

# **Dust-CIB Separation using Bayesian Inference**

*A thesis Submitted  
in Partial Fulfilment of the Requirements  
for the Degree of*

**MASTER OF SCIENCE**

*by*

**Ritabik Banerjee**

HBNI Enrollment No : **PHYS11202013039**



*to the*

**School of Physical Sciences**

**National Institute of Science Education and Research**

**Bhubaneswar**

# **Dedication**

I dedicate this work to my family, with heartfelt love and gratitude. To my mother and father—your unwavering support, boundless patience, and countless sacrifices have shaped every step of my journey. In moments of doubt and hardship, it was your belief in me that gave me the strength to carry on. You have always stood by me with quiet resilience and endless encouragement. This achievement is not mine alone; it is a shared testament to your love, guidance, and faith in my dreams.

# Declaration

I hereby declare that I am the sole author of this thesis in partial fulfillment of the requirements for a postgraduate degree from National Institute of Science Education and Research (NISER). I authorize NISER to lend this thesis to other institutions or individuals for the purpose of scholarly research.

Ritaleek Banerjee

Signature of the Student

Date: 20/5/25

Dust-CIB separation  
using Bayesian Inference

The thesis work reported in the thesis entitled ..... was carried out under my supervision, in the school of ..... at NISER, Bhubaneswar, India.

Tuhin Ghosh

Signature of the thesis supervisor

School: SPS

Date: 20-5-2025

## **Acknowledgement**

I would like to express my deep and sincere gratitude to my supervisor, Dr. Tuhin Ghosh, for his unwavering support, insightful guidance, and invaluable mentorship throughout the course of this work. His constant encouragement, patience, and depth of knowledge have greatly shaped both my academic progress and personal development, and I am truly grateful for the opportunity to work under his supervision.

I am also thankful to my group members and friends, Soham Maji and Aniket Nath, whose constant support, stimulating discussions, and collaborative spirit made this journey intellectually engaging and personally enjoyable. Their companionship and readiness to help at every step have been instrumental in overcoming challenges and maintaining motivation.

# Abstract

Galactic dust emission is a significant foreground in Cosmic Microwave Background (CMB) observations, especially at frequencies above 100 GHz, where it overlaps with the Cosmic Infrared Background (CIB). Accurate understanding of dust emission is crucial for distinguishing CMB signals and reconstructing CIB anisotropies. Dust grains emit across far-infrared and sub-millimeter wavelengths and influence CMB polarization, particularly B-mode polarization, which is vital for cosmological studies. Previous works has been done such as, using single-template models based on HI maps from the GASS survey, assuming dust emission is correlated with neutral hydrogen at high Galactic latitudes and low-column density regions. In contrast, this study uses Hamiltonian Monte Carlo sampling, providing pixel-dependent mean emissivity and global offset values for far-infrared and sub-millimeter wavelengths. In this project we utilize multiple dust extinction maps in a Bayesian framework employing Hamiltonian Monte-Carlo sampling, capturing detailed galactic dust features and line-of-sight variations, yielding residual maps that are less sensitive to noise and more accurate in isolating CIB contributions.

# Contents

	ii
	iii
	iv
	v
	Page
<b>Dedication</b>	ii
<b>Declaration</b>	iii
<b>Acknowledgement</b>	iv
<b>Abstract</b>	v
<b>1 Introduction</b>	<b>2</b>
1.1 Galactic Dust . . . . .	2
1.2 Advantages of Galactic Dust in Astrophysical Studies . . . . .	4
1.3 Previous Work on Galactic Dust Modeling . . . . .	4
1.4 Dust Extinction . . . . .	5
1.5 Correlation between Galactic dust and Dust Extinction . . . . .	5
1.6 Cosmic Infrared Background (CIB) and Its Importance . . . . .	7
<b>2 Theory and Methodologies</b>	<b>8</b>
2.1 Bayesian Inference Methodology . . . . .	8
2.1.1 Model for Dust Emission . . . . .	8
2.1.2 Hamiltonian Monte Carlo Sampling for Posterior Inference . . . . .	9
2.2 Extinction Maps from Green et al. (2015) . . . . .	11
2.3 Data Products . . . . .	18
2.3.1 Planck HFI 353 GHz Foreground Emission Map . . . . .	18
2.3.2 Sky Masking . . . . .	18
2.3.3 Dust Maps . . . . .	19
<b>3 Simulation</b>	<b>20</b>
3.1 Construction of Simulated data . . . . .	20
3.2 Fitting the Simulated Dust Map using HMC sampling . . . . .	21
3.3 Simulation Processing Time . . . . .	23

<b>4</b>	<b>Planck Data Fitting using HMC sampling</b>	<b>26</b>
4.1	Data Fitting using the HMC sampling . . . . .	26
4.2	Pixel-Dependent Dust-Extinction Coefficient maps . . . . .	27
<b>5</b>	<b>Conclusions and Future Directions</b>	<b>32</b>
5.1	CIB map and its Power spectrum . . . . .	32
5.2	Correlation between the Dust extinction coefficient Maps . . . . .	33
5.3	Comparison of the input and observed CIB map . . . . .	34
5.3.1	Map Normalization and Derivatives . . . . .	35
5.3.2	Definition of Minkowski Functionals . . . . .	35
5.3.3	Analytical Expectation for Gaussian Random Fields . . . . .	36
5.4	Conclusion . . . . .	38
5.5	Future Directions . . . . .	38

# Chapter 1

## Introduction

Galactic dust emission is a major foreground contaminant in observations of the Cosmic Microwave Background (CMB), especially at frequencies above 100 GHz where its spectral energy distribution overlaps significantly with that of the Cosmic Infrared Background (CIB). Accurately modeling and subtracting this foreground is critical to unlocking the cosmological information encoded in the CMB. In this study, we employ a Bayesian inference approach, combined with Hamiltonian Monte Carlo (HMC) sampling, to develop an improved model of Galactic dust emission. This framework captures spatial variations in dust properties and allows for probabilistic treatment of uncertainties, leading to more precise separation of the CIB from Galactic foregrounds and better constraints on the underlying cosmological signals.

### 1.1 Galactic Dust

Galactic dust is composed of submicron-sized solid particles primarily made of carbonaceous and silicate materials. These grains absorb ultraviolet (UV) and optical light from stars and re-radiate the energy as thermal emission in the far-infrared (FIR) and submillimeter (submm) wavelengths. This emission is particularly relevant at frequencies above 100 GHz, where it becomes a dominant foreground in CMB measurements. The spectral and spatial characteristics of dust emission are key to distinguishing it from the CMB and CIB signals.



Fig. 1.1: Galactic dust image from Spitzer's infrared array camera, using its 8-micron detector

The emission from Galactic dust can be described using a *modified blackbody* (or *greybody*)

spectrum, which incorporates the effects of dust emissivity and temperature:

$$I_\nu(r) = \epsilon_{\nu_0}(r) B_\nu(T_d(r)) \left( \frac{\nu}{\nu_0} \right)^{\beta(r)} \quad (1.1)$$

Here,

- $I_\nu(r)$ : Specific intensity of thermal dust emission at frequency  $\nu$  and sky location  $r$  [MJy/sr],
- $\epsilon_{\nu_0}(r)$ : Emissivity at reference frequency  $\nu_0$ ,
- $B_\nu(T_d(r))$ : Planck function at dust temperature  $T_d(r)$  [MJy/sr],
- $\nu_0$ : Reference frequency, commonly taken as 353 GHz,
- $\beta(r)$ : Dust spectral emissivity index, typically ranging between 1.5 and 2.0 in the diffuse ISM.

This expression captures the frequency scaling of the dust emission and its dependence on local physical parameters. The optical depth  $\tau_{\nu_0}(r)$  reflects the column density and absorption efficiency of dust along the line of sight. The function  $B_\nu(T_d)$  governs the blackbody thermal emission at dust temperature  $T_d$ , while the power-law term accounts for the emissivity variation with frequency.

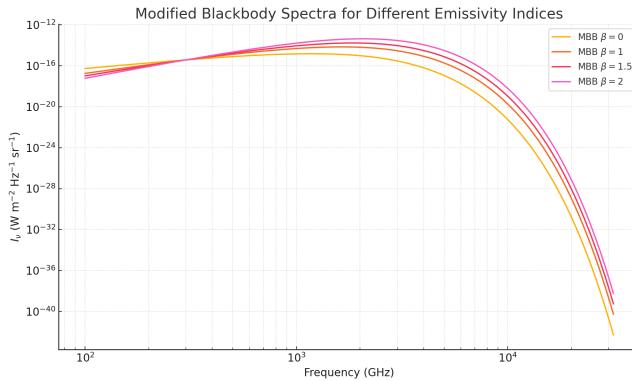


Fig. 1.2: Modified Black body spectrum (MBB) fit

To account for the three-dimensional distribution of dust in the Galaxy, the total observed emission is obtained by integrating along the line of sight:

$$I_\nu(\hat{\mathbf{n}}) = \int_0^\infty \rho_d(r) B_\nu(T_d(r)) \left( \frac{\nu}{\nu_0} \right)^\beta dr \quad (1.2)$$

In this expression,

- $\rho_d(r)$  is the dust density,
- $T_d(r)$  is the dust temperature as a function of distance  $r$ ,
- $\hat{\mathbf{n}}$  is the direction of observation on the sky.

Additionally, the optical depth  $\tau_\nu(r)$  is related to the dust opacity  $\kappa_\nu(r)$  by:

$$\tau_\nu(r) = \kappa_\nu(r) \rho_d(r) \quad (1.3)$$

This formulation provides the basis for constructing full-sky dust emission templates, which can be compared to observations from FIR and submm instruments such as *Planck* and *IRAS*.

