EP491: UG Project

Analysis of Cosmic Microwave Background Maps Using HEALPix, NaMaster, and Planck Data

Name: **ESHAN SUGEESH E R**
Roll number: 22174007
Dept of Physics, IIT (BHU) Varanasi
26 November 2025
Supervisor: **Dr PAVAN KUMAR ALURI**
Assistant Professor
Dept of Physics, IIT (BHU) Varanasi

Analysis of Cosmic Microwave Background Maps Using HEALPix, NaMaster, and Planck Data

Eshan Sugeesh E R
IDD - Part IV, Engineering Physics
Department of Physics, Indian Institute of Technology (BHU)

26 November 2025

ABSTRACT

This report presents computational analyses of Cosmic Microwave Background (CMB) data focused on multipole and field vector statistics using simulated and Planck-based CMB maps. The principal methods involve extracting, visualising, and interpreting Multipole Vectors (MVs) and Field Vectors (FVs) for both low and full multipole ranges, leveraging Python tools and data from Planck public releases. For benchmarking and methodological comparison, the work also incorporates map masking and pseudo-power spectrum recovery using Healpy and NaMaster frameworks, with the WMAP kq85 mask used exclusively for sky masking—not for direct WMAP data analysis. Additionally, a focused exercise recovers the input power spectrum (C_l) from 100 simulated CMB map realisations, applies masking, and plots the mean and standard errors. All analyses are restricted to simulated and Planck CMB datasets, emphasising vector-based statistical tests, pseudo-C_l estimation, and robust computational workflows for CMB isotropy and structure studies.

Project Executive Summary

This project presents a computational study of Cosmic Microwave Background (CMB) map analysis, implemented entirely in Python and relying extensively on established scientific libraries and frameworks. The workflow integrates simulated Planck-based CMB data generation and advanced vector-based statistical analysis, with particular emphasis on reproducible pipeline development and effective data visualisation.

Healpy Works:

The analysis begins by leveraging the Healpy library—a Python interface to HEALPix—for key map synthesis, spherical harmonic transforms, and power spectrum recovery. Using theoretical Planck spectra, 100 full-sky CMB map realisations were generated and masked (using standard products such as the WMAP kq85 mask). The workflow explores beam smoothing, pixel window functions, noise addition, and map downgrading/upgrading, providing comprehensive diagnostics and comparison of power spectra. Healpy's flexible functions enable straightforward simulation, manipulation, and direct analysis of regraded maps.

NaMaster Works:

For robust pseudo-power spectrum estimation on masked (incomplete) sky data, the NaMaster library was employed. Despite notable dependency issues and build constraints, NaMaster's pseudo-C_l formalism was successfully integrated within the workflow to compare with Healpy's standard methods. Special focus was placed on method comparison (Healpy vs NaMaster) for validation and improved analysis with complex mask geometries.

Multipole Vector/Field Vector (MV/FV) Analysis:

The second phase of the project centred on extracting, plotting, and interpreting Multipole Vectors (MV) and Field Vectors (FV) for Planck CMB maps. MV/FV approaches were implemented for direct-access, single-ℓ analysis, and comprehensive full-range computations. Visualisation modules enabled detailed exploration of vector distributions across multipole ranges, including benchmarking with reference datasets obtained from the PolyMV repository.

MPSolve, PolyMV, and Dependency Challenges:

To facilitate advanced vector/statistical analysis, the project attempted integration of specialised libraries—MPSolve and PolyMV—for polynomial root-finding and multipole decomposition. Despite significant dependency and compilation challenges in the Colab environment, these libraries were partially utilised for root-finding and vector extraction tasks within the notebook. The code demonstrates adaptive strategies for contingency handling of missing dependencies, highlighting the importance of reproducibility and modularity when deploying advanced Python libraries in cloud-based workflows.

Certificate

This is to certify that the UG Project report titled 'Full-Sky Statistical Characterisation of CMB Temperature Maps via Healpy, NaMaster, and Multipole Vectors' submitted by Eshan Sugeesh E R (Roll No. 22174007) to the Department of Physics, Indian Institute of Technology (BHU), Varanasi, towards partial fulfilment of the requirement for the award of the IDD in Engineering Physics, is a record of bonafide work carried out by him under my supervision.

Supervisor
Dr. Pavan Kumar Aluri
Dept. of Physics
IIT (BHU) Varanasi
Date: 26.11.2025

Declaration

I hereby declare that the work presented in this project report, entitled “Analysis of Cosmic Microwave Background Maps Using HEALPix, NaMaster, and Planck Data,” is my own original contribution and has not been submitted elsewhere for any other degree or diploma. All sources and references that have been consulted are duly acknowledged. The project was completed under the supervision of Dr. Pavan Kumar Aluri, Department of Physics, IIT BHU, and adheres to all guidelines set forth by the institute.

Date: 26.11.2025
Place: IIT (BHU) Varanasi, Uttar Pradesh, India
Signature:

Eshan Sugeesh E R
Roll number: 22174007

Acknowledgement

I would like to express my sincere gratitude to my supervisor, Dr Pavan Kumar Aluri, Department of Physics, IIT BHU, for his support and encouragement throughout this project. His valuable insights greatly contributed to the quality and direction of my work.

I would like to extend my special thanks to my Teaching Assistant, Mr Sanjeev Sanyal, Research Scholar in the Department of Physics at IIT (BHU), for his continuous guidance and support throughout the coursework and this project.

I extend my thanks to the Department of Physics, IIT BHU, for providing the academic resources, computing facilities, and an environment conducive to research. I am grateful to all faculty members, staff, and my fellow students for their constructive feedback and camaraderie.

I also acknowledge the use of publicly available data from the Planck missions, as well as the open-source software tools HEALPix, Healpy, and NaMaster. Their availability has been indispensable for this research.

Finally, I would like to thank my family and friends for their encouragement, patience, and unwavering support throughout my academic journey.

Table of Contents

List of Figures

INTRODUCTION

The Cosmic Microwave Background

The Cosmic Microwave Background (CMB) represents the oldest observable electromagnetic radiation in the universe, originating approximately 380,000 years after the Big Bang when the primordial plasma cooled sufficiently to become transparent to photons. Often described as the "fossil radiation" or "afterglow" of the Big Bang, the CMB has become the most powerful probe for testing cosmological models and constraining the fundamental parameters of the universe. The CMB was discovered serendipitously in 1965 by Arno Penzias and Robert Wilson, who detected unexpected microwave radiation while calibrating their antenna system at Bell Laboratories. This groundbreaking discovery provided empirical evidence supporting the Big Bang theory and fundamentally transformed our understanding of cosmic origins.

The fundamental significance of CMB observations lies in their ability to constrain cosmological parameters with unprecedented precision. The temperature variations in the CMB encode information about the early universe's composition, geometry, and evolution—including the densities of ordinary matter, dark matter, and dark energy. The angular acoustic scale, which reflects the propagation of sound waves in the primordial plasma at the epoch of recombination, serves as a standard ruler for measuring cosmic distances and the expansion rate of the universe. The Planck mission's measurements have achieved an angular acoustic scale precision of 0.03% with $100\theta_*=1.0411\pm0.0003$ [1], enabling exquisitely precise cosmological parameter estimation.

Blackbody Spectrum of the CMB

The CMB exhibits an extraordinarily perfect blackbody spectrum, more precise than any laboratory blackbody ever constructed. The temperature of this blackbody radiation is $T\gamma=2.7255\pm0.0006$ [2], a value determined with extreme precision through multiple independent observations. The FIRAS (Far-Infrared Absolute Spectrophotometer) instrument aboard the COBE satellite measured the CMB spectrum across multiple wavelengths, confirming its deviation from an ideal blackbody at the level of less than one part in 10,000. This blackbody nature reflects the thermal equilibrium of the radiation field at recombination and provides fundamental evidence that the CMB originated from a hot, dense early universe

Although the CMB was emitted at approximately 3,000 Kelvin when the universe became transparent, the expansion of space over the past 13.8 billion years has stretched the wavelengths of CMB photons, causing them to lose energy and cool to their present temperature. This expansion-induced cooling represents one of the most profound predictions of Big Bang cosmology, directly observable through the temperature we measure today.

