UNIVERSITY OF SCIENCE AND TECHNOLOGY OF HANOI
Department of Space and Application

FINAL REPORT
BACHELOR GRADUATION

By:

NGUYEN Thi Yen Binh

◇

CHARACTERIZATION OF INSTRUMENTAL EFFECTS FOR B-MODE POLARIZATION MEASUREMENTS OF COSMIC MICROWAVE BACKGROUND

◇

Supervisor: Dr. Josquin ERRARD
Laboratoire Astroparticule & Cosmologie (APC), Paris, France

Jury: Assoc. Prof. NGO Duc Thanh[†] (Chairman)
Assoc. Prof. PHAM Ngoc Diep[‡] (Reviewer)
Dr. BUI Van Tuan[†] (Reviewer)
Dr. PHAM Tuan Anh[‡]
Dr. NGUYEN Le Dung[†]
Dr. PHAN Thanh Hien[†]
Dr. NGUYEN Xuan Thanh[†]
Dr. TONG Si Son[†]
[†]*University of Science & Technology of Hanoi (USTH), Hanoi, Vietnam*
[‡]*Vietnam National Space Center (VNSC), Hanoi, Vietnam*

October 17, 2024

Contents

List of Figures

List of Tables

Acknowledgements

The past three months of my internship at APC have been one of the most memorable things in my life. It is my first time doing full-time research in a grand and frontier field like cosmology. I never anticipated the breadth of science I would have the opportunity to think about, with many people I can get to know, in an unfamiliar place that took a nearly 13-hour flight from my hometown.

First of all, I would like to sincerely thank Dr. Josquin Errard, who has advised me and guided me for the last few months from knowing nothing to being able to wrap up this 25-page thesis. He has been a caring and supportive supervisor, always cheering for me when we make little progress, answering all my scientific/technical questions, and discussing with me even if it is late at night or weekend even though he is very busy. I am inspired by his broad and deep involvement in cosmology projects while maintaining his duty to train the young generation. Given a chance, I would love to work with him again in the future.

I am grateful toward Assoc. Prof. Guillaume Patanchon, who sparked my interest in cosmology from his talk and also the one who gave me the opportunity to come to France to work with Dr. Errard.

I have been blessed to join such an amazing team `SciPol` that Dr. Errard has put together. I would like to thank Simon Biquard for helping me with the TOAST simulation, Ema Tsang King Sang for teaching me about HWP, Arianna Rizzieri for suggestions on the minimizer, and other `Scipol` members Magdy Morshed, Benjamin Beringue, Amalia Villarrubia-Aguilar, Wassim Kabalan for supports. Thanks everyone for explaining to me different new concepts like component separation, being my French professors, tolerating my pizza with ketchup, and many more than I can list here. The meeting we have every week has helped me grasp a bigger picture and allowed me to see how much effort everyone has put into `SciPol` project, which makes me greatly appreciate what I am doing. Also, going to lunch with everyone is always my favorite time of the day.

I am happy to know Tran Hoang Viet, my senior at USTH, who taught me the very first lesson in cosmology and my life adviser sometimes. Thanks for listening to me talking about my concerns as well as nonsense babbling literally all the time. I cherish all the conversations we have had, not to mention all the advice and help whenever I need. Now, as you become a `PhD` student, I wish you the best future, always.

Thanks to my Vietnamese friends: Viet, Bao, Huy, Huong who took me to every corner of Paris and had to deal with every trouble that I accidentally caused. I enjoy the time we spend together on the weekends so that I never feel alone in Paris.

Thanks to every staff and professor in APC and USTH, who assisted me with logistics and support throughout my internship.

Last but not least, I would like to express my deepest love for my family. I thank you not just for this internship but for your care and unconditional support throughout my whole life, no matter what I choose to do. Just your presence is more than enough to keep me happy and motivated to follow my passion.

In all sincerity,

Binh

Abstract

The Cosmic Microwave Background (CMB) polarization is a crucial tool for probing the mysteries of cosmic inflation through the measurement of primordial B-modes. Achieving this requires meticulous control over instrumental systematic effects. In this report, we present an instrumental model that focuses on two primary effects: intensity-to-polarization leakage and cross-polarization within a parametric component separation framework. We estimate the upper limits for each parameter to meet the scientific requirements for next-generation CMB measurement. Using Markov Chain Monte Carlo (MCMC) methods to optimize the χ^2 log-likelihood function, we constrain the parameters and estimate the data volume needed to reduce statistical errors within the upper limits of the parameters. We analyze unpolarized atmospheric data with and without the use of a half-wave plate (HWP). Our findings indicate that polarized data is essential for fully characterizing and calibrating all parameters in the HWP model. We further explore the potential of using drone-based fully polarized data for parameter constraints.

Keywords: Cosmic Microwave Background, Polarization, B-modes, Instrumental Systematics, Parametric Component Separation, MCMC.

List of abbreviations

ACT - **A**tacama **C**osmology **T**elescope
BICEP - **B**ackground **I**maging of **C**osmic **E**xtragalactic **P**olarization
CAMB - **C**ode for **A**nisotropies in the Microwave Background **CMB** - **C**osmic Microwave **B**ackground
CP - **C**ross **P**olarization
FIM - **F**isher **I**nformation **M**atrix
FURAX - **F**ramework for a **U**nified and **R**obust **A**lgebra with ja**X**
IP - **I**ntensity-to-**P**olarization
LiteBIRD - **Lite** (Light) satellite for the studies of **B**-mode polarization and **I**nflation from cosmic background **R**adiation **D**etection
MCMC - **M**arkov **C**hain **M**onte **C**arlo
PDF - **P**robability **D**ensity **F**unction
SciPol - **Sci**ence from the large scale Cosmic Microwave Background **Pol**arization structure
TNC - **T**runcated **N**ewton
TOAST - **T**ime **O**rdered **A**strophysics **S**calable **T**ools
TOD - **T**ime **O**rdered **D**ata
WMAP - **W**ilkinson **M**icrowave **A**nisotropy **P**robe

1 Introduction

1.1 Inflationary universe

Over the past 50 years, cosmology has evolved from a data-starved field into a precision science (Turner 2022). The most successful model, named the "hot Big Bang model" by physicist Steven Weinberg, describes the evolution from an early hot soup of hadrons to the current expanding space-time, populated with galaxies.

Modern cosmology's foundation dates back about 100 years when Albert Einstein introduced General Relativity. In the 1920s, solutions for both static (Einstein, de Sitter) and dynamic universes (Friedmann, Lemaitre) were proposed. Hale and Ritchey invented the modern reflecting telescope, enabling Hubble to discover the expanding universe model.

The "hot" aspect of the Big Bang model comes from the discovery of CMB radiation, estimated at 3.5K by Penzias and Wilson in 1964. Earlier, Gamow had proposed a hot beginning in 1948 (Alpher et al. 1948) to explain the origin of chemical elements. A Princeton team, including Dicke and Peebles, interpreted Penzias and Wilson's discovery, suggesting the universe started in a hot, dense state filled with a plasma of particles. As the universe expanded and cooled to ~ 3000 K, neutral atoms formed and photons traveled freely about 380,000 years after the Big Bang (Peebles 1968).

The CMB is a powerful tool for probing the early universe. Measurements of CMB anisotropy and polarization have accurately determined six cosmological parameters but also highlighted issues like the horizon and flatness problems. The horizon problem questions the CMB's homogeneity since only points within a small 2° cone in the sky are causally connected. The flatness problem arises from the universe's extraordinary flatness, with the total density parameter finely tuned to 1 at a 10^{-40} level. To address these issues, Alan Guth proposed the theory of inflation (e.g. Guth 1981).

Inflation suggests a period of exponential expansion of space around 10^{-36} seconds after the Big Bang. This expansion magnified initial quantum fluctuations into perturbations, which evolved into the cosmic structures observed today. Perturbations can be decomposed into scalar, vector, and tensor modes. Due to the universe's expansion, vector perturbations are dissipated and typically ignored in models.

The scalar mode is the main contributor to the CMB spectra as well as to the distribution of galaxy clusters and large-scale structure formation. The primordial power spectrum of scalar perturbation is parameterized in Fourier space with an amplitude A_S and scalar tilt n_S:

$$P_S = A_S \left(\frac{k}{k_0}\right)^{n_S - 1} \tag{1}$$

Tensor mode provides a much smaller amplitude than the previous scalar mode. The primordial power spectrum of tensor perturbation is parameterized similarly as:

$$P_h = A_T \left(\frac{k}{k_0}\right)^{n_T} \tag{2}$$

We can define a new parameter called the tensor-to-scalar ratio $r = \dfrac{A_T}{A_S}$, which is connected to the energy scale of the slow-rolling inflation field (Linde 2017):

$$V_{\text{inflation}} \sim (1.04 \times 10^{16} \text{GeV})^4 \left(\frac{r}{0.01}\right) \tag{3}$$
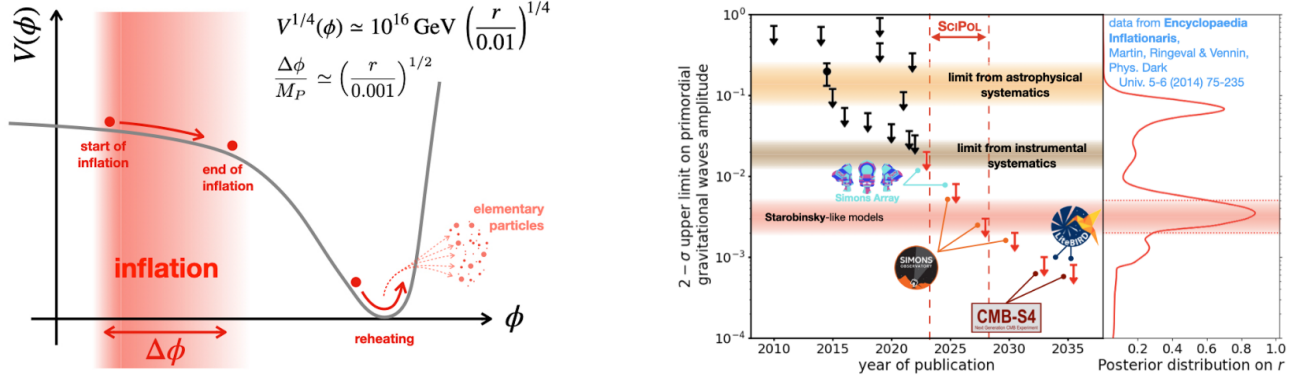
Figure 1: *Left:* Schematic of slow-rolling inflation, where potential energy converts to kinetic energy, leading to acceleration. Inflation ends at a minimum, decaying into particles, and perturbation modes evolve before the CMB last scattering. *Right:* Expected upper limits on the tensor-to-scalar ratio r from forecasting, compared with the posterior distribution of r from single-field inflation models (Martin et al. 2023). *Source:* Josquin Errard.

The parameter r is crucial in inflationary theory, as it is related to both the steepness of the inflation potential and the duration of inflation. Constraining r at the $\mathcal{O}(10^{-3})$ level could reveal new physics of inflation or potentially rule out a vast majority of theoretical models. Currently, the best constraint, $r < 0.032$, has been obtained from a joint analysis of Planck and BICEP/Keck data (Tristram et al. 2022).

When inflation ends, the perturbation modes continue to evolve in the plasma of dark matter, protons, electrons, photons, and neutrinos. Dark matter deepens the potential wells left by density perturbations, while photons and baryons, tightly coupled by Thomson scattering, create pressure that counteracts gravitational collapse. This results in acoustic oscillations between baryons and photons. The evolution of perturbation modes depends on their scale and the cosmological horizon. Sub-horizon modes evolve, while super-horizon modes remain static. This evolution ceases at recombination when photons decouple and form the last scattering surface, known as CMB.

1.2 Cosmic Microwave Background Polarization

The CMB map was first detected by the COBE satellite (Smoot et al. 1992). It represents the best-measured thermal black body radiation in nature (Fixsen 2009), with an average temperature of 2.725 K over all points in the sky. CMB has dipole distortion in temperature due to the motion of the Solar System at about 370 km/s with respect to the CMB. After subtracting the Doppler dipole, the temperature fluctuation is at the scale of 10^{-5} K, proving the fact that the universe is extremely homogeneous and isotropic.

Due to local quadrupole anisotropy during decoupling, the CMB radiation is polarized at the 10% level. The polarization can be divided into two types: rotational-free E-mode and gradient-free B-mode, analogous to electric and magnetic fields. Large-amplitude scalar perturbations give rise to CMB temperature anisotropies and E-modes. Small-amplitude primordial gravitational waves are expected to generate both E-modes and B-modes. Because only primordial gravitational waves can produce B-modes, it becomes the probe for cosmic inflation. The E-mode polarization

comes at a tiny fluctuation scale of 10^{-6} K, and B-mode polarization is expected to be $\sim 10-100$ times smaller.
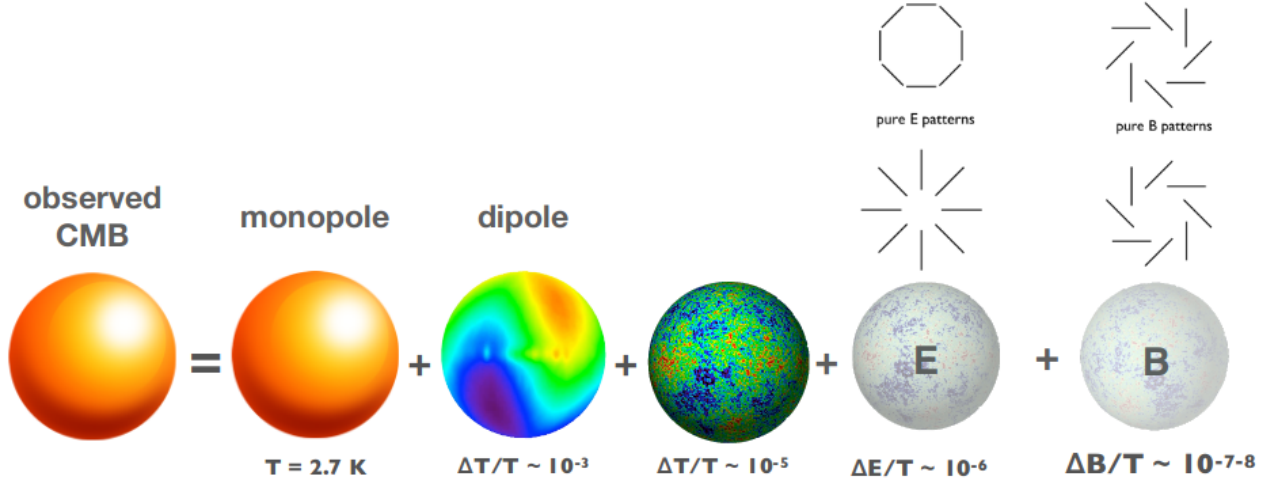
Figure 2: Illustration of CMB at different scales, from monopole to polarized anisotropies, decomposed into E-mode and B-mode. *Source:* Josquin Errard.

The observed quantities that we use to measure CMB are so-called Stokes parameters $\begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix}$, where I encodes the temperature, Q and U encode the linear polarization, and V encodes the circular polarization. Q includes the horizontal and vertical directions of polarization, and U includes the 45° direction of polarization. V is expected to vanish as Thomson scattering does not generate this type of polarization.

The statistical properties of the CMB map, which show nearly Gaussian fluctuations, are predominantly described by the 2-point correlation functions. Assuming isotropy in the underlying probability distribution, correlations between anisotropies at different points in the sky depend only on the angle between them. To analyze these correlations, temperature and linear polarization are typically expanded in terms of spherical harmonics:

$$\Delta T(\hat{n}) = \sum_{l,m} a_{lm}^T Y_{lm}(\hat{n}) \tag{4}$$

$$Q(\hat{n}) + iU(\hat{n}) = - \sum_{l,m} (a_{lm}^E + i a_{lm}^B) \,_2 Y_{lm}(\hat{n}) \tag{5}$$

The index ℓ represents the degree, corresponding to the angular separation $\theta \sim \pi\ell$ in the map, while index m represents the order linked to orientation. The intensity map is expanded using standard spherical harmonics Y_{lm}, whereas the polarization map is expanded using spin-weighted spherical harmonics $\pm_2 Y_{lm}$.