## 1.2 Advantages of Galactic Dust in Astrophysical Studies

While Galactic dust is often considered a contaminant in cosmological observations, it also provides a wealth of information about the interstellar medium (ISM), star formation, and Galactic magnetic fields. The following are key advantages of studying Galactic dust:

1. **Tracing the Interstellar Medium (ISM):** Dust is well-mixed with gas in the ISM, allowing it to serve as a proxy for gas column density. The dust column density  $\Sigma_d$  is related to the observed optical depth  $\tau_\nu$  and dust opacity  $\kappa_\nu$  by:

$$\Sigma_d = \frac{\tau_\nu}{\kappa_\nu} \quad (1.4)$$

This relation allows us to map gas distribution using infrared observations of dust.

2. **Probing Galactic Magnetic Fields:** Polarized dust emission arises from aspherical grains aligned with the magnetic field. The polarization angle and degree can reveal the orientation and structure of the Galactic magnetic field, described by:

$$P = p_0 I_\nu \cos^2(\gamma) \quad (1.5)$$

where  $p_0$  is the maximum polarization fraction and  $\gamma$  is the angle between the magnetic field and the line of sight.

3. **Tracing Star-Forming Regions:** Dense dust clouds serve as the birthplaces of stars. The temperature  $T_d$  and optical depth  $\tau_\nu$  help identify cold, dense regions such as molecular clouds where star formation is ongoing.
4. **Foreground Characterization for CMB and CIB Studies:** Dust modeling enables accurate subtraction of foregrounds in cosmological datasets. Understanding its emission spectrum improves component separation techniques and enhances the fidelity of residual CMB and CIB maps.

Overall, Galactic dust is not only an obstacle in CMB studies but also a valuable astrophysical tool that provides insights into the composition, structure, and evolution of the Milky Way.

## 1.3 Previous Work on Galactic Dust Modeling

Recent advancements in Galactic dust modeling have leveraged sophisticated statistical methods to improve our understanding of the structure and composition of the interstellar medium. Notably, Adak et al. (2024) [3] developed a Bayesian inference framework to characterize dust emissivity at far-infrared and submillimeter wavelengths using Markov Chain Monte Carlo (MCMC) techniques. Their model captures spatial variations in dust properties and includes rigorous uncertainty quantification, enabling more accurate modeling of Galactic foregrounds.

Despite these advancements, a persistent challenge remains in the accurate separation of the Cosmic Infrared Background (CIB) from the CMB signal. Many models primarily focus on fitting dust

parameters such as temperature and emissivity, but do not fully address the issue of isolating the CIB, which is superimposed on Galactic dust emission in observational data. The CIB, originating from high-redshift galaxies, often contributes residual structures that are mistaken for Galactic features, complicating foreground removal efforts.

To address this issue, our approach builds upon previous work by incorporating multiple dust extinction templates derived from the 3D dust mapping efforts of Green et al. (2015) [7]. Rather than relying on a single HI template, we use spatially-resolved extinction maps that reflect the varying properties of dust along different lines of sight. This enables a more nuanced modeling of the dust emission, improving the accuracy of CIB extraction and reducing systematic errors in CMB analysis. By integrating multi-template modeling within a Bayesian framework, we enhance both the fidelity and interpretability of dust-corrected sky maps.

## 1.4 Dust Extinction

Dust extinction refers to the attenuation of electromagnetic radiation as it traverses the interstellar medium, primarily due to scattering and absorption by dust grains. This phenomenon significantly alters the apparent brightness and color of astrophysical sources, particularly in the ultraviolet (UV), optical, and near-infrared bands. Quantifying dust extinction is essential for correcting astronomical observations and for probing the structure of the ISM.

One of the most widely used metrics for quantifying dust extinction is the color excess  $E(B - V)$ , defined as:

$$E(B - V) = (B - V)_{\text{observed}} - (B - V)_{\text{intrinsic}} \quad (1.6)$$

where  $B$  and  $V$  represent the magnitudes in the blue and visual bands, respectively. The value of  $E(B - V)$  reflects the amount of reddening caused by dust along the line of sight and correlates with the column density of interstellar dust.

To relate extinction at a specific wavelength  $\lambda$  to  $E(B - V)$ , the total-to-selective extinction ratio  $R_V$  is used:

$$A_\lambda = R_V \cdot E(B - V) \quad (1.7)$$

where  $A_\lambda$  is the extinction at wavelength  $\lambda$ , and  $R_V = A_V/E(B - V)$  is typically around 3.1 in the diffuse ISM, but may vary depending on dust grain size and composition.

In this study, we employ the Planck 2013 scaling relation [1] to convert the modeled dust optical depth to extinction values. This allows for direct comparison between emission-based models and extinction observations, thereby linking radiative transfer models with empirical measurements of dust attenuation.

## 1.5 Correlation between Galactic dust and Dust Extinction

Galactic dust grains are crucial components of the interstellar medium (ISM), influencing both the propagation of starlight and the observed thermal emission at submillimeter wavelengths. These aspherical grains are partially aligned with the Galactic magnetic field, and this alignment causes anisotropic interactions with electromagnetic radiation. This dual interaction—absorption in the optical and emission in the infrared/submillimeter—forms the physical basis for using dust

extinction maps as reliable tracers of Galactic dust.

In the optical regime, dust grains cause extinction and polarization of background starlight through selective absorption. Light from distant stars, initially unpolarized, passes through dust grains that preferentially absorb electric field components aligned with their long axes. As a result, the transmitted light becomes linearly polarized with the electric vector  $\vec{E}$  oriented perpendicular to the grain alignment, and hence perpendicular to the local magnetic field.

In contrast, in the submillimeter regime, the same grains emit thermally polarized radiation. Since thermal emission is strongest along the grain's long axis, and grains are aligned with their short axes along the magnetic field, the emitted radiation becomes polarized **perpendicular** to the magnetic field. Therefore, extinction-induced polarization and emission-induced polarization are orthogonal in orientation.

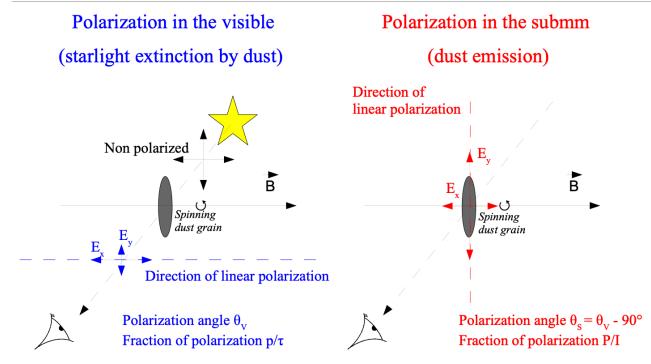


Fig. 1.3: Schematic representation of dust-induced polarization in extinction (left) and emission (right). In extinction, polarization angle  $\theta_V$  is aligned with the magnetic field, while in emission,  $\theta_S = \theta_V - 90^\circ$ . Fractional polarization is given by  $p/\tau$  in extinction and  $P/I$  in emission.

Mathematically, the observed extinction polarization can be expressed as:

$$\frac{p}{\tau} = \frac{C_{\parallel} - C_{\perp}}{C_{\parallel} + C_{\perp}},$$

where  $C_{\parallel}$  and  $C_{\perp}$  are the extinction cross-sections parallel and perpendicular to the grain alignment. Similarly, for thermal emission, the polarization fraction becomes:

$$\frac{P}{I} = \frac{j_{\perp} - j_{\parallel}}{j_{\perp} + j_{\parallel}},$$

where  $j_{\perp}$  and  $j_{\parallel}$  are the emissivities orthogonal and parallel to the grain alignment axis.

This physical connection implies that dust extinction maps—derived from optical observations—encode the spatial distribution of dust column density and magnetic alignment that also govern far-infrared and submillimeter dust emission. Therefore, extinction templates serve as valuable priors for dust emission models used in cosmological foreground separation.

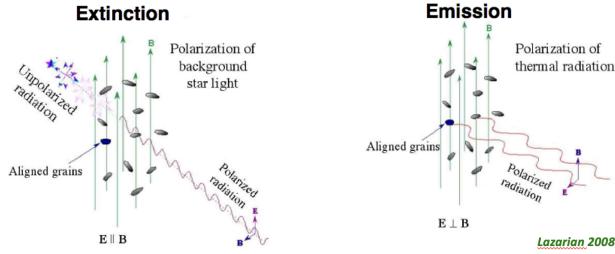


Fig. 1.4: Illustration from Lazarian (2008) showing how aligned grains lead to polarization in extinction (left) and in thermal emission (right). Both phenomena are governed by the same dust grain alignment with respect to the magnetic field.

As depicted in Figure 1.4, aligned grains modulate both extinction and emission in a coherent, physically consistent manner. This unified framework supports the central idea of this work that distance-resolved extinction maps can serve as templates to decompose Galactic dust emission, enabling the recovery of cosmological signals like the Cosmic Infrared Background (CIB).

## 1.6 Cosmic Infrared Background (CIB) and Its Importance

The **Cosmic Infrared Background (CIB)** is a diffuse background radiation field composed of the integrated infrared emission from distant, dust-enshrouded galaxies across cosmic time. It is a relic of galaxy formation and evolution, tracing the buildup of stellar mass and dust in the early universe. The CIB originates primarily from thermal emission by dust heated by young, massive stars in high-redshift galaxies, redshifted into the far-infrared and submillimeter regime.

Studying the CIB provides crucial insights into the star formation history of the universe, the evolution of dust-obscured galaxies, and the growth of large-scale cosmic structures. It serves as a powerful probe of early galaxy formation, complementing optical and radio surveys by accessing obscured star-forming regions invisible at other wavelengths [10].

