Making Maps with ACT data

by

Felipe Rojas Aracena

Thesis report presented to the Physics Faculty of Pontificia Universidad Católica de Chile for the Degree of Master in Sciences, Astrophysics

Supervisor : Prof. Rolando Dünner
Correctors : Prof. Leopoldo Infante
            : Prof. Felipe Barrientos

August 2014
Santiago, Chile
Acknowledgements

I want to thank my family and Javiera for their support during all this time. This achievement is also yours.
I also want to thank Rolando for helping me, for his patience and for believing in me.
Finally, I thank all who have contributed to my professional development.
Abstract

In this work, we present a semi-independent mapper to produce CMB and point sources maps using data from the Atacama Cosmology Telescope, in particular, data from the 280 GHz band which has not been successfully processed until now. The data corresponds to \( \sim 30 \) square degrees area in the range \([4 \text{ hr}, 5 \text{ hr}]\) in Right Ascension and \([-51^\circ, -54.5^\circ]\) in Declination. The code was designed to run in a parallel environment, being able to work with several data files at the same time. We use a noise model based on Singular Value Decomposition (SVD) modes identification and frequency space downweightening. The modes we found are the dominant singular values of the data correlation matrix in two Time Ordered Data (TOD) frequency ranges. Then, this modes are removed from the data to compute the mean power in frequency bins; the same is done with modes expanded in time-domain. This help us to build a noise matrix per frequency bin, including the cleaned data (white noise) plus correlations between detectors given by the eigenvectors obtained during the singular value decomposition. This model is included in the map making equation to downweight the noisy modes present in the time-streams. We started by running our code with data from the 148 GHz band, season 2008, in order to compare our results with previous ACT results. The point sources fluxes in 148 GHz band are consistent with previous results, and we see an improvement in signal-to-noise ratios up to \( \sim 35\% \). The situation is different for 218 GHz fluxes, same season, with differences of \( \sim 15\% \). We found white noise levels of 28.9 \( \mu \text{K-arcmin} \) for 148 GHz and 46.1 \( \mu \text{K-arcmin} \) for 218 GHz. We explored a cross-correlation calibration technique; we calibrate our 148 GHz map to ACT, finding a calibration factor \( \alpha = 0.956 \pm 0.025 \). We also mapped the 280 GHz band. The noise level is 115.6 \( \mu \text{K-arcmin} \) for season 2008. For seasons 2009 and 2010, the white noise levels are 105.3 \( \mu \text{K-arcmin} \) and 84.6 \( \mu \text{K-arcmin} \), respectively. These 280 GHz maps correspond to the first ones of the collaboration.
# Contents

1 Introduction 3
   1.1 Cosmic Microwave Background 3
   1.2 Atacama Cosmology Telescope 4
   1.3 Signal description 5
      1.3.1 CMB 5
      1.3.2 Sunyaev-Zel’dovich Effect 8
      1.3.3 Point sources 9
   1.4 Chile ACT Ultradeep Survey 9

2 The Data 11
   2.1 Data description 11
   2.2 Data selection 14
   2.3 Data preprocessing 15

3 Noise sources 17
   3.1 Sources of contamination in ACT data 17
   3.2 Singular Value Decomposition 21

4 Map making 25
   4.1 Map making equation 25
   4.2 Noise model 26
      4.2.1 Dealing with gaps in data 28
   4.3 Preconditioned Conjugate Gradient 29
   4.4 Implementation 32
5 Results ................................................................. 35
  5.1 The Maps ...................................................... 35
  5.2 Noise ........................................................... 39
    5.2.1 Maps preprocessing ..................................... 39
    5.2.2 Noise estimation ........................................ 43
  5.3 Transfer Function and convergence ....................... 48
  5.4 Cross-calibration to ACT .................................... 49
  5.5 CMB signal ................................................... 52

6 Discussion and Conclusions .................................... 54

References ......................................................... 56
Chapter 1

Introduction

1.1 Cosmic Microwave Background

The Cosmic Microwave Background (CMB) is the dominant radiation field of the Universe. This radiation had been predicted by [1] and by [7] and was finally measured by [25] in Bell Labs when they were experimenting with an antenna to study galactic radio emission.

In the early Universe the temperature was very high (around $10^8$ Kelvin, $10^3$ seconds after the Big Bang) and the radiation was coupled with matter forming a primordial soup. As the Universe expanded, the temperature decreased, allowing electrons to combine forming neutral atoms. This is known as the epoch of recombination. The lower electron density reduced photon scattering, increasing photon’s mean free path. This happened very quickly forming a last scattering surface (LSS), with redshift $z \sim 1100$ (around 380,000 years after the Big Bang) when the temperature was around $T \sim 3000$ K. The radiation from the LSS is what constitutes the CMB today. Since the photons were strongly coupled to matter, the matter density fluctuations produced temperature fluctuations in the CMB temperature. The first measurement of these fluctuations was made by COBE in 1992, which measured $\Delta T/T \sim 10^{-5}$ for scales larger than seven degrees [27].

The CMB temperature fluctuations can be classified as primary anisotropies, if they are a consequence of primordial photon-baryon fluid density fluctuations, and as secondary anisotropies, if produced by foreground distortions imprinted in CMB photons. These dis-
tortions includes effects like: integrated Sachs-Wolfe effect, due to photons moving through a time-varying linear gravitational potential; gravitational lensing, which redistribute the power toward small scales; Rees-Sciama effect due to photons moving though a non-linear gravitational potential; Sunyaev-Zel’dovich effect (SZE) due to inverse Compton scattering of CMB photons by hot gas in clusters producing a spectral distortion away from the blackbody spectrum, as described by Sunyaev and Zel’dovich [33]. Recently, experiments like ACT and SPT have measured this effect [15], [28].

1.2 Atacama Cosmology Telescope

The Atacama Cosmology Telescope is a six meter telescope located at 5200 meters over the sea level in the Atacama Desert in northern Chile. The location is characterized by low precipitable water vapor (PWV), an important condition for millimeter experiments. The main goal is to map the CMB at arcminute scales and characterize sources naturally selected, such as millimeter galaxies and galaxy cluster through their SZ signature. The observations will help to test the predictions of the ΛCDM cosmological model through precise measurements of the damping tail ($500 \lesssim \ell \lesssim 3000$) and the composite tail ($\ell \gtrsim 3000$), which is sensitive to several phenomena, like the clustering of sub-millimeter galaxies and properties of the intracluster medium.

The telescope is composed by two ellipsoidal mirrors: the primary with a collecting diameter of 5.8 m formed by 71 aluminum panels, and a secondary mirror of 2 m in diameter, formed by 11 panels. The secondary focus the radiation into the Millimeter Bolometric Array Camera (MBAC), a 3 frequency band camera with bands centered at 148 GHz, 218 GHz, and 277 GHz, with angular resolutions of $1'.4$, $1'.0$ and $0'.9$, respectively. Each array is a set of 1024 Transient Edge Sensor (TES) bolometers organized in columns of 32 elements plus one column of dark detectors, i.e., detectors not coupled to the optical signal. The detectors are cooled down to 300 mK using a helium sorption fridge system (more details in [29]).

The observation technique is to scan in azimuth at constant elevation while the sky rotates, observing the same area while both rising (east) and setting (west). This ensures cross-linked observations of the sky. In this way, the telescope scans through a constant air
mass, and cryogenics and optics remain stable. During the scan, each detector is sampled at 400 Hz, producing data segments which are stored in time-ordered data sets (TODs).

1.3 Signal description

An instrument like ACT, together with the CMB, detects different kinds of sources emitting in millimeter wavelengths, like radio sources and sub-millimeter galaxies. Additionally, clusters of galaxies can be detected through Sunyaev-Zel’dovich effect.