We can form the angular power spectra from the expansion coefficients a_{lm}^T, a_{lm}^E, and a_{lm}^B as:

$$C_l^{XY} = \frac{1}{2l+1} \sum_m a_{lm}^X a_{lm}^{Y*} \tag{6}$$

where X and Y are either T, E, or B.

3
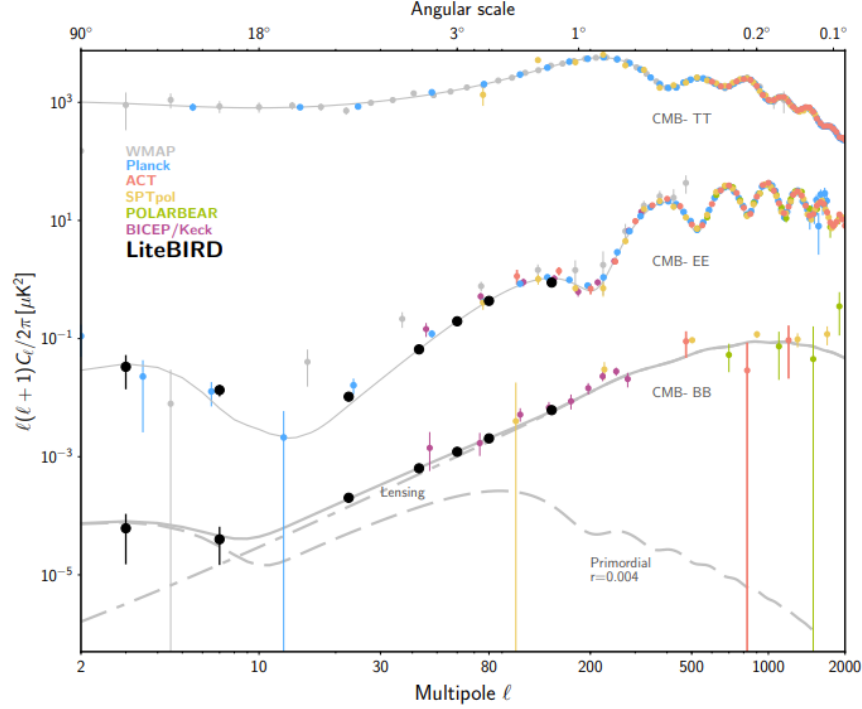
Figure 3: *Solid line:* CMB power spectra of temperature anisotropy (top), E-mode polarization (middle), and B-mode polarization (bottom) with present measurements (colored points). *Dashed line:* primordial B-mode spectrum from scale-invariant tensor perturbation with $r = 0.004$ with expected measurement of LiteBIRD (black points). *Source:* The LiteBIRD Collaboration (2022).

The best results of CMB anisotropies and E-modes including C_l^{TT}, C_l^{TE}, and C_l^{EE} come from WMAP (Bennett et al. 2013) and The Planck Collaboration (2020), covering a huge multipole l range. Lensing B-mode was discovered through cross-correlation by South Pole Telescope (Samuel Reich 2013) with the help of Herschel Space Observatory, then directly measured by POLARBEAR (P. A. R. Ade et al. 2014), ACT and BICEP/Keck (Ade et al. 2015) experiments. The primordial C_l^{BB} remains the target for the next generation of CMB experiments since it would be direct evidence for inflationary gravitational waves.

1.3 Challenge of instrumental effects for the inflationary B-modes

Upcoming projects such as ground-based SO, CMB-S4, and spaceborne LiteBIRD, with the expectation of reaching $r \sim \mathcal{O}(10^{-3})$, require us to accurately distinguish the CMB blackbody from other sources: astrophysical sources (polarized dust and synchrotron), atmosphere for ground-based telescopes, and instrumental systematic effects.

The main framework that has been proven to be successful in precedent experiments is the component separation method, including two main approaches: blind and parametric. The blind method imposes conditions on spectra and separates the components based on their statistical properties; some examples include `SMICA`, `NILC`, and `SEVEM`. The parametric method introduces a model based on the physical nature of the phenomena, such as emission laws of foregrounds or multiplication of the gains of instruments. Notable examples are `FGBuster` and `Commander`.

With the deployment of more detectors in new CMB instruments, sensitivity, and data volume
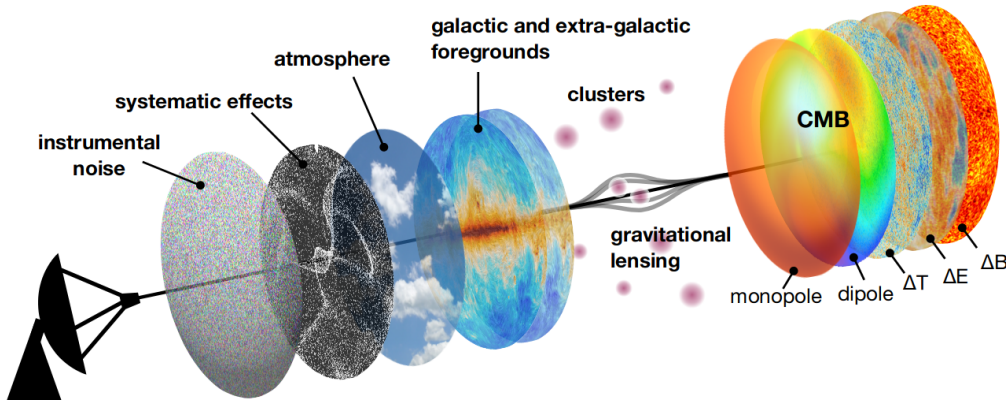
Figure 4: Illustration of various contaminants affecting the CMB signal before it reaches the detectors. These contaminants include astrophysical sources, atmospheric interference, instrumental systematic effects and noise. *Source:* Josquin Errard.

have increased significantly, leading to greater instrumental complexity. Consequently, there are challenges in characterizing and calibrating systematic effects that do not average out over time and contaminate measurements. Numerous instrumental effects must be accounted for in the calibration process, including the optical reflection on the cover of a telescope, the beam's far side lobes (Leloup et al. 2023), gains, bandpasses (Ghigna et al. 2020), polarization angles (The Simons Observatory 2021), and electronic cross-talk (Fabbian & Peloton 2021). These well-known hardware characteristics can broadly lead to two main effects on the Stokes parameters: Intensity-to-Polarization leakage (I-to-P or $T \to E, B$), and Cross-Polarization ($E \leftrightarrow B$). The characterization of these two effects under the parametric component separation method will be the primary focus of this study. The main goal is to estimate our ability to characterize the systematic effects using the atmospheric data through simulation. This research is part of the global effort undertaken by the Scipol[1] project.

The report is structured as follows. We introduce the theoretical framework for modeling instrumental effects in section 2. Our data, procedure, and mean of assessment are presented in section 3. Then, section 4 describes our results. After that, we evaluate our study and explain our future works in section 5. Finally, we summarize our findings in section 6.

---

[1]https://scipol.in2p3.fr/

2 Theoretical Framework

2.1 Jones and Mueller Formalism

Jones and Stokes/Mueller are the two most commonly used treatments for polarized signals. Jones formalism models the effects of instruments with matrix operations on electromagnetic field components or the so-called Jones vector $\begin{bmatrix} E_x & E_y \end{bmatrix}^\top$.

$$\mathbf{E}_f = \mathbf{J}\mathbf{E}_i \tag{7}$$

where Jones matrix $\mathbf{J}$ which can be a general product of matrices describing different effects like optical elements on the initial field component.

$$\begin{bmatrix} E_x \\ E_y \end{bmatrix}_f = \prod_k \begin{bmatrix} J_{xx} & J_{xy} \\ J_{yx} & J_{yy} \end{bmatrix}_k \begin{bmatrix} E_x \\ E_y \end{bmatrix}_i \tag{8}$$

Similarly, Mueller formalism describes the instrument effects on the observable quantity Stokes parameters.

$$\mathbf{S}_f = \mathbf{M}\mathbf{S}_i \tag{9}$$

where the Mueller matrix $\mathbf{M}$, is the products of individual components on Stokes parameters.

$$\begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix}_f = \prod_k \begin{bmatrix} M_{II} & M_{IQ} & M_{IU} & M_{IV} \\ M_{QI} & M_{QQ} & M_{QU} & M_{QV} \\ M_{UI} & M_{UQ} & M_{UU} & M_{UV} \\ M_{VI} & M_{VQ} & M_{VU} & M_{VV} \end{bmatrix}_k \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix}_i \tag{10}$$

The transformation from the Jones matrix $\mathbf{J}$ to the Mueller matrix $\mathbf{M}$ can be referred to Hunte (2008):

$$\mathbf{M} = \mathbf{U}(\mathbf{J} \otimes \mathbf{J}^*)\mathbf{U}^{-1} \tag{11}$$

where $\otimes$ is the Kronecker product operator, $\mathbf{U} = \dfrac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -i & i & 0 \end{bmatrix}$.

The optical elements can be described under Mueller formalism as:

Polarizer: changing the orthogonal amplitudes unequally. For example, a linear polarizer can be described by a Jones matrix:

$$\mathbf{J} = \begin{bmatrix} g_1 & 0 \\ 0 & g_2 \end{bmatrix} \longrightarrow \mathbf{M}_{\text{polarizer}} = \begin{bmatrix} \dfrac{g_1^2 + g_2^2}{2} & \dfrac{g_1^2 - g_2^2}{2} & 0 & 0 \\ \dfrac{g_1^2 - g_2^2}{2} & \dfrac{g_1^2 + g_2^2}{2} & 0 & 0 \\ 0 & 0 & g_1 g_2 & 0 \\ 0 & 0 & 0 & g_1 g_2 \end{bmatrix} \tag{12}$$

Wave plate (Retarder): introducing a phase shift between orthogonal components. For example, when the x component gets a $\delta/2$ phase shift (fast axis), and the y component gets a $-\delta/2$ phase

shift (slow axis). The phase shift between x and y is δ.

$$\mathbf{J} = \begin{bmatrix} e^{i\delta/2} & 0 \\ 0 & e^{-i\delta/2} \end{bmatrix} \longrightarrow \mathbf{M}_{\text{waveplate}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\delta & \sin\delta \\ 0 & 0 & -\sin\delta & \cos\delta \end{bmatrix} \tag{13}$$

Rotator: rotating the orthogonal components by an angle θ.

$$\mathbf{J} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \longrightarrow \mathbf{R}(2\theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(2\theta) & \sin(2\theta) & 0 \\ 0 & -\sin(2\theta) & \cos(2\theta) & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{14}$$

We can describe the rotation of optical elements by (1) transforming to a basis where the Jones matrix is diagonal to describe the Jones vector in a coordinate system whose axes align with the axes of the optical elements, (2) applying the Jones matrix, and (3) transforming back to the original basis. This is similar for Mueller matrices:

$$\mathbf{M}(2\theta) = \mathbf{R}(2\theta)^{\top}\mathbf{M}\mathbf{R}(2\theta) \tag{15}$$

2.2 Mueller matrices for Instrumental Effects

2.2.1 Intensity-to-Polarization

The leakage from intensity I to polarization $P \equiv \sqrt{Q^2 + U^2 + V^2}$ can be described by the reduction in the x and y components through amplitudes g_1 and g_2, which range from 0 to 1. Additionally, a phase shift $\phi \in [0, 2\pi)$ is included to represent any mismatch between the two field components. This effect is described by the Jones matrix:

$$\mathbf{J}_{\text{IP}}(g_1, g_2, \phi) = \begin{bmatrix} g_1 & 0 \\ 0 & g_2 e^{i\phi} \end{bmatrix} \tag{16}$$

Using the formalism introduced in Section 2.1, we can derive the Mueller matrix representing the IP effect on the Stokes parameters as:

$$\mathbf{M}_{\text{IP}}(g_1, g_2, \phi) = \begin{bmatrix} \dfrac{g_1^2 + g_2^2}{2} & \dfrac{g_1^2 - g_2^2}{2} & 0 & 0 \\ \dfrac{g_1^2 - g_2^2}{2} & \dfrac{g_1^2 + g_2^2}{2} & 0 & 0 \\ 0 & 0 & g_1 g_2 \cos\phi & g_1 g_2 \sin\phi \\ 0 & 0 & -g_1 g_2 \sin\phi & g_1 g_2 \cos\phi \end{bmatrix} \tag{17}$$

In an ideal instrument with no intensity-to-polarization leakage, we expect $g_1 = 1$, $g_2 = 1$, and $\phi = 0$.

2.2.2 Cross-Polarization

Cross-polarization is the mixing between two field components, primarily affecting the off-diagonal terms in the Jones matrix with an amplitude of $\sqrt{\epsilon} \in (0, 1]$ and a phase shift $\psi \in [0, 2\pi)$.

The diagonal terms are chosen to ensure the conservation of intensity, making the matrix unitary. This effect is described by the following Jones matrix:

$$\mathbf{J}_{\mathrm{CP}}(\epsilon, \psi) = \begin{bmatrix} \sqrt{1-\epsilon} & -\sqrt{\epsilon}e^{i\psi} \\ \sqrt{\epsilon}e^{-i\psi} & \sqrt{1-\epsilon} \end{bmatrix} \tag{18}$$

Similar to the case of IP, we derive the Mueller matrix for CP as:

$$\mathbf{M}_{\mathrm{CP}}(\epsilon, \psi) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1-2\epsilon & -2\sqrt{\epsilon(1-\epsilon)}\cos\psi & -2\sqrt{\epsilon(1-\epsilon)}\sin\psi \\ 0 & 2\sqrt{\epsilon(1-\epsilon)}\cos\psi & 1-\epsilon(1+\cos(2\psi)) & -\epsilon\sin(2\psi) \\ 0 & 2\sqrt{\epsilon(1-\epsilon)}\sin\psi & -\epsilon\sin(2\psi) & 1-\epsilon(1-\cos(2\psi)) \end{bmatrix} \tag{19}$$

In an ideal instrument without cross-polarization, we expect that $\epsilon = 0$.

2.2.3   Combined Instrumental Effects IPCP

We believe that intensity-to-polarization leakage primarily occurs due to oblique light reflection on the surface of the instrument's cover. Therefore, it will be the first effect in our model, followed by cross-polarization. The combined instrumental effects can be expressed as:

$$\mathbf{M}_{\mathrm{IPCP}}(g_1, g_2, \epsilon, \phi, \psi) = \mathbf{M}_{\mathrm{CP}}(\epsilon, \psi)\mathbf{M}_{\mathrm{IP}}(g_1, g_2, \phi) \tag{20}$$

2.3   Instrumental model

In this research, we consider an instrumental model composed of a detector connected to an antenna/grid, a half-wave plate (HWP), and an imaginary layer of instrumental effects.
The grid separates the incident light into two orthogonal polarization states and we can rotate the grid by an angle θ with rotation matrix $\mathbf{R}(2\theta)$:

$$\mathbf{M}_{\mathrm{rotating\ grid}} = \mathbf{R}^{\top}(2\theta)\mathbf{M}_{\mathrm{grid}}\mathbf{R}(2\theta) \tag{21}$$

where

$$\mathbf{M}_{\mathrm{grid}} = \frac{1}{2}\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

A half-wave plate is a retarder containing birefringent materials that introduce different reflective indices for field components leading to π phase shift difference between input and output polarization states. A HWP continuously rotates by the angle $\varphi_t = 2\pi f_{\mathrm{HWP}}t$ can be modeled as:

$$\mathbf{M}_{\mathrm{rotating\ HWP}} = \mathbf{R}^{\top}(2\varphi_t)\mathbf{M}_{\mathrm{HWP}}\mathbf{R}(2\varphi_t) \tag{22}$$

where

$$\mathbf{M}_{\mathrm{HWP}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

The complete instrumental system can be written as:

$$\mathbf{M}_{\mathrm{inst}} = \mathbf{M}_{\mathrm{rotating\ grid}}\mathbf{M}_{\mathrm{rotating\ HWP}}\mathbf{M}_{\mathrm{IPCP}}(g_1, g_2, \epsilon, \phi, \psi) \tag{23}$$

8
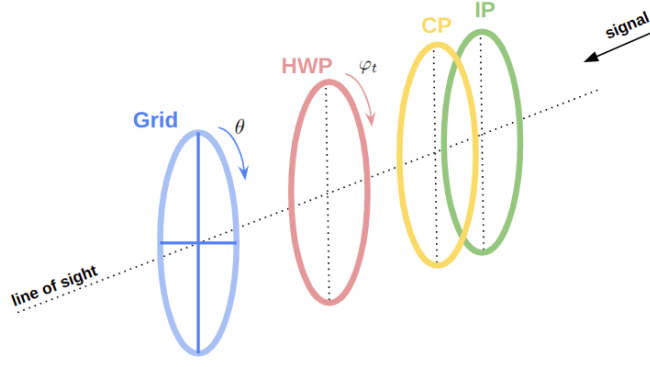
Figure 5: Illustration of the instrumental model (Equation 23).

