# Towards cosmological parameters with the SPT-3G 19/20 cosmic microwave background data

Etienne Camphuis

## To cite this version:

Etienne Camphuis. Towards cosmological parameters with the SPT-3G 19/20 cosmic microwave background data. Astrophysics [astro-ph]. Sorbonne Université, 2023. English. NNT : 2023SORUS365 . tel-04347707

HAL Id: tel-04347707
https://theses.hal.science/tel-04347707

Submitted on 15 Dec 2023

SORBONNE UNIVERSITÉ

ÉCOLE DOCTORALE 127: ASTRONOMIE ET ASTROPHYSIQUE D'ÎLE-DE-FRANCE

INSTITUT D'ASTROPHYSIQUE DE PARIS

---

# Towards cosmological parameters with the SPT-3G 19/20 cosmic microwave background data

---

## DOCTORAL THESIS IN ASTROPHYSICS

*Author:*
Etienne Camphuis

*Supervisors:*
Karim Benabed
Silvia Galli
Eric Hivon

PRESENTED AND PUBLICLY DEFENDED ON THE 13 SEPTEMBER 2023
IN FRONT OF THE FOLLOWING JURY

| | | |
|---|---|---|
| **François Bouchet** | Research Director, IAP, CNRS | President |
| **Erminia Calabrese** | Deputy Director of Research, Cardiff University | Referee |
| **Olivier Doré** | Principal Scientist, Caltech | Referee |
| **Karim Benabed** | Research Director, IAP, CNAP | Examiner |
| **Thomas Crawford** | Research Professor, University of Chicago | Examiner |
| **Sophie Henrot-Versillé** | Research Director, IJCLab, CNRS | Examiner |
| **Valeria Pettorino** | Research Director, CEA | Examiner |
| **Silvia Galli** | Researcher, IAP, CNRS | Invited member |

INSTITUT D'ASTROPHYSIQUE DE PARIS
Unité mixte de recherche 7095 CNRS - Sorbonne Université

# Abstract

Observations of primary anisotropies in the cosmic microwave background (CMB) represent a crucial milestone in modern cosmology, offering a unique window into the physics of the early Universe. However, tensions in measurements compared to other probes, such as the Hubble tension, have emerged suggesting potential discrepancies within the cosmological concordance. Although these discrepancies may be attributed to systematic effects or statistical fluctuations, the consistent disparity between early and late Universe measurements might hint at new physics. Additionally, the temperature anisotropy measurements have not reached yet the cosmic variance limit at high multipoles, while the polarized anisotropies hold substantial untapped cosmological information. In this context, investigating a complementary range of multipoles and exploring polarization become vital avenues for potentially uncovering fundamental physics.

The South Pole Telescope third-generation (SPT-3G) experiment is undertaking precisely this endeavor, aiming to provide state-of-the-art observations that enhance our understanding of the Universe. The objective of this thesis is to present a detailed analysis of the SPT-3G 2019-2020 primary anisotropies data. This dataset achieves a low noise level of 5$\mu$K-arcmin in temperature at 150GHz, making it cosmic variance limited for almost all the power spectrum multipole range. Forecasted constraints on cosmological parameters are comparable to the reference *Planck* experiment but from an independent and complementary dataset. Combining these constraints is expected to yield at least twice the statistical power, and further improvements can be achieved by incorporating information from extended fields, lensing measurements, and other CMB datasets such as those from the Atacama Cosmology Telescope.

To ensure the accuracy and precision of the analysis, the pipeline requires additional development. This thesis includes an in-depth study of the map-making pipeline and power spectrum estimation, along with the development of the necessary tools for a curved sky analysis. I introduce and develop a novel solution to the source masking problem, reducing the impact of the mask on power spectrum variance through the use of high precision Gaussian constrained realizations of CMB anisotropies in temperature and polarization called inpainting. Extensive testing is conducted to demonstrate the robustness of this method. Additionally, reliable estimation of parameters necessitates accurate covariance matrices. For this analysis, I propose a new semi-analytical framework, adapted from an analytical approximation developed for small footprint CMB surveys, tailored to the specificities of the SPT-3G experiment. While this method is computationally efficient, it is also flexible and can be adapted to other experiments and cosmological probes. Such improvements are essential for the analysis of the SPT-3G 2019-2020 data, but also pave the way for future CMB experiments, such as the Simons Observatory and CMB-S4.

The scientific results I present in this thesis are done in preparation for the analysis of the SPT-3G 2019-2020 data, which is currently being finalized and of which I am one of the leading authors. The analysis pipeline is being developed in parallel, and the results presented here are preliminary. The final results are expected to be published in the coming months.

## Résumé

Les observations des anisotropies primaires du fond diffus cosmologique (CMB en anglais) sont primordiales dans l'établissement de la cosmologie moderne et offrent une fenêtre unique sur la physique de l'Univers primitif. Cependant, des divergences au sein de la concordance cosmologique, notamment la tension de Hubble, sont mises en avant en comparant avec d'autres sondes cosmologiques. Bien que ces divergences puissent être attribuées à des effets systématiques ou à des fluctuations statistiques, la différence entre les mesures de l'Univers primitif et récent pourrait indiquer de nouveaux principes physiques. De plus, les mesures d'anisotropie de température n'ont pas encore atteint la limite de la variance cosmique à hauts multipôles, tandis que la polarisation du signal contient des informations cosmologiques inexploitées. Ainsi, explorer une gamme complémentaire de multipôles et la polarisation peut mener à des découvertes majeures en physique fondamentale.

L'instrument de troisième génération du télescope du pôle sud (SPT-3G en anglais) entreprend précisément cette démarche, visant à fournir des observations de pointe. L'objectif de cette thèse est de présenter une analyse détaillée des données 2019-2020 d'anisotropies primaires de SPT-3G. Le niveau de bruit atteint est de $5\mu$K-arcmin en température à 150GHz. Ainsi, la variance du signal est limité par la variance cosmique sur presque toute la gamme de multipôles du spectre de puissance. Les contraintes prévues sur les paramètres cosmologiques sont comparables à celles de l'expérience Planck, mais proviennent d'un jeu de données indépendant et complémentaire. La combinaison de ces contraintes fournira au moins deux fois la puissance statistique, et des améliorations supplémentaires seront obtenues en intégrant des informations provenant de champs étendus, de mesures de l'effet de lentillage gravitationnelle et d'autres jeu de données CMB, tels que ceux provenant du télescope de cosmologie d'Atacama.

Pour garantir l'exactitude et la précision de l'analyse, les méthodes d'analyse doivent être améliorées. Cette thèse comprend une étude approfondie de la création de cartes et de l'estimation du spectre de puissance, ainsi que le développement des outils nécessaires pour une analyse sphérique du ciel. J'introduis et développe une nouvelle solution au problème de masquage des sources, réduisant l'impact du masque sur la variance du spectre de puissance grâce à l'utilisation de réalisations gaussiennes contraintes de haute précision. Cette méthode, appelée « inpainting», est parfaitement adaptées aux anisotropies du CMB en température et en polarisation. La fiabilité de cette méthode est démontrée par une série des tests approfondis. De plus, une estimation fiable des paramètres nécessite un matrice de covariance précise. Pour cette analyse, je propose une nouvelle méthode d'estimation semi-analytique. Cette méthode repose sur une approximation analytique développée pour les relevés CMB sur petite zone de ciel et est adaptée aux spécificités de l'expérience SPT-3G. Bien que cette méthode soit efficace sur le plan computationnel, elle est également flexible et peut être adaptée à d'autres expériences et sondes cosmologiques. Ces améliorations sont essentielles pour l'analyse des données SPT-3G 2019-2020, mais ouvrent également la voie aux futures expériences CMB, telles que le Simons Observatory et CMB-S4.

Les résultats scientifiques que je présente dans cette thèse sont réalisés en préparation de l'analyse des données SPT-3G 2019-2020, qui est actuellement en cours de finalisation et dont je suis l'un des auteurs principaux. L'analyse est en développement parallèle, et les résultats présentés ici sont préliminaires. Les résultats finaux devraient être publiés dans les prochains mois.

## Acknowledgements

My first thanks go to Silvia, Karim, and Eric, my supervisors, for their kind guidance throughout the thesis. I discovered and explored cosmology thanks to them, and I am very grateful for their help and their commitment to the guidance of the Ph.D. Not only they are excellent scientists from whom I learned a lot, but they are also very kind and caring people who supported me and gave me confidence in my work. They challenged me to produce the best work I could, and I wouldn't have achieved a tenth of this thesis without them.

I would like to thank a lot Wei, with whom we are leading the cosmological constraints. Working with him was an amazing experience, as I greatly appreciated his dedication to the project and his excellent scientific skills. Our weekly meetings were fun and interesting moments, I am inspired by his organization and presentation skills, as well as his open-mindedness and his commitment to making research an inclusive and diverse environment. Good luck in the future, I am sure you will achieve great science!

I am grateful to the NeuCosmos junior team, Ali, Aristide, Federica, and Lennart for the great times shared together. I really enjoyed working with them, and my Ph.D. would have been much less fun and less productive without them. Group projects were a great way to learn and share knowledge, and I am excited to continue collaborating with them in the future.

Many thanks to the South Pole Collaboration, and especially the SPT-3G team. I am grateful for the opportunity to work with a great team of scientists. I really appreciated the effort of the board to ensure inclusive and safe exchanges during meetings and teleconferences. Special thanks to Thomas Crawford for his dedication to the TT TE EE 19-20 project and the great discussions we had together.

My sincere thanks go to the jury members, François Bouchet, Erminia Calabrese, Thomas Crawford, Olivier Doré, and Valeria Pettorino, for accepting to evaluate this work. I am honored to present my work to great scientists. Special thanks to Erminia Calabrese and Olivier Doré for the reading and evaluation of this manuscript and the feedback.

Special thanks to the *Institut d'Astrophysique de Paris* for hosting me during my Ph.D. I am grateful to *Ecole Polytechnique* for funding my Ph.D. through the *bourse AMX*.

I am very grateful to Benjamin Wandelt for his support during the Ph.D. and the excellent lectures on Bayesian statistics.

Sur une touche plus personnelle, la thèse est une aventure passionnante que j'ai eu la chance de vivre, grâce aux soutiens de ma famille et de mes proches. Je voudrais particulièrement remercier Aurélien, Benoît, Célia, Clément, Corentin, Gaspard, Pierre, Pierre et Valentin qui ont partagé avec moi l'expérience de la thèse, tout autant pour l'amour de la science que pour le plaisir de rester étudiants ensemble. Je nous reconnais dans cette citation: « Trop évolués pour se ressembler parfaitement, mais, sans doute, pas encore assez pour ne pas s'imiter plus ou moins consciemment, ils passaient une grande partie de leur vie en échanges. Ils s'en irritaient souvent ; ils s'en amusaient plus souvent encore. » (George Perec, *Les choses*).

# Contents

# Chapter 1

# Introduction

---

Observational cosmology is a field of study that examines the large-scale structure of the Universe through observational data. It has played a crucial role in numerous scientific breakthroughs. The era of precision cosmology began with the successful launch of the *Cosmic Background Explorer* (COBE) satellite in 1992, which provided the first full-sky observations of the **Cosmic Microwave Background** (CMB) radiation. Advancements in theoretical modeling and observational techniques have significantly improved our understanding of the Universe using CMB observations. With the *Wilkinson Microwave Anisotropy Map* (WMAP), the CMB has emerged as a powerful tool for studying the Universe, culminating in the landmark results from the *Planck* satellite in 2018. These results provided the most precise independent measurements of cosmological parameters, ultimately solidifying the standard model of cosmology. However, the availability of high-precision data sets has also revealed tensions between different cosmological probes. One well-known example is the Hubble tension, which refers to the discrepancy in the measurement of the Hubble rate, denoted as $H_0$, between early Universe observations (such as the CMB) and late Universe observations (such as measurements from supernovae IA using the distance ladder method). In addition to addressing tensions, the CMB still holds a wealth of untapped knowledge, particularly in the small scales and the weakly polarized signal. It can be used to test the inflationary theory by studying large-scale B-modes, investigate the large-scale structure of the Universe through precise lensing and secondary anisotropies measurements, and explore spectral distortions, among other phenomena. To tackle these challenges and explore new observables, new high-precision data sets are required. The **South Pole Telescope third generation** (SPT-3G) experiment is one such endeavor, aiming to provide cutting-edge observations to improve our understanding of the Universe.

## A brief introduction to cosmology

The standard cosmology is built upon two fundamental assumptions (Narlikar and Padmanabhan, 2001; Turner, 2022): first, the large-scale structure of the Universe is governed by the gravitational interaction as described by Einstein's theory of general relativity, and second, the distribution of components in the Universe is homogeneous and isotropic on large scales, known as the cosmological principle. These assumptions provide a powerful framework for describing the Universe using a simple model known as the Friedmann-Lemaître-Robertson-Walker metric, which is a solution to Einstein's field equations. The FLRW metric is given by the equation (Dodelson and Schmidt, 2020):

$$ds^2 = c^2 dt^2 - a^2(t)\left[\frac{dr^2}{1-kr^2} + r^2 d\Omega^2\right], \tag{1.1}$$

where $a(t)$ represents the scale factor, $k$ denotes the curvature of the Universe, $c$ represents the speed of light, $t$ is the cosmic time, $r$ is the comoving radial distance, and $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$ represents the solid angle. The scale factor is a function of time and describes the expansion of the Universe, with its present value set to unity. It is also related to the redshift of radiation traveling in the expanding Universe, as

$$1 + z = \frac{a(t_0)}{a(t)}, \tag{1.2}$$

where $t_0$ denotes the present time. The evolution of the scale factor is governed by the Hubble rate, denoted $H(t) \equiv \dot{a}/a$. It is connected to the energy densities in the Universe through the Friedmann equations derived from Einstein's field equations. The Friedmann equations are given by:

$$H^2(t) = \frac{8\pi G}{3}\rho(t) - \frac{kc^2}{a^2(t)}, \quad \dot{H}(t) + H^2(t) = \frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho(t) + 3\frac{p(t)}{c^2}\right), \tag{1.3}$$

where $\rho$ represents the total energy density of the Universe, $G$ is the gravitational constant, and $p$ denotes the total pressure. Given those, it is customary to define the critical density, which is the total energy density in a flat Universe today, $\rho_{\rm cr} = 3H(t_0)^2/8\pi G$. Particle physics plays a crucial role in cosmology by providing details of the energy content and interactions between particles, as well as their equation of state $w_s$ which depicts the relation between the pressure and density of that constituent. One can thus relate the energy density of each component to the scale factor as

$$\rho_s = \Omega_s \rho_{\rm cr} a^{-3(1+w_s)}, \quad \text{where} \quad w_s = \frac{p_s}{\rho c^2} \quad \text{and} \quad \Omega_s = \rho_s(t_0)/\rho_{\rm cr}. \tag{1.4}$$

The Universe is composed of various components, including photons, neutrinos, baryonic matter, cold dark matter, and dark energy represented by a cosmological constant, $\Lambda$. By specifying the densities parameters $\Omega_s$ and equations of state for each component $w_s$, along with the present value of the Hubble rate $H_0 \equiv H(t_0)$ and the total energy density, the evolution of the Universe can be determined, as depicted in Fig. 1.1. This model explains disparate observations, such as the large-scale structure in the distribution of galaxies, the observed abundances of light elements, the accelerated expansion of the Universe, and, of particular interest in this thesis, the existence of the CMB, but it is not without its problems (Baumann et al., 2009):

- The horizon problem: The Universe appears to be isotropic and homogeneous on large scales, but regions that are far apart from each other were not in causal contact since the Big Bang. This means that light or information has not had enough time to travel between these distant regions to establish thermal equilibrium. Yet, the observed isotropy suggests a common origin.
- The flatness problem: The Universe is observed to be extremely close to spatially flat, which is unexpected because a flat Universe is an unstable solution in the standard model of cosmology. In other words, the initial conditions required for the Universe to be so flat would require an extreme level of fine-tuning.
- The relic problem: According to high-energy physics theories, the early Universe should have produced topological defects and magnetic monopoles as relics. However, these relics have not been observed in the present Universe.

Inflationary cosmology offers a solution to these problems by proposing a period of exponential expansion in the early Universe, which also accounts for the observed homogeneity and isotropy of the Universe. According to this model, the observable Universe originated from an incredibly tiny patch of space, much smaller than $10^{-26}$ meters. During inflation, the Universe underwent rapid exponential expansion, causing modes to grow at a faster rate than the horizon size. Consequently, the initially small-scale quantum fluctuations were stretched to cosmological scales. These fluctuations gave rise to adiabatic perturbations, resulting in nearly identical fractional density perturbations across all species. Later, these perturbations served as the source of the temperature variations observed in the CMB and the large-scale structure of the Universe.

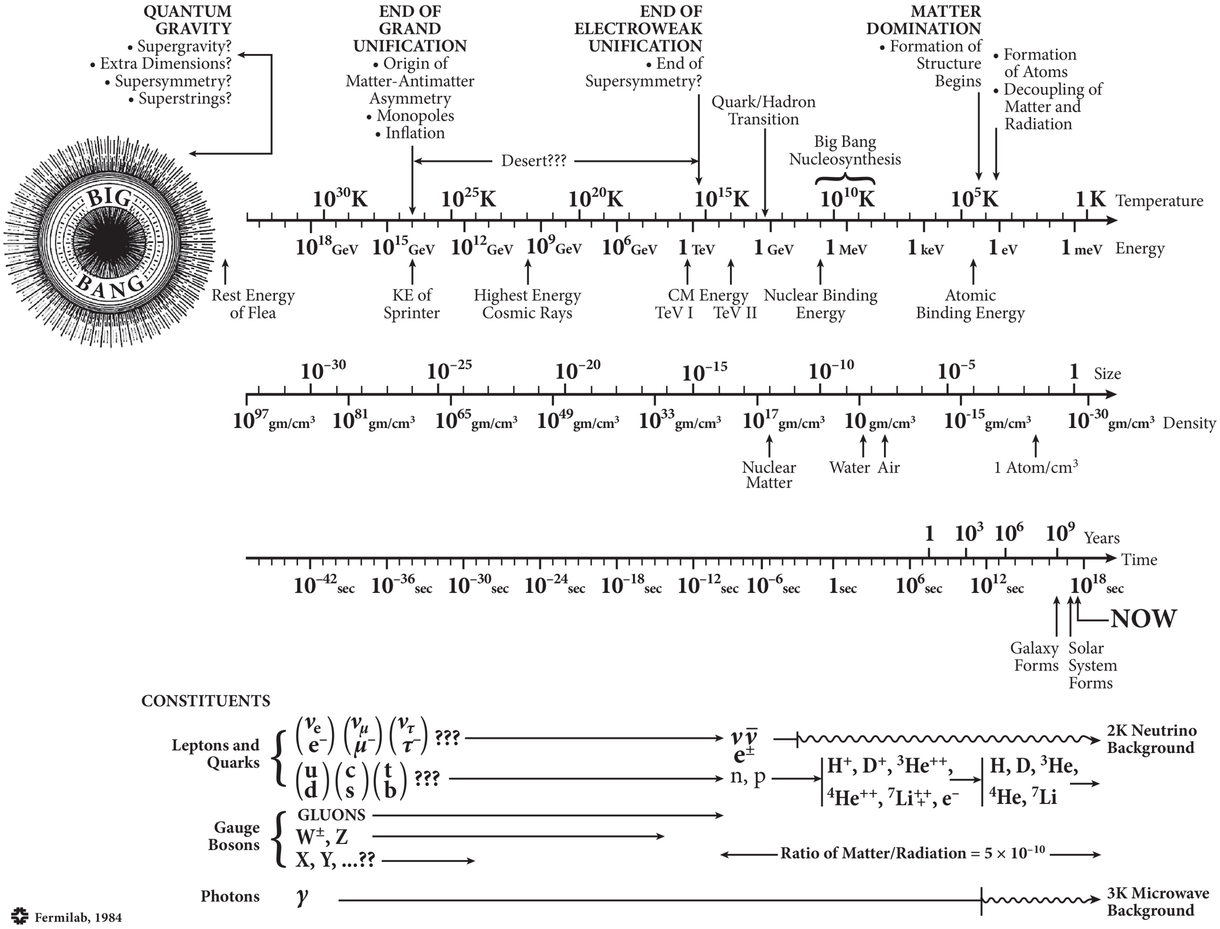
In the literature, this model is commonly referred to as the $\Lambda$CDM model, where $\Lambda$ denotes the cosmological constant and CDM stands for cold dark matter. $\Lambda$CDM is a six-parameter model, with the parameters representing the baryon density ($\omega_{\rm b}$), the cold dark matter density ($\omega_{\rm c}$), the angular size of the sound horizon at recombination ($\theta_{\rm MC}$), the optical depth to reionization ($\tau$), the amplitude of the primordial scalar power spectrum ($A_{\rm s}$), and the scalar spectral index ($n_{\rm s}$). The curvature parameter, $\Omega_k$ is sometimes considered as a free parameter, with the constraint $\Omega_k = 0$ corresponding to a spatially flat Universe. The radiation density is very precisely measured by the COBE-FIRAS measurements of the CMB energy spectrum and thus considered as a fixed parameter. One writes

$$\Omega_k = 1 - \Omega_{\rm m} - \Omega_\Lambda, \quad \text{with} \quad \Omega_{\rm m} = \Omega_{\rm b} + \Omega_{\rm c}. \tag{1.5}$$

Extensive observations have shown that the $\Lambda$CDM model is in excellent agreement with the data, and the values of these parameters have been tightly constrained by the *Planck* satellite (Planck Collaboration et al., 2020c), as presented in Table 1.1. While the free parameters completely specify the $\Lambda$CDM model, numerous physical quantities can be derived from them, such as the Hubble constant or the age of the Universe. Additional parameters are also considered in the literature, such as the sum of the neutrino masses, the effective number of relativistic species, or the running of the scalar spectral index, and can be set free in the analysis for probing extensions to the $\Lambda$CDM model.

## The cosmic microwave background

In this thesis, we will focus on the study of cosmic microwave background (CMB) anisotropies, which are of paramount importance in modern cosmology, and the measurements of cosmological parameters from this observation. A review of CMB analysis can be found in Hu and Dodelson (2002). The CMB is the relic



Turner MS. 2022
*Annu. Rev. Nucl. Part. Sci.* 72:1–35

Figure 1.1: History of the Universe in logarithmic comoving time. While the very early Universe is still poorly understood, the evolution of the Universe is well known from the electroweak era ($10^{-12}s$) to the present day. The CMB was emitted at the decoupling of matter and radiation. Image credit: Turner (2022).

| **Group** | **Parameter** | **Description** | **Value** |
|---|---|---|---|
| Free | $\omega_{\rm B} \equiv \Omega_{\rm c} h^2$ | Baryon density | $0.02237 \pm 0.00015$ |
| | $\omega_{\rm b} \equiv \Omega_{\rm b} h^2$ | Cold dark matter density | $0.1200 \pm 0.0012$ |
| | $\theta_{\rm MC}$ | Angular size of the sound horizon at recombination | $1.04092 \pm 0.00031$ |
| | $\tau$ | Reionization optical depth | $0.0544 \pm 0.0073$ |
| | $\log(10^{10} A_{\rm s})$ | Amplitude of the primordial power spectrum | $3.033 \pm 0.014$ |
| | $n_{\rm s}$ | Primordial spectral index | $0.9649 \pm 0.0042$ |
| Derived | $\omega_{\rm m} = \Omega_{\rm m} h^2$ | Matter density | $0.1430 \pm 0.0011$ |
| | $H_0$ [kms$^{-1}$Mpc] | Hubble constant | $67.36 \pm 0.54$ |
| | $t$ [Gyr] | Age of the Universe | $13.797 \pm 0.023$ |
| | $\sigma_8$ | Amplitude of matter density fluctuations | $0.8111 \pm 0.0060$ |
| | $z_{\rm re}$ | Redshift of reionization | $7.67 \pm 0.73$ |
| | $r_{\rm drag}$ | Comoving sound horizon at recombination | $147.09 \pm 0.26$ |
| | $S_8 \equiv \sigma(\Omega_{\rm m}/0.3)^{0.5}$ | Renormalized amplitude of matter density fluctuations | $0.832 \pm 0.013$ |
| Fixed | $\Omega_{\rm k}$ | Curvature density | 0 |
| | $A_L$ | Lensing amplitude relative to physical value | 1 |
| | $\sum_\nu m_\nu$ [eV] | Sum of neutrino masses | 0.06 |
| | $w$ | Dark energy equation of state parameter | -1 |
| | $N_{\rm eff}$ | Effective number of relativistic degrees of freedom | 3.046 |
| | $Y_{\rm p}$ | Primordial helium fraction | 0.24 |
| | $\alpha_s$ | Running of the spectral index | 0 |
| | $A_{\rm t}$ | Tensor amplitude | 0 |
| | $n_{\rm t}$ | Tensor spectral index | 0 |
| | $r$ | Tensor-to-scalar ratio | 0 |

Table 1.1: Baseline ΛCDM cosmological parameters from *Planck* 2018 TT,TE,EE +lowE +lensing (Planck Collaboration et al., 2020c). The free parameters can be combined into additional derived parameters, while the fixed parameters are frozen during the standard ΛCDM analysis. They can be set free when probing extensions to the ΛCDM model.

radiation of the Big Bang, a uniform thermal sea of photons, with small anisotropies which give an insight into the primordial conditions of the Universe. The CMB was predicted in pioneering work by Gamow (1948); Alpher and Herman (1948) and later serendipitously discovered by Penzias and Wilson (1965) and Dicke et al. (1965). This discovery led to great advances in cosmology, consolidating the Big Band model since the origin of CMB is very well understood within this framework. The characterization of the CMB anisotropies temperature and polarization spectrum has been of primary importance since their discovery by Smoot et al. (1992) with COBE for temperature and Kovac et al. (2002) with DASI for polarization. Independent experiments, such as WMAP (Bennett et al., 2003, 2013) and *Planck* (Planck Collaboration et al., 2014, 2015, 2020a), have yielded exquisite measurements of cosmological parameters and probed the physics of the early Universe.

## Recombination and decoupling

The early history of the Universe can be described as a sequence of phase transitions driven by the cooling resulting from its expansion (Dodelson and Schmidt, 2020). At around $1s$ after the Big Bang, the Universe consisted of decoupled relativistic neutrinos and relativistic particles in equilibrium, as well as non-relativistic baryons that were tightly coupled to the relativistic particles through photons. The temperature gradually decreased until it reached a point where chemical reactions could no longer remain in kinetic equilibrium, leading to the breaking of the reaction symmetry. This decoupling process first occurred during Big Bang nucleosynthesis (BBN), when the thermal bath of relativistic particles in equilibrium and non-relativistic baryons decoupled. This resulted in the observed abundance of primordial hydrogen and helium. After BBN, the thermal bath consisted of a "baryon-photon" fluid composed of protons, electrons, photons, and helium nuclei, which remained coupled through Thomson scattering and electromagnetic interactions. In this fluid, the temperature was uniform, with slight deviations at the locations of dark matter overdensities arising from inflation, as density perturbations caused small perturbations in the gravitational field. As long as the temperature remained above $T_{\rm dec} \sim 3000$K, the photons had sufficient energy to ionize hydrogen atoms, maintaining the coupling of the fluid. The decoupling occurred when the temperature dropped below this threshold, which happened at a redshift $z_{\rm dec} \sim 1100$. Electrons and protons combined to form neutral hydrogen during this phase transition. As a result, the Universe became transparent as photons were no longer Thomson scattered. Since then, photons have been "free streaming" through the expanding Universe from the surface of the last scattering, and we observe them today as the CMB. As the CMB radiation originates from a thermal bath, it follows a black-body spectrum characterized by the temperature at the time of decoupling, denoted as $T_{\rm dec}$. The black-body spectrum is given by the Planck law

$$I_\nu = \frac{2h\nu^3}{c^2} \frac{1}{\exp\left(\frac{h\nu}{k_B T}\right) - 1} \tag{1.6}$$

Photons have traveled through the expanding Universe and have been redshifted and cooled down, while still

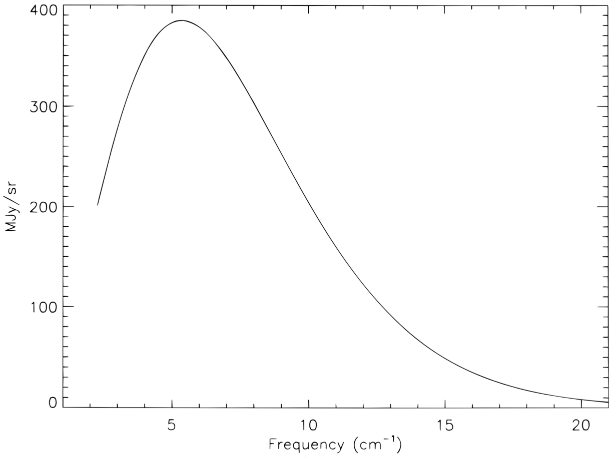


Figure 1.2: The CMB black-body spectrum measured by the FIRAS instrument of the COBE satellite. Uncertainties are a small fraction of the line thickness. Figure taken from (Fixsen et al., 1996).

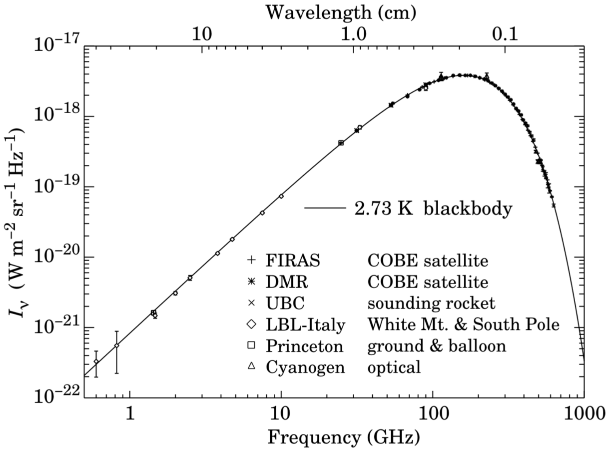


Figure 1.3: Black-body spectrum measured by different intruments (FIRAS (Fixsen et al., 1996), DMR (Kogut et al., 1993), UBC (Gush et al., 1990), LBL (Bersanelli et al., 1995), Princeton (Johnson et al.), Cyanogen (Roth and Meyer, 1995)) with associated uncertainties. Figure taken from Smoot et al. (1992).

retaining their thermal properties. The temperature of the cosmic radiation has been shifted to $T_{\rm CMB} = T_{\rm dec}/(1 + z_{\rm dec})$. The FIRAS instrument on the COBE satellite measured the CMB temperature and obtained the most precise observation of a black-body spectrum in nature, providing valuable constraints on energy injection in the early Universe (White, 1999). The measured temperature of the CMB is $T_{\mathrm{CMB}} = 2.72548 \pm 0.00057$ K. The measured black-body spectrum is shown in Fig. 1.2 and Fig. 1.3. Additionally, the CMB exhibits dipole anisotropy due to the Doppler effect related to the earth motion at the $10^{-3}$ level, but also small-scale anisotropies at the $10^{-5}$ level, which are the result of primordial fluctuations, that we describe in the next section.

## Primary anisotropies

The photon-baryon fluid before recombination exhibits a very uniform temperature, but small perturbations are present due to the gravitational potential wells. These perturbations are sourced by primordial curvature fluctuations, which are generated by inflation. The observed anisotropies on the sphere are well described using Gaussian statistics and further characterized by their angular power spectrum, $C_\ell^{\mathrm{XY}}\ \forall \mathrm{XY} \in [\mathrm{T}, \mathrm{E}, \mathrm{B}]$, where T refers to the temperature anisotropies while E and B refer respectively to the curl-free and gradient-free decomposition of the polarization modes. We introduce briefly the generation of anisotropies in the following paragraphs but refer to White et al. (1994); Hu and Sugiyama (1995); Hu et al. (1995); Hu and White (1996); Hu and Sugiyama (1996); Hu and White (1997); Hu et al. (1997); Sugiyama (2014); Dodelson and Schmidt (2020) for a detailed treatment.

**Temperature anisotropies** As an overview, we separate the generation of temperature fluctuations into main physical processes. Let us first introduce the gravitational shifting of photon frequencies, the so-called Sachs-Wolfe effect. At the time of recombination, photons are located inside gravitational over-densities or under-densities and are redshifted or blueshifted accordingly as they move out of those. The Sachs-Wolfe effect is only dominant on large scales. Similarly, the integrated Sachs-Wolfe (ISW) effect arises from the time variation of those gravitational potential wells along the line of sight. It can be separated into early ISW, which arises during the transition from the radiation-dominated era to the matter-dominated, and late ISW, happening after the transition to dark energy domination. Then, the physics of the photon-baryon fluid before recombination generates distinct features in the anisotropy statistics. The free electrons tightly couples the baryons and the photons together through Thomson and Coulomb scattering and density perturbations propagate in the form of acoustic oscillations. The evolution of oscillations is linear, since the primordial perturbations are small, and carries distinct signatures of the Universe components. An important effect is radiation driving, which arises from the domination of the radiation energy density over the matter-energy density, implying that the baryon-photon plasma will not be oscillating in a fixed potential well. The potential decays at scales that entered the horizon during the radiation era, driving up the amplitudes of the acoustic oscillations. Moreover, in the fluid, photons diffuse at scales smaller than their mean free path, where the coupling between photons and electrons is weak. This leads to the exponential damping of acoustic oscillations on small scales. This damping process is known as Silk damping (Silk, 1968). Finally, the traveling photons are scattered by the free electrons in the intergalactic medium ionized by the first stars, and subsequently, the temperature fluctuations are damped. This damping depends on the optical depth at the reionization of the intergalactic medium. All those processes carry cosmological information and can be parametrized by the model parameters.

The aforementioned physical processes have a distinct impact on the power spectrum, and their main contribution can be pinpointed on the renormalized power spectrum as indicated in Fig. 1.4. It is customary to plot $D_\ell \equiv \frac{\ell(\ell+1)}{2\pi} C_\ell^{\mathrm{TT}}$, as $D_\ell$ would be constant with pure Sachs-Wolfe effect and $n_s = 1$ (Sugiyama, 2014), and such a renormalization highlights relevant features in the power spectrum. The classical and integrated Sachs-Wolfe effects mainly impact the low multipoles, which are the equivalent of the large scales. The acoustic oscillations are visible in the intermediate and large multipoles through the peaks and troughs of the power spectrum, whose position and amplitude are sensitive to the characteristics of the acoustic oscillations. Lastly, the Silk damping affects the large multipoles. Cosmological parameters impact those processes and therefore determine the shape of the power spectrum. For example, let us take the case of the baryon density. As the pressure is primarily carried by photons, baryons provided mass inertia. The inertial mass of baryon drags photons towards overdense regions, and enhances the compression phase, and thus odd harmonic series of the acoustic oscillations. This enhancement of compression implies an additional rarefaction at the under-density locations, therefore the even peaks are decreased. This effect is depicted in panel (c) of Fig. 1.5, where $\Delta_T = D_\ell$ are the renormalized temperature power spectra. In this figure, the effects of varying the spatial curvature, the density of dark energy, and the amount of matter are also displayed for indication. As cosmological parameters have varying impacts on different scales, probing a large range of multipoles allows for the reduction of degeneracies when measuring parameters. Including polarization measurements further aids in reducing these degeneracies. As primary anisotropies are small, their computation can be handled by linear perturbation theory thus allowing a very accurate computation. The detailed computation of the power spectrum is done using linear line-of-sight integration of the Boltzmann equation. Codes such as `CAMB` Lewis et al. (2000) or `CLASS` (Lesgourgues, 2011) show exquisite accuracy and now integrate a wide range of models, allowing to study the impact of new physics

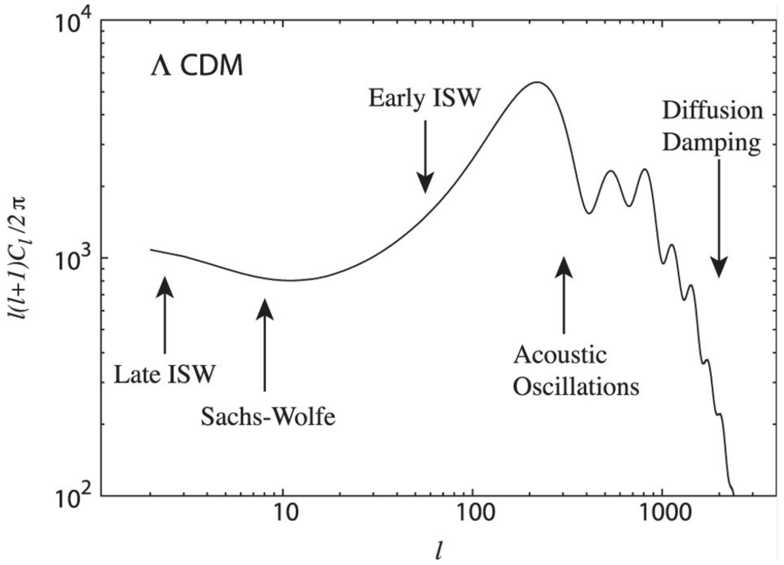


Figure 1.4: The CMB angular power spectrum of $\Lambda$CDM in $\mu$K$^2$. Physical processes which provide dominant contributions on different scales are quoted. Figure taken from Sugiyama (2014).

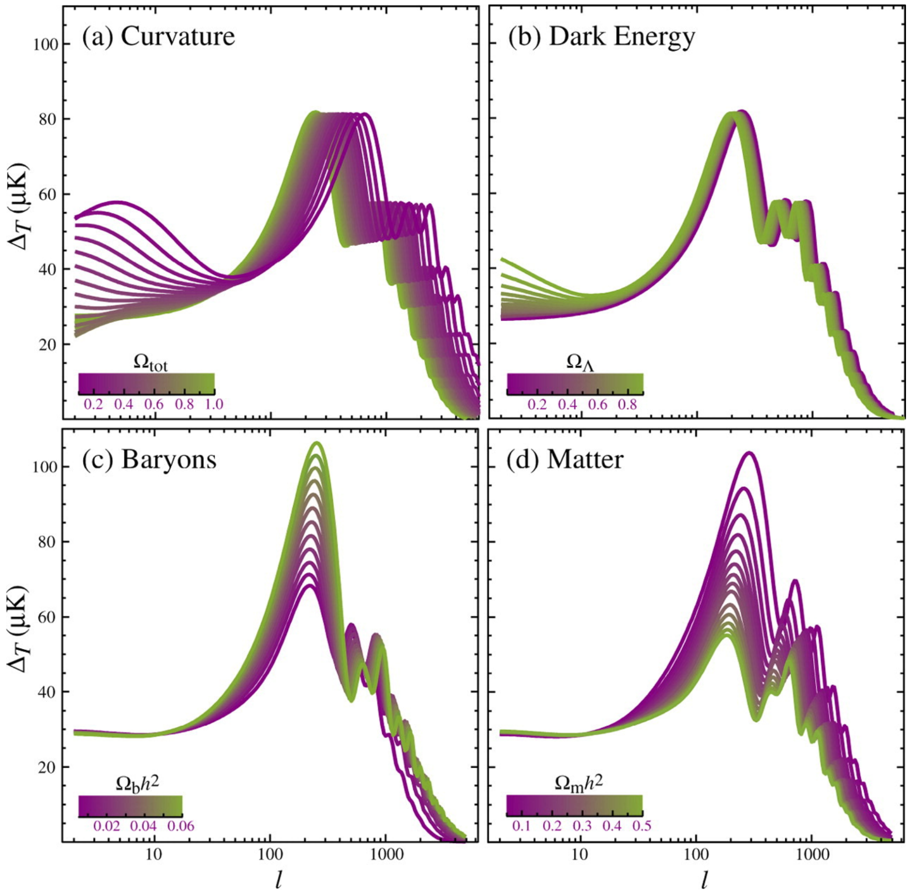


Figure 1.5: Impact of varying cosmological parameters on the CMB temperature power spectrum. Figure taken from Hu et al. (1997). Top left: varying curvature as quantified by $\Omega_{\text{tot}}$. Top right: varying dark energy content $\Omega_\Lambda$. Bottom left: varying baryon density $\Omega_b$. Bottom right: varying matter density $\Omega_m$. All parameters are varied around values $\Omega_{\text{tot}} = 1, \Omega_\Lambda = 0.65, \Omega_b = 0.02, \Omega_m = 0.147$ and $n_s = 1$.

on CMB observations. We refer to Howlett et al. (2012) for an example of CMB modeling study using `CAMB` and for discussions about well-known degeneracies of parameter constraints derived from angular power spectra.

**Additional anisotropies** In addition to the primary anisotropies generated during recombination and reionization, two additional channels of signal must be considered: secondary anisotropies and foregrounds. Secondary anisotropies arise from the interaction of photons with cosmic structures. For example, the Sunyaev-Zel'dovich (SZ) effect, which includes both kinetic (kSZ) and thermal (tSZ) components, occurs due to the inverse Compton scattering of CMB photons off hot electrons in the intra-cluster medium. This signal depends on the large-scale structure of the Universe. Another effect, gravitational lensing, occurs as CMB photons are deflected by gravitational potential wells, providing additional sensitivity to the large-scale structure. While secondary anisotropies are generally smaller in magnitude compared to primary anisotropies, they contain significant cosmological information. However, this thesis will primarily focus on the primary anisotropies. Foreground signals, on the other hand, originate from extragalactic radio or dusty sources, and galactic dust and synchrotron. Although these signals also offer interesting scientific opportunities, they are considered contaminants to the primary CMB signal.

**Polarization** The polarization of the CMB is also a powerful tool to constrain cosmology and is expected to be the primary source of information for current and future experiments (Galli et al., 2014). For a detailed treatment of the generation of polarized signals, please refer to Hu and White (1997); White (1998) and the references therein. From the linear theory of scalar perturbations, polarization is only generated from Thomson scattering by a quadrupole temperature anisotropy. Only 10% of the CMB signal is expected to be polarized. This explains the need for additional detectors to measure polarization. The polarization signal can be decomposed into non-local E and B modes, representing curl-free and divergence-free modes, respectively. Scalar perturbations generate only E modes, while vector and tensor perturbations produce both E and B modes. Therefore, B modes serve as a direct probe of the inflationary era, which might generate tensor perturbations. Additionally, the B mode can be generated through lensing of the E mode, making it a probe of large-scale structure (Seljak and Hirata, 2004). However, this lensing-induced B mode dominates at small scales, making the detection of primordial B modes particularly challenging and a major goal of current and future experiments (Baumann et al., 2009). Since the E-mode polarization signal is mostly generated by density fluctuations, it is correlated with the temperature signal, and the two are usually analyzed together to extract cosmological information, serving as consistency checks to one another and breaking degeneracies between parameters. In this thesis, we will focus on the T and E signal and will not consider the B signal.

The full-sky temperature and polarization SMICA maps produced by the *Planck* satellite are displayed in Figs. 1.6 and 1.7.

**CMB data analysis** Analyzing temperature and polarization primary anisotropies is relatively straightforward due to the Gaussian nature of the signal and noise. However, the analysis becomes computationally expensive due to the immense size of the current data sets required for achieving high sensitivity. Despite this challenge, the Gaussianity of the signal allows for the construction of nearly-optimal estimators. The standard pipelines for analyzing the CMB data consist of several key steps, as depicted in Fig. 1.8. Firstly, the raw detector data is calibrated and binned to form maps of the CMB temperature and polarization anisotropies. These maps are then compressed into their angular power spectra, which capture the complete statistical properties of the field. The angular power spectrum is subsequently utilized to constrain cosmological parameters through the use of a likelihood function. Monte Carlo Markov Chain (MCMC) methods are employed to explore the parameter space, while Boltzmann codes are employed to compute the theoretical predictions necessary for the analysis. It is important to note that the large size of the data sets and the computational complexity of the analysis pose significant challenges in this process. Therefore, the analysis pipelines are constantly being improved to achieve higher sensitivity and efficiency. In this thesis, we will focus on improvements to the analysis pipeline for the upcoming SPT-3G cosmological constraints, which are paving the way for stage-4 CMB experiments CMB-S4 (CMB-S4 Collaboration, 2019).

## Prospects and challenges of high-precision cosmology: the SPT-3G experiment

Ongoing Stage-3 CMB experiments, such as the South Pole Telescope (SPT) (Dutcher et al., 2021; Balkenhol et al., 2021; Balkenhol et al., 2022), the Atacama Cosmology Telescope (ACT) (Choi et al., 2020; Aiola et al., 2020), POLARBEAR (POLARBEAR Collaboration et al., 2014), BICEP/Keck (BICEP/Keck Collaboration et al., 2022) and upcoming Simons Observatory (SO) (Simons Observatory Collaboration et al., 2019), aim to improve *Planck* constraints in Table 1.1. Independent constraints are necessary to ensure the robustness of the

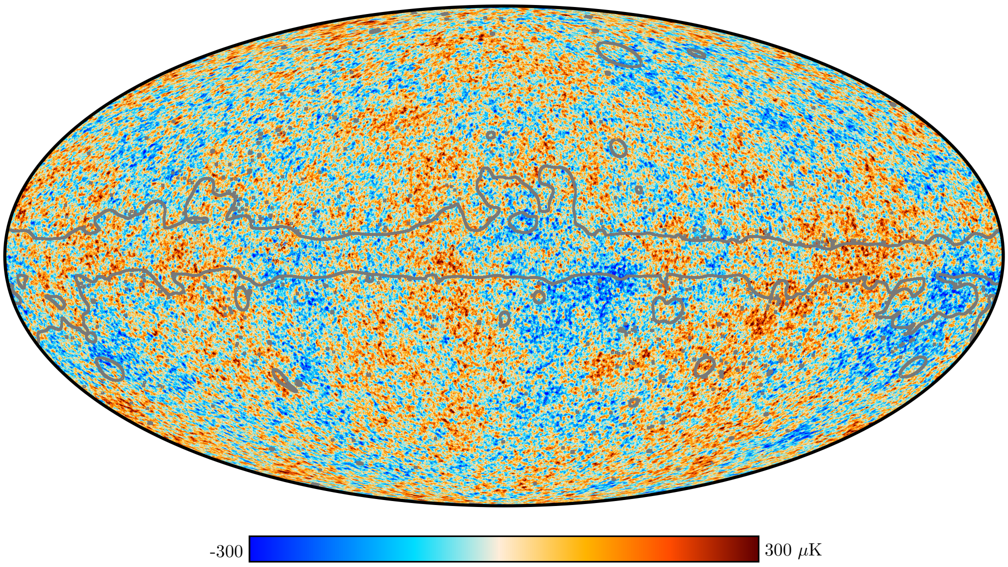


Figure 1.6: *Planck* SMICA component separated temperature map, displaying CMB anisotropies. Red zones indicate locations of the sky with higher photon temperatures. The galactic cuts are indicated with grey lines and the signal inside is generated using a constrained realization, and not used in *Planck* analysis pipeline. In this thesis, we will depict a similar technique to fill empty regions in CMB maps with high precision, so that the constrained signal can be used in the analysis.

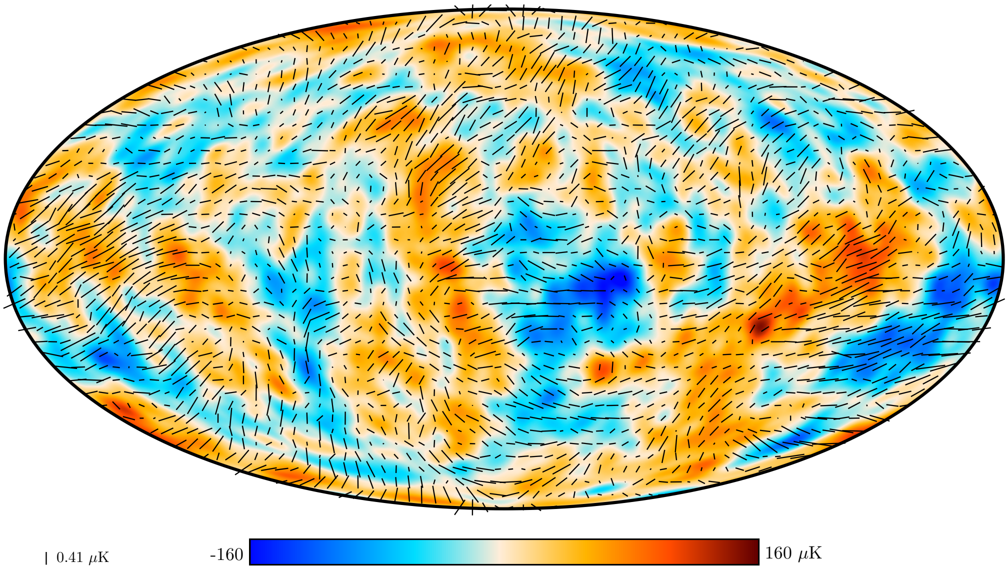


Figure 1.7: *Planck* SMICA component separated polarization map, displaying CMB polarization anisotropies. The black lines indicate the polarization direction, while the color indicates the polarization intensity.

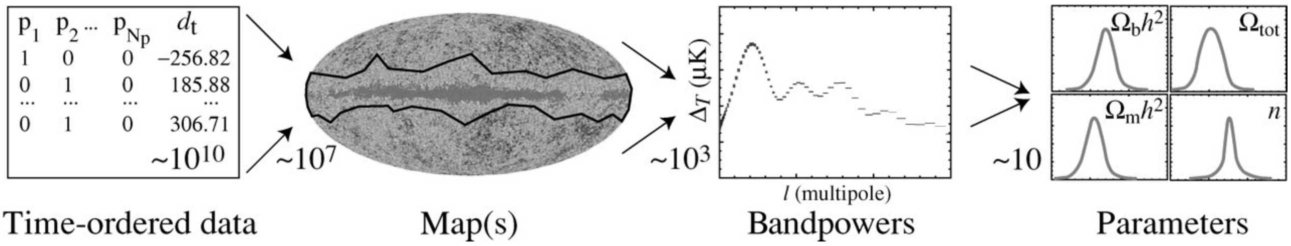


Figure 1.8: Schematics of CMB analysis pipeline. The raw data is compressed into maps, which are then decomposed into band powers. They are used to constrain cosmological parameters using a likelihood analysis. Figure taken from Hu et al. (1997).

cosmological models, investigate systematic effects in the data and probe new physics. This thesis will primarily focus on the endeavors of the SPT.

## SPT-3G and its science goals

The SPT is a 10-meter telescope observing the millimeter-wave sky from the geographic South Pole. Atmospheric conditions are ideal for such observations, and the SPT has already produced two surveys, SPT-SZ (Story et al., 2013) and SPTpol (Austermann et al., 2012), with associated scientific results: discovering new galaxy clusters with the SZ effect (Staniszewski et al., 2009), the discovery of strongly lensed high-redshift star-forming galaxies (Vieira et al., 2013), CMB temperature anisotropy measurement (Story et al., 2013) and the first detection of B modes (Hanson et al., 2013). The next generation SPT-3G camera provides additional sensitivity with improved optical design and multichroic pixels, which allows a deep observation of a relatively wide field starting from 2018 (Benson et al., 2014). This high-sensitivity instrument is used among the collaboration for a wide range of scientific objectives including the delensing of the Bicep/Keck field (Bicep/Keck Collaboration et al., 2022); the study of the CMB lensing (van Engelen et al., 2012; Story et al., 2015; Wu et al., 2019), the study of the CMB secondary anisotropies, the search for the Sunyaev-Zel'dovich (SZ) effect and high-multipole CMB temperature anisotropies (Keisler et al., 2011; Reichardt et al., 2012; Story et al., 2013; George et al., 2015); the analysis of low-multipole polarization anisotropies in the search for the primordial B-modes; the study of cosmic birefringence and axion-like particles (Bianchini et al., 2020; Ferguson et al., 2022); the study of galaxy clusters (Santos et al., 2008; Bleem et al., 2015; Raghunathan et al., 2019; Bleem et al., 2020); the joint analysis with DES-Y3 (Raghunathan et al., 2019; Abbott et al., 2022b); the search for asteroids (Chichura et al., 2022). This list gives a general overview of the decisive science inputs that can be achieved with the instrument. In this work, we will focus on the constraints on cosmological models and associated parameters from CMB primary anisotropies by the SPT-3G experiment. This work was already started for SPTpol (Henning et al., 2018), and also for the first four months of the SPT-3G survey, for which half the focal plane was available (Dutcher et al., 2021; Balkenhol et al., 2021; Balkenhol et al., 2022), labeled **SPT-3G 2018** for the rest of this work. The full SPT-3G survey is expected to be completed in 2023 and will cover 1700 square degrees of the sky. While the full five-year survey will soon be available for analysis, we will focus in this thesis on the preparation of the cosmological results with the first two years of this updated survey, and refer to this data set as **SPT-3G 19/20**. Additional data is collected on extended fields which will add 2800 $\mathrm{deg}^2$ of observations, improving the parameter constraints.

## Tensions and unkowns

In Fig. 1.9, we present the most recent compilation of CMB power spectrum measurements, including the latest results from the SPT-3G 2018 survey. While the *Planck* satellite mission focused on precisely measuring the large and intermediate angular scales of the CMB temperature field, ground-based telescopes such as ACT, POLARBEAR, and SPT have achieved high sensitivity by increasing the number of detectors on their focal planes. This enables them to probe the high multipoles of the CMB power spectrum in both temperature and polarization. Although these ground-based experiments cover a reduced sky fraction, resulting in decreased constraining power due to cosmic variance, their additional sensitivity compensates for this limitation. Exploring a complementary range of multipoles is crucial for confirming the consistency of the cosmological model, especially in light of the ΛCDM model's inability to fully explain the observed data.

Indeed, recent measurements of the CMB power spectrum by the ACT and SPT ground-based experiments, as reported in Balkenhol et al. (2022); Aiola et al. (2020), have confirmed and amplified the so-called **Hubble tension**. This tension arises from the observed discrepancy between the inferred values obtained from early-Universe data sets and local measurements of the Hubble constant. The former relies on fitting the ΛCDM model to CMB or baryon acoustic oscillations measurements, while the latter uses observations at low redshifts.

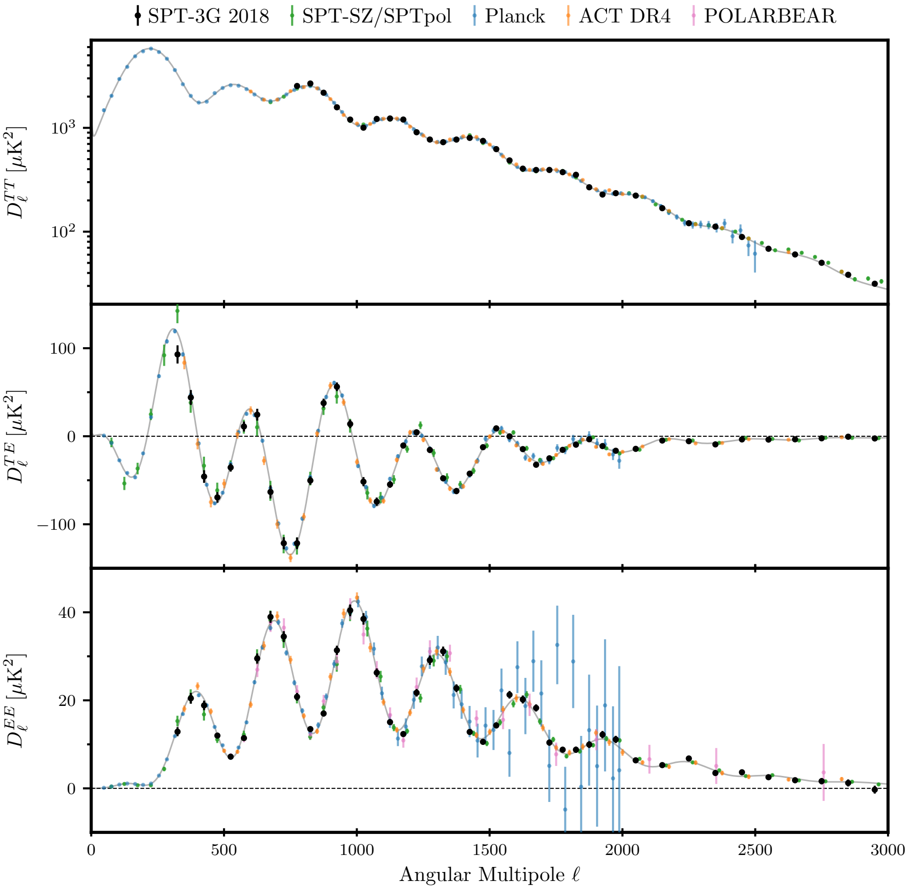


Figure 1.9: Compilation of CMB power spectrum measurements taken from SPT (Balkenhol et al., 2022). SPT-3G 2018 data points are displayed with their error bars in black, while the previous SPTpol and SPT-SZ data sets are shown in green. The *Planck* 2018, ACT DR4 and POLARBEAR data points are shown respectively in blue, orange, and pink.

The value derived from the high-redshift data, as determined by the *Planck* collaboration, is $H_0 = 67.4 \pm 0.5$ km/s/Mpc (Planck Collaboration et al., 2020c). On the other hand, the low-redshift value, based on the SH0ES collaboration's measurement using a Cepheid-calibrated supernova IA distance ladder, is $H_0 = 73.29 \pm 0.90$ km/s/Mpc (Murakami et al., 2023). The Hubble tension represents a $5\sigma$ discrepancy between these two measurements, and it is a significant problem given that alternative CMB data sets and combined measurements of Big Bang nucleosynthesis with baryon acoustic oscillation data strongly support the *Planck* value (Bennett et al., 2013; Aubourg et al., 2015; Addison et al., 2018; Blomqvist et al., 2019; Macaulay et al., 2019; Aiola et al., 2020; Balkenhol et al., 2022), while the distance ladder measurement is robust against systematic investigations. Additionally, alternative calibration methods of the distance ladder, such as the tip of the red giant branch (Freedman et al., 2019), yield larger values for $H_0$ than the *Planck* measurement. For a comprehensive review of the Hubble tension, we refer the reader to Verde et al. (2019) and Schöneberg et al. (2022). Similar discrepancies have also emerged in the measurement of the structure growth rate, which can be parameterized either by the amplitude of matter fluctuations within a comoving volume of 8Mpc$^{-1}$ or by the combined structure growth parameter $S_8 \equiv \sigma_8\sqrt{\Omega_m/0.3}$, to which joint galaxy clustering and weak lensing analyses are more sensitive. Current measurements indicate a 2 to $3\sigma$ inconsistency between *Planck* and measurements from the DES-Y3 and KIDS-1000 data sets (Heymans et al., 2020; Abbott et al., 2022a). A review of structure growth is provided in Huterer et al. (2015); Huterer (2022). In Fig. 1.10, we display the measured values of $H_0$ and $S_8$ from the SPT-3G 2018 survey, *Planck*, and ACT DR4. We also include the reference values of $H_0$ from Riess et al. (2022) and $S_8$ from Abbott et al. (2022a). The SPT-3G 2018 data set is in strong agreement with *Planck* and ACT DR4, but the $H_0$ value shows tension with the measurement of the SH0ES team. While the *Planck* value of $S_8$ is in tension with the DES Y3 and KIDS-1000 measurements, the SPT-3G 2018 data set agrees with these measurements, albeit with larger error bars.

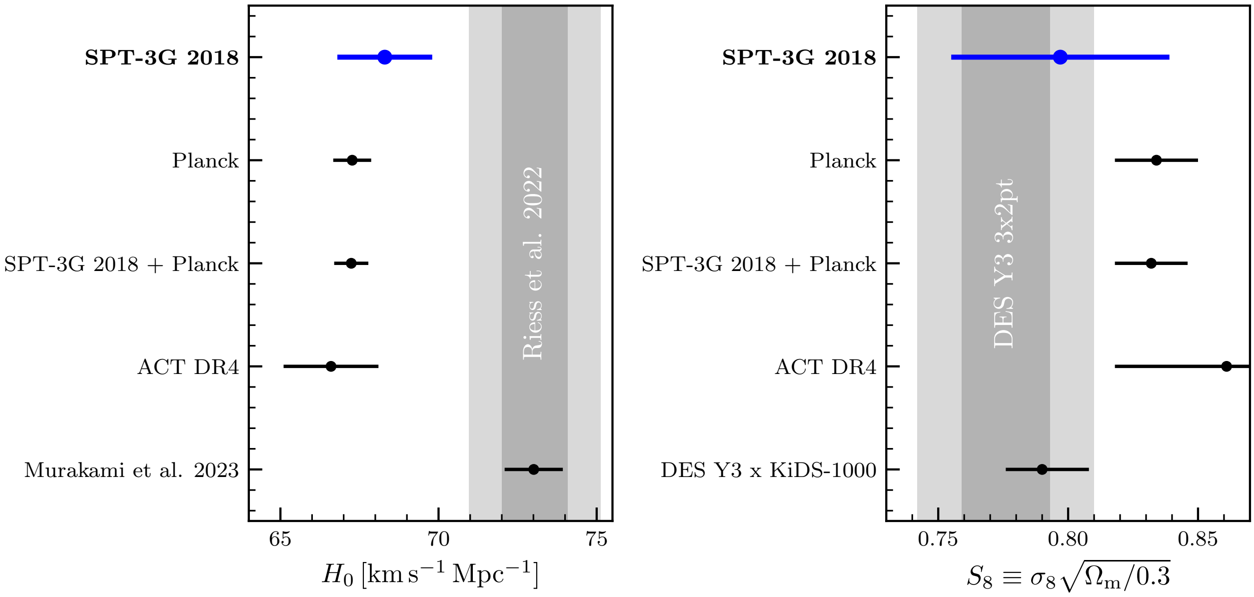


Figure 1.10: Constraints on the Hubble constant $H_0$ and the structure growth parameter $S_8$ from *Planck* 2018 (Planck Collaboration et al., 2020c), ACT DR4 (Aiola et al., 2020) and SPT-3G 2018 (Balkenhol et al., 2022). The $H_0$ reference value from Riess et al. (2022) is displayed with grey error bars on the left panel with an updated value with improved systematics treatment from Murakami et al. (2023). The $S_8$ reference value from the DES Y3 3x2pt (Abbott et al., 2022a) is displayed with grey error bars on the right panel. A combined constraint of DES Y3 and KiDS-1000 is also given (Survey et al., 2023). Note that ACT values are different from the reference paper since different $\tau$ priors were chosen for the analysis, in a consistent approach with SPT-3G 2018. Plot adapted from Balkenhol et al. (2022).

While systematic errors in the analysis pipeline could potentially explain the observed discrepancy, the wide array of probes suggests that these tensions stem from a failure of the ΛCDM model. Currently, there is a dedicated effort underway to comprehend the origins of these tensions and propose new physics that aligns with the observations (Schöneberg et al., 2022). Although one possible explanation for the Hubble tension is to increase the dark energy density at low redshift, such solutions are contradicted by measurements of supernovae and baryon acoustic oscillations at low redshift. Consequently, the most promising avenues involve modifying the sound horizon at recombination through new early-Universe physics, such as varying electron mass models (Hart and Chluba, 2020) or early dark energy models (EDE) (Doran and Robbers, 2006; Poulin et al., 2019; Niedermann and Sloth, 2020), but such models could eventually aggravate the $S_8$ measurements (Jedamzik et al., 2021).

The SPT-3G 19/20 data set is projected to have four times lower noise compared to the 2018 data set, resulting in achieved noise levels of $6/5/16\mu$K-arcmin in temperature at 95/150/220GHz. These noise levels are approximately 6 times lower than those of *Planck*. As a consequence, the SPT-3G 19/20 data set will significantly enhance measurements of $H_0$ and $S_8$ and provide valuable insights into the nature of both tensions. Fig. 1.11 showcases the current ΛCDM constraints using *Planck* and SPT-3G data sets, along with projected constraints obtained using mock band powers, considering the expected noise levels for SPT-3G 19/20 and the full five-year survey. The SPT-3G 19/20 data set will be able to constrain parameters within 20% of the full five-year survey and will be in direct competition with *Planck* 2018. Due to the high resolution of the SPT-3G experiment, it will introduce degeneracy directions orthogonal to *Planck*, for example for $\omega_{\rm b}$. The combined constraints from both experiments will break degeneracies and lead to significant improvements in the measurements of parameters. In addition, the data set will mostly be independent, as *Planck* probes the largest angular scales of temperature, while SPT-3G probes the smallest angular scales of polarization. Combined constraints are expected to be twice as tight as the *Planck* 2018 constraints, thus reaching the tightest constraints on ΛCDM parameters to date.

Ongoing stage-3 experiments (SPT, ACT, POLARBEAR) have produced significant results and will continue to do so, but they are also paving the way for Stage 4 technology experiments, which will be able to reach the cosmic variance limit. The Simons Observatory (SO) (Simons Observatory Collaboration, 2019) and CMB-S4 (CMB-S4 Collaboration, 2019) experiment have as a primary scientific goal to unveil the imprint of primordial gravitational waves on the CMB polarization as quantified by the tensor-to-scalar ratio, and output constraints on ΛCDM extensions 3 to 10 times narrower. In this work, we seek to improve the accuracy and

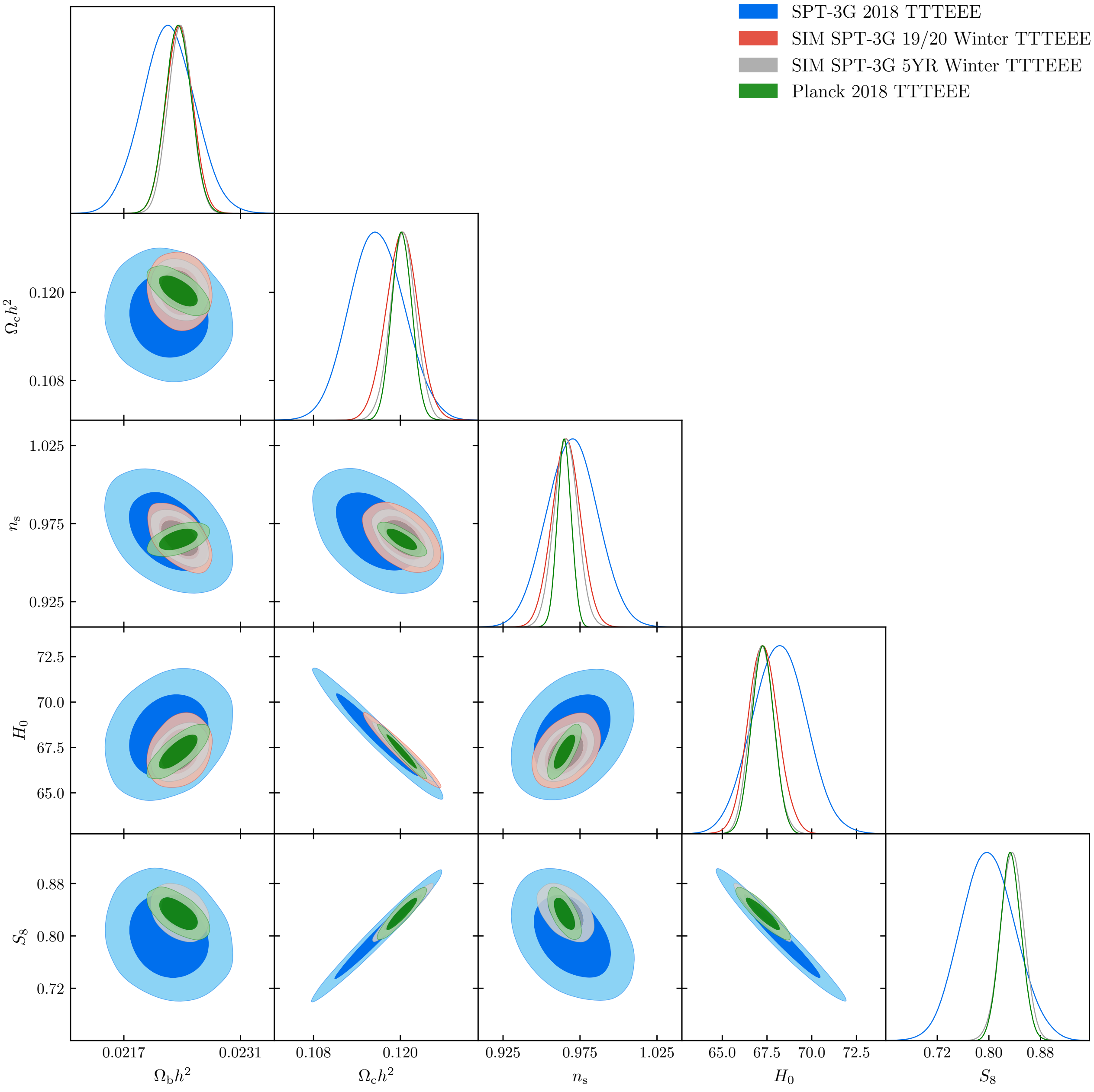


Figure 1.11: *Planck* 2018 TT TE EE (green) and SPT-3G 2018 TT TE EE (blue) cosmological constraints, with forecasted constraints around *Planck* best fit for SPT-3G 19/20 (red) and SPT-3G full five years survey (grey). Both SPT-3G 19/20 will yield comparable constraints to *Planck*, further improved with the additional three years of observation.

efficiency of analysis pipelines of the SPT-3G experiment, which will be used as a testbed for the analysis of the Stage-4 experiments. As the total amount of detectors will be likely to increase by a factor of 50, it is crucial to develop efficient analysis pipelines to extract the cosmological information from the data.

# Outline of the thesis

The objective of this work is to provide a comprehensive analysis of the SPT-3G 19/20 data set, highlight the enhancements made to the analysis pipeline, and discuss the implications of the findings. This research will yield competitive and robust cosmological constraints for an upcoming data release, as indicated in Fig. 1.11. To accomplish this, the thesis will be structured around the analysis pipeline of the SPT-3G experiment, following the main compression steps illustrated in Fig. 1.8. In Chapter 2, I will introduce the SPT-3G experiment and outline the map-making pipeline. A thorough understanding of this pipeline is crucial for accurately analyzing the statistics of the CMB. Subsequently, in Chapter 3, I will present the power spectrum estimation based on the maps. Unlike the previous analysis that was conducted on a flat sky, this analysis will be performed on the entire sky, necessitating certain improvements. Notably, I have developed key features in the power spectrum estimation pipeline and provided a framework for the generation of Gaussian-constrained realizations of the data. Moving forward, in Chapter 4, I will delve into the likelihood estimation process. This chapter focuses on constructing a highly precise semi-analytical covariance framework for the Gaussian likelihood, enhancing the accuracy of our results. Finally, I will thoroughly examine the implications of these findings and conclude in Chapter 5.

# Chapter 2

# Making maps of the CMB

---

**Contributions** In this chapter, I aim to introduce the first compression step of the analysis pipeline, which is map-making. This work has been done with the SPT-3G collaboration and led by collaborator Wei Quan, relying on earlier work that is cited throughout this chapter. My contributions to this work are the following:

- Participation in the data quality monitoring and the data selection for the 19/20 season data.
- Section 2.1.4: Study of the impact of point source masking threshold on maps, band powers and covariance matrix.
- Section 2.3.1: Development of the analysis pipeline of curved sky maps, including an in-depth study of the filtering impact on the signal and noise levels computations from noise maps.
- Section 2.3.2: Development of code to create the analysis mask from weights and apodize it. Analysis of the impact of the mask apodization on the power spectra and the covariance matrix of the power spectrum. Analysis of the impact of source masking on the various data products.
- Section 2.3.3: Investigation on the noise maps and analysis of specific features that led to additional data cuts.
- Section 2.3.4: Production and analysis of pure Gaussian CMB simulations. Study of the mock observation pipeline and outputs. Integration of resulting maps in inpainting and covariance validation pipelines.

The main purpose of this section is to provide the reader with the necessary knowledge on map-making to understand the rest of the analysis.

**Introduction** The cosmic microwave background radiation light can be observed in every direction of the sky and is characterized by its temperature and polarization fields. The electromagnetic wave is decomposed into its Stokes parameters $T(\hat{\mathbf{n}}), Q(\hat{\mathbf{n}}), U(\hat{\mathbf{n}}), V(\hat{\mathbf{n}})$, where $\hat{\mathbf{n}}$ is the direction of the line-of-sight (Perrin, 1942; Chandrasekhar, 1960). The first parameter is the temperature, or intensity, of the radiation, and the three others are the linear and circular polarization. A linear horizontal polarized field would have, with arbitrary units, $T = 1, Q = 1, U = 0, V = 0$, while a right-hand circularly polarized field would have $T = 1, Q = 0, U = 0, V = 1$. $T$ and $V$ parameters are independent of the local basis while the $Q$ and $U$ parameters are dependent on the definition of the angle on the sky, since they are related to linear horizontal and vertical polarization, respectively. Under a rotation of the local basis of angle $\alpha$, the Stokes parameters transform as

$$\begin{pmatrix} T \\ Q \\ U \\ V \end{pmatrix} \to \begin{pmatrix} T \\ Q\cos 2\alpha + U \sin 2\alpha \\ -Q \sin 2\alpha + U \cos 2\alpha \\ V \end{pmatrix}. \tag{2.1}$$

$T$ and $V$ are hence spin-0 fields, and it is convenient to define a new polarization parameter $P \equiv Q + iU$, and its independent complex conjugate $P^*$. They are spin-2 fields (Zaldarriaga and Seljak, 1997; Seljak and Zaldarriaga, 1997; Chon et al., 2004) and, under a rotation $\alpha$, transform as

$$P(\hat{\mathbf{n}}) \to e^{-2i\alpha} P(\hat{\mathbf{n}}), P^*(\hat{\mathbf{n}}) \to e^{2i\alpha} P^*(\hat{\mathbf{n}}). \tag{2.2}$$

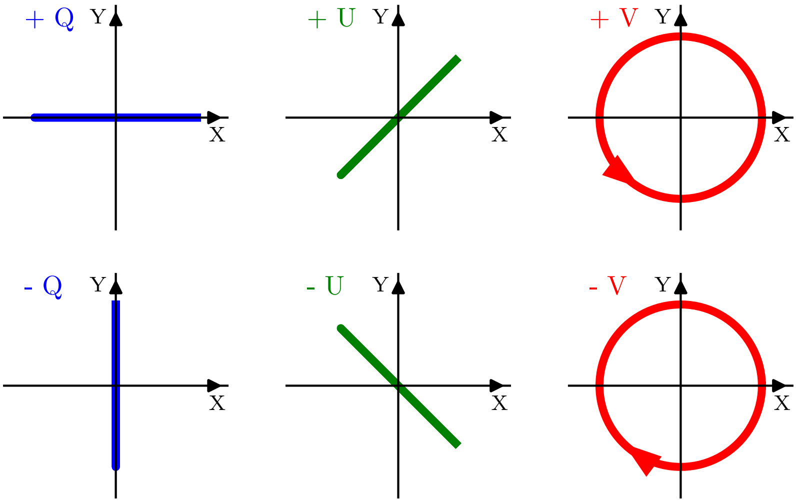


Figure 2.1: Representation of the Stokes parameters $T, Q, U$ in the horizontal-vertical basis. A rotation of the local basis would mix $Q$ and $U$ while leaving $T$ and $V$ unchanged.

Throughout the rest of this thesis, we will ignore the $V$ parameter. It cannot be generated by Thomson scattering and hence is not detectable in the CMB in classical models. In addition, the SPT-3G instrument is not sensitive to circular polarization, see Section 2.1.1.

The main data product of the SPT-3G experiment are maps of the $T, Q, U$ fields of the CMB anisotropies, that are built from the raw output of detectors. This chapter aims to describe this first compression step and its implementation in the SPT-3G data analysis pipeline. In Section 2.1, we will introduce the SPT telescope, its current instrument, and the measured detectors data, then in Section 2.2, we will describe the map-making algorithm applied to the previous time streams to obtain maps of the sky. We will present the maps and associated simulations in Section 2.3.

## 2.1 The SPT experiment

### 2.1.1 Telescope, optics and focal plane

The South Pole Telescope (SPT) is a 10-meter diameter telescope situated at the Amundsen-Scott South Pole Station in Antarctica. Its strategic location at an altitude of 2835 m and with low atmospheric humidity provides optimal conditions for millimeter wavelength observations. Additionally, the absence of a diurnal cycle ensures exceptionally stable atmospheric conditions with minimal Sun contamination (Radford and Holdaway, 1998; Radford, 2011). The telescope's optics, as depicted in Fig. 2.2, are specifically designed to maximize its efficiency in capturing millimeter wavelengths. A complete report on the telescope instrumentation can be found in Sobrin et al. (2022), hereafter S22.

Deployed in 2017 on the focal plane, the SPT-3G instrument offers an expanded field of view compared to its predecessors SPT-SZ (Bleem et al., 2015) and SPTpol (Bleem et al., 2012), increasing from $1\,\mathrm{deg}^2$ to around $2.8\,\mathrm{deg}^2$. As shown in Fig. 2.2, the instrument is divided into two cryostats for optics, both cooled to a temperature of 4K using pulse tube coolers. After traveling through the optical path, the light hits the focal plane that consists of 14266 transition edge sensor (TES) polarization-sensitive bolometers. These bolometers are further cooled to around 300mK using a 3He/4He dilution refrigerator. Coupled with the large aperture of the telescope, this extensive array of detectors enables high sensitivity and high-resolution observations of the cosmic microwave background. The focal plane, divided into 10 hexagonal modules known as wafers, each contains 269 trichroic pixels, as illustrated in Fig. 2.3.

The pixel architecture, used generally among CMB experiments as in POLARBEAR (Suzuki et al., 2012), SO (Galitzki et al., 2018), and LiteBIRD (Suzuki et al., 2018), is illustrated in Fig. 2.4. The incoming radiation is coupled to the sinuous antenna through a silicon lenslet (O'Brient et al., 2010). Antennas have an intrinsic orientation, that we characterize by the time-dependent sky-respective polarization angle $\psi_t$. The log-periodic antenna design aims at maximizing the polarization efficiency $\gamma$, which characterizes the response of the antenna to an input signal polarized along the detector plane. On each wafer, half of the antennas are mirrored and another half is rotated by 45 deg to minimize the effect of the instrumental systematics and to evenly sample the linearly polarized sky. The incoming radiation is then coupled to the TES bolometer through a microstrip transmission line. The TES bolometer is a superconducting device that is biased at its transition temperature.

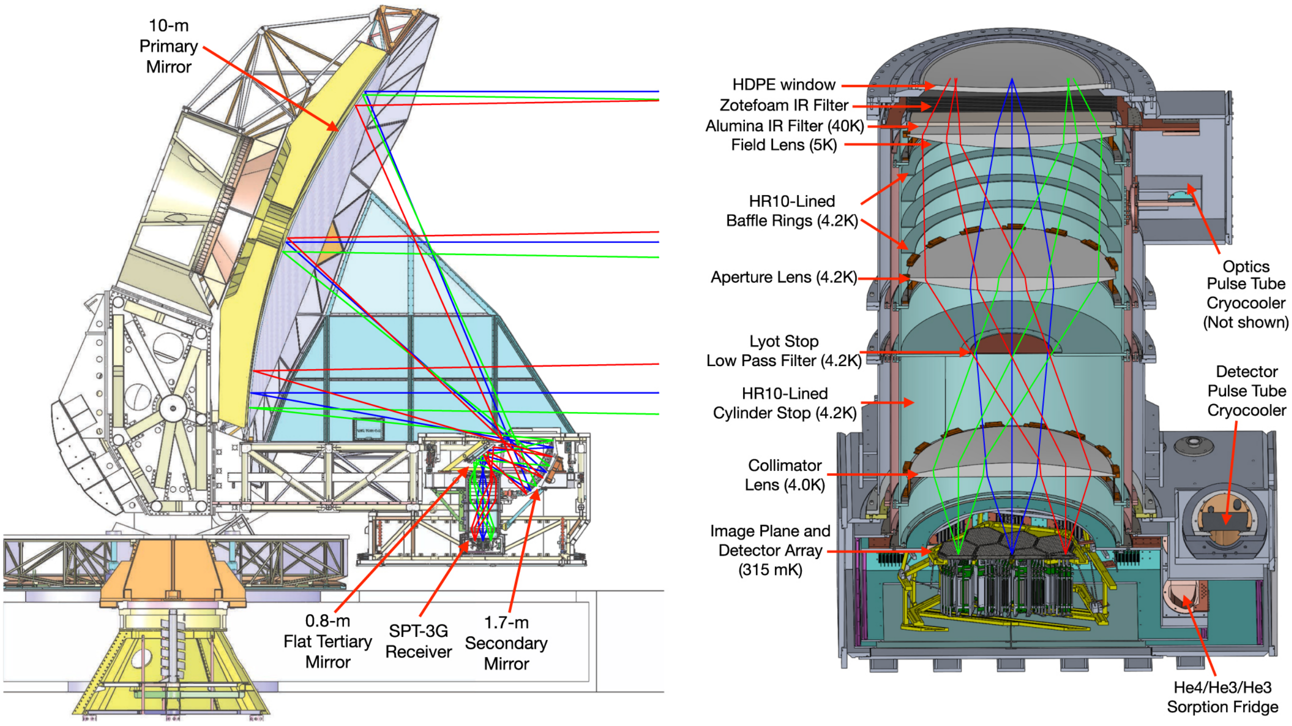


Figure 2.2: Schematic of the SPT-3G optics, from S22.