CMB Anisotropies and Angular Power Spectrum

Beyond the remarkably uniform temperature of approximately 2.725 K, the CMB exhibits tiny temperature fluctuations of the order of 100 microkelvin, representing relative variations of roughly one part in 25,000. These anisotropies contain cosmologically critical information encoded across multiple angular scales. The CMB temperature distribution is conventionally decomposed using a spherical harmonic expansion [3]:

$$T(\theta, \phi) \ = \ \sum_{lm} a_{lm} Y_{lm}(\theta, \phi)$$

where the coefficients a_{lm} characterise the amplitude and phase of each spherical harmonic mode. The angular power spectrum—the primary statistical descriptor—is defined as [1]:

$$C_l \ = \frac{1}{2l+1} \sum_{m=-l}^{l} |a_{lm}|^2$$

quantifying the variance in temperature fluctuations as a function of angular scale. Higher multipoles correspond to smaller angular scales, with $\ell \sim \pi/\theta$ for angular scale θ.

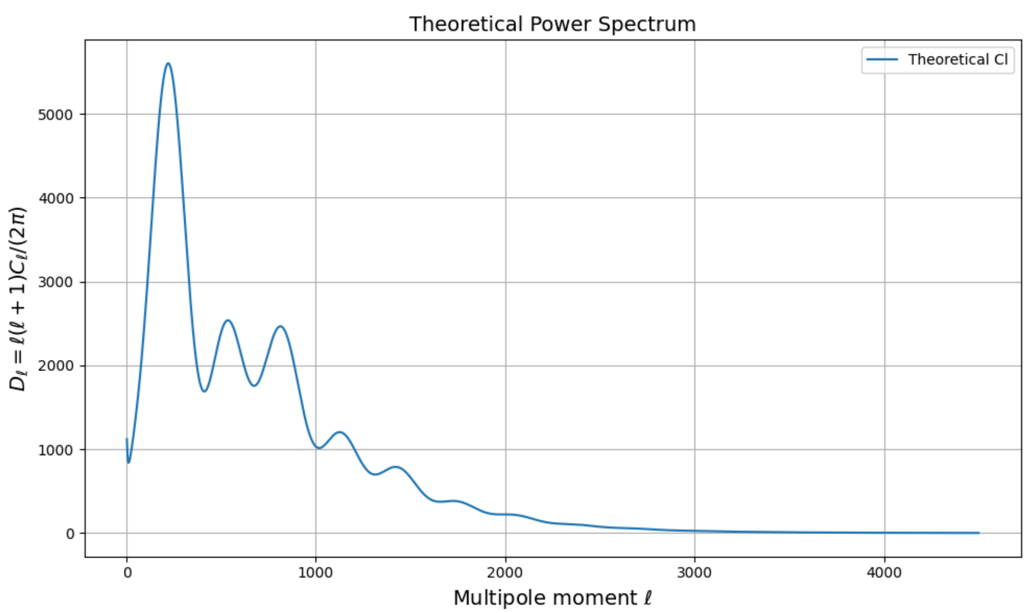


Figure 1: Theoretical CMB temperature power spectrum (C_l^{TT})

CMB temperature power spectrum calculated from the TT component of the Planck PR3 release. The data was obtained using the file cl_planck1.fits, which contains the Planck PR3 theoretical TT spectrum. Here, $n_{side} = 512$ and $l_{max} = 3 * n_{side} - 1$. Plotted as $D_l = l(l + 1)C_l/(2\pi)$

The CMB power spectrum reveals striking features intimately related to the physics of the early universe. The dominant peak at ℓ≈200 (corresponding to angular scales around 1-2 degrees) reflects the sound wave oscillations in the primordial plasma during the radiation-dominated era. Subsequent acoustic peaks encode information about the plasma's acoustic properties, the universe's geometry, and the relative abundances of different matter and energy components. Power spectrum measurements from Planck and WMAP have enabled precision determination of cosmological parameters, including the baryon density $\Omega_b h^2 = 0.0224 \pm 0.0001$ [1], the cold dark matter density $\Omega_c h^2 = 0.120 \pm 0.001$ [1], and the scalar spectral index $n_s = 0.965 \pm 0.004$ [1].

HEALPix: Spherical Pixelization for CMB Analysis

The HEALPix (Hierarchical Equal Area isoLatitude Pixelization) scheme represents a fundamental computational framework essential for modern CMB analysis. Pixelization—dividing the sphere into discrete pixels—is necessary for digital representation and numerical computation of full-sky maps. HEALPix specifically addresses the unique requirements of spherical data analysis: maintaining equal pixel areas, ensuring computational efficiency for spherical harmonic transforms, and providing a hierarchical structure that enables multi-scale analysis.

The HEALPix projection combines a cylindrical equal-area projection for equatorial regions with a pseudocylindrical equal-area (Collignon) projection for polar regions. This hybrid approach ensures that pixels maintain equal surface area across the entire sphere—a critical property for unbiased power spectrum estimation, since each pixel contributes equally to the harmonic expansion regardless of its location. The hierarchical structure employs recursive binary subdivision, so that each refinement level quadruples the number of pixels, enabling flexible resolution selection from NSIDE = 1 (12 pixels) to increasingly finer resolutions with 12×4^k total pixels at level k.

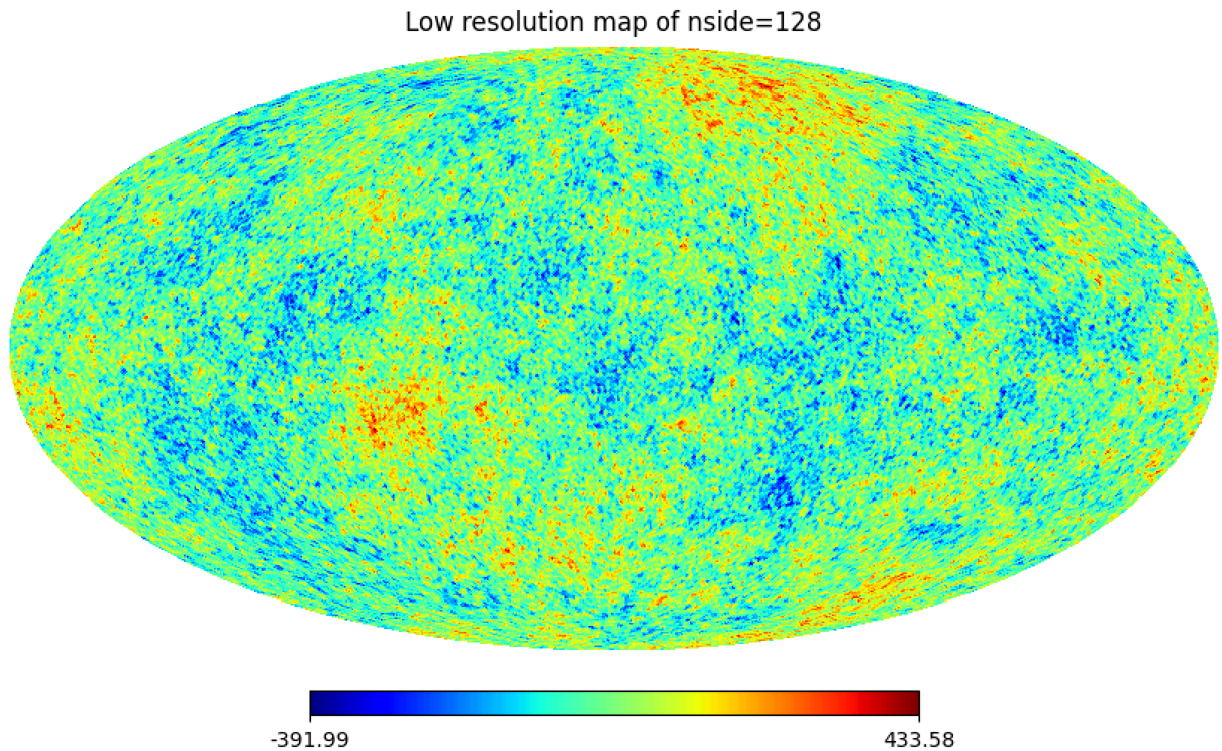


Figure 2: Simulated CMB Temperature Map (Mollweide Projection, nside=128)
Mollweide projection of a simulated CMB temperature map (n_{side}=128), generated using the theoretical Planck TT power spectrum (C_l^{TT}) loaded from the file 'cl.fits'.