However, a major observational challenge lies in isolating the CIB signal from Galactic foregrounds, especially thermal dust emission that dominates the same frequency bands. This overlap introduces uncertainties in measurements of the CMB and hinders precise estimation of cosmological parameters related to inflation, dark energy, and reionization.

In this work, we isolate the CIB using a probabilistic Bayesian modeling framework that builds upon the dust emission model of Adak et al. (2024) [3]. By incorporating multi-template extinction maps and applying Hamiltonian Monte Carlo sampling, we enhance the separation of Galactic dust from the CIB. This refined extraction methodology improves the fidelity of residual maps and enhances the interpretability of the CIB signal as a cosmological tracer. As emphasized in [11], understanding the anisotropies and power spectrum of the CIB is essential for uncovering the structure and evolution of the early universe.

# Chapter 2

## Theory and Methodologies

### 2.1 Bayesian Inference Methodology

We employ a Bayesian statistical framework to separate the Cosmic Infrared Background (CIB) from Galactic dust emission, utilizing line-of-sight-dependent dust extinction templates derived from 3D dust extinction datacube. This method jointly infers spatially resolved pixel-dependent dust extinction coefficients  $\epsilon_{j,k}^\nu$  parameters and a global offset  $O_\nu$  at Planck HFI frequency  $\nu = 353$  GHz. The framework enables the separation of large-scale foreground structures from the underlying extragalactic signal by modeling dust as a superposition of six distance-stratified extinction components.

Bayes' theorem governs the inference:

$$P(\theta|D_\nu) = \frac{P(D_\nu|\theta)P(\theta)}{P(D_\nu)}$$

where  $\theta = \{\epsilon_{j,k}^\nu, O_\nu\}$  denotes the set of model parameters,  $P(D_\nu|\theta)$  is the likelihood,  $P(\theta)$  is the prior, and  $P(D_\nu)$  is the Bayesian evidence.

#### 2.1.1 Model for Dust Emission

The observed far-infrared or submillimeter intensity  $I_\nu(\Omega_{j,i})$  at frequency  $\nu$  and angular direction  $\Omega_{j,i}$  (corresponding to subpixel  $i$  within superpixel  $j$ ) is modeled as a combination of dust emission, global background offsets, instrumental noise, CIB fluctuations, and residual Galactic foregrounds. Formally, the observed signal is expressed as:

$$I_\nu(\Omega_{j,i}) = \sum_k \epsilon_{j,k}^\nu E(B - V)_k(\Omega_{j,i}) + O_\nu + I_N^\nu(\Omega_{j,i}) + I_{\text{CIB}}^\nu(\Omega_{j,i}) + I_R^\nu(\Omega_{j,i})$$

Here:

- $\epsilon_{j,k}^\nu$  denotes the dust extinction coefficient at frequency  $\nu$ , associated with the  $k$ -th extinction template within superpixel  $j$ ,
- $E(B - V)_k(\Omega_{j,i})$  is the extinction value at the corresponding subpixel from the  $k$ -th line-of-sight distance bin,
- $O_\nu$  is a global offset, accounting for isotropic components such as the monopole of the CIB or

zero-level systematics,

- $I_N^\nu(\Omega_{j,i})$ ,  $I_{\text{CIB}}^\nu(\Omega_{j,i})$ , and  $I_R^\nu(\Omega_{j,i})$  represent the contributions from instrumental noise, unresolved CIB anisotropies, and residual Galactic foregrounds, respectively.

We define the modeled signal as:

$$S_\nu(\Omega_{j,i}) = \sum_k \epsilon_{j,k}^\nu E(B-V)_k(\Omega_{j,i}) + O_\nu$$

This signal encapsulates the deterministic part of the emission, which we aim to fit to the data.

The observational data  $D_\nu$  are assumed to follow a multivariate Gaussian distribution centered on  $S_\nu$  with a covariance matrix  $\Sigma_\nu$  that incorporates pixel-dependent uncertainties and instrumental correlations. The likelihood function thus takes the form:

$$P(D_\nu | \epsilon_{j,k}^\nu, O_\nu) = \frac{1}{(2\pi)^{N/2} \sqrt{|\Sigma_\nu|}} \exp \left[ -\frac{1}{2} (D_\nu - S_\nu)^T \Sigma_\nu^{-1} (D_\nu - S_\nu) \right]$$

For computational simplicity and tractability, we assume that  $\Sigma_\nu$  is block-diagonal within superpixels, allowing the inversion of smaller covariance matrices independently for each region. In the pixel-wise approximation, the likelihood becomes:

$$L(\epsilon_{j,k}^\nu, O_\nu | D_\nu) = \frac{1}{(2\pi)^{D/2} \sqrt{|C_\nu|}} \exp \left( -\frac{1}{2} \sum_\nu \sum_{j,i} \left( \frac{D_\nu(\Omega_{j,i}) - \sum_k \epsilon_{j,k}^\nu E(B-V)_k(\Omega_{j,i}) - O_\nu}{\sigma_\nu(\Omega_{j,i})} \right)^2 \right)$$

This likelihood function is equivalent to minimizing the following chi-squared function:

$$\chi^2 = \sum_\nu \sum_{j,i} \left( \frac{D_\nu(\Omega_{j,i}) - \sum_k \epsilon_{j,k}^\nu E(B-V)_k(\Omega_{j,i}) - O_\nu}{\sigma_\nu(\Omega_{j,i})} \right)^2$$

This  $\chi^2$  formulation quantifies the goodness of fit between the observed data and the model prediction, accounting for the spatial variations in extinction and observational uncertainties. The chi-squared is minimized when the modeled dust emission best matches the observed intensity within the bounds of the measurement errors.

The above formalism enables robust estimation of the dust extinction coefficients  $\epsilon_{j,k}^\nu$  and the global offset  $O_\nu$  by incorporating physical prior information and accounting for data covariance. The posterior distributions of these parameters are obtained using Hamiltonian Monte Carlo, as discussed in the following section.

## 2.1.2 Hamiltonian Monte Carlo Sampling for Posterior Inference

To sample efficiently from the high-dimensional posterior distribution of the dust extinction coefficients and global offsets, we employ the Hamiltonian Monte Carlo (HMC) algorithm. Unlike traditional Markov Chain Monte Carlo (MCMC) techniques that rely on random walks and often suffer from slow convergence in large parameter spaces, HMC introduces auxiliary momentum variables and leverages the principles of classical mechanics to propose new states in a more informed and globally coherent manner.

The core idea is to interpret the target posterior as a potential energy surface and simulate the evolution of a fictitious particle on this surface. The Hamiltonian governing the dynamics is defined

as:

$$H(p, q) = \sum_{j,k} \frac{p_{\epsilon_{j,k}^\nu}^2}{2\mu_{\epsilon_{j,k}^\nu}} + \frac{p_{O_\nu}^2}{2\mu_{O_\nu}} - \ln P(\epsilon_{j,k}^\nu, O_\nu | D_\nu)$$

where:

- $q = \{\epsilon_{j,k}^\nu, O_\nu\}$  are the model parameters (dust extinction coefficients and global offset),
- $p = \{p_{\epsilon_{j,k}^\nu}, p_{O_\nu}\}$  are the corresponding momentum variables,
- $\mu$  denotes the mass terms controlling the momentum scales,
- $P(\epsilon_{j,k}^\nu, O_\nu | D_\nu)$  is the posterior distribution of the model parameters given the observed data.

The system evolves through Hamilton's equations of motion:

$$\frac{dq}{dt} = \frac{\partial H}{\partial p}, \quad \frac{dp}{dt} = -\frac{\partial H}{\partial q}$$

These equations are integrated using the symplectic leapfrog method, which preserves the total energy and ensures time-reversibility and volume conservation in phase space, both of which are essential for maintaining detailed balance.

**Mass Matrix Construction :** The choice of the mass matrix  $\mu$  is crucial for achieving efficient exploration. In our implementation, the mass terms are computed analytically from the data and extinction templates within each superpixel, as follows:

$$\begin{aligned} \mu_{\epsilon_{j,k}^\nu} &= \sum_{i,i' \subset j} E(B-V)_k(\Omega_{j,i}) [\Sigma_\nu^{-1}]_{i,i'} E(B-V)_k(\Omega_{j,i'}) \\ \mu_{O_\nu} &= \sum_j \sum_{i,i' \in j} [\Sigma_\nu^{-1}]_{i,i'} \end{aligned}$$

These expressions reflect the curvature of the log-likelihood with respect to the parameters and are derived from the second derivative (Fisher information) of the chi-squared term in the Gaussian likelihood.

The covariance matrix  $\Sigma_\nu$ , representing the noise properties of the observed maps, is taken directly from the Planck 353 GHz variance data. It accounts for instrumental noise, residual Galactic foregrounds, and CIB anisotropies, and is assumed to be block-diagonal across superpixels for computational efficiency.

**Convergence and Diagnostics :** To ensure reliable posterior estimates, multiple HMC chains are initialized from independent starting points. Convergence is assessed using the Gelman–Rubin statistic  $\hat{R}$ , which compares the variance between chains to the variance within chains. A threshold of  $\hat{R} < 1.01$  is used to confirm that the chains have mixed well and represent a stationary distribution.