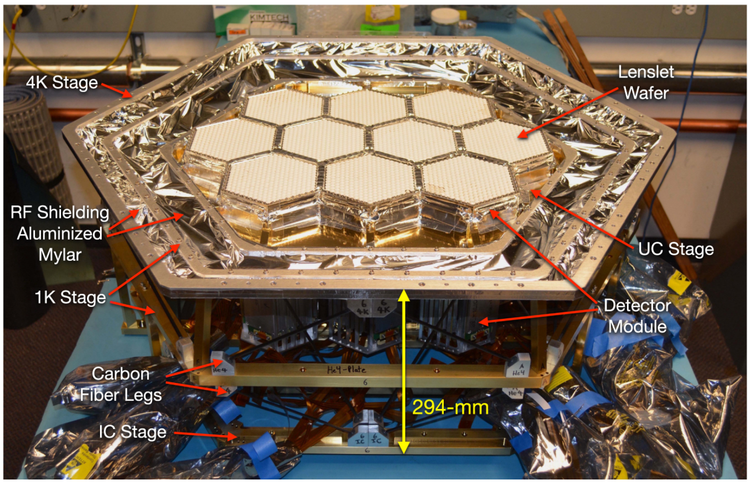


Figure 2.3: Photo of the SPT-3G focal plane, taken from S22. The ten hexagonal units are the wafers.

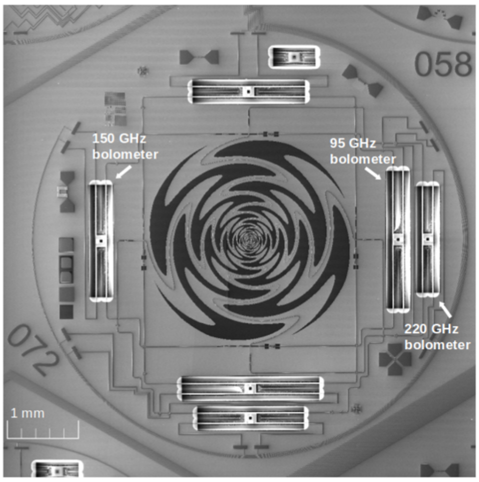


Figure 2.4: Photo of the SPT-3G detector, from S22. The log-periodic antenna is optimized to capture the incoming light polarization.

The incoming radiation is absorbed by the TES bolometer, which changes its temperature and thus its resistance. The change in resistance is multiplexed and then measured by a SQUID (Superconducting Quantum Interference Device) amplifier, a very sensitive magnetometer. These operations are repeated as detectors operate in three frequency bands centered at 95, 150, and 220GHz. For details on the readout, please refer to S22 and Stiehl et al. (2011); Doriese et al. (2015); Bender et al. (2014, 2016, 2020).

### 2.1.2 Footprint and scanning strategy

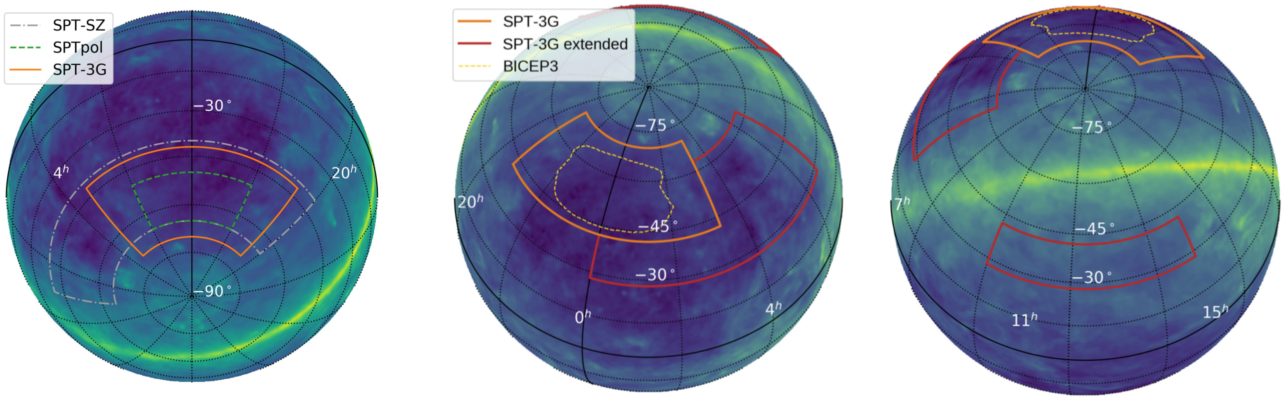
**Footprint** One of the primary scientific objectives of the SPT-3G instrument is to the precise removal of lensing effects from the B-mode signal observed by the large aperture BICEP/Keck experiment. Therefore, the careful selection of the survey footprint is of paramount importance to maximize the spatial overlap between the two experiments. The BICEP3 footprint was deliberately chosen to minimize contamination from galactic emissions (BICEP/Keck Collaboration et al., 2022). The SPT-3G deep survey footprint, or **winter field** extends over a range of declination (DEC) from $-42\,\mathrm{deg}$ to $-70\,\mathrm{deg}$, and right ascension (RA) from $20^{\mathrm{h}}40'$ to $3^{\mathrm{h}}20'$. Encompassing an area of approximately $1700\,\mathrm{deg}^2$, the footprint is illustrated in Fig. 2.5a. The overlap between the SPT-3G and BICEP/Keck survey footprints is depicted in Figs. 2.5b and 2.6. While Fig. 2.6 provides a comprehensive view of the full extent of the BICEP 3 field, Fig. 2.5b showcases the scientific footprint defined at the mean survey weight. Notably, the latter is fully contained within the SPT-3G survey footprint, enabling accurate delensing procedures. Additionally, during austral summer, when significant signal contamination arises from the Sun's side-lobes in the main or winter field, extended survey observations are conducted in the **summer fields**, as depicted in Fig. 2.5b, on a sky area of $2800\,\mathrm{deg}^2$. These extended fields participate in yielding even more precise constraints on cosmological parameters thanks to additional data. Due to the limited duration of the fridge cycles, it is impractical to observe the entirety of the fields in a single observation. Consequently, the SPT-3G fields are divided into **subfields** covering the full RA range and centered around specific DEC values. This division ensures minimal variations in atmospheric height during observations, thereby mitigating the impact of atmospheric loading and detector responsivity. The individual subfields are illustrated with varying colors in Fig. 2.7. For the winter field, the subfields are centered at DEC values of $-44.75\,\mathrm{deg}$, $-52.25\,\mathrm{deg}$, $-59.75\,\mathrm{deg}$, and $-67.25\,\mathrm{deg}$. From now on, we will call **observation** each ensemble of data acquired during 1 fridge cycle on a single subfield. When no failure, one observation covers the entire subfield.

**Observing seasons and SPT-3G 19/20 data set** The current SPT-3G focal plane has been acquiring data starting from 2019 and is expected to continue until the end of 2023. The winter field is observed 8 months per year, from March to November, while the summer fields are observed the rest of the year. In this thesis, we will focus on the winter data acquired between March 2019 and December 2020, which we label **SPT-3G 19/20**. Although this data has lower constraining power than the complete survey, it will already provide exquisite constraints on cosmological parameters.

**Scanning strategy** The SPT-3G telescope is affixed to an azimuth-elevation mount, enabling rapid azimuth scans for efficient sky coverage. The telescope follows a raster pattern while scanning the sky, as depicted in Fig. 2.8. This scanning process involves azimuthal movement at a constant elevation, followed by shifting to the next elevation step and conducting another azimuthal scan. For one observation, this procedure is repeated until the chosen subfield has been observed. To ensure the production of high-quality maps without undesired artifacts such as stripes, the elevation steps are carefully chosen to be sufficiently small and executed in a series of dither steps. Each scan represents the acquisition of data at a constant elevation and in a consistent azimuthal direction. Consequently, the telescope captures the celestial sphere through a sequence of right-going and left-going scans, carried out at various elevations, while maintaining a constant angular speed of $\omega_{\mathrm{az}} = 1\,\mathrm{deg}/\mathrm{s}$. Employing higher scanning speeds effectively shifts the CMB signal to a higher temporal frequency, thereby aiding in its separation from the $1/f$ instrumental noise characteristic. This scanning strategy yields a typical individual subfield observation time of approximately 2.5 hours.

### 2.1.3 Calibration

For subsequent cosmological investigations, it is essential to calibrate the **time-ordered data** (TOD), or **time stream**, generated by the detectors, thereby converting them into temperature units. We employ an external calibration approach utilizing two well-established galactic star-forming regions referred to as HII regions, namely RCW38 and MAT5A. These regions possess known flux values at the frequencies of interest, determined through a combination of measurements conducted by various experiments, including *Planck*, SPT-SZ, BOOMERANG (de Bernardis et al., 1997), and ACBAR (Runyan et al., 2003). Rigorous observational campaigns targeting these regions allow us to establish the calibration for each detector. The calibration process involves fitting a model to the acquired time-ordered data. In addition to the primary calibration,



(a) SPT-SZ (de Haan et al., 2016), SPTPOL (George et al., 2012) and SPT-3G. The figure is taken from Dutcher et al. (2021). While SPT-SZ probed temperature on a larger sky fraction, SPTPOL observed polarization on a deep patch and SPT-3G combined high-resolution, deep and relatively wide patch.

(b) SPT-3G winter field (orange) and extended summer fields (red) The winter field is the baseline deep field observed during 8 months covering 1700 deg$^2$, whereas summer extended fields are observed during the remaining 4 months to avoid Sun contamination. They cover a total area of 2800 deg$^2$. Figure taken from S22.

Figure 2.5: Survey footprints overplotted on a map of the thermal dust emission as measured by *Planck*.

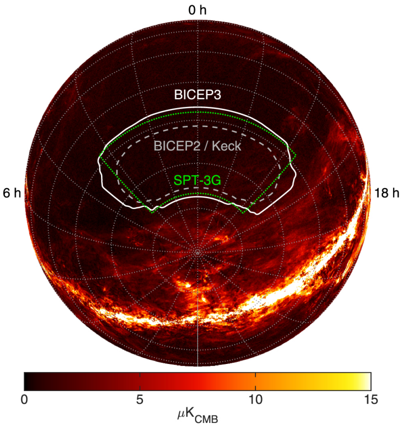


Figure 2.6: BICEP2, BICEP 3 and SPT-3G footprints overplotted on a map of polarized dust emission from *Planck*. Figure taken from BICEP/Keck Collaboration et al. (2022)

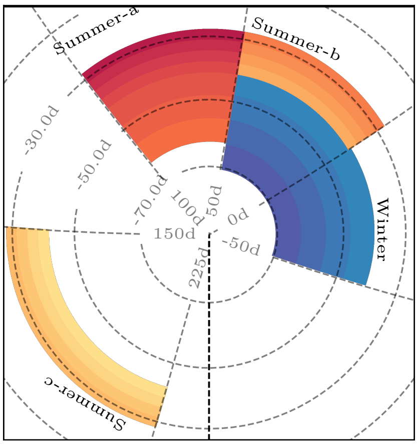


Figure 2.7: Fields of the SPT-3G experiments for winter field (blue), summer-a field (red), summer-b field (orange), and summer-c field (yellow). Subfields are highlighted using varying colors. Courtesy of Wei Quan.

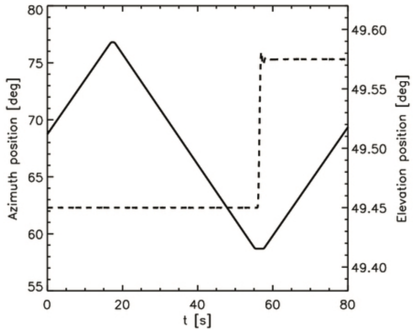


Figure 2.8: Portion of the observation scanning strategy. The full line indicates the azimuthal position while the dashed one indicates the elevation position. Figure taken from Carlstrom et al. (2011).

three supplementary calibration steps are essential to account for atmospheric loading and weather-induced effects. Firstly, given that changes in elevation result in variations in atmospheric height and subsequent alterations in atmospheric loading, it becomes crucial to measure the elevation-dependent gain of each detector. This necessitates the use of a chopped thermal source positioned at the aperture of the telescope, enabling a comparison of the detector gain between on-source and off-source positions. Secondly, the calibration of detectors, an expensive procedure conducted once per week, is susceptible to potential drift caused by changing weather conditions. To counteract this potential drift, a rapid calibration procedure is performed at the onset of each observation, utilizing either RCW38 for winter subfields at lower declinations or MAT5A for winter subfields at higher declinations. Finally, the demodulation of the signal at the output of the TES is contingent upon the airmass observed by the detector. This dependency is corrected by measuring the detector response at both elevation angles of $90\,\mathrm{deg}$ and $30\,\mathrm{deg}$, and subsequently interpolating to determine the response at the elevation of the targeted source. These three calibration steps hold significance as they contribute to the generation of reliable CMB maps while effectively mitigating potential systematic errors. The absolute calibration uncertainty associated with this experiment is estimated to be 10%, and further refinement will be achieved during the subsequent analysis through the cross-correlation of coadded maps with the maps provided by the *Planck* mission.

The uncalibrated TOD of a given detector can be modeled as a $n_{\mathrm{tod}}$ series of values indexed by the time $t$

$$y_t = g\left[(2-\gamma)T_t + \gamma\cos\left(2\psi_t\right)Q_t + \gamma\sin\left(2\psi_t\right)U_t\right] + n_t, \tag{2.3}$$

where $T_t$ is the intensity observed by the detector at time $t$, while $Q_t$ and $U_t$ are Stokes linear polarization parameters. The detectors are not sensitive to the circular $V$ polarization. Here we used the overall detector gain $g$, the polarization angle $\psi_t$, and the polarization efficiency $\gamma$. Ground-based experiments suffer from large atmospheric fluctuations that dominate the noise at low frequencies and noise caused by a drifting cold stage temperature, which has a typical $1/f$ shape. The time stream is dominated by this $1/f$ noise $n_t$ with a characteristic shape, as illustrated in the first panel of Fig. 2.9. Let us also point out that the scanning strategy introduced in Section 2.1.2 makes the polarization angles of the detectors time-independent. We will thus drop the time dependency for the rest of this work. Assuming calibrated data, we set the overall gain to its unity value, redefine the noise and recast the previous equation as

$$y_t = T_t + \frac{\gamma}{2-\gamma}\left[\cos\left(2\psi\right)Q_t + \sin\left(2\psi\right)U_t\right] + n_t \tag{2.4}$$

The ultimate goal is to compress the information contained in the sets of calibrated detector time streams into maps of the CMB anisotropies. The calibrated data consists of subfield observations of associated left-going and right-going scans for each detector, where data is kept only when the telescope speed is constant, while the data acquired when changing direction is discarded. Before map-making, we perform operations on the time stream, see Section 2.1.4. First, we remove misbehaving detectors. Then, we mask sources to prepare the filtering step. Finally, ground-based experiments TOD are dominated by complex structure noise which we filter out to simplify the map-making process.

### 2.1.4 TOD processing

**Data cuts** To mitigate data contamination arising from pathological detectors, we implement a series of stringent data cuts. During observations, detectors may exhibit unexpected behavior, such as saturation, poor calibration, excessive noise, or readout errors, leading to their identification and flagging. Additionally, time streams containing glitches or excessive line power in the frequency domain are removed. In Table 2.1 is displayed the number of observations per subfield and the associated average number of active detectors for SPT-3G 19/20 data set. The number of observations decreases as the center elevation of the subfields increases, in order to reduce noise for subfields observed through more atmosphere. Approximately 1000 detectors per frequency band are removed prior to the map-making.

**Point source masking** During the scanning process, the telescope inevitably traverses regions containing galactic sources, extragalactic active galactic nuclei (AGNs), and clusters of galaxies. These sources possess sufficient brightness to be detected by the instrument and must be masked prior to filtering. Failure to do so would result in the subsequent removal of low-frequency components inducing filtering artifacts originating from the point sources, thereby contaminating the signal and introducing biases in the generated maps. Point source masking is accomplished by applying a mask to the time stream. The threshold for masking is determined such that any point source brighter than 6 mJy at 150 GHz and any cluster with a signal-to-noise ratio (SNR) across all bands greater than 10 are masked, for a total masked area of $\sim 25\,\mathrm{deg}^2$. The masking radius for point sources, converted from physical distance to the time domain based on the angular speed as described in Eq. (2.5), is set by the radius at which the beam response (see Section 2.2.4) multiplied by the SNR falls below

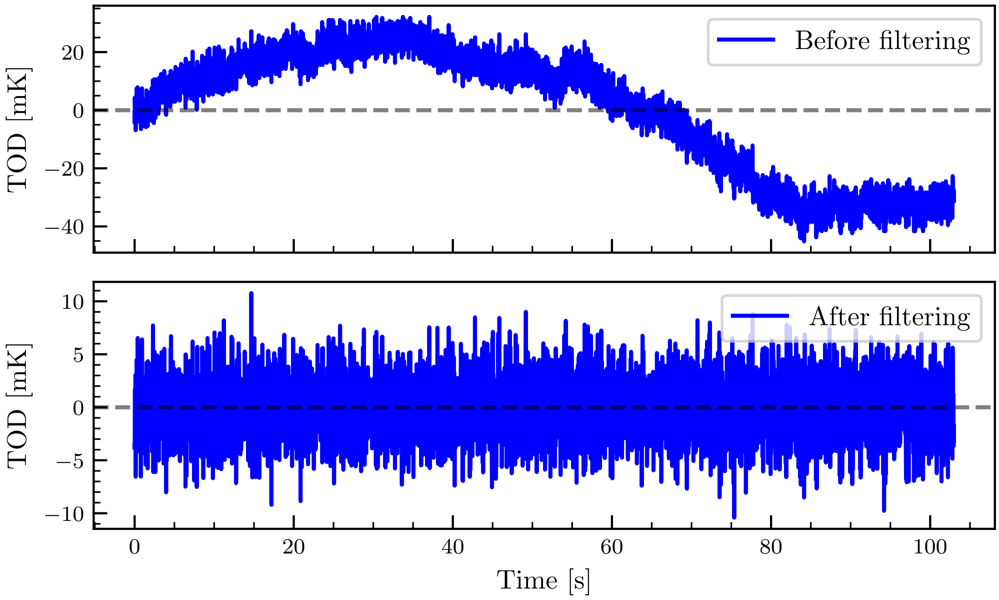


Figure 2.9: Time-ordered data for a single detector. The top panel shows the calibrated unfiltered data, the bottom panel shows the data after filtering. The filtering step removes most of the $1/f$ noise, thus allowing to model the noise as Gaussian.

| **Subfield** | **Number of observations** | **Average number of active detectors** |
|---|---|---|
| ra0hdec-44.75 | 1037 | 11158 |
| ra0hdec-52.25 | 903 | 11164 |
| ra0hdec-59.75 | 772 | 11171 |
| ra0hdec-67.25 | 579 | 11164 |

Table 2.1: Number of observations per subfield in the SPT-3G 19/20 data set and the average number of detectors used for the analysis per subfield. As a reminder, the focal plane is composed of 14266 detectors. An observation is the data acquired during 1 fridge cycle on a single subfield. When no failure, one observation covers the entire subfield.

1. For clusters, it is set by the radius at which the beam convolved with the cluster profile falls below 1. The mask is applied to the time stream by setting the data to zero within the masked region. In Fig. 2.10, I show a histogram of the number of sources masked per radius bin. The beam size (see Section 2.2.4) is indicated for comparison, and as expected, all masked sources are resolved. The flux cut is chosen to optimize lensing reconstruction performance and to decrease the contamination of the signal spectrum. In Fig. 2.11 is displayed the radio galaxy power amplitude ratio as a function of the number of masked sources and associated masking threshold in milli-jansky, with an indication of the flux cuts of SPT-3G 2018 and SPT-3G 19/20. As around 10 times more sources are masked the power ratio is reduced by a factor of 10. Additional sources with different physical origins (clusters, dusty star-forming galaxies, AGNs) are also masked, for a total of 2118 radio sources and 537 clusters.

**Low-pass and high-pass filtering** To mitigate contamination from $1/f$ and atmospheric noise, as well as prevent aliasing of high-frequency signal and noise, the time stream undergoes a sequence of linear processing steps. Filters are employed in the time domain, with the choice made to align with specific spatial multipoles rather than time frequencies. The conversion between the multipole $\ell_x$ and frequency $f$ is given by

$$\ell_x = \frac{2\pi}{\omega_{\rm ra}\cos\theta_{\rm DEC}} f, \tag{2.5}$$

where $\omega_{\rm ra}$ represents the angular speed of the telescope in the right ascension direction relative to the sky. It is computed as the sum of the angular speed of the Earth's rotation and the angular speed of the telescope, i.e., $\omega_{\rm ra} = \omega_{\rm az} + \omega_{\rm er}$ in radian per seconds. Neglecting the latter term would introduce discrepancies in the filtering process between left-going and right-going scans, thereby biasing the resultant maps. By employing a multipole-based filtering approach, the resulting maps possess angular properties that are independent of declination. The

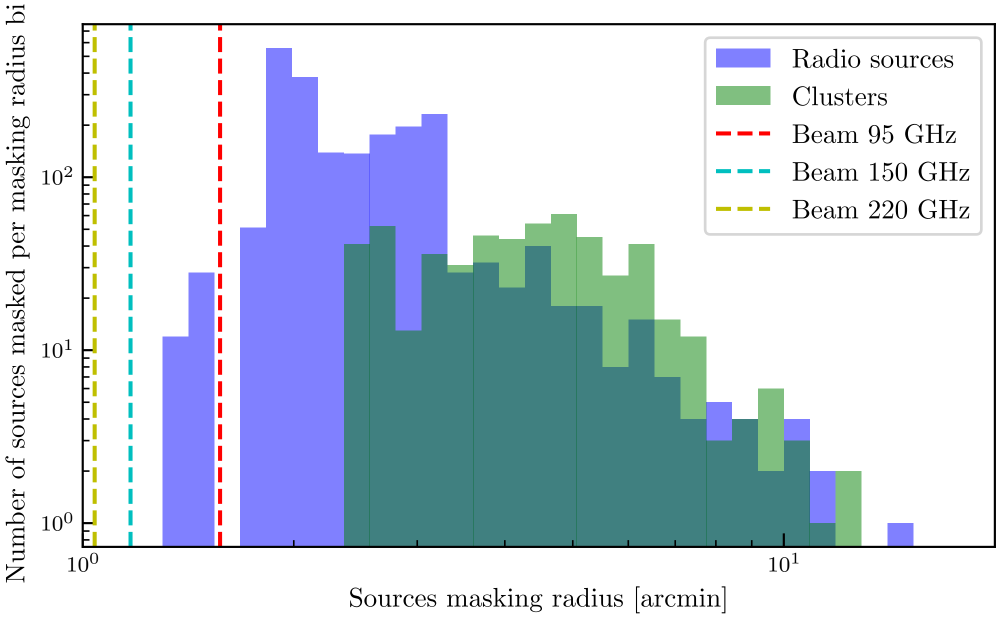


Figure 2.10: Histogram of the number of sources masked per radius bin, for a total of 2118 radio sources and 537 clusters. Beam full widths at half maximum are also indicated. A large number of resolved sources are masked.

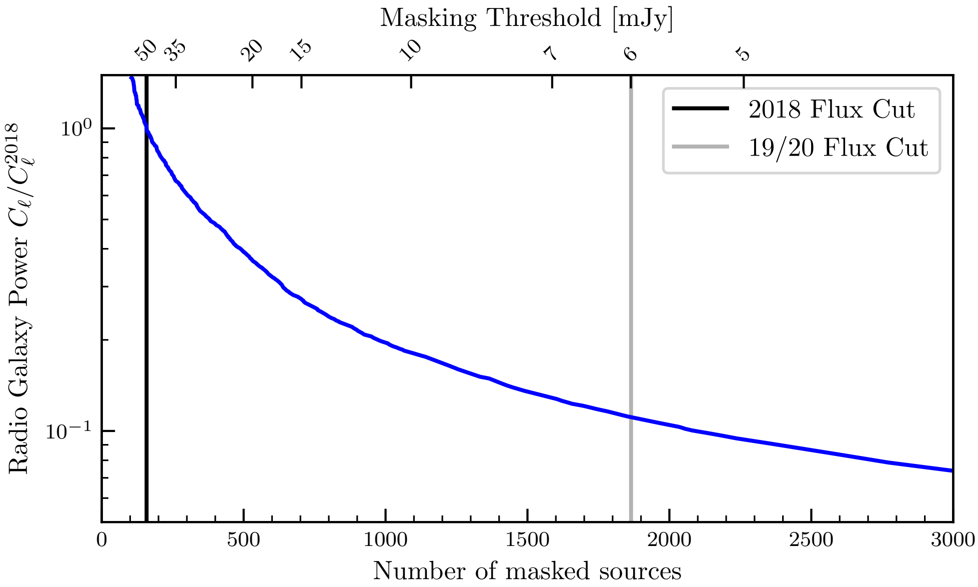


Figure 2.11: Effect of masking radio sources on the power spectrum Poisson radio source contamination. The associated source flux cut is displayed on the top axis. Masking more sources allows for the reduction of the contamination from Poisson power.

time stream undergoes a series of filtering operations. Firstly, a low-pass filter is applied, utilizing exponential damping of the form $\exp\left[-(\ell_x/\ell_{\rm high})^6\right]$, with a cutoff multipole $\ell_{\rm high}$ set to 13000. The choice of this cutoff is influenced by the pixelation scheme employed to represent the maps (see later in Section 2.2.1). Next, the high pass filter removes a combination of Legendre polynomials of order less than 30, and Fourier modes of frequencies such that $\ell_x < 300$ according to Eq. (2.5). The filtering procedure is visualized in Fig. 2.9, where the first panel displays the calibrated time stream and the second panel shows the data after filtering. As a result, the $1/f$ noise is significantly reduced, and the noise exhibits greater consistency with white noise characteristics.

## 2.2 Map-making

After applying the necessary filtering procedures, we obtain calibrated data. The subsequent step involves binning the time streams into pixels, a critical and computationally intensive process that lays the foundation for the subsequent analysis. Ground-based experiments often face challenges such as high noise levels, primarily caused by atmospheric contamination. The map-making process plays a crucial role in mitigating these effects, as it aims to produce maps that are free from systematic errors while keeping computational costs manageable.

### 2.2.1 `HEALPix` pixelation scheme

Prior to delving into the map-making procedure, it is essential to introduce the pixelation scheme utilized for binning the time stream. In the previous analysis of the SPT-3G 2018 data (Dutcher et al., 2021; Balkenhol et al., 2021), the maps were represented on a flat-sky azimuthal equal-area projection, employing a base pixel size of $2'$. This choice of pixelation was advantageous in terms of reducing the computational complexity, thanks to the fast Fourier transform (FFT). However, it also introduced a projection effect on the resulting data products (Balkenhol et al., 2022). For this analysis, we adopt a curved-sky representation and rely on the `HEALPix` (Górski et al., 2005) pixelation scheme[1]. `HEALPix` is a hierarchical pixelation scheme that enables the representation of the curved sky as a collection of pixels with equal area, parametrized by the variable $n_{\rm side}$. The total number of pixels is given by $N_{\rm pix} = 12 \times n_{\rm side}^2$. The `HEALPix` pixelation scheme offers computational advantages with an efficient implementation of the spherical harmonic transform. The Nyquist frequency of the `HEALPix` pixelation scheme in terms of spatial multipole is given by $\ell_{\rm max} = 2n_{\rm side}$. In our analysis, we employ a high-resolution pixelation scheme with $n_{\rm side} = 8192$, corresponding to a pixel size of $0.43'$ and a Nyquist frequency of $\ell_{\rm max} = 16384$. It is important to note that pixelation introduces a convolution effect on the data, resulting in a square window function of known size, the so-called **pixel window function**. This convolution effect, along with the beam response, needs to be accurately modeled in subsequent analysis steps. It is worth mentioning that the `HEALPix` software has also been employed in the analysis of data from the WMAP mission (Bennett et al., 2013) and the *Planck* mission (Planck Collaboration et al., 2020d).

### 2.2.2 Formalism

The map-making process aims to reconstruct the data by combining the calibrated data, pointing information, and noise modeling using a maximum-likelihood approach. The map-making methodology employed for the SPT-3G 19/20 analysis is based on the approach used in the SPTpol analysis (Crites et al., 2015; Keisler et al., 2015; Henning et al., 2018) and is described in detail in Jones et al. (2007). In this analysis, we consider the calibrated time-ordered data as a vector $\mathbf{y}$ of size $n_{\rm tod}$, representing the total intensity and polarization measurements. The pixelated map vector, denoted as $\mathbf{m}$, represents the CMB signal at each pixel and has a size of $n_{\rm pix}$. The pointing matrix, denoted as $P$, is an $n_{\rm tod} \times n_{\rm pix}$ matrix that provides information about the instrument pointing at each time step. Additionally, we model the noise as a random variable vector $\mathbf{n}$ of size $n_{\rm tod}$. While the other quantities have fixed values at each time step, the noise follows a random distribution, and it is necessary to develop a noise model to reconstruct the map vector $\mathbf{m}$ from the time-ordered data vector $\mathbf{y}$. In our analysis, we assume that the noise follows a Gaussian distribution with a time covariance matrix $C_n$. The modeling of the map-making process can be summarized as follows:

$$\mathbf{y} = P\mathbf{m} + \mathbf{n} \quad \Leftrightarrow \quad \mathbf{n} \equiv P\mathbf{m} - \mathbf{y} \sim \mathcal{N}\left(0, C_n\right). \tag{2.6}$$

We can describe a general class of unbiased map estimators $\tilde{\mathbf{m}}$ and the corresponding pixel-to-pixel noise correlation matrix $\mathbf{N}$ using symmetric positive definite matrices $M$. These matrices are defined in the time domain and have dimensions of $n_{\rm tod} \times n_{\rm tod}$. The estimators and noise correlation matrix can be expressed as follows:

$$\tilde{\mathbf{m}} = W^{-1}\tilde{\mathbf{m}}^W, \quad \tilde{\mathbf{m}}^W = P^\top M\mathbf{y}, \qquad \mathbf{N} = W^{-1}P^\top M C_n M P W^{-1} \quad \text{with} \quad W \equiv \left[P^\top M P\right] \tag{2.7}$$

In this process, the time stream is first weighted by the matrix $M$ and then transformed into the pixel domain. Finally, an informed weight matrix $W$ is applied to "unweight" the data in pixels. The minimum variance solution to this problem, which also corresponds to the maximum-likelihood solution, is obtained by setting $M = C_n^{-1}$, and this equation governs the relationship between the matrices

$$\tilde{\mathbf{m}} = \left[P^\top C_n^{-1} P\right]^{-1} P^\top C_n^{-1}\mathbf{y} \quad \text{and} \quad \mathbf{N} = \left[P^\top C_n^{-1} P\right]^{-1} \tag{2.8}$$

Ideally, the inverse-variance weights should be set to the noise inverse variance prior to the pixel domain transformation. This approach allows for optimal reconstruction of the map (Janssen et al., 1993; Tegmark and

---

[1] https://healpix.sourceforge.io/

de Oliveira-Costa, 2001). However, in practice, constructing and inverting the full noise covariance matrix in the time domain can be technically challenging and computationally demanding. For example, SPT-3G scans typically involve a large number of time samples on the order of $\mathcal{O}(10^4)$, making it impractical to handle the full noise covariance matrix without simplifications if this one is dense. To address this issue, the SPT-3G analysis follows the recommendation of Hivon et al. (2002) to apply a filtering step to the time stream. This filtering helps reduce the $1/f$ noise and approximate the effective noise as Gaussian white noise, as described in detail in Section 2.1.4 and illustrated in Fig. 2.9. By mitigating the impact of $1/f$ noise, the effective noise properties are improved, enabling more tractable and accurate map reconstruction as the noise covariance matrix becomes diagonal.

**Comparison with ACT and *Planck*** Different experimental setups require different map-making pipelines. The *Planck* satellite, with its full-sky circular scanning strategy, encounters noise that is structured along scans meeting at the poles. To address this, the experiment utilizes a method called *destriping*, which is further explained in Planck Collaboration et al. (2020d). It relies on jointly solving the detector gains, the foreground component, and the stripping template predicted from TOD information. Such an implementation would be impossible for SPT-3G since the scans are not crossing. On the other hand, the ACT data presents its challenges, including varying seasons, a large fraction of the sky to cover, and cross-linked scans. Despite these challenges, the experiment employs an iterative optimal maximum-likelihood map-making pipeline without prior TOD filtering. The detector-detector noise structure is caught by modeling of the noise correlations ($\mathbf{N}$ in Eq. (2.8)) as described in Dünner et al. (2013); Naess et al. (2014); Louis et al. (2017); Aiola et al. (2020). By avoiding TOD filtering, the resulting CMB maps are unbiased, and the analysis does not require simulations to compute the transfer function. However, the map noise properties are not optimal and need to be modeled using refined techniques, as explained in Atkins et al. (2023). Simulations are used in order to obtain the noise power spectrum covariance matrix. While the ACT and *Planck* methods provide more optimal results, they require significantly more computing time compared to the SPT-3G method. Therefore, the SPT-3G method strikes a good balance between computational cost and optimality, and the obtained noise properties are easy to model. The main drawback is the need for simulations to compute the transfer function, as detailed in Section 2.3.4 and Section 3.3.

### 2.2.3 Application to SPT-3G

In this section, we consider the maps $\mathbf{m}$ to be $3 \times n_{\rm pix}$, expanding the vectors to include the extra polarization dimension $T \to (T, Q, U)$. The pointing matrix is then given by $n_{\rm tod} \times n_{\rm dect} \times n_{\rm pix}$, where $n_{\rm dect}$ denotes the number of detectors. Let's focus on the contribution of a single detector $i$. The elements of the pointing matrix are denoted as $P^i_{t\alpha} = 1$ when the given detector observes pixel $\alpha$ at time $t$. To account for the mixing of temperature and polarization in the detectors, a projection matrix $A^i$ is introduced. This matrix incorporates the cosine and sine terms defined in Eq. (2.4), which can be expressed using the following notation

$$A^i = \begin{pmatrix}1 & \gamma_i c_i & \gamma_i s_i\end{pmatrix} \quad \forall \alpha \in [1, n_{\rm pix}]. \tag{2.9}$$

By applying this projection matrix, we can account for the mixing of temperature and polarization signals within each detector. The raster scanning strategy employed by SPT-3G ensures that the projection matrix remains time-independent. To account for the mixing of temperature and polarization signals in the detectors, we rewrite the map-making problem by replacing $P$ with $AP$. The noise after filtering is approximately Gaussian white noise, and a reasonable approach is to parametrize the map-making process using a diagonal matrix $M$, where the weights are constructed from the variance of the noise of the associated detector's time stream in the multipole range $\ell \in [300, 4000]$ (refer to Eq. (2.5) for the conversion between multipole order and frequency). This multipole range is chosen as it is most sensitive to the underlying cosmological signal. Let us denote the averaged noise value as $n_i$ for the corresponding indices where $P^i_{t\alpha} = 1$. As mentioned earlier, although this choice of $M$ may not yield the optimal reconstruction, it results in unbiased non-optimal maps. Those maps are expected to be close to optimal due to the filtering applied to the time-ordered data, which makes the noise properties approach Gaussian characteristics. We can express the map-making solution for these indices as follows:

$$M^i = \frac{1}{n_i^2}\mathrm{Id}_{n_{\rm tod}\times n_{\rm tod}}, \quad \left[[AP]^{i\top} M^i\right]_\alpha = \frac{1}{n_i^2}\begin{pmatrix}1\\ \gamma_i c_i\\ \gamma_i s_i\end{pmatrix}\mathrm{Id}_{n_{\rm tod}\times n_{\rm tod}} \tag{2.10}$$

$$\text{and} \quad W^i_\alpha \equiv \left[[AP]^{i\top} M^i [AP]^i\right]_{\alpha\alpha} = \frac{1}{n_i^2}\begin{pmatrix}1 & \gamma_i c_i & \gamma_i s_i\\ \gamma_i c_i & \gamma_i^2 c_i^2 & \gamma_i^2 c_i s_i\\ \gamma_i s_i & \gamma_i^2 c_i s_i & \gamma_i^2 s_i^2\end{pmatrix}. \tag{2.11}$$

To gain physical intuition, let's examine the weight matrix in two extreme cases. In both cases, we assume perfect detectors, meaning they have the same weight, the polarization efficiency is 1, and the angles are perfectly known. In the first case (1), we consider polarization angles that are randomized across the focal plane. In this scenario, we compute the average weight matrix as if the angle $\psi$ is equal to $\pi/4$. In the second case (2), we assume that all detectors are aligned, resulting in $\psi = 0$. Let's denote the average weight matrix for these cases as follows:

$$(1): \quad \bar{W}_\alpha = \frac{1}{n^2}\begin{pmatrix}1 & 0 & 0\\ 0 & \frac{1}{2} & 0\\ 0 & 0 & \frac{1}{2}\end{pmatrix} \quad \text{or} \quad (2) \quad \bar{W}_\alpha = \frac{1}{n^2}\begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{pmatrix}. \tag{2.12}$$

In the second case, we encounter a situation where we are unable to reconstruct one of the polarization parameters due to an ill-defined weight matrix. This observation helps explain why the detectors are intentionally not aligned on the SPT-3G focal plane. Aligning the detectors would result in a loss of information and render polarization reconstruction impossible.

The final step in the map-making process is to average the signal over all detectors, indexed by $i$. This yields the following equations for the weighted maps at pixel $\alpha$, considering the associated weight matrix:

$$\left(T^W, Q^W, U^W\right)_\alpha = \sum_i \left[P_i^\top A^{i\top} M^i \mathbf{y}^i\right]_\alpha \quad \text{and} \quad W_\alpha^i = \sum_i P_\alpha^{i\top} A^{i\top} M^i A^i P_\alpha^i \quad \text{of size } 3\times 3. \tag{2.13}$$

The weight matrix has a physical interpretation as the inverse noise variance of the pixels in the map. They are the sum of the inverse variance of the detectors pointing at the pixel $\alpha$, see Eq. (2.13), and are effectively similar to hit maps with varying weights for each detector. Detectors with weights deviating more than $3\sigma$ from the median are excluded, in order to remove spurious contaminations. The unweighted $(T, Q, U)$ and weighted maps $(T^W, Q^W, U^W)$ are related by

$$(T, Q, U)_\alpha^\top = W_\alpha^{-1}\left(T^W, Q^W, U^W\right)_\alpha^\top. \tag{2.14}$$

In our analysis, we generate a map by co-adding the entire observation set. Additionally, we utilize different data splits to create a set of $n_{\text{BUN}} = 150$ maps, which we refer to as **bundles** throughout this work, following the approach used in Dutcher et al. (2021) and Balkenhol et al. (2022). It is essential to ensure that each bundle has similar weights, achieves uniform coverage, and is constructed using a sufficient number of observations to ensure noise uniformity. These bundles serve two purposes: computing the data band powers, as described in Chapter 3 and estimating the noise covariance matrix, as described in Chapter 4.

### 2.2.4 Beam and time constants

In the previous problem set, the pointing matrix $P$, which indicates the pointing of detectors at a given time $t$, was obtained based on the telescope's scanning strategy. However, detectors have a finite extent and exhibit a frequency-dependent angular response to point-like sources, known as the beam. The observed data is a convolution of the sky signal with the average beam window function that depends on the telescope's optical design. This convolution results from the finite extent of the antenna and the non-ideal nature of optics and detectors. In general, larger primary mirrors yield smaller beams. Angular scales below the beam size are averaged out and cannot be accurately reconstructed. As a consequence of this convolution, the pointing matrix is no longer sparse but becomes a dense matrix. In Eq. (2.13), we could directly include the beam in the pointing matrix $P$ and deconvolve the maps. However, this approach may encounter computational scaling issues since the weight matrix would no longer be diagonal ($W_{\alpha\alpha'} \neq W_\alpha \delta_{\alpha\alpha'}$) and could potentially be ill-conditioned. Instead, we adopt the standard approach of deconvolving the average beam from the maps after the map-making process, at the spectrum level. Consequently, the reconstructed maps are biased maps that effectively account for the convolution with the beam. We replace the sky signal $T$ with $T'(\hat{\mathbf{n}}) = \int \mathrm{d}\hat{\mathbf{u}} B(\hat{\mathbf{n}} - \hat{\mathbf{u}}) T(\hat{\mathbf{u}})$, where $B$ represents the beam response function in real space. A complete description of beam analysis can be found in Hivon et al. (2017) and references therein.

Detectors not only exhibit a finite spatial response but also a finite temporal response. The time constant of a bolometer refers to the time it takes to reach 1/e of its final response to a step function and determines the maximum scanning speed of the telescope. If we were to include time constants in the pointing matrix, similar considerations as those for the beam would arise. Therefore, we choose to deconvolve the time constants at the spectrum level, accounting for their effects within the effective beam characterized by the data.

Accurate characterization of the beam is a critical step in the data analysis process since beam mismatch can introduce systematic errors in the resulting data products. For the SPT-3G, the beam is estimated using observations of bright point sources and dedicated raster scans of Saturn. Bright point sources are utilized to estimate the beam near its peak and to integrate the time constants, while the high signal-to-noise ratio of the Saturn scans enables reliable measurement of the beam response at angular distances of tens of arcminutes. The resulting beam maps, obtained for each frequency, are displayed in Fig. 2.12. In our analysis, we employ a deconvolution technique to remove the effects of the beam from the data at the spectrum level, as outlined in Section 3.1. We assume that the beam exhibits axial symmetry, as suggested

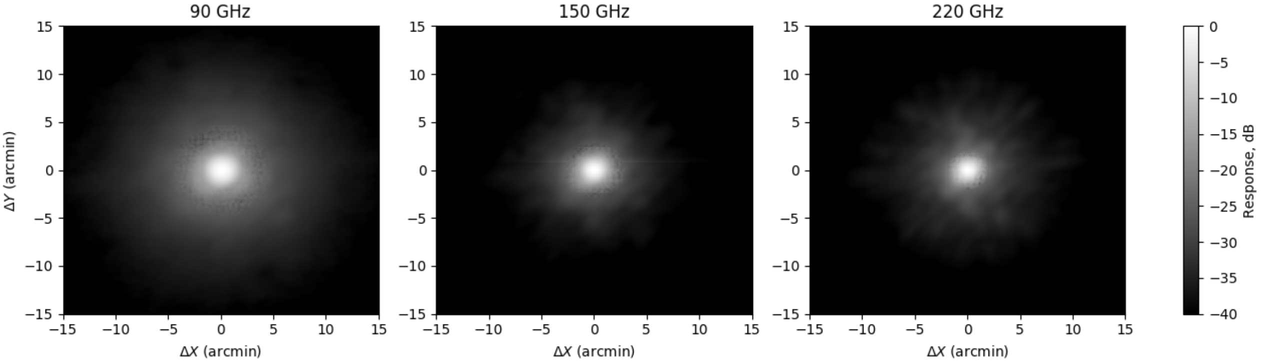


Figure 2.12: Beam maps of the SPT-3G experiment at 95, 150, and 220GHz. Some visible features in the maps are due to the stitching of radio point source maps and planet maps together. Courtesy of Neil Goeckner-Wald. The characterization of the beam is crucial to model the data.

in Wu et al. (2001), and average its response over circular regions with equal angular distances. The effective beam response is described as a function of the angle $\theta$ between the pointing direction and the response angle. To characterize the beam, we utilize its full width at half maximum (FWHM), denoted as $\theta_{\mathrm{FWHM}}$. In harmonic space, an approximation of the beam using a Gaussian profile with the same FWHM yields the well-known expression $B_\ell \propto \exp\left[-\ell(\ell+1)\theta^2_{\mathrm{FWHM}}/16\ln 2\right]$ (Challinor et al., 2000). The radial profile of the beam, along with its harmonic transform, is illustrated in Fig. 2.13, while associated FWHM is given in Table 2.2. SPT-3G has a narrower beam than the *Planck* satellite thanks to the large first mirror, thus allowing to probe high-$\ell$ modes. The ACT beam is larger than the SPT-3G beam and significantly varies across the season as the telescope is subjected to day/night and temperature variations affecting the optics. Although SPT-3G has higher resolution, it must be emphasized that the ACT survey is much wider, covering around 40% the sky, and thus has more constraining power at intermediate scales.

| Frequency | SPT-3G FWHM [′] | *Planck* FWHM [′] | ACT FWHM [′] |
|---|---|---|---|
| 95GHz | 1.57 | 9.7 | 2.0 |
| 150GHz | 1.17 | 7.2 | 1.3 to 1.46 |
| 220GHz | 1.04 | 5.0 | NA |

Table 2.2: Full width at half maximum (FWHM) of the beam in arcminutes for the SPT-3G, *Planck*, and ACT experiments, obtained from S22, Planck Collaboration et al. (2020d) and Lungu et al. (2022). Beams are given for the frequency closest to the central SPT-3G one. For ACT 150GHz, the FWHM varies across the multiple arrays and the season.

Uncertainties associated with the beam profile are quantified by comparing the computed beam responses for different subfields. These uncertainties are subsequently incorporated into the final data products through corrections applied to the covariance matrix, see Section 4.2.2 for details on the correction process. Given that detectors possess varying properties, they are subject to different beam shapes, which can result in temperature-to-polarization leakage. To address this issue, a temperature template is fitted and subtracted from the polarized maps. Furthermore, detectors may exhibit differential ellipticity, which can introduce additional spurious effects requiring correction. Further investigation and analysis of these effects are required.

### 2.2.5 Source of errors in map-making

In this section, we aim to examine potential errors that may arise during the map-making process, to anticipate systematic errors that could impact the final maps. More detail about beam systematic effects can be found in Rosset et al. (2007); Hivon et al. (2017).

**Pointing errors** The pointing matrix $P$, which is derived from the telescope scanning strategy, is assumed to be perfect. However, in practical situations, the telescope's stability is not ideal, resulting in pointing errors. These errors constitute a significant source of systematic errors. While time-independent absolute pointing errors can be corrected by observing point-like sources, time-dependent relative pointing errors introduce additional stochastic smearing, which is incorporated into the beam definition for scientific analysis, see S22.

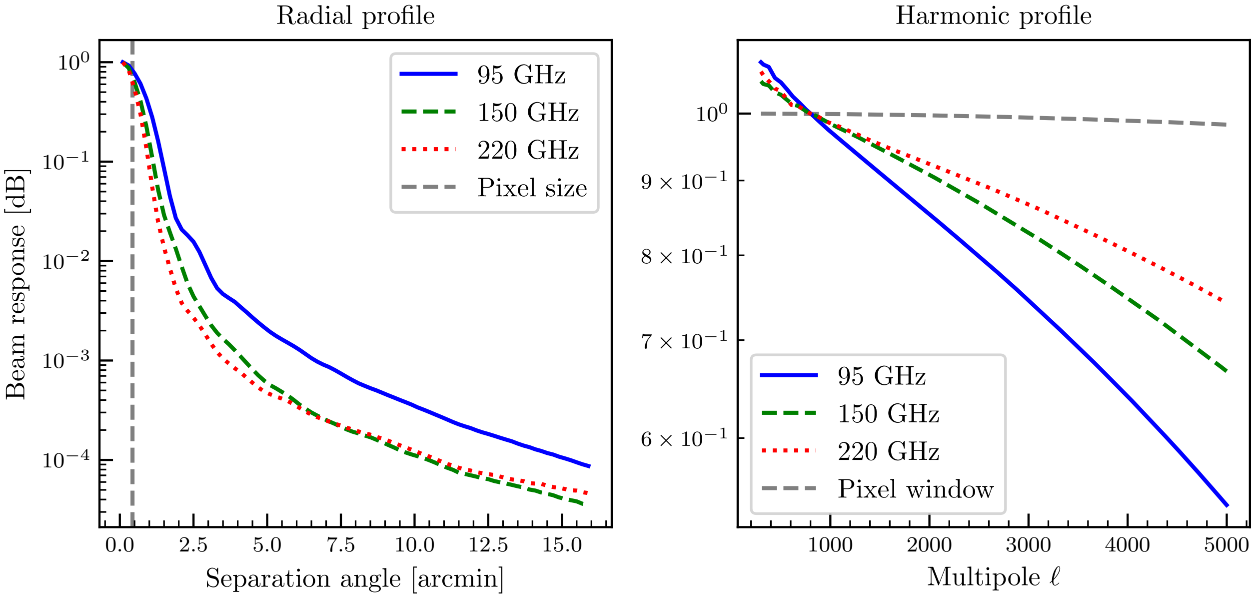


Figure 2.13: Beam response of the SPT-3G experiment at 95, 150, and 220GHz in real space (left) and harmonic domain (right). For comparison, the typical pixel size is $0.43'$ and plotted as a grey vertical line on the left plot. The harmonic equivalent, the pixel window function, is shown on the right plot.

**Polarization efficiency** In the previous map-making equations, we assumed that the polarization efficiency $\gamma_i$ is known with perfect accuracy. However, let us consider the scenario where the true polarization efficiency deviates slightly from the assigned value, given by $\gamma_i \rightarrow \gamma_i(1-\delta_i)$ while keeping the polarization angles fixed. In this case, the effective sky observed by detector $i$ can be approximated as

$$(T, Q, U)^\top \rightarrow (T, (1+\delta_i)Q, (1+\delta_i)U)^\top . \tag{2.15}$$

From this equation, we observe that stochastic errors ($\delta_i$ different for each detector) in the polarization efficiency result in additional noise in the polarization maps. On the other hand, systematic errors lead to an overall rescaling of the polarization maps, which can be corrected through the calibration procedure.

**Polarization angles** Detectors are positioned in the telescope's focal plane with specific orientations. However, the assigned polarization angles of the detectors may slightly deviate from their nominal values. In this section, we model this shift as $\psi_i \rightarrow \psi_i - \delta_i/2$, while keeping the detector efficiency fixed. Consequently, the effective sky observed by detector $i$ can be approximated as

$$(T, Q, U)^\top \rightarrow \left(T, Q + \delta_i U c_i^2, U - \delta_i Q s_i^2\right)^\top . \tag{2.16}$$

Errors in the polarization angles result in a mixing of the $Q$ and $U$ maps, thus creating E to B modes leakage. However, these errors are expected to be small since they are averaged out over the focal plane. We can verify this assumption by calculating polarized cross-spectra, as described in Section 3.1.2.

## 2.3 SPT-3G 19/20 data products

### 2.3.1 CMB maps

Following the map-making procedure outlined in Section 2.2, we have generated the SPT-3G 19/20 coadd of the CMB temperature and polarization anisotropies. The resulting temperature coadds are presented in Lambert azimuthal equal-area projection with a uniform color scheme in Figs. 2.14 to 2.16. Similarly, the polarization coadds are depicted in Figs. 2.17 to 2.22, employing an adapted uniform color scheme. Within the defined patch, the temperature maps exhibit a dominant signal, while the borders appear noisier due to fewer pixels pointing in those regions. Subsequently, these border regions are discarded during the map analysis process, as detailed in Section 2.3.2. The application of TOD filtering along the scan removes the iso-latitude co-scan large-scale modes and leaves the iso-longitude cross-scan features in the maps. In the subsequent analysis, these features will be characterized and described using transfer functions which are used to model the loss of power. TOD values around bright point sources are ignored during the filtering process, but the associated signal is projected in the maps. It must be removed prior to any further analysis as the pixel values will have different statistical properties than the rest of the map. This is accomplished through a masking and inpainting procedure, as detailed in Section 3.6. Notably, a zoom-in of a $2' \times 2'$ region centered on the brightest polarized source in the map is depicted in Fig. 2.23. Around the source location, we see the signal and noise levels of the maps. The latter exhibits an expected increase from 150 to 95 to 220GHz, in line with the properties of the detectors. The polarized noise exhibits a similar trend. Notably, the noise at 220GHz completely obscures the underlying signal.

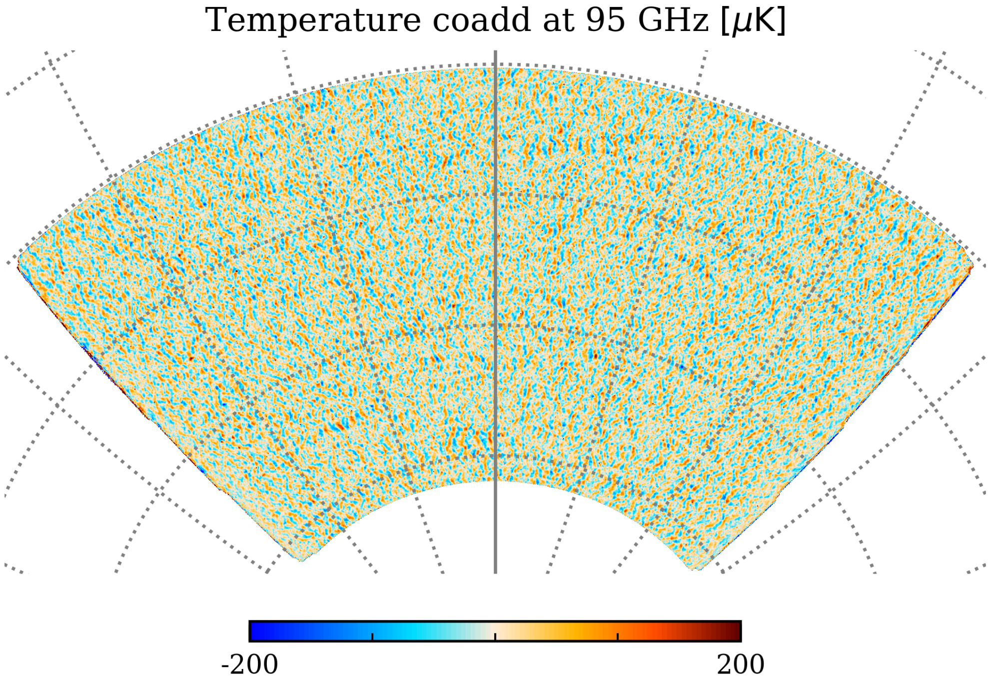

Temperature coadd at 95 GHz [$\mu$K]

Figure 2.14: temperature coadded maps in $\mu$K at 95GHz. The map is displayed in Lambert's azimuthal equal-area projection. Meridian graticules are plotted every 20 deg while parallel graticules are plotted every 10 deg.

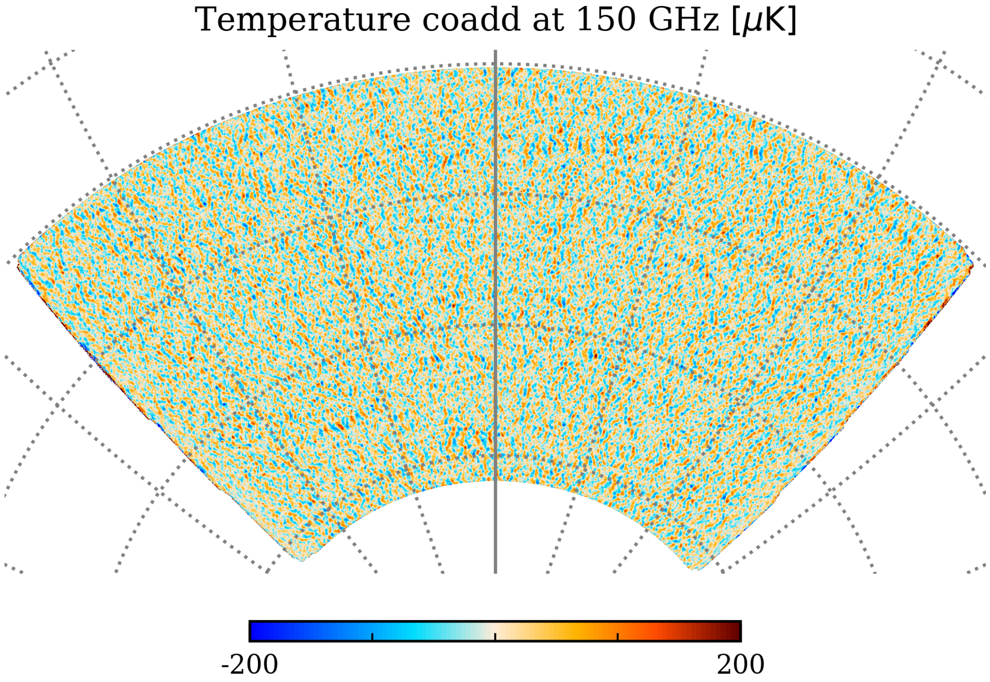

Temperature coadd at 150 GHz [$\mu$K]

Figure 2.15: temperature coadded maps in $\mu$K at 150GHz. The map is displayed in Lambert's azimuthal equal-area projection. Meridian graticules are plotted every 20 deg while parallel graticules are plotted every 10 deg.

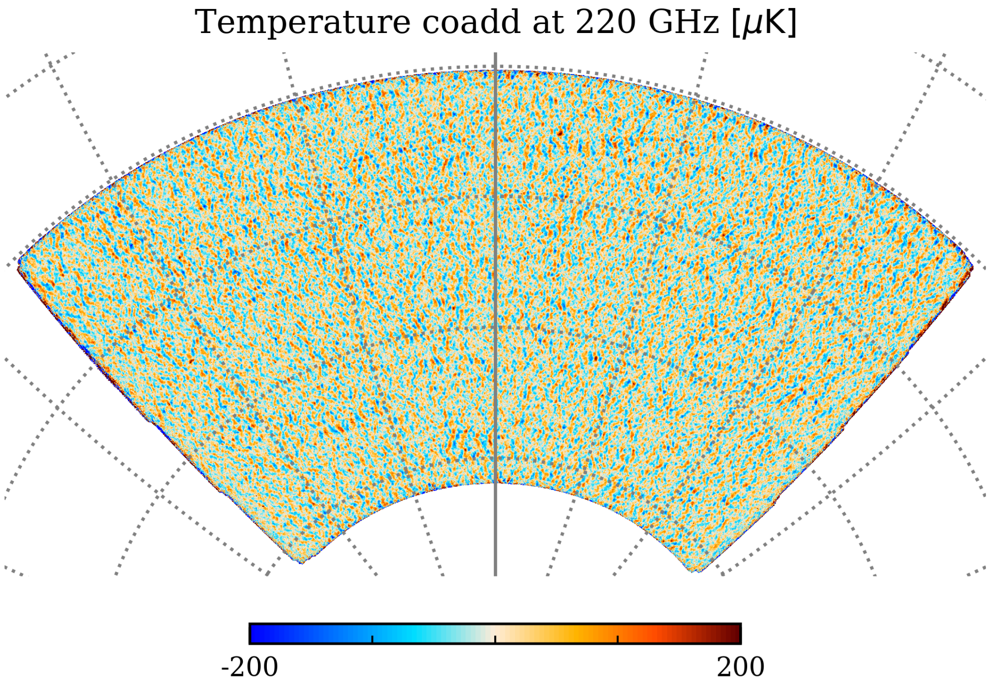


Figure 2.16: temperature coadded maps in $\mu$K at 220GHz. The map is displayed in Lambert's azimuthal equal-area projection. Meridian graticules are plotted every 20 deg while parallel graticules are plotted every 10 deg.

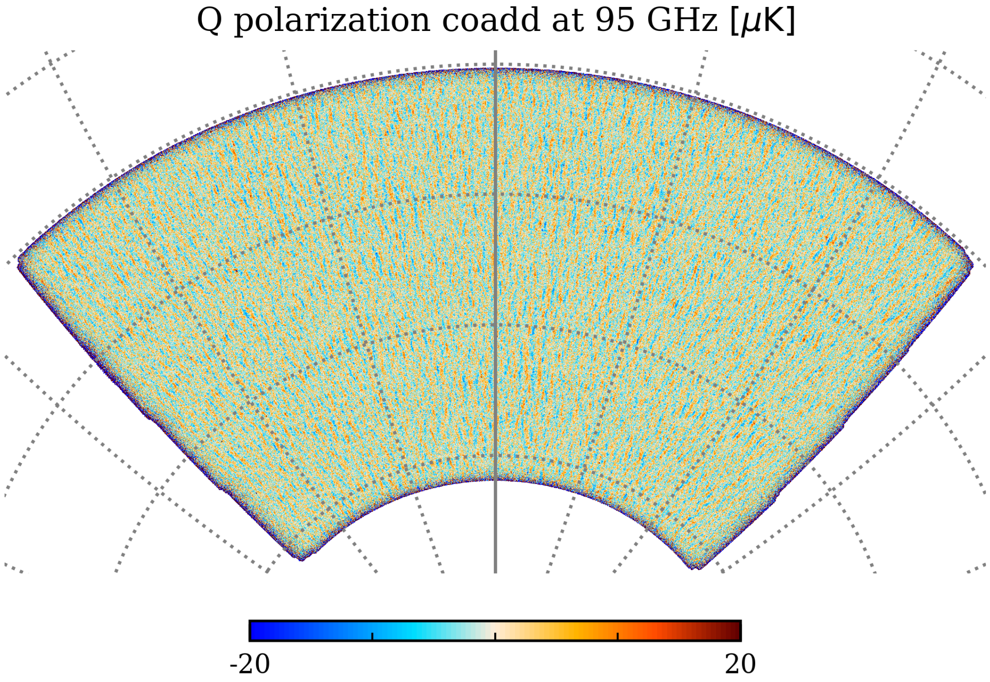


Figure 2.17: Q polarization coadded maps in $\mu$K at 95GHz. The map is displayed in Lambert's azimuthal equal-area projection. Meridian graticules are plotted every 20 deg while parallel graticules are plotted every 10 deg.

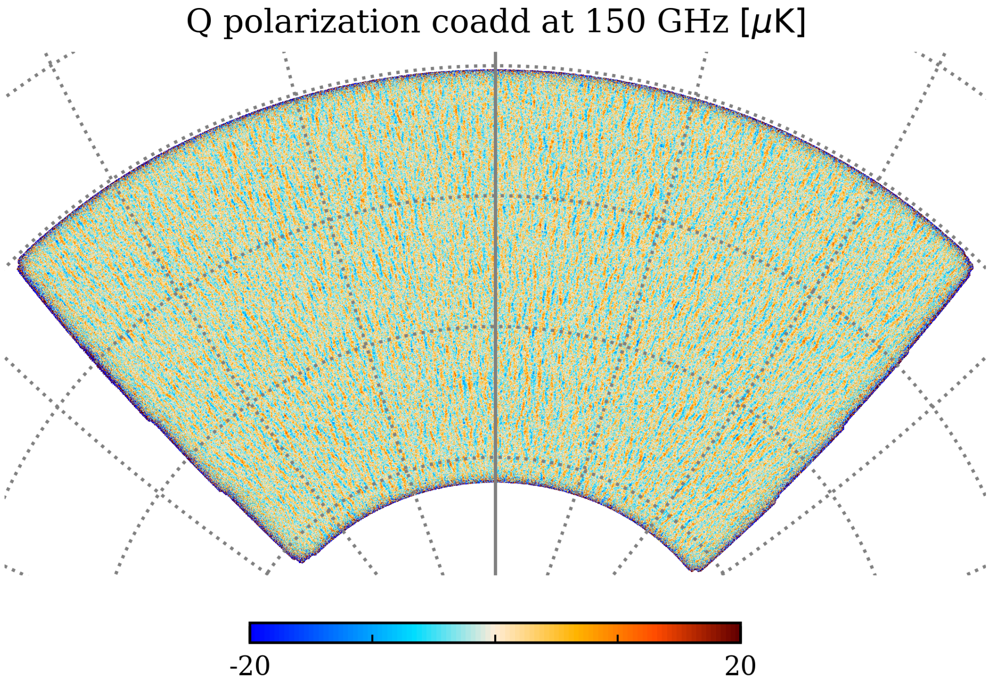


Figure 2.18: Q polarization coadded maps in $\mu$K at 150GHz. The map is displayed in Lambert's azimuthal equal-area projection. Meridian graticules are plotted every 20 deg while parallel graticules are plotted every 10 deg.

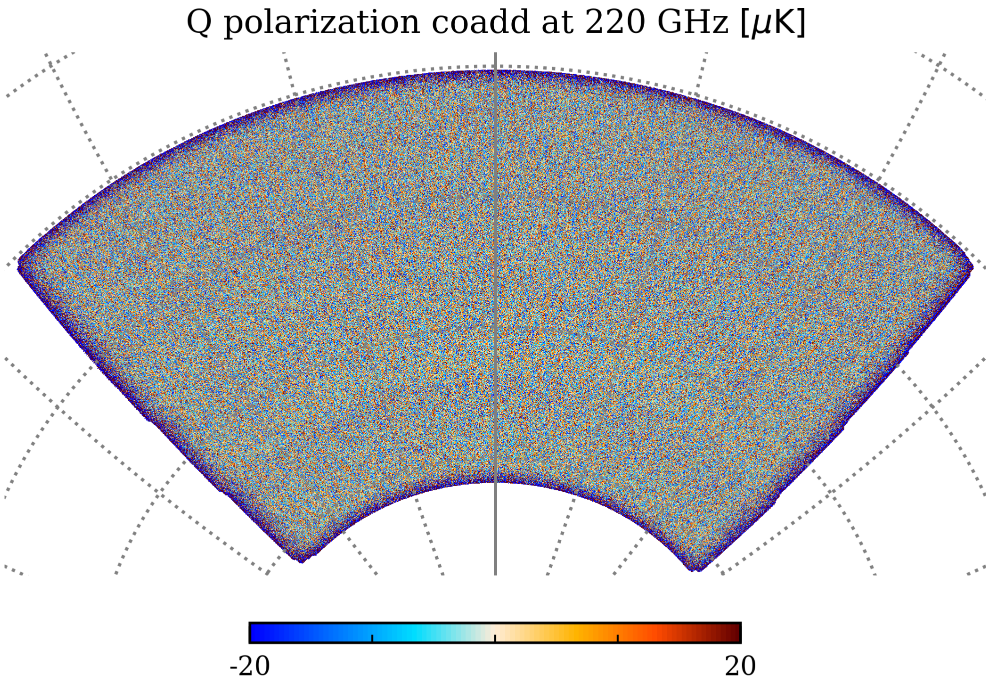


Figure 2.19: Q polarization coadded maps in $\mu$K at 220GHz. The map is displayed in Lambert's azimuthal equal-area projection. Meridian graticules are plotted every 20 deg while parallel graticules are plotted every 10 deg.

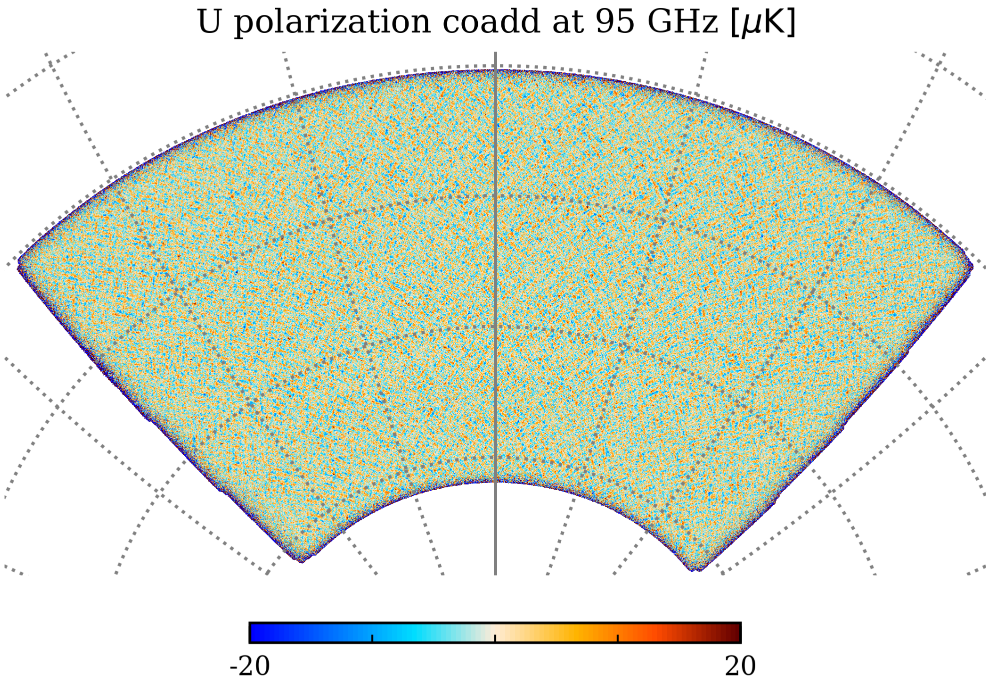


Figure 2.20: U polarization coadded maps in $\mu$K at 95GHz. The map is displayed in Lambert's azimuthal equal-area projection. Meridian graticules are plotted every 20 deg while parallel graticules are plotted every 10 deg.

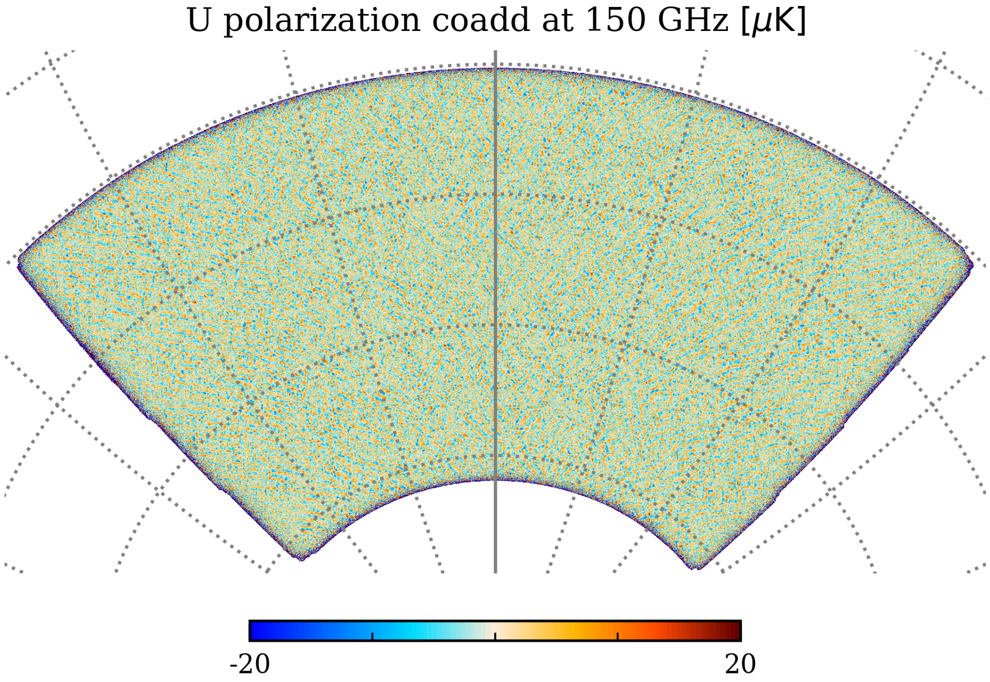


Figure 2.21: U polarization coadded maps in $\mu$K at 150GHz. The map is displayed in Lambert's azimuthal equal-area projection. Meridian graticules are plotted every 20 deg while parallel graticules are plotted every 10 deg.

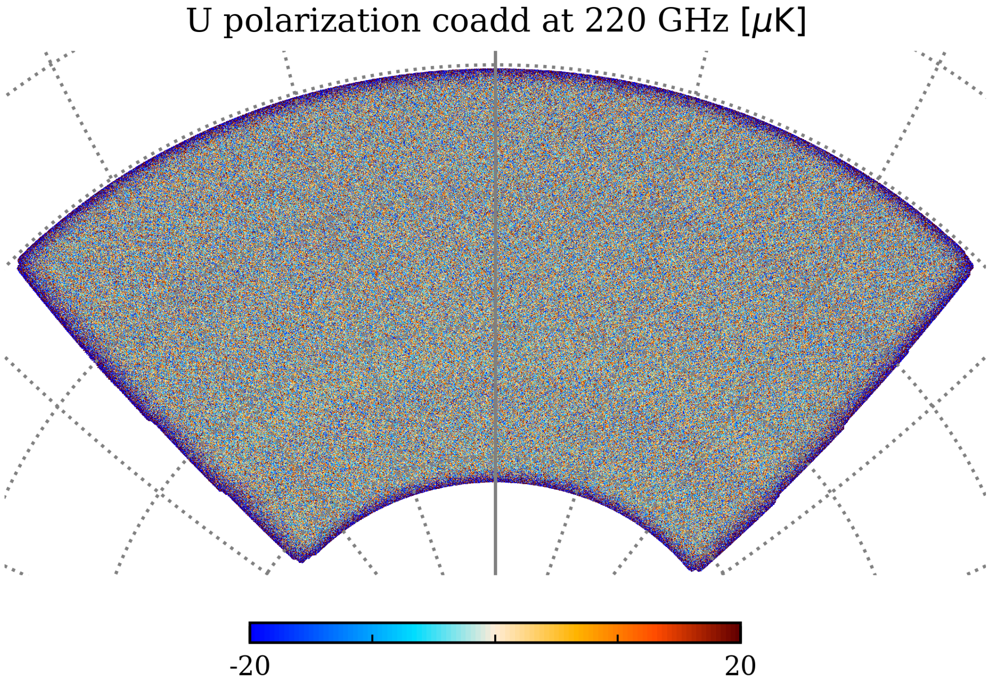


Figure 2.22: U polarization coadded maps in $\mu$K at 220GHz. The map is displayed in Lambert's azimuthal equal-area projection. Meridian graticules are plotted every 20 deg while parallel graticules are plotted every 10 deg.

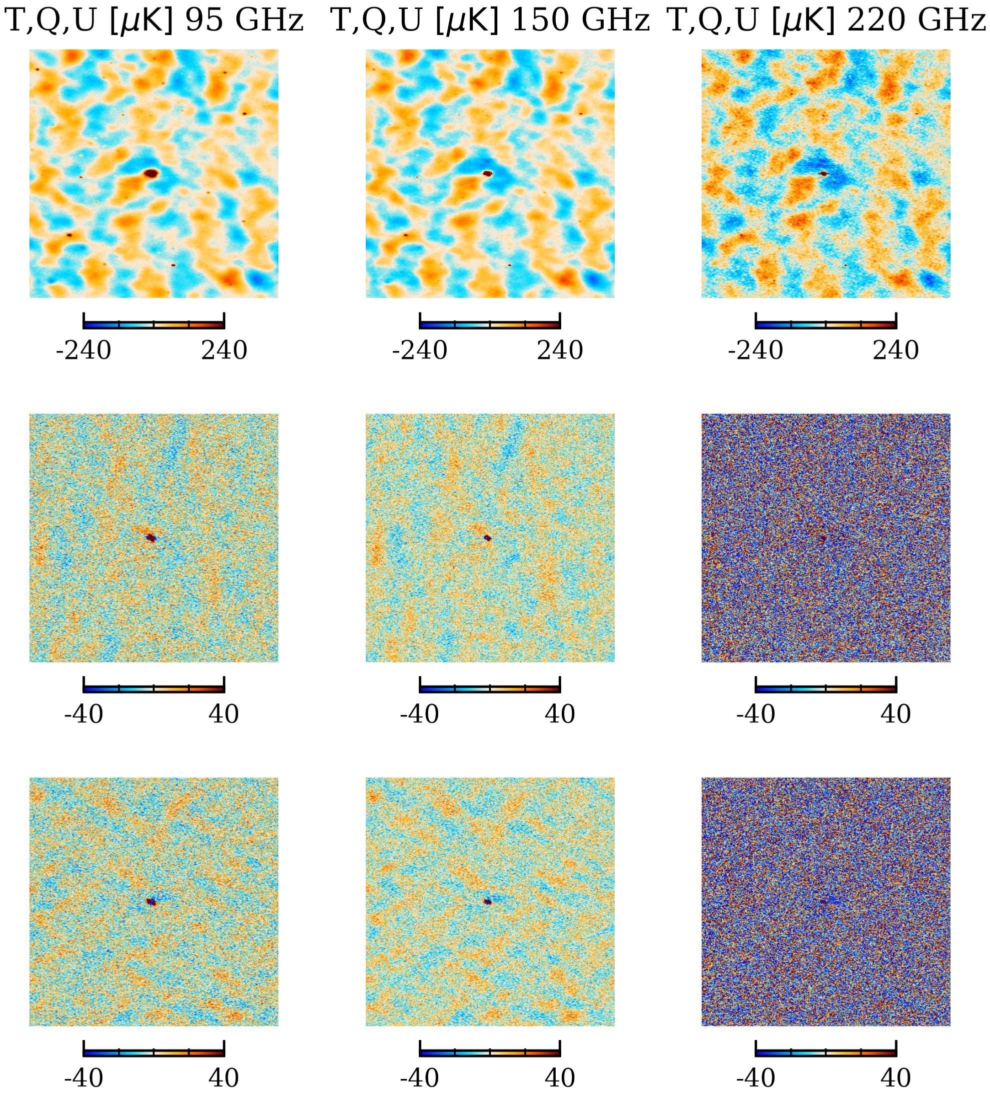


Figure 2.23: Zoom on the brightest source in the maps, located at $\mathrm{RA, DEC} = 32\,\mathrm{deg}\,39', -51\,\mathrm{deg}\,0'$. The patch is $2' \times 2'$. The source is visible in the temperature maps, and it is also visible in the polarization maps, but with lower contrast. Other less bright sources can be observed in the temperature patch. The polarized signal is clear at 95 and 150GHz, but it is almost fully hidden by noise at 220GHz.

### 2.3.2 Weights and mask

The weight map defined in Eq. (2.13) is displayed for temperature 150GHz in Fig. 2.24. It demonstrates uniform noise distribution across the patch, which is a crucial requirement for subsequent assumptions. We utilize this weight map to define a binary mask that restricts the analysis to a specific patch, ensuring maximal effective sky area while avoiding regions with non-uniform noise. In Fig. 2.25, we compare the standard deviation of temperature 95GHz coadd pixels at right-ascension (RA) to the average weight at those RA values. The former serves as a proxy for noise uniformity. We choose a weight threshold of 10% of the median weight, $W_{\rm thr} = 0.1 W_{\rm med}$, to delineate the patch for our subsequent analysis, removing $200\,\mathrm{deg}^2$ which represents 0.16% of the full sky area, and 5% of the area probed by the full data set. As depicted in Fig. 2.25, this threshold ensures uniform pixel variance and associated noise levels within the analysis patch. To optimize lensing reconstruction and reduce the contribution of map edges, I smooth the mask in a process known as **apodization**. This helps mitigate Fourier ringing when transitioning to harmonic space. This is done by first computing the distance map, which is the distance to the closest masked pixel, and smoothing it with a narrow kernel of size $5'$ to remove small-scale features. We then apply a Gaussian profile apodization with a width of 1 deg to the distance map. The same procedure is applied for holes due to point sources but with a reduced width of size $10'$. Obtained masks are displayed in Figs. 2.26 and 2.27. We evaluated the impact of the apodization procedure on the coupling between modes of the power spectrum measured through the mask. The smoothing of the distance map is crucial to removing small-scale features in the mask and reducing long-range coupling. We compared the Gaussian apodization profile to a cosine profile and Gauss error function profile, but the former led to a minimal amount of coupling between the modes.

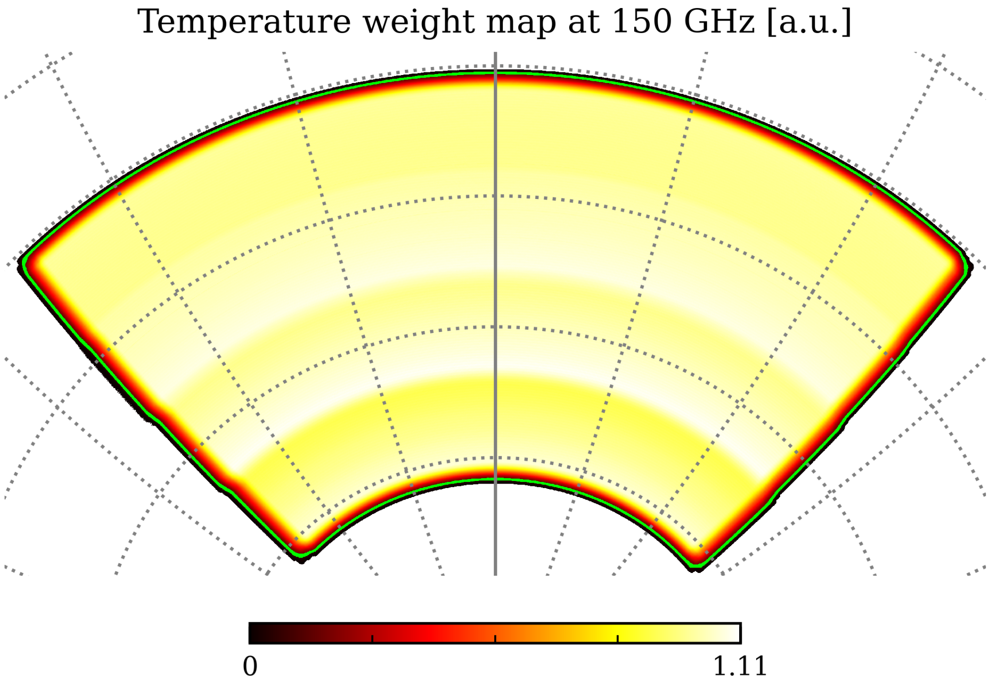


Figure 2.24: Weight map for temperature 150GHz in arbitrary units. Weights indicate the inverse variance of the noise at a given pixel and are similar to a hit count map. The lime boundary indicates the extent of the analysis patch, chosen to go until weights reach 10% of the median weight across the map. The map is displayed in Lambert's azimuthal equal-area projection. Meridian graticules are plotted every 20 deg while parallel graticules are plotted every 10 deg.