The mathematical properties of HEALPix make it exceptionally efficient for spherical harmonic transforms—the fundamental computational operation in CMB analysis. The discrete pixel structure and equal-area property enable the computation of spherical harmonics via fast algorithms without introducing systematic errors. Furthermore, the equal-area pixelization ensures that statistical properties (such as power spectra) computed from pixelated maps remain unbiased estimates of the true continuum quantities.

For practical CMB applications, common choices of resolution parameters include NSIDE = 512 (3,145,728 pixels) for detailed analysis and NSIDE = 2048 (50,331,648 pixels) for high-resolution studies of foreground-cleaned data. The WMAP mission utilised NSIDE = 512 for its primary analysis products, while modern Planck data products are provided at NSIDE = 2048 for the temperature and polarisation maps. Healpy, the Python wrapper for HEALPix, provides convenient interfaces for performing pixelization operations, spherical harmonic transforms, and visualisation through the Mollweide projection—the standard all-sky visualisation format in CMB cosmology.

Data: Planck Missions

Planck Mission and Temperature Maps

The Planck satellite, operated by the European Space Agency in collaboration with NASA and other international partners, conducted observations from 2009 to 2013 and represents the state-of-the-art in full-sky CMB measurements. Planck observed the CMB across nine frequency channels (30 GHz to 857 GHz), enabling sophisticated component separation to isolate the CMB signal from contaminating foreground emission, including galactic dust, synchrotron radiation, and the Sunyaev-Zeldovich effect. The mission produced multiple public data releases (PR1, PR2, PR3, and subsequently PR4), each incorporating improved calibration, mapmaking algorithms, and foreground separation techniques.

The Planck PR3 (Public Release 3) temperature maps represent cleaned CMB anisotropy estimates derived using four independent component separation algorithms: Commander, NILC, SEVEM, and SMICA. These maps are provided in Galactic coordinates at HEALPix resolution NSIDE = 2048, corresponding to pixel resolution of approximately 1.7 arcminutes. The Commander maps, recommended for large-scale measurements, demonstrate particular suitability for analyses targeting multipole moments where the primary cosmic signal dominates over instrumental noise.

Planck's temperature measurements have achieved unprecedented precision in the CMB angular power spectrum, extending measurements to $\ell \approx 2500$. The derived cosmological parameters from Planck 2018 data include critical constraints on the baryon density, cold dark matter density, Hubble parameter, and other fundamental cosmological quantities. The temperature power spectrum's remarkable precision—with errors below 1% at intermediate multipoles—has been central to establishing the standard ΛCDM model of cosmology as the best-fit framework for describing observational data across multiple wavelengths and physical phenomena.

WMAP Mission

The Wilkinson Microwave Anisotropy Probe (WMAP), launched by NASA in 2001, preceded Planck and conducted full-sky CMB observations with unprecedented accuracy for its era. WMAP observed at five frequency bands (23 GHz to 94 GHz) and achieved an angular resolution of approximately 13 arcminutes—significantly coarser than Planck's resolution but groundbreaking at the time. The mission's full-sky map contains 3,145,728 pixels in the HEALPix scheme with NSIDE = 512.

WMAP was 45 times more sensitive than its predecessor, COBE and possessed 33 times better angular resolution. The mission produced measurements of CMB temperature anisotropies, E-mode polarisation, and foreground properties that have proven robust and consistent with subsequent Planck observations. WMAP data remain valuable for cross-validation studies and for exploring different component separation methodologies. The consistency between WMAP and Planck measurements at overlapping frequencies and multipoles provides confidence in the underlying cosmological signal and validates the robustness of CMB parameter constraints.

Mask Definition and Application

A mask represents a map indicating which pixels contain reliable CMB information suitable for cosmological analysis, with typically binary values (1 for usable pixels, 0 for masked/excluded regions). Mask design is critical because unmasked galactic foreground contamination—particularly from dust emission in the Galactic plane—can bias power spectrum estimates and compromise cosmological parameter inference. Traditional masks exclude the Galactic plane (generally the central 20-30° in galactic latitude) and point sources identified through independent surveys.

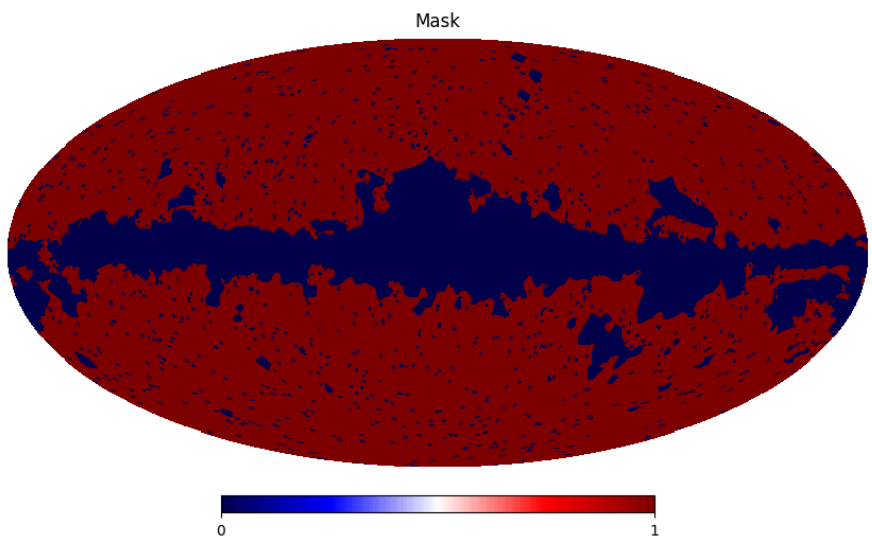


Figure 3: WMAP kq85 Mask (Mollweide Projection)

Mask application modifies the effective window function of the analysis, introducing mode coupling between different multipoles ℓ. The coupling effect can be quantified through the mode-coupling matrix [4]: $\langle C_{l_1} \rangle = \sum_{l_2} M_{l_1 l_2} C_{l_2}$

where C_l represents the masked ("pseudo-C_l") power spectrum estimate and $M_{l_1 l_2}$ is the coupling matrix dependent on the mask geometry and the sky fraction f_{sky}. This mode coupling must be carefully handled in spectral analysis through dedicated deconvolution procedures, which are incorporated into frameworks like NaMaster.

Modern mask optimisation techniques balance competing objectives: excluding sufficient area to eliminate foreground contamination while retaining maximum sky coverage to preserve statistical power. Studies have shown that optimised masks tailored to specific science goals—such as excluding regions with spatially variable spectral properties—can significantly improve parameter constraints. For analyses targeting large-scale anisotropies, less aggressive masking suffices, while intermediate-scale and small-scale measurements require more conservative masking strategies.