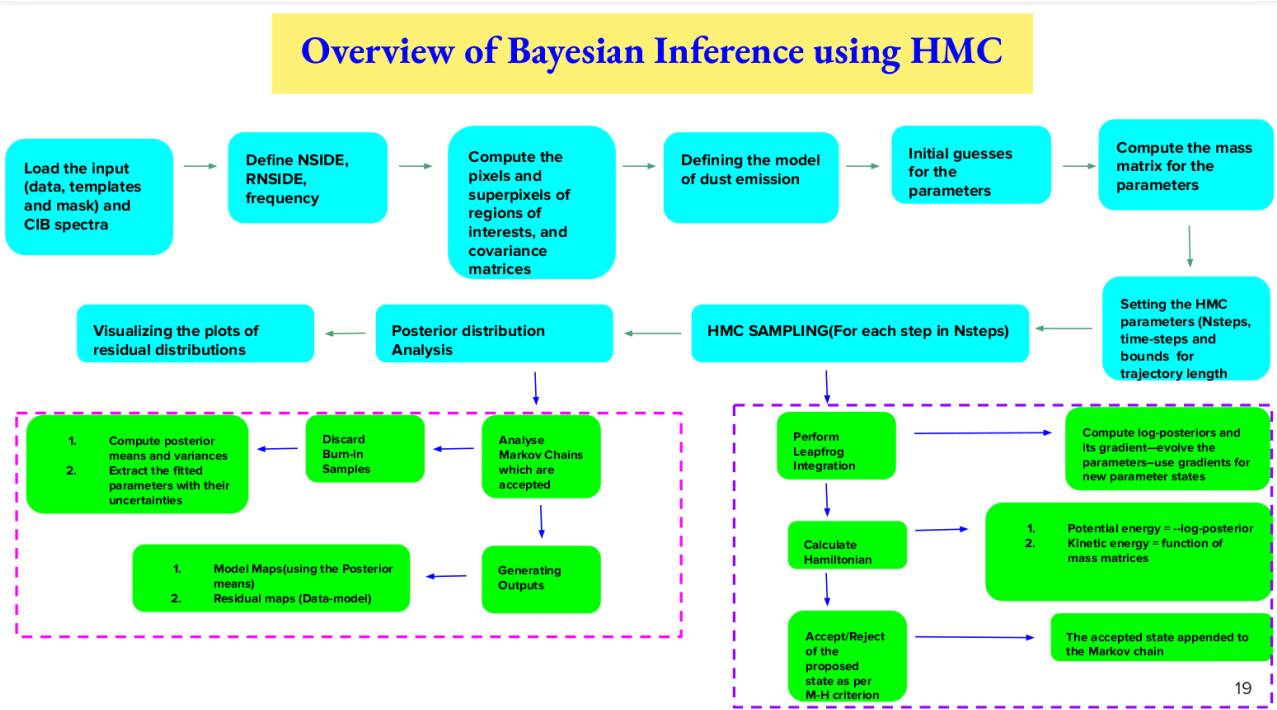


Fig. 2.1: Bayesian Methodology

This flowchart illustrates the full Bayesian inference pipeline employed in this study to estimate spatially varying dust extinction coefficients and recover the Cosmic Infrared Background (CIB). The procedure begins with the preprocessing of observational data and extinction templates, followed by the definition of the emission model, mass matrix construction, and initialization of HMC parameters. HMC sampling is performed to explore the posterior distribution of parameters using Hamiltonian dynamics. The resulting Markov chains are then analyzed to derive fitted parameters, compute model and residual maps, and visualize uncertainties. This workflow demonstrates the integration of physics-based modeling with statistical inference for robust dust foreground separation.

## 2.2 Extinction Maps from Green et al. (2015)

We utilize dust extinction data from Green et al. (2015) [7], which provides a 3D model of dust extinction in the Milky Way. The dataset includes  $E(B-V)$  values for each sky pixel at multiple resolutions ( $nside = 2048-64$ ). These maps are used to model the dust emission and separate the Cosmic Infrared Background (CIB) from the Cosmic Microwave Background (CMB). First, we un-sample the  $E(B-V)$  values at single resolution of  $nside = 64$  for the following distance moduli bins: 4, 7, 10, 13, 16, and 19 (corresponding to distances of 63, 251, 1000, 3881, 15849, and 63096 pc). These bins represent different layers of interstellar dust reddening and after generating maps we calculate the fraction  $f_i(p)$  of dust emission at each layer .

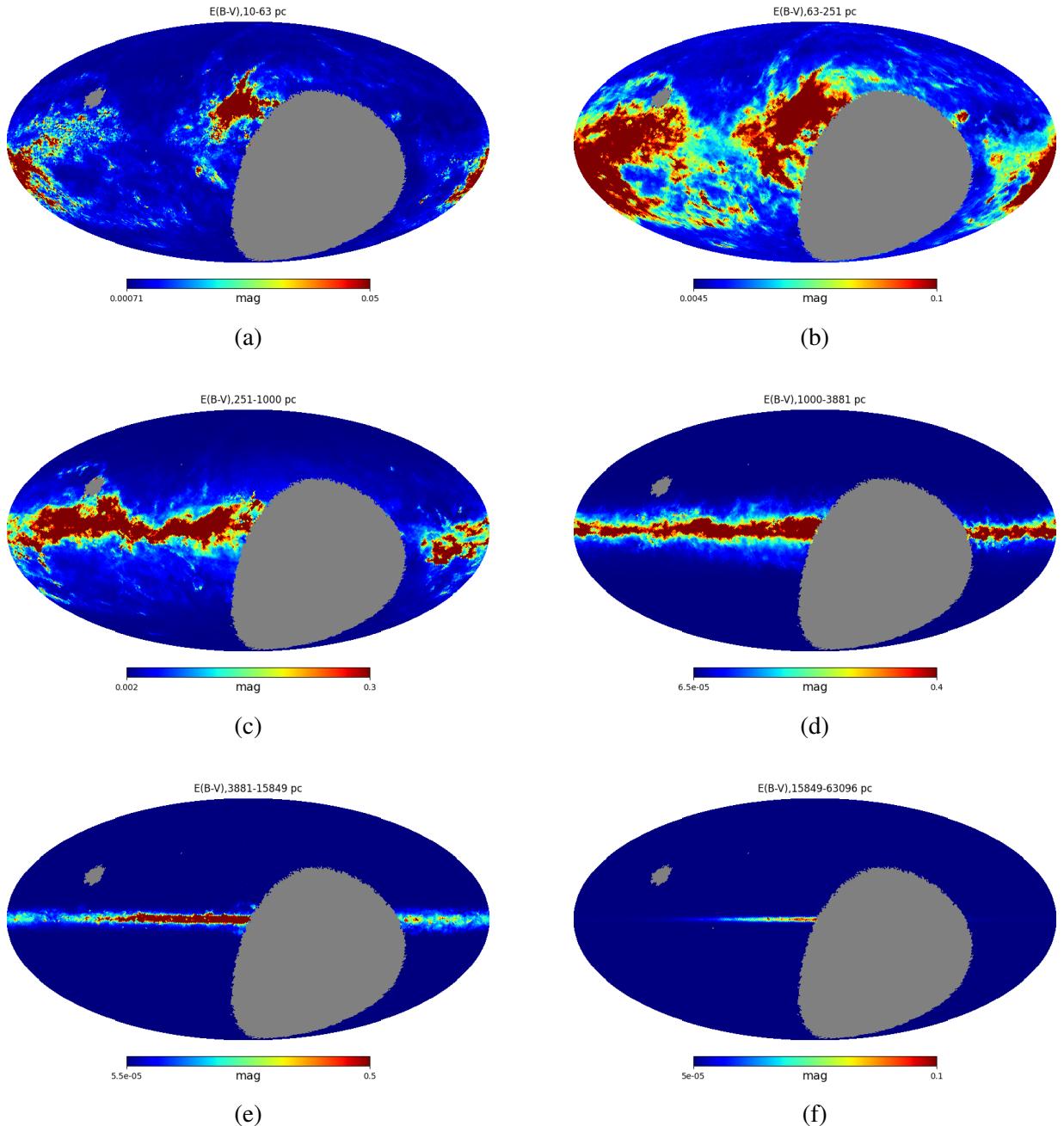


Fig. 2.2: Dust extinction maps at different distance ranges. Grey areas correspond to regions that have not been observed.

We divide the total  $E(B-V)$  extinction along each line of sight into six discrete distance bins, with each bin corresponding to a specific layer of Galactic dust at increasing distances from the observer. These layers are chosen to reflect meaningful astrophysical structures and distance scales within the Milky Way. The fractional contribution of dust in each bin, denoted by  $f_i(p)$ , is normalized such that:

$$\sum_{i=1}^6 f_i(p) = 1 \quad \text{for all } p$$

This normalization ensures that the full line-of-sight dust extinction is properly partitioned among the six layers at each sky pixel. The choice of six bins balances computational efficiency with sufficient resolution to capture depth-dependent variations in dust distribution.

