Sunyaev-Zeldovich power spectrum measurements predictions for upcoming ground-based telescopes

Dissertation
zur
Erlangung der Doktorgrades (Dr. rer. nat.)
der
Mathematisch-Naturwissenschaftlichen-Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von
Maude Charmetant
aus
Mâcon, Frankreich

14th May 2024

Bonn 2023

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der
Rheinischen Friedrich-Wilhelms-Universität Bonn

1. Gutachter: Prof. Dr. Frank Bertoldi
2. Gutachter: Prof. Dr. Cristiano Porciani

Tag der Promotion: 29.02.2024
Erscheinungsjahr: 2024

To my aunt Nadine Poncet, who provided help and tranquillity in the most challenging times.

À ma tante Nadine Poncet, qui m'a apporté aide et quiétude durant les moments les plus difficiles.

Abstract

This work aimed to assess the detectability of the Sunyaev-Zeldovich (SZ) effects for upcoming high-resolution ground-based telescopes such as the Simons Observatory (SO) and Fred Young Submillimeter Telescope (FYST). SZ effects describe how the energy of photons from the Comic Microwave Background (CMB), varies when they encounter galaxy clusters. The thermal Sunyaev-Zeldovich (tSZ) effect is the inverse Compton-scattering of the CMB photons on the hot electrons in the Intra-Cluster Medium (ICM). The kinematic Sunyaev-Zeldovich (kSZ) is a Doppler shift of the CMB photons due to the cluster's motion compared to the CMB. The SZ signals are crucial for galaxy cluster detection, and measurements of cluster characteristics such as pressure, and shock events and to constrain cosmological parameters. High signal-to-noise (S/N) detection and purity of the SZ signals are essential to avoid biased measurements. SO LAT and FYST should provide unprecedented high-resolution maps of the microwave sky. This work estimates their performance on SZ power spectrum detection and purity.

To predict how good, simulations of the microwave sky had to be set up to reproduce their characteristics. Existing template maps from the Websky, SO, or Sehgal et al. (2010a) simulation were used to generate extra-galactic emissions, and the PySM Python package to generate the Galatic emissions. These maps are then brought to specific frequencies, through interpolation or the use of their Spectral Energy Distribution (SED) and processed to account for instrumental noise, beam, and atmospheric noise. This was built as an easy-to-use, versatile, and publicly available Python package to simulate high-resolution maps of the microwave sky, called the Skymodel.

Using the Skymodel simulations of the microwave sky as seen by SO or FYST were generated. The goal was to make predictions of the detectability of the tSZ and purity, compared to its main contaminant, the Cosmic Infrared Background (CIB). A component separation method called Internal Linear Combination (ILC) was used to extract back the tSZ signal from the mock sky. It was found that combining SO and FYST capabilities would reduce the CIB noise residual in tSZ by $\sim 35\%$. When applying a Constrained ILC (CILC) to null the CIB using its SED, it was found that the CIB noise could be further reduced by $\sim 7\%$ when combining SO and FYST but more importantly that the tSZ signal could be detected with an S/N above 1, on a harmonic space scales window ($\ell \in [1700, 3400]$).

To detect the weak kSZ effect signature, it is cross-correlated with the stronger tSZ effect, to form a new estimator $C_\ell^{y y^2_{\rm kSZ}}$. The Skymodel is used to simulate a very simple sky approximation, containing only tSZ, kSZ, CMB, and instrumental noise. A CILC is applied, nulling CMB, to extract a tSZ map without kSZ, and another CILC deprojecting tSZ to extract a CMB map without tSZ and avoid spurious cross-correlation. A Wiener filter is used to separate the kSZ signal out of the CMB map. The result in such a simple case scenario is that a few sigma detections of the signal's power spectrum should be possible with an experiment such as SO.

List of publications

Chapter 1 includes concepts from,

- Basu, Kaustuv ; Erler, Jens ; Chluba, Jens ; Delabrouille, Jacques ; Hill, J. Colin ; Mroczkowski, Tony ; Niemack, Michael D. ; Remazeilles, Mathieu ; Sayers, Jack ; Scott, Douglas ; Vavagiakis, Eve M. ; Zemcov, Michael ; Aravena, Manuel ; Bartlett, James G. ; Battaglia, Nicholas ; Bertoldi, Frank ; **Charmetant, Maude**; Golwala, Sunil ; Herter, Terry L. ; Klaassen, Pamela ; Komatsu, Eiichiro ; Magnelli, Benjamin ; Mantz, Adam B. ; Meerburg, P. Daniel ; Melin, Jean-Baptiste ; Nagai, Daisuke ; Parshley, Stephen C. ; Pointecouteau, Etienne ; Ramos-Ceja, Miriam E. ; Ruszkowski, Mateusz ; Sehgal, Neelima ; Stacey, Gordon G. ; Sunyaev, Rashid (2019), *"SZ spectroscopy" in the coming decade: Galaxy cluster cosmology and astrophysics in the submillimeter*, arXiv:1903.04944

Author's contribution: K.B. wrote the manuscript and M.C. commented it before submission.

Chapter 5 presents results published in,

- CCAT-Prime Collaboration ; Aravena, Manuel ; Austermann, Jason E. ; Basu, Kaustuv ; Battaglia, Nicholas ; Beringue, Benjamin ; Bertoldi, Frank ; Bigiel, Frank ; Bond, J. Richard ; Breysse, Patrick C. ; Broughton, Colton ; Bustos, Ricardo ; Chapman, Scott C. ; **Charmetant, Maude** ; Choi, Steve K. ; Chung, Dongwoo T. ; Clark, Susan E. ; Cothard, Nicholas F. ; Crites, Abigail T. ; Dev, Ankur ; Douglas, Kaela ; Duell, Cody J. ; Dünner, Rolando ; Ebina, Haruki ; Erler, Jens ; Fich, Michel ; Fissel, Laura M. ; Foreman, Simon ; Freundt, R. G. ; Gallardo, Patricio A. ; Gao, Jiansong ; García, Pablo ; Giovanelli, Riccardo ; Golec, Joseph E. ; Groppi, Christopher E. ; Haynes, Martha P. ; Henke, Douglas ; Hensley, Brandon ; Herter, Terry ; Higgins, Ronan ; Hložek, Renée ; Huber, Anthony ; Huber, Zachary ; Hubmayr, Johannes ; Jackson, Rebecca ; Johnstone, Douglas ; Karoumpis, Christos ; Keating, Laura C. ; Komatsu, Eiichiro ; Li, Yaqiong ; Magnelli, Benjamin ; Matthews, Brenda C. ; Mauskopf, Philip D. ; McMahon, Jeffrey J. ; Meerburg, P. Daniel ; Meyers, Joel ; Muralidhara, Vyoma ; Murray, Norman W. ; Niemack, Michael D. ; Nikola, Thomas ; Okada, Yoko ; Puddu, Roberto ; Riechers, Dominik A. ; Rosolowsky, Erik ; Rossi, Kayla ; Rotermund, Kaja ; Roy, Anirban ; Sadavoy, Sarah I. ; Schaaf, Reinhold ; Schilke, Peter ; Scott, Douglas ; Simon, Robert ; Sinclair, Adrian K. ; Sivakoff, Gregory R. ; Stacey, Gordon J. ; Stutz, Amelia M. ; Stutzki, Juergen ; Tahani, Mehrnoosh ; Thanjavur, Karun ; Timmermann, Ralf A. ; Ullom, Joel N. ; van Engelen, Alexander ; Vavagiakis, Eve M. ; Vissers, Michael R. ; Wheeler, Jordan D. ; White, Simon D. M. ; Zhu, Yijie ; Zou, Bugao (2023), *CCAT-prime Collaboration: Science Goals and Forecasts with Prime-Cam on the Fred Young Submillimeter Telescope*, ApJS, 264, 7

Author's contribution: M.C. wrote part of section 8.2 and contributed with Fig. 13; M.C. is a CCAT collaboration member and therefore participates in the development of the SZ science for the FYST telescope.

Chapter 6 presents results published in,

- **Charmetant M.** and Erler J, (2023), *Thermal Sunyaev-Zeldovich measurements and cosmic infrared background leakage mitigation combining upcoming ground-based telescopes*, Astronomy & Astrophysics, 677, A87 (2023)

Author's contribution: M.C. performed the scientific analysis presented in the paper and wrote the text. J.E. coded half of the Skymodel used for the analysis and provided comments on the draft.

Contents

CHAPTER 1

Introduction

Contribution

This chapter presents a general introduction to observational cosmology, some of the themes covered here can be linked to the following publication:

- Basu, Kaustuv ; Erler, Jens ; Chluba, Jens ; Delabrouille, Jacques ; Hill, J. Colin ; Mroczkowski, Tony ; Niemack, Michael D. ; Remazeilles, Mathieu ; Sayers, Jack ; Scott, Douglas ; Vavagiakis, Eve M. ; Zemcov, Michael ; Aravena, Manuel ; Bartlett, James G. ; Battaglia, Nicholas ; Bertoldi, Frank ; **Charmetant, Maude**; Golwala, Sunil ; Herter, Terry L. ; Klaassen, Pamela ; Komatsu, Eiichiro ; Magnelli, Benjamin ; Mantz, Adam B. ; Meerburg, P. Daniel ; Melin, Jean-Baptiste ; Nagai, Daisuke ; Parshley, Stephen C. ; Pointecouteau, Etienne ; Ramos-Ceja, Miriam E. ; Ruszkowski, Mateusz ; Sehgal, Neelima ; Stacey, Gordon G. ; Sunyaev, Rashid (2019), *"SZ spectroscopy" in the coming decade: Galaxy cluster cosmology and astrophysics in the submillimeter*, arXiv:1903.04944

K.B wrote the paper, M.C provided comments on the manuscript before submission.

This chapter presents the notions, definitions, historical background, and context of this work starting from the first seconds of the Universe, through the first light to the formation of the first complex structures and elements (see Sec.1.1). Subsequently, this chapter describes how as the Universe itself developed, some matter aggregates became complex enough to reflect on their own existence and attempted to observe, characterize, and model the framework in which they evolved through observations (see Sec. 1.2). Furthermore, the theories hinting at the existence of the Universe's first light, the Cosmic Microwave Background (CMB) and its discovery (see Sec. 1.3), how under gravity the first complex structures formed (see Sec. 1.4) and how these structures are influencing the CMB light (see Sec. 1.5) are discussed.

1.1 The basics of cosmology

The study of the Universe as a whole, of its history and properties is known as cosmology. The term originates from the Greek 'cosmos' meaning world and 'logia' meaning the study of. Cosmologies have always existed but were initially composed of beliefs, either pagan or religious. Their goal was sometimes

to explain factual events related to stars and planets but were often lacking the necessary scientific method or techniques to observe the phenomena and dissociate them from mystics. Scientific and modern cosmology in astrophysics is commonly accepted as having started with Einstein's theory of general relativity (Einstein, 1916). Technical advancements allowed observations and scientific measurements that supported cosmological theories and principles. Thus, allowing cosmology to become an intrinsic part of modern science.

1.1.1 History of the Universe

The current and most widely accepted model supposes that the Universe was originally in a very dense and very hot state approximately 13.77 ± 0.04 billion years ago (Planck Collaboration et al., 2020). (Guth, 1981) posited that an exponential period of expansion followed, henceforth called the inflation era. After the inflation, the Universe continued to grow and expand slowly. The primordial plasma, density, and temperature remained so high that the light was coupled to the particles. This means that the light particles, photons, had a mean free path (i.e. length between two interactions) so short, that they would immediately be absorbed by the free electrons of the plasma, and after their re-emission, would be immediately absorbed again. However, with time and the expansion of the Universe, the primordial plasma would eventually become cold enough for atoms to form, this very moment is called 'recombination'. As the electrons became bound around the nucleus of the atom, the mean-free path of the photons became much larger allowing light to propagate throughout the Universe for the very first time. This first light can still be observed to this day and is known as the Cosmic Microwave Background (CMB). The initial tiny density fluctuations, brought to bigger scales by inflation are visible in the CMB and eventually lead to gravitational potential wells, accreting matter as they evolve leading to the formation of the first stars. The radiation coming from the first stars is going to create ionised regions around their host and slowly cover more and more areas of the Universe. This period is known as the reionization epoch. Over time, more gas was accreted and more complex systems started to form, galaxies, groups of galaxies, and eventually galaxy clusters. The structures are supposed to form following a hierarchical building pattern, meaning small structures grow in size with time through the accretion of more and more structure, gas, and matter. An illustration of this timeline can be seen in Fig. 1.1. More details about each specific epoch are given in the following subsections.

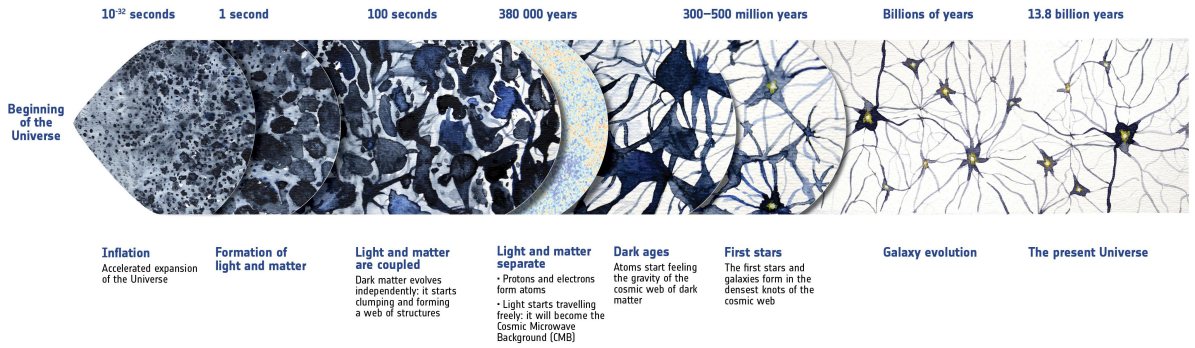

Figure 1.1: Schematic of the History of the Universe [1].

[1] https://www.esa.int/ESA_Multimedia/Images/2015/02/The_history_of_the_Universe

The Big Bang

The idea of the entire Universe originating from a singular, point-like, dense state was first formulated by Georges Lemaître, a Belgian physicist (Lemaître, 1931). His idea emerged from several observations of many of his colleagues and the general questions raised during the previous decade or so. In particular, in 1910 Vesto Slipher, an American scientist, was the first to claim that galaxies were receding away from Earth. A few years later the theory of general relativity developed by Albert Einstein was found to only admit solutions for an expanding or shrinking Universe. In 1922, a Russian scientist Alexander Friedmann was the first to derive the solution of Einstein's equations for a non-static Universe (Friedmann, 1922). In 1929, the American scientist Edwin Hubble formulated his law, that galaxies are moving away from Earth, and the velocity at which they are moving away is proportional to their distance from Earth,

$$v = H_0 D, \tag{1.1}$$

where v is the velocity at which the object is moving away from Earth, H_0 is the Hubble constant at the current time, and D is the distance to the object, all in S.I. units. The debate between a static Universe that always existed and a Universe originating from a much denser state through the *Big Bang* theory was set. Only recently with modern techniques and observations did the paradigm shift toward a wide acceptance of the Big Bang theory by the scientific community.

The inflation period

The inflation theory was developed in the 70s and 80s with significant contributions from Alexei Starobinsky, Alan Guth, and Andrei Linde (Guth, 1981). Inflation was proposed as a solution to a number of the problems raised by the Big Bang hypothesis of the Universe originating from a point-like state. Among the biggest problems, are the horizon and the flatness problem raised by observations. The idea behind the horizon problem is that because light has a finite speed and everything is receding away from Earth faster and faster as it goes away If the Universe did not originate from a smaller state, then two regions of the sky sufficiently spaced would originate from regions of the Universe that were not causally connected. However, this contradicts several observations such as the CMB, where the temperature all over the sky is the same up to 10^{-5}. It seems extremely unlikely that this would happen over the full sky randomly. However, if the Universe was previously much smaller and dense, then there is no causality problem between different regions of the sky as they all originate from the same source and were only brought apart by a fast expansion phase. This inflation phase would also explain how tiny density perturbations were brought up to much bigger scales to form the large structures that can be observed nowadays.

Big Bang nucleosynthesis

At first, the formation of the chemical elements in the Universe was thought to have been fully produced in stars or star explosions called supernovae (Burbidge et al., 1957). The problem with this theory is that it would have produced too little Helium compared to what was observed. Alpher–Bethe–Gamow had earlier postulated that some elements were produced at the time of the Big Bang (Alpher, 1948), which was later confirmed by observations. In the early stages of the Universe, the primordial plasma was composed of photons, neutrinos, neutrons, protons, and electrons. During the first few minutes of the Universe, following inflation, the plasma became cold enough for some atoms to start forming.

Proton and neutrons formed Deuterium atoms, from there, Tritium and Helium-3 could be formed, and later Helium-4. Heavier elements such as Beryllium and Lithium were also formed. However, those elements are not formed at the same time, while the formation of Deuterium is pretty straightforward, the high-temperature photons destroy the newly formed Deuterium thus preventing heavy elements from forming. For this reason, Deuterium is known as the bottleneck of nucleosynthesis. Only a few minutes after the Big Bang the primordial plasma was cold enough for Deuterium to persist and start building Helium. The Deuterium and Helium content increases around $t \sim 3$min after the Big Bang and starts to decrease after that as heavy elements such as Lithium and Beryllium start to form out of Helium (see Fig. 1.2). Three minutes after the Big Bang, the Universe's matter content was 25% Helium, 75% of Hydrogen, and some traces of Lithium, Deuterium, and Helium 3.

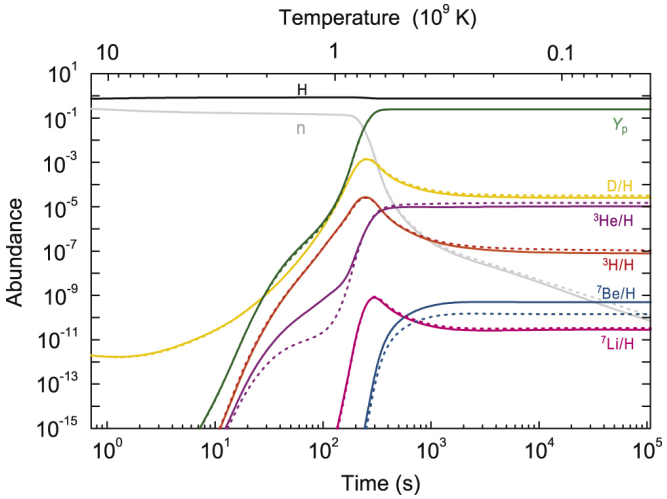

Figure 1.2: Evolution of the abundances of elements in the Universe as a function of time and temperature. Extracted from (Hou et al., 2017).

Light-matter decoupling

After the supposed inflation phase, the Universe is in a hot and dense plasma state. The mean free path, meaning the length between two interactions, of photons is shorter than the size of an atom. Therefore, photons immediately scatter on the free electrons. Due to the expansion of the Universe the plasma that composes it, cooled down with time. When it became cold enough for atoms to form, the free electrons became bound into atoms. They freed the photons, from immediately scattering off the electrons. At that time, around 380,000 years after the Big Bang the Universe became "transparent", the decoupling of the matter-radiation fluid allowed for the photons to travel freely. This first observable light shows the density perturbations in the early Universe (see Fig.1.3) and is called the Cosmic Microwave Background (CMB). While the light emitted at this time was not in the microwave regime, it is presently detected at these wavelengths due to being redshifted due to the expansion of the universe. It is called Background because it is the oldest visible light, meaning every other light is in the foreground in comparison. George Gamov was the first to mention the "condensation" of the "primordial gas" and hint at the possibility of the last scattering surface (Gamow, 1948).

[2] https://www.esa.int/ESA_Multimedia/Images/2013/03/Planck_CMB

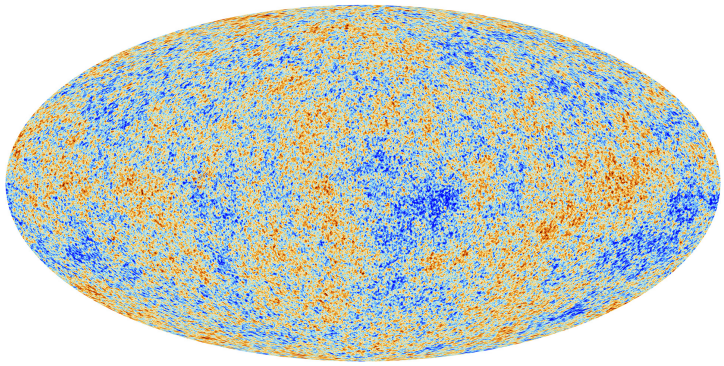

Figure 1.3: CMB map, showing the density fluctuations in the Universe projected on a 2D map. Extracted from measurements of the *Planck* space satellite, courtesy of ESA. The red areas represent the regions of the sky that are hotter and the blue ones colder [2].

From the Dark Ages to the firsts galaxies

Following the temperature fluctuations visible in the CMB, matter started to aggregate into potential wells. After millions of years, the first stars started to form and with their stellar winds, ionize the surrounding area. In their core, through thermonuclear fusion, new heavy elements not formed during Big Bang Nucleosynthesis started to form. In astrophysics, all elements heavier than hydrogen are called 'metals'. Once the star dies, those metals are ejected into the surrounding environment, thus enriching the Universe. The stars formed out of this enriched gas are more metal-rich and can produce heavier metals than their progenitors. With the number of stars increasing and their surrounding ionization bubble with them, the Universe is slowly being reionized. This period is known as the reionization period. Structures form by accretion over time, meaning from small to large. This is known as the hierarchical structure formation model in opposition to other models that suppose that bigger structures form first and then are separated into smaller clumps. The existence of this large-scale structure of the Universe was proved through observations of galaxy distribution. The largest to-date galaxy distribution survey is the Sloan Digital Sky Survey (SDSS) (Lundgren and et al., 2015). The distribution of galaxies clearly exhibits a particular structure (see Fig. 1.4) resembling a web-like pattern.

1.1.2 Formalism of cosmology

In this part, all the basic equations, derivations, and formalism on which cosmology is based are presented and interpreted. The interpretation of those equations gives information on the History of the Universe and its future.

Derivation of Friedmann's equations

All of modern cosmology formalism can be derived from Einstein's equations of general relativity

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R - \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}, \tag{1.2}$$

where $R_{\mu\nu}$, the Ricci tensor, characterizes how much a volume element changes in a curved space in comparison to a flat Euclidian space. A Euclidian space is a space where the geometry is the classical one

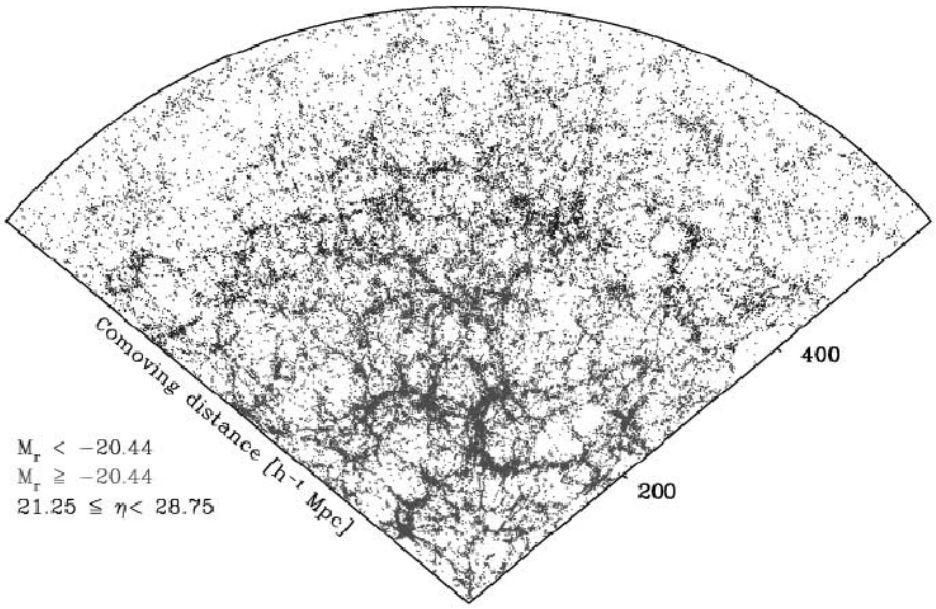

Figure 1.4: Galaxy distribution within a cone, as measured by the SDSS survey. Extracted from Matsubara (2006).

of a 2-dimensional paper sheet. $g_{\mu\nu}$ is the metric tensor, it defines the scalar product of two vectors and especially how distances and angles are measured in the particular space represented. It can be understood as a generalization of Pythagoras' theorem. R is the scalar value of the Ricci tensor. Λ is called the cosmological constant. G is the universal gravitational constant and $T_{\mu\nu}$ is the energy-momentum tensor, its expression is,

$$T_{\mu\nu} = \left(\rho + \frac{p}{c^2}\right) u_\mu u_\nu + p g_{\mu\nu}, \tag{1.3}$$

where ρ is the density of the medium, p is the pressure, and u_a is the macroscopic speed of the medium. Even though modern cosmology started with Einstein's equations, some premises of it were formulated by Newton. One of those premises is the cosmological principle stating that the Universe must be isotropic and homogeneous (Newton, 1687). Alexander Friedmann decided to place himself in the cosmological principle framework to solve Einstein's equation, additionally supposing that the Universe is in a period of expansion. A solution to the equation is then provided by the following metric:

$$ds^2 = c^2 dt^2 - a(t)^2 \left(\frac{1}{1 - kr^2} dr^2 + r^2 d\theta^2 + r^2 sin^2\theta d\phi^2\right), \tag{1.4}$$

where ds^2 is the Friedmann-Lemaître-Robertson-Walker (FLRW) metric that describes the geometry of the space, It is a 4-dimensional vector, with one dimension of time and 3 of space. c is the speed of light. t is the time coordinate. $a(t)$ is the scale factor that measures how the distance between two objects varies with time purely due to the Universe's expansion. k is the curvature of space. (r, θ, ϕ) are spherical coordinates used (see Fig. 1.5). Through the computation of the Ricci tensor, Ricci scalar, and the hypothesis of isotropy, the Friedmann equations can be derived:

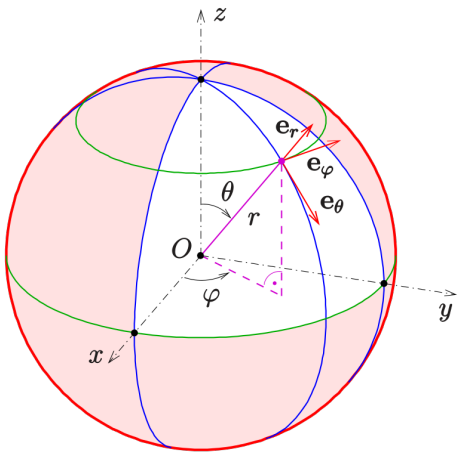

Figure 1.5: Spherical coordinates parameters. Suppose a purple point on the sphere, is defined by its radius r to the centre of the sphere O, through its angle θ and ϕ. e_r, e_ϕ, e_θ are the units vectors of the spherical coordinates.[3]

$$\left(\frac{\dot{a}(t)}{a(t)}\right)^2 = \frac{8\pi G}{3}\rho(t) - \frac{kc^2}{a^2(t)} + \frac{\Lambda c^2}{3}, \tag{1.5}$$

and

$$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{3}\left(\rho(t) + \frac{3p(t)}{c^2}\right) + \frac{\Lambda c^2}{3}. \tag{1.6}$$

Equation of state

By differentiating Eq. 1.5, injecting it back into Eq. 1.5 to get rid of k, and using Eq. 1.6 to get rid of $\ddot{a}(t)$ and Λ, the following continuity equation can be derived:

$$\dot{\rho}(t) + 3\frac{\dot{a}(t)}{a(t)}\left(\rho(t) + \frac{p(t)}{c^2}\right) = 0. \tag{1.7}$$

If one assumes the following equation of state $p(t) = w\rho(t)c^2$ and injects it into the continuity equation, one gets:

$$\rho(t) \propto \frac{1}{a(t)^{3(1+w)}}. \tag{1.8}$$

There are three particular solutions to this equation,

- w=0: this is the case of pressureless matter, this can be assumed in the non-relativistic approximation. It yields $\rho(t) \propto 1/a(t)^3$, which means that the density decreases as the volume of the Universe increases.

[3] https://en.wikipedia.org/wiki/Spherical_coordinate_system#/media/File:Kugelkoord-lokb-e.svg

- w=1/3: is the radiation case with $\rho(t) \propto 1/a(t)^4$, the density of radiation in the universe is also inversely proportional to volume but has an additional $1/a$ term coming from the photon energy E=h/λ being dimensionally equivalent to an inverse distance.
- w=-1: is the case of a negative pressure according to the equation of state, as expected from the cosmological constant. For this fluid, the density is $\rho(t) \propto a(t)^0$, meaning constant.

Simplification of the Friedmann equations

The following parameter change can be made to simplify the third term on the right of the equal sign in the Friedmann equations

$$\rho_\Lambda(t) = \frac{\Lambda c^2}{8\pi G}, \tag{1.9}$$

and

$$p_\Lambda(t) = -\rho_\Lambda(t)c^2. \tag{1.10}$$

Equation 1.10 shows that the cosmological constant can be interpreted as a fluid of negative pressure. These transformations, simplify the equations to

$$\left(\frac{\dot{a}(t)}{a(t)}\right)^2 = \frac{8\pi G}{3}\rho(t) - \frac{kc^2}{a^2(t)}, \tag{1.11}$$

and

$$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{3}\left(\rho(t) + \frac{3p(t)}{c^2}\right). \tag{1.12}$$

Dynamics of the Universe

The first reduced Friedmann equation (see Eq. 1.11) describes the expansion speed as a function of the scale factor, the density, and the curvature. There are several cases that can be derived and that give us the evolution of the universe

- k=1: in this case, the universe would increase in size, each a maximum, and eventually start decreasing in size and recollapse. This is known as the closed Universe (see Fig. 1.6).
- k=-1: known as the spatially open Universe (see Fig. 1.6), in this case, the Universe would expand forever.
- k=0: the Friedmann equation gives,

$$\rho_c(t) = \frac{3H^2(t)}{8\pi G}. \tag{1.13}$$

  where $H(t) = \dot{a}(t)/a(t)$ is the Hubble parameter. Describing how fast the Universe grows with respect to its size at a time, t, and ρ_c, the critical density of the Universe. In this case, the Universe is described as flat (see Fig. 1.6).

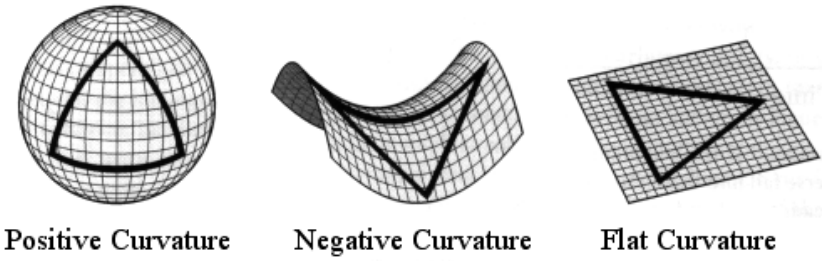

Figure 1.6: Representation of the curvature of a 3-dimensional space[4].

1.1.3 Cosmological paragidm

From the theories and observations emerged the Λ Cold Dark Matter (ΛCDM) model, also called the standard model of cosmology. It describes the content and history of the Universe, which is composed of Dark Energy, Dark Matter, Baryonic or classical matter with a flat geometry, an inflation period after the Big Bang, and currently, an accelerated expansion.

Dark energy

The first parameter of the standard model of cosmology, Λ, is Einstein's cosmological constant. This constant was first added to provide static Universe solutions to Einstein's equation (see Eq. 1.2). After Lemaître, Friedman, and Hubble's contributions to the field, the consensus became that the constant was null because of the expansion of the Universe. With modern observations (Riess et al., 1998; Perlmutter and et al., 1998) showing that the expansion is accelerating, the cosmological constant was reintroduced as a positive constant embodying the acceleration of the expansion. The physical origin of what is driving the accelerated expansion is not known and was therefore named 'Dark Energy'. Current measurements find that ~69% of the known Universe is composed of Dark Energy (see Fig. 1.7).

Cold dark matter

The Cold Dark Matter (CDM) component of the standard model emerged through observations that the galaxies' rotation curves did not behave as predicted. Specifically, the velocity of material in a given galaxy did not decrease with distance to their respective galactic centres, as expected from observations of light coming from ordinary matter. This indicates that the density is decreasing. The first measurement of such an anomaly was made in the galaxy named M31 (van de Hulst et al., 1957). Interpretations hinting that the galaxies might have a more extended mass distribution, which does not emit light were made by several scientists (Schmidt, 1957). Even though much earlier, Jan Oort had already hinted at the fact that the motions of stars in the Milky Way could not be explained purely by visible matter (Oort, 1932). Only a year later, Fritz Zwicky made similar observations in the Coma galaxy cluster (Zwicky, 1933). He was the first to name this additional matter "dunkle Materie" (Dark Matter). Dark matter is called 'dark' because it has not been observed to interact with the electromagnetic field, meaning it does not interact with light, neither absorption nor emission. It is seen to only interact gravitationally and form haloes in which the baryonic matter clumps together. Its physical origin is, however, unknown despite

[4] https://pages.uoregon.edu/jschombe/cosmo/lectures/lec15.html

various existing theories. Current measurements find that ~ 25% of the Universe is composed of Dark Matter (DM) (see Fig. 1.7). In the standard model of cosmology DM is supposed to be 'cold', i.e. of low speed, to match with structure formations theories and observations. A consequence of the CDM is the hierarchical formation model, where baryonic matter is accreted in the potential wells formed by dark matter halos, forming small structures first and developing into bigger ones later on through accretion and gravitational collapse. Simulations of hot dark matter, with high relativistic velocities, would prevent the formation of structures smaller than superclusters because of the high velocities counterbalancing the gravitation. This contradicts observations of the Universe and therefore the CDM model is preferred.

Baryonic matter

Ordinary matter, also called baryonic matter, is the matter that not only composes us humans, but all atoms. Baryonic matter interacts with the electromagnetic spectrum and with gravity. Current observations find that ~5% of the Universe is composed of this matter (see Fig. 1.7). Baryonic matter follows gravitational potentials and accretes itself over time to form structures such as stars, solar systems, galaxies, galaxies clusters, and superclusters.

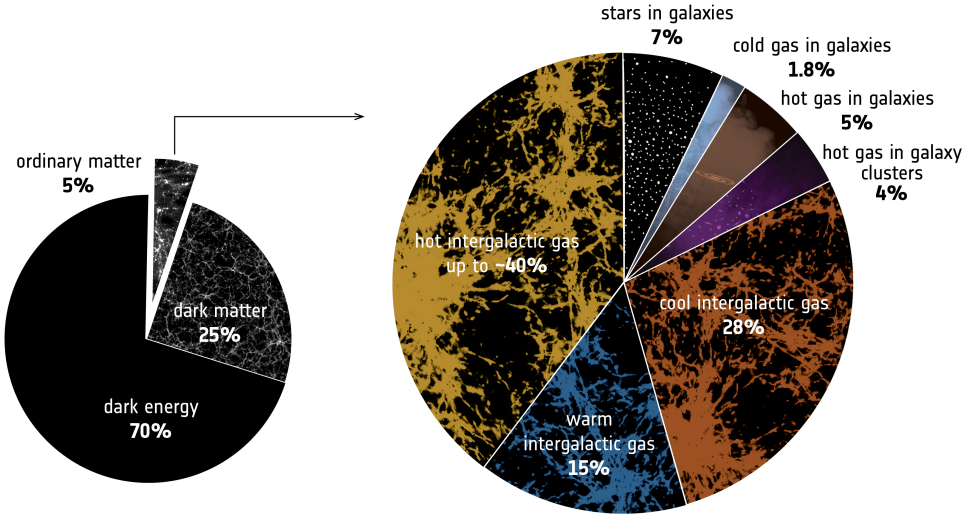

Figure 1.7: Composition of the Universe[5].

ΛCDM conclusion

The success of the standard model is seen in its ability to explain the existence and formation of the large-scale structures of the Universe but also of the CMB, and of the observed abundance of the first

[5] https://www.esa.int/var/esa/storage/images/esa_multimedia/images/2018/06/the_cosmic_budget_of_ordinary_matter/17555274-1-eng-GB/The_cosmic_budget_of_ordinary_matter_pillars.jpg

elements present after the Big Bang nucleosynthesis and the accelerated expansion of the Universe. It is based on 6 independent parameters, the age of the Universe (t), the density of baryons (Ω_b), and dark matter (Ω_c), the optical depth at the reionization period τ, the spectral index (n_s) that describe the variations of the primordial fluctuations with the scales and the curvature fluctuations amplitude (Δ^2_R). ΛCDM is the current widely accepted model of cosmology but it is not free from tensions and challenges.

1.1.4 Dynamic evolution of the Universe

Using the ΛCDM model (see Sec. 1.1.3) and the Friedmann equations (see Eq. 1.11 & 1.12) the dynamics of the Universe can be deduced. It is assumed that the total density, can be decomposed into the sum of the matter (baryonic+dark matter), dark energy, curvature, and radiation densities $\rho(t) = \rho_m + \rho_\Lambda + \rho_k + \rho_r$. Using the critical density (see Eq. 1.13), the first Friedmann equation (see Eq. 1.11), and the definition of the Hubble parameter, the following equation can be deduced

$$H^2(t) = H_0^2 E(a), \tag{1.14}$$

with

$$E(a) = \left(\frac{\Omega_{m,0}}{a^3} + \frac{\Omega_{r,0}}{a^4} + \frac{\Omega_{k,0}}{a^2} + \Omega_{\Lambda,0} \right), \tag{1.15}$$

where $\Omega_{i,0} = \rho_i(t)/\rho_{c,0}(t)$ is the dimensionless density of the quantity i as of today $t = 0$. From this equation, several epochs and models of the Universe can be computed.

Matter dominated Universe

For a Universe containing only matter (baryonic and dark matter), $\Omega_m = 1$ and $\Omega_r = \Omega_k = \Omega_\Lambda = 0$. Then Eq. 1.14 becomes

$$\frac{\dot{a(t)}}{a(t)} = H_0 \frac{1}{a(t)^{3/2}}, \tag{1.16}$$

giving

$$a(t) = \left(\frac{3}{2} H_0 t \right)^{2/3}. \tag{1.17}$$

This yields a solution of the Universe expanding with time, with an expansion rate slowing down. This model of the Universe is known as the Einstein-de Sitter.

Radiation dominated Universe

If the Universe would contain only radiation, meaning $\Omega_r = 1$ and $\Omega_m = \Omega_k = \Omega_\Lambda = 0$. Then Eq. 1.14 would yield, the following expression of the scale factor,

$$a(t) = \left(2 H_0 t \right)^{1/2}. \tag{1.18}$$

In this case, the Universe also expands with time and the expansion decelerates faster than in the Einstein-de Sitter case.

Dark Energy dominated Universe

For a Universe containing only Dark Energy $\Omega_\Lambda = 1$ and $\Omega_m = \Omega_k = \Omega_r = 0$. Equation 1.14, gives the following scale factor:

$$a(t) = e^{H_0 t}. \tag{1.19}$$

In this case, the Universe's expansion is accelerated. The different area can be seen on Fig. 1.8[6].

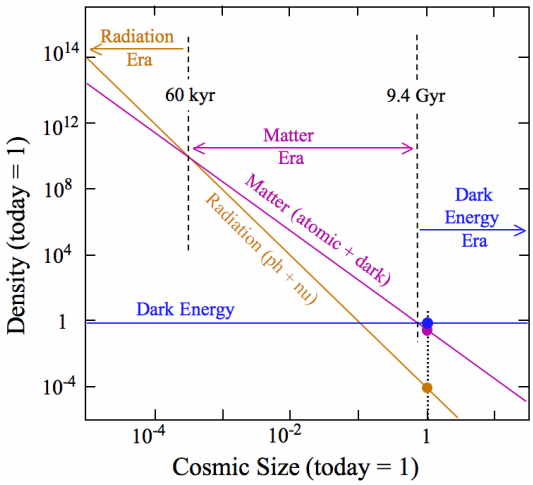

Figure 1.8: The evolution of the different epochs in the Universe depending on the density and as a function of time. For reference see footnote.

1.2 Observations of the Universe

Theories and formalism must be confirmed or excluded by observations. This work is included in a sub-field of cosmology called observational cosmology, i.e. based on observations of the Universe, with the goal of improving the understanding of processes in the Universe. Until recently, light from the electromagnetic spectrum (see Fig. 1.9) was the only tool astrophysicists had to gain information on the Universe. Recently gravitational waves, which are ripples of the space-time metric (see Eq. 1.4) can also be used to observe astrophysical phenomena. In this work, only light-based observations will be discussed. In particular, this work focuses on microwaves and sub-millimetre waves.

[6] https://physics.stackexchange.com/questions/184045/why-is-the-ratio-dark-matter-normal-matter-bigger-today-than-in-the-past-is-i

[7] https://www.japanistry.com/electromagnetic-spectrum/

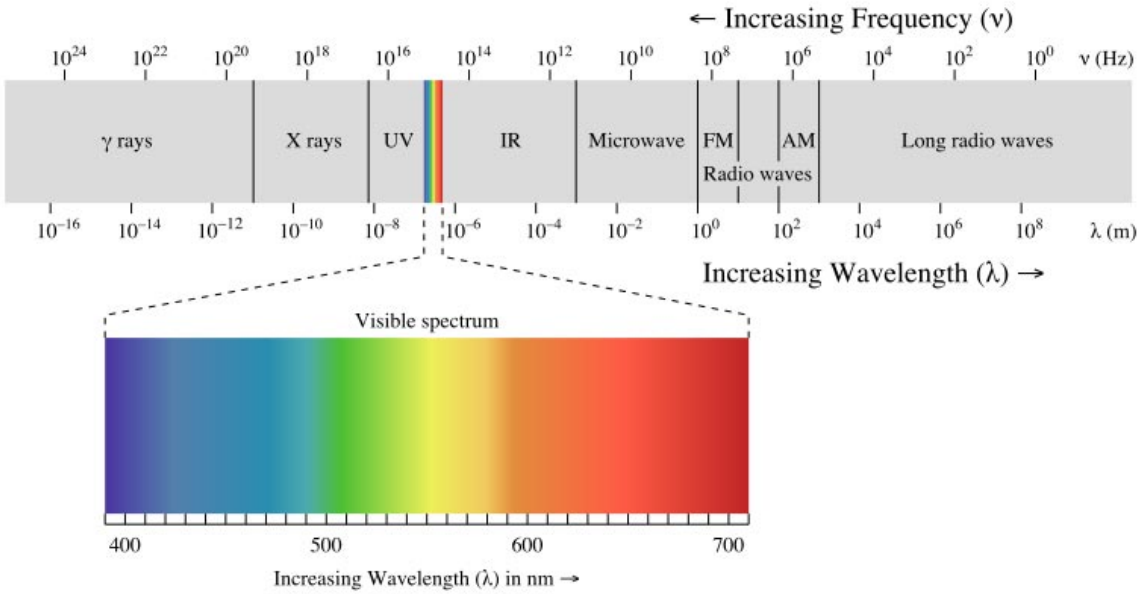

Figure 1.9: Electromagnetic spectrum[7].

1.2.1 Radio telescopes

The millimetre and sub-millimetre part of the electromagnetic spectrum is probed with antennas called radio telescopes. Those antennas are composed of a parabolic dish that focuses all the radio wave light toward a secondary reflector and then toward the receiver (see Fig. 1.10).

Theory of radio telescopes

The radio waves measured by antennas are expressed in spectral flux density which can be interpreted as the rate of energy transferred per unit area and per wavelength. It is measured in Watts per area per Herz [$Wm^{-2}Hz^{-1}$] or more commonly in Jansky [Jy], because

$$1\text{Jy} = 10^{-26}\text{Wm}^{-2}\text{Hz}^{-1}. \tag{1.20}$$

However, objects observed in the sky are not always point-like, they are often extended. In this case, the spectral flux density per steradian [sr] is measured and denoted I_ν [Jy sr^{-1}]. A steradian is the unit of a solid angle Ω defined as follows

$$\Omega = \frac{A}{r^2}, \tag{1.21}$$

where A is the area of the sphere cut by the angle and r is the sphere's diameter. This definition of the solid angle only applies to a flat Euclidian space and is not used as such for cosmology. The power the

[8] https://i1.wp.com/lh6.googleusercontent.com/-DCJXdAqrO-k/T9G4eRfvHvI/AAAAAAAAAsA/bdXVhUqvnFQ/s640/altazimuth%2520mounting%2520-%2520radio%2520telescope.jpg?ssl=1

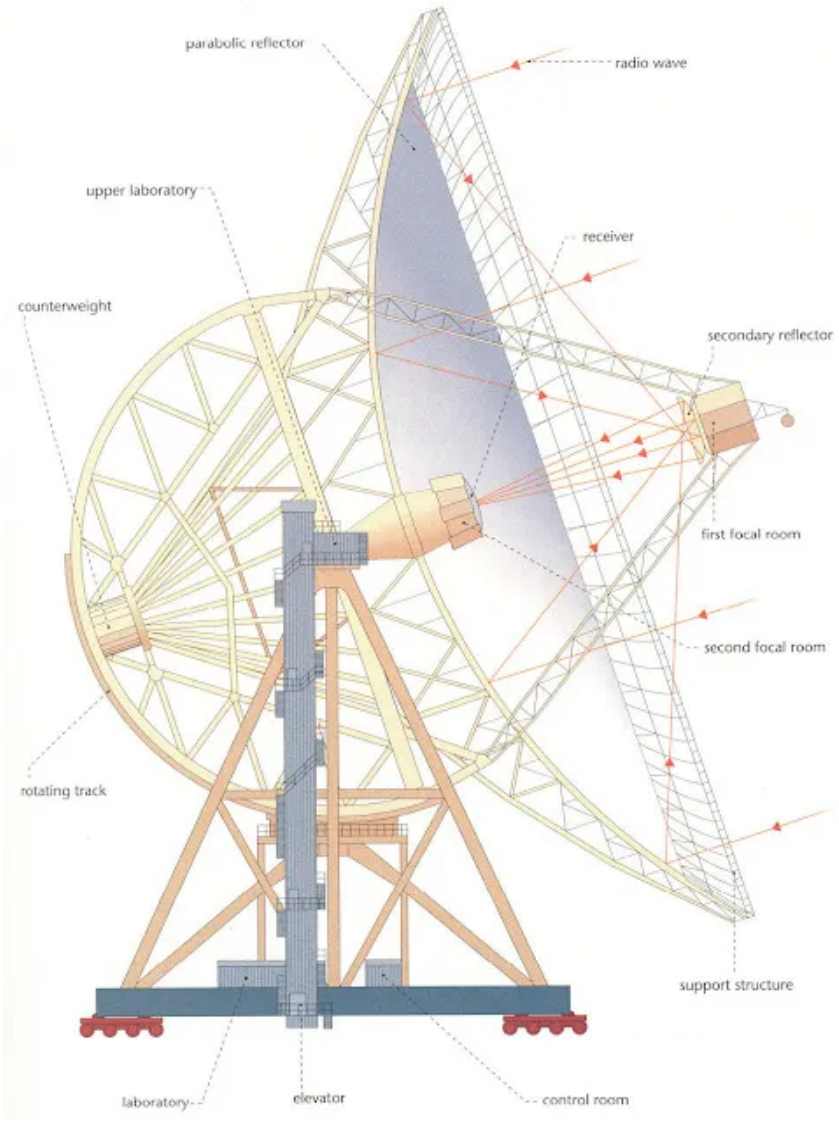

Figure 1.10: Example schematic of a radio telescope. Precisely, a schematic of the steerable Effelsberg telescope in Germany[8].

dish receives depends on the incidence of the rays. The solid angle of the telescope is not as simple as defined in Eq. 1.22 but exhibits a pattern (see Fig. 1.11). Therefore, the antenna solid angle is given by

$$\Omega_A = \int_{4\pi} P(\theta, \phi) d\Omega, \tag{1.22}$$

where θ and ϕ are the spherical coordinates and Ω the 'perfect' solid angle defined in Eq. 1.22. Because the telescope's beam exhibits such a pattern, a point-like object will not be seen as a point but will exhibit a diffraction pattern (see Fig. 1.12).

[9] https://greenbankobservatory.org/wp-content/uploads/2020/06/Radio-Telescope-Fundamentals.pdf

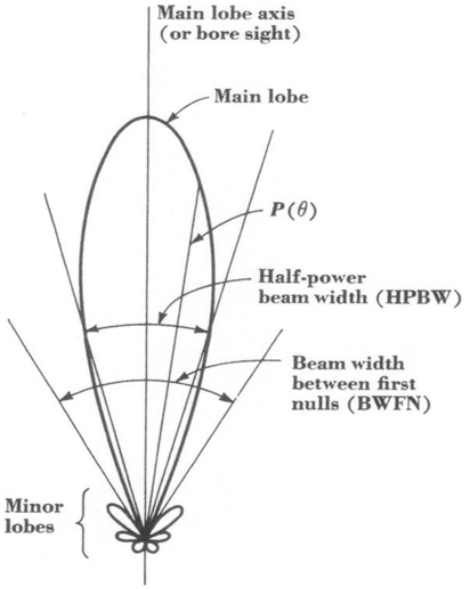

Figure 1.11: Beam pattern schematic. The beam width between the two nulls is defined at the beam tangent to the minor lobes[9].

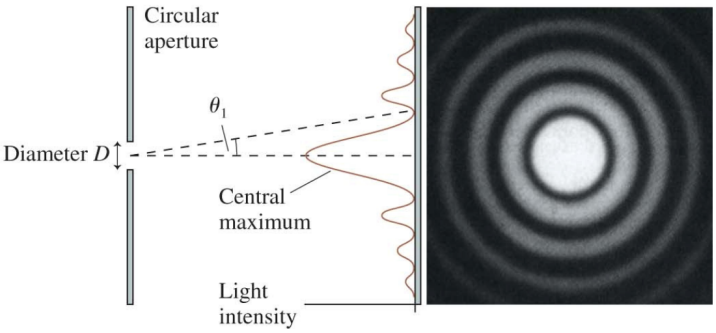

Figure 1.12: Diffraction pattern generated by the beam of the telescope when observing a point-like source. Courtesy of Pearson Education, Inc.

The angular resolution of a telescope, meaning the smallest angle two objects can be separated from each other in order not to be mixed up by the telescope, is given by

$$\theta \approx 1.22\frac{\lambda}{\mathrm{D}}, \tag{1.23}$$

where θ is the angular resolution, λ the wavelength at which the signal is being observed and D the diameter of the telescope. This formula shows the interest in having large-diameter telescopes, the larger the diameter, the smaller the angular resolution for a given wavelength. Smaller angular resolution means the ability to clearly distinguish between two sources that are very close to each other, meaning finer details can be observed. When observing the sky, there are two types of astrophysical sources. Firstly, point sources are objects that are not resolved by the telescope beam, meaning they appear like a point source even though they might be more extended in reality. Secondly, diffuse components are

extended objects, such as the InterStellar Medium (ISM).

History of radio telescopes

The field of radio astronomy started with Karl Jansky who observed the first radio emission, originating from outer space, in 1933. While working at the Bell Telephone Laboratories, he measured the radio emission of the Milky Way (Jansky, 1933). In 1937 an amateur astronomer Grote Reber built the first radio telescope antenna in his backyard. The development of radio astronomy and antennas was greatly accelerated during World War II. Some notable radio telescopes include the 100-m diameter radio telescope, built in 1971 in Effelsberg in Germany (see Fig. 1.10). It is the second largest fully-steerable telescope after the slightly larger Green Bank Telescope (GBT) built in West Virginia, USA. The GBT is elliptical instead of circular with a 100×110 m size, making it slightly larger than the Effelsberg telescope. Much larger telescopes exist, such as the Arecibo Radio Telescope built in a sinkhole in Puerto Rico and exhibiting a diameter of 305 m. It was built in 1963 and collapsed in 2020 due to a lack of maintenance and interest in maintaining it. The largest radio telescope to date is the Five-hundred-meter Aperture Spherical Radio Telescope (FAST) with a 500 m diameter dish built in a natural basin in China and started operating in 2016.

Building large diameter telescopes is complicated, and above $\sim$ 100 m they are so far never steerable which limits the possible observations. A solution to that is to build what is called an 'array' of antennas, meaning many smaller diameter telescopes that are steerable, space them far from each other, and use interferometry. The space between the antennas is equivalent to the diameter of a 'virtual' larger telescope. A notable example of such interferometer arrays is the Atacama Large Millimeter Array (ALMA) located in the Atacama desert in Chile. It started observations in 2011 and is the result of a collaboration between Europe, Chile, the United States of America, Canada, Japan, Taiwan, and South Korea. ALMA is composed of 66 antennas, 54 of which have a diameter of 12 m and 12 a diameter of 7 m. Its most compact configuration gives a baseline (distance between a given pair of antennas) diameter of 150 m and its most extended configuration of 16 km. Interferometry is not used in this project but illustrates an important aspect of radio telescopes, which is the larger the antenna diameter, the finer the resolution.

1.2.2 Measuring distances

Once the light emitted by an object has been measured, an important piece of information is the distance to this object. Measuring and defining distances in an expanding Universe is one of the challenges in cosmology.

Spectral shift

In 1842, Christian Doppler an Austrian physicist, predicted that for both acoustic and electromagnetic waves, there should be a wavelength shift when the source is in movement with respect to the observer. This effect is known as the Doppler effect and can be easily experienced in everyday life. When an ambulance is passing by, as it is getting closer one hears the frequency of the alarm increasing, and once it passes the observer and is driving away, the observer can hear the frequency decreasing. In 1845 the effect was observed experimentally by the Dutch physician Christoph Buys Ballot (Ballot, 1845). It is observed for the first time on stars by Huggins (1868). Because of the first observations that galaxies

are receding away from us (see Sec. 1.1.1), it became apparent that the wavelength of their light would increase. The term redshift (z) was defined as,

$$z = \frac{\lambda_{\text{obs}} - \lambda_{\text{emi}}}{\lambda_{\text{emi}}}, \tag{1.24}$$

where λ_{obs} is the observed wavelength of the object and λ_{emi} the intrinsic, emitted, wavelength of the object. The intrinsic wavelength of an object, can not be observed obviously but can be determined through the emissions lines of particular molecules. Indeed, for a given molecule the line emission is found at a precise wavelength (λ_{emi}) value which can be determined on Earth in a laboratory. In finding the spectrum of the observed object in the sky, the difference between the observed wavelength and its laboratory value ($\lambda_{\text{obs}} - \lambda_{\text{emi}}$) experienced by the line due to the Doppler effect can give us the source's redshift value. In cosmology, this redshift can have various origins, it can come from the peculiar velocity of the object itself, e.g., a galaxy is moving away from us due to its velocity. The redshift can also be due to the expansion of the Universe, sometimes called, cosmological redshift. In that case, the redshift of the galaxy observed is due to the universe expanding and pushing the galaxy away from the observer, independently of the galaxy's own velocity. Gravitational redshift originates from massive objects distorting the space-time metric and therefore increasing the wavelength of light of the object observed. The Hubble's equation (see Eq. 1.1) can be rewritten with the redshift as,

$$zc = H_0 D. \tag{1.25}$$

The redshift expression can also be derived directly from the metric expression (see Eq. 1.4) by integrating it over space and time, it yields,

$$1 + z = \frac{1}{a(t)} \tag{1.26}$$

However, the redshift is not a perfect indicator of the distance to an object because it does not take into account peculiar velocities (the velocity of an object relative to a rest frame).

Comoving distance

The comoving distance between two objects is always the same, independently of the expansion of the Universe, it is defined as,

$$D_{\text{com}}(z) = \frac{c}{H_0} \int_0^z \frac{1}{E(z')} dz'. \tag{1.27}$$

Proper distance

The proper distance is the distance to an object measured at a given time. The proper distance to any object will increase with time, as the universe expands, it is expressed as,

$$D_{\text{pro}}(t) = a(t) \int_0^r \frac{1}{\sqrt{1 - kr^2}} dr. \tag{1.28}$$

Angular distance

If the geometry of the Universe is flat, then the distance to an object can be measured through the ratio of its true intrinsic size with respect to the angular size that the object appears to be in the sky.

$$D_{\text{ang}}(z) = \frac{S}{\Omega}, \tag{1.29}$$

where S is the true surface of the object and Ω is the observed surface, the solid angle of the object projected on a sphere. When the true surface of the object is not known, the redshift and the comoving distance can be used to compute the angular distance,

$$D_{\text{ang}}(z) = \frac{D_{\text{com}}(z)}{1+z}. \tag{1.30}$$

Luminosity distance

Analogous to the angular distance, the luminosity distance can be defined by the flux measured from an object, which is equal to the surface brightness multiplied by the surface divided by the angular distance. Each of those three component scaled respectively by $a(t)$, $a^2(t)$ and $a(t)$. In the end,

$$D_{\text{L}}(z) = \frac{1}{a^2(t)} D_{\text{ang}}. \tag{1.31}$$

1.3 Cosmic Microwave Background

The CMB is the oldest light that can be observed from the Universe, as it was only 380,000 years old. It shows the temperature or intensity fluctuations at a time when no structures existed yet which is of the utmost interest for cosmology. Those intensity fluctuations can give information on the initial conditions for the formation of structures in the Universe. Its frequency spectrum is given by a blackbody function

$$\mathrm{B}_\nu(\mathrm{T}) = \frac{2h\nu^3}{c^2} \frac{1}{e^{h\nu/k_{\mathrm{B}}T} - 1}, \tag{1.32}$$

where $\mathrm{B}_\nu(\mathrm{T})$ is the surface brightness at a given frequency ν and temperature T. h is the Planck constant, c the speed of light and k_B the Boltzmann constant.

1.3.1 Angular power spectrum Formalism

The amplitude of the temperature fluctuations in the CMB can constrain cosmological parameters, i.e. give information on the initial conditions or the Universe. The CMB angular power spectrum can be derived through temperature fluctuations

$$\frac{T(\theta,\phi) - \overline{T}}{\overline{T}} = \frac{\Delta T}{T}(\theta,\phi), \tag{1.33}$$

where T is the temperature at a given θ and ϕ angle on the sky (see Fig. 1.5) and $\overline{T}$ the mean temperature over the full-sky. These temperature fluctuations can be decomposed into spherical harmonics on the

sphere's surface, giving

$$\frac{\Delta T}{T}(\theta,\phi) = \sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{m=\ell} a_{\ell m} Y_{\ell m}(\theta,\phi) \tag{1.34}$$

The $a_{\ell m}$ coefficients represent the amplitude of the temperature fluctuations on the sphere (see Eq. 1.36) and the spherical harmonics $Y_{\ell m}$ the orthogonal base (see Eq. 1.35) on which the fluctuations can be decomposed. It is analogous to a Fourier decomposition.

$$\int_{4\pi} Y^*_{\ell m} Y_{\ell m} d\Omega = \delta_{\ell\ell'}\delta_{mm'}. \tag{1.35}$$

Where Ω is the solid angle and δ is the Kronecker symbol and $Y^*_{\ell m}$ is the conjugate of the $Y_{\ell m}$ coefficients.

$$a_{\ell m} = \int_{4\pi} \frac{\Delta T}{T}(\theta,\phi)\, Y^*_{\ell m}(\theta,\phi)\, d\Omega. \tag{1.36}$$

To compute the power spectrum, is to look at the average distribution of the temperature over the sky per ℓ-scale. We have,

$$\langle a_{\ell m} a_{\ell' m'}\rangle = \delta_{\ell\ell'}\delta_{mm'} C_\ell, \tag{1.37}$$

where C_ℓ is the angular power spectrum and for a fixed ℓ scale it is equal to,

$$\hat{C}_\ell = \langle a_{\ell m} a_{\ell' m'}\rangle_m = \langle |a_{\ell m}|^2\rangle_m = \frac{1}{2\ell+1}\sum_{m=-\ell}^{m=\ell} |a_{\ell m}|^2. \tag{1.38}$$

Because, for each m we have $2\ell+1$ possible ℓ values, and in order to normalize the power spectrum estimator $\hat{C}_\ell$, we simply divide by this value. Looking at the angular power spectrum represented by C_ℓ, we can understand that the temperature fluctuations depend on the scale of the sphere that represents the observed sky.

1.3.2 Interpretation of the power spectrum

The power spectrum of the CMB (C_ℓ^{TT}, see Fig. 1.13) is a vast source of information about the initial conditions of the Universe. The angular power spectrum represents the amplitude of the temperature fluctuations in the function of the multipole, scale ℓ or angle θ. The link between the two is given by the relation in degrees,

$$\theta \approx \frac{180}{\ell}. \tag{1.39}$$

The information that can be extracted and interpreted from the CMB power spectrum is presented in the following parts.

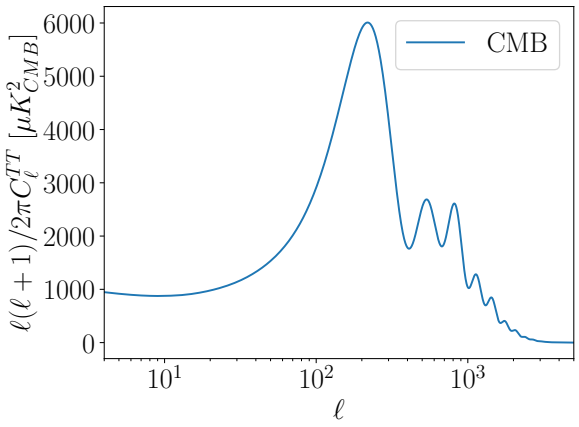

Figure 1.13: Power spectrum of the Cosmic Microwave Background in function of the scales on the sphere ℓ. The higher the ℓ value, the smaller the scales. Generated using CAMB algorithm (Lewis, 2005) with its default cosmological parameters.

Cosmic Microwave Background plateau

The CMB power spectrum is often plotted as $\ell(\ell+1)/2\pi C_\ell^{TT}$ because this exhibits the almost constant plateau between the scales $\ell \in [2, 30]$. This plateau is predicted by the Sachs-Wolfe effect (Sachs and Wolfe, 1967). This effect, explains why certain zones of the CMB appear cold today even though they were hot originally. The hot regions, due to their high density, were subject to a strong gravitational redshift. This leads to an evening out of the CMB fluctuations on sub-horizon scales ($\ell < 100$). The horizon is defined as the maximum distance in the Universe from which information can be retrieved. Therefore, even though the evened-out fluctuations have then evolved, this is not something we can observe because of the Horizon. The large-scale fluctuations are 'frozen' to a constant, which is what they were at a time when their information could still reach an earth-based observer. The slight uprise of the power on the first few scales is due to the Integrated Sachs-Wolfe (ISW) effect. The ISW is also due to gravitational redshift but this time happening between the CMB emission and Earth. This occurs because in a Universe dominated by Dark Energy, the potential wells are modified thus influencing the CMB photons on the few very large multipoles (ℓ). Because only one Universe can be observed, from only one unique position and within the limit of the horizon, the largest multipoles suffer from poor statistics. Their measurements are, therefore, limited by what is called the cosmic variance uncertainty.

Acoustic peaks

Several acoustic peaks are visible in the CMB angular power spectrum. The first peak, located at $\ell \sim 2000$ (1°) is the fundamental of the signal, followed by several harmonics. This means that higher amplitude fluctuations are found around 1° scales. The acoustic peaks formed through Baryonic Acoustic Oscillations (BAOs) in the primordial plasma. Before recombination, as the universe was one coupled fluid of matter and radiation, the tiny density fluctuations encouraged gravity to pull the baryons-photons fluid together, but the heating generated a pressure that pushed them apart. As the Universe expanded and cooled, gravity was able once more to pull the baryons-photon fluid together. This interplay between gravity and pressure on the baryons-photon fluid created acoustic oscillations in the primordial plasma.

When the matter and radiation decoupled, at the recombination era, those primordial acoustic oscillations stopped and became 'frozen', imprinted, and carried on by the now, free photons. The BAOs give the harmonic positions of the acoustic peaks. The amplitude of the peaks themselves is influenced by the baryon loading or content in the Universe. During the BAOs, the more baryons there are, the more the primordial plasma will compress due to gravity. The odd-numbered acoustics peaks in the CMB represent how much the fluid falls within the potential wells and the even-numbered peaks how much the pressure pushes it out. Therefore, the more baryons in the Universe, the more the odd peaks are enhanced in comparison to the even ones (see Fig. 1.14).

Damping tail

Above $\ell > 1000$, the acoustics oscillations in the CMB power spectrum are reduced and their amplitude decreases exponentially as the scales diminish. This happens because the recombination process is not instantaneous, the CMB itself is a given thickness. This means the largest scales, and largest fluctuations, are instantaneously frozen out by atoms forming and photons being freed while photon-matter interaction can still happen on the smallest scales. At those scales, the photons can still transport energy from the hot regions to the cold ones, evening out the temperature fluctuations and the acoustics peaks (Silk, 1968). This effect is referred to as the Silk-damping of the CMB or the part of the CMB power spectrum as the damping tail.

Influence of the cosmology

The fractional density content of the Universe, such as of Dark Energy (Ω_Λ), of baryons (Ω_b), or of matter as a whole (baryons + dark matter) (Ω_m) but also the geometry of the Universe through the curvature (Ω_k) directly influences the shape of the CMB power spectrum. Their influence is displayed in panel (a) of Fig. 1.14, for example, the more the Universe distances itself from a flat geometry to go toward a fully curved ($\Omega_k = 1$) geometry, the more the acoustics speak are shifted to smaller scales. The Sachs-Wolfe plateau also disappears to be replaced by a peak on a very large scale and a very steep increase of the power toward larger scales meaning a strong ISW effect. Panel (b), shows that as the Dark Energy content in the Universe (Ω_Λ) increases, the more the CMB power is shifted to larger scales and the more the ISW effect is important. In panel (c), as already discussed above, if there are more baryons in the Universe (Ω_b), the more the odd acoustics peaks are enhanced and the even ones reduced in amplitude. This is especially true for the fundamental peak. The total matter content (Ω_m), conversely, tends to decrease the peak amplitude as it increases.

1.3.3 Conclusion

The Cosmic Microwave Background angular power spectrum shape and amplitude is full of information about the initial conditions and parameters of the Universe. That is why being able to precisely measure the CMB signal with the lowest error possible is essential. The CMB being the oldest light observable from the Universe, is in the background of every physical process that has happened after recombination. This background light, the CMB photons, is influenced by the formation history of the Universe and its structure. The CMB in itself, originally, exhibits the imprint of the over-dense and under-dense regions of the Universe, from which emerged the first structures, stars, galaxies, and galaxy clusters.

[10] http://background.uchicago.edu/~whu/araa/node15.html

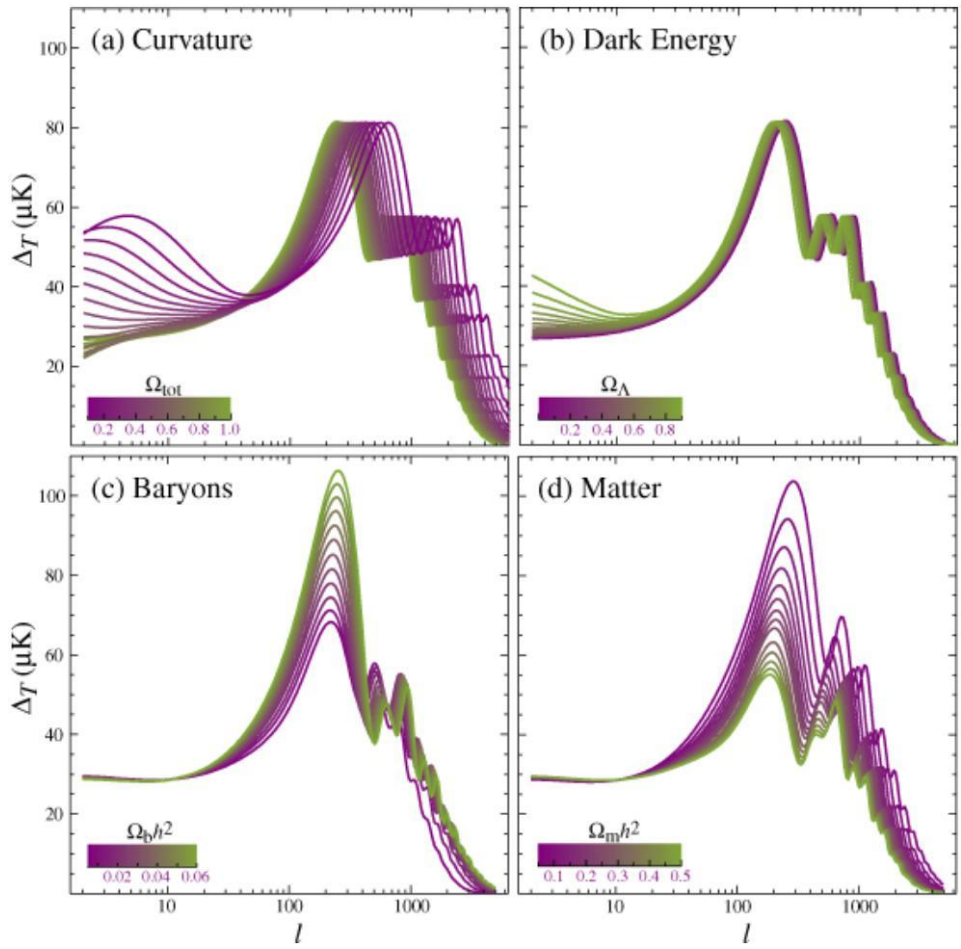

Figure 1.14: Modifications of the Cosmic Microwave Background power spectrum depending on densities of the various components of the Universe (see Eq. 1.15) [10].

1.4 Structures in the Universe

Structures in the Universe formed from the tiny density fluctuations already present in the CMB. With time, gravity clumps the matter, both baryonic and dark matter. The potential wells accrete the nearby matter making the structures grow bigger. Indeed, structures follow a hierarchical formation pattern, meaning that the smallest structures form first and through accretion grow larger to form larger structures. One could wonder how any structure can form within an expanding Universe. When the matter content, within a given radius, exceeds the critical density of the Universe (see Eq. 1.13) then, the gravitational force is stronger than the expansion meaning the system enclosed within this radius becomes gravitationally bounded. This radius is called the Virial radius (r_{vir}). It is the radius within which the structure obeys the Virial theorem. This theorem expresses that for an isolated system in equilibrium, the kinetic energy (E_{kin}) is equal to the minus half of the potential energy (E_{pot}) of the

system,

$$E_{kin} = -\frac{1}{2}E_{pot}. \tag{1.40}$$

The Virial theorem (Clausius, 1870) was formalized in 1870 by the German physicist Rudolf Clausius.

1.4.1 Dark matter halos

In the ΛCDM model, Dark Matter (DM) composes 25% of the content of the Universe (see Fig. 1.7) and only interacts with gravity. Therefore, dark matter can accumulate within potential wells and accrete the neighbouring dark matter to grow further until it virializes and becomes bound by gravity. Dark matter halos are in fact, not a halo but rather a collection of dark matter clumps of different sizes (see Fig. 1.15).

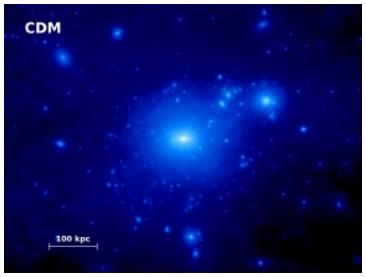

Figure 1.15: Simulation of a dark matter halo [11].

The density of Dark matter haloes has been shown using numerical simulations of structure formation to follow a universal profile (Navarro et al., 1997),

$$\rho(r) = \frac{\rho_0}{\left(\frac{r}{r_s}\right)\left(1+\frac{r}{r_s}\right)^2}, \tag{1.41}$$

where $\rho(r)$ is the density of the DM halo at a radius r, ρ_0 is a parameter different for each halo, and $r_s = r_{vir}/c$ is the scale radius. This profile is called the Navarro-Frenk-White (NFW) profile.

On much larger scales, the DM exhibits a Web-like structure referred to as the Cosmic Web. The Cosmic Web is composed of, filaments, sheets (2D planes) and nodes of Dark matter, which are clumps at the intersections of sheets and filaments (See Fig. 1.17).

1.4.2 Galaxy clusters

Baryonic matter is accreted by gravity and starts to form structures inside the potential wells created by DM halos. In early times the Universe was metal-poor, and the gas principally composed of Hydrogen and Helium first collapsed under gravity to form stars. Through nuclear fusion in their core, heavier elements started to form enriching the interstellar medium (ISM). With time, gas and dust collapsed to start forming solar systems and galaxies. Eventually, as time progresses, these systems also agglomerate into what is called galaxy clusters. This formation scenario, of galaxy clusters being composed of smaller

[11] https://www.astronomy.com/science/interactive-dark-matter-could-explain-milky-ways-missing-satellite-galaxies/

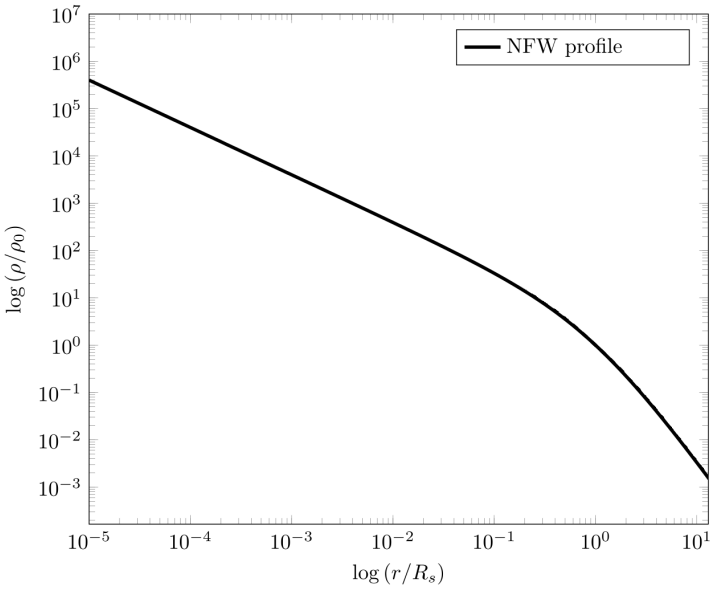

Figure 1.16: Navarro-Frenk-White profile, adapted from Wikipedia.