1.3.1 CMB

Since the CMB radiation comes from all directions, a natural way to describe it is using spherical harmonics. Consider a scalar field $T(\mathbf{x})$ defined on the sphere. It can be decomposed in spherical harmonic coefficients such that

$$a_{\ell m} = \int T(\mathbf{x}) Y_{\ell m}^*(\mathbf{x}) d\mathbf{x},$$

(1.1)

where

$$T(\mathbf{x}) = \sum_{\ell \geq 0} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\mathbf{x}).$$

(1.2)

If $T(\mathbf{x})$ is an isotropic and homogeneous field following a normal distribution, then each coefficient $a_{\ell m}$ is an independent Gaussian variable with

$$\langle a_{\ell m} \rangle = 0$$

(1.3)

$$\langle a_{\ell m} a_{\ell' m'}^* \rangle = \delta_{\ell \ell'} \delta_{mm'} \langle C_\ell \rangle.$$  (1.4)

In the last expression, $\langle C_\ell \rangle \equiv C^{th}_\ell$ is the theoretical power spectrum estimated using cosmological models. An unbiased estimator of $C^{th}_\ell$ is given by

$$C_\ell = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2.$$  (1.5)

This is the angular power spectrum of the CMB in terms of spherical harmonic coefficients. Figure 1.1 shows the power spectrum measured by Planck [26], where three regions
can be distinguished:

1. $\ell < 20$. Corresponds to angular scales larger than the Hubble radius ($\sim 1^\circ$), where the fluctuations never evolve. At this scales, the Sachs-Wolfe effect is the dominant source of anisotropy. Data from COBE shows that at scales larger than $10^\circ$ the anisotropy is around $28 \mu K$ [27]. In addition to this gravitational effect, the Doppler effect due to the motion of the Solar System with respect to the reference frame determined by the LSS, produces a dipole anisotropy ($\ell = 1$ or angular scale of $180^\circ$) with $\Delta T/T = 1.2 \times 10^{-3}$.

2. $200 < \ell < 1000$ or scales smaller than the horizon. This region is characterized by acoustic oscillations due to baryons falling into gravitational wells, and interacting with photon radiation. These oscillations produce CMB anisotropies according to $\Delta T/T = \delta \rho_\gamma / \rho_\gamma$, i.e., regions with density higher (lower) than the mean density corresponds to an increment (decrement) of the temperature with respect to the mean temperature. The modes related to this maximum (minimum) density generate a maximum in the power spectrum, called acoustic peaks. The first peak corresponds to a compression during the combination. The second one is a density minimum; it represents an oscillation mode that completed a full compression and rarefaction cycle. The third one represents a maximum density associated with a mode that completed a compression-rarefaction-compression cycle, and so on. The magnitude of this anisotropies is $\Delta T/T \sim 10^{-5}$.

3. $\ell > 1000$. At this scales, the spectrum is damped due to diffusion of photons, smoothing structures with scales smaller than the mean free path. At even smaller scales ($\ell > 3000$), the power spectrum is dominated by SZ effects and emission from radio sources and dusty star forming galaxies.

In the case of a measurement which does not cover the full sky, like the ACT strips, it is necessary to approximate the $C_\ell's$ by introducing a weighting function $W(x)$ and rewriting the power spectrum estimator. Therefore, the expression (1.1) becomes

$$\tilde{a}_\ell m = \int T(x)W(x)Y_{\ell m}^*(x)dx,$$

\[ (1.6) \]
and the pseudo power spectrum is defined as

$$\tilde{C}_\ell = \frac{1}{2\ell+1} \sum_{m=\ell}^{\ell} |\tilde{a}_{\ell m}|^2. \quad (1.7)$$

The relation between this pseudo power spectrum and the true power spectrum can be written as a mode-mode coupling matrix $M_{\ell\ell'}$ resulting from the power spectrum of the window $W(x)$, such that

$$\tilde{C}_\ell = \sum_{\ell'} M_{\ell\ell'} C_{\ell'}. \quad (1.8)$$

If the power spectrum is binned,

$$\tilde{C}_b = \sum_{\ell} P_{b\ell} \tilde{C}_\ell. \quad (1.9)$$

Here we have introduced the binning operator $P_{b\ell}$ defined as

$$P_{b\ell} = \frac{w_\ell}{\sum_\ell w_\ell}. \quad (1.10)$$
where $w_\ell$ is the bin weight, which is described in [6]. Introducing the equation (1.9) we have

$$\tilde{C}_b = P_{b\ell} M_{\ell \ell'} C_{\ell'}$$

(1.11)

$$= P_{b\ell} M_{\ell \ell'} B_{\ell'\ell'} P_{\ell'\ell} C_{\ell'},$$

(1.12)

where $B_{\ell'\ell'}$ is the reciprocal operation of $P_{b\ell}$. Defining the binned mode coupling matrix as $\tilde{M} = P^T M B$, we can write the relation between the binned pseudo spectrum and the true binned spectrum as

$$\tilde{C}_b = M_{bb'} C_{\ell'}.$$  

(1.13)

Then, the true binned power spectrum will be given by

$$C_b = M_{bb'}^{-1} \tilde{C}_{b'}.$$  

(1.14)

The last expression tells us how to recover the power spectrum when a small sky patch is considered. In this case, the orthogonality between modes is lost and the inverse of the mode-mode coupling matrix must be included in the calculations. Since we are not considering the full sky, we can simplify the analysis and use the flat-sky approximation, which consider the scalar field $T(x)$ defined on the sphere and projected on a tangent plane, then Equations 1.1 and 1.2 can be written in terms of Fourier coefficients.

### 1.3.2 Sunyaev-Zel’dovich Effect

Galaxy clusters are considered one of the most powerful tools in the study of the large scale structure. This objects are the largest virialized structures in the Universe with masses between $10^{13}$ and $10^{15} M_\odot$.

X-Ray observations have revealed the presence of intrachuster plasma with temperatures around $10^8$ K. The electrons in this hot plasma interact with ions and radiation. The latter produces inverse Compton scattering of the cosmic microwave background radiation, producing a distortion in the spectrum (SZ Effect) which can be measured and used to detect clusters of galaxies. The brightness temperature change in the direction of the
1.4 Chile ACT Ultradeep Survey

Cluster can be expressed analytically as

\[
\frac{\Delta T_{SZ}}{T_{CMB}} = y \left( x \coth \left( \frac{x}{2} \right) - 4 \right),
\]

where \( x = \frac{h \nu}{k_B T_e} \) and \( y \) is the Compton parameter defined as the electron pressure integrated over the line of sight, or

\[
y = \int_0^\infty \sigma_T n_e \frac{k_B T_e}{m_e c^2} dl.
\]

A second effect arises if the cluster is moving with respect to the CMB, called the kinetic Sunyaev-Zel’dovich (kSZ) effect and it is due to Doppler effect from motion in the line of sight. It is described by

\[
\frac{\Delta T}{T} = - \int \hat{n} \cdot \vec{v} \frac{n_e \sigma_T}{c} dl.
\]

A recent detection of this effect was done from ACT maps, consistent with simulations [13].

1.3.3 Point sources

As mentioned before, the ACT bands are sensitive to galaxies emitting radiation in millimeter wavelengths, allowing us to study their populations. Most of the sources in ACT maps corresponds to galaxies powered by a central super massive black hole, showing a spectral energy distribution dominated by synchrotron radiation [21]. Also, sub-millimeter galaxies (SMGs) can be detected at 220 and 280 GHz. These galaxies are very massive, luminous and dusty, responsible of most of star formation in the early Universe. By studying their angular correlation function, it is possible to verify their correspondence with present day massive ellipticals, their believed descendants [2].

1.4 Chile ACT Ultradeep Survey

The Chile ACT Ultradeep Survey (CACTUS) is a multiwavelength study of a 30 deg\(^2\) area of the sky using 10% of ACT’s time, addressing properties of different galaxies populations, as well as galaxy clusters, using the three ACT bands. To achieve this goal, it is necessary to implement a reduction pipeline, with special emphasis on the map making process.
Chapter 1. Introduction

In this context, the main goal of this work was to setup a complete pipeline based on the one already implemented in Princeton by the ACT collaboration. We were able to generate our own maps, in particular of the 280 GHz band, which had not been processed before. The final software is able to work with data subsets, allowing us to focus on small patches containing, for example, a galaxy cluster or an SMG.
Chapter 2

The Data

2.1 Data description

The data were taken during seasons 2008 (S2), 2009 (S3), 2010 (S4) as part of the ACT southern strip, covering the area between [4 hr, 5 hr] in Right Ascension and [-51°, -54.5°] in Declination. There are a few data from season 2007 (S1) for the 148 GHz band, but they were not included in this work due to their low density. The season 2008 includes a maximum of 98 usable nights from August 29 to December 24. For seasons 2009 and 2010 the maximum number of usable nights available are 65 from August 25 to December 18, and 99 from August 5 to December 24, respectively. The weather conditions are summarized in Figure 2.1, which shows Precipitable Water Vapor (PWV) histograms for all three seasons considered in this work. The median values are 0.51 mm, 0.91 mm and 0.53 mm for S2, S3 and S4, respectively.