2.4 Parametric maximum likelihood component separation

We introduce the parametric component separation method (Eriksen et al. 2006) for a linear data model of instrumental effects as follows:

$$\mathbf{d} = \mathbf{M}_{\text{inst}}\mathbf{s} + \mathbf{n} \tag{24}$$

- $\mathbf{d}$ is a data vector.

- $\mathbf{M}_{\text{inst}}$ is a mixing matrix representing the instrumental effects on the input sky signal.

- $\mathbf{s}$ is a sky signal vector, which contains signals from CMB, foreground, or atmosphere.

- $\mathbf{n}$ is a noise vector assumed to be Gaussian and uncorrelated with dispersion $\mathbf{N} = \langle \mathbf{n}\mathbf{n}^{\top} \rangle$.

2.4.1 Time Ordered Data

Because detectors are total-power sensitive, we can only retain the first element I of vector $\mathbf{d}$, we can rewrite the formula as:

$$d_t = \frac{1}{2}\underbrace{\begin{bmatrix} 1 & \cos(2\theta) & \sin(2\theta) & 0 \end{bmatrix}}_{\text{rotating grid}} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(4\varphi_t) & \sin(4\varphi_t) & 0 \\ 0 & \sin(4\varphi_t) & -\cos(4\varphi_t) & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}}_{\text{rotating HWP}} \underbrace{\begin{bmatrix} M_{II} & M_{IQ} & M_{IU} & M_{IV} \\ M_{QI} & M_{QQ} & M_{QU} & M_{QV} \\ M_{UI} & M_{UQ} & M_{UU} & M_{UV} \\ M_{VI} & M_{VQ} & M_{VU} & M_{VV} \end{bmatrix}}_{\text{IPCP}} \underbrace{\begin{bmatrix} I_t \\ Q_t \\ U_t \\ V_t \end{bmatrix}}_{\text{input}} + \underbrace{n_t}_{\text{noise}}$$

$$\tag{25}$$

We introduce the matrix $\mathcal{A} = \begin{bmatrix} A_1 & A_2 & A_3 & A_4 \end{bmatrix}$ so that:

$$d_t \equiv \frac{1}{2}(A_1 I_t + A_2 Q_t + A_3 U_t + A_4 V_t) + n_t \tag{26}$$

where

$$A_1 \equiv M_{II} + M_{QI}\cos(4\varphi_t - 2\theta) + M_{UI}\sin(4\varphi_t - 2\theta) \tag{27}$$
$$A_2 \equiv M_{IQ} + M_{QQ}\cos(4\varphi_t - 2\theta) + M_{UQ}\sin(4\varphi_t - 2\theta) \tag{28}$$
$$A_3 \equiv M_{IU} + M_{QU}\cos(4\varphi_t - 2\theta) + M_{UU}\sin(4\varphi_t - 2\theta) \tag{29}$$
$$A_4 \equiv `M_{IV} + M_{QV}\cos(4\varphi_t - 2\theta) + M_{UV}\sin(4\varphi_t - 2\theta) \tag{30}$$

9

2.4.2 Maximum Likelihood Solution

Applying the model from section 2.4.1, where we have a series of measurements with time t, where we can absorb the factor $1/2$ into $\mathbf{s}_t = \frac{1}{2}\begin{bmatrix} I_t & Q_t & U_t & V_t \end{bmatrix}^\top$:

$$d_t = \mathcal{A}(\xi, t, \theta)\mathbf{s}_t + n_t \tag{31}$$

where $\xi \in \{g_1, g_2, \epsilon, \phi, \psi\}$ is a set of instrumental parameters, n_t is the noise that follows a Gaussian distribution $\mathcal{N}(0, \sigma^2)$.

Following the formalism introduced in Appendix B, we can simplify our likelihood into a χ^2 log-likelihood function as the sum over all the likelihood of time and polarization angle. Optimizing the following log-likelihood function leads to an estimation of the expected parameters $\hat{\xi}$:

$$-2\log\mathcal{L}(\hat{\xi}|\mathbf{d}) = \sum_{t,\theta} \frac{(d_t - \mathcal{A}\mathbf{s}_t)^2}{\sigma^2} \tag{32}$$

2.5 The second approach to the problem: A model without HWP

In another way, assuming that the instrumental effects are the same in all the detectors defined by polarization angle θ, we can do the calibration using the data after HWP demodulation (e.g. Rashid et al. 2023).
The ideal data we get from the rotation grid after modulation should be:

$$\mathbf{s}' = \mathbf{R}^\top(2\theta)\mathbf{M}_{\text{grid}}\mathbf{R}(2\theta)\mathbf{s} \tag{33}$$

Then, the linear data model for component separation can be written as:

$$\mathbf{d} = \mathbf{M}_{\text{IPCP}}(g_1, g_2, \epsilon, \phi, \psi)\mathbf{s}' + \mathbf{n} \tag{34}$$

Following the framework of Appendix B, we formalize the log-likelihood as:

$$-2\log\mathcal{L} = (\mathbf{d} - \mathbf{M}_{\text{IPCP}}\mathbf{s}')^\top\mathbf{N}^{-1}(\mathbf{d} - \mathbf{M}_{\text{IPCP}}\mathbf{s}') \tag{35}$$

where $\mathbf{N} \equiv \langle\mathbf{nn}^\top\rangle$ is a Gaussian covariance.

3 Data and Methodology

3.1 Overall procedure

The overall procedure for our analysis is as follows:
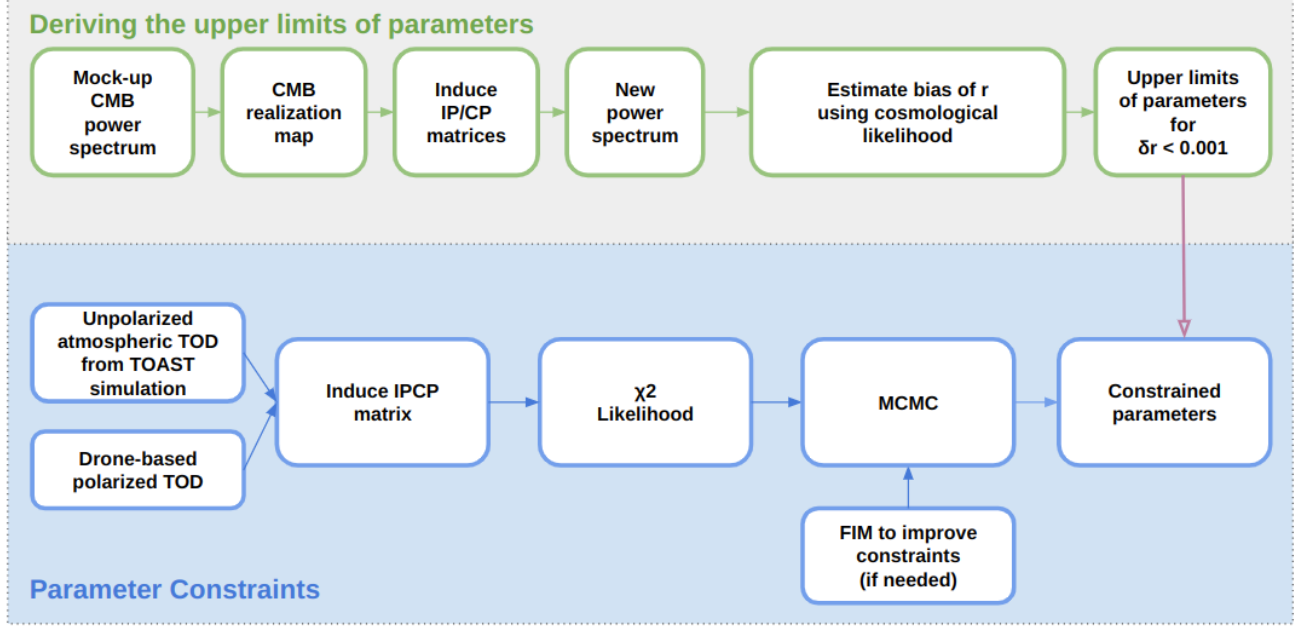
Figure 6: Block diagram summarizing the main analysis steps

3.2 Data

3.2.1 Mock-up CMB power spectra from CAMB

We choose to use CMB realization maps, which have the same statistical properties as the CMB, for our research. This involves generating the angular power spectra as the first step, which can be achieved in two ways: utilizing real data measured by Planck[2] or generating them theoretically using `CAMB`[3].

In this work, we use `CAMB`, a cosmology code for calculating CMB, lensing, galaxy count, dark-age 21cm power spectra, matter power spectra, and transfer functions. `CAMB` allows for precise theoretical calculations of the expected power spectra based on a set of cosmological parameters. The result from `CAMB` is shown in Figure 7 and 8.

With package `HEALpy`, we use the function `synalm` to generate a realization of the spherical harmonics coefficients a_{lm} from CAMB angular spectra and transform them to a map with the function `alm2map` with N_{side} = 1024. We can double-check with the published Planck map that the temperature has a $\mathcal{O}(10^{-4})$ magnitude and the polarization on a $\mathcal{O}(10^{-5})$ magnitude. From the map, we can use the function `anafast` to perform spherical harmonic analysis up to the maximum order ℓ_{max} and compare it with the input power spectra.

---

[2]https://pla.esac.esa.int/
[3]https://camb.readthedocs.io/en/latest/

(a) Temperature   **(b)** E-mode polarization   **(c)** B-mode polarization
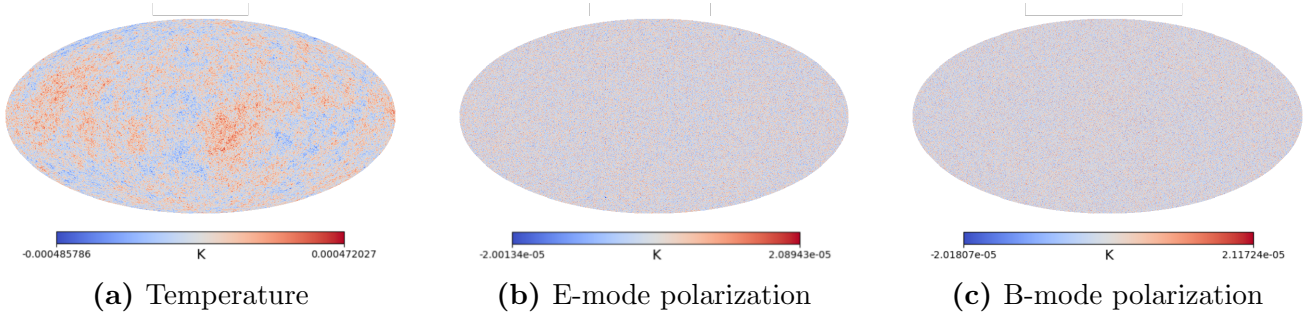
Figure 7: CMB realization maps
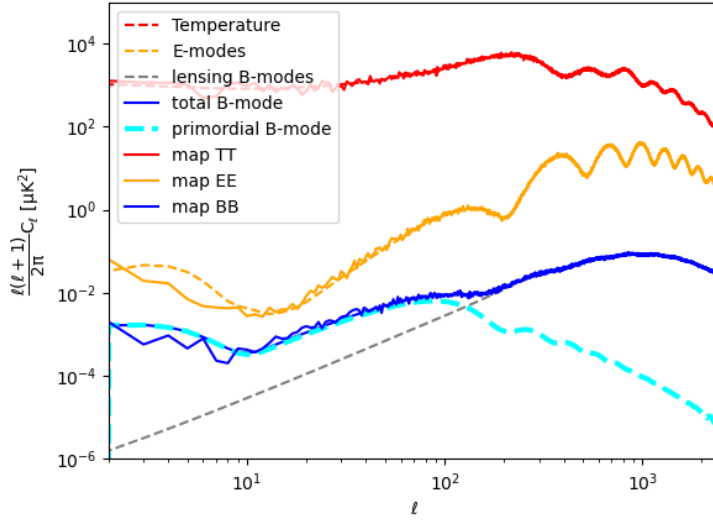
Figure 8: *Dashed lines:* CAMB angular power spectra for temperature anisotropy, E-mode, and B-mode (total, lensing, and primordial with r=0.1). *Solid lines:* angular power spectra from spherical harmonic analysis of CMB realization maps (Figure 7) shows a small level of cosmic variance/fluctuation.

3.2.2 Unpolarized atmospheric data from TOAST

`TOAST`[4] is a framework to simulate the TOD that can be collected by telescopes. In our study, we use a simulation from a fictitious Small Aperture Telescopes (SAT) located at the South Pole, doing approximately 1h24' scanning with the sample rate f_{sample} = 20 Hz. The telescope remains at the same azimuth and elevation most of the time, so the temperature is pretty much constant.

For gain calibration purposes, we also add one el-node (elevation node), a tiny moving up and down of 2° in elevation of the telescope, so that we have a small fluctuation of about 300 mK in our data at the beginning. The scanning trajectory of the telescope and the measured atmospheric TOD are illustrated in Figure 9.

---

[4] https://toast-cmb.readthedocs.io/en/latest/intro.html
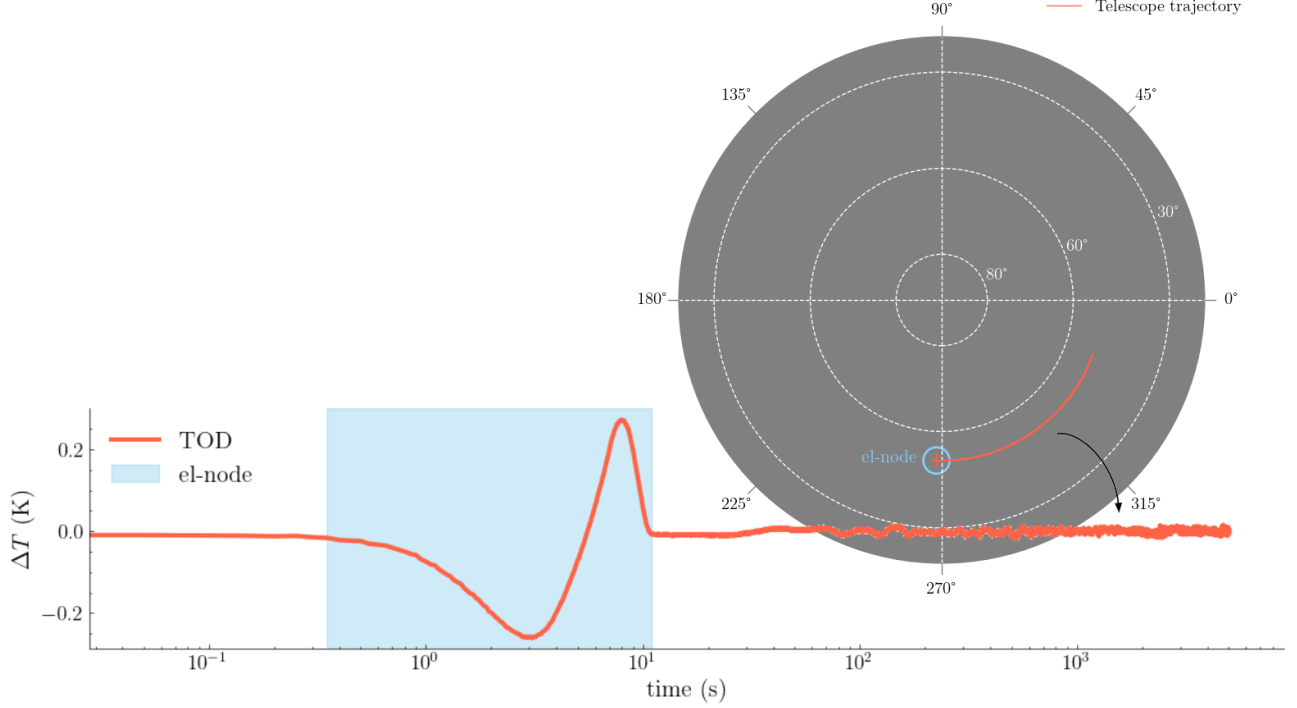
Figure 9: The trajectory of a simulated telescope shown in a polar plot with azimuth and elevation (co-eccentric circles) and atmospheric TOD.