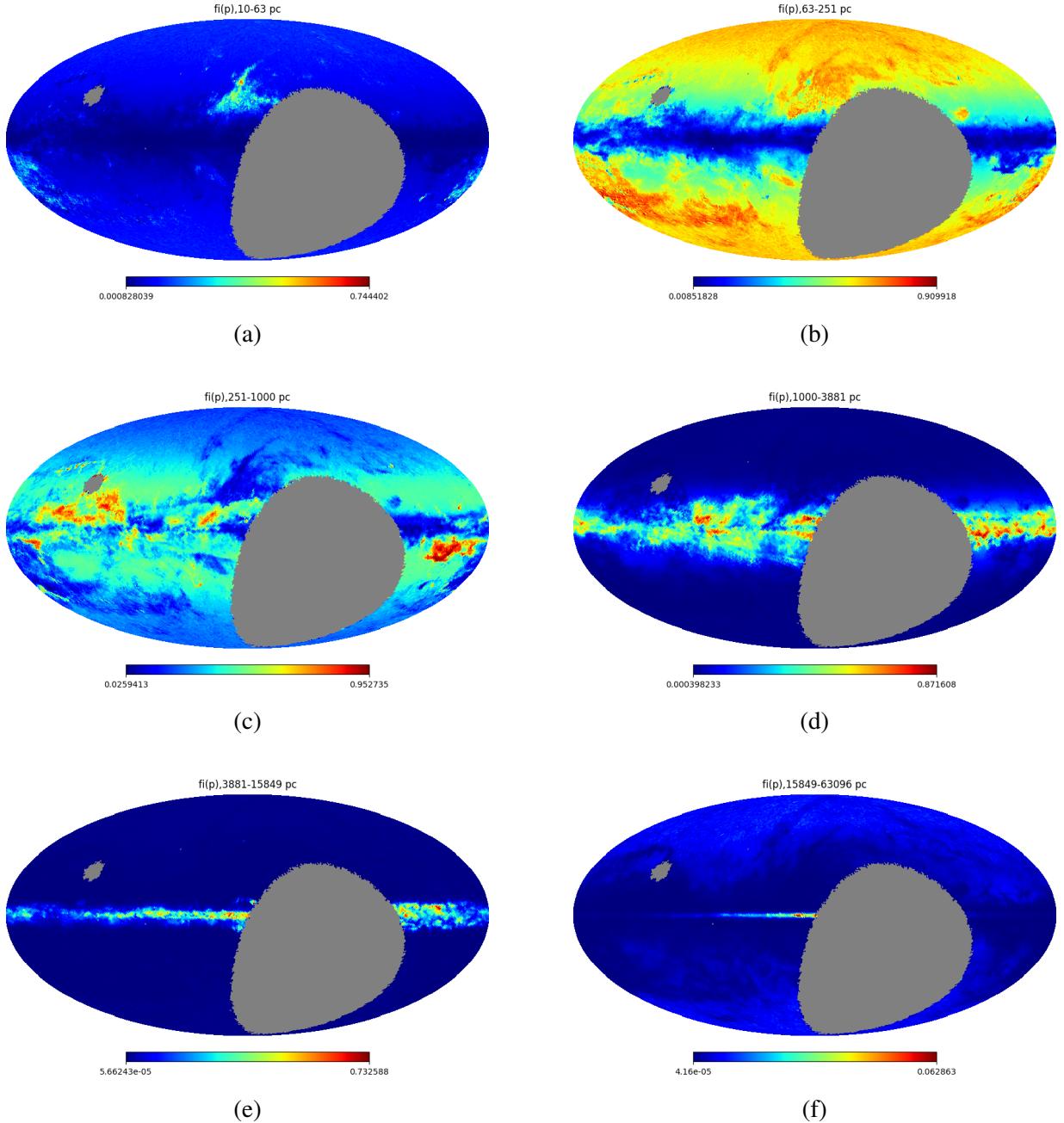


Fig. 2.3: Fraction of Dust Extinction maps at different distance ranges and grey areas correspond to regions that have not been observed.

In full-sky dust maps, certain regions often suffer from missing or unreliable data due to masking, low signal-to-noise ratios, or contamination by strong foregrounds. These incomplete regions pose a challenge to statistical modeling and map-based inference. To ensure continuity in our dust decomposition framework and allow for uniform Bayesian inference across the sky, we employ a method to estimate missing dust fraction values using symmetry arguments.

The approach leverages the approximate symmetry of Galactic structures about the Galactic equator. It is based on the assumption that dust distribution patterns are, to first order, symmetric with respect to the Galactic plane, particularly at high latitudes. This physical symmetry provides a practical basis for inferring missing values from available data.

For each pixel  $p_{\text{missing}}$  with a missing dust fraction value, we identify a symmetric counterpart  $p_{\text{sym}}$  across the Galactic equator. In spherical Galactic coordinates, this corresponds to reflecting the pixel across latitude  $b = 0^\circ$ , which is implemented by the following transformation:

$$\theta_{\text{sym}} = \pi - \theta, \quad \phi_{\text{sym}} = \begin{cases} 2\pi - \phi & \text{if } \phi > \pi, \\ \phi & \text{otherwise.} \end{cases}$$

Here,  $\theta$  and  $\phi$  denote the colatitude and longitude of the pixel, respectively. If the symmetric pixel  $p_{\text{sym}}$  contains a valid dust fraction value  $f(p_{\text{sym}})$ , we assign it directly to the missing pixel:

$$f(p_{\text{missing}}) = f(p_{\text{sym}})$$

This substitution maintains the large-scale symmetry of the dust map and is often sufficient for recovering values in high-latitude regions.

However, in some cases the symmetric pixel may also lack valid data—for example, due to masking in both hemispheres or severe contamination. In such cases, we employ a secondary fallback strategy based on local spatial averaging. Specifically, we identify the set of neighboring valid pixels  $N_{\text{valid}}(p_{\text{missing}})$  around the missing pixel and assign the average of their dust fractions:

$$f(p_{\text{missing}}) = \frac{1}{|N_{\text{valid}}(p_{\text{missing}})|} \sum_{p_{\text{valid}} \in N_{\text{valid}}(p_{\text{missing}})} f(p_{\text{valid}})$$

This ensures that missing pixels are filled in a statistically consistent manner, while also preserving local structural continuity.

Overall, this symmetry-based filling technique allows us to generate continuous full-sky maps of dust fractions for each distance bin, facilitating downstream analyses such as extinction coefficient inference and CIB residual estimation. To ensure physical consistency, the dust fractions across all six extinction layers are normalized at each pixel.

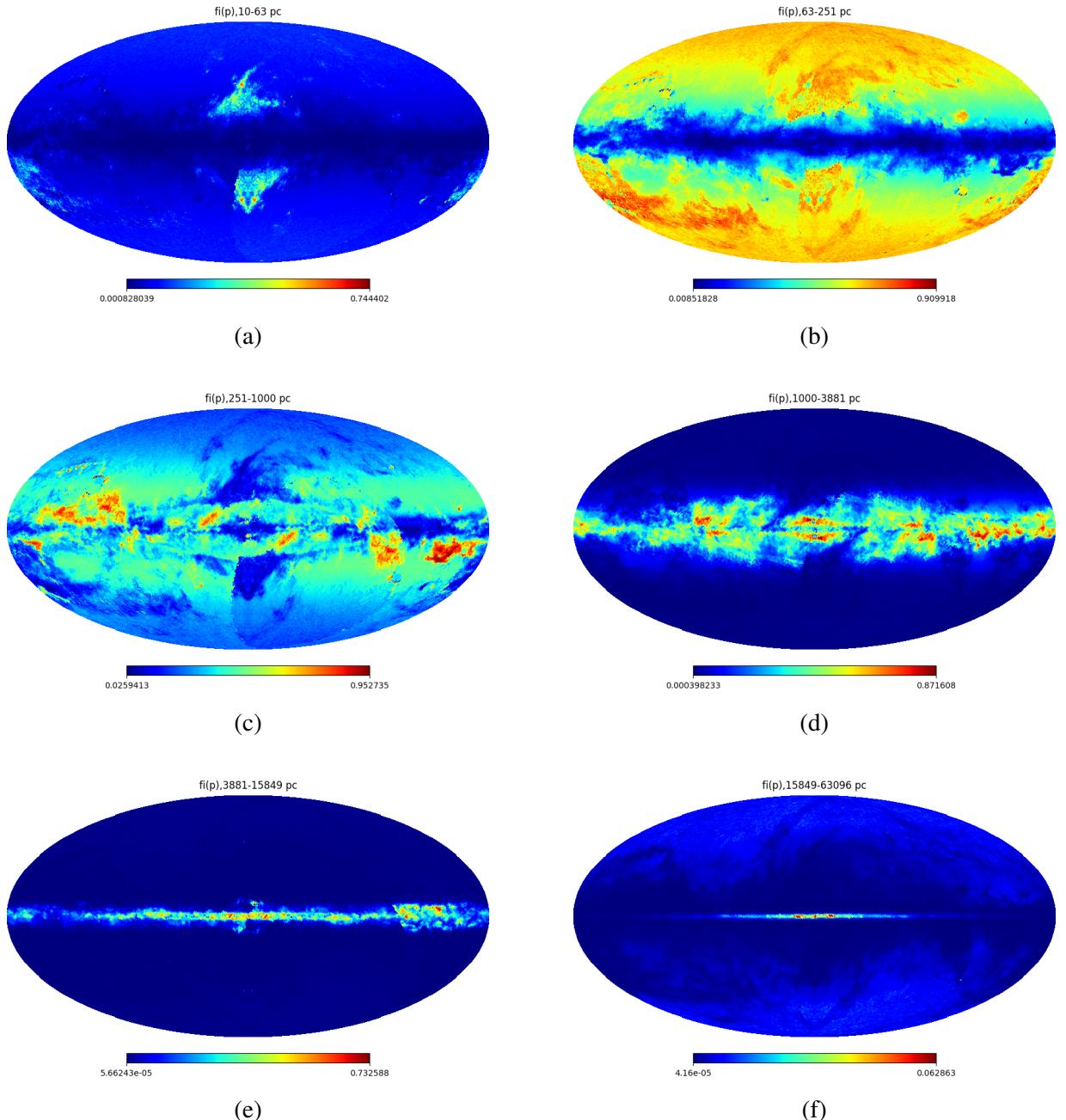


Fig. 2.4: Fraction of Dust Extinction maps at different distance ranges. Missing regions are filled using symmetry arguments.

To relate these fractional contributions to observable quantities, the maps are scaled to derive opacity values using the following transformation:

$$\tau_i(\nu_0) = f_i(p) \tau(\nu_0)$$

where  $\tau(\nu_0)$  is the total optical depth at reference frequency  $\nu_0 = 353$  GHz, obtained from Planck Modified Blackbody (MBB) fits. The result is a set of tomographic opacity maps indicating how much dust lies along each line-of-sight shell.