Figure 2.2 shows hit maps for season 2008, including all three bands after data selection. Note that the effective integration time is not the same in all three bands; this is due to selection performed as first step in data reduction (see Section 2.2).

Table 2.1 summarizes the total number of TODs and their equivalence in disk space and hours of observation after data selection for each band and season. The data is stored in a small server to make analysis and code testing. For the multi-TOD maps, we use a 512-core computer cluster with a local copy of the data, containing 2 TB of usable data.
Figure 2.1. Precipitable water vapor histogram for each season considering data used in this work. The label is showing the median value. These values are taken from the Atacama Pathfinder Experiment (APEX) weather station.
2.1. Data description

(a) Season 2008.

(b) Season 2009.

(c) Season 2010.

Figure 2.2. Time maps for the three seasons considered. The top, middle and bottom panels of each figure corresponds to 148 GHz, 218 GHz and 280 GHz, respectively.
2.2 Data selection

Before mapping, the data is selected to ensure its quality using certain criteria. These criteria are designed to reject a complete data file and also remove unwanted effects within a data file, such as: mechanical vibrations of the coupling layer placed in front of detectors, appearing as lines in the TOD power spectrum at harmonics of the scan frequency; changes in signal voltage due to quantum change in biasing flux of the readout; excess of noise due to biasing problems; and electromagnetic pick up. The complete description of data selection methods and criteria can be found in [9]. Here we summarize the method to justify the final data selection used for our maps.

There are two types of data cuts. The first one determines the total observing hours by rejecting a complete 15-minutes data file. The criteria to reject a data file are: few effective detectors, bad weather conditions, bad cryogenic performance and bad calibration. The second data cut is related to different kinds of pathologies affecting detectors, resulting in a complete detector TOD cut or a section of it (partial-cut). In this context, different tests are performed to determine the number of effective detectors, defined as

\[
N_{\text{eff}} = \sum_{i=1}^{1024} \frac{T_{i,\text{uncut}}}{T_{\text{total}}},
\]

(2.1)

where \(T_{i,\text{uncut}}\) is the available time after data selection for the detector \(i\), and \(T_{\text{total}}\) is the total time before data selection [8].

A master list is generated including the uncut data files. Also, a partial cuts file is generated for each TOD in the master list. This process was done by the ACT collaboration to the whole ACT data set, therefore we skip the data selection step and use the already computed data cuts. Table 2.1 is showing the amount of TOD files after data selection for CACTUS area and the equivalence in observation hours and size. Additionally, the percentage of cut observation time is shown in the last three columns. The total amount of usable data corresponds roughly to 15% of the total ACT data base for S2 and 7% for
2.3. Data preprocessing

Table 2.1. Data summary after data selection.

<table>
<thead>
<tr>
<th>Band</th>
<th>Total Nights</th>
<th>Total Time (Hrs)</th>
<th># Selected TOD</th>
<th>% Cut</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$S_2$</td>
<td>$S_3$</td>
<td>$S_4$</td>
<td>$S_2$</td>
</tr>
<tr>
<td>148 GHz</td>
<td>98</td>
<td>65</td>
<td>93</td>
<td>128.14</td>
</tr>
<tr>
<td>220 GHz</td>
<td>93</td>
<td>51</td>
<td>85</td>
<td>108.96</td>
</tr>
<tr>
<td>280 GHz</td>
<td>97</td>
<td>53</td>
<td>99</td>
<td>162.94</td>
</tr>
</tbody>
</table>

S3 and $S_4$.

Additional cuts

In the map making step, some unwanted effects arise due to spectral leakage and poorly sampled regions, specially in map edges. For these reasons, additional cuts were considered to avoid artifacts, reducing slightly the amount of samples available for making maps. In the case of 148 GHz band, we cut the first and last seconds of data to let them evolve during mapping; for the other two bands, we also cut the turnarounds because they produced powerful lines in the maps. The width of the turnaround cut in map-space is typically $0.3^\circ$.

2.3 Data preprocessing

Before using TODs to estimate noise or to make maps, we process them as follows:

1. **Data selection.** The partial cuts mentioned in Section 2.2 are applied to every TOD file. We interpolate linearly to fill the discontinuities. If any additional cut is needed, it is applied in this step.

2. **Calibration.** An overall calibration is computed using observations of Uranus. This is applied to every detector TOD in a data file. Also a flatfield is included. Then, an overall map calibration is computed by cross correlating the map’s power spectrum with WMAP for multipoles in the range $400 < \ell < 1000$, with 2% of uncertainty in temperature [12].

3. **Pointing solution.** The solution is computed by mapping planets and point sources.
The relative position of detectors is determined by fitting the beam in the time-streams. The boresight position is computed by comparing the positions of point sources to their catalog position. The final pointing uncertainty is 4.\textquotesingle 8 [8].

4. *Downsampling.* The signal band width is limited to around 30 Hz by the beam size so it is safe to downsample the time-streams from 400 Hz to 200 Hz. This is done using a time-domain triangular kernel. The downsampling process results in a lower memory usage during the noisy modes estimation and the map making stage, where the memory management is crucial.

5. *Mean removal.* The DC level of each detector is arbitrary. In order to make them comparable, we remove the mean value of each detector TOD.

6. *Trend removal.* The slow atmospheric drift produce a mismatch between the beginning and the ending of a TOD. This is a problem when calculating the Fourier transform of a time-stream, since the FFT assumes periodicity of the signal. We remove a linear fit which set the edges at the same level.

7. *Time constant deconvolution.* The limited time response of the detectors act as a low pass filter; the effect can be reverted deconvolving with the inverse filter for each time constant [9].

8. *Dark modes removal.* Some electronic and thermal effects can be detected using dark detectors, which are not coupled to the sky signal. We remove ten correlated modes to reduce these effects.
Chapter 3

Noise sources

ACT is a ground-based instrument, therefore the measurements are affected by the atmosphere emission and absorption in the bands of interest. Additionally, systematics and noise contamination from the telescope are present. Then, before making maps, it is important to identify those contaminants and model them to include them properly in the map solution.

3.1 Sources of contamination in ACT data

We distinguish three main sources of noise: the atmosphere, an important absorber and emitter at the science bands, systematics, related to properties of the instrument, and random noise, which is associated to properties of the detectors and the readout system. For a detailed discussion of atmospheric, systematic and random noises, see [9].

Atmosphere

The ACT site has a very low humidity, meaning low opacity in the ACT bands, which is a necessary condition for millimeter astronomy. A variety of experiments take advantage of this fact: MAT/TOCO [31], ABS [24], Polarbear [17], ALMA [32], APEX [11]. Nonetheless, the atmospheric effects are still the major source of contamination in the ACT data and need to be addressed carefully in the map solution.
The atmosphere attenuates the incoming CMB signal by a factor $e^{-\tau(\nu)}$, where $\tau(\nu)$ is the optical depth or opacity which depends on the frequency. The ACT bands are affected mainly by continuum emission as revealed in Figure 3.1, which is showing the atmosphere brightness temperature as a function of the amount of precipitable water vapor (PWV), a key variable to understand the atmospheric contamination. Atmosphere opacity at 148 GHz band is dominated by a 120 GHz $0^2$ line and 183 GHz water line. For the other two bands, the attenuation is higher due to the 183 GHz resonance and another water line at 325 GHz. The atmosphere emission can be modeled as a gray body

$$I_{\lambda atm} = \frac{2kT_{atm}}{\lambda^2} \left(1 - e^{-\tau_A}\right),$$

(3.1)

where $T_{atm}$ is the atmosphere temperature and the optical depth is given by $-\tau A$, where $A$ is the air mass and $\tau$ is the zenith optical depth. The latter can be expressed as a function of PWV ([20], [30], [9]) as well as the equivalent atmosphere temperature $T_R = T_{atm} (1 - e^{-\tau A})$.

Atmospheric structures larger than the field of view (24') appear as a common mode, while smaller features produce correlations between subsets of detectors. The atmospheric signature is shown in Figure 3.2, where the effects of different PWV values are clear: the atmosphere signal is higher when the PWV is high, moving the knee towards higher frequencies.