3.3 Cosmological maximum likelihood

Assuming the scientific goal of $\delta r < 0.001$ for the next generation of CMB experiments, this likelihood is crucial to assess the limits of the errors of parameters and the volume of data we need to use. The instrumental systematic effects can introduce residuals in the CMB maps. To estimate the bias δr, we maximize the cosmological likelihood as a function of r:

$$-2\log \mathcal{L}_{\text{cosmo}} = f_{\text{sky}} \sum_{\ell} (2\ell + 1) \left(\log C_{\ell}^{\text{th}}(r) + \frac{C_{\ell}^{\text{obs}}}{C_{\ell}(r)^{\text{th}}} \right) \tag{36}$$

where f_{sky} is the observed fraction of sky ($f_{\text{sky}} = 1$ for full sky), C_{ℓ}^{obs} and C_{ℓ}^{th} are the observed and theoretical CMB B-mode power spectrum respectively, defined as:

$$C_{\ell}^{\text{obs}} = C_{\ell}^{\text{th}} + C_{\ell}^{\text{res}} \quad \text{and} \quad C_{\ell}^{\text{th}} = r C_{\ell}^{\text{prim}} + C_{\ell}^{\text{lens}} \tag{37}$$

Here, C_{ℓ}^{res} is the residual B-mode power spectrum induced by instrumental systematics, C_{ℓ}^{prim} is the B-modes from primordial gravitational waves and C_{ℓ}^{lens} is the power spectrum from lensing.

3.4 Model of unpolarized atmospheric data

One of our goals is to assess the ability to constrain the instrumental parameters using only unpolarized atmospheric data. We try to answer the question of how many parameters could we possibly fit and the volume of data that we need to reach the limit for each parameter.

3.4.1 Model

Following Appendix B, we present the linear parametric model for both effects using $\mathbf{M}_{\text{IPCP}}$ induced on $\mathbf{s}_{\text{atm}} = \begin{bmatrix} I_t & 0 & 0 & 0 \end{bmatrix}^{\top}$:

$$d_t = A_1(\xi, t, \theta)\frac{I_t}{2} + n_t \tag{38}$$

where $\xi = \{g_1, g_2, \epsilon, \phi, \psi\}$.
We can write down the full form of the expression as:

$$d_t = \left(\frac{g_1^2 + g_2^2}{2} + \frac{g_1^2 - g_2^2}{2}(1 - 2\epsilon)\cos(4\psi_t - 2\theta) + (g_1^2 - g_2^2)\sqrt{\epsilon(1-\epsilon)}\cos\psi\sin(4\psi_t - 2\theta) \right)\frac{I_t}{2} + n_t \tag{39}$$

Note that in this case, we lose the information on parameter ϕ as it does not exist in the expression. We will constrain the set of parameters that maximize the likelihood function using the MCMC method introduced in section 3.5 by adopting the framework in section 2.4.2.

3.4.2 Fisher information matrix

We introduce the FIM $\mathbf{F}$ as described by Tegmark et al. (1997) for degeneracy breaking. FIM elements can be defined as:

$$F_{ij} = \left\langle \frac{\partial^2(-2\log\mathcal{L})}{\partial\xi_i\partial\xi_j} \right\rangle \tag{40}$$

where $\langle . \rangle$ is the operator that estimates the function at the true values.

FIM is a Hessian matrix that measures the curvature of the likelihood function at its peak, corresponding to the true parameter values if the model is correct. It is intrinsically linked to the covariance matrix representing the uncertainty due to instrumental noise.

$$\Sigma = \mathbf{F}^{-1} \tag{41}$$

We build an analytical model of FIM to deal with the nearly degenerate directions in parameter space for the unpolarized atmospheric case in Appendix C.

3.5 Markov Chain Monte Carlo

In our work, the main method is MCMC, which can be used to sample from a continuous random variable with a PDF. By generating samples from this distribution, MCMC enables the evaluation of statistical properties such as the expected value and variance of the random variable. This is particularly useful in cases where direct sampling is difficult/impractical.

The implementation is carried out by `emcee`[5] (Foreman-Mackey et al. 2013), a Python library that implements the affine-invariant ensemble sampler for MCMC (Goodman & Weare 2010). The affine-invariant approach offers several advantages over traditional MCMC methods, including improved convergence properties and the ability to handle complicated posterior distributions more efficiently.

The primary goal of using `emcee` in our analysis is to maximize the likelihood function. Although this task can also be performed using optimization techniques such as those available in

---

[5] https://emcee.readthedocs.io/en/stable/

`scipy`, MCMC provides a more comprehensive understanding of the parameter space by exploring a distribution of possible values rather than a single-point estimate. This distribution can then be used to infer confidence intervals and other statistical measures.

To further improve the convergence of the parameters with less number of walkers, we will use the `minimizer` from module `optimization` of `scipy` package with the TNC (Nocedal & Wright 2006) method to approximate the initials for the MCMC run.

4 Results

4.1 Visualization of instrumental effects

From the mock-up CMB realization maps of I, Q and U, we inject the effect of $\mathbf{M}_{IP}(g_1, g_2, \phi)$ and $\mathbf{M}_{CP}(\epsilon, \psi)$ to $\mathbf{s}_{\mathrm{CMB}} = \begin{bmatrix} I & Q & U & 0 \end{bmatrix}$ and perform the spherical harmonic expansion using `anafast` function of package `healpy`. We can demonstrate the effects on power spectra when we compare them with the original power spectra from CAMB.

(a) Intensity-to-Polarization leakage (b) Cross Polarization
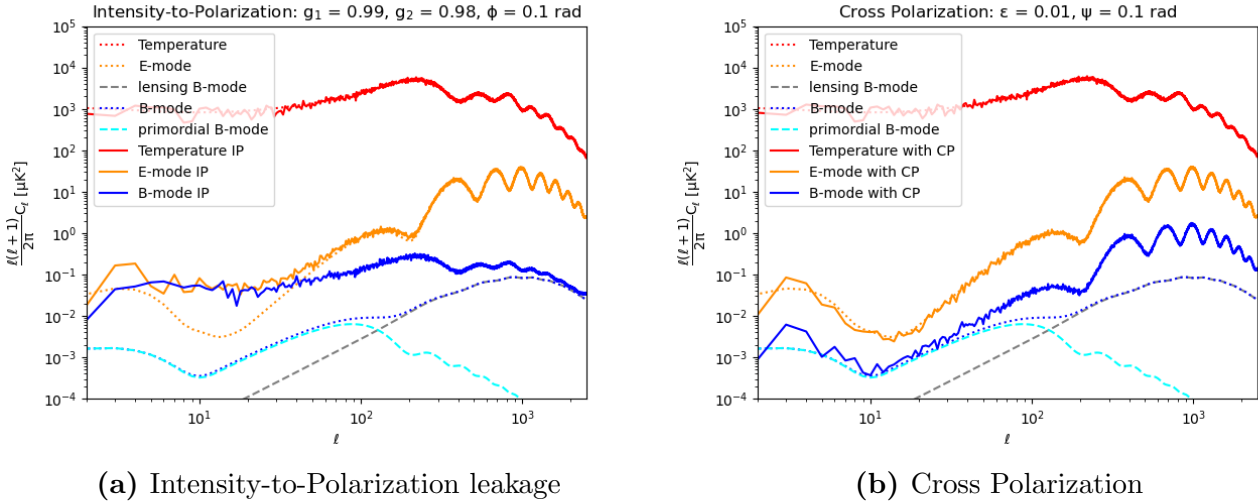
Figure 10: Visualization of instrumental effects on CMB power spectra

For the IP effect, from Figure 10a, we observe significant leakage from temperature anisotropies contaminating both the E-mode and B-mode polarizations, particularly on large angular scales (small l). Due to the inherently small magnitude of the B-mode, its spectrum becomes heavily contaminated, with the amplitude increasing by more than 100 times even with only 1% deviation of g_1, 2% deviation of g_2, and 0.1 rad deviation of ϕ from ideal case $g_1 = 1$, $g_2 = 1$ and $\phi = 0$. Consequently, the B-modes adopt the shape of the temperature spectrum, making accurate measurement of the primordial B-modes extremely challenging.

For the CP effect, there is a noticeable leakage from the E-mode to the B-mode, especially at small angular scales (high l). In this demonstration, we inject a 1% deviation of ϵ and 0.1 rad deviation of ψ. The B-mode spectrum adopts the shape of the E-mode spectrum with an amplitude increase of more than 10 times the original spectrum, as shown in Figure 10b.

As we can see, the B-mode spectrum is very sensitive to imperfections due to the change of each parameter. Therefore, good quantification and constraint on the parameters are crucial for the calibration of the instrumental systematic effects.

4.2   Deriving upper limits on instrumental parameters

By varying each parameter while keeping the others fixed at their ideal values and maximizing the cosmological likelihood with MCMC, we can assess the sensitivity of each parameter on the power spectrum and estimate the upper limits needed to achieve $\delta r < 0.001$.

First, we evaluate the level of cosmic variance, which represents the statistical uncertainty inherent in observations of the power spectra from the CMB realization map (see Figure 8). The estimation of r in the power spectra without IPCP effects on full-sky is of the order of $\mathcal{O}(10^{-5})$ deviation, which is significantly smaller (about 100 times) than our required precision, indicating that this cosmic variance is small enough for our study.

4.2.1   Intensity-to-Polarization

In the case of no cross-polarization effect, after the implementation of the "upper limits" procedure for g_1 and g_2 respectively for IP, we show the result for g_1 in Figure 11, the rest can be found in Appendix A.

(a) B-mode power spectrum dependence on g_1 **(b)** Bias on r with respect to g_1

Figure 11: Illustration of the impact of the parameter g_1, the parameter of IP effect, on cosmological likelihood using MCMC, compared with the $\delta r < 0.001$ requirement.

Note that the percentage denotes how much a parameter deviates from the ideal value. We summarize our findings in Table 1.

Table 1: The upper limits of IP ensuring $\delta r < 0.001$.

Parameter	Limit	Effect
g_1	$\lesssim 0.02\%$	$I \to E, B$
g_2	$\lesssim 0.016\%$	$I \to E, B$
ϕ	$\lesssim 5°$ or $\lesssim 0.087$ rad	$E \to B$

We observe that the two amplitudes, g_1 and g_2, for IP exhibit almost similar behaviors on the large angular scales of the B-mode power spectrum. However, parameter g_2 is slightly more sensitive than parameter g_1. As the deviation percentage of g_1 and g_2 increases, the bias on r also

increases. The phase shift ϕ introduces leakage that predominates on the small angular scales. For upper limits of g_1 and g_2, the value of ϕ should not deviate from $0°$ by more than $5°$.

4.2.2 Cross Polarization

Similarly, with the assumption of no IP effects, we obtained the result for ϵ shown in Figure 12, the result for ψ is in Appendix A. The conclusions drawn from the results are outlined in Table 2.

(a) B-mode power spectrum dependence on ϵ **(b)** Bias on r with respect to ϵ
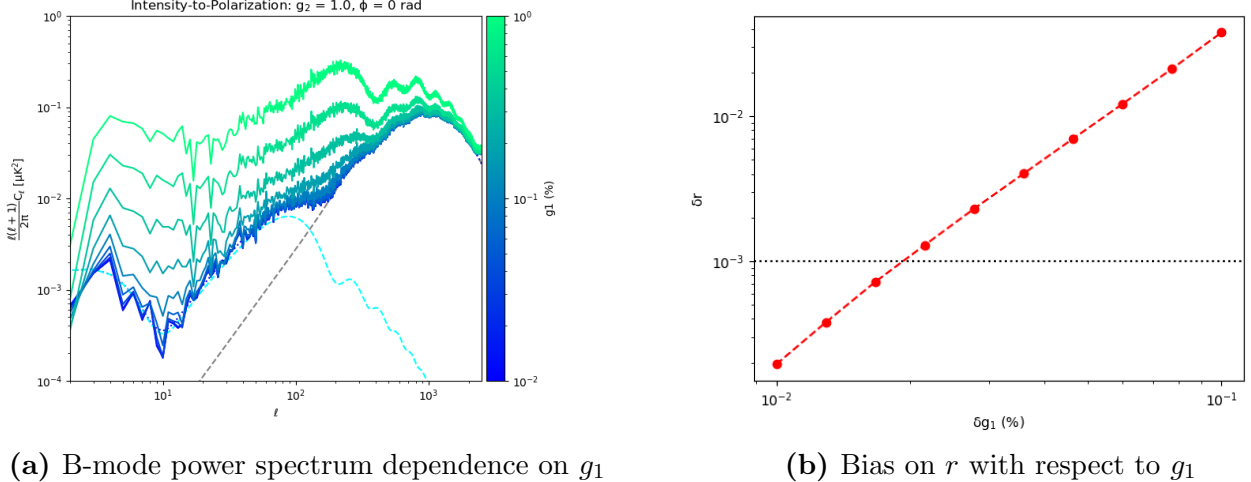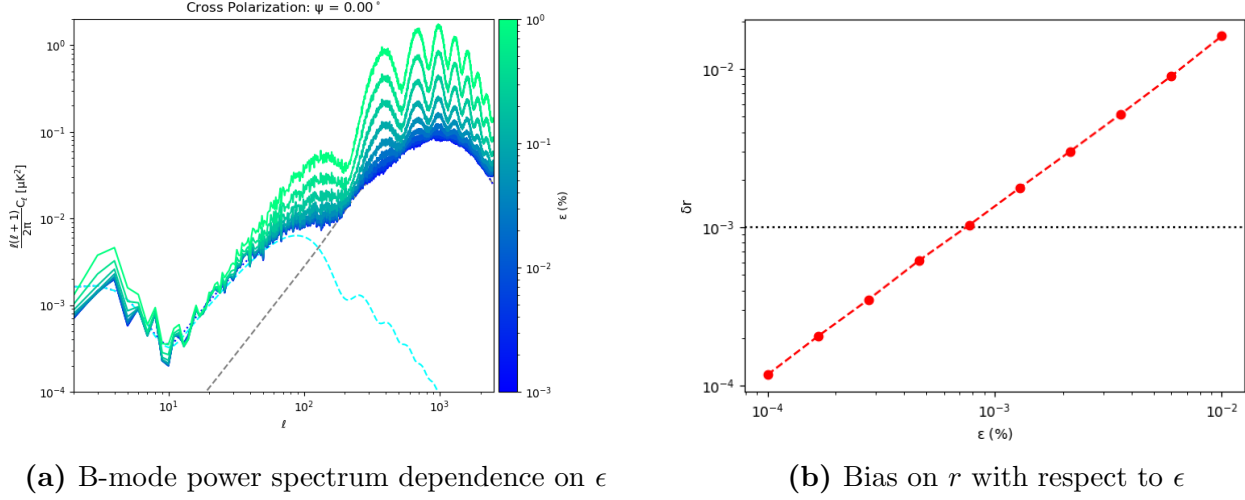
Figure 12: Illustration of the impact of the parameter ϵ, the parameter of CP effect, on cosmological likelihood using MCMC, compared with the $\delta r < 0.001$ requirement.

For the case of CP, the amplitude ϵ is about 10 times more sensitive to the bias on r compared to g_1 and g_2. To correctly remove the bias on r, we need to reach an error of no more than 0.001% $\delta\epsilon$. Interestingly, the parameter ψ depends significantly on ϵ. If we can get to a certain level of $\delta\epsilon$, at some point, ψ would no longer introduce more bias than our requirement.