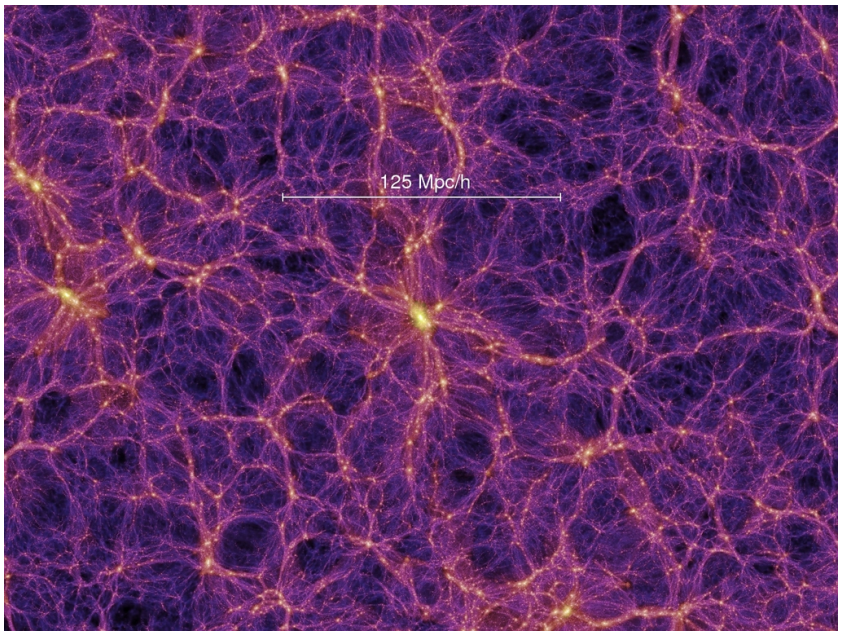

Figure 1.17: Cutout from, the Dark Matter only Millennium simulations (Springel et al., 2005) exhibiting the cosmic structure on the large scales. Here, only the clustering of Dark matter particles that have evolved under gravity is simulated, exhibiting a web-like structure.

similar systems is called the self-similar model (Kaiser, 1986). The baryonic matters that compose galaxy clusters follow the Dark Matter cosmic web-like distribution as can be seen by comparing Fig. 1.4 to Fig. 1.17. In particular galaxy clusters can be found in nodes of the dark matter distribution. Galaxy clusters are a collection of galaxies bound by gravity.

Properties of galaxy clusters

Even though galaxy clusters are different from one another, they are grouped under a few sets of characteristics defining them,

- **Size:** Galaxy cluster sizes vary but are, on average, of the order of megaparsecs (Mpc) with a diameter comprised between 1 to 5 Mpc.
- **Temperature:** The typical temperature of galaxy clusters is between $T_{cluster} \in [5, 10]$ keV.
- **Content:** They contain between 50 to 1000 galaxies. Objects containing less than 50 galaxies are usually called galaxy groups.
- **Weight:** Galaxy cluster masses ranges between $M_{cluster} \in [5 \times 10^{13}, 10^{15}] M_{\odot}$. Lighter objects, hosting between $M \in [10^{12}, 5 \times 10^{13},] M_{\odot}$ are called galaxy groups.

However, galaxy clusters are composed of more than just the visible baryonic galaxies (see Fig.1.18),

- In fact, only $\sim$ 1% of a cluster total mass is composed of galaxies, in the form of stars. The velocity dispersion of those galaxies is between 800 km/s to 1000 km/s.
- The IntraCluster Medium (ICM) is composed of hot ionized gas, mainly composed of hydrogen and helium, that makes up $\sim$ 14% of a galaxy cluster total mass content. Its temperature is usually comprises between $T_{gas} \sim [10^7, 10^8]$ K. The temperature of the ICM comes from the strong potential well attracting infalling matter into the cluster and creating shocks. At these temperatures, the ICM gas is detectable only via X-ray emission.
- The main contributor to the mass of galaxy clusters is dark matter $\sim$ 85%. The DM halo (see Sec. 1.4.1 and Fig. 1.15) is more extended than the galaxy cluster content detectable via the visible part of the electromagnetic spectrum. The dark matter halo is composed of clumps and its potential well attract baryonic matter into the gravitational bound of the cluster heating the ICM.

Formalism, mass and profiles

The mass of galaxy clusters is often referred to in M_{Δ} $[M_{\odot}]$. This unit is defined as,

$$M_{\Delta} = \frac{4\pi}{3} \Delta \rho_c R_{\Delta}^3 \tag{1.42}$$

Formalism, mass and profiles

The mass of galaxy clusters is often referred to in M_{Δ} $[M_{\odot}]$. This unit is defined as,

$$M_{\Delta} = \frac{4\pi}{3} \Delta \rho_c R_{\Delta}^3 \tag{1.43}$$

where ρ_c is the critical density of the Universe as defined in Eq. 1.13 and R_{Δ} the radius that encloses Δ times the critical density of the Universe. For virialized structures, a common approximation is that

[12] https://www.mpe.mpg.de/2040034/clusters_and_groups_of_galaxies

Figure 1.18: Galaxy cluster MACS1206 observed with optical light and overlaid with the X-ray contours [12].

$M_{\text{vir}} \approx M_{200}$, meaning the virial mass encloses approximately 200 times the critical density of the Universe. When measuring the mass of a cluster, hydrostatic equilibrium is usually assumed, however, this is not the case because of the constant gas accretion, shocks, and mergers. The fractional difference between the true mass of a cluster and the mass deduced from its gas content under the hydrostatic assumption is called the hydrostatic mass bias,

$$b = \frac{M_{hydro}}{M_{true}}. \tag{1.44}$$

The pressure profile $P(r)$, i.e., the pressure in function of the radial distance to the centre of the cluster is usually modelled as a Generalized Navarro-Frenk-White (GNFW) profile (Nagai et al., 2007) as a generalization of Eq. 1.41,

$$P(r) = \frac{P_{500} P_0}{\left(\frac{r}{r_s}\right)^{\gamma} \left(1 + \left(\frac{r}{r_s}\right)^{\alpha}\right)^{(\beta-\gamma)/\alpha}} \tag{1.45}$$

where P_{500} is the pressure contained within a radius containing 500 times the critical density of the Universe, P_0 is a free parameter, it is the cluster's pressure amplitude, r_s is the scale radius, and α, β, γ are the profile slopes and free parameters.

Interest of galaxy clusters

Galaxy clusters are interesting objects to study because they are among the largest structures in the universe. They also host galaxies, which themselves host solar systems, stars, and planets. There is

therefore an intrinsic interest in studying their properties. Such as their temperature, pressure, proper velocity, size, and masses. However, galaxy clusters are also an essential tool to understand the properties of the Universe, as a link of the chain of structure formation. They are also the largest tracers of matter in the Universe, whether baryonic or dark matter and therefore clusters can be used as probes for cosmology. The number of clusters per mass can be used to constrain or verify cosmological parameters as seen in Fig. 1.19. The curves of the number of clusters in the function of M_{500} can be computed for two different redshift intervals $z \in [0.025, 0.25]$ in black and $z \in [0.50, 0.90]$ in blue for two cosmologies, one for ΛCDM and one for a universe where there would be no dark matter ($\Omega_\Lambda = 0$). Overplotted are real measurements from galaxy cluster surveys, the models only fit the data in the ΛCDM case (Giodini et al., 2013).

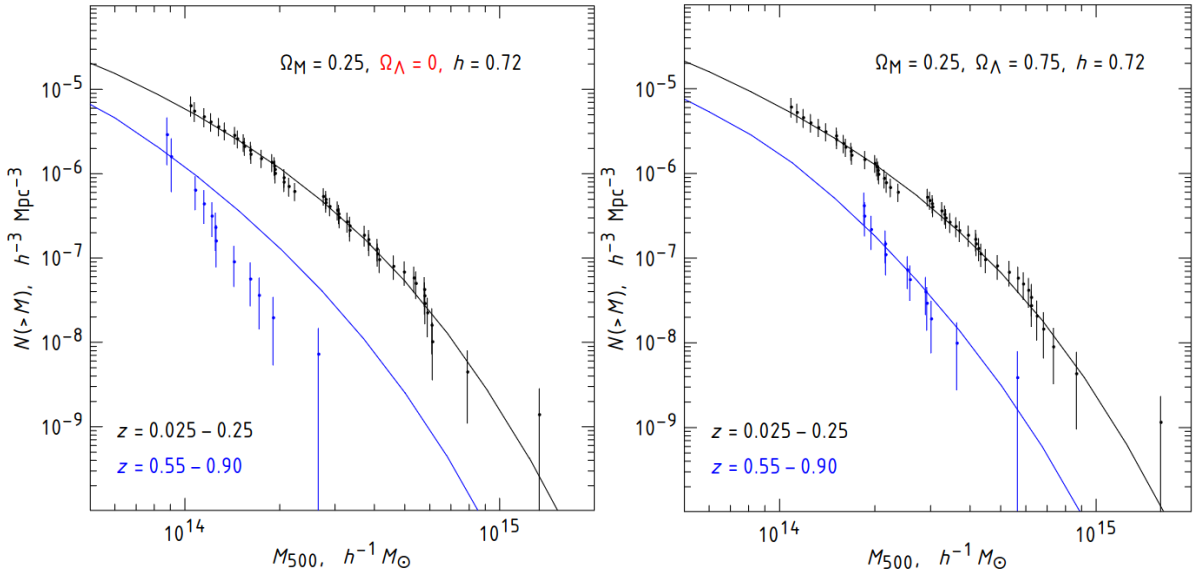

Figure 1.19: Number of galaxy clusters in function of M_{500} for two redshift ranges [0.025,0.25] (black) and [0.50,0.90] (blue) and two models, universe without dark energy (left) and ΛCDM (right plot). Extracted from (Giodini et al., 2013).

1.4.3 Cosmic Infrared Background

Inside galaxy clusters, galaxies themselves formed through gas accretion falling into potential wells. The star formation as a function of redshift clearly exhibits a peak at $z \approx 2$, meaning that is when most of the stars formed in the Universe. Around that redshift, a majority of galaxies in the sky were forming stars (see Fig. 1.20). Because of the finite value of the speed of light, observing deeper into the sky means observing the past. Therefore, when observing anything in the sky, a fair amount of star-forming galaxies are also observed but not always resolved depending on the instrument's angular resolution (see Eq. 1.23). If they are not resolved they are point sources smeered-out by the beam. In stars-forming galaxies there is a lot of gas and dust gathered through accretion and collapsing to produce stars, those galaxies are called Dusty Star Forming Galaxies (DSFG). Because the star formation is taking place at $z \sim 2$, not all the dust and gas content has yet been transformed into stars. The light produced by the already existing stars is therefore absorbed by the dust grains, which are heated up by this interaction and

emit infrared radiation. That is why when observing the sky, there is always in the background some unresolved DSFG bright in the infrared domain. This is known as the Cosmic Infrared Background (CIB), it is the cumulative, dust-absorbed, infrared emission produced by the stars in galaxies, either at the moment of their formation or throughout their lifetime.

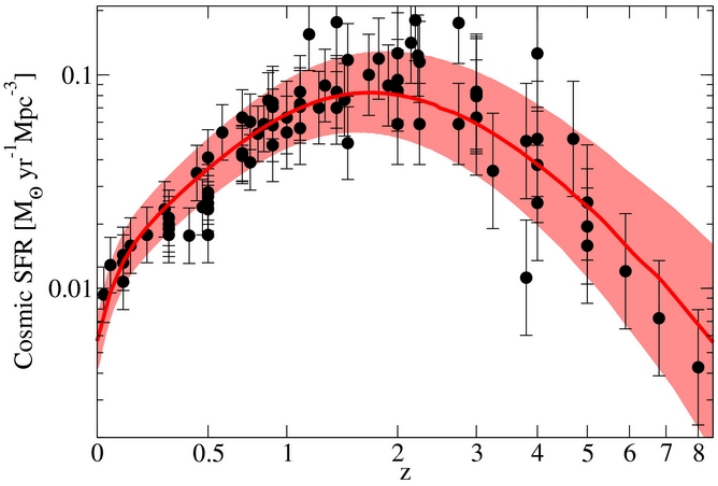

Figure 1.20: Star Formation Rate (SFR) in function of the redshift. The black data points are from observations and the red curve is the best-fit model, with the 1σ confidence interval (Behroozi et al., 2013).

1.5 Secondary anisotropies of the CMB

As previously discussed (see Sec. 1.1.1), the CMB is the oldest light observable and is in the background of processes that followed after recombination. The CMB contains information about the temperature anisotropies at the time of recombination. As the photons from the CMB travel freely in the Universe, they sometimes interact with the forming structures, gas, and dust. This interaction creates new anisotropies in the CMB, not originating from the primordial fluctuations but from the evolution of the Universe's content. Therefore, these are called secondary anisotropies of the CMB. Many physical processes can create secondary anisotropies of the CMB, but this work concentrates on the ones induced by the Sunyaev-Zeldovich effects. There are several effects describing how the presence of galaxy clusters influences the CMB photons.

1.5.1 thermal Sunyaev-Zeldovich effect

Cosmic Microwave Background (CMB) photons have a small probability of passing through galaxy clusters and a $\sim 1\%$ probably to interact with an electron (Carlstrom et al., 2002). When this happens, the CMB photons, inverse Compton scatter (Compton, 1923) on the hot electrons resulting in a transfer of energy from the electron to the photons (see Fig. 1.21). This transfer happens under the conservation of momentum in the electron's rest frame. In the rest frame of the observer the photon that appeared redshifted, appears more energetic after inverse Compton scattering. This is the thermal Sunyaev-Zeldovich effect (Sunyaev and Zeldovich, 1969, 1970). The energy transfer from the electron population to the CMB photons in the ICM by the Sunyaev-Zeldovich effects depends on the reflection angle. The

maximum energy transfer occurs when the photons are scattered in the same direction as the incoming electrons and the minimum energy is when the photons are scattered in the opposite direction as the incoming electrons. However, statistically, on average, the photons scatter with an angle of 90deg with respect to the incoming electrons.

Consequences and formalism of the tSZ

A consequence of this energy transfer is a shift of the CMB photon population, at the location of galaxy clusters, from low frequency (low energy) to high frequency (high energy), the photon energy being characterized by

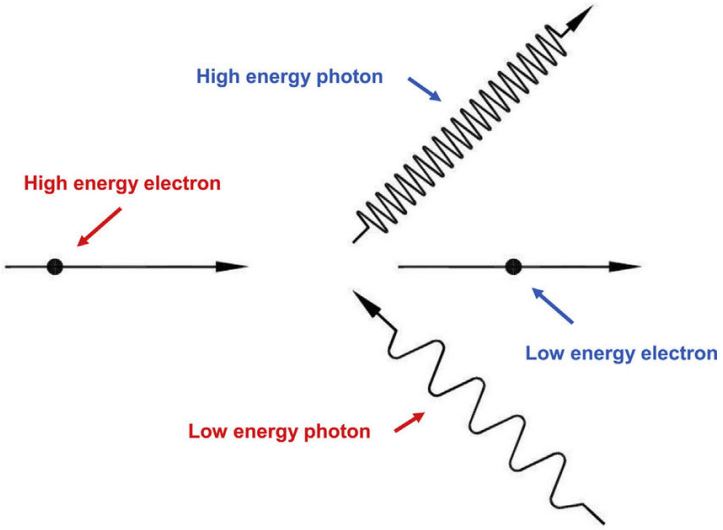

Figure 1.21: Schematic of the inverse Compton scattering process. Extracted from Wright (2015).

$$E_\gamma = h\nu, \tag{1.46}$$

where h is the Planck constant and ν the frequency. When looking at the CMB, at the position of a galaxy cluster, a decrement of photons will be seen, and at higher frequencies an increment of photons at higher frequency (see Fig. 1.22).

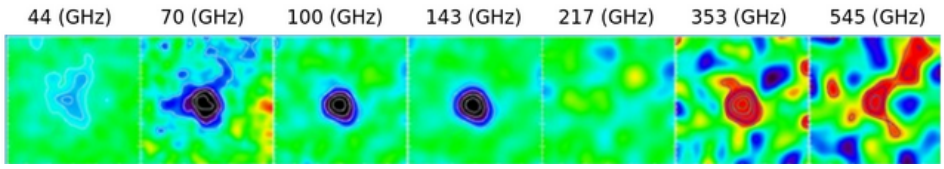

Figure 1.22: Galaxy cluster Abell 2319 observed in 7 frequency bands [13].

[13] http://public.planck.fr/outils/astrophysique/effet-sz

Indeed the low-frequency CMB photon population has been shifted to higher frequency through the tSZ effect. The amplitude of this energy change can be plotted as a function of the frequency giving the SED of the tSZ effect (see Fig. 1.23). The tSZ effect produces a negative contribution at frequencies below $\nu <\sim 217$ GHz, as a null contribution at $\nu \approx 217$ GHz, and a positive one at frequencies $\nu >\sim 217$ GHz. Therefore, the tSZ effect induces a distortion of the CMB blackbody spectrum (see Fig. 1.24) At frequencies below $\nu <\sim 217$ GHz, we observe a deficit of power in the CMB spectrum, because part of the CMB photon population, has inverse Compton scattered on the electrons in galaxy clusters and is now at higher energy (frequency). Around $\nu \approx 217$ GHz, the distorted CMB spectrum power is equal to the non-distorted one. At frequencies $\nu >\sim 217$ GHz, there is excess power in the CMB spectrum, because the lower energy (frequency) photons have been shifted to higher energy through the tSZ effect.

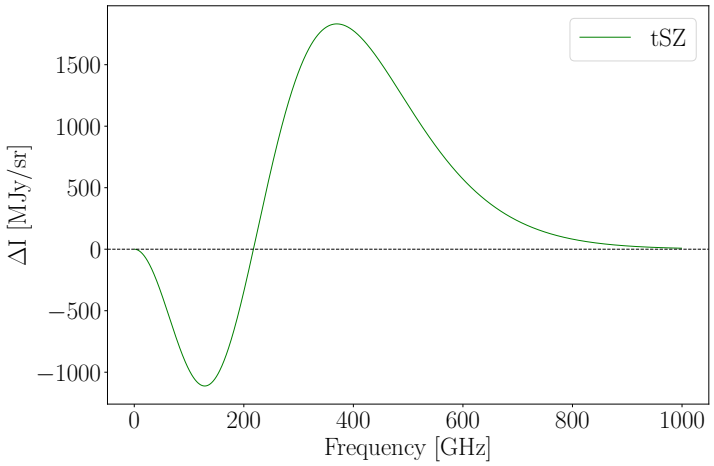

Figure 1.23: Frequency spectrum of the thermal Sunyaev-Zeldovich effect.

The temperature change in the CMB, generated by the tSZ effect is formalized as such,

$$\frac{\Delta T_{\mathrm{CMB}}}{T_{\mathrm{CMB}}} = y f(x_\nu), \tag{1.47}$$

where $T_{\mathrm{CMB}} = 2.72548 \pm 0.00057$ K (Fixsen, 2009), y is called the Compton-y parameter, representing the integrated pressure along the Line-of-Sight (LoS) and defined such as,

$$y = \frac{\sigma_{\mathrm{T}}}{m_{\mathrm{e}} c^2} \int_{\mathrm{los}} P_{\mathrm{e}}(l) \mathrm{d}l, \tag{1.48}$$

where σ_{T} is the Thomson cross-section, m_{e} the electron mass, c the speed of light, and $P_{\mathrm{e}}(l) = n_{\mathrm{e}}(l) k_{\mathrm{B}} T_{\mathrm{e}}(l)$ the electron pressure, composed of the electron density n_{e}, the Boltzmann constant k_{B} and the electron temperature T_{e}. $f(x_\nu)$ is the frequency spectrum of the tSZ effect, defined as,

$$f(x_\nu) = x_\nu \coth\left(\frac{x_\nu}{2}\right) - 4 = \left(x_\nu \frac{e^{x_\nu}+1}{e^{x_\nu}-1} - 4\right), \tag{1.49}$$

where $x_\nu = h\nu / k_{\mathrm{B}} T_{\mathrm{CMB}}$, h the Planck constant and ν the frequency. The CMB temperature change caused by the tSZ effect can also be expressed as an intensity change. The first temperature derivative of

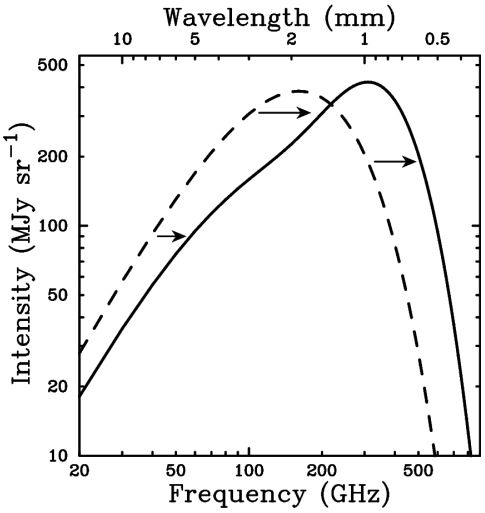

Figure 1.24: Blackbody spectrum of the CMB (dotted line and the spectrum distorted by the tSZ effect (full line). For visualization purposes, the distortion is exaggerated, it is plotted for a galaxy cluster 1000 times more massive than a typical massive cluster. The distorted spectrum is also shifted to the right, as indicated by the arrows. Figure extracted from Carlstrom et al. (2002).

the Planck blackbody spectrum ($B_\nu(T)$) directly gives us the conversion factor between temperature and intensity,

$$\Delta I_\nu = \frac{\partial B_\nu(T)}{\partial T}\Delta T, \tag{1.50}$$

with,

$$\frac{\partial B_\nu(T)}{\partial T} = \frac{I_0}{T_{CMB}}\frac{x_\nu^4 e^{x_\nu}}{(e^{x_\nu}-1)^2}, \tag{1.51}$$

where, $I_0 = 2(k_B T_{CMB})^3/(hc)^2 \approx 270$ MJysr^{-1}. Therefore by multiply Eq. (1.49) by Eq. (1.51), the tSZ intensity change is given by,

$$\frac{\Delta I_\nu}{I_0} = y h(x_\nu) f(x_\nu), \tag{1.52}$$

with $h(x_\nu) = x_\nu^4 e^{x_\nu}/(e^{x_\nu}-1)^2$.

Interest of the tSZ effect

One of the principal characteristics and advantages of the tSZ effect is its redshift independence. Indeed, the effect is fully formalized without any redshift dependence. Despite its weak amplitude, the tSZ effect has many applications and can be used for measurements in clusters astrophysics,

- It can be used to detect galaxy clusters, by looking into multi-frequency CMB data for its characteristic shape, a negative contribution at low frequency, and a positive contribution at high

frequency (Planck collaboration et al., 2016). Because of the redshift independence of the effect, clusters can be detected at very high redshift as long as they are resolved by the telescope's beam.

- It can probe a wide range of dynamic physical phenomena in clusters, shocks (Erler et al., 2015), cold fronts, energy dissipation, and pressure (see Eq. 1.48).
- It can be used in combination with X-rays to determine mass-weighted temperatures of galaxy clusters. The tSZ effect is proportional to the temperature and electron density tSZ $\sim n_e T_e$ and the X-ray signal from the cluster is proportional to $\sim n_e^2 T_e$. The linear dependence of tSZ on the gas density makes it better than X-ray to study shocks and the high temperature generated.

The tSZ effect can also be used for cosmology:

- The amplitude of the angular tSZ power spectrum is proportional to the clustering of matter and the density of matter $C_\ell^{yy} \propto b^{-3.2}\sigma_8^{8.1}\Omega_m^{3.2}$ (Komatsu and Seljak, 2002; Bolliet, 2018). Its strong dependence on σ_8 gives it strong constraining power but depends on the mass bias (b, see Eq. 1.44).
- Measure the expansion rate in combination with X-ray data through the angular diameter distance (Mroczkowski et al., 2019). Indeed, using tSZ and X-ray, the surface brightness of clusters can be deduced, and with that, how the size of a cluster appears with time, to deduce the size of the Universe and its growth.
- Cluster number count can give information on the baryons content of the Universe.

Astrophysical contamination of the tSZ effect

The tSZ effect can be difficult to detect, not only because of its weak amplitude but also because of contamination originating from other signals. For example,

- Many clusters host in their core a source emitting strong radio waves, it is for the most part unresolved by telescopes and therefore called a radio point source.
- The CMB is a strong contaminant of the tSZ.
- The Cosmic Infrared Background is the main contaminant of the tSZ on frequencies above 350 GHz, where the dust contamination dominates. Moreover, the CIB and tSZ signals are localized in the same celestial object, galaxy clusters, making them spatially correlated. Indeed, even though the peak of star formation is around $z \sim 2$ (see Fig. 1.20) and the peak of the tSZ signal should be around $z \sim 0.1$, originating from massive galaxy clusters at low redshift, a non-zero correlation between the two signals is still expected making it particularly difficult to separate them.
- Another contaminant is known as the SZ confusion, Because the SZ signal is not dependent on the redshift it is difficult to say if the signal is coming only from one cluster or from other sources in the foreground/background, some of them may not even be clusters but could be, for example, Cosmic Web filaments. This contamination should be only a small contributor to the SZ signal of massive clusters but for small clusters and groups, this could be a major contaminant.

1.5.2 The kinematic Sunyaev-Zeldovich effect

Another SZ effect is the kinematic Sunyaev-Zeldovich (kSZ) effect. It is also an energy of intensity change of the CMB photons but this time is caused by the bulk motion of the galaxy cluster with respect to the CMB rest frame. Indeed, galaxy clusters have their own proper ensemble motion, with respect to the fixed CMB. As the CMB photons going through the cluster interact with the electron population, they are Doppler boosted by the motion of the electrons. In particular, when the cluster is moving toward the observer they get blueshifted and when the cluster is moving away from the observer they get redshifted. This is visible in the SED of the kSW effect which can be either negative when the cluster is moving away from the observer or positive when it is moving toward the observer.

Formalism of the kSZ

The frequency spectrum of the kSZ effect is a Planck spectrum (see Fig. 1.25),

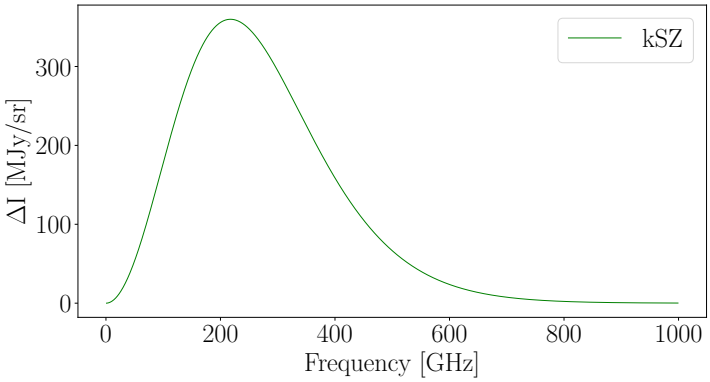

Figure 1.25: Frequency spectrum of the kinematic Sunyaev-Zeldovich effect.

It can be formalized, as the temperature change imposed on the CMB photons,

$$\frac{\Delta T_{\rm CMB}}{T_{\rm CMB}} = \int_{los} \sigma_{\rm T} n_{\rm e} \vec{n}.\frac{\vec{v}_{\rm pec}}{c} dl, \tag{1.53}$$

where, $n_{\rm e}$ is the electron density, $\vec{n}$ the line of sight direction, and $\vec{v_{\rm pec}}$ the bulk motion of the cluster. This can be expressed as an intensity change in the CMB, by multiplying by Eq. (1.51),

$$\frac{\Delta I_\nu}{I_0} = -y_{\rm kSZ} h(x_\nu), \tag{1.54}$$

where,

$$y_{\rm kSZ} = -\tau_{\rm e}\left(\frac{v_{\rm pec}}{c}\right), \tag{1.55}$$

with $\tau_{\rm e}$ is the optical depth. The temperature shift generated by the kSZ effect is positive when the cluster is moving toward the observer and negative when it is moving away from the observer.

Interest of the kSZ effect

As for the tSZ, the kSZ effect is redshift-independent and has many applications for astrophysics through its use to measure clusters' velocities and optical depth. However, the two are degenerate, therefore, some prior are required from another method or measurements. The kSZ effect can also be used for cosmology,

- Contrary to the tSZ effect, the kSZ effect probes all the baryons in the cluster, not only the hot ones. Therefore, it can be used to probe the baryon content at the outskirts of the cluster, which is still mostly unknown territory but also in the Universe in general. The missing baryon problem, in particular, exists because only part of the baryons predicted is actually observed in the Universe.
- The velocities of galaxy clusters are driven by potential wells and gravity which originates from, the primordial perturbations. The kSZ effect can either be used as a proxy for those perturbations. Especially the bulk motion reflects the large-scale density perturbations while the small scales are represented by the peculiar motions and arise from the local environment in galaxy clusters.
- Velocities should be unbiased on the large scales, therefore, kSZ should be an unbiased tracer of the large-scale electron distribution (n_e).
- Using the continuity equation,

$$\dot{\delta} + \frac{1}{a}\vec{\nabla}.\vec{v}(1+\delta) = 0, \tag{1.56}$$

  The following Fourier space expression of the velocities field can be deduced (Alonso et al., 2016),

$$\vec{v}(t,k) = \frac{Hf}{a}\frac{i\vec{k}}{k^2}\delta(t,\vec{k}), \tag{1.57}$$

  where t is the time, k the scale, H the Hubble parameter, f the growth of structure, a the scale factor, and i the complex number.

Difficulties linked to the kSZ effect

The kSZ effect amplitude is one order of magnitude smaller than the tSZ which makes it much harder to detect. Other difficulties linked to the usage of the kSZ effect are,

- The relation between the observed galaxy density distribution and the true underlying matter distribution is the following $\delta_g = b\delta_m$ where b is called the galaxy bias. The galaxy observations are a biased tracer of the matter distribution. From the relation between the galaxy's velocity field and the matter distribution (see Eq. 1.57) can be deduced, $v_g \sim \frac{f}{b}\delta_g$.
- The equation linking the velocity field to the growth factor f (see Eq. 1.57) is only valid in the linear regime, on large scales, on small scales, where a lot of non-linear physical processes happen, this is not valid anymore.
- The observable velocity field of galaxies is distorted by their radial peculiar velocities which induces redshift space distortions. In particular, the Finger of God (FoG) generates an elongation

of the galaxy's distribution along the line-of-sight in redshift space and the Kaiser effect (Kaiser, 1987) which flattened the velocities distribution because of the coherent motions of galaxies toward the cluster's centre.

- If the number of detected galaxies is too low then some shot noise term will increase the variance of the measurements.
- The kSZ effect has the same spectral shape as the CMB, meaning it will appear either as a cold spot or as a hot spot, depending on the velocity direction of the cluster and not on the frequency. It is, therefore, impossible to distinguish from original CMB anisotropies only based on frequency observations.

1.5.3 Other flavours of SZ

In this part are presented the other existing flavours of SZ effects, which are not used in this work, but are interesting to keep in mind for completeness.

Relativistic tSZ

The tSZ effect as described in (see sec. 1.5.1) supposes that the electrons are moving at non-relativistic speeds. As a first approximation, this hypothesis is good enough, but in reality, some electrons in the ICM are relativistic. Therefore, a relativistic correction has to be added to the tSZ effect expression, without this correction, there is a risk of under-predicting the amplitude of the tSZ power spectrum by $\sim 30\%$ (Remazeilles et al., 2018). The SED of the relativistic Sunyaev-Zeldovich (rSZ) effect is formalized by

$$f(x_\nu, T_e) = f(x_\nu) + \delta(x_\nu, T_e), \tag{1.58}$$

where $f(x_\nu)$ is defined in Eq. 1.49 and $\delta(x_\nu, T_e)$ are the relativistic corrections. This modifies the power spectrum of the tSZ effect in function of the temperature of the electron population in the cluster. In particular, the hotter the electron, the lower the SED peaks and the more the null is shifted towards higher frequencies while the amplitude of the tail increases.

non-thermal Sunyaev-Zeldovich effect

The non-thermal Sunyaev-Zeldovich (ntSZ) effect (Ensslin and Kaiser, 2000), originates from the scattering of CMB photons on the non-thermal electrons present in galaxy clusters. Non-thermal electrons have a temperature much higher than the other particle species in the ICM plasma, meaning the fluid is not in equilibrium. The effect is even weaker than the kSZ effect, as its amplitude is approximately only 1% of tSZ effect.

polarized Sunyaev-Zeldovich effect

The polarized Sunyaev-Zeldovich (pSZ) effect is an even smaller effect, four orders of magnitude smaller than the tSZ effect in amplitude. It occurs from the presence of a quadrupole anisotropy in the galaxy cluster local radiation field, generating polarized light when the CMB photons scatter on the electrons. Like the kSZ effect and the CMB, it is frequency-independent, making it the hardest SZ effect to measure.

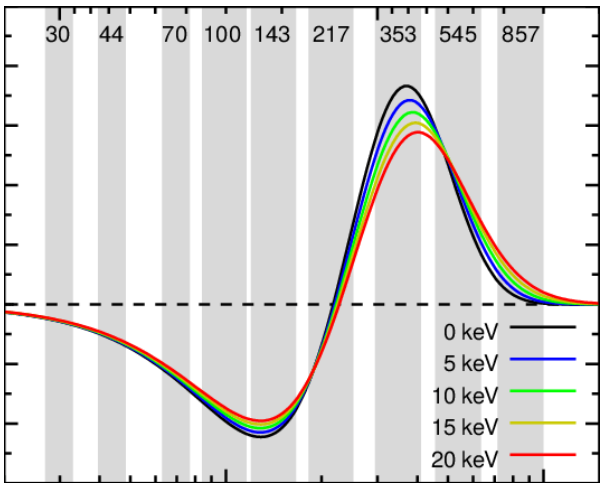

Figure 1.26: Frequency spectrum of the relativistic Sunyaev-Zeldovich effect. Overplotted are the *Planck* space telescope observation frequency bands. Extracted from Erler et al. (2018).

1.6 Conclusion

Over the course of a century, cosmology achieved the building of a model that is robust against a large variety of observations, the ΛCDM model. Of course, many questions are still open, and as we enter the area of precision cosmology some discrepancies begin to appear. The most famous and concerning one is the H_0 tension, which can be summarized as the local, late-type measurements of the H_0 value being in disagreement with the early-type, early-Universe measurements, by more than 4σ. Roughly 95% of the Universe's content, Dark Energy and Dark Matter, remain unknown and unseen, and their origin is not fully understood. That is why it is good to use many different probes of the Universe to assess the robustness of the models of the Universe but also to maybe discover the reasons for those mismatches and in the end, adjust the models. To better study the history of the Universe, and of structure formations, the Cosmic Microwave Background is an incredible tool, shedding light on all that has happened after the Universe recombination. That is why the different flavours of the Sunyaev-Zeldovich effects are such important probes for galaxy cluster astrophysics but also cosmology. The technical progress of telescopes has allowed measuring ever smaller effects, highlighting tensions in the cosmological paradigm and also opening a new window that will allow refining the models and better grasping the Universe properties.

CHAPTER 2

Experiments and data processing tools

Contribution

In this chapter, is addressed:

- the discovery, correction, and implementation of a missing feature of the smoothing function (`healpy.sphtfunc.smoothing()`) of the Python package HEALPy.
- the development of an ILC/CILC Python package for signal extraction.

This chapter gives an overview of space and ground-based experiments that have shaped Cosmic Microwave Background (CMB) science. But it also goes into the different data sets, simulations, basic tools, and Python packages used in this project to simulate upcoming CMB missions and predict their outputs on CMB science.

2.1 The *Planck* satellite

The *Planck* space satellite data and results were used to verify the goodness of our pipelines and codes. This was done by comparing the results obtained with them, to the already published results obtained by several other groups on the vastly explored *Planck* data.

2.1.1 Context of space CMB science observation

The theoretical predictions (see Sec 1.1.1) and the first detection of the CMB temperature by Penzias and Wilson (1965) motivated the scientific community to start an area of space measurements of the CMB. In an attempt to measure the CMB temperature with better precision, the first CMB space mission, RELIKT-1, was launched by the Academy of Sciences of the Soviet Union in 1983. It observed the sky for 7 months at 37 GHz and a resolution of 5.8°. It measured the CMB dipole moment at T = 2.1 ± 0.5 mK, an upper limit on the quadrupole at T = 0.2 mK and for the anisotropies at 6° and 90° of respectively, $\delta T/T = 4 \times 10^{-4}$ and $\delta T/T = 7 \times 10^{-5}$ (Strukov and Skulachev, 1984). In parallel, a rocket mission named Cobra was launched, in 1982, by the University of British Columbia to measure the at-the-time called Cosmic Background Radiation (CBR) frequency spectrum. Their result, published only in 1990 found the CBR to be a well-fitted black body spectrum with a temperature of T = 2.736 K

(Gush et al., 1990). The National Aeronautics and Space Administration (NASA), launched, in 1989 the COsmic Background Explorer (*COBE*), to measure the shape of the blackbody spectrum of the CMB and the temperatures anisotropies. COBE observed the full sky at 31.5 GHz, 53 GHz, and 90 GHz with a 10° resolution. The two goals were achieved by the Far Infrared Absolute Spectrophotometer (FIRAS) instrument. The CMB spectrum was measured to be the closet thing to a blackbody spectrum ever measured (see Fig. 2.1) and the temperature of the CMB was measured to be $T_{CMB} = 2.725 \mp 0.002$ K (Mather et al., 1999). The *COBE* satellite also measured the temperature fluctuations, the anisotropies of the CMB, at $\Delta T/T = 11 \times 10^{-6}$ (Smoot et al., 1992). Though successful the mission showed that more precision and angular resolution were needed to better measure cosmological parameters and the anisotropies of the CMB. For those reasons, NASA developed a new space telescope called the Wilkinson Microwave Anisotropy Probe (*WMAP*), which was launched in 2001. Its goal was to measure the anisotropies of CMB with more precision and on smaller angular scales. The *WMAP* satellite observed the sky at 23 GHz, 33 GHz, 41 GHz, 61 GHz, and 94 GHz. For all 5 frequency bands, it achieved a sub-degree resolution. Giving it a resolution up to 33 times better than *COBE*. Its sensitivity was also up to 45 times better than *COBE*. *WMAP* contributed to show that the Universe's composition was dominated by Dark Energy, followed by Dark Matter, and finally by baryonic matter. It measured the age of the Universe to be 13.772 ± 0.059 Gyr and established constraints on cosmological parameters (Bennett et al., 2013).

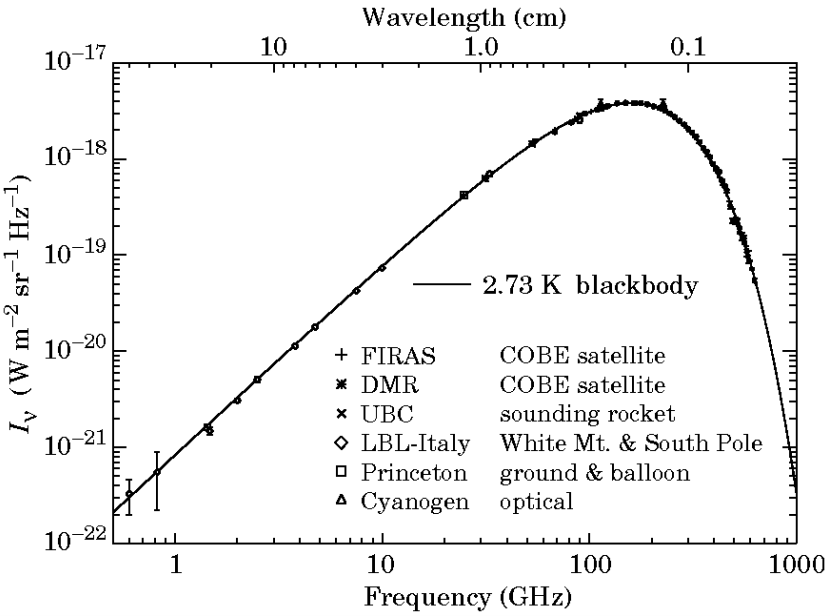

Figure 2.1: CMB blackbody spectrum with measurements data points from various experiments (Smoot).

In this context, the *Planck* satellite was proposed, with the goals to improve on *WMAP* and *COBE* while proposing for the first time measurements of the secondary anisotropies of the CMB.

2.1.2 The *Planck* mission

The *Planck* space telescope (Tauber and et al., 2010) was a 1,5 m diameter telescope developed by the European Space Agency (ESA) and with participation from the National Aeronautics and Space

Administration (NASA). The skylight enters the satellite and is reflected on the primary mirror (see LoS on Fig. 2.2), it is then, reflected on the secondary mirror, opposite of the primary, and down on the figure before being reflected onto the detector plane. The *Planck* satellite was composed of two detectors, the Low-Frequency Instrument (LFI) (Bersanelli et al., 2010) and the High-Frequency Instrument (HFI) (Lamarre and et al., 2010), they both observed at different frequencies, with different sensitivities (see Tab. 2.1) and band transmissions. The band transmissions, from which an example can be seen in Fig. 2.3, for the HFI, are broad, they cover several frequencies. Their transmission is also not equal to 1 and is not a Dirac function, it varies with frequency within the interval covered by the band. The complete transmission profile of each band has to be taken into account to properly interpret and process the data.

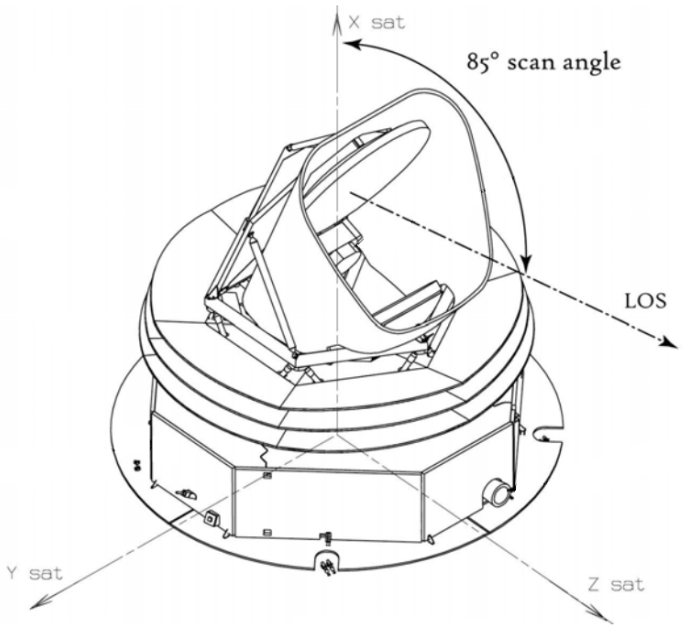

Figure 2.2: Sketch of the *Planck* satellite, showing its Line-Of-Sight (LOS) departing from the primary mirror. Adapted from Villa et al. (2003)

	LFI			HFI					
Frequency [GHz]	30	44	70	100	143	217	353	545	857
Beam size [arcmin]	33.10	27.94	13.08	9.66	7.27	5.01	4.86	4.84	4.63
Sensitivity [μK-arcmin]	11	17	29	80	34	56	167	1414	1702×10^2

Table 2.1: Array containing the frequencies, beam sizes, and sensitivities of the *Planck* satellite for the Low-Frequency Instrument (LFI) and High-Frequency Instrument (HFI). Values Extracted from (Planck Collaboration et al., 2011; Planck collaboration et al., 2013)

The *Planck* satellite was launched on the 14th of Mai 2009 by an Ariane 5 rocket along with the *Heschel* satellite to study the origin and parameters of the Universe as a whole. It surveyed the full sky seven times before the exhaustion of its cooling fuel and was shut down on the 23rd of October 2013. The *Planck* mission had several goals,

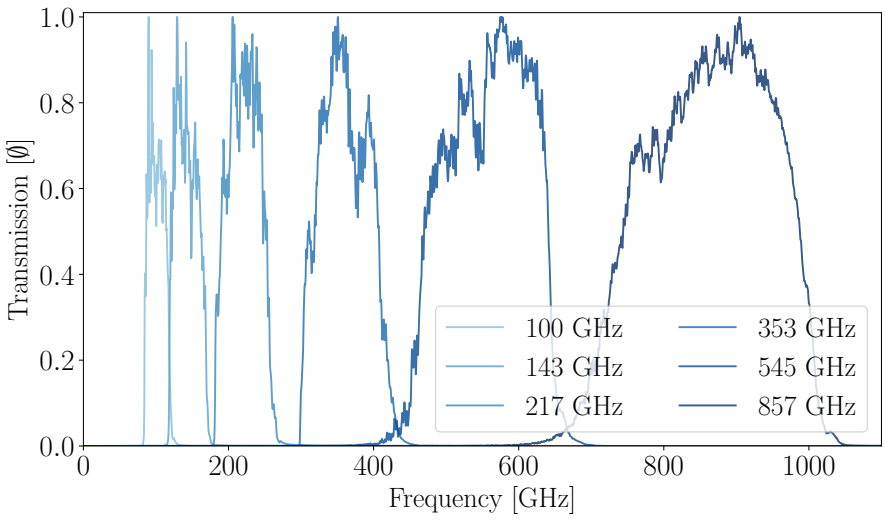

Figure 2.3: High-Frequency Instrument (HFI) transmission bands for the *Planck* satellite.

- the measurement of the anisotropies of the Cosmic Microwave Background (CMB) to the highest precision to date.
- the measurement of the Sunyaev-Zeldovich (SZ) effect.
- to put constraints on the Hubble constant (H_0).
- to test the inflation models for the Universe.
- the measurement of the amplitude of structures in the CMB.

2.1.3 The *Planck* data

This part presents the specificities of the *Planck* satellite measurements of the sky emissions. Since its launch in 2013, *Planck* had several Public data Release (PR). PR1 in 2013, PR2 in 2015, PR3 in 2018 and PR4 in 2020.

Units

The *Planck* satellite data is measurements of the sky emission in various units, the first one is in Jansky per steradian (Jy/sr), which is called the spectral variance (see Sec. 1.2.1). The spectral radiance is the flux emitted by a given surface, per unit solid angle and per frequency or wavelength. The maps can also be expressed in K_{CMB} units, which is a characteristic unit of CMB science originating from the observation of CMB. The CMB spectrum is a blackbody spectrum, meaning its spectral radiance ($B(\nu, T)$) is entirely defined by its temperature and frequency. To look at the temperature variations of the CMB is to look at the derivate of the blackbody spectrum at a given frequency and for a temperature

$T=T_{CMB}$.

$$B_\nu(T) = \frac{2h\nu^3}{c^2}\frac{1}{e^{x_\nu}-1}, \tag{2.1}$$

where h is the Planck constant, ν the frequency, c the speed of light and,

$$x_\nu = \frac{h\nu}{k_B T}, \tag{2.2}$$

where k_B is the Boltzmann constant and T the temperature. Giving,

$$\left.\frac{\partial B_\nu(T)}{\partial T}\right|_{T=T_{CMB}} = \frac{2h^2\nu^4}{k_B c^2 T_{CMB}^2}\frac{e^{x_\nu}}{(e^{x_\nu}-1)^2}, \tag{2.3}$$

with $x_\nu = h\nu/k_B T_{CMB}$.

Maps

The type of data from the *Planck* satellite, used in this work, full-sky maps (see Fig. 2.4). Those maps represent the total temperature or intensity variation over the sky at the given frequencies of observation (see Table 2.1). The maps are composed of the cumulative emission of all astrophysical processes over the sky. The horizontal bar in the middle of the maps is the strong emission coming from the Milky Way (MW). Because the solar system is located in the plan of the MW, *Planck* observed this strong close-to-home emission. It is clearly visible that the intensity and temperature variations of the MW emissions vary with frequencies, and notably increase with increasing frequency. This is due to, low frequencies ($\nu < 44$ GHz) to the Synchrotron and free-free emission and high frequencies ($\nu > 217$ GHz) to the radiating dust in the MW. This will be extensively commented on later on.

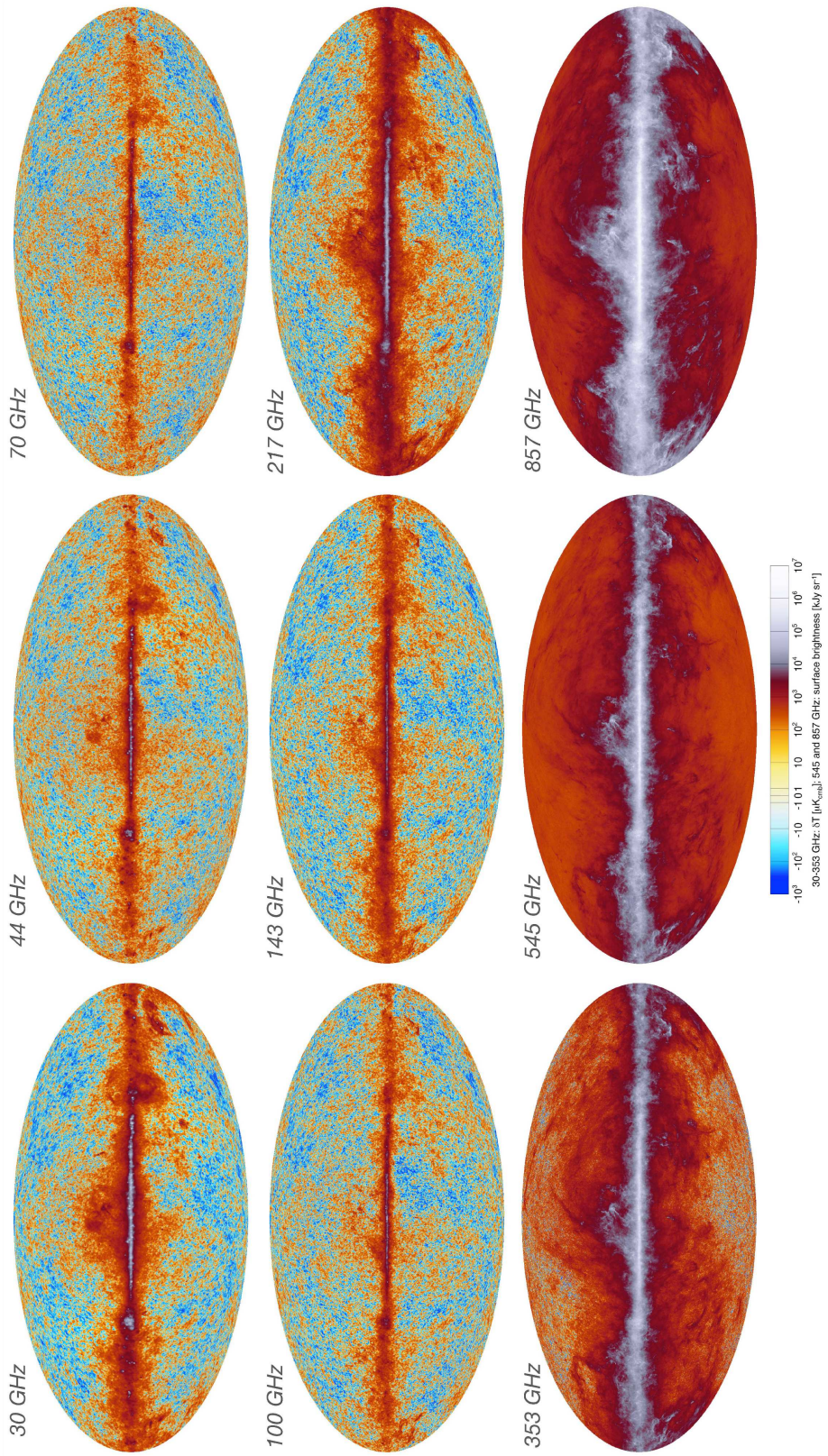

Figure 2.4: Maps of the total temperature and intensity variations over the full-sky as measured by the **Planck** satellite. Extracted from the Public data Release 3 (PR3) (Planck collaboration et al., 2020a).

2.1.4 Conclusion and outputs of the *Planck* mission

The *Planck* mission is to date, the state-of-the-art space telescope mission for CMB science. It improved on COBE and WMAP, by a factor of more than 2.5 in terms of resolution (see Fig. 2.5).

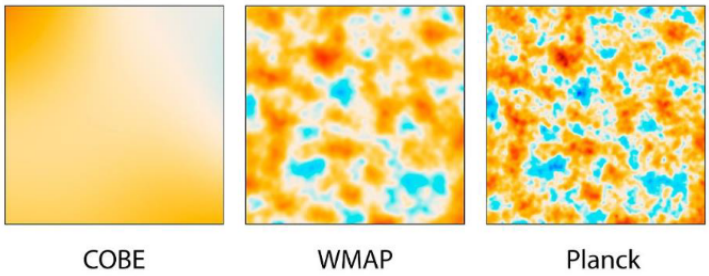

Figure 2.5: The same 10x10 degree cutout of the Cosmic Microwave Background (CMB) map, as measured by the COsmic Background Explorer (*COBE*), the Wilkinson Microwave Anisotropy Probe (*WMAP*) and *Planck*. Courtesy of NASA/JPL-Caltech/ESA.

CMB measurements

The *Planck* space satellite pushed to lower scales the measurements of the CMB power spectrum, from $\ell = 2$ to $\ell = 2508$ (Planck collaboration et al., 2020b) compared to up to $\ell = 1150$ for *WMAP* (Larson and et al., 2011) which was already an improvement compared to the range from *COBE*, of $\ell \sim 70$ to $\ell \sim 600$ (Signore and Puy, 2009). The CMB power spectrum fit was used to deduce the most precise constraints on cosmological parameters to date. The *Planck* team found that the Universe was composed of 4.82% ± 0.05% of ordinary matter, 25.8% ± 0.4% of dark matter, and 69% ± 1% of dark energy (Planck collaboration et al., 2014). It also found an age of 13.798 ± 0.037 Gyr for the Universe. The Universe is found to be consistent with a flat geometry by the *Planck* results, with a precision of 0.4% (95% C.L.) (Planck collaboration et al., 2020c).

The SZ effect measurements

The *Planck* space satellite made the first-ever measurements of the Sunyaev-Zeldovich (SZ) effect on its galaxy clusters sample. The cluster catalogue of *Planck* was the first ever blindly detected cluster catalogue using the SZ effect (Planck collaboration et al., 2011). It also made the first-ever full-sky Compton-y map of the thermal Sunyaev-Zeldovich (tSZ) effect (Planck Collaboration et al., 2014) and from it measured the tSZ power spectrum over angular scales between $\ell \sim 10$ to $\ell \sim 1000$. Using the relation $C_\ell^{yy} \propto \sigma_8^{8.1}\Omega_\mathrm{m}^{3.2}$ between the tSW power spectrum, the matter clustering σ_8 and the matter density Ω_m (Komatsu and Seljak, 2002; Planck collaboration et al., 2014; Bolliet, 2018) they could deduce the following constrain $\sigma_8(\Omega_m/0.28)^{0.4} = 0.784 \pm 0.016$ (68% C.L.). In practice this was done by fitting the extracted tSZ power spectrum data point with a model of the signal and residual contaminant in the extracted data, assuming a Gaussian likelihood function, and applying a Markov Chain Monte Carlo (MCMC) analysis.

The Hubble constant measurements

The *Planck* team found a value of $H_0 = 67.80 \pm 0.77$ kms^{-1}Mpc^{-1} for the Hubble constant (Planck collaboration et al., 2014). This is in agreement with the previous results from WMAP of $H_0 = 70 \pm 2.2$ kms^{-1}Mpc^{-1} (Hinshaw and et al., 2013).

2.2 Ground-based CMB science

This part explores the telescope doing CMB science from the ground. First, the historical context of ground-based telescopes is exposed followed by the description of two future telescopes for which predictions were made in this work.

2.2.1 History of Ground-based CMB observations

CMB science exploration and measurements started from the ground. In 1955, Émile Le Roux, surveyed the sky at a wavelength of $\lambda = 33$ cm at the Nançay observatory in France and found a highly isotropic temperature of $T = 3 \pm 2$ K (Kragh, 1999). Followed by Tigran Shmaonov who reported in his Ph.D. thesis published in 1957, a background radiation measured at $T = 4 \pm 3$ K (Durrer, 2018). The first detection of the CMB temperature was reported in 1965 by Arno Penzias and Robert Woodrow Wilson. Using a Horn antenna to study satellite communication, they found an excess temperature of T = 3.5 K, isotropic and unpolarized (Penzias and Wilson, 1965). Several decades later an era of ground-based telescopes specifically designed for CMB science emerged. The first of those was the Tenerife Experiment installed at the Observatorio del Teide which started operation in 1984 as a collaboration between the Jodrell Bank Observatory of the University of Manchester and the Instituto de Astrofisica de Canarias in Tenerife. Its goal was to measure the CMB power spectrum and observe the sky at 10 GHz, 15 GHz, and 33 GHz. This was followed by several other ground-based telescopes[1], some notable ones being the Background Imaging of Cosmic Extragalactic Polarization (BICEP) from 2006 to 2008 located in the South pole, followed in 2009 by BICEP2 with an upgrade to the detectors of BICEP. The South Pole Telescope (SPT) started in 2006 intending to measure small-scale temperature fluctuations of the CMB. The Atacama Pathfinder Experiment (APEX) which started observation in 2007 in the Atacama desert in Chile aimed for the CMB anisotropies and the SZ effect. The Atacama Cosmology Telescope (ACT) in 2008, the Q U I JOint TEnerife (QUIJOTE) experiment in Tenerife that started in 2012 and POLARBEAR in Atacama, which also started in 2012, mainly targetting CMB polarisation and B-modes, are all currently making important contributions to the field.

2.2.2 The Simons Observatory

The Simons Observatory is part of the next generation of ground-based CMB science telescopes called CMB-Stage 4 (CMB-S4). In the past ground-based were based on one telescope at one observing site. CMB-stage 1 was defined by telescopes containing the order of 100 detectors, stage 2 in the order of 1,000, and stage 3 in the order of 10 000. The CMB-S4 project aims to achieve the best-to-date CMB measurements both on large and small scales. This work makes predictions for the Simons Observatory (SO). This part presents the SO goals and technicals detailed exploited through this thesis.

[1] https://en.wikipedia.org/wiki/List_of_cosmic_microwave_background_experiments

Overview of the project

The SO (Ade et al., 2019) are upcoming ground-based telescopes that will be located on Cerro Toco at 5300 m, in the Atacama desert in Chile. Its main science goals are the following:

- Measure the tensor-to-scalar ratio.
- Constraints on inflations models.
- Measure the sum of neutrino masses.
- Look for deviations from the cosmological constant paradigm.
- Measure duration and redshift of the Epoch of reionization.
- Measure the thermal and kinematic Sunyaev-Zeldovich effect.

Predictions of results regarding those science goals can be found in Ade et al. (2019). The SO telescopes will be composed of 3 Small Aperture Telescopes (SAT) of a diameter of 0.4 m (Kiuchi and et al., 2020) and one Large Aperture Telescope (LAT) of 6 m diameter (Xu and et al., 2021). The LAT will be used to achieve high angular resolution and the SAT to address the need for very low noise on large angular scales. This work only uses the SO-LAT telescope and therefore the rest of the thesis will mention only this telescope. The survey strategy planned by the SO team can be seen in Fig. 2.6. The SATs will observe some fields avoiding the galactic plane and overlapping with BICEP and another experiment called SPIDER. While the LAT will observe a band of the sky a bit larger than the one that will be observed by the Large Synoptic Survey Telescope (LSST).

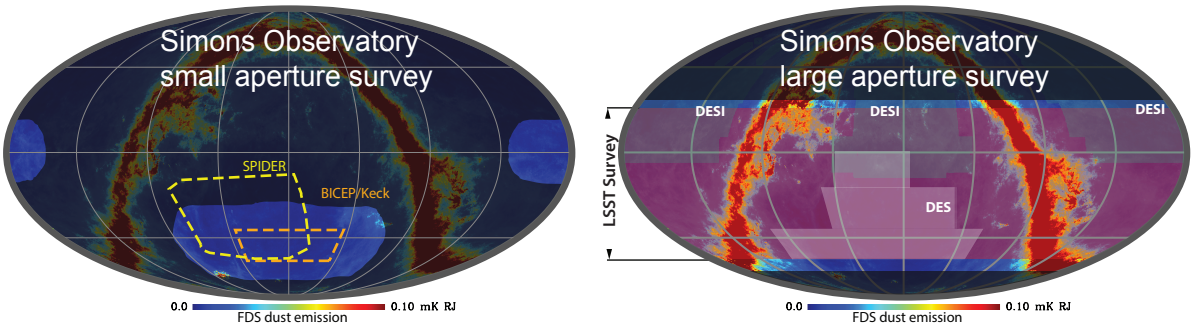

Figure 2.6: Map of the Galactic dust sky emission, in mK_{RJ} units, with overlaid in blue the planned survey strategy for the SATs on the left and the LAT on the right. Overlaid is also the observation strategies of other experiments, SPIDER and BICEP on the left and LSST on the right. Extracted from Ade et al. (2019).

Technical overview of SO-LAT

The LAT telescope was fabricated by CPI Vertex Antennentechnik GmbH in Xanten Germany and was transported by boat to Chile during early 2023. It is now being assembled in the Atacama desert. First-light is expected in 2024. The SO-LAT is a Cross-Dragone telescope meaning the light is reflected on the primary mirror onto a secondary mirror, crossing the light path of the entering light, before entering the detectors (see Fig. 2.7). This design was pioneered by Dragone (1978) and adapted by

Frequency [GHz]	27	39	93	145	225	280
FWHM [arcmin]	7.4	5.1	2.2	1.4	1.0	0.9
Sensitivity baseline [μK-arcmin]	71	36	8	10	22	54
Sensitivity goal [μK-arcmin]	52	27	5.8	6.3	15	37

Table 2.2: Array containing the frequencies, beam sizes, and sensitivities of the SO-LAT. Values Extracted from (Ade et al., 2019).

Niemack (2016) for SO-LAT. The advantage of such a design is a large Field-of-View (FoV) in a compact configuration with less systematics due to diffraction at millimetre and sub-millimetre wavelength because the secondary mirror is not obstructing the incoming light.

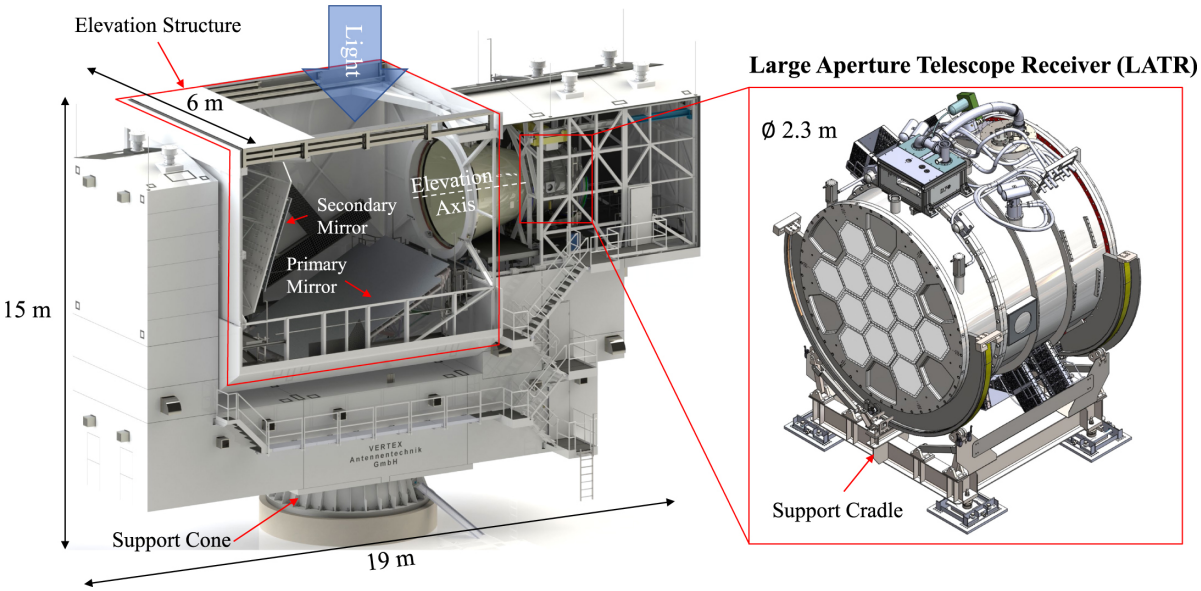

Figure 2.7: Schematic of the SO-LAT telescope, and the detector module LATR on the right. The telescope can turn around the support cone axis and his pointing at the zenith. The elevation structure in red and the LATR turn around the elevation axis. Extracted from Xu and et al. (2021)

The SO-LAT cryostat Large Aperture Telescope Receiver (LATR) will be able to host 6 frequency bands, with a given beam and sensitivity (see Tab. 2.2), with the possibility for extensions in the future. The SO-LAT baseline sensitivity is what is expected from the telescope at the very start of the observations, under minimum conditions and minimum technical development. The goal sensitivity is what is expected from the satellite design capabilities and full technical development.

2.2.3 AtLAST: concept idea for a large single dish telescope

The Atacama Large Aperture Submillimeter Telescope (AtLAST[2]), is a concept idea for a 50 m single dish telescope that would survey the sky at millimetre and sub-millimetre wavelength in the Atacama desert in Chile. Its large dish would provide incredible sensitivity. At this stage, AtLAST is an idea

[2] https://www.atlast.uio.no/

starting to get momentum, but not yet a financed project. Its relevance for this work will be to show an ideal case of a large single-dish telescope with lots of high-precision detectors, a concept idea without limitations for the moment. Its goals would be:

- To image the Milky Way plane with the highest resolution, mapping down, to molecular clouds, proto-planetary disks, dusts and proto-stars.
- Detect the cold material in galaxy halos and map the full life-cycle of baryons in and out of galaxies.
- To perform the deepest, most complete and highest spectroscopic survey to date.
- Measure the thermal and kinematic Sunyaev-Zeldovich effects down to single galaxies and in filaments between galaxies clusters.

The AtLAST concept is also targeting a telescope entirely operated by renewable energies. The design study to assess the feasibility of the telescope has received funding from the European Union Horizon 2020 research and innovation programme. This study is supposed to end in 2024.

2.3 The CCAT Observatory

This part introduces the Cerro Chajnantor Atacama Telescope (CCAT) observatory project, first by going through the history and then through the technical specificities of this telescope. This upcoming telescope is at the centre of this work, which consists of making predictions of what this telescope will achieve regarding the Sunyaev-Zeldovich (SZ) effects.

2.3.1 History of the CCAT project

The Cornell Caltech Atacama Telescope (CCAT) was supposed to be a 25 m diameter dish located at 5 600 m on the Cerro Chajnantor mountain in Chile (Sebring et al., 2006). The NASA Jet Propulsion Lab (JPL) was also part of the project. The goal of CCAT was to combine a high sensitivity, up to 2 arcsec with a wide Field-of-View (FoV) of up to 20 arcmin on a high-altitude site, one of the dryest spots on earth. Its concept design study was completed in 2005 and the telescope was thought to become complementary to the Atacama Large Millimeter/submillimeter Array (ALMA) thanks to its larger FoW.

The ALMA is an array of 66 radio-telescopes, composed of 7 to 12 m antennas functioning as an interferometer. The antennas are on tracks and their spatial configurations can be adjusted from a compact 150m one to a 16 km spaced, which renders it possible to emulate a much larger dish telescope and reach a fine resolution of 0.1 arcsec.

The CCAT was ranked as the highest priority medium-sized mission by the US Astro2010 decadal survey, an American call for project proposals to decide the upcoming decade's science funding. However, the project did not move forward due to a lack of funding. In early 2016, the CCAT project was reformed into a more modest 6 m diameter dish telescope and renamed the Cerro Chajnantor Atacama Telescope prime (CCAT-p). Along this journey the collaborating entities also changed, to be composed of as of today, Cornell University, the Universities of Bonn and Cologne, a consortium of Canadian universities (Waterloo, Toronto/CITA, British Columbia, Calgary, Dalhousie, McGill, Western Ontario) and their Chilean partners. Recently the telescope was renamed one last time to the Fred Young Submillimeter

Telescope (FYST) to honour the main donator and important collaborator of the project, Fred Young. The first light, i.e., the start date of observation is planned for late 2024. The capabilities, characteristics and goals of the telescope have evolved along the project and are extensively discussed in the next part.

2.3.2 The current FYST telescope characteristics

The Fred Young Submillimeter Telescope (FYST) will be a 6 m diameter cross-Dragone telescope (Dragone, 1978; Niemack, 2016). It will be located in one of the dryest places on earth, in the Atacama desert, in Chile. Its design will be similar to the SO-LAT (see Sec. 2.2.2 and Fig. 2.7). It is currently under construction in Xanten, Germany, by Vertex Anntennentechnik, the same constructor as for SO-LAT. The advantages of this configuration are the same as for SO (see Sec 2.2.2). The light path is also the same, the light will be entering the top part toward the primary mirror, onto the secondary mirror to be reflected to the instrument module which is called Prime-cam or Mod-cam (see Fig. 2.8). Another instrument called the CCAT Heterodyne Array Instrument (CHAI), will be located below the Prime-cam module, its goal is to serve as a spectrometer.

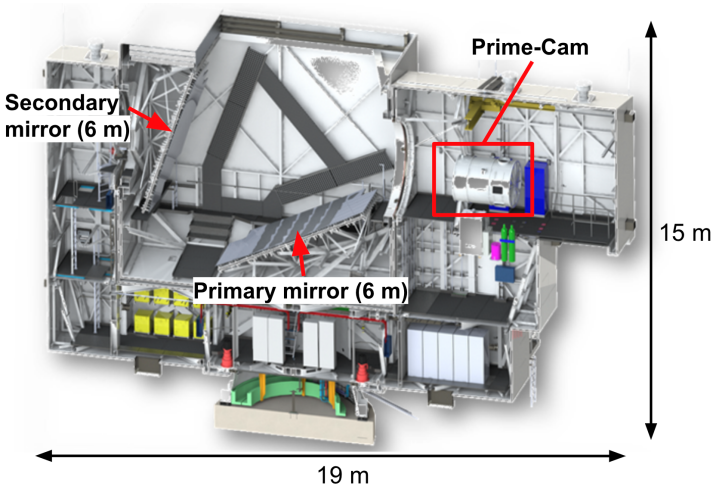

Figure 2.8: Schmeatic of the Fred Young Submillimeter Telescope (FYST) showing a cut view of the telescope and positions of the primary and secondary mirrors and of the receiver module, Prime-cam (Vavagiakis and et al., 2022).

Prime-cam

Prime cam is a cryogenic tube composed of seven tube receivers. Each of them is composed of three wafer arrays and will observe at one frequency (see Fig. 2.9). Two optics tubes will be reserved for the Epoch-of-Reionization (EoR) spectrometers (Nikola et al., 2023). The goal of those spectrometers will be to produce tomographic maps of [CII] and [OIII] lines, which are tracers of the star formation covering redshift between z ~ 3.5 up to z ~ 15 for [OIII]. This will allow FYST to trace the ionization bubbles at the reionization epoch. The frequency is 1.9 THz for [CII] and 3.4 THz for [OIII]. The Mod-cam cryostat is already built and is currently located at Cornell University in Ithaca, New York, where they are conducting tests on it. The first light configuration should be composed only of the 280 GHz band.

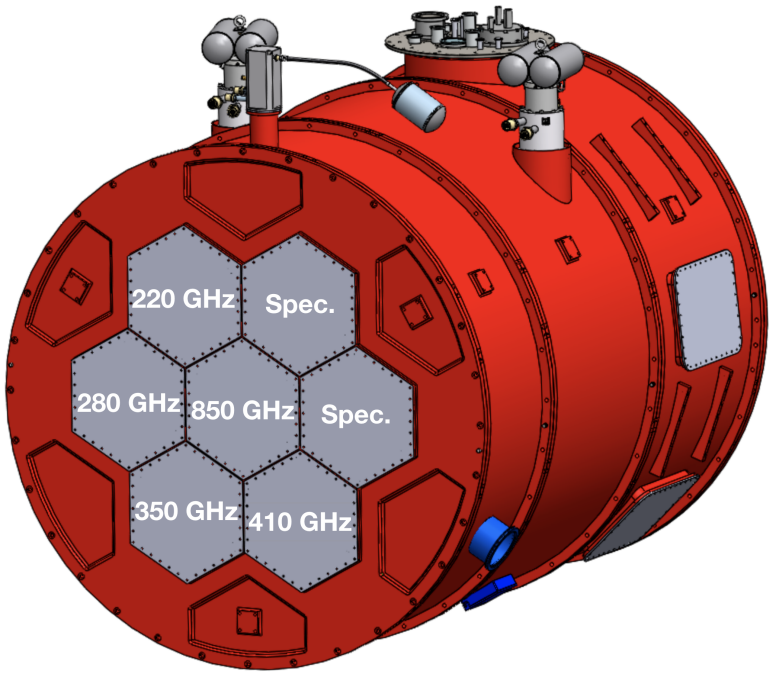

Figure 2.9: Prime-cam cryostat module with the 7 optics tube and their respective frequency. Extracted from Choi and et al. (2020).

Frequency [GHz]	220	280	350	405	860
FWHM [arcmin]	1.0	0.9	0.6	0.5	0.2
Sensitivity [μK-arcmin]	15	27	105	372	5.7×10^5

Table 2.3: Array containing the frequencies, beam sizes, and sensitivities of the FYST. Values Extracted from Choi and et al. (2020).