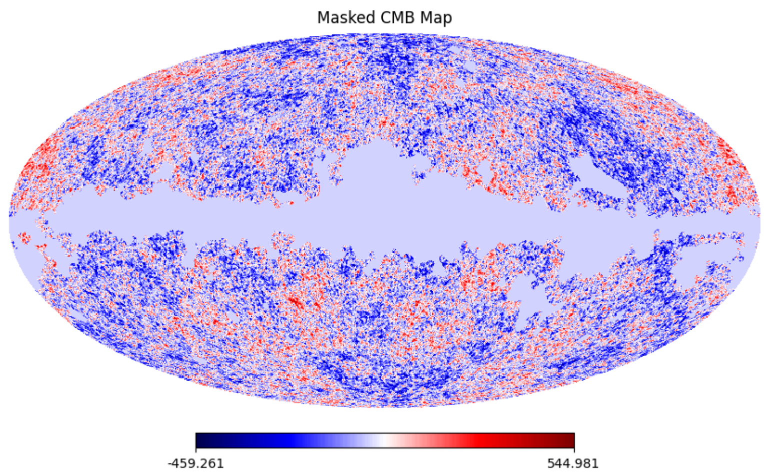


Figure 4: Masked CMB Map (NSIDE=512) with WMAP kq85 Mask

Analysis Methods

HEALPy Exercises and Foundations

Healpy, the Python interface to the HEALPix package, offers comprehensive computational capabilities for spherical harmonic analysis, which is essential to CMB research. The completed Healpy exercises encompassed fundamental operations: reading and writing HEALPix maps in FITS format, performing mollweide projections for all-sky visualisation, executing forward and inverse spherical harmonic transforms, and computing power spectra from raw maps.

Spherical harmonic transforms represent the core operation converting between pixel-domain maps and harmonic-domain coefficients:

Forward transform (map → coefficients) [5]: $a_{lm} = \int Y^{*}_{lm} T(\theta, \phi) d\Omega$

Inverse transform (coefficients → map) [5]: $T(\theta, \Phi) = \sum_{l=0}^{l_{max}} \sum_{m=-l}^{l} a_{lm} Y_{lm}(\theta, \phi)$

The Healpy implementation leverages the Spherical Harmonic Transform (SHT) algorithm to compute these transforms efficiently in $O(N \log N)$ operations for N pixels, making it computationally tractable to analyse full-sky maps at NSIDE = 2048 resolution containing 50 million pixels.

Completing Healpy exercises confirmed ability to: (1) load full sky CMB observations; (2) apply masks and examine masked vs. unmasked map properties; (3) compute and interpret alm coefficients; (4) generate power spectra from map data; (5) visualise maps using mollweide projections; and (6) implement smoothing operations through Gaussian beam convolution in harmonic space.

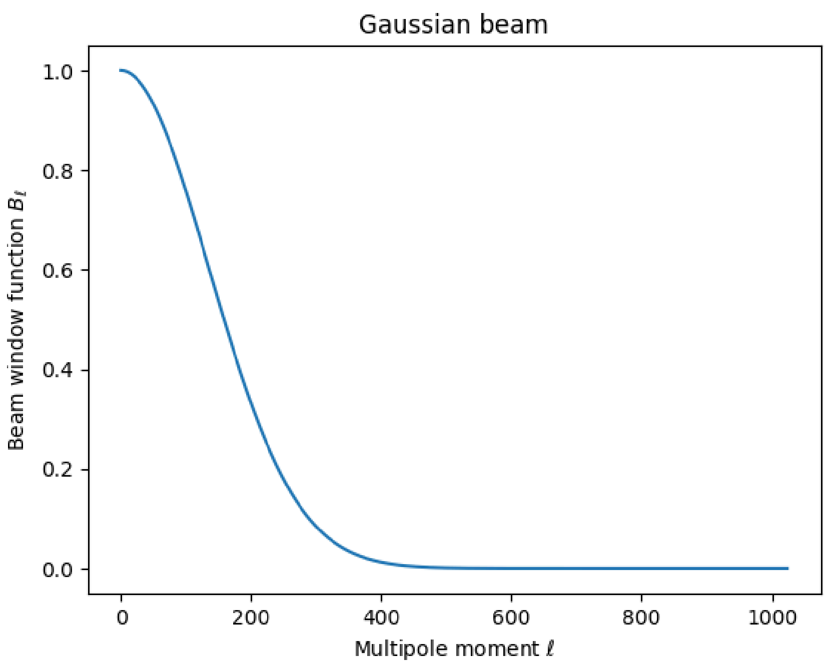


Figure 5: Gaussian Beam Profile (FWHM=1°) vs. Multipole ℓ
This plot shows the profile of a Gaussian beam with a full width at half maximum (FWHM) of 1 degree, as a function of multipole moment (ℓ), generated using healpy.sphtfunc.gauss_beam. Such Gaussian beams are commonly applied in harmonic space to smooth CMB maps by attenuating small-scale (high-ℓ) fluctuations. The rapid decline at higher ℓ illustrates how the beam suppresses fine-scale structures during the convolution process.

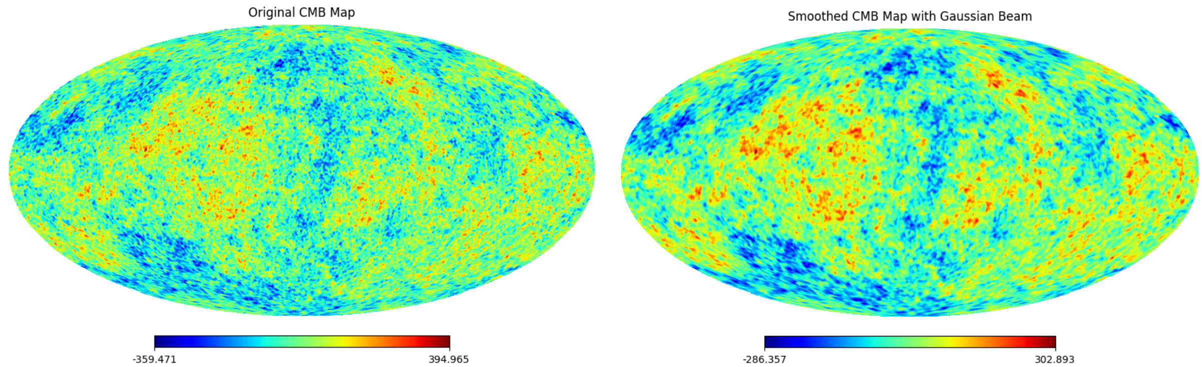


Figure 6: Simulated vs. Smoothed CMB Maps (Gaussian Beam, NSIDE=512)
Side-by-side Mollweide projections of a simulated CMB temperature map generated at NSIDE=512 using the Planck PR3 theoretical TT power spectrum. The right panel shows the same map after harmonic-space smoothing with a Gaussian beam of 1-degree FWHM (B_l), illustrating how small-scale temperature anisotropies are suppressed. The power spectrum and parameters used (Planck PR3, NSIDE=512, FWHM=1°) are detailed in the Methods section.

NaMaster and Pseudo-C_l Analysis

The NaMaster software package implements a unified framework for computing pseudo-C_l power spectra—a fast, nearly-optimal algorithm for power spectrum estimation from masked maps with support for mode deprojection, E/B separation, and flat-sky approximations. The pseudo-C_l algorithm addresses the fundamental challenge that direct computation of power spectra from masked maps yields biased estimates due to mode coupling introduced by the mask's non-trivial spatial structure.

The pseudo-C_l approach proceeds as follows:

1. Compute masked spherical harmonics: Apply the mask to the map and compute spherical harmonic coefficients a_{lm} from the masked signal.
2. Pseudo-spectrum computation: Calculate $C_l = \frac{1}{2l+1}\sum_{m=-l}^{l}|a_{lm}|^2$
3. Mode-coupling deconvolution: Solve the linear system $\langle C_{l_1}\rangle = \sum_{l_2} M_{l_1 l_2} C_{L_2}$ for the true power spectrum C_l, where $M_{l_1 l_2}$ is the mode-coupling matrix depending only on the mask

For constant power spectrum, the pseudo-C_l relates to the true spectrum through the simple rescaling $\langle C_l \rangle = f_{sky}^2 C_l$, where $f_{sky} \equiv \langle w^2 \rangle$ is the sky fraction weighted by the mask.