**Uncorrelated noise**

Random noise dominates above the atmosphere knee frequency. It is related to fundamental properties of the detectors and the readout system, and to the quantum nature of light. One of the main properties of random noise is that it is uncorrelated between detectors, therefore, the signal-to-noise ratio can be improved by adding more observations, since the variance decreases as $\sqrt{N}$. To quantify the noise, the TOD power spectral density is averaged in a frequency range, typically 5-20 Hz (since the signal band is below 30 Hz), and it is expressed as the noise equivalent temperature (NET) or the temperature needed to produce a SNR of unity for an integration of one second. Table 3.1 summarizes the NET values for AR1 and AR2; the values for AR3 comes from the noise estimation given in Chapter 5.
3.1. Sources of contamination in ACT data

Figure 3.1. Atmosphere brightness temperature as a function of the amount of precipitable water vapor. Taken from [20].

Table 3.1. NET summary.

<table>
<thead>
<tr>
<th>Season</th>
<th>Total NET $(\mu K/\sqrt{s})$</th>
<th>Typical NET $(\mu K/\sqrt{s})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td></td>
<td></td>
</tr>
<tr>
<td>148 GHz</td>
<td>31.9 ± 2.9</td>
<td>789.3 ± 35.6</td>
</tr>
<tr>
<td>220 GHz</td>
<td>41.1 ± 4.8</td>
<td>1151 ± 89.4</td>
</tr>
<tr>
<td>280 GHz</td>
<td>115.6</td>
<td>-</td>
</tr>
<tr>
<td>2009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>148 GHz</td>
<td>28.7 ± 2.2</td>
<td>728.8 ± 47.9</td>
</tr>
<tr>
<td>220 GHz</td>
<td>43.1 ± 7.4</td>
<td>1149 ± 182.7</td>
</tr>
<tr>
<td>280 GHz</td>
<td>105.3</td>
<td>-</td>
</tr>
<tr>
<td>2010</td>
<td></td>
<td></td>
</tr>
<tr>
<td>148 GHz</td>
<td>29.2 ± 1.7</td>
<td>676.5 ± 29.6</td>
</tr>
<tr>
<td>220 GHz</td>
<td>41.4 ± 11.7</td>
<td>1005 ± 142.7</td>
</tr>
<tr>
<td>280 GHz</td>
<td>84.6</td>
<td>-</td>
</tr>
</tbody>
</table>
Correlated noise

In addition to the above noise sources, there is noise contamination that affect more than one detector at the time, appearing as correlations between them. This effects can produce artifacts in map space.

The main examples are: cryogenic drift, magnetic noise, electromagnetic noise and mechanical accelerations. The first one is related to temperature fluctuations in the 3K stage, affecting the SQUID series array used to amplify the signal from bolometers. These fluctuations produce a slow drift in the TODs, which is important when PWV is low, but it is sub-dominant most of the time. The drift is also present in the dark detectors, which can used to identify the signal. The magnetic pickup is due to the Earth’s magnetic field and it is synchronous with the scan. This signal gets downweighted in the map making step. The electromagnetic pickup is another source of contamination which couples to the readout circuit, producing correlations between rows of detectors; it was particularly important in 2008. The effect is non stationary produced by electromagnetic glitches coupling in the
time-multiplexed readout system. Thus, this noise appeared correlated in detectors from the same row in the array, as rows were read simultaneously. We clean this contamination by removing modes obtained by averaging detectors TODs of the same row and selecting the main modes using the SVD method described below. Finally, mechanical accelerations at the scan turnaround produce mechanical vibrations of the detector coupling layer, which translates into multiple lines in the TOD power spectrum. This effect is stronger for 280 GHz because the coupling layer is more flexible and thinner.

### 3.2 Singular Value Decomposition

Contaminant signals in the data can be identified using Singular Value Decomposition. This method allows us to decompose a \(m \times n\) matrix \(d\) as the product of three matrices, as follows

\[
d = USV^T,
\]

where \(U\) is an \(m \times r\) column-orthonormal matrix \((U^TU = I)\) with \(r\) the rank of matrix \(d\) (number of linearly independent columns), \(V\) is an \(r \times n\) orthogonal matrix \((V^TV = I)\) and \(S\) is a \(r \times r\) diagonal matrix with \(\sigma_{ij} = 0\) if \(i \neq j\) and \(\sigma_{ii} \geq 0\). By doing this factorization, we are changing to a basis where the matrix \(d\) becomes diagonal. If we multiply Equation 3.2 by \(V\) from the right we get

\[
dV = US.
\]

The last expression indicates how the columns in \(d\) must be combined through vectors in \(V\) to get the columns in \(U\) with a relevance factor \(S\). This idea can be used in the case \(d\) is a data matrix with each column representing a detector time-stream. If we decompose a data matrix \(d\) in its singular values, then it is possible to obtain a representation of the signals present in the data by combining linearly the time-streams using vectors \(V\). We can express a normalized mode as follows

\[
\hat{m}_i = \frac{d\hat{v}_i}{\sigma_{ii}},
\]

where \(\hat{v}_i\) is the column \(i\)-th of matrix \(V\) and \(\sigma_{ii}\) is the corresponding singular value in \(S\).

An important property is that vectors in \(V\) are the eigenvectors of the data covariance.
matrix

\[ C = d^T d, \quad (3.5) \]

which is a square matrix with dimensions \( n \times n \); if \( d \) is a data matrix, the dimensions corresponds to the number of detectors. Introducing the SVD factorization given by Equation 3.2 in the last expression we get

\[ C = VSU^T USV^T \quad (3.6) \]

\[ = VS^2 V^T, \quad (3.7) \]

where we see that it has the form of an SVD, sharing the same eigenvectors \( V \), and with singular values equal to the singular values of the data matrix squared. Obtaining \( V \) from the covariance matrix is very useful as it significantly reduces the computational time.

The relevance of a particular mode is determined by its singular value. In the case of the ACT data, the strongest signal present in the data is the atmosphere, which produces high correlations between detectors; also, there are correlations among subsets of detectors produced by structures in atmosphere at scales of \( \sim 10' \). Then, we expect to see a very steep distribution of singular values, where the higher values are likely to represent the atmospheric signal. Figure 3.3 shows a typical distribution for the singular values of a TOD. We see that the power is mostly concentrated in the first five modes, in this example. Once we have identified the important modes, we would like to remove them from the data in a proper way. Following [9], the modes can be deprojected using the equation

\[ \tilde{d} = (I - \hat{m}_{i}\hat{m}_{i}^T) d \quad (3.8) \]

\[ = d (I - \hat{v}_{i}\hat{v}_{i}^T) \quad (3.9) \]

For the last equality we use Equation 3.6. This result is easily generalized to the case of a matrix \( V' \) containing \( n \) vectors \( \hat{v} \)

\[ \tilde{d} = d (I - V'V'^T) \quad (3.10) \]

In order to verify the performance of the method, we applied the previous definitions (Equations 3.6 and 3.10) to a time-stream and check how the correlations looks like after
removing modes. The TOD is split at 4 Hz by low/high pass filtering the data, obtaining two TODs with data below/above 4 Hz. We can quantify the correlations with a quality factor defined as the mean value of the squared off-diagonal elements [8]. Before removing a set of modes, this quality factor is almost unity in case of low frequency data. Figure 3.4 shows the effect of Equation 3.10 in correlation matrix for low frequency data. Note that the quality factor is lower than unity, meaning that the correlations between detectors were suppressed. In the low frequency correlation matrix there are still correlations, specially between columns. This is due to modes below 0.5 Hz ($\ell \lesssim 200$) that could be removed adding more modes to matrix $\mathbf{V}'$, but instead of doing that, we give them zero weight during map making. On the other hand, the high frequency correlations need only a few modes to improve. There are some remaining correlations between rows (bottom right panel) due to electromagnetic noise; this situation can be improved by using row correlated modes or adding more high frequency modes.

**Figure 3.3.** Singular value distribution for low and high frequency data (see text). For low frequency data, the first singular value is out of the plot.
Chapter 3. Noise sources

Figure 3.4. Correlation matrices for low frequency (top panels) and high frequency (bottom panels) after removing 20 low frequency modes and 5 high frequency modes.
Chapter 4

Map making

In this chapter, we review how to make a map that is the best representation of the sky given a noise model, pointing solution, calibration, and cuts.

The data are mapped using the pointing solution which determine the corresponding pixel in a map for each sample in the time-streams. This can be written as

\[ s = M^T d, \] (4.1)

where \( s \) is an \( n_{\text{pix}} \) vector representing the TOD projection to map-space, \( M \) is a projection matrix with dimensions \( n_{\text{data}} \times n_{\text{pix}} \) and \( d \) is the data vector with \( n_{\text{data}} \) elements.