In this work, the EoR-Spectrometers and the CHAI instrument will not be considered. When referring to FYST, it will always imply the 5 frequency channels of the Prime-cam module with the following sensitivities and beam sizes (see Table. 2.3).

Site

The Fred Young Submillimeter Telescope (FYST) will be located at 5 612 m on the Cherro Chajnantor peak in the Atacama desert, in Chile. The Atacama desert is the driest place on Earth not being part of the poles and the second driest overall. Dry places are a crucial advantage for Astronomical observations, the low Precipitable Water Vapor (PWV) content of the atmosphere increases the atmospheric transmission (see Fig. 2.10). The PWV is defined as the depth of the water of a 1 m^2 atmospheric column in which the total water vapour would be precipitated into liquid water. In the very dry places in which Astronomers build telescopes, the unit of this quantity is in mm. The FYST location has another advantage, which is that it is placed at a high altitude in the Atacama desert. In comparison, the ALMA interferometer, located at 5,000 m, has a much lower PWV, considering the same conditions of a 50% best weather quartile. Atmosphere contamination is one of the downsides of ground-based telescopes compared

to space telescopes such as *Planck*. Atmospheric contamination limits the recovery of large-scale information in the signal. Contamination can be reduced through filtering and modelling of atmospheric noise.

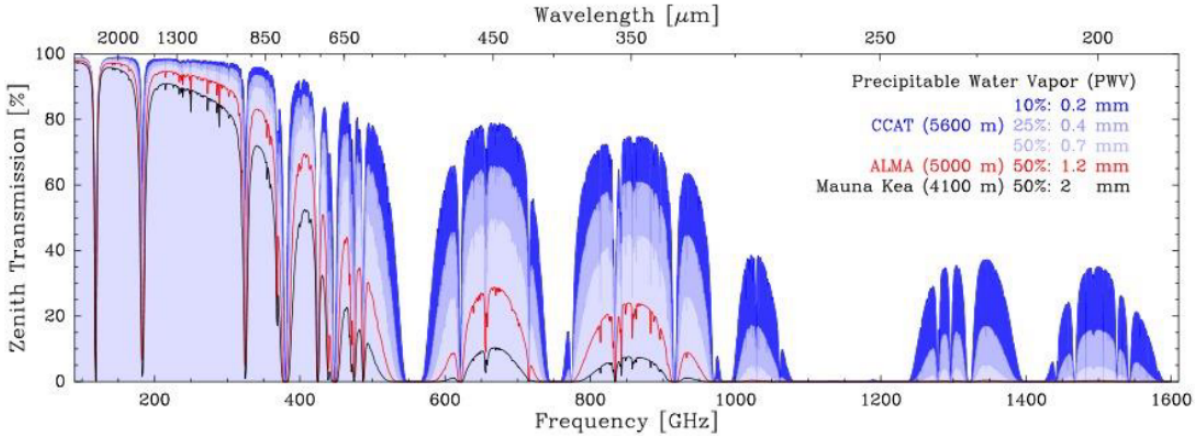

Figure 2.10: Atmospheric transmission at the Zenith in % in function of the frequency in GHz for various telescopes. In black for the Mauna Kea telescopes located in Hawaii at 4 100 m of altitude. In red for the ALMA interferometers, located in the Atacama Desert in Chile at 5,000 m. In various blue colours for the CCATp telescope located at 5 600 in the Atacama desert in Chile. The percentages indicate the PWV values for the best weather median quartiles. Courtesy of Martha P. Haynes.

2.3.3 Science goals

The FYST telescope, and its two instruments Prime-cam and CHAI, have been thought to address many different scientific goals (CCAT-Prime Collaboration et al., 2023).

- Probe the Epoch of Reionization (EoR). Using EoR-Spec two frequencies to observe [CII] (1.9 THz) and [OIII] (3.4 THz) lines. These two lines are tracers of the star formation, and therefore the reionization bubbles at redshift z $\in$ [3.5, 15].
- Tracing the 3D galaxy evolution over cosmic time. Using FYST high-frequency bands to trace the dust peak emission heated by stellar photons during the peak of star formation activity between a redshift of $1 < z < 3$.
- Measuring the Galactic-foregrounds. FYST's high-frequency bands will help measure the Galactic foregrounds contaminating CMB-B mode measurements by several orders of magnitude. Because thermal dust dominates the energy budget well above CMB-B at frequencies $\nu > 150$ GHz.
- Detect Rayleigh scattering. FYST broad frequency coverage should permit the statistical detection of the Rayleigh scattering of neutral hydrogen and helium atoms during the recombination epoch. Subsequently, this scattering will be used as a probe of the EoR.
- Explore time domain astrophysics. With its large FoV and frequency coverage FYST should be able to detect transient Astrophysical events of short duration such as Tidal Disruption Events

(TDEs), SuperNovae (SNe), Gamma-Ray Bursts (GRBs), merging neutron stars or X-ray binaries systems.

- Probing Galactic polarization. Galactic dust polarisation is perpendicular to the Galactic magnetic field lines. FYST will probe the dust thanks to its high-frequency channels which trace these field lines.
- Probe the Galaxy-cluster evolution. The Multi-frequencies channels, the exceptional site, and the low sensitivity of the FYST will allow us to derive some of the cleanest maps, to date, of the Sunyaev–Zeldovich (SZ). As a consequence, FYST will detect more clusters and more low-mass high-redshift clusters than previous telescopes. The detected clusters will improve the understanding of their formation and evolution.

2.4 Data and simulations processing

The results presented in this thesis revolve around the data processing of CMB telescopes observing large portions of the sky. Such as the past *Planck* satellite, or upcoming ones such as the Simons Observatory (SO) and the Fred Young Submillimeter Telescope (FYST). These CMB telescopes observe the sky in all directions, forming a sphere of intensity measurements. Analogously to the map of the planet Earth, we can visualize the sphere on a plane by projecting it. This is exactly as is done for CMB data (as visible in Fig. 2.4). This part describes how the projected intensity maps measured by CMB telescopes are dealt with using a software called HEALPix, which extracts physical information from the maps through the computation of the power spectrum and how to still compute it when part of the sky sphere is not surveyed using PyMaster.

2.4.1 HEALPix and HEALPy

This subsection presents the main CMB data processing package called HEALPix and its Python implementation HEALPy. First, by going through its characteristics and then through some of the problems encountered and solved when using HEALPy.

Pixelisation of the sphere

Hierarchical Equal Area isoLatitude Pixelation (HEALPix) is an IDL software that was developed by Gorski et al. (2005) to project *WMAP* CMB data onto a 2D-plane but also to develop a set of tools and functions to process these data. The idea of HEALPix is to tesselate the sphere into pixels of equal areas. An intensity measurement can later be attributed to those pixels. The pixelization scheme of HEALPix has been constructed on a few main points (see Fig. 2.11):

- The pixels are quadrilaterals (rhombus-like) of equal areas
- The pixels are all located on lines of constant latitude.
- The sphere's smallest tesselation is 12 pixels.
- The side of each pixel can be further divided into 2 to create more pixels of equal areas.

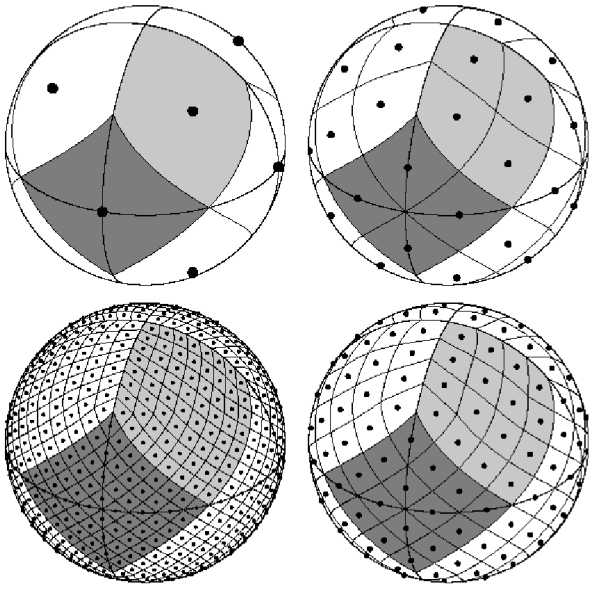

Figure 2.11: HEALPix sphere tesselation. Upper-left: the base 12 pixels tesselating the sky. Upper-right: The sides of the 12 basic pixels have been cut in two ($N_{\rm side}$), sub-diving each pixel into 4 smaller ones. Lower-right: subdivision with $N_{\rm side}$

The number of times (N) the side of the original rhombus is sub-divided is called

$$N_{\rm side} = 2^N. \tag{2.4}$$

This parameter gives the number of pixels covering the sky through the relation,

$$N_{\rm pix} = 12 \times N_{\rm side}^2. \tag{2.5}$$

Because the base resolution is 12 pixels if their side is divided $N_{\rm side}$ time, the number of pixels per basic pixel is $N_{\rm side}^2$, that for each of the 12 basic pixels. Using the number of pixels, one can deduce the resolution of each pixel knowing the surface of the sphere. The total surface of the sky can be computed, by first taking the surface of a sphere which is given by 4πr, where r, the radius can be deduced by looking at the length of a circle which is given by 2πr and represents 360 deg. Therefore, r=360/2π. This gives us that the total surface of the sky is, $S_{\rm sky} = 4\pi(360/2\pi)^2 = 41253$ deg^2. Therefore, the pixel resolution in deg is,

$$pix_{reso} = \sqrt{41253/N_{\rm pix}}. \tag{2.6}$$

Once the sphere is tesselated into pixels, one problem remains to do computations on those pixels, which is to number the pixel's position. HEALPix proposed two ways to index the pixels on the sphere. One is called RING ordering and consists of indexing the pixels one by one along their iso-latitude rings. The second consists in indexing the pixels within their original base pixel, in a NESTED kind one format (see Fig. 2.12).

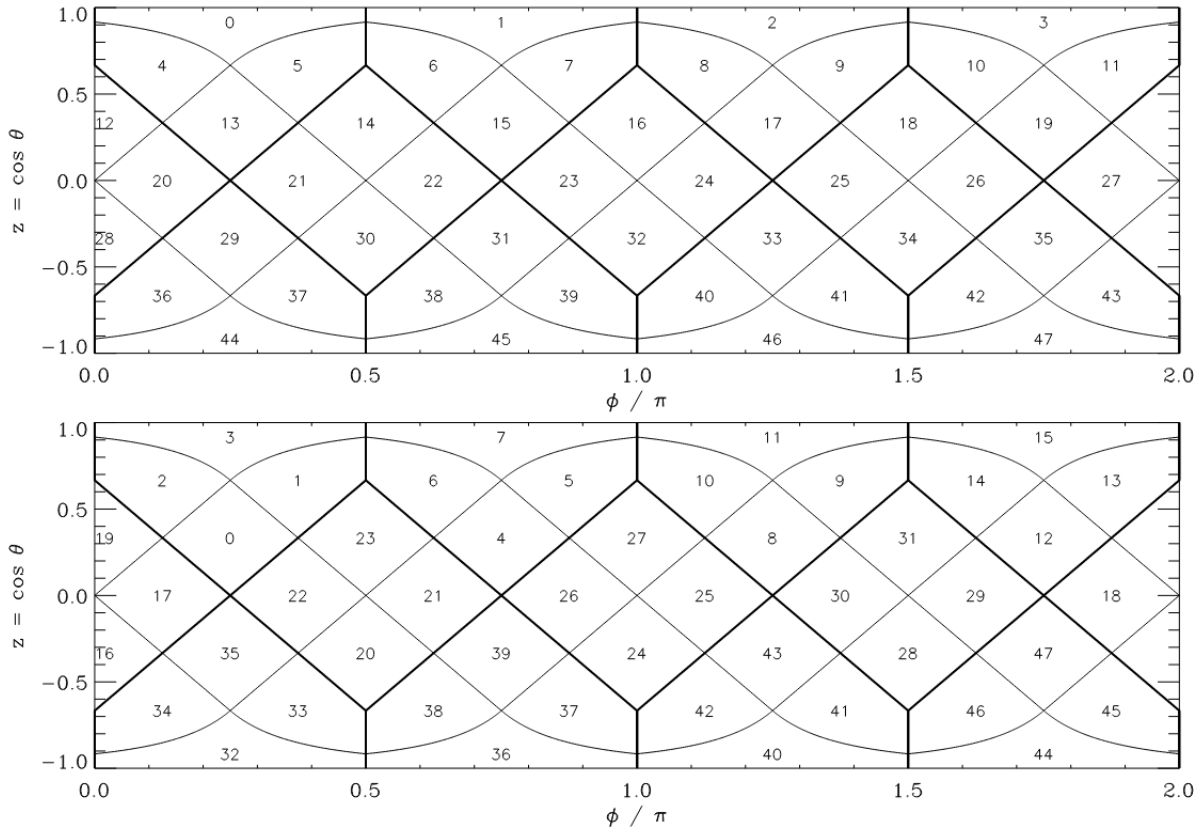

Figure 2.12: Example of the two pixels indexing schemes of HEALPix, at the top, the RING ordering, and at the bottom the NESTED ordering.

HEALPy

A Python implementation of HEALPix was developed by Zonca et al. (2019) and is called HEALPy. The HEALPy package is a state-of-the-art package for CMB data processing. It is composed of many basic functions that allow the processing of CMB data in the form of Flexible Image Transport System (FITS) maps. It is used extensively throughout this thesis.

However, the HEALPy smoothing function (`healpy.sphtfunc.smoothing()`) was incomplete. The function did not specify in its header that it could only function properly on a RING-ordered map. The function also did not display any messages or user feedback, nor did contain an automatic re-ordering of the map in case the provided input was wrongly ordered. This loophole first leads to higher-order mistakes in some analyses of this work, that were especially difficult to identify.

In practice, when creating an empty $N_{\text{side}} = 2048$ HEALPix map of NESTED ordering, and selecting a subset of pixels within the rhombus associated with $N_{\text{side}} = 4$. Using the HEALPy function `healpy.sphtfunc.smoothing()` to smooth with a 10-degree Gaussian resulted in the wrong result (see Fig. 2.13). The expected result would be a smoothed-out rhombus. This was a bigger problem[3] than it looked because some of the *Planck* maps were provided in NESTED ordering. Meaning, that when applying the state-of-the-art HEALPy smoothing function directly on them one would introduce a

[3] https://github.com/healpy/healpy/issues/677

systematic error in the result without any notice (see Fig. 2.14).

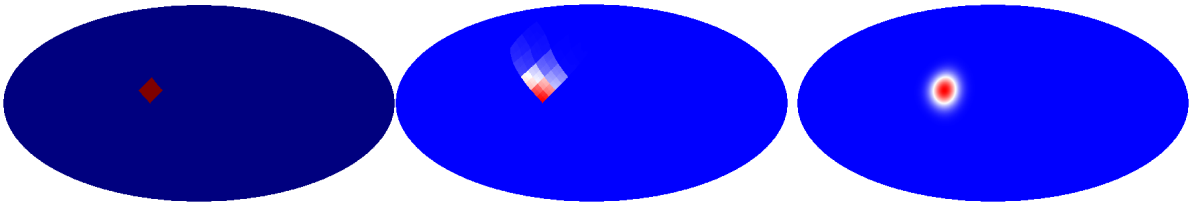

Figure 2.13: Three $N_{\text{side}} = 2048$ HEALPix maps. Left: NESTED map full of 0 except for a rhombus of size $N_{\text{side}} = 4$ full of ones. Middle: the smoothed version of the left map obtained using `healpy.sphtfunc.smoothing()` with a Gaussian of 10 degrees. Right: expected correct result when smoothing the left map with a 10-degree Gaussian.

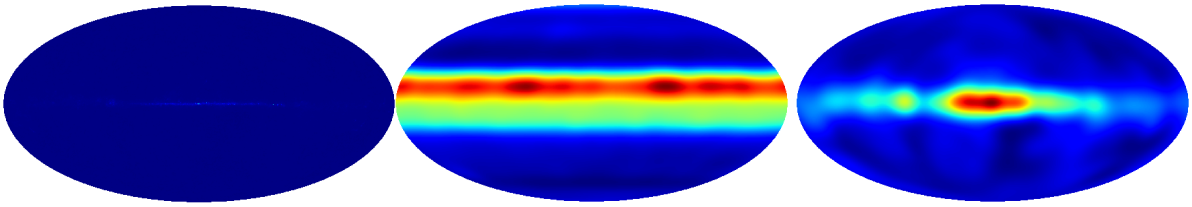

Figure 2.14: Three $N_{\text{side}} = 2048$ HEALPix. Left: *Planck* NESTED map for the 100 GHz frequency channel. Middle: the smoothed version of the left map obtained using `healpy.sphtfunc.smoothing()` with a Gaussian of 10 degrees. Right: expected correct result when smoothing the left map with a 10-degree Gaussian.

I addressed[4] the incompleteness of the HEALPy smoothing function, by implementing in this function a new parameter called 'nest' which is by default set up to False. This makes it clear in the function header that the ordering of the maps matters. Another addition was made to the core of the function. If the user feeds a NESTED map, the smoothing function will temporarily re-order it to RING ordering to perform the correct smoothing of the map before re-ordering it to NESTED. Therefore, the `healpy.sphtfunc.smoothing()` function will now return the correct output no matter the ordering of the input map. Those implementations were accepted by the HEALPy team and are now available in the latest versions of the package.

2.4.2 Angular power spectrum computation

Once the data is formatted into HEALPix maps it can be processed to extract scientific information. This work essentially looks at an estimator called the angular power spectrum of the maps. It tells us the evolution of the power of the sky emissions depending on the scales.

HEALPy anafast

As stated earlier (see Sec. 2.1.3), when observing the sky with CMB telescopes, the physical quantity observed is a temperature variation or an intensity variation over the full sky, this is what is represented by the power spectrum. The HEALPy software has a function called `healpy.sphtfunc.anafast()` that computes the power spectrum or cross-correlation of any full sky map of a HEALPix format. The

[4] https://github.com/healpy/healpy/pull/678

function simply performs the spherical harmonics analysis (see Sec. 1.3.1) on the full map. If some information is missing on the map, or part of the map is masked HEALPy will attribute the missing or mask pixels the value '0'. This, if not properly accounted for, can create spurious correlations between different regions of the maps which now have the same value without any physical meaning.

Pseudo-C_ℓ: PyMaster

There is a caveat to using the angular C_ℓ estimator to compute the power spectrum over the sky, which is, that we do not always observe the full sky area (see for example the SO sky strategy in Fig. 2.6). There are multiple reasons for not observing the full sky, ranging from pure survey strategy like trying to avoid the Galactic plane which is highly contaminated by strong emission (see Fig. 2.4) to limitations coming from the telescope itself. Indeed, ground-based telescopes, such as SO, located at a given point on the surface of the Earth will not be able to observe the full sky because of ground obscuration. These limitations mean that the $a_{\ell m}$ coefficients are not fully defined for all ℓ, m on the sky sphere, and the survey cut generates mode coupling. To circumvent that, several methods were constructed to correct for the incomplete sky coverage and estimate the true underlying power spectrum even in a case where only part of the sky is available. In this work, the Python package NaMaster (Alonso et al., 2019) and in particular its PyMaster package is used to compute the angular power spectrum of the maps for future ground-based telescopes with incomplete sky coverage, such as SO, FYST, and for concepts ideas like AtLAST. NaMaster is using the (Monte Carlo Apodized Spherical Transform Estimator) MASTER (Hivon et al., 2002) to compute the Pseudo-C_ℓ and retrieve an estimation of the true underlying angular power spectrum. PyMaster also deals with apodization around the survey area when it is not the full sphere. Apodization consists of smoothing the edges with a function so that it gradually goes to 0 instead of having a sharp cut. Sharp cuts can create ringing artefacts in Fourier space, thus messing up the power spectrum.

2.5 Signal extraction

When observing the microwave sky with telescopes like *Planck*, SO, FYST, or concept ideas like AtLAST, the total sky emission is measured. This means only the sum of all emissions coming potentially from different sources and processes is measured. For example, on the *Planck* maps, we see a band located in the middle of the maps (see Fig. 2.4) which is emission coming essentially from the Milky Way. In this work, and in CMB science in general we want to separate the different emissions that compose the raw data maps to study them independently. Indeed the power spectrum of some emissions, such as the thermal Sunyaev-Zeldovich (tSZ) effect, for example, is proportional to the clustering and density of matter in the Universe. Being able to obtain a pure map of this effect or of the CMB to obtain their true underlying power spectrum is essential to constrain such cosmological parameters. A category of signal processing methods called 'component separation methods' focuses on extracting a target signal out of the multi-component microwave sky.

2.5.1 Internal Linear Combination methods

A popular class of methods to separate components from the Microwave sky is called Internal Linear Combination (ILC). The ILC method was originally developed to process *WMAP* by Bennett et al. (2003). The ILC methods are multi-frequency 'blind' component separation methods. The term 'blind'

refers to the fact that they do not assume anything about the unwanted astrophysical emissions present in the sky and are treated as contaminants. To function, ILC methods, only need to know the Spectral Energy Distribution (SED) of the component to retrieve. Because ILC methods use the knowledge of the intensity variations of the target signal with frequency in the observation to extract it. That is why ILC methods are multi-frequency methods, meaning they function with observations of the sky at many frequencies to identify the component of interest.

Internal Linear Combination formalism

The Internal Linear Combination (ILC) (Bennett et al., 2003; Eriksen et al., 2004) is the original and simplest form of all the ILC flavours. It is the most popular and widely used method in CMB science. It is based on two hypotheses:

- Each observation map of the sky can be represented as a linear combination of the different astrophysical emissions that compose the sky.
- The Astrophysical components composing the microwave sky are uncorrelated.

An observation of the sky at a frequency i and pixel p can be decomposed as the sum between a signal of interest s and its frequency dependence information contained into a_i and a component containing all the noise present in the observation at this frequency i. The undesired astrophysical emissions are factored in the noise component. We have

$$m_i(p) = a_i s(p) + n_i(p). \tag{2.7}$$

This can be generalized to $\mathrm{N_{obs}}$ observations of the sky each at a different frequency, giving

$$\vec{M}(p) = \vec{A}s(p) + \vec{N}(p), \tag{2.8}$$

where $\vec{M}$ is the vector containing all the observed maps, $\vec{A}$ is called the mixing vector and encapsulates the SED information of the target signal at each frequency, and $\vec{N}$ the noises.

$$\vec{M} = \begin{pmatrix} m_1 \\ \vdots \\ m_{\mathrm{N_{obs}}} \end{pmatrix}; \vec{A} = \begin{pmatrix} a_1 \\ \vdots \\ a_{\mathrm{N_{obs}}} \end{pmatrix}; \vec{N} = \begin{pmatrix} n_1 \\ \vdots \\ n_{\mathrm{N_{obs}}} \end{pmatrix}. \tag{2.9}$$

The signal of interest $s(p)$ can be extracted from the multi-component multi-frequency sky through an estimator $\hat{s}_{\mathrm{ILC}}$ defined as,

$$\hat{s}_{\mathrm{ILC}}(p) = \sum_i \omega_i m_i(p) = \vec{W}^T.\vec{M}(p), \tag{2.10}$$

where ω_i is the a weight at the frequency i and,

$$\vec{W} = \begin{pmatrix} \omega_1 \\ \vdots \\ \omega_{\mathrm{N_{obs}}} \end{pmatrix}. \tag{2.11}$$

is the vector containing the ILC weights. The signal estimator constructed as a weighted linear combination of the signal and contaminants is able to extract the signal of interest by constructing the weights with two conditions:

- The weights must have a unit response to the spectral shape of the signal of interest, represented by the mixing vector A. This condition ensures recovery of the signal with the same amplitude and without distortions over the frequency range. This condition is formalised by,

$$\vec{W}^T.\vec{A} = \vec{A}^T.\vec{W} = 1; \tag{2.12}$$

$$\sum_i \omega_i a_i = 1. \tag{2.13}$$

- The weights must ensure that the variance of the estimated signal $\hat{s}_{\rm ILC}(p)$ is minimum.

The condition of minimal variance has appeared in the ILC method through empirical tests and the intuition that minimizing the estimator's variance will minimize its scatter around the true underlying value. However

$$VAR(\hat{s}_{\rm ILC}(p)) = VAR(\vec{W}^T.\vec{M}(p)), \tag{2.14}$$

and

$$VAR(\vec{W}^T.\vec{M}(p)) = \frac{1}{\mathrm{N_{obs}}-1}\sum_{i=1}^{\mathrm{N_{obs}}}(\vec{W}_i^T.\vec{M}_i(p) - \overline{\vec{W}^T.\vec{M}(p)})(\vec{W}_i^T.\vec{M}_i(p) - \overline{\vec{W}^T.\vec{M}(p)})^T. \tag{2.15}$$

This gives,

$$VAR(\vec{W}^T.\vec{M}(p)) = COV(\vec{W}^T.\vec{M}(p), \vec{W}^T.\vec{M}(p)) = \vec{W}^T C\vec{W}. \tag{2.16}$$

Where, VAR() is the variance, COV() is the covariance, and C is the covariance matrix of the maps $\vec{M}(p)$. Minimizing the variance is equivalent to finding the mull of its derivate, which means that the condition of minimum variance is equivalent to

$$\frac{\partial}{\partial\vec{W}}VAR(\vec{W}^T.\vec{M}(p)) = \frac{\partial}{\partial\vec{W}}\vec{W}^T C\vec{W} = 0. \tag{2.17}$$

The expression of the ILC weights needed to recover the signal of interest can be deduced using the two conditions. First, Eq. 2.13 gives,

$$\vec{W}^T.\vec{A} = 1 \Leftrightarrow 1 - \vec{W}^T.\vec{A} = 0. \tag{2.18}$$

The second condition given by Eq. 2.17 can be solved using Lagrange's multipliers. Suppose λ is the Lagrange multiplier whose goal is to find the minimum of the function, it can be written such as,

$$\frac{\partial}{\partial\vec{W}}(\vec{W}^T C\vec{W} + \lambda(1 - \vec{W}^T.\vec{A})) = 0. \tag{2.19}$$

Computing the derivates gives,

$$2C\vec{W} - \lambda\vec{A} = 0 \tag{2.20}$$

$$\Leftrightarrow \vec{W} = C^{-1}\frac{\lambda}{2}\vec{A}. \tag{2.21}$$

By multiplying on both sides with $\vec{A}^T$, it becomes,

$$\vec{A}^T\vec{W} = \vec{A}^T C^{-1}\vec{A}\frac{\lambda}{2}, \tag{2.22}$$

$$\Leftrightarrow \frac{\lambda}{2} = \frac{1}{\vec{A}^T C^{-1}\vec{A}}. \tag{2.23}$$

Injecting Eq. 2.28 into Eq. 2.21 gives the expression of the weights,

$$\vec{W}^T = \frac{\vec{A}^T C^{-1}}{\vec{A}^T C^{-1}\vec{A}}. \tag{2.24}$$

The ILC code gives an estimation of a signal of interest by summing the product between the set of multi-frequency observed maps and their frequency-associated ILC weights, weights that are solely constructed using the SED information of the signal contained in the mixing vector $\vec{A}$ and the covariance matrices of the data C (Eq. 2.24 injected into Eq. 2.10). A schematic of the method is given in Fig. 2.15. However, an offset to the ILC estimator can be introduced if the set of multi-frequency maps has a non-zero mean, as this does not modify the variance of the maps, it will not be dealt with by the ILC. To correct that, the median of the input maps should be subtracted from the estimated map.

ILC example on *Planck* data

The way an ILC function can be illustrated on *Planck* data (see Sec. 2.1.3). Let's consider the *Planck* DR2 data maps (See an example on Fig. 2.4) and take a cutout around the Coma cluster. It is one of the largest clusters known, to date, in the sky and it is located at (RA, DEC)=($12^h59^m48.7^s$,+27°58'50"). Before running the ILC on the cutouts, they have to be put to the same resolution, so that the maps can be linearly combined by the ILC. Each frequency map has to be smoothened down (see App. A.1) to the same resolution as the lowest frequency map, and therefore, lower resolution (see Eq. 1.23). The frequency bands at which each cutout is observed are not a Dirac function but are extended (see for example Fig. 2.3), the shape and transmission of the frequency band have to be taken into account. The ILC weights to extract the tSZ signal coming from the coma cluster out of the set of multi-frequency (see Table 2.1) *Planck* data can be computed using Eq. 2.24. Each cutout is multiplied by its associated ILC weight and the results are summed resulting in the ILC-extracted tSZ map of the Coma cluster (see Fig. 2.16). This shows, that already with only the frequency channels up to 353 GHz, the tSZ signal coming from the Coma cluster at the centre of the cutout is visible. Meaning the frequency channels above 353 GHz, i.e., 545 GHz and 857 GHz only contribute to minimizing the noise of the recovered tSZ map. It is visible in Fig. 2.16, where the background noise greatly diminishes when including the high-frequency channel bands. The same method can be applied to the recovery of any signal of known

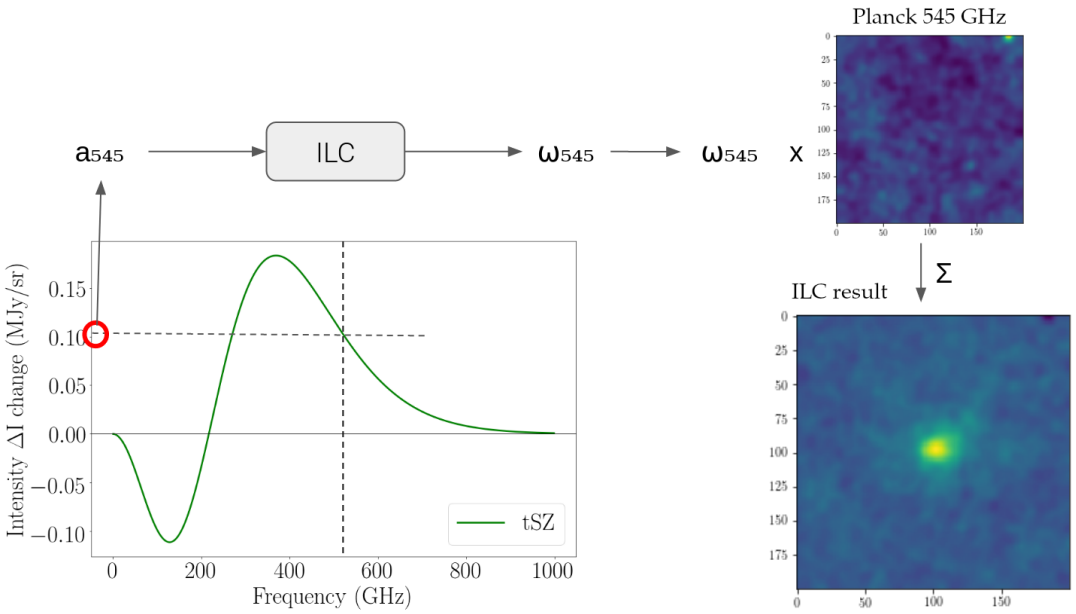

Figure 2.15: Schematic of the ILC method. Example on *Planck* data, around the Coma cluster, for the Frequency 545 GHz. The ILC takes as input the intensity change generated by the tSZ effect at the Frequency 545 GHz and the covariance matrix of the data to compute the ILC weight at this particular frequency (ω_{545}, see Eq. 2.24). The *Planck* 545 GHz map is then multiplied by ω_{545} and added to all the previously computed maps at lower frequency. The result is a Compton-y map of the Coma cluster.

SED, for example, the CMB.

Constrained Internal Linear Combination

The ILC recovers an estimation of the underlying target signal under a few assumptions. One is that the sky components are uncorrelated between them and the other is that the variance of the recovered map must be minimum. Meaning in cases where signals and contaminants are correlated, or when the variance of a contaminant does not impact the final map, it is not reduced by the ILC. Therefore, the recovered ILC map can be heavily biased against such contaminants. One way to circumvent that is to make use of the same technique as for the ILC. The idea behind Constrained ILC (CILC), is that if the SED of a contaminating signal is known, then the CILC can target its variation through multi-frequency data and remove it. This idea was first introduced by Remazeilles et al. (2011). With this extra information on one of the contaminants, the content of the observed maps can be represented as,

$$m_i(p) = a_i s(p) + b_i f(p) + n_i(p), \tag{2.25}$$

where m is a map of the sky at the frequency i, a_i the intensity change of the signal of interest $s(p)$ at this frequency, and b_i the intensity change of the contaminant to suppress $f(p)$. $n_i(p)$ contain all the other contaminants and noise at the frequency i. The generalization to N_{obs} frequency observations of

the sky gives,

$$\vec{M}(p) = \vec{A}s(p) + \vec{B}f(p) + \vec{N}(p), \tag{2.26}$$

where $\vec{M}$ is the vector containing all the observed maps, $\vec{A}$ is the mixing vector of the component of interest, and $\vec{B}$ is the mixing vector of the component we want to suppress. The same estimator as for the ILC remains (see Eq. 2.10), The CILC also tries to find a set of weights whose linear combination with the set of multi-frequency maps returns the target signal. However, the CILC adds an additional constrain on the construction of the weights, on top of Eq. 2.13 and Eq. 2.17, the CILC weights ($\vec{W}$) must verify

$$\vec{W}^T.\vec{B} = \vec{B}^T.\vec{W} = 0; \tag{2.27}$$

This means that the CILC weights must null the known contaminant using its spectral variation with frequency. Mathematically, this is a projection of the vector $\vec{W}^T$ onto the vector $\vec{B}$. Because this projection is constructed to return 0, it is often referred to as a deprojection of the contaminant through its SED. Finding the expression of the weights is done through the same method as for the ILC, using Eq. 2.17 and the two conditions on the weights (Eq. 2.13 & Eq. 2.27) but this time with two Lagrange multipliers λ and μ. For the CILC Eq. 2.19 is transform into,

$$\frac{\partial}{\partial \vec{W}}(\vec{W}^T C\vec{W} + \lambda(1 - \vec{W}^T.\vec{A}) - \mu\vec{W}^T.\vec{B}) = 0. \tag{2.28}$$

Where C is the covariance matrix of the maps. Applying the same technique as for the ILC, and multiplying by both sides by $\vec{A}^T$ or $\vec{B}^T$, a system of two-equation can be deduced and solved to obtain the expression of λ and μ. Injecting their expression back yields the expression of the CILC weights,

$$\vec{W}^T = \frac{(\vec{B}^T C^{-1}\vec{B})\vec{A}^T C^{-1} - (\vec{A}^T C^{-1}\vec{B})\vec{B}^T C^{-1}}{(\vec{A}^T C^{-1}\vec{A})(\vec{B}^T C^{-1}\vec{B}) - (\vec{A}^T C^{-1}\vec{B})^2}. \tag{2.29}$$

Of course, if the SED of more contaminants is known one can generalize the CILC to deproject more than one contaminant. However, there is an important caveat to that, which is every deprojection is an additional condition on the weights, which reduces the solution space, or said differently there are fewer degrees of freedom, leading to what is known as a noise penalty, or an increased noise in the recovered signal map (Hurier et al., 2013).

Residual noises

This work aims to study the leakage of one component into the target signal. To extract the target signal, this work uses ILC methods. However, those methods do not retrieve a fully pure, noise-free signal. Some residual noise, leftover from the presence of other Astrophysical emissions, considered contaminants are present in the recovered map. This can be seen through Eq. 2.10, expanding it gives,

$$\hat{s}_{\rm ILC}(p) = \sum_i \omega_i a_i s(p) + \sum_i \omega_i n_i(p), \tag{2.30}$$

Therefore, with the ILC condiction on the weights (see Eq. 2.13),

$$\hat{s}_{\rm ILC}(p) = s(p) + \sum_i \omega_i n_i(p). \tag{2.31}$$

The total residual noise, left over by the ILC is defined as,

$$r_{\rm ILC} = \hat{s}_{\rm ILC}(p) - s(p) = \sum_i \omega_i n_i(p). \tag{2.32}$$

The exact same formula holds for the CILC case because of the conditions on the weights. n_i can also be expanded into each of the known astrophysical components in the sky, making it possible in the case where the true underlying sky emission is known to compute how much this component will contribute to the noise residual by simply multiplying it at each frequency by the ILC or CILC weights used to recover the signal of interest. Let's suppose an astrophysical emission X, the ILC noise residual of X (r_X) is,

$$r_X = \sum_i \omega_{{\rm ILC},i} X_i(p). \tag{2.33}$$

Where $\omega_{{\rm ILC},i}$ is the ILC weight obtained to retrieve the signal of interest $s(p)$ at a frequency i and $X_i(p)$ the map of the contaminant X.

Advantages and disadvantages of the ILC/CILC

The advantages of the ILC/CILC are,

- assumptions are simple
- only the SED of the component of interest needs to be known or assumed. In the CILC case, the SED of the component to be deprojected also needs to be known.

The disadvantages of the ILC are,

- the noise residual bias of the recovered signal, and its contribution to the variance make the ILC algorithm perform poorly when the signal-to-noise of the data is not so good.
- The ILC and CILC are only based on the knowledge of frequency variation of the signals and are blind to their spatial distribution or angular spectra variation. When the SED taken is not fully representative of the signal behaviour, a bias is introduced in the estimated signal.
- the assumption of the component being uncorrelated is rarely verified. This can bias the estimation of the signal

Other ILCs flavors

Apart from the extension of deprojecting many contaminants, over the years many other flavours of ILC were developed. Among the notable ones are the Needled ILC (NILC), first used on *WMAP* data (Delabrouille et al., 2009) and later used by the *Planck* team for their Compton-y map or CMB map. Needlets are wavelet functions on the sphere, the idea behind NILC is to filter the set of multi-frequency maps using needlets, which are filtered in ℓ-space, and then apply the ILC on each ℓ-space decomposition

of the set of maps. This has the advantage, compared to an ILC, that the decomposition of the maps onto some scale functions makes it possible to apply independent ILC to them allowing for the ILC to better target and or minimize signals that dominate over a special scale range. Modified ILC Algorithm (MILCA) (Hurier et al., 2013) is another method used by the *Planck* team. The MILCA filter the maps spatially by defining zones in the sky based on the HEALPix Nested scheme on which to apply their ILC and filtering the maps in ℓ-space using Gaussian kernels functions. Additionally, the MILCA not only constrains its weights to minimize the variance of the recovered signal map but also of the noise component. Finally, it is merged with a CILC as it allows to deproject components accordingly. Another called Partially Constrained ILC (PCILC) is based on, such as the CILC, component deprojection but instead of setting them to 0, this method allows to constrain them by any value between [0,1]. Therefore, it partially constrains the contaminants in the maps. This is useful in cases where the contaminant does not necessarily need to be nullified but simply controlled under a given threshold, this mitigates the noise penalty and reduces the variance of the recovered map by a factor between 2 to 3 compared to CILC (Abylkairov et al., 2021). A more complex version of NILC, called Sale-discretised ILC (SILC), also proposes to filter the set of maps according to some ℓ window functions but this time with a direction on the sphere, these added parameters seem to yield higher Signal-to-Noise CMB maps (Rogers et al., 2016). This is not an extensive list but gives an overview of the possibilities, modifications, and added complexity that can be encompassed in the ILCs methods.

2.5.2 Full-sky strategies

The ILC or CILC can be applied to arrays, patches of the sky, or onto full sky HEALPix maps such as was the case for *Planck* data (see Fig. 2.4). However, the data exhibit emissions which strongly associated with specific regions of the sky, especially the emission coming from the Milky Way (MW). Therefore, it is unoptimized to apply only one ILC and get only one set of weights for the full sky, the same weight for regions that are highly contaminated by the MW, and those that are not. Partitioning the sky in several regions, to apply many separate ILCs and obtain several sets of weights each optimized in function of the degree of contamination by the MW makes sense and was considered early on in the signal extraction area of CMB science (Tegmark et al., 2003). This study implements a simple tesselation of the sky into equal surface area regions using the HEALPix nested scheme (see Fig. 2.11). $N_{\rm side} = 4$ is taken, resulting in 192 rhombi of equal area covering the sky. This number was the best combination to avoid having too small sky tesselation or too big.

2.5.3 Python package

To ensure that the ILC and CILC code created could be reused easily a Python package publicly available on GitHub[5] was created and called CILC. The correct functioning of the code was tested on the *Planck* PR2 data and compared to the literature. The test was performed on the recovery of the Compton-y signal from the *Planck* data.

Test of the ILC/CILC wieghts

The first thing that can be checked, to ensure the ILC/CILC is properly functioning, is the ILC/CILC weights. They are simply constructed from the covariance matrix of the data and the mixing vector (see

[5] https://github.com/MaudeCharmetant/CILC

Eq. 2.24 and Eq. 2.29). The weights can be displayed in the function of their associated frequency, which gives information on which frequency band contributes to recovering the signal and minimizing the noise because those are the two sole conditions on the weights (see Eq. 2.12). To estimate in which proportion each frequency band contributes to the recovery of the COmpton-y signal, in this case, one can look at $\omega_i a_i$ where ω_i is the weight at the frequency i and a_i the value of the tSZ SED at this frequency. This comes from Eq 2.13 and

$$\sum_i \omega_i a_i s(p) = s(p), \tag{2.34}$$

thus giving for each frequency the factor of each band to the tSZ reconstruction. In the case of *Planck*, 143 GHz, 353 GHz, and 545 GHz are the three main contributors to the Compton-y reconstruction. Respectively with a factor of $\omega_{143} a_{143} \approx 0.65$, $\omega_{353} a_{353} \approx 0.4$ and $\omega_{545} a_{545} \approx -0.05$ (see top panel of Fig. 2.17). Those values are in complete agreement with the literature (Erler et al., 2018). The 217 GHz was attributed the second largest weight in absolute value by the ILC (see Bottom panel Fig. 2.17) but does not contribute to the Compton-y reconstruction. This can be explained by the absence of a tSZ signal at this frequency close to the null of SED (see Fig. 1.23). At 217 GHz the *Planck* map contains only the noise contaminants, making it a very useful frequency band for the ILC to target and minimize the noise. The exact same conclusions can be drawn for the CILC weights, which do not differ much from the ILC weights, this was also previously reported in the literature (Alonso et al., 2018).

A second check is to compare the extracted Compton-y angular power spectrum (C_ℓ^{yy}) to the one from *Planck*. The *Planck* team used more advanced ILC methods such as NILC and MILCA (see Sec. 2.5.1). Therefore, we do not expect a similar result with the ILC/CILC package developed in this work but results should be and are consistent. The ILC and CILC extracted Compton-y power spectrum is exactly within the *Planck* NILC and MILCA ones on large scales ($\ell \in [100, 1000]$) and exhibits a bit more noise on smaller scales ($\ell > 1000$) as expected from the less refined techniques (see Fig. 2.18). The *Planck* NILC and MILCA power spectrum were computed from the publicly available *Planck* and MILCA maps. The noise penalty of deprojecting CIB, in this case, is visible in the CILC-extracted power spectra, which on small scales has more power than the ILC-extracted one.

2.6 Conclusion

CMB science is still a very active field of research with many upcoming ground-based telescopes such as the Simons Observatory (SO), and the Fred Young Submillimeter Telescope (FYST). The field is also rich in concepts and ideas that might develop later as projects if their feasibility is verified, such as the Atacama Large Aperture Submillimeter Telescope (AtLAST), and many others. They all have different characteristics and objectives. Their comparison to find the best way to combine them is essential to understanding how to make the most of those cutting-edge telescopes. Many Python package already exists to process their data easily and consistently between the different research team, maximizing the possibilities of independent cross-checking. The next Chapter concentrates on one important aspect to bring the most out of that experiment and existing tools, which is how to set up realistic forecasts.

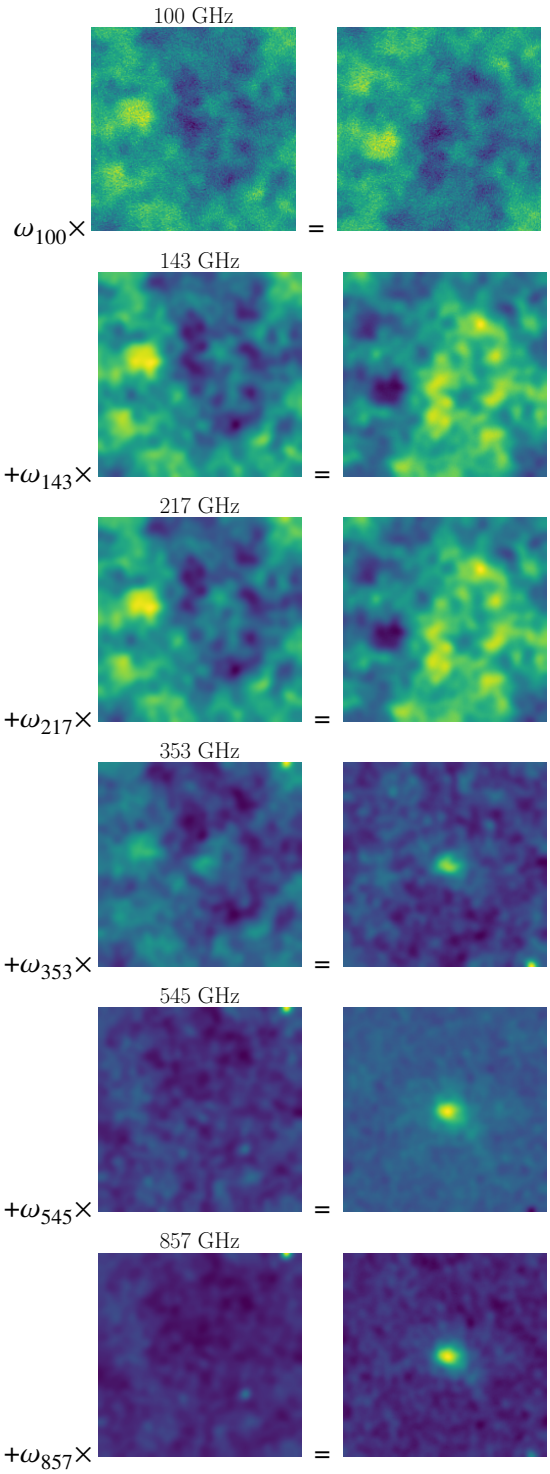

Figure 2.16: $5.6° \times 5.6°$ maps. Left column: Each map is a cutout of a *Planck* DR2 full sky HEALPix map at a given frequency and around the Coma cluster. The maps are in units of MJy/sr. Each cutout is multiplied by the frequency associated with its ILC weights. Right column: the cumulative sum of the ILC weights multiplication with the *Planck* cutouts, in a case where the tSZ signal is being recovered. The maps are in Compton-y units.

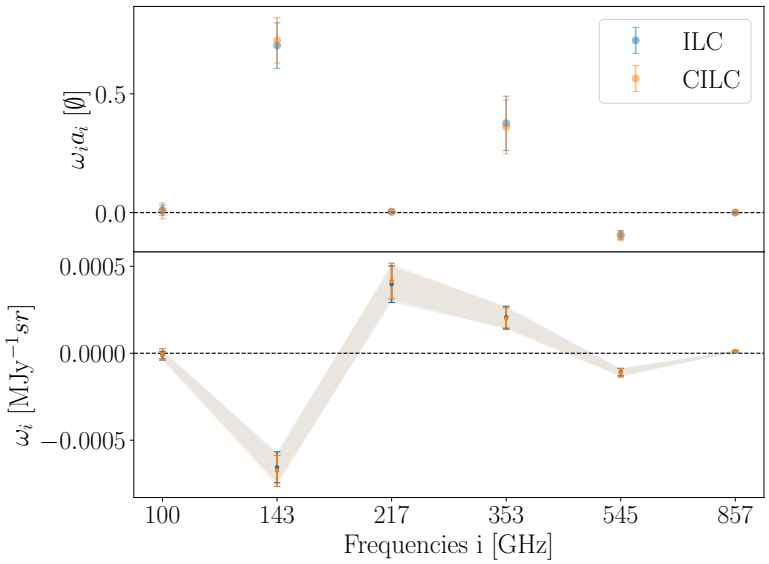

Figure 2.17: Weights returned by the ILC in blue or CILC in orange when extracting the Compton-y signal on *Planck* data, in the function of their associated frequency band. Top panel: the fractional contribution of each frequency to the recovered Compton-y. Bottom panel: the values of the weights, representing how much a band contributes to the recovery of the signal and minimization of the noise.

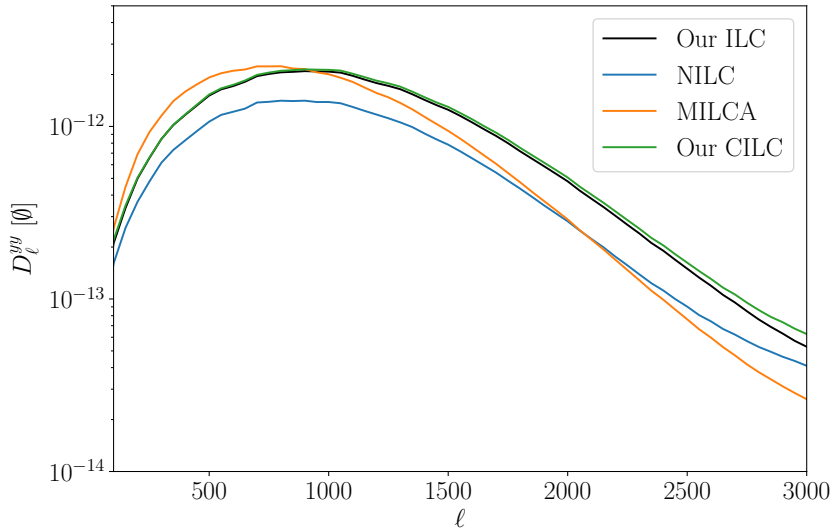

Figure 2.18: Angular power spectrum of the Compton-y maps extracted using the ILC developed in this work in black and the CILC in green. Blue is the Compton-y power spectrum of the *Planck* NILC-extracted map and in orange using the MILCA method.

CHAPTER 3

Goals and context of this work

This work focuses on the constraining power of upcoming ground-based telescopes such as the Fred Young Submillimeter Telescope (FYST), the Simons Observatory Large Aperture Telescope (SO LAT) for the two main Sunyaev-Zeldovich effects, the thermal Sunyaev-Zeldovich (tSZ) and the kinematic Sunyaev-Zeldovich (kSZ). Some results for the concept idea for the Atacama Large Aperture Submillimeter Telescope (AtLAST) are also presented.

3.1 Context of the study

This work is part of the observational cosmology field, meaning its end goal is a better understanding of the history and processes of the Universe. In particular, this work focuses on making predictions for the Sunyaev-Zeldovich effects measurements with upcoming telescopes. This raises the question of the past and current state-of-the-art of SZ observations.

3.1.1 Observations of the SZ effects

This part summarizes the observations of tSZ and kSZ effects by chronological order.

- **1973-1978:** The first attempts to measure SZ effects occurred right after its predictions by Sunyaev and Zeldovich (1969, 1970) and where target at the strongest signal, the tSZ, and the easiest part of the signal to measure, which is the tSZ decrement at frequencies below $\nu < 220$ GHz (see Fig. 1.49) because of its characteristic negative signal (Gull and Northover, 1976; Lake and Partridge, 1977; Birkinshaw et al., 1978).
- **1984:** Using a 40 m dish, stronger detection of the tSZ intensity change at the decrement was achieved (Birkinshaw et al., 1984) at $\nu \sim 20$ GHz.
- **1989-1993:** The first interferometric observations of the tSZ decrement, gives the first tSZ image (Moffet and Birkinshaw, 1989; Jones and et al., 1993; Grainge et al., 1993).
- **1994:** Measurement of the tSZ effect at 150 GHz on one cluster (Wilbanks et al., 1994).
- **1996:** Higher frequency interferometer measurements $\nu \in [30, 90]$ GHz (Grego et al., 1996).

- **1998:** New measurements of the tSZ decrement at 150 GHz at a first measurement of the tSZ increment at 350 GHz (Désert et al., 1998; Glenn et al., 1998; Lamarre and et al., 1998).
- **1999-2001:** First SZ clusters images using interferometers and first imaging of the SZ increment (Komatsu et al., 1999; Pointecouteau and et al., 2001).
- **2003:** The APEX-SZ telescope delivers the first large sample of tSZ images (Schwan and et al., 2003).
- **2007-2010:** First large sample of tSZ sources detected through interferomety (Muchovej et al., 2007; Zwart and et al., 2008). First high-quality SZ increment measurements (Zemcov et al., 2007; Nord and et al., 2009) and the first detection of tSZ intensity change at $\nu > 600$ GHz (Zemcov and et al., 2010). The first image of shock-heated gas in galaxy clusters made through tSZ (Mason and et al., 2010) and the first galaxy cluster catalogue detected by tSZ effect (Staniszewski and et al., 2009; Menanteau and et al., 2010).
- **2012:** High resolution tSZ measurements of the increment (Zemcov and et al., 2012) and first indication of a non-zero kSZ effect at 2σ (Mroczkowski and et al., 2012).
- **2013:** measurements of the tSZ at high-resolution, indicating the need for relativistic corrections and measurement of the kSZ effect (Sayers and et al., 2013).
- **2016:** the *Planck* satellite delivers the largest sample of galaxy cluster detected through tSZ effect (Planck collaboration et al., 2016) and measurement of shocks in galaxy clusters Basu et al. (2016).
- **2018:** *Planck* data is used to measure the tSZ spectra and shocks (Erler et al., 2018).

3.1.2 Up-coming telescopes

Over the past 50 years, many measurements of the SZ effects have been realized and have vastly improved. However much remains to be done, especially in the domain of the power spectra of the tSZ and kSZ effects, that have not yet been directly detected by experiments over all scales. The difficulty in detecting their power spectra is due to the many other astrophysical emissions present in the sky, the weak amplitude of both effects and for the kSZ, its spectral shape is similar to the CMB, making their separation especially difficult. Even with good observation, the question of the purity of the signal remains, especially for the tSZ effect, as mentioned in sec. 1.5.1, its spatial correlation with the CIB makes it difficult to get biased measurements of the tSZ effect.

Up-coming high-resolution ground-based telescopes such as SO LAT, and FYST could radially improve the measurements of both the tSZ and kSZ effect because of their high number of detectors, an exceptional site on one of the earth's dryest places, allowing for very low water vapour and good atmospheric transmission (see Fig. 2.10). In the AtLAST concept idea case, the large 50 m dish could allow the measurement of pressure, temperature, and velocities on single individual galaxies through the tSZ and kSZ effects.

Making predictions on what those telescopes could achieve in terms of tSZ and kSZ power spectra detection is essential to motivate their construction, and their schedule and optimize their frequency coverage. SO LAT is already fixed at frequencies $\nu < 280$ GHz which does not really probe the frequencies where the CIB, main bias of the tSZ measurements lies. Indeed, CIB contamination

dominates over frequencies $\nu > 350$ GHz, which are the frequencies probed by the FYST. Therefore, one can expect that one good constraining of the CIB contamination into tSZ measurements should come from combining SO LAT and FYST data.

3.2 Objectives of the project

This part lists the different objectives of this work within the context of SZ science. Three projects were explored, all revolving around predictions of SZ measurements for upcoming telescopes.

3.2.1 Setting realistic microwave sky simulations

To make predictions for upcoming telescopes, i.e., telescopes that are not yet built, that have not yet started observing or in the case of AtLAST that are in their concept phase, it is important to set up simulations of what those telescopes would see. At the time of this work, only low-resolution or incomplete codes were publicly available to simulate the microwave sky. This lack, motivated this project, to create a Python package to generate high-resolution maps of the microwave sky including all the known microwave emissions. In Chap. 4, is given a complete description of how these sky simulations Python package was developed. For all three telescopes explored in this project, SO LAT (see sec. 2.2.2), FYST (see sec. 2.3) and the concept idea AtLAST (see sec. 2.2.3), their observation frequencies ranges from 30 GHz to 860 GHz, which covers the submillimeter, microwave part of the electromagnetic spectrum (see Fig. 1.9). A complete description of the different emissions that compose the microwave sky is given, along with their origin and modelling. It is important to note that no simulation is realized in this project, HEALPix template maps from existing simulations are used and adapted to reproduce stat-of-the-art realistic mock sky maps for SO LAT, FYST, or for concept ideas such as AtLAST-like. The characteristic of each telescope, its beam, instrumental noise, and frequency of observation is simulated. Finally, the chapter touches on the creation of a Python package to automatize the transformations of the based templates to generate with a simple function any mock sky map with the desired emissions, sensitivities, resolution, and frequency.

3.2.2 Prediction of tSZ observation with upcoming telescopes

The tSZ effect is an essential tool for cluster astrophysics and cosmology (see Sec. 1.5.1). Therefore retrieving an unbiased tSZ signal is essential to deducing the proper values of things such as cluster pressure or cosmological parameters. The tSZ power spectrum is buried under two main contaminants, the CMB power spectrum on large scales and up to frequencies $\nu < 280$ GHz and above those frequencies, on small scales by the CIB leakage. The CMB power spectrum is very well-known and well-constrained. New high-resolution telescopes will probe smaller scales than before, and therefore the CIB bias in tSZ will become a major issue in measurements. FYST's high-frequency coverage seems to be a very good way to probe the dominant CIB contamination, constrain it, and eventually remove it. The CIB is composed of different unresolved sources at different redshifts, each described by a different modified blackbody SED, it can not be fully removed by a CILC assuming only one SED. This is why CIB contamination is difficult to get rid of, and why making predictions and combining data from different telescopes to optimize the frequency coverage might help mitigate the CIB contamination. Using the simulation scheme set up in Chap. 4, the purity and detection of the tSZ power spectrum can be estimated for telescopes like SO LAT, and FYST, or concept ideas such as AtLAST. This is done in Chap. 5,

through the progressive build-up of a realistic sky as seen by those telescopes, to understand how each component affects the tSZ signal. the tSZ signal is extracted using an ILC or CILC component separation method over the full sky in map-space, using some additional filtering to optimize the signal extraction. The detectability of the tSZ power spectrum and the CIB ILC or CILC noise residual is then computed for different cases, SO LAT alone, SO LAT and FYST combined with different frequency bands, to see if the CIB noise can be reduced. This is tested in different sky configurations, with only the signal and noise, without atmospheric noise, and with atmospheric noise.

3.2.3 Detection of the kSZ power spectrum with future telescopes

Detection of the kSZ effect has been historically difficult due to the weak amplitude of the signal. One of the main motivations of the effect is to potentially detect the missing baryons, that could be located at the outskirts of galaxy clusters. To increase the chances of detecting the effect, even with the new high-resolution telescope the idea of this new estimator $C_\ell^{y y_{\rm kSZ}^2}$, based on the cross-correlation between the tSZ signal and the kSZ signal squared emerged. Some previous studies already did exactly that but between the cross-correlation of tSZ and the galaxy density field. tSZ and kSZ are spatially correlated and the tSZ signal being much stronger one could expect to use it to grab the kSZ signal out of the noise. The goal is to estimate if, in a very ideal mock sky configuration containing only the signal and noise, detection of the power spectrum could be possible with upcoming telescopes such as SO LAT and FYST. The tSZ signal is obtained through a CILC on the mock sky to avoid kSZ and CMB leakage, which would bias the cross-correlation. kSZ is obtained through CILC to avoid tSZ contamination and then through Wiener filtering of the kSZ+CMB signal. This filtering step is essential as the ILC/CILC can not spectrally separate kSZ and CMB. The resulting map is then squared to avoid having a null, as kSZ can be positive and negative its correlation with the tSZ is supposed to be 0. This method has several caveats, the squaring of the map also squares the noise residual in the map, and the Wiener filter returns a biased estimation of the true underlying kSZ signal.

3.3 Conclusion

The general goal of this work, which encompasses all three projects, is to make predictions of the detection and purity of the tSZ and kSZ power spectra with future upcoming telescopes such as SO LAT, and FYST, and concept ideas such as AtLAST. This is done through the creation of a Python package that generates high-resolution realistic HEALPix mock maps of the Microwave sky with all of its various emissions in Chapter 4. Using this package to simulate the sky as seen by the above-mentioned telescopes, the detectability of the tSZ signal and the amount of CIB contamination leftover after ILC or CILC component separation is estimated in Chapter 5. Using the same method, a map of the tSZ and a map of the kSZ+CMB as seen by those telescopes are extracted. The kSZ+CMB map is filtered to remove the CMB which is a contaminant. The remaining kSZ map is then squared and cross-correlated with the tSZ map in the hope of pick-up the kSZ signal through its cross-correlation with the stronger, tSZ signal in Chapter 7.

CHAPTER 4

Skymodel: Simulating the microwave sky

Contribution

This chapter presents the development of a Python package made to generate high-resolution maps of the microwave from existing simulation template maps taken from Sehgal et al. (2010b); Thorne et al. (2017); Stein et al. (2018); Simons Observatory Collaboration (2019); Stein et al. (2020). The Skymodel package is publicly available on GitHub (`https://github.com/MaudeCharmetant/skymodel`) and was designed to be easy to use through the calling of a single function to generate the microwave sky with the desired components, units, resolution, and noise properties.

This chapter presents the history of microwave sky simulation and the current state-of-the-art of the field. It highlights the importance of modelling the microwave sky through the detailed astrophysical emissions that compose it. It presents the Python package developed in this thesis to generate maps of the microwave sky using existing template maps taken from Sehgal et al. (2010b); Thorne et al. (2017); Stein et al. (2018); Simons Observatory Collaboration (2019); Stein et al. (2020). This chapter provides a comparison of the different existing simulations the package is based on and the modifications that were done to the map to adapt them to any experiment. Finally, the problem of development, documentation, and publication of Astrophysical codes are addressed.

4.1 Emissions of the Microwave sky

The microwave sky is usually modelled as the sum of several emissions. Those emissions are separated into two categories. The extra-galactic emissions are composed of emissions coming from outside of the Milky Way galaxy, and the Galactic foregrounds, the emissions coming from the Milky Way. They are called 'foreground' because to us inside the Milky Way, they are the foreground compared to extra-galactic emission (see Fig. 4.1).

4.1.1 Galactic foregrounds

The Galactic foreground emissions coming from the Milky Way are composed of several emissions coming from different mechanisms. Free-free, Synchrotron, Galactic dust, and Anomalous Microwave

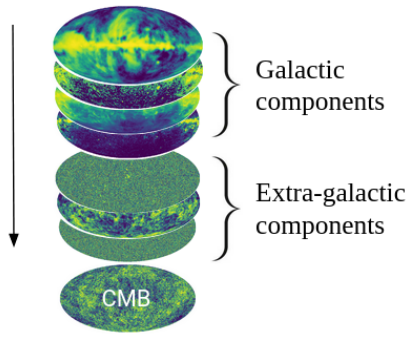

Figure 4.1: Illustration of the sky 'layered' composition. The CMB is in the Background and several extra-galactic and Galactic emissions are in the foreground. Only the total, summed-up sky emission is measured.

Emissions, are described below. Their polarization can be considered but was not used in this work and its description is, therefore, left out.

Galactic dust emission

The radiation from stars is absorbed by the dust grains. The heated dust grains then re-emit the light in the far-infrared domain and cool down. This re-emitted light is called Galactic thermal dust emission. In the first approximation, the intensity spectrum of the thermal dust emission is modelized by a modified blackbody

$$I_\nu = A\nu^\beta B_\nu(T_d), \tag{4.1}$$

where A is the amplitude, β the spectral index, B_ν the Planck spectrum and T_d the dust temperature. In reality, not all the dust grains will have the same temperature, this will depend on the Line of Sight (LoS) in our observation. Of course, dust grains along the LoS will also have different temperatures, but because we only observe the sky projected, this information is not often modelized. The same limitations exist for the spectral index (β) and the amplitude (A).

Free-free emission

Free-free emission occurs when, a free electron, in the sense not bound to a nuclei, is deflected by a charged particle. Causing the electron to decelerate and emit radiation. This electron remains free and unbounded, which is why this emission is called free-free. This is a particular type of Bremsstrahlung emission. It is modelled by a power law,

$$I_\nu = A\nu^\beta, \tag{4.2}$$

where A is the amplitude and β the spectral index.

Synchrotron emission

Synchrotron radiation is due to relativistic cosmic rays spiraling into the Galactic magnetic field emitting radiation (see Fig. 4.2).

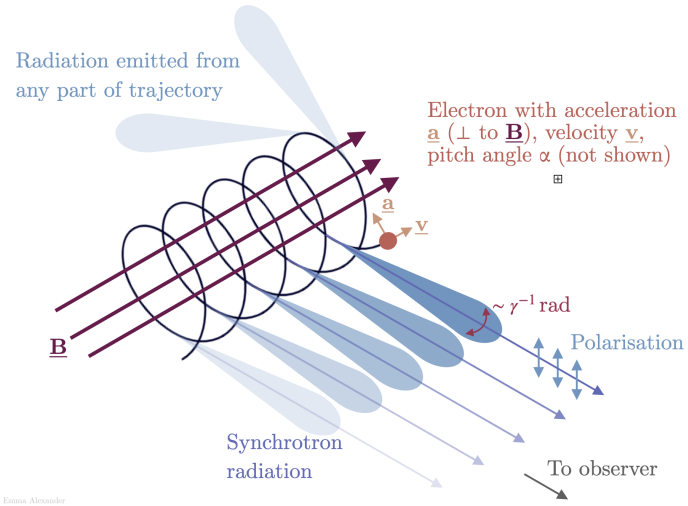

Figure 4.2: Illustration by Emma Alexander of the Synchrotron emission process.

The Synchrotron emission for only one electron is modelized as a broken power law. However, in the Galaxy, there are many different electron populations, all with slightly different amplitude and frequency cuts. The added contribution of all the electrons leads to modelizing the synchrotron emission as a power law (see Fig. 4.3), like for the free-free emission (see Eq. (4.2)). The spectral index and amplitude are different. Note that this modelization is also a simplification, the spectral index of the Synchrotron emission is expected to vary spatially and steepened with frequency due to Cosmic ray ageing (Platania et al., 1998; Kogut et al., 2011; Strong et al., 2011). However, only the projected Synchrotron emission along the Line-of-Sight is observed. This sum of all the different spectral indexes is expected to lead to a flattening of the Synchrotron spectral index.

Anomalous Microwave Emission

Anomalous Microwave Emission (AME) is supposed to be originating from spinning dust grains (see Fig. 4.4). The electric or magnetic dipole moment of dust grains produces radiation according to Larmor's formula:

$$P = \frac{2}{3}\frac{\omega^4 \mu^2 \sin\theta^2}{c^3}, \tag{4.3}$$

where P is the total power emitted, ω the angular frequency, μ the electric or magnetic dipole moment, and θ the angle between them. The emission law of AME can be estimated by integrating the Larmor law (Eq. (4.3)) over all the electric or magnetic dipole moments (Ali-Haïmoud et al., 2009).

4.1.2 Extra-galactic components

Emissions not coming from the Milky Way are called extra-galactic components. It is composed of all the extra-galactic foregrounds, the thermal and kinematic Sunyaev-Zeldovich effects (tSZ & kSZ),

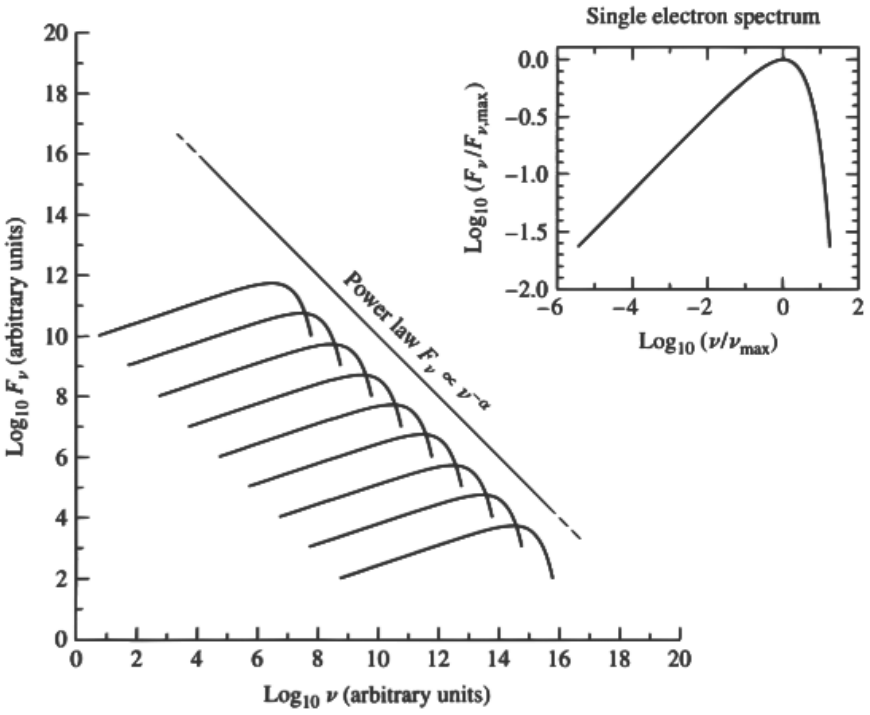

Figure 4.3: Synchrotron frequency spectrum, on the right for only one electron and the left for many electrons, leading the represent the frequency spectrum of the synchrotron emission as a power law. Extracted from Ertley (2014).

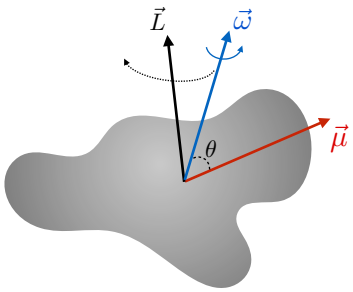

Figure 4.4: A spinning dust grain with its angular momentum $\vec{L}$, angular velocity $\vec{\omega}$, and electric or magnetic dipole moment $\vec{\mu}$, and θ the angle between them. Extracted from Dickinson et al. (2018).

the Cosmic Infrared Background (CIB), and the Radio Point Sources (RPS). The Cosmic Microwave Background (CMB) is the oldest extra-galactic emission able to reach us.

Cosmic Microwave Background

The Cosmic Microwave Background emission originates from the decoupling between the light and matter fluid following the formation of the first atoms, ~380 000 years after the big bang (see Sec. 1.1.1). Its SED is the closest observed thing to a perfect blackbody emission (see Eq. 1.32). Since its prediction,

it has been measured by several experiments, in space, with balloons, or on the ground (see Fig. 2.1 and Sec. 2.1.1).

thermal Sunyaev-Zeldovich effect

The thermal Sunyaev-Zeldovich (tSZ) effect is the intensity shift experienced by Cosmic Microwave Background (CMB) photons when their inverse Compton scatter on the hot electron present in the IntraCluster Medium (ICM). The tSZ effect is therefore a simple shift of part of the CMB photons population to higher energy because of their passage through galaxy clusters (1.5.1). The temperature change experienced by the CMB photons due to the tSZ effect is of the order of $\sim 100\mu K$ (Mroczkowski et al., 2019), making the tSZ effect a weak effect in comparison to the other components of the sky. The SED of the effect can be easily computed (see Eq. 1.23). The Compton-y parameter (see Eq. 1.48) can be approximated using the typical cluster temperature between $T_{\text{cluster}} \sim [1, 10]$ KeV, which gives $kT/mc^2 \approx 10^{-2}$. Clusters optical depth usually is around,

$$\tau_{\text{cluster}} = \int_{los} \sigma_{\text{T}} n_{\text{e}} dl \approx 10^{-2}, \tag{4.4}$$

giving $y \approx 10^{-4}$, which is the value assumed throughout this work. In this project, for simplicity, we do not consider the relativistic corrections to the tSZ effect.

kinematic Sunyaev-Zeldovich effect

The kSZ effect, is the Doppler boost experienced by CMB photons scattering on the electron, in a galaxy cluster, undergoing the bulk motion of the cluster compared to the CMB rest frame (see Sec. 1.5.2). Its SED can be easily computed (see Eq. 1.25). The average bulk velocity of galaxy clusters is $\vec{v} \approx 1000$ km/s.

Cosmic Infrared Background

Radiation emitted by stars, during their formation and through their lives is absorbed by dust and re-emitted into the infrared. This background radiation is called Cosmic Infrared Background (CIB). It is formalized as the modified blackbody or grey body spectrum (see Fig. 4.5).

$$I_\nu = \text{A}_\text{d} \nu^\beta \text{B}_\nu(\text{T}_\text{d}), \tag{4.5}$$

where, I_ν is the intensity at a frequency ν, A_d is an amplitude factor. β is the spectral index and $\text{B}_\nu(\text{T}_\text{d})$ the Planckian (see Eq. (1.32)) evaluated a the dust temperature T_d.

Radio Point Sources

Radio Point Sources (RPS) are unresolved sources emitting in radio. Their frequency spectrum is modelized as a power law,

$$I_\nu = \text{A}_\text{RPS} \nu^\beta, \tag{4.6}$$

where I_ν is the intensity at a frequency ν, A_RPS is an amplitude factor and β is the spectral index.

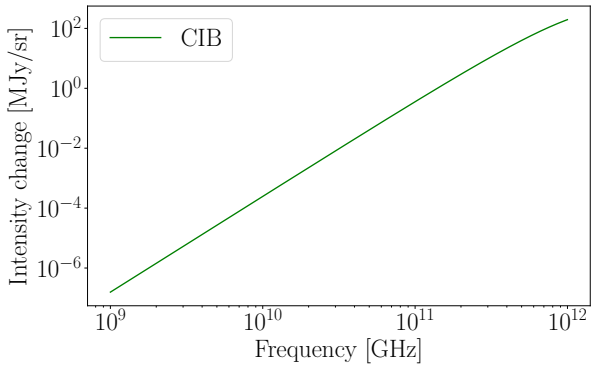

Figure 4.5: Frequency spectrum of the Cosmic Infrared Background, with $A_d = 15$, $\beta = 1.2$ and $T_d = 24$ K.

4.2 Context of microwave emissions research

The Python package developed in this work, to simulate microwave sky emissions is put into context through the history behind the first model developments. How this, led to nowadays state-of-the-art simulations. Their limitations are discussed, and a point is made about where this work fits into the bigger picture. An outlook is given on what are the next steps for the field.