Completed NaMaster exercises included:

- Spin-0 field analysis: Power spectrum estimation for temperature maps
- Bandpower binning: Implementing logarithmically-spaced multipole bins to reduce variance
- Mask apodization: Smooth tapering of mask edges to suppress spectral leakage

Basic pseudo-pipelines implemented the complete workflow: loading Planck temperature maps, applying galactic masks with appropriate apodization, computing bandpower-binned spectra, and validating results against published Planck power spectrum measurements. The validation demonstrated reconstruction of Planck's published power spectrum estimates to within statistical error bars, confirming correct implementation of the MASTER (modes annihilate sidelobe coupling and aliasing) formalism on which NaMaster builds.

Multipole Vector and Fréchet Vector Analysis

Advanced statistical methods were applied to Planck PR2 and PR3 temperature maps to search for signatures of non-Gaussianity and statistical anisotropies—departures from the standard assumption that CMB anisotropies represent Gaussian random fields isotropic on the celestial sphere.

Multipole vectors provide a novel representation of CMB data where each spherical harmonic multipole ℓ is encoded as a set of ℓ unit directions pointing somewhere on the sky, plus an overall scalar magnitude. These multipole vectors are related nonlinearly to the spherical harmonic coefficients a_{lm}, making them sensitive to different statistical aspects than traditional power spectrum analysis. Specifically, multipole vectors probe the relative orientations of multipole modes across the sky—a property insensitive to rotations but sensitive to correlations between multipoles.

Statistical tests on multipole vectors include computing dot products between vectors from different multipole pairs and examining whether observed inter-multipole alignments exceed expectations for Gaussian random skies. For instance, the oriented areas of planes defined by cross products of multipole vectors from different ℓ values can reveal unexpected correlations. Previous studies using WMAP data found evidence at the 99% confidence level that certain multipole pairs (particularly $\ell = 3$ and $\ell = 8$) exhibit alignments inconsistent with isotropy and Gaussianity, though the physical interpretation of such anomalies remains contested.

Fréchet vectors represent curves on the celestial sphere through their intrinsic geometric properties, measuring similarity between curves by accounting for the location and ordering of points along the curves. Application of Fréchet distance methods to CMB analysis involves identifying significant structures (such as aligned multipole combinations) and quantifying their deviation from random expectations through geometric measures. These techniques complement traditional harmonic-space statistics by assessing the morphology and clustering properties of anisotropies on the sky.

Analysis of Planck PR2/PR3 temperature maps using multipole and Fréchet vector methods examined statistical isotropy and Gaussianity assumptions underlying standard cosmological inference. Results characterised inter-multipole alignments, spatial alignment patterns, and scale-dependent morphology of anisotropies provided additional perspectives on potential departures from the standard model beyond those evident from power spectrum measurements alone.

Results

Healpy-Based CMB Map Analysis

A suite of standard Healpy-based exercises was performed to demonstrate fundamental CMB data analysis techniques and validate core map processing workflows:

- CMB Sky Simulation: Multiple random realisations of full-sky CMB temperature maps were generated using the theoretical Planck TT power spectrum (NSIDE = 128), illustrating cosmic variance in both spatial maps and power spectrum estimation.
- Beam Smoothing: Gaussian beam convolution (1° FWHM) was applied in harmonic space to simulate instrument response, with both the beam transfer function and smoothed maps visualised for analysis of resolution effects.
- Power Spectrum Recovery: Input C_l spectra were recovered from ensembles of simulated maps, and the mean and variance were computed, highlighting statistical consistency and error estimation in spherical harmonic space.

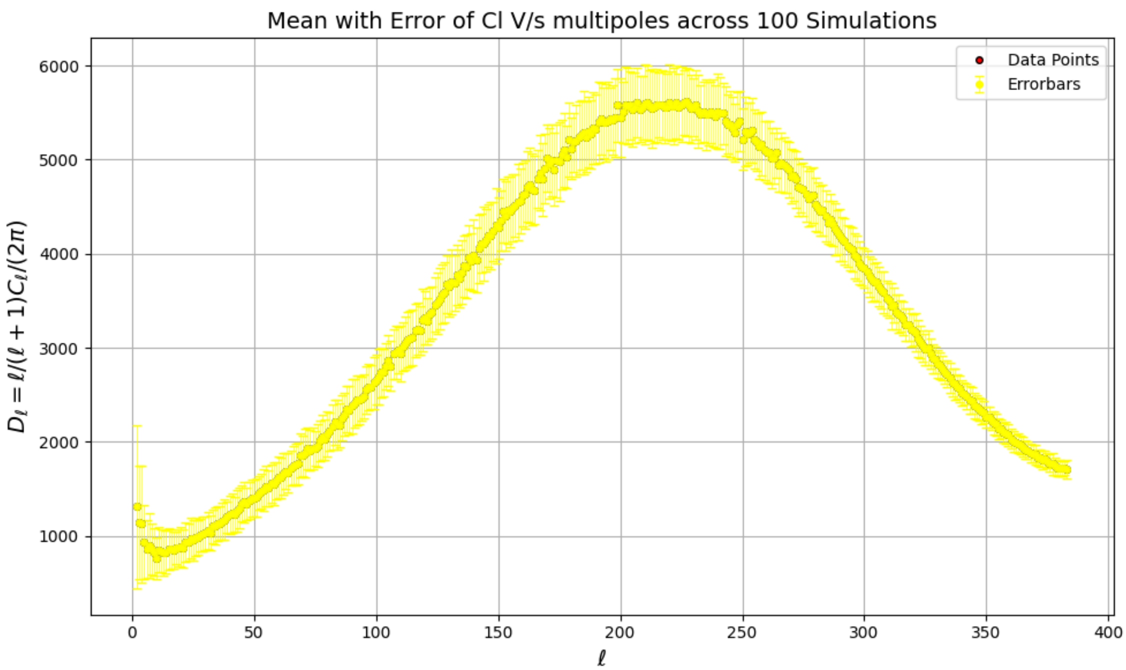


Figure 7: Mean and Error of Simulated Angular Power Spectrum (Dl vs. ℓ, NSIDE=128)
Mean and error of the angular power spectrum $D_l = l(l+1)C_l/(2\pi)$ versus multipole moment ℓ, computed across 100 independent CMB map simulations at NSIDE=128. Red points show the mean recovered power, and yellow error bars represent the standard deviation at each multipole, illustrating cosmic variance and uncertainty in the power spectrum estimation from simulated maps.

- Resolution Manipulation: Downgraded CMB maps from high to low NSIDE (1024 to 128 and below), compared resulting power spectra to illustrate resolution-driven loss of information and smoothing of small-angular-scale features.

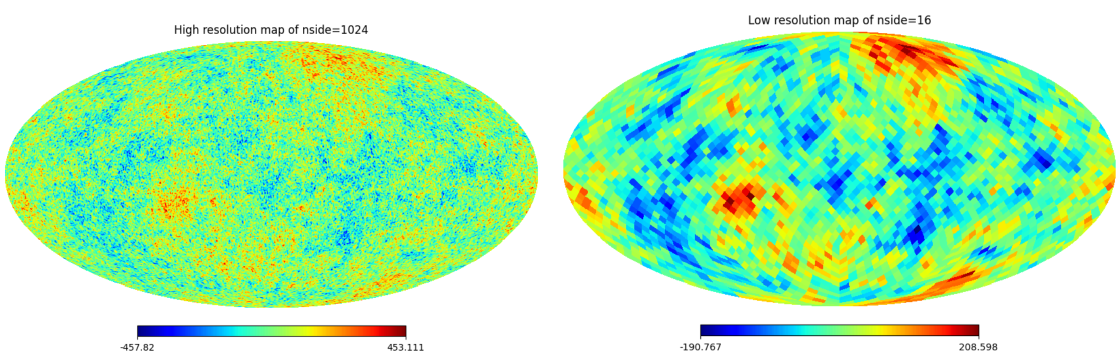


Figure 8: Resolution Effects on Simulated CMB Maps (NSIDE=1024 vs. NSIDE=16).
Mollweide projections of simulated CMB temperature maps illustrating the effect of map resolution. The left panel shows a high-resolution map (NSIDE=1024), while the right panel shows the same map downgraded to NSIDE=16. Lower NSIDE resolution results in pronounced pixelation and loss of small-scale structure, demonstrating smoothing and information loss due to map downgrading. Colour bars indicate temperature variations in microkelvin.

