [76] Dongwoo Chung, Patrick Breysse, Kieran Cleary, Håvard T. Ihle, Hamsa Padmanabhan, Marta Silva, and The COMAP Collaboration. Early Science Results from the CO Mapping Array Pathfinder: V. CO Emission Models at $z \sim 3$. To be submitted, 2021.
- [77] M. R. Calabretta and E. W. Greisen. Representations of celestial coordinates in FITS. *Astronomy and Astrophysics*, 395:1077–1122, December 2002.
- [78] E Keihänen, H Kurki-Suonio, T Poutanen, D Maino, and C Burigana. A maximum likelihood approach to the destriping technique. *Astronomy and astrophysics (Berlin)*, 428(1):287–298, 2004.
- [79] Jonathan R Shewchuk. An Introduction to the Conjugate Gradient Method Without the Agonizing Pain. Technical report, USA, 1994.
- [80] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, USA, 3 edition, 2007.
- [81] Nikos Alexandris, Sandeep Gupta, and Nikos Koutsias. Remote sensing of burned areas via PCA, Part 1; centering, scaling and EVD vs SVD. *Open geospatial data, software and standards*, 2(1):1–11, 2017.
- [82] David C Lay. *Linear algebra and its applications*. Pearson, Harlow, 5th ed. edition, 2016.