4.2.1 History of microwave emissions

This part explores the history of microwave sky simulations, where the need for simulations came from and what was the first effort to simulate the microwave sky.

Cosmic Microwave Background

After (Gamow, 1948) prediction of the CMB and his estimated temperature of T= 340 K (see Sec. 1.1.1). Ralph A. Alpher, a former student of Gamov, found a few errors, in the computation of the matter density (ρ_{mat}) of the Universe. This led Alpher, to predict a temperature of T = 600 K, for the primordial gas at the condensation time and a current temperature of T = 5 K (Alpher, 1948). Which is only a factor two above the current Cosmic Microwave Background (CMB) temperature given by the most recent observations at $T_{\text{CMB}} = 2.72548 \pm 0.00057$ K (Fixsen, 2009). Those predictions started a series of tentative ground-based observations of this "primordial light" for the decade to come (see Sec. 2.2.1), culminating in 1965 with the first official measurement of the CMB by Penzias and Wilson (1965). After this groundbreaking discovery, the CMB temperature is measured at various wavelengths by the scientific community. Those measurements indicated the CMB spectrum to be consistent at each of those wavelengths to a blackbody spectrum (Partridge and Wilkinson, 1967; Weinberg, 1977). And led to the first studies investigating the possible anisotropies of the CMB (Partridge and Wilkinson, 1967). Peebles and Yu (1970) first predicted that at a 1' resolution, the rms temperature fluctuations of the CMB should be at least $\delta T/T = 0.00015$. The anisotropies and CMB autocorrelation predictions of Peebles and Yu (1970) and the first measurements of the RELIKT-1 satellite, lead to the first simulations of the CMB. Using Monte Carlo (MC), Scaramella and Vittorio (1990) gave upper limits on the maximum amplitude of the density fluctuations at their horizon crossing η_{H}, on the quadrupole and the shape of the autocorrelation function. All those predictions were followed by more precise measurements from space

(see Sec. 2.1.1). The need to design space experiments with targeted science goals is what gave rise to the first CMB simulation software, like CMBFAST (Seljak and Zaldarriaga, 1996) and later CAMB (Lewis, 2005).

Galactic foregrounds

The first simulations of the free-free Galactic emission were realized based on existing survey data, simply rescaling the map to account for the dust absorption (Dickinson et al., 2003). These simulations were then refined, in the following studies, notably using more data from the WMAP satellite to get a better estimation of the signal (Miville-Deschênes et al., 2008). It is also the first observation, from COBE/FIRAS, which, by giving some hints on the spectral index of dust (see Eq. 4.1), started the first series of Galactic dust simulations (Finkbeiner et al., 1999). Using the first free-free and synchrotron maps simulations, (Miville-Deschênes et al., 2008) subtracted them from WMAP data to obtain the first map of the AME.

Extra-galactic foregrounds

Point sources-like backgrounds and radio point sources were first simulated using the Poisson distributions, which were found to reproduce the statistics of the point sources distribution (Argüeso et al., 2003; González-Nuevo et al., 2005)

$$f(k,\lambda) = \frac{\lambda^k e^{-\lambda}}{k!}, \tag{4.7}$$

where, f is the probability distribution function of the number of occurrences k and λ the variance of the expected value. Regarding Sunyaev-Zeldovich effects, the first simulated maps of the tSZ effect were obtained using X-ray cluster data (Markevitch et al., 1992).

Conclusion

The necessity of simulating the different emission that composes the microwave sky naturally emerged from the necessity to make predictions for the first upcoming Microwave experiments. General codes to model all emissions from the microwave sky were set set-up before the first large CMB experiments. Two important examples of that were the Global Skymodel (de Oliveira-Costa et al., 2008) and the Planck Sky Model (Delabrouille et al., 2013). They were first based on observations or parametric modeling but nowadays powerful algorithms can simulate the microwave sky from Dark Matter simulations where the physics is then implemented on the collapsed large-scale structure.

4.2.2 Necessity for microwave sky simulations

Simulations of the microwave sky are essential to make predictions for upcoming telescopes that need clear science goals and a list of their possible achievement to secure financing. But also because modeling the sky allows us to compare our current understanding to real data. To simulate the microwave sky, there are three different techniques.

Dark matter only simulations

They consist of populating a cube with several particles. That particle, depending on the resolution of the simulation can represent, a galaxy cluster, a few galaxies, a galaxy, or a star depending on the weight of the particle in solar mass. Those particles are left to evolve under gravity to reproduce the clustering of matter in the given volume of the cube. Then, the clustered halos need to be painted on, meaning the mass and structure of those clustered particles have to be related to physical quantities such as density, luminosity, and effects produced by the 16% of baryonic matter composing the total mass. This is done through Halo Ocuppation Distribution (HOD) and empirical relations. An example of such simulations for CMB science is the Sehgal et al. (2010a) used in this work. Or the famous Millennium simulations[1] (see Fig. 1.17) (Springel et al., 2005).

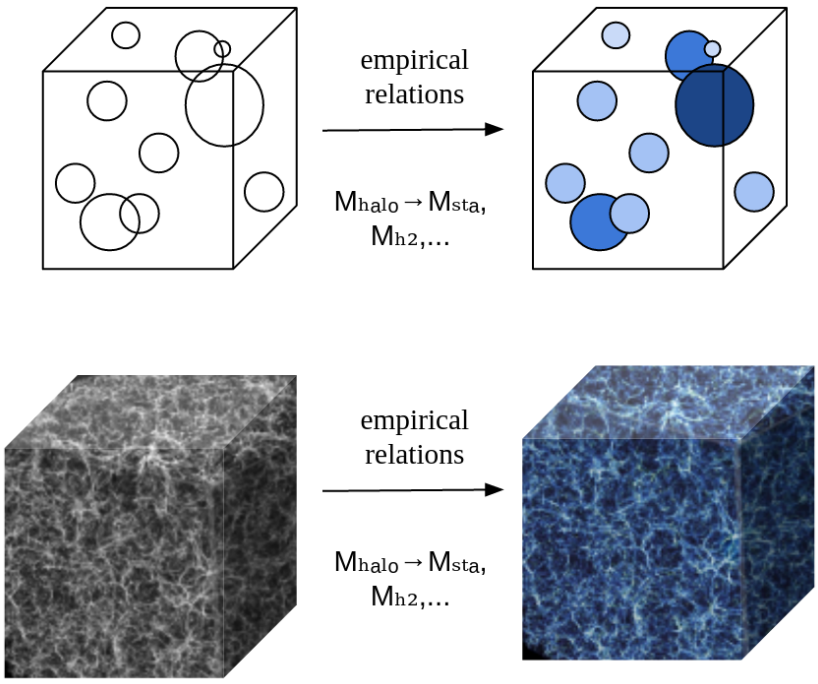

Figure 4.6: Schematic of two types of Dark Matter only simulations. Top: Dark Matter halos of a given mass (M_{halo}) in a box, using empirical relations, the halos are "painted on" to attribute them a star mass (M_{sta}), gas mass and other parameters. Bottom: Same as top, but with a fully collapsed cosmic web structure. Plot adapted from Karolina Garcia.

Hydrodynamical

Hydrodynamical simulations are also generated from particles in a cube, like dark-matter ones. The difference is that they contain several types of particles, for example, some particles might be stars and some black holes. The particles are left to evolve under gravity but because their type is known, hydrodynamics laws are also applied to their evolution. The Magneticum[2] simulations explored in

[1] https://wwwmpa.mpa-garching.mpg.de/galform/virgo/millennium/
[2] http://www.magneticum.org/

this work are an example of such simulations (Hirschmann et al., 2014; Dolag, 2015) (see App. B.1). Another example is the Illustris TNG[3] simulations (Vogelsberger et al., 2014).

data based and model based

Data-based simulation consists of taking existing observations of the component and rescaling it, debiasing it, or processing it with some prior knowledge of how it should look like to obtain simulations (see 4.2.1). Sometimes, various observations from different experiments are used, sometimes data from other emissions is substracted to clean the sky and obtain only the data of interest. Model-based simulations are solely based on the theory associated with the emission, often using its SED or source distribution (see 4.2.1). The two techniques were of course combined on multiple occasions, the SED being used to extrapolate the simulations to frequencies at which the experiments did not observe.

4.2.3 State-of-the-art

In 2019 when this project started the state-of-the-art of Microwave sky simulation was PySM (Thorne et al., 2017) for Galactic-foregrounds and the Websky (Stein et al., 2018) for the extra-galactic foregrounds. The Planck Sky Model (PSM) (Delabrouille et al., 2013) simulating both Galactic and extra-galactic foregrounds was previously the state-of-the-art used by the *Planck* team but was not publicly available in its latest version. This is why Jens Erler and I decided to work on developing a Python package to simulate mock maps of the microwave sky tailored for the Fred Young Submillimeter Telescope (FYST) predictions. At the time of this thesis, PySM3 (Zonca et al., 2021) has been released, it can simulate both Galactic and extra-galactic foregrounds up to very high resolution ($N_{\mathrm{side}} = 4096$). It is still based on *Planck* data and modelling for the Galactic foreground and is based on the latest version of the Websky simulation (Websky 0.4) for the extra-galactic foregrounds. The PSM became fully available online via a web interface[4] on the Planck Legacy Archive and offers a large variety of *Planck* data to base its mock maps on, but also several empirical models. It also goes to the current higher possible resolution with HEALPix $N_{\mathrm{side}} = 8192$.

4.2.4 Future of Microwave sky simulations

The Galactic-foreground field is pushing for more modelization of the components especially in 3D, because the SED is not necessarily the same along the Line-of-Sight (LoS) or varies spatially whether close or away from the Galactic plane. Better modelization is essential in order to do Astrophysics with the Galactic foregrounds. For example to probe the Galactic magnetic field with Synchrotron or dust. Another interest of better modelization is to be able to remove the foregrounds better. They are contaminants for CMB science (see Fig. 4.7) or for extra-galactic emissions, such as tSZ and kSZ that are greatly subdominant to the CMB in amplitude. For the extra-galactic components, the CMB intensity is fairly well modelized, its polarization components are the current science drivers. SZ science is fairly well modelized. The CIB component however is poorly constrained and the dominant contaminat on frequencies above 300 GHz. The SED varies spatially and along the line-of-sight, making the component really difficult to constrain and efficiently remove.

[3] https://www.tng-project.org/
[4] https://pla.esac.esa.int/

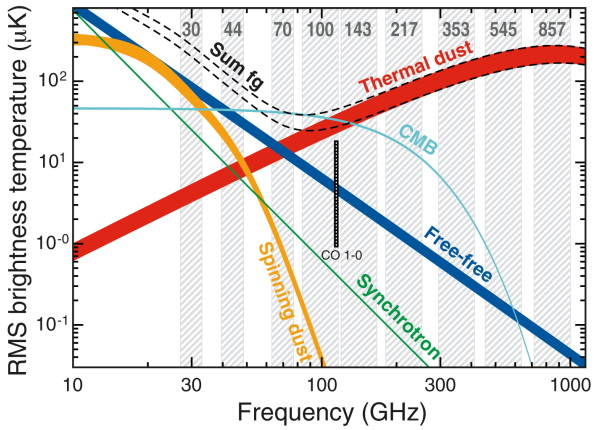

Figure 4.7: Illustration of the Galatic foreground contamination to the CMB. The brightness temperature of each component is displayed as a function of its frequency. Image extracted from (Planck Collaboration at al., 2020)

4.3 Overview of the Skymodel's base simulations from the literature

The Skymodel python package presented in this work uses existing full-sky intensity templates HEALPix maps of each of the microwave sky components provided by existing codes and simulations (Sehgal et al., 2010b; Thorne et al., 2017; Stein et al., 2018; Simons Observatory Collaboration, 2019; Stein et al., 2020). To later, adapt them to the characteristics of the Fred Young Submillimeter Telescope (FYST) and of the Simons Observatory Large Aperture Telescope (SO-LAT). Below, is given an extensive description of those based template maps.

4.3.1 Galactic emission

The Galactic emissions are generated using a Python package called the Python SkyModel (PySM). This was set up by Benjamin Thorne and his collaborators (Thorne et al., 2017), as an effort to offer an easy-to-use code to generate the different microwave emissions coming from the Galactic plane. PySM simulates those components both in intensity and polarization over the full sky based on WMAP and *Planck* data. The goal of PySM was to provide modelization of the galactic foreground components, which dominate several orders of magnitude above the CMB polarization signals, to better extract them. PySM also simulates the CMB and instrumental noise. To generate the galactic sky, PySM requires the desired frequencies, beam width, bandpass width, noise values, and the preferred units for the output as inputs. Phenomenological models and template maps are used and extrapolated to produce maps of the components at any frequency. An extensive description of them, for each Galactic component is given below. Polarization is left out, as it was not used in this thesis.

Thermal dust

The thermal dust, frequency spectrum, is simulated by PySM using a modified blackbody with different assumptions possible.

- 'd1' the nominal index: uses the 545 GHz map from Planck Collaboration et al. (2016). The

originally $N_{\rm side} = 2048$ map is downgraded to $N_{\rm side} = 512$. The frequency spectrum is given by

$$I_\nu(\vec{n}) = A_{\nu_0}(\vec{n}) \left(\frac{\nu}{\nu_0}\right)^{\beta_d(\vec{n})} B_\nu(T_d(\vec{n})), \tag{4.8}$$

where $\nu_0 = 545$ GHz, A_{ν_0} is the amplitude of the 545 GHz map for each pixel and direction $\vec{n}$. β_d is the dust emission spectral index map (see Fig. 4.8). $B_\nu(T_d)$ the Planck spectrum at a frequency ν and dust temperature of T_d, which map can be seen on Fig. 4.8. All the maps were extracted using the `Commander` software.

- 'd2' & 'd3' for spatially variable indexes: the spectral index is expected to vary spatially, this can be modelized by PySM on drawing the spectral index from a Gaussian distribution around $\beta_d = 1.6$ and with a dispersion of $\sigma = 0.2$ for 'd2' and $\sigma = 0.3$ for 'd3'. This approximation matched with *Planck* data (Planck Collaboration et al., 2015). The input β_d map can be seen in Fig. 4.8 under 'Thermal dust β_{alt}'.

- 'd4' two dust temperature: it is expected that not all dust grains will have the same temperature, in this model PySM allows parametrizing the thermal dust emission for two different temperature values

$$I_\nu(\vec{n}) = \sum_{i=1}^{2} I_a(\vec{n}) \left(\frac{\nu}{\nu_0}\right)^{\beta_i} \frac{B_\nu(T_i(\vec{n}))}{B_{\nu_0}(T_i(\vec{n}))}, \tag{4.9}$$

with

$$I_a(\vec{n}) = A_{\nu_1}(\vec{n}) \frac{\left(\frac{\nu_0}{\nu_1}\right)^{\beta_i} f_i q_i B_{\nu_0}(T_i(\vec{n}))}{\sum_{j=1}^{2} f_j q_j B_{\nu_1}(T_j(\vec{n}))} \tag{4.10}$$

where, A_{ν_1} is the thermal dust emission map at $\nu_1 = 3000$ GHz, ν_0 the frequency at 545 GHz and β_i the spectral index of the first dust temperature considered and constant over the full sky. f_i is the fraction of interstellar light absorbed by the dust grains and re-emitted into the infrared domain for the i temperature. q_i is the ratio between infrared and optical light for the i temperature. $B_{\nu_0}(T_i(\vec{n}))$ is the Planckian at the frequency ν_0, temperature T_i and direction $\vec{n}$.

Free-free

The free-free emission is implemented with only one emission model in the formed of an analytic power law (Draine, 2011)

$$I_\nu(\vec{n}) = A_{\nu_0}(\vec{n}) \left(\frac{\nu}{\nu_0}\right)^{\beta_s}, \tag{4.11}$$

Where $A_{\nu_0}(\vec{n})$ is the amplitude at the frequency $\nu_0 = 30$ GHz and direction $\vec{n}$. The spectral index is considered constant over the full sky at $\beta_s = -2.14$ in accordance with *WMAP* and *Planck* measurements (Bennett et al., 2013; Planck Collaboration et al., 2016). The free-free emission map at 30 GHz can be seen in Fig. 4.8.

Synchrotron

PySM simulates the Synchrotron intensity using the template maps of Haslam et al. (1981, 1982), coming from observations of the whole sky made at 408 MHz by the 100-m Effelsberg radio telescope but reprocessed by Remazeilles et al. (2015). The input map used by PySM can be seen in Fig. 4.8. The frequency spectrum taken by PySM to extrapolate this map to other frequencies is the following

$$I_\nu(\vec{n}) = A_{\nu_0}(\vec{n}) \left(\frac{\nu}{\nu_0}\right)^{\beta_s(\vec{n})}, \tag{4.12}$$

where I_ν is the intensity of the Synchrotron emission in the direction $\vec{n}$ at a frequency ν, A_{ν_0} the amplitude and ν_0 the cutout frequency. β_s is called the spectral index and PySM offers 3 different models for this slope parameter.

- 's1' the nominal index: uses, as input, the spectral index map ($\beta_s(\vec{n})$) from the 'Model 4' of Miville-Deschênes et al. (2008) display in Fig. 4.8 under the name 'Synchrotron β'.
- 's2' the spatially steepening index: this model the steepening of the Synchrotron spectral index away from the galactic plane. This is expected because processes energizing cosmic rays, like supernovae are thought to be more frequent in the galactic plane. Thus, the further the Synchrotron emission from the Galactic plane, the higher the chance it originates from older cosmic rays diffused by the galactic plane. Which are expected to be of lower energy, steepening the spectral index (Kogut, 2012; Ichiki, 2014). PySM models that by injecting in Eq. (4.12):

  $$\beta_s = \beta_{s,\mathrm{b=0}} + \delta_\beta \sin|\mathrm{b}|, \tag{4.13}$$

  where, $\beta_{s,\mathrm{b=0}}(\vec{n})$ is the spectral index at the Galactic latitude $b = 0$ and $\delta_\beta = -0.3$ in accordance with *WMAP* data (Kogut et al., 2007). Resulting in the spectral index map called 'Synchrotron β_{alt}' in Fig. 4.8.
- 's3' the curvature of index: In this model, the Synchrotron spectral index emission is supposed to flatten or steepen depending on the frequency. In this case, in Eq. (4.12):

  $$\beta_s(\vec{n}) = \beta_s(\vec{n}) + \mathrm{C} \ln\left(\frac{\nu}{\nu_\mathrm{c}}\right), \tag{4.14}$$

  where C= -0.052 at $\nu_\mathrm{c} = 23$ GHz extracted from Kogut (2012) by fitting *WMAP* data and where $\beta_s(\vec{n})$ is the same map as in 's1'.

Anomalous Microwave Emission

Anomalous Microwave Emission (AME), or spinning dust, is modelized by PySM as the sum of two spinning dust populations based on *Planck* data templates (Planck Collaboration et al., 2016) and represented by a power law. They both have spatially varying emission templates but one has a spatially varying peak frequency (see Spinning dust ν_p in Fig. 4.8) and the other a constant peak frequency. Both emission laws are computed using `SpDust2` (Ali-Haïmoud et al., 2009; Silsbee et al., 2010).

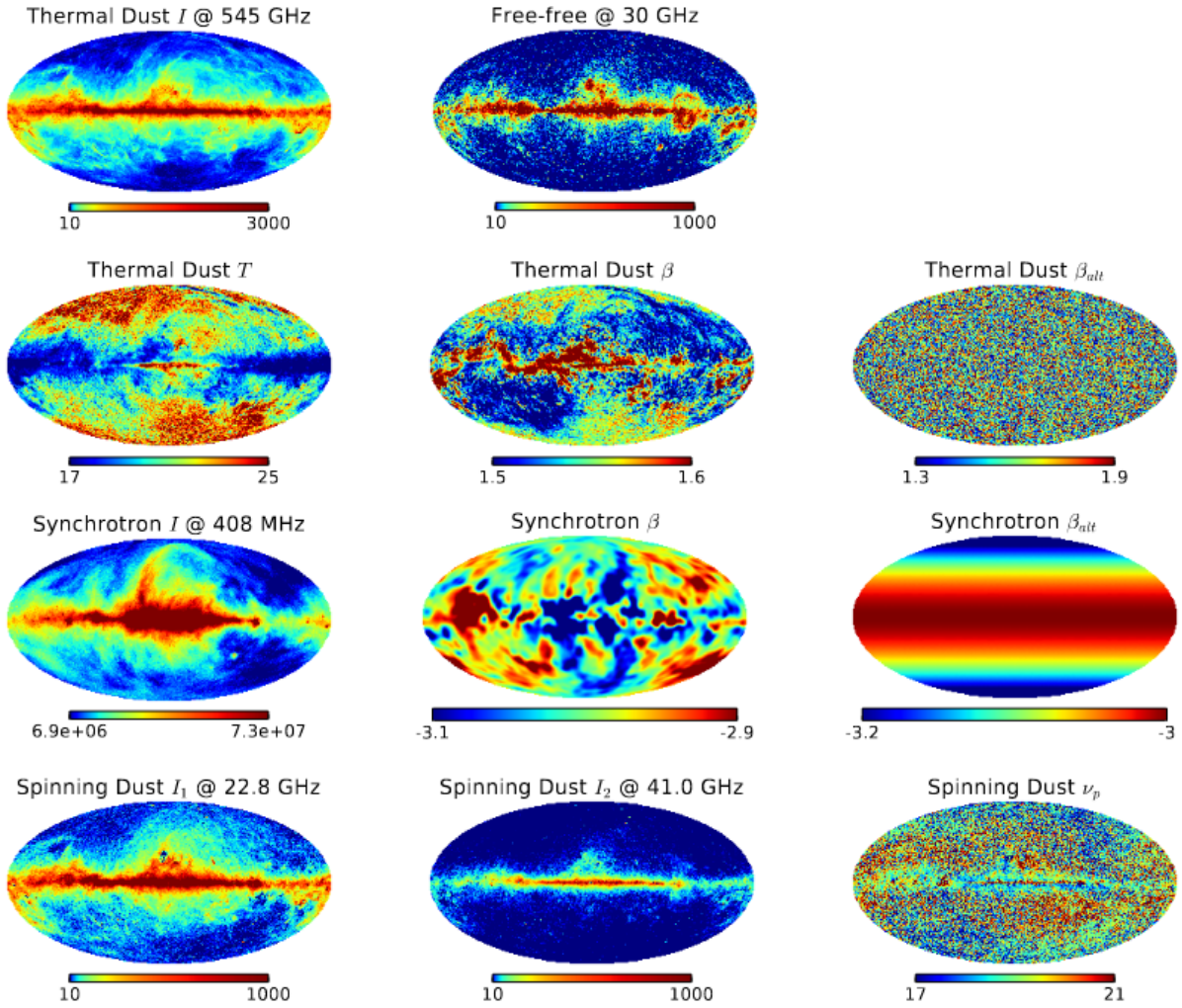

Figure 4.8: Adapted from Thorne et al. (2017), template HEALPix maps derived from *Planck* used by PySM to generate the galactic foreground maps at any frequency. They are in units of μK_{RJ} except for the dust temperature that is in K. The details of each map can be found in 4.3.1.

4.3.2 Extra-galactic emission

Emissions coming from outside the Milky Way were generated using HEALPix template maps from simulations. The Skymodel can select between several simulations to base itself on. Allowing it to make the most of the characteristics of each simulation and giving it versatility. The specificities of each simulation are described below. As for the galactic foreground, polarized emissions are left out.

Sehgal simulations

The Sehgal et al. (2010a) simulations are dark matter-only simulations of 1024^3 particles in a 1 h^{-1}Gpc box, with a resolution of $6.82 \times 10^{10} h^{-1} M_{\odot}$. A tree-particle-mesh (Bode et al., 2001) code is run on them to simulate the large-scale structure (see Sec. 4.2.2). A flat ΛCDM cosmology is assumed with the

following cosmological parameters $h = 0.71$, $\Omega_{\rm cdm} = 0.22$, $\Omega_b = 0.049$, $\Omega_\Lambda = 0.736$, $n_s = 0.96$, and $\sigma_8 = 0.80$, which is in good agreement with the WMAP5 cosmology Komatsu et al. (2009). The halos are painted on and the full-sky maps are assembled from eight identical boxes projected over the sky. They are high resolution $N_{\rm side} = 8192$ ($N_{\rm pix} = 0.43'$) maps in units of Jy/sr, available at frequencies ranging between 30 GHz and 350 GHz (see Tab. 4.1). The extra-galactic template maps provided are Cosmic Infrared Background (CIB), lensed and unlensed Cosmic Microwave Background (CMB), and thermal and kinematic Sunyaev-Zeldovich (tSZ & kSZ) effects and Radio Point Sources (RPS).

The characteristical features of this set of maps are:

- the high resolution ($N_{\rm pix} = 0.43'$) which is necessary to comply with the Nyquist sampling theorem for experiments of 1' resolution.
- radio and infrared galaxies are correlated with the galaxy clusters population.
- The box in which the simulation was executed is reproduced 8 times to fill up the full sky. A consequence of that is, that at the pole, the patch-up is visible in the maps.

The Sehgal et al. (2010a) simulations also provide template maps for the Galactic dust emission which was produced fitting COBE and *WMAP* data. This is not used in the Skymodel, instead the PySM Galactic dust map, based on the more recent *Planck* data is used.

Emission maps	**Frequency [GHz]**	**Units**
Unlensed CMB Lensed CMB tSZ kSZ RPS CIB Galactic dust	{30, 90, 148, 219, 277, 350}	Jy/sr

Table 4.1: Array containing the detailed available emission maps in the Sehgal et al. (2010a) set, their frequency and units. Each template map can be downloaded from `https://lambda.gsfc.nasa.gov/simulation/full_sky_sims_ov.html`.

SO simulations

The Simon Observatory (SO) simulations are a downgraded resolution and rescaled version of the Sehgal et al. (2010a) simulations. The resolution of the template maps are $N_{\rm side} = 4016$ ($N_{\rm pix} = 0.86'$). They are provided at the SO Large Aperture Telescope (SO-LAT) frequencies between 27 GHz and 353 GHz (see Tab. 4.2). The extension of the frequency range of the existing base (Sehgal et al., 2010a) CIB and RPS maps were performed using pixel-by-pixel interpolation of the maps flux as a function of frequency using a piecewise linear spline in log-log space. All maps are provided in frequency-independent units of $\mu{\rm K}_{\rm CMB}$ except for the tSZ map provided in Compton-y units and the lensing convergence map. Provided

On top of having the characteristics of the Sehgal et al. (2010a) templates maps they also have the following particularities:

- their tSZ maps have been rescaled by a factor of 0.75 to match the Planck Collaboration et al. (2014) Compton-y map power spectrum.
- their CIB maps were rescaled by a factor bringing down the power spectrum of the 148 GHz map in agreement with J Dunkley and et al. (2013).
- the agreement between the CIB maps and data such as *Planck* at other frequencies is not good because the SED used in the simulations not being compatible with the one measured by *Planck*.
- an unleased CMB map is not provided.
- a flux cut of 7 mJy at 148 GHz was applied to the RPS map to remove the brightest point sources.
- the Sehgal et al. (2010a) maps were downgraded using the HEALPix function `alter_alm()` to create the SO maps. This creates ringing artefacts in the output map because of the sharp cut in the $a_{\ell m}$. They are set to zero above a given scale and their value in harmonic space is used to produce a map in real space, generating rings. This can be easily understood by looking at the Fourier transform of a high-pass filter, the sharps cuts generate wavelets, that look like ringing on a 2D surface.

This set of template maps was prepared and used for Simons Observatory Collaboration (2019) predictions. templates maps of Galactic dust and low-frequency Galactic foregrounds (i.e. Synchrotron, Free-Free, AME) are also provided they are based on the *Planck* `Commander` analysis. The Skymodel rather uses the PySM template maps for the Galactic foregrounds because they are an updated version of the *Planck* analysis.

Emission maps	**Frequency [GHz]**	**Units**
CIB RPS Low-Frequency Galactic Foregrounds Galactic dust	{27, 30, 39, 44, 70, 93, 100, 143, 217, 225, 280, 353}	$\mu K_{\rm CMB}$
tSZ	$\emptyset$	Compton-y
Lensed CMB Lensing convergence kSZ	$\emptyset$	$\mu K_{\rm CMB}$

Table 4.2: Array containing the detailed available emission maps in the SO (Simons Observatory Collaboration, 2019) set, their frequency, and units. Each template map can be downloaded from `https://lambda.gsfc.nasa.gov/simulation/full_sky_sims_ov.html`.

Websky simulations

The Websky CMB mock templates (Stein et al., 2020) are generated through dark matter-only mass-peak patch (Stein et al., 2018) simulations containing 12288^3 particles in a box of 15.4 h^{-1}Gpc. Assuming a ΛCDM with the cosmological parameters $h = 0.68$, $\Omega_{\rm cdm} = 0.261$, $\Omega_b = 0.049$, $\Omega_\Lambda = 0.69$, $n_s = 0.965$, and $\sigma_8 = 0.81$. The mass-peak patch algorithm functions by,

1. generates a linear density field at z=0 and evolves it
2. some halos are selected using top-hats function filtering and density thresholds
3. overlapping halos are removed using exclusions techniques to keep only topologically disconnected mass-peaks
4. The remaining halos are evolved using 2nd order Lagrangian perturbation theory.

The CMB mocks have the following characteristics:

- a frequency de-correlation is implemented between tSZ and CIB.
- a Radio Point Sources map is missing from the simulations.

The resolutions, units, and frequencies at which the maps are available vary with each version of the CMB mock and will be detailed below. The template maps can be downloaded from https://mocks.cita.utoronto.ca/index.php/WebSky_Extragalactic_CMB_Mocks.

Websky 2018: the 2018 simulations came with a resolution of $N_{\text{side}} = 2048$ ($N_{\text{pix}} = 1.72'$). They are not available anymore for download. The detailed maps available are summarized in Tab. 4.3.

Emission maps	Frequency [GHz]	Units
CIB field CIB halo	{100, 143, 217, 353, 545, 857}	MJy/sr
tSZ	∅	Compton-y
kSZ field kSZ halo CMB unlensed CMB lensed	∅	μK_{CMB}

Table 4.3: Array containing the detailed available emission maps in the Wensky CMB mock from 2018 (Stein et al., 2018, 2020), their frequency, and units.

Websky 2019: are the simulations used by default in the Skymodel package. They are at a resolution of $N_{\text{side}} = 4096$ ($N_{\text{pix}} = 0.86'$). The details of the template maps can be found in Tab. 4.4. Those are the ones used as a baseline for the Skymodel. They are not available for download anymore, in this format.

Websky 2022: on the 18th of February 2022, new CIB maps were added (see Tab. 4.5) to the already existing Websky 2019 (see Tab. 4.4).

Websky 0.4 Websky 0.4 is Websky latest's version dated the 29th of October 2022. It is the version used by state-of-the-art software such as PySM3. It provides the most extensive frequency coverage for the template maps of all the Websky versions (see Tab. 4.6)

Emission maps	Frequency [GHz]	Units
CIB	{143, 217, 353, 545, 857}	MJy/sr
tSZ	∅	Compton-y
kSZ patchy kSZ κ $a_{\ell,m}$ lensed $a_{\ell,m}$ unlensed	∅	μK_{CMB}

Table 4.4: Array containing the detailed available emission maps in the Wensky CMB mock from 2019 (Stein et al., 2018, 2020), their frequency, and units.

Emission maps	Frequency [GHz]	Units
CIB	{93, 100, 143, 217, 353, 545, 857}	MJy/sr

Table 4.5: Array containing the detailed available emission maps in the Wensky CMB mock from 2022 (Stein et al., 2018, 2020), their frequency, and units.

4.4 This work take on the existing simulations

The first important point of this work was to compare the different existing mock maps of the sky to determine which one suited our projects the best. This original work was left out by the literature so far despite the need for reliable mock maps to make predictions. The investigation of shortcomings of the SO simulations resulting from contaminated maps were investigated and a correction was provided and implemented in the Skymodel.

4.4.1 SO reproduced

To circumvent the problem of SO having a sharp cut in ℓ-space (see Sec. 4.3.2). The Skymodel implemented a 'SO_reproduced' model which is derived from the Sehgal et al. (2010a) in the same way as the SO. Except that the maps are downgraded using the HEALPy routine UD_GRADE() this avoids the sharp cut in Fourier space and ringing artifact that could create unwanted correlations in the map. This is an entirely new set of sky simulations, original to the Skymodel correcting the defaults of the SO mocks.

4.4.2 A comparison of the Extra-galactic components

The Skymodel can base itself on the template maps of the three simulations, Sehgal et al. (2010a), SO, and Websky 2019 (here simply denoted Websky). This part investigates the differences between the three simulations in terms of power spectra for the microwave components and compares them with one another. For all components, the SO power spectra exhibit a sharp ℓ-scale cut on small scales due to it being the Sehgal et al. (2010a) maps simply downgraded to a lower N_{side} by setting the $a_{\ell m}$ coefficients to 0 above a given ℓ threshold (see Sec. 4.3.2).

Cosmic Microwave Background: The CMB power does not differ much between the three base simulations (see Fig. 4.9).

Emission maps	Frequency [GHz]	Units
CIB	{18.7, 21.6, 24.5, 27.3, 30, 35.9, 41.7, 44, 47.4, 63.9, 67.8, 70, 73.7, 79.6, 90.2, 100, 111, 129, 143, 153, 164, 189, 210, 217, 232, 256, 275, 294, 306, 314, 340, 353, 375, 409, 467, 525, 545, 584, 643, 729, 817, 857, 906, 994, 1080}	MJy/sr
RPS	{18.7, 21.6, 24.5, 27.3, 30, 35.9, 41.7, 44, 47.4, 63.9, 67.8, 70, 73.7, 79.6, 90.2, 100, 111, 129, 143, 153, 164, 189, 210, 217, 232, 256, 275, 294, 306, 314, 340, 353, 375, 409, 467, 525, 545, 584, 643, 729, 817, 857, 906}	MJy/sr
tSZ ($N_{\rm side}$ = 8192)	$\emptyset$	Compton-y
kSZ	$\emptyset$	$\mu K_{\rm CMB}$
κ		
$a^L_{\ell,m}$ $a^L_{\ell,m}$ s=0 $a^L_{\ell,m}$ s=1 I^L (s=1, $N_{\rm side}$ = 512) I^L (s=1, $N_{\rm side}$ = 4096)	$\emptyset$	$\mu K_{\rm CMB}$
$a^{UL}_{\ell,m}$ s=0 $a^{UL}_{\ell,m}$ s=1 I^{UL} (s=1, $N_{\rm side}$ = 512) I^{UL} (s=1, $N_{\rm side}$ = 4096)	$\emptyset$	$\mu K_{\rm CMB}$

Table 4.6: Array containing the detailed available templates maps in the Wensky 0.4 CMB mocks (Stein et al., 2018, 2020), their frequency, and units.

Cosmic Infrared Background: The SO and Sehgal CIB power spectrums are identical. This is expected as the SO power is simply the Sehgal power rescaled by a factor 0.75 to better match the *Planck* observations (see Sec. 4.3.2). The Websky simulations exhibit a lower CIB power spectrum on large scales (up to $\ell < 300$) and a higher power on smaller scales ($\ell > 300$).

thermal Sunyaev-Zeldovich: The tSZ power spectrum of SO and Sehgal is, like in the CIB case, the same. SO being a scaled-down version of the Sehgal one by a factor 0.75 to better match the *Planck* measurements. The Websky power spectrum is in agreement with Sehgal on large scales and with SO on small scales.

kinematic Sunyaev-Zeldovich: For the kSZ power spectrum, Sehgal and SO are in complete agreement. The Websky power spectrum is more than three orders of magnitude above SO & Sehgal on large scales ($\ell < 300$) and a lit lower but in generally good agreement on smaller scales ($\ell > 300$). According to the Websky team, this extra power on large scales is due to a Doppler effect from the velocity fluctuations.

Radio Point Sources: The RPS are not available for the Websky 2019 simulations. The SO RPS power spectrum is around 2 orders of magnitude below the Sehgal power.

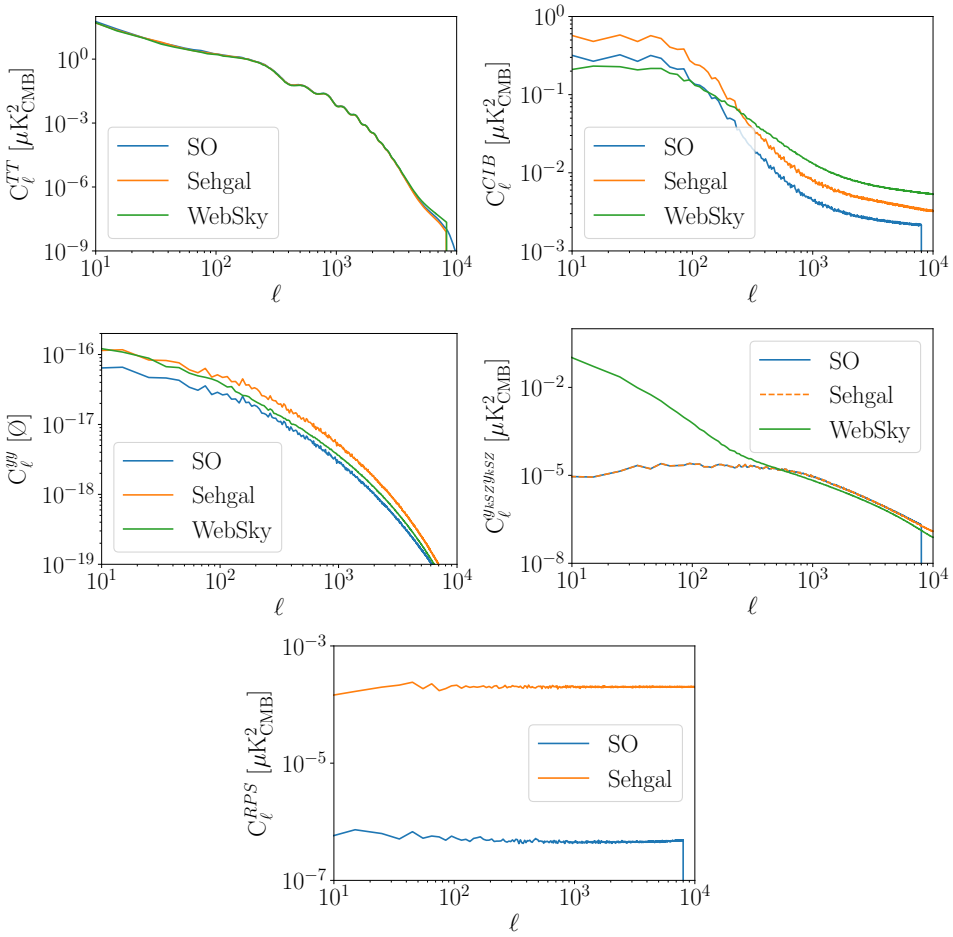

Figure 4.9: Comparison of the power spectra of the different extra-galactic template maps provided by the three simulations available in the Skymodel. Top row left: for the Cosmic Microwave Background. Top row right: for the Cosmic Infrared Background. Middle row left: for the thermal Sunyaev-Zeldovich. Middle row right: for the kinematic Sunyaev-Zeldovich. Lower row: for the Radio Point Sources.

4.5 This work processing of the existing templates

The Skymodel package does not generate any simulations of the microwave sky emissions but uses existing HEALPix template maps. Those maps are provided by the existing Python packages, like PySM (Thorne et al., 2017) for the Galactic-foreground emissions. Or they were created from dark matter-only simulations like for the extra-galactic emission templates Websky (Stein et al., 2020). The templates are then processed to create high-resolution maps at any frequency, sensitivity, and beam size for upcoming ground-based experiments. The way in which they are processed is explained in the following sections.

4.5.1 Galactic foregrounds

The galactic foreground template maps are generated using PySM (Thorne et al., 2017). Which is itself based on processed *Planck* data (Planck Collaboration et al., 2016). The Planck data are originally low resolution $N_{\text{side}} = 256$ ($N_{\text{pix}} = 13.74'$). PySM upgrades them to $N_{\text{side}} = 512$ ($N_{\text{pix}} = 6.9'$) using the HEALPy function `ud_grade()`. This functions to upgrade the maps to a higher pixel resolution by subdividing existing pixels into smaller ones with the exact same value as the initial pixel. This creates pixelization artefacts in the PySM maps.

In particular, upcoming ground-based telescopes such as the Simons Observatory Large Aperture Telescope (SO-LAT) and the Fred Young Submillimeter Telescope (FYST) will have a resolution of around 1' per pixel. This requires the Skymodel to upgrade the PySM maps to at least $N_{\text{side}} = 4096$ ($N_{\text{pix}} = 0.86'$) and would result in highly pixelated maps with a large-scale pixel-to-pixel correlation of 1.

To address this issue it was considered to use the HEALPix function `alter_alm()`, to directly modify and extrapolate the map in harmonics space, by first converting them into their $a_{\ell,m}$ coefficients. However, because of the sharp cuts in pixel space, the transformation of the map into harmonic space generates ring-like structures (see Fig. 4.10). Those undesired artifacts could create spurious correlations in the maps. This was avoided, by upgrading the maps to the desired resolution using `ud_grade()` and smoothing them, using a Gaussian beam close to the original PySM pixel size, of 10' (see Fig. 4.10). This does not lower the information contents of the maps. As they did not have small-scale variations below the original PySM pixel size. However, the resulting maps might have a high pixel resolution they do not contain the physical scale information in those pixels. They are limited to 10' scales.

The galactic foreground maps generated by the Skymodel are the PySM-generated maps, upgraded to the desired N_{side} and smoothed with a 10' Gaussian beam.

4.5.2 Extra-galactic foregrounds

The Skymodel uses by default the Websky 2019 template maps (see Tab. 4.4) to generate the extra-galactic foreground emissions. But the Simons Observatory (SO) or Sehgal et al. (2010a) templates maps can also be selected.

Cosmic Infrared Background

This section details the post-processing of the Cosmic Infrared Background (CIB) maps for each of the available base templates in the Skymodel.

Websky: The CIB maps are provided by the Websky simulations at some of the *Planck* HFI frequencies. In order for the Skymodel to produce CIB maps based on those templates, at any frequency ($\nu \in [27, 860]$ GHz), we need to interpolate and extrapolate. This is done by fitting the following modified blackbody spectrum to each pixel of the HEALPix template maps for all available frequencies (see Fig. 4.11),

$$I_\nu = A_{\text{CIB},\nu_0} \left(\frac{\nu}{\nu_0}\right)^{\beta+3} \frac{e^{\frac{h\nu_0}{k_B T_{\text{dust}}}} - 1}{e^{\frac{h\nu}{k_B T_{\text{dust}}}} - 1}, \tag{4.15}$$

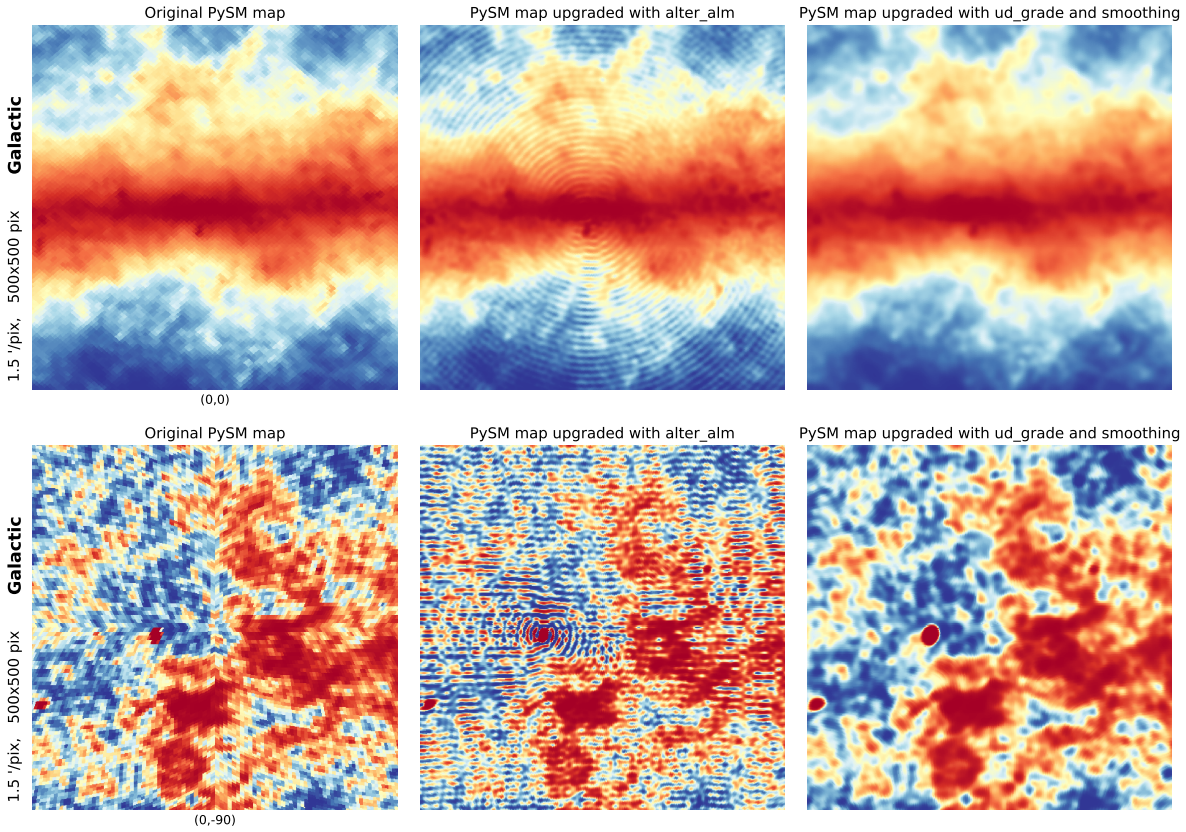

Figure 4.10: All panels are $12.5^\circ \times 12.5^\circ$ gnomic projection of the sum of all Galactic foreground emissions generated by PySM at a frequency of 353 GHz. They are plotted in histogram-equalized colours. Upper row: cutout centred on the Galactic centre. Lower row: centred on the Galactic south pole. Left panel: unmodified PySM maps at a pixel resolution of 6.9' ($N_{\mathrm{side}} = 512$) featuring pixelization artefacts. Middle panel: PySM map upgraded to a pixel resolution of 1.72' ($N_{\mathrm{side}} = 2048$) using `alter_alm()`, the ringing artefacts are clearly visible. Right panel: PySM maps upgraded to 1.72' ($N_{\mathrm{side}} = 2048$) using `ud_grade()` and smoothed using a 10' Gaussian kernel. Which is the upgrade technique used in the Skymodel. Courtesy of Jens Erler.

where $\mathrm{A_{CIB}}$ is the amplitude of the modified blackbody for each pixel. ν is the desired frequency. β is the spectral index. $h = 6.62 \times 10^{-34}$ Js is the Planck constant, $k_{\mathrm{B}} = 1.38 \times 10^{-23}$ JK^{-1} is the Boltzmann constant. T_{dust} is the temperature of the dust. $\mathrm{A_{CIB}}$, β, and T_{dust} are the parameters extracted by the fitting procedure for each pixel. The result of the fit is three HEALPix full-sky maps. One for $\mathrm{A_{CIB}}$ containing in each pixel the amplitude parameter. One β map containing the slope parameter for each pixel and one T_{dust} containing the temperature of the CIB for each pixel. Those three maps ensure that each pixel will follow a modified blackbody distribution. The fit is done using the Python Scipy function `curve_fit()`. This function uses the Trust Region Reflective algorithm, particularly suitable for large sparse problems with bounds and a generally robust fitting method.

The three parameters maps $\mathrm{A_{CIB}}$, β and T_{CIB} are injected back into Eq. (4.15) to produce new CIB maps at any frequency $\nu \in [27, 860]$ GHz, ensuring that every pixel of those new maps will follow the original Websky map structure and the modified blackbody described in Eq. (4.15). This is done via the Skymodel function `simulate_cib()`.

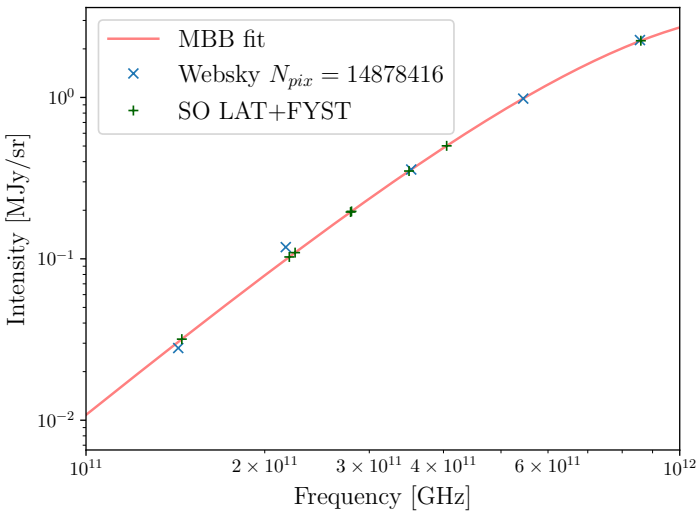

Figure 4.11: Fit of one randomly selected pixel ($N_{pix} = 14878416$) for all the available Websky 2019 CIB template maps (see Tab. 4.4). In red are the fitted Modified Black Body (MBB) (see Eq. 4.15), and the blue crosses are the Websky intensity values for each given CIB template map. The dark green '+' is the extracted intensity value for this particular pixel at the SO LAT and FYST frequencies (see Tab. 4.2 2.3). In this example, the fit extracted a value of $T_{dust} = 16.94$ and $\beta = 1.09$

SO & Sehgal et al. (2010a): The exact same process is done for the SO and Sehgal et al. (2010a) CIB templates. Only the pivot frequency is different for the SO maps, at $\nu_0 = 353$ GHz.

Radio Point Sources

Template maps of Radio Point Sources (RPS) are implemented in the Skymodel only for SO and Sehgal et al. (2010a) as no maps were available from Websky 2019.

The Sehgal et al. (2010a): RPS template maps are fitted by the following emission law

$$I_\nu = I_{\nu_0} \left(\frac{\nu}{\nu_0}\right)^{\beta} \exp\left(-\frac{\nu}{\nu_{\text{cut}}}\right), \tag{4.16}$$

where I_{ν_0} is the intensity in each pixel at the frequency $\nu_0 = 30$ GHz, β the spectral index, and ν_{cut} is the cut frequency at which the spectrum starts to show curvature.

SO: template maps contain ringing artifacts because the Sehgal et al. (2010a) resolution downgrade was done through the HEALPix function `alter_alm()` (see Sec. 4.3.2 and 4.5.1). This messes up their frequency spectrum. To avoid wrong modelling, the SO template maps are only linearly pixel-by-pixel interpolated for frequencies between $\nu \in [27, 353]$ GHz. Above 353 GHz a null map is returned by the Skymodel.

Cosmic Microwave Background

The Cosmic Microwave Background (CMB) template maps are processed differently by the Skymodel depending on the simulation they are based on. The Skymodel also offers the possibility to generate a random CMB realization from an input power spectrum.

Websky: provide two files (see Tab. 4.4), one containing the unlensed CMB $a_{\ell m}$ and another containing the lensed CMB $a_{\ell m}$. These are obtained through the unlensed $a_{\ell m}$ using the CMB lensing convergence (κ) map from $0 < z < 1100$ in the Born approximation. From those $a_{\ell m}$ two maps are created using the HEALPy function `alm2map()`, one CMB lensed and one CMB unlensed map.

SO: offers a lensed CMB map (see Tab. 4.2). Which was produced by Ferraro and Hill (2018), by applying LensPix (Lewis, 2005) to an unlensed CMB realization generated by CMB at $N_{\text{side}} = 4096$ and only including scales up to $\ell = 10000$. To lensed the CMB, LensPix used the convergence map (κ_{CMB}) provided by Sehgal et al. (2010a) (see Tab. 4.1).

Sehgal et al. (2010a): provide frequency-dependent maps of the lensed or unlensed CMB in intensity units (see Tab. 4.1). The 30 GHz map is multiplied by the inverse of Eq. (1.51) to obtain a frequency-independent CMB map in temperature units. To obtain a CMB map at a frequency that was not provided originally this frequency-independent CMB map can be multiplied by Eq. (1.51) evaluated at the desired frequency ν.

CAMB: The Skymodel offers the possibility to use a text file containing the spherical harmonics for the CMB generated by CAMB. By plugging in the C_ℓ coefficients, it will generate one realization of a CMB map. This can be used for null tests where CMB needs to be randomized or to test the effects of different cosmological parameters.

thermal Sunyaev-Zeldovich

The processing of the tSZ maps by the Skymodel is relatively straightforward. When the simulation selected provides a Compton-y map (Websky & SO), extrapolating it to any desired frequency is achieved by multiplying the y-map by the frequency spectrum of the tSZ and I_0 as shown in Eq. (1.52). When the tSZ is provided at various temperatures (Sehgal et al. (2010a)), the first step is to transform it into a Compton-y map by dividing it by I_0 and the spectral shape of the tSZ. Once the Compton-y map is produced the above-mentioned procedure can be applied to obtain a tSZ map at any frequency. Note that the Skymodel does not take into consideration relativistic corrections of the tSZ effect.

kinematic Sunyaev-Zeldovich effect

Skymodel processes the kSZ map in an analogous way to the tSZ map. When the map is provided in temperature units (Websky & SO), the Skymodel multiplies it by Eq. (1.51) at the desired frequency ν to obtain kSZ maps at this new frequency. When only frequency-dependent maps of the kSZ are provided by the simulation (Sehgal et al. (2010a)), the Skymodel first converts it to a temperature map, by diving it by Eq. (1.51) and then applies the above-mentioned procedure to produce kSZ maps at any desired frequencies.

4.5.3 Unit conversion

Once the based simulations and the Astrophysical emissions have been selected to produce a mock sky, the Skymodel ensures units are dealt with appropriately.

It ensures the units of every component of the mock sky are consistent with each other, using Eq. 1.50, it converts them appropriately either to MJyr/sr or to K_{CMB} units (See Sec. 2.1.3). Skymodel can also return the mock sky in Rayleigh-Jeans units, which approximates the Planck law in the low-frequency regime. It can be formalized as such

$$\Delta T_{RJ} = \frac{c^2}{2k_B \nu^2} \Delta I_\nu, \tag{4.17}$$

where ΔT_{RJ} is the Temperature variation in Rayleigh-Jeans units, c the speed of light, k_B Boltzmann constant, ν the frequency and ΔI_ν the intensity variation for the given frequency.

The Rayleigh-Jeans temperature can also be linked directly to the K_{CMB} temperatures units, by injecting Eq. (1.50) in Eq. (4.17) which gives this relation

$$\Delta T_{RJ} = \frac{x_\nu^2 e^{x_\nu}}{(e^{x_\nu} - 1)^2} \Delta T, \tag{4.18}$$

where $x_\nu = h\nu / k_B T_{CMB}$, with h the Planck constant and T_{CMB} the CMB temperature.

4.6 Study limitations and noises

When observing the sky with an instrument, one does not only measure Astrophysical signals coming from the Milky Way or from extra-galactic sources. The measurements also contain contaminants such as noise and beam relative to the instrument. But also limitations due to the choice of observations. Ground-based observation will be contaminated with atmospheric noise and any observation is limited by its survey strategy. All of those components have to be modeled and included in sky simulations to simulate future measurements.

4.6.1 Instrumental noise and telescope beam

In the Skymodel the instrumental white noise is computed for each pixel of the map using the normal distribution,

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{4.19}$$

where, $p(x)$ is the probability distribution of the variable x, μ the mean of the distribution and σ its standard deviation. The mean in this case is $\mu = 0$ and the standard deviation is given by the sensitivity per pixel size. Telescope sensitivities are often given in μK-arcmin, but the pixel size of your map might not be 1'. In this case, the sensitivity needs to be rescaled. Indeed, for a noise of X μK-arcmin, if the pixel size is 1/2 arcmin then the noise in this pixel should be X/2 μK-arcmin. The power spectrum of white noise is constant over all scales at the value of the telescope sensitivity for a given frequency.

The beam of the telescope is represented by a Gaussian kernel (7.3) and it is implemented using the HEALPy smoothing function (see Sec. 2.4.1).

4.6.2 Atmospheric noise

Ground-based telescopes, such as the Simons Observatory Large Aperture Telescope (SO LAT) and the Fred Young Submillimeter Telescope (FYST) are subject to contamination from the Atmosphere fluctuations and absorption (see Sec. 2.3 & Fig. 2.10). This additional source of noise coming from the Atmosphere is implemented in the Skymodel only for SO and FYST frequencies (see Tab. 4.2 & 2.3). The sensitivity values are based on Ade et al. (2019) for SO and for FYST on Choi and et al. (2020) which is just a version of the noise curves of Ade et al. (2019) adapted for FYST. Noise curves include instrumental white noise (see Fig. 4.12) and are parametrized by,

$$N_\ell = N_{\rm red}\left(\frac{\ell}{\ell_{\rm knee}}\right)^{\alpha_{\rm knee}} + N_{\rm white}, \tag{4.20}$$

where, N_ℓ is the total atmospheric and instrumental noise spectrum, $N_{\rm red}$ the atmospheric red noise values, $\ell_{\rm knee} = 1000$, $\alpha_{\rm knee} = -3.5$ and $N_{\rm white}$ the white noise values. The noise power spectrum is used to generate a random map realization in the Skymodel using the HEALPy routine SYNFAST(). This means when running the Skymodel multiple times, or for multiple frequencies, each map will contain a different realization of the noise, totally uncorrelated from the previous one.

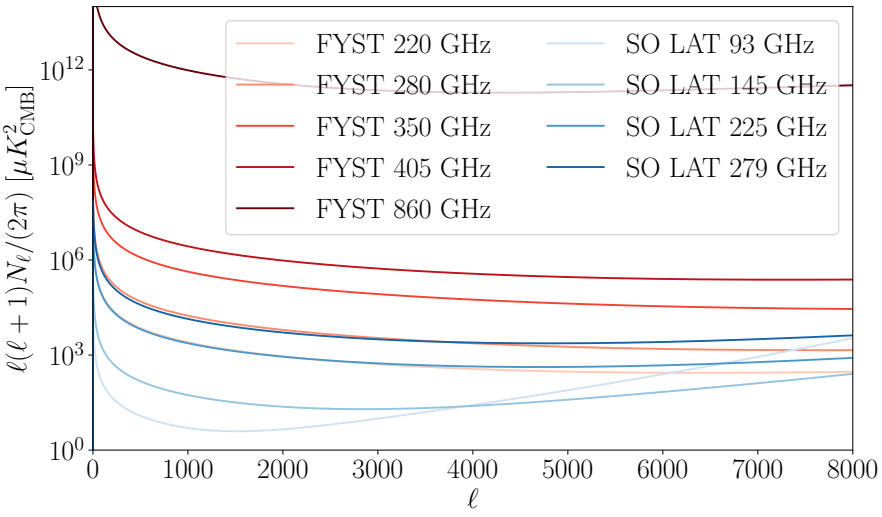

Figure 4.12: Noises curves at each frequency for SO LAT and FYST based on Ade et al. (2019); Choi and et al. (2020). The noise powers are beam-corrected.

4.6.3 survey masks

When commissioning an instrument to observe the sky, the question of surveying arises. Should the full sky the surveyed? is that necessary? is that optimal for the scientific goals? or could one simply observe a patch of the sky? which could be less expensive and time-consuming. Does it make sense to observe

regions that are highly contaminated by signals that are not our target signal? for example, if the target signal is extra-galactic does it make sense to observe the galactic plane?

When making predictions using simulations, one can reproduce the survey strategies using masks. Masks are arrays containing only 1 and 0 following the survey pattern so as to 'null' or 'mask' the simulation part that is not observed by the instrument and unaffected by the parts that are observed. In practice, the simulation maps are simply multiplied by the mask. The fraction of the sky covered sky a survey is defined by $f_{\rm sky}$,

$$f_{\rm sky} = \frac{\mathrm{S}_{\rm survey}}{\mathrm{S}_{\rm sky}}, \tag{4.21}$$

where, $\mathrm{S}_{\rm survey}$ is the total surface observe by the survey in deg^2 and $\mathrm{S}_{\rm sky}$ the total surface of the sky in deg^2 (see Sec. 2.4.1 for the computation of the sky surface). The Skymodel package offers several masks to reproduce the most common CMB experiment survey strategies.

advanced Atacama Cosmology Telescope

The advanced Atacama Cosmology Telescope (advACT) survey (see Fig. 4.13) covers an area of ~ 15 000 deg^2 ($f_{\rm sky} \approx 0.4$). The ACT telescope consists of a 6 m diameter primary mirror and is located at 5190 m in Cerro Toco in the Atacama desert in Chile. It has two science goals,

1. Improve the measurements of the early universe cosmological parameters.
2. Observe distant and large galaxy clusters and their environment.

advACT is the third generation of cameras mounted on the ACT telescope, it extended the frequency coverage of ACT and survey surface. It has been observing the sky since 2016.

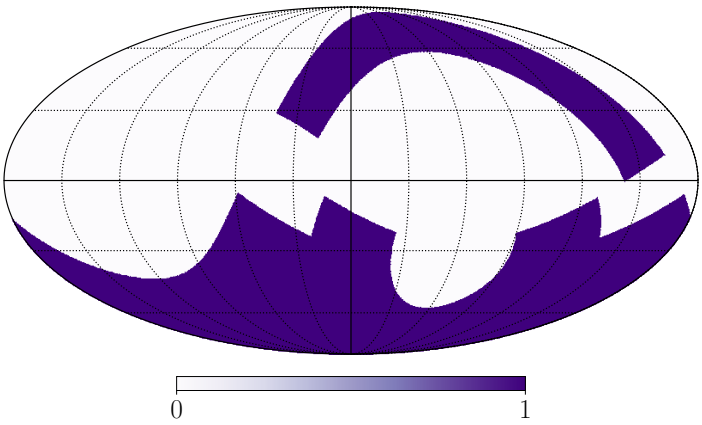

Figure 4.13: The advanced Atacama Cosmology Telescope (advACT) survey mask.

South Pole Telescope Sunyaev-Zeldovich

The South Pole Telescope Sunyaev-Zeldovich (SPT-SZ) survey (see Fig. 4.14) covers an area of 2 500 deg^2 ($f_{\rm sky} \approx 0.06$). The SPT-SZ survey was completed by the South Pole Telescope in 2011. Its goal

was to detect galaxy clusters. The SPT is a 10 m diameter telescope located at the National Science Foundation Amundsen-Scott South Pole station in Antarctica. It observes at 95 GHz, 150 GHz, and 220 GHz frequencies.

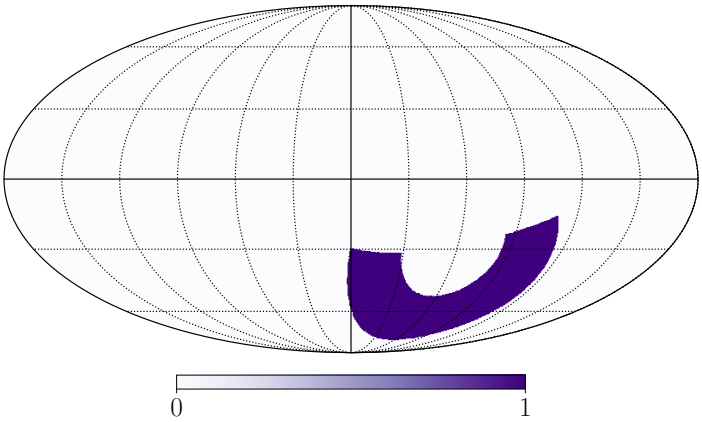

Figure 4.14: The South Pole Telescope Sunyaev-Zeldovich (SPT-SZ) survey mask.

Planck masks

The Skymodel also offers two survey masks derived from *Planck* observations (see Sec. 2.1.3). The first one, 'Dust' (see Fig. 4.15), was derived by Erler et al. (2018) by masking the 40% brightest pixels of the *Planck* galactic emission mask. Meaning this mask, masks 40% of the brightest emission coming from the galactic plane and of the sky ($f_{\mathrm{sky}} = 0.6$). The second mask (see Fig. 4.16) is derived from the first with the additional masking of a part of the sky that was not observed by the National Radio Astronomy Observatory Very Large Array Sky Survey (NRAO VLA Sky Survey → NVSS). Resulting in $f_{\mathrm{sky}} = 0.45$. The NVSS survey was a 1.4 GHz continuum survey completed in 1998 by the NRAO using the VLA. It is located in New Mexico in the United States and is composed of 28, 25 m diameter radio-telescopes.

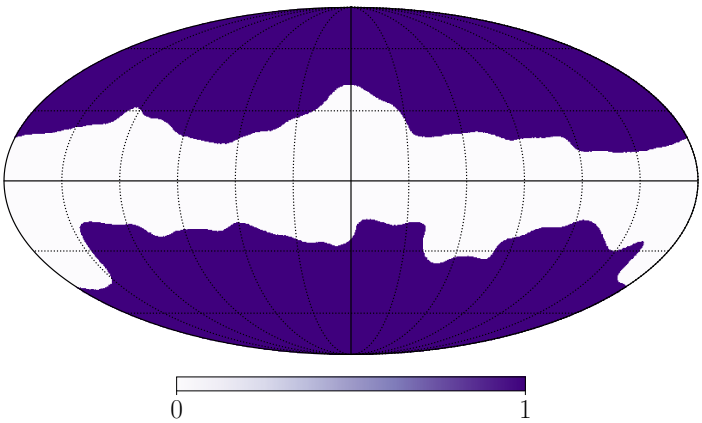

Figure 4.15: 'Dust' mask derived from *Planck* observation.

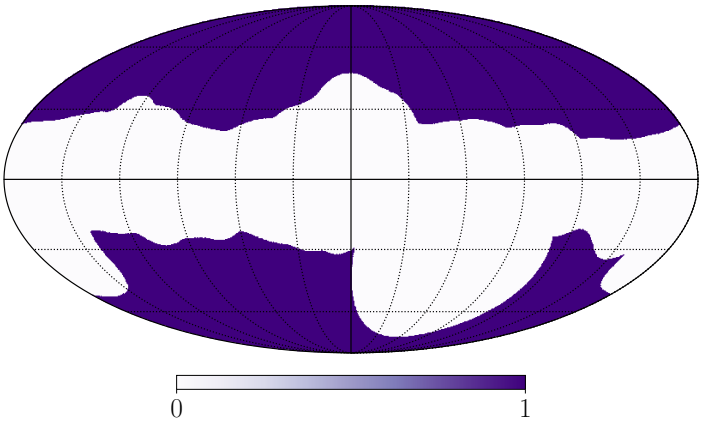

Figure 4.16: 'NVSS' mask derived from *Planck* observation.

4.7 Creation of a python package

The Skymodel was created to be used as a downloadable Python package that could be used like a black box to realistically simulate a high-resolution Microwave sky with desired components. This was achieved through clean coding and documentation, easy access, and team development on GitHub. Codes and packages are essential for rapid and reliable work in Astrophysics, but the discipline originally based on pen-and-paper calculation or observations, is struggling to make the shift to accept codes as serious scientific contributions by themselves. Journal publication is per se not a thing yet without a combination with an application of the code but options are emerging and require scientists to adapt to their specific requirements.

4.7.1 Good coding practices

When creating a Python package for the Astrophysics community, the code has to be clear and clean. Including comments in the code is also crucial to ease the reading by collaborators and future users (see Lis. 4.1). Everything has to be in the form of a function and apart from the usual astrophysical constants, nothing should be hardcoded. To allow full versatility of the code and to avoid mistakes that could occur from a small hidden variable that was hard coded to its most common value but different in a very particular case. When one of the variables, has a fixed value in most cases, a good practice coding solution is instead to attribute it a default value in the function header. It should also be advised to not hesitate to code exceptions to be printed while executing the code, to remind the user of special default parameters. An example of that is the Skymodel function to simulate Radio Point Sources (RPS), the header of the function contains several variables with an attributed default value. The function selects by default the Websky templates simulations. Note that it is not hard coded in the body of the function to avoid selecting the Websky in all cases, but only preselected as a default input parameter if the user does not specify anything. The Websky simulations do not provide RPS templates. This function will print a message to inform the user that no RPS will be included (see Lis. 4.1). The path leading to the folder was the required input files are stored should also not be hard coded in every function, but instead be

defined as a global variable at the beginning of the code. This will allow users to change it to their own local computer path without having to dive through every function and reduce the possibility of mistakes.

```
#Several variables have a default value attributed in the header of the function.
def simulate_radio_ps(freq, nside_out = 4096, lmax = None, beam_FWHM = None,
    template = 'WebSky', unit = 'cmb'):
...
#Raises user's feedback message that Websky does not have RPS.
    if template == 'WebSky':
        print('Warning: No radio PS template provided by the WebSky simulations')
...
```

Listing 4.1: Snipet of the Skymodel function to simulate RPS.

4.7.2 Documentation

Extended code documentation is not limited to comments and is quickly achieved through Read the Docs[5]. It is an automated website, that reads the 'Doc Strings' of every function of your code and automatically formatates them into an interactive webpage. The only condition for it to work is to write each of your functions with the appropriate Doc Strings. Doc Strings, are found directly under the header of the function and should follow strict formatting (see Lis. 4.2 and webpage[6]). They should always start with an indent, followed by ''', then a short description of what the function does. Then, every variable of the function has to be listed under 'parameters', with the exact same name followed by the type of the variable. Directly under it, with an indent, the variable has to be explained. Lastly, the outputs of the function have to be listed under 'returns', with their name first, directly followed by their type. Again, directly below this, starting with an indent, the return parameters have to be explained.

```
def sigmoid_filter(l_0, d_l, lmax):
    '''Computes a sigmoid-shaped filer high-pass window in spherical harmonic
    space.

    Parameters
    ----------
    l_0: float
        Determines the spatial scale above which the sky signal will be filtered.
    d_l: float
        Determines the slope of the widow.
    lmax: float
        Maximum value of ell to be used

    Returns
    -------
    window: float array
        Computed window in spherical harmonic space
    '''

    ell = np.arange(lmax)
    window = 1/(1+np.exp((-ell+l_0)/d_l))

```

[5] https://readthedocs.org/

[6] https://skymodel.readthedocs.io/en/latest/generated/skymodel.sigmoid_filter.html

```
22      return(window)
```

Listing 4.2: Skymodel function returning a Sigmoid filter.

Doc Strings not only help generate automatic documentation of the code, but they also provide direct clarification of the function inside the code itself, minimizing the need for unformatted messy comments. One constraint of Read the Doc is that it requires a GitHub account and the code to be uploaded on GitHub.

4.7.3 Github

GitHub[7] is the website to develop code with collaborators. This can be done privately or publicly. The code host account owns the "master branch", collaborators can create a 'fork' meaning a copy of the master version, and develop it independently on their GitHub page without influencing the master branch. They can then submit a 'pull request' to the master branch, GitHub will then provide a line-by-line comparison of the two code versions, highlighting the changes, and will allow the fork to merge and update the master branch only if it did not find any conflicts. GitHub also keeps track of every modification, they are accessible and every previous code version can be redeemed at any time. The 'Issues' tab functions like a discussion channel allowing contributors to express their problems with the code or desired modification. It also allows the code to be easily downloaded and installed as a Python package through the 'git clone url' command. The Skymodel is fully accessible on GitHub[8].

GitHub is not only a practical platform for developing code, but it has also become a requirement for publication in some of the upcoming Journals dedicated to code. The Journal Of Open Source Software (JOSS) chosen by the PySM3 team to publish their updated version of the package, requires the paper to be included in the GitHub folder and submitted to them through the GitHub link. The editorial board also takes into account the history of the GitHub page of the package. In particular, its longevity, the length of the code, which must be above 300 lines, but also its development. JOSS encourages and desires a long and extensive history of 'commits', meaning code modification on the GitHub page, showing that the code was transparently developed there and that every step of its development is still available.

4.7.4 Journals for code and packages

Codes are published in Astrophysics Journals only through their application to a given science question. In this case, the code itself, functionality, usability, and documentation are not reviewed at all. This raises two problems, First, the lack of recognition for the creation of packages that simplify, accelerate, and render less susceptible to mistakes in research work. Second, it removes the interest and reward of proper coding and documentation of codes, making them more prone to mistakes and difficult to re-use or build on, thus de-optimizing hours of work. Efforts have been made to make a shift, and several possibilities exist.

The HEALPix package (Gorski et al., 2005), which is of utmost importance for CMB science analysis was published on ArXiv as a non-peer-reviewed package. Websites like the Astrophysics Source Code Library (ASCL[9]), offer to reference a code package with a number. It also tries to provide a go-to library to find and reference codes for Astronomy. GitHub is another example of such a library (see Sec. 4.7.3)

[7] https://github.com/
[8] https://github.com/MaudeCharmetant/skymodel
[9] http://ascl.net/

but is not specific to Astronomy. Another option is to assign a Digital object identifier (DOI) number to the code, this DOI can then be used to cite the code within academic papers and collect citations. As discussed in the previous section (Sec. 4.7.3), PySM3 (Zonca et al., 2021) opted for JOSS, this has the advantage of being a peer-reviewed journal, even though it is multidisciplinary and not specific to Astronomy. Meaning, the code quality and documentation are reviewed but the technical consequences of the code for Astrophysics can not be accessed. Even though JOSS does require a clear justification of what the code brings to the community and whether it has already been used in many publications.

4.8 Conclusion and outlook

This part reviews all that was done in regard to the modelling of the microwave sky and what could be added in regard to the current state-of-the-art.