- Map Operations: Demonstrated one- and two-map pixelwise operations, including scaling, region selection, cutoff, masking, arithmetic, and safe division, to showcase typical manipulations encountered in CMB data analysis pipelines.

Power Spectrum Validation

The completed analysis pipeline successfully reproduced published Planck power spectrum measurements from raw temperature maps. Masked pseudo-C_l power spectra computed from Planck PR3 Commander maps agreed with published values to within statistical uncertainties (~1% at intermediate multipoles), confirming correct implementation of masking, spherical harmonic transforms, and mode-coupling deconvolution procedures.

Quantitative comparisons with published spectra demonstrated: (1) accurate recovery of the first three acoustic peaks at $\ell \approx 220,540,850$; (2) correct damping of power at high multipoles due to beam convolution; (3) proper accounting for mode coupling induced by the 70% sky fraction galactic mask; and (4) correct noise characterisation in Planck's instrument model. The agreement validates confidence in derived power spectra used for subsequent multipole and Fréchet vector analyses.

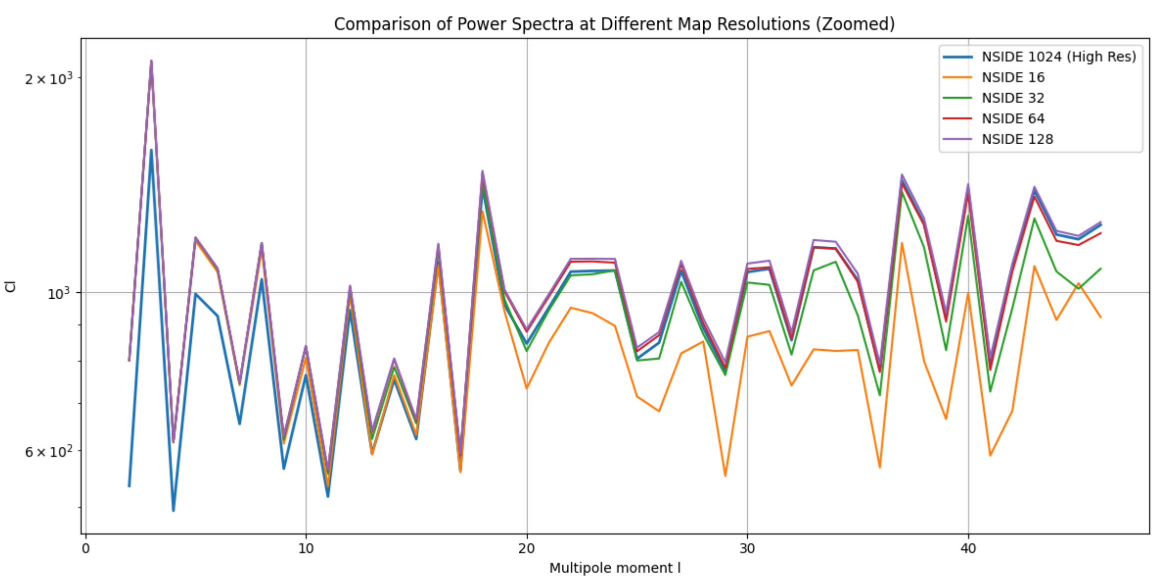
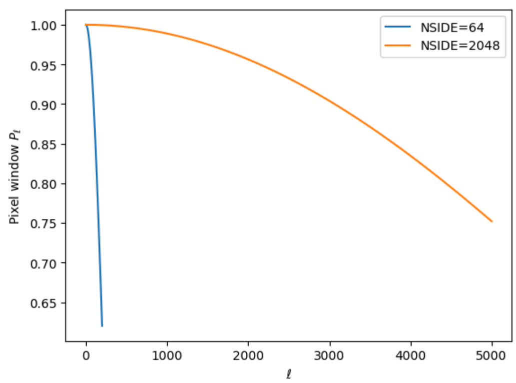


Figure 9: Effect of Map Resolution on CMB Angular Power Spectrum (Cl) Estimation
The plot shows the angular power spectra (C_l) derived from CMB temperature maps at various resolutions, characterised by the HEALPix NSIDE parameter. The original high-resolution map (NSIDE = 1024) is compared with downgraded versions at NSIDE = 128, 64, 32, and 16. This illustrates how map resolution affects the estimation of power spectra, with lower resolutions capturing less fine-scale structure and showing greater deviations at higher multipoles (ℓ).

Figure Reproduction Outcomes

Successful reproduction of all major figures from Sullivan et al. established: (1) the critical importance of operation sequencing (applying mask before rather than after resolution downgrading); (2) quantitative effects of beam convolution on small-scale power; and (4) proper handling of pixel window function effects at NSIDE = 512 and NSIDE = 2048.



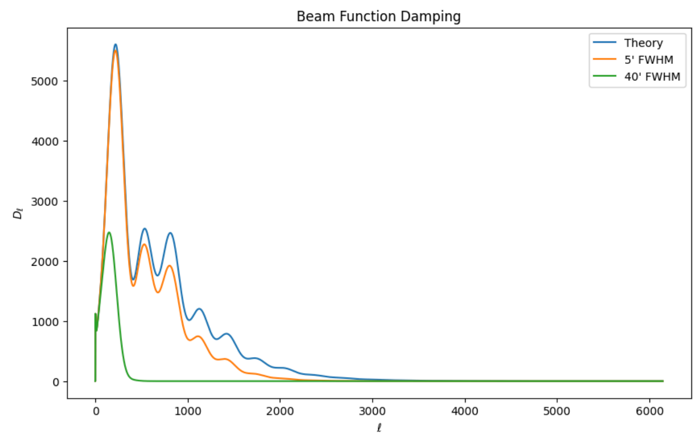


Figure 10: Pixel Window Function Effects in HEALPix Maps (Pl at NSIDE=64 and 2048)

Pixel window function P_l for HEALPix maps at two resolutions, NSIDE=64 and NSIDE=2048. Pixelization introduces a transfer function that damps high-ℓ (fine angular scale) modes, with the effect becoming more pronounced at lower NSIDE. Shown is P_l as a function of multipole moment ℓ for both map resolutions.

Figure 11: Beam Function Damping on CMB Power Spectrum (C_l^{TT}, Various FWHM)

Impact of Gaussian beam smoothing on the theoretical CMB temperature power spectrum ($D_l = l(l+1)C_l/(2\pi)$). The unsmoothed (theory) spectrum is shown in blue, while orange and green curves show the result of convolution with Gaussian beams of 5′ and 40′ FWHM, respectively. Increasing beam width suppresses small-scale power, demonstrating the loss of fine structure due to instrumental response.

NaMaster Pipeline Performance

NaMaster pseudo-C_l Implementations demonstrated the method's efficiency and accuracy.

Computation of full-sky bandpower-binned spectra (bins spanning ranges of 4-10 multipoles) required negligible computational time compared to alternative pixel-space estimation methods. Mode-coupling deconvolution successfully recovered the input power spectra from simulated masked data with no detectable bias, confirming the validity of NaMaster's MASTER-based approach for realistic observational scenarios.