4.1 Map making equation

The goal of the map making process is to find the best combination of samples such that each pixel in the map represents the sky signal. This can be done by minimizing \( \chi^2 \) for a given noise model. The data can be written as a linear process

\[ d = Mx + n, \] (4.2)

where \( x \) is the map solution and \( n \) is noise. A good approximation is to say that the noise follows a normal distribution, such that the likelihood, or the probability to obtain the
data $d$ when its true value is $x$ given a Gaussian distribution of $n$, is written as

$$L \propto \exp \left( -\frac{1}{2} n^T N^{-1} n \right) = \exp \left( -\frac{1}{2} (d - Mx)^T N^{-1} (d - Mx) \right),$$

(4.3)

where $N = \langle nn^T \rangle$ is the covariance matrix of the noise. The likelihood is maximum when $\partial \ln L / \partial x = 0$, i.e.,

$$M^T N^{-1} Mx = M^T N^{-1} d.$$

(4.4)

The last equation is known as Map making equation. It is important to note that it has the form $Ax = b$, where $A = M^T N^{-1} M$ and $b = M^T N^{-1} d$, so we can find the solution $x$ by inverting $A$. In the context of making maps, the inversion of $A$ is technically very difficult due to the huge dimensions of $A$. Instead an iterative method is used. A known method to solve the linear problem 4.4 is the Conjugate Gradient method [14], whose convergence can be accelerated by introducing a preconditioner $P$, which is an approximation of $A^{-1}$. In the context of map making, a naive preconditioner is the inverse of the weight matrix $M^T M$, which contains the number of samples falling in each pixel. After applying the preconditioner, Equation 4.4 becomes

$$PM^T N^{-1} Mx = PM^T N^{-1} d.$$

(4.5)

### 4.2 Noise model

The map solution depends on the correct modeling of the noise matrix. The noise term in equation 4.4 allows us to incorporate the noise correlations discussed in Chapter 3. Here we describe the noise model used in this work.

We model the noise following [8]. As mentioned previously, the noise structure is complex due to presence of atmospheric noise affecting mainly large angular scales and detector noise acting on small scales. For this reason, we split the data in frequency domain by applying a low pass and a high pass filter, before computing two covariance matrices: $C_l$ and $C_h$ where subscripts indicate low and high frequency, respectively. After decomposing these matrices using SVD, the main modes are selected by setting a threshold on the singular values, which is ten times the mean value of the singular values with index higher than 50. This selection criteria leaves typically 20 low frequency modes and 5 high frequency modes.
4.2. Noise model

Then, the final set of modes is conformed by $i$ low frequency modes and $j$ high frequency modes, grouped in a matrix $V$ of dimensions $n_{\text{det}} \times (i + j)$.

The next step is to model the noise in Fourier space. To do that, we decorrelate the time-streams using Equation 3.10 and the modes previously found. Then, we compute the power spectrum of each detector TOD in $\tilde{d}$ and each mode; to write a particular mode in time domain we use Equation 3.4. The power spectrum is then binned in 2.5 Hz steps and the mean power in each bin is computed; the frequency step may vary to increase the resolution in frequency space and improve the model in cases like, for example, powerful lines due to coupling layer vibrations. This power is used to construct a noise matrix per each frequency bin $f$ by summing the white noise contribution (from $\tilde{d}$) and the correlated noise given by the mean power of modes properly combined with $V$, written as

$$N_f = N_{\tilde{d}} + VN_mV^T.$$  \hspace{1cm} (4.6)

The matrices $N_{\tilde{d}}$ and $N_m$ are diagonal of dimensions $n_{\text{det}} \times n_{\text{det}}$ and $n_{\text{modes}} \times n_{\text{modes}}$ respectively, containing the mean power of $\tilde{d}$ and of the modes per frequency bin $f$, then we have $n_{\text{bins}}$ matrices $N_f$. Note that the second term in Equation 4.6 has the form of Equation 3.6. The inverse of Equation 4.6 can be computed using the Woodbury identity

$$N_f^{-1} = N_{\tilde{d}}^{-1} - N_{\tilde{d}}^{-1}V\left(N_m^{-1} + V^TN_m^{-1}V\right)^{-1}V^TN_{\tilde{d}}^{-1}. \hspace{1cm} (4.7)$$

This is the noise model for the frequency bin $f$. The matrix $N_f^{-1}$ has dimensions $n_{\text{det}} \times n_{\text{det}}$ and it is applied to the data by multiplying it with the Fourier transform of the time-stream $d$, $D \equiv FWd$, where $W$ is a window in time-domain of the form $(1 - \cos(x))/2$ to reduce spectral leakage. Then, for each frequency bin $f$ we have

$$\tilde{D}_f = N_f^{-1}D_f,$$  \hspace{1cm} (4.8)

where $D_f$ is a segment of $D$ of dimensions $n_{\text{det}} \times n_f$, with $n_f$ the number of elements of the bin $f$. The filtered time-stream $d_F$ is obtained by stitching together all frequency bins and applying the inverse Fourier transform

$$d_F = W^TF^T [\tilde{D}_{f_1} \tilde{D}_{f_2} \cdots \tilde{D}_{f_n}] \hspace{1cm} (4.9)$$

27
Finally, the entire inverse noise operator is defined as

$$N^{-1} = W^T F^T N^{-1}_F F W$$

(4.10)

where the operator $N^{-1}_F = \begin{bmatrix} N^{-1}_{f_1} & N^{-1}_{f_2} & \ldots & N^{-1}_{f_n} \end{bmatrix}$ contains the $N^{-1}_f$ matrices for $n$ frequency bins $f_n$.

### 4.2.1 Dealing with gaps in data

The noise operator 4.10 acting in Fourier space requires a continuous function. As described in Section 2.2, data segments are rejected according to various selection criteria, meaning that only a fraction of a time-stream could be used for making maps. This fact is included in the operator $M$ as a cutting operator defined as

$$\tau_i = \begin{cases} 0, & \text{if } i \in \text{gaps} \\ 1, & \text{otherwise,} \end{cases}$$

(4.11)

where $i$ identify a sample in time domain and $\text{gaps}$ contains the cuts for each detector time-stream. When projecting a time-stream to map-space using $M_T$, the samples $i$ in $gaps$ are not projected; the inverse operation add zeros to $i$ in time domain, generating discontinuity in the TOD. This is a problem since we need Fourier transform the time-streams to apply the noise model. One way to solve this problem, is creating a second map $x_c$ containing a suitable solution for gaps and defining a projection operator $M_c$ with the inverse of the cutting operator 4.11, defined as

$$\tau_c^i = \begin{cases} 1, & \text{if } i \in \text{gaps} \\ 0, & \text{otherwise.} \end{cases}$$

(4.12)

Therefore, we can rewrite Equation 4.2 including the above definitions:

$$d = Mx + M_c x_c + n,$$

(4.13)

or

$$d = \begin{bmatrix} M & M_c \end{bmatrix} \begin{bmatrix} x \\ x_c \end{bmatrix} + n = \bar{M} \bar{x} + n.$$

(4.14)
4.3. Preconditioned Conjugate Gradient

The final preconditioned map making equation is

\[ P \tilde{M}^T N^{-1} \tilde{M} \tilde{x} = P \tilde{M}^T N^{-1} d, \]  
\[(4.15)\]

\[ P \begin{bmatrix} M^T \\ M_c^T \end{bmatrix} N^{-1} \begin{bmatrix} M \\ M_c \end{bmatrix} \begin{bmatrix} x \\ x_c \end{bmatrix} = P \begin{bmatrix} M^T \\ M_c^T \end{bmatrix} N^{-1} d. \]  
\[(4.16)\]

This is the equation that we actually solve.

Another method to solve for gaps is to start with a solution in TOD-space and let it evolve without projecting into map-space. We found that this method produce slow convergence and sometimes diverges.