4.8.1 Conclusion

The Skymodel python package developed in this work simulates high resolution ($N_{\rm side}$ = 4096, $N_{\rm pix}$ = 0.86) mock maps of the microwave sky components at any frequency $\nu \in [27, 853]$ GHz. Those components are the Galactic foregrounds (see Sec. 4.1.1 & 4.3.1), Synchrotron emission, Galactic dust, Anomalous Microwave Emission (AME), free-free emission, and the extra-galactic components (see Sec. 4.1.2), Cosmic Microwave Background (CMB), Cosmic Infrared Background (CIB), Radio Point Sources (RPS), and the thermal and kinematic Sunyaev-Zeldovich effects (tSZ & kSZ). It can work like a black box for non-experts through the use of one single function. It is extremely versatile, allowing for the choice of microwave emissions to include, the pixel resolution of the map through the $N_{\rm side}$ parameter, units, the beam resolution (see Sec. 1.2.1), the inclusion of noises such as instrumental white noise (see Sec. 4.6.1) or atmospheric noise (see Sec. 4.6.2), and survey strategies through the available masks (see Sec. 4.6.3). The simulation on which, the extra-galactic components can be based, can also be selected depending on their characteristics, among, the Sehgal et al. (2010a), the Websky, or the Simons Observatory (SO) (see Sec. 4.3.2). It is public and fully available on GitHub (see Sec. 4.7.3). It is easy to use or get a hand on through the extensive documentation, which is composed of a full PDF document, a ReadTheDoc webpage, and example Jupyter notebooks for every main function of the code (see Sec. 4.7.2).

The Skymodel code filled up a need for mock simulations in Bonn at a given time. It provided skill and in-house knowledge. It is not outdated, but with the recent releases of PySM3 and PSM, it is not state-of-the-art anymore. This does not mean that it is or was not useful, in science it is important to have independent groups reproducing similar results. It also does not have a full overlap with PySM3 or PSM, it is much more versatile in terms of simulations for the extra-galactic components and is much faster to execute than PySM. It takes $\sim$ 14 min for PySM to compute a microwave sky containing all the Galactic foregrounds and CMB at a given frequency, and $\sim$ 45 s for an additional frequency. While takes only $\sim$ 28 s for the Skymodel to produce the mock sky with the exact same component, for each frequency. If only the Galactic components are simulated, i.e., without the CMB, it takes $\sim$ 1 min 43s for PySM to compute a map at a given frequency. It seems that the treatment of the CMB done by the Skymodel is much faster than PySM.

4.8.2 Future of the Skymodel

The Skymodel can be built on and improved to be once again state-of-the-art. It should be completed to simulate the polarization component for all the microwave emissions Galactic and extra-galactic alike. It should also include the new Galactic foreground models and template maps from PySM3 and update the extra-galactic templates map by using the latest version of the Websky simulations, Websky 04 (see Sec. 4.3.2). It should be updated to include relativistic correction of the SZ effects. The Skymodel code could be updated to work on arrays of any size, not necessarily HEALPix format maps, allowing it to base itself on hydrodynamical simulations.

CHAPTER 5

Exploring Cosmic Microwave Background Leakage into Thermal Sunyaev-Zeldovich maps

Contribution

The section 5.2.1 of this chapter presents the work published in:

- CCAT-Prime Collaboration ; Aravena, Manuel ; Austermann, Jason E. ; Basu, Kaustuv ; Battaglia, Nicholas ; Beringue, Benjamin ; Bertoldi, Frank ; Bigiel, Frank ; Bond, J. Richard ; Breysse, Patrick C. ; Broughton, Colton ; Bustos, Ricardo ; Chapman, Scott C. ; **Charmetant, Maude** ; Choi, Steve K. ; Chung, Dongwoo T. ; Clark, Susan E. ; Cothard, Nicholas F. ; Crites, Abigail T. ; Dev, Ankur ; Douglas, Kaela ; Duell, Cody J. ; Dünner, Rolando ; Ebina, Haruki ; Erler, Jens ; Fich, Michel ; Fissel, Laura M. ; Foreman, Simon ; Freundt, R. G. ; Gallardo, Patricio A. ; Gao, Jiansong ; García, Pablo ; Giovanelli, Riccardo ; Golec, Joseph E. ; Groppi, Christopher E. ; Haynes, Martha P. ; Henke, Douglas ; Hensley, Brandon ; Herter, Terry ; Higgins, Ronan ; Hložek, Renée ; Huber, Anthony ; Huber, Zachary ; Hubmayr, Johannes ; Jackson, Rebecca ; Johnstone, Douglas ; Karoumpis, Christos ; Keating, Laura C. ; Komatsu, Eiichiro ; Li, Yaqiong ; Magnelli, Benjamin ; Matthews, Brenda C. ; Mauskopf, Philip D. ; McMahon, Jeffrey J. ; Meerburg, P. Daniel ; Meyers, Joel ; Muralidhara, Vyoma ; Murray, Norman W. ; Niemack, Michael D. ; Nikola, Thomas ; Okada, Yoko ; Puddu, Roberto ; Riechers, Dominik A. ; Rosolowsky, Erik ; Rossi, Kayla ; Rotermund, Kaja ; Roy, Anirban ; Sadavoy, Sarah I. ; Schaaf, Reinhold ; Schilke, Peter ; Scott, Douglas ; Simon, Robert ; Sinclair, Adrian K. ; Sivakoff, Gregory R. ; Stacey, Gordon J. ; Stutz, Amelia M. ; Stutzki, Juergen ; Tahani, Mehrnoosh ; Thanjavur, Karun ; Timmermann, Ralf A. ; Ullom, Joel N. ; van Engelen, Alexander ; Vavagiakis, Eve M. ; Vissers, Michael R. ; Wheeler, Jordan D. ; White, Simon D. M. ; Zhu, Yijie ; Zou, Bugao (2023), *CCAT-prime Collaboration: Science Goals and Forecasts with Prime-Cam on the Fred Young Submillimeter Telescope*, ApJS, 264, 7

M.C. is a CCAT collaboration member and contributed to the SZ-science predictions in the section of this paper. M.C. produced parts of section 8.2 and Fig. 13 using simulations from the Skymodel code.

This chapter compiles all the preliminary work done around CIB leakage into tSZ maps. The goal was to make forecasts for upcoming ground-based telescopes. This chapter includes the preliminary work, null tests, and simple cases explored. As well, as some simple published predictions and results made for the Simons Observatory Large Aperture Telescope (SO LAT) and the Fred Young Submillimeter Telescope (FYST). Finally, additional forecasts and tests for the AtLAST concept idea are presented.

5.1 Introduction

This part gives an overview of the work done in this chapter, the astrophysical context surrounding it, and the motivations. The results are listed and further details and explanations are given in the following sections.

5.1.1 Overview

This work investigates the residual CIB noise left over by the ILC or CILC (see sec. 2.5.1) in the extracted tSZ map for upcoming ground-based telescopes. To emulate the microwave sky that will be seen for those upcoming telescopes, the Skymodel code is used (see Chap. 4). A set of multi-frequency full sky HEALPix maps (see Sec. 2.4.1) based on the Websky simulations (see Sec. 4.3.2) is generated. A tesselation of the sky with $N_{\rm side} = 4$, into 192 equal area (see Sec. 2.4.1) is done and a Galactic dust mask is applied (see Fig. 4.15). An ILC or CILC method deprojecting CIB is applied to each of the 192 fields and a map of the full sky Compton-y is reconstructed. The angular power spectrum of this map is computed using PyMaster (see Sec.). The CIB residual noise (see Eq. 2.32) is then computed for several configurations. It is found in a simple test case, with only the signal and dust noises, that the more dust contamination in the sky, the more SO will benefit from being combined with FYST data (CCAT-Prime Collaboration et al., 2023). The frequency bands probing the decrement of the tSZ effect are the most important channel for its reconstruction, in all cases, independently of the sky composition. The frequency bands close to the tSZ null are essential for noise minimisation, except when atmospheric contamination becomes too strong. Atmospheric contamination also makes the CMB residual noise blow-up, this can be dealt with by using prior knowledge of the CMB power spectrum to estimate the CMB noise residual and debias the result from it. In the most contaminated case, combining FYST to SO should reduce the CIB contamination by $\sim 35\%$, and using a CILC to deproject CIB, SO and FYST combined should be able to directly detect the tSZ effect with a signal-to-noise above 1.

5.1.2 Context of the study

The Cosmic Infrared Background (CIB) is spatially correlated with the thermal Sunyaev-Zeldovich (tSZ) signal (Maniyar et al., 2021). Indeed, the tSZ signal comes from the hot gas in galaxy clusters, and the CIB signal is from the gas remission in star-forming regions in galaxies and clusters of galaxies. Even though, the tSZ signal mainly comes from clusters around $z \approx 0.5$ and the CIB around $z \approx 1.5$, a small correlation between the two is still expected (Addison et al., 2012). The correlation between the two effects does not comply with the hypothesis of the ILC (see Sec. 2.5.1) that the components are uncorrelated. The consequence of this correlation between the two signals is the incapacity of the ILC to fully separate them. Another reason why the ILC can not separate the CIB and why the recovered tSZ signal is biased by the CIB originates from the fact that the CIB is not well constrained by one SED (see Sec. 1.4.3). The CIB is composed of infrared emission coming from sources at different

Table 5.1: List of emissions contained in the simulated mock sky for each specific test case.

	Skymodel composition \ Cases	S1	S2	S3	S4
Extra-galactic	Comic Microwave Background (CMB)			✓	✓
	Cosmic Infrared Background (CIB)	✓	✓	✓	✓
	thermal Sunyaev-Zeldovich (tSZ)	✓	✓	✓	✓
	kinematic Sunyaev-Zeldovich (kSZ)			✓	✓
	Radio Point Sources (RPS)				
Galactic	Galactic Dust		✓	✓	✓
	free-free			✓	✓
	Synchrotron			✓	✓
	Anomalous Microwave Emission (AME)			✓	✓
Noises	Instrumental Noise (IN)	✓	✓	✓	✓
	Atmospheric noise (Atmo) 0% correlated				
	Atmo 90% correlated				✓

redshifts, meaning the SED of those populations are different and the frequency bands at which the data is observed will be sensitive to different CIB sources (Hurier et al., 2013). The ILC recovered tSZ power spectrum is, therefore, biased by the CIB. The tSZ can be used for cosmology or for cluster Astrophysics (1.5.1). The CIB would bias the results deduced from the tSZ signal. Therefore, the CIB contamination must be quantified and if possible mitigated.

5.1.3 Motivations

In this chapter, the CIB leakage is investigated, in several test cases. Each sky configuration for the different test cases is summarized in Tab. 5.1. The motivations for looking at SO LAT and FYST separately but also combined are, that their frequency is complementary SO LAT scanning at frequencies $\nu \leq 280$ and FYST at $\nu \geq 280$, their resolution is also similar a few arcmin over their frequency range (see Tab. 4.2 & 2.3). Another important motivation is that the CIB emission dominates the energy budget in the sky at high frequencies ($\nu \geq 350$ GHz) making it highly interesting to see if FYST and its high frequencies channels could help mitigate the CIB contamination because it see the CIB emission much better than SO LAT. A nonscientific but nevertheless important motivation is our affiliation with the CCAT collaboration. Therefore, forecasting the possible results of our instrument to help refine the science goal is an absolute necessity. Many CCAT collaboration members are also members of the SO team, which allowed for privileged access to some of the SO characteristics like the noise curves, crucial to make realistic simulations.

5.2 Simple sky tests

This part explores several mock skies that are simplified and far from containing all the emissions and subtilities of the real expected sky. This is used to build trust in the pipeline developed in this work and

to better understand the behaviour of the dynamic component vs. telescope characteristics with isolated variables.

5.2.1 Sky with only signal and noise

The simplest test case investigated in this work is when the sky is only composed of the signal of interest tSZ or the dust contaminant, CIB, instrumental white noise and eventually Galactic dust (see S1 & S2 of Tab. 5.1). The mock sky maps are generated using the Skymodel (see Sec. 4) and the tSZ signal is extracted using the ILC pipeline (see Sec. 2.5.1). The small goal investigated was to check how the cumulative noise residual alone would behave in such a sky for SO alone and SO and FYST combined. It was found (see Fig. 13 of CCAT-Prime Collaboration et al. (2023)) that residual noise was reduced by ∼16% when combining SO and FYST rather than for SO alone in the S1 case. Meaning that the noise residual is composed only of CIB and instrumental noise because they are the two other component target signals excluded. If Galactic dust is added to the mock sky budget (S2), then the advantage of combining FYST and SO increases, because in this case, even though the noise residual is higher, the leverage of adding FYST is better. The noise is reduced by ∼ 22% compared to SO alone. This leads to several possible interpretations, either SO performs worst when Galactic dust is present in the sky, or another possibility is that FYST is simply more performant than SO at reducing Galactic dust in particular, which would make sense because of its high-frequency coverage. The last possibility is of course a combination of the two. This is further investigated in the following parts, in particular, a large overview is given of the performance of every component of the microwave sky.

5.2.2 Simple sky: without atmospheric noise

We place ourselves in the simple sky scenario without atmospheric noise (see Tab. 5.1).

ILC weights

Fig. 5.1 top panel shows the relative contribution of each frequency band weight to the final reconstructed y map. For SO alone, the orange circles of Fig. 5.1 indicate that the 93 GHz frequency channel is by far the greatest contributor to the reconstruction of the tSZ signal, but it is overestimating the signal. The 145 GHz and 279 GHz channels are correcting that by removing the fractional excess signal. The 225 GHz channel contributes only marginally (∼ 10%) to the reconstruction of the tSZ signal, which is expected, as it is close to the tSZ SED null (220 GHz), the band does not contain much of the signal. The FYST telescope alone, probing none of the SZ decrement channels (below 217 GHz) uses almost exclusively the 280 GHz channel for the signal reconstruction (see green circles of Figure 5.1). For SO+FYST the main contributors driving the signal reconstruction are the frequencies band probing the SZ decrement, this is expected as this negative signal at frequencies below 220 GHz is characteristic of the tSZ signal. Thus we see that the SO 93 GHz and 145 GHz channels are the main drivers of the tSZ reconstruction, they respectively contribute by ∼ 55% and ∼ 50% to the signal reconstruction. We see that combining SO+FYST avoids the overestimation of the tSZ signal by the 93 GHz channel that was happening for SO alone and the tSZ signal being reconstructed almost exclusively by the 280 GHz channel for FYST alone. This was only the case because for FYST alone this is the lowest frequency channel containing some SZ signal when one goes to higher frequencies dust contamination becomes too important.

The bottom panel of Figure 5.1 shows the weights alone, which are constrained by the only two conditions of Eq. (2.13) and the condition of minimum variance. Therefore when the product $\omega_i a_i$ in the upper panel is null but the weight value is high in the bottom panel, we can conclude that the associated frequency band is a strong contributor to the noise minimization, not being one for signal reconstruction. This is especially true for the 220 GHz frequency band, the weight associated with this band is high, for SO+FYST combined it is the highest weight given by the ILC. It is not used to reconstruct the tSZ signal, the signal being absent from this frequency channel. The map at this frequency contains all the other astrophysical components, which are the noise contaminants of our signal. Thus, making this map the most useful for noise minimization in the ILC. For SO alone, we see that even though the 225GHz channel was not contributing much to the tSZ reconstruction (see top panel) it has the second-highest weight in absolute value and is an important channel for noise minimization. For FYST alone, we see the same behaviour but for the 220GHz channel. Combining SO+FYST rebalances the weights given to 93GHz and 145GHz for tSZ reconstruction, and uses the 220GHz for noise minimization and the higher frequencies channel for dust suppression with their small weights.

The addition of FYST reduces the standard deviation of the weights, even for the SO channels. This means that the different fields of the maps are treated more uniformly, the ILC is less affected by non-uniformly spatially distributed contaminants, and it better targets the tSZ signal. This could be due to the lower sensitivities of FYST.

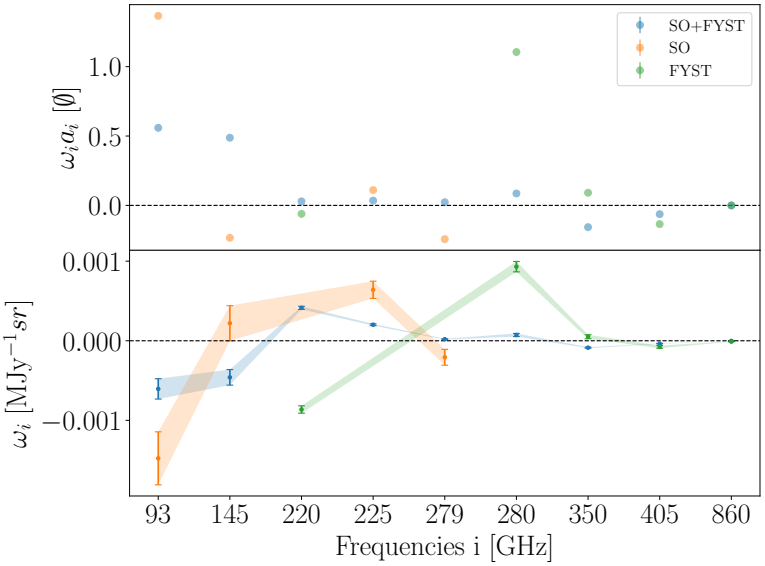

Figure 5.1: Comparative plot of the ILC-weights for telescopes like SO, FYST and SO+FYST combined. **Top panel** : fractional contribution of each frequency channel to the recovered y map. **Bottom figure**: represent the channel that contributes the most to maximizing the SZ signal and minimizing the noise. The error bars show the standard deviation around the mean weight of all of the 192 fields for each frequency map.

Power spectra

In Figure 5.2 and Figure 5.3 one can see the ILC-extracted power spectrum compared to the various Astrophysical contaminants residual noises power spectra, for SO and SO+FYST respectively. For

SO alone, on the top panel of Figure 5.2, the residual Instrumental Noise (IN) and the CIB residual noise power spectra ($r_{\rm IN}$ & $r_{\rm CIB}$) both start to dominate above the input y signal around $\sim l = 500$, the instrumental noise residual growing much faster and being the dominant contaminant of the reconstructed Compton-y map. The other residual power spectra, including the CMB ($r_{\rm CMB}$), kSZ $r_{\rm kSZ}$ and the cumulative residual noise of all the galactic foreground r_f are much lower. The bottom panel of Figure 5.2 shows that the SO Signal-to-Noise (S/N) ratio decreases as the scales get smaller, its peak being at around $S/N \sim 0.25$. For SO+FYST, in Figure 5.3, the residual IN power spectrum starts to dominate above the input y signal around $\sim l = 1000$ and the CIB residual noise power spectrum around $\sim l = 1500$. The other residual power spectra, including the CMB, kSZ, and the cumulative residual noise of all the galactic foreground are much lower. The bottom panel of Figure 5.3 shows that the SO+FYST Signal-to-Noise (S/N) ratio decreases as the scales are getting smaller, its peak being at around $S/N \sim 0.90$.

Comparing Figure 5.2 and Figure 5.3 shows that combining SO+FYST push back the limit at which the $C_\ell^{r_{\rm IN} r_{\rm IN}}$ dominates over the input y by $\sim l = 500$ and the limit by which $C_\ell^{r_{\rm CIB} r_{\rm CIB}}$ dominates over the input y by $\sim l = 1000$. The S/N ratio is almost 4 times higher on large scales when SO+FYST are combined. This is due to all the noise residuals being lower for SO+FYST combined rather than SO alone. The addition of the FYST bands and sensitivities lower the CIB residual, but also all the other noise residuals compared to SO alone. In particular, we see that combining SO+FYST lowers the residual kSZ, CMB, and cumulated galactic foregrounds by almost an order of magnitude.

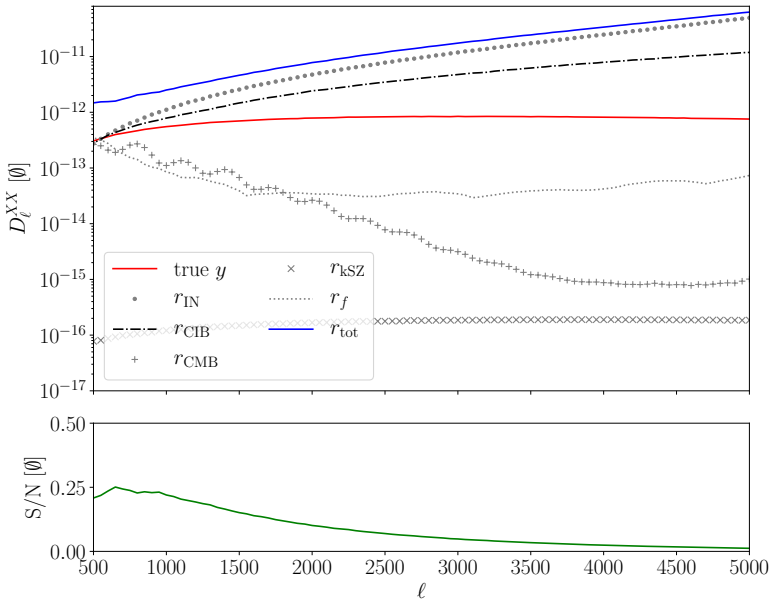

Figure 5.2: For SO alone. **Top panel**: In red, target signal, the power spectrum of the input Compton-y map coming from the WebSky simulations. In blue the power spectrum of the total residual noise in the ILC-extracted tSZ map. All the other lines are the power spectra of the associated residual noises $r_{\rm X}$. X being the Instrumental Noise (IN), kSZ, CMB, or the foregrounds (f) in grey and black for the CIB. **Bottom panel** : Signal-to-Noise (S/N) ratio. All power spectra are beam-corrected for comparison with the input y. They are also all binned and averaged over a window $\Delta\ell = 50$.

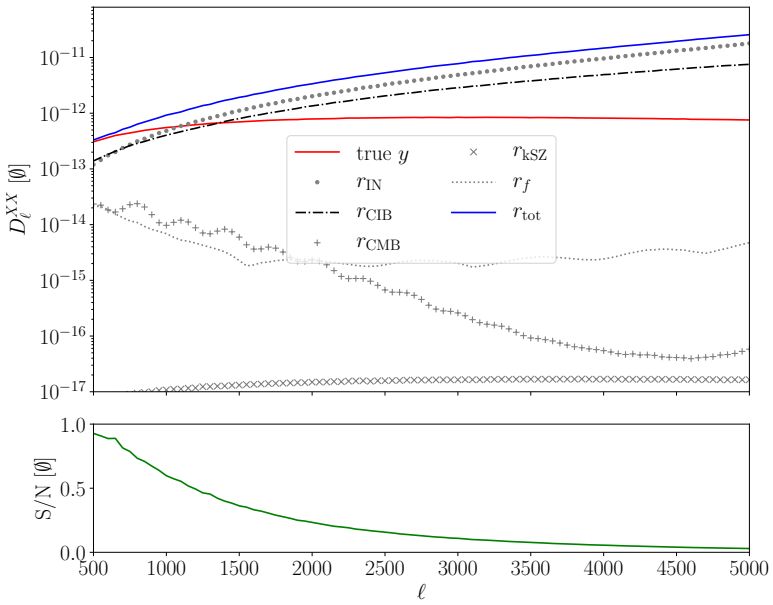

Figure 5.3: For SO+FYST combined. **Top panel**: In red, target signal, the power spectrum of the input Compton-y map coming from the WebSky simulations. In blue the power spectrum of the total residual noise in the ILC-extracted tSZ map. All the other lines are the power spectra of the associated residual noises $r_{\rm X}$. X being the Instrumental Noise (IN), kSZ, CMB or the foregrounds (f) in grey and in black for the CIB. **Bottom panel** : Signal-to-Noise (S/N) ratio.All power spectra are beam-corrected for comparison with the input y. They are also all binned and averaged over a window $\Delta\ell = 50$.

5.3 Realistic sky

In this section, a realistic sky was built, meaning containing all the Galactic foreground emissions and extra-galactic ones (see Tab. 5.1 S3 & S4).

5.3.1 Method

The composition of the Microwave sky was made as realistic as possible. The extragalactic components were generated using the Skymodel code (see Sec. 4) based on the Websky (Stein et al., 2018, 2020) templates from 2019 (see Tab. 4.4). The Galactic foregrounds were generated using the PySM (**?**) and upgraded to the same resolution as the Websky maps ($N_{\rm side} = 4096$, see details in Sec. 4.3.1) by using the HEALPy (see Sec. 2.4.1) function `ud_grade()` followed by smoothing with a 10' Gaussian kernel. The Skymodel code converted all the maps to MJy/sr units, added instrumental white noise and atmospheric red noise based on the noise curves of Fig. 4.12, and smoothed each map to the Gaussian beam. The noises and beam sizes are generated using the specificities of the SO LAT telescope and FYST (Tab. 2 of App. C.3). Another feature is added to the noise to make it more realistic. In reality, when SO LAT or FYST will be observing the sky with their cryostat (see Fig. 2.7 & 2.9) it is commonly accepted that the two neighbouring channels optics tubes will observe the same portion of the sky at the same time, meaning the atmospheric noise they observe will be correlated. It is classically assumed that the atmospheric noise will be 90% correlated between the two neighbouring frequency channels

(Simons Observatory Collaboration, 2019). The implementation of this correlation of the atmospheric noise can be seen in (Sec. 4.3 of App. C.3). The set of simulated maps for SO LAT alone, SO LAT, and FYST combined, or FYST alone are then fed into the ILC code or CILC to extract back the tSZ effect with or without CIB deprojection. The code smooths all the maps down to the lowest resolution, and tesselates the sky using the $N_{\rm side} = 4$ nested scheme from HEALPix, separating the full sky map in 192 fields. A Galactic dust map is applied and ILC or CILC is applied to each of the 192 fields separately. Because of the sky tesselation, many ILCs are applied to the 192 field covering the sphere. For each of those fields, a set of weights optimized for the contamination in this region is computed. A schematic illustrating the method can be seen in Fig. 5.4. Once the ILC or CILC extracted tSZ map is available, it is possible to estimate the total residual noise () left over by all the other astrophysical emissions, by subtracting the input tSZ map from the simulation from the extracted one. The power spectrum of the input tSZ map, and of the residual noise map using PyMaster (see Sec. 2.4.2). The power spectrum of the noise is then compared to the input power spectrum (see Fig. 5.5). Because it all simulations, it is possible to further decompose the total noise residual into the contribution of each Astrophysical contaminant to the total noise residual. This is done to assess the CIB leakage, the CIB noise residual left into the recovered tSZ map.

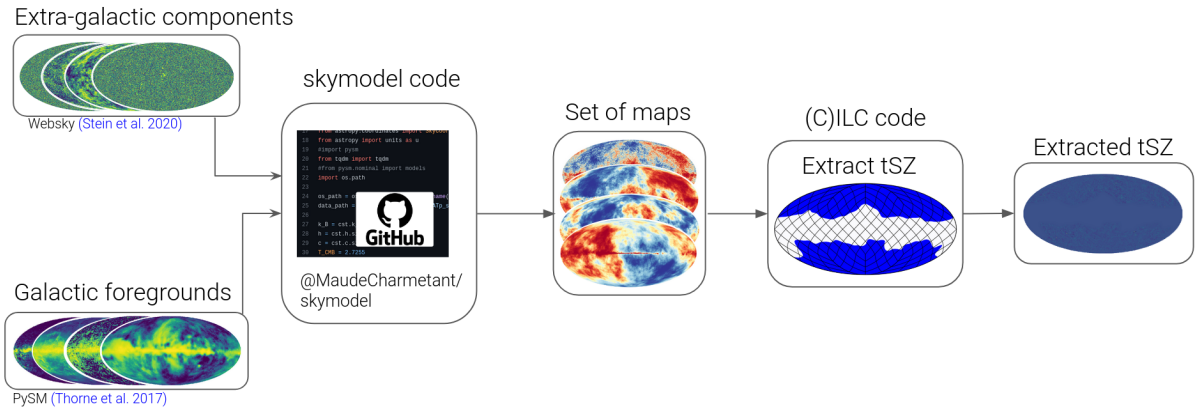

Figure 5.4: Schematic of the method used in this work to extract a map of the tSZ effect. The inputs are template maps of the extragalactic component from the Websky simulations and of Galactic foregrounds from PySM. The Skymodel code adds the noise and produces the maps at the correct specificities to reproduce the Sky as seen by SO LAT, FYST, or FYST and SO LAT combined. The set of maps produced is fed into an ILC or CILC, the sky tesselated and masked with a Galactic dust mask. An ILC or CILC is applied to each of the sky regions and an extracted map of the tSZ effect is retrieved.

debiasing

Once the total noise residual power spectrum is computed, information on some contaminants can be used to reduce the noise. The instrumental white noise coming from the instrument can be measured or deduced from the characteristics and calibration of the instrument. The intensity of the noise per pixel over the sky can be used to generate random Gaussian realization of the noise using HEALPy. First, the power spectrum of the map is computed using the function `healpy.sphtfunc.anafast()`. Then this power spectrum is used to produce a new realization of the instrumental noise with the same properties as the original one using `healpy.sphtfunc.synfast()`. The ILC residual power spectrum of this

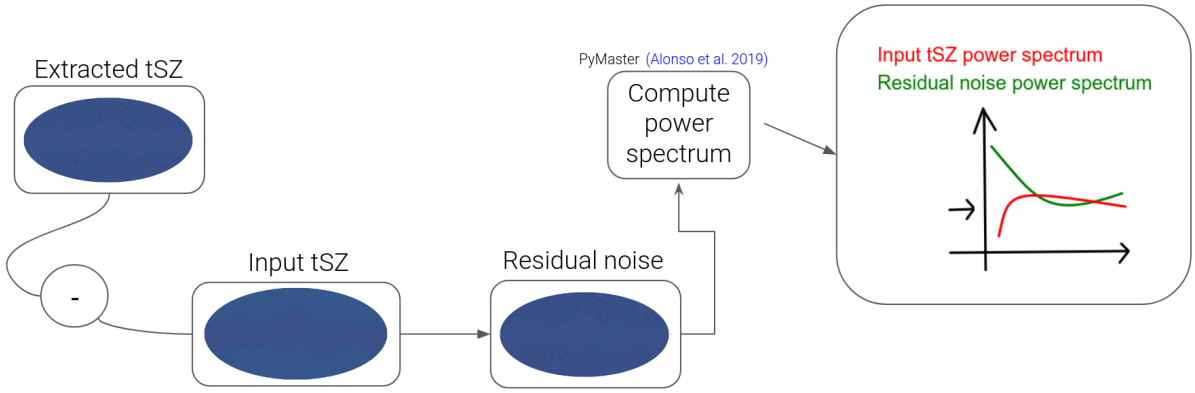

Figure 5.5: Schematic of the method used to compute the total noise residual and compare it to the input target signal. The input tSZ map from the simulations templates is substracted to the ILC or CILC extracted tSZ map to obtain the total noise residual. The power spectrum of the total noise residual is then computed using PyMaster and compared to the one of the input map.

new map realization and be computed by multiplying the maps of each frequency by the ILC weights used to recover the signal of interest. This residual instrumental noise estimating is then subtracted from the final total noise residual of the ILC or CILC. The same can be done for the CMB because its power spectrum is very well constrained by observations. One might think that because it is all simulations, the noise is perfectly known and could all be removed. This is correct, but the goal is to try to eliminate only the noise that can be estimated in a real experiment when only the total sky emission is measured and the noise content, proportion, and respective intensities are unknown. In this case, only a few noises can be estimated and removed, only the one we can estimate to a very good precision, which is the case of the instrumental noise and CMB.

5.3.2 Results without debiasing

The ILC output is a map containing the Compton-y signal and residual noises (eq. 2.32). Using PyMaster, the power spectra of this map and the residual noises can be computed and compared to the input pure Compton-y map power spectra from the simulation. We define, $D_\ell^{XX} = \ell(\ell+1)C_\ell^{XX}/2\pi$ the power spectra of the quantity X or of the residual noise r_X of the quantity X. For SO alone (see Fig. 5.6 top left panel) the dominant noise residual on small scales is the Instrumental white noise residual ($r_{\rm IN}$) which start to dominates around $\ell \sim 2300$, followed by the CIB noise residual ($r_{\rm CMB}$) which dominates on $\ell \sim 3800$. The cumulative foreground residual (r_f) and the kSZ residual ($r_{\rm kSZ}$) are both subdominants to the input signal by around one order of magnitude over the scales $\ell \in [500, 5000]$. The atmospheric noise residual ($r_{\rm Atmo}$) dominates the input signal for $\ell < 1400$ and $\ell > 4500$. On large scales, up to $\ell > 2800$, CMB is the dominant noise residual ($r_{\rm CMB}$) by more than two orders of magnitude. This is an increase of three orders of magnitude compared to the case without atmospheric noise (see Fig. 5.6). This can be explained by the condition of minimum variance on the ILC weights. Without atmosphere the ILC can tailor the weights to minimize the CMB variance, with atmospheric noise, the ILC weights will adapt to minimize this additional strong noise and therefore be less tailored to minimize CMB contamination.

In the case of SO+FYST combined (see right panel of Fig. 5.6), the situation is similar to the SO

alone case except that for SO+FYST the CMB residual power spectrum ($r_{\rm CMB}$) is around a factor ~ 2 lower than for SO alone. The instrumental noise residual ($r_{\rm IN}$) power spectrum and the atmospheric one ($r_{\rm Atmo}$) are higher when FYST is present, this could be due to FYST's higher noise because of its higher frequencies channels. The CIB noise residual power spectrum is significantly lower for SO+FYST combined. It starts to dominate above the input signal around $\ell \sim 4200$ while for SO alone it was around $\ell \sim 3700$. We conclude that adding FYST to SO helps mitigate CIB, CMB, and kSZ noise residuals, and even though the residual atmospheric and instrumental noise are increased the overall total noise residual is smaller in amplitude for SO+FYST combined.

The lower panel of Fig. 5.6 shows to which extent the input Compton-y power spectra are dominated by the total ILC noise residual as a function of the scale. In terms of simply recovering the input signal, SO+FYST does a better job than SO alone up to $\ell \sim 2700$.

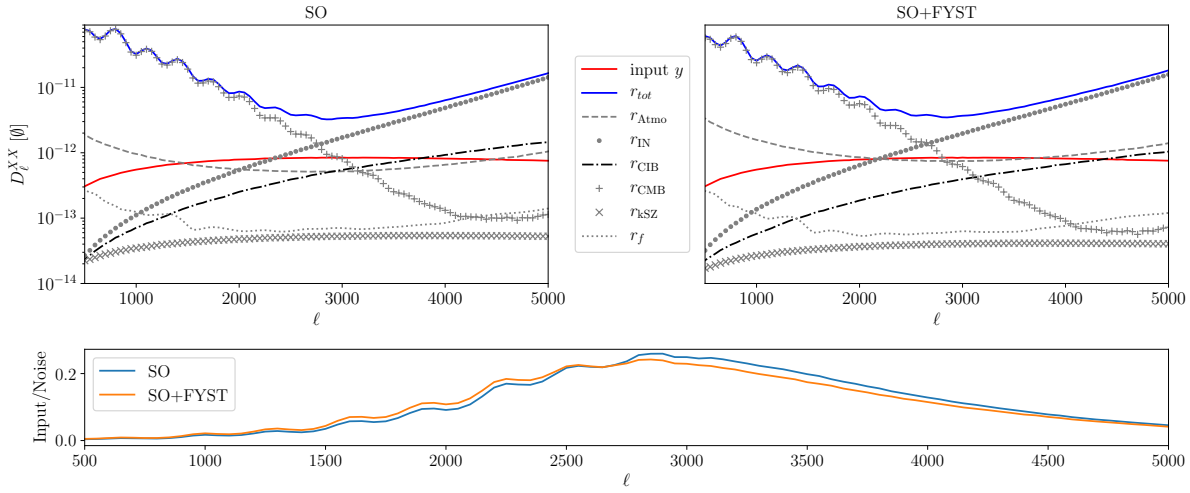

Figure 5.6: Power spectra of the ILC residual noises compared to the input Compton-y power spectrum (red). The simulated sky contains all extra-galactic components (see Tab. 5.1), galactic foregrounds, and Atmospheric noise. In red, is the target signal, the power spectrum of the input Compton-y map coming from the WebSky simulations.Top panels: grey curves are the ILC noises residual power spectra, denoted by $r_{\rm X}$ X being the Instrumental Noise (IN), kSZ, CMB, or the foregrounds (f). In black the power spectrum of the CIB residual noise. In blue the sum of all grey and black curves, the total ILC noise residual. The left panel is for SO alone and the right panel is for SO+FYST combined.Bottom panel: the ratio between the input Compton-y power spectrum (red) and the total noise residual (blue) for SO and SO+FYST.All power spectra are beam-corrected for comparison with the input y. They are also all binned and averaged over a window $\Delta\ell = 50$.

Some intermediate results with only the instrumental noise debiased can be seen in Fig. C.4.

5.4 Predictions for the AtLAST concept idea

Some predictions were made using our Skymodel pipeline and CILC code for the concept idea of the Atacama Large Aperture Submillimeter Telescope (AtLAST) (see Sec. 2.2.3). This concept is as of now not yet funded, but this work serves as an illustration of the capabilities of our pipeline for a theoretical telescope that so far has virtually no limit. An overview of how the results were generated, and of the extraction process is presented, followed by the results and interpretation. The goal was to test what is

the best selection of frequency channels to retrieve the tSZ effect using an ILC.

5.4.1 Simulations of the data

The mock set of maps generated by the Skymodel (see Sec. 4), to emulate the kind of data the AtLAST concept would produce, was based on the (Stein et al., 2018, 2020) templates from 2019 for the extra-galactic components (see Tab. 4.4). A list of emissions simulated in this case can be seen in Tab. 5.1, the sky contains, all the extra-galactic components, radio point sources excluded, and all Galactic components. Regarding the noise, no atmospheric noise is present because the Skymodel can only process atmospheric noise at the SO LAT and FYST frequencies. In this study, we did not investigate how to generate atmospheric noise curves for other frequencies/telescopes. The instrumental noise is computed from the Noise-Equivalent-Temperature (NET) values in $\mu K_{CMB} s^{1/2}$. Those values were provided by Kenichi Karatsu. The conversion to $\mu K_{CMB}-$arcmin is done through,

$$\sigma_N = \sqrt{\frac{(\sqrt{2}NET)^2\Omega}{N_{\rm pixel}T_{\rm obs}}}, \tag{5.1}$$

where Ω is the size of the observing patch in radians and $T_{\rm obs}$ the observation time in seconds. σ_N must be divided by the beam size to obtain the noise that will be seen by the telescope per pixel. The size of the observation patch was considered to the similar to the SPT one, which is around 2 500 deg^2. The SPT sky mask was applied to the mock sky maps (see Fig. 4.14). Therefore, $\Omega = 4\pi 0.06 \times 3600 \times (180/\pi)^2$. The number of pixels is given by the number of detectors 250 000 detectors per band were considered and 3 000 for the observing time ($T_{\rm obs} = 10800 \times 1000$). For the frequencies, 3 cases were tested:

- **Case 1:** No high-frequency channels $\nu < 345$ GHz
- **Case 2:** A few high-frequency channels $\nu < 680$ GHz
- **Case 3:** Going to very high frequency with up to $\nu = 860$ GHz.

The values of the frequencies, noises, and beam sizes for each of the three cases can be found in Tab. 5.2. The sky was also tesselated using $N_{\rm side} = 4$, to separate the sphere into 192 fields of equal area. One important caveat of this study is that it was conducted on maps of $N_{\rm side} = 4096$, meaning, with a pixel resolution of pix$_{\rm size} \approx 0.86'$. However, the AtLAST beam is more than three times smaller than the pixel resolution. Which does not comply with the Nyquist-Shannon sampling theorem, the requirement would be to have a pixel size two times smaller than the beam, i.e. of pix$_{\rm size} \approx 0.12'$. This pixel resolution would be equivalent to $N_{\rm side} = 32768$, which is currently not attainable by HEALPix, the $a_{\ell,m}$ being only defined up to $N_{\rm side} = 8192$ in the code. Therefore, one should not trust the scales above $\ell \sim 5000$ in this study.

5.4.2 Results

This part summarizes the results an AtLAST-like experiment could achieve regarding tSZ recovery with an ILC and in a simulated mock sky without atmospheric noise. The goal was to see in 3 test cases, the impact of including more or less high-frequency channels. First, the impact on the ILC weights is shown followed by the impact on the recovered power spectrum and the dominant noise residuals.

Table 5.2: Values of the frequency bands, noise, and beam sizes for the 3 AtLAST test cases.

			Frequency [GHz]	NET [$\mu K_{CMB} s^{1/2}$]	beam [']	σ_N [μK_{CMB}−arcmin]
Case 1	Case 2	Case 3	100	885.09	0.25	0.25
			110	1148.53	0.22	0.32
			130	1276.73	0.19	0.39
			140	1276.62	0.17	0.42
			150	1421.52	0.16	0.48
			160	1647.01	0.15	0.55
			170	2126.93	0.14	0.67
			200	2129.80	0.12	0.79
			220	2432.86	0.11	0.93
			240	3273.26	0.10	1.17
			260	4240.32	0.09	1.45
			280	5862.58	0.08	1.84
			300	8331.27	0.08	2.35
			345	20008.53	0.07	4.18
			405	59122.44	0.06	8.45
			433	280615.94	0.05	19.69
			462	526077.03	0.05	28.76
			482	977971.49	0.05	40.92
			495	1482768.63	0.05	51.74
			640	11675268.20	0.03	187.74
			680	12252847.54	0.03	204.34
			815	2.54×10^8	0.03	1117.16
			870	3.24×10^8	0.02	1345.41

ILC weights

Looking at the ILC weights gives information on which frequency channel contributes the most to the recovery of the tSZ signal and depreciation of the noise residual (see Sec. 2.5.1). For an AtLAST-like experiment, with the conditions of this study and in case 1, 110 GHz, 130 GHz, 140 GHz, and 345 GHz are the main contributors to the tSZ recovery (see top panel of Fig. 5.7), by a factor of respectively, $\omega_{110}a_{110} \approx 4.4$, $\omega_{130}a_{130} \approx 2.4$, $\omega_{140}a_{140} \approx 0.8$ and $\omega_{345}a_{345} \approx 0.9$. The tSZ signal is overestimated by these frequency channels. The ILC balances it out using 100 GHz, 150 GHz, 160 GHz, 170 GHz, 200 GHz, 260 GHz, 280 GHz and 300 GHz, which are recovering the signal with a factor of respectively $\omega_{100}a_{100} \approx -2$, $\omega_{150}a_{150} \approx -0.6$, $\omega_{160}a_{160} \approx -1.3$, $\omega_{170}a_{170} \approx -1$, $\omega_{200}a_{200} \approx -0.2$, $\omega_{260}a_{260} \approx -0.3$, $\omega_{280}a_{280} \approx -0.8$ and $\omega_{300}a_{300} \approx -0.7$. For the concept idea that is AtLAST, as such as, for SO and FYST (Charmetant and Erler, 2023), the frequency band contributing the most to the tSZ reconstruction, and even overestimating it are the ones located at the tSZ decrement ($\nu <\sim 220$ GHz). Without atmospheric contamination, such as in this AtLAST concept test case or for *Planck* we also have a small contribution from the frequency bands around 300 GHz. It is interesting to note that in this AtLAST concept test case, contrary to *Planck* (see Sec. 2.5.1), the 200 GHz channel is not attributed a strong weight by the ILC, it is actually among the smallest weights values (see bottom panel of Fig. 5.7). Meaning the ILC is not using this frequency band that much for noise minimization, or at least not more than any other

channel. This is probably due to the large frequency coverage of the tSZ decrement and between the tSZ null and ~350 GHz. If the tSZ SED is largely covered below the dust contamination peak, the ILC might not require as much to focus on a map containing only the noise contaminants.

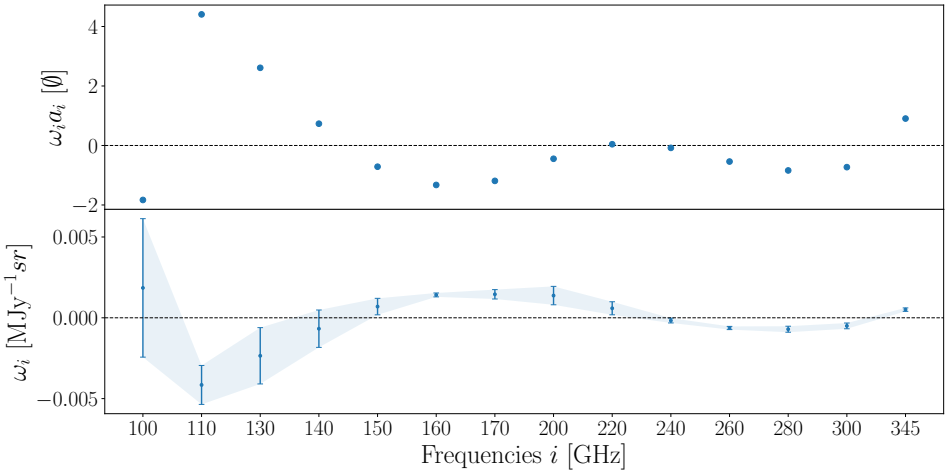

Figure 5.7: Weights return by the ILC used to extract the tSZ signal on AtLAST-like mock sky in Case 1 (see Tab. 5.2). Top panel: fractional contribution of each frequency band to the tSZ reconstruction. Bottom panel: the value of the weights, the error bar, and the shaded region representing the dispersion around the mean weight among the 192 fields.

If more high-frequency channels are added like in Case 2 (see Tab. 5.2) the overestimation of the tSZ reconstruction by the 110 GHz band is greatly diminished $\omega_{110}a_{110} \approx 3$, (see top panel Fig. 5.8). The weight behaviour is very similar to Case 1, except for the 345 GHz channel being now attributed to a negative factor rather than a positive one. However, Fig. 5.8 shows that the added high-frequency bands do not contribute to the tSZ reconstruction, nor do they contribute to the total noise minimisation in the recovered tSZ map. Those findings are further confirmed in Case 3 (see Tab. **??**), the two added high-frequency bands at 815 GHz and 870 GHz do not change the values of the weights nor do they change the fractional contribution of the frequency bands to the tSZ reconstruction. Those two frequency bands do not contribute to it themselves and are not used by the ILC for noise minimisation (see Fig. 5.9). However, as was already shown in the literature the frequency channel only gives information on how the ILC deals with the tSZ reconstruction and the total noise residual. It does not give any information on how well some of the noise contaminants might change with the addition of more frequency channels. That is why it is important to look at the power spectra to get a complementary picture.

Power spectra

The power spectra give information on the impact of the various frequency ranges on the tSZ recovery and the various noise residuals from a different and more global perspective than the weights.

Fig. 5.10 shows a comparison of the total noise residual and of the two dominant contributors of this total noise residual for the three cases. It shows that including 7 more frequency channels between $\nu \in [405, 680]$ GHz (Case 1 to Case 2) reduces the total ILC noise residual (r_{tot}) by a factor of ~ 2 (full blue curve compared to the green one). The further addition of two high-frequency channels (Case 2 to Case 3) only reduces the total noise residual by a few per cent. The results and factors are similar

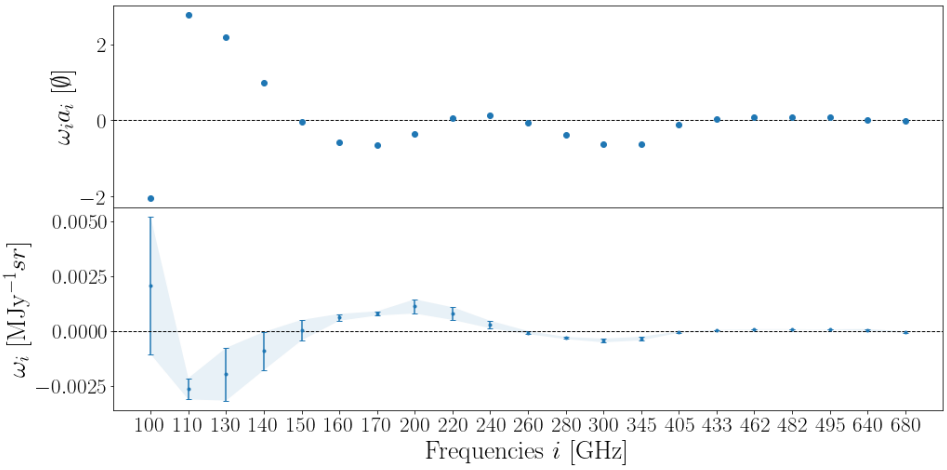

Figure 5.8: Weights return by the ILC used to extract the tSZ signal on AtLAST-like mock sky in Case 2 (see Tab. 5.2). Top panel: fractional contribution of each frequency band to the tSZ reconstruction. Bottom panel: the value of the weights, the error bar, and the shaded region representing the dispersion around the mean weight among the 192 fields.

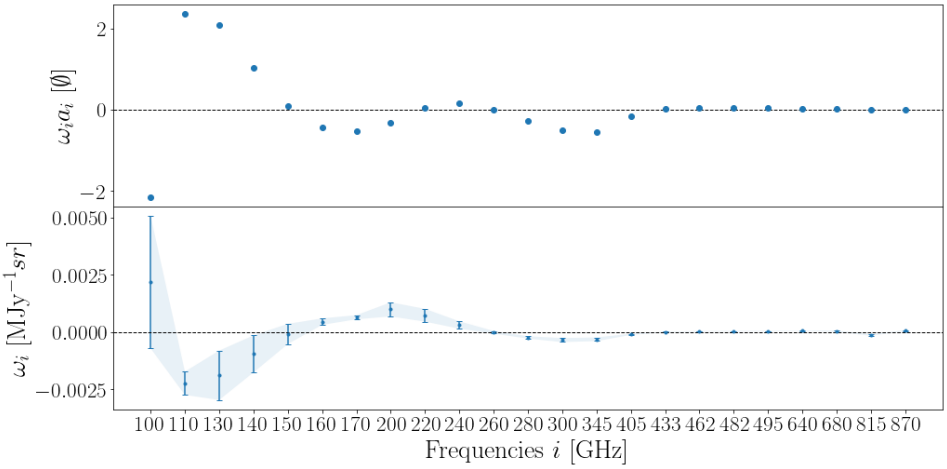

Figure 5.9: Weights return by the ILC used to extract the tSZ signal on AtLAST-like mock sky in Case 3 (see Tab. 5.2). Top panel: fractional contribution of each frequency band to the tSZ reconstruction. Bottom panel: the value of the weights, the error bar, and the shaded region representing the dispersion around the mean weight among the 192 fields.

for the CIB ($r_{\rm CIB}$) and CMB ($r_{\rm CMB}$) noise residuals. In conclusion, the AtLAST concept, having more frequency channels would help bring the noise residual down, but only by a small factor in comparison to the number of frequency bands that need to be added to make a difference. The component-separated noise residual also does not exhibit different behaviour, which would lead to a benefit in adding only one or two frequency channels to significantly reduce a component (like the CIB for example).

It is also interesting to have a general view of all the noise components of the total noise residual (r_{tot}) for a concept such as AtLAST in the three cases and compare it to the performances of SO and FYST combined. Consider an AtLAST-like experiment in Case 2 (see Tab. 5.2), when trying to extract back a

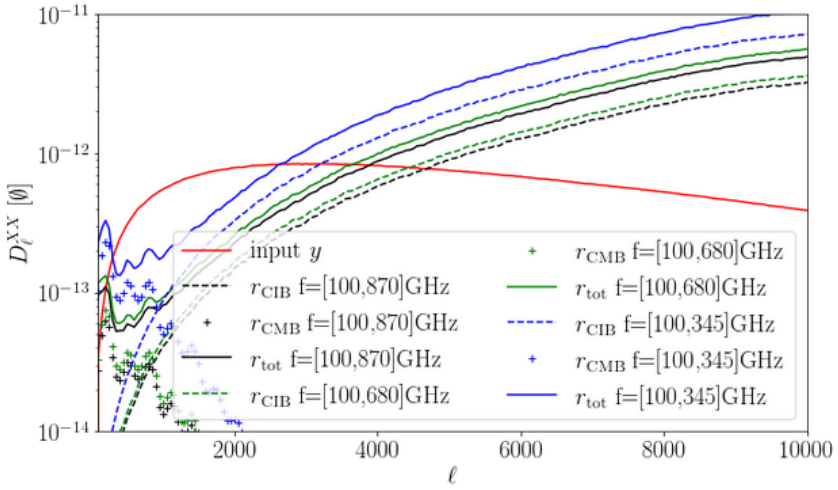

Figure 5.10: Power spectra of different quantities. In full the power spectrum of the total ILC noise residual (r_{tot}). In dotted the power spectrum of the CIB noise residual ($r_{\rm CIB}$) and in + signs, the power spectrum of the CMB noise residual ($r_{\rm CMB}$). In blue for Case 1, in green for Case 2 and in black for Case 3 (see Tab. 5.2).

tSZ signal. Fig C.5 shows that the total ILC noise residual dominates above the signal of interest over the very large scales ($\ell < 200$) because of the CMB ($r_{\rm CMB}$) and cumulative Galactic foreground noise residual (r_f). Over intermediate scales ($200 < \ell < 3500$) the extracted-tSZ is signal dominated. On small scales ($\ell > 3500$) the ILC-extracted tSZ signal becomes once more noise dominated, the main contaminant being the CIB noise residual ($r_{\rm CIB}$) directly followed by the instrumental noise residual ($r_{\rm IN}$). The kSZ noise residual ($r_{\rm kSZ}$) is subdominant to the total noise by more than four orders of magnitude over the ℓ range ($\ell \in [0, 10000]$). It seems like a concept telescope like AtLAST would provide a very large ℓ-window to measure an unbiased tSZ power spectrum. But one should not forget that this study does not include atmospheric noise, which is one of the main contaminants as shown in the SO and FYST combined case (Charmetant and Erler, 2023). Therefore, to provide a fair comparison of the AtLAST concept performance, a 25 m large dish telescope, with 250 000 detectors per band and lots of frequency bands to SO and FYST which are two smaller 6 m telescopes with a smaller frequency coverage, SO+FYST needs to be simulated without atmospheric noise. The results of simulating a SO and FYST combined kind of experiment without atmospheric noise are displayed in Fig. 5.12. In comparison to the AtLAST concept result (see Fig. C.5), SO+FYST does not achieve an unbiased detection of the tSZ power spectrum over the full ℓ-range, the total noise residual (r_{tot}) dominates on all scales. However, the CMB and cumulative Galactic foreground noise residual are smaller. The reason for these noises to be smaller is that the ILC is a trade-off, it focuses on trying to minimize the dominant noise and becomes less performant on the other sub-dominant noises. This is why, in the SO and FYST combined case, the CIB noise residual is much worse, and is the dominant noise on large scales ($\ell < 1000$). ON smaller scales ($\ell > 1000$) the total noise budget is entirely dominated by the instrumental noise. This makes sense as SO and FYST will have way fewer detectors and resolution than the AtLAST-model explored in this study (see Tab. 5.2), thus making the instrumental noise per frequency much larger (see Tab. 2.2 & 2.3). The larger beam also increases the noise on small scales in comparison to the AtLAST concept case.

Some complementary results are shown in Fig. C.5 & C.6. Especially the power spectra plots and

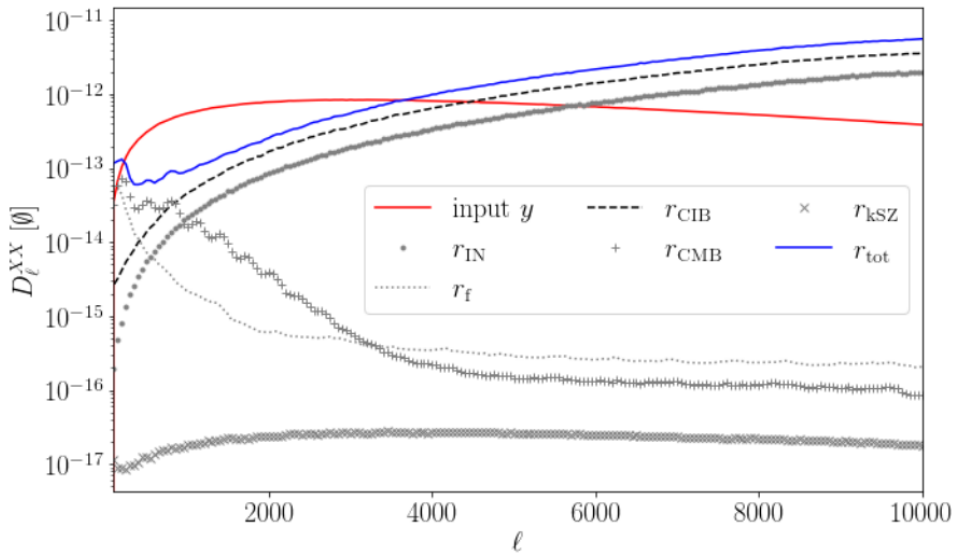

Figure 5.11: Power spectra of various components for an AtLAST-like experiment in Case 2 and S3 (see Tab. 5.2). In red is the input target tSZ signal, and in blue is the total ILC residual noise (r_{tot}). All the other curves are the decomposition of the total noise residual into the fractional contribution of each astrophysical emission in the sky. The dotted black curve is the CIB noise residual ($r_{\rm CIB}$) and all the other grey curves are the other component noise residuals. All spectra are beam deconvolved.

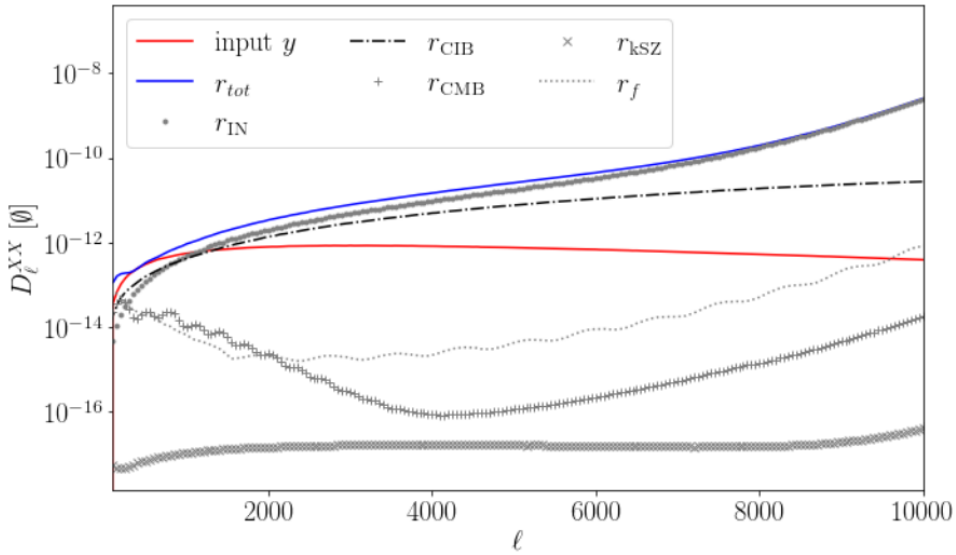

Figure 5.12: Power spectra of various components for an SO+FYST-like experiment in a Case 2 and S3 (see Tab. 5.1) . In red is the input target tSZ signal, and in blue is the total ILC residual noise (r_{tot}). All the other curves are the decomposition of the total noise residual into the fractional contribution of each astrophysical emission in the sky. The dotted black curve is the CIB noise residual ($r_{\rm CIB}$) and all the other grey curves are the other component noise residuals. All spectra are beam deconvolved.

decomposition into each separated component for Case 1 and Case 3.

5.5 Discussion and conclusion

In this chapter the recovery of the tSZ signal using an ILC or a CILC (see Sec. 2.5.1) was tested on different sky compositions simulated using the Skymodel (see Sec. 4).

5.5.1 Influence on separated components

The first result on a simple sky configuration (S1; see Tab. 5.1) was that combining SO and FYST does help reduce the total noise residual, i.e. the noise leftover by the contaminant in the recovered signal, by $\sim 16\%$. Moreover, when Galactic dust is added to the sky the advantage of combining SO and FYST compared to SO alone becomes higher, and the total noise residual is reduced by $\sim 22\%$ (CCAT-Prime Collaboration et al., 2023). These first results hinted at the fact that FYST is indeed helping lower the noise and is hinting that FYST could help to reduce efficiently not only the CIB noise residual. This is further probed and proved by the final result obtained on SO and FYST combined (see Fig. 9 of App. C.3) in the most realistic case (S4; see Tab. 5.1). Using and ILC to extract back the tSZ signal, it is found that SO and FYST combined reduces mostly the CIB noise residual by an average $\sim 35\%$ but also the CMB and kSZ noise residuals respectively by average $\sim 25\%$ and $\sim 15\%$. The cumulative Galactic foregrounds noise residual is also reduced by $\sim 10\%$ when SO and FYST are combined while the atmospheric noise is increased by $\sim 30\%$ by the addition of FYST to SO. The increase in the atmospheric noise contamination is due to FYST high-frequency coverage which is more contaminated by the atmosphere and has lower atmospheric transmission (see Fig. 2.10). Therefore it is clear that the FYST high-frequency channel mainly helps SO to reduce CIB contamination but also reduces the other noise contaminants. When a CILC is applied to deproject CIB using one modified blackbody SED (see Eq. 4.5) the improvements due to the addition of FYST to SO diminish for all components, CIB included except for the cumulative foregrounds noise residual. Probably due to the similar SED between the CIB and Galatic dust, which is anyway one of the main contributors to the Galactic foregrounds noise budget.

5.5.2 ILC and CILC weights

Information can also be extracted on the ILC and CILC weights with different sky ingredients. In all the sky configurations explored in this work and for all telescopes, SO, FYST, and a concept telescope such as AtLAST, the main conclusion is that the main contributors of the tSZ reconstruction favoured by the ILC and CILC are the frequency channels below the tSZ null ($\nu < 220$ GHz), meaning the frequencies probing the tSZ decrement, the negative part of the signal which is one of the kinds in the microwave sky. This finding is not surprising and was already reported in the literature. Whereas the frequency channels above ~ 400 GHz are in all cases never contributing to the tSZ reconstruction. In all cases, the variance of the weights among the 192 tessellated fields diminishes with increasing frequency when atmospheric noise is absent from the sky. This can be interpreted as the sky content being more and more similar between the fields as the frequency increases. This makes sense because the Galactic foreground emission becoming stronger with frequency (see Fig. 2.4) and starting to dominate the sky but also with the instrumental noise increasing. Comparing Fig. 5.1 of case S3 and Fig. 6.1 of case S4,

it is clear that the added presence of atmospheric noise in case S4 compared to case S3, is lowering the weight associated by the ILC do the 220GHz channel.

5.5.3 Power spectra

The comparison of the power spectra of Fig. 5.2 & 5.3 of case S3 to Fig. 5.6 of case S4 clearly shows that the addition of atmospheric noise to the sky increases the CMB noise residual by more than two orders of magnitude. This increase in the CMB residual is due to the ILC always trying to minimize the recovered map's variance and tailoring the weights to focus on reducing the main contaminant. When atmospheric noise is absent from the sky CMB is the main contaminant on large scales, the ILC weights can focus on reducing it. When atmospheric noise is added to the mock sky, it is the main contaminant, the ILC therefore targets this in priority, making the ILC weights less efficient at minimizing the CMB noise residual (Simons Observatory Collaboration, 2019). In all cases, CMB is the main noise along with atmospheric noise when present on large scales and CIB is the main noise residual on small scales. The instrumental noise residual can dominate the small-scale budget above the CIB noise residual in cases like SO and SO+FYST but now with low noise telescopes like the AtLAST concept idea. The kSZ and cumulative Galactive foreground noise are always subdominant by several orders of magnitude. With instrumental noise, CMB debiasing and CIB deprojection, telescopes such as SO and FYST combined should be able to measure an unbiased tSZ signal over a ℓ-window of $\ell \in [1800, 3500]$ (Charmetant and Erler, 2023). For a telescope such as the AtLAST concept, no predictions with atmospheric noise were made, but without it, the telescope achieved an unbiased detection of the tSZ power spectrum on large scales which is not the case for SO+FYST. The residual noise of the current AtlAST concept model is almost one order of magnitude below the signal, therefore a direct detection of the tSZ power spectrum is imaginable even with atmospheric noise.

CHAPTER 6

tSZ measurements and CIB leakage mitigation combining upcoming ground-based telescopes

Contribution

This chapter presents the works published in:

- **Charmetant M.** and Erler J, (2023), *Thermal Sunyaev-Zeldovich measurements and cosmic infrared background leakage mitigation combining upcoming ground-based telescopes*, Astronomy & Astrophysics, 677, A87 (2023).

The Python package (Skymodel) used to simulate the sky in this work (see Chap. 4) came from an original idea of J.E. and was coded by J.E. and M.C. The paper's content was written by M.C. valuable feedback on the manuscript content and J.E. provided one figure.

This part summarizes the results presented in (Charmetant and Erler, 2023) (see App. C.3). This presents realistic CIB leakage predictions, while Chap. 5 was only presenting results on a simplified version of the microwave sky for illustration of the concept and first-hand understanding of the behaviour of the pipeline. In this summary, is first addressed the information that can be extracted from the ILC and CILC weights. To assess the purity of the recovered tSZ map and the amount of CIB leaking, one has to look at the power spectra. Finally, a discussion about how each component of the microwave sky seems to influence the noise budget and the conclusion of the study.

6.1 Summary of the context and method

The thermal Sunyaev-Zeldovich (tSZ) effect can be a powerful probe of cluster characteristics and cosmology. However, it has the disadvantage of being a weak effect and of being correlated with the Cosmic Infrared Background (CIB). The CIB is a very strong contaminant at high frequency. Upcoming ground-based telescopes such as the Simons Observatory Large Aperture (SO LAT) and the Fred Young Submillimeter Telescope (FYST) should obtain unprecedented high-resolution low-noise maps measurements of the microwave sky (see sec. 2.2.2 and fig. 2.10). FYST covering the higher frequencies, where CIB dominates, it is expected that it can leverage and mitigate CIB contamination. To make predictions of detectability of the tSZ power spectrum and of the CIB leakage with these two upcoming

telescopes the Skymodel was used to generate realistic mock maps of the microwave sky as seen by these two telescopes. An ILC or CILC method, deprojecting one modified blackbody expression of the CIB is then used to recover the tSZ signal. The power spectra are then debiased from the CMB and instrumental noise contribution. The input tSZ power spectrum is compared to the ILC or CILC noise residual.

6.2 Summary of results

this part summarize all the results presented in Charmetant and Erler (2023) (see App. C.3).

6.2.1 ILC and CILC weights

As previously explained in Sec. 2.5.1, the ILC weights inform us on which frequency band contributes to the signal reconstruction and to the noise minimization, as those are the two sole constraints for their construction (see Eq. 2.13). Looking at the weights time the mixing vector ($\omega_i a_i$) informs on the fractional contribution of each channel to the reconstruction of the target signal. For the interpretation part, only the mean weight of the 192 is considered per frequency band, the variance is given as an error bar. In this study, we found that the 93 GHz and 145 GHz channels from SO LAT are the main contributors to the tSZ reconstruction when SO LAT is alone observing but also for SO LAT and FYST combined. The other frequency channels are mostly helpful for noise minimization (see Fig. 6 of App. C.3). One important thing to notice is that the 220 GHz and 225 GHz have a weight close to 0, despite being close to the tSZ null ($\sim$ 217 GHZ) and therefore supposed to be containing also no signal and purely noise, which should be of the utmost important to help the ILC target the noise and minimize it. This was the case in Sec. 2.5.1, but not here, because of the atmospheric red noise in the sky, undermining the usefulness of that frequency channel and of all the high-frequency channel bands. The conclusion is that SO LAT is the main driver of the tSZ reconstruction when SO LAT and FYST are combined. This is expected as only SO LAT has frequencies proving the decrement part of the tSZ spectrum (see Fig. 1.23) this is essential for signal targeting as this is the only known SED signal exhibiting this behaviour.

When using a CILC to deproject the CIB for its SED the modified blackbody given in Eq. 4.5, the weights are sensibly the same (see Fig. A.1 from App. C.3). This is consistent with Alonso et al. (2018) showing that the ILC does already a very good job at minimizing CIB, thus the CILC not making much of a difference. However, comparing Fig. 6.1 and Fig. 6 of App. C.3, it is clear that FYST alone produced ILC weights that had a smaller standard deviation among all the fields, meaning the noise was more controlled than for SO LAT. In the ILC case, Combining FYST with SO helped reduce the variance in the weights. In the CILC case, the standard variation is almost the same for SO LAT alone and with FYST added to it. This also goes in the direction of the CIB deprojection undermining the benefits FYST brings to SO LAT.

The maps of the ILC and CILC weights can be found in Appendix C.1.

6.2.2 ILC extracted power spectra

Looking at the power spectra of the total noise residual (see Eq. 2.32) and of the target input signal tells how much of the ILC-recovered tSZ power spectrum is noise-dominated or signal-dominated. The total noise residual can also be decomposed into the noise contribution of each contaminant to inform us of the bias those contaminants produce on the extracted tSZ power spectrum. Note that computing the contribution of the component to the noise budget is only possible with simulations because the

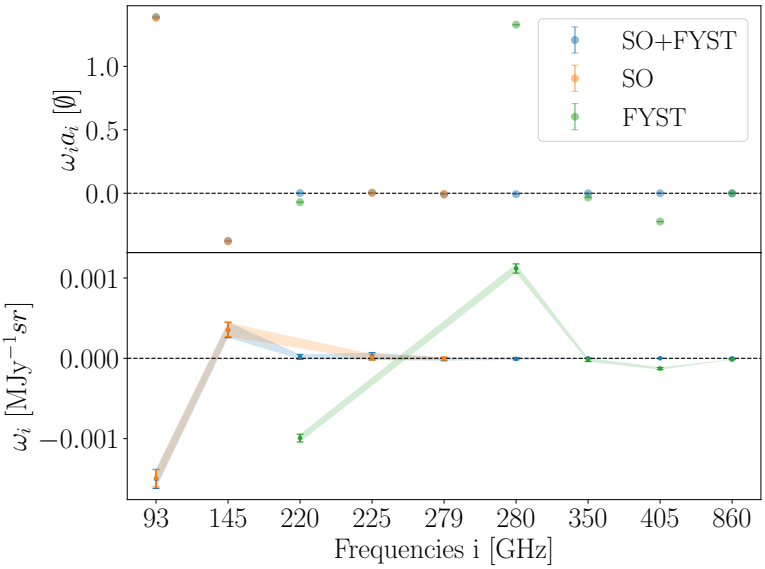

Figure 6.1: CILC weights for a case where the sky contains the emissions listed in Tab. 5.1 under 'real' and the modified blackbody SED of Eq. 4.5 was used to deproject CIB. The upper panel shows the fractional contribution of the weights to the tSZ reconstruction. The lower panel shows the values of the weights, meaning which frequency contributes to the reconstruction and noise minimization. The error bars and shared region represent the standard deviation of the mean weight computed among the 192 fields for each frequency. In orange for a SO LAT type of experiment, in blue for SO LAT+FYST and in green for FYST alone.

underlying true signal needs to be known. From the results displayed in Fig. 7 of App. C.3 can be deduced:

- the CIB noise residual ($r_{\rm CIB}$) is $\sim 35\%$ lower when SO LAT and FYST are combined rather than for SO LAT alone.
- for SO LAT the atmospheric red noise ($r_{\rm Atmo}$) dominates the noise budget on large scales up to $\ell \approx 2800$. On smaller scales, the noise budget is dominated by the CIB noise residual. The noise residual of other contaminants, the cumulative foreground (r_f), and kSZ are subdominant by more than one order of magnitude.
- for SO LAT+FYST combined, the atmospheric noise dominates the noise budget on all scales. This is not surprising, FYST is probing higher frequency channels, and the atmospheric transmission is lower than for SO (see Fig. 2.10) even at a higher site.
- looking at the ratio between the input tSZ power spectrum and the ILC noise residual, SO LAT achieves a better Signal-to-noise on large scales $\ell \in [500, 2800]$ and SO LAT and FYST combined better on small-scales $\ell \in [2800, 5000]$ due to the significantly lower CIB noise residual.
- for both SO LAT alone and SO LAT and FYST combine, a small window over which the signal and the noise are equivalent exists between $\ell \in [2100, 2700]$ but only SO LAT achieve a signal dominated extracted tSZ on a small $\ell \in [2100, 2400]$ window.

6.2.3 CILC extracted power spectra

A CILC was applied to deproject CIB using only one modified blackbody SED (see Eq. 4.5). In this case, the findings (see Fig. 8 of App. C.3) are the following:

- The CIB noise residual ($r_{\rm CIB}$) is further reduced compared to the ILC case, but because the ILC was already reducing CIB quite a lot, the further improvement is not that big.
- Because the deprojection does not manage to significantly reduce CIB, the improvement FYST was bringing to SO LAT in terms of CIB mitigation is reduced. There is only a 7% lower CIB noise residual ($r_{\rm CIB}$) for SO LAT and FYST combined compared to SO LAT alone.
- The atmospheric noise contamination increases a lot for SO LAT because of the additional noise penalty of the CIB deprojection. In the SO LAT and FYST cases, this can be mitigated because of the reduction in the CIB noise residual, lowering the atmospheric noise residual.
- In the CILC case, because of the CIB deprojection, the atmospheric noise residual is the dominant noise residual on all scales $\ell \in [500, 5000]$, the second main contaminant remains the CIB but only on small scales $\ell > 4500$ for both SO LAT and SO LAT+FYST. The noise astrophysical emissions are greatly subdominant.
- For SO LAT alone, the total noise residual (r_{tot}) dominates above the signal of interest on all scales. For SO LAT and FYST combined, the recovered tSZ spectrum is not noise-dominated over a very wide ℓ-window, $\ell \in [1700, 3400]$.