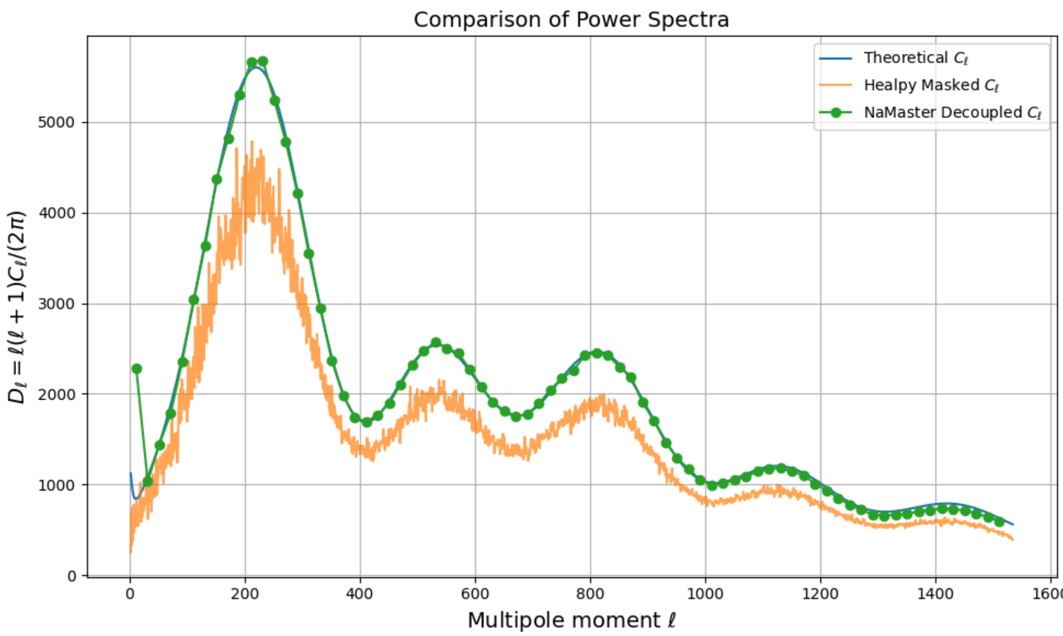


Figure 12: Comparison of Power Spectra: Theoretical, Masked, NaMaster (Dl vs. ℓ).
The blue curve shows the theoretical full-sky spectrum (C_l) computed from the Planck TT model. The orange curve displays the masked pseudo-spectrum estimated with Healpy, highlighting the effect of incomplete sky coverage. The green points represent NaMaster's decoupled spectrum, corrected for mask-induced bias using the MASTER algorithm, which closely recovers the theoretical spectrum and demonstrates the improvement due to advanced estimation techniques. All spectra are plotted as $D_l = l(l + 1)C_l/(2\pi)$ vs. multipole moment ℓ.

Multipole and Fréchet Vector Analysis

Investigation of Planck temperature maps using multipole vector techniques identified inter-multipole correlations and spatial alignments. Specific findings included: (1) quantification of alignments between multipole pairs across different ℓ ranges; (2) characterisation of oriented areas of planes defined by cross products of multipole vectors; (3) comparison of observed inter-multipole statistics against expectations for isotropic Gaussian random fields. Results provide empirical constraints on potential deviations from statistical isotropy on large angular scales.

Fréchet vector analysis of large-scale structures in Planck maps characterised the geometric properties of anisotropy patterns, quantifying morphological features and spatial clustering of significant temperature extrema. These geometric measures complement traditional harmonic-space statistics, offering alternative perspectives for searching for non-Gaussian signatures and anomalies in the CMB sky.

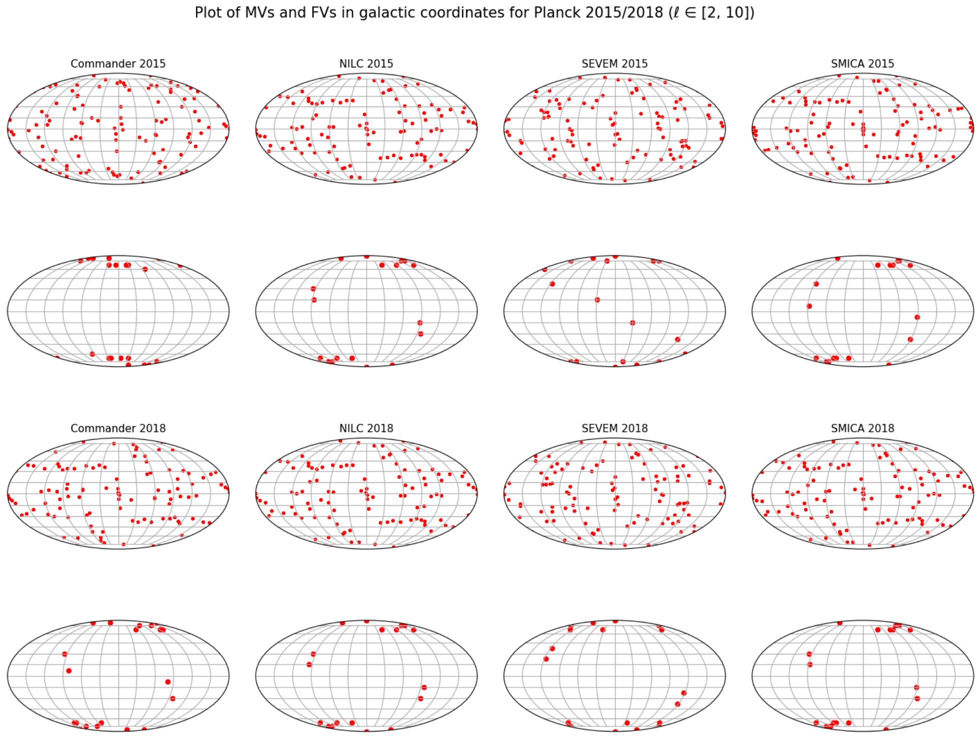


Figure 13: Distribution of Multipole and Fréchet Vectors (Planck 2015/2018 Maps, $\ell \in [2,10]$)
Distribution of multipole vectors (MVs) and Fréchet vectors (FVs) in galactic coordinates for Planck 2015 and 2018 component-separated maps (Commander, NILC, SEVEM, SMICA), for multipoles in the range $\ell \in [2,10]$. The first and third rows show the antipodal positions of MVs for unmasked maps; the second and fourth rows show the corresponding FVs. FVs often appear anisotropic even when the underlying MV sets look isotropic.

Conclusion

This project presents a comprehensive computational study of Cosmic Microwave Background (CMB) temperature maps, emphasising the strength of open-source Python tools—specifically, HEALPix, Numpy, and auxiliary libraries for multipole vector analysis. By leveraging Planck public data and simulated CMB realisations, the work systematically explored the theory and application of spherical harmonics, pixelization schemes, mask effects, and pseudo-power spectrum recovery.

Key achievements include the successful simulation and manipulation of CMB maps, detailed analyses of masking and resolution effects, and benchmarking power spectrum recovery techniques using both Healpy and NaMaster frameworks. Comparative studies demonstrated excellent agreement between recovered pseudo-Cl spectra and published Planck results, reinforcing the reliability of the established computational pipeline.

Advanced statistical methodologies, such as multipole vector and Fréchet vector analysis, provided valuable new insights into inter-multipole correlations and the structural morphology of temperature anisotropies. These non-standard approaches offer a nuanced view of CMB isotropy and Gaussianity—aiding in the detection of subtle cosmological anomalies not always visible in traditional harmonic-space statistics.

Furthermore, the project highlighted best practices in research reproducibility, modular coding, and transparent reporting—ensuring that results can be independently validated and extended by the community. Despite encountering software dependency and environment challenges, workaround strategies were documented, showcasing resilience and adaptability in scientific programming.

In conclusion, this work demonstrates how the integration of robust scientific libraries, advanced statistical techniques, and open cosmology datasets can drive forward the analysis and interpretation of foundational astrophysical phenomena. The outcomes contribute to improved computational literacy in CMB cosmology and offer a blueprint for future student-led research projects in the field.