4.3 Preconditioned Conjugate Gradient

As mentioned before, we solve the linear system 4.5 using preconditioned conjugate gradient method [14]. It is an iterative method to solve the linear system \( Ax = b \) and it is a generalization of the steepest descent method. The method starts considering a base \( p \) of \( R^n \) such that a solution \( \tilde{x} \) of \( A \tilde{x} = b \) can be expanded as follows

\[ \tilde{x} = \sum_{i=1}^{n} \alpha_i p_i. \]  
\[(4.17)\]

Introducing this solution in the linear equation we have

\[ A \tilde{x} = \alpha_1 A p_1 + \alpha_2 A p_2 + \cdots + \alpha_n A p_n = b, \]  
\[(4.18)\]

with \( \alpha_k \) given by

\[ \alpha_k = \frac{p_k^T b}{p_k^T A p_k}. \]  
\[(4.19)\]

The method requires to move in the direction of the residuals defined as \( r_k = b - Ax_k \). Also, the directions should be conjugate to each other. As the method does not use a predetermined orthonormal base, it makes use of the residual to get the new directions by considering the additional property of the succession \( r_k \): the orthogonality \( \langle r_i, r_j \rangle = 0 \) for
Chapter 4. Map making

\(i \neq j\). It can be shown that the Gram-Schmidt orthonormalization gives

\[
p_{k+1} = r_k - \sum_{i \leq k} \frac{p_i^T A r_k}{p_i^T A p_i} p_i.
\]  

(4.20)

Following that direction, the next solution will be given by

\[
x_{k+1} = x_k + \alpha_{k+1} p_{k+1}
\]  

(4.21)

and the coefficients \(\alpha_k\) are rewritten as

\[
\alpha_k = \frac{r_k^T r_k}{p_k^T A p_k}.
\]  

(4.22)

The method can be improved by reducing the matrix condition number or the ratio between the highest and smallest eigenvalue. To achieve this, we introduce a preconditioner \(P\), a symmetric, positive definite matrix that approximates \(A^{-1}\), so that the problem to solve is

\[
PAx = Pb
\]  

(4.23)

If the condition number of \(PA\) is lower than \(A\), then we solve 4.23 iteratively more quickly than the original equation. For this reason we introduced a preconditioner in the map making equation 4.5. In this work, we use the inverse of the weight matrix \(W = M^T M\).

Finally, the full algorithm is summarized as follows

1. Set the initial conditions:
   
   (a) \(r_0 = b - Ax_0\).
   
   (b) \(p_0 = r_0\).

2. Iteration over \(k\):
   
   (a) \(\alpha_k = \frac{r_k^T r_k}{p_k^T A p_k}\).
   
   (b) \(x_{k+1} = x_k + \alpha_k p_k\).
4.3. Preconditioned Conjugate Gradient

(c) \( r_{k+1} = r_k - \alpha A p_k \).

(d) If \( r_{k+1}/b \) is small enough: break

(e) \( \beta_k = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k} \).

(f) New direction: \( p_{k+1} = r_{k+1} + \beta_k p_k \).

(g) \( k = k + 1 \).

The convergence is achieved when the ratio between residuals and \( b \) is low enough (d); a typical convergence limit is \( 10^{-6} \). In general, we use the convergence ratio defined in Section 5.3 to check power convergence. In some cases, it is not necessary to achieve the convergence limit. For example, small angular scales converge faster (see Section 5.3), so for finding point sources it is possible to stop PCG or set a limit for the number of iterations.

![PCG diagram](image)

**Figure 4.1.** PCG diagram.

The basic algorithm is shown in Figure 4.1. The initial step is done in parallel, where the master process is responsible of map-space operations while slaves works with TODs.
4.4 Implementation

We developed a code that is able to work with several TODs at the same time using a parallel environment, enabling us to make maps of small regions with a few TODs in a short time or a big map including all the available data. It is based in MOBY, a data reduction software developed by ACT team, which is written in C and Python.

For software development, we used a 8-core computer with serial TOD processing, i.e., one data file at a time, which meant several hours to produce a completely converged map. For this reason, we had to improve our code in order to run in a big cluster. We used MPI for Python [3], which allowed us to processes several data files at once, with the advantages of the Python programming language. The software is currently installed in the Geryon cluster located at Center for Astro-Engineering AIUC. This cluster contains 64 nodes, adding up to 128 Intel Xeon Quad Core CPUs, which means 512 cores available for computations. The total amount of RAM is 1024 GB\(^1\), meaning 2GB per core.

The data preprocessing, noise estimation and map making processes are quite intensive in terms of CPU and memory usage. The size of a single TOD file ranges between 450MB and 600MB after applying compression, but once the file is loaded in memory, the total memory used is 1.1GB. This means that we had to limit the number of data files per node and optimize the memory intensive operations, such as Fourier transforms of time-streams. As mentioned in Section 4.2, the noise model operates in frequency space, meaning that, for each iteration described in Section 4.3, we allocate the equivalent to two data files: the map projected onto time space, and its Fourier transform; in our case \( A = M^T N^{-1} M \), where \( M \) projects a map onto TOD space and the operator \( N^{-1} \) Fourier transforms the resulting time-stream. But, the preprocessing steps described in Section 2.3 consider downsampling, reducing the used memory in data loading and FFT computations to almost half. Therefore, 2GB of memory per core is enough to process a downsampled TOD and to keep in memory the vectors (maps) associated with the PCG. The memory needed to allocate vectors (maps) stays constant along the process and its size depends on the dimensions of the mapped area. For a full area map, we found that this constant level is high enough to difficult the processing of multiple TODs in a single node. Since the memory is shared

\(^1\)http://www2.astro.puc.cl/geryon
between cores in the same node, simultaneous processes could fill up the memory. To address this, the FFTs are calculated in a serial process to prevent that memory peaks add up.

As our access to the computer cluster is limited, the noise model estimation is done in a small server so that we can dedicate our cluster time for making maps. This last process involves data reading, preprocessing and PCG initialization, for which the performance is limited mainly by the hard drive data transfer rate: for big data sets, the data reading process is inefficient, especially when several TOD files are being read at the same time, producing lags and failures. This problem can be easily solved by limiting the number of data files which are read simultaneously, improving significantly the performance.

The parallelization was made by separating map-space and TOD-space operations. In the PCG context, the map-space operations consist in calculating the new map, residuals and directions in each iteration. The TOD-space operations instead, involve Fourier transforms, doubling the amount of memory. Then, we divide the code into a master process and a slave process. The former, splits the data list, assigning tasks to the slaves (normally one TOD per slave). Each slave performs the preprocessing steps for its corresponding data list. The PCG steps that include the matrix $\mathbf{A} (\mathbf{M}^T \mathbf{N}^{-1} \mathbf{M})$ are always done by a slave process. We can summarize the parallel PCG as follows

1. Set up:
   
   (a) **Master.** Set up slaves and data list. Split data list and pass it to the slaves processes.

   (b) **Slaves.** Preprocess TODs and set up PCG vectors $p_i$, $q_i$, $r_i$ for each $i$ slave. Then, send the vectors to the master.

   (c) **Master.** Reduce PCG vectors, setup initial conditions ($r_0 = b - q$, $p_0 = r_0$), and apply preconditioner to $x_0$.

2. PCG iterations:
   
   (a) **Master.** Check convergence: $r/b$. Apply preconditioner to $q$ and compute $r^T r$, $p^T \mathbf{A} p$, and $\alpha$. Send $q$ to the slaves.

   (b) **Slave.** Receive a copy of $q$ from the master for the operation $\mathbf{A} q = \mathbf{M}^T \mathbf{N}^{-1} \mathbf{M} q$. After that, send $q$ to master.
(c) *Master.* Calculate the next solution with reduced $q$: $x = x + \alpha q$, and the residuals $r = r - \alpha q$. Then get Gram-Schmidt coefficients $\beta$ and the new direction $p = r + \beta p$.

(d) Repeat steps (a), (b), (c) until convergence is achieved or until the iteration limit set by the user.
Chapter 5

Results

In this chapter we show the maps computed using the code developed as part of this work. Also, we study the spectral properties of the noise, as well as convergence of the maps as function of the iterations, and the transfer function of the map making process.

5.1 The Maps

We generated maps of a 30 square degrees area centered at 67.5° in Right Ascension and -52.8° in Declination, considering 588 hours of data corresponding to 1748 GB. The main goal was to make 280 GHz maps, but we first made 148 GHz maps to check that our pipeline was working well by comparing those maps with ACT official data release. We did an initial validation test comparing the fluxes of point sources already detected by ACT collaboration [19], and checking the convergence of those values. In Section 5.2 we calculate the convergence level as function of angular scales.