This decomposition allows us to track dust contributions across multiple distance layers. It enhances our understanding of dust structure and its projection onto the sky. Our method for opacity decomposition is inspired by the techniques in Martinez et al. (2018) [12].

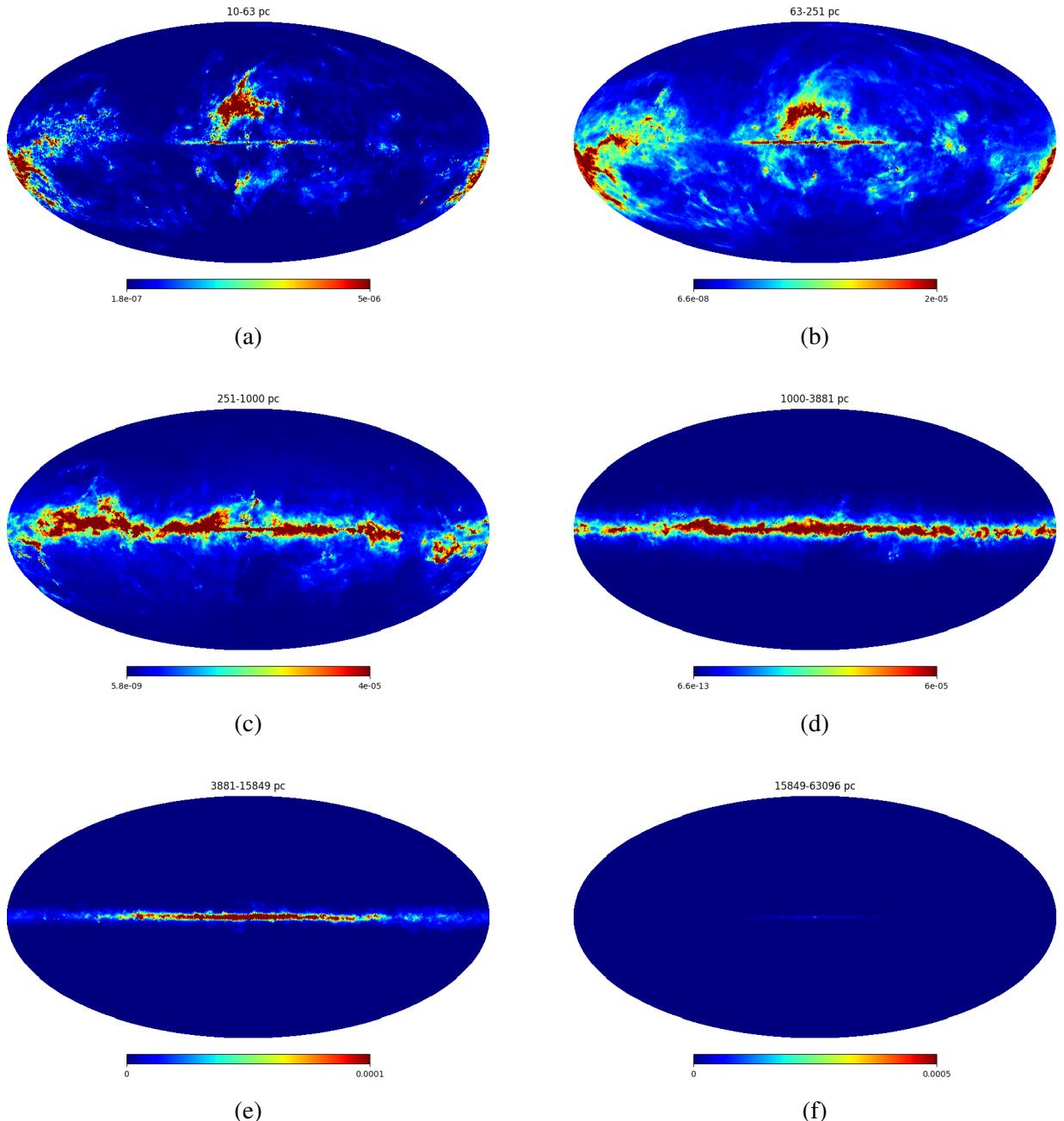


Fig. 2.5: Opacity maps at different distances at 353 GHz

After computing the opacity maps, we convert them into dust extinction maps by applying the Planck 2013 empirical conversion factor that links optical depth  $\tau(\nu)$  at 353 GHz to reddening values  $E(B - V)$ :

$$E(B - V) = \tau(\nu) \times 1.49 \times 10^4$$

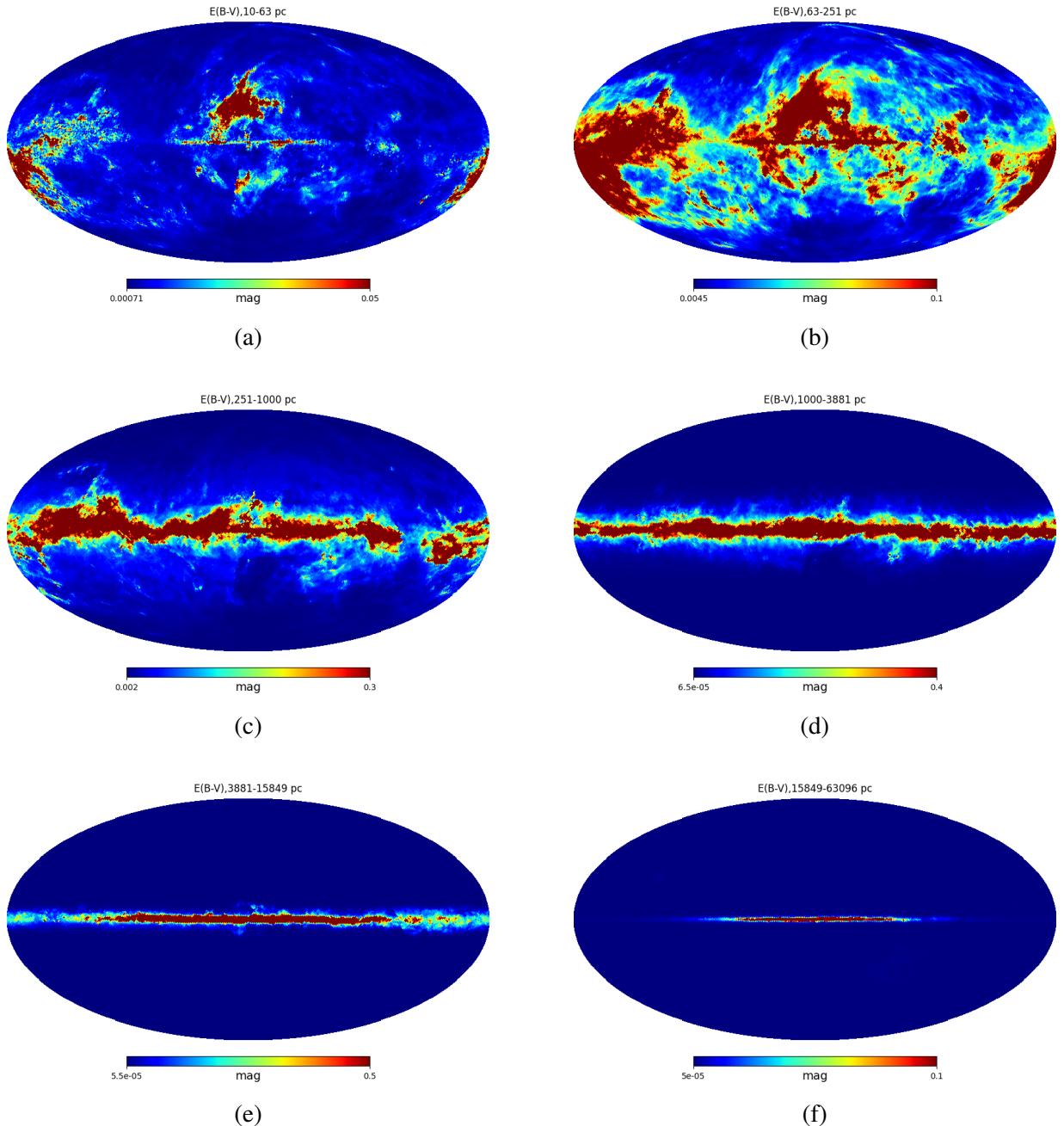


Fig. 2.6: Dust extinction maps at different line-of-sight distance ranges. Missing regions are filled using symmetry arguments.

All these maps are processed at a HEALPix resolution of  $\text{NSIDE} = 512$ , which offers sufficient angular resolution to capture fine-scale structures in Galactic dust across multiple distance layers. This high-resolution approach ensures accurate modeling of both diffuse and compact components in the line-of-sight extinction structure.

## 2.3 Data Products

### 2.3.1 Planck HFI 353 GHz Foreground Emission Map

We use the dust emission map produced by Planck's High-Frequency Instrument (HFI) at 353 GHz, with a pixel resolution of  $16'$  (arcminutes), corresponding to  $\text{NSIDE} = 512$ . It provides detailed modeling of thermal dust and which are critical for studying Galactic foregrounds and separating them from the Cosmic Microwave Background (CMB). The 353 GHz frequency is significant for observing Galactic dust, as its thermal emission peaks in the far-infrared. Using SMICA (Spectral Matching Independent Component Analysis), this map aids in accurately isolating the Cosmic Infrared Background (CIB) and other extragalactic signals.