6.2.4 Overview

On Figure 9 of App. C.3 is a Violin plot showing the ratio between the residual noise power spectrum of each component of the microwave sky for SO LAT and FYST combined compared to SO LAT alone. A ratio of one would mean no improvement in adding FYST to SO LAT and a ratio below 1 indicates a lower noise residual for the component, when SO LAT and FYST are combined. The dispersion of the violin represents the spread of power spectrum values in function of ℓs. The results are:

- In the ILC case the component that is the most reduced by the combination of FYST and SO LAT is the CIB noise residual ($r_{\rm CIB}$), followed by the CMB ($r_{\rm CMB}$) and kSZ ($r_{\rm kSZ}$) noise residual. Which have similar noise residual values because of their SED, the ILC can not distinguish between the two. Then comes the cumulative noise residual (r_f) and finally the atmospheric noise ($r_{\rm Atmo}$) which is made worse by the addition of FYST to SO LAT in the ILC case because of the lower atmospheric transmission the FYST high frequencies (see Fig. 2.10).
- In the CILC case with CIB deprojected, the improvement FYST brings to SO LAT in terms of CMB and kSZ noise residual reduction is diminished, probably because of the noise penalty due to the additional constraint. This allows the CILC to better target and remove the other noise contaminants, lowering the cumulative noise residual (r_f) and atmospheric noise ($r_{\rm Atmo}$) in the SO LAT and FYST case.

6.3 Detailed contribution

This part presents the detailed contribution that has been made to this publication:

- M.C. coded the functions generating the tSZ, kSZ, CMB, and instrumental white noise maps in the Skymodel as well as the function relative to the pixel size in the maps. The other functions were coded by J.E.
- M.C. was responsible for the testing of the function under his supervision and J.E. for the one he coded.
- M.C. was in charge of hosting the skymodel package on her Github page, and of processing the updates.
- M.C. wrote half of the documentation PDF of the Skymodel, the other half was done by J.E. .
- The full Readthedoc webpage was written by M.C., who was also in charge of creating an example notebook for each of the functions of the Skymodel.
- M.C. coded the ILC and CILC codes under J.E. supervision.
- M.C. conducted the full scientific analysis of the paper. This included setting up the pipeline, testing it, and applying it to the Skymodel's most realistic sky. This was done under the active contribution of Kaustuv Basu and also J.E. and Frank Bertoldi.
- M.C. contributed to the scientific interpretation of the results under K.B.'s supervision and J.E.'s advice.
- M.C. wrote the full paper under K.B.'s supervision and J.E.'s and F.B.'s advice.

CHAPTER 7

A new estimator tSZ-kSZ^2

Contribution

In this chapter, MC investigated the possibility of detection of the kSZ signal through its cross-correlation with the tSZ using the tools developed in other chapters. The work presented in this chapter is still ongoing and the results are preliminary.

This chapter explores some preliminary results about a new estimator to detect the kSZ effect through its cross-correlation with the tSZ. First, the context of the kSZ studies is presented, through why this effect is interesting to study in the first place and the detection status. Then, the method to explain the estimator construction, especially why this cross-correlation and the squaring. Finally, some preliminary results are commented on, results based on some mock sky generated using the Skymodel and extracted signals using ILC/CILC codes.

7.1 Introduction

The kSZ effect is another flavour of the SZ effects that can be of interest for observational cosmology but also for clusters Astrophysics. However, due to its weak amplitude and spectral shape, it is difficult to detect.

7.1.1 Context of the sutdy

The kinematic Sunyaev-Zeldovich effect (Sunyaev and Zeldovich, 1969, 1970) occurs when the CMB photons scatter on the cluster's electron population, moving compared to the CMB rest frame. Both the tSZ and kSZ effect occur at the same time and the energy change of CMB photon due to the presence of galaxy clusters can be expressed as :

$$\frac{\Delta I_\nu}{I_0} = (y f(x_\nu) - y_{kSZ}) h(x_\nu) \tag{7.1}$$

However, the two effects probe different electron populations, the tSZ effect probes the hot baryons inside galaxy clusters while the kSZ effect probes all the baryons that are undergoing the bulk motion of the cluster. They also differ in terms of amplitude $\sim 100\mu$K for the tSZ and $\sim 10\mu$K for the kSZ which

is an order of magnitude smaller. Because of its weak amplitude and its spectral shape similar to the CMB which are ~5 orders of magnitude stronger in amplitude (Mroczkowski et al., 2019), the kSZ effect is difficult to detect. It was detected for the first time by Hand and et al. (2012) using galaxies pairs momentums called "pairwise momentum estimator". It was first detected on individual clusters by Sayers and et al. (2013). A more complete summary of the methods and detections of the kSZ effect can be found in Table 1 of Planck collaboration et al. (2018). A cross-correlation between the two effects is expected because they are spatially located in galaxy clusters and partially generated by the same electron population (some electrons are both hot and undergoing the cluster's motion).

7.1.2 Interest in studying kSZ

The kSZ effect is a powerful probe to measure clusters' peculiar velocities (Zhang et al., 2008; Hurier and Tchernin, 2017), the growth rate of structures (Sugiyama et al., 2017) and is a powerful probe of the baryons content (Kusiak et al., 2021; Vavagiakis and et al., 2021) in the Universe because contrary to the tSZ effect it probes all the electron population, not only the hot ones. This makes the kSZ effect an interesting probe of potential baryonic matter on the outskirts of clusters to address the missing baryons problem Hernández-Monteagudo and et al. (2015). Therefore, having ways to efficiently detect the kSZ effect is crucial to address all of these questions. A new estimator to detect the kSZ power spectrum is proposed in this project. Through the cross-correlation between the kSZ and the tSZ effect. Because of the higher tSZ amplitude and well-known SED, it is expected that the cross-correlation can pick up the kSZ effect. This new estimator is motivated by the upcoming ground-based telescopes such as SO and FYST (see Sec. 2) which will provide unprecedented clean tSZ and kSZ maps due to their location and sensitivities.

7.2 Sky simulations specifics

For this project, the Websky 2019 maps are used (see Tab. 4.4). However, early on some potential problems were discovered. First about the shape of the kSZ power spectrum and second about the treatment of the kSZ in the Websky simulations.

7.2.1 Exploring the Websky 2019 kSZ power spectrum

The kSZ power spectrum from the Websky simulations exhibits a strong peak on the large scales ($\ell \in [0, 400]$) (see Fig. 7.1 & 4.9). This is inconsistent with the modelization of the expected kSZ power spectrum from the literature (Shaw et al., 2012). This peak in the kSZ power spectrum was described by the Websky team as a Doppler effect from the velocity fluctuations (see Fig. 5 of Hu and Dodelson (2002)). The up-turn around $\ell \sim 7000$ is supposedly due to small-scale discretization effects. This raised some discussions that were never fully solved, considering that the power at $\ell \sim 12000$ is more than double the power at $\ell \sim 2000$. The Websky team suggested not to trust the maps above $\ell \sim 5000$. Moreover, if there are some discretization effects in the simulations above a certain scale, increasing the power, we would expect it to be the same for all the emissions. This up-rising of the power on small scales is however not seen in the tSZ map provided by the Websky simulations.

To investigate this matter further, the 2018 Websky maps were used because their kSZ map is separated into two components, the kSZ field and the kSZ halo (see Tab. 4.3). The kSZ field component is the kSZ signal coming from the gas outside haloes moving compared to the CMB rest frame and on which the

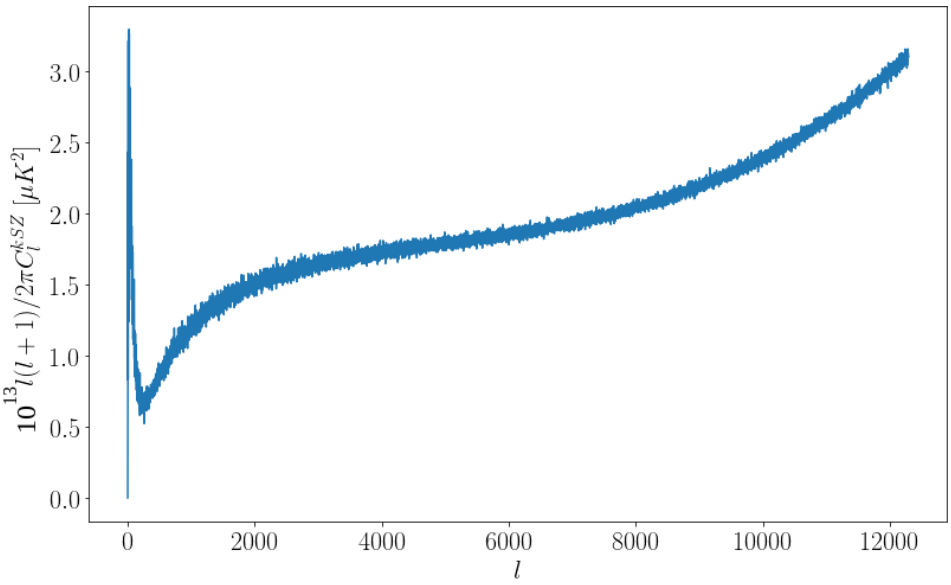

Figure 7.1: kSZ power spectrum extracted from the Websky 2019 map (see Sec. 4.4). This is a raw unbinned power spectrum, which is why there are fluctuations.

CMB photons scatter. The kSZ halo is the SZ signal occurring in galaxy clusters, because of the bulk motion of the gas bound in haloes. The maps for this Websky release have an $N_{\rm side} = 2048$ ($N_{\rm pix} = 1.6'$). The kSZ field, also sometimes called kSZ 'patchy' and kSZ halo maps already shows the difference between the two components of the kSZ effect (see Fig. 7.2). The kSZ halo is much more clustered a similar, visually, to a tSZ map than the kSZ patchy component. Their power spectrum is very different too (see Fig. 7.3). The kSZ field power spectrum dominates by on average a factor of 2 above the kSZ halo power spectrum (Stein et al., 2020). This is counter-intuitive, one would expect most of the electrons to be bounded in clusters and therefore, the kSZ halo to dominate over the field component. Moreover, previous studies such as Flender et al. (2016), showed that the halo term was dominating above the field term. The Websky team was contacted about this and raised two main justifications for this:

- The generalized Navarro-Frenk-White profile (see Eq. 1.45) fit is performed by Websky only on Haloes $M_h > 10^{13} M_\odot$. In particular, they assumed that those profiles were possibly inaccurate for the kSZ effect on haloes of Mass $10^{12} M_\odot < M_h < 10^{13} M_\odot$. Haloes of this mass are non-negligible contributors to the kSZ effect and were excluded from the "halo" contribution. Instead, they let the peak-patch algorithm paint some profiles on them and those haloes ended up being included in the "field" part of the kSZ signal. However Flender et al. (2016) also has a minimum halo mass of $M_h > 10^{13} M_\odot$ for the kSZ effect and still has a dominant kSZ halo component.

- The redshift range up to which the kSZ is computed. In Flender et al. (2016) the kSZ power is computed only up to z<1 at such redshift, the proportion of haloes above $M_h > 10^{13} M_\odot$, is large, thus inflating the halo term. The Websky computes their kSZ power up to z$\sim$ 4.5, at that redshift, there are few to no haloes for this mass, and therefore the field component dominates. The Websky kSZ power is the total integrated power up to redshift z$\sim$ 4.5, and the field component is the main component of the signal.

The difference, between the kSZ halo and field importance between their results and the literature, is due to the definition and limits or what is part of the 'field' component and what is part of the 'halo' along with how far it looked back in the history of the Universe.

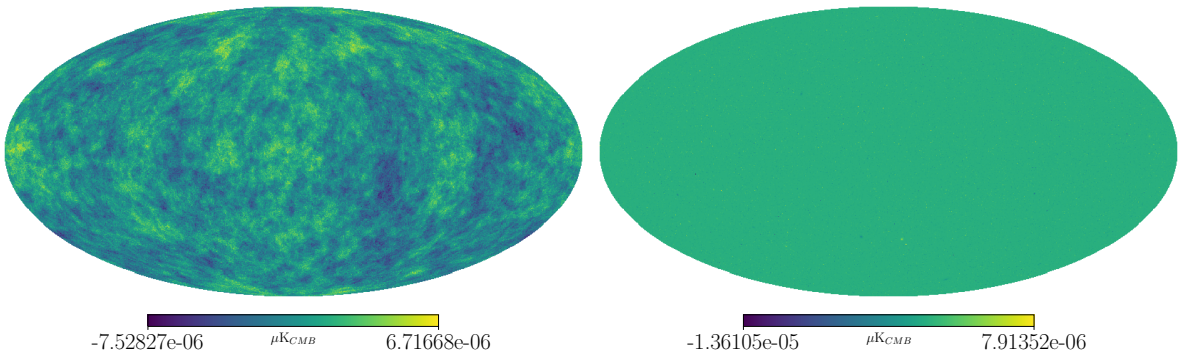

Figure 7.2: the HEALPix maps of the kSZ field component on the left and the kSZ halo component on the right. From the Websky simulations (see Tab. 4.3), with $N_{\text{side}} = 2048$.

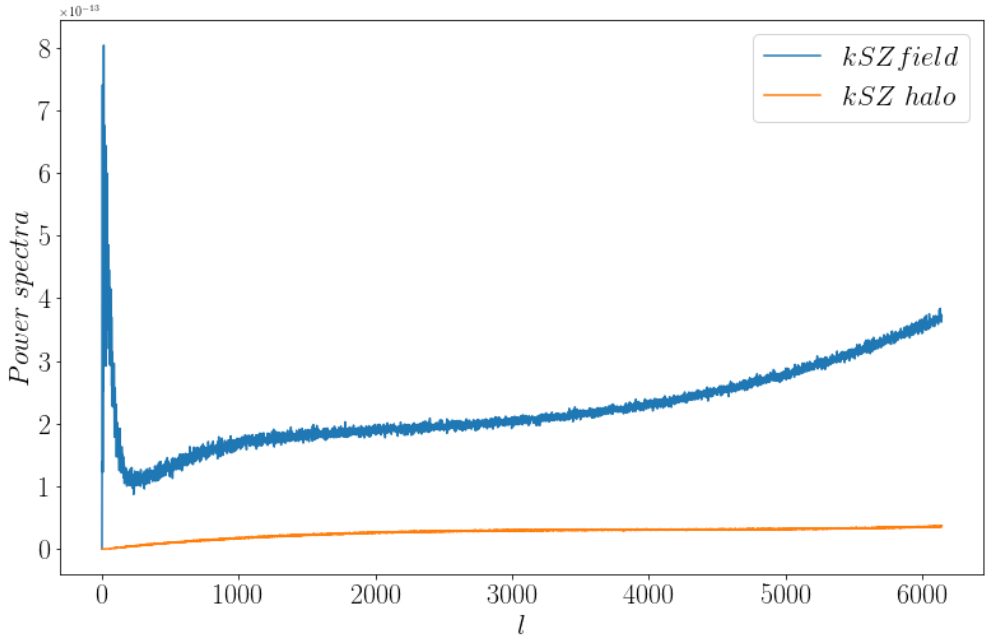

Figure 7.3: Power spectra of the kSZ field, in blue, and kSZ halo component in orange. From the Websky simulations kSZ maps (see Tab. 4.3), with $N_{\text{side}} = 2048$.

7.3 Method

The method applied in this project is the following, the Skymodel (see Sec. 4) is used to generate mock maps of the microwave sky at several frequencies and containing the desired components (see Tab. 5.1). Then, an ILC is run on the set of multi-frequency maps to extract back a map of the tSZ. A CILC is run on the same multi-frequency maps to extract back a tSZ-free CMB map, using the modified black body spectrum SED (see Eq. 1.32). A kSZ map can not be directly extracted because its spectral shape is the same as the CMB and because the CMB is so much stronger in amplitude the CILC would anyway end up extracting a CMB map. The CMB map extracted by the CILC, contains the CMB signal, the kSZ signal, and residual noises from the other components present in the Sky. Those residual noises are the reasons why a CILC needs to apply deprojecting tSZ and not a simple ILC. An ILC tuned in to extract back a CMB map would still contain some residual tSZ, which would create some spurious correlation with the ILC-extracted tSZ map. The goal is to measure the purest possible correlation between tSZ and kSZ, if the kSZ signal contains some tSZ residual then we are systematically biasing the result. Because the CMB and kSZ are the same SED they can not be separated by a simple ILC. However, they do have a different angular power spectrum, especially the CMB power spectrum dominates on low ℓ and the kSZ one on high-ℓ. A high-pass filter can therefore be constructed to filter out the CILC-extracted CMB map containing the kSZ signal.

7.3.1 Estimator tSZ-kSZ^2

The kSZ effect can be positive and negative, therefore, the cross-correlation will tSZ will be null. Indeed on average, the two effects will be positively correlated and negatively correlated, giving 0. A solution is to look at the cross-correlation between tSZ and kSZ^2. A similar idea was exploited through the cross-correlation between kSZ^2 and the galaxy overdensity δ_g (Hill et al., 2016; Kusiak et al., 2021; Plante et al., 2022). The goal of this project is to see if a tSZxkSZ^2 cross-correlation could be detected by upcoming ground-based telescopes like SO LAT or SO LAT and FYST combined.

Null-test problem

It is important to check that the cross-correlation between tSZ and kSZ is indeed null. To verify the theory, and ensure that no spurious correlation is present before the squaring. Using the Websky 2019 (see Tab. 4.4) mocks Compton-y map and $y_{\rm kSZ}$ map the cross-correlation can be computed with the HEALPy function `healpy.sphtfunc.anafast()`. The result of the cross-correlation is presented in Fig. 7.4. It seems that the cross-correlation between tSZ and kSZ exhibits a strong, non-null, negative component between $\ell \in [0, 3000]$. This result is surprising. Considering the scale extent this can not be attributed to the strong "Doppler" term present in the Websky simulations. This excess in the tSZxkSZ ($C_\ell^{y\,y_{\rm kSZ}}$) could be originating from some spurious correlation present in the Websky simulations themselves. The Sehgal (Sehgal et al., 2010a) simulations Compton-y and $y_{\rm kSZ}$ map are selected to check if this residual signal persists. Fig. 7.5 shows that this non-null cross-correlation is also present in the Sehgal simulations, to an even greater extent because this time, the signal is non-zero and negative up to $\ell = 10000$. To confirm that this negative signal is not a random event, 100 random realizations of the Sehgal kSZ maps were created. This is done by taking the Sehgal kSZ map and using HEALPy's function `healpy.sphtfunc.anafast()` to compute its power spectrum. Then using the computer power spectrum and the HEALPy function `healpy.sphtfunc.synfast()` a new kSZ

map is generated. This technique however only generates random realizations of the Gaussian part of the kSZ effect, the non-Gaussian part is left out and not reproduced. In Fig. 7.5 it is clear that the cross-correlation between the Compton-y map and the randomized kSZ map is 0 on all scales. More tests or exploration would be necessary to give a definite answer to this counterintuitive problem but it seems that tSZxkSZ ($C_\ell^{y y_{\rm kSZ}}$) does exhibit a negative correlation independently of which based simulation is used. This cross-correlation is present in the non-gaussian component of the kSZ effect.

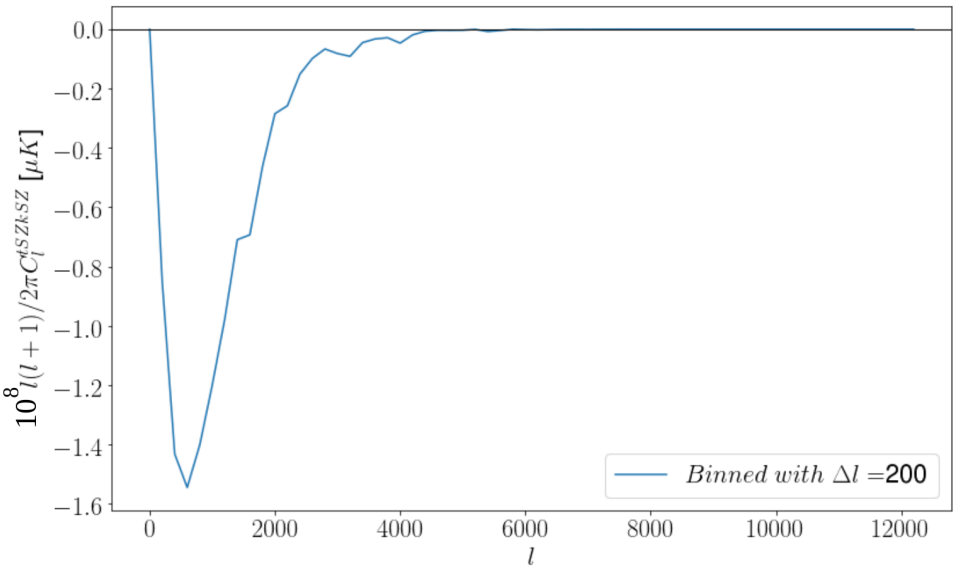

Figure 7.4: Cross-correlation between the Websky 2019 mocks (see Tab. 4.4) Compton-y map and $y_{\rm kSZ}$ map. The resulting cross-correlation is binned with a ℓ-window of $\Delta\ell = 200$.

7.3.2 Wiener filtering

One can separate CMB and kSZ using harmonic-space information instead of their frequency spectra because they do not have the same power spectrum. To get back only the kSZ signal one can use a Wiener filter (W_ℓ) (Wiener, 1949) which is equivalent to a harmonic space high-pass filter. It is constructed as follows. Suppose S and N are respectively the signal and the noise contaminating your signal. The power spectra of those two signals are altered in our map by the telescope beam b_ℓ, therefore,

$$W_\ell = \frac{C_\ell^S}{C_\ell^{S+N}} b_\ell. \tag{7.2}$$

The signal of interest in this study is the kSZ effect, ideally, its power spectrum should be constructed using its analytical shape but because its true shape is still uncertain and for simplicity, we did not use the analytical power spectrum but instead the binned power spectrum of our input signal model. This means, the spectrum given by the Websky simulations. The power spectrum is then interpolated to get back all the missing ℓ values, lost because of the binning. A consequence of using the input kSZ to construct the Wiener filter is that the Wiener filter is very efficient at targeting the kSZ signal. In reality, because we do not know the true power spectrum of the kSZ effect, the Wiener filter would have lower performances.

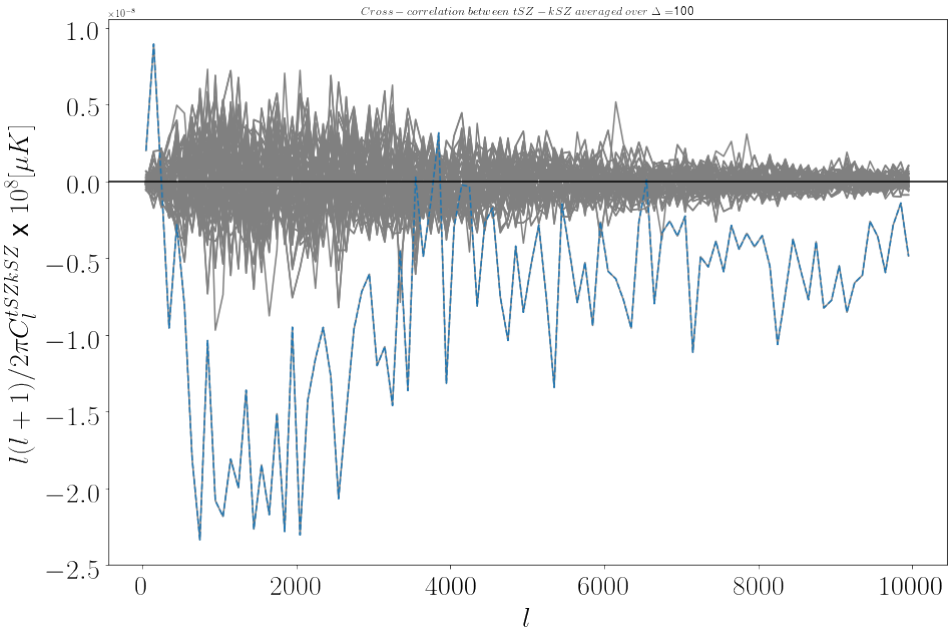

Figure 7.5: Cross-correlation between the Sehgal (see Tab. 4.1) Compton-y map and $y_{\rm kSZ}$ map, in blue. In grey, the cross-correlation between the Sehgal Compton-y map and a 100 random realization of the kSZ map. The resulting cross-correlation is binned with a ℓ-window of $\Delta\ell = 100$.

The noise component must include all the contaminants present in the CILC map, this means the CMB signal and the cumulative residual CILC noise coming from all the other unwanted astrophysical emissions present in the mock sky but also instrumental white noise. For the CMB signal, the lensed CMB power spectrum given by CAMB[1] (Lewis and Challinor, 2011) was used, this code takes as inputs a given set of cosmological parameters and solves Boltzmann equation to compute the CMB power spectrum. On the web interface, the following default cosmological parameters were used $H_0 = 70 \mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$, $T_{\rm CMB} = 2.7255$K, $\Omega_b h^2 = 0.0226$, $\Omega_c h^2 = 0.112$, $\Omega_\nu h^2 = 0.00064$ where $h = H_0/100 \mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$ and $n_s = 0.96$. The CILC residual noise can be computed by taking the difference between the CILC-extracted map and the pure signal the CILC was supposed to recover (see Eq. 2.32). Note that the residual noise power spectrum has to be beam corrected (i.e. multiplied by the inverse of the telescope beam squared) to compensate for the smoothing to the lowest channel resolution done by the CILC and be comparable to the signal power spectrum. The telescope which with the sky is observed has a finite beam size b_ℓ which is defined as

$$b_\ell = e^{-\frac{1}{2}\ell(\ell+1)\frac{\theta^2}{8\ln 2}} \tag{7.3}$$

Where θ is the FWHM of the Gaussian kernel in radians and ℓ is the multipole moment. This needs to be taken into account in the Wiener filter so that the filter can target the right signal spectrum. The filter is then normalized by dividing all the values by the maximum amplitude of the Filter see Figure 7.6.

[1] https://lambda.gsfc.nasa.gov/toolbox/camb_online.html

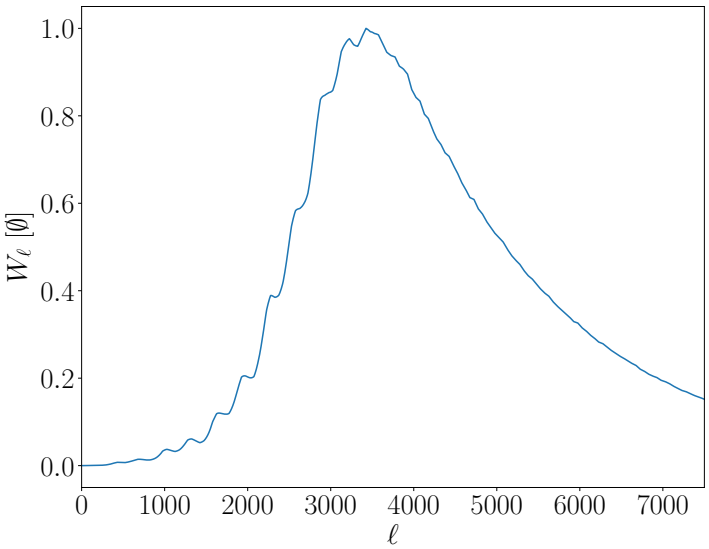

Figure 7.6: Wiener filter used to recover the kSZ effect power spectrum. On the large scales we are dominated by the inverse CMB power spectrum, we clearly see the CMB accoustic peaks and the filter goes to 0. It is maximum around $\ell \sim 3500$ and then drop on smaller scales because of the telescope beam.

$$\hat{s}_{kSZ}(p) = W_\ell \hat{s}_{CILC}(p) \tag{7.4}$$

One should note that the Wiener filter introduces a bias in our recovered kSZ map ($\hat{s}_{kSZ}$) because by definition $W_\ell < 1$, thus we can not recover C_ℓ^S except if the noise (N) is null (Delabrouille and Cardoso, 2007). The Wiener filtered signal can be debiased (Tegmark and Efstathiou, 1996) but this is not done here for simplicity. This filter is routinely used to separate kSZ and CMB, in particular in studies of the cross-correlation between kSZ and the galaxy overdensity (Ferraro et al., 2016; La Plante et al., 2022). To get the estimated kSZ signal ($\hat{s}_{kSZ}$) map one simply multiplies the CILC extracted map containing CMB, kSZ, and residual noises ($\hat{s}_{\rm CILC}$) by the Wiener filter.

Simplified scenario

A simplified version of the Wiener filter can be considered, in the first approximation, in a case where the sky maps do not have any intrinsic beam. Equation 7.2 remains, only without the b_ℓ term. The shape of the Wiener filter is in this case, purely a high-pass filter (see Fig. 7.7)

7.4 Results

Before using the Skymodel to generate a set of multi-frequency sky maps containing all the astrophysical emissions (see Tab. 5.1), some tests are done on simplified sky models where only a few components are present and no ILC/CILC is applied. Later on SO LAT and FYST-like telescopes are reproduced and a realistic prediction is made for these upcoming ground-based telescopes.

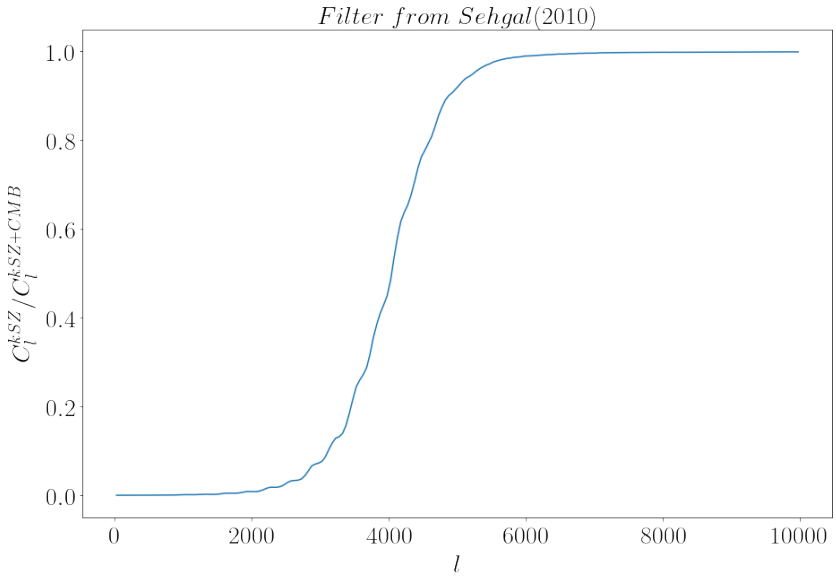

Figure 7.7: Wiener filter without any beam component. Example made using the Sehgal et al. (2010a) template maps, in a case where the only noise contaminant is the CMB.

7.4.1 Simplified sky scenario

In this section, the sky is considered to be only composed of tSZ, CMB, kSZ, and eventually instrumental white noise. All the results are based on the SO simulations template maps (see Tab. 4.2). The results presented are built step by step, going from the simplest to more realistic scenarios. The first thing to assess is whether detection of the tSZxkSZ^2 cross-correlation spectrum could be possible, this is done by comparing the signal ($C_\ell^{y y^2_{\rm kSZ}}$) to the most basic real-sky case, which CMB as the only noise source. Independently of any other Astrophysical emissions, all tests will always boil down to a cross-correlation between tSZ and CMB-contaminated kSZ (tSZx(kSZ+CMB)^2). Because ILC and CILC methods can not separate CMB and kSZ due to their similar SEDs. Figure 7.8 shows that no detection of the spectrum is possible without any filtering, especially on large scales. The sudden drop of the signal around $\ell \approx 7500$ is due to the SO resolution downgrade (see Sec. 4.3.2). The uncertainties shrinking as the scales decrease (ℓ increase) is due to the presence of less CMB contamination as the scales decrease. Therefore less noise contamination is present.

To achieve detection of the cross-correlation spectrum, a Wiener filter (see for example Fig. 7.7) is applied on the SO map containing the kSZ and unlensed CMB. As of now, no beam is taken into consideration. After filtering, a high detection of the cross-correlation spectrum seems to be possible (see Fig. 7.9). This result seems very promising, but one should not forget that this is the most ideal case. Indeed, that would be for a sky composed only of tSZ,kSZ, and CMB. Without any instrumental noise, beam limitation, atmospheric noise, or contamination from other Astrophysical components. This time the 1σ confidence interval is composed of the cross-correlation between the tSZ map from the SO simulations and Wiener-filtered random realization of the SO kSZ and unlensed CMB maps, all of that squared. This is not compared to the cross-correlation of the true cross-correlation between tSZ and kSZ^2 anymore but between the tSZ and Wiener-filtered kSZ, all of that squared. The $C_\ell^{y F(y_{\rm kSZ})^2}$

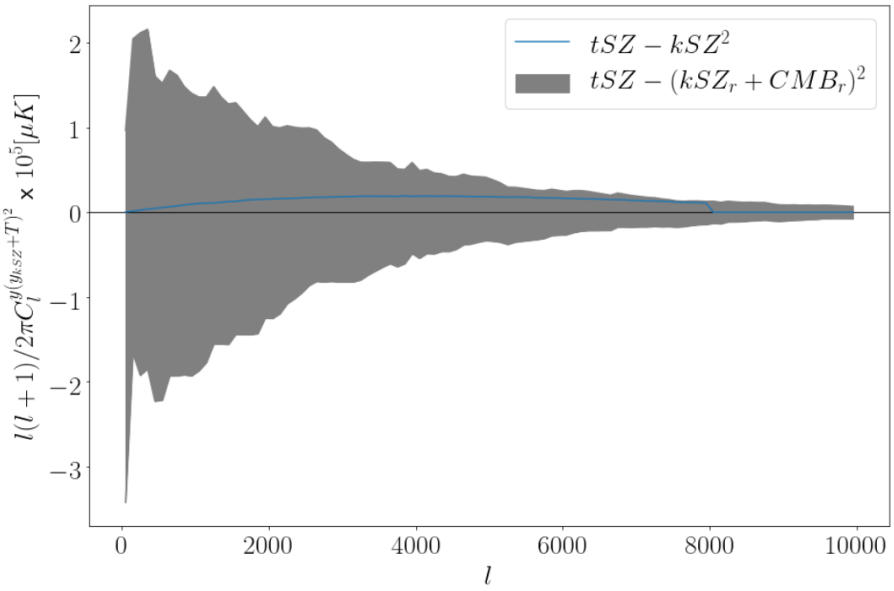

Figure 7.8: **Preliminary result:** In blue the true $C_\ell^{yy^2_{\rm kSZ}}$ signal. In gray 1σ interval of 100 cross-correlations between the Compton-y map and random Gaussian realizations of the kSZ and the unlensed CMB. The SO template maps are used (see Tab. 4.2) and the spectra are binned with $\Delta\ell = 100$.

cross-correlation is a fairer comparison. In reality, the pure kSZ signal is not known. Therefore, the signal that will be recovered and picked up is the result of the cross-correlation between tSZ and the filtered kSZ signal.

To make this test case more realistic, instrumental White Noise (WN) can be added to the SO sky map. Of course, this additional noise term is taken into account in the denominator of the Wiener filter. As a first approximation, the white noise is considered to be 7 μK−arcmin. In comparison to Fig. 7.9, the 1σ confidence interval is wider, but a high detection of the spectrum is still possible (see Fig. 7.10). One can notice that the 1σ confidence interval is widening as the scales diminishes (ℓ increases), this is opposite to the behaviour seen in Fig. 7.8. This can be explained by the presence of the Wiener filtering, which lowers the noise components between $\ell \in [0, 3800]$ and returns 1 above $\ell > 5000$. The result can be made a bit more realistic by adding the telescope beam to the picture. At first, a one 1' beam is considered. This corresponds approximatively to the SO LAT and FYST resolutions (see Tab. 2.2 & 2.3). Adding the beam results in a drop of the Wiener filter on small scales (see Fig. 7.6 in comparison to Fig. 7.7). Adding the beam has two effects, first, it lowers the signal amplitude, second, it makes the signal drop on small scales, especially after $\ell > 6000$. A high detection of the spectrum should still be possible (see Fig. 7.11). Even though this case tries to emulate an SO LAT or FYST-like experiment and exhibit a very strong detection of the cross-correlation spectrum it is still very simplified as the sky does not contain any contaminant from the other Astrophysical emissions (see Tab. 5.1). However, this simple framework can still give us a first estimation of whether its spectrum is possibly detectable or not by a given experiment. This exact framework is used to simulate a *Planck*-like an experiment, taking into account an instrumental white noise at a level of 33 μK−arcmin. In this case, a $\sim 2\sigma$ detection of the cross-correlation spectrum seems possible (see Fig. 7.10) but again, one should not forget that the beam and numerous astrophysical emissions are missing.

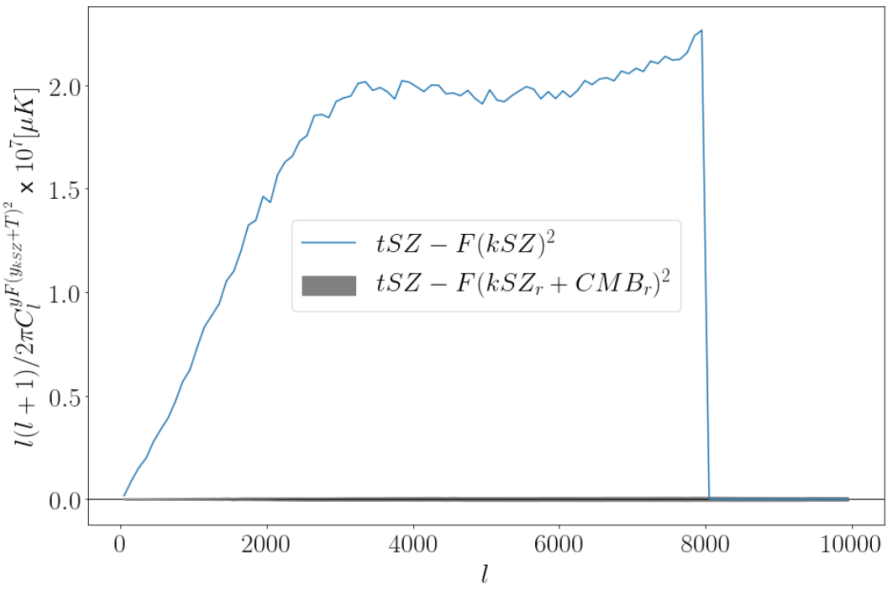

Figure 7.9: **Preliminary result:** In blue the true $C_\ell^{yF(y_{\rm kSZ})^2}$ signal, the cross-correlation between the tSZ and the Wiener-filtered kSZ. In grey 1σ interval of 100 cross-correlations between the Compton-y map and random Gaussian realizations of the kSZ and the unlensed CMB after filtering. The SO template maps are used (see Tab. 4.2) and the spectra are binned with $\Delta\ell = 100$.

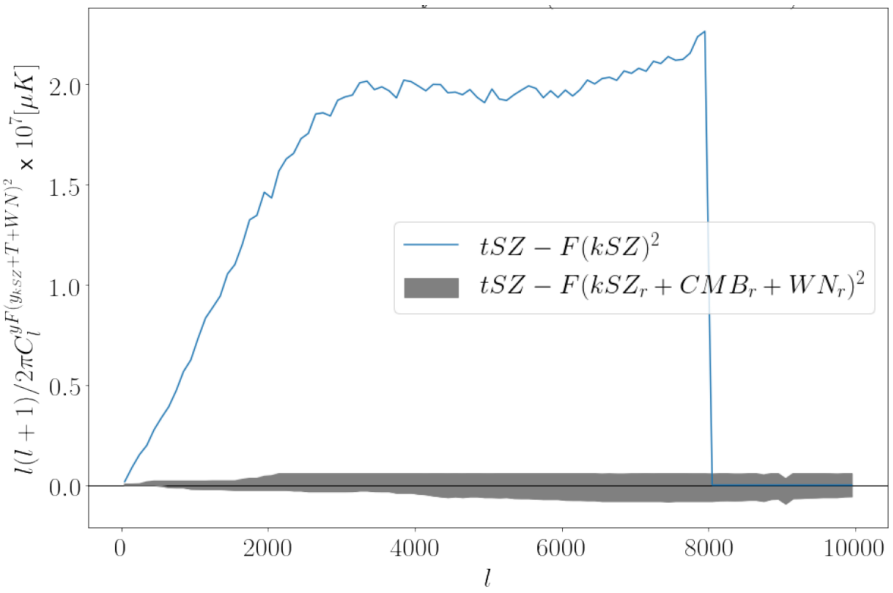

Figure 7.10: **Preliminary result:** In blue the true $C_\ell^{yF(y_{\rm kSZ})^2}$ signal, the cross-correlation between the tSZ and the Wiener-filtered kSZ. In grey 1σ interval of 100 cross-correlations between the Compton-y map and random Gaussian realizations of the kSZ, instrumental white noise at a level of 7 μK−arcmin and the unlensed CMB after filtering. The SO template maps are used (see Tab. 4.2) and the spectra are binned with $\Delta\ell = 100$.

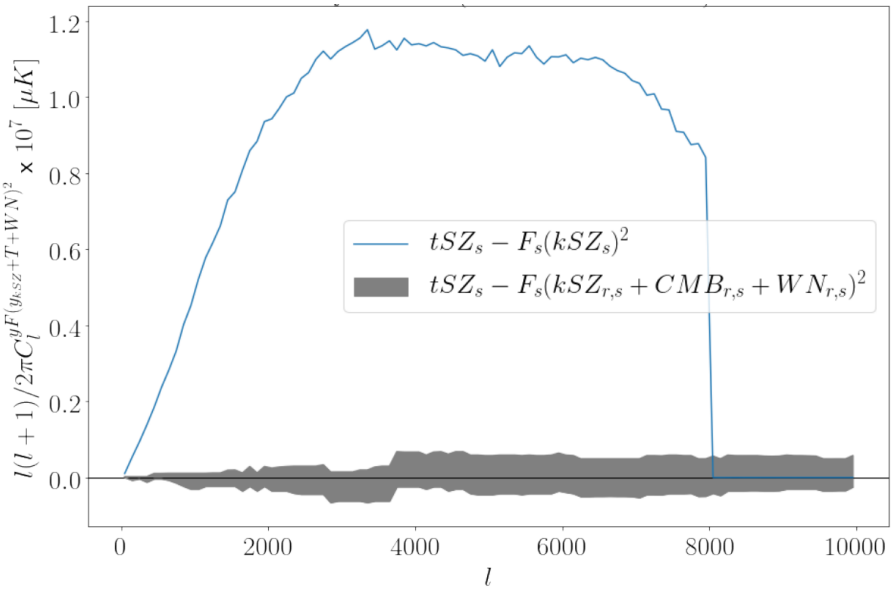

Figure 7.11: **Preliminary result:** In blue the true $C_\ell^{yF(y_{\rm kSZ})^2}$ signal, the cross-correlation between the tSZ and the Wiener-filtered kSZ. In grey 1σ interval of 100 cross-correlations between the Compton-y map and random Gaussian realizations of the kSZ, instrumental white noise at a level of 7 μK−arcmin and the unlensed CMB after filtering. A beam of 1' is applied to the maps. The SO template maps are used (see Tab. 4.2) and the spectra are binned with $\Delta\ell = 100$.

7.5 Conclusion

The goal of this project was to make predictions of detectability of the cross-correlation spectrum for a new estimator, made of the cross-correlation between tSZ and kSZ^2 ($C_\ell^{yy^2_{\rm kSZ}}$). The kSZ effect is of interest for both clusters Astrophysics and cosmology, as it can be used to measure the peculiar velocities of galaxy clusters, the optical depth of the gas, or constrain cosmological parameters such as the growth of structure or the reionization duration. However, the kSZ effect is small, it is one order of magnitude below the tSZ effect in amplitude and because of its similar SED to the CMB can not be separated from it using frequency-dependent information. Thus the interest in developing a new estimator and attempting to pick up the kSZ signal through its cross-correlation with the tSZ. The squaring of the kSZ signal seems essential, as the kSZ effect can be positive or negative, on average, its cross-correlation with the tSZ should be 0. However, with our simulations templates, Websky 2019 and SO (see Sec. 4), a significant, non-zero cross-correlation is found. This can not be explained only by the particularity of the kSZ power spectrum in the Websky simulations (see Fig. 7.1) but could still originate from some spurious correlations in the simulations. Some detection tests are realized using the SO simulations and building on the sky by starting with only a few contaminants and complexifying the sky step by step. The first test consists of checking whether statistical detection can be achieved in the most simple case when only the tSZ, kSZ, and CMB are present in the sky. Because of the strong CMB signal, no detection of the cross-correlation can be achieved (see Fig. 7.8). To separate kSZ from the CMB a Wiener filter is used. In this case, a high-significance detection of the cross-correlation would be possible (see Fig. 7.9).To make the sky more realistic, instrumental white noise can be included with various noise levels,

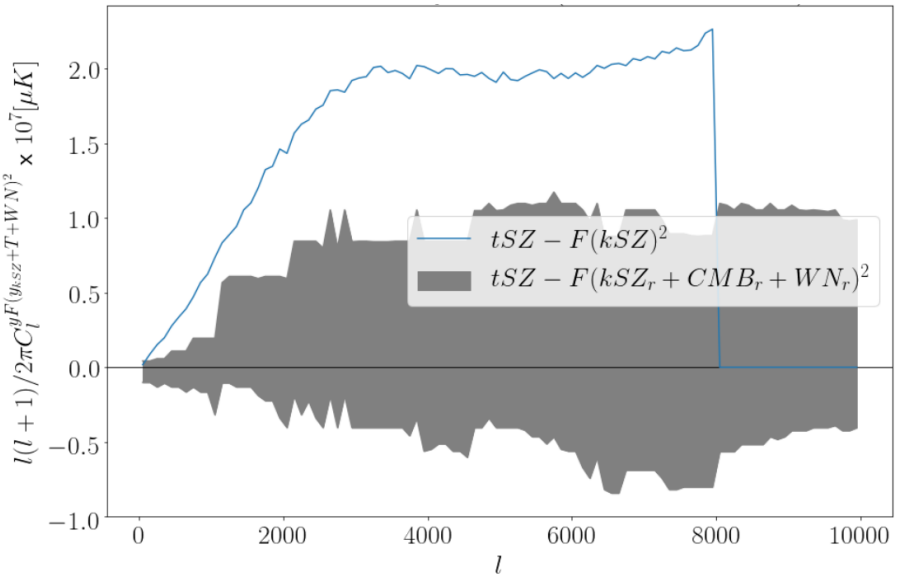

Figure 7.12: **Preliminary result:** In blue the true $C_\ell^{yF(y_{\rm kSZ})^2}$ signal, the cross-correlation between the tSZ and the Wiener-filtered kSZ. In grey 1σ interval of 100 cross-correlations between the Compton-y map and random Gaussian realizations of the kSZ, instrumental white noise at a level of 33 μK−arcmin and the unlensed CMB after filtering. A beam of 1' is applied to the maps. The SO template maps are used (see Tab. 4.2) and the spectra are binned with $\Delta\ell = 100$.

the higher the noise, the wider the confidence interval, and the lower the detection significance (see Fig. 7.10 & 7.12). Including a large telescope's beam reduces the signal amplitude and makes the signal drop on small scales (see Fig. 7.11).

In a very simplified sky, it seems that a high-significance detection of the cross-correlation could be achieved.

CHAPTER 8

Conclusion

This chapter summarizes the main findings of this research work, and their significance, and discusses the results and the possible outlooks. This work is included in the field of observational cosmology, and in particular in CMB and galaxy clusters science. The goal of this work was to make predictions of the detectability of tSZ and kSZ effects with upcoming high-resolution ground-based telescopes such as SO LAT and FYST. It was segmented into three projects. The simulations of the microwave sky, the CIB leakage into ILC-recovered tSZ maps, and a new estimator for the kSZ signal.

8.1 Simulations of the microwave sky

The first project (see Sec. 4) was about creating realistic and versatile high-resolution sky simulations of microwave sky emissions to reproduce the most realistic sky possible as seen by SO LAT and FYST.

8.1.1 Summary of the project

The context of the project was the fabrication of two new CMB experiments, the SO LAT and the FYST. Because they have not started observing yet, to make a prediction or their scientific outputs in the hope of securing funding or to make the right choice in terms of characteristics of the telescopes, for example, number of detectors, frequency bands, simulations of how they would observe the sky is essential. At the time of the project, some template maps of the various extra-galactic emissions of the microwave sky existed. The Sehgal et al. (2010a), Websky (Stein et al., 2018, 2020) and the SO (Simons Observatory Collaboration, 2019) were among those available templates and are the ones used in this work. For the Galactic emission, a Python package based on measurements and models called PySM (Thorne et al., 2017) existed. However, this was insufficient, because the extra-galactic templates were provided at frequencies that did not correspond to the planned frequencies of observation of FYST and PySM was too low resolution. The *Planck* team had one software tool called the Planck Sky Model (PSM) that was able to simulate any component of the microwave sky at any frequency, sensitivity, or beam size of a given instrument but was not publically available. Therefore, as members of the CCAT collaboration, it was essential to develop an easy-to-use Python package that would simulate the microwave sky. The state-of-the-art simulations, such as SO, Websky, and Sehgal et al. (2010a) were used for the extra-galactic templates, simply interpolating or rescaling the existing fixed frequency maps to any frequency desired. For the Galactic components, the PySM package was used, and the resolution

was improved through revitalisation and smoothing to avoid pixelisation artefacts. The characteristics of the telescope, such as its beam, sensitivity, and atmospheric noise (Simons Observatory Collaboration, 2019; Choi et al., 2020). The units between the different templates are also made uniform in the code, to ensure that there is no mixing.

8.1.2 Results and outlooks

The Python package created in this work called Skymodel (see Sec. 4), can be used through one single function to generate mock simulations of the microwave sky based on state-of-the-art existing simulation templates. The user can choose which astrophysical emission to include, at with frequency to simulate them within $\nu \in [27, 860]$ GHz. But also with which beam, unit, sensitivity, resolution, and atmospheric noise. The Skymodel code ensures in the background that the templates, units, and resolutions match between all the components. The Skymodel package is publicly available on GitHub[1] and can be used to generate mock maps of the sky as seen by FYST, but also for concept ideas such as AtLAST. It was used to make predictions of CIB contamination into the tSZ signal for FYST and SO LAT (CCAT-prime Collaboration et al., 2021; Charmetant and Erler, 2023). The Skymodel package could be further improved by including the simulations of polarized emissions from the Microwave sky. This is crucial as the next big step of CMB science is about the detection of the CMB polarisation modes, which could constrain inflation models and be a new probe for the standard model of cosmology. The Skymodel could also be improved by generalizing its input, at the moment it can only deal with HEALPix formats maps, but for precision cluster simulation on small scales, or to include hydrodynamical simulations such as Magneticum (Dolag et al., 2009) or Millenium (Springel et al., 2001). In the future, it could also be interesting to include Time Ordered Data (TOD) which is the real format in which astrophysical signal is measured, as an intensity in function of time, and refine the atmospheric noise modeling to reproduce atmospheric stripping. These additions would make the Skymodel an all-around tool for CMB science simulations. However, within the time of this study, the PSM was publicly released with an HTML webpage that allows simulation of any components of the microwave sky, at any frequency, sensitivity beam size, whether in intensity or polarisation, and either based on data or modelling. A new version of PySM called PySM3 (Zonca et al., 2021) can be seen as an in-between the Skymodel and the PSM. Contrary to its previous version it also includes the extra-galactic components based on the Websky simulations but does not offer the versatility of the Skymodel to choose other simulations such as Sehgal et al. (2010a) or SO. However, it includes polarization which is not the case for the Skymodel. The PySM Python implementation is also quite different from the Skymodel and in certain cases, one order of magnitude slower, because it always simulates the full complexity of the microwave sky before returning to one particular configuration. Overall, it is good to dispose of several tools to cross-check that no mistakes are being passed on from the simulations to the physical predictions or interpretations.

8.2 CIB leakage mitigation into tSZ

Chapter 5 explored if CIB leakage into ILC or CILC extracted tSZ maps could be mitigated by combining SO and FYST.

[1] https://github.com/MaudeCharmetant/skymodel

8.2.1 Summary of the project

The tSZ signal is an important tool for cluster astrophysics to measure cluster characteristics such as pressure, and temperature or to detect shocks. It can also be used to constrain cosmological parameters such as Ω_m and σ_8. It is important to retrieve the purest tSZ signal possible to avoid biasing astrophysical measurements or cosmological parameters deduced from the signal. The measurements of the tSZ signal can be contaminated by various other astrophysical components, in particular on low frequencies ($\nu < 30$ GHz) by Radio Point Sources (RPS) which are located in galaxy clusters, and therefore spatially correlated with the tSZ. However, in this study, such low frequencies are not used and therefore the contamination should remain minimal. The second contaminant, especially of the tSZ power spectrum is the CMB, because of its high amplitude, but it is a contaminant only on large scales and its power spectrum is well-known. Because of the high resolution of new-generation telescopes, the tSZ signal will be able to be investigated on small scales for the very first time. However, on those small scales and at high-frequencies the dominant contaminant of the tSZ is the CIB signal. The CIB signal is particularly difficult to mitigate because CIB originates from a population of stars at various redshifts, and with various properties. Therefore, the CIB emission over the sky can not be summarized by one SED, but by a sum or modified blackbody SED, that differs with frequency and redshift, making it difficult for an ILC or CILC to fully suppress it. The CIB is also located in galaxy clusters, where star formation and dust absorption happens, and has a non-zero correlation with the tSZ effect. However new high-resolution and high-frequencies telescopes such as FYST, will probe the frequencies where the CIB dominates and therefore could help mitigate the contamination better than ever before. In this chapter, the Skymodel code based on the Websky templates, was used to reproduce mock maps of the microwave sky for an SO LAT and FYST-like experiment. An ILC or CILC deprojecting CIB with one assumed SED is used to extract back the tSZ signal from the multi-frequency, multi-component sky. The power spectrum of the residual CIB noise left over by the ILC or CILC is then investigated.

8.2.2 Results and outlook

This was tested in several sky configurations, at first in a very simple case, where only the signal tSZ and some noise component CIB, instrumental noise were present. This showed that, while using a simple ILC to retrieve the tSZ signal, the addition of FYST to SO LAT would reduce the CIB noise residual by $\sim 16\%$. Hinting that the high-frequencies are indeed helping to reduce CIB contamination. When Galactic dust contamination is added to the sky, the addition of FYST to SO LAT brings an even bigger advantage and reduces the CIB noise residual power spectrum by $\sim 25\%$ (CCAT-prime Collaboration et al., 2021). Adding all the other astrophysical emissions in the sky but leaving out the atmospheric noise at first yields the following results. The ILC weights show that the main contributors to the tSZ signal reconstruction are the frequency channels probing the tSZ decrement ($\nu < 220$ GHz), i.e., the 93 and 145 GHz channels of the OS LAT. However, the is still a small contribution from the FYST 280, 350 and 405 GHz. Especially visible is the importance of FYST 220 and secondly SO LAT 225 GHz in the noise minimisation. Indeed, the maps at those frequencies contain almost no tSZ signal, mainly the noise components, which is very handy for the ILC to better target and remove those noises. When looking at the power spectra, the CIB noise residual improvement with the addition of FYST remains around $\sim 25\%$, showing that the other astrophysical contaminants, such as CMB or kSZ do not influence much the variation of the CIB mitigation. When looking at the potential signal-to-noise (S/N) ratio, i.e. the input true underlying tSZ power spectrum compared to the total ILC noise residual power spectrum, the

S/N is increased by a factor of 4 when adding FYST to SO. When atmospheric noise is added to the sky, the tSZ decrement channels remain the main contributor to the tSZ reconstruction, but this time the frequencies probing the null of the tSZ SED do not contribute much to the noise minimization because of the high-atmospheric contamination. This contamination also makes the CMB noise residual blow up by 3 orders of magnitude, becoming a very strong contaminant to the tSZ signal on large scales. One way to circumvent that is to debias the extracted tSZ power spectrum from the CMB noise residual. In this case, even though the addition of FYST to SO LAT reduces the CIB noise residual by $\sim 35\%$, the overall S/N of FYST and SO LAT combined is better than SO LAT alone only on large scales. When using a CILC to deproject one CIB SED instead of the ILC, the weights remained unchanged, and this time the advantage of adding FYST to SO is reduced to $\sim 7\%$ because of the deprojection, however, it is enough for the ILC to better minimize the other noises. In the SO LAT and FYST combined case, the tSZ power spectrum should be detectable with an S/N above 1 on a large ℓ-window ($\ell \in [1700, 3400]$). Using the exact same framework, the Skymodel generates maps of the sky but no atmospheric noise. Once again, the tSZ decrement channel, especially $\nu < 150$ GHz is the most useful to the ILC to recover the tSZ signal. Adding more frequency channels does decrease the total noise residual but not by more than a factor of 2. A concept idea such as AtLAST, with its low sensitivity and high resolution should be able to detect the tSZ power spectrum with an S/N above 1 on scales $\ell \in [100, 2500]$ without debiasing.

Those results could be improved by using more evolved ILC techniques that do not require smoothing down all the maps to the lowest resolution, such as NILC or GNILC, This would probably highlight even more the slight advantage FYST has on SO in terms of resolution. The treatment of the Atmosphere could also be improved, the model used in this work is fairly simple, and in reality, the atmospheric noise might be of lower amplitude but more stripy due to the scanning pattern. The treatment of CIB could also be improved, especially since several different SEDs could be tested for deprojection or even a combination of more than one SED. However, this would result in a higher noise penalty due to the additional constraint.

8.3 A new estimator for the kSZ signal

In chapter 7, a new estimator for the kSZ signal is constructed and explored in a few simplified cases.

8.3.1 Summary of the project

The kSZ effect can be a probe for the velocity field and optical depth of galaxy clusters and a probe for cosmology, through the large-scale velocity field, or its dependence on the duration of the reionization epoch. It also has the advantage of probing all baryons moving compared to the CMB rest frame and not only the hot ones, contrary to the tSZ. This makes the kSZ a potential probe for the missing baryons. However, its weak amplitude and its spectral shape identical to the CMB makes it difficult to detect it. Previous studies investigated the cross-correlation between the kSZ and the galaxy density, hoping to pick up the kSZ signal through its cross-correlation with a stronger amplitude signal. A similar tracer of galaxies and galaxy clusters is the tSZ effect. However, the average correlation between the tSZ and kSZ should be null, because the kSZ effect can be correlated or anti-correlated with the tSZ effect depending on the velocity of the cluster. To avoid this cancellation, the kSZ signal was squared before computing its cross-correlation with the tSZ. To make predictions of the detectability of $C_\ell^{y y^2_{\rm kSZ}}$ with upcoming experiments, the Skymodel is used and based on the SO simulations to generate mock maps

of the microwave sky. In the first approximation was considered a sky containing only tSZ, kSZ, CMB, and instrumental white noise. A CILC is used to extract a kSZ and CMB-free tSZ map, another CILC is then used to extract a tSZ-free CMB+kSZ map. Indeed ILC methods can not separate CMB and kSZ, a Wiener filter is then used on the extracted map to obtain a kSZ-only map.

8.3.2 Results and outlook

The results are that in this very simple framework, detection of the cross-correlation spectrum of several σ could be achieved for an experiment such as SO. By adding more contaminants in the sky one by one, it is clear that higher noise increases the noise 1σ exclusion zone. However, it seems that the main killer of the signal is the beam of the telescope, which reduces the amplitude of the signal by a factor of almost two. These results should be made more realistic by including the other astrophysical emissions in the sky. Some techniques to unbiased the returns of the Wiener filter should also be explored.

8.4 Outlook of this work

This work aimed at the modelization of a realistic microwave sky and using it to make predictions of detectability and purity of the tSZ and kSZ spectra. Detecting unbiased measurements of the power spectrum of these two effects is crucial to measure galaxy cluster properties and to use them as probes for cosmology. This thesis work was realized to predict the capability of new upcoming ground-based telescopes such as SO LAT and FYST. This work hints at promising detection for both the tSZ and kSZ spectra with those experiments. The CIB contamination into the tSZ effect should dominate only on very small scales $\ell > 45000$ and be well-mitigated on the larger scales. This work highlights the importance of combining experiment data to get stronger constraints and better minimize the noise. It also highlights the importance of probing the tSZ decrement for tSZ signal reconstruction. Atmospheric noise contamination is a big player in undermining signal detection and should be mitigated as much as possible. Without the presence of atmospheric noise, this work shows that high-frequency channels help reduce the CIB contamination but not by orders of magnitude but rather by a factor up to two or less.

Appendices

APPENDIX A

Other technical details

A.1 Down smoothing of the ILC maps

This part explains in detail the formalism behind the smoothing of the ILC maps to the resolution of the lowest resolution map in the case where the map is a cutout or an array that does not have the HEALPix format and therefore can not be treated using the HEALPy function `healpy.sphtfunc.smoothing()`. A Gaussian function represents the beam of the telescope (see Fig. A.1), passing over all the pixels of the array

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}, \tag{A.1}$$

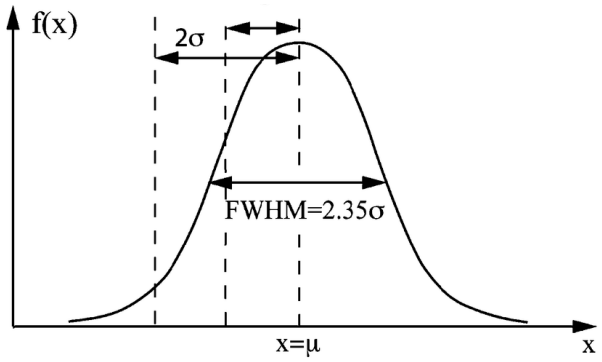

Figure A.1: Gaussian function $f(x)$, with mean μ and standard deviation σ and a the Full Width Half Maximum (FWHM) of FWHM=$2\sqrt{2\ln 2}\sigma \approx 2.35\sigma$. Extracted from Tavernier (2010).

where $f(x)$ is the Gaussian kernel applied on a pixel x, σ the standard deviation and μ the mean of the pixels. To make the Gaussian wider or thinner, the Full Width Half Maximum (FWHM) has to be modified. This can be done by multiplying with a Gaussian of the same mean but a different standard deviation. Suppose the desired Gaussian has a standard deviation of σ_w, which is the sum of the standard deviation of the Gaussian representing the resolution of the array σ_o and of another unknown Gaussian σ_u, which when added to our original resolution Gaussian gives us the wanted new Gaussian. We have

$\sigma_w^2 = \sigma_o^2 + \sigma_u^2 \Leftrightarrow \sigma_w = \sqrt{\sigma_o^2 + \sigma_u^2}$. This is equivalent, to multiplying the original Gaussian, with a Gaussian of the same mean but a different standard deviation

$$f(x) \approx e^{-\frac{1}{2}\frac{(x-\mu)^2}{\zeta}}, \tag{A.2}$$

where

$$\zeta = \frac{\mathrm{FWHM}}{2\sqrt{2\ln 2}}\frac{1}{\mathrm{pix}_{\mathrm{size}}} = \frac{\sqrt{b_o^2 - b_w^2}}{2\sqrt{2\ln 2}}\frac{1}{\mathrm{pix}_{\mathrm{size}}}, \tag{A.3}$$

with b_o the value in arcmin of the original resolution of the array and b_w the new wanted resolution in arcmin and $\mathrm{pix}_{\mathrm{size}}$ the pixel size of the array also in arcmin. The normalization by the pixel size is essential, as the pixels in the map have a finite resolution defined by the griding of the data.

APPENDIX B

Simulations

B.1 Magneticum hydrodynamical simulations

Magneticum simulations are hydrodynamical simulations that consider many processes, such as radiative cooling, star formation, heating by Ultraviolet background, and feedback. Simulations can follow the evolution of metals throughout the history of star formation. However, this does not prevent the overcooling problem: too many stars are produced in the centre of galaxies because the feedback implemented is insufficient to eject the gas and metals. Radiative cooling must be taken into account because this cools down the baryons at the centre of haloes, which alters the haloes' shape through adiabatic contraction (Gnedin et al., 2004). The Magneticum simulations are produced using GADGET-2 code (Springel et al., 2001; Springel et al., 2005). The substructures in the dark matter are found using the code SUBFIND (Dolag et al., 2009). Here is how they proceed

1. A Friend-of-Friend (FoF) algorithm is applied to identify the parent haloes. FoF simply looks at how close structures are given a threshold distance.

2. The density of each particle in a FoF group is estimated by looking at the saddles points compared to a threshold.

3. The FoF is applied only to DM particles to avoid systematically biasing the baryonic mass content of haloes. Because the baryonic matter does not trace the DM content. Each baryon is then associated with its closet DM particle. The mass of baryons is then estimated separately.

B.1.1 Boxes parameters

The Magneticum simulations are composed of several boxes of different sizes and characteristics (see Tab. B.1 & B.2).

B.1.2 Light cone production

Each box of the Magneticum simulations can be used to produce light. The principle is that the box is selected, put at different redshifts, and a random seed determines where the light starts in the first box. The content of each box is then projected to create the lightcone. Because the sizes of the boxes do not

	Box0	Box1	Box2b	Box2	Box3	Box4	Box5
Size [Mpc/h]	2688	896	640	352	128	48	18
mr	2×4536^3	2×1526^3		2×594^3	2×216^3	2×81^3	
hr				2×1584^3	2×576^3	2×216^3	2×81^3
uhr						2×5763^3	2×2163^3

Table B.1: Magneticum boxes name, sizes, and resolutions, medium resolution (mr), high resolution (hr), ultra-high resolution (uhr). Extracted from http://www.magneticum.org/.

	m_{DM} $[M_\odot h^{-1}]$	m_{gas} $[M_\odot h^{-1}]$
mr	1.3×10^{10}	2.6×10^9
hr	6.8×10^8	1.4×10^8
uhr	3.6×10^7	7.3×10^6

Table B.2: Particles mass in $M_\odot h^{-1}$, for either Dark Matter (m_{DM}) or the baryonic gas (m_{gas}) in function of the box resolution. Extracted from http://www.magneticum.org/.

match, between the different redshifts, some correlation problems could occur at the level of the patch-up between the boxes. To avoid that, it is important to have a lot of boxes, at very close redshift to minimize the size difference between the boxes (see Fig.B.1). On the middle panel, the size of the boxes is very different because of the big redshift jump between them. In the right figure, the correlation problems should be minimized. Taking a different seed will produce another entirely different light uncorrelated with the first one. However, depending on the size of the box, only a limited number of uncorrelated lightcone can be produced.

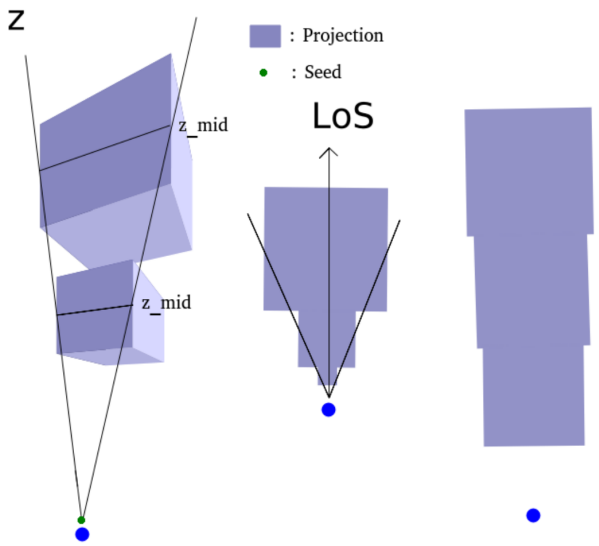

Figure B.1: Illustration of how Magneticum generates lightcones using the boxes.

B.1.3 Maps returned by Magneticum

The Magneticum simulations can return various physical information from light cones containing the projected data of a given box,

- 3D or 2D electron density.
- Mass or Emission weighted temperature.
- Volume, age, or X-ray Surface brightness of particles.
- Compton-y map.
- y_{kSZ} map.
- Q and U stocks parameters for transverse particle motions.
- Mass weighted velocity or velocity squared.
- Faraday rotation map.
- Pasquale weighted temperature.
- Emission temperature.
- Mass weighted metallicity for various components, either everything above >Fe, or C, N, O, Mg, Si, and Fe.
- Rees-Sciama effect (Rees and Sciama, 1968) by mass-weighted temperature motion.
- Mach number map, either weighted by mass, energy, or volume.
- 3D or 2D dark matter density
- Y_X parameter either mass or Pasquale weigthed.
- Deflection angle of gravitational lensing.
- Projected momentum from the Rees-Sciama effect either by Fast Fourier Transform (FFT) or by integral.

Below are a few examples of tSZ (see Fig. B.3, B.2) and kSZ maps (see Fig. B.5, B.4) returned by Magneticum.

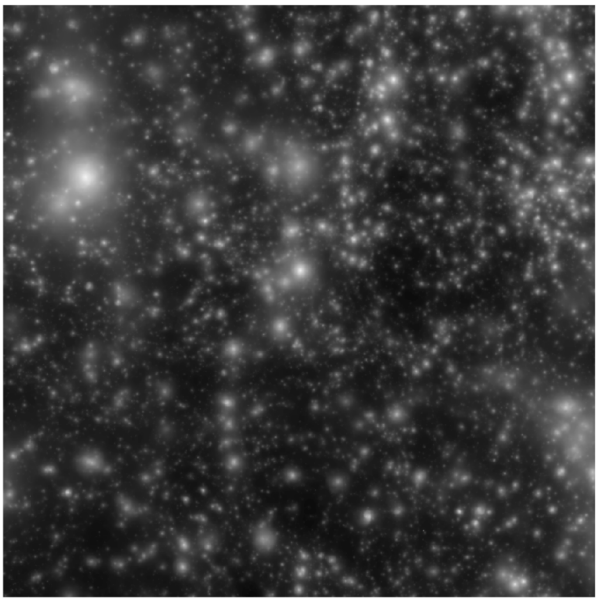

Figure B.2: Compton-y map from the lightcone generated in Box3 (see Tab. B.1). $7° \times 7°$ patch with a number of pixel $N_{\rm pix} = 1024$.

Figure B.3: Compton-y map from the lightcone generated in Box0 (see Tab. B.1). $40° \times 40°$ patch with a number of pixel $N_{\rm pix} = 1024$.

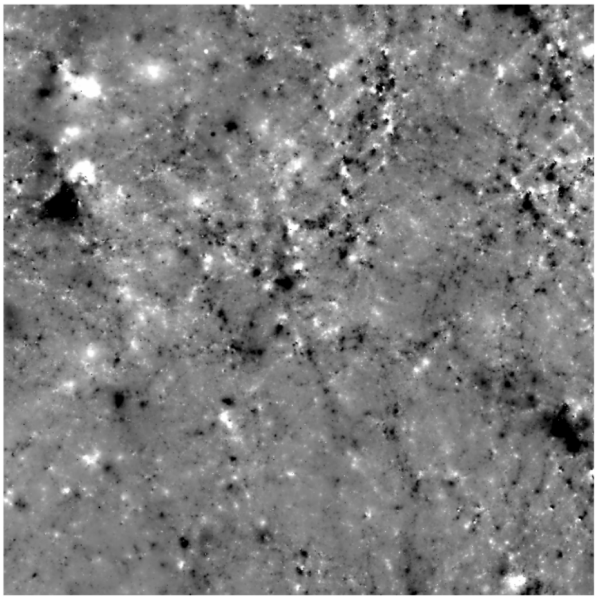

Figure B.4: $y_{\rm kSZ}$ map from the lightcone generated in Box3 (see Tab. B.1). $7° \times 7°$ patch with a number of pixel $N_{\rm pix} = 4096$.

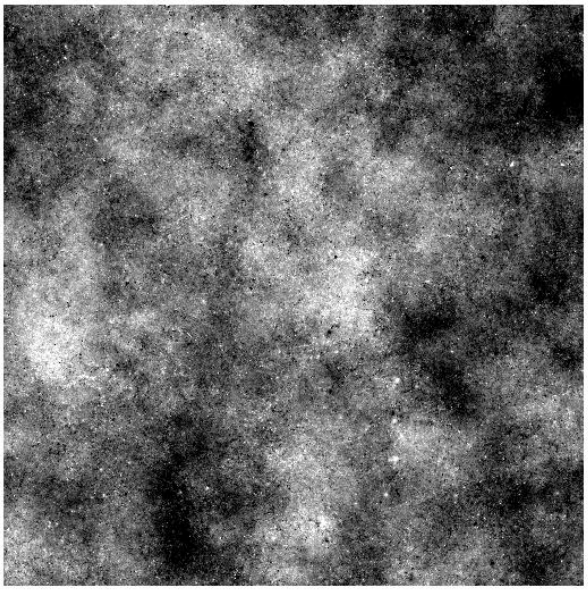

Figure B.5: $y_{\rm kSZ}$ map from the lightcone generated in Box0 (see Tab. B.1). $40° \times 40°$ patch with a number of pixel $N_{\rm pix} = 1024$.

APPENDIX C

Exploring Cosmic Microwave Background leakage into Thermal Sunyaev-Zeldovich maps

C.1 CILC weights maps

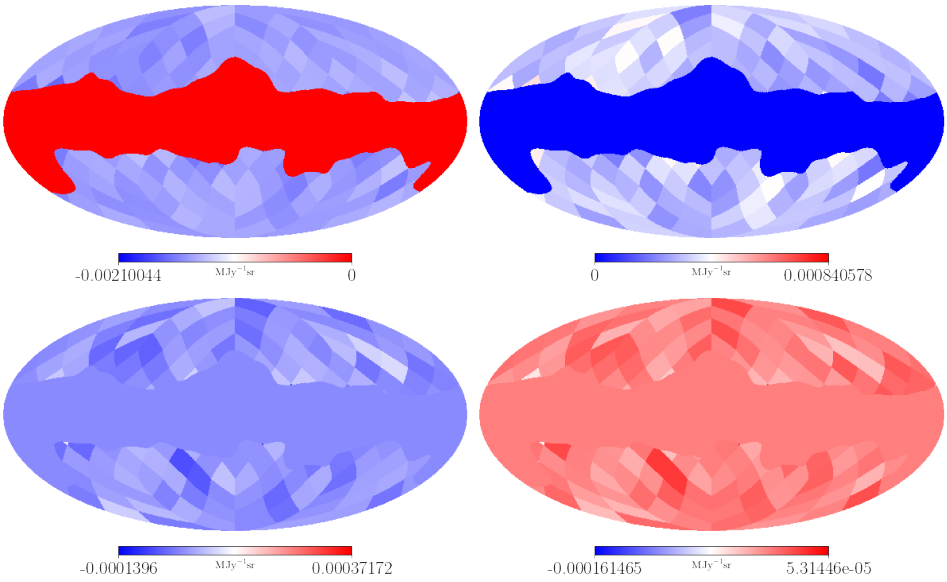

Figure C.1: The CILC weights map for the SO LAT frequencies. From left to right and up to down, 93 GHz, 145 GHz, 225 GHz, 279 GHz.