Since we have limited computational resources and the map making process is very intensive, it is not possible to produce a full area map using the entire data set in one run. For this reason, the data were divided into four subsets with roughly equal integration time. The final map is a weighted mean of the four splits or a coadded map; we proceed in the same way with other bands and seasons. We measured the fluxes of all point sources detected in our area from a coadded map and compared them to ACT measurements. Figure 5.1 shows the convergence of the flux of three point sources to the ACT value given
Chapter 5. Results

Figure 5.1. Flux convergence of three selected point sources. The red dashed line corresponds to the flux measured by the ACT collaboration [19]. The error bar corresponds to the RMS around the source, beyond 3 times the FWHM [22]. The dashed line and the shaded region correspond to the flux and the error bar reported by ACT.

in [19], showed as a dashed line. It is clear that we need a few iterations to recover point source fluxes. The fluxes were computed using a point source analysis code developed by Rolando Dünner with the help of Gustavo Morales [22].

Figure 5.2 shows our fluxes compared to the official ACT catalog, together with the signal-to-noise ratios of those detections. The fluxes are consistent with each other, except for fluxes below 30 mJy, where the map noise begins to be important. We observe an improvement in the SNR of our fluxes compared to ACT which ranges from 2% to 35%. This is due to the careful inclusion of new correlated modes in the TOD noise model for map making.
5.1. The Maps

Figure 5.2. Flux and SNR comparison for the 148 GHz band in season 2008. Our SNR show an improvement compared to the one reported by ACT.

Having validated our system for small angular scales, we continued iterating our solution to study the convergence of larger angular scales, which converge more slowly as seen
Table 5.1. Maps summary. In case of ACT maps: $S$ is for southern strip and $Eq$ for equatorial strip.

<table>
<thead>
<tr>
<th>Season</th>
<th>148 GHz</th>
<th>218 GHz</th>
<th>280 GHz</th>
</tr>
</thead>
<tbody>
<tr>
<td>2008</td>
<td>CACTUS</td>
<td><strong>Done</strong></td>
<td><strong>Done</strong></td>
</tr>
<tr>
<td>ACT</td>
<td>$S/Eq$</td>
<td>$S/Eq$</td>
<td>$S$</td>
</tr>
<tr>
<td>2009</td>
<td>CACTUS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ACT</td>
<td>$S/Eq$</td>
<td>$S/Eq$</td>
<td>$Eq$</td>
</tr>
<tr>
<td>2010</td>
<td>CACTUS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>ACT</td>
<td>$S/Eq$</td>
<td>$S/Eq$</td>
<td>$S/Eq$</td>
</tr>
</tbody>
</table>

in simulations [10]. Table 5.1 details the full TOD maps computed in this work and those computed by ACT collaboration.

Figure 5.3 shows 1-degree maps with two point sources as seen in the three bands in season 2008. The three maps were filtered below than $\ell = 300$. The measured fluxes of the central source (ACT-S J042503-533201) are $152.37 \pm 2.21$ mJy @ 148 GHz, $113.09 \pm 3.22$ mJy @ 218 GHz, and $59.51 \pm 2.32$ mJy @ 280 GHz and of the source near the edge of the map (ACT-S J042906-534945) they are $96.34 \pm 3.22$ mJy @ 148 GHz, $69.57 \pm 3.30$ mJy @ 218 GHz, and $21.65 \pm 2.49$ mJy @ 280 GHz. The central source has a spectral index of $-0.3$ [21], dominated by synchrotron emission, and, according to temporal variability studies in Morales’ thesis [22], it is consistent with a blazar. The second point source is also showing a decreasing flux with spectral index of -0.6.

The maps in Table 5.1 are shown in Figures 5.4 and 5.5. All of them were high pass filtered in order to suppress large scale modes. The first figure groups the maps from season 2008 for all the three bands: the top panel corresponds to the 148 GHz band, the middle to 218 GHz and the bottom to 277 GHz. The next figure shows $AR3$ maps for seasons 2009 and 2010. The white noise levels are shown in the next section. Figure 5.6 shows a comparison between our map and the official ACT 148 GHz map; the differences are mainly on large angular scales.
5.2 Noise

Before measuring the noise spectrum and the CMB signal, we summarize the steps we followed to preprocess the maps. The basics of the power spectrum theory are given in Section 1.3.1. The complete procedure is based on [6].

5.2.1 Maps preprocessing

Before power spectrum estimation, the maps are preprocessed as described below.

High pass filter

The large angular scales are dominated by atmospheric noise. For this reason, we apply a high pass filter equal to zero for $\ell < 100$ and has the form $\sin^2 \ell$ for $100 < \ell < 500$. The filtered map is $T_F(x) = F_\ell T(x)$.

Prewhitening or Dynamical range reduction

This step is designed to reduce the leakage of power from large to small angular scales due to point source mask. In this step we add a fraction of the map to an approximation of

Figure 5.3. Sources ACT-S J042503-533201 (center of each map) and ACT-S J042906-534945 (left of each map) in three bands, season 2008.
Chapter 5. Results

Figure 5.4. Season 2008 maps for all the three bands. Top panel: 148 GHz, Middle panel: 220 GHz, Bottom panel: 280 GHz. The white noise levels are 28.9 µK-arcmin, 46.1 µK-arcmin and 115.6 µK-arcmin, respectively (Section 5.2).
5.2. Noise

Figure 5.5. Seasons 2009 (top) and 2010 (bottom) maps, 280 GHz band. The white noise levels are 105.3 $\mu$K-arcmin and 84.6 $\mu$K-arcmin, respectively (Section 5.2).
Figure 5.6. Maps comparison. The top panel shows the 148 GHz map computed in this work. The middle panel shows the official ACT map. The bottom panel is the difference between them. The RMS values are 91 $\mu$K, 97 $\mu$K and 64 $\mu$K, respectively.
5.2. Noise

its laplacian, which is computed by taking the difference of two versions of the map, one convolved with a disk of radius 1$'$ and the other convolved with a disk of radius 3$'$.

Window

The maps are multiplied by the total weight map, containing the total number of hits for each pixel. By doing this, the regions with less samples or poorly cross-linked are downweighted. Then, we apply a point source mask and an apodization window.

\[
\tilde{T}(x) = W(x) T_F(x)
\]  

(5.1)

In this expression, \(W(x)\) is the product of the three windows mentioned before and it is the final window used to compute the mode-mode coupling matrix \(M_{\ell \ell'}\).

5.2.2 Noise estimation

The noise properties are not isotropic, therefore, a weighted average must be used to downweight the noisy region of spectra. If we assume that the auto power spectrum is composed by CMB signal plus noise, and since the cross spectrum measures the correlation between two splits with independent noise properties, then the noise can be estimated by taking the difference between the auto and cross spectrum, i.e.,

\[
\tilde{C}^{\text{auto}}_\ell = \frac{1}{4} \sum_{i=1}^{4} \tilde{C}^i \times i
\]  

(5.2)

\[
\tilde{C}^{\text{cross}}_\ell = \frac{1}{6} \sum_{i=1}^{3} \sum_{j=i+1}^{4} \tilde{C}^i \times j
\]  

(5.3)

\[
\tilde{N}_\ell = \tilde{C}^{\text{auto}}_\ell - \tilde{C}^{\text{cross}}_\ell
\]  

(5.4)

The power spectrum in Equations 5.2 and 5.3 are computed using 1.14. Assuming that each split has roughly equal noise, the noise power spectrum of a map can be approximated by \(\tilde{N}_\ell/4\).

Figures 5.7 to 5.11 show PUC and ACT noise power spectra, computed in PUC area by dividing it in four patches. Each pair of spectra correspond to the same number of
iterations. At large scales ($\ell \lesssim 500$), the noise is typically atmospheric contamination, while for small angular scales ($\ell \gtrsim 3000$) the noise is essentially white. The fit to the high multipoles represents the white noise level. For the 148 GHz band, we see a slight improvement on the noise level for $\ell > 3000$, which implies a better signal to noise ratio of the detected objects. For larger scales, i.e., $\ell < 3000$, there is excess noise in our map due to the lack of low frequency modes. Despite this excess, our noise level is still low enough to allow us to detect the CMB signal. For the other bands, there is also an excess noise for large scales, with poor signal to noise ratio, making more difficult to recover the CMB signal. The noise levels for the 280 GHz band, season 2008 and 2009, are still uncertain. In the case of season 2008, we found a dependency of calibration factor with the angular scale; we need more studies to understand this effect (see Section 5.4). In the case of season 2009, the amount of usable data is low (Table 2.1).
5.2. Noise

Figure 5.8. 218 GHz noise power spectra comparison, season 2008.

Figure 5.9. 280 GHz noise power spectra comparison, season 2008.
Chapter 5. Results

Figure 5.10. 280 GHz noise power spectra comparison, season 2009.