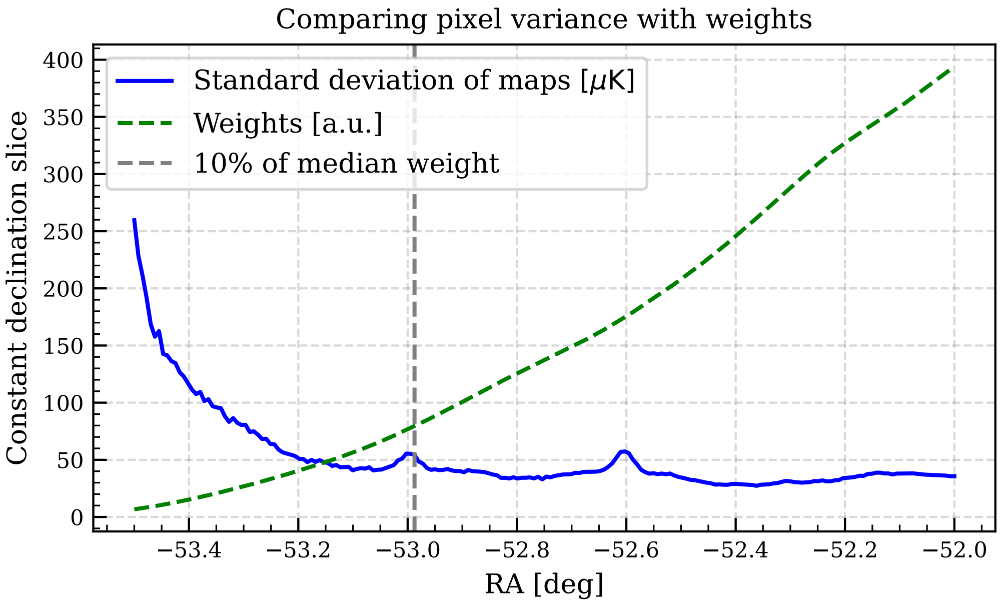


Figure 2.25: Standard deviation of the map at constant RA (blue line) compared to the average weight at constant RA (green line) for temperature 95GHz. The vertical line indicates the 10% median weight threshold. The lines are obtained by averaging over the DEC coordinate in the el1 subfield, *i.e.* between −49 deg and −55.5 deg. The peaks in the standard deviation are due to the presence of bright point sources in the map. In the region delimited by the threshold, the noise is uniform.

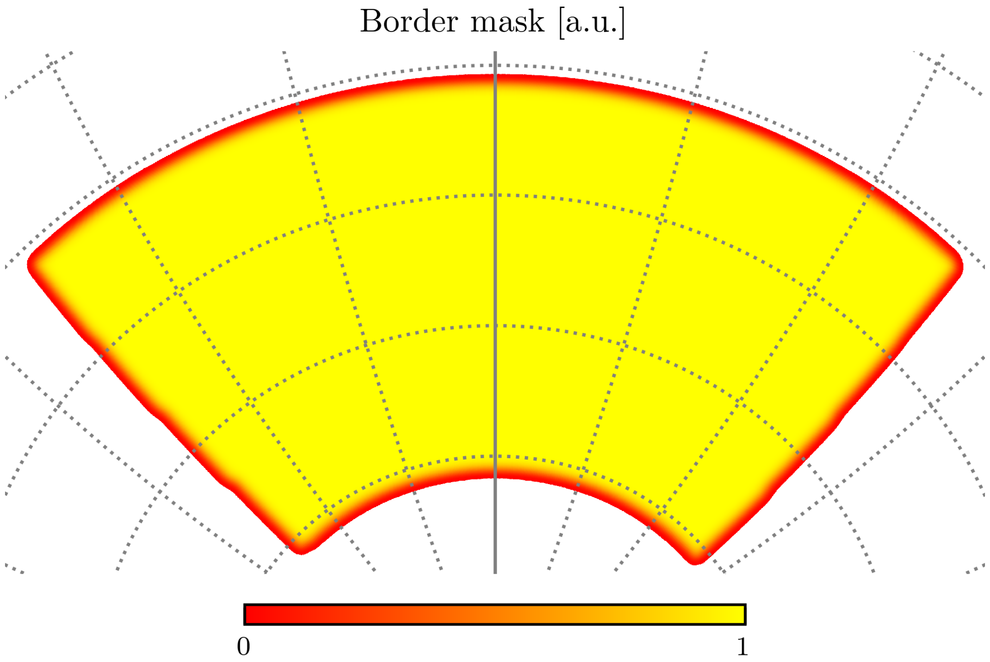


Figure 2.26: Border mask of the SPT-3G 19/20 analysis. Borders are apodized with a 1 deg Gaussian profile. The mask is displayed in Lambert's azimuthal equal-area projection. Meridian graticules are plotted every 20 deg while parallel graticules are plotted every 10 deg. Apodizing the mask helps reduce Fourier ringing in harmonic space.

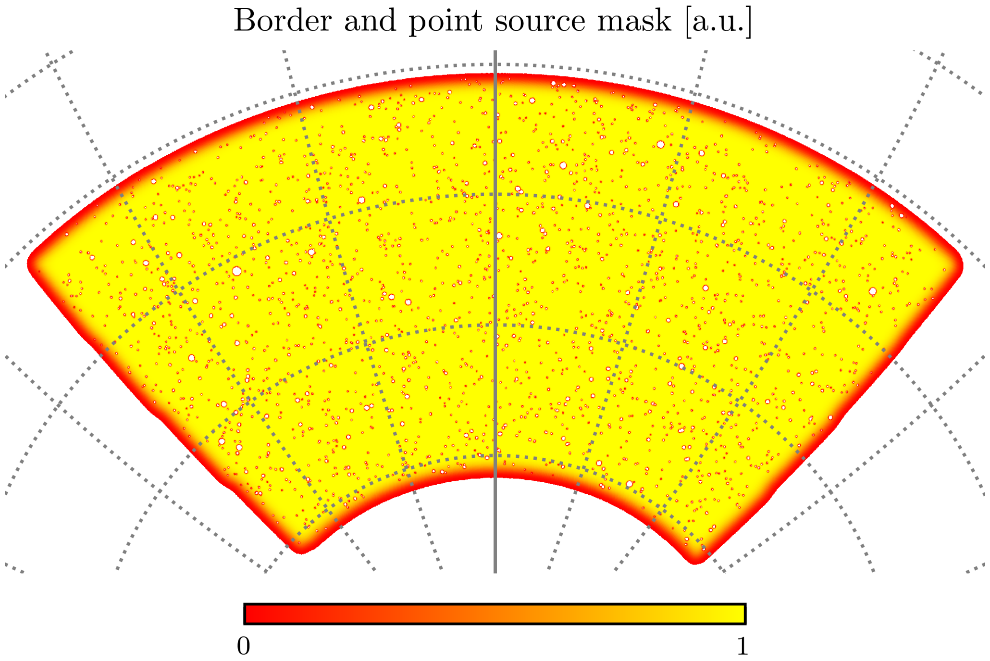


Figure 2.27: Point source mask of the SPT-3G 19/20 analysis. Borders are apodized with a 1 deg Gaussian profile. Holes are apodized with a $10'$ Gaussian profile. The mask is displayed in Lambert's azimuthal equal-area projection. Meridian graticules are plotted every 20 deg while parallel graticules are plotted every 10 deg. Apodizing the mask helps reduce Fourier ringing in harmonic space. Removing the sources is crucial for mitigating associated contamination.

### 2.3.3 Noise maps

We obtained noise maps by combining the full coadd and maps created from individual observations along with their associated weights, as for previous SPT analyses (Wu et al., 2019). To derive these noise maps, for each of the bundles $j$, each of the subfields $k$, and each of the observations $l$, we use the local temperature map $T^{jkl}$ and weight map $W^{jkl}$. We draw $\epsilon_l^{jk} = \pm 1$ with $\sum_{jkl} \epsilon_l^{jk} = 0$ and also use the full signal coadd $T^{\mathrm{coadd}}$. We build

$$N\left[\epsilon\right] = \sum_{\substack{j \text{ bundles} \\ k \text{ subfields}}} \sum_{l \text{ observations}} \epsilon_l^{jk} (W^{jkl} T^{jkl} - W^{jkl} T^{\mathrm{coadd}}) \tag{2.17}$$

By employing this procedure, we have the potential to generate $\binom{n_{obs}}{n_{obs}/2}$ noise estimations. However, it is important to consider that since they are built from the same set of observations, those maps are correlated, and only a limited amount are independent of each other. The 3290 observations imply estimations correlated below 1% of the noise power. An example of a set of resulting noise maps at 150GHz is displayed in Figs. 2.28 to 2.30. Temperature noise maps display large-scale cross-scan features, which originate from atmospheric contamination filtered along the scan. From the noise maps, we can obtain the noise levels displayed in Table 2.3. SPT-3G 19/20 are 3 to 4 times deeper than SPT-3G 2018, and much deeper than *Planck* and ACT DR4, although those experiments cover much larger sky fractions. The 150GHz channel is the deepest channel of the SPT-3G 19/20 analysis, with a noise level of 4.6$\mu$K-arcmin in temperature and 6.6$\mu$K-arcmin in polarization. Noise levels are expected to be divided by two in the full five-year survey.

| Frequency | SPT-3G 19/20 | | SPT-3G 2018 | | *Planck* | | ACT DR4 | |
|---|---|---|---|---|---|---|---|---|
| | TT | EE | TT | EE | TT | EE | TT deep | TT wide (AA) |
| 95GHz | 5.4 | 8.1 | 20.9 | 29.6 | 77.4 | 118 | > 18.4 | 72.9 |
| 150GHz | 4.6 | 6.6 | 14.9 | 21.1 | 33 | 70 | > 12.6 | 118.5 |
| 220GHz | 16 | 23 | 53 | 75 | 46.8 | 105 | NA | NA |

Table 2.3: Noise levels of the SPT-3G 19/20 analysis, compared with SPT-3G 2018, *Planck* and ACT DR4. Noise levels are given in $\mu$K-arcmin. SPT-3G 19/20 noise levels are obtained from the noise maps, while SPT-3G 2018, *Planck* and ACT DR4 noise levels are taken from Balkenhol et al. (2022), Planck Collaboration et al. (2020a) and Aiola et al. (2020), respectively.

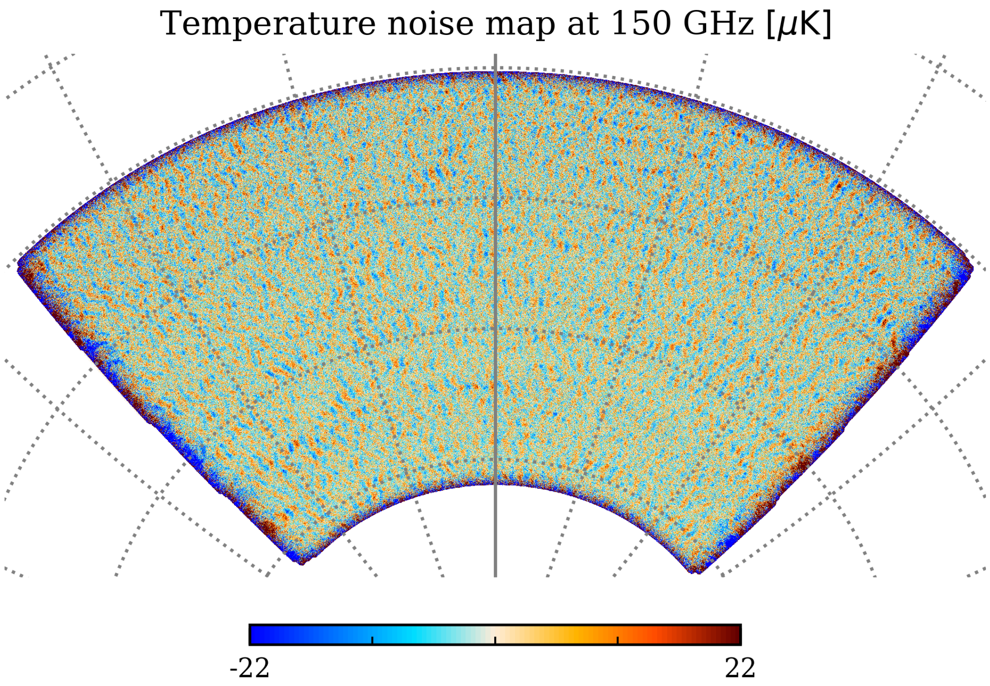


Figure 2.28: Temperature noise maps of the SPT-3G experiment at 150GHz. The map is displayed in Lambert's azimuthal equal-area projection. Meridian graticules are plotted every 20 deg while parallel graticules are plotted every 10 deg.

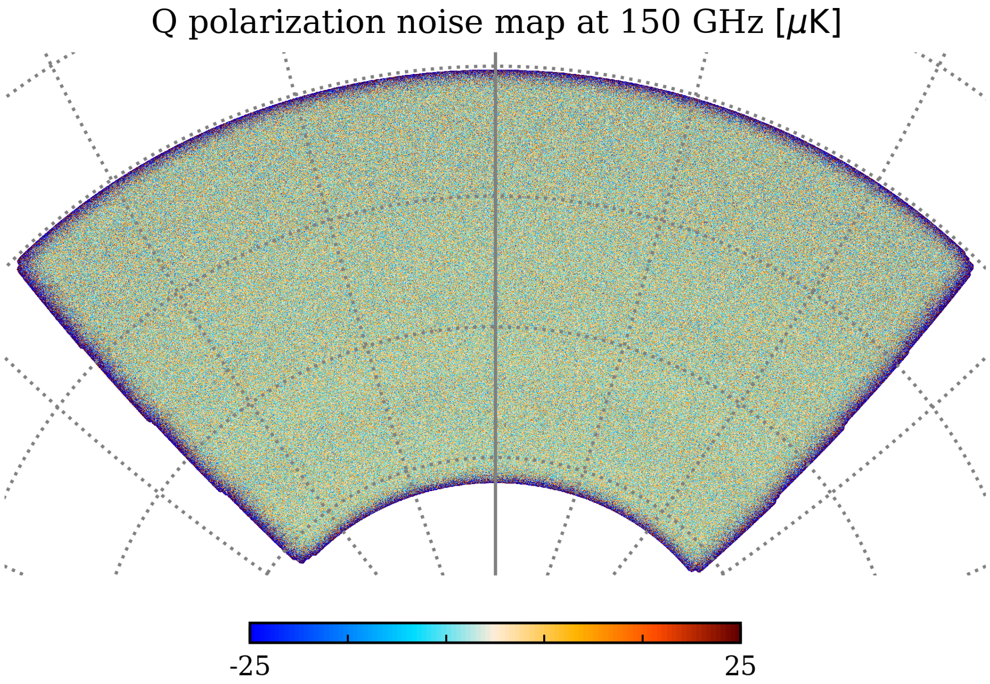


Figure 2.29: $Q$ polarization noise maps of the SPT-3G experiment at 150GHz. The map is displayed in Lambert's azimuthal equal-area projection. Meridian graticules are plotted every 20 deg while parallel graticules are plotted every 10 deg.

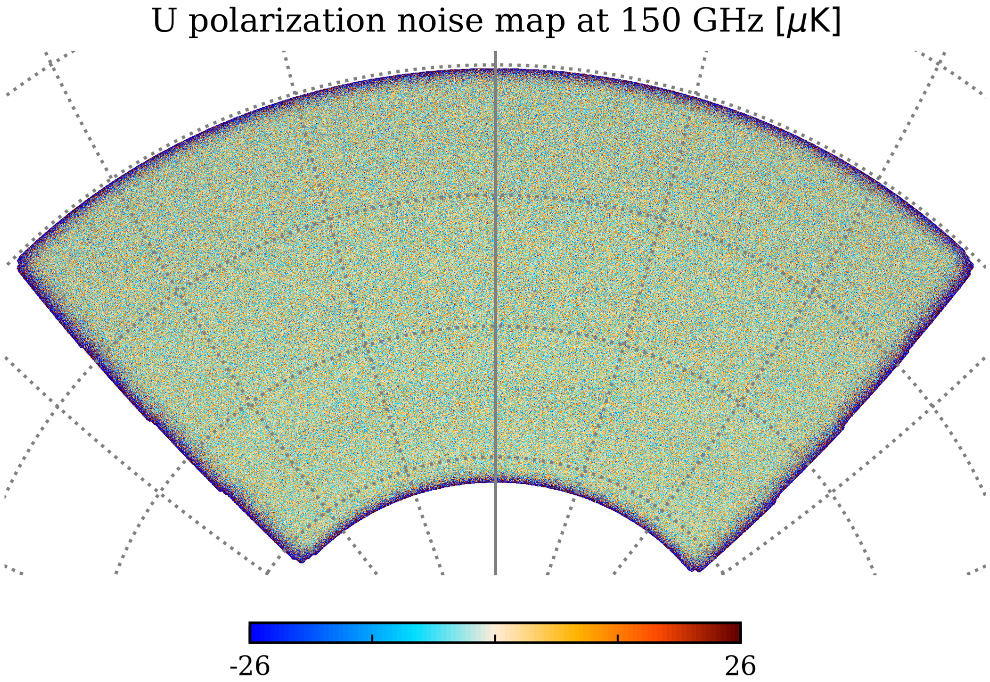


Figure 2.30: $U$ polarization noise maps of the SPT-3G experiment at 150GHz. The map is displayed in Lambert's azimuthal equal-area projection. Meridian graticules are plotted every 20 deg while parallel graticules are plotted every 10 deg.

### 2.3.4 Simulations and mock observations

For downstream analysis, we conduct a comprehensive set of high-complexity simulations known as **mock-observations**. These simulations aim to replicate the instrument and pipeline effects on the data. To generate the input skies, we utilize the AGORA software (Omori, 2022), which is specifically designed for producing realistic microwave skies with correlated foregrounds. The input CMB is assumed to be *Planck* best fit, and CMB lensing, thermal and kinetic Sunyaz-Zeldovich effects, cosmic infrared background, and radio sources are included. They are first calibrated against the observed data. Such a simulation suite enables us to investigate correlated foreground studies, such as component separation, or non-Gaussianities of the lensing potential, as well as cross-correlation with large-scale structure observables. The resulting skies, at a resolution of $n_{\mathrm{side}} = 8192$, are then interpolated into a noiseless time stream. Subsequently, the same map-making procedure is applied to these interpolated time streams using the map statistics, pointing information, and calibration information, just as in the case of the actual data. This entire procedure is repeated for multiple sky realizations to account for statistical variations in the data products. An example of the mock-observations at 150GHz is displayed exclusively on the analysis patch in Figs. 2.31 to 2.33. These figures demonstrate the same filtering characteristics as the data. Notably, since no noise was introduced during this procedure, the resulting maps solely represent the signal component of the data. Consequently, we can observe distinct polarization features, with the Q map exhibiting similar iso-longitude stripping as the temperature maps, while the U component exhibits large-scale stripes oriented at 45 deg relative to the parallels, consistent with the definition of this Stokes parameter. Importantly, these features are also observed in the actual data (see Fig. 2.23). The difference maps in Figs. 2.34 to 2.36 highlight the signal removed by the filtering, with dominant large-scale modes. To achieve a more realistic analysis, noise maps can be added from the data products in a subsequent step. Obtained maps are used to estimate the power spectrum transfer function, see Section 3.3. They are also the input simulations of the inpainting validation pipeline, see Section 3.6. They are also key elements to validate the covariance computation pipeline as detailed in Chapter 4.

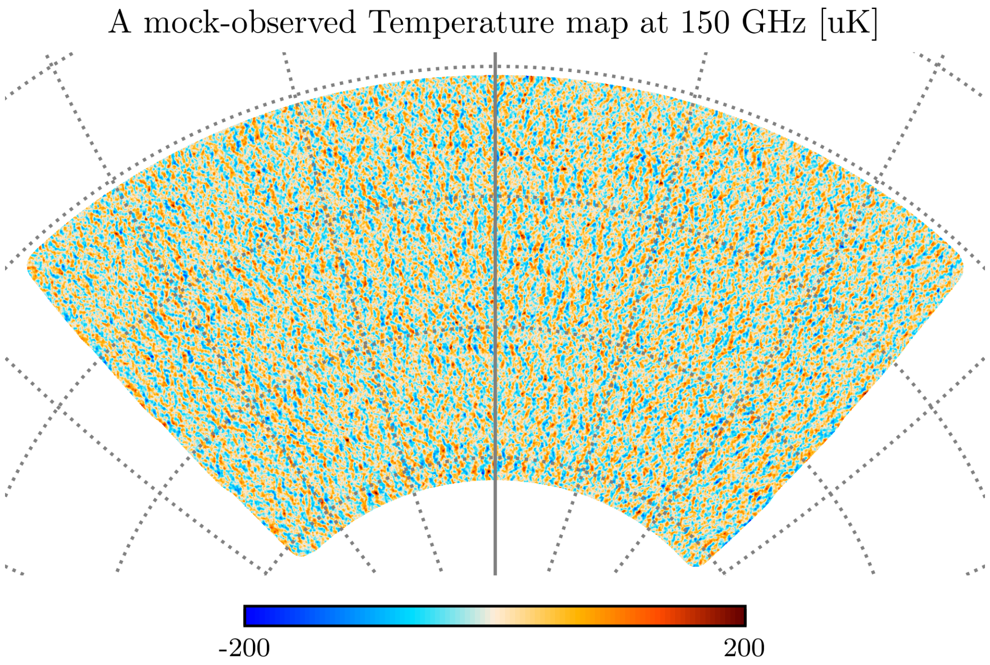


Figure 2.31: Mock temperature observations of the SPT-3G experiment at 150GHz. The map is displayed in Lambert's azimuthal equal-area projection. Meridian graticules are plotted every 20 deg while parallel graticules are plotted every 10 deg. This map shows the temperature signal after filtering, for which the iso-latitude large-scale modes have been removed.

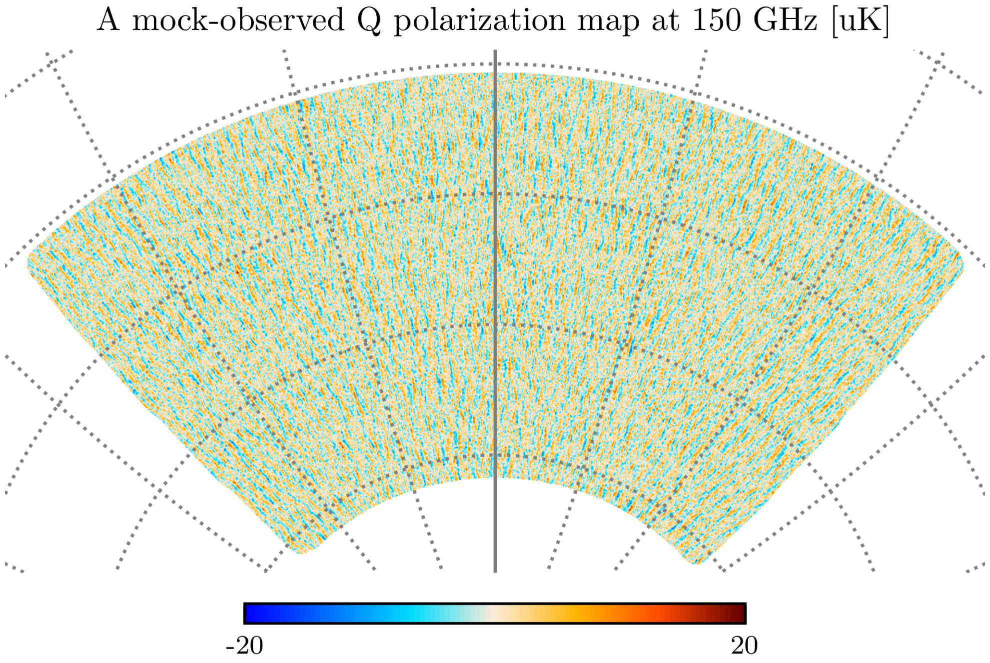


Figure 2.32: Mock Q polarization observations of the SPT-3G experiment at 150GHz. The map is displayed in Lambert's azimuthal equal-area projection. Meridian graticules are plotted every 20 deg while parallel graticules are plotted every 10 deg. This map shows the Q polarization signal after filtering, for which the iso-latitude large-scale modes have been removed.

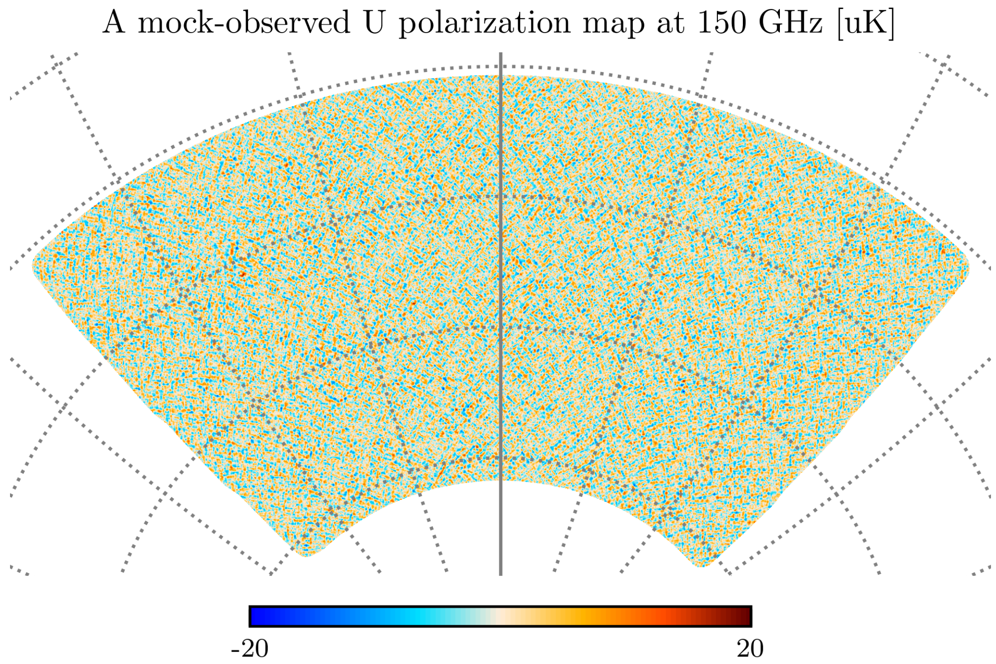


Figure 2.33: Mock U polarization observations of the SPT-3G experiment at 150GHz. The map is displayed in Lambert's azimuthal equal-area projection. Meridian graticules are plotted every 20 deg while parallel graticules are plotted every 10 deg. This map shows the U polarization signal after filtering, for which the iso-latitude large-scale modes have been removed.

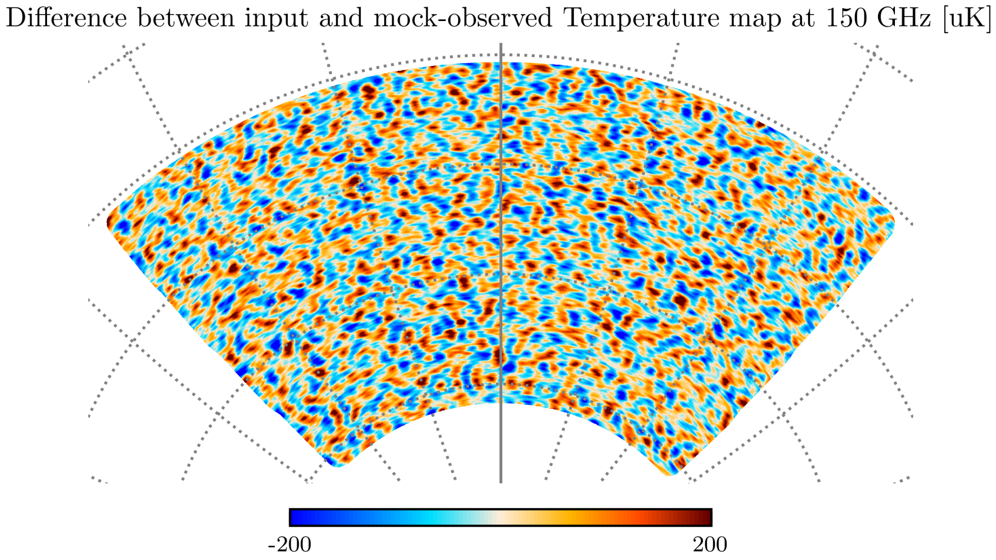


Figure 2.34: Temperature difference map at 150GHz. The map is displayed in Lambert's azimuthal equal-area projection. Meridian graticules are plotted every 20 deg while parallel graticules are plotted every 10 deg. This map shows the removed temperature signal, mostly large-scale modes.

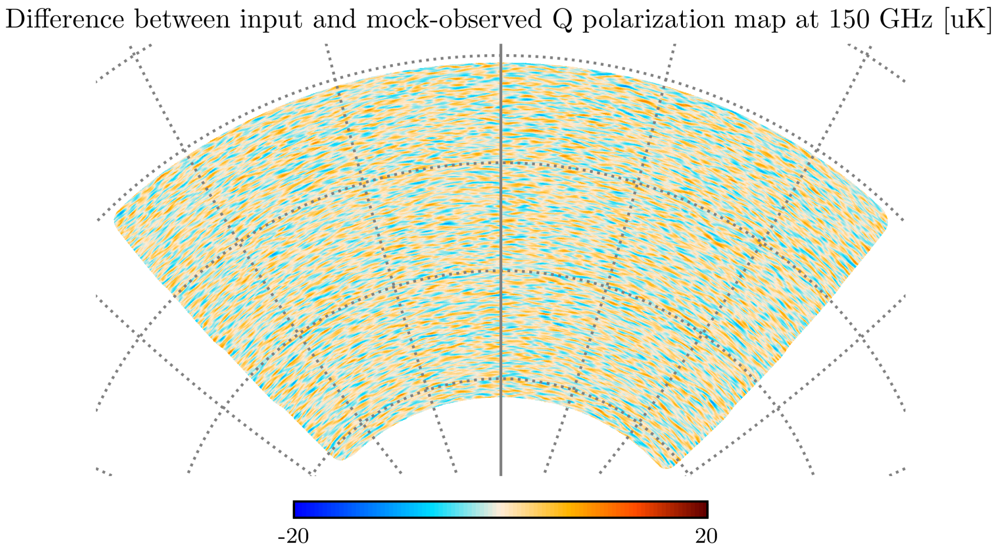


Figure 2.35: Q polarization difference map at 150GHz. The map is displayed in Lambert's azimuthal equal-area projection. Meridian graticules are plotted every 20 deg while parallel graticules are plotted every 10 deg. This map shows the removed Q polarization signal, mostly llarge-scalemodes.

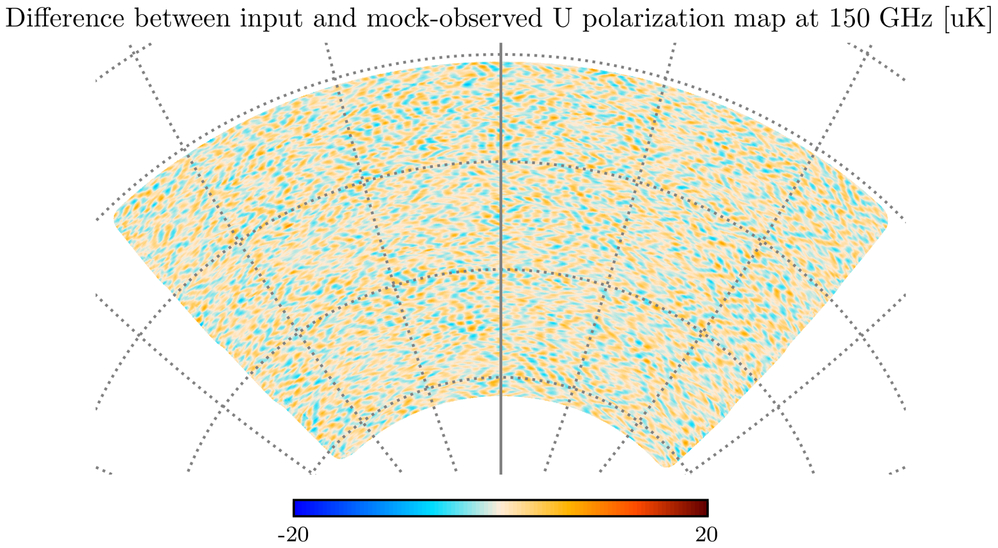


Figure 2.36: Q polarization difference map at 150GHz. The map is displayed in Lambert's azimuthal equal-area projection. Meridian graticules are plotted every 20 deg while parallel graticules are plotted every 10 deg. This map shows the removed U polarization signal, mostly large-scale modes.

# Conclusion

In this chapter, we have presented the SPT-3G 19/20 map-making pipeline and obtained data products. We first reviewed the telescope optics, scanning strategy, and pre-processing, then detailed the map-making algorithm and finally presented the resulting maps with associated secondary data products. This step allows for data reduction of the $\mathcal{O}(100)$TB of raw data to a few TB of maps. The resulting maps are then used for band powers estimation presented in the next chapter.

# Chapter 3

# Power spectrum analysis

---

**Contributions** In this section, we report the implementation of the power spectrum computation pipeline. I am leading this work inside the SPT-3G collaboration, and I obtained all of the results presented in this chapter. Results detailed in Section 3.2 have been released in (Camphuis et al., 2022). Moreover, the publication of 19/20 cosmological constraints will include a description of the power spectrum pipeline Sections 3.3 to 3.5. Finally, the inpainting procedure depicted in Section 3.6 will lead to a separate publication (Camphuis, Benabed, et al., *in prep*).

**Introduction** The primary objective of this study is to extract cosmological information from the data, which involves testing the compatibility of the cosmological model with the data and then estimating the associated parameters within the Bayesian framework. The temperature and polarization anisotropies of the CMB can be accurately modeled as statistically isotropic Gaussian random fields with predicted two-point correlation. To solve the backward problem, we could work in the map pixel domain, where the temperature map can be written as a pixel value set $T_i$, and solve the maximum likelihood problem of the map given the model. We would estimate the covariance of the pixel values for a given model $\mathcal{M}$ and its set of parameters $\vartheta$ as $\xi_{ij}[\mathcal{M},\vartheta] = \langle T_i T_j \rangle$. Assuming $\mathcal{M},\vartheta$, we can then compute the Gaussian pixel-likelihood of the data as:

$$\mathcal{L}(T_i|\mathcal{M},\vartheta) \propto \exp\left(-\frac{1}{2}\sum_{ij} T_i \left[\xi_{ij}[\mathcal{M},\vartheta]\right]^{-1} T_j\right). \tag{3.1}$$

The pixel domain is intuitive and facilitates modeling of the noise component in the covariance matrix with a Gaussian noise covariance, while masks are easy to implement. This approach is used for large-scale analysis of both WMAP and *Planck* satellite experiments (Page et al., 2007; Dunkley et al., 2009; Planck Collaboration et al., 2020c). However, the pixel space likelihood is not computationally feasible for small-scale CMB analysis due to the high-resolution maps and associated large number of pixels (on the order of $10^6$ for SPT-3G maps). This makes the inversion of the pixel covariance matrix $\xi$ impractical, as it scales as $\mathcal{O}(N_{\text{pix}}^3)$. It is customary to add a compression step before running the likelihood analysis. In this chapter, we will discuss how to transform the pixel space likelihood into a more computationally tractable form using the spherical harmonic transform and the power spectrum formalism.

## 3.1 Theoretical framework

Let us first introduce some crucial analysis tools for data analysis on the sphere and apply it to CMB maps. Random fields are functions that take random values at each point of their defined space. For their analysis, physics makes heavy use of Fourier-mode decomposition, as it allows us to infer crucial physical information by exploiting the symmetries of the problem. Such decompositions have useful properties, as they are linear, unitary, and invertible while transforming differential equations into algebraic ones. The original Fourier transform is decomposition in frequency modes of the Euclidean 3D space, but a whole class of Fourier analysis stems from it, with an equivalent of such decomposition on the 2D sphere is the spherical harmonic transform, which has profound implications in many physical fields, such as atomic physics, quantum mechanics, electromagnetism, and cosmology. In our case, the statistical properties of the field are invariant by translation thanks to CMB statistical isotropy. This implies that the Fourier modes are statistically independent, which is a powerful property that we will exploit in the following sections.

### 3.1.1 Random fields on the sphere

**Definitions** In the subsequent chapter, we will treat the CMB as a random field on the sphere $\mathbb{S}^2 = \left\{x \in \mathbb{R}^3 \,|\, ||x|| = 1\right\}$. Specifically, we postulate that the CMB anisotropies can be modeled as isotropic Gaussian random fields. While we cannot predict the exact values of the CMB anisotropies in a given direction, we can determine their statistical properties. The field follows a Gaussian distribution, implying that its probability distribution function is fully characterized by its mean and covariance. The covariance is expressed as

$$\langle T(\hat{\mathbf{u}})T(\hat{\mathbf{n}})\rangle = \xi_0(\hat{\mathbf{u}}, \hat{\mathbf{n}}), \tag{3.2}$$

where the averaging operator $\langle\cdot\rangle$ represents the expected value over all possible realizations of the field. It is worth noting that a less stringent assumption can be made by considering the possibility of departures from strict Gaussianity in the CMB. Such departures allow for the existence of higher-order correlation functions, which provide valuable insights into the properties of the early Universe and its inflationary phase. While the investigation of higher-order correlation functions is a fascinating avenue of research, it falls beyond the scope of this work. The covariance is also known as the two-point correlation function of the field. Furthermore, the CMB field exhibits statistical isotropy, which implies that its statistical properties remain unchanged under any rotation of the sphere. Consequently, the two-point correlation function of the field solely depends on the angle $\theta$ between two points on the sphere, denoted by $\hat{\mathbf{n}} \cdot \hat{\mathbf{u}} = \cos\theta$.

$$\langle T(\hat{\mathbf{n}})T(\hat{\mathbf{u}})\rangle = \xi_0(\theta). \tag{3.3}$$

**CMB polarization** The spin-2 $Q$ and $U$ fields need to be properly rotated before defining their correlation functions. Given the line-of-sights, $\hat{\mathbf{n}}$ and $\hat{\mathbf{u}}$, we use the combined parameter $P = Q + iU$ and define the rotated parameters $\bar{P}(\hat{\mathbf{n}}), \bar{P}(\hat{\mathbf{u}}) \equiv e^{2i\alpha}P(\hat{\mathbf{n}}), e^{2i\gamma}P(\hat{\mathbf{u}})$, where $\alpha$ and $\gamma$ are the angles required to rotate the local basis at, respectively, $\hat{\mathbf{n}}$ and $\hat{\mathbf{u}}$ so that they are aligned with the sphere geodesic between $\hat{\mathbf{n}}$ and $\hat{\mathbf{u}}$. Once this rotation is done, the correlation functions between the Stokes parameters are given by

$$\begin{cases} \langle \bar{P}(\hat{\mathbf{n}})\bar{P}(\hat{\mathbf{u}})\rangle & \equiv \xi_-(\theta) \\ \langle \bar{P}^*(\hat{\mathbf{n}})\bar{P}(\hat{\mathbf{u}})\rangle & \equiv \xi_+(\theta) \\ \langle T(\hat{\mathbf{n}})\bar{P}(\hat{\mathbf{u}})\rangle & \equiv \xi_\times(\theta) \end{cases} \Leftrightarrow \begin{cases} 2\langle \bar{Q}(\hat{\mathbf{n}})\bar{Q}(\hat{\mathbf{u}})\rangle & = [\xi_+ + \Re[\xi_-]]\,(\theta) \\ 2\langle \bar{U}(\hat{\mathbf{n}})\bar{U}(\hat{\mathbf{u}})\rangle & = [\xi_+ - \Re[\xi_-]]\,(\theta) \\ 2\langle \bar{Q}(\hat{\mathbf{n}})\bar{U}(\hat{\mathbf{u}})\rangle & = \Im[\xi_-](\theta) \\ \ \langle T(\hat{\mathbf{n}})\bar{Q}(\hat{\mathbf{u}})\rangle & = \Re[\xi_\times](\theta) \\ \ \langle T(\hat{\mathbf{n}})\bar{U}(\hat{\mathbf{u}})\rangle & = \Im[\xi_\times](\theta) \end{cases} \tag{3.4}$$

where we have defined the real and imaginary part operators $\Re$ and $\Im$, respectively. Due to the statistical isotropy of the temperature and polarization field $T, P$, the correlation function $\xi_0(\theta)$ and $\xi_+(\theta)$ are real. However, the other correlation functions are complex in general, but their imaginary part cancels as we assumed that the Universe is parity-invariant in mean (Chon et al., 2004). In other words, if we define the parity transformation as

$$P(\hat{\mathbf{n}}) \to P^*(-\hat{\mathbf{n}}), \tag{3.5}$$

and assume that this transformation should not affect the correlation functions on average, then

$$\langle \bar{Q}(\hat{\mathbf{n}})\bar{U}(\hat{\mathbf{u}})\rangle = 0, \quad \langle T(\hat{\mathbf{n}})\bar{U}(\hat{\mathbf{u}})\rangle = 0. \tag{3.6}$$

**Estimator** In the context of pixelated maps, it is more convenient to work with the two-point correlation function in the pixel domain. Therefore, we use vectors of size $n_{\text{pix}}$ that represent maps and we define the matrix of all possible two-point correlation functions between pixels as

$$\langle T(\hat{\mathbf{n}}_i)T(\hat{\mathbf{n}}_j)\rangle = \langle T_i T_j^\top\rangle = \xi_{ij} = \xi(\theta_{ij}), \tag{3.7}$$

where $\hat{\mathbf{n}}_i$ and $\hat{\mathbf{n}}_j$ represent the unit vectors associated with pixels $i$ and $j$, respectively, and $\theta_{ij}$ is a $n_{\text{pix}} \times n_{\text{pix}}$ matrix giving the angle between these two pixels. Since it contains all the statistical information, we aim to reconstruct the two-point correlation function of the field from a single realization of the field. In practice, it is not feasible to measure an infinite number of CMB realizations to obtain the exact covariance function, and thus we can only estimate it based on the available unique realization. We can estimate the temperature two-point correlation function of an isotropic field from a single realization of the field using the following binned estimator, defined in Szapudi et al. (2001):

$$\bar{\xi}(\theta) = \frac{1}{\sum_{ij} f_{ij}^\theta} \sum_{i,j=1}^{n_{\text{pix}}} f_{ij}^\theta T_i T_j \quad \implies \quad \langle \bar{\xi}(\theta_{ij})\rangle = \xi(\theta_{ij}) \quad \forall i, j. \tag{3.8}$$

where $f_{ij} = 1$ is a window function that is one when the angles between corresponding pixels are in the bin $\theta_{ij} = \theta$, enabling the extraction of the relevant pixels. We encapsulate the estimation of a correlation from a map into the non-invertible operator $\mathbb{G}$ as

$$\bar{\xi} = \mathbb{G}(T, T), \tag{3.9}$$

where we allowed the operator to take as input two different maps. The estimator is unbiased, meaning that its expected value is equal to the true value of the two-point correlation function. Its covariance is given by

$$\mathrm{cov}\left(\bar{\xi}(\theta_1), \bar{\xi}(\theta_2)\right) = \frac{1}{\sum_{ij} f_{ij}^{\theta_1} \sum_{kl} f_{kl}^{\theta_2}} \sum_{i,j,k,l=1}^{n_{\mathrm{pix}}} f_{ij}^{\theta_1} f_{kl}^{\theta_2} \left[\xi(\theta_{ik})\xi(\theta_{jl}) + \xi(\theta_{il})\xi(\theta_{jk})\right]. \tag{3.10}$$

This approach is not optimal because since the pixels are correlated, the estimator of the two-point correlation function is dense and highly non-diagonal, which makes it unsuitable for parameter estimation as it can propagate errors. We can solve this problem by considering the two-point correlation function in the harmonic domain, as we will see in Section 3.1.2.

In summary, by considering the CMB as an isotropic Gaussian random field, we can characterize entirely its statistical properties through its two-point correlation function. The two point-correlation introduced here is the most basic statistical property of the field, used in Eq. (3.1). We will now introduce the Legendre decomposition and the spherical harmonic transform, which allows us to express the two-point correlation function in the harmonic domain, where its covariance is diagonal and sparse, making it more suitable for parameter estimation, as we will see in Chapter 4 (Efstathiou, 2004; Percival and Brown, 2006).

### 3.1.2 Harmonic decomposition

The harmonic analysis of the CMB anisotropies naturally put forward typical scales and patterns in the temperature and polarization fields. It allows us to study spatial angular modes on the sphere, which are intrinsically related to independent primordial anisotropy modes. The harmonic analysis of the CMB anisotropies is based on the spherical harmonic transform, which we introduce in this section.

**Spin weighted spherical harmonics** Complete and orthogonal sets of polynomials on the two-dimensional sphere can be constructed for Fourier analysis. This is achieved through the spin-weighted spherical harmonic decomposition, extensively discussed in Penrose (1966); Goldberg et al. (1967). The spin-weighted spherical harmonics ${}_sY_{\ell m}(\hat{\mathbf{n}})$, characterized by their degree or multipole $\ell$, and order $m$, form an infinite set of spin-weighted orthogonal polynomials on the sphere for any spin $s$. The standard set of spherical harmonics corresponds to the spin-0 scenario and finds widespread utilization across various physics domains. For the analysis of CMB polarization, the spin-2 set of spherical harmonics emerges as particularly advantageous. By employing the appropriate spherical harmonics, we can express the temperature field $T(\hat{\mathbf{n}})$ and the polarization field $P(\hat{\mathbf{n}})$ as expansions involving their respective spherical harmonics.

$$\begin{cases} T(\hat{\mathbf{n}}) = \sum_{\ell m} a_{\ell m}^{\mathrm{T}} {}_0Y_{\ell m}(\hat{\mathbf{n}}) \\ (Q \pm iU)(\hat{\mathbf{n}}) = \sum_{\ell m} (a_{\ell m}^{\mathrm{E}} \mp i a_{\ell m}^{\mathrm{B}}) {}_{\mp 2}Y_{\ell m}(\hat{\mathbf{n}}) \end{cases} \Leftrightarrow \begin{cases} a_{\ell m}^{\mathrm{T}} = \int \mathrm{d}\hat{\mathbf{n}} T(\hat{\mathbf{n}}) {}_0Y_{\ell m}^*(\hat{\mathbf{n}}) \\ (a_{\ell m}^{\mathrm{E}} \mp i a_{\ell m}^{\mathrm{B}}) = \int \mathrm{d}\hat{\mathbf{n}} (Q \pm iU)(\hat{\mathbf{n}}) {}_{\mp 2}Y_{\ell m}^*(\hat{\mathbf{n}}) \end{cases} \tag{3.11}$$

With those definitions, the following relations are verified

$$a_{\ell m}^{\mathrm{X}} = (-1)^m a_{\ell -m}^{\mathrm{X}} \quad \forall \mathrm{X} \in \{\mathrm{T}, \mathrm{E}, \mathrm{B}\}. \tag{3.12}$$

The E and B modes are a natural decomposition of the linear spin-2 polarization field in the CMB. In the context of the CMB, polarization vectors of the E modes are directed radially around cold spots and tangentially around hot spots. These E modes are curl-free and primarily arise due to density fluctuations in the early universe, followed by the scattering of CMB photons by these fluctuations. They exhibit a gradient-like pattern and are responsible for the observed striped patterns in CMB polarization. On the other hand, the B modes of the CMB polarization field are divergence-free and exhibit a curl or vorticity around any direction of the sky. They are particularly significant as the detection of primordial B-modes might provide evidence for the inflationary theory of the early universe. B modes arise due to primordial gravitational waves generated during the inflationary phase (Baumann et al., 2009). These B modes display a vorticity-like pattern and offer valuable insights into the physics of the early universe. These physical properties of the E and B modes are related to parity transformations. Specifically, the spin-2 spherical harmonic coefficients, denoted as $a_{\ell m}^{\mathrm{E}}$ and $a_{\ell m}^{\mathrm{B}}$, undergo specific transformations under parity. As stated in Eq. (3.5), the E-mode coefficients transform as $a_{\ell m}^{\mathrm{E}} \to (-1)^\ell a_{\ell m}^{\mathrm{E}}$, while the B-mode coefficients transform as $a_{\ell m}^{\mathrm{B}} \to (-1)^{\ell+1} a_{\ell m}^{\mathrm{B}}$. This parity behavior of the spherical harmonic coefficients further highlights the distinct properties of the E and B modes in the CMB polarization field. As a result, the power spectra $C_\ell^{\mathrm{EE}}$ and $C_\ell^{\mathrm{BB}}$ are even under parity, while $C_\ell^{\mathrm{TE}}$ and $C_\ell^{\mathrm{TB}}$ are odd. Structure and parity properties of the E and B modes are illustrated in Fig. 3.1. We compact Eq. (3.11)

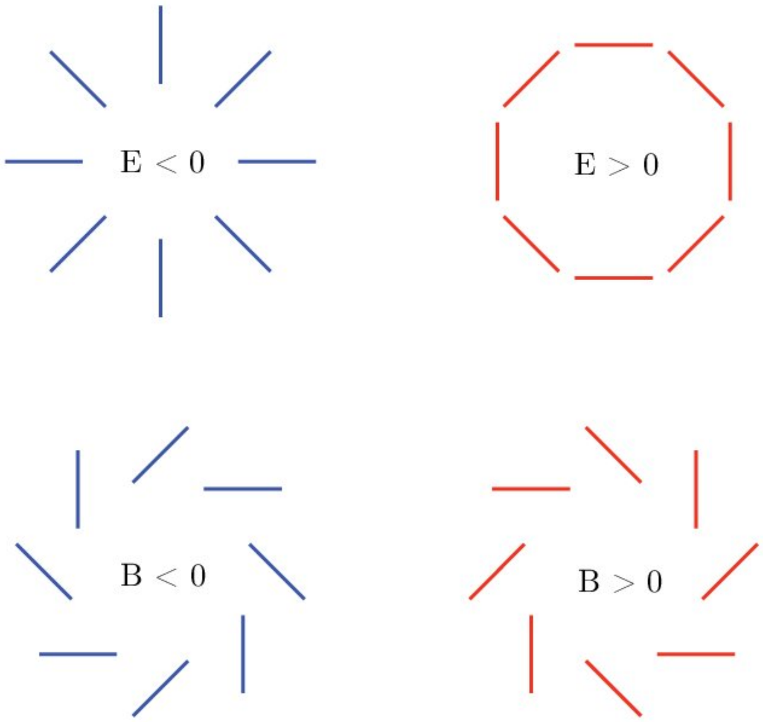


Figure 3.1: Pure E-mode and B-mode patterns of polarization. Note that if reflected across a line going through the center the E-patterns are unchanged, while the positive and negative B-patterns get interchanged. Figure from Baumann et al. (2009).

into an operator form, which we will use extensively in the following,

$$\begin{cases} T = \mathbb{SHT}_0^{-1} \cdot a^{\mathrm{T}} \\ (Q \pm iU) = \mathbb{SHT}_{\mp 2}^{-1} \cdot (a^{\mathrm{E}} \mp i a^{\mathrm{B}}) \end{cases} \Leftrightarrow \quad \begin{cases} a^{\mathrm{T}} = \mathbb{SHT}_0 \cdot T \\ (a^{\mathrm{E}} \mp i a^{\mathrm{B}}) = \mathbb{SHT}_{\mp 2} \cdot (Q \pm iU) \end{cases} \tag{3.13}$$

In the context of our analysis, harmonic coefficients are considered Gaussian random variables, as they are obtained from the harmonic decomposition of the temperature and polarization fields, which are themselves Gaussian random variables, with zero mean. As they are isotropic, their statistical properties do not depend on the order $m$, and their variance is given by the angular power spectrum.

$$\begin{pmatrix} a^{\mathrm{T}}_{\ell m} \\ a^{\mathrm{E}}_{\ell m} \\ a^{\mathrm{B}}_{\ell m} \end{pmatrix} \sim \mathcal{N}\left( \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} C_\ell^{\mathrm{TT}} & C_\ell^{\mathrm{TE}} & C_\ell^{\mathrm{TB}} \\ C_\ell^{\mathrm{TE}} & C_\ell^{\mathrm{EE}} & C_\ell^{\mathrm{BB}} \\ C_\ell^{\mathrm{TB}} & C_\ell^{\mathrm{EB}} & C_\ell^{\mathrm{BB}} \end{pmatrix} \right) \tag{3.14}$$

In other words,

$$\langle a^{\mathrm{X}}_{\ell m} \rangle = 0, \quad \langle a^{\mathrm{X}}_{\ell m} a^{\mathrm{X}'*}_{\ell' m'} \rangle = \delta_{\ell\ell'} \delta_{mm'} C_\ell^{\mathrm{XX}'}, \forall \mathrm{X}, \mathrm{X}' \in \{\mathrm{T}, \mathrm{E}, \mathrm{B}\}. \tag{3.15}$$

These properties make it straightforward to generate a realization of the harmonic coefficients, given a set of theoretical power spectra. One can first draw a unity variance noise realization and color it with the square root of the covariance matrix, here the power spectrum. This is easily obtainable using Cholesky decomposition of the spectra (Petersen and Pedersen, 2012). Going back to real space, we can then obtain a realization of the temperature and polarization fields. This is the approach we will use in the following to generate pure Gaussian simulations of the CMB sky, with the `HEALPix` software.

**From spherical harmonic transform to Legendre decomposition** In spherical harmonic analysis, Wigner $D$-matrices play a crucial role in propagating the properties of spin fields into the harmonic domain (Khersonskii et al., 1988; Ng and Liu, 2001). These matrices are representations of the SU(2) group and serve as rotation operators for spherical harmonics and can be obtained from spin-weighted spherical harmonics using the relation:

$$D^\ell_{ss'}(\alpha, \beta, \gamma) = \frac{4\pi}{2\ell + 1} \sum_m {}_sY^*_{\ell m}(\hat{\mathbf{n}}) {}_{s'}Y_{\ell m}(\hat{\mathbf{n}}'), \tag{3.16}$$

where $\alpha$, $\beta$, and $\gamma$ are the Euler angles defining the rotation from $\hat{\mathbf{n}}$ to $\hat{\mathbf{n}}'$. We can introduce a set of general 1D polynomials, known as reduced Wigner $D$-matrices:

$$d^\ell_{ss'}(\beta) = e^{is\alpha} D^\ell_{ss'}(\alpha, \beta, \gamma) e^{-is'\gamma}. \tag{3.17}$$

The reduced Wigner matrices can be used to decompose the 1D real line into series expansion, as in particular, the spin-0 case is the well-known Legendre decomposition. Concerning spin-2, the rotation of the polarization fields in Eq. (3.4) is done in an analogous way we defined Eq. (3.17), showing the intrinsic relations between the two. From equations (3.16) and (3.17), we obtain the addition theorem, which establishes a connection between spherical harmonics and the Legendre decomposition.

$$P_\ell(\hat{\mathbf{n}}\cdot\hat{\mathbf{u}}) \equiv d^\ell_{00}(\arccos(\hat{\mathbf{n}}\cdot\hat{\mathbf{u}})) = \frac{4\pi}{2\ell+1}\sum_{m=-\ell}^{\ell} {}_0Y_{\ell m}(\hat{\mathbf{n}}) {}_0Y^*_{\ell m}(\hat{\mathbf{u}}) \quad \implies \quad \frac{4\pi}{2\ell+1}\sum_{\ell m} {}_0Y_{\ell m}(\hat{\mathbf{n}}) {}_0Y^*_{\ell m}(\hat{\mathbf{n}}) = 1. \tag{3.18}$$

The last equality is known as the Unsöld theorem for spin-0 fields. We can rewrite the addition theorem by first introducing a new operator, $\mathbb{H}$, which sums spherical harmonic coefficients of the same multipole $\ell$,

$$\mathbb{H}_\ell\left(a,b\right) \equiv \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell} a_{\ell m} b^*_{\ell m} \implies P_\ell(\hat{\mathbf{n}}\cdot\hat{\mathbf{u}}) = 4\pi\mathbb{H}_\ell\left({}_0Y, {}_0Y\right). \tag{3.19}$$

In our study of isotropic fields, where statistical properties do not depend on direction, we can leverage the relationship between spherical harmonics and Legendre polynomials. Indeed, the degree $\ell$ of spherical harmonics indicates the mode number while the order $m$ is related to the direction of the associate mode on the sphere. Assuming isotropy implies that any direction is equivalent, allowing us to consider that all orders $m$ for a given degree $\ell$ have the same statistical properties. Using Eqs. (3.15) and (3.18), we can write the two-point correlation function as

$$\xi_0(\arccos\hat{\mathbf{n}}\cdot\hat{\mathbf{u}}) = \langle T(\hat{\mathbf{n}})T(\hat{\mathbf{u}})\rangle = \sum_{\ell m, \ell' m'} {}_0Y_{\ell m} {}_0Y^*_{\ell' m'} \langle a^{\mathrm{T}}_{\ell m} a^{\mathrm{T}*}_{\ell' m'}\rangle \tag{3.20}$$

$$= \sum_\ell C^{\mathrm{TT}}_\ell \sum_m {}_0Y_{\ell m}(\hat{\mathbf{n}}) {}_0Y^*_{\ell m}(\hat{\mathbf{u}}) = \sum_\ell \frac{2\ell+1}{4\pi} C^{\mathrm{TT}}_\ell P_\ell(\hat{\mathbf{n}}\cdot\hat{\mathbf{u}}). \tag{3.21}$$

$$\Leftrightarrow \xi_0 = \mathbb{L}_0^{-1}\cdot C^{\mathrm{TT}} \tag{3.22}$$

This is the Wiener-Khinchin theorem equating the power spectrum of a random field as the Fourier transform of its two-point correlation function. The same can be done for polarization, using the spin-2 spherical harmonics and the completeness relations introduced above,

$$\xi_+(\theta) = \sum_\ell \frac{2\ell+1}{4\pi}\left(C^{\mathrm{EE}}_\ell + C^{\mathrm{BB}}_\ell\right) d^\ell_{22}(\theta), \tag{3.23}$$

$$\xi_-(\theta) = \sum_\ell \frac{2\ell+1}{4\pi}\left(C^{\mathrm{EE}}_\ell - C^{\mathrm{BB}}_\ell - 2iC^{\mathrm{EB}}_\ell\right) d^\ell_{2-2}(\theta), \tag{3.24}$$

$$\xi_\times(\theta) = \sum_\ell \frac{2\ell+1}{4\pi}\left(C^{\mathrm{TE}}_\ell - iC^{\mathrm{TB}}_\ell\right) d^\ell_{20}(\theta). \tag{3.25}$$

We compact those relations into an operator form, which we will use extensively in the following,

$$C^{\mathrm{TT}} = \mathbb{L}_0\cdot\xi_0, \quad C^{\mathrm{EE}} + C^{\mathrm{BB}} = \mathbb{L}_+\cdot\xi_+, \quad C^{\mathrm{EE}} - C^{\mathrm{BB}} - 2iC^{\mathrm{EB}} = \mathbb{L}_-\cdot\xi_-, C^{\mathrm{TE}} - iC^{\mathrm{TB}} = \mathbb{L}_\times\cdot\xi_\times. \tag{3.26}$$

Note that the Legendre operator $\mathbb{L}$ and its inverse $\mathbb{L}^{-1}$ are linear operators.

### 3.1.3 Relations in harmonic and real space

In the next Section 3.2, we will write relations in real and harmonic space to build unbiased estimators of the CMB power spectrum, using tools introduced in the following. More details can be found in Hivon et al. (2002).

**From scalar multiplication to matrix multiplications** A principal property of Fourier analysis is the convolution theorem, which states that scalar multiplication is transformed into convolution in the dual space. This property is not valid anymore when working on the curved sphere, however, we can derive an equivalent class of relations. We define for a map $W$, the following matrix ${}_sI_{\ell m \ell' m'}\left[W\right]$,

$${}_sI_{\ell m \ell' m'}\left[W\right] \equiv \int \mathrm{d}\hat{\mathbf{n}}\, {}_sY_{\ell m}(\hat{\mathbf{n}}) W(\hat{\mathbf{n}}) {}_sY^*_{\ell' m'}(\hat{\mathbf{n}}) \quad \forall s\in\{0,\pm 2\}, \quad \forall \ell,\ell',m,m'. \tag{3.27}$$

We can then write the following relations, for any spin-$s$ field $F_s$,

$$\mathbb{SHT}_s\cdot(W\times F_s) = {}_sI[W] * (\mathbb{SHT}_s\cdot F_s) \tag{3.28}$$

where $*$ denotes the matrix multiplication operator and $\times$ the scalar multiplication. For a function $w$ of the polar coordinate, we define the **mode-coupling** operator $\Xi$ acting on $w$ and outputing a symmetric matrix as follows:

$$\Xi^{ss'}_{\ell\ell'}[w] \equiv \frac{1}{2}\int_0^\pi w(\theta) d^{\ell}_{ss'}(\theta) d^{\ell'}_{ss'}(\theta) \; \mathrm{d}\cos\theta. \tag{3.29}$$

If the function is set to one, then the orthogonality of the reduced Wigner matrices implies that the operator gives the identity matrix. Adding a right multiplication by a rescaling factor, we define the non-symmetric matrices as

$${}_0M_{\ell\ell'} = (2\ell'+1)\Xi^{00}_{\ell\ell'}[w], \quad {}_\pm M_{\ell\ell'} \equiv (2\ell'+1)\Xi^{2\pm2}_{\ell\ell'}[w], \quad {}_\times M_{\ell\ell'} \equiv (2\ell'+1)\Xi^{20}_{\ell\ell'}[w]. \tag{3.30}$$

For any function $f_s$, we have

$$\mathbb{L}_s \cdot (w \times f_s) = {}_sM * (\mathbb{L}_s \cdot f_s), \quad \forall s \in [0, \pm, \times]. \tag{3.31}$$

**Focus on the mode-coupling matrix** In the previous paragraph, we introduced the symmetric matrix $\Xi[w]$, which plays a crucial role in deriving useful relations in harmonic space. In Eq. (3.29), we defined $\Xi[w]$ as a function of the polar function $w$. However, we can extend the definition of the operator $\Xi$ to act on a spectrum, where:

$$\Xi^{ss'}_{\ell\ell'}[\mathcal{W}] \equiv \Xi^{ss'}_{\ell\ell'}[\mathbb{L}_0^{-1}\mathcal{W}]. \tag{3.32}$$

For any function $w$, we have:

$$\mathcal{W} \equiv \mathbb{L}_0 \cdot w \implies \Xi^{ss'}_{\ell\ell'}[w] = \sum_L \frac{2L+1}{4\pi}\mathcal{W}_L \begin{pmatrix} \ell & \ell' & L \\ s & -s & 0 \end{pmatrix} \begin{pmatrix} \ell & \ell' & L \\ s' & -s' & 0 \end{pmatrix}, \tag{3.33}$$

where we used the Wigner 3-$j$ symbols (Khersonskii et al., 1988). We can relate the mode-coupling symmetric operator of Eq. (3.29) to the coupling coefficients relying on the addition theorem of Eq. (3.18). We have, for any spin-$s$, the following orthogonality relations for coupling coefficients, see Hivon et al. (2002) for more details,

$$\sum_{\ell m} {}_sI_{\ell_1 m_1 \ell m}[W_1] {}_sI^*_{\ell_2 m_2 \ell m}[W_2] = {}_sI_{\ell_1 m_1 \ell_2 m_2}[W_1 W_2], \tag{3.34}$$

$$\sum_{m_1 m_2} {}_sI_{\ell_1 m_1 \ell_2 m_2}[W_1] {}_{s'}I^*_{\ell_1 m_1 \ell_2 m_2}[W_2] = (2\ell_1+1)(2\ell_2+1)\Xi^{ss'}_{\ell_1\ell_2}[\mathbb{G}(W_1, W_2)], \tag{3.35}$$

where $\mathbb{G}$ is the two-point correlation function estimator defined in Eq. (3.9), giving the two-point correlation function of the input maps. This leads us to extend the mode-coupling operator so that it acts on map $W_1, W_2$, with

$$\Xi^{ss'}_{\ell\ell'}[W_1, W_2] \equiv \Xi^{ss'}_{\ell\ell'}[\mathbb{G}(W_1, W_2)]. \tag{3.36}$$

The symmetric mode-coupling operator will be used throughout this chapter to expand relations between the power spectrum estimator and the expected theory. It also constitutes a building block of the covariance matrix as we will later detail in Chapter 4. We have seen through this section that one can write direct relations between quantities in real space and their equivalent in harmonic space.

$$\begin{gathered} \mathbb{L} \cdot \mathbb{G}(W_1 T_1, W_2 T_2) = \mathbb{H}\left(\mathbb{SHT} \cdot W_1 T_1,\ \mathbb{SHT} \cdot W_2 T_2\right) = M * \mathbb{H}\left(\mathbb{SHT} \cdot T_1, \mathbb{SHT} \cdot T_2\right), \\ \text{where } M_{\ell\ell'} = (2\ell'+1)\Xi_{\ell\ell'}[\mathbb{G}(W_1, W_2)]. \end{gathered} \tag{3.37}$$

This last relation is a concise way to write

$$\begin{aligned} &\int_0^1 \mathrm{d}\cos\theta P_\ell(\cos\theta) \iint \mathrm{d}\hat{\mathbf{n}}\mathrm{d}\hat{\mathbf{u}} T(\hat{\mathbf{n}}) T(\hat{\mathbf{u}}) \delta(\hat{\mathbf{n}} \cdot \hat{\mathbf{u}} - \cos\theta) \\ &\qquad = \frac{1}{2\ell+1} \sum_m \int \mathrm{d}\hat{\mathbf{n}} [Y^*_{\ell m} W T](\hat{\mathbf{n}}) \int \mathrm{d}\hat{\mathbf{u}} [Y_{\ell m} W T](\hat{\mathbf{u}}) \\ &\qquad = \sum_{\ell'} (2\ell'+1) \Xi_{\ell\ell'}[\mathcal{W}] \sum_{m'} \int \mathrm{d}\hat{\mathbf{n}} [Y^*_{\ell' m'} T](\hat{\mathbf{n}}) \int \mathrm{d}\hat{\mathbf{u}} [Y_{\ell' m'} T](\hat{\mathbf{u}}) \\ &\qquad \text{where } \mathcal{W}_L = \sum_M \int \mathrm{d}\hat{\mathbf{n}} [Y^*_{LM} W](\hat{\mathbf{n}}) \int \mathrm{d}\hat{\mathbf{u}} [Y_{LM} W](\hat{\mathbf{u}}). \end{aligned} \tag{3.38}$$

## 3.2 Estimators of the power spectrum

In the previous section, we introduced the estimators of the correlation function between the temperature and polarization fields of the CMB in real space. These quantities provide a comprehensive description of the field properties. However, in practical observations, it is easier to work in the harmonic domain, necessitating the estimation of the power spectra from the unique observation of the sky. Assuming isotropy, we recognize that the harmonic coefficients at the same multipole $\ell$ possess equal variances. Leveraging this property, an unbiased estimator for the power spectrum can be obtained by averaging the square of the coefficients over the multipole $\ell$:

$$\bar{C}_\ell^{\mathrm{XY}} \equiv \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell} a_{\ell m}^{\mathrm{X}} a_{\ell m}^{\mathrm{Y}*}, \implies \langle \bar{C}_\ell^{\mathrm{XY}} \rangle = C_\ell^{\mathrm{XY}}, \quad \forall \mathrm{X},\mathrm{Y} \in \{\mathrm{T},\mathrm{E},\mathrm{B}\}. \tag{3.39}$$

It is important to note that this estimator has a nonzero variance due to its reliance on random variables. Consequently, it is not expected to perfectly reconstruct the true underlying power spectrum. As highlighted in Percival and Brown (2006), the variance of this full-sky estimator is given by:

$$\sigma^2_{\bar{C}_\ell^{\mathrm{XY}}} = \frac{1}{2\ell+1}\left( C_\ell^{\mathrm{XY}\,2} + C_\ell^{\mathrm{XX}} C_\ell^{\mathrm{YY}} \right). \tag{3.40}$$

The limitation of information in estimating cosmological parameters is a well-known challenge faced by every cosmological experiment, known as **cosmic-variance**. To address this issue, Bayesian formalism is particularly suitable. In Chapter 4, we will delve into the details of the Bayesian framework and its application to parameter estimation. Another noteworthy characteristic of this estimator is its direct relation to the equivalent estimator in real space, as derived in Eq. (3.8), where $\bar{C} = \mathbb{L} \cdot \bar{\xi}$. In practice, to perform the estimation, we employ the `HEALPix` package introduced in Section 2.2.1 to conduct spherical harmonic decompositions, which scale as $\mathcal{O}(\ell_{\max}^3)$, where $\ell_{\max}$ is the maximum multipole of the decomposition.

### 3.2.1 Mask and pseudo-power spectrum

The sky is masked both due to the instrument scanning strategy and to avoid spurious foreground contaminations. Foregrounds can be galactic dust contamination, radio sources, AGNs, clusters, or dusty galaxies. We detailed the construction of the mask in Section 2.3.2. As a consequence, the power spectrum is measured on the partial sky. Efforts have been made to recover the power spectrum with minimal information loss using maximum-likelihood approaches (Bond et al., 2000; Wandelt et al., 1998) or minimum-variance and quadratic estimators (Tegmark, 1997; Tegmark and de Oliveira-Costa, 2001; Vanneste et al., 2018). However, the computational implementation of these methods becomes highly expensive for high-resolution data. The computational cost scales as $\mathcal{O}(\ell_{\max}^6)$ for maximum multipole $\ell_{\max}$ or $\mathcal{O}(\ell_{\max}^4)$ for optimal algorithms like Wandelt and Hansen (2003). The alternative pseudo-power spectrum approach was first used by Yu and Peebles (1969), and has gained popularity. Those algorithms preserve the scaling of $\mathcal{O}(\ell_{\max}^3)$ of spherical harmonic transforms. It can be optimized for measuring the CMB E/B-mode power spectrum (Lewis et al., 2002; Bunn et al., 2003; Challinor and Chon, 2004; Chon et al., 2004; Smith, 2006; Grain et al., 2009; Bunn, 2011). Several useful publicly available codes exist to perform these calculations, including `MASTER` (Hivon et al., 2002), `NaMaster` (Alonso et al., 2019), `PolSpice` (Szapudi et al., 2001; Chon et al., 2004), and `Xspect`, `Xpure` (Tristram et al., 2005; Grain et al., 2009). In this section and our analysis pipeline, we will focus on the `MASTER` and `PolSpice` framework, used in the *Planck* pipeline (Planck Collaboration et al., 2020b), which allows fast and reliable computation and debiasing of the pseudo-power spectrum and associated covariance matrix.

Let us consider that we only observe $(TW, QW, UW)$, where $W$ is the mask. This operation is also known as a Gabor transform and its impact in harmonic space has been extensively studied in Hansen et al. (2002). In the following and for the rest of this work, all **pseudo** quantities, *i.e.* masked quantities, will be labeled with a tilde. The pseudo-harmonic coefficients are defined as

$$\tilde{a}^T \equiv \mathbb{SHT}_0 \cdot (W \times T), \quad \left(\tilde{a}^{\mathrm{E}} \mp i\tilde{a}^{\mathrm{B}}\right) \equiv \mathbb{SHT}_{\mp 2} \cdot (W \times (Q \pm iU)). \tag{3.41}$$

We can relate the pseudo-harmonic coefficients to the harmonic coefficients of the underlying maps with the aforementioned coupling coefficients of Eq. (3.28),

$$\tilde{a}^{\mathrm{T}} = {}_0I[W] * a^{\mathrm{T}}, \quad \left(\tilde{a}^{\mathrm{E}} \pm i\tilde{a}^{\mathrm{B}}\right) = {}_{\pm 2}I[W] * \left(a^{\mathrm{E}} \pm i a^{\mathrm{B}}\right). \tag{3.42}$$

In the full-sky case, we obtain ${}_sI_{\ell m \ell' m'}[1] = \delta_{\ell\ell'}\delta_{mm'}$ through the orthonormality properties of the spin-weighted spherical harmonics. In the top-hat window case, Wandelt et al. (1998) has derived an analytical expression of such kernels, while Challinor et al. (2002) did it for the ring case. In this work, we developed a code relying on the `HEALPix` package to compute those coefficients at low $\ell, m$ writing that

$${}_sI_{\ell m \ell' m'} = \left(\mathbb{SHT}_s \cdot [W\, {}_sY_{\ell' m'}]\right)_{\ell m}. \tag{3.43}$$

This operation scales as $\mathcal{O}(\ell_{\max}^3)$ and is reasonable for low $\ell_{\max}$. Eq. (3.42) gives a full system of equations. Inverting this system is impossible as it would imply we could reconstruct the total CMB from a portion of it. We will rather use the pseudo-harmonic coefficients to estimate the power spectrum, and we will use the coupling coefficients to estimate the covariance matrix of the power spectrum, as we will see in Chapter 4. The pseudo-power spectrum $\tilde{C}_\ell^{\mathrm{XY}}$ is obtained by squaring the measured pseudo-harmonic coefficients $\tilde{a}_{\ell m}^{\mathrm{X}}$, $\mathrm{X}, \mathrm{Y} \in [\mathrm{T}, \mathrm{E}, \mathrm{B}]$ with the same multipole $\ell$, with

$$\tilde{C}_\ell^{\mathrm{XY}} = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \left| \tilde{a}_{\ell m}^{\mathrm{X}} \tilde{a}_{\ell m}^{\mathrm{Y}*} \right| . \tag{3.44}$$

Starting from Eq. (3.41), we relate the ensemble average of the pseudo-power spectrum to the underlying power spectrum using the `MASTER` mode-coupling kernel with[1]

$$\langle \tilde{C}^{\mathrm{TT}} \rangle = {}_0M * C^{\mathrm{TT}}, \qquad {}_0M_{\ell\ell'} \equiv (2\ell'+1)\Xi^{00}_{\ell\ell'}[W], \tag{3.45}$$

$$\langle \tilde{C}^{\mathrm{EE}} + \tilde{C}^{\mathrm{BB}} \rangle = {}_{+2}M * \left( C^{\mathrm{EE}} + C^{\mathrm{BB}} \right), \qquad {}_{+2}M_{\ell\ell'} \equiv (2\ell'+1)\Xi^{22}_{\ell\ell'}[W], \tag{3.46}$$

$$\langle \tilde{C}^{\mathrm{EE}} - \tilde{C}^{\mathrm{BB}} - 2i\tilde{C}^{\mathrm{EB}} \rangle = {}_{-2}M * \left( C^{\mathrm{EE}} - C^{\mathrm{BB}} - 2iC^{\mathrm{EB}} \right), \qquad {}_{-2}M_{\ell\ell'} \equiv (2\ell'+1)\Xi^{2-2}_{\ell\ell'}[W], \tag{3.47}$$

$$\langle \tilde{C}^{\mathrm{TE}} \pm i\tilde{C}^{\mathrm{TB}} \rangle = {}_{\times}M * \left( C^{\mathrm{TE}} \pm iC^{\mathrm{TB}} \right), \qquad {}_{\times}M_{\ell\ell'} \equiv (2\ell'+1)\Xi^{20}_{\ell\ell'}[W]. \tag{3.48}$$

In the full-sky case, we recover ${}_sM = \mathrm{Id}$. Looking at only the imaginary part of Eqs. (3.46) and (3.47), we have the following relation for the EB and TB cross-spectra,

$$\langle \tilde{C}^{\mathrm{EB}} \rangle = {}_{-2}M * C^{\mathrm{EB}} \quad \text{and} \quad \langle \tilde{C}^{\mathrm{TB}} \rangle = {}_{\times}M * C^{\mathrm{TB}}. \tag{3.49}$$

In other words, if the parity is preserved in the mean for the overall CMB, it will also be for the pseudo-power spectrum.

**Mask and its power spectrum** The mask determines the survey footprint but also the cuts to the foreground components, namely the point sources and clusters for the SPT-3G experiment, as detailed in Section 2.1.4. The `MASTER` mode-coupling kernels are defined from the mask, and determine the properties of the estimator, as it gives the bias on the estimated power spectrum but also will determine the structure of its covariance matrix, as discussed in Section 4.2. Eq. (3.33) puts forward that the mask power spectrum determines the shape of the kernel, and is further detailed in Louis et al. (2020). The mode-coupling kernel is close to Toeplitz, which is the expected matrix structure from the convolution theorem. To reduce the long-range mix between multipoles, we apodize the mask, as described in Section 2.3.2. The border mask and the border and point source mask are displayed in Fig. 2.27, and associated spectra are displayed in Fig. 3.2. While the low-$\ell$ features of the power spectrum are dictated by the border mask, the high-$\ell$ features are dominated by the small scale feature in the mask and thus the point source holes. The two peaks in the power spectrum of the border and point source mask originate in the typical distance separating point sources and the typical hole size. The power spectrum measured from the border and point source mask is then expected to be correlated between small and large scales due to this feature. This is further discussed in Section 4.2.

### 3.2.2 Debiasing the pseudo-power spectrum

In Eqs. (3.45) to (3.48), we highlight that the pseudo-power spectrum, obtained from the pseudo-harmonic coefficients defined in Eq. (3.41), is a biased estimator of the underlying power spectrum. This bias arises due to the lack of independence between the pseudo-harmonic coefficients. Despite this bias, the pseudo-power spectrum contains crucial information about the underlying power spectrum and can be utilized for parameter estimation, particularly when employing a likelihood-based approach with window functions. Correcting for the bias is challenging, especially when the `MASTER` mode-coupling kernel is not invertible due to a small observed sky patch. On one hand, the *Planck* survey, although limited by galactic cuts, managed to reconstruct all temperature modes up to $\ell_{\max} = 2500$ as they were all probed on the survey footprint. On the other hand, in the case of the SPT-3G experiment, the maximum angular separation between observed pixels is approximately 100 deg, leading to the inability to probe all large-scale modes. Consequently, alternative methods are required to de-bias the pseudo-power spectrum. A common practice is to bin across multipoles, enabling the inversion of the mode-coupling kernels, as we reconstruct only band-powers and not single multipoles. In this work, we employ the `PolSpice` software[2] (Szapudi et al., 2001; Chon et al., 2004), which utilizes stable and accurate real space operations to convert the pseudo-power spectrum into a regularized estimator.

---

[1] As a reminder, $*$ denotes the matrix multiplication.

[2] http://www2.iap.fr/users/hivon/software/PolSpice/

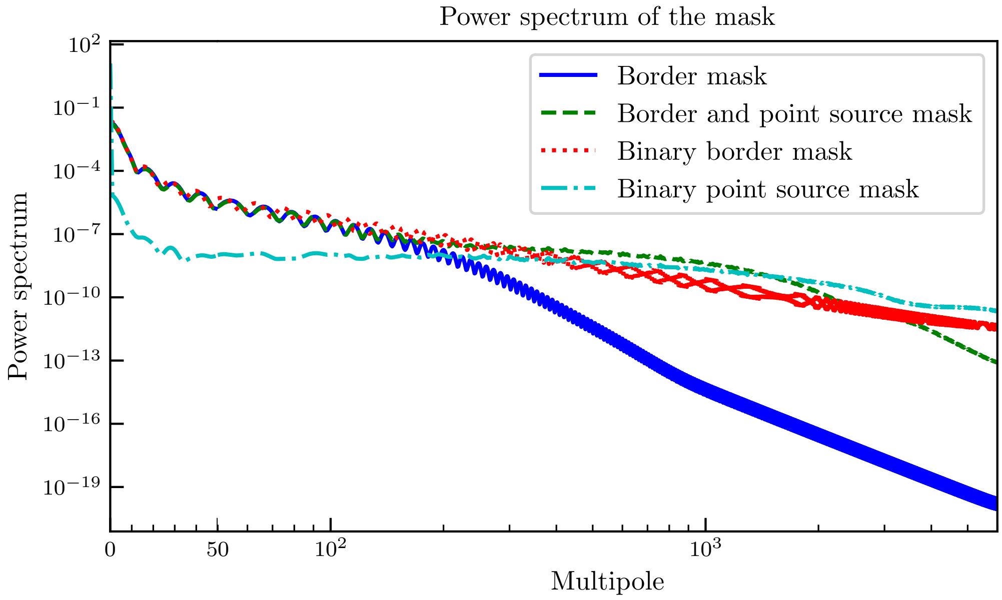


Figure 3.2: Power spectrum of the masks displayed in Fig. 2.27. They are computed as $\mathbb{H}(\mathbb{SHT}\cdot W, \mathbb{SHT}\cdot W)$. The scale is linear up to $\ell = 50$ and logarithmic thereafter. While apodization helps reduce the small-scale power, point source masking has the opposite effect. The large-scale behavior of the power spectrum is determined by the border mask.