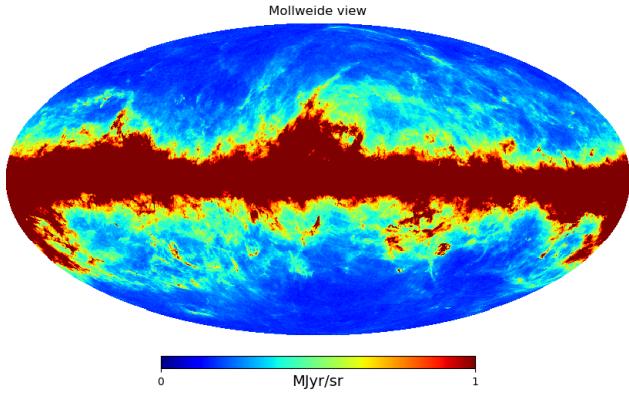


Fig. 2.7: HFI Planck Foreground Dust Map

### 2.3.2 Sky Masking

In this study, we focus on fitting regions of the sky predominantly near the Galactic poles, where the contamination from Galactic dust is minimal. By masking areas close to the Galactic plane, we exclude regions with significant foreground interference, allowing us to isolate the Cosmic Infrared Background (CIB) signal more effectively. This approach ensures that the analysis targets regions where the CIB is in the same order over Galactic dust, specifically in the unmasked area of approximately  $13,939 \text{ deg}^2$  area, covering portions of both the Northern and Southern Galactic poles which is about 33.8% of the total sky area.

- **Near the Galactic Plane:** The contribution from Galactic dust is greater than the Cosmic Infrared Background (CIB) due to the abundance of interstellar dust and molecular clouds. These regions are rich in dense matter, which emits strongly in the far-infrared and submillimeter wavelengths through thermal processes, overwhelming the faint extragalactic signals from the CIB.
- **Near the Galactic Poles:** The Cosmic Infrared Background (CIB) is in the same order over Galactic dust emission due to the significantly lower density of interstellar matter in these regions. This reduction in foreground contamination makes the poles ideal for isolating extragalactic signals, as the CIB arises from the cumulative emission of distant galaxies, providing insights into the large-scale structure of the universe.

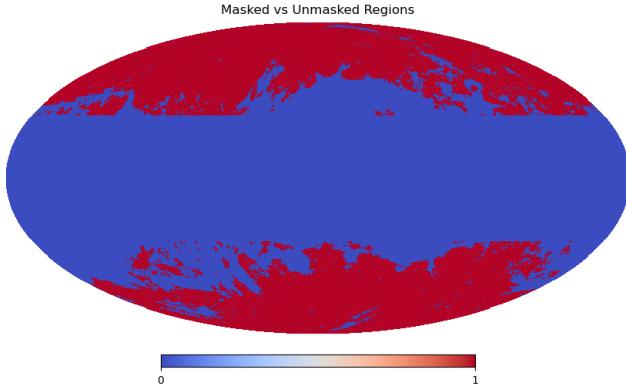


Fig. 2.8: Sky mask applied in the analysis. The **blue regions** indicate masked pixels, typically corresponding to high dust contamination or unreliable data near the Galactic plane. The **red regions** represent the **unmasked** or valid pixels used for CIB signal recovery and statistical analysis. This mask helps minimize foreground contamination and ensures accurate component separation.

### 2.3.3 Dust Maps

We incorporate multiple line-of-sight-dependent dust extinction  $E(B - V)$  maps from Green et al. 2015 [7], shown in Figure 2.5, as templates in our model. These maps effectively capture the spatial distribution and variation of dust clouds at different distances along the line of sight, providing a detailed representation of the interstellar medium. By accounting for these variations, the model can more accurately characterize the dust emission and its influence on the observed signal. This comprehensive approach significantly improves the accuracy of residual calculations, reducing contamination from foreground dust and enabling better isolation of extragalactic signals, such as the Cosmic Infrared Background (CIB), for subsequent analyses. Here Planck opacity map MBB fit is given in reference to the above section,

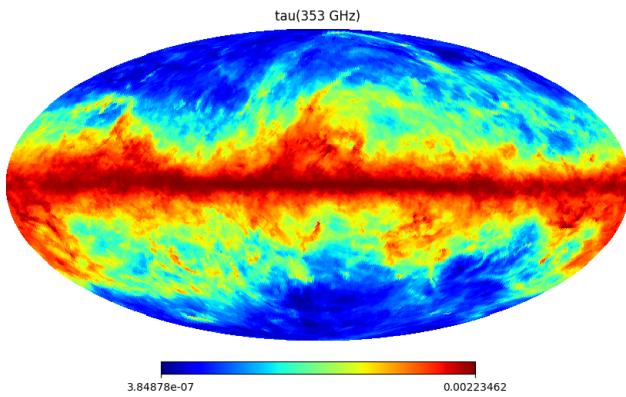


Fig. 2.9: Planck Opacity map at 353 GHz

This plot is used in the constructions from opacity map from fraction map at a reference frequency that is in our case is 353 GHz.

# Chapter 3

## Simulation

### 3.1 Construction of Simulated data

The correctly simulated map, constructed using six dust extinction templates with their respective emissivity values, shows strong morphological agreement with the Planck HFI Foregorund data. This validates the multi-template dust modeling approach in capturing the large-scale Galactic foreground structures observed in the Planck maps.

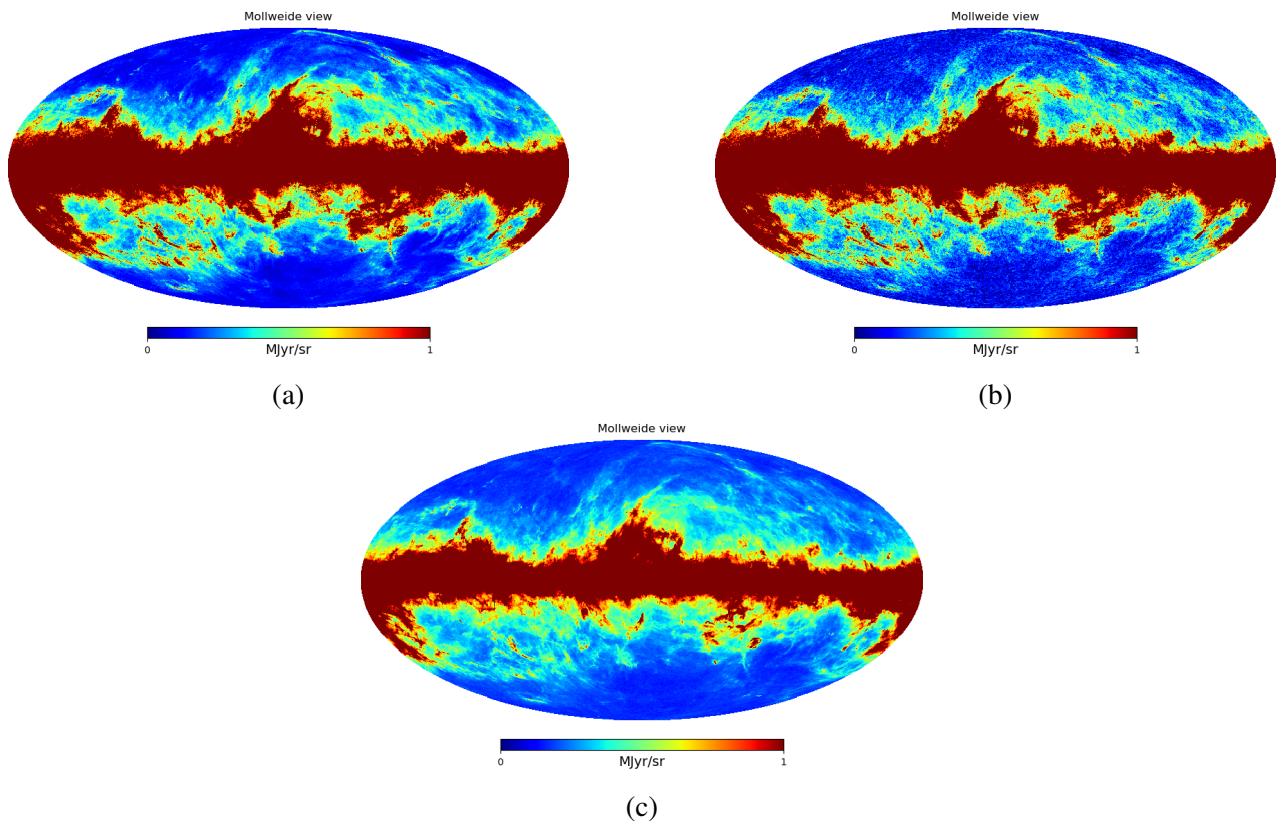


Fig. 3.1: Comparison of simulated and observed Galactic dust maps at  $N_{side}=512$  resolution. (a) Simulated map based solely on dust extinction templates and fitted emissivity values, capturing key features of the observed morphology. (b) Simulated map with added Cosmic Infrared Background (CIB) fluctuations, showing increased deviations from the observed data. (c) Simulated dust map, including thermal dust emission, CIB and instrument noise.

Figure 3.1 presents a visual comparison between simulated Galactic foreground maps and the

observed *Planck* dust map, all at a resolution of  $\text{Nside} = 512$ . The objective of this comparison is to determine which simulated model most accurately reproduces the observed Galactic dust structure and should therefore be used as input for the Hamiltonian Monte Carlo (HMC) fitting procedure.

Based on the visual alignment, the extinction-only simulated map in panel (a) provides the best match to the observed *Planck* data in terms of both structural features and dynamic range. Therefore, this version is adopted as the input model for the HMC sampling, ensuring that the Bayesian inference framework starts from a physically consistent and observationally validated foreground model.