Figure 5.11. 280 GHz noise power spectra comparison, season 2010.
Null test

Using the 4-way splits we checked if the signal in the split is consistent between independent map realizations by looking at the cross spectrum of difference maps, defined as

\[ T^{ij}(\hat{n}) = \frac{(T^i(\hat{n}) - T^j(\hat{n}))}{2}, \]  

(5.5)

with \((i, j) = (1, 2), (3, 4), (1, 3), (2, 4), (1, 4), (2, 3)\). We compute the cross correlation of these maps following the steps described in Section 5.2. Figure 5.12 shows the three spectra for each band, in season 2008, being consistent with no signal.

**Figure 5.12.** Null test for all three bands, season 2008. We see that our splits are consistent with no signal.
5.3 Transfer Function and convergence

The response of the map making process is given by a transfer function defined as the ratio between the power spectrum of the output map and the input ideal map. To measure the transfer function, we begin by projecting a known signal (simulated CMB map) to the detector TODs of the real data (sim-inject data). Then, we compute the correlated modes following Section 3.2 and use them to make a sim-inject map. The previously computed data map is removed from the sim-inject map to produce the output map. The ideal case is that this output map must contain only the injected signal. To check that, the cross-correlation between the input map (simulation) is computed. The final transfer function is given by the ratio between the output map power spectrum and this cross-correlation. Figure 5.13 shows the transfer function computed after one hundred iterations. For scales $\ell > 500$, the variations are less than 1%. This effect can be improved with a second estimation of the noise matrix performed after removing the computed data map from the time-streams.

![Figure 5.13. Transfer function of the mapper. The bias in the output is less than 1%.

In addition to transfer function, we tested convergence of the maps or the maximum change in power between the iteration $i$ and the final iteration, to see if large-scale modes
5.4 Cross-calibration to ACT

are recovered. Following [10], the convergence ratio is defined as

\[ r_c = \frac{2\sqrt{\hat{C}_b^i \hat{C}_b}}{\sigma(\hat{C}_b)}, \]  

(5.6)

where \( \hat{C}_b^i \) is the power spectrum of the difference map between iteration \( i \) and the final iteration, \( \hat{C}_b \) is the power spectrum of the final iteration, and \( \sigma(\hat{C}_b) \) is computed using Equation (9) of [10]. Figure 5.14 shows the convergence ratio for iterations 10, 50, 100, 400. We see that small angular scales (\( l \gtrsim 1500 \)) converged by iteration 50, while large angular scales need more PCG iterations to converge. According to [10], with \( r_c \leq 0.5 \) the convergence is reached, then we are well converged by iteration 400.

\[ \text{Figure 5.14. Convergence of angular scales as function of iterations.} \]

5.4 Cross-calibration to ACT

The calibration used in the time-streams comes from measurements of Uranus, with an uncertainty of 6%. The final map calibration is computed correlating multipoles with WMAP-7, with an uncertainty of 2%. We follow [12] in order to calibrate our 148 GHz
map using the official ACT data release. The calibration factor is defined as the square root of the ratio between ACT map cross spectrum and CACTUS cross spectrum, i.e.,

\[
\alpha = \sqrt{\frac{C_{\ell}^{ACT \times ACT}}{C_{\ell}^{PUC \times PUC}}}. \tag{5.7}
\]

We compute \( \alpha \) as a weighted mean in the range \( 500 < \ell < 2500 \). Figure 5.15 shows \( \alpha \) as a function of \( \ell \). The \( C_{\ell}'s \) error bars were computed following [10] and then propagated to get error bars of \( \alpha(\ell) \). The final calibration is \( \alpha = 0.956 \pm 0.025 \). We applied the same method for the 280 GHz maps, as described below.

**Figure 5.15.** Calibration factor for 148 GHz, season 2008.

**Special case: 280 GHz band**

By analyzing the cross-spectrum 148 × 280 GHz we found that the calibration factor for the 280 GHz band is function of \( \ell \) as shown in Figure 5.16. This dependency was found in both ACT and CACTUS maps. Figure 5.16 shows the calibration factor for the 280 GHz band. As a comparison, we show the calibration factor for the season 2010, which is flat, giving us some confidence on the analysis performed. For this work, we adopt a
5.4. Cross-calibration to ACT

constant calibration factor of 1.7, which is the weighted average in the multipoles range $500 < \ell < 2000$. We still do not understand this dependency, but we rejected mapping problems by computing the transfer function shown in Figure 5.17. The transfer function does not show a strong dependence with $\ell$. We need a deep analysis in order to understand this effect.

Figure 5.16. Calibration factors of 280 GHz map of both ACT and CACTUS.
Chapter 5. Results

Figure 5.17. Mapper transfer function using 280 GHz data with a CMB simulation injected on it.

5.5 CMB signal

We use the concepts discussed in Section 1.3.1 to measure the CMB power spectrum in our maps. Figure 5.18 shows the spectra for all CACTUS maps listed in Table 5.1 and the cross-spectra between 148 GHz map and all others. The error bars were computed following the appendix of [6]. We clearly see the second through the sixth peaks of the CMB in our 148 GHz spectrum. It is also remarkable that the cross spectra show the second and third peaks, letting us to cross calibrate the 218 GHz and 280 GHz maps, as we show in Section 5.4.
5.5. CMB Signal

Figure 5.18. CMB signal of CACTUS maps. The top panel shows the 148 GHz spectrum, where the acoustic peaks are clear. All the error bars were computed following [6]. We expect to improve our 280 GHz spectra in future analysis.
Chapter 6

Discussion and Conclusions

The purpose of this thesis was to develop the tools needed to produce CMB and point source temperature maps using data from the Atacama Cosmology Telescope, as part of the CACTUS Project. We begin by preprocessing the time-streams, step which involves data selection, calibration, pointing solution, detector time constant deconvolution, down-sampling, and dark modes removal. After that, we estimate the correlated noise modes to construct a frequency dependent noise model. This model is used during the map making step to produce a temperature map. This procedure is very similar to that presented in [8], but here we include more correlated modes in the matrix, especially those related to row and column correlations.

As a first check, we studied the source population of our area and we found that the 148 GHz band fluxes (season 2008) are consistent with [21]. For 218 GHz the differences are larger (~ 15%), presumably because we used more stringent cuts. Also, there are some differences in the map making procedure that could affect the fluxes: the maps used in [21] were made following [8], which considers removing the sources at the beginning of the process, and then reproject them into the maximum likelihood map, procedure which recovers the fluxes to 1% accuracy according to simulations. We do not do this. Additionally, those maps were made with non-downsampled data to avoid artifacts in the power spectrum and improve the point sources SNR as well. Then, our fluxes could be affected by the down sampling during preprocessing and the inclusion of the sources in the noise matrix calculation, downweightening the source fluxes during map making.
The noise properties of the maps are also studied by computing the angular power spectrum (Section 5.2). We found noise level of $28.9 \mu$K-arcmin for 148 GHz, comparable with $29.5 \mu$K-arcmin level found by [5]. For large angular scales, we have excess noise due to the lack of the low frequency modes. Despite this excess, our signal to noise ratio is enough to calibrate the maps using cross correlations. The official 148 GHz ACT maps were calibrated using this technique, therefore we use them to calibrate ours as showed in Section 5.4. The calibration factor for 148 GHz band is $\alpha = 0.956 \pm 0.025$, computed in range $500 < \ell < 2500$. The calibration can be improved through cross-correlations between our maps with maps from the Planck satellite. This has been done in [18] with a 0.7% uncertainty for 148 GHz and 2% for 218 GHz. For the other two bands and seasons, the noise levels that we found are $46.1 \mu$K-arcmin for 218 GHz band, season 2008, and $115.6 \mu$K-arcmin, $105.3 \mu$K-arcmin and $84.6 \mu$K-arcmin for 280 GHz band, seasons 2008, 2009, 2010, respectively.

The map making transfer function and convergence are also explored. We found that the bias in the output is less than 1%. Additionally, we showed that for $\ell \gtrsim 1500$ the convergence is achieved by iteration 50, while larger angular scales converged by iteration 400, consistent with simulations.

As we have seen in this work, our implementation is able to reproduce ACT results. The inclusion of additional row and column correlated modes improved the white noise level and the SNR of point sources. An important progress was made in the case of 280 GHz band, which has been mapped successfully, becoming the first maps of that band. We expect to improve our understanding of this band in order to get better results.
References


References


References