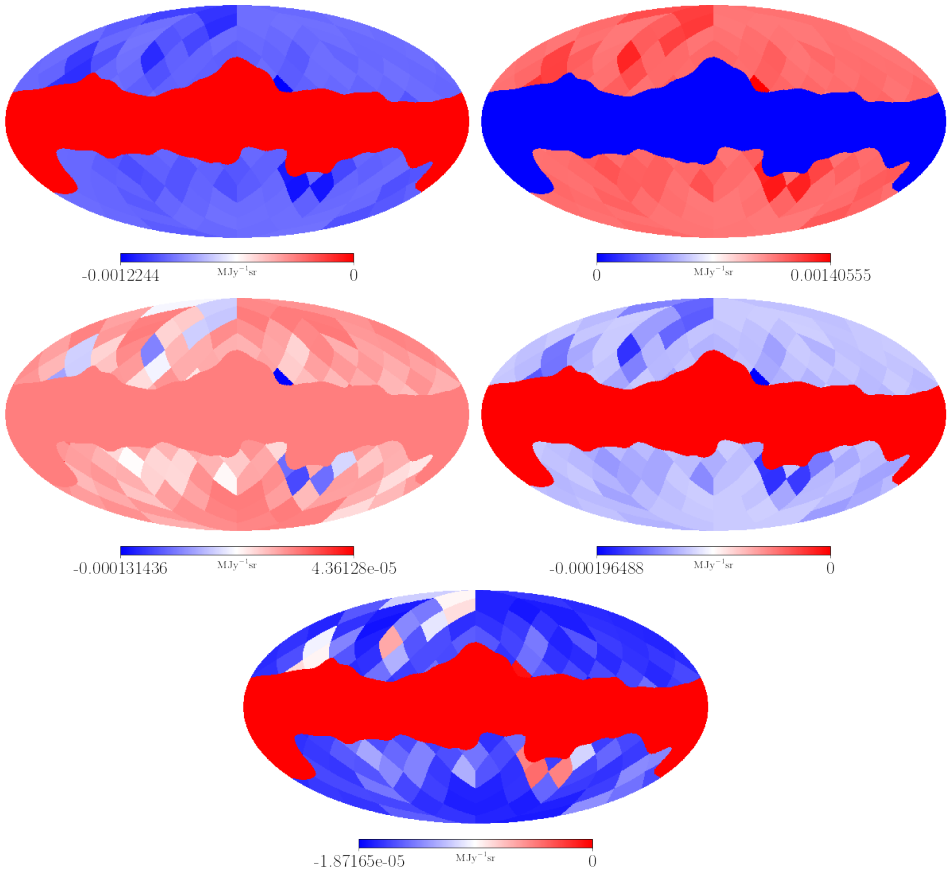

Figure C.2: The CILC weights map for the FYST frequencies. From left to right and up to down, 220 GHz, 280 GHz, 350 GHz, 405 GHz, 860 GHz.

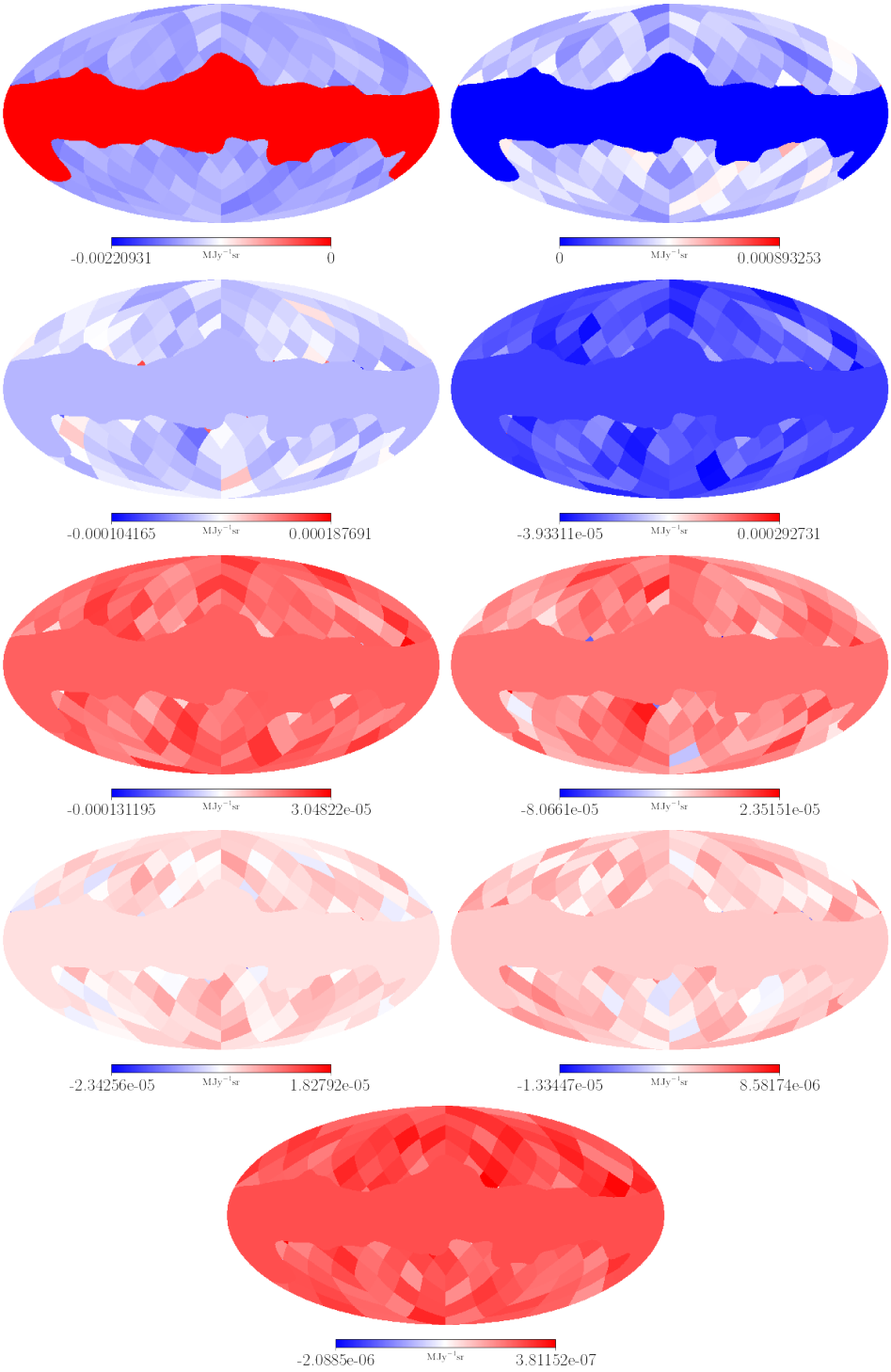

Figure C.3: The CILC weights map for the SO LAT and FYST combined frequencies. From left to right and up to down, 93 GHz, 145 GHz, 220 GHz, 225 GHz, 279 GHz, 280 GHz, 350 GHz, 405 GHz, 860 GHz.

C.2 Realistic sky scenario: instrumental noise debiased

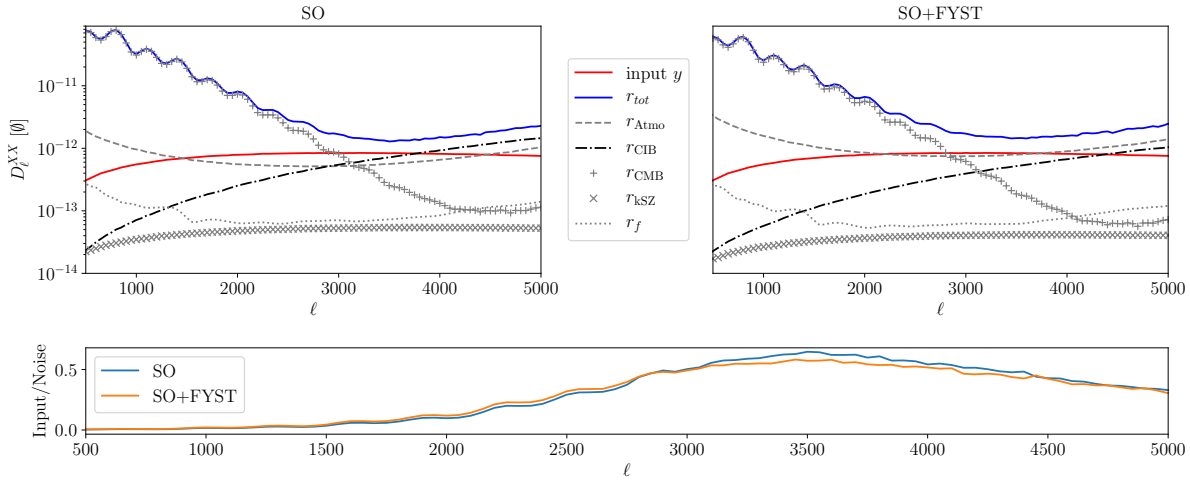

Figure C.4: Power spectra of the ILC residual noises compared to the input Compton-y power spectrum (red). The simulated sky contains all extra-galactic components (see Tab. 5.1), galactic foregrounds, and Atmospheric noise. In red, the target signal, the power spectrum of the input Compton-y map coming from the WebSky simulations.**Top panels:** grey curves are the ILC noises residual power spectra, denoted by $r_{\rm X}$ X being the kSZ, CMB, or the foregrounds (f). In black the power spectrum of the CIB residual noise. In blue the sum of all grey and black curves, the total ILC noise residual debiased for Instrumental white noise. Left panel for SO alone and right panel for SO+FYST combined.**Bottom panel:** the ratio between the input Compton-y power spectrum (red) and the total noise residual (blue) for SO and SO+FYST.All power spectra are beam corrected for comparison with the input y. They are also all binned averaged over a window $\Delta\ell = 50$.

C.3 Thermal Sunyaev-Zeldovich measurements and cosmic infrared background leakage mitigation combining upcoming ground-based telescopes

A&A 677, A87 (2023)
https://doi.org/10.1051/0004-6361/202245834

Astronomy & Astrophysics

Thermal Sunyaev–Zeldovich measurements and cosmic infrared background leakage mitigation combining upcoming ground-based telescopes

M. Charmetant[1] and J. Erler[1,2]

[1] Argelander-Institut für Astronomie, Universität Bonn, Auf dem Hügel 71, 53121 Bonn, Germany
e-mail: mcharmetant@astro.uni-bonn.de
[2] Deutsches Zentrum für Luft- und Raumfahrt e.V. (DLR) Projektträger, Joseph-Beuys-Allee 4, 53113 Bonn, Germany

Received 31 December 2022 / Accepted 5 June 2023

ABSTRACT

Context. The *Fred Young* Submillimeter Telescope (FYST) and the Simons Observatory Large Aperture Telescope (SO LAT) will deliver unprecedented high-resolution measurements of microwave sky emissions. Notably, one of those microwave sky emissions, the thermal Sunyaev–Zeldovich (tSZ) signal, is an essential probe for cluster astrophysics and cosmology. However, an obstacle to its measurement is contamination by the cosmic infrared background (CIB), especially at high frequencies.
Aims. Our goal is to assess the detection and purity of tSZ power spectrum measurements from these two telescopes. We demonstrate that FYST's high-frequency coverage helps lower CIB contamination and improves signal detection.
Methods. We simulated the various components of the microwave sky at the frequencies, sensitivities, and beam sizes of the upcoming SO LAT and FYST telescopes using full-sky Hierarchical Equal Area isoLatitude Pixelisation (HEALPix) map templates from the Websky simulations and the Python Sky Model (PySM). We used a map-based internal linear combination (ILC) and a constrained ILC (CILC) to extract the tSZ signal and compute residual noises to assess CIB contamination and signal recovery.
Results. We find that the CIB's residual noise power spectrum in the ILC-recovered tSZ is lowered by ∼35% on average over the scales $\ell \in [500, 5000]$ when SO LAT and FYST are combined compared to when SO LAT is used alone. We find that when using CILC to deproject CIB, the combined abilities of SO LAT and FYST offer a large $\ell \in [1800, 3500]$ window in which the recovered tSZ power spectrum is not noise dominated.

Key words. cosmic background radiation – cosmology: observations – galaxies: clusters: general – methods: data analysis – submillimeter: diffuse background

1. Introduction

The cosmic microwave background (CMB) photons we observe have been influenced by their interaction with the baryonic matter as they travel through the Universe. These interactions modify the CMB photons and are called secondary anisotropies of the CMB. The study of these secondary anisotropies allows for the extraction of information about the large-scale structure (LSS) of the Universe as well as about the primary anisotropies, the matter content of the Universe. One of the secondary anisotropies, the thermal Sunyaev–Zeldovich (tSZ) effect, is due to the inverse Compton scatting of the CMB photons by the hot electrons in the intracluster medium (ICM) of galaxy clusters. This effect is a crucial tool for cluster detection and was first used by the South Pole Telescope (SPT; Staniszewski et al. 2009) and the Atacama Cosmology Telescope (ACT) Collaborations (Menanteau et al. 2010) and later by the Planck Collaboration (Planck Collaboration XXVII 2016) to construct cluster catalogues. The tSZ effect can be used to measure cluster properties, such as temperature (Erler et al. 2018). It is also crucial for cosmology, as it has a strong dependence on matter clustering σ_8 and matter density $C_\ell^{yy} \propto \sigma_8^{8.1}\Omega_\mathrm{m}^{3.2}$ (Komatsu & Seljak 2002; Planck Collaboration XXII 2016; Bolliet 2018). By fitting the tSZ power spectrum extracted from data with models of the signal and its contaminants, assuming a Gaussian likelihood function, and applying a Markov chain Monte Carlo (MCMC) analysis with a fixed mass bias b, one can constrain these cosmological parameters (Planck Collaboration XXI 2014). This implies that a small improvement in the tSZ power spectra statistical significance or a reduction of the bias due to residual contamination from other signals will lead to a significant improvement in measurements of σ_8 and Ω_m. Therefore, a clean, uncontaminated tSZ map is essential.

The cosmic infrared background (CIB) is the emission from stars and galaxies absorbed and re-emitted by dust in the infrared domain. It is the main contaminant in tSZ reconstruction and is also a tracer of the matter distribution and the LSS (Hurier 2015). The CIB contamination, or leakage, originates from the spatial correlation between CIB and tSZ and is also due to the component separation techniques, such as internal linear combination (ILC; Bennett et al. 2003; Eriksen et al. 2004; Park et al. 2007; Kim et al. 2008; Delabrouille et al. 2009), not being able to perfectly separate the signal from the many other astrophysical emissions present in the sky. Previous studies on CIB leakage in the *Planck* tSZ map made by Planck Collaboration XXII (2016) showed that the amplitude of the leakage increases when the CIB originates from higher redshift populations. The dependence of the leakage on the redshift of the CIB population comes from the fact that the separation methods of components primarily focus on reducing the main contaminant, which at high frequencies is

Galactic thermal dust, and is thus less efficient at reducing the high-z CIB contaminants. This could lead to a bias in tSZ estimation for high-z clusters, as mentioned in Vikram et al. (2017). Another study by Alonso et al. (2018) showed that the error bars on the tSZ profile are not much improved by the deprojection (i.e. nulling) of the CIB with more constrained component separation methods such as constrained internal linear combination (CILC; Remazeilles et al. 2011). Dust contamination can reduce the number of detected galaxy clusters by up to ∼9% within $z \in [0.3, 0.8]$ (Melin et al. 2018). Similarly, Zubeldia et al. (2023) demonstrated that CIB contamination in the tSZ signal can lead to ∼20% fewer clusters being detected and proposed a publicly available code to produce CIB-free cluster catalogues.

In the past three decades, several space missions, such as the COsmic Background Explorer (COBE; Mather et al. 1994), the Wilkinson Microwave Anisotropy Probe (WMAP; Bennett et al. 2003), and *Planck* (Planck Collaboration I 2014), have observed the sky at microwave frequencies and extracted full-sky maps of the CMB, tSZ, CIB, and other astrophysical components. Upcoming ground-based experiments, such as those using the Simons Observatory Large Aperture Telescope (SO LAT; Ade et al. 2019) and the *Fred Young* Submillimeter Telescope (FYST; CCAT-prime Collaboration 2023), will allow for unprecedented measurements of the tSZ and CIB thanks to their high spatial resolution and high sensitivities.

In this work, our focus is on quantifying the benefits of combining future experiments, such as SO LAT and FYST to reduce the CIB leakage in ILC-reconstructed tSZ maps. A cleaner tSZ map will allow better constraints on the cosmological parameters σ_8 and Ω_m to be drawn, more accurate cluster counts and detection, and more accurate measurements of the properties of clusters, such as their temperature. Without mitigation, CIB leakage along the line of sight (LoS) could systematically bias measurements of the tSZ effect and lead to higher tSZ measurements, or it could lower such measurements by filling the characteristic tSZ decrement at frequencies below 217 GHz (Schaan et al. 2021). Obtaining less contaminated signals by combining microwave experiments is common and will be part of the observation strategies of future experiments. For example, previous studies, such as Ade et al. (2019), have shown that SO LAT will achieve precise measurement of the tSZ power spectrum when combined with *Planck* data. As *Planck* observations are not contaminated by the atmosphere, this combination allows SO LAT to mitigate such contamination on large scales. Similarly, Melin et al. (2021) constructed a bigger cluster catalogue by combining *Planck* and the South Pole Telescope (SPT-SZ) data. Our work also encourages the combination of all available data from previous and upcoming CMB experiments.

This paper is structured as follows: In Sect. 2, we formalise the tSZ and CIB description in our study. In Sect. 3, we describe the modelling of high-resolution maps of the different microwave sky emissions. In particular, we first explain how we generated Galactic foregrounds, the extragalactic components, and their processing. Then, we detail the generation of noises and the handling of beams and masks. This is followed by a conclusion of the output of the sky modelling. In Sect. 4, we summarise the methods used to extract the tSZ effect, starting with the ILC and followed by CILC. Next, we explain beam convolution handling and the treatment of full-sky maps, and we provide an overview of the methods used. We present a detailed description of our results and an overview in Sect. 5. Section 6 contains a summary and discussion of the above-mentioned results.

2. Formalism

2.1. The thermal Sunyaev–Zeldovich effect

The thermal tSZ effect, predicted by Sunyaev & Zeldovich (1970, 1972), is the inverse Compton scattering of CMB photons on the hot electrons present in the ICM. A consequence of this effect is the distortion of the CMB blackbody spectrum, due to the shift of the low frequency (corresponding to low energies) CMB photons to higher frequencies (corresponding to higher energies). The temperature change of the CMB due to the tSZ effect can be expressed as:

$$\frac{\Delta T_\mathrm{CMB}}{T_\mathrm{CMB}} = y f(x_\nu), \tag{1}$$

where $T_\mathrm{CMB} = 2.72548 \pm 0.00057$ K (Fixsen 2009) is the temperature of the CMB and y is the dimensionless Compton-y parameter measuring the electron pressure integrated over the LoS proper distance l,

$$y = \frac{\sigma_\mathrm{T}}{m_\mathrm{e} c^2} \int P_\mathrm{e}(l) \mathrm{d}l, \tag{2}$$

where σ_T is the Thomson cross-section, m_e the electron mass, and $P_\mathrm{e}(l) = n_\mathrm{e}(l) k_\mathrm{B} T_\mathrm{e}(l)$ the electron pressure composed of the electron density n_e, the Boltzmann constant k_B, and the electron temperature T_e. In Eq. (1), $f(x_\nu)$ is the spectral shape of the tSZ (Sunyaev & Zeldovich 1972, 1980):

$$f(x_\nu) = x_\nu \coth\left(\frac{x_\nu}{2}\right) - 4 = \left(x_\nu \frac{e^{x_\nu}+1}{e^{x_\nu}-1} - 4\right), \tag{3}$$

where $x_\nu = h\nu / k_\mathrm{B} T_\mathrm{CMB}$, h being the Planck constant and ν the frequency. The frequency spectrum of the tSZ is characterised by a decrement in the intensity change for frequencies lower than $\nu \sim 217$ GHz and an increment at higher frequencies. To express the frequency variations induced by the tSZ effect in terms of an intensity variation, we use the conversion factor to pass from units of thermodynamic temperature (K_CMB; Planck Collaboration IX 2014) in Eq. (1), to intensity units (Jy sr^{-1}),

$$\Delta I_\nu = \frac{\partial B_\nu(T)}{\partial T} \Delta T = \frac{I_0}{T_\mathrm{CMB}} h(x_\nu) \Delta T, \tag{4}$$

where $B_\nu(T)$ is the Planck blackbody spectrum at a fixed frequency ν, $I_0 = 2(k_\mathrm{B} T_\mathrm{CMB})^3/(hc)^2 \approx 270$ MJy sr^{-1}, c is the speed of light, and

$$h(x_\nu) = \frac{x_\nu^4 e^{x_\nu}}{(e^{x_\nu}-1)^2}. \tag{5}$$

Therefore, the tSZ intensity change is given by

$$\frac{\Delta I_\nu}{I_0} = y h(x_\nu) f(x_\nu). \tag{6}$$

More details on the Sunyaev–Zeldovich (SZ) effects can be found in the review by Mroczkowski et al. (2019). In this study, we place ourselves in a simplistic approach and do not take into account the relativistic correction of the tSZ effect (Wright 1979; Itoh et al. 1998; Pointecouteau et al. 1998; Chluba & Sunyaev 2012), even though *Planck* data have shown that this leads to an underestimation of the amplitude of its power spectrum C_ℓ^{yy} (Remazeilles et al. 2018).

2.2. The cosmic infrared background

The CIB is the extragalactic integrated emission produced by all stars and galaxies during their formation and throughout their life. The CIB frequency spectrum peaks in the infrared regime because the majority of sources are far and their light is redshifted by the expansion of the Universe but also because some of this light is absorbed by dust and re-emitted into the infrared regime. This existence of the CIB was predicted by Partridge & Peebles (1967) and first detected by Puget et al. (1996), Fixsen et al. (1998), Hauser et al. (1998). Contrary to the tSZ, one frequency spectrum is not sufficient to fully model it. The CIB spectral energy distribution (SED) depends on the redshift of sources. It can be approximated as a modified blackbody spectrum:

$$I_\nu = A_{\rm dust}\nu^\beta B_\nu(T_{\rm dust}), \tag{7}$$

where $A_{\rm dust}$ is the amplitude of the dust emission, β is the dust emissivity spectral index, and $T_{\rm dust}$ is the temperature of the dust grains. The CIB is a tracer of the entire history of star formation in the Universe. One characteristic of the CIB of particular importance for our study is its non-zero correlation with the tSZ signal. Indeed, some of the unresolved dusty star-forming galaxies constituting the CIB are found in galaxy clusters. Therefore, the spatial correlation between tSZ and CIB is non-zero, and it was measured in Reichardt et al. (2012, 2021), Dunkley et al. (2013), George et al. (2015), Choi et al. (2020b).

3. Modelling the microwave sky

We set up a pipeline called Skymodel to create mock high-resolution maps of the microwave sky that would be observed in future experiments. Our pipeline is based on Hierarchical Equal Area isoLatitude Pixelisation (HEALPix) template maps of the microwave sky from existing simulations as well as the most recent observations of the microwave sky. The microwave sky components are usually separated into two categories: the emissions coming from the Galaxy, called Galactic foregrounds, and the extragalactic components. Our Skymodel deals differently with these two categories, using an existing Python package for the Galactic foregrounds and HEALPix maps from mock simulations for the extragalactic ones.

3.1. Galactic foregrounds

Galactic foregrounds are the various emissions originating from the Milky Way and are thus in the foreground of extragalactic emissions. We considered the following Galactic foregrounds: Galactic synchrotron; free-free emission; anomalous microwave emission (AME), which is supposedly produced by spinning dust (Dickinson et al. 2018); and thermal emission from Galactic dust grains.

All of the above-mentioned components were generated using the Python Sky Model (PySM; Thorne et al. 2017). The PySM[1] package is based on Galactic foreground template maps obtained from Planck Collaboration XXV (2016) that were extracted using the COMMANDER algorithm (Eriksen et al. 2004, 2008; Planck Collaboration IX 2015; Planck Collaboration IV 2020) on *Planck* 2015 raw data, data from WMAP, and 408 MHz all-sky survey by Haslam et al. (1981, 1982), reprocessed by Remazeilles et al. (2015). The original data maps were at a HEALPix (Górski et al. 2005) pixel resolution of $\text{pix}_{\rm size}$ = 13.74′ ($N_{\rm side}$ = 256). The resolution of a HEALPix pixel is defined as the square root of the pixel area. The PySM upgrades them to 6.9′ ($N_{\rm side}$ = 512) and adds small-scale fluctuations generated randomly by extrapolating the power spectra of those maps (Miville-Deschênes et al. 2007). We generated each foreground component at the desired frequency using the basic spectral model of PySM (i.e. “s1”, “f1”, “d1”, and “a1”). These maps were then upgraded to a higher $N_{\rm side}$ = 4096 and smoothed with a 10′ full width at half maximum (FWHM) Gaussian to mitigate pixelation artefacts (see Fig. 1).

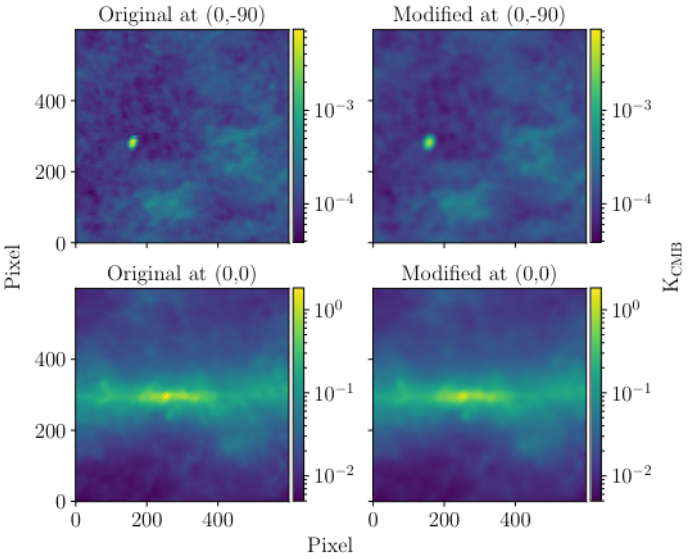

Fig. 1. Projections of the PySM sky at 353 GHz composed of free-free, synchrotron, AME, and Galactic dust. The upper row shows cutouts around the Galactic south pole, and the lower row shows cutouts around the Galactic centre. Each cutout is composed of 600 × 600 pixels of the size 0.86′, making 8.6° × 8.6° maps. The left column shows cutouts from the original PySM sky, which feature pixelation artefacts due to their pixel size of 6.9′ ($N_{\rm side}$ = 512). The right column shows modified cutouts obtained by first updating the original map to a higher resolution, 0.86′ ($N_{\rm side}$ = 4096), by oversampling using the HEALPy function ‘ud_grade()’ and then smoothing with a Gaussian kernel of a width that is close the original pixel size of the PySM maps ($FWHM$ = 10′). This reprocessing of the PySM maps removes the pixelation artefacts while only marginally lowering the spatial resolution.

3.2. The extragalactic components

The extragalactic components include the CMB, tSZ, CIB, and the kinematic Sunyaev–Zeldovich (kSZ) effect. These components were generated at various frequencies using template maps provided by the Websky extragalactic CMB mocks (Stein et al. 2020), which are HEALPix full-sky maps of the extragalactic sky at $N_{\rm side}$ = 4096 ($\text{pix}_{\rm size}$ = 0.86′). The large-scale structure of the Websky simulations was obtained through a peak-patch and 2nd-order Lagrangian perturbation theory (2LPT) simulation with 12288^3 particles and a box size of 15.4 h^{-1} Gpc. The authors assumed a flat Λ cold dark matter (ΛCDM) cosmology with h = 0.68, $\Omega_{\rm cdm}$ = 0.261, $\Omega_{\rm b}$ = 0.049, Ω_Λ = 0.69, $n_{\rm s}$ = 0.965, and σ_8 = 0.81. The simulations can be downloaded from their website[2]. An overview of the provided template maps, their frequencies, and units can be seen in Table 1.

[1] https://github.com/bthorne93/PySM_public

[2] https://mocks.cita.utoronto.ca/index.php/Websky_Extragalactic_CMB_Mocks

Table 1. Summary of the base extragalactic templates of the Websky simulations.

Extragalactic templates from Websky	Type of data	Frequencies (GHz)	Units
Cosmic infrared background (CIB)	HEALPix map	{143, 217, 353, 545, 857}	[MJy sr^{-1}]
Radio point sources (RPSs)	–	–	–
Cosmic microwave background (CMB)	$\{a_{\ell m}\}_L$ or $\{a_{\ell m}\}_{\rm UL}$	–	$[\mu K_{\rm CMB}]$
Thermal Sunyaev–Zeldovich (tSZ)	HEALPix map	Compton-y	[$\emptyset$]
Kinematic Sunyaev–Zeldovich (kSZ)	HEALPix map	–	$[\mu K_{\rm CMB}]$

Notes. The templates are generally composed of either one full-sky map or a set of full-sky maps at various frequencies. For the CMB component, the Websky simulations provided a set of $a_{\ell m}$ that are either lensed (L) or unlensed (UL).

To generate the tSZ map, Stein et al. (2020) projected the pressure profiles from hydrodynamical simulations (Battaglia et al. 2012) onto the mass-Peak Patch halo catalogue. They did not consider field contributions, and the minimum halo mass considered to compute their pressure profile was $\sim 1 \times 10^{13}\, M_\odot$. Their tSZ power spectrum is in good agreement with the *Planck* data (Stein et al. 2020). Websky provides all-sky, dimensionless Compton-y maps (linked to the tSZ effect through Eqs. (1) or (6)), which can be multiplied by the tSZ spectrum in Eqs. (3) and (5) to obtain frequency maps of the tSZ. The tSZ signal was simulated by Websky in its non-relativistic limit.

The CIB was generated using the halo occupation distribution model proposed by Shang et al. (2012) and with the parameters given in Viero et al. (2013) to fit the CIB power spectrum with the *Herschel* observations. Within this model, the rest-frame SED of a source is a modified blackbody at low frequencies and a power law at high frequencies. The CIB's SED depends on the source redshift, halo mass, and frequency at which it is observed. The Websky halo catalogue is populated with central and satellite galaxies. The field contribution was not considered. The Websky simulations generated CIB maps at the *Planck* High-Frequency Instrument (HFI) frequencies (see Table 1), and in particular, the 545 GHz CIB power spectrum was normalised to fit the *Planck* 545 GHz CIB power spectrum (Planck Collaboration I 2014) at $\ell = 500$.

Because of the redshift dependence of the CIB distribution, a decorrelation was expected between the CIB at different frequency channels. The decorrelation was determined in the Websky simulations (Stein et al. 2020) by computing the Pearson product-moment correlation coefficients on the power spectrum of the CIB maps, evaluated at the *Planck* HFI frequency channels. The correlation matrix between the different CIB maps provided by the Websky simulations is illustrated in Fig. 2. The Websky CIB decorrelation values are in agreement with Planck Collaboration I (2014) and Lenz et al. (2019).

In this work, we extrapolated the CIB model to arbitrary frequencies by fitting each pixel of the existing Websky CIB maps with the following modified blackbody:

$$I_\nu = A_{\rm CIB,\nu_0} \left(\frac{\nu}{\nu_0}\right)^{\beta+3} \frac{e^{\frac{h\nu_0}{k_{\rm B}T_{\rm dust}}} - 1}{e^{\frac{h\nu}{k_{\rm B}T_{\rm dust}}} - 1}, \tag{8}$$

where $\nu_0 = 353$ GHz. This was done using the Scipy "curve_fit()" function to derive three all-sky parameter value maps for the variables $A_{\rm CIB}$, $T_{\rm dust}$, and β. These full-sky maps were injected in Eq. (8) to produce Websky-like CIB intensity maps at new frequencies.

In the first approximation, we assumed this exact modified black body across frequencies, which is not fully realistic since the CIB is composed of different sources at different redshifts, and their SED is not the same. The Websky simulation includes cross-correlation between the tSZ and CIB, which is in agreement with the *Planck* data results (see Fig. 11 of Stein et al. 2020).

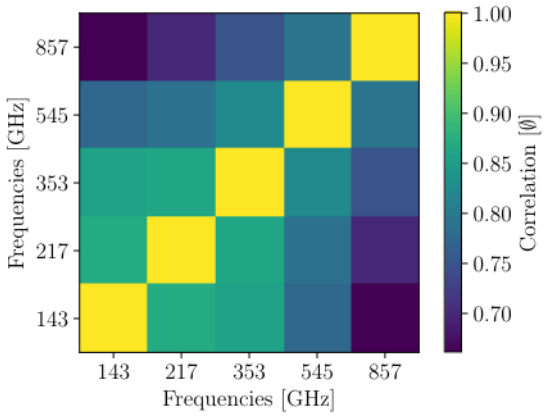

Fig. 2. Correlation matrix between the Websky CIB maps provided at the *Planck* HFI frequencies. The axis shows the frequencies of the maps in GHz, and the dimensionless colour bar shows the correlation coefficient.

For the CMB component, the Websky simulations provided us with two files containing random realisations of the primary T/Q/U $a_{\ell m}$ coefficients, one being unlensed and one being lensed. The lensed file was obtained using the unlensed $a_{\ell m}$ file and the CMB lensing convergence from $0 < z < 1100$ in the Born approximation. From those $a_{\ell m}$, we could reconstruct a lensed and an unlensed CMB map in $\mu K_{\rm CMB}$. Using those maps and the conversion factor (Eq. (4)), we could obtain a CMB map at any desired frequency. In the first approximation, we used the unlensed CMB. For the kSZ component, the situation is analogous to the tSZ one. The Websky templates provided frequency-independent maps of the kSZ, which can be seen as a '$y_{\rm kSZ}$-map'. We multiplied the maps by the conversion factor in Eq. (4) to get a kSZ map at any frequency. The Websky simulations do currently not provide radio point source (RPS) templates, but the Websky team will add them in a future release. This extragalactic component is therefore not accounted for in our study. The different extragalactic full-sky map templates were all converted to the same unit, that is, either temperature units ($K_{\rm CMB}$) or intensity units (Jy sr^{-1}) units, using Eq. (4).

3.3. Future experiments

We simulated the microwave sky as will be seen by two future ground-based telescopes. One of these telescopes is the FYST

Table 2. Frequencies, beam sizes in FWHM, and baseline instrumental sensitivities for SO LAT and FYST.

		SO LAT		FYST	
Frequency	$FWHM$	Sensitivity		Sensitivity	
(GHz)	(arcmin)	($\mu K_{\rm CMB}$-arcmin)	($\mu K_{\rm RJ}$-arcmin)	($\mu K_{\rm CMB}$-arcmin)	($\mu K_{\rm RJ}$-arcmin)
93	2.2	8	6.42	–	–
145	1.4	10	5.96	–	–
220	1.0	–	–	15	4.86
225	1.0	22	6.80	–	–
279	0.9	54	9.68	–	–
280	0.9	–	–	27	4.79
350	0.6	–	–	105	8.38
405	0.5	–	–	372	15
860	0.2	–	–	5.7×10^5	33.93

of the Cerro Chajnantor Atacama Telescope prime (CCAT-prime) Collaboration (CCAT-prime Collaboration 2023). It is an upcoming 6 m diameter telescope that will be located at 5600 m on the mountain of Cerro Chajnantor in Chile. Due to its very low precipitable water vapour, this exceptional site will provide one of the best atmospheric transmissions for ground-based microwave observations to date (Bustos et al. 2014). The crossed-Dragone optical design proposed by Niemack (2016) will deliver a high-throughput, wide-FoV telescope capable of rapidly scanning large areas of the sky. The first light is expected in 2024.

The second telescope of interest will be part of the Simons Observatory (SO; Ade et al. 2019). The SO will be composed of three refracting 0.4 m small aperture telescopes (SATs) and one cross-Dragone 6 m large aperture telescope (LAT). The four instruments will be located at 5300 m on Cerro Toco in the Atacama Desert in Chile.

We studied what benefits may result from the combination of FYST with the SO LAT. To do so, we used PySM and the Websky mock sky to simulate the microwave sky at frequencies and sensitivities, meaning the smallest signal a telescope can detect above the random background noise, as well as the beam sizes of future observations conducted with these telescopes, see Table 2.

3.4. Noises, masks, and beams

Our Skymodel pipeline includes two types of noise: white noise, to reproduce the instrumental noise, and atmospheric red noise, which is present for ground-based experiments. The noise implementation is based on the SO Collaboration (Ade et al. 2019) noise curves forecasted and adapted for CCAT-prime by Choi et al. (2020a):

$$N_\ell = N_{\rm red}\left(\frac{\ell}{\ell_{\rm knee}}\right)^{\alpha_{\rm knee}} + N_{\rm white}, \tag{9}$$

where $\ell_{\rm knee} = 1000$ and $\alpha_{\rm knee} = -3.5$. The values for $N_{\rm red}$ and $N_{\rm white}$ for the SO LAT and FYST are included in the Skymodel. For this reason, the atmospheric noise model can only be computed at valid SO LAT frequencies or FYST frequencies (see Table 2) and takes the form of a series of power spectra, as defined by Eq. (9) and shown in Fig. 3. Using those power spectra, the Skymodel generates full-sky map realisations of the atmospheric noise or atmospheric and instrumental noise. This is done through the HEALPy function 'synfast()', which generates map realisations from the power spectrum while assuming the field to be Gaussian. This was done for each frequency band. In reality, neighbouring frequency bands will observe similar regions of the sky at the same time. Thus, the atmospheric noise is partially correlated. We adopted a 90% correlation between each pair of neighbouring frequency channels as suggested in Ade et al. (2019). Fixing the 'synfast()' seed allowed for the production of a 100% correlated map for two neighbouring channels. Another seed and a correction factor can be applied to the power spectrum to produce a down-weighted correlated map and an uncorrelated map of one of the two channels. This new map and the fully correlated map can be linearly combined to produce a 90% correlated map using

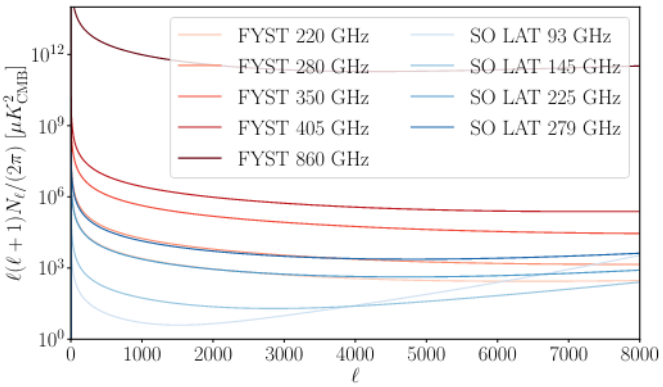

Fig. 3. SO LAT (blue lines) and FYST (red lines) sum of the atmospheric red and instrumental white noise spectra based on the noise model presented by Ade et al. (2019) and adopted by Choi et al. (2020a). The noise curves have been beam corrected.

$$M^{90\%}(\nu) = h(0.9^2 C_\ell^{n,100\%}(\nu)) + h((1-0.9)^2 C_\ell^{n,0\%}(\nu)), \tag{10}$$

where M is the final map at the frequency ν, which will be 90% correlated with its neighbouring channel. The HEALPy synfast() function is represented by h, and C_ℓ^n is the atmospheric noise curve, which is 100% correlated when produced from the same seed as its neighbouring channel and 0% when produced from another seed.

The frequency bands are broad and extended over some frequency range, and their transmission is not perfect but varies over their extent. In the first approximation, we considered the frequency bands to be delta functions with perfect transmission.

In addition to the maps of astrophysical emission and noise, the Skymodel uses survey masks. The masks are stored at $N_{\rm side}$ =

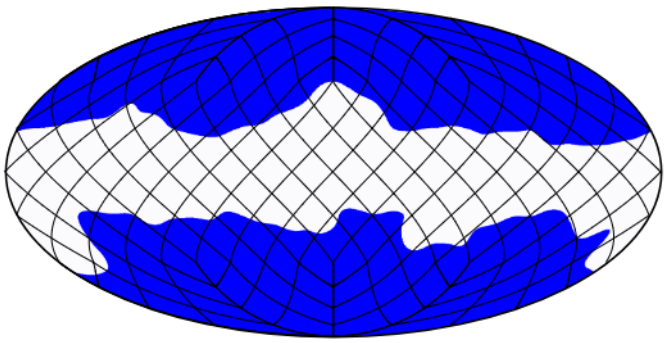

Fig. 4. Galactic foreground mask derived from *Planck* data by Erler et al. (2018; $f_{\rm sky} \approx 0.6$). The white part is the masked region, and the blue part is what was used in our analysis. Overplotted is the HEALPix tessellation scheme for $N_{\rm side} = 4$.

256 and were upgraded to $N_{\rm side} = 4096$ in order to correspond to the Websky map resolution. In this study, we used a mask that excludes the Galactic plane with a sky fraction of $f_{\rm sky} = 0.6$ (see Fig. 4). This mask was derived in Erler et al. (2018) by masking the 40% brightest pixels of a *Planck* Galactic emission map. This is an initial approximation, as in reality SO LAT and FYST will observe a smaller fraction of the sky.

The Skymodel pipeline also simulates the effect of telescope beams using the Gaussian FWHMs given in Table 2. The product of the Skymodel is a full-sky HEALPix map containing all the Galactic foreground components, the extragalactic components, instrumental white noise, and atmospheric red noise at a given frequency and assuming a given beam size (see Appendix D).

4. Method to extract thermal Sunyaev–Zeldovich map

To study the power spectrum measurement of the tSZ effect for the SO LAT and FYST telescopes, a tSZ map needs to be extracted from the simulated microwave sky. The sky is composed of various Galactic and extragalactic emissions that are considered to be contaminants of our signal of interest. Component separation methods permit the recovery of a signal from the multi-component microwave sky. Internal linear combination (ILC) is a multi-frequency component separation technique that was first used by Bennett et al. (2003) to extract CMB maps from WMAP data and later refined by Eriksen et al. (2004). Multiple flavours of ILC exist. In their paper, Tegmark et al. (2003) suggested applying it in Fourier space rather than in map space. Needlet ILC (NILC; Delabrouille et al. 2009) processes the maps in both spatial and harmonic space by proposing to filter each multi-frequency map by a set of ℓ-dependent window functions called needlets in order to apply multiple ILCs on different multipole intervals. This makes the ILC weights ℓ-dependent, allowing them to take advantage of the scale-dependent distribution of astrophysical components to minimise the noise and better extract the signal of interest. The Modified ILC Algorithm (MILCA; Hurier et al. 2013) differentiates itself by allowing multiple constraints on the Astrophysical emissions in order to reduce their contamination. Scale-discretised directional wavelet ILC (SILC; Rogers et al. 2016) filters the map with ℓ windows that are direction dependent.

In this work, we used ILC and CILC methods in a map-based space. It is also possible to use these methods directly in Fourier space, which is computationally faster but does not allow for spatial separation of signals. Fourier space ILC/CILC only has information on the power spectra of the component in ℓ-space and no spatial information on the localisation of the signal over the full-sky map.

4.1. Internal linear combination

Internal linear combination is widely used in tSZ studies (Bennett et al. 2003; Fornazier et al. 2022). Its main advantage is that it is a so-called "blind" method because only the SED of the component of interest must be known, and no information on the contaminants is needed. The ILC supposes that (1) the maps are composed of a linear combination of astrophysical components and noise, and (2) the astrophysical components are uncorrelated.

A set of $N_{\rm obs}$ observed map $m_i(p)$ at a frequency i and a pixel p can be expressed as

$$m_i(p) = a_i s(p) + n_i(p), \tag{11}$$

where a_i is the amplitude of the frequency spectrum of the signal of interest s at each frequency i. The noise at a given frequency i and pixel p is represented as $n_i(p)$.

To retrieve the signal of interest s, an estimator $\hat{s}_{\rm ILC}$ is defined as

$$\hat{s}_{\rm ILC}(p) = \sum_i \omega_i m_i(p), \tag{12}$$

where ω_i is the ILC weight at the frequency i. The ILC weights are defined such that they follow two conditions. One of them is that in order to obtain an unbiased estimate, the ILC weights must have a unit response to the frequency spectrum of the signal of interest s characterised by a_i at each frequency i (see Eq. (13)):

$$\sum_i \omega_i a_i = 1. \tag{13}$$

The other condition is that the minimisation of the presence of the unwanted astrophysical contaminants and the instrumental noise in the reconstructed map of the signal s is achieved through the condition of minimum variance of the reconstructed map, $\mathrm{VAR}(\hat{s}_{\rm ILC}(p))$. As shown in Eriksen et al. (2004), because the components are uncorrelated, we have $\mathrm{VAR}(\hat{s}_{\rm ILC}(p)) = \sum_k \sum_j \hat{c_{jk}} \omega_j \omega_k$, where $\hat{c_{jk}}$ is the value of the covariance of the observed maps between the frequencies j and k. Thus, the minimum variance condition is given by

$$\frac{\partial}{\partial \omega_i} \sum_k \sum_j \hat{c_{jk}} \omega_j \omega_k = 0. \tag{14}$$

The minimum variance under the constraint of Eq. (14) can be achieved using the Lagrange multiplier method to solve the above equation. The solution gives the following ILC weights:

$$\boldsymbol{W}^{\rm T} = \frac{\boldsymbol{A}^{\rm T} \hat{\boldsymbol{C}}^{-1}}{\boldsymbol{A}^{\rm T} \hat{\boldsymbol{C}}^{-1} \boldsymbol{A}}, \tag{15}$$

where $\boldsymbol{W}$ is the vector containing the ILC weights, and $\boldsymbol{A}$ is the vector containing the frequency spectrum information of the signal of interest and is called 'the mixing vector'. The covariance matrix of the set of maps is represented by $\hat{\boldsymbol{C}}$. The map containing our signal of interest is given by

$$\hat{s}_{\rm ILC}(p) = \sum_i \omega_i^{\rm T} m_i(p). \tag{16}$$

From that, we can define the residual noise ($r_{\rm N}$), that is, the noise left over by the ILC in the reconstructed map. This noise can be defined as the difference between the ILC-recovered map and the true signal:

$$r_{\rm N} = \hat{s}_{\rm ILC}(p) - \sum_i \omega_i^{\rm T} a_i s(p) = \sum_i \omega_i^{\rm T} n_i(p). \tag{17}$$

The contaminants are not absent from the recovered signal map. Moreover, the low variance contaminants might not be reduced in the resulting map. This becomes a problem when assumption two of the ILC is not fully verified and our target signal is spatially correlated with another emission, such as for the tSZ and CIB. This is the case between CIB and tSZ. The remaining CIB signal in the map can bias the ILC estimate of the tSZ. One way to reduce this effect is through CILC.

4.2. Constrained internal linear combination

The CILC method was first introduced by Remazeilles et al. (2011). It adds one extra constraint to the ILC approach. If the spectral shape of one of the contaminants is known, it can be taken into account in the modelling such that instead of Eq. (11), it is supposed that

$$m_i(p) = a_i s(p) + b_i q(p) + n_i(p), \tag{18}$$

where $s(p)$ is the signal of interest and $q(p)$ is the contaminant we want to suppress. Respectively, a_i and b_i are the intensity of the frequency spectrum evaluated at a frequency i of the signal and contaminant. All the remaining astrophysical contaminants and the instrumental noise at a given frequency i and pixel p are contained by $n_i(p)$. The CILC weights are defined such that

$$\sum_i \omega_i a_i = 1 \text{ and } \sum_i \omega_i b_i = 0. \tag{19}$$

The CILC weights have a unit response to the SED of the signal of interest s and a null response to the SED of the astrophysical contaminant q we want to suppress so that this contaminant is absent from the reconstructed signal.

The condition of minimum variance (see Eq. (14)) stays the same, and only the expression of the weights is changed to take into account the two conditions imposed by the CILC,

$$\boldsymbol{W}^{\rm T} = \frac{(\boldsymbol{B}^{\rm T}\hat{\boldsymbol{C}}^{-1}\boldsymbol{B})\boldsymbol{A}^{\rm T}\hat{\boldsymbol{C}}^{-1} - (\boldsymbol{A}^{\rm T}\hat{\boldsymbol{C}}^{-1}\boldsymbol{B})\boldsymbol{B}^{\rm T}\hat{\boldsymbol{C}}^{-1}}{(\boldsymbol{A}^{\rm T}\hat{\boldsymbol{C}}^{-1}\boldsymbol{A})(\boldsymbol{B}^{\rm T}\hat{\boldsymbol{C}}^{-1}\boldsymbol{B}) - (\boldsymbol{A}^{\rm T}\hat{\boldsymbol{C}}^{-1}\boldsymbol{B})^2}, \tag{20}$$

where $\boldsymbol{W}$ is the vector containing the CILC weights, $\boldsymbol{B}$ is the mixing vector of the contaminant, $\boldsymbol{A}$ is the mixing vector of the signal, and $\hat{\boldsymbol{C}}$ is the covariance matrix of the set of maps. Those weights are then applied to the set of multi-frequency maps using Eq. (16) to recover a map containing the signal of interest free from the nulled contaminant and with minimum contamination from other astrophysical signals and instrumental noise. The CILC method can be generalised to null more than one contaminant if their SED is known. However, these additional constraints reduce the degree of freedom available for the condition of minimum variance to be verified (Hurier et al. 2013), thus generating increased noise for the CILC compared to the ILC. Similar to the ILC, the residual noise can be defined as the difference between the CILC-recovered map and the signal of interest (see Eq. (17)), except that the weights are given by Eq. (20).

4.3. Beam convolution

The set of multi-frequency maps was smoothed with a Gaussian kernel to the lowest frequency channel resolution so that they could be linearly combined by the ILC. No meaningful physical information is available on scales smaller than the resolution of each map. When linearly combining the maps, the upper limit of meaningful physical information was therefore given by the lowest resolution map.

More sophisticated component separation methods, such as NILC, allows the original resolution of each frequency map to be kept. Indeed, the NILC weights are a function of the multipole ℓ for each frequency map and can therefore be set to zero when the scale exceeds the resolution of the map. Setting the weights to 0 above a given ℓ, allows the lower-resolution maps to contribute to the signal reconstruction only up to their beam-limited scale, leaving the higher-resolution maps to be the contributors of signal reconstruction on smaller scales.

We note that when simulating the microwave sky, only the astrophysical components and atmospheric noise are limited by the instrument beam. Instrumental noise is not beam limited. The inclusion of the beam was done using HEALPy to smooth the maps $a_{\ell,m}$ in the multipole space and multiply them with a ℓ-dependent Gaussian kernel. Our set of multi-frequency maps can be expressed as

$$a_{\ell,m}(m_i^{\rm s}) = a_i b_i(\ell) a_{\ell,m}(s) + b_i(\ell) a_{\ell,m}(c_i) + a_{\ell,m}(n_i), \tag{21}$$

where $a_{\ell,m}(m_i^{\rm s})$ are the $a_{\ell,m}$ coefficient of the smoothed map at the frequency i. The Gaussian beam in the multipole space at the frequency i is indicated as $b_i(l)$. Next, $a_{\ell,m}(s)$ is the $a_{\ell,m}$ decomposition of component of interest map, and $a_{\ell,m}(c_i)$ represents all astrophysical contaminants and the atmospheric noise at a frequency i. Finally, $a_{\ell,m}(n_i)$ is the instrumental noise $a_{\ell,m}$ at the frequency i. In general, to apply an ILC, all maps have to be smoothed down to the lowest resolution. This is necessary to avoid adding information from other higher resolution, higher frequencies maps that are smaller than the beam size of some of the maps into the recovered map. The smoothing of the maps to the lowest map resolution was done using HEALPy, and this process can be formalised as

$$a_{\ell,m}(m_i^{*}) = a_i b_{\rm f}(\ell) a_{\ell,m}(s) + b_{\rm f}(\ell) a_{\ell,m}(c) + \frac{b_{\rm f}(\ell)}{b_i(\ell)} a_{\ell,m}(n_i), \tag{22}$$

where $a_{\ell,m}(m_i^{*}) = (b^{\rm f}(\ell)/b_i(\ell)) a_{\ell,m}(m_i^{\rm s})$ are the $a_{\ell,m}$ coefficient of the map at a frequency i smoothed down to the lowest resolution. The Gaussian kernel at this desired 'lowest' resolution is indicated at $b^{\rm f}$. The consequence of the necessary smoothing to the lowest resolution is an increase of the instrumental white noise contribution in each frequency map by a factor $b_{\rm f}(\ell)/b_i(\ell)$ because for all the frequency channels, except the lowest one, we have $b_{\rm f}(\ell) > b_i(\ell)$.

4.4. Full-sky treatment

The microwave full-sky maps are composed of many different astrophysical signals. Some of those signals are spatially localised. The Galactic foregrounds are mainly localised around the Galactic plane. When applying the map-based ILC to a full-sky map, instead of applying it straightforwardly to the multi-frequency set of maps, thus getting only one set of weights for the full sky and equally treating regions that are strongly contaminated by Galactic foregrounds and regions that are less contaminated, we tessellated the sky into different zones and applied

an ILC on each of these zones separately. This resulted in a set of different weights for each region so that they are always optimised for the level of contamination present in their zone. This has been done previously in various studies (Tegmark et al. 2003). The most common tessellation approach is to separate the sky into zones of Galactic contamination.

To tessellate the sky, we used the native nested HEALPix tessellation (see Fig. 4). This nested scheme tessellates the sky originally into 12 equal-area fields. The sides of each of those fields can be further divided to create more equal-area cells of smaller sizes. Using HEALPix nested scheme to create subfields (Górski et al. 2005), we chose to tessellate the sky using $N_{\rm side} = 4$, which gives $N_{\rm pix} = 12 \times N_{\rm side}^2 = 192$ fields, each covering an area of $\approx$215 deg^2. In general, tessellation can create border effects along the limits of the fields. In Eriksen et al. (2004), the border effects were attenuated by smoothing the ILC weights. In this work, we do not correct the border effects.

The Galactic plane was masked (see Fig. 4) with a $f_{\rm sky} \approx 0.6$ Galactic mask map to reduce foreground contamination and better extract the tSZ effect. To compute power spectra on a masked map, we used PyMaster (Alonso et al. 2019), which corrects the power spectra for the masking of part of the sky. In this work, we do not apodise the masks, as it is computationally very intensive at our high resolution and does not make a significant difference in the resulting power spectrum (see Fig. A.2). The full treatment which includes, creation, tessellation, masking and ILC is performed at the map level. Only when a tSZ full-sky map was extracted was the power spectrum of the map computed using spherical harmonics.

4.5. Debiasing

The instrumental white noise biasing of the tSZ power spectrum extraction was estimated for a 4000 h observation time (Choi et al. 2020b). Therefore, the mean of this noise could be computed and removed to debias the tSZ reconstruction. This was done by generating random realisation of the instrumental white noise for each frequency at the sensitivities given in Table 2 and computing the projected noise by multiplying the ILC weights from the tSZ reconstruction with the noise maps and summing them. The power spectrum of this noise was then subtracted from the ILC-extracted Compton-y ($y_{\rm ILC}$) power spectrum. All the following power spectrum results were debiased for instrumental white noise.

Similarly, as the CMB power spectrum is well-known, we could generate random realisation from the Websky CMB power spectrum to debias for the CMB residual noise. In the first approximation, this study presents an idealised case where the CMB power spectrum is perfectly known and its residual is removed. With real data, this CMB debiasing will be limited by the uncertainties on the recovered CMB power spectrum.

4.6. Overview of steps

In this section, we describe our tSZ extraction method and illustrate it in Fig. 5.

First, using the Skymodel based on the Websky simulation and PySM, we generated multi-frequency maps of the microwave sky at the frequencies, sensitivities, and beam sizes of SO LAT and FYST (see Table 2). The set of maps is in the HEALPix format and uses intensity MJy sr^{-1} units.

Second, the simulated maps were all smoothed down to the resolution of the lowest frequency channel used in our analysis.

Third, a Galactic mask that excludes 40% of the sky (see Fig. 4) was applied to the set of multi-frequency maps.

Fourth, each map was tessellated into smaller equal-area regions using the HEALPix nested format.

Fifth, an ILC was applied to each multi-frequency set of the tessellated area to obtain the Compton-y patches. The patches were reassembled to obtain the full-sky Compton-y map. The covariance matrix was computed only on the unmasked data by our ILC.

Sixth, the power spectrum of the recovered Compton-y map was computed using PyMaster and debiasing, and the residual noises were computed using Eq. (17).

5. Results

To quantify how much CIB could leak into the reconstructed tSZ maps for future experiments and access the detectability of the tSZ power spectrum, we used the microwave sky model detailed in Sect. 3. We simulated a set of multi-frequency maps at the frequencies, sensitivities, and beam sizes for SO LAT and FYST (see Table 2) with atmospheric noise. We applied an ILC or CILC component separation method to extract a tSZ map and looked at the ratios between the different noises and signal power spectra. We restricted ourselves to the scales $\ell \in [500, 5000]$. We chose this lower limit to avoid the cosmic variance and the upper limit because of the 2.2′ resolution of our HEALPix maps, which allowed us to resolve scales up to $\ell \sim 4909$.

5.1. Internal linear combination results

By construction, the weights of an ILC must verify the first condition, which is to have a unit response to the SED of the tSZ effect (see Eq. (13)). This first condition means that the element-wise product of weights (ω_i) times the mixing vector (a_i) is an indicator of the fractional contribution of each frequency band to the recovery of the tSZ signal. It can be negative since the frequency spectrum of the tSZ effect is negative below 217 GHz. The weights attributed by the ILC can themselves be negative for optimisation purposes, with the only constraint being that the sum over all frequencies must be one. The value of the weights is given by the mean of the 192 weights (one per tessellated field), and the error bars are given by the standard deviation of those weights. The top panel of Fig. 6 shows that for SO LAT alone (orange dots), the 93 GHz channel is the main contributor to the tSZ reconstruction by a factor $\omega_{93}a_{93} \approx 1.4$. This channel overestimates the tSZ signal, and the SO LAT 145 GHz channel was used by the ILC to correct this. As a result, the 145 GHz channel contributes roughly by a factor $\omega_{145}a_{145} \approx -0.4$ to the tSZ reconstruction. This result is not surprising, as those frequency bands are located at the tSZ decrement that is characteristic of the signal in the microwave sky and where the values of the mixing vector are negative. When SO LAT and FYST are combined (SO LAT+FYST; blue dots), the SO LAT 93 GHz and 145 GHz channels remain the main contributors to the tSZ reconstruction and have similar values, $\omega_{93}a_{93} \approx 1.45$ and $\omega_{145}a_{145} \approx -0.45$. Other channels from SO LAT and FYST do not contribute to the reconstruction. When FYST is alone (green dots), the main contributor to the tSZ reconstruction is the 280 GHz channel, by a factor $\omega_{280}a_{280} \approx 1.1$, and the 350 GHz channel, by $\omega_{350}a_{350} \approx 0.05$. The 220 GHz and 405 GHz channels were used to mitigate the over evaluation of the tSZ signal with the respective factors $\omega_{220}a_{220} \approx -0.05$ and $\omega_{405}a_{405} \approx -0.1$.

The second condition that has to be verified by the ILC weights is to ensure that the variance of the recovered map has

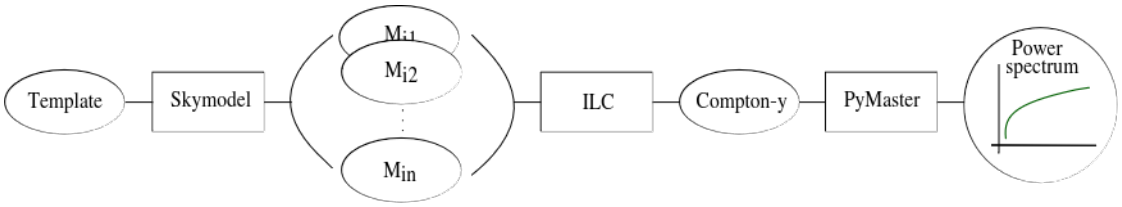

Fig. 5. Scheme of the method we use in this paper to extract the tSZ signal. The input of the sky model is template maps coming from the Websky simulations and PySM. The sky model processes them to produce maps of the different sky components at a given frequency, sensitivity, and beam. This set of multi-frequency maps $\{M_{i1}, \ldots, M_{in}\}$ was smoothed to a common resolution and tessellated, and the Compton-y signal was extracted by the ILC. The PyMaster algorithm was used on the resulting masked map to extract the power spectrum.

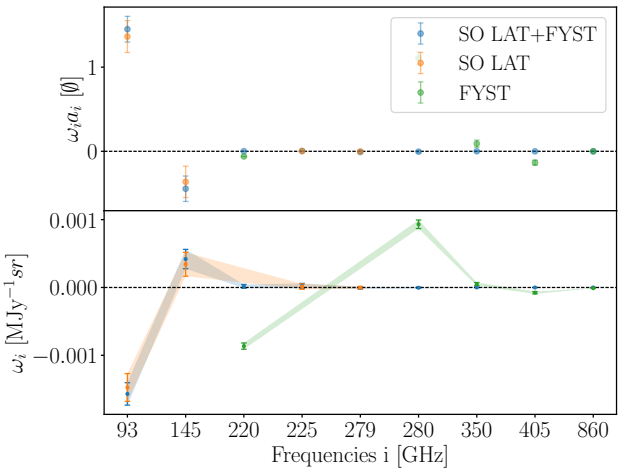

Fig. 6. Internal linear combination weights for SO LAT-like (orange) and FYST-like (green) experiments and SO LAT+FYST (blue) when the sky contains all extragalactic components (see Table 1), Galactic foregrounds, and atmospheric noise. Top figure: fractional contribution of each frequency channel to the recovered y-map. Bottom figure: frequency channels that contribute to maximising the tSZ signal and minimising the noise. The error bars show the standard deviation around the mean weight of all of the 192 tessellated fields for each frequency map.

been minimised (see Eq. (14)). The bottom panel of Fig. 6 shows the values of the weights. In combination with the top panel, we concluded that the weights associated with some frequency bands that do not contribute to the tSZ signal reconstruction but have a high absolute value contribute to the noise minimisation. In particular, the 220 GHz channel located close to the tSZ null is often attributed a high weight by the ILC even though it does not contribute to the tSZ reconstruction. This is because it is the only map containing all the noise contaminants and almost no signal. It is thus the perfect map for noise minimisation. But this is not the case in Fig. 6, except for when FYST is alone (green dots), because of the presence of atmospheric noise in our mock sky. The strong atmospheric noise contaminates the 220 GHz band, making it less exploitable by the ILC (see the comparison with Fig. A.1 for a case without atmospheric noise). Figure 6 shows that in the presence of all the microwave emissions and atmospheric noise, the SO LAT frequency channels that probe the tSZ decrement, 93 GHz and 145 GHz, drive both the tSZ reconstruction and the noise minimisation of the tSZ signal.

However, the weights do not tell us anything about the CIB contamination or the detectability of the tSZ signal. For that, we turned to the power spectrum and the residual ILC noise power spectra. The ILC output is a map containing the Compton-y signal and residual noises (Eq. (17)). Using PyMaster, the power spectra of this map and the residual noises can be computed and compared to the input-pure Compton-y map power spectra from the Websky simulations. We defined $D_\ell^{XX} = \ell(\ell+1)C_\ell^{XX}/2\pi$ as the power spectra of the quantity X or of the residual noise r_X of the quantity X. For SO LAT alone (see Fig. 7 top-left panel) the dominant noise residual on small scales is the CIB residual ($r_{\rm CIB}$), which starts to dominate around $\ell \sim 3700$. The cumulative foreground residual (r_f) and the kSZ residual ($r_{\rm kSZ}$) are both subdominants to the input signal by around one order of magnitude over the scales $\ell \in [500, 5000]$. The atmospheric noise residual ($r_{\rm Atmo}$) dominates the input signal for $\ell < 1400$ and $\ell > 4500$.

In the combined case of SO LAT+FYST (see right panel of Fig. 7), the situation is similar to the case where SO LAT is alone except that the atmospheric noise residual ($r_{\rm Atmo}$) is higher when FYST is present. This could be due to higher atmospheric noise being present in the high-frequency channels of FYST (see Fig. 3). The CIB noise residual power spectrum is significantly lower for SO LAT+FYST. It starts to dominate above the input signal around $\ell \sim 4200$, while for SO LAT alone it was around $\ell \sim 3700$. We concluded that adding FYST to SO LAT reduces the CIB noise residual.

The lower panel of Fig. 7 shows the extent the input Compton-y power spectrum is dominated by the total ILC noise residual as a function of the scale. The total noise residual ($r_{\rm tot}$) power spectrum dominates over the input signal on nearly all scales except around $\ell \in [2100, 2700]$ for SO LAT alone. This is because of our debiasing of the CMB on those critical scales. Overall, SO LAT alone has a better input/noise ratio on scales $\ell < 2500$ because of its slightly lower atmospheric noise residual, and SO LAT+FYST does better than SO LAT alone on scales $\ell > 3000$ because of its lower CIB noise residual.

5.2. Constrained internal linear combination results

Constrained internal linear combination can be used to 'null' the contribution of a component of the microwave sky using its frequency spectrum. Nullifying several components was tested.

Before any debiasing, the CMB is the main contaminant of the recovered Compton-y power spectra on large scales, so it seemed natural to deproject (i.e. 'null') it. However, the additional constraints on the ILC weights (see Eq. (18)) led to a smaller solution space for the weights to extract the signal, which resulted in a 'noise penalty', that is, higher noise in the recovered signal map. In this particular case of a mock sky containing all Galactic foregrounds and extragalactic emission, excluding RPSs, and applying a CILC to deproject CMB, the ILC weights are so constrained to null CMB that they become less efficient at reducing the second main contaminant, the atmospheric noise contamination, which blows up. As a consequence, the total residual noise dominates the input signal even more, by up to one

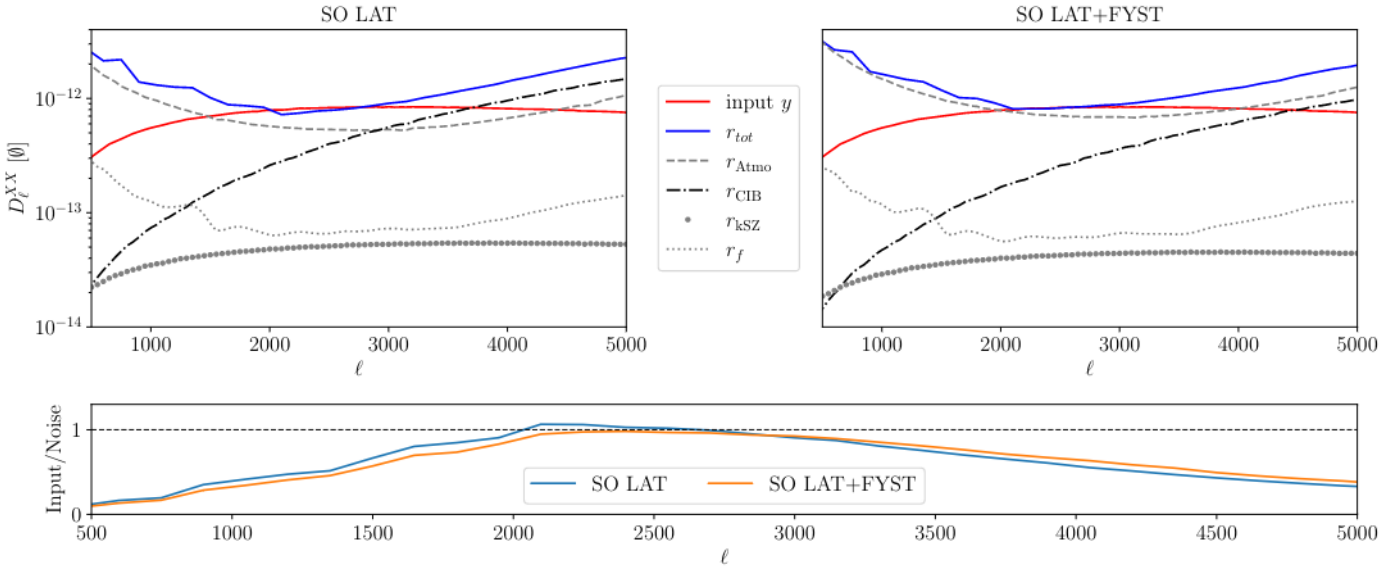

Fig. 7. Power spectra of the ILC residual noises compared to the input Compton-y power spectrum from the Websky simulations (red). The simulated sky contains all extragalactic components (see Table 1), Galactic foregrounds, and atmospheric noise. Top panels: power spectra of various quantities. The grey curves show the ILC noises residual power spectra, denoted by r_X (X being the kSZ); the cumulative Galactic foregrounds (f); and the atmospheric noise (Atmo). The black curve shows the power spectrum of the CIB residual noise. The blue curve shows the sum of all grey and black curves, that is, the total ILC noise residual debiased for instrumental white noise and the CMB. The top-left panel shows the results for the simulated SO LAT data, while the top-right panel shows the results for the simulated data for SO LAT and FYST combined. Bottom panel: ratio of the input Compton-y power spectrum (red) and the total noise residual (blue) for SO LAT and SO LAT+FYST. All power spectra have been beam corrected and bin averaged over a window $\Delta\ell = 50$.

order of magnitude on small scales (see Appendix B). Deprojecting the CMB is therefore not a recommended option for our study. This effect was already reported in Fig. 36 of Ade et al. (2019), but in their study, the blowing up of the noise was mitigated on large scales by the addition of *Planck* data.

Because of its high-frequency coverage, we expected FYST to better probe the CIB, which is the dominant contaminant above $\nu > 300\,\mathrm{GHz}$, and therefore better target CIB contamination and remove it. Deprojecting the CIB using CILC seemed the natural option to take full advantage of the potential constraining power of the FYST high-frequency channels. We note that a perfect deprojection of the CIB is not possible because its signal is not composed of only one SED. We adopted the modified blackbody spectrum given in Eq. (8), with $A_{\rm CIB} = 1$, $T_{\rm dust} = 10\,\mathrm{K}$, and $\beta = 1.6$. Parameter values were calculated by computing the average for each of the parameter maps derived in Sect. 3.2.

The top panel of Fig. 8 shows that the CIB noise residual ($r_{\rm CIB}$) power spectrum is reduced by the CILC deprojection for both SO LAT and SO LAT+FYST, but it is not zero because the CIB is composed of multiples spectra and only one is used for the deprojection. It also shows that the atmospheric noise residual ($r_{\rm Atmo}$) power spectrum increases for SO LAT and decreases for the combined SO LAT+FYST when the CIB is deprojected (see Fig. 7 for comparison). This can be explained by the higher degrees of freedom available with the higher number of frequency channels for SO LAT+FYST, making the CILC noise penalty have a less penalising effect in comparison to when SO LAT is alone. The lower panel of the figure shows that SO LAT+FYST offers a large $\ell \in [1700, 3500]$ window over which the Compton-y signal is not noise dominated, while for SO LAT alone, the signal remains noise dominated (exclusively by the atmospheric noise). For SO LAT alone or for SO LAT+FYST combined, compared to the ILC case, the ratio between the input target signal and the resulting noise is not much higher in general (Fig. 7). This was also seen when applying our pipeline to *Planck* data, as the recovered Compton-y map amplitude is roughly the same as the ILC and CILC. This is not surprising because we also saw that the ILC and CILC weights are very similar. This observation of the CILC not providing a much lower noise bias Compton-y map compared to the ILC was already reported by Alonso et al. (2018).