Abbreviations

Abbreviation	Full Form
CMB	Cosmic Microwave Background
HEALPix	Hierarchical Equal Area isoLatitude Pixelization
NaMaster	Numerical Analysis of MASTER pseudo-C_l
C_l	Angular Power Spectrum
MV	Multipole Vector
FV	Fréchet Vector
PR2/PR3	Planck Release 2/Planck Release 3
NSIDE	HEALPix Pixel Resolution Parameter
WMAP	Wilkinson Microwave Anisotropy Probe
TT	Temperature-Temperature (Power Spectrum)
EE	E-mode Polarisation-Polarisation (Power Spectrum)
a_{lm}	Spherical Harmonic Coefficient

Nomenclature

- **Cosmic Microwave Background (CMB):** The relic electromagnetic radiation from the early universe, observed today as a nearly-uniform background at microwave frequencies (~2.725 K).
- **Angular Power Spectrum (C_l):** Statistical description of temperature/polarisation fluctuations in the CMB as a function of angular scale (multipole moment ℓ).
- **Multipole Moment (ℓ):** Integer index representing angular scale; smaller ℓ = larger angular scales.
- **Spherical Harmonic Coefficient (a_{lm}):** Amplitude for each spherical harmonic mode in the expansion of the sky map.
- **HEALPix Pixelization:** A method for dividing a sphere into equal-area pixels for efficient map storage and analysis.
- **NaMaster:** Software package for pseudo-C_l power spectrum estimation, especially for masked sky maps, by implementing the MASTER algorithm.
- **Mode-Coupling Matrix (MCM):** Matrix describing how masking distorts the coupling between spherical harmonic modes, requiring correction in power spectrum analysis.
- **Multipole Vectors (MV):** Set of unit vectors uniquely representing each spherical harmonic mode (ℓ,m), used in isotropy and non-Gaussianity analysis.
- **Fréchet Vectors (FV):** Vectors quantifying geometric features and clustering of multipole components on the sphere.
- **Mask Apodization:** Smoothing the edges of a mask to reduce artificial spectral leakage.
- **Power Spectrum Decoupling:** Correcting C_l for mode-coupling effects induced by masking.
- **NSIDE:** Parameter determining the pixel resolution in HEALPix (higher NSIDE = higher resolution).
- **Planck/WMAP:** Satellite missions providing full-sky measurements of the CMB.

References

1. **[1] -** ***Planck Collaboration, N. Aghanim et al. "Planck 2018 results. VI. Cosmological parameters." Astronomy & Astrophysics, 641, A6 (2020). arXiv: 1807.06209***
2. **[2] -** ***D.J. Fixsen (2009), "The Temperature of the Cosmic Microwave Background." The Astrophysical Journal, Volume 707, Number 2, p. 916–920. arXiv: 0911.1955***
3. **[3] - Dodelson, Scott. Modern Cosmology. San Diego, Calif.: Academic Press, 2003.**
4. **[4] - Sullivan, R. M., Hergt, L. T., & Scott, D. (2024, October 16). Methods for CMB map analysis. arXiv.org. https://arxiv.org/abs/2410.12951**
5. **[5] - Górski, K. M., et al. "HEALPix – A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere." The Astrophysical Journal, vol. 622, no. 2, 2005, pp. 759–771. https://ui.adsabs.harvard.edu/abs/2005ApJ...622..759G**
6. **[6] - Oliveira, R. A., Pereira, T. S., & Quartin, M. (2020). CMB statistical isotropy confirmation at all scales using multipole vectors. *Physics of the Dark Universe*, *30*, 100608. https://doi.org/10.1016/j.dark.2020.100608**
7. **[7] - Scott, D., Smoot, G. F., & M. Tanabashi et al. (Particle Data Group). (2019). *Cosmic microwave background*. https://pdg.lbl.gov/2019/reviews/rpp2019-rev-cosmic-microwave-background.pdf**
8. **[8] - Lsstdesc. (n.d.). *GitHub - LSSTDESC/NaMaster: A unified pseudo-Cl framework*. GitHub: https://github.com/LSSTDESC/NaMaster**
9. **[9] - Robol. (n.d.). *GitHub - robol/MPSolve: Multiprecision Polynomial Solver*. GitHub: https://github.com/robol/MPSolve**
10. **[10] - Oliveirara. (n.d.). *GitHub - oliveirara/polymv: Python package that converts multipolar coefficients into Multipole Vectors and Fréchet Vectors*. GitHub: https://github.com/oliveirara/polymv**
11. **[11] - *HEALPix - Getting HEALPix*. (n.d.). HEALPix. https://healpix.sourceforge.io/downloads.php**
12. **[12] - *LAMBDA - CMB Experiments*. (n.d.). https://lambda.gsfc.nasa.gov/product/expt/**
13. **[13] - *LAMBDA - WMAP DR5 Processing, Intensity and Polarisation Exclusion Mask downloads*. (n.d.). https://lambda.gsfc.nasa.gov/product/wmap/dr5/masks_get.html**
14. **[14] - *Planck Legacy Archive*. (n.d.). https://pla.esac.esa.int/#home**
15. **[15] - *Welcome to the healpy documentation — healpy 0.1.dev158+g266430192 documentation*. (n.d.). https://healpy.readthedocs.io/en/latest/index.html**
16. **[16] - Oliveira, R. A., Pereira, T. S., & Quartin, M. (2020b). MVs and FVs for Planck 2015 and 2018 temperature maps obtained by polyMV. *Zenodo*. https://doi.org/10.5281/zenodo.3866410**
17. **[17] - *NumPy documentation — NumPy v2.3 Manual*. (n.d.). https://numpy.org/doc/2.3/**
18. **[18] - *Matplotlib documentation — Matplotlib 3.10.7 documentation*. (n.d.). https://matplotlib.org/stable/index.html**

Appendix A: Code Notebooks

- **Notebook 1: CMB Map Generation and Map Operations (healpy & NaMaster)**
  Google Colab: healpix_exercises_v1.ipynb
  This notebook covers:
    - *Mounting Google Drive for file access and persistent storage of CMB simulation data and masks*
    - *Installing and building Healpy and NaMaster from source, including system dependencies and Python version management for compatibility*
    - *Generating 100 random realisations of the CMB sky using theoretical Cℓ*
    - *Applying Gaussian (beam) transfer functions and visualising the effect on CMB maps*
    - *Statistical analysis: recovering input Cℓ from 100 simulations and plotting their mean and error*
    - *Upgrading/downgrading CMB maps: comparing power spectra at different map resolutions (nside = 16, 32, 64, 128, 1024)*
    - *Map operations (single and dual map):*
        - *Scaling, division, cutoff-based region selection (setting pixels to zero or Healpix bad value)*
        - *Multiplication (masking), addition, subtraction, safe division (with zero denominator handling) between two maps with proper nside matching and regrading*
    - *Visualisation: mollview plots for original, scaled, downgraded, and masked maps, and detailed comparisons of resulting spectra*
    - *Mask construction and application (including apodization)*
    - *Masked power spectrum estimation (Healpy, pseudo-C_l)*
    - *MASTER/NaMaster workflow: coupling matrix, binning, and decoupling*
    - *Error analysis and binning choices for power spectra*
    - *Visualisation and comparison of theoretical, masked, and decoupled spectra*

- **Notebook: MV/FV-Related Exercises for CMB Map Analysis**
  Google Colab *Mv/Fv_related _exercises_UG.ipynb*

  This notebook covers:
    - *Environment setup: Initial imports, configuration of paths, and basic function definitions for multipole and field vector analysis.*
    - *Loading CMB maps and defining core functions: Utility routines to prepare and process CMB datasets for further analysis.*
    - *Approach 1 — Direct access from results dictionary: Rapid plotting and analysis using pre-computed Multipole Vectors (MV) and Field Vectors (FV) for fast visualisation.*
    - *Approach 2 — Computation for single ℓ: Targeted calculation of MVs/FVs for individual multipole moments, ideal for focused studies.*
    - *Approach 3 — Full range computation and storage: Comprehensive extraction of MVs/FVs across a user-defined or standard multipole range, with output saved for future reuse and reproducibility.*
    - *MV/FV plotting for PR4 SEVEM map: Visualise and interpret MV/FV distributions across all relevant multipoles for the PR4 SEVEM CMB map.*
    - *Reference plotting with Planck CMB maps: Benchmarking MV/FV results against Planck public data for validation and comparison.*