Table 2: The upper limits of CP ensuring $\delta r < 0.001$.

Parameter	Limit	Effect
ϵ	$\lesssim 0.001\%$	$E \to B$
ψ	depends on ϵ, ignorable if $\epsilon \lesssim 0.0007\%$	$B \to E$

4.3 Parameter Constraints with MCMC

Following the framework and method introduced in section 2 and 3, we implemented the MCMC to maximize the χ^2 likelihood, so that we can constrain samples of parameters as follows. We assess the volume of data we need to reach our requirement for δr.

4.3.1 Unpolarized atmospheric data

We utilized the TOD of the atmosphere from the TOAST simulation, where the parameter ϕ is completely absent, rendering it unconstrained as previously mentioned. After running MCMC

17
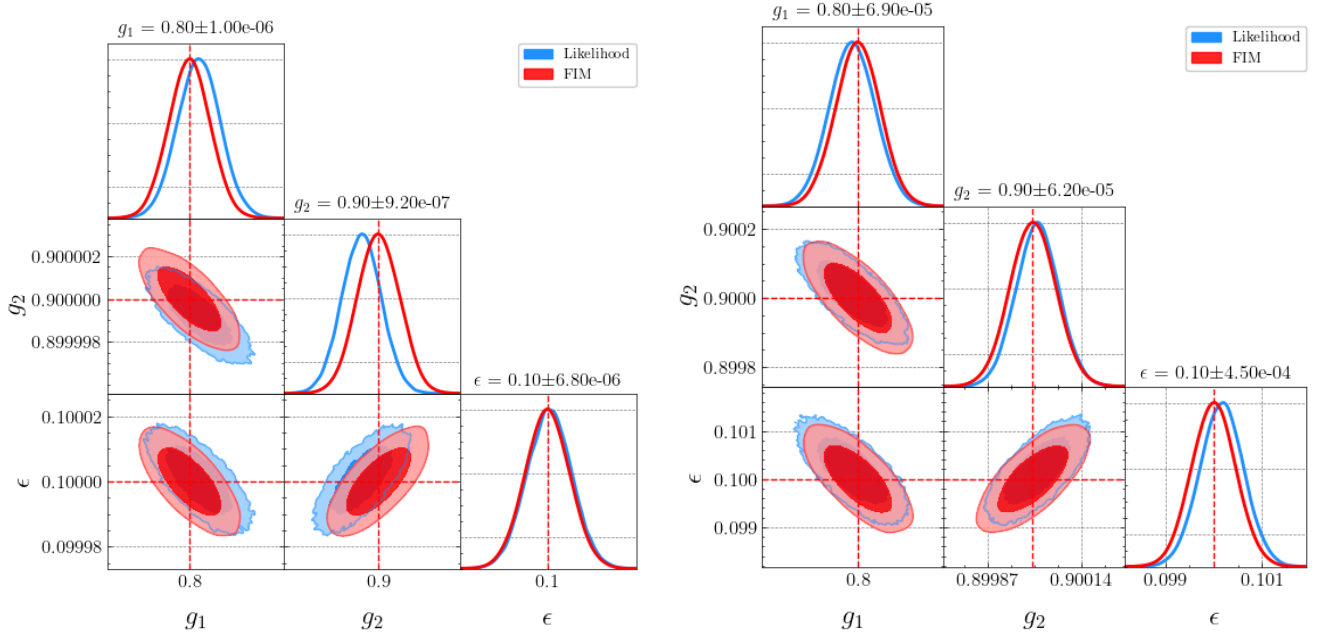
(a) 180 detectors with polarization angles separated by 1° on 1000 time series.

(b) One detector with polarization angle of 45° on 100000 time series.

Figure 13: The MCMC run for unpolarized atmospheric data from the TOAST simulation, using 80 walkers and 5000 iterations on both single and multiple detectors, resulted in Gaussian distributions and correlations with 1-sigma and 2-sigma contour (blue). An overlay of samples from FIM is shown with the same contour (red).

for the remaining four parameters, we detected a degeneracy among them. To address this, we used the FIM to characterize the degeneracy.

Firstly, the eigen-decomposition of the FIM revealed that the eigenvalue corresponding to ψ is zero, indicating that the likelihood function does not change concerning the change of ψ. This aligns with our findings that the CP effect is dominated by ϵ, making ψ negligible when $\epsilon \lesssim 0.0007$. Therefore, we focused on running the model with only three amplitudes: g_1, g_2, and ϵ. We also overlaid the covariance derived from the FIM to confirm our analysis.

Secondly, we noticed that the standard deviation of parameter ϵ from the likelihood function is always $\mathcal{O}(100)$ higher than for g_1 and g_2. However, as stated previously, δr is more sensitive to variations in ϵ than in g_1 and g_2. As a result, achieving the required precision for ϵ automatically fulfills the requirement for g_1, g_2, and ψ. For TOD simulation containing about 10^5 data points, we can reach $\sigma_\epsilon = 0.00045$, which is 45 times higher than needed. We then used the multi-detector option, with a 180-degree rotation angle θ and 1000 data points, giving us a total of 1.8×10^5 data points. We can then constrain $\sigma_\epsilon \approx 0.00068\% < \delta\epsilon_{\text{upper limit}}$, making ψ negligible, with σ_{g_1} and σ_{g_2} confirmed to be $\mathcal{O}(100)$ smaller than necessary. Therefore, 1.8×10^5 points, equivalent to about two and a half hours of observation on one detector or about 50 seconds on 180 detectors, should be the minimum data volume ensuring that $\delta r < 0.001$.

4.3.2 Second approach: a model without HWP

This model, built under the assumption that a demodulation operation for the HWP was carried out, results in the following form for the unpolarized data after going through the grid:

$$\mathbf{d}_t = \frac{I_t}{2}\begin{bmatrix} 1 \\ \cos(2\theta) \\ \sin(2\theta) \\ 0 \end{bmatrix} + \mathbf{n}_t \tag{42}$$

In this approach, we no longer work on the observable TOD, the polarized data reconstructed
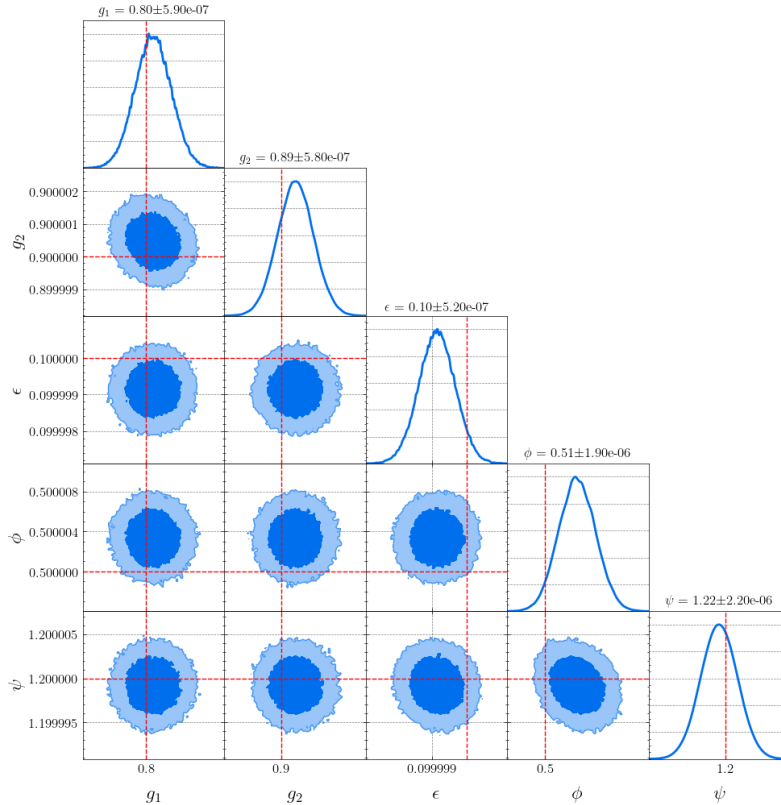
Figure 14: The result of MCMC simulation for the model without HWP on 15 detectors of different polarization angles on 5000 time-series data.

after doing HWP demodulation. The result of maximizing the log-likelihood function is shown in Figure 14.

Assuming that perfect demodulation of HWP was performed, we observe that the parameters are more easily constrained because they are no longer correlated with each other. In addition, we require a much smaller volume of data $\sim 0.75 \times 10^5$ (about half of the HWP model) points or about 1 hour of observation with one detector in our TOAST simulation.

5 Discussion

Overall, we successfully achieved our goals of constraining the instrumental parameters according to our models satisfying the necessary sensitivity needed for the next generation of CMB

experiments. There are some notable points we need to discuss to gain more understandable knowledge.

5.1 Effects of each parameter on the power spectrum

The demonstration of the effects of each parameter shows that g_1 and g_2 significantly impact the B-mode spectrum at small ℓ values, where the primordial B-mode dominates. In contrast, parameters ϕ, ϵ, and ψ are more influential at large ℓ values. This difference can be attributed to the distinct shapes of temperature anisotropies and E-mode polarization. Interestingly, the impact of the phase shift ϕ on the power spectrum is quite similar to the CP effects. Additionally, the IP effect influences not only the B-mode polarization but also the E-modes, particularly at small multipoles. This region is crucial for future efforts to better constrain the optical depth τ.

Another interesting point is that g_1 and g_2 exhibit similar behavior, yet they are not given the same limit. This discrepancy is caused by the term $(g_1^2 - g_2^2)$. When $g_2 = 1$ and g_1 vary, the term is negative, but when $g_1 = 1$ and g_2 vary, the term is positive.

5.2 Degeneracy in the case of unpolarized atmospheric data
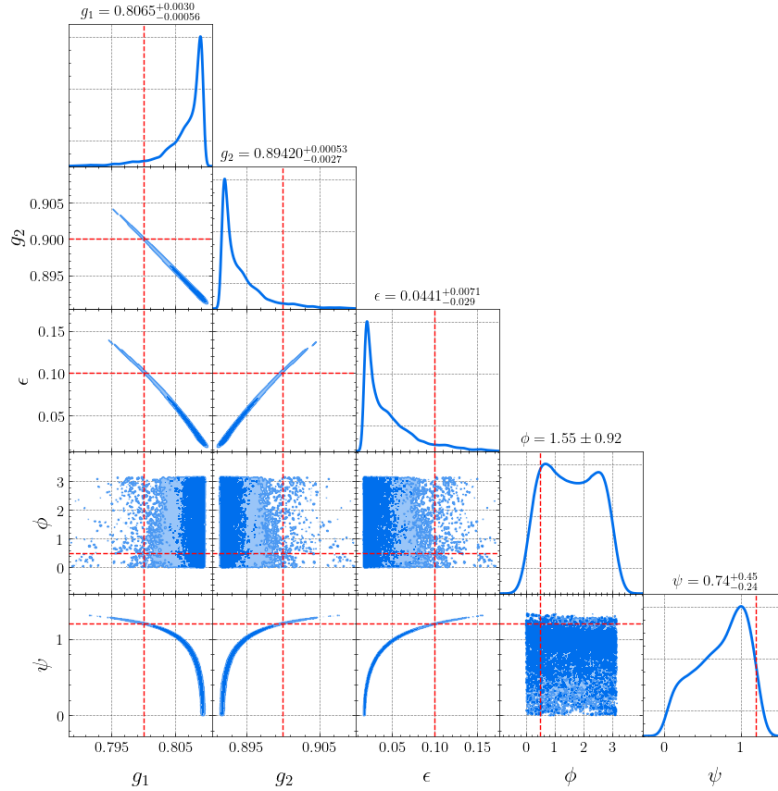
Figure 15: Degeneracies of 4 parameters g_1, g_2, ϵ, ψ are shown by the curves between each parameter pair; ϕ disappears in the model results in the uniform posterior distribution.

The shape curves in Figure 15 show that there should be less than four degrees of freedom. Therefore, we decided to use FIM analysis, which results in one of the four eigenvalues being equal to zero, indicating that there are only three degrees of freedom that can be characterized in our model. This can be accounted for by the fact that ψ depends greatly on ϵ. Additionally, from

the diagonal elements in the covariance derived from FIM, we obtained the standard errors of the rest of the three parameters g_1, g_2, and ϵ. As mentioned previously, the FIM analysis agrees with MCMC sampling that σ_ϵ is about $\mathcal{O}(100)$ times larger than σ_{g_1} and σ_{g_2}.

5.3 Model of drone polarized data

One problem with using unpolarized atmospheric data is that we lost the information on ϕ which should be strictly constrained to completely remove the bias on r. Therefore, polarization data is necessary. The plan is to use a drone that provides a fully polarized signal $\mathbf{s}_{\text{drone}} =$
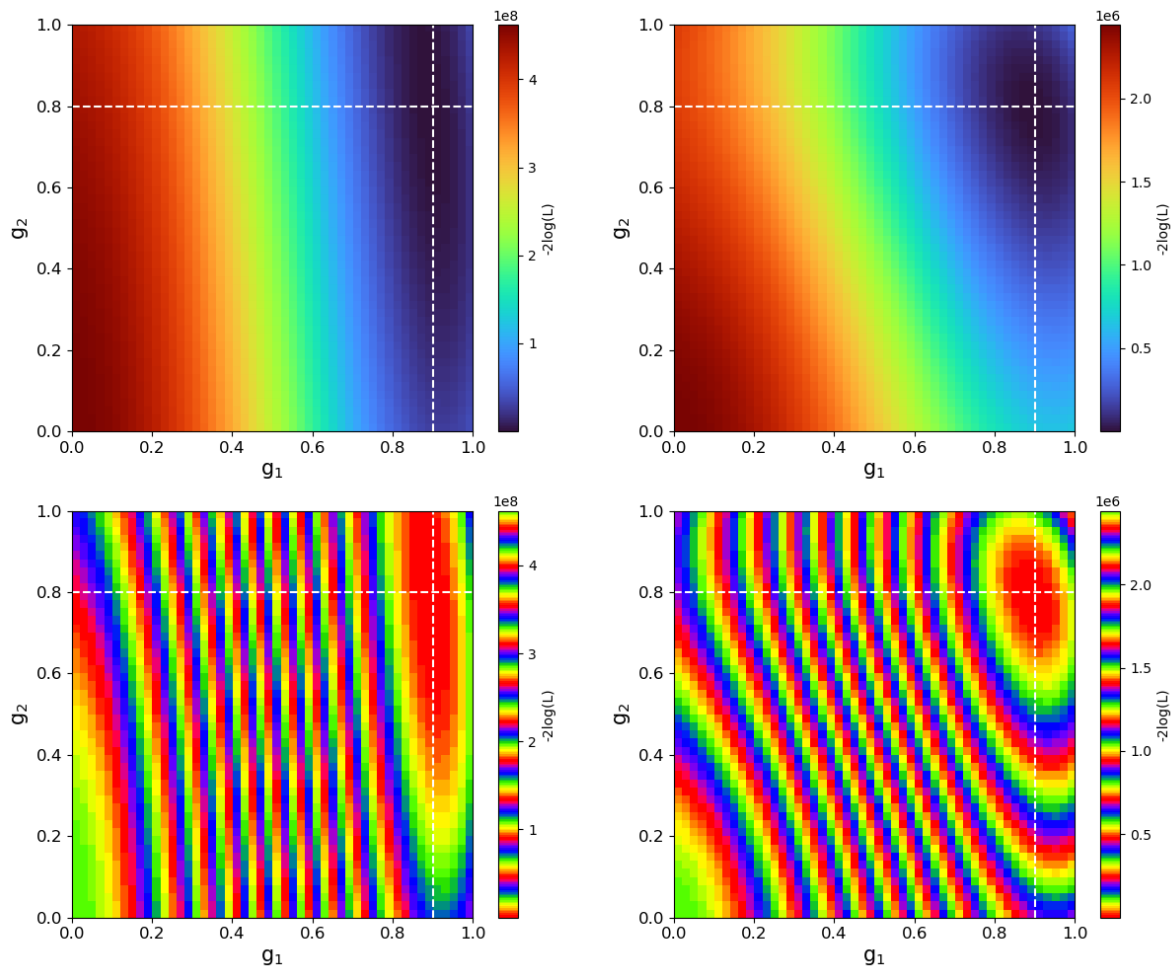
Figure 16: The log-likelihood as the function of g_1 and g_2. *Left:* constant drone signal, the likelihood function is flatted out. *Right:* constant drone signal with a modulated signal of uniform distribution between [0,1] constraint the minimum much better.