5.3. Overview

Figure 9 gives an overview of the relative amplitude changes for the power spectrum of residual noises of each astrophysical component when combining SO LAT and FYST and shows a comparison to SO LAT alone in the two previous cases applying an ILC (in blue) and a CILC deprojecting the CIB (in orange). To see how much SO LAT+FYST reduces residual noise power spectra in comparison to SO LAT alone, we took the ratio of the residual power spectra of each component in both cases:

$$R = \frac{C_\ell^{XX}(\mathrm{SO\,LAT{+}FYST})}{C_\ell^{XX}(\mathrm{SO\,LAT})}, \tag{23}$$

where $C_\ell^{XX}(\mathrm{SO\,LAT{+}FYST})$ is the power spectrum of the component X or of the residual noise r_X for SO LAT+FYST and $C_\ell^{XX}(\mathrm{SO\,LAT})$ is the same quantity but for SO LAT alone. For the extracted Compton-y map ($y_{\rm ILC}$), the blue violin plot shows a mean improvement of around 4% for SO LAT+FYST compared to SO LAT alone but with a large scatter. Especially at $\ell < 2800$, SO LAT performs better (see bottom panel of Fig. 7). When deprojecting the CIB (orange violin plot), we moved to a mean improvement of around 8% in the ILC-extracted Compton-y map for SO LAT+FYST compared to SO LAT alone. In this case, the scatter is also much smaller than before because of the high number of degrees of freedom offered by the combined SO LAT+FYST frequency bands, and the CILC noise

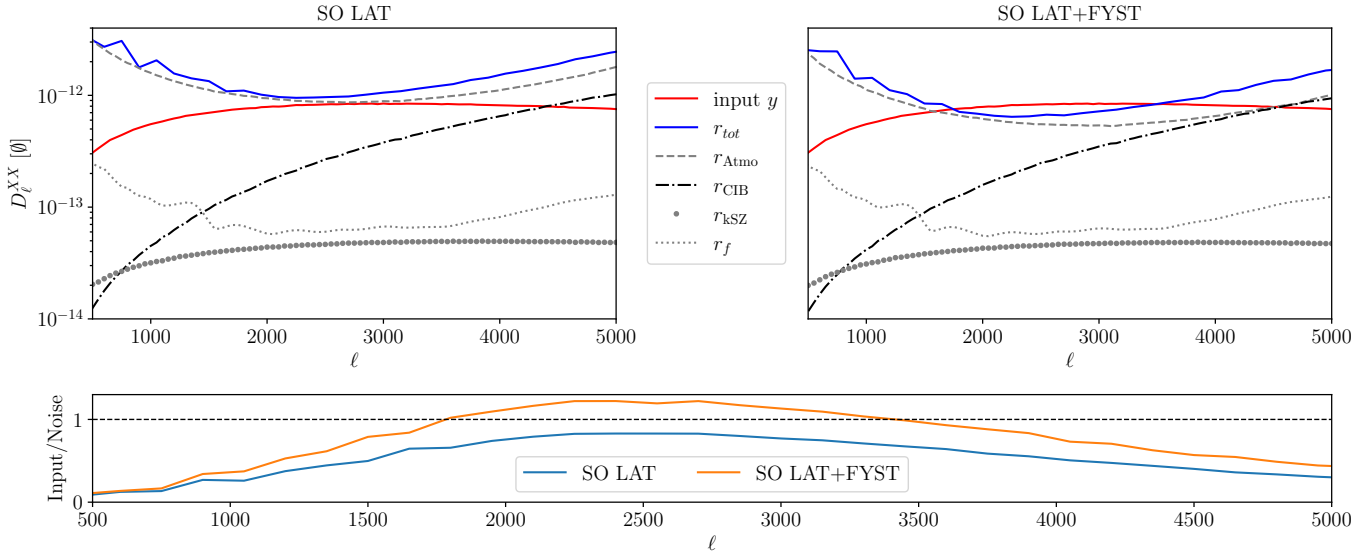

Fig. 8. Power spectra of the CILC residual noises when deprojecting CIB compared to the input Compton-y power spectrum from the Websky simulations (red). The simulated sky contains all extragalactic components (see Table 1), Galactic foregrounds, and atmospheric noise. Top panels: power spectra of various quantities. The grey curves show the ILC noise residual power spectra, denoted by r_X (X being the kSZ); the Galactic foregrounds (f); and the atmospheric noise (Atmo). The black curve shows the power spectrum of the CIB residual noise. The blue curve shows the sum of all grey and black curves, that is, the total ILC noise residual debiased for instrumental white noise and the CMB. The top-left panel shows the results for the simulated SO LAT data, while the top-right panel shows the results for the simulated data for SO LAT and FYST combined. Bottom panel: ratio between the input Compton-y power spectrum (red) and the total noise residual (blue) for SO LAT and SO LAT+FYST. All power spectra have been beam corrected and bin averaged over a window $\Delta\ell = 50$.

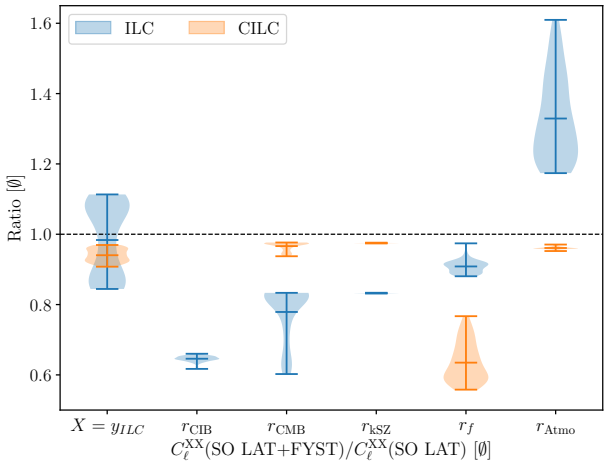

Fig. 9. Value of the ratios of the power spectra of X for SO LAT+FYST compared to SO LAT(C_ℓ^{XX}(SO LAT+FYST)/C_ℓ^{XX}(SO LAT)). The ILC-extracted Compton-y signal ($y_{\rm ILC}$) and r_X represent the residual noise of the quantity X. The shaded regions are the density distribution of the ratios over the ℓ range. The bar is the mean averaged over the ℓ range of the ratio distribution. In blue is the ILC case (see Sect. 5.1) and in orange is the CILC case when CIB is deprojected (see Sect. 5.2).

penalty is much lower for SO LAT+FYST compared to SO LAT alone, making the combination strictly superior to extract a less biased tSZ power spectrum in the CILC case.

For the residual CIB noise ($r_{\rm CIB}$), the power spectrum is on average around 35% lower when combining SO LAT and FYST in the ILC case. In the CILC case, it drops to ∼7%, which might not be surprising because the CILC focusses on nulling as much CIB as possible no matter the frequencies. The addition of FYST does reduce the CIB residual even more but not as much as without the null constraint. Another reason for the drop in improvement when combining SO LAT and FYST in the CILC case might be the higher atmospheric noise at the high frequencies where the CIB signal dominates, undermining the advantage of FYST to constrain and remove CIB.

In the ILC case, the CMB residual noise ($r_{\rm CMB}$) power spectrum is reduced by ∼22% when combining SO LAT with FYST compared to SO LAT alone. Because the ILC cannot separate CMB and kSZ due to their identical spectral frequency distribution, the residual kSZ noise ($r_{\rm kSZ}$) is also suppressed when combining SO LAT and FYST, on average by around 18%. When deprojecting CIB with a CILC, adding FYST to SO LAT also reduces the CMB and kSZ residual noises but only by ∼5% over all scales on average.

The sum of all galactic foreground noise residuals is defined as the cumulative foreground residual ($r_{\rm f}$). Its noise power spectrum is improved on average by ∼10% for SO LAT+FYST compared to SO LAT in the ILC case. In the CILC case, the cumulative foreground residual power spectrum is reduced by ∼30% when combining FYST with SO LAT. This could be due to the additional constraint on the CIB dust spectrum, nulling part of the Galactic dust contribution, which is more easily nulled with FYST high-frequency coverage.

The FYST high-frequency channels are more contaminated by the atmosphere. For this reason, the atmospheric red noise residual ($r_{\rm Atmo}$) power spectrum is ∼30% higher when SO LAT and FYST are combined rather than when SO LAT is alone in the ILC case. When deprojecting the CIB, the atmospheric noise residual becomes slightly lower ∼1% for SO LAT+FYST compared to SO LAT alone. This is due to the noise penalty generated by the additional constraint in the CILC worsening the atmospheric noise for LAT alone compared to SO LAT+FYST. Because SO LAT+FYST benefits from additional frequency bands and therefore more degrees of freedom, it is more resilient to the inevitable noise increase caused by additional spectral constraints.

6. Summary and discussion

Using full-sky HEALPix template maps from the Websky simulations (Stein et al. 2020) and PySM, we created a pipeline that allowed us to simulate the high-resolution maps of the different microwave sky emissions at any beam, frequency, and sensitivity. Based on the Choi et al. (2020a) noise modelling, our pipeline also simulates atmospheric red noise at the locations of the Simons Observatory and FYST.

We simulated the sky that would be seen during experiments at these locations and used a map-based ILC and CILC to extract back the thermal Sunyaev–Zeldovich (tSZ) effect from this multi-frequency, multi-component sky. To minimise Galactic foreground contamination, we used a Galactic mask derived from *Planck* (see Fig. 4). To optimise our ILC weights in order to take advantage of the spatial distribution of some of the microwave emissions, we tessellated the sky using the HEALPix equal-area nested tessellation scheme. We made use of CILC to deproject the CIB by using a simple modified blackbody spectrum to represent the CIB frequency variation and constrain the ILC weights to nullify it. We then used PyMaster to compute the power spectrum and its associated residual noise power spectra, which are the noise coming from all the other microwave sky emissions left by ILC and CILC in the recovered map.

We note that our study does have certain shortcomings that should be taken into account. For example, the Websky simulations do not yet offer a map of RPSs, and these extragalactic components are therefore absent from our study. The RPSs are an important contaminant of the tSZ signal, as they are located in galaxy clusters. If RPSs are not considered, they can fill up the tSZ decrement and bias an estimation towards a lower estimated tSZ signal. A solution to this problem is to mask the RPSs, but this also lowers the estimated tSZ power spectrum, as it removes parts of the tSZ signal. This effect especially impacts measurements of the tSZ signal below 90 GHz, which we do not study here. A study from Holder (2002) showed that RPS subtraction can lead to an underestimation of the tSZ power spectrum of between ∼30% and ∼50%, depending on the scale. We did not consider the relativistic correction of the tSZ effect, which can lead to a significant underestimation of the signal (Remazeilles et al. 2018). This correction is left to future studies. Regarding the method, our tessellation scheme can lead to border effects that were reported by Eriksen et al. (2004) but not accounted for. Our pipeline was tested on *Planck* data with and without tessellation and only a marginal (i.e. <5%) difference could be seen in the power spectrum amplitude. Applying ILC in map-based space also has the disadvantage of the necessary resolution downgrade of all maps to the lower beam channel, which is 2.2′ for SO LAT, thus reducing the benefits of FYST's higher resolution channels. This also prevented us from using *Planck* data that are essential to lowering atmospheric and CMB contamination on large scales. This was addressed by performing CMB debiasing of the residual CMB noise in the extracted Compton-y map. A Fourier space or needlet ILC implementation is left for future work.

We find that when using an ILC to extract tSZ, the combination of SO LAT with FYST reduces the CIB residual noise power spectrum by ∼35%, on average, on scales $\ell \in [500, 5000]$ when compared to SO LAT alone. For the CMB and kSZ components, the addition of FYST to SO LAT reduces the residual noise by ∼22% on average, with a strong scale dependence for the CMB, as the addition of FYST mainly helped on large scales. For the kSZ component, the reduction observed when adding FYST to SO LAT is not scale dependent. Adding FYST also helps reduce the Galactic foreground contamination by, on average, ∼10% on scales $\ell \in [500, 5000]$. However, FYST's high-frequency channels are more affected by atmospheric noise due to poorer transmission. Therefore, the residual noise power spectrum coming from the atmosphere is on average 30% higher when SO LAT and FYST are combined compared to when SO LAT is alone. This greatly affects the total residual noise budget, lowering the benefits of adding FYST to SO LAT for the tSZ recovery to only an average ∼4%, and on larger scales, SO LAT alone performs better. We find that over $\ell \in [2100, 2700]$, the ILC-extracted tSZ power spectrum is not noise dominated for SO LAT alone. But on smaller scales, it is extremely biased, and its noise budget is dominated by the CIB. When SO LAT is combined with FYST, the CIB noise residual dominates much later $\ell \sim 4200$, but because the atmospheric noise is higher, there is no window over which the ILC-extracted tSZ power spectrum is not noise dominated.

We note that the CIB residual noise reduction that FYST brings to SO LAT is not something that SO LAT could achieve on its own by simply having a longer survey time. This was tested by reducing the instrumental white noise in SO LAT bands by two, and we observed a reduction of the ILC residual instrumental noise in the reconstructed Compton-y map, but the other noise components remained the same. This 'boosted' version of SO LAT alone did not have a lower CIB residual power spectrum compared to the planned SO LAT or SO LAT plus FYST combination.

Because tSZ and CIB have a non-zero correlation in the Websky simulations and because FYST's high frequencies probe the CIB better, we used CILC to deproject the CIB component, modelling it with a simple one-component modified blackbody spectrum. Compared to the ILC case, adding FYST to SO LAT lowers the CIB noise residual by only ∼7%. The CMB and kSZ noise residuals are also reduced by ∼5% when combining SO LAT and FYST when using CILC. However, the cumulative foregrounds are suppressed by ∼30%, and atmospheric noise is a bit lower for the SO LAT and FYST combination. Despite a lower reduction of the CIB residual noise when combining SO LAT and FYST, the CIB deprojection case offers a large $\ell \in [1800, 3500]$ window where the CILC-extracted Compton-y map is not noise dominated after instrumental noise and CMB debiasing.

Acknowledgements. The authors would like to thank Kaustuv Basu for his guidance and advice throughout the project. Srinivasan Raghunathan, Jacques Delabrouille, and Frank Bertoldi for their time, valuable insights, and numerous helpful discussions. The referee for useful comments and numerous improvements to the draft. The editor Laura Pentericci for her advice and help on the submission process. Michael Coronado for his language editing work which significantly helped improve the manuscript. The CCAT-prime Collaboration for their feedback on this work and Colin Hill for his help. Laila Linke for her useful comments on this draft. We furthermore acknowledge support from the International Max-Planck Research School (IMPRS) and the Bonn-Cologne Graduate School of Physics and Astronomy (BCGS). The simulations of the CIB used in this paper were developed by the Websky Extragalactic CMB Mocks team, with the continuous support of the Canadian Institute for Theoretical Astrophysics (CITA), the Canadian Institute for Advanced Research (CIFAR), and the Natural Sciences and Engineering Council of Canada (NSERC), and were generated on the GPC supercomputer at the SciNet HPC Consortium. SciNet is funded by the Canada Foundation for Innovation under the auspices of Compute Canada, the Government of Ontario, Ontario Research Fund – Research Excellence, and the University of Toronto.

References

Ade, P., Aguirre, J., Ahmed, Z., et al. 2019, JCAP, 2019, 056
Alonso, D., Colin Hill, J., Hložek, R., et al. 2018, Phys. Rev. D, 97, 093514

Alonso, D., Sanchez, J., Slosar, A., et al. 2019, MNRAS, 484, 4127
Battaglia, N., Bond, J. R., Pfrommer, C., et al. 2012, ApJ, 758, 75
Bennett, C. L., Bay, M., Halpern, M., et al. 2003, ApJ, 583, 1
Bolliet, B. 2018, arXiv e-prints [arXiv:1806.04786]
Bustos, R., Rubio, M., Otárola, A., et al. 2014, PASP, 126, 1126
CCAT-prime Collaboration (Aravena, M., et al.) 2023, ApJS, 264, 7
Chluba, J., & Sunyaev, R. A. 2012, MNRAS, 419, 1294
Choi, S. K., Austermann, J., Basu, K., et al. 2020a, J. Low Temp. Phys., 199, 1089
Choi, S. K., Hasselfield, M., Ho, S.-P. P., et al. 2020b, JCAP, 2020, 045
Delabrouille, J., Cardoso, J.-F., Le Jeune, M., et al. 2009, A&A, 493, 835
Dickinson, C., Ali-Haïmoud, Y., Barr, A., et al. 2018, New Astron. Rev., 80, 1
Dunkley, J., Calabrese, E., Sievers, J., et al. 2013, JCAP, 2013, 025
Eriksen, H. K., Banday, A. J., Górski, K. M., et al. 2004, ApJ, 612, 633
Eriksen, H. K., Jewell, J. B., Dickinson, C., et al. 2008, ApJ, 676, 10
Erler, J., Basu, K., Chluba, J., et al. 2018, MNRAS, 476, 3360
Fixsen, D. J. 2009, ApJ, 707, 916
Fixsen, D. J., Dwek, E., Mather, J. C., et al. 1998, ApJ, 508, 123
Fornazier, K. S. F., Abdalla, F. B., Remazeilles, M., et al. 2022, A&A, 664, A18
George, E. M., Reichardt, C. L., Aird, K. A., et al. 2015, ApJ, 799, 177
Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Haslam, C. G. T., Klein, U., Salter, C. J., et al. 1981, A&A, 100, 209
Haslam, C. G. T., Salter, C. J., Stoffel, H., et al. 1982, A&AS, 47, 1
Hauser, M. G., Arendt, R. G., Kelsall, T., et al. 1998, ApJ, 508, 25
Holder, G. P. 2002, ApJ, 580, 36
Hurier, G. 2015, A&A, 575, A11
Hurier, G., Macías-Pérez, J. F., & Hildebrandt, S. 2013, A&A, 558, A118
Itoh, N., Kohyama, Y., Nozawa, S., et al. 1998, J. Phys. Condens. Matter, 10, 11273
Kim, J., Naselsky, P., & Christensen, P. R. 2008, Phys. Rev. D, 77, 103002
Komatsu, E., & Seljak, U. 2002, MNRAS, 336, 1256
Lenz, D., Doré, O., & Lagache, G. 2019, ApJ, 883, 75
Mather, J. C., Cheng, E. S., Cottingham, D. A., et al. 1994, ApJ, 420, 439
Melin, J.-B., Bartlett, J. G., Cai, Z.-Y., et al. 2018, A&A, 617, A75
Melin, J.-B., Bartlett, J. G., Tarrío, P., et al. 2021, A&A, 647, A106
Menanteau, F., González, J., Juin, J.-B., et al. 2010, ApJ, 723, 1523
Miville-Deschênes, M. A., Lagache, G., Boulanger, F., et al. 2007, A&A, 469, 595
Mroczkowski, T., Nagai, D., Basu, K., et al. 2019, Space Sci. Rev., 215, 17
Niemack, M. D. 2016, Appl. Opt., 55, 1686
Park, C.-G., Park, C., & Gott, J. R., III 2007, ApJ, 660, 959
Partridge, R. B., & Peebles, P. J. E. 1967, ApJ, 148, 377
Planck Collaboration I. 2014, A&A, 571, A1
Planck Collaboration IX. 2014, A&A, 571, A9
Planck Collaboration XXI. 2014, A&A, 571, A21
Planck Collaboration IX. 2015, A&A, 594, A9
Planck Collaboration XXII. 2016, A&A, 594, A22
Planck Collaboration XXV. 2016, A&A, 594, A25
Planck Collaboration XXVII. 2016, A&A, 594, A27
Planck Collaboration IV. 2020, A&A, 641, A4
Pointecouteau, E., Giard, M., & Barret, D. 1998, A&A, 336, 44
Puget, J.-L., Abergel, A., Bernard, J.-P., et al. 1996, A&A, 308, L5
Reichardt, C. L., Shaw, L., Zahn, O., et al. 2012, ApJ, 755, 70
Reichardt, C. L., Patil, S., Ade, P. A. R., et al. 2021, ApJ, 908, 199
Remazeilles, M., Delabrouille, J., & Cardoso, J.-F. 2011, MNRAS, 410, 2481
Remazeilles, M., Dickinson, C., Banday, A. J., et al. 2015, MNRAS, 451, 4311
Remazeilles, M., Bolliet, B., Rotti, A., & Chluba, J. 2018, MNRAS, 483, 3459
Rogers, K. K., Peiris, H. V., Leistedt, B., et al. 2016, MNRAS, 460, 3014
Schaan, E., Ferraro, S., Amodeo, S., et al. 2021, Phys. Rev. D, 103, 063513
Shang, C., Haiman, Z., Knox, L., & Oh, S. P. 2012, MNRAS, 421, 2832
Staniszewski, Z., Ade, P. A. R., Aird, K. A., et al. 2009, ApJ, 701, 32
Stein, G., Alvarez, M. A., Bond, J. R., et al. 2020, JCAP, 2020, 012
Sunyaev, R. A., & Zeldovich, Y. B. 1970, Comments Astrophys. Space Phys., 2, 66
Sunyaev, R. A., & Zeldovich, Y. B. 1972, Comments Astrophys. Space Phys., 4, 173
Sunyaev, R. A., & Zeldovich, Y. B. 1980, ARA&A, 18, 537
Tegmark, M., de Oliveira-Costa, A., & Hamilton, A. J. 2003, Phys. Rev. D, 68, 123523
Thorne, B., Dunkley, J., Alonso, D., & Næss, S. 2017, MNRAS, 469, 2821
Viero, M. P., Wang, L., Zemcov, M., et al. 2013, ApJ, 772, 77
Vikram, V., Lidz, A., & Jain, B. 2017, MNRAS, 467, 2315
Wright, E. L. 1979, ApJ, 232, 348
Zubeldia, I., Chluba, J., Battye, R., et al. 2023, MNRAS, 522, 5123

Appendix A: *Planck* data

A.1. Internal linear combination and constrained internal linear combination weights

We applied our pipeline to *Planck* High-Frequency instrument (HFI) Data Release 2[3] (DR2; Planck Collaboration XXII 2016) in order to see how well we could recover the true y using an ILC. We also considered a CILC to deproject CIB by constraining the weights so that they have a null response to the frequency spectrum of the CIB (see Eq. (20)). However, the CIB deprojection is not perfect, because the CIB frequency spectrum depends on the population of the sources and their redshift.

Looking at the ILC weights gave us information about which channel contributes the most to extracting the tSZ signal and minimising the noise. These were the two conditions that the weights had to satisfy (see Eq. (13) and (14)). The ILC weights ponderated by their respective frequency mixing vector value showed the fractional contribution of each frequency channel to the recovered y. The fractional contribution of the ∼220 GHz band was null because the value of the mixing vector at this frequency is null.

For a map-based ILC applied to *Planck* data (see blue points in Fig. A.1), we concluded from the weight that the 143 GHz channel is the one that contributes the most, with a factor of $\omega_{143}a_{143} \approx 0.7$, to the reconstructed final map as well as to minimising the noise. The 217 GHz channel does not contribute to the final reconstructed y map because the tSZ signal is null at this frequency. However, its high ω_{217} value makes it the main contributor to noise minimisation. This was expected, as the tSZ signal is absent at this frequency and the map contains all the noise contaminants. Next are the contributions of the 353 GHz channel and the 545 GHz channel. The contribution values of those channels are consistent with previous studies (Erler et al. 2018). In the case where a CILC is applied to *Planck* data to deproject the CIB (orange points, Fig. A.1), we saw that the weights are almost identical to the ILC ones. This is expected, as Alonso et al. (2018) also found that deprojecting CIB with a CILC does not result in much less CIB contamination in the reconstructed tSZ profile, arguing that this shows the ILC already does a sufficient job of cleaning the CIB.

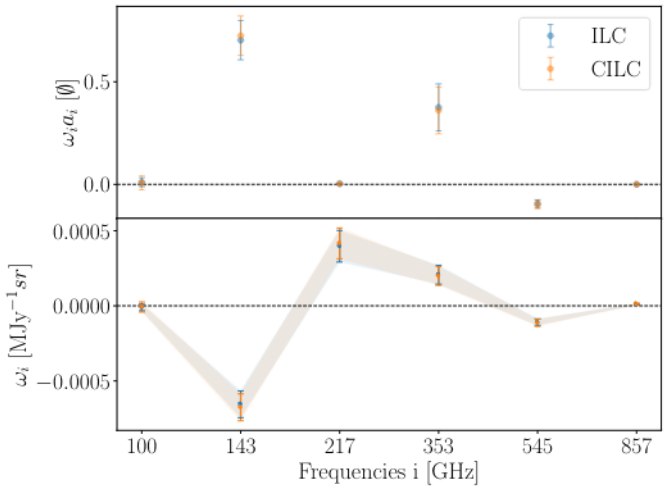

Fig. A.1. Internal linear combination (blue) and CILC (orange) weights with CIB deprojected on *Planck* data. Top figure: Fractional contribution of each frequency channel to the recovered y-map. Bottom figure: Frequency channels that contribute to maximising the tSZ signal and minimising the noise. The error bars show the standard deviation around the mean weight of all of the 192 tessellated fields for each frequency map.

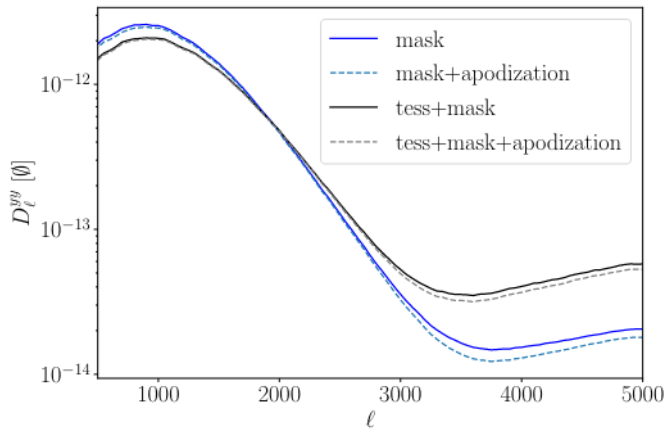

Fig. A.2. Power spectra of the Compton-y map using our ILC pipeline on *Planck* DR2 with different configurations. In blue is the power spectrum of the extracted map with a simple $f_{sky} = 0.6$ Galactic mask. In black is the same thing but with a tessellation of the sky. The dotted lines (blue and black) indicate the inclusion of apodisation in each case.

A.2. Processing of the maps

We tested the consistency and effect of each of the post-processing steps applied on the map (see Fig. A.2). The blue curve in the figure shows the ILC-extracted power spectrum computed with a Galactic mask and without tessellation (see Fig. 4), while the lighter blue dotted curve shows the same but with apodisation. The black curve shows the case with the tessellation of the sky and the ILC being applied to each rhombus separately. The grey dotted line shows when apodisation was also applied. For our particular case of tSZ extraction, the mask apodisation only lowers the power spectra of the maps by less than 5%. Comparing the power spectrum with tessellation (black) and without (blue) showed that tessellation helps reduce the noise on large scales, up to $\ell \approx 2000$, by a factor of approximately two. At $\ell = 2500$, it increases the noise by around 15%. This can be understood as the size of the rhombus being small enough to break down regions highly contaminated by unwanted astrophysical components on large scales and to weigh down such regions in order to better recover the tSZ signal. But the rhombus size might be too big to optimally deal with smaller region noise. Moreover, with 192 total rhombuses over the sky, we might be creating border effects on smaller scales and thus increasing the noise residual on those scales. The scales ($\ell > 2200$) are also affected by the ∼9.66' beam.

[3] https://irsa.ipac.caltech.edu/data/Planck/release_2/all-sky-maps/

Appendix B: Constrained ILC: Deprojecting the cosmic microwave background

When applying an ILC to extract back a Compton-y map, the CMB is the dominant noise residual ($r_{\rm CMB}$) on large scales (see Fig. B.1). Using CILC to deproject the CMB, the residual CMB noise is suppressed by three orders of magnitude. However, the additional constraints imposed on the weights to deproject the CMB led to a smaller solution space and therefore increased noise. This led to a one order of magnitude increase of the

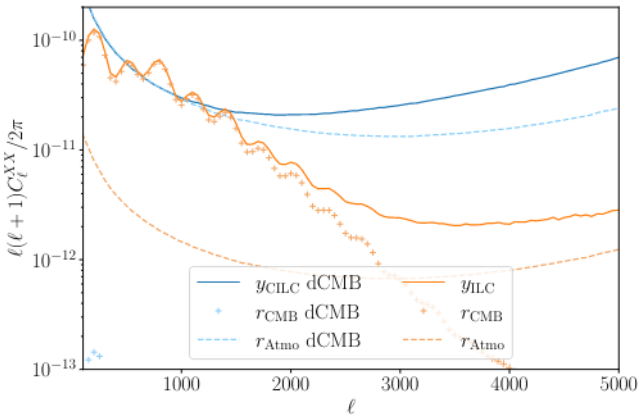

Fig. B.1. Power spectra of ILC/CILC (full lines) and their noise residuals (dotted lines). Specifically, the orange curve is the power spectrum of the ILC-extracted Compton-y map. The lighter orange lines are the noise residuals. Specifically, they are the ILC residual CMB noise ($r_{\rm CMB}$) and $r_{\rm Atmo}$ for the atmospheric noise. The blue lines represent the case with CILC where the CMB is deprojected. The solid, darker blue curve is the power spectrum of the CILC-extracted Compton-y map. The lighter blue curves are the noise residuals.

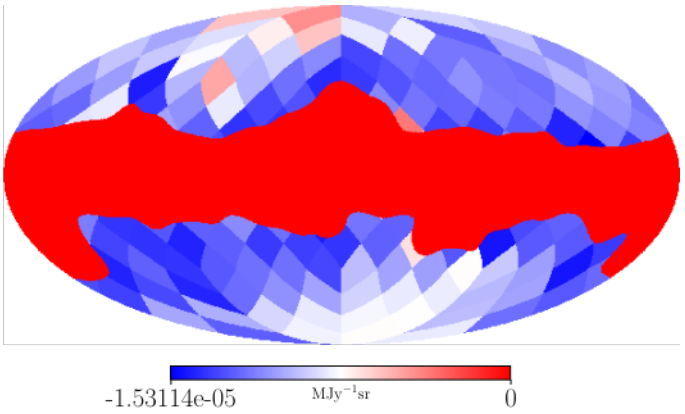

Fig. C.1. Maps of the weights for each patch of the sky where the ILC was applied for the 860 GHz frequency band. The weights were computed from simulations that included all Galactic foregrounds and extragalactic components (see Table 1) at the sensitivities, beam sizes, and frequencies of SO LAT and FYST (see Table 2) and with 90% correlated atmospheric red noise generated using the noise curves given by Choi et al. (2020a), (see Fig. 3).

atmospheric residual noise ($r_{\rm Atmo}$), which worsens the estimate of the extracted Compton-y power spectrum, especially on small scales, by more than one order of magnitude. This is commented on extensively in Sec. 5.2 and was tested using our Skymodel pipeline (see Sec. 3).

Appendix C: ILC weights

In our study, we tesselated the sky into 192 rhombi of equal area using the HEALPix nested scheme. This sky separation allowed us to better adapt to the spatially dependant contaminant of some Astrophysical emissions, such as the Galactic foregrounds. The ILC we applied on each zone resulted in one weight per frequency and per rhombus. An example of a full sky weights map produced and used by our ILC can be seen in Fig. C.1.

Appendix D: Simulated maps

In this work, we used the Skymodel pipeline based on PySM and the Websky simulations to generate mock maps of the microwave sky at the frequencies, sensitivities, and instrument beam size of the upcoming SO LAT and FYST (see Table 2). The atmospheric red noise was generated using the noise curves of Choi et al. (2020a) and the red noise of the neighbouring frequency channels of a given telescope were correlated at 90% using Eq. 10. The resulting simulated microwave sky maps used can be seen in Fig. D.1.

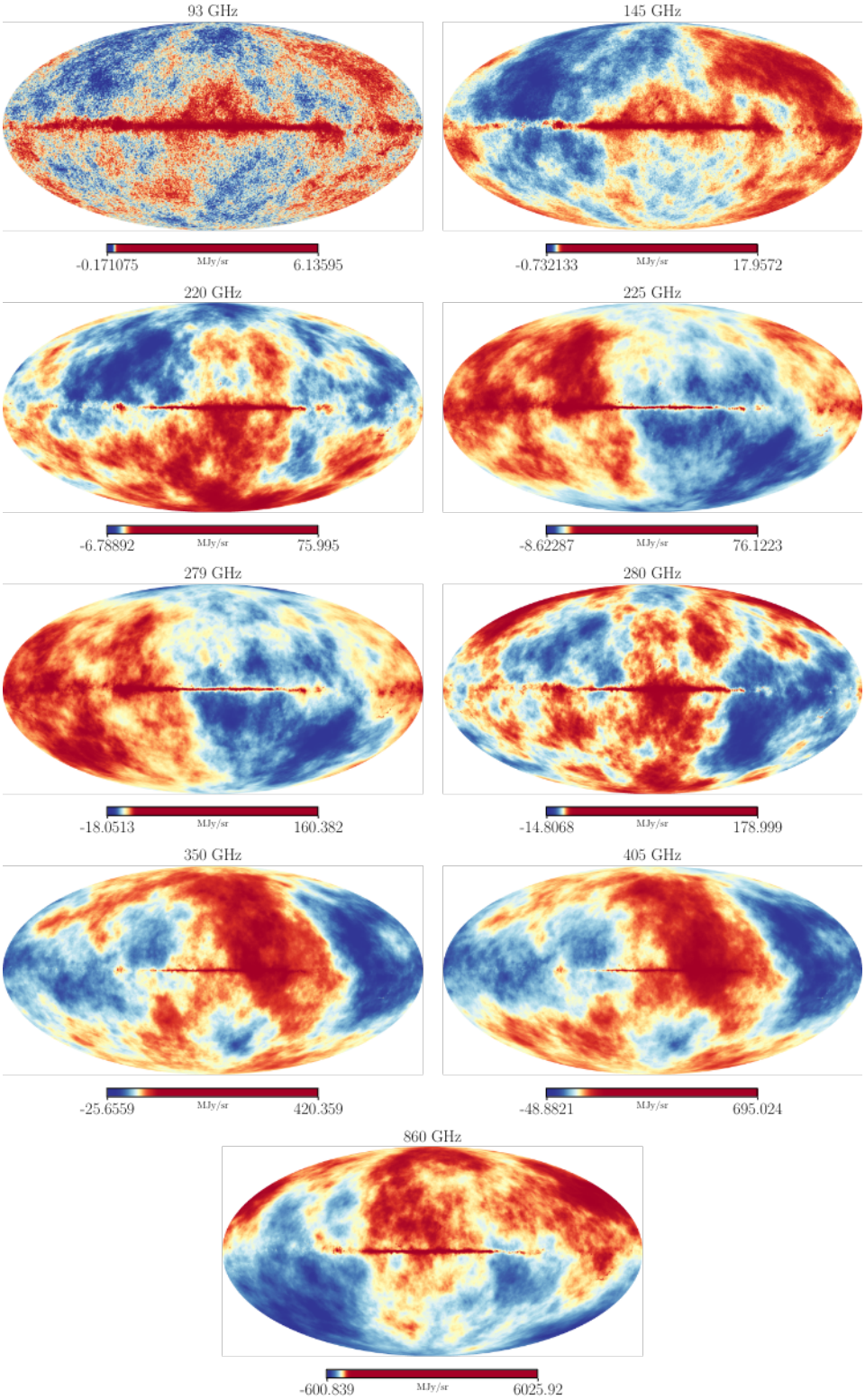

Fig. D.1. Mock HEALPy maps of the microwave sky simulated using the Skymodel pipeline (see Sec. 3). All Galactic foregrounds and extragalactic components (see Table 1) at the sensitivities, beam sizes, and frequencies of SO LAT and FYST (see Table 2) and the 90% correlated atmospheric red noise between neighbouring channels generated using the noise curves given by Choi et al. (2020a) were included (see Fig. 3).

C.4 AtLAST power spectra

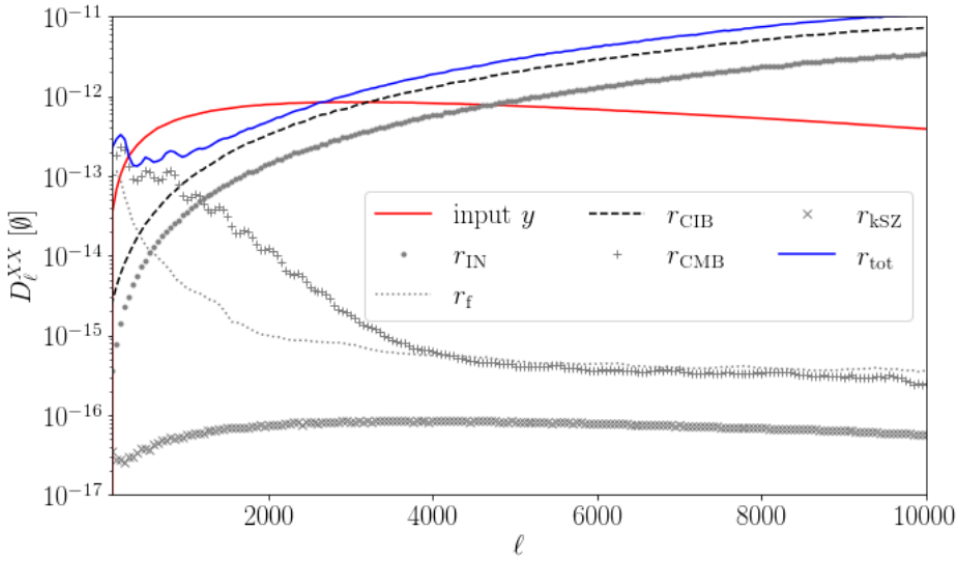

Figure C.5: Power spectra of various components for an AtLAST-like experiment in Case 1 (see Tab. 5.2). In red is the input target tSZ signal, and in blue is the total ILC residual noise (r_{tot}). All the other curves decompose the total noise residual into the fractional contribution of each astrophysical emission in the sky. The dotted black curve is the CIB noise residual ($r_{\rm CIB}$) and all the other grey curves are the other component noise residuals. All spectra are beam deconvolved.

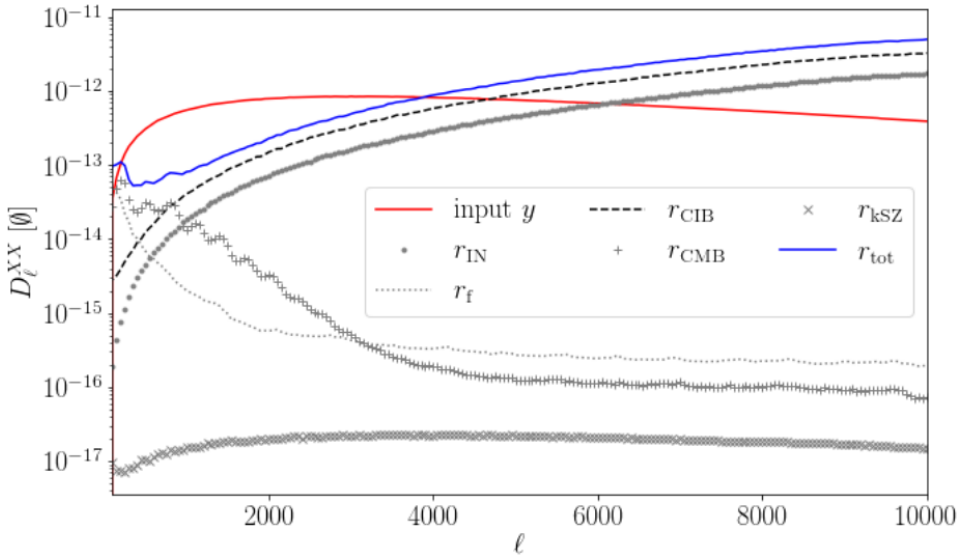

Figure C.6: Power spectra of various components for an AtLAST-like experiment in Case 3 (see Tab. 5.2). In red is the input target tSZ signal, and in blue is the total ILC residual noise (r_{tot}). All the other curves are the decomposition of the total noise residual into the fractional contribution of each astrophysical emission in the sky. The dotted black curve is the CIB noise residual ($r_{\rm CIB}$) and all the other grey curves are the other component noise residuals. All spectra are beam deconvolved.

List of Figures

List of Tables

Listings

Acknowledgements

During the PhD:

I would first like to thank Frank Bertoldi and Kaustuv Basu for offering me the opportunity to join their group and conduct research work under their supervision. The knowledge they shared with me and their working style has and will continue to influence my own work.

A huge thanks to Jens Erler, who was my PhD student role model, and who helped supervise me throughout my own work. As a fellow group member at first and later as a collaborator and invested co-author. For all of those evenings, you invested in discussing with me and investigating the open questions of my projects. I am also very fortunate to be able to call you my friend, for all of the evening fighting for victory on the badminton court, for all the fun at parties but also the comfort you provided to share the more personal things of life whether good or bad.

This scientific work received very useful input from Srinivasan Raghunathan and Jacques Delabrouille who both found the time to meet several times to discuss interpretation and open questions. I would like to thank them for that.

I would like to thank Rainer Mauersberger and Cristiano Porciani, who supported me as members of my TAC meeting and provided advice, and insights for the conduction of my PhD.

I would like to thank the International Max-Planck Research School (IMPRS) of Bonn and Cologne and the Bonn Cologne Graduate School (BCGS) for their regular funding to attend conferences, this was always an inspiring moment of my scientific work and allowed me to expand my experience of the world.

A huge thank you to Hans Nguyen for all the fun on the badminton court, the crazy discussions with Henning Hilmarsson, which always lightened my day, and his friendship, support and feedback on the introduction of this thesis.

I would like to thank the institute secretaries Christina Stein-Schmitz who took the time to accompany me to register at the city hall when I could not speak any German and Sabine Derdau who consistently helped me through the ups and downs of bureaucratic work. Zum Andreas Bödewig für seine Hilfe bei Computerproblemen, aber auch für die Menge Spaß beim Kicker während die Mittagspause. Dank ihm bin ich jetzt viel besser beim Kicker!

Thank you to Gunther Witzel for his support through the IMPRS and for initiating the Friday social.

I would like to thank Manuel Drees and Dennis Lehmkuhl for agreeing to be members of my committee.

A huge thanks to Sandra Unruh, who is such an inspiration to me, as a scientist, woman, and mother and who supported me when times were hard. I love your sarcasm, your sense of a nice pun, and your cynicism at times. We share so much, I am fortunate to have you as a friend.

A huge thanks to Cosima Eibensteiner, for all the nice and fun evening spent discussing together, for her support through difficult times, her softness, her humor, her advice, and for being such a good person

and dear friend of mine with whom I share so much.

I would like to thank Kevin Weber for all the fun and crazy discussions on Discord, the evening playing Borderland 3 and the introduction to the Warhammer Universe! thanks also to the ALARM group, which was a nice proof that you can put a bunch of people from all over the world together, and if no one takes themselves seriously, joke about anything and end up loving each other.

I would like to thank my amazing colleagues and friends: Christos Karoumpis for all the crazy fun discussions but also for his reassuring take on life. Sven Heydenreich for always being so nice and funny to be around, and for his epic parties at their place with Davit Alkhanshvili. David Aguilera Dena for his support and constant fun attitude. Markus Keil for his humor and my first-ever Dungeon & Dragon game. Hannah Zohren for her softness and all the incredible nights playing board games, especially Love Letter with Fatimah Raihan and Beatriz Hernandez Martin. Jakob den Brok for all the super interesting discussions about politics and mental support. Kostantinos Migkas for the fun we had in Sesto during my first-ever conference week, and his super nice late-summer BBQs. Vanessa Frohn for her support. Ana Paola Mikler for her support and the nice evening spent together in Amsterdam. Vyoma Muralidhara and Ankur Dev for the nice week we spent in New York and their mental support. Rovina Pinto for the nice discussions. Ralf Timmermann for his support.

Thank you to, Franzi Leupold, Nikolaus Sulzenauer, Abel Schootemeijer, Aaron Beyer, Angie Veronica, and Jonah Wagenveld I always have a nice time whenever I speak with you!

Herzlichen dank an Sebastian meine Tandempartner, Jabok den Brok, Cosima Eibensteiner, Nikolaus Sulzenauer, Kevin Weber, Andreas Bödewig, Ralf Timmermann, Reinhold Schaaf, Lukas Fleischmann, Sabine Derdau, Sandra Unruh, Jens Erler, David Ohse und alle deutschen, österreichischen, bayerischen oder schweizerischen, die mir durch ihre Geduld und Hilfe geholfen haben, mein Deutsch zu verbessern. Ohne euch wäre nichts! Maude? was? OBJEKT! *Ohne euch wäre ES nichts!

Ich möchte mich auch bei Deutschland bedanken, wegen die Möglichkeit hier eine Doktorarbeit zu führen und hier zu leben. Ich habe die letzte vier Jahren eine neue Kultur entdeckt und das war eine riesige Chance.

Early science work:

I would like to Thank Hervé Dole for offering me the chance to have my first-ever research experience in his group. A huge thank you to Benjamin Clarenc for supervising me and helping me through this internship and to the Institut d'Astrophysique Spaciale.

I would like to thank Nobuhiko Katayama and Anne Ducout, for allowing me to do my second internship at the Institut for the Physics and Mathematics of the Universe in Japan. There, I would like to Thank Ben Thorne, Louis Papin, Robin Schneider, Tomislav Vladisavljević, Hiroko, and Li Xiangchong for making my stay and everyday work in Japan fun.

I would like to thank Sylvain de la Torre and l'Institut d'Astrophysique de Marseille, for the opportunity of working there. In particular to my colleagues, who made my stay so fun Eva Daoud, Peter Lustig, Romain Paviot also for his help in moving to Germany, Junais.

I would also like to thank some notable teachers, who made a positive difference in my science work and life: M. Martinez my French teacher, my CM1 teacher, Linoel Girard Maths teacher at the Lycée Claude Bernard in Paris, Patrick Puzo director of the Magistère de physique of Paris XI and Claire Marrache-Kikuchi.

Throughout my life:

Dans les comptes pour enfants, on dit souvent que l'héroïne est née sous une bonne étoile ou a une bonne fée, ça a été mon cas, en la personne de ma tante, Nadine Poncet, que je ne pourrais jamais remercier suffisament, pour le calme qu'elle a su m'apporter dans les moments les plus difficiles de ma vies mais aussi pour son soutien face à l'adversité, face à des gens pour qui la remise en cause et le doute n'existe pas. Un grand merci à Christophe Poncet dont l'hummour m'a toujours apporté beaucoup de bonheur.

To Jacques Prost, who was my partner for 10 years, and whom I have admired for his "sense de la répartie", his humor, and his practical intelligence. You brought joy to my life, a first try at love, tons of nice holiday memories around Europe, and support in some ways, and for that I am forever grateful. Without your help, I could not have done it.

To the members of the EgocentriK guild on Flyff, Destiné server who unknowingly helped me during the most difficult period of my life, by providing a fun space to escape to online, every evening and throughout the many sleepless nights. In particular, to Nicolas Luisi, who offered me a home when I had nowhere else to go. Alexandre Tertre who kept my belongings when I needed it. Anne-Sophie Rayes who helped me discover the Parisian shopping life. Kévin Couvreur, Frédéric Batault, Pauline Amor, Zakaria Chbicheb, Luis Marcelino.

À mes grands-parents, Monique et Michel Couturier, pour tout le bonheur qu'ils m'ont apporté enfant et tous les soutiens durant mon adolescence, ces moments de calme et d'amour ont été une douce pause durant la tempête.

To Anais Schwab, who was my best friend for so long, thank you so much, for all the fun we had. To Emeline Defalque my best friend in high-school, for all the fun English classes where we spoke about Manga and Animés instead of listening. To Pauline Josserand whom I have always admired and with whom I had so much fun as a kid, for all the Diddl letters we sent to others. To Roxane Taleb, a wonderful person who supported me through the prépa. But also to many other women whom I am so grateful to have counted among my friends, Florianne Grollau, Sophie Roudier, and Sonia Bouquet.

À ma sœur Éloïse Charmetant, mon frère Tristan Charmetant, mes cousins et cousines, Isabelle Zerez, Damien Poncet, Hélène Poncet, Manon Poncet, Ophélie Morey, Arnaud Morey.

I would also like to thank, all the people with whom I disagreed, who told me it was impossible, who mocked me, or who hurt me. All of that is a part of life and a learning experience that made me the person I am.

I acknowledge the financial support of my studies from September 2012 to July 2017 from François et Fabienne Charmetant.

Bibliography

Y. S. Abylkairov, O. Darwish, J. C. Hill, and B. D. Sherwin. *Physical Review D*, 103(10), 2021. doi: 10.1103/physrevd.103.103510.

G. E. Addison, J. Dunkley, and D. N. Spergel. *MNRAS*, 427(2):1741–1754, 2012. doi: 10.1111/j.1365-2966.2012.21664.x.

P. Ade, J. Aguirre, Z. Ahmed, and S. Aiola. *Journal of Cosmology and Astroparticle Physics*, (02): 056–056, 2019. doi: 10.1088/1475-7516/2019/02/056.

Y. Ali-Haïmoud, C. M. Hirata, and C. Dickinson. *MNRAS*, 395(2):1055–1078, 2009. doi: 10.1111/j.1365-2966.2009.14599.x.

D. Alonso, T. Louis, P. Bull, and P. G. Ferreira. *Physical Review D*, 94(4), 2016. doi: 10.1103/physrevd.94.043522.

D. Alonso, J. C. Hill, R. Hložek, and D. N. Spergel. *Phys. Rev. D*, 97:063514, 2018. doi: 10.1103/PhysRevD.97.063514.

D. Alonso, J. Sanchez, and A. S. and. *MNRAS*, 484(3):4127–4151, 2019. doi: 10.1093/mnras/stz093.

R. A. Alpher. *nature*, 162:774–775, 1948. doi: 10.1038/162774b0.

F. Argüeso, J. González-Nuevo, and L. Toffolatti. *The Astrophysical Journal*, 598(1):86, 2003. doi: 10.1086/377731.

B. Ballot. *Annalen der Physik*, 142(11):321–351, 1845. doi: 10.1002/andp.18451421102.

K. Basu et al. , 829(2):L23, 2016. doi: 10.3847/2041-8205/829/2/L23.

P. S. Behroozi, R. H. Wechsler, and C. Conroy. *The Astrophysical Journal*, 770(1):57, 2013. doi: 10.1088/0004-637X/770/1/57.

C. L. Bennett et al. *The Astrophysical Journal Supplement Series*, 148(1):97–117, 2003. doi: 10.1086/377252.

C. L. Bennett et al. *The Astrophysical Journal Supplement Series*, 208(2):20, 2013. doi: 10.1088/0067-0049/208/2/20.

M. Bersanelli et al. *Astronomy and Astrophysics*, 520:A4, 2010. doi: 10.1051/0004-6361/200912853.

M. Birkinshaw, S. F. Gull, and K. J. E. Northover. *Nature*, 275:40,41, 1978. doi: 10.1038/275040a0.

M. Birkinshaw, S. F. Gull, and H. Hardebeck. *Nature*, 309:34–35, May 1984. doi: 10.1038/309034a0.

P. Bode, J. Ostriker, and G. Xu. *The Astrophysical Journal Supplement Series*, 128, 2001. doi: 10.1086/313398.

B. Bolliet. *arXiv e-prints*, art. arXiv:1806.04786, 2018.

E. M. Burbidge, G. R. Burbidge, W. A. Fowler, and F. Hoyle. *Rev. Mod. Phys.*, 29:547–650, 1957. doi: 10.1103/RevModPhys.29.547.

J. E. Carlstrom, G. P. Holder, and E. D. Reese. *Annual Review of Astronomy and Astrophysics*, 40(1): 643–680, 2002. doi: 10.1146/annurev.astro.40.060401.093803.

CCAT-prime Collaboration et al., 2021.

CCAT-Prime Collaboration et al. , 264(1):7, 2023. doi: 10.3847/1538-4365/ac9838.

M. Charmetant and J. Erler. *A&A*, 677:A87, 2023. doi: 10.1051/0004-6361/202245834.

S. K. Choi and et al. *jcap*, 2020(12):045, 2020. doi: 10.1088/1475-7516/2020/12/045.

S. K. Choi et al. *Journal of Low Temperature Physics*, 199(3-4):1089–1097, 2020. doi: 10.1007/s10909-020-02428-z.

R. Clausius. *Annalen der Physik*, 217:124–130, 1870.

A. H. Compton. *Phys. Rev.*, 21:483–502, 1923. doi: 10.1103/PhysRev.21.483.

A. de Oliveira-Costa, M. Tegmark, B. M. Gaensler, J. Jonas, T. L. Landecker, and P. Reich. *MNRAS*, 388 (1):247–260, 2008. doi: 10.1111/j.1365-2966.2008.13376.x.

J. Delabrouille and J. F. Cardoso. *arXiv e-prints*, art. astro-ph/0702198, 2007.

J. Delabrouille, J. F. Cardoso, M. Le Jeune, M. Betoule, G. Fay, and F. Guilloux. *A & A*, 493(3):835–857, 2009. doi: 10.1051/0004-6361:200810514.

J. Delabrouille et al. *A & A*, 553:A96, 2013. doi: 10.1051/0004-6361/201220019.

F.-X. Désert et al. *New Astronomy*, 3(8):655–669, 1998. doi: 10.1016/s1384-1076(98)00032-3.

C. Dickinson, R. D. Davies, and R. J. Davis. *MNRAS*, 341(2):369–384, 2003. doi: 10.1046/j.1365-8711.2003.06439.x.

C. Dickinson et al. *New Astronomy Reviews*, 80:1–28, 2018. doi: 10.1016/j.newar.2018.02.001.

K. Dolag. In *IAU General Assembly*, volume 29, page 2250156, 2015.

K. Dolag, S. Borgani, G. Murante, and V. Springel. *MNRAS*, 399(2):497–514, 2009. doi: 10.1111/j.1365-2966.2009.15034.x.

C. Dragone. *The Bell System Technical Journal*, 57:2663–2684, 1978.

B. T. Draine. 2011.

R. Durrer. *General Relativity and Gravitation*, 50, 2018. doi: 10.1007/s10714-018-2440-y.

A. Einstein. *Annalen der Physik*, 354(7):769–822, 1916. doi: 10.1002/andp.19163540702.

T. A. Ensslin and C. R. Kaiser, 2000.

H. K. Eriksen, A. J. Banday, K. M. Gorski, and P. B. Lilje. *The Astrophysical Journal*, 612(2):633–646, 2004. doi: 10.1086/422807.

J. Erler, K. Basu, M. Trasatti, U. Klein, and F. Bertoldi. , 447(3):2497–2502, 2015. doi: 10.1093/mnras/stu2750.

J. Erler, K. Basu, J. Chluba, and F. Bertoldi. *MNRAS*, 476(3):3360–3381, 2018. doi: 10.1093/mnras/sty327.

C. Ertley. In *American Astronomical Society Meeting Abstracts #223*, volume 223 of *American Astronomical Society Meeting Abstracts*, page 118.03, 2014.

S. Ferraro and J. C. Hill. , 97(2):023512, 2018. doi: 10.1103/PhysRevD.97.023512.

S. Ferraro, J. C. Hill, N. Battaglia, J. Liu, and D. N. Spergel. , 94(12):123526, 2016. doi: 10.1103/PhysRevD.94.123526.

D. P. Finkbeiner, M. Davis, and D. J. Schlegel. *ApJ*, 524(2):867–886, 1999. doi: 10.1086/307852.

D. J. Fixsen. *The Astrophysical Journal*, 707(2):916–920, 2009. ISSN 1538-4357. doi: 10.1088/0004-637x/707/2/916.

S. Flender, L. Bleem, H. Finkel, S. Habib, K. Heitmann, and G. Holder. *The Astrophysical Journal*, 823 (2):98, 2016. doi: 10.3847/0004-637x/823/2/98.

A. Friedmann. 10:377–386, 1922. doi: 10.1007/BF01332580.

G. Gamow. *nature*, 162:680–682, 1948. doi: 10.1038/162680a0.

S. Giodini, L. Lovisari, E. Pointecouteau, S. Ettori, T. H. Reiprich, and H. Hoekstra. *Space Science Reviews*, 177(1-4):247–282, 2013. doi: 10.1007/s11214-013-9994-5.

J. Glenn et al. In *Proceedings of SPIE - The International Society for Optical Engineering*, volume 3357, pages 326–334, 1998. doi: 10.1117/12.317418. Advanced Technology MMW, Radio, and Terahertz Telescopes ; Conference date: 26-03-1998 Through 26-03-1998.

O. Y. Gnedin, A. V. Kravtsov, A. A. Klypin, and D. Nagai. *The Astrophysical Journal*, 616(1):16–26, 2004. doi: 10.1086/424914.

J. González-Nuevo, L. Toffolatti, and F. Argüeso. *ApJ*, 621(1):1–14, 2005. doi: 10.1086/427425.

K. M. Gorski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke, and M. Bartelmann. *The Astrophysical Journal*, 622(2):759–771, 2005. doi: 10.1086/427976.

K. Grainge, M. Jones, G. Pooley, R. Saunders, and A. Edge. *MNRAS*, 265:L57–L58, 1993. doi: 10.1093/mnras/265.1.L57.

L. Grego, J. Carlstrom, and M. Joy. In *American Astronomical Society Meeting Abstracts #188*, volume 188 of *American Astronomical Society Meeting Abstracts*, page 21.03, 1996.

S. F. Gull and K. J. E. Northover. *Nature*, 263(5578):572–573, 1976. doi: 10.1038/263572a0.

H. P. Gush, M. Halpern, and E. H. Wishnow. *Phys. Rev. Lett.*, 65:537–540, 1990. doi: 10.1103/PhysRevLett.65.537.

A. H. Guth. *Phys. Rev. D*, 23:347–356, 1981. doi: 10.1103/PhysRevD.23.347.

N. Hand and et al. , 109(4):041101, 2012. doi: 10.1103/PhysRevLett.109.041101.

C. G. T. Haslam, U. Klein, C. J. Salter, H. Stoffel, W. E. Wilson, M. N. Cleary, D. J. Cooke, and P. Thomasson. *A & A*, 100:209–219, 1981.

C. G. T. Haslam, C. J. Salter, H. Stoffel, and W. E. Wilson. , 47:1, 1982.

C. Hernández-Monteagudo and et al. *Phys. Rev. Lett.*, 115:191301, 2015. doi: 10.1103/PhysRevLett.115.191301.

J. C. Hill, S. Ferraro, N. Battaglia, J. Liu, and D. N. Spergel. *Phys. Rev. Lett.*, 117:051301, 2016. doi: 10.1103/PhysRevLett.117.051301.

G. Hinshaw and et al. *The Astrophysical Journal Supplement Series*, 208(2):19, 2013. doi: 10.1088/0067-0049/208/2/19.

M. Hirschmann, K. Dolag, A. Saro, L. Bachmann, S. Borgani, and A. Burkert. *MNRAS*, 442(3):2304–2324, 2014. doi: 10.1093/mnras/stu1023.

E. Hivon, K. M. Gorski, C. B. Netterfield, B. P. Crill, S. Prunet, and F. Hansen. *The Astrophysical Journal*, 567(1):2–17, 2002. doi: 10.1086/338126.

S. Q. Hou, J. J. He, and et al. *The Astrophysical Journal*, 834(2):165, 2017. doi: 10.3847/1538-4357/834/2/165.

W. Hu and S. Dodelson. *Annual Review of Astronomy and Astrophysics*, 40(1):171–216, 2002. doi: 10.1146/annurev.astro.40.060401.093926.

W. Huggins. *Philosophical Transactions of the Royal Society of London*, 158:529–564, 1868. ISSN 02610523.

G. Hurier and C. Tchernin. *A & A*, 604:A94, 2017. doi: 10.1051/0004-6361/201629993.

G. Hurier, J. F. Macías-Pérez, and S. Hildebrandt. *A & A*, 558:A118, 2013. doi: 10.1051/0004-6361/201321891.

K. Ichiki. *Progress of Theoretical and Experimental Physics*, 2014(6), 2014. ISSN 2050-3911. doi: 10.1093/ptep/ptu065. 06B109.

J Dunkley and et al. *Journal of Cosmology and Astroparticle Physics*, 2013(07):025–025, 2013. doi: 10.1088/1475-7516/2013/07/025.

K. G. Jansky. , 132(3323):66, 1933. doi: 10.1038/132066a0.

M. Jones and et al. *Nature*, 365(6444):320–323, 1993. doi: 10.1038/365320a0.

N. Kaiser. 222:323–345, 1986. doi: 10.1093/mnras/222.2.323.

N. Kaiser. *MNRAS*, 227(1):1–21, 1987. ISSN 0035-8711. doi: 10.1093/mnras/227.1.1.

K. Kiuchi and et al. In H. K. Marshall, J. Spyromilio, and T. Usuda, editors, *Ground-based and Airborne Telescopes VIII*. SPIE, 2020. doi: 10.1117/12.2562016.

A. Kogut. *The Astrophysical Journal*, 753(2):110, 2012. doi: 10.1088/0004-637x/753/2/110.

A. Kogut, D. J. Fixsen, S. M. Levin, M. Limon, P. M. Lubin, P. Mirel, M. Seiffert, J. Singal, T. Villela, E. Wollack, and C. A. Wuensche. *ApJ*, 734(1):4, 2011. doi: 10.1088/0004-637X/734/1/4.

A. Kogut et al. *The Astrophysical Journal*, 665(1):355–362, 2007. doi: 10.1086/519754.

E. Komatsu and U. Seljak. *MNRAS*, 336(4):1256–1270, 2002. doi: 10.1046/j.1365-8711.2002.05889.x.

E. Komatsu et al. *The Astrophysical Journal*, 516(1):L1–L4, 1999. doi: 10.1086/311983.

E. Komatsu et al. *The Astrophysical Journal Supplement Series*, 180(2):330–376, 2009. doi: 10.1088/0067-0049/180/2/330.

H. Kragh. Princeton University Press, 1999.

A. Kusiak, B. Bolliet, S. Ferraro, J. C. Hill, and A. Krolewski. *Physical Review D*, 104(4), 2021. doi: 10.1103/physrevd.104.043518.

P. La Plante, J. Sipple, and A. Lidz. *ApJ*, 928(2):162, 2022. doi: 10.3847/1538-4357/ac5752.

G. Lake and R. B. Partridge. *Nature*, 270:502, 1977. doi: 10.1038/270502a0.

J. M. Lamarre and et al. , 507(1):L5–L8, 1998. doi: 10.1086/311678.

J. M. Lamarre and et al. *A & A*, 520:A9, 2010. doi: 10.1051/0004-6361/200912975.

D. Larson and et al. *The Astrophysical Journal Supplement Series*, 192(2):16, 2011. doi: 10.1088/0067-0049/192/2/16.

A. G. Lemaître. , 128(3234):704–706, 1931. doi: 10.1038/128704a0.

A. Lewis. , 71(8):083008, 2005. doi: 10.1103/PhysRevD.71.083008.

A. Lewis and A. Challinor, 2011.

B. Lundgren and et al. 127(954):776, 2015. doi: 10.1086/682387.

A. Maniyar, M. Béthermin, and G. Lagache. *A & A*, 645:A40, 2021. doi: 10.1051/0004-6361/202038790.

M. Markevitch, G. R. Blumenthal, W. Forman, C. Jones, and R. A. Sunyaev. *ApJ*, 395:326, 1992. doi: 10.1086/171655.

B. S. Mason and et al. *ApJ*, 716(1):739–745, 2010. doi: 10.1088/0004-637X/716/1/739.

J. C. Mather, D. J. Fixsen, R. A. Shafer, C. Mosier, and D. T. Wilkinson. *ApJ*, 512(2):511–520, 1999. doi: 10.1086/306805.

T. Matsubara. *Journal of Physics: Conference Series*, 31:27, 2006. doi: 10.1088/1742-6596/31/1/004.

F. Menanteau and et al. *The Astrophysical Journal*, 723(2):1523–1541, 2010. doi: 10.1088/0004-637x/723/2/1523.

M.-A. Miville-Deschênes et al. *A & A*, 490(3):1093–1102, 2008. doi: 10.1051/0004-6361:200809484.

A. T. Moffet and M. Birkinshaw. , 98:1148, 1989. doi: 10.1086/115204.

T. Mroczkowski and et al. *ApJ*, 761(1):47, 2012. doi: 10.1088/0004-637X/761/1/47.

T. Mroczkowski et al. *Space Science Reviews*, 215(1), 2019. doi: 10.1007/s11214-019-0581-2.

S. Muchovej et al. *The Astrophysical Journal*, 663(2):708–716, 2007. doi: 10.1086/511971.

D. Nagai, A. V. Kravtsov, and A. Vikhlinin. *ApJ*, 668(1):1–14, 2007. doi: 10.1086/521328.

J. F. Navarro, C. S. Frenk, and S. D. M. White. *ApJ*, 490(2):493–508, 1997. doi: 10.1086/304888.

I. Newton. 1687. doi: 10.3931/e-rara-440.

M. D. Niemack. *Applied Optics*, 55(7):1688, 2016. doi: 10.1364/ao.55.001688.

T. Nikola et al. In *Physics and Chemistry of Star Formation: The Dynamical ISM Across Time and Spatial Scales*, page 352, 2023.

M. Nord and et al. *A & A*, 506(2):623–636, 2009. doi: 10.1051/0004-6361/200911746.

J. H. Oort. , 6:249, Aug. 1932.

R. B. Partridge and D. T. Wilkinson. *Phys. Rev. Lett.*, 18:557–559, 1967. doi: 10.1103/PhysRevLett.18.557.

P. J. E. Peebles and J. T. Yu. *ApJ*, 162:815, 1970. doi: 10.1086/150713.

A. A. Penzias and R. W. Wilson. *ApJ*, 142:419–421, 1965. doi: 10.1086/148307.

S. Perlmutter and et al. *Nature*, 391(6662):51–54, 1998. doi: 10.1038/34124.

Planck collaboration et al. *A & A*, 536:A8, 2011. doi: 10.1051/0004-6361/201116459.

Planck collaboration et al. *A & A*, 550:A128, 2013. doi: 10.1051/0004-6361/201219361.

Planck collaboration et al. *A & A*, 571:A1, 2014. doi: 10.1051/0004-6361/201321529.

Planck collaboration et al. *A & A*, 594:A27, 2016. doi: 10.1051/0004-6361/201525823.

Planck collaboration et al. *A & A*, 617:A48, 2018. doi: 10.1051/0004-6361/201731489.

Planck collaboration et al. *A & A*, 641:A1, 2020a. doi: 10.1051/0004-6361/201833880.

Planck collaboration et al. *A & A*, 641:A5, 2020b. doi: 10.1051/0004-6361/201936386.

Planck collaboration et al. *A & A*, 641:A10, 2020c. doi: 10.1051/0004-6361/201833887.

Planck Collaboration at al. *A & A*, 641:A4, 2020. ISSN 1432-0746. doi: 10.1051/0004-6361/201833881.

Planck Collaboration et al. *A & A*, 536:A3, 2011. doi: 10.1051/0004-6361/201116480.

Planck Collaboration et al. *A & A*, 571:A21, 2014. doi: 10.1051/0004-6361/201321522.

Planck Collaboration et al. *A & A*, 576:A107, 2015. doi: 10.1051/0004-6361/201424088.

Planck Collaboration et al. *A & A*, 594:A10, 2016. doi: 10.1051/0004-6361/201525967.

Planck Collaboration et al. *A & A*, 641:A6, 2020. doi: 10.1051/0004-6361/201833910.

P. L. Plante, J. Sipple, and A. Lidz. *The Astrophysical Journal*, 928(2):162, 2022. doi: 10.3847/1538-4357/ac5752.

P. Platania, M. Bensadoun, M. Bersanelli, G. D. Amici, A. Kogut, S. Levin, D. Maino, and G. F. Smoot. *The Astrophysical Journal*, 505(2):473–483, 1998. doi: 10.1086/306175.

E. Pointecouteau and et al. *ApJ*, 552(1):42–48, 2001. doi: 10.1086/320447.

M. J. Rees and D. W. Sciama. *Nature*, 217(5128):511–516, 1968. doi: 10.1038/217511a0.

M. Remazeilles, J. Delabrouille, and J.-F. Cardoso. , 410(4):2481–2487, 2011. doi: 10.1111/j.1365-2966.2010.17624.x.

M. Remazeilles, C. Dickinson, A. J. Banday, M. A. Bigot-Sazy, and T. Ghosh, 2015.

M. Remazeilles, B. Bolliet, A. Rotti, and J. Chluba. *MNRAS*, 483(3):3459–3464, 2018. ISSN 0035-8711. doi: 10.1093/mnras/sty3352.

A. G. Riess et al. *The Astronomical Journal*, 116(3):1009–1038, 1998. doi: 10.1086/300499.

K. K. Rogers, H. V. Peiris, B. Leistedt, J. D. McEwen, and A. Pontzen. *MNRAS*, 460(3):3014–3028, 2016. doi: 10.1093/mnras/stw1121.

R. K. Sachs and A. M. Wolfe. *ApJ*, 147:73, 1967. doi: 10.1086/148982.

J. Sayers and et al. *ApJ*, 778(1):52, 2013. doi: 10.1088/0004-637X/778/1/52.

R. Scaramella and N. Vittorio. *ApJ*, 353:372, 1990. doi: 10.1086/168623.

M. Schmidt. , 14:17, 1957.

D. Schwan and et al. *New Astronomy Reviews*, 47(11-12):933–937, 2003.

T. Sebring, R. Giovanelli, S. Radford, and J. Zmuidzinas. *Proceedings of SPIE - The International Society for Optical Engineering*, 6267, 2006. doi: 10.1117/12.668735.

N. Sehgal, P. Bode, S. Das, C. Hernandez-Monteagudo, K. Huffenberger, Y.-T. Lin, J. P. Ostriker, and H. Trac. *ApJ*, 709(2):920–936, 2010a. doi: 10.1088/0004-637X/709/2/920.

N. Sehgal, P. Bode, S. Das, C. Hernandez-Monteagudo, K. Huffenberger, Y.-T. Lin, J. P. Ostriker, and H. Trac. *ApJ*, 709(2):920–936, 2010b. doi: 10.1088/0004-637X/709/2/920.

U. Seljak and M. Zaldarriaga. *ApJ*, 469:437, 1996. doi: 10.1086/177793.

L. D. Shaw, D. H. Rudd, and D. Nagai. *ApJ*, 756(1):15, 2012. doi: 10.1088/0004-637X/756/1/15.

M. Signore and D. Puy. *European Physical Journal C*, 59:117–172, 2009. doi: 10.1140/epjc/s10052-008-0807-z.

J. Silk. *ApJ*, 151:459, 1968. doi: 10.1086/149449.

K. Silsbee, Y. Ali-Haïmoud, and C. M. Hirata. *MNRAS*, 411(4):2750–2769, 2010. doi: 10.1111/j.1365-2966.2010.17882.x.

Simons Observatory Collaboration. *jcap*, 2019(2):056, 2019. doi: 10.1088/1475-7516/2019/02/056.

G. F. Smoot. URL https://api.semanticscholar.org/CorpusID:18444018.

G. F. Smoot et al. , 396:L1, 1992. doi: 10.1086/186504.

V. Springel, N. Yoshida, and S. D. M. White. *Nature*, 6(2):79–117, 2001. doi: 10.1016/S1384-1076(01)00042-2.

V. Springel et al. *Nature*, 435(7042):629–636, 2005. doi: 10.1038/nature03597.

Z. Staniszewski and et al. *ApJ*, 701(1):32–41, 2009. doi: 10.1088/0004-637X/701/1/32.

G. Stein, M. A. Alvarez, and J. R. Bond. *MNRAS*, 483(2):2236–2250, 2018. doi: 10.1093/mnras/sty3226.

G. Stein, M. A. Alvarez, J. R. Bond, A. van Engelen, and N. Battaglia. *Journal of Cosmology and Astroparticle Physics*, 2020(10):012–012, 2020. doi: 10.1088/1475-7516/2020/10/012.

A. W. Strong, E. Orlando, and T. R. Jaffe. *A & A*, 534:A54, 2011. doi: 10.1051/0004-6361/201116828.

I. A. Strukov and D. P. Skulachev. *Soviet Astronomy Letters*, 10:1–4, 1984.

N. S. Sugiyama, T. Okumura, and D. N. Spergel. *jcap*, 2017(1):057, 2017. doi: 10.1088/1475-7516/2017/01/057.

R. A. Sunyaev and Y. B. Zeldovich. *Nature*, 223(5207):721–722, 1969. doi: 10.1038/223721a0.

R. A. Sunyaev and Y. B. Zeldovich. *Comments on Astrophysics and Space Physics*, 2:66, 1970.

J. A. Tauber and et al. *A & A*, 520:A1, 2010. doi: 10.1051/0004-6361/200912983.

S. Tavernier. 2010. ISBN 978-3-642-00828-3. doi: 10.1007/978-3-642-00829-0.

M. Tegmark and G. Efstathiou. *MNRAS*, 281(4):1297–1314, 1996. doi: 10.1093/mnras/281.4.1297.

M. Tegmark, A. de Oliveira-Costa, and A. J. Hamilton. , 68(12):123523, 2003. doi: 10.1103/PhysRevD.68.123523.

B. Thorne, J. Dunkley, D. Alonso, and S. Næss. *MNRAS*, 469(3):2821–2833, 2017. doi: 10.1093/mnras/stx949.

H. C. van de Hulst, E. Raimond, and H. van Woerden. , 14:1, 1957.

E. M. Vavagiakis and et al. , 104(4):043503, 2021. doi: 10.1103/PhysRevD.104.043503.

E. M. Vavagiakis and et al., 2022.

F. Villa, M. Sandri, M. Bersanelli, R. C. Butler, N. Mandolesi, A. Mennella, J. Marti-Canales, and J. Tauber, 2003.

M. Vogelsberger, S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G. Snyder, D. Nelson, and L. Hernquist. *MNRAS*, 444(2):1518–1547, 2014. doi: 10.1093/mnras/stu1536.

S. Weinberg. Basic Books, 1977. ISBN 9780465024353.

N. Wiener, 1949.

T. M. Wilbanks, P. A. R. Ade, M. L. Fischer, W. L. Holzapfel, and A. E. Lange. , 427:L75, 1994. doi: 10.1086/187368.

M. Wright. *Physica Medica*, 7, 2015. doi: 10.1016/j.ejmp.2015.04.003.

Z. Xu and et al. *Research Notes of the AAS*, 5(4):100, 2021. doi: 10.3847/2515-5172/abf9ab.

M. Zemcov and et al. *A & A*, 518:L16, 2010. doi: 10.1051/0004-6361/201014685.

M. Zemcov and et al. *ApJ*, 749(2):114, 2012. doi: 10.1088/0004-637X/749/2/114.

M. Zemcov, C. Borys, M. Halpern, P. Mauskopf, and D. Scott. *MNRAS*, 376(3):1073–1098, 2007. doi: 10.1111/j.1365-2966.2007.11443.x.

P. Zhang, H. A. Feldman, R. Juszkiewicz, and A. Stebbins. *MNRAS*, 388(2):884–888, 2008. doi: 10.1111/j.1365-2966.2008.13454.x.

A. Zonca, L. Singer, D. Lenz, M. Reinecke, C. Rosset, E. Hivon, and K. Gorski. *Journal of Open Source Software*, 4(35):1298, 2019. doi: 10.21105/joss.01298.

A. Zonca, B. Thorne, N. Krachmalnicoff, and J. Borrill. *Journal of Open Source Software*, 6(67):3783, 2021. doi: 10.21105/joss.03783.

J. T. L. Zwart and et al. The Arcminute Microkelvin Imager. *MNRAS*, 391(4):1545–1558, 2008. doi: 10.1111/j.1365-2966.2008.13953.x.

F. Zwicky. *Helvetica Physica Acta*, 6:110–127, 1933.