**`MASTER` relations in real space** In Section 3.1.3, we have discussed the translation of relations from harmonic space to real space. The advantage of performing calculations in real space is that multiplications in this domain are equivalent to matrix multiplications in harmonic space, resulting in more efficient computations. To this end, we define the real space correlation functions $\tilde{\xi}_s = \mathbb{L}^{-1} \cdot \tilde{C}$ and express the `MASTER` relations in real space as:

$$\langle\tilde{\xi}_0\rangle = w\xi_0, \quad \langle\tilde{\xi}_+\rangle = w\xi_+, \quad \langle\tilde{\xi}_-\rangle = w\xi_-, \quad \langle\tilde{\xi}_\times\rangle = w\xi_\times. \tag{3.50}$$

Here, $w$ represents the mask angular correlation function, and the $\xi_s$ correlation function captures the true underlying properties of the field. The mask angular correlation function can be directly computed from the mask or obtained from the mask power spectrum through Legendre decomposition. It is important to note that this relation is exact and does not rely on any approximation. These relations provide a means to debias the pseudo-power spectrum by computing the real space correlation functions and dividing them by the mask angular correlation function. However, it should be noted that this approach is only valid for angular separations where the mask correlation function does not vanish, i.e., $w(\theta) \neq 0$. It is worth mentioning that the mask correlation function becomes zero for angular separations larger than the maximum angular separation between two pixels in the map. The presence of vanishing eigenvalues in ${}_sM$ directly relates to the angular separations where the mask correlation function also vanishes. To overcome this issue, we employ regularization techniques to stabilize the vanishing elements in real space

**`PolSpice` regularization** In line with the `PolSpice` regularization method proposed by Szapudi et al. (2001), we introduce a smoothed real space function, denoted as $\hat{\xi}$, which is corrected for the mask bias and simultaneously was smoothed out from the angular scales that are not probed by the mask. This smoothing is achieved by reducing the contribution of angular scales beyond the maximum angular separation between two pixels in the map. To accomplish this, we define the scalar apodizing function $f^{\rm apo}$, which transitions from 1 to 0 over the angular scale $\theta_{\rm max}$. The function $g$ is then defined as the ratio between the apodizing function and the mask angular correlation function. The smoothed correlation function in temperature, denoted as $\hat{\xi}_0$, is given by:

$$\hat{\xi} \equiv g\tilde{\xi}, \quad \text{with} \quad g(\theta) = \begin{cases} f^{\rm apo}(\theta)/w(\theta), & \forall\theta \in [0, \theta_{\rm max}] \\ 0, & \forall\theta \in [\theta_{\rm max}, \pi] \end{cases} \tag{3.51}$$

As a result of this operation, we can write the average of the smoothed temperature correlation function as:

$$\langle\hat{\xi}_0\rangle = f^{\rm apo}\xi_0. \tag{3.52}$$

This equation can be transformed back to harmonic space using Eq. (3.31) to obtain the average of the `PolSpice` estimator

$$\hat{C}^{\rm TT} \equiv \mathbb{L}_0 \cdot \hat{\xi}_0, \quad \Longrightarrow \langle\hat{C}^{\rm TT}\rangle = {}_0K * C^{\rm TT} \quad \text{with} \quad {}_0K_{\ell\ell'} = (2\ell'+1)\Xi^{00}_{\ell\ell'}[f^{\rm apo}]. \tag{3.53}$$

The `PolSpice` estimator is biased, however as $f^{\mathrm{apo}}(0) = 1$, it is properly normalized, *i.e.* $\sum_{\ell'} {}_0K_{\ell\ell'} = 0$, so the direct comparison of the measured estimator with the theory spectra does not need to be corrected for the normalization. In addition, if the apodizing kernel is large enough, the `PolSpice` kernel $K$ is localized in $\ell$-space and has a similar effect as a binning kernel on the theory spectra. Concerning polarization, the same procedure can be applied. However, as indicated in Eqs. (3.46) to (3.48), the polarized pseudo-power spectrum has an expected value that depends on both EE and BB underlying spectra, with

$$\langle \tilde{C}^{\mathrm{EE}} \rangle = \frac{1}{2} \left( {}_2M * (C^{\mathrm{EE}} + C^{\mathrm{BB}}) + {}_{-2}M * (C^{\mathrm{EE}} - C^{\mathrm{BB}}) \right). \tag{3.54}$$

E and B modes get coupled through the mask, as they are non-local quantities in real space, and thus cannot be separated correctly if we only have access to a partial zone of the sky. To obtain pure EE and BB spectra, we would need to invert the spin-2 mode-coupling matrices as in Challinor and Chon (2004). However, as we discussed earlier, the mode-coupling matrices are not invertible due to the presence of vanishing eigenvalues. To circumvent this issue, `PolSpice` uses non-local integrations in real space to obtain a decoupled biased version of the EE and BB spectra, exploiting relations between reduced Wigner-$D$ matrices $d_{22}$ and $d_{2,-2}$. Starting from $\tilde{\xi}_+$, we can define the following real space function:

$$\hat{\xi}_+(\theta) \equiv f^{\mathrm{apo}}(\theta) \int \mathrm{d}\cos\theta' \frac{\tilde{\xi}_+(\theta')}{w(\theta)} \sum_{\ell} \frac{2\ell+1}{2} d^{\ell}_{2-2}(\theta) d^{\ell}_{22}(\theta'), \tag{3.55}$$

and further show that this integration can be performed only on the range $\theta \in [0, \theta_{\mathrm{max}}]$ without introducing any bias while already correcting for the effect of the mask. Another real space function can be defined as $\hat{\xi}_- = g\tilde{\xi}_-$. The `PolSpice` estimator for the EE and BB spectra are then given by $\hat{C}^{\mathrm{EE}} \pm \hat{C}^{\mathrm{BB}} = \mathbb{L} \cdot \hat{\xi}_\pm$ and follow

$$\langle \hat{C}^{\mathrm{EE}} \rangle = {}_{-2}K * C^{\mathrm{EE}}, \quad \langle \hat{C}^{\mathrm{BB}} \rangle = {}_{-2}K * C^{\mathrm{BB}}, \quad \text{with} \quad {}_{\pm 2}K_{\ell\ell'} = (2\ell'+1)\Xi^{2,\pm 2}_{\ell\ell'}[f^{\mathrm{apo}}]. \tag{3.56}$$

While ${}_{+2}K$ is normalized ($\sum_{\ell'} {}_{+2}K_{\ell\ell'} = 1$), ${}_{-2}K$ is not (see Fig. 3.4). Indeed, as spin $+2$ is associated with the sum of EE and BB spectrum, it relates to the power conservation in polarization, while the spin $-2$ applies to the difference of EE and BB, thus indicating our ability to distinguish the two separate modes. When decoupling, we get rid of the mixing by discarding the ambiguous modes, eventually leading to a loss of power. Pushing further, the same operation can be run for the TE estimator, which is given by $\hat{C}^{\mathrm{TE}} = \mathbb{L} \cdot \hat{\xi}_\times$ and follows

$$\langle \hat{C}^{\mathrm{TE}} \rangle = {}_\times K * C^{\mathrm{TE}}, \quad \text{with} \quad {}_\times K_{\ell\ell'} = (2\ell'+1)\Xi^{2,0}_{\ell\ell'}[f^{\mathrm{apo}}]. \tag{3.57}$$

`PolSpice` not only provides a way to correct for the mask bias but also allows us to decouple the polarization estimator correctly while efficiently operating in real space. The shape of the `PolSpice` kernel depends on the shape of the apodization function $f^{\mathrm{apo}}$, namely its width $\theta_{\mathrm{max}}$, and the choice of the apodization function itself. In the following, we will use a cosine apodization function with a width of $\theta_{\mathrm{max}} = 30\,\mathrm{deg}$, given by

$$f^{\mathrm{apo}}(\theta) = \begin{cases} \frac{1}{2}\left(1 + \cos\left(\frac{\pi\theta}{\theta_{\mathrm{max}}}\right)\right), & \forall\theta \in [0, \theta_{\mathrm{max}}] \\ 0, \quad \forall\theta \in [\theta_{\mathrm{max}}, \pi]\,. \end{cases} \tag{3.58}$$

The apodizing function and associated kernels are illustrated in Fig. 3.3. A shallow apodization function gives a narrow diagonal-like kernel. The `PolSpice` kernels will be a building block of later window functions for the likelihood analysis. In the following, I will demonstrate that one can relate the `PolSpice` spectrum estimator to the pseudo-spectrum estimator $\tilde{C}$. Taking only the case of temperature, Eq. (3.51) can be translated in harmonic space as

$${}_0\hat{\xi} = g_0\tilde{\xi} \to \hat{C} = {}_0G * \tilde{C}, \quad \text{with} \quad {}_0G_{\ell\ell'} = (2\ell'+1)\Xi^{00}_{\ell\ell'}[g] \tag{3.59}$$

The decoupling kernel ${}_0G$ is equivalent to a pseudo-inversion of the mask mode-coupling matrix ${}_0M$. Were the mode-coupling matrix invertible, we would have ${}_0G = {}_0K * {}_0M^{-1}$. For the polarization case, due to the decoupling operation, the expression of the decoupling kernels is more complicated. I can however show that the decoupling kernels are given by

$${}_{-2}G_{\ell\ell'} = (2\ell'+1)\Xi^{2-2}_{\ell\ell'}[g],\ {}_\times G_{\ell\ell'} = (2\ell'+1)\Xi^{20}_{\ell\ell'}[g],\ {}_{+2}G_{\ell\ell'} = \frac{2\ell'+1}{2}\int \mathrm{d}\cos\theta\, g(\theta) d^{\ell}_{22}(\theta) d^{\ell'}_{2-2}(\theta). \tag{3.60}$$

While the $\Xi$ operator is efficient to compute on arbitrary correlation function, the non-orthogonality of the reduced Wigner-$d$ matrices makes the computation of the decoupling kernel ${}_{+2}G$ more computationally expensive but can be done efficiently using the `PolSpice` code as an emulator. This is a direct consequence of the polarization decoupling procedure of `PolSpice`. I can then show that

$$\hat{C}^{\mathrm{EE}} \pm \hat{C}^{\mathrm{BB}} = {}_{\pm 2}G * \left(\tilde{C}^{\mathrm{EE}} \pm \tilde{C}^{\mathrm{BB}}\right), \quad \text{and} \quad \hat{C}^{\mathrm{TE}} = {}_\times G * \tilde{C}^{\mathrm{TE}}. \tag{3.61}$$

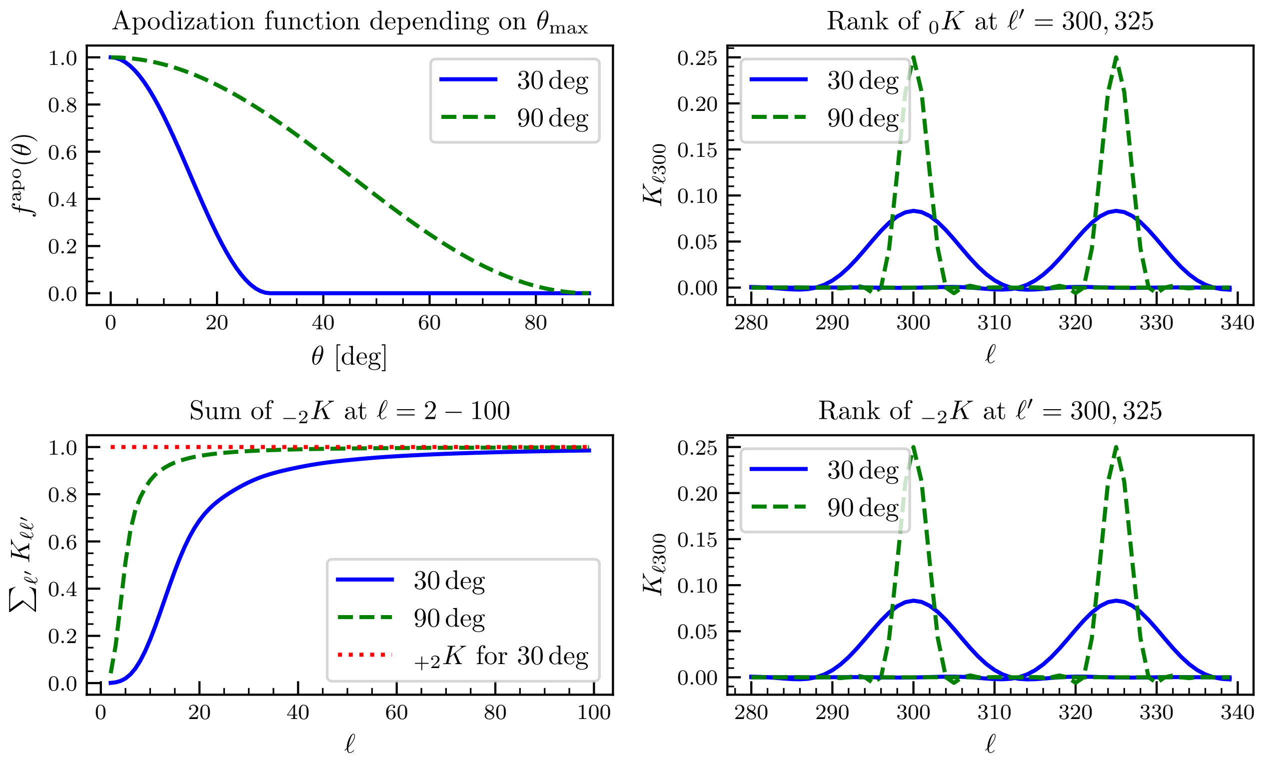


Figure 3.3: Illustration of the apodization function and associated kernels. Top left: apodization function $f^{\rm apo}$. Top right: `PolSpice` kernel ${}_0K$. Bottom left: Sum over rows of ${}_{-2}K$. Bottom right: `PolSpice` kernel ${}_{+2}K$. The shape of the apodization function determines the width and normalization of the `PolSpice` kernel.

The kernels are illustrated in Fig. 3.4, with on the left panel the apodized function $g = f^{\rm apo}/w$ and on the right panel the decoupling kernels. The decoupling kernels are localized around the diagonal, implying that the decoupling operation is quasi-local in harmonic space. I will use the decoupling kernels to build the window functions for the likelihood analysis and to compute the spectrum covariance. In the next section, I will show how to de-bias the band powers from instrumental effects. From now on, we will drop the notations $*, \times$ that differentiate the scalar and matrix multiplication, and assume that only matrix multiplications are performed when not specified otherwise.

## 3.3 Transfer function and beam deconvolution

The filtering procedure applied to construct the maps in Chapter 2, specifically discussed in Section 2.1.4, results in a filtered map that possesses different statistical properties compared to the underlying CMB signal. In particular, due to the removal of iso-latitude large-scale modes, the power spectrum of the filtered map becomes biased. We make the reasonable assumption that the filtering procedure modifies the harmonic coefficients with an effective 2D transfer function denoted as $f$, such that

$$a_{\ell m} \to f_{\ell m} a_{\ell m}. \tag{3.62}$$

This assumption holds as most of the filtering is performed in Fourier space, where the full sky modes are not coupled. The measured modes are coupled due to the masking, see Eq. (3.41). The transfer function can be determined either using the mock observations introduced in Section 2.3.4 or with a faster-emulated version of the mock observations. In the latter approach, each ring of a generated `HEALPix` map is filtered using an equivalent procedure as described in Section 2.1.4. While this method is efficient and allows for testing different filtering configurations, it does not encompass all the map-making procedures as comprehensively as the full suite of mock observations. The harmonic coefficients averaged from 30 mock observations, denoted as $f_{\ell m}$, are displayed in Fig. 3.5. As expected, the imaginary part cancels out since it is related to the localization of the associated mode on the sphere. The transfer functions exhibit consistency across frequencies, which is to be expected as the filtering procedure is independent of frequency. The significant suppression observed for large $m/\ell$ values is a consequence of the survey mask. The SPT-3G survey footprint is situated low in the sky, and as a result, ${}_0Y_{\ell m}$ modes with $m/\ell \sim 1$ are located close to the equator. Therefore, these modes are typically not probed by the survey footprint, leading to a sharp cutoff in the transfer function. On the other hand,

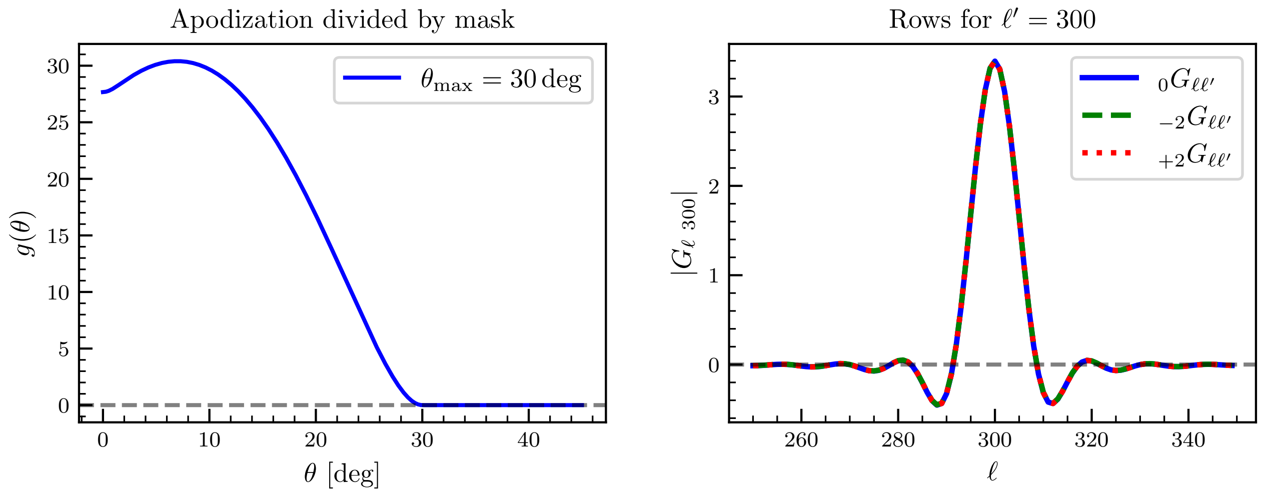


Figure 3.4: Left: Function $g = f^{\mathrm{apo}}/w$ where $w$ is the two-point correlation function of the mask. At null angular separations, $g(0)$ indicates the inverse of the sky fraction. Increasing $\theta$ is first associated with a rise in $g$ due to the mask contribution, further damped at larger angular separations thanks to the apodization function and reaching zero at $\theta_{\max} = 30\,\mathrm{deg}$. Right: Row $\ell' = 300$ of associated decoupling kernels. The decoupling kernels are localized around the diagonal. The behavior of $g$ determines the shape of $G$, which relates the pseudo-power spectrum estimator to the `PolSpice` one.

spherical harmonics with low-order $m$ are suppressed by the filtering strategy since they correspond to modes with large-scale variations in the scanning direction, which are filtered out by the high-pass filter employed during scan filtering. A low-$m$ real spherical harmonic is illustrated in the left panel of Fig. 3.6, while a $m/\ell \sim 1$ real spherical harmonic is depicted in the right panel, demonstrating the delineation of the survey footprint in lime green. For low $m$, the transfer function exhibits nearly linear growth from 0 to 1 between $m = 100$ and $m = 200$. Across the rest of the $\ell, m$ range, the transfer function remains close to unity, as the filtering procedure has minimal impact on these modes. It is important to note that the estimation of the transfer function shows statistical deviations, even when measured on a noiseless sky. This arises because the transfer function is estimated from a finite number of mock observations, making it susceptible to statistical fluctuations. To reduce the statistical uncertainty and facilitate error propagation from the transfer function to the power spectrum, we will characterize the transfer function further using a larger number of mock observations.

**Power spectrum transfer function** As put forward in Hivon et al. (2002), the transfer function is used to de-bias the power spectrum measured on the maps. The debiased power spectrum is given by $\hat{C}_\ell^{\mathrm{u}} = \hat{C}_\ell / F_\ell$ where $F_\ell$ is the transfer function. I use mock observations to measure the transfer function representing the effect of the map-making on the measured power spectrum. The AGORA software is used to produce input skies in all frequency channels, of which I compute the `PolSpice` power spectrum estimator $\hat{C}_{\ell;i}^{\nu\times\mu;\mathrm{input}}$. The same estimator is applied to the output of the mock observation pipeline, to obtain $\hat{C}_{\ell;i}^{\nu\times\mu;\mathrm{output}}$. The transfer function is estimated through

$$F_\ell^{\nu\times\mu;\mathrm{est}} \equiv \frac{\sum_i \hat{C}_{\ell;i}^{\nu\times\mu;\mathrm{output}}}{\sum_i \hat{C}_{\ell;i}^{\nu\times\mu;\mathrm{input}}}. \tag{3.63}$$

The averaged $F_\ell$ estimator shows statistical fluctuations, which are reduced by increasing the number of mock observations and the error needs to be propagated to the covariance matrix of the power spectrum.

**Alternative pipeline** From the filtering settings, a quick and efficient analytical pipeline has been developed by a collaborator to account for the effect of filtering on the signal. This pipeline outputs the transfer function $F_\ell^{\mathrm{ana};\mu\times\nu}$. This work will be published in a scientific journal (Hivon, Doussot, Camphuis, et al., *in prep.*). In Fig. 3.7 are displayed the transfer functions measured from the mock observations and obtained with this emulated pipeline. The two transfer functions are consistent. Their remarkable agreement demonstrates that we understand precisely the impact of the filtering on the signal.

**Beam deconvolution** In Section 2.2.4, we described the effect of the beam on the maps by a convolution with a quasi-axisymmetric Gaussian beam of frequency-dependent width. We deconvolve the effect of the beam

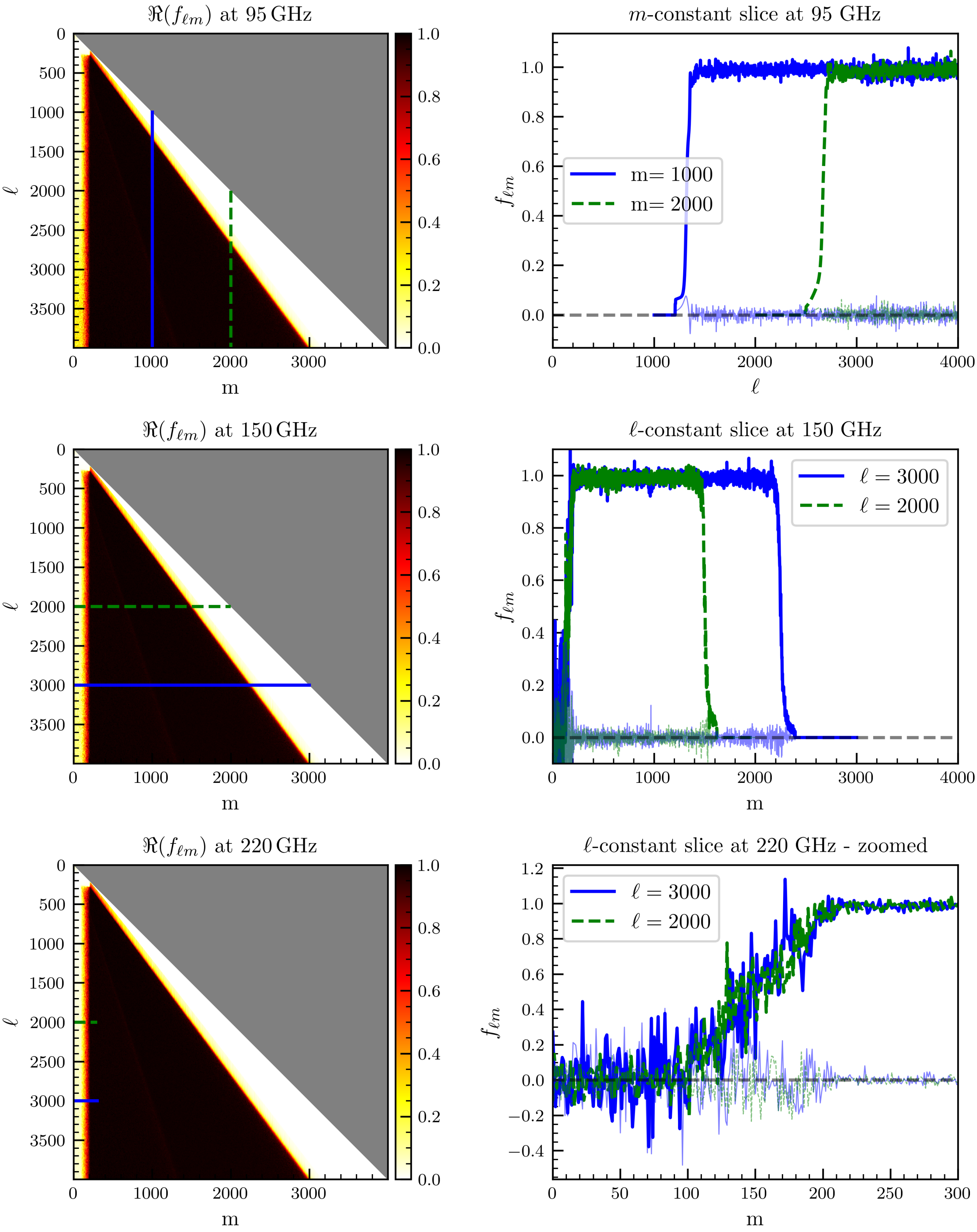


Figure 3.5: Left: 2D $f_{\ell m}$ transfer function for different frequencies. Right: Slices of $f_{\ell m}$ for constant $m$ at 95GHz (top), for constant $\ell$ at 150GHz (middle) and for constant $\ell$ at 220GHz (bottom). Real parts are plotted as thick lines while imaginary parts are plotted as thin lines. The transfer functions are plotted only for the temperature but are similar for the $E$ polarization. The 2D transfer function indicates which $\ell m$ modes are suppressed due to filtering at low $m$, as the high $m$ suppression is due to the real space masking.

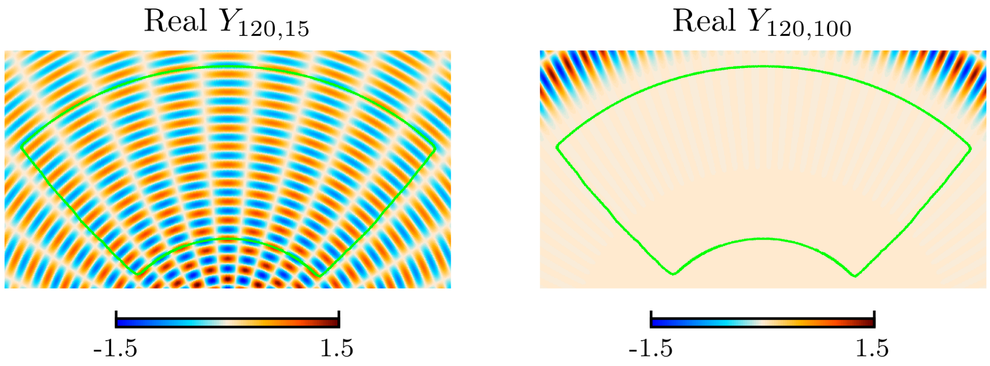


Figure 3.6: Real spherical harmonics ${}_0Y_{\ell m}$ for $\ell, m = 120, 15$ (left) and $\ell, m = 120, 100$ (right). While low-order $m$ spherical harmonics show large-scale components in the scan direction, high-order $m$ spherical harmonics are localized on the equator, hence justifying the structure of the 2D transfer function displayed in Fig. 3.5. Maps are displayed in Lambert's azimuthal equal-area projection.

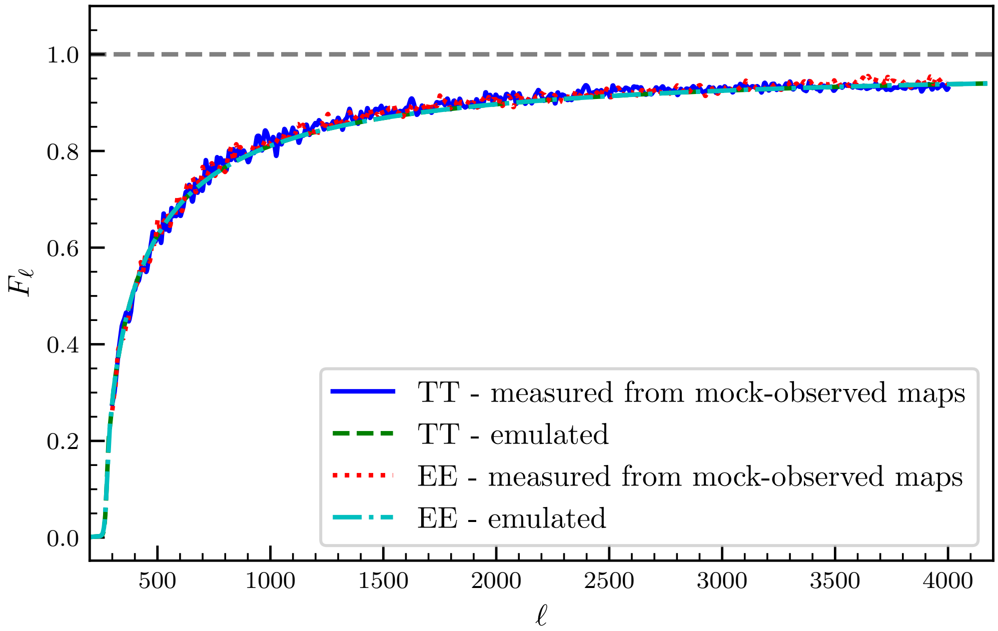


Figure 3.7: Transfer function $F_\ell^{220\times220;\mathrm{est}}$ measured from 30 mock-observations and $F_\ell^{220\times220;\mathrm{ana}}$ from the analytical pipeline in temperature and polarization. The power is suppressed on all scales due to the filtering, especially on large scales, which are filtered to remove the atmosphere contribution.

by dividing the harmonic coefficients by the beam window function

$$B_\ell^\nu \equiv \mathbb{L} \cdot B^\nu(\theta), \tag{3.64}$$

where $B^\nu(\theta)$ is the beam profile at frequency $\nu$ and $\mathbb{L}$ is the Legendre transform. The beam transfer function is displayed in Fig. 2.13. We also compute the pixel-window function $\mathcal{P}$, which depends on the field spin.

## 3.4 Noise spectra

The noise maps are essential data products generated by the SPT-3G map-making pipeline. These maps play a crucial role in characterizing the noise properties of the resulting maps and modeling the noise in the power spectrum estimation. To obtain the noise maps, a series of steps described in Section 2.3.3 are followed. Then, I compute the pseudo-harmonic coefficients $\tilde{n}_{\ell m}$ by applying a spherical harmonic transform after masking with a mask including point source masking. Next, I use the `PolSpice` approach to compute the mask-debiased power spectra $N_\ell^{\mu\times\nu}$. In Fig. 3.8, I present the real and imaginary parts of $n_{\ell m}$ measured at 95GHz on a single noise map estimation. Modes $\ell, m < 300$ are all suppressed due to the filtering strategy, while high $m/\ell$ modes are

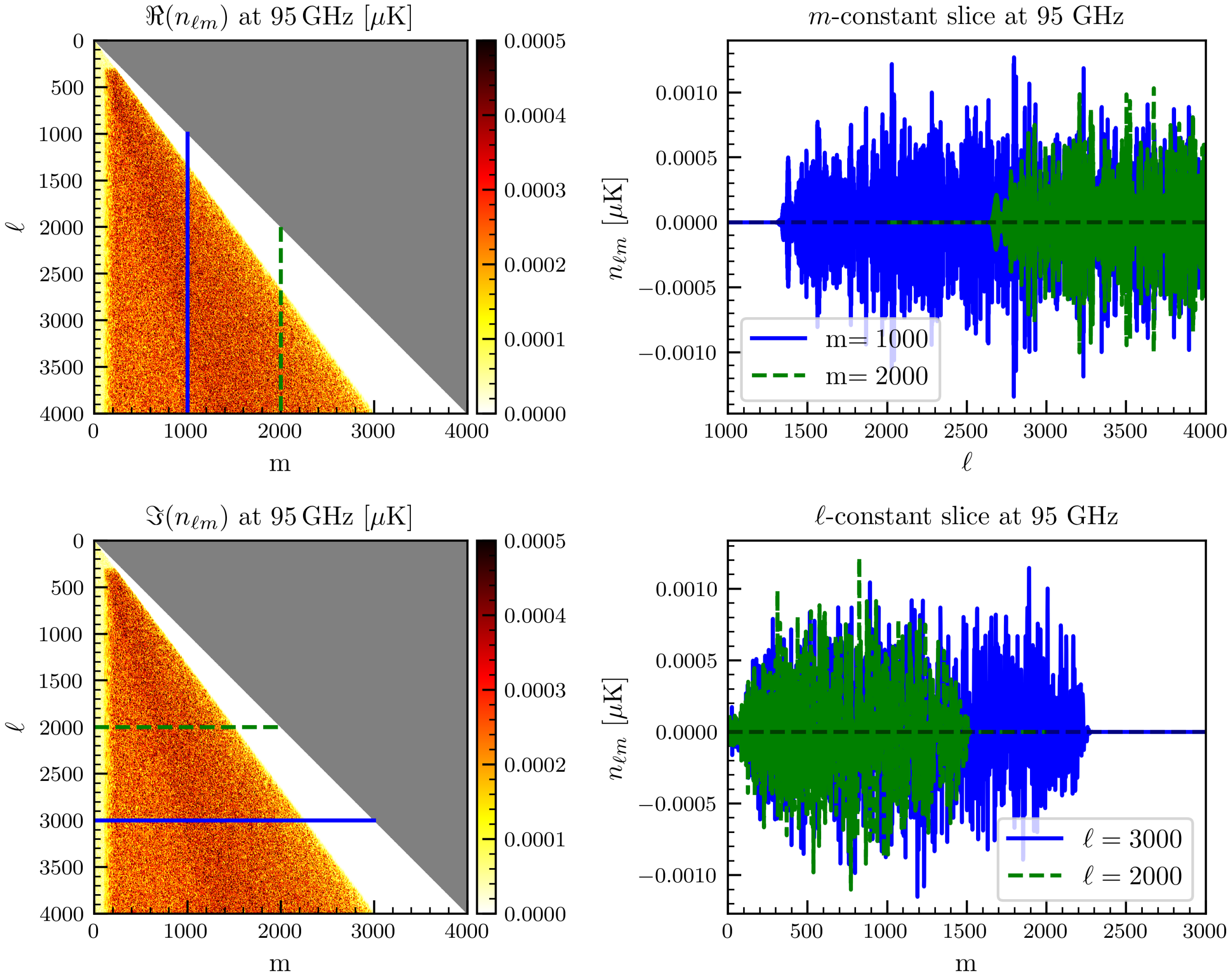


Figure 3.8: On the left, the real and imaginary part of the noise spherical harmonic coefficients $n_{\ell m}$ measured on a single noise map estimation at 95GHz. Blue and green lines correspond to the slice plotted with corresponding colors on the right plot. The noise is uniform on all $\ell m$ modes, suppressed at low $m$ due to the filtering, and at high $m$ due to the real space masking.

suppressed by the masking. The noise coefficients are complex, and the real and imaginary parts have similar power, as expected for a white noise process. The noise power spectra measured and averaged over 63 noise maps estimations are displayed in Fig. 3.9, with *Planck* best fit TT and EE spectrum as indications. At larger scales, the dominant source of temperature noise arises from the $1/f$ contribution caused by atmospheric contamination. However, at smaller scales, the noise is flat, similar to the noise observed in polarization measurements. It is important to note that the experiment noise exhibits correlations across different frequencies. This correlation arises due to the shared use of detectors across frequencies, leading to correlated electronic and line-of-sight atmospheric noise. To ensure accurate analysis, the noise power spectra undergo correction for the transfer function. This correction is particularly significant at larger scales where the transfer function values are small. The excedent large-scale power for $\ell < 300$ is due to residual large-scale noise boosted by the transfer function inversion. It is not significant since those modes are not used in the analysis. The noise power spectra play a crucial role in both the inpainting procedure and the estimation of covariance. In the inpainting procedure, I use these spectra to inform the reconstruction of missing data points. Additionally, in covariance estimation, I use the noise power spectra to accurately model the noise contribution to the covariance matrix.

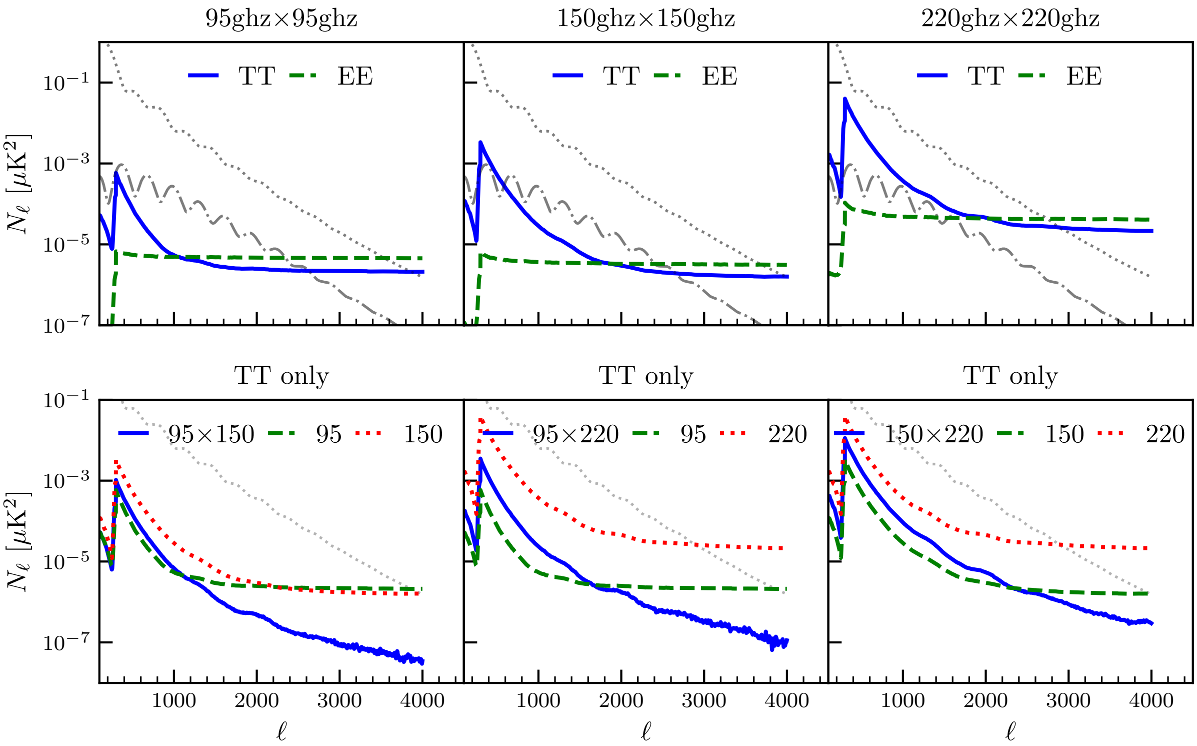


Figure 3.9: Top: Auto-frequency noise power spectra, in TT (blue) and EE (green). Bottom: Cross-frequency noise power spectra in temperature only, plotted against auto-frequency noise power spectra. The power spectra have been measured on 63 noise map estimations, and corrected by the emulated transfer function $F^{\text{ana}}$. The *Planck* best fit TT (dot) and EE spectra (dash-dot) are plotted as grey lines. The noise exhibits its expected behavior of residual $1/f$ at large scales in temperature, and flat white noise at high multipoles. The polarization noise is white and larger than the temperature noise. The temperature noise is correlated across frequencies on the large scales due to the common observed atmosphere, but decorrelates on the small scales. The polarization noise exhibits no significant correlation.

## 3.5 Bundles cross-spectrum analysis

### 3.5.1 Band powers

The map-making procedure outputs a set of $n_{\text{BUN}}$ maps $(T, Q, U)^i$ for $i \in [1, n_{\text{BUN}}]$, see Section 2.2.3, each of them being the CMB and foreground filtered signal at each frequency with different noise realizations. I compute the cross-spectrum between each bundle $i$ and $j$ for every frequency channel as

$$\tilde{C}^{\text{X}i\nu,\text{Y}j\mu} = \mathbb{H}\left[\tilde{a}^{\text{X}i\nu}, \tilde{a}^{\text{Y}j\mu*}\right], \quad \forall i,j \in [1, n_{\text{BUN}}] \text{ s.t. } i \neq j, \forall \nu, \mu \in [95, 150, 220]\,\text{GHz}, \forall \text{X}, \text{Y} \in [\text{T}, \text{E}], \tag{3.65}$$

and remove the mask bias using the `PolSpice` procedure. I also correct the power spectrum using the by-products of the analysis pipeline

$$\hat{C}_\ell^{\mathrm{X}i\nu,\mathrm{Y}j\mu} \to \frac{1}{A^{\mathrm{X}\nu,\mathrm{Y}\mu}} \hat{C}_\ell^{\mathrm{X}i\nu,\mathrm{Y}j\mu}, \tag{3.66}$$

where

$$A_\ell^{\mathrm{X}\nu,\mathrm{Y}\mu} = F_\ell^{\mathrm{X}\nu,\mathrm{Y}\mu} B_\ell^\nu B_\ell^\mu \mathcal{P}_\ell^{\mathrm{X}} \mathcal{P}_\ell^{\mathrm{Y}}. \tag{3.67}$$

In this definition, $F_\ell^{\mathrm{X}\nu,\mathrm{Y}\mu}$ is the transfer function obtained from mock observations as in Eq. (3.63), the beam is obtained as the Legendre decomposition of the measured real-space profile see Eq. (3.64) and $\mathcal{P}$ is the pixel window function obtained with the `HEALPix` software. This operation is done so that we have the following equation

$$\left\langle \hat{C}_\ell^{\mathrm{X}i\nu,\mathrm{Y}j\mu} \right\rangle = C_\ell^{\mathrm{X}\nu,\mathrm{Y}\mu} + \delta_{ij} \frac{N_\ell^{\mathrm{X}\nu,\mathrm{Y}\mu}}{A_\ell^{\mathrm{X}\nu,\mathrm{Y}\mu}}, \tag{3.68}$$

where $C^{X\nu,Y\mu}$ is the underlying power spectrum of the observed CMB anisotropies. The bundles are chosen to have a consistent signal-to-noise ratio, and this explains why the second term of the right-hand side of the previous equation does not depend on the bundle index. For two different bundles, the signal is identical while the noise is uncorrelated. The cross-spectrum approach is motivated to remove the noise bias by taking only the combinations $i \neq j$, see Polenta et al. (2005), and trades bias correction for more variance than the auto-spectrum approach. However, the additional variance is scaling as $n_{\mathrm{BUN}}^{-1}$ and is thus reduced when the number of bundles is large. The number of bundles is chosen so that the noise uniformity is ensured across the sky while $n_{\mathrm{BUN}}$ is maximal, allowing for small additional variance and an accurate estimation of the noise covariance matrix of the band powers, see Section 4.2.2. The cross-spectrum between two distinct bundles has no noise bias, but still has the noise variance as further discussed in Section 4.2. I then obtain the final estimator by averaging the cross bundles spectra defined in Eq. (3.68), as

$$\hat{C}^{\mathrm{X}\nu,\mathrm{Y}\mu} = \frac{1}{n_{\mathrm{BUN}}(n_{\mathrm{BUN}}-1)} \sum_{i,j\neq i} \hat{C}^{\mathrm{X}i\nu,\mathrm{Y}j\mu} \quad \Longrightarrow \quad \left\langle \hat{C}^{X\nu,Y\mu} \right\rangle = C^{X\nu,Y\mu}. \tag{3.69}$$

Before further analysis and to reduce the correlation between multipoles as well as errors on the final estimator, I apply a binning procedure to the power spectrum, by averaging over a width of $\Delta\ell = 25$. The binning operator is defined as

$$P_{b\ell} \propto \frac{1}{2\pi} \frac{\ell(\ell+1)}{\Delta\ell} \text{ if } \ell \in [\ell_{b-1}, \ell_b] \text{ and 0 otherwise,} \quad \text{with} \sum_\ell P_{b\ell} = 1 \text{ and } \hat{\mathcal{C}}_b = \sum_\ell P_{b\ell} \hat{C}_\ell. \tag{3.70}$$

The renormalization by $\ell(\ell+1)$ is applied to flatten the power spectrum, and corresponds to having uniform map variance per multipole. The band powers $\hat{\mathcal{C}}$ are the main observable of our analysis. They are the result of the data processing pipeline and are the input of the likelihood function. A preliminary version is displayed in Fig. 3.22, with conservative multipole cuts $\ell_{\min} = 300$ and $\ell_{\max} = 4000$. This version is obtained using the apodized mask including source masking, and will be updated after running the Gaussian-constrained realization on the data as described in the next chapter. Both binning scheme and multipole cuts will be improved after completing null tests and consistency checks, see Section 4.3.

## 3.6 Gaussian constrained realization

The microwave sky is populated by numerous point sources, which not only pose challenges for accurate power spectrum estimation but also affect the reconstruction of the CMB temperature map by introducing significant flux contamination in the time stream. To mitigate the impact of these point sources, they are masked out during the map-making procedure, see Section 2.2. However, this leaves us with a masked map that is unsuitable for accurate power spectrum covariance estimation. Although the MASTER debiasing procedure introduced in Section 3.2.2 can potentially be applied to any mask, it is important to note that the statistical properties of the estimator are dependent on the specific mask used. As discussed further in Chapter 4, the variance of the power spectrum estimation is influenced by the mask. We demonstrate in Section 3.6.1 that the variance of the power spectrum is larger if measured with a mask including point source masking as opposed to pure border masking. This discrepancy arises from the fact that modes get mixed when analyzed through a mask with holes, as large-scale modes are effectively seen as small-scale ones. Moreover, it is challenging to compute analytically this mixing at the covariance level and standard approaches for covariance computation fail to estimate the covariance. A way to solve this problem is to fill in the masked region with a physically motivated model of the CMB, and this is the technique I have decided to perform on the SPT-3G data. Let me emphasize that I obtained all the results mentioned in this chapter and that this work will be submitted for publication (Camphuis, Benabed, et al., *in prep*).

Inpainting techniques for reconstructing missing regions of the CMB sky have been extensively studied in the literature as presented in this paragraph. Simply masking the sky by assigning zero values to contaminated pixels results in a loss of information and reduced statistical power. Diffusive inpainting, which fills in missing pixels with the average values of their neighbors, has been proposed as an alternative method (Zacchei et al., 2011). However, this approach assumes that the signal is smooth at the scale of the inpainting region, which is not valid in the presence of numerous sources and may introduce biases at the power spectrum level. Hybrid approaches that reconstruct a biased and coupled field have also been proposed (Plaszczynski et al., 2012). Techniques specifically designed to preserve the Gaussian properties of the CMB field have been developed (De Oliveira-Costa and Tegmark, 2006; Bucher and Louis, 2012), and advanced methods have been proposed in idealized scenarios (Ramanah et al., 2017). In this study, we focus on the Gaussian constrained realization method originally proposed for temperature map reconstruction by Lavaux and Wandelt (2010); Benoit-Lévy et al. (2013) and extend it to polarization. This approach is particularly advantageous for our analysis as we assume the Gaussianity of the CMB, allowing us to utilize the CMB two-point correlation to constrain the inpainting procedure. One limitation of this approach is the assumption of perfect knowledge of the underlying CMB power spectrum. However, we will demonstrate that our procedure remains robust even in the presence of reasonable deviations from the true power spectrum.

### 3.6.1 Motivations

To perform power spectrum likelihood analysis, the use of a covariance matrix is necessary to estimate the variance and coupling between band powers. In Chapter 4, I will provide a detailed explanation of how the covariance is analytically approximated for the analysis of the 19/20 data. However, major complications arise when dealing with masked sources in the field.

Firstly, the presence of small-scale features introduces additional variance and coupling between the small scales of the band power. Holes in the mask result in an increased power at high multipoles of the mask power spectrum, as indicated in Fig. 3.2. This additional power leads to a mixing of the large and small scales of the signal when computing the signal pseudo-power spectrum, resulting in additional variance and coupling between band powers. This can be understood from Eq. (3.45), as the mask power spectrum is the only ingredient of the `MASTER` mode-coupling matrix. As a consequence, the small multipoles of the power spectrum impact the shape of the covariance. I compute simulated covariances using 1000 Monte Carlo simulations generated from the *Planck* 2018 best fit, with three separate cases : (1) masking the sources for power spectrum computation, (2) cutting the power spectrum at $\ell < 200$ when generating simulations to reduce the contamination and masking the sources for power spectrum computation and finally (3) not masking the sources. The masks used are displayed in Figs. 2.26 and 2.27. In Fig. 3.10, I plot on the left panel the diagonal of the covariances, and on the right panel the ratio of those over the reference case where the sources are not masked. Masking sources imply additional variance, but this can be mitigated by cutting the low multipoles of the power spectrum. Furthermore, the right panel of Fig. 3.11 displays a comparison of rows of the correlation matrices, which are the covariance matrices rescaled by their diagonals. This plot clearly illustrates the additional coupling between band powers due to holes in the mask. This coupling depends on the low multipole of the power spectrum, as it is greatly reduced when cutting the first multipoles of the power spectrum before generating the sky.

Secondly, the analytical covariance framework developed in the next chapter does not accurately predict the additional variance and coupling. In the left panel of Fig. 3.11, I compare the diagonal of covariances estimated through simulations to the one computed with an analytical approximation. I plot the ratio of covariance diagonals as

$$\Sigma_{\ell\ell}^{\text{simulations}} / \Sigma_{\ell\ell}^{\text{analytical approximation}} - 1. \tag{3.71}$$

In the left panel of Fig. 3.11, the red line compares analytical estimation and simulation in the case of no source masking, showing perfect agreement. Simulations only match analytical prediction in the unmasked case, both for the variance and the coupling in this test case. This failure occurs because the approximated covariance relies on the hypothesis that the power spectrum of the mask does not exhibit long-range power.

The mismatch of simulations to the analytical approximation approach is a well-known problem of covariance computation. I showed that it depends on the low multipole cut, but it also depends on the point source density in the mask. As described in Section 3.3, the low multipoles of the power spectrum are suppressed by filtering in SPT-3G. This suppression mitigates the long-range coupling, as illustrated in Fig. 3.12. I generate 1000 simulations with 1D filtering of the power spectrum using the emulated transfer function $F^{\text{ana}}$, and compare on the left panel of the figure the diagonal of the covariance with masked sources or unmasked sources. On the right panel, I compare the coupling levels by displaying a row of the covariances. While the *Planck* experiment low point source density resulted in minimal contamination of order 10%, the low noise levels of SPT-3G imply that approximately 2500 sources are masked on 4% of the sky, leading to 50% more variance at small scales due to the point sources' long-range leakage. However, additional couplings are suppressed by the low multipole filtering due to map-making for SPT-3G. Nevertheless, we expect point source masking density to increase for future

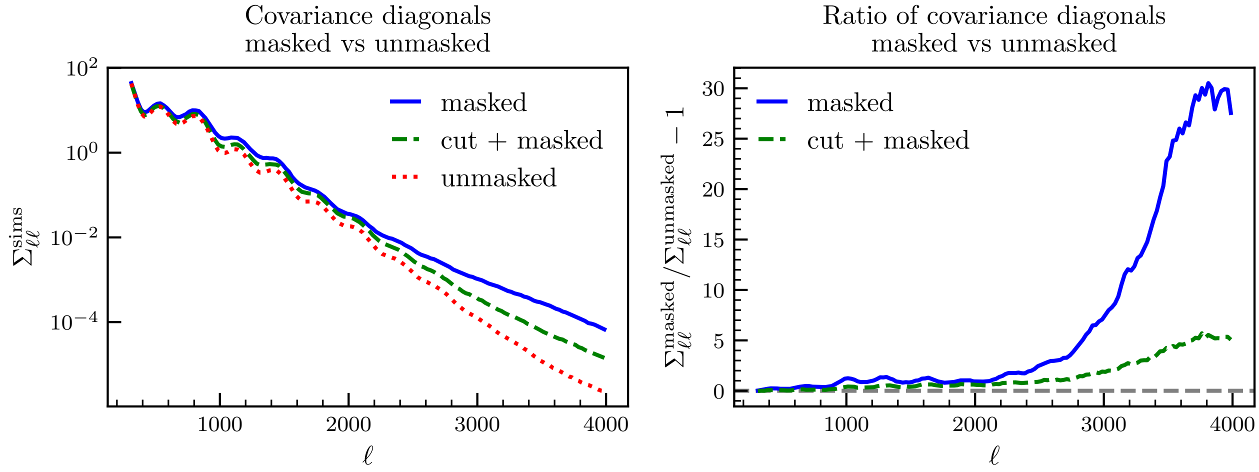


Figure 3.10: Left: diagonals of Monte Carlo covariance in case (1) masked sources, (2) $\ell < 200$ cut and masked sources, or (3) unmasked sources. The covariances were computed using 1000 simulations. Right: ratio of the diagonals to the unmasked case. In the presence of point sources, the power spectrum estimator covariance is boosted at small scales, and this boost depends on the large-scale power in the maps.

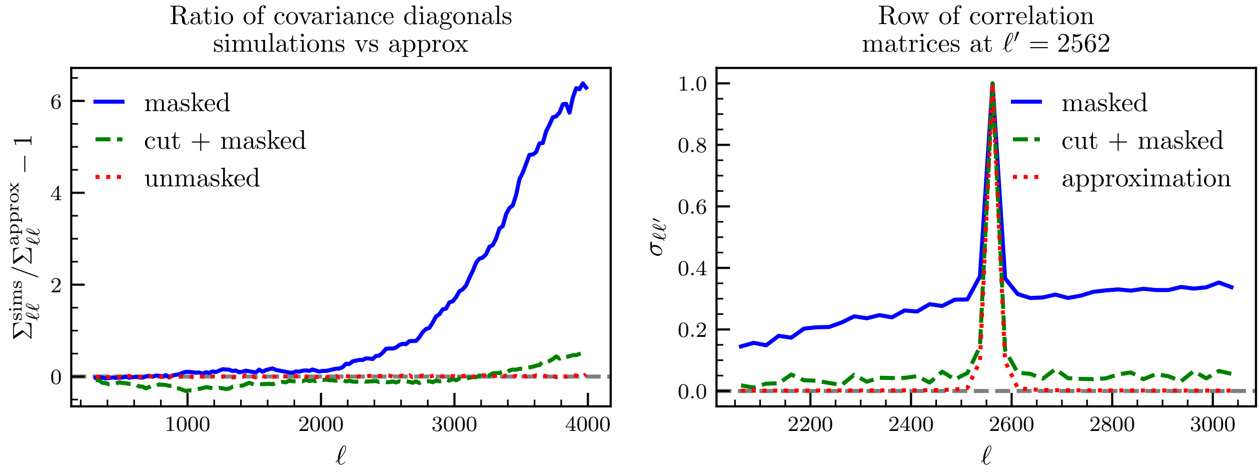


Figure 3.11: Left: ratio of Monte Carlo covariance diagonals to the analytical approximation. Right: Row of the correlation matrices, at $\ell' = 2562$. Additional variance and coupling stem from source masking and are not accurately modeled by approximation.

surveys, and thus additional coupling might emerge. *Planck* relied on simulations to correct the additional variance, but the increase in the case of SPT-3G settings cannot be accurately corrected using simulations. For the SPT-3G pipeline, we will address the issue of holes in the mask by using constrained realization and inpainting techniques, which will be described in the following sections. The advantage is twofold: first, it allows me to use the analytical covariance framework since the holes in the mask are no longer present. Second, the estimator variance will be reduced, thereby improving the final constraints on parameters.

### 3.6.2 Temperature case

In this section, I will detail the ideal Gaussian-constrained realization of the temperature map. I assume a simple Gaussian isotropic signal of known statistical properties, entirely characterized by the power spectrum $C_\ell^{\rm th}$.

Let us consider a dataset comprising of CMB temperature and noise maps, wherein region (1) contains missing data. It is possible to predict the signal in the region (1) based on the data available in a **constraining region** denoted as (2). The CMB temperature data map can be represented as a vector $T^D = \left(T_1^D, T_2^D\right)$, where $T_1^D$ corresponds to the unknown part. The primary objective is to obtain a new inpainted map $T^I$ that possesses the same statistical properties as the original data $T^D$ while incorporating predicted signal inside holes, as the data $T_1^D$ is not accessible. For a given CMB temperature map $T^D = \left(T_1^D, T_2^D\right)$, it is known that

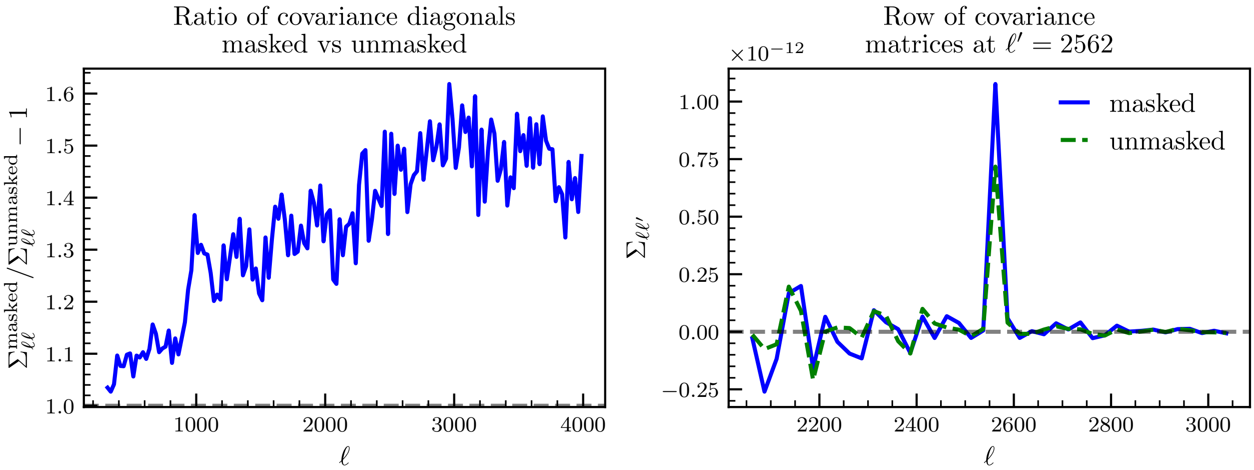


Figure 3.12: Left: ratio of Monte Carlo covariance diagonals for simulations drawn from *Planck* 2018 best fit filtered by SPT-3G 1D transfer function $F^{\rm ana}$. With the realistic SPT-3G simulations, I demonstrate that masking sources leads to 50% more variance in the power spectrum estimator, but no significant additional correlations.

the pixels follow a Gaussian distribution with a known covariance, as depicted in Eq. (3.1)). This covariance can be obtained from the true underlying power spectrum $C_\ell^{\rm th}$ utilizing Eq. (3.22)). Consequently, the joint probability of $(T_1^D, T_2^D)$ can be expressed as:

$$\mathcal{P}(T_1^D, T_2^D) \propto \exp\left(-\frac{1}{2} T^{D\top} \xi^{-1} T^D\right), \quad \xi^{\rm th} = \begin{pmatrix} \xi_{11}^{\rm th} & \xi_{12}^{\rm th} \\ \xi_{21}^{\rm th} & \xi_{22}^{\rm th} \end{pmatrix} \equiv \begin{pmatrix} \langle T_1^D T_1^{D\top} \rangle & \langle T_1^D T_2^{D\top} \rangle \\ \langle T_2^D T_1^{D\top} \rangle & \langle T_2^D T_2^{D\top} \rangle \end{pmatrix}. \tag{3.72}$$

Based on Eq. (3.72)), it can be inferred that the conditional distribution $T_1^D | T_2^D$ follows a Gaussian distribution with a known mean, commonly referred to as the Wiener filter of the data (Wiener, 1964). The variance of this distribution is determined using the Schur complement, resulting in:

$$\mathcal{P}(T_1^D | T_2^D) \sim \mathcal{N}(\xi_{12}^{\rm th} \xi_{22}^{\rm th-1} T_2^D, \xi_{11}^{\rm th} - \xi_{12}^{\rm th} \xi_{22}^{\rm th-1} \xi_{21}^{\rm th}). \tag{3.73}$$

Remarkably, in this expression, only the mean of the field $T_1^D$ depends on the constraints $T_2^D$, while the variance remains unaffected. Exploiting this crucial property, we propose a method for inpainting the missing data in region (1). First, we draw a random CMB map $T^R = \left(T_1^R, T_2^R\right)^\top$ using a given power spectrum $C_\ell^R$, as detailed in Section 3.1.2. We then adjust the average values in the constraining zone according to the constraints provided by $T_2^D$. Utilizing a fiducial spectrum $C^{\rm fid}$, we construct the correlation functions $\xi^{\rm fid}$ and the corresponding Wiener filter. The inpainted map $T_1^I$ is then obtained through the following expression, put forward in Benoit-Lévy et al. (2013).

$$T_1^I = T_1^R + \xi_{12}^{\rm fid} \xi_{22}^{\rm fid-1} (T_2^D - T_2^R). \tag{3.74}$$

This procedure effectively fills in the missing data in region (1). I show in the next paragraph that it preserves the statistical properties of the original data $T^D$ under the assumption of $\xi^R = \xi^{|rmfid} = \xi^{\rm th}$. By leveraging the known covariance structure of the CMB temperature field and utilizing the constraining information from the region (2), we obtain an inpainted map suitable for subsequent analyses and power spectrum computations. Let me rewrite the inpainting procedure in the following equation:

$$T^I = (\mathbb{1} + \mathbb{X}) T^D - \mathbb{X} T^R, \tag{3.75}$$

where $\mathbb{1}$ is the identity matrix, and $\mathbb{X}$ is an operator representing the masking and the Wiener filtering operation. We can express this operator in matrix form, acting on the regions (1) and (2) of the sky as:

$$\mathbb{X} \equiv \begin{pmatrix} -1 & X \\ 0 & 0 \end{pmatrix} \equiv \begin{pmatrix} -1 & \xi_{12}^{\rm th} \xi_{22}^{\rm th-1} \\ 0 & 0 \end{pmatrix}. \tag{3.76}$$

**Correlation function of the inpainted map** As indicated in Eq. (3.75), the inpainted map is obtained from the data $T^D$ with a two-point correlation $\xi^{\rm th}$, a random realization $T^R$ with a two-point correlation $\xi^R$, and the Wiener operation $\mathbb{X}[C^{\rm fid}]$. I can write the correlation function of the inpainted map as follows:

$$\begin{aligned} \langle T^I T^{I\top} \rangle &= (\mathbb{1}+\mathbb{X})\langle T^D T^{D\top}\rangle(\mathbb{1}+\mathbb{X}^\top) + \mathbb{X}\langle T^R T^{R\top}\rangle \mathbb{X}^\top, && (3.77)\\ &= (\mathbb{1}+\mathbb{X})\xi^{\rm th}(\mathbb{1}+\mathbb{X}^\top) + \mathbb{X}\xi^R\mathbb{X}^\top, && (3.78)\\ &= \xi^{\rm th} + \mathbb{X}\xi^{\rm th} + \xi^{\rm th}\mathbb{X}^\top + \mathbb{X}\xi^{\rm th}\mathbb{X}^\top + \mathbb{X}\xi^R\mathbb{X}^\top. && (3.79) \end{aligned}$$

I show that if the data spectrum is perfectly known, i.e., $\xi^{\rm fid} = \xi^{\rm th}$, then we can set $\mathbb{X} = \xi^{\rm th}_{12}\xi^{\rm th-1}_{22}$, and the Wiener operator cancels out the off-diagonal term of the two-point correlation matrix as $X\xi^{\rm th}_{22} = \xi^{\rm th}_{12}$. We then have:

$$\mathbb{X}\xi^{\rm th} = \xi^{\rm th}\mathbb{X}^\top = -\mathbb{X}\xi^{\rm th}\mathbb{X}^\top. \tag{3.80}$$

By assuming perfect knowledge of the data spectrum, we can also generate a random realization of the data spectrum $\xi^R = \xi^{\rm th}$, and thus $\mathbb{X}\xi^R\mathbb{X}^\top = \mathbb{X}\xi^{\rm th}\mathbb{X}^\top$. The right term of Eq. (3.79) then simplifies to $\xi^{\rm th}$. Hence, the Gaussian constrained realization Eq. (3.75) is a valid inpainting procedure under the assumption of perfect data spectrum knowledge:

$$\xi^{\rm th} = \xi^{\rm fid} = \xi^R \implies \langle T^I T^{I\top}\rangle = \xi^{\rm th}. \tag{3.81}$$

**Harmonic space** Maps are inpainted so that we can reconstruct an unbiased two-point correlation function and thus an unbiased power spectrum from them. I relate the power spectrum from the maps to the correlation function using the Legendre decomposition of in Eq. (3.31).

$$C^{(\mathbb{1}+\mathbb{X})D} \equiv \mathbb{L}\cdot\left[(\mathbb{1}+\mathbb{X})\left(T^D T^{D\top}\right)(\mathbb{1}+\mathbb{X}^\top)\right], \quad C^{\mathbb{X}R} = \mathbb{L}\cdot\left[\mathbb{X}\left(T^R T^{R\top}\right)\mathbb{X}^\top\right]. \tag{3.82}$$

Eqs. (3.75), (3.77) and (3.81) allow me to write that

$$C^I = C^{(\mathbb{1}+\mathbb{X})D} + C^{\mathbb{X}R} \qquad \implies \qquad \langle C^I\rangle = \langle C^{(\mathbb{1}+\mathbb{X})D}\rangle + \langle C^{\mathbb{X}R}\rangle = C^{\rm th}. \tag{3.83}$$

The final measured power spectrum is the sum of the Wiener-filtered component and the random realization. Those two power spectra quantities are uncorrelated. We can then compare the two contributions to the final power spectrum, and this is done in Fig. 3.13 by plotting

$$1-\rho_\ell \equiv \frac{C_\ell^{(\mathbb{1}+\mathbb{X}){\rm D}}}{C_\ell^I}. \tag{3.84}$$

The Wiener-filtered component is measured by taking the power spectrum of the Wiener-filtered map. Most of the power in the final spectrum is due to the Wiener-filtered component, and the random realization is a few percent. This is expected as the sky fraction of the sources covers only $\sim 1\%$ of the footprint. Assuming that the Wiener filter is fixed, *i.e.* that we have chosen a fiducial power spectrum for its construction, all inpainting operations are linear in the correlation function. The correlation matrices of the data and the random component can be decomposed as $\xi^D = \sum_{\ell'} P^{\ell'} C^D_{\ell'}$, $\quad \xi^R = \sum_{\ell'} P^{\ell'} C^R_{\ell'}$, where I relied on Eq. (3.21) and defined the matrix $P^L_{ij} \equiv \frac{2L+1}{4\pi} P_L(\cos\theta_{ij})$. $\theta_{ij}$ is the matrix of angular distance between the two pixels $i$ and $j$. One can write

$$\begin{aligned} C^{\mathbb{X}R} &= \sum_{\ell'} Z^{(2)}_{\ell\ell'} C^R_{\ell'}, \quad \text{where} \quad Z^{(2)}_{\ell'} \equiv \mathbb{L}\cdot\left[\mathbb{X}P^{\ell'}\mathbb{X}^\top\right], && (3.85)\\ C_\ell^{(\mathbb{1}+\mathbb{X})D} &= C^D_\ell + \sum_{\ell'}\left[2Z^{(1)}_{\ell\ell'} + Z^{(2)}_{\ell\ell'}\right] C^D_{\ell'}, \quad \text{where} \quad Z^{(1)}_{\ell'} \equiv \mathbb{L}\cdot\left[\mathbb{X}P^{\ell'}\right]. && (3.86) \end{aligned}$$

Then, the inpainted power spectrum is given by

$$C^I = C^{(\mathbb{1}+\mathbb{X})D} + C^{\mathbb{X}R} = C^D + 2Z^{(1)}C^D + Z^{(2)}\left(C^D + C^R\right), \tag{3.87}$$

As it stands, the Wiener-filtered data could be used to estimate parameters, but one should characterize precisely its mean and variance, and that would require characterizing precisely the matrices $Z^{(1)}, Z^{(2)}$, either by an analytical approach or by emulating those matrices from simulations. Our inpainting is done such that $C^{\mathbb{X}R}$ corrects for the bias of the Wiener-filtered data, and we can then use the unbiased power spectrum $C^I$ to estimate cosmological parameters. However, we need to make some assumptions to perform those operations, thus the inpainting procedure is not perfect, and we need to make sure that the bias on the power spectrum is small enough to not affect the cosmological parameters. We will discuss this in the next section. The addition of the random component needs to be taken into account when computing the covariance matrix of the power spectrum, as we will see in Section 4.2.

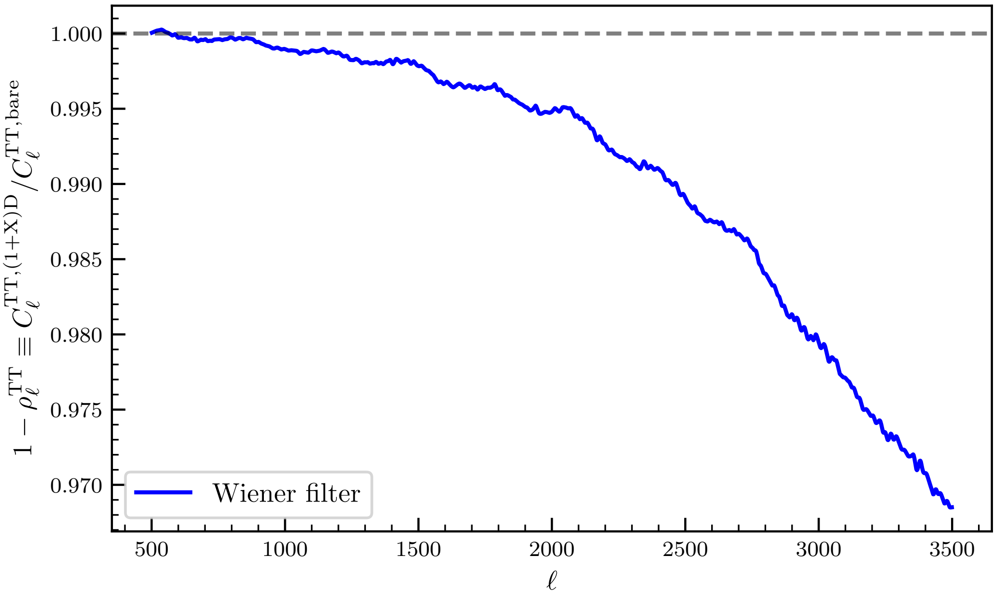


Figure 3.13: Ratio of Wiener filtered component over input power spectrum. 97% of the power in the final power spectrum is due to the Wiener filtered component at $\ell = 3500$. This quantity indicates how much the random realization contributes to the power depending on the scale.

### 3.6.3 Implementing inpainting

In practice, the underlying power spectrum of the maps is not perfectly known, and we need to make assumptions for running the inpainting procedure. Additionally, the Wiener operator $X$ can not be constructed for the full data map due to its large size. I will now discuss how to address these issues.

**Wiener filter computation** To build the operator $\mathbb{X}$, I use the knowledge of the signal power spectrum and power spectrum of the noise. The power spectrum of the signal can be computed from a theoretical mean, for instance, *Planck* 2018 best fit with foreground contributions, convolved by the transfer function of the experiment, the beam, and the pixel window function to obtain $C_\ell^{\mathrm{fid}}$. We apply this method to run our tests with pure Gaussian simulations, see Section 3.6.5. Another option is to use the power spectrum directly estimated from the masked-source maps using the `PolSpice` estimator. This estimator is minimally biased, see Eq. (3.53), hence giving an agnostic estimation of the power spectrum of the maps, even though it has intrinsic scatter due to cosmic variance. We apply this procedure for tests in Section 3.6.6 and show that it produces unbiased results.

Concerning the noise, we consider pure Gaussian noise, which is a good assumption in the case of SPT-3G. In practice, it is estimated from the data itself (see Section 3.4). It is separated into a white noise component and a colored noise component as $N_\ell = N^{\mathrm{white}} + N_\ell^{\mathrm{col}}$, and is integrated into the computation of the Wiener filter with

$$\xi^{\mathrm{th}} = \mathbb{L}^{-1} \cdot C_\ell^{\mathrm{fid}} \rightarrow \xi^{\mathrm{th}} = \mathbb{L}^{-1} \cdot \left[C_\ell^{\mathrm{fid}} + N_\ell^{\mathrm{c}}\right] + \sigma \mathrm{Id}, \quad \text{where} \quad \sigma^2 = {N^{\mathrm{white}}}^2/\Omega_{\mathrm{pix}}. \tag{3.88}$$

The white noise component is important and must be integrated as is, as it contributes to the variance of the inpainted map, and regularizes the inversion of the correlation matrix by rescaling the diagonal.

In the following, we assume that the Wiener filtering always takes properly the signal and noise into account.

**Random realization** The random realizations are generated with the same ingredients as the Wiener filter. Indeed, we can use prior knowledge of the problem using theoretical estimation of the power spectrum, or use directly the power spectrum estimated from the data. In the case of our realistic inpainting, I adapted the code to take as random realizations the mock observations, which are our very realistic simulations designed to have the same statistical properties as the signal. This procedure ensures that the inpainted map has the same anisotropic statistical properties as the input sky. The noise contribution comes either from noise realization drawn for the noise power spectrum or from noise estimation maps presented in Section 2.3.3, which are added to the mock observations before running the inpainting procedure. Both procedures are reliable are give unbiased results.

**Multiple holes and partial constraining zone** The formalism introduced in Section 3.6.2 works for a single hole in the map, and can not be extended to multiple holes. While the SPT-3G map contains thousands of holes, a mitigation strategy is to consider that each hole is independent of the others. This is a good approximation in our case since the holes cover only around 1% of our effective sky. The constraining zone is built so that it excludes the other holes in the map. The missing data in the constraining zone is small enough so that no strong impact can be found on the final constraints. The final inpainted map is then obtained by combining the results of the different holes.

In addition, for a typical resolution of $n_{\mathrm{side}} = 8192$, a SPT-3G map contains approximately $\mathcal{O}(30 \times 10^6)$ pixels, and the two-point correlation matrix of the constraining region $\xi_{22}$ consists of trillions of elements. Inverting such a large matrix is computationally infeasible. However, we can leverage the correlation length of the CMB, which is typically around 1 degree, to assume that only pixels in a surrounding region can be used to constrain the unknown pixels. Let us consider the case where we have only one hole in our map that we want to fill using a Gaussian-constrained realization. We divide the map into three parts: a **constraining region** $T_2$, an unknown region $T_1$, and the remaining part of the map $T_3$, where $\mathbf{T} = (T_1, T_2, T_3)^\top$. Those regions are illustrated in Fig. 3.14 for 3 different holes with increasing size. As the masked source becomes larger, the width of the constraining region is adapted. This figure highlights that even restricted to 1 deg, the constraining zone is incomplete, as other sources need to be masked. The inpainting map is constructed similarly as in Eq. (3.75),

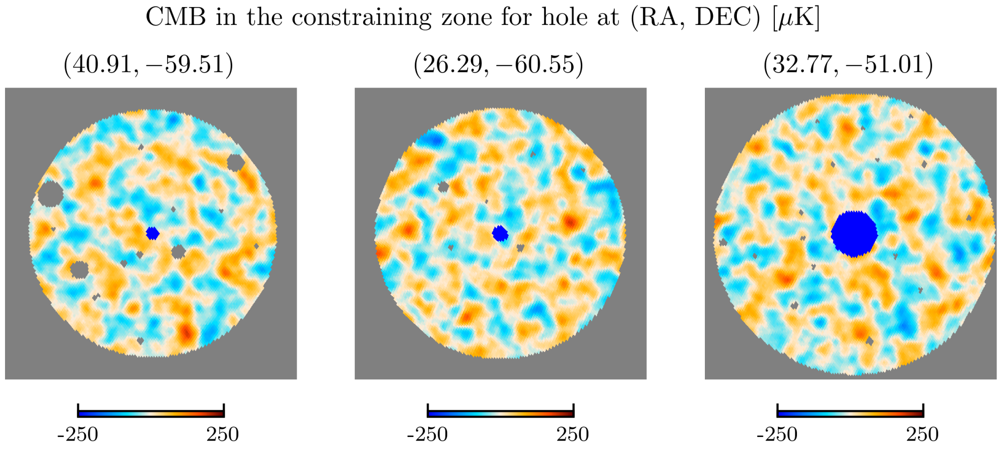


Figure 3.14: Illustration of the three regions of the sky used to build the inpainted map. The CMB-like zone is the constraining zone $T_2$, with a Gaussian realization of CMB temperature field where large-scale modes have been filtered accordingly to the SPT-3G filtering with a 1D transfer function. The constraining region is used to constrain the unknown pixels $T_1$ in dark blue, and its width increases as the masking radius increases. The rest of the map $T_3$, composed of the exterior region and additional sources, in grey, is not used.

however, the inpainting operator is modified to

$$\mathbb{X} \equiv \begin{pmatrix} -1 & X & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \equiv \begin{pmatrix} -1 & \xi_{12}^{\mathrm{fid}} \xi_{22}^{\mathrm{fid}-1} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{3.89}$$

It has a zero block on the last columns, indicating that we are not using information from the rest of the map ($T_3$) to constrain the unknown pixels in the region (1). In Fig. 3.15, I show a row of the pixel correlation matrix $\xi_{22}^{\mathrm{fid}}$ for a hole in the map, extending on the constraining region of size 1 deg. It is displayed for every Stokes combination, anticipating the polarization part. Note that the representation is 2D, to better put forward the spatial properties of the field. The temperature covariance on the top left panel is peaked around the correlated pixel and is extremely narrow. The essential part of the correlation matrix is encapsulated in this limited correlation region.

I compute the final correlation function of the inpainted map under the assumption of separating the map into three different parts:

$$\langle T^I T^{I\top} \rangle = (\mathbb{1} + \mathbb{X}) \xi^{\mathrm{th}} (\mathbb{1} + \mathbb{X}^\top) + \mathbb{X} \xi^R \mathbb{X}^\top, \tag{3.90}$$

$$= \begin{pmatrix} \xi_{11}^R + X \left( \xi_{22}^{\mathrm{th}} + \xi_{22}^R \right) X^\top - X \xi_{21}^R - \xi_{12}^R X^\top & X \xi_{22}^{\mathrm{th}} & X \xi_{23}^{\mathrm{th}} \\ \xi_{22}^{\mathrm{th}} X^\top & \xi_{22}^{\mathrm{th}} & \xi_{23}^{\mathrm{th}} \\ \xi_{32}^{\mathrm{th}} X^\top & \xi_{23}^{\mathrm{th}} & \xi_{33}^{\mathrm{th}} \end{pmatrix} \tag{3.91}$$

When setting $\xi^R = \xi^{\mathrm{fid}} = \xi^{\mathrm{th}}$, the first term of the correlation matrix simplifies to $\xi_{11}^{\mathrm{th}}$ as for the full inpainting. However, the measured covariance block $\langle T_1^I T_3^{I\top} \rangle$ is not equal to $\xi_{13}^{\mathrm{th}}$, but $X \xi_{23}^{\mathrm{th}} = \xi_{12}^{\mathrm{th}} \xi_{22}^{\mathrm{th};-1} \xi_{23}^{\mathrm{th}}$. As a consequence,

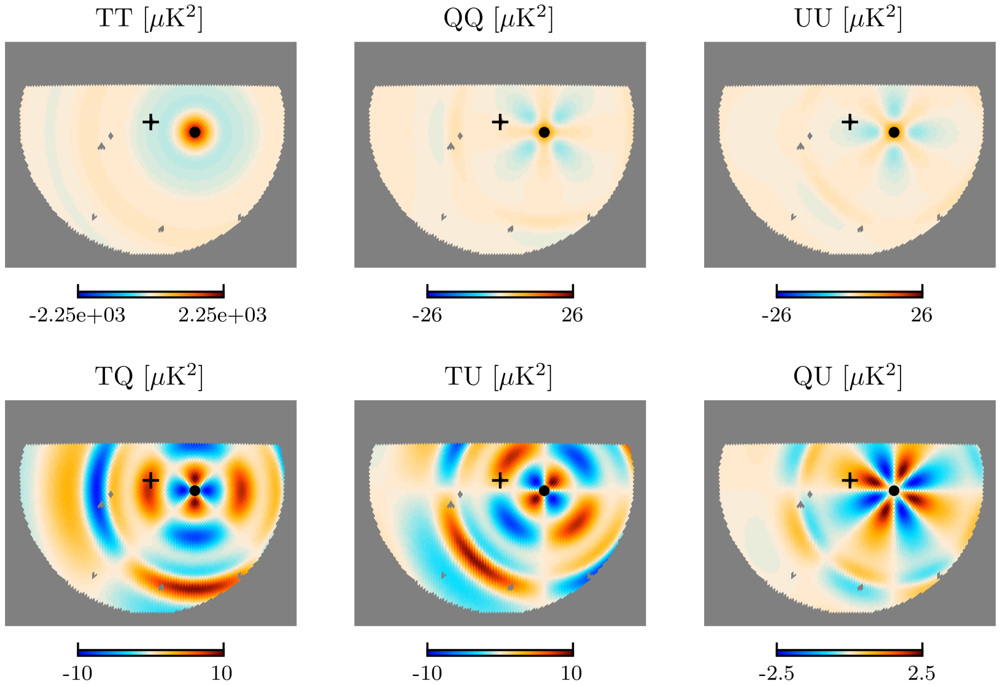


Figure 3.15: Row of pixel correlation matrix $\xi$ for a hole in the SPT-3G mask, located at $(\mathrm{RA}, \mathrm{DEC}) = (-1.43\,\mathrm{deg}, -41.11\,\mathrm{deg})$, for every combination of Stokes parameter. The hole to be inpainted is located under the cross. The correlated pixel is located under the black dot. The hard border on the top of the map is due to the mask. The rest of the pixel cut-off corresponds to the maximal extent of the correlation used to constrain the CMB inside the hole. The bottom row shows the correlation across Stokes parameters, with the left Stokes parameter being the one correlated. For example, the leftmost plot shows the correlation between the central temperature pixel with the $Q$ pixels. The transposed matrix has a similar structure.

even if the underlying spectrum is perfectly known, restricting the constraining zone to a small subset implies that the correlation function of the inpainted map will be biased. However, when the constraining region is large enough, the bias is negligible. We assert that this assumption is robust with a Monte Carlo test suite, first by building a suite of 100 realistic simulations, filtered with the 2D transfer function measured in Section 3.3. We measure the spectra of the input skies $\hat{C}_\ell^{\mathrm{ba}}$ and then apply the inpainting procedure multiple times, varying the radius $r$ of the constraining zone (2) in $[14', 31', 52', 69']$. The power spectra of the inpainted maps are measured as $\hat{C}_\ell^{\mathrm{fi}}$ and we compare them to the input spectra by taking the difference weighted by the standard deviation of the simulations. The latter is a proxy for cosmic variance and allows us to put forward significant features in the power spectrum. Results are obtained in Fig. 3.16, and show that the bias is negligible for $r \geq 1\,\mathrm{deg}'$, while it reaches $\sim 0.1\sigma_\ell$ when the constraining zone is too small. Indeed, the CMB is correlated on a typical scale of 1 deg, see Fig. 3.15. Since regions (1) and (3) are separated from more than 1 deg, their correlation contributes negligibly to the power spectrum computation. When the radius of the constraining zone is too small, the effect is a smoothing of the peaks and troughs of the power spectrum, which contains cosmological information. Such a bias should be avoided or precisely modeled. I conclude that we can use the assumptions of independent holes and local constraining zones to build the Wiener filter as long as the latter is large enough to contain the relevant information.