$\begin{bmatrix} I & I/\sqrt{2} & I/\sqrt{2} & 0 \end{bmatrix}$ that has been used for calibration at Simons Observatory (Nati et al. 2017). However, there is a weak constraint between parameters as shown in Figure 16 for g_1, g_2 as an example.

We suggest that this weak constraint can be attributed to the coupling/similarity between the

coefficients A_1 and A_2 of I and Q:

$$A_1 = \frac{g_1^2 + g_2^2}{2} + \frac{g_1^2 - g_2^2}{2}(1 - 2\epsilon)\cos(4\psi_t - 2\theta) + (g_1^2 - g_2^2)\sqrt{\epsilon(1 - \epsilon)}\cos\psi\sin(4\psi_t - 2\theta) \tag{43}$$

$$A_2 = \frac{g_1^2 - g_2^2}{2} + \frac{g_1^2 + g_2^2}{2}(1 - 2\epsilon)\cos(4\psi_t - 2\theta) + (g_1^2 + g_2^2)\sqrt{\epsilon(1 - \epsilon)}\cos\psi\sin(4\psi_t - 2\theta) \tag{44}$$

There needs to be a way to modulate the drone signals, to distinguish I and Q for regularizing χ^2 likelihood function. In this preliminary result shown in Figure 17, we demonstrated the use of randomized signals on top of the constant drone signal for regularization. In practice, there is a chopper with a frequency of approximately 40 Hz on board the drone. This may be achieved by simulating a regular boxcar on top of the TOD. Alternatively, we can combine atmospheric data with drone data to resolve this problem.

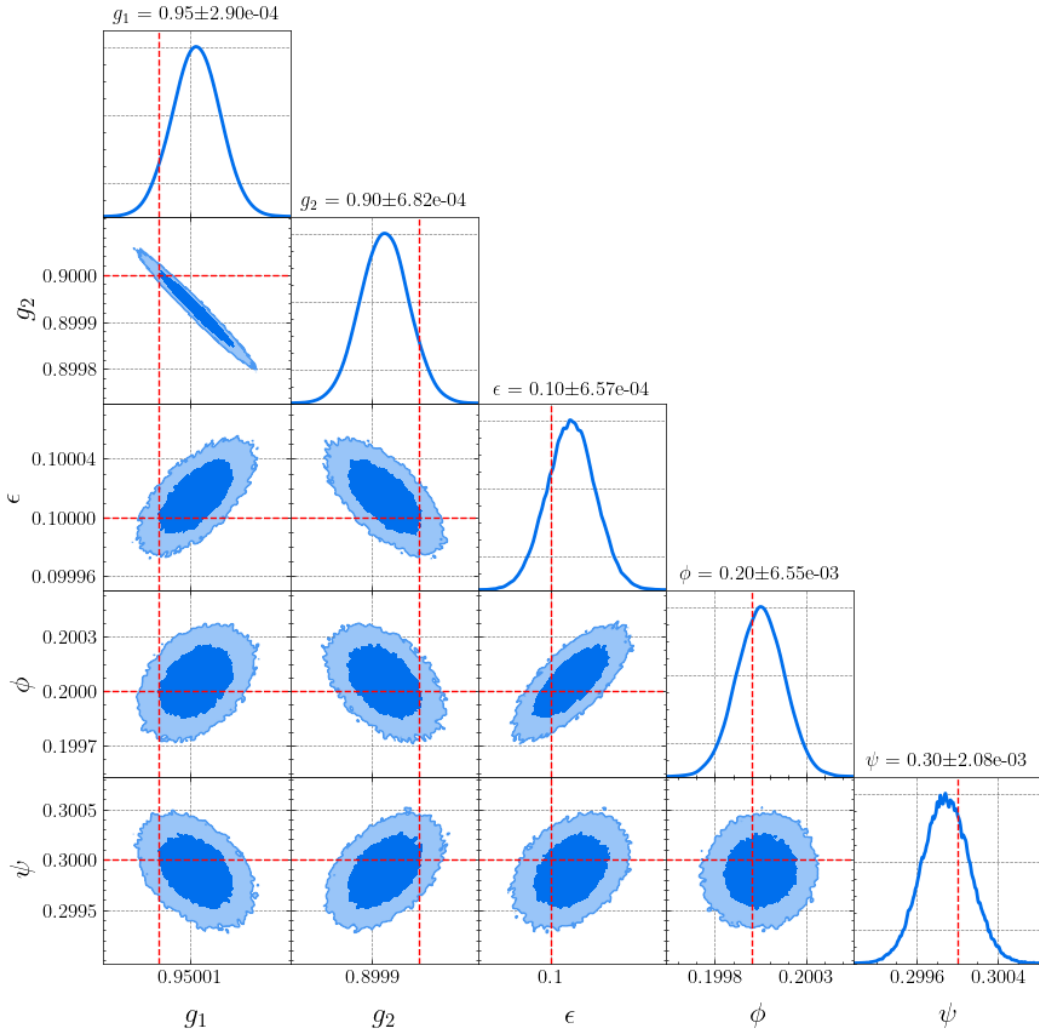**Figure 17:** MCMC simulation results for drone-based fully polarized data with uniform modulated signal.

In Figure 17, we observe a correlation between g_1 and g_2, which appears to be nearly linear. Additionally, other pairs of parameters show varying degrees of correlation and anti-correlation. The phases do not exhibit a significant correlation with each other. Since we only need to constrain

the ϕ for calibration under ~ 0.087 rad deviation, a much smaller data volume can be used (~ 7-minute observation on 1 detector). Other parameters can be cross-checked with results from atmospheric data.

On the other hand, we could also try to fix the constrained g_1, g_2 and ϵ from the unpolarized atmospheric data and run MCMC for only ϕ.

5.4   The model without HWP

This model proves more effective at constraining data using only atmospheric measurements, showing no correlations among its parameters. However, its effectiveness inevitably depends on our demodulation process. During this internship, we haven't had enough time to thoroughly cover the demodulation signal of HWP. Additionally, the HWP introduces intensity-to-polarization leakage, which the model without HWP does not address during demodulation. In contrast, the first model includes these effects. Therefore, further investigation into the HWP demodulation process is necessary to evaluate the usefulness of this approach.

5.5   Future works

In the next few months, we plan to closely examine the polarized drone data to better understand how to modulate the data for a practical degeneracy breaking. Following this, we will focus on enhancing and integrating our work into the broader framework of the SciPol project. In particular, we intend to incorporate our model into the modeling of the HWP and implement it within the JAX[6]-powered SciPol framework so-called FURAX, which is a high-performance array computing framework. In addition, we plan to explore the use of spectral likelihood (Stompor et al. 2009) in this model and combine it into a generalized framework with a pointing matrix on a 2D atmospheric frozen sheet simulation using the Kolmogorov model. We will improve the TOD by incorporating the pointing matrix $\mathbf{s}_t = \mathbf{P}_t^k \cdot \mathbf{s}_{\mathrm{atm}}$ for each detector k.

One of the key questions is how to apply our methods to real data sets. This can be resolved with the Simons Observatory which is currently collecting data, we hope to gain access to this data in the future.

6   Conclusion

CMB B-mode polarization is crucial in probing cosmic inflation and is a primary focus of upcoming cosmological experiments. Achieving high sensitivity requires precise instrumental modeling for calibration. Our project addresses two main systematic effects: intensity-to-polarization leakage and cross-polarization. Using the Jones/Mueller formalism within the parametric component separation method, we aimed to constrain model parameters with an accuracy ensuring calibration to $\delta r < 0.001$. Key findings are summarized below:

(i) We developed a model that accounts for intensity-to-polarization leakage with three parameters g_1, g_2 and ϕ and cross-polarization effect with two parameters ϵ and ψ useful for the calibration of the upcoming CMB instruments.

(ii) We demonstrated that the g_1 and g_2 parameters have effects primarily on large angular scales while ϕ, ϵ and ψ are mostly on small angular scales.

---

[6]https://jax.readthedocs.io/en/latest/index.html

(iii) We derived the upper limits on each parameter using power spectra from CMB realization and setting the requirement on residuals to be $\delta r < 0.001$. ϵ is $\mathcal{O}(10)$ more sensitive to δr compared to g_1 and g_2. The limits for phase shift $\phi < 5°(\approx 0.087 \text{ rad})$ are also estimated. Parameter ψ is less sensitive to δr and can be negligible when we reach $\delta_\epsilon < 0.0007\%$.

(iv) With unpolarized data from the atmosphere, we use FIM to characterize poorly-constrained ψ and successfully constrain for the amplitudes g_1, g_2 and ϵ. There are correlations between parameters and the error level of ϵ is $\mathcal{O}(100)$ times more than g_1 and g_2. To meet our requirements, we estimate that approximately $\mathcal{O}(180000)$ data points (~ 2.5 hours of observation on one detector) are necessary.

(v) To constrain ϕ, we need to rely on the fully polarized drone reference source, with an additional modulated signal on top to improve the decoupling of I and Q.

(vi) In our second approach, which accounts for instrumental effects after HWP demodulation, we successfully constrain all of the parameters with a smaller data quantity of about $\mathcal{O}(75000)$ (~ 1 hour of observation on one detector) due to the lack of correlations between them.

References

Ade, P. A. R., Ahmed, Z., Aikin, R. W., et al. 2015, The Astrophysical Journal, 811, 126

Alpher, R. A., Bethe, H., & Gamow, G. 1948, Phys. Rev., 73, 803

Bennett, C. L., Larson, D., Weiland, J. L., et al. 2013, The Astrophysical Journal Supplement Series, 208, 20

Eriksen, H. K., Dickinson, C., Lawrence, C. R., et al. 2006, The Astrophysical Journal, 641, 665–682

Fabbian, G. & Peloton, J. 2021, Journal of Open Source Software, 6, 3022

Fixsen, D. J. 2009, The Astrophysical Journal, 707, 916–920

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, Publications of the Astronomical Society of the Pacific, 125, 306–312

Ghigna, T., Matsumura, T., Patanchon, G., Ishino, H., & Hazumi, M. 2020, Journal of Cosmology and Astroparticle Physics, 2020, 030–030

Goodman, J. & Weare, J. 2010, Communications in Applied Mathematics and Computational Science, 5, 65

Guth, A. H. 1981, Phys. Rev. D, 23, 347

Hunte, C. 2008, Fizika A, 17, 51

Kunnen, B., Macdonald, C., Doronin, A., et al. 2015, Journal of Biophotonics, 8, 317

Leloup, C., Patanchon, G., Errard, J., et al. 2023, Impact of beam far side-lobe knowledge in the presence of foregrounds for LiteBIRD

Linde, A. 2017, Journal of Cosmology and Astroparticle Physics, 2017, 006–006

Martin, J., Ringeval, C., & Vennin, V. 2023, Encyclopaedia Inflationaris

Nati, F., Devlin, M. J., Gerbino, M., et al. 2017, Journal of Astronomical Instrumentation, 06

Nocedal, J. & Wright, S. J. 2006, Numerical Optimization, 2nd edn. (New York, NY, USA: Springer)

P. A. R. Ade, T. P. C., Akiba, Y., Anthony, A. E., et al. 2014, The Astrophysical Journal, 794, 171

Peebles, P. J. E. 1968, , 153, 1

Rashid, M., Brown, M. L., & Thomas, D. B. 2023, CMB Polarisation Signal Demodulation with a Rotating Half-Wave Plate

Samuel Reich, E. 2013, Nature

Smoot, G. F., Bennett, C. L., Kogut, A., et al. 1992, , 396, L1

Stompor, R., Leach, S., Stivoli, F., & Baccigalupi, C. 2009, Monthly Notices of the Royal Astronomical Society, 392, 216–232

Tegmark, M., Taylor, A. N., & Heavens, A. F. 1997, The Astrophysical Journal, 480, 22–35

The LiteBIRD Collaboration. 2022, Progress of Theoretical and Experimental Physics, 2023

The Planck Collaboration. 2020, Astronomy amp; Astrophysics, 641, A5

The Simons Observatory. 2021, Journal of Cosmology and Astroparticle Physics, 2021, 032

Tristram, M., Banday, A., Górski, K., et al. 2022, Physical Review D, 105

Turner, M. S. 2022, Annual Review of Nuclear and Particle Science, 72, 1–35

APPENDIX

A   Upper limits of parameters
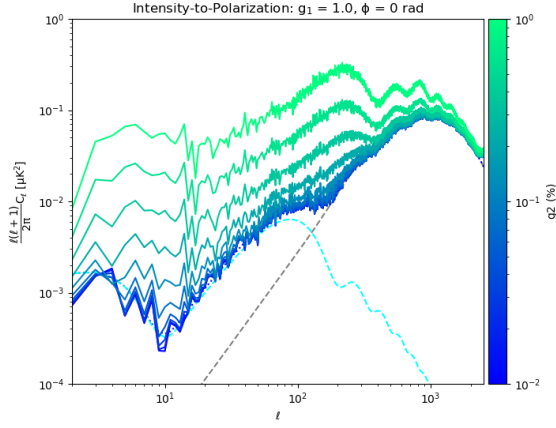
(a) B-mode power spectrum dependence on g_2
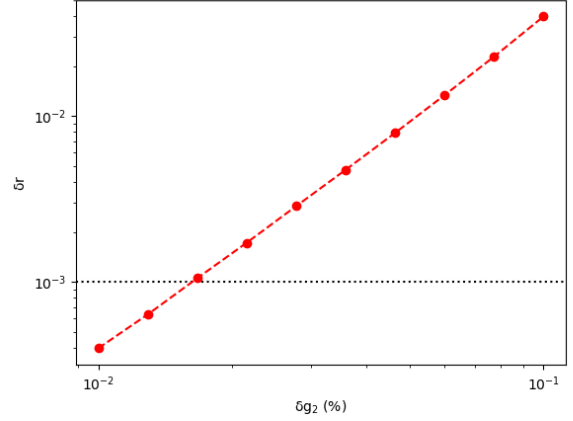
(b) Bias on r with respect to g_2
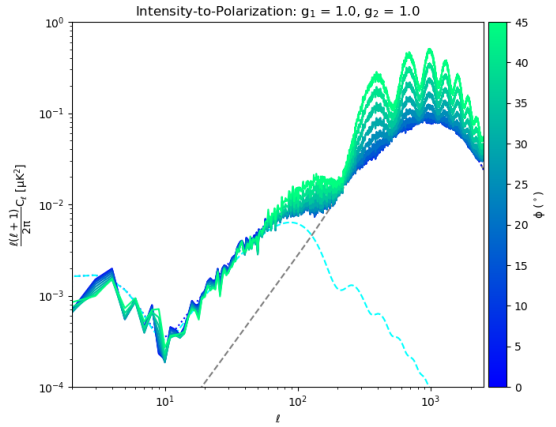
(c) B-mode power spectrum dependence on ϕ
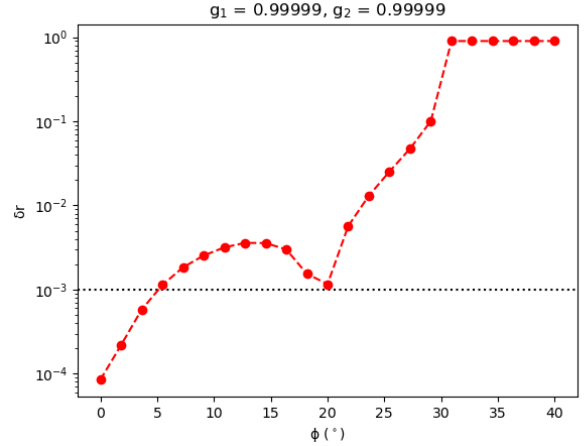
(d) Bias on r with respect to ϕ

Figure 18: Estimation of bias on tensor-to-scalar ratio r for intensity-to-polarization effect's parameter using MCMC on cosmological likelihood, comparing with the requirement $\delta r < 0.001$ for the upcoming CMB experiments.
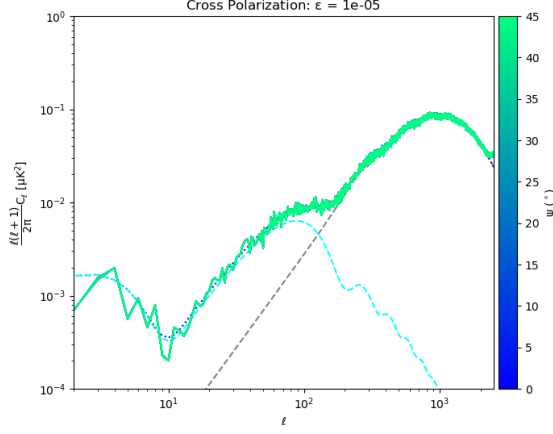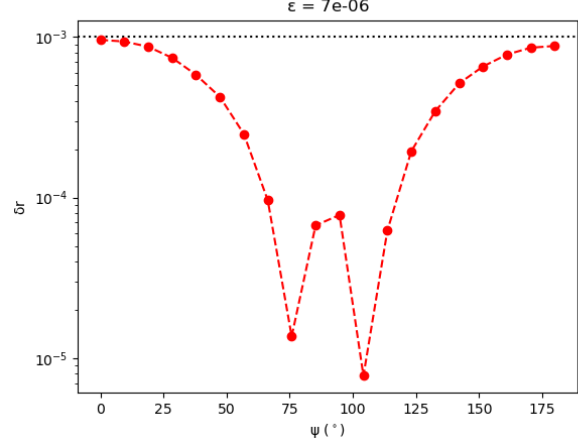
(a) B-mode power spectrum dependence on ψ

(b) Bias on r with respect to ψ

Figure 19: Estimation of bias on tensor-to-scalar ratio r for cross polarization's parameter ψ using MCMC on cosmological likelihood, comparing with the requirement $\delta r < 0.001$ for the upcoming CMB experiments.