### 3.6.4 Inpainting polarization

I showed above that the procedure is reliable for temperature. In order to extend this procedure to polarization, we first start by writing that CMB is a correlated isotropic Gaussian random field, and that, similarly to Eq. (3.72), I can write

$$\mathcal{P}(T^D, Q^D, U^D) \propto \exp\left(-\frac{1}{2}(T^D\ Q^D\ U^D)\xi^{-1}\begin{pmatrix}T^D\\Q^D\\U^D\end{pmatrix}\right). \tag{3.92}$$

The two-point correlation matrix $\xi$ is now a $3n_{\mathrm{pix}} \times 3n_{\mathrm{pix}}$ matrix and is more complicated to compute than the pure temperature one. We assumed, consistently with the rest of this thesis, that the Universe is invariant under parity, and starting from Eq. (3.4) we obtain the polarization correlation function from the spectra

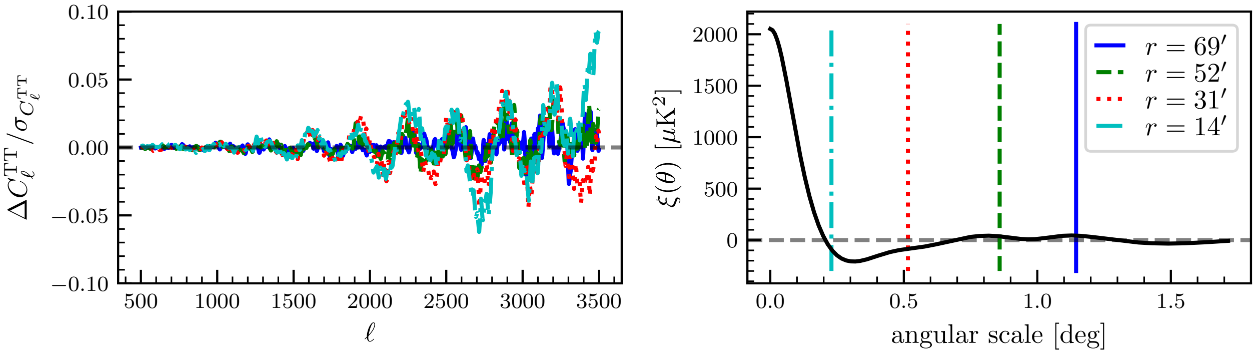


Figure 3.16: Left: Bias on the power spectrum as a function of the radius of the constraining zone. The bias is computed as the difference between the inpainted power spectrum and the input power spectrum, weighted by the standard deviation of the simulations. The bias is computed $r = [14', 31', 52', 69']$. Right: Correlation function of the input sky, with indications of the radii of the constraining zones. When all significant correlation is taken into account ($\theta > 1\,$deg), the inpainted power spectrum is unbiased.

$C^{\rm TT}, C^{\rm EE}, C^{\rm BB}, C^{\rm TE}$, (Kamionkowski et al., 1997)

$$\langle \bar{Q}(\hat{\mathbf{n}})\bar{Q}(\hat{\mathbf{u}})\rangle = \sum_\ell \frac{2\ell+1}{4\pi}\left[C_\ell^{\rm EE}F_{1,\ell 2}\left(\hat{\mathbf{n}}\cdot\hat{\mathbf{u}}\right) - C_\ell^{\rm BB}F_{2,\ell 2}\left(\hat{\mathbf{n}}\cdot\hat{\mathbf{u}}\right)\right] \tag{3.93}$$

$$\langle \bar{U}(\hat{\mathbf{n}})\bar{U}(\hat{\mathbf{u}})\rangle = \sum_\ell \frac{2\ell+1}{4\pi}\left[-C_\ell^{\rm EE}F_{2,\ell 2}\left(\hat{\mathbf{n}}\cdot\hat{\mathbf{u}}\right) + C_\ell^{\rm BB}F_{1,\ell 2}\left(\hat{\mathbf{n}}\cdot\hat{\mathbf{u}}\right)\right] \tag{3.94}$$

$$\langle T(\hat{\mathbf{n}})\bar{Q}(\hat{\mathbf{u}})\rangle = -\sum_\ell \frac{2\ell+1}{4\pi}C_\ell^{\rm TE}F_{1,\ell 0}\left(\hat{\mathbf{n}}\cdot\hat{\mathbf{u}}\right) \tag{3.95}$$

$$\langle T(\hat{\mathbf{n}})\bar{U}(\hat{\mathbf{u}})\rangle = 0, \tag{3.96}$$

where

$$F_{1,\ell m}(\theta) = A_{\ell m}\left[-\left(\frac{\ell - m^2}{\sin^2\theta} + \frac{1}{2}\ell(\ell-1)\right)P_\ell^m(\cos\theta) + (\ell+m)\frac{\cos\theta}{\sin^2\theta}P_{\ell-1}^m(\cos\theta)\right], \tag{3.97}$$

$$F_{2,\ell m}(\theta) = A_{\ell m}\frac{m}{\sin^2\theta}\left[-(\ell-1)\cos\theta P_\ell^m(\cos\theta) + (\ell+m)P_{\ell-1}^m(\cos\theta)\right], \tag{3.98}$$

and $A_{\ell m} = 2\sqrt{\frac{(\ell-2)!(\ell-m)!}{(\ell+2)!(\ell+m)!}}$ and $P_\ell^m$ are the associated Legendre polynomials. The correlation matrix is then obtained by computing the correlation function on the pixel grid. However, those values are obtained for the rotated Stokes parameters $\bar{Q}, \bar{U}$ and need to be rotated back to the original frame, which is done by applying, for each combination of pixels $(i,j)$, $e^{2i\alpha}(\bar{Q}_i + i\bar{U}_i) = Q_i + iU_i$ and $e^{2i\gamma}(\bar{Q}_j + i\bar{U}_j) = Q_j + iU_j$, here $\alpha$ and $\gamma$ are the angles required to rotate the local basis in pixels $i$ and $j$ so that they are aligned with the sphere geodesic between the two pixels. This operation is expensive and is one of the bottlenecks of the inpainting procedure. We verify that the correlation matrices are properly computed by comparing them to one obtained from Monte Carlo simulations. The correlation matrices $\xi$ are illustrated in Fig. 3.15, where a row is displayed for each Stokes combination on the map. While the temperature and polarization have expected narrow correlations, the cross-correlation between temperature and polarization is much more spread out, as expected from the $C^{\rm TE}$ spectrum, and also much weaker, but are all non-zero as induced by the rotation of the polarization frame.

The rest of the inpainting procedure is the same as for temperature, but it should be noted that we are solving jointly for the temperature and polarization maps, as the polarization maps are not independent of the temperature ones because of the non-zero $C^{\rm TE}$ spectrum. The inverted matrix size is multiplied by three, which induces a notable scaling factor as the inversion scales as $\mathcal{O}(n_{\rm pix}^3)$.

### 3.6.5 Asserting robustness

The inpainting procedure has a direct effect on the mean value of the power spectrum, as illustrated in Eq. (3.87). A bias on the power spectrum will propagate on the cosmological parameters, and we need to make sure that our procedure is robust to this bias. In this section, we perform a series of tests demonstrating that our procedure is unbiased and highlights results for temperature, but the results are similar for polarization. We perform those tests on a set of 800 simulations that are filtered with a 2D transfer function $f_{\ell m}$, see Section 3.3, and use as input for the Wiener filter and the random realization the power spectrum used to generate those realizations, filtered with a 1D transfer function $F_\ell$.

**Response function** In this paragraph, I evaluate the robustness of the code to an error on the fiducial input power spectrum used to generate the Wiener filtering and the random realization. I therefore assume $\xi^R = \xi^{\mathrm{fid}} \neq \xi^{\mathrm{th}}$. First, I take advantage of the linear properties of the Gaussian-constrained realization to evaluate its response function. Let me detail the mathematical formalism. I compute derivatives of the Wiener operator with respect to the fiducial spectrum. For a given multipole $L$, I first compute the derivative of $X$ with respect to $C_L^{\mathrm{fid}}$. I write

$$\frac{\partial X}{\partial C_L^{\mathrm{fid}}} = \frac{\partial \xi_{12}^{\mathrm{fid}}}{\partial C_L^{\mathrm{fid}}} \xi_{22}^{-1} + \xi_{12}^{\mathrm{fid}} \frac{\partial \xi_{22}^{\mathrm{th}-1}}{\partial C_L^{\mathrm{fid}}} = P_{12}^L \xi_{22}^{\mathrm{fid}-1} - \xi_{12}^{\mathrm{fid}} \xi_{22}^{\mathrm{fid}-1} P_{22}^L \xi_{22}^{\mathrm{fid}-1} = \left(P_{12}^L - X P_{22}^L\right) \xi_{22}^{\mathrm{fid}-1}. \tag{3.99}$$

With this expression, the derivatives of the inpainted correlation matrix $\xi^I = \langle T^I T^{I\top} \rangle$ with respect to the fiducial power spectrum can be computed. Let's take the ideal case where the constraining zone covers all the available data, which is equivalent to ignoring the region (3). From Eq. (3.91), the only non-vanishing elements of the derivative of the correlation evaluated at the theoretical value are

$$\left.\frac{\partial \xi_{11}^I}{\partial C_L^{\mathrm{fid}}}\right|_{\xi^{\mathrm{fid}}=\xi^{\mathrm{th}}} = P_{11}^L - X P_{22}^L X^\top, \tag{3.100}$$

$$\left.\frac{\partial \xi_{12}^I}{\partial C_L^{\mathrm{fid}}}\right|_{\xi^{\mathrm{fid}}=\xi^{\mathrm{th}}} = \left(P_{12}^L - X P_{22}^L\right). \tag{3.101}$$

I can then conclude that the derivative of the inpainted correlation matrix with respect to the fiducial power spectrum is obtained by applying the Wiener filter operation on the Legendre correlation matrix as

$$\left.\frac{\partial \xi^I}{\partial C_L^{\mathrm{fid}}}\right|_{\xi^{\mathrm{fid}}=\xi^{\mathrm{th}}} = P^L - (\mathbb{1} + \mathbb{X}) P^L (\mathbb{1} + \mathbb{X}^\top) \tag{3.102}$$

Assuming an input error on the power spectrum $\epsilon_L$, I obtain a first-order expansion of the inpainted correlation matrix

$$\xi^I \approx \xi^{\mathrm{th}} + \sum_L \epsilon_L \left.\frac{\partial \xi^I}{\partial C_L^{\mathrm{fid}}}\right|_{\xi^{\mathrm{fid}}=\xi^{\mathrm{th}}}. \tag{3.103}$$

Following this computation, in Fig. 3.17, I show the response of inpainting procedure to a deviation of a hundred $\sigma$ at multipole $L = 1000$, *i.e.* $\epsilon_L = 100\sigma_{1000}\delta_{L1000}$, compared with the residual measured on 800 realistic simulations. This deviation on the input spectrum would induce a residual bias on the power spectrum of the inpainted maps of around $0.02\sigma_{1000}$. This is expected from Fig. 3.13 since most of the power spectrum of the inpainted maps is determined by the data, and hence errors made on the Wiener filter or the random realization are damped. Nevertheless, we need to evaluate the impact of an integrated deviation on the total range of multipoles.

**Error on the fiducial power spectrum** To further evaluate the impact of an error on the fiducial power spectrum, I generate 800 realistic filtered simulations following *Planck* cosmology and inpaint them following the above procedure, but choosing as fiducial spectrum a different spectrum. The rescaled difference $(C_\ell^{\mathrm{fid}} - C_\ell^{\mathrm{th}})/10$ is displayed in Fig. 3.18 with the blue line and is chosen to be of the order of the statistical error on the power spectrum, reaching $\sim 0.6\sigma_\ell$ at $\ell = 3000$. I thus evaluate how the integrated error impacts the resulting power spectrum. I compute the difference between the inpainted power spectrum and the input power spectrum, weighted by the standard deviation of the simulations, for the correct inpainting procedure and the wrong fiducial spectrum. Results are displayed on the same plot. I show that the bias is negligible, of order 1% of the statistical error on the power spectrum, even when the input spectrum is significantly wrong on a wide range of multipoles. We then conclude that the inpainting procedure is robust to any mismatch of the power spectrum.

**Impact on the covariance matrix** The impact of the inpainting procedure on the covariance matrix is expected to be minimal. Since $C^{(\mathbb{1}+\mathbb{X})D}$ and $C^{\mathbb{X}R}$ are uncorrelated, their covariances add up to form the covariance of $C^I$. Then, as the inpainted power spectrum is unbiased thanks to Eq. (3.87), this property also propagates to the covariance. Nevertheless, as several assumptions were made, we verify this property by taking a set of 800 realistic simulations and comparing their sample variance before and after inpainting, in Fig. 3.19.

### 3.6.6 Realistic inpainting

The former framework was introduced to detail the inpainting procedure. In practice, the data set is not composed of a single realization of the CMB sky, but of a set of observations across multiple frequencies that have varying noise and instrument properties, and foreground contaminations. Filling the holes should be done in a way that is consistent with the data set.

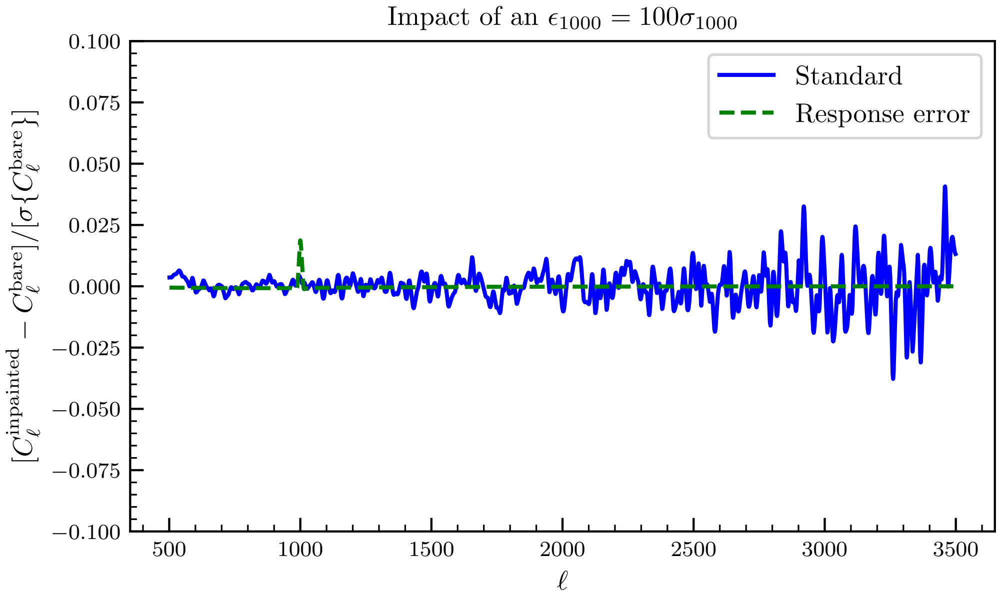


Figure 3.17: Comparison of the residual spectra after inpainting to a response error resulting from applying an error $\epsilon_L = 100\sigma_{1000}\delta_{L1000}$. A very significant error on the fiducial power spectrum leads to a small error.

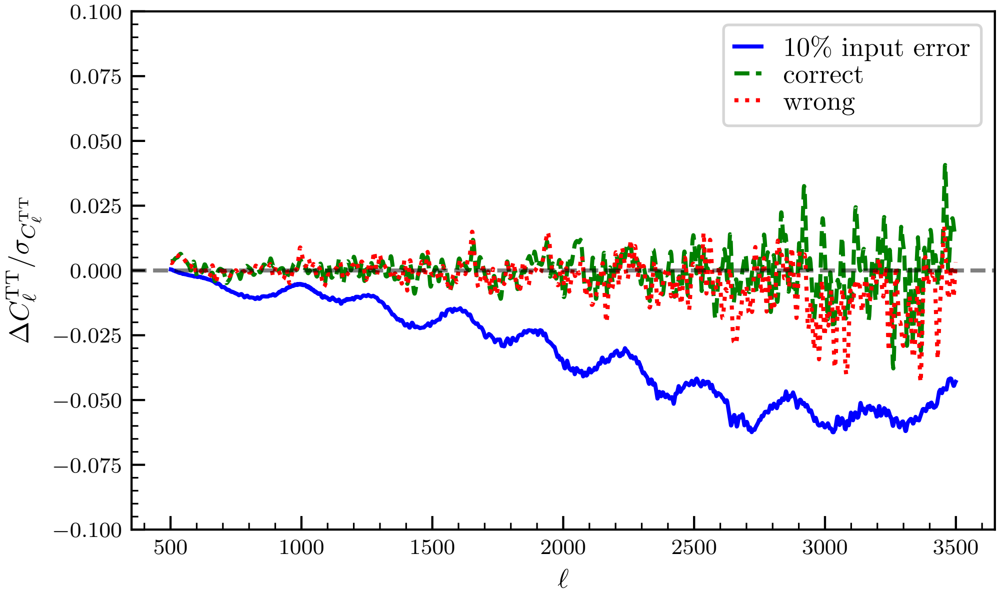


Figure 3.18: The blue line indicates 10% of the error on the fiducial power spectrum used to generate the simulations, displayed along the residual power for inpainting with the correct input (green) and the wrong spectrum (red). A significant input error has no strong impact on the inpainted power spectrum.

**Influence of map-making** While the formalism for Gaussian constrained realization is obtained from the assumption of isotropic signal, we know from Section 3.3 that SPT-3G 19/20 data set exhibits a statistical anisotropy stemming from the filtering of the time streams and the scanning strategy. In the tests of the previous section, we filtered the maps by a 2D transfer function, thus showing that even if the maps display statistical anisotropies, the inpainting procedure correctly reconstructs the power spectrum. Indeed, Eq. (3.83) implies that the Gaussian-constrained realization is designed to reconstruct the averaged isotropic two-point function, even though the field exhibits anisotropic variation. Subsequently, it also implies that the non-Gaussian statistics are not reconstructed. In order to show that this procedure can be applied to real data, we apply it to one of our realistic mock observations and display the results in Fig. 3.20. Note that the figure displays results in polarization, which are as accurate as temperature.

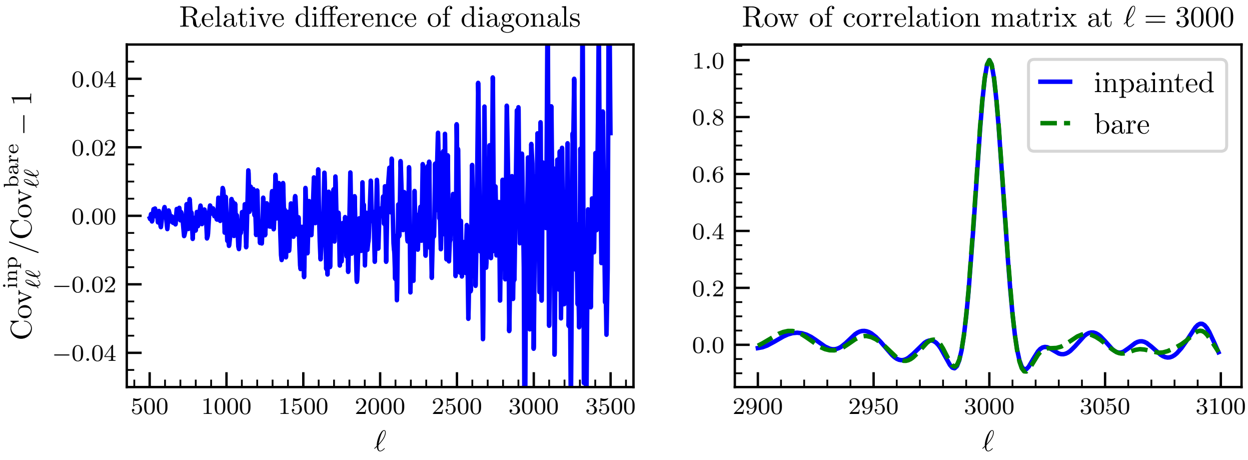


Figure 3.19: Left: relative difference of covariance diagonals. The covariances are obtained by taking sample variance of power spectra computed over 800 hundred realistic simulations, computed before and after applying the inpainting procedure to the maps. Right: the row at $\ell = 3000$ of the obtained correlation matrices. No significant deviation due to inpainting is found at the level of the covariance matrix.

**Multi-frequency inpainting** The SPT-3G band powers cover 3 frequency channels. The reconstructed signal inside the holes should be coherent across frequencies so that the measured cross-frequency spectra are not biased, as they are a crucial block of the data vector. Let us illustrate that problem with temperature, writing the data vector as $T = (T^{95}, T^{150}, T^{220})^\top$. The Wiener filter should then be computed as a dense matrix so that correlations among frequencies are taken into account. However, the region (2) two-point correlation function would then be going from $3n_{\mathrm{pix}} \times 3n_{\mathrm{pix}} \to 9n_{\mathrm{pix}} \times 9n_{\mathrm{pix}}$, thus having its size multiplied by 3. While we allowed ourselves to scale the matrix in order to include polarization, this is not feasible for frequency. Indeed, the computational cost scales as the cube of the size of the matrix, and the matrices inversion is the most expensive part of the algorithm. To solve this problem, I compute the Wiener filter for every single frequency and inpaint the maps independently. I assume that as long as the random realization carries the same statistical properties as the data and is correlated between frequencies, the inpainted maps will be nearly unbiased. Effectively, we replace the Wiener filter $\mathbb{X}$ by a block-diagonal matrix

$$\mathbb{X} = \begin{pmatrix} \mathbb{X}^{95,95} & \mathbb{X}^{95,150} & \mathbb{X}^{95,220} \\ \mathbb{X}^{150,95} & \mathbb{X}^{150,150} & \mathbb{X}^{150,220} \\ \mathbb{X}^{220,95} & \mathbb{X}^{220,150} & \mathbb{X}^{220,220} \end{pmatrix} \to \begin{pmatrix} \mathbb{X}^{95,95} & 0 & 0 \\ 0 & \mathbb{X}^{150,150} & 0 \\ 0 & 0 & \mathbb{X}^{220,220} \end{pmatrix} . \tag{3.104}$$

This approach has the advantage of being computationally cheaper, as the inversion of the Wiener filter is done on smaller matrices. To ensure consistency across frequencies, we use the same mock observation as a random realization for all frequencies and correlated noise estimated from the data. A suite of 10 realistic multi-frequency signal simulations with noise estimations from Section 2.3.3 is performed. Inpainting is applied in every frequency channel independently. The power spectra are computed on the full patch for the inpainted skies, but sources are masked on the input skies since the noise data inside holes are not filtered as the rest of the map. In Fig. 3.21 is displayed the averaged residual of the temperature power spectrum. This average is performed on 10 simulations. The variance is comparable to the test in Fig. 3.20 despite averaging over 10 simulations since the input skies have masked sources. This test demonstrates that the assumption of neglecting cross-frequency components of the Wiener filter is a valid approximation.

**Inpainting bundles** As described in Section 3.5, we use the bundles cross-spectrum approach to get rid of the noise bias. We thus need to inpaint consistently the $n_{\mathrm{BUN}}$ maps of the data set. I do so by inpainting the bundles with consistent Wiener filtering. The random realizations are obtained with a unique signal realization and different noise realization generated from the noise power spectrum measured in Section 3.4. This approach allows no bias in the cross-spectra, as the noise is not correlated across bundles.

## Conclusion

In this chapter, I detailed the power spectrum estimation of the SPT-3G 2019-2020 data set. I first described the correspondence between real space and harmonic analysis, in order to introduce the pseudo-power spectrum framework and the `PolSpice` de-biasing procedure. I detailed a new formulation in harmonic space of such

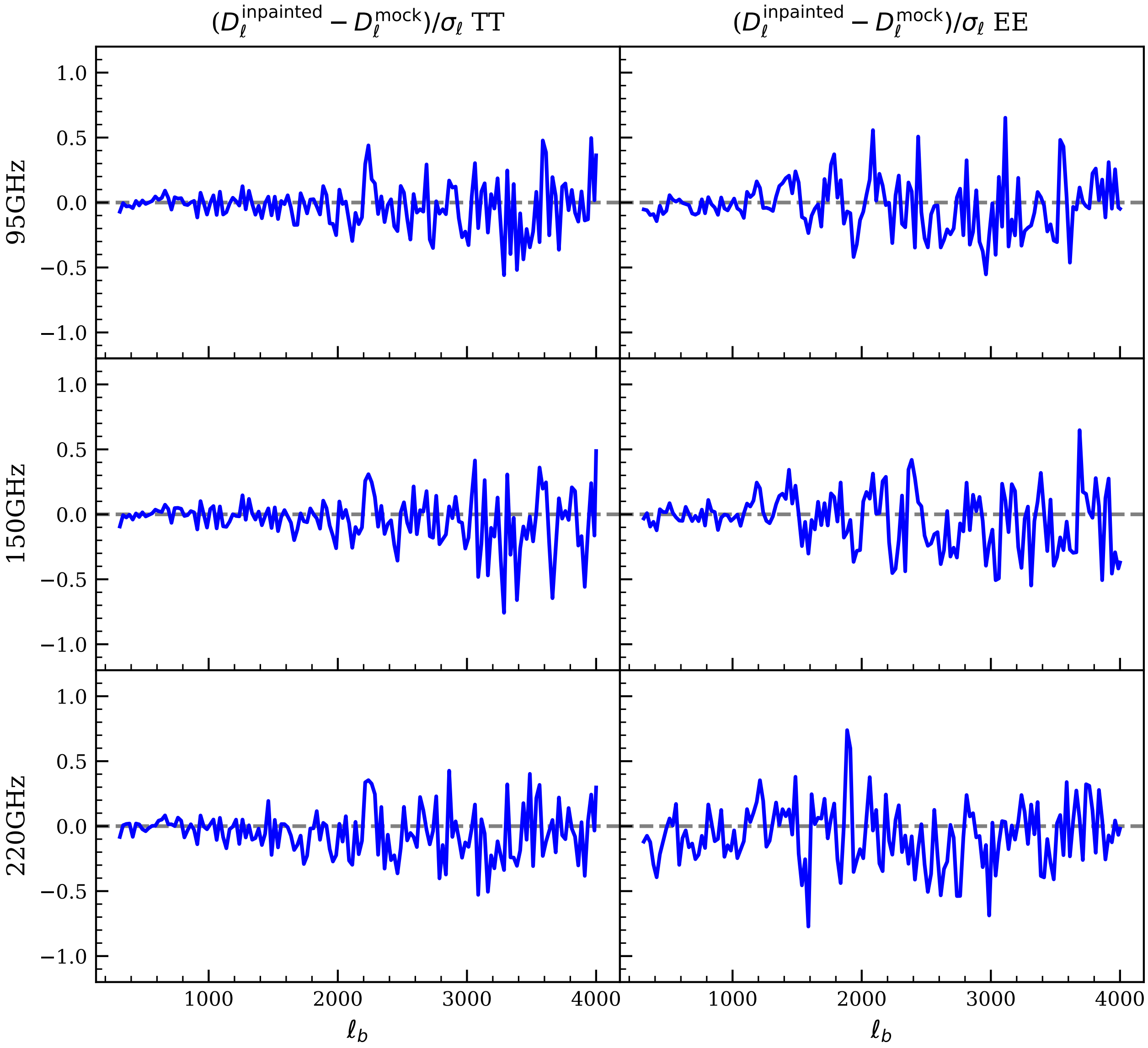


Figure 3.20: Binned relative difference in units of the standard deviation. $D^{\rm inpainted}$ is the spectrum computed on the inpainted mock observations, while $D^{\rm mock}$ is computed on the bare mock observations. The binning applied is defined in Eq. (3.70). No significant deviation is found after applying our inpainting procedure to the mock observations.

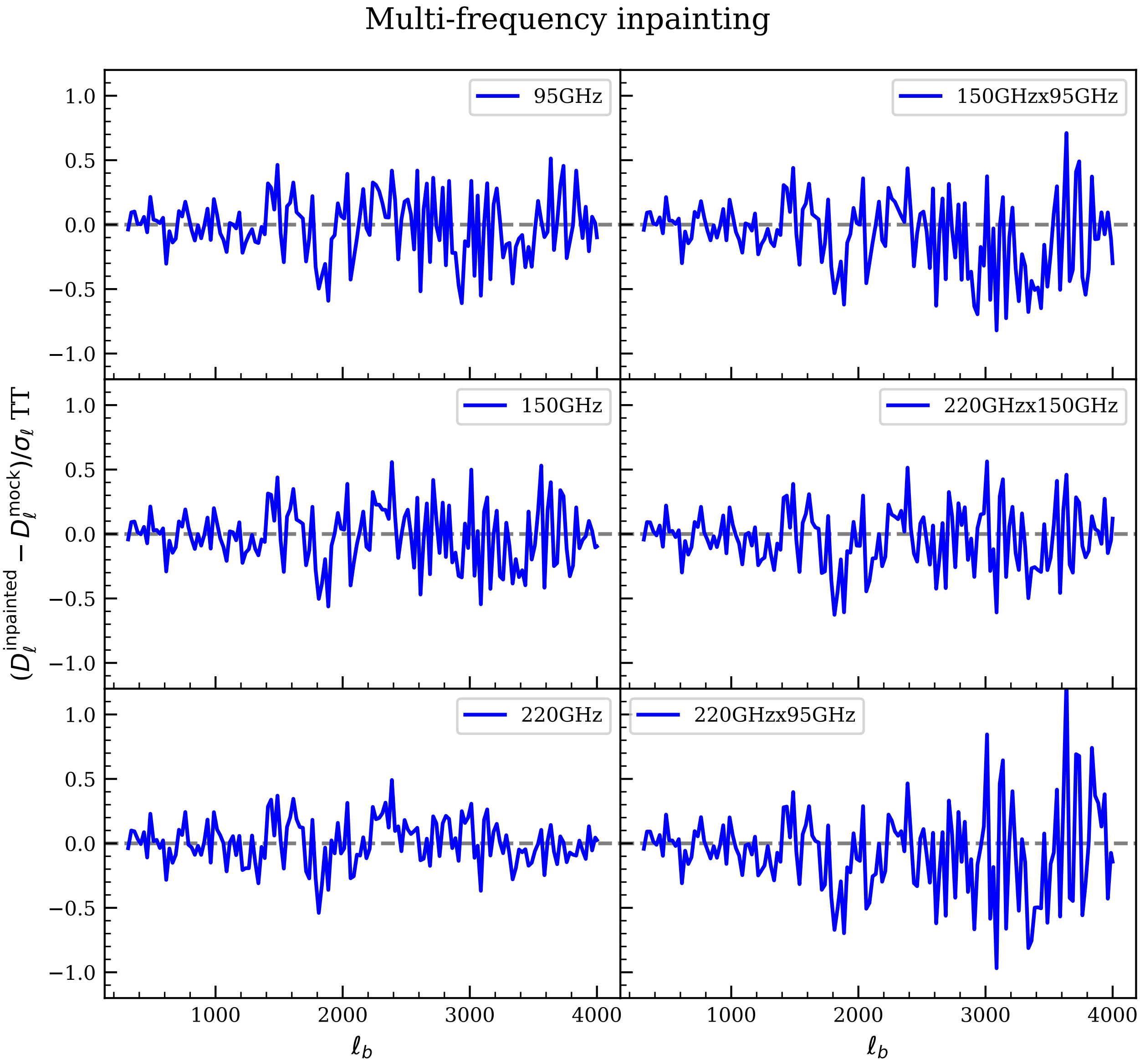


Figure 3.21: Binned relative difference in units of the standard deviation. Inuts skies are obtained with realistic simulations and noise map estimations from data. $D^{\rm inpainted}$ is the spectrum computed on the inpainted skies, while $D^{\rm mock}$ is computed on the input skies with source masked. The binning applied is defined in Eq. (3.70). Our procedure accurately reconstrucs the cross-frequency power spectra.

an operation and applied this procedure to the noise maps to estimate the noise power spectra, and to the data maps to estimate the auto and cross-frequency power spectra using the cross-bundles method. I then described the inpainting procedure that is used to fill the holes in the data maps. The inpainting procedure is unbiased as long as certain assumptions are met. I performed a series of tests to show its precision. Inpainting is implemented in the SPT-3G 2019-2020 analysis pipeline. Preliminary band powers are displayed on the following page.

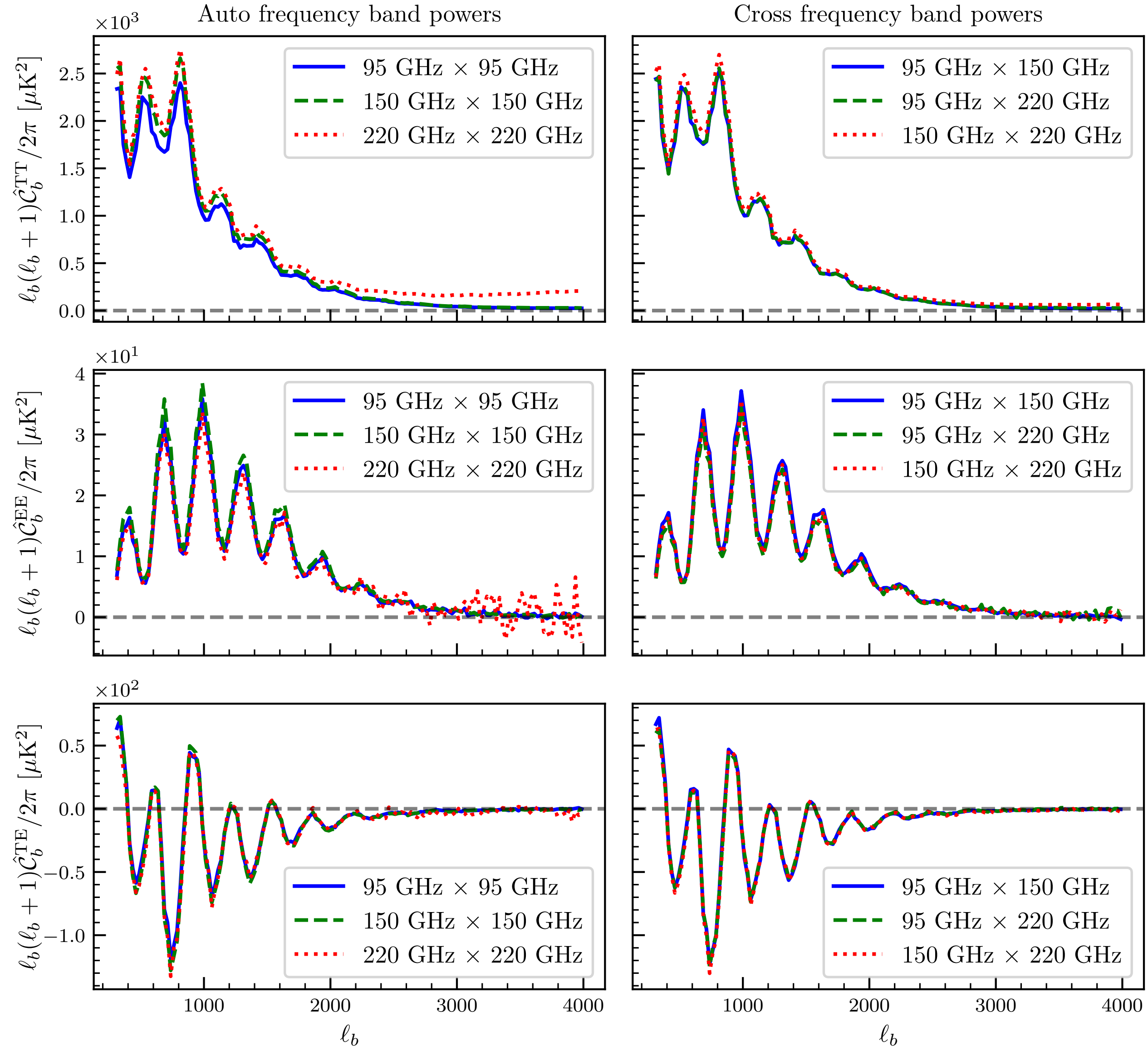


Figure 3.22: Uncalibrated binned SPT-3G 19/20 preliminary band powers obtained with cross-spectrum approach, using an analysis mask with source masking. They are corrected for the beam $B$ and the emulated transfer function $F^{\mathrm{ana}}$. The binning width is chosen to be $\Delta\ell = 25$, but is likely to evolve in the analysis. Left plots: auto frequency spectra. Right plots: cross-frequency spectra. First row: TT, second row: TE, third row: EE.

# Chapter 4

# Likelihood analysis

---

**Contributions** This chapter focuses on the final compression step of the analysis pipeline: the likelihood analysis. I have taken a leadership role in this work within the SPT-3G collaboration, where I have been actively involved in developing new components of the pipeline. Specifically, my contributions involve the computation of the covariance matrix and building the likelihood. Here is a breakdown of my contributions:

- In Section 4.1, I am responsible for the computation of the window functions and the band powers.
- In Section 4.2, I have introduced a high-precision analytical approach to compute the signal CMB covariance matrix on small survey areas. This new method, described in the paper by Camphuis et al. (2022), is an integral part of the analysis presented in this manuscript.
- In Section 4.2.2, I have performed a high-precision analytical estimation of the noise covariance matrix. Ongoing work is done to compare it with an estimator derived from the data.
- In Sections 4.2.2 and 4.2.3, I have adapted the analytical approach to account for the SPT-3G combination of frequencies and analysis choices. Specifically, I have implemented a correction for filtering bias at the covariance level, developed covariance rescaling techniques for marginalizing over the inpainting random realization, and integrated the lensing contribution into the main covariance matrix.

The implementation of analytical covariance correction to account for filtering loss of modes will be presented in an upcoming publication (Hivon, Doussot, Camphuis, et al., *in prep*). Furthermore, ongoing work on the likelihood analysis will be summarized in a future publication, which will present the cosmological results of the SPT-3G 19/20 data.

**Introduction** The measured band powers encompass all the necessary statistical information to constrain cosmological parameters. As discussed in the preceding chapter, significant effort is required to accurately model the bias introduced by the instrument and the analysis pipeline. In order to extract cosmological information, it is imperative to construct a likelihood function that describes the probability of observing the band powers given a specific set of cosmological and nuisance parameters. In this chapter, we detail the construction of the likelihood function, which will later be used to test various cosmological models and explore their associated parameter space using the Markov Chain Monte Carlo (MCMC) method.

The CMB band powers consist of sets of power spectra (TT, TE, ET, EE) for all six frequency combinations, and they are treated as random variables. If the signal were all-sky and isotropic, we could assume that the band powers follow a gamma distribution (Percival and Brown, 2006). However, in practice, the signal is not all-sky, and the posterior distribution is affected by the coupling between modes originating in sky masking. On small scales, the pseudo-power spectrum follows a gamma distribution, but with a reduced number of degrees of freedom denoted by $\nu_\ell = (2\ell+1) f_{\mathrm{sky}} \frac{w_2^2}{w_4}$, where $w_i$ represents the $i$-th moment of the employed mask, and $f_{\mathrm{sky}}$ is the fraction of the sky covered by the observation (Percival and Brown, 2006). To approximate the signal, we rely on the central limit theorem, which states that the signal can be approximated by a multivariate Gaussian distribution. This assumption holds particularly true at sufficiently large multipoles. We are not probing the low multipoles with fewer degrees of freedom as the filtering strategy employed introduces a cutoff at $\ell = 300$, see Fig. 3.7. Under this assumption, the statistical properties of the band powers $\hat{\mathcal{C}}^{\mathrm{X}\nu\mathrm{Y}\mu}$ can be fully characterized by their mean $\mathcal{C}^{\mathrm{X}\nu\mathrm{Y}\mu}$ and their covariance matrix $\hat{\Sigma}^{\mathrm{X}\nu\mathrm{Y}\mu;\mathrm{Z}\alpha\mathrm{V}\beta}$, which we will derive in this section. Consequently, we will be able to formulate the likelihood function for the band powers given a specific model and a set of cosmological parameters $\vartheta$ as

$$\mathcal{L}\left(\hat{\mathcal{C}}|\mathcal{C}^{\rm th}(\vartheta)\right) = \frac{1}{\sqrt{\det\left(2\pi\Sigma\right)}} \exp\left(-\frac{1}{2}\Delta\mathcal{C}\ \hat{\Sigma}^{-1}\ \Delta\mathcal{C}\right), \tag{4.1}$$

where $\Delta\mathcal{C} = \hat{\mathcal{C}} - \mathcal{C}^{\rm th}(\vartheta)$ is the difference between the measured band powers and the theoretical prediction depending on the model and its parameters $\vartheta$.

## 4.1 Modeling of theory power spectrum

The band powers represent the final output of the analysis pipeline, as explained in detail in Section 3.5. In order to estimate their mean value, we rely on the parameters described in Eq. (4.1). Currently, theoretical predictions of cosmological models for the pure CMB can be obtained through different methods. One approach involves using Boltzmann line-of-sight solvers like `CAMB` (Lewis et al., 2000) or `CLASS` (Lesgourgues, 2011). Alternatively, state-of-the-art neural network emulators like `CosmoPower` (Spurio Mancini et al., 2022) can also be employed. These emulators provide a significant speed-up for the MCMC analysis, but they require training on a large set of predictions first. To account for foreground components in the power spectrum, nuisance parameters are employed in the likelihood function, similar to the approach used in the SPT-3G 2018 study and earlier works (De Zotti et al., 2005; Dunkley et al., 2013; George et al., 2015; Reichardt et al., 2020). This is preliminary work and will be improved shortly.

We will model foreground contamination on the temperature power spectrum by computing contribution from galactic dust, the Poisson-distributed unresolved point sources, the CIB clustering, the thermal Sunyaev-Zel'dovich effect (tSZ), the kinetic Sunyaev-Zel'dovich (kSZ) effect and their cross-correlations.

$$\begin{cases} D^{\rm cirrus}_{\ell,\nu\times\mu} = A^{\rm cirrus}_{80}\dfrac{g(\mu)g(\nu)}{g(\nu_0^{\rm cirrus})^2}\left(\dfrac{\nu\mu}{(\nu_0^{\rm cirrus})^2}\right)^{\beta^{\rm cirrus}}\left(\dfrac{\ell}{80}\right)^{\alpha^{\rm cirrus}+2} \quad \text{with varying } \left[A^{\rm cirrus}_{80}, \beta^{\rm cirrus}, \alpha^{\rm cirrus}\right], \\ D^{\rm TT,Poisson}_{\ell,\nu\times\mu} = D^{\rm TT,Poisson}_{3000,\nu\times\mu}\left(\dfrac{\ell}{3000}\right)^2 \quad \text{with varying } \left[D^{\rm TT,Poisson}_{3000,\nu\times\mu},\ \forall\nu,\mu\right], \\ D^{\rm CIB-cl.}_{\ell,\nu\times\mu} = A^{\rm CIB-cl.}_{80}\dfrac{g(\mu)g(\nu)}{g(\nu_0^{\rm CIB-cl.})^2}\left(\dfrac{\nu\mu}{(\nu_0^{\rm CIB-cl.})^2}\right)^{\beta^{\rm CIB-cl.}}\left(\dfrac{\ell}{80}\right)^{0.8} \quad \text{with varying } \left[A^{\rm CIB-cl.}_{80}, \beta^{\rm CIB-cl.}\right], \\ D^{\rm tSZ}_{\ell,\nu\times\mu} = A^{\rm tSZ}\dfrac{f(\nu)f(\mu)}{f(\nu_0^{\rm tSZ})^2}D^{\rm tSZ,template}_{\ell} \quad \text{with varying } \left[A^{\rm tSZ}\right], \\ D^{\rm tSZ-CIB}_{\ell,\nu\times\mu} = -\xi\left(\sqrt{D^{\rm tSZ}_{\ell,\nu\times\nu}D^{\rm CIB-cl.}_{\ell,\nu\times\nu}} + \sqrt{D^{\rm tSZ}_{\ell,\mu\times\mu}D^{\rm CIB-cl.}_{\ell,\mu\times\mu}}\right) \quad \text{with varying } [\xi], \\ D^{\rm kSZ}_{\ell} = A^{\rm kSZ}D^{\rm kSZ,template}_{\ell} \quad \text{with varying } \left[A^{\rm kSZ}\right]. \end{cases} \tag{4.2}$$

The foreground frequency dependence is characterized by the functions $g(\nu)$ and $f(\nu)$. The former is computed as $g(\nu) = B_\nu(T)\left(\partial B_\nu(T)/\partial T\right)^{-1}|_{T_{\rm CMB}}$, where $T$ is set to the appropriate temperature, $T_{\rm dust} = 19.6$K or $T_{\rm CIB} = 25$K. The latter is computed as $f(h\nu/k_BT_{\rm CMB}) = h\nu/k_BT_{\rm CMB}\coth(h\nu/2k_BT_{\rm CMB}) - 4$. The templates for the tSZ and kSZ effects are obtained from Reichardt et al. (2020).

We will adopt the same modeling of polarization foreground contamination in TE and EE by adding contributions from Poisson-distributed sources and polarized galactic dust, with mathematical formulation equivalent to temperature, but independent nuisance parameters. We will also ignore higher-order statistics of the foreground, namely the four-point contribution to the covariance.

The theoretical mean is obtained by combining the CMB and foreground components. Since the data band powers have already been corrected for instrumental effects, the theoretical prediction involves convolving the theoretical power spectrum with the window function. The window function is derived by binning the `PolSpice` kernel and can be expressed as follows:

$$\mathcal{C}^{\rm th}(\vartheta) = W\left(C^{\rm cmb}(\vartheta) + C^{\rm foregrounds}(\vartheta)\right), \quad \text{with} \quad {}_sW_{b\ell} = \sum_{\ell'} P_{b\ell'\,s}K_{\ell'\ell}. \tag{4.3}$$

Additional effects such as calibration or aberration need additional work to be included. The SPT-3G 2018 results paved the way for the choices we will make in this analysis.

## 4.2 Covariance matrix

The covariance matrix plays a crucial role in characterizing the statistical properties of our estimator. It captures the variance of the band powers as well as the correlations between different modes. A precise estimation of the covariance matrix is vital to ensure unbiased parameter estimation and accurate interpretation of the data (Hartlap et al., 2007; Dodelson and Schneider, 2013; Sellentin and Starck, 2019; Beck et al., 2022). The final covariance matrix is obtained by combining the signal and noise covariance matrices.

**State-of-the-art** CMB experiments adapt their covariance estimation depending on their noise structure and mask. On one hand, the *Planck* covariance is estimated using an analytical approximation described in Efstathiou (2004); Challinor and Chon (2004). The signal part of the covariance is obtained from the theoretical power spectrum and the mask, while the noise part is correctly integrated by evaluating the impact of noise statistical anisotropies on the covariance. Appendix C of Planck Collaboration et al. (2016) provides more details on this method. The large sky fraction covered by *Planck* allows for minimal coupling between modes of the estimated power spectrum, making the use of the approximation method feasible. However, I will demonstrate that this assumption is less accurate for small sky footprints. Additionally, as discussed in Section 3.6.1, the presence of holes in the mask is expected to introduce additional variance. The *Planck* experiment's point source density results in minimal contamination, which is corrected using a rescaling factor computed from simulations. This approach is reasonable because the maximum deviation is on the order of 15% and can be effectively captured by simulations. On the other hand, ACT DR4 uses multiple subfields with varying map depths, and the covariance is computed using realistic simulations to obtain a robust estimate with low Monte Carlo variance. Analytical estimations are employed for off-diagonal terms, as described in Section 6.4 of Choi et al. (2020). It is important to note that both experiments utilize maximum-likelihood map-making, eliminating the need to account for filtering at the covariance level.

For SPT-3G 2018, the signal covariance matrix is estimated through simulations that fully account for instrument effects and the analysis pipeline, as outlined in Henning et al. (2018); Dutcher et al. (2021). However, conducting a comprehensive simulation suite is computationally demanding due to the complexity of generating input skies and performing mock map-making. The limited availability of simulations can potentially result in a biased covariance matrix (Dodelson and Schneider, 2013). While regularization techniques can be employed to mitigate this issue, their effectiveness relies on assumptions of a common correlation structure between covariance blocks. In SPT-3G 2018, the correlation structure is dominated by the flat-sky projection coupling. For SPT-3G 19/20, we use curved sky representations of the maps. The correlation structure is now dominated by mask effects, and the assumption of common correlation does not hold anymore. In addition, as noise levels decrease, the precision of the signal covariance estimation has a stronger impact on parameters. For this analysis, I will derive the signal covariance matrix analytically. This approach allows for precise modeling of the signal covariance matrix without relying on simulations and models accurately the variance and the coupling between modes. Moreover, the computation is cheap and the input ingredients can be easily adjusted to accommodate variations in the model. However, this computation presents challenges as every step of the analysis needs to be accurately modeled. In Section 4.2.1, I will detail the baseline computation of the covariance matrix given a power spectrum and a mask, and in Sections 4.2.2 and 4.2.3, I will detail the additional implementations required to build the full covariance.

### 4.2.1 High precision computation of the signal covariance matrix

Analytical approximations of the signal covariance matrix have been developed for previous surveys as primary covariance estimators or for verification purposes (Efstathiou, 2004; Challinor and Chon, 2004; Chon et al., 2004; Louis et al., 2017; Planck Collaboration et al., 2020b; Nicola et al., 2021; Friedrich et al., 2021). However, these approaches are based on the assumption that the survey mask is sufficiently large to neglect certain mode couplings. To account for the small sky coverage of the SPT-3G survey, these approximations need to be extended. In addition, it is worth noting that a complete and exact computation of the covariance matrix has never been implemented for a CMB experiment. In our work (Camphuis et al., 2022), we computed, for the first time, the exact covariance matrix and proposed a new approximation for the covariance computation. In this section, we will summarize the main results before including this work.

To enhance clarity, we will use the notation $a^{\mathrm{X}}_{\ell m} \to \mathrm{X}_{\ell m}$ to represent the spherical harmonic coefficients, and for this section, we will disregard the instrumental corrections to the signal, and the binning, which will be addressed in the subsequent section. The covariance matrix of the pseudo-power spectrum estimator, when considering a pure CMB signal observed through a mask $W$, can be expressed as follows (Hansen et al., 2002)

$$\tilde{\Sigma}^{\mathrm{X}\nu\mathrm{Y}\mu;\mathrm{Z}\alpha\mathrm{V}\beta}_{\ell\ell'} \equiv \mathrm{cov}\left(\tilde{C}^{\mathrm{X}\nu\mathrm{Y}\mu}_{\ell}, \tilde{C}^{\mathrm{Z}\alpha\mathrm{V}\beta}_{\ell'}\right) = \frac{1}{(2\ell+1)(2\ell'+1)} \sum_{mm'} \left[\langle \tilde{\mathrm{X}}^{\nu}_{\ell m} \tilde{\mathrm{Z}}^{\alpha *}_{\ell' m'} \rangle \langle \tilde{\mathrm{Y}}^{\mu *}_{\ell m} \tilde{\mathrm{V}}^{\beta}_{\ell' m'} \rangle + \tilde{\mathrm{V}}^{\beta} \leftrightarrow \tilde{\mathrm{Z}}^{\alpha}\right]. \quad (4.4)$$

The second term inside the sum is identical to the first but exchanges the Stokes and frequency indices as indicated. To streamline the notation, let's begin by examining the temperature pseudo-power spectrum at a single frequency, with X = T, temporarily omitting the $\nu$ index. Utilizing the decomposition of pseudo-harmonic coefficients outlined in Eq. (3.41), we obtain the following expression:

$$\langle \tilde{\mathrm{T}}_{\ell m} \tilde{\mathrm{T}}^*_{\ell' m'} \rangle = \sum_{\ell_1 m_1 \ell_2 m_2} \langle \mathrm{T}_{\ell_1 m_1} \mathrm{T}^*_{\ell_2 m_2} \rangle {}_0 I_{\ell m \ell_1 m_1}[W] {}_0 I^*_{\ell' m' \ell_2 m_2}[W] = \sum_{\ell_1 m_1} C^{\mathrm{TT}}_{\ell_1} {}_0 I_{\ell m \ell_1 m_1}[W] {}_0 I^*_{\ell' m' \ell_1 m_1}[W]. \tag{4.5}$$

Subsequently, we can express the covariance matrix of the temperature pseudo-power spectrum as follows:

$$\tilde{\Sigma}_{\ell\ell'} = \frac{2}{(2\ell+1)(2\ell'+1)} \sum_{mm'} \sum_{\ell_1 m_1} \sum_{\ell_2 m_2} C^{\mathrm{TT}}_{\ell_1} C^{\mathrm{TT}}_{\ell_2} {}_0 I_{\ell m \ell_1 m_1}[W] {}_0 I^*_{\ell' m' \ell_1 m_1}[W] {}_0 I^*_{\ell m \ell_2 m_2}[W] {}_0 I_{\ell' m' \ell_2 m_2}[W]. \tag{4.6}$$

The polarization pseudo-power spectra involve a combination of both the EE and BB contributions, requiring the consideration of additional combinations. However, for our parameter constraints, we will solely utilize the EE power spectrum. Given that $C^{\mathrm{BB}} \ll C^{\mathrm{EE}}$ and assuming the EB and TB power spectra are negligible, we can simplify the expression by setting $C^{\mathrm{BB}} = C^{\mathrm{TB}} = C^{\mathrm{EB}} = 0$. We can write the covariance matrix of the pseudo-power spectrum

$$\begin{aligned} \tilde{\Sigma}^{\mathrm{X}\nu\mathrm{Y}\mu;\mathrm{Z}\alpha\mathrm{V}\beta}_{\ell\ell'} = \frac{1}{(2\ell+1)(2\ell'+1)} \sum_{mm'} \sum_{\ell_1 m_1} \sum_{\ell_2 m_2} \Big[ & C^{\mathrm{X}\nu\mathrm{Z}\alpha}_{\ell_1} C^{\mathrm{Y}\mu\mathrm{V}\beta}_{\ell_2} {}_{s(\mathrm{X})} I_{\ell m \ell_1 m_1} {}_{s(\mathrm{Z})} I^*_{\ell' m' \ell_1 m_1} \\ & {}_{s(\mathrm{Y})} I^*_{\ell m \ell_2 m_2} {}_{s(\mathrm{V})} I_{\ell' m' \ell_2 m_2} \\ & + (\mathrm{V}\beta) \leftrightarrow (\mathrm{Z}\alpha) \Big] \quad \forall \mathrm{XYZV} \in \{\mathrm{T}, \mathrm{E}\}. \end{aligned} \tag{4.7}$$

Where $s(\mathrm{T}) = 0$ for temperature and $s(\mathrm{E}) = +$ for polarization, and ${}_+I = \frac{1}{2}({}_2I + {}_{-2}I)$. For simplicity, we have omitted the mask dependence of the coupling coefficients in the notation. The covariance matrix of the pseudo-power spectrum only depends on the underlying power spectrum and the mask. However, the contribution of these two factors becomes intertwined through the coupling coefficients. In the case of a full-sky observation, the coupling coefficients simplify drastically thanks to the relations in Eq. (3.27) with $I_{\ell m \ell' m'}[1] = \delta_{\ell\ell'} \delta_{mm'}$. Consequently, in the full-sky scenario, the modes are uncorrelated, and we can express this by including the Stokes and frequency indices as follows:

$$\tilde{\Sigma}^{\mathrm{X}\nu\mathrm{Y}\mu;\mathrm{Z}\alpha\mathrm{V}\beta}_{\ell\ell'} = \frac{1}{2\ell+1} \delta_{\ell\ell'} (C^{\mathrm{X}\nu,\mathrm{Z}\alpha}_{\ell} C^{\mathrm{Y}\mu\mathrm{V}\beta}_{\ell} + C^{\mathrm{X}\nu\mathrm{V}\beta}_{\ell} C^{\mathrm{Y}\mu,\mathrm{Z}\alpha}_{\ell}). \tag{4.8}$$

We recovered the general formula for cosmic variance as in Eq. (3.40). In the case of partial sky coverage, a widely used approximation for the diagonal elements of the covariance matrix after binning can be obtained by rescaling the previous equation and writing the following formula (Knox, 1995; Hivon et al., 2002):

$$\Sigma^{\mathrm{X}\nu\mathrm{Y}\mu;\mathrm{Z}\alpha\mathrm{V}\beta}_{bb'} \equiv \mathrm{cov}\left( \tilde{\mathcal{C}}^{\mathrm{X}\nu\mathrm{Y}\mu}_b, \tilde{\mathcal{C}}^{\mathrm{Z}\alpha\mathrm{V}\beta}_{b'} \right) \approx \frac{1}{\nu_b} \delta_{bb'} (\mathcal{C}^{\mathrm{X}\nu,\mathrm{Z}\alpha}_b \mathcal{C}^{\mathrm{Y}\mu\mathrm{V}\beta}_b + \mathcal{C}^{\mathrm{X}\nu\mathrm{V}\beta}_b \mathcal{C}^{\mathrm{Y}\mu,\mathrm{Z}\alpha}_b), \tag{4.9}$$

Here, $b$ and $b'$ represent the bin indices, and $\nu_b$ denotes the effective number of degrees of freedom in the sky. Specifically, $\nu_b$ is determined by $(2b+1)\Delta\ell_b f_{\mathrm{sky}} \frac{w_2^2}{w_4}$, where $\Delta\ell_b$ corresponds to the bin width in multipole space. The binning process is crucial as it enables the reduction of coupling between the band powers. This approximation applies to a wide range of masks and provides a reasonable first-order estimation of the error bars if the binning scheme is wide enough. However, in the case of the SPT-3G winter field, the sky coverage is small, encompassing only 4% of the total sky. Consequently, we anticipate a high level of correlation between spectral modes, leading to an accuracy of the Knox formula. In this section, we introduce the work of Camphuis et al. (2022), which proposes a method to accurately compute the covariance matrix of the pseudo-power spectrum. This work not only assesses the accuracy of existing approximations but also describes the procedure for transforming the covariance matrix of the pseudo-power spectrum into the covariance matrix of the `PolSpice` estimator.

**Exact computation** To estimate the precision of the covariance matrix, a large number of simulations or an exact computation is required. The terms of the covariance matrix can be estimated from $n_{\mathrm{sim}}$ simulations with a variance given by (Lueker et al., 2010)

$$\mathrm{var}\ (\tilde{\Sigma}_{\ell\ell'}) = \frac{\tilde{\Sigma}^2_{\ell\ell'} + \tilde{\Sigma}_{\ell\ell} \tilde{\Sigma}_{\ell'\ell'}}{n_{\mathrm{sim}}}. \tag{4.10}$$

The diagonal is dominant for pseudo-power spectrum covariance matrices, and the variance of Monte-Carlo estimated off-diagonal terms is high, necessitating a challenging number of simulations. In our work (Camphuis et al., 2022), we proposed an exact method to compute the covariance matrix. The standard computation scales as $\mathcal{O}(\ell^6_{\mathrm{max}})$, but we can accelerate it using the speed gain from utilizing the `HEALPix` fast harmonic decomposition. The algorithm is based on the observation that the left-hand side of Eq. (4.5) can be computed using harmonic decomposition.

Before any further computations, let us remind the reader that we defined the spherical harmonic transform operator $\mathbb{SHT}$ in Eq. (3.13) acting on a map, and its inverse $\mathbb{SHT}^{-1}$ acting on a set of spherical harmonic coefficients. The dot $\cdot$ refers to the application of the operator to the left of the dot on the object to the right, and $\times$ refers to term-to-term multiplication, which needs to be differentiated from matrix multiplication. Furthermore, $\mathbb{H}$ is an operator defined in Eq. (3.18) and performs the `HEALPix` routine `alm2cl`, applied on two separate sets of harmonic coefficients.

Let's define $\delta^{\ell' m'}$ as the set of harmonic coefficients such that $\zeta^{\ell' m'}_{\ell m} \equiv \delta_{\ell\ell'}\delta_{mm'}$, and extend the power spectrum notation $C_{\ell m} \equiv C_\ell$. We can then write the two-point harmonic function as a series of spherical harmonic transforms

$$\langle \tilde{\mathrm{X}}_{\ell m} \tilde{\mathrm{Y}}^*_{\ell' m'} \rangle = \int \mathrm{d}\hat{\mathbf{n}} \mathcal{Y}_{\ell m}(\hat{\mathbf{n}}) W(\hat{\mathbf{n}}) \mathcal{Y}^*_{LM}(\hat{\mathbf{n}}) C^{\mathrm{XY}}_L \int \mathrm{d}\hat{\mathbf{u}} \mathcal{Y}^*_{LM}(\hat{\mathbf{u}}) W(\hat{\mathbf{u}}) \mathcal{Y}^*_{\ell' m'}(\hat{\mathbf{u}}) \tag{4.11}$$

$$\implies \langle \tilde{\mathrm{X}}_{\ell m} \tilde{\mathrm{Y}}^*_{\ell' m'} \rangle = \left[\mathbb{SHT} \cdot W \times \mathbb{SHT}^{-1} \cdot C^{\mathrm{XY}} \times \mathbb{SHT} \cdot W \times \mathbb{SHT}^{-1} \cdot \zeta^{\ell' m'}\right]_{\ell m}. \tag{4.12}$$

In this equation only, we denoted spherical harmonics with $\mathcal{Y}$ to avoid mixing with the harmonic coefficients. We assume that the spherical harmonic transforms are performed with the appropriate spin-weighting. To simplify this equation, let's introduce a new operator $\mathbb{I}$ defined as $\mathbb{I} \equiv \mathbb{SHT} \cdot W \times \mathbb{SHT}^{-1}$, hence

$$\langle \tilde{\mathrm{X}}_{\ell m} \tilde{\mathrm{Y}}^*_{\ell' m'} \rangle = \left[\mathbb{I} \cdot C^{\mathrm{XY}} \times \mathbb{I} \cdot \zeta^{\ell' m'}\right]_{\ell m}. \tag{4.13}$$

With the help of a suite of `HEALPix` routines, we can implement the previous operation and obtain the set of harmonic coefficients for given indices input $\ell', m'$. The $\ell'$ row of the covariance matrix is then obtained by computing the sum over all $m, m'$ index

$$\tilde{\Sigma}^{\mathrm{X}\nu\mathrm{Y}\mu;\mathrm{Z}\alpha\mathrm{V}\beta}_{\ell\ell'} = \sum_{m'} \mathbb{H}\left\{\mathbb{I} \cdot C^{\mathrm{XZ}} \times \mathbb{I} \cdot \zeta^{\ell' m'}, \mathbb{I} \cdot C^{\mathrm{XZ}} \times \mathbb{I} \cdot \zeta^{\ell' m'}\right\} \tag{4.14}$$

Applying $\mathbb{I}$ with `HEALPix` has a computational cost scaling as $\mathcal{O}(\ell^3_{\max})$, and the sum over the $m'$ index introduces an additional factor of $\ell_{\max}$ to obtain a single row $\ell'$ of the covariance matrix. Performing this operation for multiple rows $\ell'$ implies that the exact covariance computation scales as $\mathcal{O}(\ell^5_{\max})$. This method is exact and does not rely on any approximation, as only relies on the assumption of a homogeneous isotropic Gaussian field. It can be expanded to anisotropic fields and the computation of the polarization B spectrum combinations. However, the computational cost remains high, with around $10^4$CPU-hours for $\ell_{\max} \sim 1000$. In the next section, we will explore approximations that can reduce the computational burden.

**Approximations** In order to reduce the computational complexity associated with exact computations, approximations are commonly employed in the literature. These approximations are valid for masks that cover a significant fraction of the sky but are known to perform less well for smaller masks. In this study, we propose a new approximation specifically tailored for masks of size comparable to the SPT-3G mask. To begin, let us focus on temperature. Starting from Eq. (4.6), we define the renormalized covariance coupling kernel as the sum over the multipole orders $(m, m', \dots)$ of the coupling coefficients. This kernel, denoted as $\bar{\Theta}^{\ell_1\ell_2}_{\ell\ell'}[W]$, can be expressed as follows:

$$\bar{\Theta}^{\ell_1\ell_2}_{\ell\ell'}[W] \equiv \frac{1}{(2\ell+1)(2\ell'+1)\Xi_{\ell\ell'}[W^2]} \sum_{mm_1m'm_2} \left({}_0I_{\ell m\ell_1 m_1}\ {}_0I^*_{\ell' m'\ell_1 m_1}\ {}_0I_{\ell' m'\ell_2 m_2}\ {}_0I^*_{\ell m\ell_2 m_2}\right)[W]. \tag{4.15}$$

The covariance coupling kernel depends on the mask and quantifies the coupling between modes $\ell_1$ and $\ell_2$ for a given pair of modes $\ell$ and $\ell'$. It is symmetric under $\ell_1 \leftrightarrow \ell_2$ or $\ell \leftrightarrow \ell'$ and is normalized such that $\sum_{\ell_1\ell_2} \bar{\Theta}^{\ell_1\ell_2}_{\ell\ell'}[W] = 1$. Finally, the covariance matrix is obtained by summing over the coupling kernel and the power spectrum, as expressed by the equation:

$$\tilde{\Sigma}^{\mathrm{TTTT}}_{\ell\ell'} = 2\Xi_{\ell\ell'}[W^2] \sum_{\ell_1\ell_2} C^{\mathrm{TT}}_{\ell_1} \bar{\Theta}^{\ell_1\ell_2}_{\ell\ell'}[W] C^{\mathrm{TT}}_{\ell_2}. \tag{4.16}$$

The covariance coupling kernel plays a central role in covariance computations. Existing literature leverages its simple structure to propose approximations. When the mask is sufficiently large, the mask power spectrum exhibits rapid variations as the multipole order changes, see Fig. 3.2, while the input CMB power spectrum and the coupling kernel can be approximated as remaining relatively constant. This allows for an approximation of the coupling kernel as a set of Dirac delta functions (Efstathiou, 2004; Friedrich et al., 2021). A more refined approximation utilizes the `MASTER` mode-coupling kernel which captures the convolution effects of the kernel on the power spectrum (Hivon et al., 2002; Nicola et al., 2021). In our work, we exploit the symmetries of the covariance coupling kernel and demonstrate that, given a certain set of $\Delta_{i\in[1,2,3]}$, the following approximation holds

$$\forall\ell,\lambda,\ \bar{\Theta}^{(\ell+\Delta_2)(\ell+\Delta_3)}_{\ell(\ell+\Delta_1)}[W] \approx \bar{\Theta}^{(\lambda+\Delta_2)(\lambda+\Delta_3)}_{\lambda(\lambda+\Delta_1)}[W]. \tag{4.17}$$

The kernel is invariant by simultaneous translation of all four indices. Consequently, it is sufficient to compute the covariance coupling kernel for a specific set of $\ell$ and $\Delta_i$ in order to accurately calculate the covariance matrix. This approximation proves particularly useful when the mask is small. The covariance matrix can be obtained by expressing

$$\Theta^{\ell_1\ell_2}_{\ell\ell'}[W] = \sum_{mm'} \mathbb{H}_{\ell_1}\left\{\mathbb{I}\cdot\delta^{\ell m},\mathbb{I}\cdot\delta^{\ell' m'}\right\}\mathbb{H}_{\ell_2}\left\{\mathbb{I}\cdot\delta^{\ell m},\mathbb{I}\cdot\delta^{\ell' m'}\right\}, \tag{4.18}$$

and relying on the `HEALPix` decompositions for the exact computation. This approximation is compared to others in Fig. 6 of Camphuis et al. (2022) and is shown to be the most accurate for the SPT-3G mask. Details can be found in the article included below. We will refer to this approximation as the ACC (accurate covariance coupling) approximation and use it to compute the covariance matrix of the SPT-3G band powers.

**From pseudo-covariance to `PolSpice` covariance** A final step is required to derive the covariance matrix of the `PolSpice` estimator, which models the mask de-biasing procedure and polarization decoupling. The covariance matrix of the `PolSpice` estimator is obtained by relying on the transformation defined in Eq. (3.59), and we finally obtain

$$\hat{\Sigma}^{\mathrm{X}\nu\mathrm{Y}\mu;\mathrm{Z}\alpha\mathrm{V}\beta} \equiv {}_{s(\mathrm{XY})}G\tilde{\Sigma}^{\mathrm{X}\nu\mathrm{Y}\mu;\mathrm{Z}\alpha\mathrm{V}\beta}{}_{s(\mathrm{ZV})}G^{\top}, \tag{4.19}$$

where we used $s(\mathrm{TT}) = 0$, $s(\mathrm{TE}) = \times$ and $s(\mathrm{EE}) = +$, and defined ${}_{+}G \equiv \frac{1}{2}\left({}_{+2}G + {}_{-2}G\right)$. This correction depends on the mask and should be computed in agreement with the `PolSpice` real space apodization setup chosen for the analysis.

We include Camphuis et al. (2022) in the thesis as published. The notations are consistent with previously defined equations, except for the definitions of the `PolSpice` operator ${}_{+2}G$, which is labeled ${}_{\mathrm{dec}}G$ in the article. It is important to note that this framework fails when the analysis mask includes small-scale features, such as point source masking holes. To mitigate this issue, we introduced the inpainting procedure which is detailed in Section 3.6.

A&A 668, A62 (2022)
https://doi.org/10.1051/0004-6361/202243948

Astronomy & Astrophysics

# Accurate cosmic microwave background covariance matrices: Exact calculation and approximations

E. Camphuis[1], K. Benabed[1], S. Galli[1], E. Hivon[1], and M. Lilley[2]

[1] Sorbonne Université, UMR7095, Institut d'Astrophysique de Paris, 98 bis Boulevard Arago, 75014 Paris, France
e-mail: `etienne.camphuis@iap.fr`
[2] SYRTE, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, LNE, 61 avenue de l'Observatoire, 75014 Paris, France

Received 3 May 2022 / Accepted 29 September 2022

## ABSTRACT

*Context.* A reliable estimation of cosmological parameters from pseudo-power spectrum estimators requires accurate covariance matrices.

*Aims.* We focus on the analytical calculation of covariance matrices. We consider the case of observations of the cosmic microwave background (CMB) in temperature and polarization on a small footprint such as in the South Pole Telescope third-generation (SPT-3G) experiment, which observes 4% of the sky. Power spectra evaluated on small footprints are expected to have strong correlations between modes, and these need to be accurately modeled.

*Methods.* We present for the first time an algorithm that allows an efficient (but computationally expensive) exact calculation of analytic covariance matrices. Using it as our reference, we tested the accuracy of existing fast approximations of the covariance matrix. Furthermore, we propose a new approximation that is designed to be more precise. Finally, we derived the covariance matrices for mask-corrected power spectra estimated by the `PolSpice` code. In particular, in the case of a small sky fraction, we included the effect of the apodization of the large-scale modes.

*Results.* We find that when the power spectrum is binned in wide bandpowers, current approximations of the covariance matrix are correct up to the 5% level on the SPT-3G small sky footprint. Our new approximation improves the previous approximations and reaches a precision of 1% for the wide bandpowers. It is generally more than four times more accurate than current approaches.

*Conclusions.* While we considered the specific case of the CMB, our results are applicable to any other cosmological probe that requires the calculation of pseudo-power spectrum covariance matrices.

**Key words.** cosmic background radiation – cosmology: observations – cosmological parameters – methods: data analysis

## 1. Introduction

One of the most powerful probes of cosmology is the observation of the cosmic microwave background (CMB) anisotropies. The NASA WMAP and the ESA Planck satellite CMB measurements marked the entry into the era of precision cosmology, in which many $\Lambda$ cold dark matter ($\Lambda$CDM) matter cosmological parameters are measured with uncertainties smaller than 1% (Hinshaw et al. 2013; Planck Collaboration V 2020). Ongoing and upcoming ground-based and satellite experiments such as the Atacama Cosmology Telescope (ACT; Aiola et al. 2020), the South Pole Telescope (SPT; Dutcher et al. 2021), the Simons Observatory (SO; Ade et al. 2019), CMB-Stage 4 (CMB-S4; Gallardo et al. 2022), and Litebird (Hazumi et al. 2012) will provide yet more information about the nature of our Universe.

Because primary CMB anisotropies in intensity and polarization are distributed as a Gaussian random field, most of the cosmological information is contained in the angular power spectrum of the CMB anisotropies. As the evolution of the primary anisotropies is linear, the multipoles of the angular power spectrum are uncorrelated when the full sky is observed. However, any realistic experiment requires masking parts of the sky, either to avoid regions that are highly contaminated by foregrounds (e.g., galactic emission or point sources) or because the scanning strategy is designed to observe specific regions of the sky. The estimation of the power spectrum on the masked sky, the so-called pseudo-power spectrum, is biased, and different multipoles become correlated (Hivon et al. 2002). An unbiased estimator of the spectra can then be obtained through the `MASTER` approach (Hivon et al. 2002), as implemented in the `PolSpice`[1] software, for instance (Szapudi et al. 2001; Chon et al. 2004). A robust inference of cosmological parameters requires accurate covariance matrices that describe the variance of the spectra along their diagonal, as well as the correlations between multipoles in the off-diagonal terms. Pseudo-$C_\ell$ covariance matrices are corrected for the effect of the mask using `MASTER` to obtain the covariance matrices for the unbiased $C_\ell$ estimator. Inaccuracies in the covariance matrix estimation can lead to the misestimation of cosmological parameters and of their uncertainties (Dodelson & Schneider 2013; Sellentin & Starck 2019).

Covariance matrices can be calculated through the use of simulations. The number of simulations determines the accuracy of the estimator. As the simulations are expensive to produce, the obtained noisy realization of the covariance has to be regularized (Balkenhol & Reichardt 2022)[2]. Alternatively, it is possible to calculate pseudo-$C_\ell$ covariance matrices analytically. However,

---

[1] `http://www2.iap.fr/users/hivon/software/PolSpice/`

[2] While this work focuses on covariance estimates obtained through empirical estimators, the conditioning schemes it presents can similarly be applied to estimates from simulations.

these depend on integrals whose exact numerical implementation is computationally expensive. Thus, approximations have been proposed in previous works to make these calculations efficient; see, for example, Efstathiou (2004), Nicola et al. (2021) and Friedrich et al. (2021).

We analyze the problem of computing accurate analytical covariance matrices. We take the specific case of the South Pole Telescope third-generation (SPT-3G) experiment, which observes the CMB anisotropies in temperature and polarization on a small sky patch that corresponds to about 4% of the sky. On such a small sky region, the calculated power spectra has strong correlations between multipoles. The existing approximations of the covariance matrix can be less accurate in these conditions. Considering this particular case is thus a particularly stringent test of the validity of analytical algorithms.

We implement for the first time the exact computationally expensive calculation of the covariance matrices, which we find to be numerically feasible at multipoles smaller than $\ell \lesssim \ell_{\rm max,ex} \equiv 1000$ through a new algorithm that gains one order of numerical complexity over the brute-force approach, resulting in a thousandfold speed improvement. Then, we test the existing approximations, and find that they are accurate at the 5% level in the case of the SPT-3G footprint when the power spectrum is averaged in wide bandpowers. We then propose a new approximation that improves the existing algorithms to attain an accuracy of 1% in the same case. Finally, we describe how the covariance matrix of the `PolSpice` $C_\ell$ estimator can be calculated from the pseudo-$C_\ell$ covariance matrix.

While in this work, we focused on the specific case of the SPT-3G CMB experiment, many considerations can be applied more broadly to any probe relying on the calculation of power spectra and covariance matrices. We also highlight that in this work we only consider the signal-signal part of the covariance matrix. Analytical approaches to modeling the noise contribution have already been developed and used in CMB experiments, such as *Planck*; see Appendix C.1 of Planck Collaboration XI (2016). They rely on assumptions on the noise properties, such as isotropy or whiteness. When these assumptions do not hold, the noise-noise and signal-noise components can be obtained directly from the data, as was done in the SPT-3G analysis (Lueker et al. 2010; Dutcher et al. 2021). The integration of the noise part in our exact formalism and new approximation is beyond the scope of this paper.

The paper is organized as follows. In Sect. 2 we introduce the pseudo-power spectrum estimator and its covariance. In Sect. 3 we perform the exact calculation of the covariance matrix. In Sect. 4 we describe the existing approximations for the calculation of the covariance matrix, and we test their accuracy against the exact computation. Section 5 presents our new approximation, which is more accurate. Section 6 describes how the covariance matrix of the `PolSpice` estimator can be calculated. We conclude in Sect. 7, and some detailed calculations are given in the appendices.

## 2. Covariance of the pseudo-power spectrum

### 2.1. Pseudo-power spectrum

Cosmic microwave background anisotropies in intensity and polarization can be described as maps of Stokes parameters $T(\hat n), Q(\hat n), U(\hat n)$ for a direction $\hat n$ of the sky. They are Gaussian random fields, fully characterized by their angular power spectra $(C_\ell^{\rm TT}, C_\ell^{\rm EE}, C_\ell^{\rm BB}, C_\ell^{\rm TE})$, which are the variances of the harmonic coefficients $a^{\rm T}_{\ell m}, a^{\rm E}_{\ell m}, a^{\rm B}_{\ell m}$ obtained by spherical harmonic decomposition of the maps. Cosmological models allow the computation of the expectation of the different power spectra in an ideal full-sky case. However, data only ever cover a part of the sky. We describe the partial coverage with the weight map $W(\hat n)$. The power spectrum of masked maps, labeled $\tilde C_\ell^{\rm XY}$, is usually called throughout pseudo-power spectrum. Its expression for temperature is given in Eq. (A.6). It can be computed from the masked harmonic coefficients $\tilde a^{\rm X}_{\ell m}$, which are directly related to the unmasked ones, $a^{\rm X}_{\ell' m'}$, by the application of the mode-coupling kernels ${}_sI_{\ell m \ell' m'}[W]$. In the case of temperature, we write

$$\tilde a^{\rm T}_{\ell m} = \sum_{\ell' m'} a^{\rm T}_{\ell' m'}\; {}_0I_{\ell m \ell' m'}[W], \tag{1}$$

where we have defined the mode coupling kernels[3]

$${}_sI_{\ell m \ell' m'}[W] \equiv \int \mathrm{d}\hat u\, {}_sY_{\ell m}(\hat u) W(\hat u)\, {}_sY^*_{\ell' m'}(\hat u). \tag{2}$$

These coupling kernels are an important component in the following discussions[4]. In the full-sky case, the orthonormality properties of spin-weighted spherical harmonics ensure that ${}_sI_{\ell m \ell' m'}[1] = \delta_{\ell\ell'}\delta_{mm'}$. We recall in Appendix A some summation properties of products of coupling matrices that appear in the computation of pseudo-power spectra and their covariance. In particular, they are related to the symmetric coupling kernel acting on a power spectrum $\mathcal{A}$, labeled $\Xi[\mathcal{A}]$, with

$$\Xi^{ss'}_{\ell\ell'}[\mathcal{A}] \equiv \sum_L \frac{2L+1}{4\pi}\mathcal{A}_L \begin{pmatrix} \ell & \ell' & L \\ s & -s & 0 \end{pmatrix}\begin{pmatrix} \ell & \ell' & L \\ s' & -s' & 0 \end{pmatrix}. \tag{3}$$

This operator, introduced in Efstathiou (2004), can also be seen as acting on a map $A$ with power spectrum $\mathcal{A}_\ell$. In the following, we use the notation $\Xi[A] \equiv \Xi[\mathcal{A}]$. We recall that the average of the pseudo-spectrum is related to the underlying power spectrum by the application of the asymmetric coupling kernel computed for the mask $W$, also known as the `MASTER` mode-coupling matrix $M$. In the case of temperature, we have

$$\langle \tilde C_\ell^{\rm TT} \rangle = \sum_{\ell'} {}_0M_{\ell\ell'}[W] C_{\ell'}^{\rm TT} \tag{4}$$

$${}_0M_{\ell\ell'}[W] \equiv (2\ell'+1)\Xi^{00}_{\ell\ell'}[W]. \tag{5}$$

In this work, without loss of generality, we develop the computations for the intensity case (i.e., $s = s' = 0$), the polarization case being similar. We also assume that a single mask is used for both temperature and polarization. When required, we highlight the differences between the temperature and polarization cases and give insight into the importance of the single-mask assumption.

### 2.2. Covariance

Estimating the covariance of the measured power-spectrum is crucial to assess the agreement between data and model predictions and to constrain cosmological parameters from CMB maps. As discussed in Hivon et al. (2002) and demonstrated in Appendix A, masking breaks the statistical isotropy and induces

---

[3] The complex conjugate is denoted with a star.

[4] This coupling matrix is often denoted $K$ in the literature, such as in Hivon et al. (2002). In this work, we modified the notation for it to be consistent with the notation of Sect. 6.

correlations between the modes of the pseudo-spectrum. The details of the derivation of the analytical expression of the pseudo-spectrum covariance can be found in Appendix B. We give here the expression in terms of the coupling matrices ${}_0I$ and the true underlying intensity power spectrum $C_\ell$, for the temperature case,

$$\begin{aligned}\tilde{\Sigma}_{\ell\ell'} &\equiv \mathrm{cov}(\tilde{C}_\ell, \tilde{C}_{\ell'}),\\ &= \frac{2}{(2\ell+1)(2\ell'+1)} \sum_{mm'} \sum_{\ell_1 m_1} \sum_{\ell_2 m_2} C_{\ell_1} C_{\ell_2}\\ &\quad \times {}_0I_{\ell m \ell_1 m_1}[W]\, {}_0I^*_{\ell' m' \ell_1 m_1}[W]\, {}_0I^*_{\ell m \ell_2 m_2}[W]\, {}_0I_{\ell' m' \ell_2 m_2}[W].\end{aligned} \tag{6}$$

As shown in Eq. (1), the mode-coupling coefficient ${}_0I$ kernels relate the underlying harmonic coefficients to the harmonic coefficients measured on the sky through the mask. In the analytic expression of the covariance, they represent the coupling between modes due to partial sky coverage. An expression similar to Eq. (6) can be written for polarization, using spin-2 spherical harmonics, that is, $s, s' = \pm 2$. These expressions mix the EE and BB power spectra.

The expression in Eq. (6) involves several convolutions, and its evaluation is computationally expensive. The full computation scales as $\mathcal{O}(\ell^6_{\max})$, $\ell_{\max}$being the largest multipole, making the exact computation of this covariance a daunting task given the currently available computation power. We have developed an algorithm that allows the computation of the covariance matrix at low multipoles with a gain of an order of magnitude in computational time. We discuss this result in Sect. 3.

This approach was previously unavailable, therefore existing work has relied on approximations of Eq. (6). In Sect. 4 we present different approximations that have been proposed in previous works, and we then validate them against our full computation. This validation was performed for a small survey footprint, where spectral modes are highly correlated. These correlations can challenge the assumptions made in the different approximations. Throughout this work, we use a test-case inspired by SPT-3G. The footprint of the first year of the survey presented in Dutcher et al. (2021) covers roughly 4% of the sky and is displayed in Fig. 1 along with the mask power spectrum $\mathcal{W}_\ell$. We apodized the mask with a Gaussian window function of 30 arcmin full width at half maximum, using an algorithm similar to the one used in *Planck* (Planck Collaboration VI 2020). We also show in Fig. 1 the power spectrum of one of the masks used in the *Planck* cosmological analysis, which covers a much larger patch of the sky, around 70% before apodization. The precision of the standard approximation of the covariance was validated in the latter case, but it needs to be assessed for a smaller survey area.

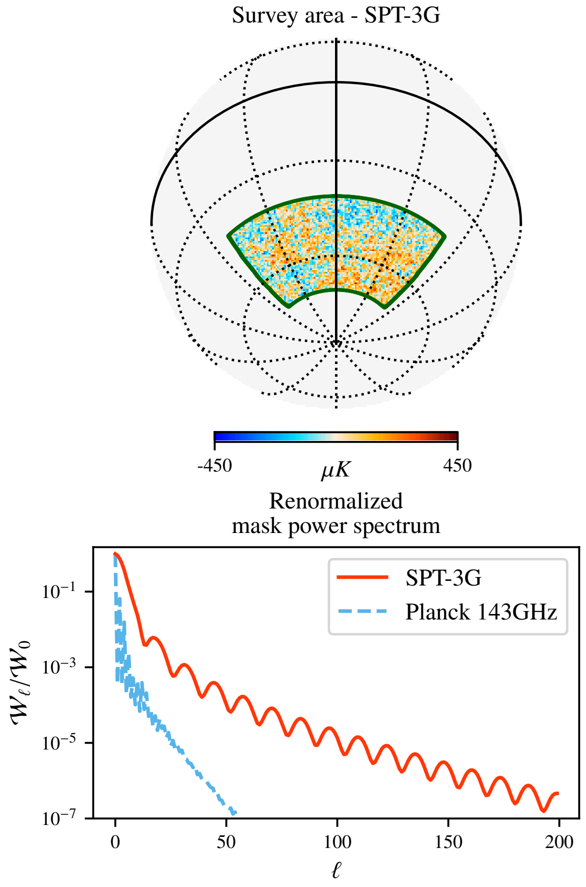


**Fig. 1.** Survey area and the mask power spectrum. *Top panel*: CMB temperature anisotropies on the SPT-3G patch in galactic coordinates. The dark green line delimits the survey footprint. The vertical and horizontal bold black lines are the zero-longitude and zero-latitude coordinates, respectively. The SPT-3G patch covers roughly 4% of the sky. *Bottom panel*: mask power spectra as defined in Eq. (A.4) for SPT-3G and for the 143 GHz map used in the *Planck* cosmological analysis, which covers around 70% of the sky. The spectra have been renormalized by their first value for comparison purposes. Masks corresponding to small sky fractions, such as that of SPT-3G, have a shallower power spectrum than large ones. We emphasize that the mask used here does not include point-source masking.

## 3. Exact computation

An exact calculation of the pseudo-power spectrum covariance matrix can be obtained by integrating Eq. (6). We propose an algorithm that performs the computation in $\mathcal{O}(\ell^5_{\max})$, typically gaining a thousandfold speed-up compared to the direct implementation in $\mathcal{O}(\ell^6_{\max})$. This is achieved with the fast harmonic transform tools implemented in the `HEALPix` library[5]. It enables the exact computation of the covariance matrix, albeit on a limited range of multipoles. In this work, we have computed the full covariance up to $\ell_{\max,\mathrm{ex}} \equiv 1000$, and calculated a few rows of the matrix at $\ell > \ell_{\max,\mathrm{ex}}$. This allows the direct comparison of the various analytic covariance approximation formulae with the exact calculation.

In the following, we describe the algorithm we developed to perform this computation. We validate it with Monte-Carlo estimates of the covariance for the reference SPT-3G survey.

### 3.1. Algorithm description

We focus on the computation of a given row of the covariance matrix $\tilde{\Sigma}$. This allows us either to compute a full covariance matrix at low multipoles or to test our approximations on a selection of rows. We first start by describing the computation of the covariance of the intensity spectrum. A diagrammatic implementation of our calculation is presented in Fig. 2. For the polarization spectra, the calculation follows a similar pattern, and the difference between the two cases are discussed in the next section.

---

[5] `https://healpix.sourceforge.io`

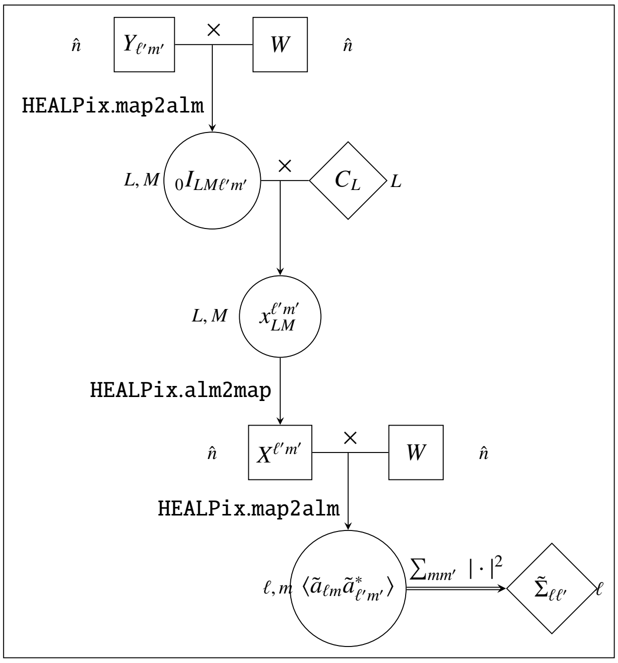


**Fig. 2.** Algorithm used to compute one row of the covariance $\tilde{\Sigma}_{\ell\ell'}$ using the `HEALPix` tools for a fixed $\ell'$ and varying $\ell$. Square, diamond, or circle boxes are arrays representing maps, power spectra, or spherical harmonic coefficients, respectively. Operations are symbolized by arrows and described alongside. The indices of the arrays are indicated on the side of the corresponding boxes. For example, near the top of the diagram, the `HEALPIX.map2alm` operation applied on the product $Y_{l'm'}W$ produces the array ${}_0I_{LMl'm'}$ with indices $L, M$. At the bottom of the diagram, the operations before the final summation produce the array $\langle\tilde{a}_{\ell m}\tilde{a}^*_{\ell'm'}\rangle$ with indices $\ell, m$ for a fixed $\ell', m'$ pair. This part of the algorithm scales as $O(\ell'^3)$ because this is the scaling of `HEALPix` operations (which we applied three times) when the resolution of the map is comparable to the maximum multipole index considered. The last operation, summing over the indices $m, m'$, requires repeating the precedent steps for all $m' \in [-\ell', \ell']$, thus repeating them $2\ell' + 1$ times. Therefore, the final complexity for producing a single row with fixed multipole index $\ell'$ is $O(\ell'^4)$. Computing the covariance matrix for all $\ell'$ then increases the computational time to $O(\ell'^5)$.

We derive in Appendix B the expression of the covariance of the pseudo-spectrum that leads to Eq. (6). In particular, we express the covariance as a sum over $m$ and $m'$ of the square of the correlation $\langle\tilde{a}_{\ell m}\tilde{a}^*_{\ell'm'}\rangle$,

$$\tilde{\Sigma}_{\ell\ell'} = \frac{2}{(2\ell+1)(2\ell'+1)} \sum_{mm'} |\langle\tilde{a}_{\ell m}\tilde{a}^*_{\ell'm'}\rangle|^2. \tag{7}$$

The harmonic coefficients correlation can be written as

$$\begin{aligned}\langle\tilde{a}_{\ell m}\tilde{a}^*_{\ell'm'}\rangle &= \sum_{LM} C_L\, {}_0I_{\ell mLM}\, {}_0I^*_{\ell'm'LM},\\ &= \int \mathrm{d}\hat{u}\, {}_0Y_{\ell m}(\hat{u})W(\hat{u}) \sum_{LM} \left\{C_L\, {}_0I^*_{\ell'm'LM}\right\} {}_0Y^*_{LM}(\hat{u}),\end{aligned} \tag{8}$$

where we have used Eq. (2) to expand one of the ${}_0I$ kernels, reorganized the equation, and used the fact that the power spectrum $C_L$ and the mask $W$ are real quantities (we also dropped the explicit $W$ dependences of the kernel to simplify notations). For fixed $\ell'$ and $m'$, the rightmost part of the equation can be seen as the complex conjugate of the backward spherical harmonic transform of a set of spherical harmonic coefficients into a map $X^{\ell'm'}$, defined as

$$X^{\ell'm'}(\hat{u}) \equiv \sum_{LM} x^{\ell'm'}_{LM} Y_{LM}(\hat{u}), \tag{9}$$

where we defined the spherical harmonic coefficients with

$$x^{\ell'm'}_{LM} \equiv C_L\, {}_0I_{\ell'm'LM}. \tag{10}$$

Here, we emphasize that the map $X^{\ell'm'}$ is a complex map, thus it needs special care when it is decomposed into harmonic coefficients. With these definitions, Eq. (8) reduces to

$$\langle\tilde{a}_{\ell m}\tilde{a}^*_{\ell'm'}\rangle = \int \mathrm{d}\hat{u}\, {}_0Y_{\ell m}(\hat{u})W(\hat{u})X^{\ell'm'*}(\hat{u}), \tag{11}$$

where we recognize the forward harmonic transform of the map $X^{\ell'm'}$, masked by $W$. Thus, a spherical harmonic transform of a masked map can produce the correlation $\langle\tilde{a}_{\ell m}\tilde{a}^*_{\ell'm'}\rangle$ for all $\ell, m$ and a fixed pair of $\ell', m'$. As we discussed, this $X^{\ell'm'}$ map is defined by a set of spherical harmonic coefficients whose expression is given in Eq. (10).

The computation of the $x^{\ell'm'}_{LM}$ coefficients requires the evaluation of the ${}_0I$ kernel. When Eq. (2) is used for a fixed $\ell', m'$, the ${}_0I_{LM\ell'm'}$ kernel can be computed as the spherical harmonic transform of a masked ${}_0Y_{\ell'm'}$ map for all the $L, M$ indices. When everything is added, we see that for a choice of $\ell', m'$, the computation of $\langle\tilde{a}_{\ell m}\tilde{a}^*_{\ell'm'}\rangle$ for all $\ell, m$ reduces to two forward and one backward spherical harmonic transforms, as summarized in Fig. 2.

In practice, these decompositions can be performed with `HEALPix`, which takes advantage of a specific pixelation scheme to make the computation more efficient. This is where the gain announced at the beginning of this section comes from, and it allows us to implement the exact computation. `HEALPix` decompositions typically scale as $O(\ell'^3)$[6]. We repeated the decompositions resulting in $\langle\tilde{a}_{\ell m}\tilde{a}^*_{\ell'm'}\rangle$ for all $m' \in [-\ell', \ell']$ to perform the summation in Eq. (7). Thus, the computation of a single row $\tilde{\Sigma}_{\ell\ell'}$ for all $\ell$ and fixed $\ell'$ scales as $O(\ell'^4)$. Finally, the computation of a full covariance matrix for all $\ell'$ scales as $O(\ell^5_{\mathrm{max,ex}})$.

Additional optimizations can be implemented in the algorithm by degrading maps and running the algorithm at a lower `HEALPix` resolution, $n_{\mathrm{side}}$, for small multipoles. `HEALPix` computations are precise up to $\ell \sim 2n_{\mathrm{side}}$, hence choosing a map resolution on the order of the multipole is sufficient to precisely compute the close-to-diagonal elements of the covariance. This operation requires a degraded version of the mask, which must be computed while avoiding aliasing from small-scale features. This can be done by implementing a hard cutoff of the mask harmonic coefficients before degrading its resolution. This allowed us to compute the exact covariance up to multipole $\ell_{\mathrm{max,ex}} = 1000$. The algorithm requires 300 h of CPU-time to compute a row of the intensity (TTTT) and polarization (EEEE) matrices at multipole $\ell = 950$ with map resolution $n_{\mathrm{side}} = 1024$. It is also well suited to a potential GPU implementation, which could lead to more speed-ups.

---

[6] Details about the computation scaling of `HEALPix` can be found on the website `https://healpix.sourceforge.io` or in Gorski et al. (2005).

### 3.2. Polarization

The polarized case is very similar to the intensity case detailed in the previous subsection. We only describe the EEEE case, which gives a general template to the other polarization and temperature×polarization cases.

The polarized version of Eq. (8) is given by Eq. (6) of Challinor & Chon (2005), which reads

$$\left\langle \tilde{a}^{\mathrm{E}}_{\ell m} \tilde{a}^{\mathrm{E}*}_{\ell' m'} \right\rangle = \sum_{LM} \left[ C_L^{\mathrm{EE}} {}_{+}I_{\ell m LM} {}_{+}I^{*}_{\ell' m' LM} + C_L^{\mathrm{BB}} {}_{-}I_{\ell m LM} {}_{-}I^{*}_{\ell' m' LM} \right], \tag{12}$$

where we defined the Hermitian coupling coefficients

$${}_{\pm}I_{\ell m LM} = \frac{1}{2}\left({}_{+2}I_{\ell m LM} \pm {}_{-2}I_{\ell m LM}\right), \tag{13}$$

with the spin-weighted coupling coefficients ${}_{\pm 2}I$ defined as for intensity, see Eq. (2). Reordering the terms and repeating the same operations as in the previous section, the final harmonic coefficient correlation can be seen as the masked forward spherical harmonic decomposition of two maps $Z_1^{\ell' m'}, Z_2^{\ell' m'}$,

$$\left\langle \tilde{a}^{\mathrm{E}}_{\ell m} \tilde{a}^{\mathrm{E}*}_{\ell' m'} \right\rangle = \frac{1}{2}\left[ \int \mathrm{d}\hat{u} W(\hat{u}) Z_1^{\ell' m'}(\hat{u}) \left({}_{+2}Y^{*}_{\ell m} + {}_{-2}Y^{*}_{\ell m}\right)(\hat{u}) \right. \tag{14}$$
$$\left. - i \int \mathrm{d}\hat{u} W(\hat{u}) Z_2^{\ell' m'}(\hat{u}) \left({}_{+2}Y^{*}_{\ell m} - {}_{-2}Y^{*}_{\ell m}\right)(\hat{u}) \right].$$

The maps $Z_1^{\ell' m'}, Z_2^{\ell' m'}$ are obtained using a backward spherical harmonic decomposition of the coefficients $x_{LM}^{\ell' m';\mathrm{E,B}}$,

$$\left(Z_1^{\ell' m'} - i Z_2^{\ell' m'}\right)(\hat{u}) \equiv \sum_{LM} \frac{x_{LM}^{\ell' m';\mathrm{E}} + x_{LM}^{\ell' m';\mathrm{B}}}{2} {}_{+2}Y_{LM}(\hat{u}), \tag{15}$$

$$\left(Z_1^{\ell' m'} + i Z_2^{\ell' m'}\right)(\hat{u}) \equiv \sum_{LM} \frac{x_{LM}^{\ell' m';\mathrm{E}} - x_{LM}^{\ell' m';\mathrm{B}}}{2} {}_{-2}Y_{LM}(\hat{u}). \tag{16}$$

The set of harmonic coefficients $x_{LM}^{\mathrm{E,B}}$ is defined similarly to the temperature case (Eq. (10)), and it was obtained by filtering the coefficients computed with a masked forward harmonic decomposition of the spin-2 spherical harmonics ${}_{\pm 2}Y_{\ell' m'}$,

$$\begin{cases} x_{LM}^{\ell' m';\mathrm{E}} = C_L^{\mathrm{EE}} {}_{+}I_{\ell' m' LM}, \\ x_{LM}^{\ell' m';\mathrm{B}} = C_L^{\mathrm{BB}} {}_{-}I_{\ell' m' LM}. \end{cases} \tag{17}$$

This algorithm can be extended for any combination of spectra for the other polarization cases, including the cross-correlation between temperature and polarization that we do not treat here.

### 3.3. Validation on simulations

We compared the results of our implementation of the exact computation with a MC estimate of the covariance, obtained with $N_{\mathrm{sim}}$ simulations. The MC covariance terms are expected to be Wishart distributed with $N_{\mathrm{sim}}$ degrees of freedom, as explained in Lueker et al. (2010). We can estimate their variance to be

$$\left\langle \left(\tilde{\Sigma}^{\mathrm{sim}}_{\ell\ell'} - \left\langle \tilde{\Sigma}^{\mathrm{sim}}_{\ell\ell'} \right\rangle\right)^2 \right\rangle = \frac{\tilde{\Sigma}^2_{\ell\ell'} + \tilde{\Sigma}_{\ell\ell}\tilde{\Sigma}_{\ell'\ell'}}{N_{\mathrm{sim}}}. \tag{18}$$

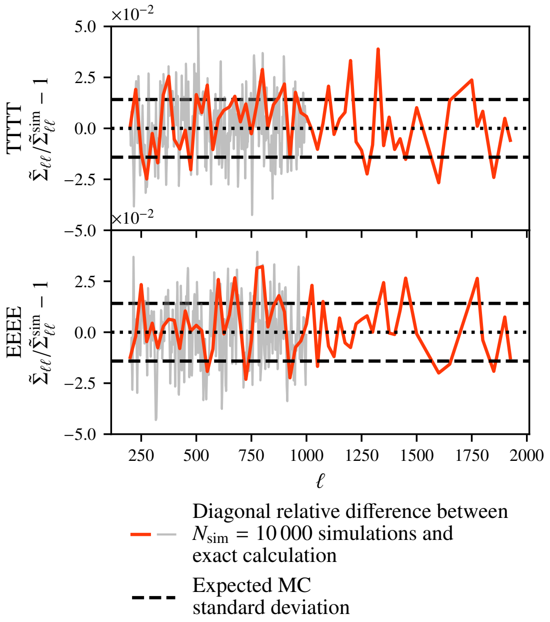


**Fig. 3.** Relative difference of diagonals $\tilde{\Sigma}_{\ell\ell}/\tilde{\Sigma}^{\mathrm{sim}}_{\ell\ell} - 1$ for temperature TTTT (*top*) and polarization EEEE (*bottom*). In red we plot the relative differences every 25 multipoles until $\ell = 1500$ and a few well-chosen ones (at the locations of peaks and troughs of the spectra) up to $\ell = 2000$. In gray, the same quantity is plotted for all multipoles for $\ell \in [\ell_{\mathrm{cut}} = 200, \ell_{\mathrm{max,ex}} = 1000]$. We are able to compute the covariance exactly only for a limited number of rows, and it is computationally cheaper for lower multipoles, justifying our choice of full calculation at $\ell < \ell_{\mathrm{max,ex}}$ and partial calculation for larger multipoles. This plot shows the agreement between the two approaches and validates our exact calculation.

$N_{\mathrm{sim}} = 10\,000$ allows us to reach a percent-level accuracy on the diagonal. This is the number of realizations that we use for the MC covariance. For this validation, we used the mask shown in Fig. 1. In this idealized setting, we did not include a point source mask.

We performed an exact computation of the TTTT and EEEE covariance up to $\ell_{\mathrm{max,ex}} = 1000$, using our algorithm and degrading the mask to smaller resolutions. Furthermore, we computed the rows every 25 multipoles of the matrix up to $\ell_{\mathrm{max}} = 1500$, as well as at a few well-chosen multipoles that correspond to peaks and troughs of the spectra up to $\ell_{\mathrm{max}} = 2000$.

We first focus on the diagonal of the covariance. Figure 3 presents the comparison between the exact computation of the diagonal, obtained by selecting the corresponding value in the rows we computed, and the MC evaluation. The two agree within the MC noise expected for $N_{\mathrm{sim}}$.

For the off-diagonal terms, we show in Fig. 4 a few rows of the exact and MC covariance. The rows agree within the MC noise. The correlation between the modes falls to the percent level within a distance $|\ell - \ell'| \sim 25$ bands around the diagonal. The correlation matrix is defined as the covariance renormalized by its diagonal,

$$\sigma_{\ell\ell'} \equiv \frac{\Sigma_{\ell\ell'}}{\sqrt{\Sigma_{\ell\ell}\Sigma_{\ell'\ell'}}}. \tag{19}$$

We display the exact and MC correlation matrix in the same multipole range in Fig. 5. Our tests demonstrate that our implementation of the exact computation of the pseudo-spectrum

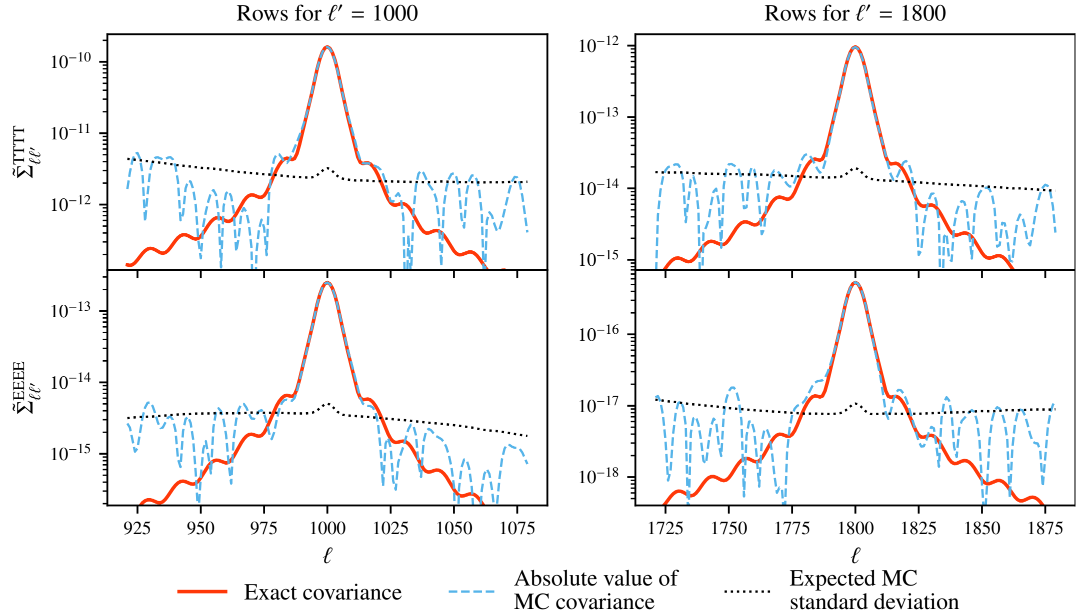


**Fig. 4.** Rows of the exact covariance (solid red) and of the simulated one (dashed blue) for multipole $\ell' = 1000$ (*left-hand side*) and $\ell' = 1800$ (*right-hand side*). The dotted black line shows the expected MC uncertainty on the simulated covariance from Eq. (18), using $N_{\mathrm{sim}} = 10\,000$ realizations. The two approaches agree very well. The MC standard deviation is as large as the covariance off-diagonal terms at $|\ell - \ell'| \sim 25$.

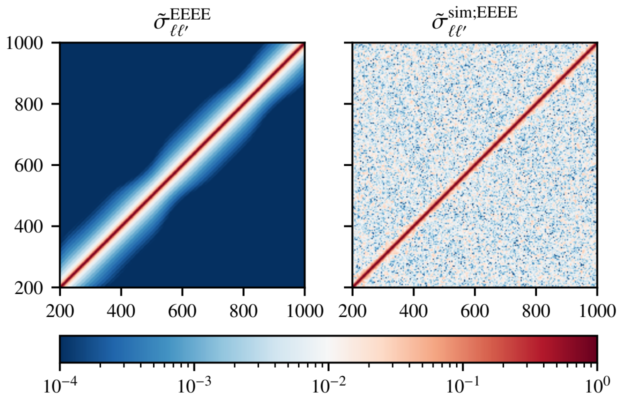


**Fig. 5.** EEEE correlation matrix (see Eq. (19)) obtained from the exact (*left*) and MC (*right*) calculations of the covariance matrix. The exact computation allows us to have the full correlation matrix, but the MC approach is limited by numerical noise. All terms below $10^{-4}$ are plotted in dark blue. Results for TTTT are comparable.

covariance is correct, at least at the level of accuracy that can be reached by MC estimation.

While massive MC estimates as we performed here in this idealized case can produce accurate of-diagonal term estimates, performing numerous MC estimates can be more challenging in the case of a realistic experiment. It requires regularization approaches, where the number of possible simulations is limited by the computational cost of mock-observations.

We also stress that our covariance matrix cannot be directly compared to the one used in Dutcher et al. (2021) because the presence of a point source mask (which we did not include in our simple example), the complications brought about by introducing realistic noise and scanning strategy, and projection effects (the analysis in Dutcher et al. (2021) is performed using a flat-sky approach) can all yield different levels of correlations between the modes.

## 4. Existing approximations and their accuracy on a small patch of the sky

Our algorithm allows us to obtain the exact covariance only for $\ell < 1000$ or for a few rows at higher $\ell'$s due to the expensive computing resources required. The usual analytical approach consists of using approximations of Eq. (6) to decrease the computational cost. In this section, we introduce a new framework to express the approximations of the covariance matrix and use it to list the different methods proposed in the literature. Then, we test and discuss their accuracy against our exact computation.

### 4.1. General framework

Before discussing the approximations of the covariance, we define a few quantities that help relate the various approximations to each other. We rewrite Eq. (6) as

$$\tilde{\Sigma}_{\ell\ell'} = \frac{2}{(2\ell+1)(2\ell'+1)} \sum_{\ell_1\ell_2} C_{\ell_1} \Theta^{\ell_1\ell_2}_{\ell\ell'}[W] C_{\ell_2}, \tag{20}$$

introducing $\Theta^{\ell_1\ell_2}_{\ell\ell'}[W]$. The covariance coupling kernel, defined as the sum over the multipole orders $(m, m', \dots)$ of the coupling coefficients, reads

$$\Theta^{\ell_1\ell_2}_{\ell\ell'}[W] \equiv \sum_{mm_1m'm_2} \left({}_0I_{\ell m\ell_1 m_1}\ {}_0I^*_{\ell' m'\ell_1 m_1}\ {}_0I_{\ell' m'\ell_2 m_2}\ {}_0I^*_{\ell m\ell_2 m_2}\right)[W]. \tag{21}$$

The covariance coupling kernel $\Theta$ represents the coupling between the modes of the theoretical underlying power spectrum

$C_{\ell_i}$ for $i = 1, 2$ depending on the index of the pseudo-covariance $(\ell, \ell')$. We chose to show the two indices related to the covariance $(\ell, \ell')$ as subscripts and the two summing indices $(\ell_1, \ell_2)$ as superscripts. We considered a single-mask temperature case, for which the coupling kernel is symmetric with respect to the exchange of multipole indices $\ell \leftrightarrow \ell'$ or $\ell_1 \leftrightarrow \ell_2$. In the following, we write our results in this case for the sake of simplicity, but they are valid regardless of the choice of single or multiple masks. In the case of spectra obtained from maps with different masks, or in the case of cross-spectra, the kernel is not symmetric. While the results of this work apply to both cases, considering multiple masks increases computing cost.

Using the completeness relations of spherical harmonics, given in Eqs. (A.10) and (A.11), we can write

$$\sum_{\ell_1\ell_2} \Theta^{\ell_1\ell_2}_{\ell\ell'}[W] = (2\ell + 1)(2\ell' + 1)\Xi^{00}_{\ell\ell'}[W^2]. \tag{22}$$

We can now define the reduced covariance coupling kernel as

$$\bar{\Theta}^{\ell_1\ell_2}_{\ell\ell'}[W] \equiv \frac{\Theta^{\ell_1\ell_2}_{\ell\ell'}[W]}{(2\ell + 1)(2\ell' + 1)\Xi^{00}_{\ell\ell'}[W^2]}, \tag{23}$$

for which

$$\sum_{\ell_1\ell_2} \bar{\Theta}^{\ell_1\ell_2}_{\ell\ell'}[W] = 1. \tag{24}$$

With these notations, we can rewrite the covariance as

$$\tilde{\Sigma}_{\ell\ell'} = 2\Xi^{00}_{\ell\ell'}[W^2] \sum_{\ell_1\ell_2} C_{\ell_1} \bar{\Theta}^{\ell_1\ell_2}_{\ell\ell'}[W] C_{\ell_2}. \tag{25}$$

The symmetric mode-coupling kernel $\Xi[W^2]$ provides the purely geometric coupling due to sky masking and is common to all approximations of the covariance. It only depends on the power spectrum of the squared mask, and it is easy to be convinced that up to a normalization, it corresponds to the exact covariance in the case of a constant power spectrum. Its computation scales as $O(\ell^3_{\max})$. This could be improved by noting, as was done by Louis et al. (2020), that at small enough scales, $\Xi[W^2]$ is close to a Toeplitz matrix, allowing us to further reduce the scaling to $O(\ell^2_{\max})$ for a wide range of modes. The sum on the right-hand side of Eq. (25) describes the contribution of the signal power spectrum modulated by the kernel $\bar{\Theta}$, which depends on the mask. This is the sum that all approximations try to simplify, replacing the kernel $\bar{\Theta}$ with a simpler ansatz. In the following, we describe all approximations in terms of this redefinition of the covariance matrix. Since the reduced covariance coupling kernel is normalized, every approximation formulated in this formalism yields the exact covariance for a constant underlying power spectrum $C_\ell = N$.

### *4.2. Approximations*

#### 4.2.1. Narrow-kernel approximation

Based on the observation that the coupling coefficients ${}_0I$ in Eq. (1) are narrow and peak at their first multipole indices $\ell$ or $\ell'$, Efstathiou (2004) introduced the following approximation of Eq. (25), taking the convolving spectra $C_{\ell_i}$, $i = 1, 2$ out of the sum, and replacing them by the power spectrum evaluated at the first multipole index of the coupling coefficients (i.e., the covariance indices of $\bar{\Theta}$), $C_\ell$ or $C_{\ell'}$. Following the notation introduced in García-García et al. (2019), we refer to this approximation of the covariance as the narrow-kernel approximation (NKA),

$$\begin{aligned}\tilde{\Sigma}_{\ell\ell'} &\approx 2C_\ell C_{\ell'} \Xi^{00}_{\ell\ell'}[W^2] \sum_{\ell_1\ell_2} \bar{\Theta}^{\ell_1\ell_2}_{\ell\ell'}[W] \\ &= 2C_\ell C_{\ell'} \Xi^{00}_{\ell\ell'}[W^2] \;\equiv\; \tilde{\Sigma}^{\mathrm{NKA}}_{\ell\ell'}.\end{aligned} \tag{26}$$

In terms of the reduced covariance coupling kernel, the NKA uses

$$\bar{\Theta}^{\ell_1\ell_2}_{\ell\ell'}[W] \approx \bar{\Theta}^{\ell_1\ell_2;\mathrm{NKA}}_{\ell\ell'}[W] \equiv \frac{\delta_{\ell\ell_1}\delta_{\ell'\ell_2} + \delta_{\ell'\ell_1}\delta_{\ell\ell_2}}{2}. \tag{27}$$

The approximation is exact for the full sky. It provides an accurate estimator whenever the underlying power spectrum $C_\ell$ varies slowly as a function of $\ell$ compared to the typical size of the operators ${}_0I$. This condition is fulfilled when the amplitude of the mask power spectrum drops quickly with multipole $\ell$, which is the case for large sky fractions observed with a mask that contains no small-scale features. This is shown in Fig. 1, for example, where we plot the power spectrum of one of the masks used in the *Planck* analysis. In this case, the above approximation holds for multipoles much larger than those for which the mask spectrum contains power.

The NKA was first introduced in intensity by Efstathiou (2004) and extended to polarization in Challinor & Chon (2005). As in the temperature case, the approximated covariances in polarization are expressed as a function of the polarization spectra EE and BB and the symmetric coupling kernels $\Xi^{\pm2,\pm2}_{\ell,\ell'}$. The expressions of the approximated polarization covariances mix EE and BB due to leakage that appears because the sky is masked; see Eqs. (25)–(27) of Challinor & Chon (2005).

The NKA has been widely used, for instance, in the *Planck* cosmological analysis, which masked only small portions of the full sky; see Planck Collaboration XI (2016). However, it has never been thoroughly tested on small sky fractions. As shown in Fig. 1, the mask power spectrum in the case of the small survey footprint of SPT-3G drops much more slowly than the large *Planck* one. From this observation, we expect the mode-coupling kernels ${}_sI$ to be wider, as can be deduced from Eq. (2). As a result, the theoretical underlying spectrum $C_\ell$ might not be treated as constant compared to the covariance coupling kernels in the sums of Eq. (25), and SPT-3G may be outside the regime of validity of the NKA assumption. This is tested at the end of this section. We now list some proposed improvements to the NKA.

#### 4.2.2. Friedrich approximation

A straightforward extension of the NKA has been proposed in Friedrich et al. (2021). It is based on the observation that the reduced covariance-coupling kernel $\bar{\Theta}$ has four maxima at $[\ell = \ell_1, \ell' = \ell_1]$, $[\ell = \ell_1, \ell' = \ell_2]$, $[\ell = \ell_2, \ell' = \ell_1]$, and $[\ell = \ell_2, \ell' = \ell_2]$. This suggests the following form of the reduced covariance coupling matrix:

$$\bar{\Theta}^{\ell_1\ell_2}_{\ell\ell'}[W] \approx \bar{\Theta}^{\ell_1\ell_2;\mathrm{FRI}}_{\ell\ell'}[W] \equiv \frac{\delta_{\ell\ell_1} + \delta_{\ell'\ell_1}}{2}\,\frac{\delta_{\ell\ell_2} + \delta_{\ell'\ell_2}}{2}. \tag{28}$$

Thus, the approximated covariance is

$$\tilde{\Sigma}^{\mathrm{FRI}}_{\ell\ell'} \equiv 2\Xi^{00}_{\ell\ell'}[W^2]\left(\frac{C_\ell + C_{\ell'}}{2}\right)^2. \tag{29}$$

We refer to this approximation as FRI in the rest of the article.

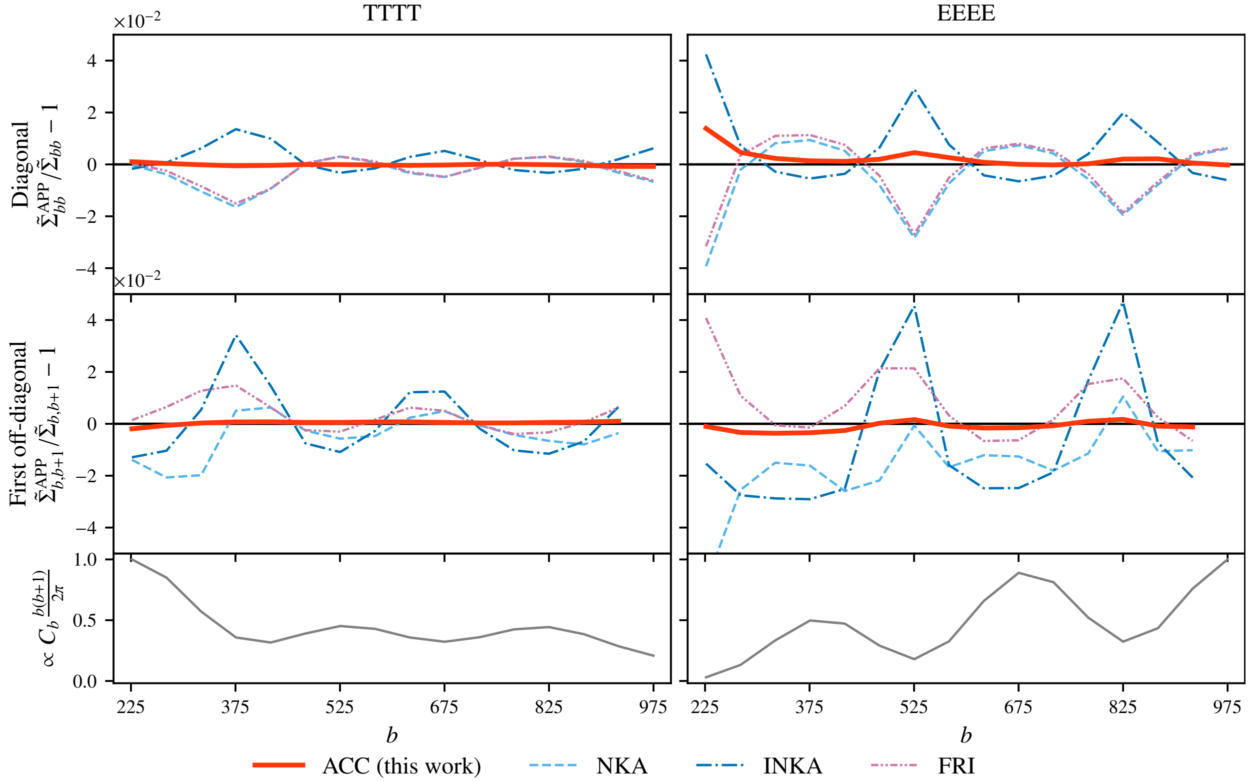


**Fig. 6.** Relative differences of binned approximations with respect to the exact binned covariance: $\tilde\Sigma^{\mathrm{APP}}_{bb'}/\tilde\Sigma_{bb'}-1$ for TTTT (*left*) and EEEE (*right*), with binning $\Delta\ell = 50$ . In the *first row*, we plot the relative differences for the diagonal, i.e., $b = b'$, while in the *second row*, we plot those of the first off-diagonal, i.e., $b' = b+1$. The NKA (light blue dashed), FRI (purple dashed-double-dot) and INKA (dark blue dashed-dot) approximations are accurate at the 5% level, whereas the ACC approximation (solid red) reaches the 1% level, both in intensity and polarization for multipoles larger than $\ell_{\mathrm{cut}} = 200$. The relative differences are plotted for bins that include multipoles up to $\ell_{\mathrm{max,ex}} = 1000$ because it is the maximum multipole for which we computed all the rows of the exact covariance. The third row displays the corresponding binned underlying renormalized spectrum TT or EE to show that the difference of the covariances lies in the peaks and troughs of the spectra.

### 4.2.3. Improved narrow-kernel approximation

Nicola et al. (2021) proposed an improved version of the NKA, the improved narrow-kernel approximation (INKA). In this approximation, the Dirac functions in Eq. (27) are replaced by ${}_0\bar M$, the renormalized `MASTER` mode-coupling kernel, as defined in Appendix A.3. It reads

$$\bar\Theta^{\ell_1\ell_2}_{\ell\ell'}[W] \approx \bar\Theta^{\ell_1\ell_2;\mathrm{INKA}}_{\ell\ell'}[W] \equiv \frac{{}_0\bar M_{\ell\ell_1}\,{}_0\bar M_{\ell'\ell_2} + {}_0\bar M_{\ell'\ell_1}\,{}_0\bar M_{\ell\ell_2}}{2}. \tag{30}$$

The convolution in Eq. (6) indeed averages the power spectra $C_{\ell_i}, i = 1,2$ over multipoles close to $\ell$ and $\ell'$. We can take advantage of this by replacing the convolution by a multiplication with a smoothed power spectrum. When $\bar C \equiv \bar M C$, the resulting covariance can be written as

$$\tilde\Sigma^{\mathrm{INKA}}_{\ell\ell'} \equiv 2\bar C_\ell \bar C_{\ell'}\, \Xi^{00}_{\ell\ell'}[W^2]. \tag{31}$$

All the NKA, FRI, and INKA scale as $O(\ell^3_{\mathrm{max}})$, which are the computing resources needed to obtain the coupling kernels $\Xi$ and $\bar M$. This is a significant improvement over the $O(\ell^5_{\mathrm{max}})$ scaling of our full computation. We now validate the approximations in the case of small surveys using the expensive exact computation of the covariance matrix.

### 4.3. Accuracy

We tested the accuracy of the NKA, FRI and INKA using the exact computation in the case of the SPT-3G small survey footprint shown in Fig. 1. For this mask, the correlations between modes are significant, as we showed in Fig. 4. In this case, it is customary to bin the individual multipoles into wider bandpowers. For this reason, we performed all of our tests on a binned version of the covariance. Given the shape of the power spectrum of the mask and the correlations that we expect from it, we adopted a $\Delta\ell = 50$ binning with $\ell(\ell+1)/(2\pi)$ weights to flatten the dynamics of the spectra in each bin. With this bin size, we expect that most of the correlations between bandpowers are concentrated in the first off-diagonal bin. We also conservatively excluded the first $\ell_{\mathrm{cut}} = 200$ multipoles from our analysis. They are more challenging to measure on a small survey footprint as they can suffer from leakage from the super-survey scales. We restrict our comparison to the multipoles between $\ell_{\mathrm{cut}}$ and $\ell_{\mathrm{max,ex}} = 1000$, where we have carried out the exact computation of all the matrix rows.

We present in Fig. 6 a comparison between the exact computation and the NKA, FRI, and INKA for the diagonal and first off-diagonal of the TTTT and EEEE binned covariances. We discuss the performance of our new accurate covariance coupling (ACC) approximation, also shown in the figure, in the next section. The existing approximations provide good

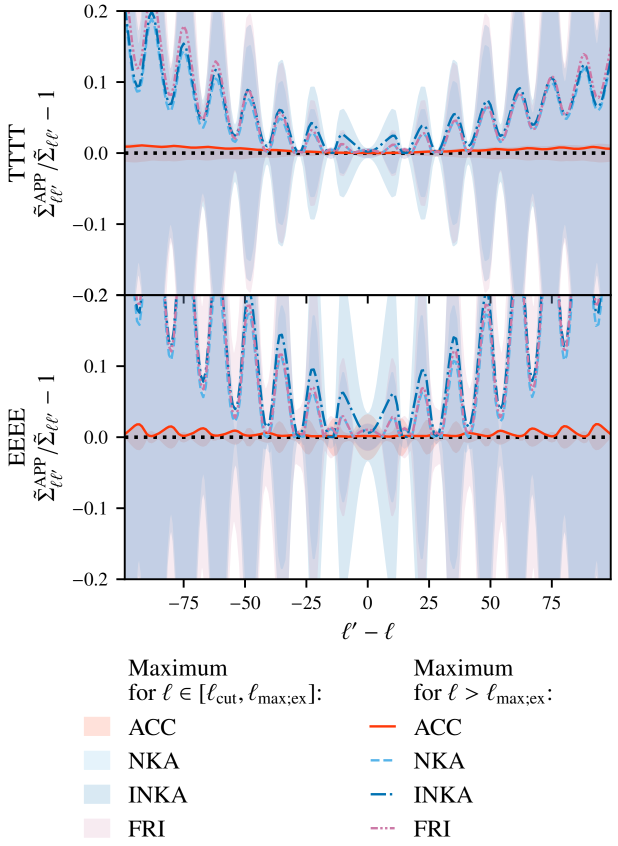


**Fig. 7.** Relative difference between the unbinned approximated covariance matrices compared to the exact calculation, $\tilde{\Sigma}^{\rm APP}_{\ell\ell'}/\tilde{\Sigma}_{\ell\ell'} - 1$, as a function of $\Delta = \ell - \ell'$. We show the NKA (light blue), INKA (dark blue), and FRI (purple) approximations for TTTT (*top*) and EEEE (*bottom*). We only show the relative differences for the $\ell$ at which the deviation is largest, with $\ell \in [\ell_{\rm cut} = 200, \ell_{\rm max,ex} = 1000]$ (shaded regions) or $\ell \in [\ell_{\rm max,ex}, 2000]$ (lines).

estimates of these elements of the covariance as they fall within 5% of the accuracy. The amplitude of the errors varies at different multipoles. Even though the FRI and INKA schemes were implemented to improve upon the simple NKA, their errors are of similar amplitude for this choice of binning. However, all approximations fail to recover the binned covariance at the percent level. On the third row, we show that the difference of the covariances lies in the peaks and troughs of the spectra. As the covariance coupling kernel acts as a symmetric convolution on the spectrum (see Eq. (25)), it is sensitive to the second-order derivative of the spectrum rather than the first. An error on the covariance coupling kernels leads to a larger relative difference in the covariance at multipoles where the curvature of the spectrum is maximum. In Sect. 5 we use the knowledge gained from the exact computation to propose an improved approximation scheme.

Because we cannot easily compute the full matrix to present binned results, we only compare some unbinned rows in Fig. 7 at higher multipoles. The shaded regions in this figure give the lowest values of the relative difference for the approximation within multipoles $\ell \in [\ell_{\rm cut} = 200, \ell_{\rm max,ex} = 1000]$ and $\ell' \in [\ell - 2\Delta\ell, \ell + 2\Delta\ell]$. These are the covariance terms for which we can calculate the full binned covariance; their accuracy is shown in Fig. 6.

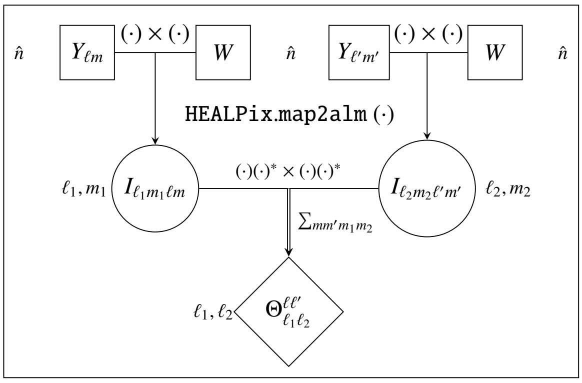


**Fig. 8.** Algorithm for computing the reduced covariance -coupling kernels using `HEALPix` tools. We use the same notation as in the diagram of Fig. 2. The `HEALPix` functions require $O(n_{\rm side}^3)$, with $n_{\rm side}$ the chosen resolution of maps $W$ and $Y_{\ell m}$. As we choose the resolution $n_{\rm side}$ to be on the order of the multipoles indices $\ell, \ell'$, it is equivalent to say that they require $O((\ell + \ell')^3)$. As operations are done $O(\ell + \ell')$ times, the whole operation of computing $\bar{\Theta}_{\ell\ell'}$ is $O((\ell + \ell')^4)$ . Finally, it is clear in this diagram that the computing time of this kernel is at least doubled when multiple masks are used. Different masks would be used as inputs in the first line. As a result, the coupling coefficients would need to be computed for each of the masks, as shown in Eq. (D.3).

Furthermore, the lines in Fig. 7 show the same quantity as the shaded regions, that is, the maximum relative difference, but for multipoles $\ell \in [\ell_{\rm max,ex}, 2000]$, estimated over the sparse number of rows for which we computed the matrix exactly. The difference with the exact covariance for all approximations at large multipoles is always within the same error range as for lower multipoles. This shows that the approximations still work with the same precision at higher multipoles, both for temperature and polarization, and that the accuracy of the approximations in the unbinned case quickly falls below 20% when $\Delta = 50$.

### 4.4. *Structure of the reduced covariance-coupling kernel*

Our expression of the covariance matrix approximations in terms of the normalized coupling kernel $\bar{\Theta}$ in Eq. (27) gives us a very efficient tool for examining the validity of each approximation and for better understanding their differences. We designed an algorithm to calculate this kernel exactly, similar to the one described in Sect. 3 for the exact calculation of the matrix. We show a diagram of the algorithm in Fig. 8.

The reduced covariance-coupling kernel is then displayed in Fig. 9 for the INKA, NKA, and FRI approximations compared to the exact computation for a fiducial multipole $\ell = 200$. The kernels are represented as matrices as a function of $\ell_1, \ell_2$ for different fixed choices of the indices $\ell, \ell'$. Columns show the results for different choices of $\ell' = \ell - \Delta$, with $\Delta = 0$ (i.e., the kernels for the diagonal terms of the covariance matrix, e.g., $\tilde{\Sigma}_{200,200}$), or $\Delta = 10, 50$ (i.e., the kernels for the off-diagonal terms separated by 10 or 50 multipoles, e.g., $\tilde{\Sigma}_{200,190}$). We recall that the reduced kernel is multiplied to $C_{\ell_1}, C_{\ell_2}$ and summed over the indices $\ell_1, \ell_2$ in Eq. (25). Hence, Fig. 9 directly shows the weight of the $\ell_1, \ell_2$ power spectra that contribute to the $\tilde{\Sigma}_{\ell\ell'}$ element of the covariance matrix.

We first focus on the kernels for the diagonal of the covariance matrix, $\Delta = 0$, shown in the first column of Fig. 9. All kernels peak at $\ell_1 = \ell_2 = \ell$, as expected. However, it is clear from the exact calculation that the kernel has a significant width

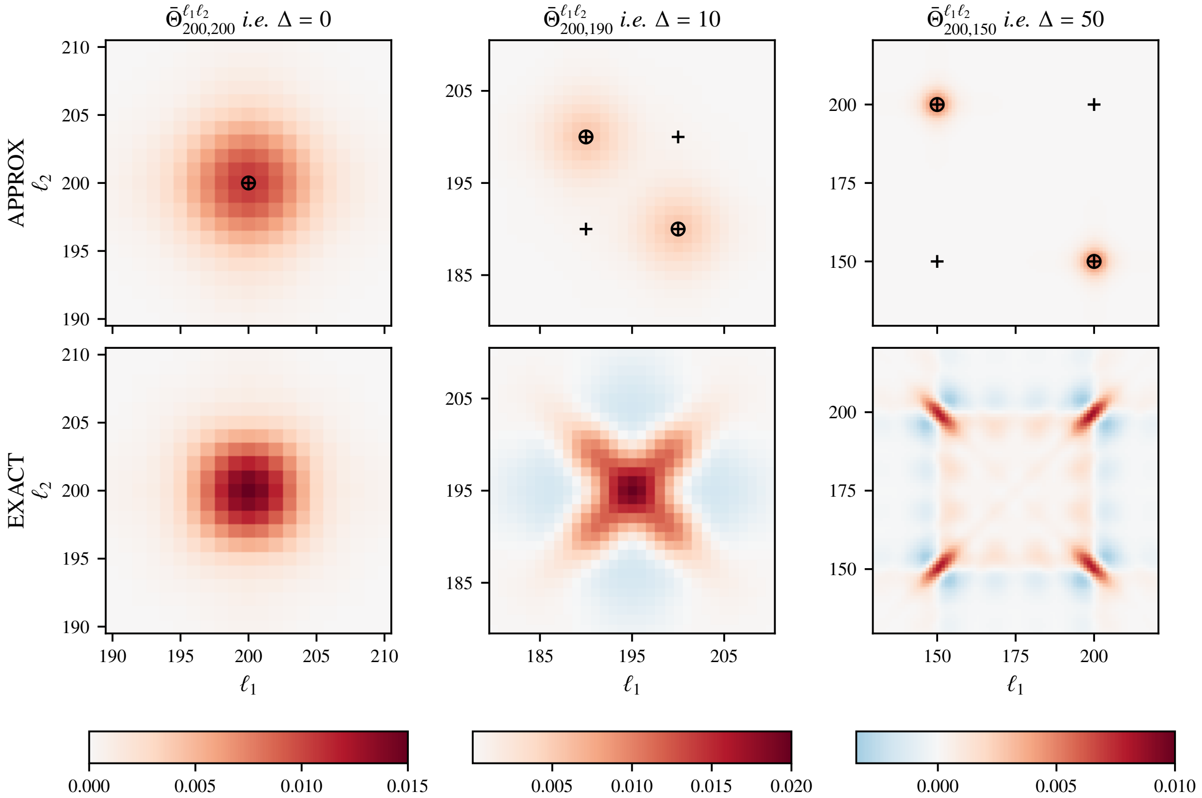


**Fig. 9.** Reduced covariance coupling kernels $\bar{\Theta}^{\ell_1\ell_2}_{\ell\ell'}$ for $\ell = 200$ and $\ell' = \ell + \Delta$, with $\Delta = [0, 10, 50]$ shown in the three different columns. All kernels in one column share the same color map. All of the matrices are shown as a function of $\ell_1$ and $\ell_2$ and are centered on $\ell_1, \ell_2 = (\ell + \ell')/2$. Each of the displayed kernels is properly normalized according to Eq. (22). The plots are restricted to the elements that have a significant value. *Top row*: approximated INKA kernels and the positions at which the delta distributions of the NKA (circles) and FRI (crosses) approximation peak. *Bottom row*: exact kernels. The comparison between the two highlights how much of the structure of the exact kernel is missed by the different approximations.

compared to the CMB power spectrum. This is more clearly shown in Fig. 10, where we plot a slice of the coupling kernels for $\ell_1 = 200$. The width of the kernel cannot be neglected compared to the slope of the CMB power spectrum. This justifies the INKA, which replaces the Dirac $\delta$ functions of NKA and FRI by renormalized mode coupling kernels, see Eq. (31). However, as shown in this figure, the INKA kernel is slightly broader than the exact calculation and its amplitude is smaller. This explains why INKA underestimates the covariance diagonal in the peaks of the power spectrum and overestimates it on the troughs, as shown in Fig. 6: it averages the underlying power spectrum in a wider range of multipoles. Conversely, the NKA/FRI kernels are much thinner than in the exact computation, and so they overestimate the diagonal in the peaks and underestimate it in the troughs of the power spectrum.

Second, we focus on the off-diagonal terms, $\Delta = 10, 50$, shown in the second and third column of Fig. 9. The difference between the exact computation and all the existing approximations is striking, and it is clear that the kernel has more structure than the simple approximated forms. For close off-diagonal terms such as $\Delta = 10$, the true kernel peaks at its central index $\ell_1 = \ell_2 = \bar{\ell} \equiv (\ell + \ell')/2$. For far off-diagonal terms such as $\Delta = 50$, there are four maxima, as predicted by the FRI approximation, which are partially missed by the INKA. Moreover, the true coupling has more dynamics and also covers negative values. Therefore the different approximations, even the INKA approximation, fail to correctly represent the off-diagonal terms of the covariance, as observed in Fig. 7.

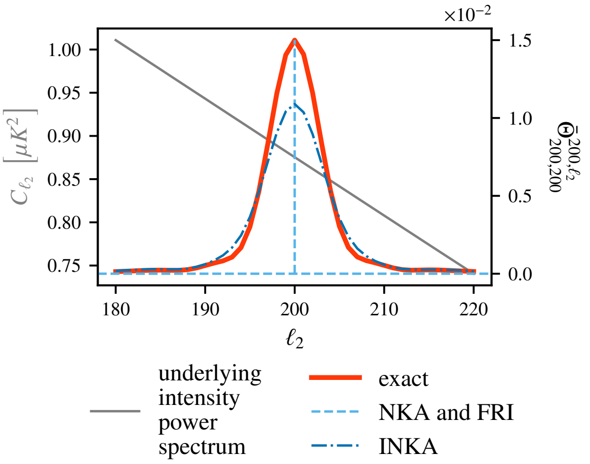


**Fig. 10.** Slices of the exact covariance coupling kernel (solid red) vs. the approximated NKA and FRI (dashed light blue) and INKA (dot-dashed dark blue) ones (*right scale*). We show for comparison the CMB intensity power spectrum (*left scale*, solid gray line).

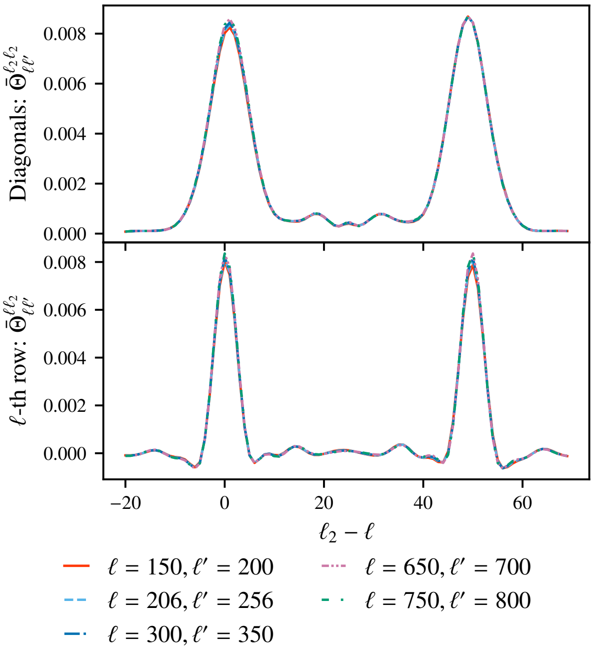


**Fig. 11.** Diagonal (*top*) and row (*bottom*) of the reduced covariance coupling kernels $\bar\Theta^{\ell_1,\ell_2}_{\ell,\ell'}$ for $\ell \in [150, 206, 300, 650, 750]$ and $\Delta = \ell' - \ell = 50$ as a function of $\Delta_2 = \ell_2 - \ell$. The row (*bottom*) is plotted for $\ell_1 = \ell$, i.e. $\Delta_1 = \ell_1 - \ell = 0$. The plots show that for different $\ell$ but the same $\Delta$, the kernels are very similar, differing only at the 5% level. A similar result can be shown for other values of $\Delta$. This leads us to formulate our new approximation, where we assume $\bar\Theta$ to depend only on the multipole separations $\Delta, \Delta_1, \Delta_2$.

## 5. New approximation for the covariance

### 5.1. Improved approach: Approximated covariance coupling

Our ability to calculate the exact reduced covariance coupling matrix $\bar\Theta$, described in Sect. 4, allows us to introduce a new approximation for the computation of the pseudo-power spectrum covariance matrix. We note that for a fixed $\Delta = \ell' - \ell$, the structure of $\bar\Theta_{\ell\ell'}$ appears to be invariant. In other words, the coupling matrices contributing to the $\tilde\Sigma_{\ell\ell'}$ term of the covariance matrix only depend on the distance $\Delta$ from the diagonal. This is demonstrated in Fig. 11, where we plot diagonal and rows of the exact calculation of $\bar\Theta_{\ell,\ell+\Delta}$ for $\Delta = 50$ and different $\ell$, for $\ell_1 = \ell$. The plot reveals that the kernels are nearly identical when they are plotted as a function of $\Delta_2 = \ell_2 - \ell$. We thus infer that in general, when the $\bar\Theta$ matrices are written as a function of $\Delta_1 = \ell_1 - \ell$ and $\Delta_2 = \ell_2 - \ell$, they only depend on $\Delta = \ell' - \ell$ for any $\ell$ and $\ell'$. The difference between kernels computed at different $\ell$s for same $\Delta$ is small, at the 5% percent level. We can thus assume that for any choice of multipole $\ell, \lambda$,

$$\forall(\ell,\lambda),\ \bar\Theta^{(\ell+\Delta_1)(\ell+\Delta_2)}_{\ell(\ell+\Delta)} \approx \bar\Theta^{(\lambda+\Delta_1)(\lambda+\Delta_2)}_{\lambda(\lambda+\Delta)}. \tag{32}$$

An analytical justification of this approximation is provided in Appendix E using the asymptotic expansion of the Wigner-3j symbols when $\ell$ is large.

This suggests a new approximation where the coupling kernel just has to be computed at a given fiducial $\ell$ for all relevant values of $\Delta, \Delta_1$, and $\Delta_2$. We call this the new approximated covariance coupling method, ACC. More precisely, the ACC kernel is given by

$$\bar\Theta^{\ell_1\ell_2}_{\ell\ell'} \approx \bar\Theta^{(\ell+\Delta_1)(\ell+\Delta_2)}_{\ell(\ell+\Delta);\mathrm{ACC}} \equiv \bar\Theta^{(\ell^*+\Delta_1)(\ell^*+\Delta_2)}_{\ell^*(\ell^*+\Delta)}, \tag{33}$$

where we perform the exact and costly computation only for the $\bar\Theta^{(\ell^*+\Delta_1)(\ell^*+\Delta_2)}_{\ell^*(\ell^*+\Delta)}$. We are free to choose the reference $\ell^*$ multipole. Because there are no significant long-range correlations in our case, we can pick a low[7] $\ell^*$ and use a low $n_{\mathrm{side}}$ map resolution. We have to ensure, however, that $\ell^*$ is larger than $\ell_{\mathrm{cut}}$, the low-$\ell$ cutoff that was introduced to avoid issues with large-scale leakages. Close to $\ell_{\mathrm{cut}}$, the exact computation can be used. With a small mask, large-scale modes are difficult to measure and are usually excluded from the cosmological analysis. We can also restrict the range of $\Delta$ to the number of off-diagonal terms of interest in the covariance matrix. The correlation falls quickly (see Figs. 4 and 5), and in practice, we can restrict it to $|\Delta| < d_{\max}$, with $d_{\max}$ being on the order of a few times the correlation length. Similarly, the kernels fall quickly in $\Delta_1, \Delta_2$, so that we can also restrict ourselves to a small region of a similar order, and in the case of the single-mask analysis, use the symmetry around $\Delta \leftrightarrow -\Delta$ to reduce the computational cost.

While we only presented temperature coupling kernels in Fig. 9, the situation is identical in polarization, and a similar approximation can be built; see Appendix E. We used this approximation with $\ell^* = 300$, $n_{\mathrm{side}} = 512$, and $d_{\max} = 100$ to compute the ACC results in Fig. 6.

### 5.2. Accuracy and scaling

We validate the accuracy of our ACC approximation and compare it to the other approximations in Fig. 6. Our new approximation succeeds at estimating the covariance within an error of 1% for all multipoles larger than $\ell_{\mathrm{cut}}$ in intensity and polarization, which is an improvement of a factor of ∼4 over previous approximations. This is also shown in Figs. 7 and 12, which is just the same as Fig. 7, but focused on the EEEE ACC residuals with respect to INKA. This figure shows that the ACC approximation estimates both the diagonal and off-diagonal terms of the covariance matrix far better.

Figure 8 shows the computations needed to obtain the covariance-coupling kernels. Following the same argumentation as for the exact computation of the Sect. 3, we can show that the computation of a single kernel $\bar\Theta$ scales as $O((\ell+\ell')^4)$. As a result, because we need to compute one for each diagonal index $\Delta \in [0, d_{\max}]$, the final ACC approximation scales as $O(n_{\mathrm{side}}^4 d_{\max})$, where $n_{\mathrm{side}}$ is the map resolution chosen to compute the kernels. The computing resources needed to obtain $\bar\Theta$ for all approximations are summarized in Table 1. They add up to the resources needed to compute the symmetric coupling kernel $\Xi$ in Eq. (25). In practice, the kernel $\bar M$ needed to build INKA is often already known for the sky analysis because it has the same structure as $\Xi$, so that the effective complexity for this approximation is $O(1)$.

### 5.3. Point-source mask

We did not include a point-source mask in the survey footprint. Point-source masks significantly complicate the problem as the power spectrum of the mask will have power at large multipoles, hence it will extend the correlation length. This has been an issue for all analysis thus far. Apodizing the point-source masks

---

[7] In the limit where the asymptotic justification of Appendix E remains valid.

**Table 1.** Summary of computation methods to obtain the pseudo-power spectrum covariance matrix.

| Method | Equation numbers | $\bar{\Theta}_{\ell\ell'}^{\ell_1\ell_2}$ | Precision | Complexity |
|---|---|---|---|---|
| Exact (this work) | Eq. (25) | $\bar{\Theta}_{\ell\ell'}^{\ell_1\ell_2}$ computed $\forall \ell, \ell', \ell_1, \ell_2$ | N/A | $\mathcal{O}(\ell_{\max}^5)$ (Using `HEALPix` pixelation) |
| NKA | Eq. (26) | $(\delta_{\ell\ell_1}\delta_{\ell'\ell_2} + \delta_{\ell'\ell_1}\delta_{\ell\ell_2})/2$ | 4% | $\mathcal{O}(1)$ |
| FRI | Eq. (29) | $(\delta_{\ell\ell_1} + \delta_{\ell'\ell_1})(\delta_{\ell\ell_2} + \delta_{\ell'\ell_2})/4$ | 4% | $\mathcal{O}(1)$ |
| INKA | Eq. (31) | $({}_0\bar{M}_{\ell\ell_1}{}_0\bar{M}_{\ell'\ell_2} + {}_0\bar{M}_{\ell'\ell_1}{}_0\bar{M}_{\ell\ell_2})/2$ | 4% | $\mathcal{O}(\ell_{\max}^3)$ or $\mathcal{O}(\ell_{\max}^2)$ with Louis et al. (2020) |
| ACC (this work) | Eq. (33) | $\bar{\Theta}_{\ell\ell'}$ invariant for $\Delta \equiv \lvert\ell - \ell'\rvert$ =cst | 1% | $\mathcal{O}(d_{\max} n_{\mathrm{side}}^4)$ |

**Notes.** First column: name of approximation. Second column: equation to which they are referred. Third column: expression of $\bar{\Theta}$ in this approximation. Fourth column: precision determined by the maximum values of the relative difference of the EEEE binned covariance on diagonal for multipoles $\ell_{\mathrm{cut}} \leq \ell \leq \ell_{\mathrm{max,ex}}$ in Fig. 6. For larger multipoles, the approximation are expected to be in this range of precision as shown in Fig. 12. Fifth column: summary of computing resources needed to obtain $\bar{\Theta}$ in each approximation. Let us here specify that for INKA, the kernel $\bar{M}$ is often already known, thus the practical complexity is $\mathcal{O}(1)$. $\ell_{\max}$ is the multipole range of the covariance, $d_{\max}$ is the number of diagonal computed in the ACC approximation, $n_{\mathrm{side}}$is the resolution chosen to compute the covariance coupling kernels in the ACC approximation (closest to $\ell_{\mathrm{cut}}$ is sufficient).

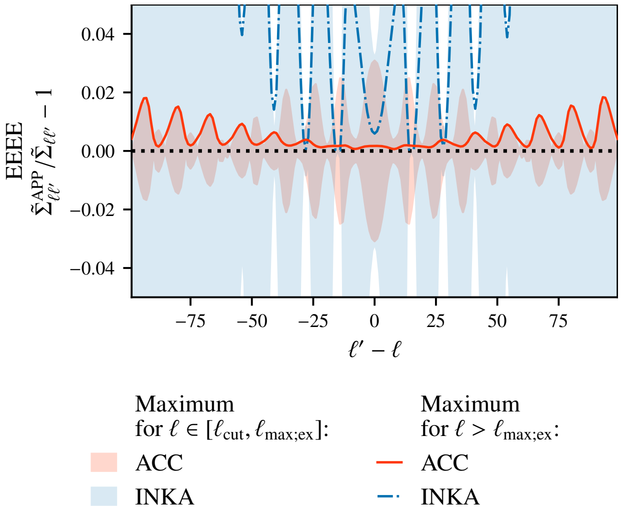


**Fig. 12.** Zoom of Fig. 7. We focus on the relative differences of the ACC and INKA covariance matrices with respect to the exact covariance for EEEE. The deviations found at $\ell < 1000$ (shaded regions) are similar to those found at the higher multipoles (lines), showing that the approximations work at the same level of accuracy in the two cases.

helps to alleviate the problem, but at the price of discarding a significant area of the usable sky. Even in the case of a large survey footprint, such as *Planck*, the point-source masks have been shown to break the NKA. In this case, the issue was mitigated using a simulation-based correction. For FRI, INKA, and ACC, preliminary work that we have performed also suggests that they fail when sources are included. We expect these approximations to perform poorly because the improvements over NKA are focused on the central shape of the reduced covariance coupling matrix $\bar{\Theta}$, while the sources tend to affect the far off-diagonal terms, which increases the correlation between distant multipoles. Particularly for the ACC approximation, the stronger correlations at distant multipoles will break the asymptotic behavior of the Wigner-3j symbols shown in Appendix E. We can expect that this reduces the validity of the approximated invariance by translation along the covariance diagonal, which is at the core of the ACC approximation. More work is required to assess the accuracy of ACC and other approximations in this case. However, different approaches can be adopted to mitigate the effect of a point-source mask. For example, analytical solutions might be found (Gratton et al., in prep.), the maps might be inpainted (Benoit-Lévy et al. 2013), or a MC correction such as the one used in *Planck* might be employed (Planck Collaboration XI 2016).

## 6. Covariance of the `PolSpice` estimator

The pseudo-power spectrum is a biased estimator of the true underlying spectrum of the masked CMB maps. To recover an unbiased estimator, we can apply the `MASTER` (Hivon et al. 2002) formalism, which inverts the mode-coupling matrix and applies it to the biased estimator. Similarly, we can use the `PolSpice` (Szapudi et al. 2001; Chon et al. 2004) algorithm, which corrects the two-point correlation function for the effect of the mask in real space and then converts the result back into harmonic space. However, when the sky footprint is small and no large angular scales are observed, the unbinned mode-coupling matrix becomes singular. Analogously, the `PolSpice` conversion of the two-point correlation function into a power spectrum cannot be performed. In the first case, the mode-coupling matrix must be binned to allow the inversion. In the second case, we must apodize (i.e., cut gradually) the large angular scales of the two-point correlation function before calculating the corresponding power spectrum. This introduces a small bias in the final estimator that cannot be corrected for.

In this section, we explain in detail how the covariance matrix is calculated for the `PolSpice` estimator starting from a pseudo-power spectrum covariance matrix, which we produce through our ACC approximation. We show how the effect of the correction of the mask is included, as well as the small bias introduced by the `PolSpice` apodization of the two-point correlation function. In particular, we show that this apodization can be expressed in harmonic space, allowing us to relate the `PolSpice` spectrum covariance matrix to that of the pseudo-spectrum with a convolution.

### *6.1. MASTER equation*

The pseudo-power spectrum is related to the true spectrum through the well-known `MASTER` equation introduced in Hivon et al. (2002),

$$\left\langle \tilde{C}_\ell^{\mathrm{TT}} \right\rangle = \sum_{\ell'} {}_0M_{\ell\ell'} C_{\ell'}^{\mathrm{TT}}, \tag{34}$$

with similar equations for polarization; see Appendix A.2. This bias comes from the information that is missing due to the masked sky. Given the weighted mask $W(\hat{n})$, we can compute ${}_0M$ using Eq. (A.15). Provided that ${}_0M$ is invertible, an unbiased estimator can be constructed. These relations can be expressed in real space using the two-point correlation functions $\xi$, which for a statistically isotropic sky depend only on the relative angle between two directions,

$$\langle T(\hat{n}_1)T(\hat{n}_2)\rangle = \xi(\arccos(\hat{n}_1\cdot\hat{n}_2)). \tag{35}$$

They can be related to the power spectrum $C_\ell$ using a Legendre series, with

$$\xi(\theta) = \sum_\ell \frac{2\ell+1}{4\pi} C_\ell P_\ell(\cos\theta), \tag{36}$$

$$C_\ell = 2\pi \int_{-1}^{1} \mathrm{d}\cos\theta\, \xi(\theta) P_\ell(\cos\theta). \tag{37}$$

When we define the correlation function $\tilde{\xi}$ of the masked sky in the same manner, associated with the pseudo-spectrum $\tilde{C}_\ell$, we obtain from Eq. (34) by applying the decomposition in a Legendre series the following relation:

$$\langle\tilde{\xi}(\theta)\rangle = w(\theta)\xi(\theta),\ \forall\theta\in[0,\pi], \tag{38}$$

where $w(\theta)$ is the mask angular correlation function (more details can be found in Appendix C). From this relation, we can establish that the `MASTER` mode-coupling matrix ${}_0M$ is invertible only when the correlation function of the mask $w(\theta)$ is nonzero for all $\theta\in[0,\pi]$, which implies that the survey area explores all angular separations on the sky. While this is valid for an almost full-sky analysis such as *Planck*, it does not hold for experiments observing small patches, such as SPT-3G, where angular scales larger than $\theta\sim 30\,\mathrm{deg}$ are unexplored. As a result, ${}_0M$ is not invertible. Binning allows the regularization of the `MASTER` matrix and thus to build a nearly unbiased estimator of the bandpowers. This approach is described in Hivon et al. (2002) and is adopted in `NaMaster`[8] (Alonso et al. 2019). Similarly, we show in the next section how the unobserved large angular scales are handled in the `PolSpice` estimator.

## *6.2. Regularizing with PolSpice*

### 6.2.1. Temperature

The pseudo-power spectrum estimator can be regularized in real space following the `PolSpice` approach in Szapudi et al. (2001). The pseudo-correlation function $\tilde{\xi}$ is smoothed with a scalar apodizing function $f^{\rm apo}(\theta)$, which cuts out large $\theta$ and then corrects for the bias coming from the weighted mask described in Eq. (38). The scalar apodizing function goes smoothly from $f^{\rm apo}(0) = 1$ to $f^{\rm apo}(\theta_{\rm max}) = 0$ to avoid Fourier ringing. $\theta_{\rm max}$ should be chosen as the maximum angular size of the weighted mask. A new correlation function estimator $\hat{\xi}(\theta)$ is defined as

$$\hat{\xi}(\theta) \equiv g(\theta)\tilde{\xi}(\theta), \tag{39}$$

with

$$g(\theta) = \begin{cases} f^{\rm apo}(\theta)/w(\theta) & \forall\theta\in[0,\theta_{\rm max})\ , \\ 0 & \forall\theta\in[\theta_{\rm max},\pi]\,. \end{cases} \tag{40}$$

The function $g$ is well defined and smooth for all angles through the apodization $f^{\rm apo}$. As a consequence of Eqs. (38) and (39), the `PolSpice` estimator of the correlation function can be related on average to the true underlying correlation function with

$$\langle\hat{\xi}(\theta)\rangle = f^{\rm apo}(\theta)\xi(\theta)\ \ \forall\theta\in[0,\pi]\,. \tag{41}$$

Returning to harmonic space using a Legendre transform, this operation can be expressed as

$$\langle\hat{C}_\ell\rangle = \sum_{\ell'} {}_0K_{\ell\ell'} C_{\ell'}. \tag{42}$$

$${}_0K_{\ell\ell'} = (2\ell'+1)\Xi^{00}_{\ell\ell'}[f^{\rm apo}]. \tag{43}$$

The `PolSpice` kernel ${}_0K$ is obtained from the scalar apodizing function $f^{\rm apo}$ with an extended definition of the operators $\Xi$ (see Appendix C and Eq. (C.7) for more details). The operator acts on the Legendre transform of $f^{\rm apo}$.

The advantage of `PolSpice`, which performs the regularization in real space rather than in harmonic space, is that it replaces an $\ell$-space convolution by an integration and a multiplication, which are faster and numerically more stable. This produces an estimator for all multipoles $\ell$. We denote this estimator with a hat, for instance, $\hat{C}^{\rm XY}_\ell$. This regularization (which is only required for small sky patches) introduces a bias in the `PolSpice` estimator that cannot be corrected for because the coupling is not invertible generally as the apodizing function $f^{\rm apo}$ reaches zero. The bias is small because ${}_0K$ is properly normalized, that is, $\sum_\ell {}_0K_{\ell\ell'} = 1$. Furthermore, the regularization increases the correlations between unbinned modes. The `PolSpice` kernels behave as window functions, mixing multipoles of the pseudo-power spectrum. The lack of information at large scales induces the inability to distinguish multipoles that are close to each other. For this reason, the spectrum estimator is binned in ranges larger than the typical correlation between multipoles for a cosmological analysis.

### 6.2.2. Polarization

`PolSpice` allows correcting for the bias introduced by the cut sky in the same manner as for the polarized spectra. It also allows decoupling the EE and BB estimator; see Challinor & Chon (2005) or Appendix C. Similarly to the intensity case, we can express the effect of the `PolSpice` real-space regularization in spherical harmonics by defining the polarized `PolSpice` kernels, ${}_{\pm2}K_{\ell\ell'} = (2\ell'+1)\Xi^{2\pm2}_{\ell\ell'}[f^{\rm apo}]$. The `PolSpice` estimator follows for X ∈ [E, B]

$$\left\langle\hat{C}^{\rm XX}_\ell\right\rangle = \sum_{\ell'} {}_{-2}K_{\ell\ell'} C^{\rm XX}_{\ell'}. \tag{44}$$

For the temperature×polarization case, we can show that

$$\langle\hat{C}^{\rm TE}_\ell\rangle = \sum_{\ell'} {}_\times K_{\ell\ell'} C^{\rm TE}_{\ell'}, \tag{45}$$

$$\text{with } {}_\times K_{\ell\ell'} = (2\ell'+1)\Xi^{20}_{\ell\ell'}[f^{\rm apo}]. \tag{46}$$

## *6.3. Relating PolSpice and MASTER in harmonic space*

We can translate the relations Eq. (39) into harmonic space in temperature and polarization to obtain the `PolSpice` estimator as a harmonic convolution of the pseudo-power spectrum estimator,

$$\hat{C}^{\rm TT}_\ell = \sum_{\ell'} {}_0G_{\ell\ell'}\tilde{C}^{\rm TT}_{\ell'}, \tag{47}$$

---

[8] https://github.com/LSSTDESC/NaMaster

$$\hat{C}^{\mathrm{EE}}_{\ell} - \hat{C}^{\mathrm{BB}}_{\ell} = \sum_{\ell'} {}_{-2}G_{\ell\ell'} \left(\tilde{C}^{\mathrm{EE}}_{\ell'} - \tilde{C}^{\mathrm{BB}}_{\ell'}\right). \tag{48}$$

$$\hat{C}^{\mathrm{EE}}_{\ell} + \hat{C}^{\mathrm{BB}}_{\ell} = \sum_{\ell'} {}_{\mathrm{dec}}G_{\ell\ell'} \left(\tilde{C}^{\mathrm{EE}}_{\ell'} + \tilde{C}^{\mathrm{BB}}_{\ell'}\right), \tag{49}$$

$$\hat{C}^{\mathrm{TE}}_{\ell} = \sum_{\ell'} {}_{\times}G_{\ell\ell'} \tilde{C}^{\mathrm{TE}}_{\ell'}. \tag{50}$$

The $G$ kernels are constructed in the same manner as the PolSpice kernels, with the operator $\Xi$ acting on the function $g = f^{\mathrm{apo}}/w$ according to Eq. (C.7) (or equivalently, on the associated power spectrum of $g$ via Legendre transform). They are given by

$$ {}_{0}G_{\ell\ell'} = (2\ell' + 1)\Xi^{00}_{\ell\ell'}[g], \tag{51}$$

$$ {}_{-2}G_{\ell\ell'} = (2\ell' + 1)\Xi^{2-2}_{\ell\ell'}[g], \tag{52}$$

$$ {}_{\times}G_{\ell\ell'} = (2\ell' + 1)\Xi^{20}_{\ell\ell'}[g], \tag{53}$$

$$ {}_{\mathrm{dec}}G_{\ell\ell'} = \frac{2\ell' + 1}{2} \int_{-1}^{1} g(\theta) d^{\ell}_{22}(\theta) d^{\ell'}_{2-2}(\theta) \mathrm{d}\cos(\theta). \tag{54}$$

The first three equations above reduce to the inverse of the master kernels when the PolSpice apodization function is set to 1, that is, no apodization. The last kernel, referred to as dec, is the kernel that allows the decoupling of the EE and BB spectra. Appendix C gives more details on this point. ${}_{\mathrm{dec}}G$ is associated with integral relations in real space, thus its harmonic expression is not straightforward. This expression requires more numerical resources to be computed because no closed relations exist for the Wigner d-matrix with different multipole indices. It can be obtained with PolSpice for all $\ell$ by setting $\tilde{C}^{\mathrm{EE}}_{\ell'} = \tilde{C}^{\mathrm{BB}}_{\ell'} = \delta_{\ell\ell'}$ as input.