B   Maximum likelihood formalism

Our goal is to find the set of parameters $\hat{\xi} = \{g_1, g_2, \epsilon, \phi, \psi\}$ for the mueller matrix $\mathbf{A}(\hat{\xi})$, so that we can maximize the probability of obtaining the observed data given a sky signal:

$$\mathbf{d} = \mathbf{A}(\hat{\xi})\mathbf{s} + \mathbf{n} \tag{45}$$

From the Bayes's theorem:

$$P(\mathbf{s}|\mathbf{d}) = \frac{P(\mathbf{d}|\mathbf{s})P(\mathbf{s})}{P(\mathbf{d})} \tag{46}$$

where

- $P(\mathbf{s}|\mathbf{d})$ is the probability of recovering the original sky signal given the observable data, referred to as the *posterior*.

- $P(\mathbf{d}|\mathbf{s})$ is the probability of having the data given a sky signal, referred to as the *likelihood*.

- $P(\mathbf{s})$ is the probability of the sky signal, referred to as the *prior*.

- $P(\mathbf{d})$ is the probability of the data, referred to as the *marginalization*.

Finding the maximum probability for the posterior is similar to answering the question, "Given the observations, what is the most likely model?". In other words, we try to find $\hat{\xi}$ that can maximize $\frac{P(\mathbf{d}|\mathbf{s})P(\mathbf{s})}{P(\mathbf{d})}$. We can ignore $P(\mathbf{d})$ since it does not depend on $\hat{\xi}$. Assuming a uniform prior for $P(\mathbf{s})$, the only thing left is to maximize the likelihood $P(\mathbf{d}|\mathbf{s})$.

In our model, $\mathbf{As}$ is well-characterized; therefore, the likelihood probability can be described solely by the probability of the noise $\mathcal{L}(\mathbf{n})$. Assuming that the noise $\mathbf{n}$ added to the input maps is drawn from a multivariate Gaussian distribution with zero mean and covariance $\mathbf{N} \equiv \langle \mathbf{nn}^\top \rangle$:

$$\mathcal{L}(\mathbf{n}) = \mathcal{N}(0, \mathbf{N}) = \frac{1}{\sqrt{2\pi|\mathbf{N}|^2}}e^{-\frac{1}{2}\mathbf{n}^\top\mathbf{N}^{-1}\mathbf{n}} \tag{47}$$

Taking the logarithm of both sides of the equation:

$$\log \mathcal{L}(\mathbf{n}) = -\frac{1}{2}\mathbf{n}^\top \mathbf{N}^{-1}\mathbf{n} + \log\left(\frac{1}{\sqrt{2\pi|\mathbf{N}|^2}}\right) \tag{48}$$

Because the term $\log\left(\dfrac{1}{\sqrt{2\pi|\mathbf{N}|^2}}\right)$ is a constant, maximizing the expression $\log \mathcal{L}(\mathbf{n})$ similar to minimizing the expression:

$$-2\log \mathcal{L}(\mathbf{n}) = \mathbf{n}^\top \mathbf{N}^{-1}\mathbf{n} = (\mathbf{d}-\mathbf{As})^\top \mathbf{N}^{-1}(\mathbf{d}-\mathbf{As}) \tag{49}$$

The best estimate of the sky signal, $\hat{\mathbf{s}}$, can be determined by minimizing the log-likelihood function given in Equation 49:

$$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}} -2\log \mathcal{L}(\mathbf{n}) = \arg\min_{\mathbf{s}} (\mathbf{d}-\mathbf{As})^\top \mathbf{N}^{-1}(\mathbf{d}-\mathbf{As}) \tag{50}$$

$$= \arg\min_{\mathbf{s}} \left[\mathbf{d}^\top \mathbf{N}^{-1}\mathbf{d} + \mathbf{s}^\top \mathbf{A}^\top \mathbf{N}^{-1}\mathbf{As} - \mathbf{d}^\top \mathbf{N}^{-1}\mathbf{As} - \mathbf{s}^\top \mathbf{A}^\top \mathbf{N}^{-1}\mathbf{d}\right] \tag{51}$$

We assume that the noise $N = \langle \mathbf{nn}^\top \rangle$ is zero-mean and uncorrelated (white noise); therefore, $\mathbf{N}$ is a symmetric matrix. We know that given a symmetric matrix $\mathbf{O}$ and any arbitrary vector $\mathbf{m}$, $\mathbf{m}^\top \mathbf{O} = \mathbf{Om}$. Additionally, the inverse of a symmetric matrix is also symmetric. Thus:

$$\hat{\mathbf{s}} = \arg\min_{\mathbf{s}} \left[\mathbf{d}^\top \mathbf{N}^{-1}\mathbf{d} + \mathbf{s}^\top \mathbf{A}^\top \mathbf{N}^{-1}\mathbf{As} - 2\mathbf{s}^\top \mathbf{A}^\top \mathbf{N}^{-1}\mathbf{d}\right] \tag{52}$$

The minimum $\hat{\mathbf{s}}$ can be evaluated by taking the derivative of log-likelihood equal to 0:

$$\frac{\partial\left[\mathbf{d}^\top \mathbf{N}^{-1}\mathbf{d} + \mathbf{s}^\top \mathbf{A}^\top \mathbf{N}^{-1}\mathbf{As} - 2\mathbf{s}^\top \mathbf{A}^\top \mathbf{N}^{-1}\mathbf{d}\right]}{\partial \mathbf{s}} = 0 \tag{53}$$

$$\Rightarrow \quad 2\mathbf{A}^\top \mathbf{N}^{-1}\mathbf{A}\hat{\mathbf{s}} - 2\mathbf{A}^\top \mathbf{N}^{-1}\mathbf{d} = 0 \tag{54}$$

$$\Rightarrow \quad \hat{\mathbf{s}} = (\mathbf{A}^\top \mathbf{N}^{-1}\mathbf{A})^{-1}\mathbf{A}^\top \mathbf{N}^{-1}\mathbf{d} \tag{55}$$

Note that the solution only holds when all the columns of $\mathbf{A}$ are linearly independent, meaning that $\mathbf{A}$ is of full column rank. Therefore, $\operatorname{rank}(\mathbf{A}^\top \mathbf{N}^{-1}\mathbf{A}) = \operatorname{rank}(\mathbf{A})$, ensuring that the matrix $\mathbf{A}^\top \mathbf{N}^{-1}\mathbf{A}$ is invertible.

Replace $\hat{\mathbf{s}}$ into the second term of $-2\log \mathcal{L}(\mathbf{n})$, we have:

$$\mathbf{s}^\top \mathbf{A}^\top \mathbf{N}^{-1}\mathbf{As} = \mathbf{s}^\top (\mathbf{A}^\top \mathbf{N}^{-1}\mathbf{A})(\mathbf{A}^\top \mathbf{N}^{-1}\mathbf{A})^{-1}\mathbf{A}^\top \mathbf{N}^{-1}\mathbf{d} = \mathbf{s}^\top \mathbf{A}^\top \mathbf{N}^{-1}\mathbf{d} \tag{56}$$

Thus,

$$-2\log \mathcal{L}(\mathbf{n}) = \mathbf{d}^\top \mathbf{N}^{-1}\mathbf{d} - \mathbf{s}^\top \mathbf{A}^\top \mathbf{N}^{-1}\mathbf{d} \tag{57}$$

Because $\mathbf{d}^\top \mathbf{N}^{-1}\mathbf{d}$ is just a constant, we can ignore it in the likelihood we need to minimize:

$$-2\log \mathcal{L}(\mathbf{n}) = -\mathbf{d}^\top \mathbf{N}^{-1}\mathbf{A}(\mathbf{A}^\top \mathbf{N}^{-1}\mathbf{A})^{-1}\mathbf{A}^\top \mathbf{N}^{-1}\mathbf{d} \tag{58}$$

This is the *spectral likelihood.*

Note that, since the matrix $\mathbf{A}$ is full column rank, it implies that if $\mathbf{A}$ is also a square matrix, then $\mathbf{A}$ is invertible. Therefore, the spectral likelihood approach would not be effective if $\mathbf{A}$ is a square matrix.

C FIM for unpolarized atmospheric model

The log-likelihood model is defined as:

$$-2\log\mathcal{L} = \sum_{t,\theta} \frac{\left(d_t - A_1(\hat{\xi},t)\frac{I_t}{2}\right)^2}{\sigma^2} \tag{59}$$

where $\xi \in \{g_1, g_2, \epsilon, \psi\}$.

The Fisher Information Matrix **F** elements can be defined as:

$$F_{ij} = \left\langle \frac{\partial^2(-2\log\mathcal{L})}{\partial\xi_i\partial\xi_j} \right\rangle \tag{60}$$

where $\langle . \rangle$ is the operator that estimates the function at the true values.

The derivative of the log-likelihood function:

$$\frac{\partial(-2\log\mathcal{L})}{\partial\xi_i} = -\sum_t \frac{I_t}{\sigma^2}(d_t - A_1\frac{I_t}{2})\frac{\partial A_1}{\partial\xi_i} \tag{61}$$

$$\frac{\partial^2(-2\log\mathcal{L})}{\partial\xi_i\partial\xi_j} = -\sum_t \left[\frac{I_t}{\sigma^2}(d_t - A_1\frac{I_t}{2})\frac{\partial^2 A_1}{\partial\xi_i\partial\xi_j} - \frac{I_t^2}{2\sigma^2}\frac{\partial A_1}{\partial\xi_i}\frac{\partial A_1}{\partial\xi_j} \right] \tag{62}$$

Evaluating the second derivative of the log-likelihood function at the true values reduces the first term to zero. Then, the FIM elements are given in terms of:

$$F_{ij} = \frac{1}{2\sigma^2}\left\langle \sum_t I_t^2 \frac{\partial A_1}{\partial\xi_i}\frac{\partial A_1}{\partial\xi_j} \right\rangle \tag{63}$$

$$\tag{64}$$

where the partial derivatives of $\mathcal{M}$ are given by:

$$\frac{\partial A_1}{\partial g_1} = g_1 + g_1(1-2\epsilon)\cos(4\varphi_t - 2\theta) + 2g_1\sqrt{(1-\epsilon)\epsilon}\cos\psi\sin(4\varphi_t - 2\theta) \tag{65}$$

$$\frac{\partial A_1}{\partial g_2} = g_2 - g_2(1-2\epsilon)\cos(4\varphi_t - 2\theta) - 2g_2\sqrt{(1-\epsilon)\epsilon}\cos\psi\sin(4\varphi_t - 2\theta) \tag{66}$$

$$\frac{\partial A_1}{\partial\epsilon} = -(g_1^2 - g_2^2)\cos(4\varphi_t - 2\theta) + (g_1^2 - g_2^2)\frac{1-2\epsilon}{2\sqrt{(1-\epsilon)\epsilon}}\cos\psi\sin(4\varphi_t - 2\theta) \tag{67}$$

$$\frac{\partial A_1}{\partial\psi} = -(g_1^2 - g_2^2)\sqrt{(1-\epsilon)\epsilon}\sin\psi\sin(4\varphi_t - 2\theta) \tag{68}$$

D   Jones and Mueller Formalism

D.1   Jones Vector and Polarization States

Starting with the Maxwell's equation of electromagnetism in the vacuum without any charges or currents:

$$\nabla \cdot \mathbf{E} = 0 \tag{69}$$

$$\nabla \cdot \mathbf{B} = 0 \tag{70}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{71}$$

$$\nabla \times \mathbf{B} = \frac{1}{c^2} \frac{\partial \mathbf{E}}{\partial t} \tag{72}$$

These four equations can be reduced into the wave equation for $\mathbf{E}$ and $\mathbf{B}$:

$$\nabla^2 \mathbf{E} = \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} \qquad \nabla^2 \mathbf{B} = \frac{1}{c^2} \frac{\partial^2 \mathbf{B}}{\partial t^2} \tag{73}$$

This means that each component (x, y, z) of E and B also satisfies the scalar wave equation, for example, $E_x(\mathbf{r}, t)$ can be expressed as:

$$\nabla^2 E_x(x, y, z, t) - \frac{1}{c^2} \frac{\partial E_x(x, y, z, t)}{\partial t^2} = 0 \tag{74}$$

The separation solution for each component can be represented in a plane wave expression:

$$E_x(\mathbf{r}) = A_x e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \tag{75}$$

We note that similar equations hold for E_y, E_z, B_x, B_y, B_z.

Assuming that an electromagnetic field propagates in the z-direction, which means $A_z = 0$ and the field oscillates in the xy-plane. Polarization is the direction of oscillation, which can be characterized by a two-dimensional *Jones vector*:

$$J = \begin{bmatrix} A_x \\ A_y \end{bmatrix} \tag{76}$$

where A_x and A_y are independent complex-valued amplitudes, defining the state of polarization. We can present some simple polarization states, for instance: horizontal linear polarization $J = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, vertical linear polarization $J = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, left circular polarization $J = \begin{bmatrix} 1 \\ i \end{bmatrix}$

D.2   Jones matrices and Manipulating polarization states

How can we manipulate the polarization state of a light beam? In reality, we will use optical elements:

- *Polarizer*: setting the amplitude of one component to 0.

- *Wave plate*: introducing a phase shift between two components by birefringence.

Mathematically, we can describe the changing of 2×1 Jones vector by 2×2 Jones matrices.

Polarizers: A horizontal or vertical polarizer transmits the corresponding field component and blocks the other. The Jones matrices can be described as:

$$\text{Horizontal polarizer:} \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_x \\ A_y \end{bmatrix} = \begin{bmatrix} A_x \\ 0 \end{bmatrix} \tag{77}$$

$$\text{Vertical polarizer:} \quad \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A_x \\ A_y \end{bmatrix} = \begin{bmatrix} 0 \\ A_y \end{bmatrix} \tag{78}$$

In general, a linear polarized field with field amplitude E_0 can be represented as:

$$\begin{bmatrix} A_x \\ A_y \end{bmatrix} = E_0 \begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} \tag{79}$$

The intensity of an electric field is given by the square of its magnitude:

$$I = |\mathbf{E}^2| = |E_x^2| + |E_y^2| + |E_z^2| \tag{80}$$

Before a linear polarized field goes through a polarizer the intensity, the intensity can be expressed as:

$$I_{\text{in}} = |\mathbf{E}^2| = E_0^2 \tag{81}$$

If we send a linear polarized field through a horizontal polarizer, the amplitude of the y-component becomes 0, so the intensity will become:

$$I_{\text{out}} = E_0^2 \cos^2(\theta) = I_0 \cos^2(\theta) \tag{82}$$

This is the Malus's law.

Wave plates Wave plates contain birefringent materials that introduce different refractive indices on each component of the electromagnetic field, meaning that when light propagates through such a medium with the same thickness d, the two field components E_x and E_y see different phase shifts e^{iknd} described by the Jones matrix:

$$\begin{bmatrix} e^{ikn_x d} & 0 \\ 0 & e^{ikn_y d} \end{bmatrix} \begin{bmatrix} A_x \\ A_y \end{bmatrix} \tag{83}$$

What we care about is the phase difference between A_x and A_y, so we can ignore the global phase factoring out from the Jones matrix as it is irrelevant:

$$e^{ikn_x d} \begin{bmatrix} 1 & 0 \\ 0 & e^{ik(n_x - n_y)d} \end{bmatrix} \longrightarrow \begin{bmatrix} 1 & 0 \\ 0 & e^{ik\Delta n d} \end{bmatrix} \tag{84}$$

Setting values for k, Δn and d can be done by changing the birefringent materials or the thickness d. For example, to convert a linearly polarized light to circularly polarized light, we can use this Jones matrix:

$$\text{Quarter wave plate:} \quad \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ i \end{bmatrix} \tag{85}$$

We called such wave plate the *quarter wave plate* because $i = e^{i\frac{\pi}{2}}$ where $\frac{\pi}{2}$ is a quarter of 2π, analogous to rotating a quarter of a circle or shifting a quarter of one full wave oscillation.

Similarly, a *half-wave plate* introduces a phase of π, which has been used in the new generation of CMB experiments as a polarization modulator:

$$\text{Half wave plate:} \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \tag{86}$$

So the half-wave plate mirrors the polarization vector about the x-axis. By rotating the wave plate, the polarization vector can be mirrored around any axis, so a half-wave plate can be used to arbitrarily rotate the direction of linearly polarized light without loss of intensity.

D.3 Stokes parameters

D.3.1 Coherency

In a very famous Young's double-slit experiment which shows the wave nature of light, people usually use a laser beam which is monochromatic light instead of using ordinary sunlight or a normal flashlight. Why? Because those light sources do not generate any interference fringes (or not as clearly as the case of the laser beam). This fact leads us to the conclusion that not all light reveals its polarized nature, for some types of light, no matter which wave plate or polarizer you use, in all cases 50% of all light will be transmitted.

The polarization depends on the degree of coherence, the more coherent (like a laser beam) the more possibility that interference can happen. This could be explained by phase fluctuation and because the phase difference varies rapidly with time, we can only observe the time-average. Mathematically, we describe the degree of coherence using time-average correlation between two field components:

$$\langle E_x(t)^* E_y(t) \rangle \tag{87}$$

By taking the complex conjugate of E_x, we are looking at the phase difference between E_x and E_y.

$$E_x(t)^* E_y(t) = |E_x(t)||E_y(t)|e^{i(\phi_x(t)-\phi_y(t))} = |E_x(t)||E_y(t)|e^{i(\Delta\phi)} \tag{88}$$

If the phase difference $\Delta\phi$ fluctuates randomly, the time average of the correlation converges to 0, so this quantity can be used to measure the degree of polarization. If one of the components equals 0 the time-average also returns to 0, but this is simply a linearly polarized state. To keep track of all the correlations of the field, we define a *coherency matrix* G:

$$G = \left\langle \begin{bmatrix} E_x(t) \\ E_y(t) \end{bmatrix} \begin{bmatrix} E_x^*(t) & E_y^*(t) \end{bmatrix} \right\rangle = \begin{bmatrix} \langle E_x^2(t) \rangle & \langle E_x(t)E_y^*(t) \rangle \\ \langle E_x^*(t)E_y(t) \rangle & \langle E_y^2(t) \rangle \end{bmatrix} \tag{89}$$

Note that the field components matrix can be related to the Jones matrix:

$$\begin{bmatrix} E_x(t) \\ E_y(t) \end{bmatrix} = \begin{bmatrix} A_x \\ A_y \end{bmatrix} e^{-i\omega t} = \mathbf{J}e^{-i\omega t} \tag{90}$$

Therefore, the coherency matrix G can be described as:

$$G = \mathbf{J}\mathbf{J}^\dagger \tag{91}$$

Some examples of the coherency matrices of different states of polarization:

- Horizontally linearly polarized light: $\mathbf{J} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \longrightarrow G_{\text{horizontal}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

- Diagonally linearly polarized light: $\mathbf{J} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \longrightarrow G_{\text{diagonal}} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

- Right circularly polarized light: $\mathbf{J} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix} \longrightarrow G_{\text{circular}} = \frac{1}{2} \begin{bmatrix} 1 & i \\ -i & 1 \end{bmatrix}$

- Unpolarized light: $G_{\text{unpolarized}} = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

D.3.2 Stokes parameters

From those examples, we can see that in general the diagonal elements are positive, real-valued and the sum of them is the total average intensity. The off-diagonal elements are complex-valued and are each other's complex conjugates. Therefore, with the average intensity I, we can write an arbitrary coherency matrix as:

$$G = \frac{1}{2} \begin{bmatrix} I+Q & U+iV \\ U-iV & I-Q \end{bmatrix} = \frac{1}{2}\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} I + \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} Q + \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} U + \begin{bmatrix} 0 & i \\ -i & 0 \end{bmatrix} V \right) \tag{92}$$

Those four matrices are called *Pauli matrices*, the coefficients I, Q, U, V are known as *Stokes parameters* which describe the polarization state of light. They can be put into a *Stokes vector* S:

$$\begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} \tag{93}$$

We can rewrite the transformation of the Stokes vectors by expanding the Pauli matrices as:

$$\begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & i & -i & 0 \end{bmatrix} \begin{bmatrix} \langle E_x^2(t) \rangle \\ \langle E_x(t) E_y^*(t) \rangle \\ \langle E_x^*(t) E_y(t) \rangle \\ \langle E_y^2(t) \rangle \end{bmatrix} = \begin{bmatrix} \langle E_x^2 \rangle + \langle E_y^2 \rangle \\ \langle E_x^2 \rangle - \langle E_y^2 \rangle \\ 2E_x E_y \cos(\phi) \\ 2E_x E_y \sin(\phi) \end{bmatrix} \tag{94}$$

Some examples of how Stokes parameters are used to define polarization states:

$$G_{\text{horizontal}} \longrightarrow S = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad G_{\text{diagonal}} \longrightarrow S = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad G_{\text{circular}} \longrightarrow S = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \quad G_{\text{unpolarized}} \longrightarrow S = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\tag{95}$$

As you can see, each parameter characterizes a state: Q for a horizontal linearly polarized state, U for a diagonal linearly polarized state, and V for circularly polarized state.

D.3.3 Poincaré Sphere

In practice, we can measure the first component of the Stokes vector - the intensity I. The three other Stokes parameters Q, U and V can be interpreted geometrically as the coordinates of a point in the Poincare sphere. If the light is fully polarized, then this point lies on the surface of the sphere with radius I. If the light is fully unpolarized, this point lies at the origin. The length of the vector $\begin{bmatrix} Q \\ U \\ V \end{bmatrix}$ is expressed as $\sqrt{Q^2 + U^2 + V^2}$ which is a measure of the degree of polarization.
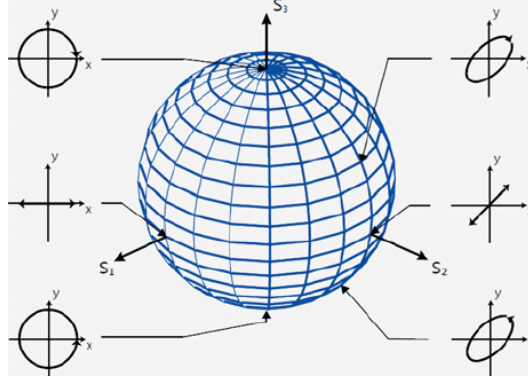

Figure 20: The Poincaré sphere illustrates all possible polarization states. Each point on the sphere corresponds to a specific polarization state: linear polarizations are located on the equator, while elliptical polarizations occupy the remaining surface. *Source:* Kunnen et al. (2015).

E   Full form of muller matrices

E.1   TOD

$$d_t = \frac{1}{2} \underbrace{\begin{bmatrix} 1 & \cos(2\theta) & \sin(2\theta) & 0 \end{bmatrix}}_{\text{rotating grid}} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos(4\varphi_t) & \sin(4\varphi_t) & 0 \\ 0 & \sin(4\varphi_t) & -\cos(4\varphi_t) & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}}_{\text{rotating HWP}} \underbrace{\begin{bmatrix} M_{II} & M_{IQ} & M_{IU} & M_{IV} \\ M_{QI} & M_{QQ} & M_{QU} & M_{QV} \\ M_{UI} & M_{UQ} & M_{UU} & M_{UV} \\ M_{VI} & M_{VQ} & M_{VU} & M_{VV} \end{bmatrix}}_{\text{IPCP}} \underbrace{\begin{bmatrix} I_t \\ Q_t \\ U_t \\ V_t \end{bmatrix}}_{\text{input}} + \underbrace{n_t}_{\text{noise}}$$

(96)

$$d_t = \frac{1}{2} \begin{bmatrix} 1 & \cos(4\varphi_t - 2\theta) & \sin(4\varphi - 2\theta) & 0 \end{bmatrix} \begin{bmatrix} M_{II} & M_{IQ} & M_{IU} & M_{IV} \\ M_{QI} & M_{QQ} & M_{QU} & M_{QV} \\ M_{UI} & M_{UQ} & M_{UU} & M_{UV} \\ M_{VI} & M_{VQ} & M_{VU} & M_{VV} \end{bmatrix} \begin{bmatrix} I_t \\ Q_t \\ U_t \\ V_t \end{bmatrix} + n_t$$

(97)

We introduce the matrix $\mathcal{A} = \begin{bmatrix} A_1 & A_2 & A_3 & A_4 \end{bmatrix}$ so that:

$$d_t = \frac{1}{2} \begin{bmatrix} A_1 & A_2 & A_3 & A_4 \end{bmatrix} \begin{bmatrix} I_t \\ Q_t \\ U_t \\ V_t \end{bmatrix} + n_t$$

(98)

where

$$A_1 = M_{II} + M_{QI}\cos(4\varphi_t - 2\theta) + M_{UI}\sin(4\varphi_t - 2\theta) \tag{99}$$

$$A_2 = M_{IQ} + M_{QQ}\cos(4\varphi_t - 2\theta) + M_{UQ}\sin(4\varphi_t - 2\theta) \tag{100}$$

$$A_3 = M_{IU} + M_{QU}\cos(4\varphi_t - 2\theta) + M_{UU}\sin(4\varphi_t - 2\theta) \tag{101}$$

$$A_4 = {}^{\text{`}}M_{IV} + M_{QV}\cos(4\varphi_t - 2\theta) + M_{UV}\sin(4\varphi_t - 2\theta) \tag{102}$$

We can express them in full form as:

$$A_1 = \frac{g_1^2 + g_2^2}{2} + \frac{g_1^2 - g_2^2}{2}(1 - 2\epsilon)\cos(4\psi_t - 2\theta) + (g_1^2 - g_2^2)\sqrt{\epsilon(1-\epsilon)}\cos\psi\sin(4\psi_t - 2\theta) \tag{103}$$

$$A_2 = \frac{g_1^2 - g_2^2}{2} + \frac{g_1^2 + g_2^2}{2}(1 - 2\epsilon)\cos(4\psi_t - 2\theta) + (g_1^2 + g_2^2)\sqrt{\epsilon(1-\epsilon)}\cos\psi\sin(4\psi_t - 2\theta) \tag{104}$$

$$A_3 = -2g_1 g_2\sqrt{\epsilon(1-\epsilon)}\cos(\phi + \psi)\cos(4\varphi_t - 2\theta) + g_1 g_2(\cos(\phi) - 2\epsilon\cos(\psi)\cos(\phi + \psi))\sin(4\varphi_t - 2\theta) \tag{105}$$

$$A_4 = -2g_1 g_2\sqrt{\epsilon(1-\epsilon)}\sin(\phi + \psi)\cos(4\varphi_t - 2\theta) + g_1 g_2(\sin(\phi) - 2\epsilon\cos(\psi)\sin(\phi + \psi))\sin(4\varphi_t - 2\theta) \tag{106}$$

E.2 Inverse matrices of instrumental Mueller matrices

$$\mathbf{M}_{\text{IP}}^{-1} = \begin{bmatrix} \frac{1}{2}\left(\frac{1}{g_1^2} + \frac{1}{g_2^2}\right) & \frac{1}{2}\left(\frac{1}{g_1^2} - \frac{1}{g_2^2}\right) & 0 & 0 \\ \frac{1}{2}\left(\frac{1}{g_1^2} - \frac{1}{g_2^2}\right) & \frac{1}{2}\left(\frac{1}{g_1^2} + \frac{1}{g_2^2}\right) & 0 & 0 \\ 0 & 0 & \frac{\cos\phi}{g_1 g_2} & -\frac{\sin\phi}{g_1 g_2} \\ 0 & 0 & \frac{\sin\phi}{g_1 g_2} & \frac{\cos\phi}{g_1 g_2} \end{bmatrix} \tag{107}$$

$$\mathbf{M}_{\text{CP}}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 - 2\epsilon & 2\sqrt{\epsilon(1-\epsilon)}\cos\psi & 2\sqrt{\epsilon(1-\epsilon)}\sin\psi \\ 0 & -2\sqrt{\epsilon(1-\epsilon)}\cos\psi & 1 - 2\epsilon\cos^2\psi & -\epsilon\sin(2\psi) \\ 0 & -2\sqrt{\epsilon(1-\epsilon)}\sin\psi & -\epsilon\sin(2\psi) & 1 - 2\epsilon\sin^2\psi \end{bmatrix} \tag{108}$$

$$\mathbf{M}_{\text{IPCP}}^{-1} = (\mathbf{M}_{\text{CP}}\mathbf{M}_{\text{IP}})^{-1} = \mathbf{M}_{\text{IP}}^{-1}\mathbf{M}_{\text{CP}}^{-1} =$$

$$\begin{bmatrix} \frac{1}{2}\left(\frac{1}{g_1^2} + \frac{1}{g_2^2}\right) & \frac{1 - 2\epsilon}{2}\left(\frac{1}{g_1^2} - \frac{1}{g_2^2}\right) & \sqrt{\epsilon(1-\epsilon)}\left(\frac{1}{g_1^2} - \frac{1}{g_2^2}\right)\cos\psi & 2\sqrt{\epsilon(1-\epsilon)}\left(\frac{1}{g_1^2} - \frac{1}{g_2^2}\right)\sin\psi \\ \frac{1}{2}\left(\frac{1}{g_1^2} - \frac{1}{g_2^2}\right) & \frac{1 - 2\epsilon}{2}\left(\frac{1}{g_1^2} + \frac{1}{g_2^2}\right) & \sqrt{\epsilon(1-\epsilon)}\left(\frac{1}{g_1^2} + \frac{1}{g_2^2}\right)\cos\psi & 2\sqrt{\epsilon(1-\epsilon)}\left(\frac{1}{g_1^2} + \frac{1}{g_2^2}\right)\sin\psi \\ 0 & -\frac{2\sqrt{\epsilon(1-\epsilon)}\cos(\psi + \phi)}{g_1 g_2} & \frac{2\epsilon\sin\psi\sin(\phi + \psi) + (1 - 2\epsilon)\cos\phi}{g_1 g_2} & \frac{-2\epsilon\sin\psi\cos(\phi + \psi) - \sin\phi}{g_1 g_2} \\ 0 & -\frac{2\sqrt{\epsilon(1-\epsilon)}\sin(\psi + \phi)}{g_1 g_2} & \frac{2\epsilon\sin\psi\cos(\phi + \psi) + (1 - 2\epsilon)\sin\phi}{g_1 g_2} & \frac{-2\epsilon\sin\psi\sin(\phi + \psi) + \cos\phi}{g_1 g_2} \end{bmatrix} \tag{109}$$