### *6.4. Covariance of the PolSpice estimator*

Given the previous relations in Eqs. (47)–(50), the covariance of the PolSpice estimator can be written as a convolution of the covariance of the pseudo-power spectrum, with

$$\begin{aligned}\hat{\Sigma}^{\mathrm{TTTT}}_{\ell\ell'} &\equiv \mathrm{cov}(\hat{C}^{\mathrm{TT}}_{\ell}, \hat{C}^{\mathrm{TT}}_{\ell'}) \\ &= \sum_{LL'} {}_{0}G_{lL}\, \tilde{\Sigma}^{\mathrm{TTTT}}_{LL'}\, {}_{0}G_{l'L'}.\end{aligned} \tag{55}$$

For polarization, there is mixing between the EE and BB components in the covariance. We write the polarized EEEE PolSpice covariance after defining ${}_{\pm}G \equiv \frac{1}{2}\left({}_{\mathrm{dec}}G \pm {}_{-2}G\right)$ as

$$\begin{aligned}\hat{\Sigma}^{\mathrm{EEEE}} =& {}_{+}G\tilde{\Sigma}^{\mathrm{EEEE}}{}_{+}G^{\top} + {}_{-}G\tilde{\Sigma}^{\mathrm{BBEE}}{}_{+}G^{\top} \\ &+ {}_{+}G\tilde{\Sigma}^{\mathrm{EEBB}}{}_{-}G^{\top} + {}_{-}G\tilde{\Sigma}^{\mathrm{BBBB}}{}_{-}G^{\top}.\end{aligned} \tag{56}$$

The polarized PolSpice covariance is built on the polarized pseudo-covariance, mixing the components EE and BB, thanks to the kernel ${}_{\pm}G$; see Fig. 13. This figure displays a row of the kernels that were computed on the mask SPT-3G used in our analysis. It shows the window functions that are applied to the pseudo-power spectra to produce the PolSpice spectra.They correct for the bias due to the mask, but introduce a small bias due to the lack of information at large scales. The temperature kernel ${}_{0}G$ and the polarization ${}_{+}G$ kernel are almost identical. The leakage kernel ${}_{-}G$ (all negative), which accounts for the mixing of the $E$ and $B$ polarization pseudo-spectra in the PolSpice spectrum, is orders of magnitudes smaller than the other two. Hence, the BB covariance terms do not affect the EEEE covariance. On the other hand, the EE terms affect the BBBB covariance because the EE spectrum is a few orders of magnitude larger than BB. The PolSpice apodizing function of the correlation function we used is

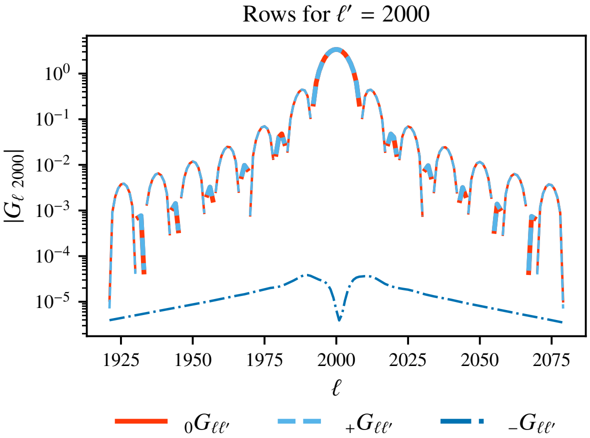


**Fig. 13.** Amplitude of the PolSpice convolution kernels $G$ for the SPT-3G footprint. The negative terms are plotted with thinner lines.

$$f^{\mathrm{apo}}(\theta) = \begin{cases} \frac{1}{2}\left(1 + \cos\frac{\pi\theta}{\theta_{\max}}\right) & \forall\theta < \theta_{\max}, \\ 0 & \text{otherwise.}\end{cases} \tag{57}$$

Here we have, $\theta_{\max} = \pi/6$. Without apodization, but with partial sky, as for *Planck*, the decoupling kernel ${}_{\mathrm{dec}}G$ is not null, still resulting in a nonzero ${}_{-}G$ kernel. However, it is orders of magnitude smaller than ${}_{+}G$, hence we can compute EEEE ignoring the leakage from covariance terms that include BB.

### *6.5. Accuracy of the covariance approximations for the PolSpice estimator*

We can build estimates of the PolSpice spectrum covariance by convolving the pseudo-spectrum covariance with the appropriate kernels following Eqs. (55) and (56). To calculate the pseudo-spectrum covariance, we can use the NKA, INKA, FRI, and ACC approximations or the exact computation. Figure 14 shows the accuracy of the binned PolSpice covariance calculated with the approximations compared to the exact calculation. The results are similar to those we found for the accuracy of the pseudo-spectrum covariances shown in Fig. 6. The NKA, INKA, and FRI approaches provide a good estimate of the PolSpice covariance. However, the ACC approach improves the existing approximations dramatically. This shows that our results for the accuracy of the pseudo-covariance also hold for the PolSpice.

## 7. Summary and conclusions

One of the key ingredients of cosmological analysis based on power spectra are covariance matrices. Accurate covariance matrices ensure precise error bars and an unbiased estimation of cosmological parameters. The analytical estimation of these matrices can be difficult when small sky fractions are observed because existing approximations might fail. We have considered the specific example of estimating accurate analytical signal-signal covariance matrices for the SPT-3G CMB experiment,

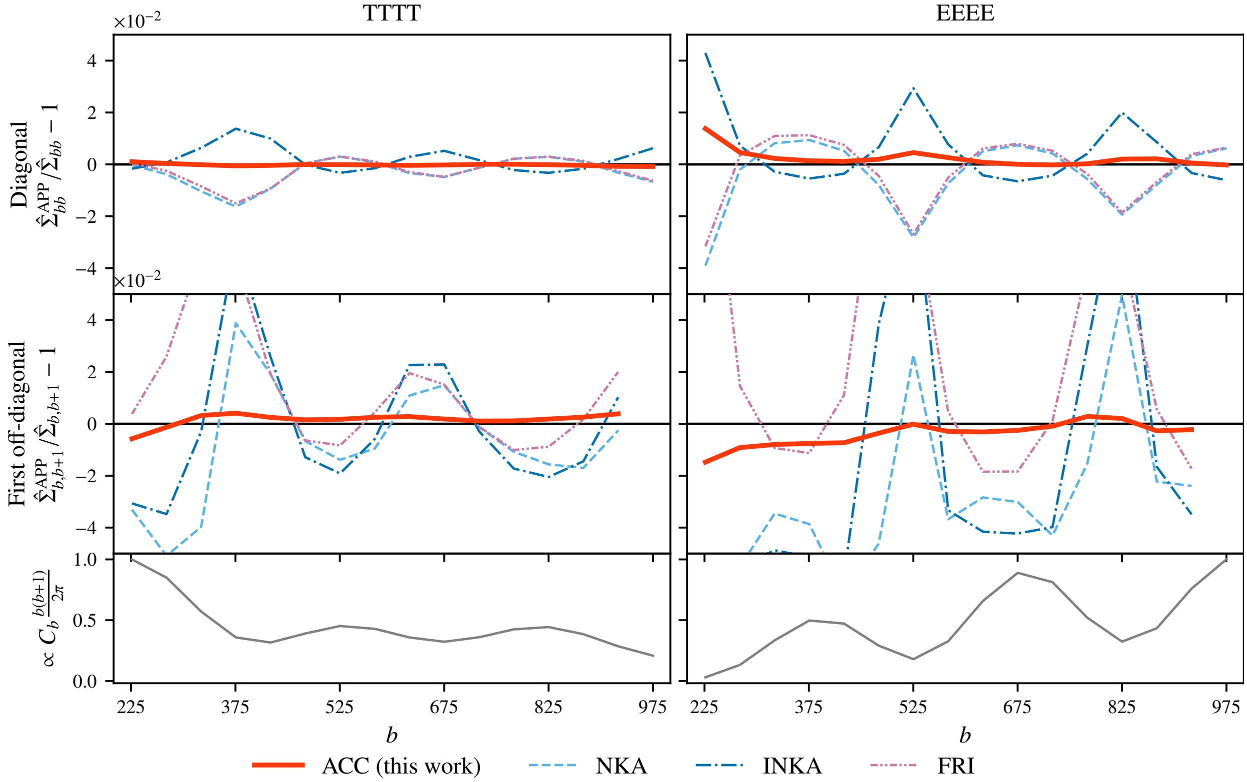


**Fig. 14.** Relative differences of binned `PolSpice` covariance matrices calculated using approximations of the pseudo-spectrum covariance with respect to the exact computation: $\hat{\Sigma}^{\rm APP}_{bb'}/\hat{\Sigma}_{bb'} - 1$, for TTTT (*left-hand side*) and EEEE (*right-hand side*), with binning $\Delta\ell = 50$ . On the first row we plot the relative differences for the diagonal, i.e. $b = b'$, while on the second row we plot the differences for the first off-diagonal, i.e. $b' = b + 1$. Similarly to the case of pseudo-covariances in Fig. 6, we find acceptable accuracy for the NKA (dash light blue), INKA (dash-dot dark blue) and FRI (dash-double-dot purple) approximations, while our ACC approximation (solid red) improves overall the others. The `PolSpice` covariance matrices have been calculated using Eqs. (55) and (56). The last row displays the corresponding binned underlying renormalized spectrum TT or EE, to highlight the fact that the differences in the covariances are on the peaks and in the troughs of the spectra, i.e. where the curvature of the spectrum is maximal.

whose survey covers about 4%, without masking the contribution of point sources. We considered the cases of estimating the matrix for pseudo-power spectrum and for the `PolSpice` power spectrum estimator.

First, we implemented in Sect. 3 an expensive exact calculation of the covariance of the pseudo-power spectrum in intensity and polarization for the first time. We used a map-based algorithm that is accelerated thanks to the `HEALPix` pixelation tools. We were thus able to compute the entire covariance matrix up to $\ell_{\rm max,ex} = 1000$ exactly. We also obtained a selection of rows of the covariance of particular interest up to $\ell = 2000$.

Based on this result, we were able to estimate the accuracy of the existing approximations in Sect. 4 precisely by comparing them to the binned exact covariances of the pseudo-power spectra measured on the SPT-3G patch. The approximations were found to be precise to the 5% level.

Then, using the code we developed for an exact computation of the covariance matrix, we estimated the covariance-coupling kernel $\bar{\Theta}$, which determines how the CMB power spectrum couples into the covariance matrix. We were able to understand why the existing approximations in the literature fail to achieve a precision better than 5%. We then proposed a new approximation in Sect. 5, the ACC, which is more computationally expensive than the existing approximations, but allows a more precise estimation of the covariance matrix at the 1% level.

Finally, we showed in Sect. 6 that we are able to build the covariance of the `PolSpice` power spectrum in both temperature and polarization using a harmonic correction. This computation is exact and based on the `PolSpice` algorithm real-space corrections that we translated into harmonic space. Through this correction, we produced estimates of the `PolSpice` covariance matrix based on the previous approximations of the pseudo-power spectrum covariance. The accuracy of the resulting `PolSpice` covariance approximations is the same as for the pseudo-power spectrum.

While this paper considered the particular example of the SPT-3G experiment, the results can be extended to a non-CMB power spectrum analysis such as that of weak-lensing shear or photometric catalogs. We would also like to stress that the accuracy of any of the approximations presented in this paper (existing or new) is reduced when a point-source mask is included in the sky footprint. Nevertheless, the exact computation of the covariance matrices still holds in this particular case. While previous experiments have included the effect of point-source masks through the use of simulations (see, e.g., Planck Collaboration XI 2016) or by inpainting the holes with constrained realizations (see, e.g., Benoit-Lévy et al. 2013), additional work is required to find an analytical calculation of this contribution.

*Acknowledgements.* We are grateful to the SPT collaboration for discussions and suggestions. We thank Lennart Balkenhol for discussions and insightful comments on the manuscript. This work has received funding from the French Centre National d'Etudes Spatiales (CNES). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 101001897).

## References

Ade, P., Aguirre, J., Ahmed, Z., et al. 2019, J. Cosmol. Astropart. Phys., 2019, 056
Aiola, S., Calabrese, E., Maurin, L., et al. 2020, J. Cosmol. Astropart. Phys., 2020, 047
Alonso, D., Sanchez, J., & Slosar, A. 2019, MNRAS, 484, 4127
Balkenhol, L., & Reichardt, C. L. 2022, MNRAS, 512, 4394
Benoit-Lévy, A., Déchelette, T., Benabed, K., et al. 2013, A&A, 555, A37
Challinor, A., & Chon, G. 2005, MNRAS, 360, 509
Chon, G., Challinor, A., Prunet, S., Hivon, E., & Szapudi, I. 2004, MNRAS, 350, 914
Dodelson, S., & Schneider, M. D. 2013, Phys. Rev. D, 88, 063537
Dutcher, D., Balkenhol, L., Ade, P. A. R., et al. 2021, Phys. Rev. D, 104, 022003
Efstathiou, G. 2004, MNRAS, 349, 603
Friedrich, O., Andrade-Oliveira, F., Camacho, H., et al. 2021, MNRAS, 508, 3125
Gallardo, P. A., Benson, B., Carlstrom, J., et al. 2022, Proc. SPIE, 12190, 121900C
García-García, C., Alonso, D., & Bellini, E. 2019, J. Cosmol. Astropart. Phys., 2019, 043
Gorski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759
Hazumi, M., Borrill, J., Chinone, Y., et al. 2012, Proc. SPIE, 8442, 844219
Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19
Hivon, E., Gorski, K. M., Netterfield, C. B., et al. 2002, ApJ, 567, 2
Khersonskii, V., Moskalev, A., & Varshalovich, D. 1988, Quantum Theory Of Angular Momemtum (Singapore: World Scientific Publishing Company)
Louis, T., Naess, S., Garrido, X., & Challinor, A. 2020, Phys. Rev. D, 102, 123538
Lueker, M., Reichardt, C. L., Schaffer, K. K., et al. 2010, ApJ, 719, 1045
Nicola, A., Garcia-Garcia, C., Alonso, D., et al. 2021, J. Cosmol. Astropart. Phys., 2021, 067
Planck Collaboration XI. 2016, A&A, 594, A11
Planck Collaboration V. 2020, A&A, 641, A5
Planck Collaboration VI. 2020, A&A, 641, A6
Sellentin, E., & Starck, J.-L. 2019, J. Cosmol. Astropart. Phys., 2019, 021
Szapudi, I., Prunet, S., Pogosyan, D., Szalay, A. S., & Bond, J. R. 2001, ApJ, 548, L115

## Appendix A: Analysis of the curved sky

This appendix describes the mathematical tools that are used on curved sky for a CMB analysis. We make use of the spherical harmonic decomposition of Gaussian fields. We introduce various operators that allow us to express the couplings and the covariance of the power spectra. We make use of some geometrical relations of spherical harmonics to obtain our results. In this appendix, we introduce formulae that can either be used in the temperature or in the polarization case.

### *A.1. Temperature*

We first consider the case of a map of the CMB intensity anisotropies $T(\hat{n})$ observed in direction $\hat{n}$. The anisotropies are distributed as a Gaussian random field with a corresponding power spectrum $C_\ell^{\mathrm{TT}}$, observed through a mask $W(\hat{n})$.

#### Harmonic coefficients and underlying power spectrum

The intensity map can be decomposed with spin-0 spherical harmonics to obtain the harmonic coefficients and their variance, the intensity power spectrum, which fully characterizes the physical properties of the field,

$$a_{\ell m}^{\mathrm{T}} = \int \mathrm{d}\hat{u} T(\hat{u})\, {}_0Y_{\ell m}^*(\hat{u}), \tag{A.1}$$

$$\langle a_{\ell m}^{\mathrm{T}} a_{\ell' m'}^{\mathrm{T}*} \rangle = C_\ell^{\mathrm{TT}} \delta_{\ell\ell'} \delta_{mm'}. \tag{A.2}$$

Here the brackets $\langle\rangle$ indicate an average over many realizations of the maps.

#### Weighted mask $W(\hat{n})$

The weighted mask is a real map with weights from 0 to 1 that is used to taper the data on the border of the survey area in order to reduce Fourier ringing when harmonic decomposition is used. We define the mask harmonic coefficients and its power spectrum as

$$w_{\ell m} \equiv \int \mathrm{d}\hat{u} W(\hat{u})\, {}_0Y_{\ell m}^*(\hat{u}), \tag{A.3}$$

$$\mathcal{W}_\ell \equiv \frac{1}{2\ell+1} \sum_m w_{\ell m} w_{\ell m}^*. \tag{A.4}$$

#### Pseudo-power spectrum estimator

One way to obtain a biased estimator of the power spectrum in CMB experiments is to define the pseudo-harmonic coefficients and the pseudo-power spectrum,

$$\tilde{a}_{\ell m}^{\mathrm{T}} \equiv \int \mathrm{d}\hat{u} W(\hat{u}) T(\hat{u})\, {}_0Y_{\ell m}^*(\hat{u}), \tag{A.5}$$

$$\tilde{C}_\ell^{\mathrm{TT}} \equiv \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} |\tilde{a}_{\ell m}^{\mathrm{T}}|^2. \tag{A.6}$$

#### Relation between harmonic coefficients

We relate the masked pseudo-harmonic coefficients to the unmasked coefficient with

$$\tilde{a}_{\ell m}^{\mathrm{T}} = \sum_{\ell' m'} a_{\ell' m'}^{\mathrm{T}}\, {}_0I_{\ell m \ell' m'}[W]. \tag{A.7}$$

The $I[W]$ couplings are defined below and can be expressed in terms of sums over Wigner-3j symbols and the $w_{\ell m}$, with

$${}_sI_{\ell m \ell' m'}[W] \equiv \int \mathrm{d}\hat{u}\, {}_sY_{\ell m}(\hat{u}) W(\hat{u})\, {}_sY_{\ell' m'}^*(\hat{u}), \tag{A.8}$$

$$= \sum_{LM} w_{LM} (-1)^{m'} \left[ \frac{(2\ell+1)(2\ell'+1)(2L+1)}{4\pi} \right]^{1/2} \begin{pmatrix} \ell & \ell' & L \\ -s & s & 0 \end{pmatrix} \times \begin{pmatrix} \ell & \ell' & L \\ m & -m' & M \end{pmatrix}. \tag{A.9}$$

Here, we anticipated the extension of this notation to the polarized case, which deals with spin-2 fields. The mask-dependent ${}_sI_{\ell m \ell' m'}[W]$ coupling coefficients relate the underlying harmonic coefficients to the measured pseudo-harmonic coefficients. In the full-sky case, we obtain ${}_sI_{\ell m \ell' m'}[1] = \delta_{\ell\ell'}\delta_{mm'}$ through the orthonormality properties of the spin-weighted spherical harmonics.

We introduce some useful relations that are demonstrated in Hivon et al. (2002),

$$\sum_{\ell m} {}_sI_{\ell_1 m_1 \ell m}[W]\, {}_sI_{\ell_2 m_2 \ell m}^*[W] = {}_sI_{\ell_1 m_1 \ell_2 m_2}[W^2], \tag{A.10}$$

$$\sum_{m_1 m_2} \frac{{}_sI_{\ell_1 m_1 \ell_2 m_2}[W]\, {}_{s'}I_{\ell_1 m_1 \ell_2 m_2}^*[W]}{(2\ell_1+1)(2\ell_2+1)} = \Xi_{\ell_1 \ell_2}^{ss'}[W]. \tag{A.11}$$

Here we introduced the symmetric operator $\Xi^{ss'}$ acting on a power spectrum $\mathcal{A}_\ell$,

$$\Xi_{\ell\ell'}^{ss'}[\mathcal{A}] \equiv \sum_L \frac{2L+1}{4\pi} \mathcal{A}_L \begin{pmatrix} \ell & \ell' & L \\ s & -s & 0 \end{pmatrix} \begin{pmatrix} \ell & \ell' & L \\ s' & -s' & 0 \end{pmatrix}. \tag{A.12}$$

We extend this definition to an operator acting on a map $A(\hat{n})$, with

$$\Xi_{\ell\ell'}^{ss'}[A] \equiv \Xi_{\ell\ell'}^{ss'}[\mathcal{A}], \tag{A.13}$$

where we defined the power spectrum $\mathcal{A}_\ell$ of the map $A$ as in Eqs. (A.3) and (A.4).

#### `MASTER` relation between estimated and true spectra

By inserting Eq. (A.5) into Eq. (A.6), using the relations of Eq. (A.11) and the definition of Eq. (A.12)-(A.13), we relate the ensemble average of the pseudo-power spectrum to the underlying power spectrum using the `MASTER` mode-coupling kernel with

$$\langle \tilde{C}_\ell^{\mathrm{TT}} \rangle = \sum_{\ell'} {}_0M_{\ell\ell'}[W] C_{\ell'}^{\mathrm{TT}}. \tag{A.14}$$

The `MASTER` mode-coupling matrix is given by

$${}_0M_{\ell\ell'}[W] \equiv (2\ell'+1) \Xi_{\ell\ell'}^{00}[W]. \tag{A.15}$$

### *A.2. Polarization*

We consider the case of the CMB intensity and polarization anisotropies, represented by maps $T(\hat{n}), Q(\hat{n}), U(\hat{n})$ in direction $\hat{n}$ of the sky. These are Gaussian random fields, fully characterized by their power spectra $(C_\ell^{\mathrm{TT}}, C_\ell^{\mathrm{EE}}, C_\ell^{\mathrm{BB}}, C_\ell^{\mathrm{TE}})$ observed through a mask $W(\hat{n})$. The definitions and relations of the previous section can be extended to polarization spectra. First we compute the pseudo-harmonic coefficients on the masked sky with spin-weighted spherical harmonics, given in the following inverse relation from Chon et al. (2004):

$$(Q \pm iU)(\hat{n}) = \sum_{\ell m} (\tilde{a}^{\mathrm{E}}_{\ell m} \mp i\tilde{a}^{\mathrm{B}}_{\ell m})_{\mp 2}Y_{\ell m}. \tag{A.16}$$

The pseudo-power spectrum $\tilde{C}^{\mathrm{XY}}_{\ell}$ is obtained by summing over the measured pseudo-harmonic coefficients $\tilde{a}^{\mathrm{X}}_{\ell m}$, X, Y $\in$ [T, E, B] with the same multipole $\ell$, with

$$\tilde{C}^{\mathrm{XY}}_{\ell} = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} \tilde{a}^{\mathrm{X}}_{\ell m} \tilde{a}^{\mathrm{Y}*}_{\ell m}. \tag{A.17}$$

More details can be found in Challinor & Chon (2005). Following the same approach as in the previous section, we can write the `MASTER` relation in polarization,

$$\langle \tilde{C}^{\mathrm{EE}}_{\ell} + \tilde{C}^{\mathrm{BB}}_{\ell} \rangle = \sum_{\ell'} {}_{+2}M_{\ell\ell'} \left( C^{\mathrm{EE}}_{\ell'} + C^{\mathrm{BB}}_{\ell'} \right), \tag{A.18}$$

$${}_{+2}M_{\ell\ell'} = (2\ell' + 1)\Xi^{22}_{\ell\ell'}[W], \tag{A.19}$$

$$\langle \tilde{C}^{\mathrm{EE}}_{\ell} - \tilde{C}^{\mathrm{BB}}_{\ell} \rangle = \sum_{\ell'} {}_{-2}M_{\ell\ell'} \left( C^{\mathrm{EE}}_{\ell'} - C^{\mathrm{BB}}_{\ell'} \right), \tag{A.20}$$

$${}_{-2}M_{\ell\ell'} = (2\ell' + 1)\Xi^{2-2}_{\ell\ell'}[W], \tag{A.21}$$

$$\langle \tilde{C}^{\mathrm{TE}}_{\ell} \rangle = \sum_{\ell'} {}_{\times}M_{\ell\ell'} C^{\mathrm{TE}}_{\ell'}, \tag{A.22}$$

$${}_{\times}M_{\ell\ell'} = (2\ell' + 1)\Xi^{20}_{\ell\ell'}[W]. \tag{A.23}$$

### A.3. Renormalized kernels

We define the renormalized `MASTER` kernels, which are used in the INKA. They are written as

$${}_{\mathrm{k}}\bar{M}_{\ell\ell'} \equiv \frac{1}{\sum_{\ell'} {}_{\mathrm{k}}M_{\ell\ell'}} {}_{\mathrm{k}}M_{\ell\ell'}, \quad \forall \mathrm{k} \in [0, -2, +2, \times]. \tag{A.24}$$

Summing over $\ell'$ yields

$$\sum_{\ell'} {}_{\mathrm{k}}\bar{M}_{\ell\ell'} = 1, \tag{A.25}$$

which ensures that the approximated covariance-coupling kernel defined in Eq. (30) is properly normalized.

## Appendix B: Covariance of the pseudo-power spectrum

In this appendix, we outline how the formula of the covariance matrix of the pseudo-power spectrum is obtained in the temperature case. Our goal is to introduce Eq. (B.9). The covariance matrix of the pseudo-power spectrum reads

$$\begin{aligned} \tilde{\Sigma}_{\ell\ell'} &= \langle \tilde{C}_{\ell}\tilde{C}_{\ell'} \rangle - \langle \tilde{C}_{\ell} \rangle \langle \tilde{C}_{\ell'} \rangle, \\ &= \sum_{mm'} \frac{\langle |\tilde{a}_{\ell m}|^2 |\tilde{a}_{\ell' m'}|^2 \rangle - \langle |\tilde{a}_{\ell m}|^2 \rangle \langle |\tilde{a}_{\ell' m'}|^2 \rangle}{(2\ell + 1)(2\ell' + 1)}. \end{aligned} \tag{B.1}$$

As the intensity map $T(\hat{n})$ is real and the spherical harmonics follow

$${}_{0}Y^{*}_{\ell m} = (-1)^{m} {}_{0}Y_{\ell(-m)}, \tag{B.2}$$

the spherical harmonic coefficients of $T(\hat{n})$ follow

$$a^{*}_{\ell m} = (-1)^{m} a_{\ell(-m)}. \tag{B.3}$$

When the four-point function is computed and thanks to Wick's theorem,

$$\begin{aligned} &\sum_{mm'} \langle |\tilde{a}_{\ell m}|^2 |\tilde{a}_{\ell' m'}|^2 \rangle - \langle |\tilde{a}_{\ell m}|^2 \rangle \langle |\tilde{a}_{\ell' m'}|^2 \rangle \\ &\quad = \sum_{mm'} \langle \tilde{a}_{\ell m} \tilde{a}_{\ell' m'} \rangle \langle \tilde{a}^{*}_{\ell m} \tilde{a}^{*}_{\ell' m'} \rangle + \langle \tilde{a}_{\ell m} \tilde{a}^{*}_{\ell' m'} \rangle \langle \tilde{a}^{*}_{\ell m} \tilde{a}_{\ell' m'} \rangle. \end{aligned} \tag{B.4}$$

Based on the first term of this sum and using the change of variable $m'' = -m'$, it is straightforward to show that

$$\begin{aligned} &\sum_{mm'} \langle \tilde{a}_{\ell m} \tilde{a}_{\ell' m'} \rangle \langle \tilde{a}^{*}_{\ell m} \tilde{a}^{*}_{\ell' m'} \rangle \\ &\qquad = \sum_{mm'} (-1)^{(2m')} \langle \tilde{a}_{\ell m} \tilde{a}^{*}_{\ell'(-m')} \rangle \langle \tilde{a}^{*}_{\ell m} \tilde{a}_{\ell'(-m')} \rangle, \\ &\qquad = \sum_{mm''} \langle \tilde{a}_{\ell m} \tilde{a}^{*}_{\ell' m''} \rangle \langle \tilde{a}^{*}_{\ell m} \tilde{a}_{\ell' m''} \rangle. \end{aligned} \tag{B.5}$$

Finally, we have

$$\begin{aligned} &\sum_{mm'} \langle |\tilde{a}_{\ell m}|^2 |\tilde{a}_{\ell' m'}|^2 \rangle - \langle |\tilde{a}_{\ell m}|^2 \rangle \langle |\tilde{a}_{\ell' m'}|^2 \rangle, \\ &\qquad = 2 \sum_{mm'} \langle \tilde{a}_{\ell m} \tilde{a}^{*}_{\ell' m'} \rangle \langle \tilde{a}^{*}_{\ell m} \tilde{a}_{\ell' m'} \rangle \\ &\qquad = 2 \sum_{mm'} |\langle \tilde{a}_{\ell m} \tilde{a}^{*}_{\ell' m'} \rangle|^2. \end{aligned} \tag{B.6}$$

Then, we have for the covariance matrix,

$$\tilde{\Sigma}_{\ell\ell'} = \frac{2}{(2\ell + 1)(2\ell' + 1)} \sum_{mm'} |\langle \tilde{a}_{\ell m} \tilde{a}^{*}_{\ell' m'} \rangle|^2. \tag{B.7}$$

When we use the decomposition of pseudo-harmonic coefficients, we obtain

$$\begin{aligned} \langle \tilde{a}_{\ell m} \tilde{a}^{*}_{\ell' m'} \rangle &= \sum_{\ell_1 m_1 \ell_2 m_2} \langle a_{\ell_1 m_1} a^{*}_{\ell_2 m_2} \rangle I_{\ell m \ell_1 m_1}[W] I^{*}_{\ell' m' \ell_2 m_2}[W], \\ &= \sum_{\ell_1 m_1} C_{\ell_1} I_{\ell m \ell_1 m_1}[W] I^{*}_{\ell' m' \ell_1 m_1}[W]. \end{aligned} \tag{B.8}$$

Inserting Eq. (B.8) into Eq. (B.7) gives

$$\begin{aligned} \tilde{\Sigma}_{\ell\ell'} = &\frac{2}{(2\ell + 1)(2\ell' + 1)} \sum_{mm'} \sum_{\ell_1 m_1} \sum_{\ell_2 m_2} C_{\ell_1} C_{\ell_2} \\ &I_{\ell m \ell_1 m_1}[W] I^{*}_{\ell' m' \ell_1 m_1}[W] I^{*}_{\ell m \ell_2 m_2}[W] I_{\ell' m' \ell_2 m_2}[W]. \end{aligned} \tag{B.9}$$

## Appendix C: Expansion in Legendre series

In this section, we introduce the Legendre transforms of the harmonic quantities used in this work. Each relation in harmonic space has a corresponding expression in real space. The `PolSpice` software relies on the relations in real space.

### *C.1. From spin-0 spectra to two-point correlation functions*

Given a spectrum $\mathcal{A}_\ell$, we can associate a real two-point correlation function $a$ with it,

$$a(\theta) = \sum_\ell \frac{2\ell+1}{4\pi} \mathcal{A}_\ell P_\ell(\cos\theta) \;\; \forall\theta \in [0,\pi]. \tag{C.1}$$

The inverse relation is

$$\mathcal{A}_\ell = 2\pi \int_0^\pi \mathrm{d}\theta \sin\theta a(\theta) P_\ell(\cos\theta) \;\; \forall \ell \geq 0. \tag{C.2}$$

The two-point function gives the correlations between two directions of the sky, for instance, in the CMB anisotropy full-sky case, we can write

$$\langle T(\hat{n}_1) T(\hat{n}_2)\rangle = \sum_\ell \frac{2\ell+1}{4\pi} C_\ell^{\mathrm{TT}} P_\ell(\hat{n}_1 \cdot \hat{n}_2). \tag{C.3}$$

### *C.2. From convolution to multiplication*

A convolution with a square matrix $A$ in harmonic space, such as in Eq. (A.14), is equivalent to a multiplication in real space with the correlation function $a(\theta)$ given by

$$a(\theta) \frac{2\ell'+1}{4\pi} P_{\ell'}(\cos\theta) = \sum_\ell \frac{2\ell+1}{4\pi} A_{\ell\ell'} P_\ell(\cos\theta) \;\; \forall\theta \in [0,\pi]. \tag{C.4}$$

The inverse relation is

$$A_{\ell\ell'} \equiv \frac{2\ell'+1}{2} \int_0^\pi a(\theta) P_\ell(\cos\theta) P_{\ell'}(\cos\theta) \sin(\theta) \mathrm{d}\theta. \tag{C.5}$$

The last relation is equivalent to

$$A_{\ell\ell'} = (2\ell'+1) \Xi^{00}_{\ell\ell'}[\mathcal{A}], \tag{C.6}$$

with $\mathcal{A}$ the power spectrum associated with the two-point function $a$ through a Legendre transform. We can thus extend the definition of the operator $\Xi$ to an operator acting on a correlation function $a$, with

$$\Xi^{ss'}_{\ell\ell'}[a] \equiv \Xi^{ss'}_{\ell\ell'}[\mathcal{A}]. \tag{C.7}$$

Here we have already extended the definition to be used in the spin-2 case, which we discuss in the next subsection.

### *C.3. Spin-2*

Similar rules to those introduced above can be written for spin-2 quantities by replacing the Legendre polynomials by the more generally reduced Wigner $d$-matrix $d^\ell_{2\pm2}$. Details can be found in Challinor & Chon (2005). The spin-2 relations for the $\mathcal{A}^\pm_\ell$ power spectra are associated with a spin-2 field,

$$a_\pm(\theta) = \sum_\ell \frac{2\ell+1}{4\pi} \mathcal{A}^\pm_\ell d^\ell_{2\pm2}(\cos\theta), \tag{C.8}$$

$$\mathcal{A}^\pm_\ell = 2\pi \int_0^\pi \mathrm{d}\theta \sin\theta a_\pm(\theta) d^\ell_{2\pm2}(\cos\theta). \tag{C.9}$$

We can associate a spin-0 correlation function with its spin-2 convolution matrix,

$$\begin{aligned} {}_{\pm2}A_{\ell\ell'} &= \frac{2\ell'+1}{2} \int_0^\pi a(\theta) d^\ell_{2\pm2}(\cos\theta) d^{\ell'}_{2\pm2}(\cos\theta) \sin(\theta) \mathrm{d}\theta, \\ &= (2\ell'+1) \Xi^{2\pm2}_{\ell\ell'}[a]. \end{aligned} \tag{C.10}$$

We can also compute the matrix associated with the spin-0 cross spin-2 case,

$${}_\times A_{\ell\ell'} = (2\ell'+1) \Xi^{20}_{\ell\ell'}[a]. \tag{C.11}$$

### *C.4. Applying this formalism to the `MASTER` matrix*

In our case, we can write for $s \in [0,2]$,

$${}_{\pm s}M_{\ell\ell'} = (2\ell'+1) \Xi^{s\pm s}_{\ell\ell'}[w], \tag{C.12}$$

with $w(\theta)$ the correlation function of the mask.

We apply the previous formalism and particularly the Eq. (C.12) to the `MASTER` relation. We use Legendre series expansion of the true power spectrum $C_\ell$ and the pseudo-power spectrum estimator $\tilde{C}_\ell$ to define the correlation functions $\xi$ and $\tilde{\xi}$, respectively. It gives

$$\tilde{\xi}(\theta) \equiv \sum_\ell \frac{2\ell+1}{4\pi} P_\ell(\cos\theta) \tilde{C}_\ell, \tag{C.13}$$

$$\xi(\theta) \equiv \sum_\ell \frac{2\ell+1}{4\pi} P_\ell(\cos\theta) C_\ell. \tag{C.14}$$

Starting from the right-hand side of Eq. (34) and going to real space using a Legendre transform at an angle $\theta \in [0,\pi]$, we have

$$\begin{aligned} \sum_{\ell\ell'} \frac{2\ell+1}{4\pi} P_\ell(\cos\theta) {}_0M_{\ell\ell'} C_{\ell'} & \\ &= \sum_{\ell'} w(\theta) \frac{2\ell'+1}{4\pi} P_{\ell'}(\cos\theta) C_{\ell'}, \\ &= w(\theta)\xi(\theta), \end{aligned}$$

$$\text{which implies } \langle\tilde{\xi}(\theta)\rangle = w(\theta)\xi(\theta), \tag{C.15}$$

with $w(\theta)$ the correlation function of the mask.

### *C.5. `PolSpice` in polarization*

This section describes the regularization technique that is used for polarization by `PolSpice`. One of the main advantages of `PolSpice` is that it allows the possibility of eliminating EE to BB (and BB to EE) mixing, using nonlocal relations between Wigner $d$-matrices; see Sec. 5 of Chon et al. (2004). The obtained estimator $\hat{C}^{\mathrm{EE}}_\ell$ ( $\hat{C}^{\mathrm{BB}}_\ell$) depends only on the average of $C^{\mathrm{EE}}_\ell$ ( $C^{\mathrm{BB}}_\ell$) and the scalar apodizing function $f$.

Using the Legendre transforms of spin-2 quantities (Eq. (C.9)), we associate the correlation functions $\xi_\pm$ with the spectra $C^{\mathrm{EE}}_\ell \pm C^{\mathrm{BB}}_\ell$ and $\tilde{\xi}_\pm$ to $\tilde{C}^{\mathrm{EE}}_\ell \pm \tilde{C}^{\mathrm{BB}}_\ell$. `PolSpice` builds two correlation functions $\hat{\xi}_{\mathrm{dec}}$ and $\hat{\xi}_-$ to produce an estimator of the true underlying polarized power spectrum. This spectrum is defined similarly to $\hat{\xi}$ in Eq. (39),

$$\hat{\xi}_-(\theta) = g(\theta)\tilde{\xi}_-(\theta). \tag{C.16}$$

The first is built on integral relations. `PolSpice` eliminates the mixing inherent in $\tilde{\xi}_+$ with the following relation in real space:

$$\hat{\xi}_{\mathrm{dec}}(\theta) = f^{\mathrm{apo}}(\theta) \int_{-1}^{1} \mathrm{d}\cos\theta' \frac{\tilde{\xi}_+(\theta')}{w(\theta')} \sum_\ell d^\ell_{2-2}(\theta) d^\ell_{22}(\theta'). \tag{C.17}$$

This integration can be shown to depend only on $\xi_+$ in the range $\theta \in [0, \theta_{\max}]$; see Chon et al. (2004). This allows decoupling the correlation functions from an incomplete range of angular separations that are missing due to the mask. This correlation function is noted with the subscript *dec* to emphasize that it is the crucial step allowing the decoupling of the polarized estimator. When averaging out Eq. (C.16) and Eq. (C.17) on multiple realizations,

$$\langle \hat{\xi}_{\rm dec}(\theta)\rangle = f(\theta)\int_{-1}^{1} {\rm d}\cos\theta' \xi_+(\theta')\sum_\ell d^\ell_{2-2}(\theta) d^\ell_{22}(\theta'), \tag{C.18}$$

$$\langle \hat{\xi}_-(\theta)\rangle = f(\theta)\xi_-(\theta). \tag{C.19}$$

In harmonic space, transforming Eq. (C.18) and Eq. (C.19) using $d^\ell_{2-2}$, this gives with ${}_{-2}K_{\ell\ell'} = (2\ell'+1)\Xi^{2-2}_{\ell\ell'}[f^{\rm apo}]$,

$$\hat{C}^{\rm EE}_\ell + \hat{C}^{\rm BB}_\ell \equiv 2\pi \int_0^\pi {\rm d}\cos\theta\, \hat{\xi}_{\rm dec}(\theta) d^\ell_{2-2}(\theta), \tag{C.20}$$

$$\hat{C}^{\rm EE}_\ell - \hat{C}^{\rm BB}_\ell \equiv 2\pi \int_0^\pi {\rm d}\cos\theta\, \hat{\xi}_-(\theta) d^\ell_{2-2}(\theta), \tag{C.21}$$

which implies

$$\langle \hat{C}^{\rm EE}_\ell \pm \hat{C}^{\rm BB}_\ell\rangle = \sum_{\ell'} {}_{-2}K_{\ell\ell'}(C^{\rm EE}_{\ell'} \pm C^{\rm BB}_{\ell'}). \tag{C.22}$$

Summing or subtracting the last equation for + or − allows us to build unmixed estimators of the polarization power spectra. If we had chosen not to decouple the correlation functions and had built $\hat{C}'^{\rm EE}_\ell + \hat{C}'^{\rm BB}_\ell$ as the Legendre transform of $\hat{\xi}_+ = g\tilde{\xi}_+$, the output `PolSpice` spectra would follow

$$\langle \hat{C}'^{\rm EE}_\ell \pm \hat{C}'^{\rm BB}_\ell\rangle = \sum_{\ell'} {}_{\pm 2}K_{\ell\ell'}(C^{\rm EE}_{\ell'} \pm C^{\rm BB}_{\ell'}), \tag{C.23}$$

which would leave some mixing in the polarization spectra because ${}_{+2}K \neq {}_{-2}K$.

## Appendix D: Multimask analysis

In this section, we generalize our analysis to multiple masks. This situation occurs when a cross-power spectrum analysis with maps with different masks is performed. We restrict ourselves to the study of the intensity case. Writing the expression of the covariance explicitly as in García-García et al. (2019), we obtain the following expression:

$$\begin{aligned}
&{\rm cov}(\tilde{C}^{i,j}_\ell, \tilde{C}^{p,q}_{\ell'}) \\
&= \frac{1}{(2\ell+1)(2\ell'+1)}\sum_{mm'}\left[\langle \tilde{a}^i_{\ell m}\tilde{a}^{p*}_{\ell' m'}\rangle\langle \tilde{a}^{j*}_{\ell m}\tilde{a}^q_{\ell' m'}\rangle + (p\leftrightarrow q)\right], \\
&= \frac{1}{(2\ell+1)(2\ell'+1)}\sum_{\ell_1\ell_2}\sum_{m_1m_2mm'}\Big[C^{i,p}_{\ell_1} I_{kk_1}[W^i] I_{k_1k'}[W^p] \\
&\qquad C^{j,q}_{\ell_2} I_{k_2k}[W^j] I_{k'k_2}[W^q] + (p\leftrightarrow q)\Big], \\
&= \Xi^{00}_{\ell\ell'}[W^i, W^p, W^j, W^q]\sum_{\ell_1\ell_2}\Big[C^{i,p}_{\ell_1}\bar{\Theta}^{\ell_1\ell_2}_{\ell\ell'}[W^i, W^p, W^j, W^q] C^{j,q}_{\ell_2} + \\
&\qquad (p\leftrightarrow q)\Big].
\end{aligned}$$

We noted $k_i = (\ell_i, m_i)$. We also extended the definition of the kernels to the case multiple masks as

$$\Xi^{00}_{\ell\ell'}[W^i, W^p)(W^j, W^q] \equiv \Xi^{00}_{\ell\ell'}[\mathcal{V}^{(ip)\times(jq)}], \tag{D.1}$$

$$\text{where } \mathcal{V}^{(ip)\times(jq)}_\ell \equiv \frac{1}{2\ell+1}\sum_m \left[W^iW^p\right]_{\ell m}\left[W^jW^q\right]^*_{\ell m}, \tag{D.2}$$

$$\Theta^{\ell_1\ell_2}_{\ell\ell'}[W^i, W^p, W^j, W^q] = \sum_{mm'm_1m_2} I_{\ell m\ell_1m_1}[W^i] \tag{D.3}$$

$$I_{\ell_1m_1\ell'm'}[W^p]\, I_{\ell'm'\ell_2m_2}[W^j]\, I_{\ell_2m_2\ell m}[W^q].$$

As long as the considered masks have similar properties, for instance, that none of them includes point sources, the computations developed in this work will hold. In Eq. (E.2), the assumptions hold as long as the mask harmonic coefficients fall quickly enough, which is the case even when the survey area varies a little from one map to the next.

## Appendix E: Details of the ACC approximation

### *E.1. Mathematical validation*

We explore the mathematical justification of the ACC approximation. From Fig. 9, the $\bar{\Theta}^{\ell_1\ell_2}_{\ell\ell'}$ kernel appears only to depend on $\Delta \equiv \ell' - \ell$, $\Delta_1 \equiv \ell_1 - \ell$, and $\Delta_2 \equiv \ell_2 - \ell$. We recall that the normalization of the reduced coupling kernel (Eq. (22)) already approximately only depends on $\Delta$ because $\Xi^{00}_{\ell\ell'}$ is close to a Toeplitz matrix (Louis et al. 2020). The kernel itself is given by a summation of products of four coupling coefficients ${}_0I_{\ell m\ell_1m_1}$. They are expressed as the sum of the mask window function with a product of two Wigner-3j symbols, as shown in Eq. (A.9). We remark that because the mask power spectrum falls relatively fast (Fig. 1), the terms with low $L$ in the sum in Eq. (A.9) contribute mostly. However, we are interested in the cases in which all the other multipoles $\ell, \ell', \ell_1$, and $\ell_2$ are significantly larger than the width of the mask spectrum. For this reason, all the Wigner-3j symbols in Eq. (A.9) can be replaced by their asymptotic behavior, where in the limit $\ell_i, \ell_j \gg L_i$, we have

$$\begin{pmatrix}\ell_i & \ell_j & L_i\\ m_i & m_j & M_i\end{pmatrix} \approx \frac{(-1)^{\ell_j+m_j}}{\sqrt{2\ell_j+1}} d^{L_i}_{M_i,(\ell_j-\ell_i)}(\theta) \tag{E.1}$$

(Khersonskii et al. 1988). Here, $\theta = \arccos(-m_j/(\ell_j(\ell_j+1))^{1/2})$ and $d^j_{k,m}$ are reduced Wigner rotation matrices.

When this approximation is introduced in Eq. (A.9), Eq. (21) reads

$$\begin{aligned}
\Theta^{\ell_1\ell_2}_{\ell_3\ell_4} \approx{}& \frac{1}{(4\pi)^2}\sum_{L_iM_i}\Pi_i w_{L_iM_i}\sqrt{2L_i+1}\, d^{L_i}_{0,(\ell_{i+1}-\ell_i)}(\pi/2) \\
&\left[\sum_{m_i} d^{L_i}_{M_i,(\ell_{i+1}-\ell_i)}\left(\arccos\frac{-m_{i+1}}{(\ell_{i+1}(\ell_{i+1}+1))^{1/2}}\right)\right].
\end{aligned} \tag{E.2}$$

Here we have defined $\ell_5 \equiv \ell_1$ for notational purposes. We note that when $\ell_i$ is large enough, which is the case because $\ell_i, \ell_j \gg L_i$, $\arccos\frac{-m_{i+1}}{(\ell_{i+1}(\ell_{i+1}+1))^{1/2}}$ explore the $[0,\pi]$ range, and the expression in brackets in the last equation can be seen as a Riemann sum over $\theta \in [0,\pi]$. This expression can be approximated by the integral $\int d^{L_i}_{M_i,(\ell_{i+1}-\ell_i)}(\theta){\rm d}\theta$, which only depends on $L_i, M_i$, and $\ell_{i+1} - \ell_i$. Because $L_i, M_i$ are summed over in Eq. (E.2), we directly see that as soon as $\ell, \ell', \ell_1$, and $\ell_2$ are significantly larger than the width of the mask spectrum, the coupling kernel behaves as a function of their difference, and we expect that this approximation improves in accuracy with $\ell$.

### *E.2.* EE *expression*

For now, when we ignore BB to EE leakage, we can write

$$\tilde{\Sigma}^{\rm EEEE}_{\ell\ell'} \approx 2\Xi^{22}_{\ell\ell'}[W^2]\left[C^{\rm EE}\cdot\bar{\Theta}^{++++}_{\ell\ell'}[W]\cdot C^{\rm EE}\right], \tag{E.3}$$

where we defined the polarized covariance coupling

$$\Theta^{\ell_1\ell_2;++++}_{\ell\ell'} = \sum_{mm_1m'm_2} {}_+I_{\ell m\ell_1m_1}\, {}_+I_{\ell_1m_1\ell'm'}\, {}_+I_{\ell'm'\ell_2m_2}\, {}_+I_{\ell_2m_2\ell m}. \tag{E.4}$$

Based on the relations in Challinor & Chon (2005), this kernel $\Theta^{++++}$ can be approximately normalized to

$$\sum_{\ell_1\ell_2} \Theta_{\ell\ell'}^{\ell_1\ell_2;++++}[W] \approx \Xi_{\ell\ell'}^{22}[W^2], \tag{E.5}$$

hence the relation (E.3).

### 4.2.2 Covariance of the SPT-3G band powers

The high-precision analytical framework developed in Camphuis et al. (2022) allows us to compute the covariance matrix on the SPT-3G mask. In this section, we describe the adaptation of the framework to the SPT-3G band powers analysis.

**Applying ACC** Let me first define the following notation for the analytical estimation of the covariance. To obtain the covariance $\mathrm{cov}\left(\hat{C}^{i,j}, \hat{C}^{k,l}\right)$, we need to provide theoretical estimations of the power spectrum $C^{i,k}, C^{i,l}, C^{j,k}$ and $C^{j,l}$. I combine those using the ACC approximation to obtain

$$\begin{aligned}\hat{\Sigma}(\hat{C}^{i,j}, \hat{C}^{k,l}) \equiv G^{ij} \big(&\Xi^{ik,jl}[W^2] \times \left[C^{i,k} * \bar{\Theta}^{\mathrm{ACC}}_{ik,jl}[W] * C^{j,l}\right] \\ &+ \Xi^{il,jk}[W^2] \times \left[C^{i,l} * \bar{\Theta}^{\mathrm{ACC}}_{il,jk}[W] * C^{j,k}\right]\big)\, {G^{kl}}^{\top}.\end{aligned} \tag{4.20}$$

In this equation only, I used × to denote the scalar term-to-term multiplication, and $*$ to denote the matrix product. $\bar{\Theta}$ is a four-dimension matrix, thus the two scalar contractions output a 2 dimension matrix, that is multiplied term-to-term to $\Xi$. Let me emphasize that the $\Xi$ and $\bar{\Theta}$ contributions must be computed accordingly with the spin of the associated power spectrum, hence their dependence on $i, j, k, l$. I have defined the operator $\hat{\Sigma}(\cdot,\cdot)$, which uses the ACC approximation to compute the covariance of a pair of power spectra. The approximation used can be summarized as

$$\mathrm{cov}\left(\hat{C}^{i,j}, \hat{C}^{k,l}\right) \approx \hat{\Sigma}(\hat{C}^{i,j}, \hat{C}^{k,l}) \tag{4.21}$$

**Covariance matrix of cross-bundles spectra** In Section 3.5, we introduced the final estimator of the power spectrum, obtained by averaging over the cross-bundles spectra. The bundles are chosen to have uniform noise, and the cross-bundles approach is chosen to remove noise bias from the band powers. In Polenta et al. (2005); Lueker et al. (2010) is presented the expression of the cross-bundles spectra covariance matrix in the uniform noise case across bundles, which is summarized in the following. The cross-bundles spectra covariance matrix prior to any instrumental beam or transfer function correction is given by

$$\hat{\Sigma}^{\mathrm{mean}} = \hat{\Sigma}^{\mathrm{C,C}} + \hat{\Sigma}^{\mathrm{C,N}} + \hat{\Sigma}^{\mathrm{N,C}} + \frac{n_{\mathrm{BUN}}}{n_{\mathrm{BUN}} - 1} \hat{\Sigma}^{\mathrm{N,N}}, \tag{4.22}$$

where $n_{\mathrm{BUN}} = 150$ is the number of cross-bundles spectra, and $\hat{\Sigma}^{C,C}$, $\hat{\Sigma}^{C,N}$ and $\hat{\Sigma}^{N,N}$ are the signal-signal, signal-noise and noise-noise covariance matrices, respectively. The noise-noise covariance matrix is rescaled by a factor $\frac{n_{\mathrm{BUN}}}{n_{\mathrm{BUN}}-1}$ to account for the additional variance coming from ignoring the auto spectra contributions. Under the assumption of uniform Gaussian noise, all those components are obtained using the ACC approximation, by applying

$$\hat{\Sigma}^{\mathrm{mean}} \approx \hat{\Sigma}(C, C) + \hat{\Sigma}(C, N) + \hat{\Sigma}(N, C) + \frac{n_{\mathrm{BUN}}}{n_{\mathrm{BUN}} - 1} \hat{\Sigma}(N, N), \tag{4.23}$$

Let me put forward that this framework correctly catches the noise correlations across frequencies. Indeed, applying Eq. (4.20) works for computing the noise correlation as long as the noise power spectrum is known. I measured the noise correlations across frequencies in Section 3.4. Nevertheless, an additional validation can be run to demonstrate that the noise-noise covariance matrix can be obtained using the analytical framework by comparing it to an estimator obtained from the data (Balkenhol and Reichardt, 2022). This comparison will validate the Gaussian assumption describing the noise. I will build

$$\hat{\Sigma}^{\mathrm{N,N;data}}_{bb'} \equiv \frac{n+1}{n-1} \sum_{i,j\neq i} \left[\delta\hat{C}^{ij}_{b}\,\delta\hat{C}^{ij}_{b'} + 2\left(\sum_{k\neq j,i} \delta\hat{C}^{ij}_{b}\,\delta\hat{C}^{ik}_{b'}\right)\right], \tag{4.24}$$

where $\delta\hat{C}$ are the mean-subtracted cross-bundles spectra, and the sum is over the different cross-bundles spectra. Given that definition, I can show that

$$\langle \hat{\Sigma}^{\mathrm{N,N;data}}_{bb'} \rangle = \hat{\Sigma}^{\mathrm{C,N}} + \hat{\Sigma}^{\mathrm{N,C}} + \frac{n(n+1)}{(n-1)^2} \hat{\Sigma}^{\mathrm{N,N}}. \tag{4.25}$$

As the number of bundles is limited, $n_{\mathrm{BUN}} = 150$, this estimator will eventually be regularized using assumptions on the covariance structure, to avoid any Monte Carlo bias of around $\sqrt{2/n_{\mathrm{BUN}}} \sim 12\%$, as in (Balkenhol and Reichardt, 2022). Further work is still needed to complete that comparison and is not reported in this thesis.

**Instrumental beam and transfer function correction** We construct the covariance matrix by block, relying on the ACC approximation and the `PolSpice` correction, and adding a correction step to account for the beam and transfer function. We build the covariance matrix as

$$\hat{\Sigma}_{\ell\ell'}^{\mathrm{X}\nu\mathrm{Y}\mu;\mathrm{Z}\alpha\mathrm{V}\beta} = \frac{1}{A_\ell^{\mathrm{X}\nu\mathrm{Y}\mu} A_{\ell'}^{\mathrm{Z}\alpha\mathrm{V}\beta}} \hat{\Sigma}_{\ell\ell'}(A^{\mathrm{X}\nu\mathrm{Y}\mu} C^{\mathrm{X}\nu\mathrm{Y}\mu} + N^{\mathrm{X}\nu\mathrm{Y}\mu}, A^{\mathrm{Z}\alpha\mathrm{V}\beta} C^{\mathrm{Z}\alpha\mathrm{V}\beta} + N^{\mathrm{Z}\alpha\mathrm{V}\beta}), \tag{4.26}$$

where $A$ models the instrumental and pixelation effect on the power spectrum, see Eq. (3.67) for the definition. The input power spectrum is obtained from a high accuracy `CAMB` run starting from the *Planck* parameters (Planck Collaboration et al., 2020c), the foreground components are obtained from the SPT-3G 2018 best-fit nuisance parameters (Balkenhol et al., 2022), the beam, pixel window function, and transfer function are obtained as described in Chapter 3 and the noise spectra are measured directly from the data as described in Section 3.4. The covariance matrix is corrected to match the band power definition. As the computation of the covariance implies inside convolution, this method yields different results compared to the naive approach for which the unbiased power spectrum is directly fed into the covariance algorithm.

In Fig. 3.5, we showed that the filtering strategy leads to a significant loss of $\ell m$ modes. The covariance needs to be further corrected to account for the lost power spectrum modes and the subsequent impact on its variance. Similarly, as for the power spectrum correction, we build a correction factor for the four points using an analytical computation obtained from an emulated filtering pipeline. This work is done by a collaborator and is not detailed in this thesis. The computation is exact in the case of white noise but is an approximation in the case of the true CMB. We validate this technique in the case of CMB over simulations as shown in Fig. 4.1, for which we drew 1000 simulations of CMB skies on which we applied anisotropic filtering using a 2D transfer function. In this specific set-up, I do not need more simulations to assert the precision of the correction since the compared realizations are correlated, hence the cosmic variance contribution to the statistical fluctuations is negligible. The latter was analytically estimated from the filtering setup, with emulated mock observations. We then computed the variance of the spectra computed on the filtered maps and compared it to the same quantity computed on the unfiltered maps. We found that the correction factor is precise at the 5% level and always overestimates the variance, which is a conservative approach that does not bias the parameters but only overestimates the final error bars (Sellentin and Starck, 2019). This work will be further detailed in an upcoming work (Hivon, Doussot, Camphuis, et al., *in prep.*). We correct the covariance matrix by multiplying it by the factor,

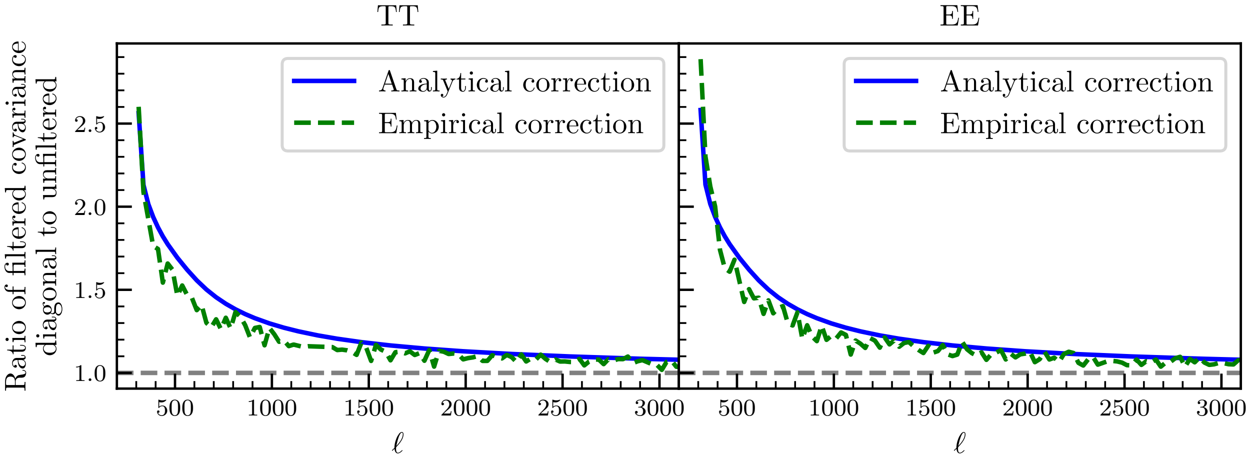


Figure 4.1: Impact of the correction on the power spectrum variance. The empirical correction is obtained with a set of 1000 simulations filtered with a set of $f_{\ell m}$. The analytical prediction is computed from the emulated mock observation pipeline. Analytical and empirical correction match within 10% level.

$$\hat{\Sigma}_{\ell\ell'}^{\mathrm{X}\nu\mathrm{Y}\mu;\mathrm{Z}\alpha\mathrm{V}\beta} \to \hat{\Sigma}_{\ell\ell'}^{\mathrm{X}\nu\mathrm{Y}\mu;\mathrm{Z}\alpha\mathrm{V}\beta} H_\ell H_{\ell'}. \tag{4.27}$$

This correction is done by rescaling the diagonal, thus assuming a fixed correlation structure. This technique is particularly adapted in this case since the harmonic space filtering does not introduce additional coupling between modes. Nevertheless, I assert that no spurious couplings appear due to filtering with simulations and validate that approach. The binning operator from Eq. (3.70) is finally applied and the covariance matrix is displayed in Fig. 4.2.

The conservative multipole cuts and binning scheme $(\ell_{\min}, \ell_{\max}, \Delta\ell) = (300, 4000, 25)$ lead to 148 band powers per spectrum combination. The low multipole cut is chosen from the transfer function cut-off, while the high multipole one is chosen to maximize cosmological information. The binning scheme is chosen to

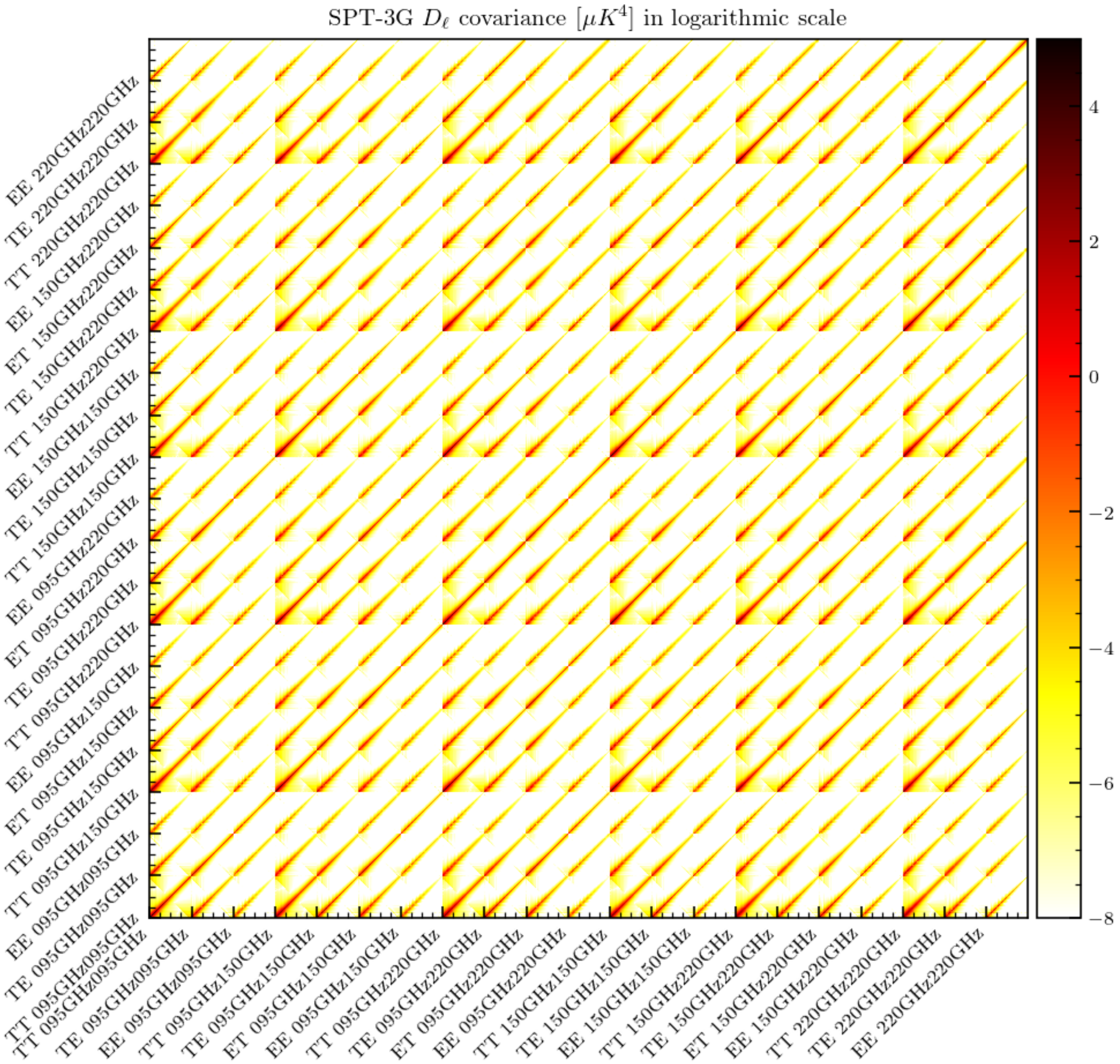


Figure 4.2: Pure Gaussian covariance matrix for the SPT-3G band powers.

minimize the correlation between band powers, while keeping a sufficient number of band powers to constrain the cosmological parameters, and will be updated when running likelihood tests. Auto-frequency block has 3 frequency combinations, while cross-frequency has 4 since ET differs from TE, which gives a total of 21 spectral combinations. The covariance size is then $21 \times 148 = 3108$. The covariance is positive definitive and its eigenvalues are displayed in Fig. 4.3, for the whole covariance, but also the TTTT, TETE, EEEE blocks (including all frequency combinations) and the frequency blocks (including all available Stokes combinations). Diagonals of the covariance per Stokes index are displayed in Fig. 4.4 for TTTT (top), EEEE (middle), and TETE (bottom). The covariance of the 220GHz × 220GHz is larger at high multipoles due to the high noise levels of this frequency channel, see Table 2.3.

**Forecasted constraints** The obtained covariance is used by the collaboration to produce forecasted constraints on the cosmological parameters by running a MCMC analysis on a set of simulated band powers. The forecasts are displayed in Fig. 4.5, showing the predicted constraining power of the SPT-3G 19/20 data and comparing it to *Planck* and SPT-3G 2018 results. The excellent noise levels of the data set will allow us to put constraints three times smaller than SPT-3G 2018 and comparable to the *Planck* experiment. In this figure, we also display the forecasts for the full five-year survey, which is obtained by rescaling the 19/20 spectra. The constraints obtained with the first two years are only 20% wider than the total survey. Let me emphasize that I did not produce those constraints and show them to demonstrate the full capacity of this data set.

### 4.2.3 Additional covariance contribution

**Inpainting contribution to the covariance matrix**

In Section 3.6, we described the Gaussian constrain realization procedure that is used to fill the masked sources of the map. This is a crucial step to the pipeline as it allows the use of the ACC framework from Camphuis et al. (2022) for the covariance matrix estimation since the latter is accurate in the absence of small-scale features in the mask. For a detailed motivation of inpainting, please refer to Section 3.6.1. However, the inpainting procedure makes use of a simulated CMB sky to fill the masked sources and to correct for the bias induced by Wiener filtering. This implies that the final sky contains a mixture of true data and fake CMB that we need to account for in the covariance matrix estimation, as we want to marginalize the fake data. The final band powers are obtained as in Eq. (3.87), where we defined two matrices, $Z^{(1)}$ and $Z^{(2)}$, that represent the effect of inpainting on the power spectrum, and we have

$$\hat{C}^I = (\mathrm{Id} + 2Z^{(1)} + Z^{(2)})\hat{C}^D + Z^{(2)}\hat{C}^R. \tag{4.28}$$

We marginalize over the random realization of the CMB sky, relying on the fact that the latter is a Gaussian random field, and thus obtain that the additional variance induced by this marginalization is given by

$$\hat{\Sigma}^I = Z^{(2)}\hat{\Sigma}Z^{(2)\top}. \tag{4.29}$$

The final covariance matrix is then obtained by adding the two contributions, and further assuming that binning allows to write $Z^{(2)}$ as a diagonal matrix, and $\left(PZ^{(2)}\hat{C}^R\right)_b = \rho_b(\hat{C}^R)_b$. We finally obtain, using the binned rescaling factor defined in Eq. (3.84),

$$\hat{\Sigma}_{bb'} \to (1+\rho_b)(1+\rho_{b'})\hat{\Sigma}_{bb'}. \tag{4.30}$$

We measure the inpainting rescaling with 100 simulations as in Fig. 3.13. This procedure will increase by $\sim 10\%$ the error bars on the band powers at high multipoles, but do not impact the large scales. Indeed, Gaussian-constrained realization affects only signals on scales of the size of the holes in the map. I display the ratio of covariance diagonal after the inpainting correction in Fig. 4.6, and show that it is minimal compared to the additional variance expected from masked sources, see Section 3.6.1 for details. This further justifies the inpainting procedure, as it allows for a 40% reduction of the variance at small scales.

**Lensing contribution**

In this section, the lensing covariance has been computed by a collaborator. The lensing of large-scale structure on CMB photons mixes the modes leading to non-zero four-point functions and thus induces a contribution to the non-diagonal covariance matrix. The full-sky lensing contribution to the covariance matrix can be analytically modeled as in Benoit-Lévy et al. (2012); Hotinli et al. (2022), using the unlensed power spectra $C^{\mathrm{XY,u}}$ and the lensing potential power spectrum $C^{\phi\phi}$. The lensing contribution to the covariance matrix is given by

$$\Sigma^{\mathrm{lens}}_{\ell_1\ell_2} = \left\{ \sum_\ell \left( \frac{\partial C^{\mathrm{XY}}_{\ell_1}}{\partial C^{\mathrm{XY,u}}_\ell} \Sigma^{\mathrm{XuYu,ZuVu}}_{\ell\ell} \frac{\partial C^{\mathrm{ZV}}_{\ell_2}}{\partial C^{\mathrm{ZV,u}}_\ell} \right) (1-\delta_{\ell_1\ell_2}) + \sum_\ell \left( \frac{\partial C^{\mathrm{XY}}_{\ell_1}}{\partial C^{\phi\phi}_\ell} \Sigma^{\phi\phi,\phi\phi}_{\ell\ell} \frac{\partial C^{\mathrm{ZV}}_{\ell_2}}{\partial C^{\phi\phi}_\ell} \right) \right\}. \tag{4.31}$$

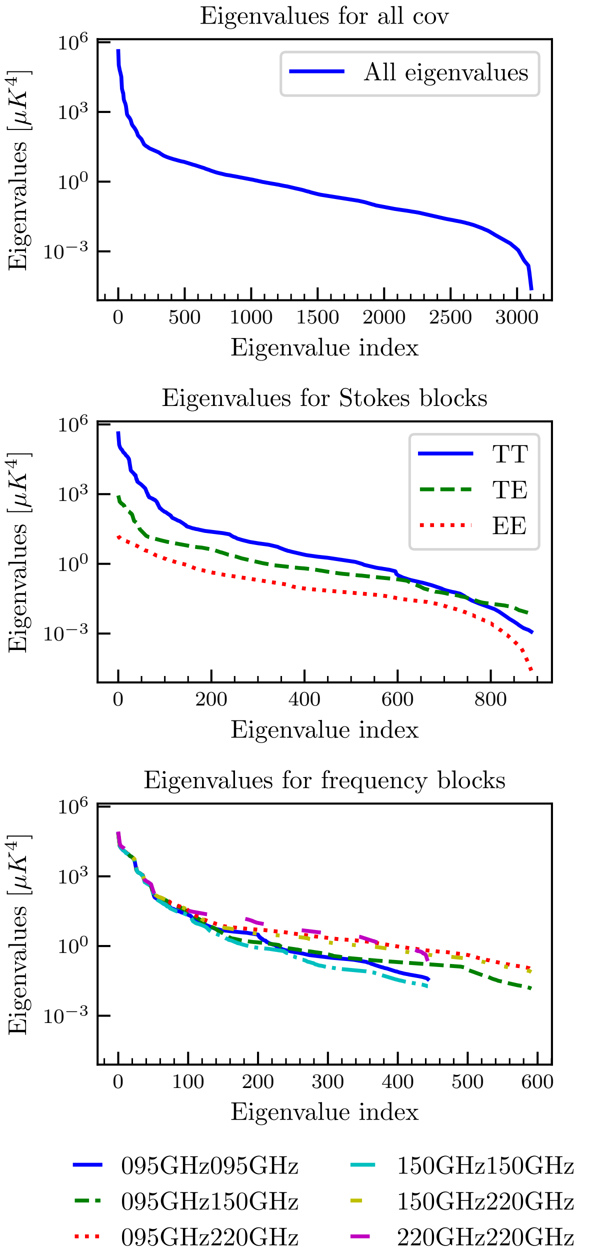


Figure 4.3: Eigenvalues of the covariance matrix. Top: all covariance matrix. Middle: TTTT, TETE, EEEE blocks. Bottom: frequency blocks. In this last case, the number of eigenvalues is larger for cross-frequency blocks since both ET and TE are included.

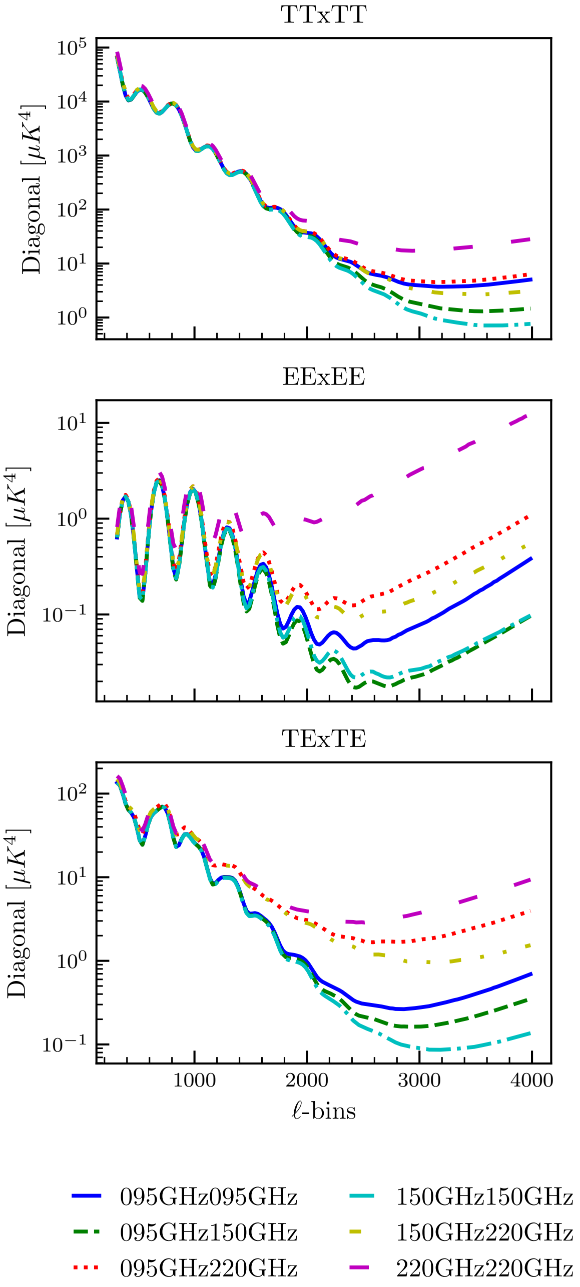


Figure 4.4: Diagonals of the covariance matrix. Top: TTTT block. Middle: EEEE block. Bottom: TETE block. The diagonals have the expected hierarchical ordering from noise levels and beam deconvolution.

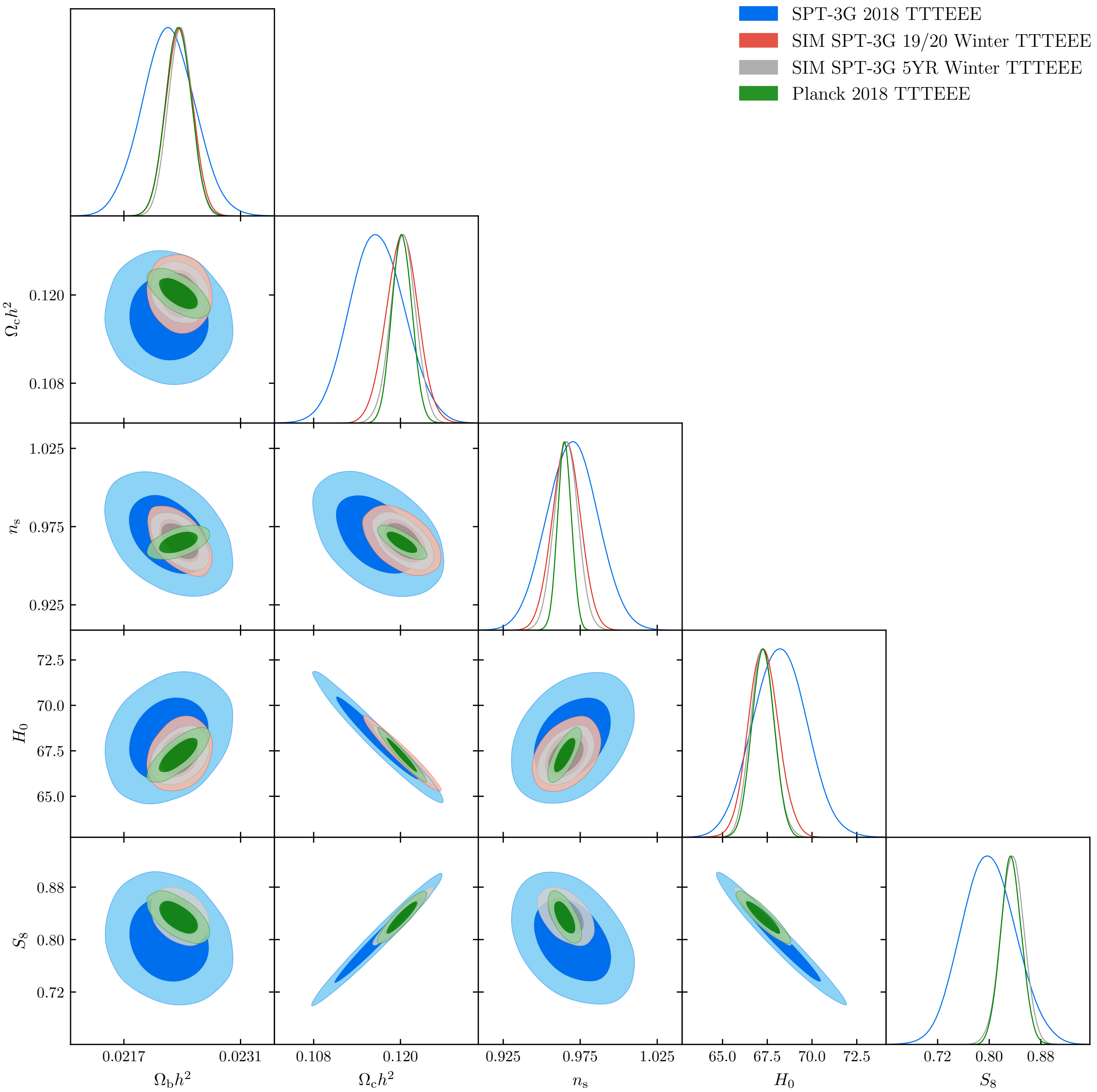


Figure 4.5: *Planck* 2018 TT TE EE (green) and SPT-3G 2018 TT TE EE (blue) cosmological constraints, with forecasted constraints around *Planck* best fit for SPT-3G 19/20 (red) and SPT-3G full five years survey (grey). The analysis of the primary anisotropies of the full five-year patch will yield comparable constraints to *Planck*.

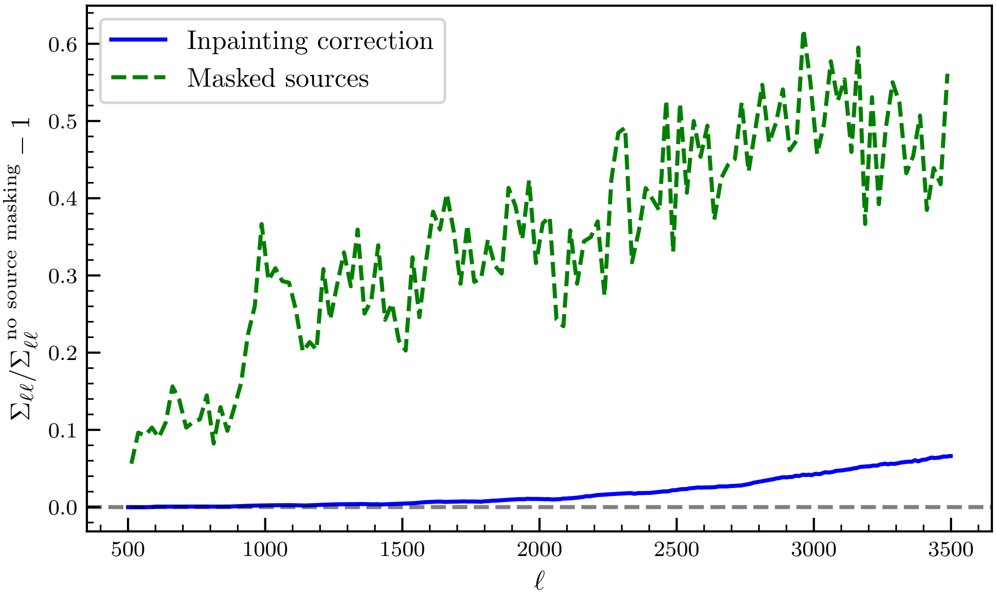


Figure 4.6: Ratio of covariance diagonals. The blue line displays the inpainting correction of the covariance. The green line is the ratio of the covariance diagonal with masked sources to the one with unmasked sources. Inpainting allows us to reduce the variance of the power spectrum estimator by 40% at small scales.

The non-Gaussian covariance effectively reduces the number of independent modes and thus increases the error bars on the band powers. We compute the non-Gaussian contribution using the `FisherLens` code [1], relying on the `CLASS` Boltzmann solver, using as input *Planck* cosmological parameters, consistently with the computation of the pure Gaussian contribution of the covariance. The lensing contribution is then rescaled by the fraction of the sky to capture the loss of spectral modes due to masking. We ignore the mode correlation induced by the mask as the lensing signal is subdominant. We finally correct the covariance to match the `PolSpice` estimator and write

$$\hat{\Sigma}_{bb'} \to \hat{\Sigma}_{bb'} + \hat{\Sigma}^{\text{lens}}_{bb'}, \quad \text{where} \quad \hat{\Sigma}^{\text{lens}}_{bb'} = \frac{1}{f_{\text{sky}}} \left( PG\Sigma^{\text{lens}}G^{\top}P^{\top} \right). \tag{4.32}$$

The resulting contribution to the covariance matrix is presented in Fig. 4.7. Although this contribution is more densely populated, it remains subdominant compared to the Gaussian component. To evaluate this contribution, we compute the lensing correlation matrix by normalizing the lensing covariance with the square root of the Gaussian covariance diagonal. The resulting correlation is shown in Fig. 4.8. It is worth noting that lensing introduces correlations at the percent level. Finally, I will include the super-sample lensing by following the procedure laid out by Crites et al. (2015); Henning et al. (2018); Dutcher et al. (2021) and modify the theoretical estimation of the power spectrum at the likelihood level with a nuisance parameter, see (Manzotti et al., 2014) or Eq. 16 of Dutcher et al. (2021).

### Beam covariance

The beam covariance accounts for uncertainty arising from a statistical error in the beam measurements as well as systematic effects. The beam covariance has not been computed yet, but we will follow the method in Henning et al. (2018); Dutcher et al. (2021); Balkenhol et al. (2022) and build a correlation matrix a sum of correlation matrices measured by varying input parameters of the beam model for each of the potential sources of error. The final beam correlation matrix will be built as

$$\rho^{\text{beam}}_{bb'} = \left[1 - \left(1 + \frac{\delta B_b}{B_b}\right)^{-2}\right]\left[1 - \left(1 + \frac{\delta B'_b}{B'_b}\right)^{-2}\right]. \tag{4.33}$$

The beam covariance matrix will be constructed from the correlation matrix by multiplying the diagonal elements by the band power variance and adding the output to the covariance matrix, as

$$\Sigma^{\text{beam}}_{bb'} = \rho^{\text{beam}}_{bb'}\sqrt{\hat{\Sigma}_{bb}\hat{\Sigma}_{b'b'}} \quad \text{and} \quad \hat{\Sigma}_{bb'} \to \hat{\Sigma}_{bb'} + \Sigma^{\text{beam}}_{bb'}. \tag{4.34}$$

This work is in progress.

---

[1] https://github.com/ctrendafilova/FisherLens

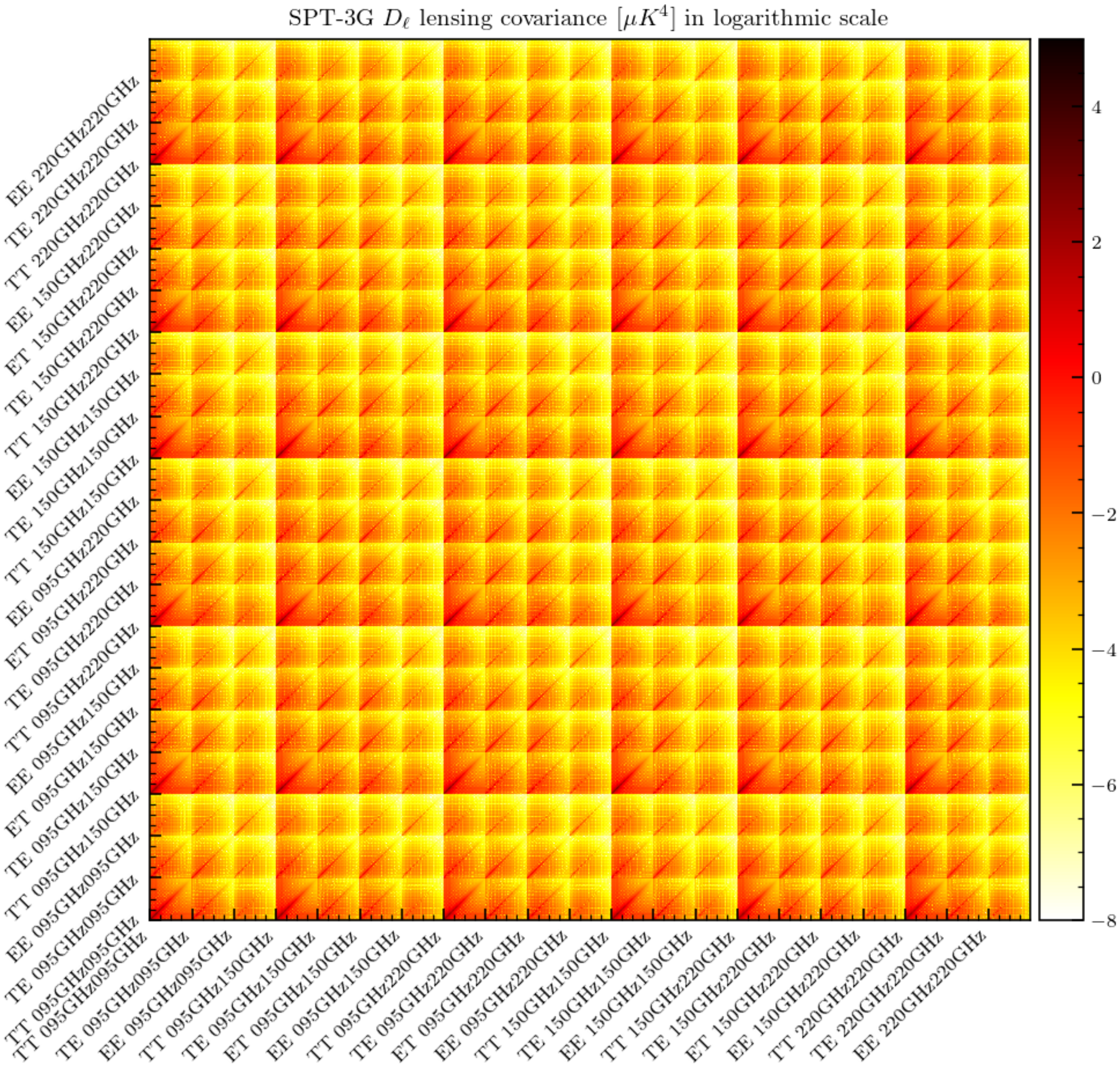


Figure 4.7: Non-Gaussian lensing contribution to the covariance matrix for the SPT-3G band powers.

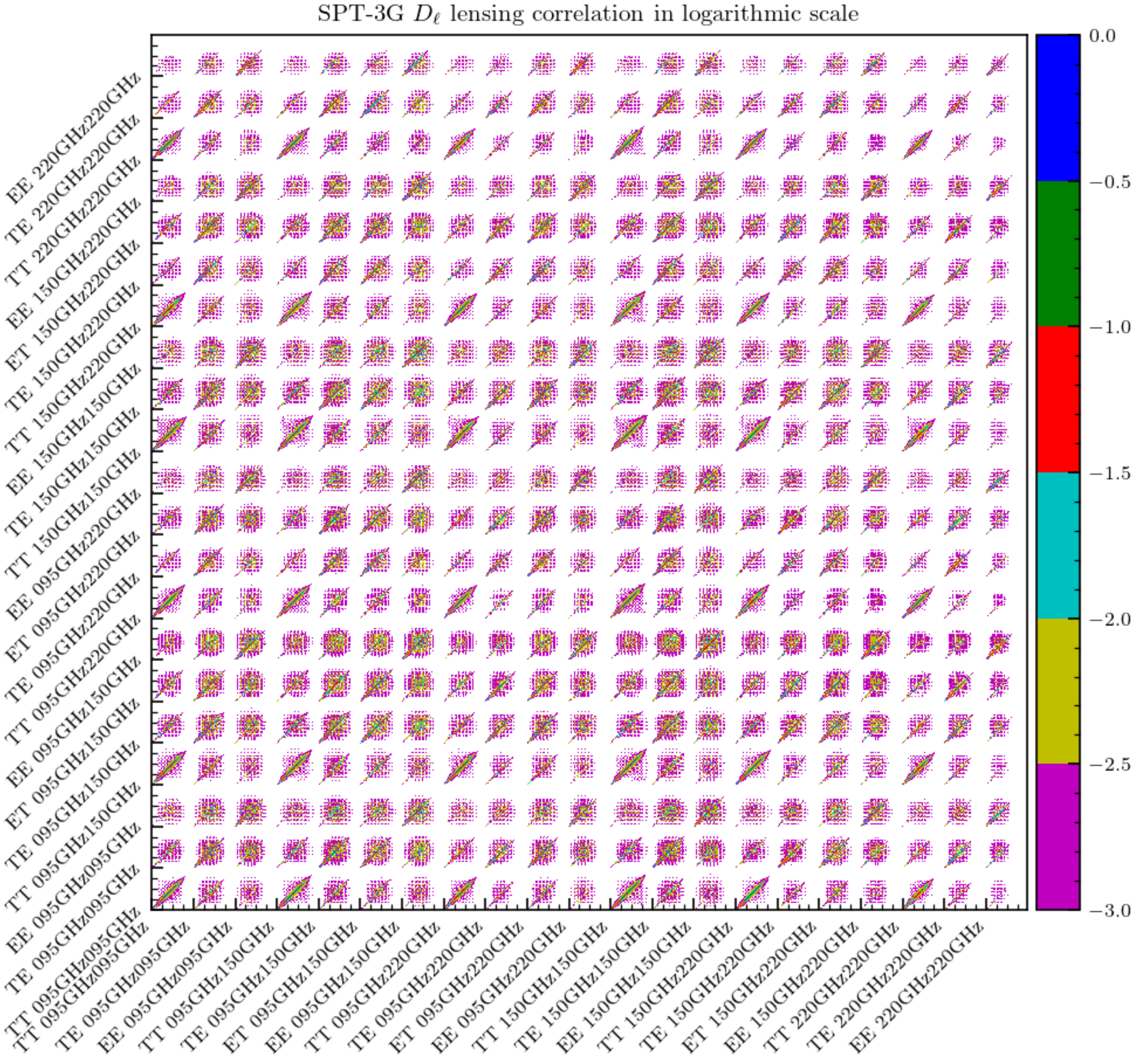


Figure 4.8: Non-Gaussian lensing correlation matrix for the SPT-3G band powers. The matrix is obtained by dividing the lensing covariance matrix by the square root of the Gaussian covariance (Fig. 4.2) diagonal elements.

## 4.3 Consistency tests

### 4.3.1 Null tests

In this subsection, I briefly report the work conducted by a collaborator on null tests. I do not present the results, as the work is not finalized yet.

The band powers of the inpainted maps, obtained by averaging the cross-bundle spectra as described in Section 3.5 and displayed in Fig. 3.22 (prior to the inpainting procedure), serve as the final output of the analysis pipeline. To validate the analysis pipeline, we conduct a set of consistency tests referred to as **null tests** on the band powers. These tests aim to demonstrate that the data is not significantly affected by systematic biases. Data splits are created based on potential sources of systematic errors, including azimuth cuts to test for ground pick-up, chronological cuts to examine time-dependent effects, moon up/moon down cuts to evaluate beam sidelobe pickup, sun up/sun down cuts to assess sun contamination, scan-direction cuts to test for left-right direction-dependent effects, and wafer cuts to isolate the impact of detector groups. Null maps are generated by taking the difference between map bundles created from the data split criteria. The null spectra are computed from those maps, and the error bars are determined from the noise levels. We verify the consistency of the spectra distribution with zero expectation using a $\chi^2$ test.

### 4.3.2 Likelihood consistency tests

To validate the likelihood and the analysis pipeline, we will conduct a series of consistency tests at the band power, likelihood, and parameter constraint levels. These tests are designed to evaluate the precision and the consistency of the results while keeping them blinded, to avoid cognitive bias. In this section, I report the plan for likelihood consistency tests, which are entirely inspired by the methodology developed in Balkenhol et al. (2022). This work has not been performed yet.

**Power spectrum tests** We will verify that the analysis pipeline produces consistent band powers across different frequency combinations by performing power spectrum consistency checks similar to those in Planck Collaboration et al. (2016); Mocanu et al. (2019); Balkenhol et al. (2022). Firstly, a minimum variance combination of the power spectrum is computed using the formula:

$$\hat{\mathcal{C}}^{\mathrm{MV}} = \left(X^T \hat{\Sigma} X\right)^{-1} X \hat{\Sigma}^{-1} \hat{\mathcal{C}}, \tag{4.35}$$

where $X$ is the frequency combination matrix and $\hat{\mathcal{C}}$ represents the vector of band powers. The minimum variance combination is then compared to the band powers of each frequency combination using a $\chi^2$ test after removing the best-fit foreground model.

Additionally, we will conduct a conditional spectrum test, in which the band powers of a specific frequency combination are compared to the band powers of other frequency combinations, conditioned on the band powers of the best-fit combination. The conditional band powers are calculated as follows:

$$\hat{\mathcal{C}}^{\nu\times\mu,\mathrm{cond}} = \mathcal{C}^{\nu\times\mu,\mathrm{BF}} + \hat{\Sigma}^{\nu\times\mu;\mathrm{others}} \left(\hat{\Sigma}^{\mathrm{others}\times\mathrm{others}}\right)^{-1} \left(\hat{\mathcal{C}}^{\mathrm{others}} - \mathcal{C}^{\mathrm{others;BF}}\right), \tag{4.36}$$

where a covariance matrix for the conditional band powers is constructed as the Schur complement of the covariance matrix at the specific frequency combination. The resulting conditional band powers are compared to the band powers of the best-fit combination using a $\chi^2$ test. This test will be repeated to ensure consistency between TT, TE, and EE band powers.

**Parameter-level tests** The internal consistency of the dataset needs to be assessed at the parameter level through tests. Cosmological parameters will be estimated for selected subsets of the data, and consistency will be evaluated using the approach developed in Gratton and Challinor (2019). This involves transforming the difference of parameters $\Delta\vartheta$ between two data subsets into a $\chi^2$ test statistic by utilizing the difference of parameter covariance matrices of the full dataset and the subset, as

$$\chi^2 = \Delta\vartheta^T \left(\Sigma_\vartheta^{\mathrm{full}} - \Sigma_\vartheta^{\mathrm{subset}}\right)^{-1} \Delta\vartheta. \tag{4.37}$$

This test is model-dependent and can be run for multiple models

**Robustness of cosmological constraints** To further assess the accuracy of the likelihood, cosmological constraints will be performed with variations to the pipeline. As mentioned in Balkenhol et al. (2022), this can involve varying the priors of foreground nuisance parameters, rescaling the beam covariance to model possible systematic errors, or adjusting the prior for the optical depth $\tau$.

In addition, since significant improvements were made to the pipeline, additional tests are required. Firstly, the updated analysis pipeline should be compared to the former version, in which the covariance matrix was obtained from simulations and the noise covariance from the data. The robustness of the cosmological parameter constraints after the inpainting procedure will be assessed by producing a new set of inpainted maps with a modified fiducial power spectrum and running the full analysis pipeline on the new band powers. Furthermore, we will also evaluate the effect of not considering the inpainting rescaling of the covariance matrix. This approach is expected to reduce the constraints on cosmological parameters but may introduce bias to the results if the inpainting procedure is not accurate.

## Conclusion

In this chapter, we reviewed the detailed computation of the SPT-3G covariance matrix, which relies on the framework developed in Camphuis et al. (2022). Further implementations were necessary to adapt the framework to the SPT-3G analysis. We also presented the consistency tests to be performed on the data to validate the likelihood and the analysis pipeline. The results of these tests and the obtained constraints on ΛCDM parameters will be presented in a future publication.

# Chapter 5

# Perspectives and conclusion

The new upcoming SPT-3G 19/20 data set studied in this thesis will provide maps with noise levels three times lower than in 2018. The measured band powers will put an independent and more competitive measurement of the cosmological parameters and will be a valuable addition to the current CMB data sets such as *Planck*, ACT, Bicep. ΛCDM constraints will be as tight as *Planck*, and the complementary range of angular scales will help to break degeneracies between parameters and probe new physics beyond the standard model. We summarize the forecasted constraints in Table 5.1, to be compared with Table 1.1. Obtaining tight constraints from high-resolution and low noise levels experiments requires the development of new analysis methods to extract cosmological information from the data with high precision. Indeed, as noise levels decrease, the likelihood gets more sensitive to systematic effects and errors are more likely to propagate through the analysis pipeline. Constraints must be robust to instrumental systematics, analysis choices, and cognitive biases. In this thesis, we developed new tools for CMB analysis and implemented them in the pipeline of the SPT-3G 19/20 data set. Such tools will be crucial to obtaining robust constraints from the data, and pave the way for the complete SPT-3G 2019-2023 analysis, which will be signal-dominated on all the multipole range, but also Stage-4 CMB experiments that will be also limited by systematics (SO, CMB-S4). Additional work is still required to perform the complete SPT-3G 19/20 analysis.

| | | ΛCDM | | |
|---|---|---|---|---|
| $\Omega_b h^2$ | $\Omega_c h^2$ | $H_0$ | $n_s$ | $\ln(10^{10} A_s)$ |
| 0.00018 | 0.0024 | 0.91 | 0.0099 | 0.016 |
| 0.00015 | 0.0012 | 0.54 | 0.0042 | 0.014 |

Table 5.1: Forecasted error bars on the ΛCDM model parameters from the SPT-3G 19/20 data set on the first row, with *Planck* error bars on the second row for comparison. The constraints are obtained from the primary anisotropies TT,EE, and EE spectrum with a $\tau$ prior.

**Power spectrum** In Section 3.2, we examined the pseudo-power spectrum method, which is a nearly optimal estimator that efficiently and accurately characterizes the mean. However, this estimator is biased, and its polarization component is coupled due to the masking of the sky. To address this, I introduced the de-biasing procedure in `PolSpice`, which is necessary to correct the bias caused by the mask and decouple the polarization. I proposed an enhancement to the `PolSpice` operation that writes the equivalent operation in harmonic space, to simplify the correction of the pseudo-power spectrum covariance matrix. An additional improvement would be to perform the full computation of the decoupling kernel $_{+2}G$ without relying on emulation from the software. This development would help reduce computational requirements. The decoupling kernel is crucial for modeling the coupling of polarization, particularly when using the `PolSpice` method to measure B modes of a masked sky, as significant contamination arises from the coupling with E modes, and could be of interest for analyses relying on real space de-biasing for measuring B modes. However, obtaining a complete computation of the kernel necessitates the analytical development of non-orthogonal masked spin-2 spherical harmonics, which is a challenging task. Furthermore, it's important to note that the `MASTER` approach assumes a fixed and uncorrelated mask with the underlying signal. As noted in recent work Fabbian et al. (2021); Surrao et al. (2023), power spectrum analyses are expected to be biased when the masked regions are correlated with the signal of interest. While this does not impact the analysis of the SPT-3G 19/20 data set, given that the masked source signal is uncorrelated with CMB, it is important to address this issue for CMB lensing or foreground analysis. Developing a method to unbias the power spectrum in the presence of correlated masked regions using real space de-biasing would be of interest for future analyses.

Understanding the impact of filtering on the power spectrum is crucial for obtaining unbiased constraints. In order to achieve this, we conducted a comprehensive analysis of the transfer function at both the harmonic coefficients and power spectrum levels, as described in Section 3.3. We compared the impact of the estimated values of the transfer functions obtained from the mock-observations pipeline with those from an emulated pipeline. The emulated pipeline provided a cost-effective way to test the effects of different filtering choices on the time streams and assess their impact on the final signal. Our analysis revealed that the transfer function was well estimated by the pipeline, and we gained a good understanding of how filtering affects the power spectrum. Prior to the final analysis, we will evaluate the impact of changing the fiducial cosmology on the transfer function to ensure the robustness of the estimation. Further work could be undertaken to explore filtering choices using the emulated pipeline and propagate them to the cosmological constraints, to optimize the analysis. This approach would be particularly useful in preparation for CMB-S4, where systematic effects are expected to dominate the errors and can be mitigated by adapting filtering choices.

In Section 3.4 and Section 3.5, I evaluated the signal and noise power spectra derived from the data products of the map-making pipeline. The temperature noise power spectrum exhibited the expected behavior, showing frequency-correlated atmospheric contamination at large scales across the three frequency channels, and flatness due to white noise at small scales. The polarized noise spectrum demonstrated the anticipated flatness, consistent with the properties of white noise. These two key data products were properly debiased using the transfer function, as well as the deconvolution of the beam and pixel window function.

Finally, in Section 3.6, I presented a new inpainting pipeline for temperature and polarization maps, which plays a crucial role in mitigating the impact of source masking on the coupling of the power spectrum. As the noise levels decrease, more sources are detected with higher signal-to-noise ratios and subsequently masked to avoid artifacts introduced by filtering. Although the power spectrum remains unbiased with source masking, the presence of additional small-scale features in the mask leads to undesired coupling between modes of the power spectrum and increased variance. To address this issue, I developed and implemented a pipeline that replaces the masked regions with Gaussian-constrained realizations of the CMB anisotropies. This approach utilizes a fiducial power spectrum and a realistic CMB simulation at the location of the masked regions as prior knowledge. The method is unbiased under some assumptions, as shown by applying it to realistic mock observations. Additionally, the inpainting method is robust to variations in the fiducial power spectrum. Although the current version of the pipeline is already efficient, implementing GPU acceleration for the inpainting process would be beneficial in reducing computational costs and enabling the inpainting of larger regions or utilizing more complex constraints. Further improvements, such as reducing the map resolution in the constraining region, could also be explored. Additionally, a more physically motivated development of this work would involve incorporating the lensing contribution into the constrained realizations by utilizing a lensing CMB potential as prior knowledge. Indeed, previous referred research Fabbian et al. (2021) has shown that point-source masking introduces a small bias in the measurement of the lensing power spectrum due to the correlation between masked sources and the lensing signal, which might become significant as noise levels are reduced. Reconstructing the full-sky map with prior lensing knowledge could help mitigate this bias.

**Covariance matrix** In previous SPT analyses, the estimation of the covariance matrix relied on mock observations, which involved regularization techniques and significant computational costs. To address this, we developed a new semi-analytical framework in Chapter 4 to compute the covariance matrix for the SPT-3G 19/20 likelihood. This is particularly important because power spectrum modes are expected to exhibit increasing coupling with decreasing sky coverage, and the SPT-3G 19/20 survey only covers 4% of the sky.

To tackle this challenge, I implemented the first exact computation of the harmonic four-point function of a masked Gaussian isotropic signal on the sphere. To reduce computational costs, we used this exact computation to evaluate the accuracy of existing approximations and implemented the most precise approximation of the covariance matrix to date, ACC, which can be obtained at a low computational cost. This comprehensive study provided a profound understanding of the coupling effects in the pseudo-power spectrum estimation. Furthermore, we extended the framework to incorporate the use of the `PolSpice` spectrum estimator, leveraging the decoupling kernels introduced in the previous chapter. This new framework, used to compute the covariance matrix for the SPT-3G 19/20 likelihood, is discussed in detail in Camphuis et al. (2022). There are potential improvements that can be proposed for this framework. Firstly, while our focus was on CMB anisotropies, the framework itself is general and can be applied to any isotropic signal on the sphere, such as weak lensing or galaxy clustering, which also rely on the pseudo-power spectrum. Secondly, the exact computation can be further adapted to study homogeneous but anisotropic signals by modifying Eq. (4.5) to account for a $m$-dependence of the power spectrum as

$$\langle \tilde{\mathrm{T}}_{\ell m} \tilde{\mathrm{T}}^*_{\ell' m'} \rangle = \sum_{\ell_1 m_1} C^{\mathrm{TT}}_{\ell_1 m_1} {}_0 I_{\ell m \ell_1 m_1}[W] {}_0 I^*_{\ell' m' \ell_1 m_1}[W].$$

Additionally, GPU implementations are highly suitable for this problem and would enable fast and exact computation of the covariance matrix, building upon existing implementations such as Schaeffer (2013) and Tian et al. (2022). Finally, the exact computation could be utilized for large-scale analyses as it is computationally cheaper for small multipoles. It could be used to analytically model the additional variance of B modes from E modes in the presence of a mask, and potentially incorporate the anisotropy of the signal.

The analytical covariance ACC approximation forms the core of the SPT-3G covariance matrix. We extended the pipeline to account for multi-frequency analysis and signal correction, involving beam and pixel window functions. The signal anisotropies due to filtering are modeled and included, accounting for excess variance resulting from the loss of modes. The resulting covariance matrix exhibits the expected structure. We also consider additional contributions, such as the impact of beam uncertainty and the marginalizing of inpainting random realization. While each step of the analysis is validated against simulations, further work is required to compare the complete covariance estimator pipeline with a comprehensive suite of simulations. This procedure also needs to be conducted for the noise covariance matrix.

We finally detailed the additional work to be conducted on the likelihood prior to obtaining cosmological parameters. This includes extended consistency checks and validation procedures, which are crucial validations of the robustness of the results and need to be validated prior to unblinding to avoid cognitive biases.

**Additional perspectives for covariance computation** Analytical approximations of the covariance matrix are known to be inaccurate when the analysis mask displays small-scale features, such as holes that mask point sources. I am participating in a collaborative project to propose an analytical estimation of the covariance in such a case. This computation relies on the assumption that holes in the mask are Poisson distributed on the sphere, with a known density, and relies on a flat-sky computation of the coupling.

Additionally, Chartier et al. (2021); Chartier and Wandelt (2022) shows that optimal covariance estimator can be obtained from pairs of correlated simulations. As highlighted by Eq. (4.10), a high number of Monte Carlo simulations are necessary to estimate the coupling between modes accurately. The aforementioned technique can help us to improve the precision of the covariance matrix estimations by exploiting the correlations between sets of simulations. I am collaborating with the author of the previously cited work to demonstrate the efficiency of such a method in the case of CMB analysis.

In this thesis, I have outlined the efforts undertaken to establish the likelihood of the SPT-3G 19/20 dataset and outlined plans for analysis. The findings of this study will impose tight constraints on cosmological parameters but also delve into exciting new physics. The cosmological analysis of the ΛCDM model and extensions will be performed, alongside exotic models. Furthermore, this research serves as a stepping stone toward analyzing the full SPT-3G five-year survey and introduces innovative techniques for forthcoming experiments, which promise to illuminate previously uncharted aspects of cosmology.

# Appendix

---

I acknowledge the use of Grammarly[1] and ChatGPT[2] for language and grammar corrections of this manuscript. The analysis and visualization of data in this manuscript was made with `ipython` (Pérez and Granger, 2007), `matplotlib` (Hunter, 2007), `numpy` (van der Walt et al., 2011) and `scipy` (Virtanen et al., 2020).

## Thoughts on the carbon footprint of the Ph.D.

This section aims to provide a preliminary assessment of the environmental impact of a Ph.D. program in terms of carbon footprint, employing simple assumptions and rough estimates. It is not intended to be a comprehensive or rigorous study, but rather an initial attempt to quantify the environmental impact associated with research activities. The objectives of this assessment are twofold. Firstly, on a personal level, it is meant to familiarize oneself with the tools and considerations related to the environmental impact of research. Secondly, on a broader scale, it aims to raise awareness about sustainable research practices and encourage the research community to adapt their projects, tools, and mindset accordingly. It is understood that choices made in this approach are open to discussion and critique. Feedback and discussions on this topic are welcomed and appreciated. I would like to thank Benoît Sauty and Jean-François Cardoso for inspiring this approach and providing the tools necessary to carry out this assessment.

| | | Emissions (kg CO2-eq) | ~ % of emissions |
|---|---|---|---|
| Plane | 2 x Paris-Chicago | $4204 \pm 2872$ | 26 |
| | Paris-Venise | $486 \pm 340$ | 3 |
| | Paris-Ferrara | $486 \pm 340$ | 3 |
| Train and bus | Paris-Trento | $28 \pm 17$ | <0.2 |
| | Paris-Cambridge | $16 \pm 9$ | <0.1 |
| | Paris-Grenoble | $4 \pm 1$ | <0.1 |
| | Venise-Sesto | $13 \pm 8$ | <0.1 |
| | Trento-Tonale | $4 \pm 3$ | <0.1 |
| Computing ressources | CPU Argonne ($10^4$h) | $1200 \pm 300$ | 8 |
| | CPU Paris ($7 \times 10^5$h) | $4000^{+2000}_{-1000}$ | 25 |
| Office | Electricity | $1500 \pm 300$ | 8 |
| | Heating/AC | $2250 \pm 300$ | 14 |
| | IT hardware | $1400 \pm 1400$ | 9 |
| Total | | $16^{+8}_{-7} \times 10^3$ | 100 |

Table A.1: Estimation of the carbon footprint associated with my Ph.D. The total emissions are estimated to be approximately 16 tons of CO2-eq, *i.e.* around 5 tons per year. Further work would be needed to refine the estimations.

The latest report from the Intergovernmental Panel on Climate Change (IPCC) reinforces the scientific consensus on anthropogenic causes of global warming, the severe consequences for biodiversity, human life, and the environment, and the urgent need to limit the temperature rise to 1.5 degrees Celsius. The synthesis report published in March 2023 states, "Without urgent, effective, and equitable mitigation and adaptation actions, climate change increasingly threatens ecosystems, biodiversity, and the livelihoods, health, and well-being of current and future generations." The estimation of the carbon footprint associated with my research activities is divided into three main categories: travel, computing resources, and office work. Emissions related

---

[1] https://app.grammarly.com

[2] https://chat.openai.com/

to transportation are calculated using data from the French database ADEME[3] and checked againt computation from labos1point5[4], taking into account the distance traveled for each mode of transportation (long-haul flights, short-haul flights, and high-speed trains). Errors associated result mostly from uncertainties on carbon impact of airplane contrails and energy efficiency of transportation. Workplace-home commuting is not taken into account as almost carbon-free thanks to bicycle usage. Estimations resulting from computing resources were conducted using the MachineLearning Impact calculator[5] presented in Lacoste et al. (2019), which takes into account CPU-time, the energy consumption of the hardware used (in kW), and the carbon intensity of electricity (measured in CO2-eq/kWh). Errors on this measure come mostly from unkowns of overall offset of building emission and variation in the carbon intensity of electricity, as well as intrinsic uncertainties to the emission of CPU-time due to lack of data. Estimates for office-related emissions are based on annual data provided by an internal commission at IAP, removing emissions already included in computing resources. Emissions of IT hardware are computed using labos1point5 tools, using a total of 4000€ of buyings. The results are summarized in Table A.1.

The need for plane travel arises directly from the international nature of my Ph.D., which necessitated trips to Chicago for collaboration meetings. It represents 1/3 of the overall footprint. It is important to note that long-haul flights contribute significantly to carbon emissions and should be minimized whenever feasible. Short-haul flights contribution is negligible, however it is worth emphasizing that choosing to travel from Paris to Tonale using train and bus, as opposed to plane, led to $\sim$ 400kg CO2-eq less release. Remote conferences or a combination of in-person and remote attendance (where only a few individuals are present on-site while others participate remotely) serve as viable alternatives to traditional in-person meetings. While it is customary in the field to attend international conferences a few times a year, it is crucial to acknowledge that this practice is unsustainable given the current global warming scenario. While face-to-face interactions are invaluable for fostering exchanges and achieving scientific breakthroughs, it is imperative to explore alternative approaches to reduce the carbon footprint associated with frequent travel. The second emission source is linked to the computational nature of my research, which necessitates access to high-performance computing facilities and represents 1/3 of the overall footprint. It is worth emphasizing that even though CPU computations conducted in Paris may take 100 times longer than those in Chicago, the significantly lower carbon intensity of French electricity helps to minimize the carbon impact of computing activities. Reducing the computational infrastructure would have a profound impact on research endeavors and may not be a readily implementable solution. Nevertheless, it is crucial to consider these factors when proposing new projects and actively seek alternatives to computationally intensive simulations whenever possible. Finally, office-related emissions represent the last 1/3 of the overall footprint. The laboratory is undergoing major renovations, including the replacement of all windows. These changes will significantly reduce the carbon footprint associated with office work. In addition, the laboratory is committed to reducing its carbon footprint by 7% per year, in line with the Paris Agreement's objectives.

---

[3] https://bilans-ges.ademe.fr/

[4] https://apps.labos1point5.org

[5] https://mlco2.github.io/impact#compute

# Bibliography

---

T. M. C. Abbott, M. Aguena, A. Alarcon, S. Allam, et al. Dark Energy Survey Year 3 results: Cosmological constraints from galaxy clustering and weak lensing. *Phys. Rev. D* , 105(2):023520, January 2022a. doi: 10.1103/PhysRevD.105.023520.

T. M. C. Abbott, M. Aguena, A. Alarcon, O. Alves, et al. Joint analysis of DES Year 3 data and CMB lensing from SPT and Planck III: Combined cosmological constraints. *arXiv e-prints*, art. arXiv:2206.10824, June 2022b.

G. E. Addison, D. J. Watts, C. L. Bennett, M. Halpern, et al. Elucidating ΛCDM: Impact of Baryon Acoustic Oscillation Measurements on the Hubble Constant Discrepancy. *Ap. J.* , 853:119, February 2018. doi: 10.3847/1538-4357/aaa1ed.

Simone Aiola, Erminia Calabrese, Loïc Maurin, Sigurd Naess, Benjamin L. Schmitt, Maximilian H. Abitbol, et al. The Atacama Cosmology Telescope: DR4 Maps and Cosmological Parameters. *arXiv e-prints*, art. arXiv:2007.07288, July 2020.

David Alonso, Javier Sanchez, and Anže Slosar. A unified pseudo-c framework. *Monthly Notices of the Royal Astronomical Society*, 484(3):4127–4151, 2019. ISSN 13652966. doi: 10.1093/mnras/stz093. Citation Key: Alonso2019.

R. A. Alpher and R. Herman. Evolution of the Universe. *Nature*, 162:774–775, November 1948. doi: 10.1038/162774b0.

Zachary Atkins, Adriaan J. Duivenvoorden, William R. Coulton, Frank J. Qu, et al. The atacama cosmology telescope: Map-based noise simulations for dr6. (arXiv:2303.04180), Mar 2023. arXiv:2303.04180 [astro-ph].

Éric Aubourg, Stephen Bailey, Julian E. Bautista, Florian Beutler, et al. Cosmological implications of baryon acoustic oscillation measurements. *Phys. Rev. D*, 92:123516, Dec 2015. doi: 10.1103/PhysRevD.92.123516.

J. E. Austermann, K. A. Aird, J. A. Beall, D. Becker, et al. SPTpol: an instrument for CMB polarization measurements with the South Pole Telescope. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 8452 of *Proc. SPIE*, September 2012. doi: 10.1117/12.927286.

L. Balkenhol and C. L. Reichardt. The parameter-level performance of covariance matrix conditioning in cosmic microwave background data analyses. *MNRAS* , 512(3):4394–4403, May 2022. doi: 10.1093/mnras/stac573.

L. Balkenhol, D. Dutcher, P. A. R. Ade, Z. Ahmed, et al. Constraints on Λ CDM extensions from the SPT-3G 2018 E E and T E power spectra. *Phys. Rev. D* , 104(8):083509, October 2021. doi: 10.1103/PhysRevD.104.083509.

L. Balkenhol, D. Dutcher, A. Spurio Mancini, A. Doussot, et al. A measurement of the cmb temperature power spectrum and constraints on cosmology from the spt-3g 2018 tt/te/ee data set. (arXiv:2212.05642), Dec 2022. arXiv:2212.05642 [astro-ph].

D. Baumann, M. G. Jackson, P. Adshead, A. Amblard, et al. Probing Inflation with CMB Polarization. In S. Dodelson, D. Baumann, A. Cooray, J. Dunkley, A. Fraisse, M. G. Jackson, A. Kogut, L. Krauss, M. Zaldarriaga, and K. Smith, editors, *American Institute of Physics Conference Series*, volume 1141 of *American Institute of Physics Conference Series*, pages 10–120, June 2009. doi: 10.1063/1.3160885.

Dominic Beck, Ari Cukierman, and W. L. Kimmy Wu. Bias on tensor-to-scalar ratio inference with estimated covariance matrices. *MNRAS* , 515(1):229–236, September 2022. doi: 10.1093/mnras/stac1775.

A. N. Bender, J.-F. Cliche, T. de Haan, M. A. Dobbs, et al. Digital frequency domain multiplexing readout electronics for the next generation of millimeter telescopes. In *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII*, volume 9153 of *Proc. SPIE*, page 91531A, July 2014. doi: 10.1117/12.2054949.

A. N. Bender, P. A. R. Ade, A. J. Anderson, J. Avva, et al. Integrated performance of a frequency domain multiplexing readout in the SPT-3G receiver. In *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VIII*, volume 9914 of *Proc. SPIE*, page 99141D, July 2016. doi: 10.1117/12.2232146.

A. N. Bender, A. J. Anderson, J. S. Avva, P. A. R. Ade, Z. Ahmed, et al. On-Sky Performance of the SPT-3G Frequency-Domain Multiplexed Readout. *Journal of Low Temperature Physics*, 199(1-2):182–191, April 2020. doi: 10.1007/s10909-019-02280-w.

C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, et al. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Preliminary Maps and Basic Results. *Ap. J. Suppl.* , 148:1–27, September 2003.

C. L. Bennett, D. Larson, J. L. Weiland, N. Jarosik, et al. Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Final Maps and Results. *Ap. J. Suppl.* , 208:20, October 2013. doi: 10.1088/0067-0049/208/2/20.

A. Benoit-Lévy, K. M. Smith, and W. Hu. Non-Gaussian structure of the lensed CMB power spectra covariance matrix. *Phys. Rev. D* , 86(12):123008, December 2012. doi: 10.1103/PhysRevD.86.123008.

A. Benoit-Lévy, T. Déchelette, K. Benabed, J.-F. Cardoso, D. Hanson, et al. Full-sky CMB lensing reconstruction in presence of sky-cuts. *Astron. Astrophys.* , 555:A37, July 2013. doi: 10.1051/0004-6361/201321048.

B. A. Benson, P. A. R. Ade, Z. Ahmed, S. W. Allen, K. Arnold, et al. SPT-3G: A Next-Generation Cosmic Microwave Background Polarization Experiment on the South Pole Telescope. In *Millimeter, Submillimeter, and Far-Infrared Detectors and Instrumentation for Astronomy VII*, volume 9153 of *Proc. SPIE*, page 91531P, July 2014. doi: 10.1117/12.2057305.

M. Bersanelli, G. F. Smoot, M. Bensadoun, G. de Amici, et al. Measurements of the cmb spectrum at centimeter wavelengths. *Astrophysical Letters and Communications*, 32:7, Jan 1995. ISSN 0888-6512.

F. Bianchini, W.?L.?K. Wu, P.?A.?R. Ade, A.?J. Anderson, et al. Searching for anisotropic cosmic birefringence with polarization data from sptpol. *Physical Review D*, 102(8), Oct 2020. ISSN 2470-0029. doi: 10.1103/physrevd.102.083504.

L. Bleem, P. Ade, K. Aird, J. Austermann, et al. An Overview of the SPTpol Experiment. *Journal of Low Temperature Physics*, page 196, January 2012. doi: 10.1007/s10909-012-0505-y.

L. E. Bleem, B. Stalder, T. de Haan, K. A. Aird, et al. Galaxy Clusters Discovered via the Sunyaev-Zel'dovich Effect in the 2500-Square-Degree SPT-SZ Survey. *Ap. J. Suppl.* , 216:27, February 2015. doi: 10.1088/0067-0049/216/2/27.

L. E. Bleem, S. Bocquet, B. Stalder, M. D. Gladders, et al. The SPTpol Extended Cluster Survey. *Ap. J. Suppl.* , 247(1):25, March 2020. doi: 10.3847/1538-4365/ab6993.

Michael Blomqvist, Hélion du Mas des Bourboux, Nicolás G. Busca, Victoria de Sainte Agathe, James Rich, Christophe Balland, et al. Baryon acoustic oscillations from the cross-correlation of Ly$\alpha$ absorption and quasars in eBOSS DR14. *Astron. Astrophys.* , 629:A86, September 2019. doi: 10.1051/0004-6361/201935641.

J. R. Bond, A. H. Jaffe, and L. Knox. Radical Compression of Cosmic Microwave Background Data. *Ap. J.* , 533:19–37, April 2000. doi: 10.1086/308625.

Martin Bucher and Thibaut Louis. Filling in cosmic microwave background map missing data using constrained gaussian realizations: Filling in missing cmb data. *Monthly Notices of the Royal Astronomical Society*, 424(3):1694–1713, Aug 2012. ISSN 00358711. doi: 10.1111/j.1365-2966.2012.21138.x.

E. F. Bunn, M. Zaldarriaga, M. Tegmark, A. de Oliveira-Costa, et al. E/B decomposition of finite pixelized CMB maps. *Phys. Rev. D* , 67:023501–+, January 2003.

Emory F. Bunn. Efficient decomposition of cosmic microwave background polarization maps into pure e , pure b , and ambiguous components. *Physical Review D*, 83(8):083003, Apr 2011. ISSN 1550-7998, 1550-2368. doi: 10.1103/PhysRevD.83.083003.

E. Camphuis, K. Benabed, S. Galli, E. Hivon, et al. Accurate cosmic microwave background covariance matrices: Exact calculation and approximations. *Astronomy & Astrophysics*, 668:A62, Dec 2022. ISSN 0004-6361, 1432-0746. doi: 10.1051/0004-6361/202243948.

J. E. Carlstrom, P. A. R. Ade, K. A. Aird, B. A. Benson, et al. The 10 Meter South Pole Telescope. *Pub. Astron. Soc. Pac.* , 123:568–581, May 2011. doi: 10.1086/659879.

A. D. Challinor, D. J. Mortlock, F. van Leeuwen, A. N. Lasenby, et al. Harmonic analysis of cosmic microwave background data ii. from ring sets to the sky. *MNRAS* , 331(4):994–1010, Apr 2002. ISSN 0035-8711, 1365-2966. doi: 10.1046/j.1365-8711.2002.05255.x.

Anthony Challinor and Gayoung Chon. Error analysis of quadratic power spectrum estimates for cmb polarization: sampling covariance. *Monthly Notices of the Royal Astronomical Society*, 360(2):509–532, Oct 2004. doi: 10.1111/j.1365-2966.2005.09076.x. arXiv: astro-ph/0410097 Citation Key: Challinor2004a.

Anthony Challinor, Pablo Fosalba, Daniel Mortlock, Mark Ashdown, et al. All-sky convolution for polarimetry experiments. *PHYSICAL REVIEW D*, page 8, 2000.

Subrahmanyan Chandrasekhar. *Radiative transfer.* 1960.

Nicolas Chartier and Benjamin D Wandelt. Bayesian control variates for optimal covariance estimation with pairs of simulations and surrogates. *Monthly Notices of the Royal Astronomical Society*, 515(1):1296–1315, Jul 2022. ISSN 0035-8711. doi: 10.1093/mnras/stac1837.

Nicolas Chartier, Benjamin Wandelt, Yashar Akrami, and Francisco Villaescusa-Navarro. Carpool: fast, accurate computation of large-scale structure statistics by pairing costly and cheap cosmological simulations. *Monthly Notices of the Royal Astronomical Society*, 503(2):1897–1914, Feb 2021. ISSN 0035-8711. doi: 10.1093/mnras/stab430.

P. M. Chichura, A. Foster, C. Patel, N. Ossa-Jaen, et al. Asteroid measurements at millimeter wavelengths with the south pole telescope. *The Astrophysical Journal*, 936(2):173, Sep 2022. ISSN 0004-637X, 1538-4357. doi: 10.3847/1538-4357/ac89ec.

Steve K. Choi, Matthew Hasselfield, Shuay-Pwu Patty Ho, Brian Koopman, Marius Lungu, et al. The Atacama Cosmology Telescope: A Measurement of the Cosmic Microwave Background Power Spectra at 98 and 150 GHz. *arXiv e-prints*, art. arXiv:2007.07289, July 2020.

G. Chon, A. Challinor, S. Prunet, E. Hivon, and I.and others Szapudi. Fast estimation of polarization power spectra using correlation functions. *MNRAS* , 350:914–926, May 2004. doi: 10.1111/j.1365-2966.2004.07737.x.

CMB-S4 Collaboration. CMB-S4 Science Case, Reference Design, and Project Plan. *arXiv e-prints*, art. arXiv:1907.04473, Jul 2019.

Bicep/Keck Collaboration, P. A. R. Ade, Z. Ahmed, M. Amiri, D. Barkats, et al. Bicep/keck xv: The bicep3 cosmic microwave background polarimeter and the first three-year data set. *The Astrophysical Journal*, 927 (1):77, Mar 2022. ISSN 0004-637X, 1538-4357. doi: 10.3847/1538-4357/ac4886.

A. T. Crites, J. W. Henning, P. A. R. Ade, K. A. Aird, et al. Measurements of E-Mode Polarization and Temperature-E-Mode Correlation in the Cosmic Microwave Background from 100 Square Degrees of SPTpol Data. *Ap. J.* , 805:36, May 2015. doi: 10.1088/0004-637X/805/1/36.

P. de Bernardis, P. Ade, E. Aquilini, J. Bock, A. Boscaleri, et al. Boomerang : Ballon Observations of Millimetric Extragalactic Radiation and Geophysics. In *Microwave Background Anistropies*, pages 155+, 1997.

T. de Haan, B. A. Benson, L. E. Bleem, S. W. Allen, et al. Cosmological Constraints from Galaxy Clusters in the 2500 Square-degree SPT-SZ Survey. *Ap. J.* , 832:95, November 2016. doi: 10.3847/0004-637X/832/1/95.

Angélica De Oliveira-Costa and Max Tegmark. Cmb multipole measurements in the presence of foregrounds. *Physical Review D*, 74(2):023005, Jul 2006. ISSN 1550-7998, 1550-2368. doi: 10.1103/PhysRevD.74.023005.

G. De Zotti, R. Ricci, D. Mesa, L. Silva, et al. Predictions for high-frequency radio surveys of extragalactic sources. *Astron. Astrophys.* , 431:893–903, March 2005. doi: 10.1051/0004-6361:20042108.

R. H. Dicke, P. J. E. Peebles, P. G. Roll, and D. T.and others Wilkinson. Cosmic Black-Body Radiation. *Ap. J.* , 142:414–419, July 1965. doi: 10.1086/148306.

S. Dodelson and M. D. Schneider. The effect of covariance estimator error on cosmological parameter constraints. *Phys. Rev. D* , 88(6):063537, September 2013. doi: 10.1103/PhysRevD.88.063537.

Scott Dodelson and Fabian Schmidt. *Modern Cosmology.* 2020. doi: 10.1016/C2017-0-01943-2.

M. Doran and G. Robbers. Early dark energy cosmologies. *JCAP*, 6:26, 2006. doi: 10.1088/1475-7516/2006/06/026.

W. B. Doriese, K. M. Morgan, D. A. Bennett, E. V. Denison, et al. Developments in time-division multiplexing of x-ray transition-edge sensors. *Journal of Low Temperature Physics*, 184(1-2):389–395, December 2015. doi: 10.1007/s10909-015-1373-z.

J. Dunkley, E. Komatsu, M. R. Nolta, D. N. Spergel, et al. Five-Year Wilkinson Microwave Anisotropy Probe Observations: Likelihoods and Parameters from the WMAP Data. *Ap. J. Suppl.* , 180:306–329, February 2009. doi: 10.1088/0067-0049/180/2/306.

J. Dunkley, E. Calabrese, J. Sievers, G. E. Addison, et al. The Atacama Cosmology Telescope: likelihood for small-scale CMB data. *JCAP*, 7:025, July 2013. doi: 10.1088/1475-7516/2013/07/025.

R. Dünner, M. Hasselfield, T. A. Marriage, J. Sievers, et al. The Atacama Cosmology Telescope: Data Characterization and Mapmaking. *Ap. J.* , 762:10, January 2013. doi: 10.1088/0004-637X/762/1/10.

D. Dutcher, L. Balkenhol, P. A. R. Ade, Z. Ahmed, et al. Measurements of the E -mode polarization and temperature-E -mode correlation of the CMB from SPT-3G 2018 data. *Phys. Rev. D* , 104(2):022003, July 2021. doi: 10.1103/PhysRevD.104.022003.

G. Efstathiou. Myths and truths concerning estimation of power spectra: the case for a hybrid estimator. *Monthly Notices of the Royal Astronomical Society*, 349(2):603–626, Apr 2004. ISSN 00358711. doi: 10.1111/j.1365-2966.2004.07530.x. arXiv: astro-ph/0307515 Citation Key: Efstathiou2004.

Giulio Fabbian, Julien Carron, Antony Lewis, and Margheritaand others Lembo. Lensed cmb power spectrum biases from masking extragalactic sources. *Physical Review D*, 103(4):043535, Feb 2021. ISSN 2470-0010, 2470-0029. doi: 10.1103/PhysRevD.103.043535.

K. R. Ferguson, A. J. Anderson, N. Whitehorn, P. A. R. Ade, et al. Searching for axionlike time-dependent cosmic birefringence with data from spt-3g. *Physical Review D*, 106(4):042011, Aug 2022. ISSN 2470-0010, 2470-0029. doi: 10.1103/PhysRevD.106.042011.

D. J. Fixsen, E. S. Cheng, J. M. Gales, J. C. Mather, R. A. Shafer, et al. The Cosmic Microwave Background Spectrum from the Full COBE FIRAS Data Set. *Ap. J.* , 473:576, December 1996.

Wendy L. Freedman, Barry F. Madore, Dylan Hatt, Taylor J. Hoyt, In-Sung Jang, et al. The Carnegie-Chicago Hubble Program. VIII. An Independent Determination of the Hubble Constant Based on the Tip of the Red Giant Branch. *arXiv e-prints*, art. arXiv:1907.05922, Jul 2019.

O Friedrich, F Andrade-Oliveira, H Camacho, O Alves, et al. Dark energy survey year 3 results: Covariance modelling and its impact on parameter estimation and quality of fit. *Monthly Notices of the Royal Astronomical Society*, 3165:3125–3165, 2021. ISSN 0035-8711. doi: 10.1093/mnras/stab2384. arXiv: 2012.08568 Citation Key: Friedrich2021.

Nicholas Galitzki, Aamir Ali, Kam S. Arnold, Peter C. Ashton, Jason E. Austermann, et al. The Simons Observatory: instrument overview. In *Proc. SPIE*, volume 10708 of *Proc. SPIE*, page 1070804, Jul 2018. doi: 10.1117/12.2312985.

S. Galli, K. Benabed, F. Bouchet, J.-F. Cardoso, et al. CMB polarization can constrain cosmology better than CMB temperature. *Phys. Rev. D* , 90(6):063504, September 2014. doi: 10.1103/PhysRevD.90.063504.

G. Gamow. The evolution of the universe. *Nature*, 162(4122):680–682, Oct 1948. ISSN 0028-0836, 1476-4687. doi: 10.1038/162680a0.

E. M. George, P. Ade, K. A. Aird, J. E. Austermann, et al. Performance and on-sky optical characterization of the SPTpol instrument. In *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, volume 8452 of *Proc. SPIE*, September 2012. doi: 10.1117/12.925586.

E. M. George, C. L. Reichardt, K. A. Aird, B. A. Benson, et al. A Measurement of Secondary Cosmic Microwave Background Anisotropies from the 2500-Square-degree SPT-SZ Survey. *Ap. J.* , 799:177, February 2015. doi: 10.1088/0004-637X/799/2/177.

J N Goldberg, A J MACFARLANEt, F Rohrlich, and E C Gand others SUDARSHANt. Spin-s spherical harmonics and 5. 1967.

K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, et al. HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. *Ap. J.* , 622:759–771, April 2005. doi: 10.1086/427976.

J. Grain, M. Tristram, and R. Stompor. Polarized cmb power spectrum estimation using the pure pseudo-cross-spectrum approach. *Physical Review D*, 79(12):123515, Jun 2009. ISSN 1550-7998, 1550-2368. doi: 10.1103/PhysRevD.79.123515.

Steven Gratton and Anthony Challinor. Understanding parameter differences between analyses employing nested data subsets. *arXiv e-prints*, art. arXiv:1911.07754, November 2019.

H. P. Gush, M. Halpern, and E. H. Wishnow. Rocket measurement of the cosmic-background-radiation mm-wave spectrum. *Physical Review Letters*, 65(5):537–540, Jul 1990. ISSN 0031-9007. doi: 10.1103/PhysRevLett.65.537.

F. K. Hansen, K. M. Gorski, and E. Hivon. Gabor transforms on the sphere with applications to cmb power spectrum estimation. *Monthly Notices of the Royal Astronomical Society*, 336(4):1304–1328, Nov 2002. ISSN 0035-8711. doi: 10.1046/j.1365-8711.2002.05878.x. Citation Key: Hansen2002.

D. Hanson, S. Hoover, A. Crites, P. A. R. Ade, et al. Detection of B-Mode Polarization in the Cosmic Microwave Background with Data from the South Pole Telescope. *Physical Review Letters*, 111(14):141301, October 2013. doi: 10.1103/PhysRevLett.111.141301.

Luke Hart and Jens Chluba. Updated fundamental constant constraints from planck 2018 data and possible relations to the hubble tension. *Monthly Notices of the Royal Astronomical Society*, 493(3):3255–3263, Apr 2020. ISSN 0035-8711, 1365-2966. doi: 10.1093/mnras/staa412.

J. Hartlap, P. Simon, and P. Schneider. Why your model parameter confidences might be too optimistic. Unbiased estimation of the inverse covariance matrix. *Astron. Astrophys.* , 464:399–404, March 2007. doi: 10.1051/0004-6361:20066170.

J. W. Henning, J. T. Sayre, C. L. Reichardt, P. A. R. Ade, et al. Measurements of the Temperature and E-mode Polarization of the CMB from 500 Square Degrees of SPTpol Data. *Ap. J.* , 852:97, January 2018. doi: 10.3847/1538-4357/aa9ff4.

Catherine Heymans, Tilman Tröster, Marika Asgari, Chris Blake, Hendrik Hildebrandt, et al. KiDS-1000 Cosmology: Multi-probe weak gravitational lensing and spectroscopic galaxy clustering constraints. *arXiv e-prints*, art. arXiv:2007.15632, July 2020.

Eric Hivon, Krzysztof M. Górski, C. Barth Netterfield, Brendan P. Crill, Simon Prunet, and Frodeand others Hansen. MASTER of the Cosmic Microwave Background Anisotropy Power Spectrum: A Fast Method for Statistical Analysis of Large and Complex Cosmic Microwave Background Data Sets. *Ap. J.* , 567(1):2–17, March 2002. doi: 10.1086/338126.

Eric Hivon, Sylvain Mottet, and Nicolas Ponthieu. Quickpol: Fast calculation of effective beam matrices for cmb polarization. *Astron. Astrophys.* , 598:A25, Feb 2017. ISSN 0004-6361, 1432-0746. doi: 10.1051/0004-6361/201629626.

Selim C. Hotinli, Joel Meyers, Cynthia Trendafilova, Daniel Green, et al. The benefits of cmb delensing. *Journal of Cosmology and Astroparticle Physics*, 2022(04):020, Apr 2022. ISSN 1475-7516. doi: 10.1088/1475-7516/2022/04/020.

C. Howlett, A. Lewis, A. Hall, A. Challinor, et al. CMB power spectrum parameter degeneracies in the era of precision cosmology. *JCAP*, 4:027, April 2012. doi: 10.1088/1475-7516/2012/04/027.

W. Hu and S. Dodelson. Cosmic Microwave Background Anisotropies. *Annu.Rev.Astron.Astrophys.*, 40:171–216, 2002.

W. Hu and N. Sugiyama. Anisotropies in the cosmic microwave background: an analytic approach. *Ap. J.* , 444:489–506, 1995. doi: 10.1086/175624.

W Hu and M. White. Acoustic signatures in the cosmic microwave background. *Ap. J.* , 471:30, 1996.

W. Hu and M. White. The damping tail of cosmic microwave background anisotropies. *Ap. J.* , 479:568, 1997.

W. Hu, D. Scott, N. Sugiyama, and M.and others White. Effect of physical assumptions on the calculation of microwave background anisotropies. *Phys. Rev. D* , 52:5498–5515, 1995. doi: 10.1103/PhysRevD.52.5498.

W. Hu, N. Sugiyama, and J. Silk. The physics of microwave background anisotropies. *Nature*, 386:37–43, 1997.

Wayne Hu and Naoshi Sugiyama. Small-Scale Cosmological Perturbations: an Analytic Approach. *Ap. J.* , 471: 542, November 1996. doi: 10.1086/177989.

Wayne Hu and Martin White. A CMB polarization primer. *New Astronomy*, 2(4):323–344, October 1997. doi: 10.1016/S1384-1076(97)00022-5.

John D. Hunter. Matplotlib: A 2D Graphics Environment. *Computing in Science and Engineering*, 9(3):90–95, May 2007. doi: 10.1109/MCSE.2007.55.

D. Huterer, D. Kirkby, R. Bean, A. Connolly, et al. Growth of cosmic structure: Probing dark energy beyond expansion. *Astroparticle Physics*, 63:23–41, March 2015. doi: 10.1016/j.astropartphys.2014.07.004.

Dragan Huterer. Growth of cosmic structure. (arXiv:2212.05003), Dec 2022. arXiv:2212.05003 [astro-ph, physics:gr-qc, physics:hep-th].

Michael A. Janssen, Samuel Gulkis, Charles L. Bennett, and Alan J. Kogutand others. Design and results of differential microwave radiometers on cobe. In *Optics & Photonics*, 1993.

Karsten Jedamzik, Levon Pogosian, and Gong-Bo Zhao. Why reducing the cosmic sound horizon alone can not fully resolve the hubble tension. *Communications Physics*, 4(1):123, Jun 2021. ISSN 2399-3650. doi: 10.1038/s42005-021-00628-x.

David G Johnson, David T Wilkinson, and Joseph Henry Laboratories. A 1% measurement of the temperature of the cosmic microwave radiation at a = 1.2 centimeters. *Ap. J.* .

W. C. Jones, T. E. Montroy, B. P. Crill, C. R. Contaldi, et al. Instrumental and analytic methods for bolometric polarimetry. *Astron. Astrophys.* , 470:771–785, August 2007. doi: 10.1051/0004-6361:20065911.

Marc Kamionkowski, Arthur Kosowsky, and Albert Stebbins. Statistics of cosmic microwave background polarization. *Phys. Rev. D* , 55(12):7368–7388, June 1997. doi: 10.1103/PhysRevD.55.7368.

R. Keisler, C. L. Reichardt, K. A. Aird, B. A. Benson, et al. A Measurement of the Damping Tail of the Cosmic Microwave Background Power Spectrum with the South Pole Telescope. *Ap. J.* , 743:28, December 2011. doi: 10.1088/0004-637X/743/1/28.

R. Keisler, S. Hoover, N. Harrington, J. W. Henning, P. A. R. Ade, et al. Measurements of Sub-degree B-mode Polarization in the Cosmic Microwave Background from 100 Square Degrees of SPTpol Data. *Ap. J.* , 807 (2):151, July 2015. doi: 10.1088/0004-637X/807/2/151.

V. K. Khersonskii, A. N. Moskalev, and D. A. Varshalovich. *Quantum Theory Of Angular Momentum*. World Scientific Publishing Company, 1988. ISBN 978-981-4415-49-1, 978-9971-5-0107-5. doi: 10.1142/0270.

L. Knox. Determination of inflationary observables by cosmic microwave background anisotropy experiments. *Phys. Rev. D* , 52:4307–4318, October 1995.

A. Kogut, C. Lineweaver, G. F. Smoot, C. L. Bennett, et al. Dipole anisotropy in the cobe differential microwave radiometers first-year sky maps. *The Astrophysical Journal*, 419:1, Dec 1993. ISSN 0004-637X, 1538-4357. doi: 10.1086/173453.

J. M. Kovac, E. M. Leitch, C. Pryke, J. E. Carlstrom, et al. Detection of polarization in the cosmic microwave background using DASI. *Nature*, 420:772, December 2002.

Alexandre Lacoste, Alexandra Luccioni, Victor Schmidt, and Thomasand others Dandres. Quantifying the carbon emissions of machine learning. *arXiv preprint arXiv:1910.09700*, 2019.

Guilhem Lavaux and Benjamin D. Wandelt. Fast and optimal cosmic microwave background lensing using statistical interpolation on the sphere. *The Astrophysical Journal Supplement Series*, 191(1):32–42, Nov 2010. ISSN 0067-0049, 1538-4365. doi: 10.1088/0067-0049/191/1/32.

Julien Lesgourgues. The cosmic linear anisotropy solving system (class) i: Overview. (arXiv:1104.2932), May 2011. arXiv:1104.2932 [astro-ph].

A. Lewis, A. Challinor, and A. Lasenby. Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models. *Ap. J.* , 538:473–476, August 2000. doi: 10.1086/309179.

A. Lewis, A. Challinor, and N. Turok. Analysis of CMB polarization on an incomplete sky. *Phys. Rev. D* , 65: 023505–+, January 2002.

T. Louis, E. Grace, M. Hasselfield, M. Lungu, et al. The Atacama Cosmology Telescope: two-season ACTPol spectra and parameters. *JCAP*, 6:031, June 2017. doi: 10.1088/1475-7516/2017/06/031.

Thibaut Louis, Sigurd Naess, Xavier Garrido, and Anthonyand others Challinor. Fast computation of angular power spectra and covariances of high-resolution cosmic microwave background maps using the toeplitz approximation. *Physical Review D*, 102(12):123538, 2020. ISSN 24700029. doi: 10.1103/PhysRevD.102.123538. arXiv: 2010.14344.

M. Lueker, C. L. Reichardt, K. K. Schaffer, O. Zahn, et al. Measurements of Secondary Cosmic Microwave Background Anisotropies with the South Pole Telescope. *Ap. J.* , 719:1045–1066, August 2010. doi: 10.1088/0004-637X/719/2/1045.

Marius Lungu, Emilie R. Storer, Matthew Hasselfield, Adriaan J. Duivenvoorden, et al. The atacama cosmology telescope: measurement and analysis of 1d beams for dr4. *Journal of Cosmology and Astroparticle Physics*, 2022(05):044, May 2022. ISSN 1475-7516. doi: 10.1088/1475-7516/2022/05/044.

E Macaulay, R C Nichol, D Bacon, D Brout, et al. First cosmological results using type ia supernovae from the dark energy survey: measurement of the hubble constant. *Monthly Notices of the Royal Astronomical Society*, 486(2):2184–2196, Jun 2019. ISSN 0035-8711, 1365-2966. doi: 10.1093/mnras/stz978.

A. Manzotti, W. Hu, and A. Benoit-Lévy. Super-sample CMB lensing. *Phys. Rev. D* , 90(2):023003, July 2014. doi: 10.1103/PhysRevD.90.023003.

L. M. Mocanu, T. M. Crawford, K. Aylor, B. A. Benson, L. E. Bleem, et al. Consistency of cosmic microwave background temperature measurements in three frequency bands in the 2500-square-degree SPT-SZ survey. *JCAP*, 2019(7):038, Jul 2019. doi: 10.1088/1475-7516/2019/07/038.

Yukei S. Murakami, Adam G. Riess, Benjamin E. Stahl, W. D'Arcy Kenworthy, et al. Leveraging sn ia spectroscopic similarity to improve the measurement of $h_0$. (arXiv:2306.00070), May 2023. arXiv:2306.00070 [astro-ph].

S. Naess, M. Hasselfield, J. McMahon, M. D. Niemack, et al. The Atacama Cosmology Telescope: CMB polarization at $200 < \mathrm{l} < 9000$. *JCAP*, 10:007, October 2014. doi: 10.1088/1475-7516/2014/10/007.

Jayant V. Narlikar and T. Padmanabhan. Standard cosmology and alternatives: A critical appraisal. *Annual Review of Astronomy and Astrophysics*, 39(1):211–248, Sep 2001. ISSN 0066-4146, 1545-4282. doi: 10.1146/annurev.astro.39.1.211.

Kin-Wang Ng and Guo-Chin Liu. Correlation functions of cmb anisotropy and polarization. 2001.

Andrina Nicola, Carlos García-García, David Alonso, Jo Dunkley, et al. Cosmic shear power spectra in practice. *Journal of Cosmology and Astroparticle Physics*, 2021(3), 2021. ISSN 14757516. doi: 10.1088/1475-7516/2021/03/067. arXiv: 2010.09717 Citation Key: Nicola2021.

Florian Niedermann and Martin S. Sloth. Resolving the hubble tension with new early dark energy. *Physical Review D*, 102(6):063527, Sep 2020. ISSN 2470-0010, 2470-0029. doi: 10.1103/PhysRevD.102.063527.

Yuuki Omori. Agora: Multi-component simulation for cross-survey science. (arXiv:2212.07420), Dec 2022. arXiv:2212.07420 [astro-ph].

Roger O'Brient, Peter Ade, Kam Arnold, Jennifer Edwards, et al. A dual-polarized multichroic antenna-coupled tes bolometer for terrestrial cmb polarimetry. page 77410J, Jul 2010. doi: 10.1117/12.857801.

L. Page, G. Hinshaw, E. Komatsu, M. R. Nolta, et al. Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Polarization Analysis. *Ap. J. Suppl.* , 170:335–376, June 2007. doi: 10.1086/513699.

R. Penrose. Note on the bondi-metzner-sachs group. *Journal of Mathematical Physics*, 7(5):863–870, May 1966. ISSN 0022-2488, 1089-7658. doi: 10.1063/1.1931221.

A. A. Penzias and R. W. Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. *Ap. J.* , 142: 419–421, July 1965.

Will J. Percival and Michael L. Brown. Likelihood methods for the combined analysis of cmb temperature and polarisation power spectra. *Monthly Notices of the Royal Astronomical Society*, 372(3):1104–1116, Apr 2006. doi: 10.1111/j.1365-2966.2006.10910.x. arXiv: astro-ph/0604547 Citation Key: Percival2006.

F. Perrin. Polarization of light scattered by isotropic opalescent media. *Journal of Chemical Physics*, 10: 415–427, 1942.

K. B. Petersen and M. S. Pedersen. The matrix cookbook, nov 2012. Version 20121115.

Planck Collaboration, P. A. R. Ade, N. Aghanim, M. I. R. Alves, et al. Planck 2013 results. I. Overview of products and scientific results. *Astron. Astrophys.* , 571:A1, November 2014. doi: 10.1051/0004-6361/201321529.

Planck Collaboration, R. Adam, P. A. R. Ade, N. Aghanim, et al. Planck 2015 results. I. Overview of products and scientific results. *ArXiv e-prints*, February 2015.

Planck Collaboration, N. Aghanim, M. Arnaud, M. Ashdown, et al. Planck 2015 results. XI. CMB power spectra, likelihoods, and robustness of parameters. *Astron. Astrophys.* , 594:A11, September 2016. doi: 10.1051/0004-6361/201526926.

Planck Collaboration, N. Aghanim, Y. Akrami, F. Arroja, M. Ashdown, et al. Planck 2018 results. I. Overview and the cosmological legacy of Planck. *Astron. Astrophys.* , 641:A1, September 2020a. doi: 10.1051/0004-6361/201833880.

Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, et al. Planck 2018 results. V. CMB power spectra and likelihoods. *Astron. Astrophys.* , 641:A5, September 2020b. doi: 10.1051/0004-6361/201936386.

Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, et al. Planck 2018 results. VI. Cosmological parameters. *Astron. Astrophys.* , 641:A6, September 2020c. doi: 10.1051/0004-6361/201833910.

Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, et al. Planck 2018 results. III. High Frequency Instrument data processing and frequency maps. *Astron. Astrophys.* , 641:A3, September 2020d. doi: 10.1051/0004-6361/201832909.

S. Plaszczynski, A. Lavabre, L. Perotto, and J.-L.and others Starck. A hybrid approach to cosmic microwave background lensing reconstruction from all-sky intensity maps. *Astron. Astrophys.* , 544:A27, Aug 2012. ISSN 0004-6361, 1432-0746. doi: 10.1051/0004-6361/201218899.

POLARBEAR Collaboration, P. A. R. Ade, Y. Akiba, A. E. Anthony, et al. Measurement of the Cosmic Microwave Background Polarization Lensing Power Spectrum with the POLARBEAR experiment. *Physical Review Letters*, (2):021301, July 2014.

G. Polenta, D. Marinucci, A. Balbi, P. de Bernardis, et al. Unbiased estimation of an angular power spectrum. *JCAP*, 11:1–+, November 2005. doi: 10.1088/1475-7516/2005/11/001.

Vivian Poulin, Tristan L. Smith, Tanvi Karwal, and Marcand others Kamionkowski. Early dark energy can resolve the hubble tension. *Physical Review Letters*, 122(22), Jun 2019. ISSN 1079-7114. doi: 10.1103/physrevlett.122.221301.

Fernando Pérez and Brian E. Granger. Ipython: A system for interactive scientific computing. *Computing in Science & Engineering*, 9(3):21–29, 2007. doi: 10.1109/MCSE.2007.53.

S. J. E. Radford. Observing Conditions for Submillimeter Astronomy. In *Revista Mexicana de Astronomia y Astrofisica Conference Series*, volume 41 of *Revista Mexicana de Astronomia y Astrofisica Conference Series*, pages 87–90, November 2011.

Simon J. Radford and Mark A. Holdaway. Atmospheric conditions at a site for submillimeter-wavelength astronomy. page 486–494, Kona, HI, Jul 1998. doi: 10.1117/12.317382.

S. Raghunathan, S. Patil, E. Baxter, B. A. Benson, et al. Mass Calibration of Optically Selected DES Clusters Using a Measurement of CMB-cluster Lensing with SPTpol Data. *Ap. J.* , 872:170, February 2019. doi: 10.3847/1538-4357/ab01ca.

Doogesh Kodi Ramanah, Guilhem Lavaux, and Benjamin D. Wandelt. Wiener filter reloaded: fast signal reconstruction without preconditioning. *MNRAS* , 468(2):1782–1793, Jun 2017. ISSN 0035-8711, 1365-2966. doi: 10.1093/mnras/stx527. arXiv:1702.08852 [astro-ph].

C. L. Reichardt, L. Shaw, O. Zahn, K. A. Aird, et al. A Measurement of Secondary Cosmic Microwave Background Anisotropies with Two Years of South Pole Telescope Observations. *Ap. J.* , 755:70, August 2012. doi: 10.1088/0004-637X/755/1/70.

C. L. Reichardt, S. Patil, P. A. R. Ade, A. J. Anderson, J. E. Austermann, et al. An Improved Measurement of the Secondary Cosmic Microwave Background Anisotropies from the SPT-SZ + SPTpol Surveys. *arXiv e-prints*, art. arXiv:2002.06197, Feb 2020.

Adam G. Riess, Wenlong Yuan, Lucas M. Macri, Dan Scolnic, et al. A Comprehensive Measurement of the Local Value of the Hubble Constant with 1 km $s^{-1}$ $Mpc^{-1}$ Uncertainty from the Hubble Space Telescope and the SH0ES Team. *Ap. J. Lett.* , 934(1):L7, July 2022. doi: 10.3847/2041-8213/ac5c5b.

C. Rosset, V. B. Yurchenko, J. Delabrouille, J. Kaplan, et al. Beam mismatch effects in cosmic microwave background polarization measurements. *Astronomy & Astrophysics*, 464(1):405–415, Mar 2007. ISSN 0004-6361, 1432-0746. doi: 10.1051/0004-6361:20042230.

Katherine C. Roth and David M. Meyer. Cyanogen excitation in diffuse interstellar clouds. *The Astrophysical Journal*, 441:129, Mar 1995. ISSN 0004-637X, 1538-4357. doi: 10.1086/175343.

M. C. Runyan, P. A. R. Ade, R. S. Bhatia, J. J. Bock, et al. ACBAR: The Arcminute Cosmology Bolometer Array Receiver. *Ap. J. Suppl.* , 149:265–287, December 2003.

J. S. Santos, P. Rosati, P. Tozzi, H. Böhringer, et al. Searching for cool core clusters at high redshift. *Astron. Astrophys.* , 483:35–47, May 2008. doi: 10.1051/0004-6361:20078815.

Nathanael Schaeffer. Efficient spherical harmonic transforms aimed at pseudospectral numerical simulations. *Geochemistry, Geophysics, Geosystems*, 14(3):751–758, 2013. doi: 10.1002/ggge.20071.

Nils Schöneberg, Guillermo Franco Abellán, Andrea Pérez Sánchez, Samuel J. Witte, et al. The h 0 olympics: A fair ranking of proposed models. *Physics Reports*, 984:1–55, Oct 2022. ISSN 03701573. doi: 10.1016/j.physrep.2022.07.001.

U. Seljak and C. M. Hirata. Gravitational lensing as a contaminant of the gravity wave signal in the CMB. *Phys. Rev. D* , 69(4):043005–+, February 2004.

U. Seljak and M. Zaldarriaga. Signature of Gravity Waves in the Polarization of the Microwave Background. *Physical Review Letters*, 78:2054–2057, March 1997. doi: 10.1103/PhysRevLett.78.2054.

Elena Sellentin and Jean Luc Starck. Debiasing inference with approximate covariance matrices and other unidentified biases. *Journal of Cosmology and Astroparticle Physics*, 2019(8), 2019. ISSN 14757516. doi: 10.1088/1475-7516/2019/08/021. arXiv: 1902.00709 Citation Key: Sellentin2019.

J. Silk. Cosmic black-body radiation and galaxy formation. *Ap. J.* , 151:459, 1968.

Simons Observatory Collaboration. The Simons Observatory: science goals and forecasts. *JCAP*, 2019(2):056, Feb 2019. doi: 10.1088/1475-7516/2019/02/056.

Simons Observatory Collaboration, Peter Ade, James Aguirre, Zeeshan Ahmed, Simone Aiola, et al. The Simons Observatory: science goals and forecasts. *JCAP*, 2019(2):056, Feb 2019. doi: 10.1088/1475-7516/2019/02/056.

K. M. Smith. Pseudo-Cell estimators which do not mix E and B modes. *Phys. Rev. D* , 74(8):083002–+, October 2006. doi: 10.1103/PhysRevD.74.083002.

G. F. Smoot et al. Structure in the COBE Differential Microwave Radiometer First-Year Maps. *Ap. J.* , 396: L1–L5, 1992.

J. A. Sobrin, A. J. Anderson, A. N. Bender, B. A. Benson, et al. The Design and Integrated Performance of SPT-3G. *Ap. J. Suppl.* , 258(2):42, February 2022. doi: 10.3847/1538-4365/ac374f.

Alessio Spurio Mancini, Davide Piras, Justin Alsing, Benjamin Joachimi, et al. COSMOPOWER: emulating cosmological power spectra for accelerated Bayesian inference from next-generation surveys. *MNRAS* , 511 (2):1771–1788, April 2022. doi: 10.1093/mnras/stac064.

Z. Staniszewski, P. A. R. Ade, K. A. Aird, B. A. Benson, et al. Galaxy Clusters Discovered with a Sunyaev-Zel'dovich Effect Survey. *Ap. J.* , 701:32–41, August 2009. doi: 10.1088/0004-637X/701/1/32.

G. M. Stiehl, H. M. Cho, G. C. Hilton, K. D. Irwin, et al. Time-division squid multiplexers with reduced sensitivity to external magnetic fields. *IEEE Transactions on Applied Superconductivity*, 21(3):298–301, 2011. doi: 10.1109/TASC.2010.2091483.

K. T. Story, C. L. Reichardt, Z. Hou, R. Keisler, et al. A Measurement of the Cosmic Microwave Background Damping Tail from the 2500-Square-Degree SPT-SZ Survey. *Ap. J.* , 779:86, December 2013. doi: 10.1088/0004-637X/779/1/86.

K. T. Story, D. Hanson, P. A. R. Ade, K. A. Aird, et al. A Measurement of the Cosmic Microwave Background Gravitational Lensing Potential from 100 Square Degrees of SPTpol Data. *Ap. J.* , 810:50, September 2015. doi: 10.1088/0004-637X/810/1/50.

N. Sugiyama. Introduction to temperature anisotropies of cosmic microwave background radiation. *Progress of Theoretical and Experimental Physics*, 2014(6):6B101–0, Jun 2014. ISSN 2050-3911. doi: 10.1093/ptep/ptu073.

Kristen M. Surrao, Oliver H. E. Philcox, and J. Colin Hill. Accurate estimation of angular power spectra for maps with correlated masks. *Phys. Rev. D*, 107:083521, Apr 2023. doi: 10.1103/PhysRevD.107.083521.

Dark Energy Survey, Kilo-Degree Survey Collaboration, T. M. C. Abbott, M. Aguena, et al. Des y3 + kids-1000: Consistent cosmology combining cosmic shear surveys. (arXiv:2305.17173), May 2023. arXiv:2305.17173 [astro-ph].

A. Suzuki, K. Arnold, J. Edwards, G. Engargiola, et al. Multi-chroic dual-polarization bolometric focal plane for studies of the cosmic microwave background. *Journal of Low Temperature Physics*, 167:852–858, 2012. ISSN 0022-2291. doi: 10.1007/s10909-012-0602-y.

A. Suzuki, P. A. R. Ade, Y. Akiba, D. Alonso, K. Arnold, et al. The LiteBIRD Satellite Mission: Sub-Kelvin Instrument. *Journal of Low Temperature Physics*, 193(5-6):1048–1056, December 2018. doi: 10.1007/s10909-018-1947-7.

I. Szapudi, S. Prunet, D. Pogosyan, A. S. Szalay, and J. R.and others Bond. Fast Cosmic Microwave Background Analyses via Correlation Functions. *Ap. J. Lett.* , 548:L115–L118, February 2001. doi: 10.1086/319105.

M. Tegmark. How to measure cmb power spectra without losing information. *Phys. Rev. D* , 55:5895–5907, May 1997.

M. Tegmark and A. de Oliveira-Costa. How to measure CMB polarization power spectra without losing information. *Phys. Rev. D* , 64:063001 (15 pages), 2001.

Chi Tian, Siyu Li, and Hao Liu. Accelerating spherical harmonic transforms for a large number of sky maps. *The Astrophysical Journal Supplement Series*, 263(2):21, Dec 2022. ISSN 0067-0049, 1538-4365. doi: 10.3847/1538-4365/ac958b.

M. Tristram, J. F. Macías-Pérez, C. Renault, and D.and others Santos. XSPECT, estimation of the angular power spectrum by computing cross-power spectra with analytical error bars. *MNRAS* , 358:833–842, April 2005. doi: 10.1111/j.1365-2966.2005.08760.x.

Michael S. Turner. The road to precision cosmology. *Annual Review of Nuclear and Particle Science*, 72(1):1–35, Sep 2022. ISSN 0163-8998, 1545-4134. doi: 10.1146/annurev-nucl-111119-041046.

Stéfan van der Walt, S. Chris Colbert, and Gaël Varoquaux. The numpy array: A structure for efficient numerical computation. *Computing in Science & Engineering*, 13(2):22–30, 2011. doi: 10.1109/MCSE.2011.37.

A. van Engelen, R. Keisler, O. Zahn, K. A. Aird, et al. A Measurement of Gravitational Lensing of the Microwave Background Using South Pole Telescope Data. *Ap. J.* , 756:142, September 2012. doi: 10.1088/0004-637X/756/2/142.

S. Vanneste, S. Henrot-Versillé, T. Louis, and M.and others Tristram. Quadratic estimator for cmb cross-correlation. *Physical Review D*, 98(10):103526, Nov 2018. ISSN 2470-0010, 2470-0029. doi: 10.1103/PhysRevD.98.103526.

Licia Verde, Tommaso Treu, and Adam G. Riess. Tensions between the early and late universe. *Nature Astronomy*, 3(10):891–895, Oct 2019. ISSN 2397-3366. doi: 10.1038/s41550-019-0902-0. Citation Key: Verde2019.

J. D. Vieira, D. P. Marrone, S. C. Chapman, C. De Breuck, et al. Dusty starburst galaxies in the early Universe as revealed by gravitational lensing. *Nature*, 495:344–347, March 2013. doi: 10.1038/nature12001.

Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, et al. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020. doi: https://doi.org/10.1038/s41592-019-0686-2.

B. D. Wandelt and F. K. Hansen. Fast, exact CMB power spectrum estimation for a certain class of observational strategies. *Phys. Rev. D* , 67(2):023001–+, January 2003.

Benjamin D. Wandelt, Eric Hivon, and Krzysztof M. Gorski. Cosmic microwave background anisotropy power spectrum statistics for high precision cosmology. *Phys. Rev. D* , 64(8):14, Aug 1998. doi: 10.1103/PhysRevD.64.083003.

M. White. Polarization of the cmb anisotropy. In *Proceedings: Evolution of Large Scale Structure*, 1998.

M. White, D. Scott, and J. Silk. Anisotropies in the cosmic microwave background. *Annu.Rev.Astron.Astrophys.*, 32:319, 1994.

Martin White. Anisotropies in the cmb. (arXiv:astro-ph/9903232), Mar 1999. arXiv:astro-ph/9903232.

N. Wiener. *Extrapolation, Interpolation, and Smoothing of Stationary Time Series: With Engineering Applications.* Massachusetts Institute of Technology: Paperback series. M.I.T. Press, 1964. ISBN 978-0-262-73005-1.

J. H. P. Wu, A. Balbi, J. Borrill, P. G. Ferreira, S. Hanany, et al. Asymmetric Beams in Cosmic Microwave Background Anisotropy Experiments. *Ap. J. Suppl.* , 132:1–17, January 2001. doi: 10.1086/318947.

W. L. K. Wu, L. M. Mocanu, P. A. R. Ade, A. J. Anderson, J. E. Austermann, et al. A Measurement of the Cosmic Microwave Background Lensing Potential and Power Spectrum from 500 deg$^2$ of SPTpol Temperature and Polarization Data. *Ap. J.* , 884(1):70, Oct 2019. doi: 10.3847/1538-4357/ab4186.

J. T. Yu and P. J. E. Peebles. Superclusters of galaxies? *The Astrophysical Journal*, 158:103, Oct 1969. ISSN 0004-637X, 1538-4357. doi: 10.1086/150175.

A. Zacchei, D. Maino, C. Baccigalupi, M. Bersanelli, et al. Planck early results. v. the low frequency instrument data processing. *Astron. Astrophys.* , 536:A5, Dec 2011. ISSN 0004-6361, 1432-0746. doi: 10.1051/0004-6361/201116484.

Matias Zaldarriaga and Uroš Seljak. All-sky analysis of polarization in the microwave background. *Phys. Rev. D* , 55(4):1830–1840, February 1997. doi: 10.1103/PhysRevD.55.1830.