### 3.2 Fitting the Simulated Dust Map using HMC sampling

For initial validation, the simulated map was fitted using Hamiltonian Monte Carlo (HMC) sampling at  $\text{nsamples}=5000$  to ensure that the Bayesian inference pipeline behaves as expected under ideal conditions. This step was crucial for verifying that the model setup, extinction templates, and emissivity scaling were correctly implemented. By assigning the known input parameters used to generate the simulated map, we confirmed that the HMC-based fitting process is capable of accurately recovering these values. This serves not only as a sanity check for the pixel-based segmentation and noise handling, but also establishes a robust foundation before applying the same methodology to real Planck data.

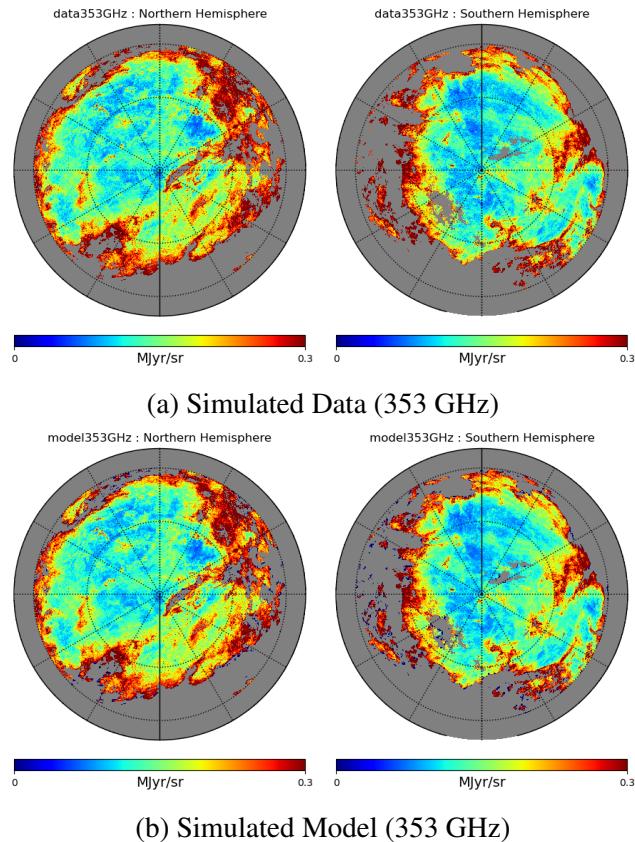


Fig. 3.2: Comparison between simulated data and model outputs at 353 GHz. Each subfigure shows Northern and Southern Hemisphere

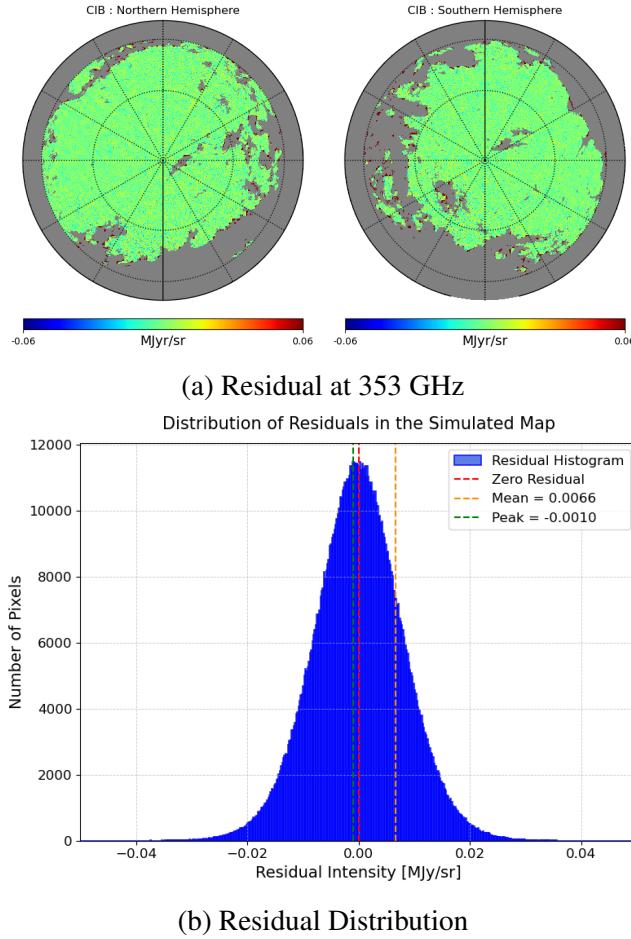


Fig. 3.3: Residual Map and its distribution (Simulated data - Model)

These results show a notable resemblance to the outcomes obtained during the fitting procedure performed on the actual Planck data using Hamiltonian Monte Carlo (HMC) sampling. The close similarity between the simulated and real-data-based fittings reinforces the reliability of the simulation pipeline. Furthermore, the distribution of the residuals exhibits an approximately Gaussian profile, which is in good agreement with theoretical expectations. This behavior suggests that the noise properties and model deviations are statistically well-behaved, further validating the robustness of the implemented method.

Here are the results where we try to fit the Planck data without HMC and then we have used the fitted mean dust-extinction coefficient and mean offset values in the HMC algorithm to fit the simulated map to check whether the values are correctly retrieved or not.

Model Parameters	Input(MJy/sr/mag)	Output(MJy/sr/mag)
$\epsilon_1$	10.03	$10.05 \pm 1.25$
$\epsilon_2$	7.93	$7.94 \pm 0.306$
$\epsilon_3$	3.67	$3.65 \pm 0.608$
$\epsilon_4$	6.10	$6.08 \pm 0.92$
$\epsilon_5$	36.82	$37.04 \pm 5.63$
$\epsilon_6$	147.76	$147.88 \pm 14.45$
$\emptyset_\nu$	0.0087	$0.008788 \pm 0.00004568$

Table 3.1: Comparison between input and output fitted parameters for each extinction template.

From these values, we observe that the input and output parameters are in very close agreement across all six templates. This consistency strongly indicates that the fitting procedure has been implemented correctly and that the HMC code is functioning as expected, accurately recovering the underlying parameters used in the simulation.

### 3.3 Simulation Processing Time

We thus compute the processing time of the Hamiltonian Monte Carlo code using simulated data at different nside and rnside configurations.

Samples	Time (minutes)					
	nside = 512		nside = 256		nside = 128	
	rnside = 32	rnside = 16	rnside = 32	rnside = 16	rnside = 32	rnside = 16
1000	282	450	113	67	101	39
2000	540	886	263	132	207	66
3000	796	1393	353	203	292	81
4000	1133	1661	530	275	416	103
5000	1320	2604	583	336	488	168

Table 3.2: Processing Time for Different nsides and rnside Configurations (in minutes)

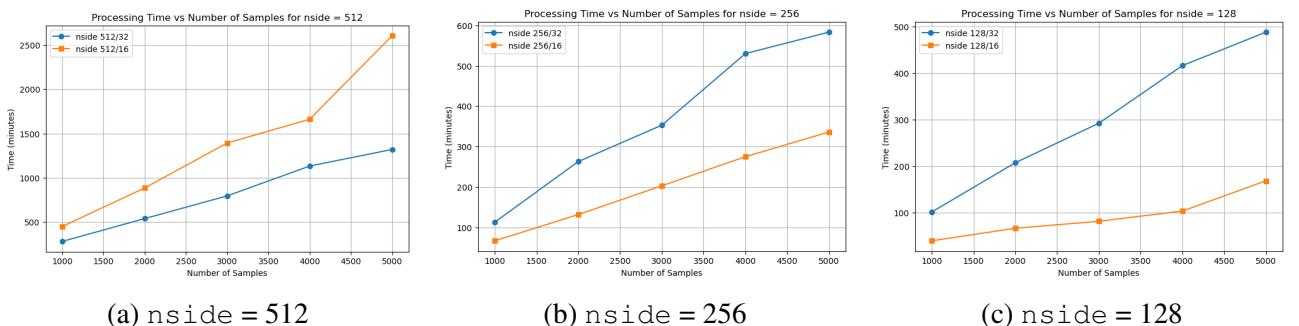


Fig. 3.4: Scaling of HMC processing time with number of samples for different HEALPix resolutions ( $nside = 512, 256, 128$ ) at fixed  $rnside = 32$  and  $16$ .

- **Processing Time Increases with Number of Samples:**

- As expected, the processing time increases as the number of samples grows from 1000 to 5000.
- This suggests that the computational cost scales approximately **linearly or slightly super-linearly** with the number of samples.
- The increase in processing time is consistent across all configurations, indicating that the number of samples is a significant factor in determining execution time.

- **Higher Nside Leads to Higher Processing Time:**

- Comparing the three plots, we observe that **higher nside values** result in significantly longer processing times.
- The total computation time follows the trend:

$$\text{nside } 128 < \text{nside } 256 < \text{nside } 512$$

- This is expected because increasing `nside` leads to a higher pixel resolution in HEALPix maps, increasing both memory usage and computational complexity.

- **Difference Between rNside 32 and rNside 16:**

- For a given `nside`, `rNside 32` consistently takes more time than `rNside 16`.
- The performance gap between the two configurations widens as the number of samples increases.
- This indicates that the choice of `rNside` also significantly affects the processing time, with `rNside 32` being computationally more expensive.

- **Scaling Behavior:**

- The processing time for `nside 512/16` exhibits **steeper growth** compared to `nside 512/32`.
- This suggests that for **higher nside**, the `rNside 16` configuration incurs a disproportionately higher cost compared to `rNside 32`.
- The processing time growth rate appears to be **linear**, likely due to factors such as increased memory access costs and numerical computation overhead.

- **Scaling of HMC Runtime at Nside = 512 and rNside = 32:**

- To characterize the computational performance of the Hamiltonian Monte Carlo (HMC) implementation, we analyzed how the total runtime scales with the number of samples for a fixed map resolution configuration (`nside = 512, rNside = 32`).
- The plot below shows runtime vs. sample size:

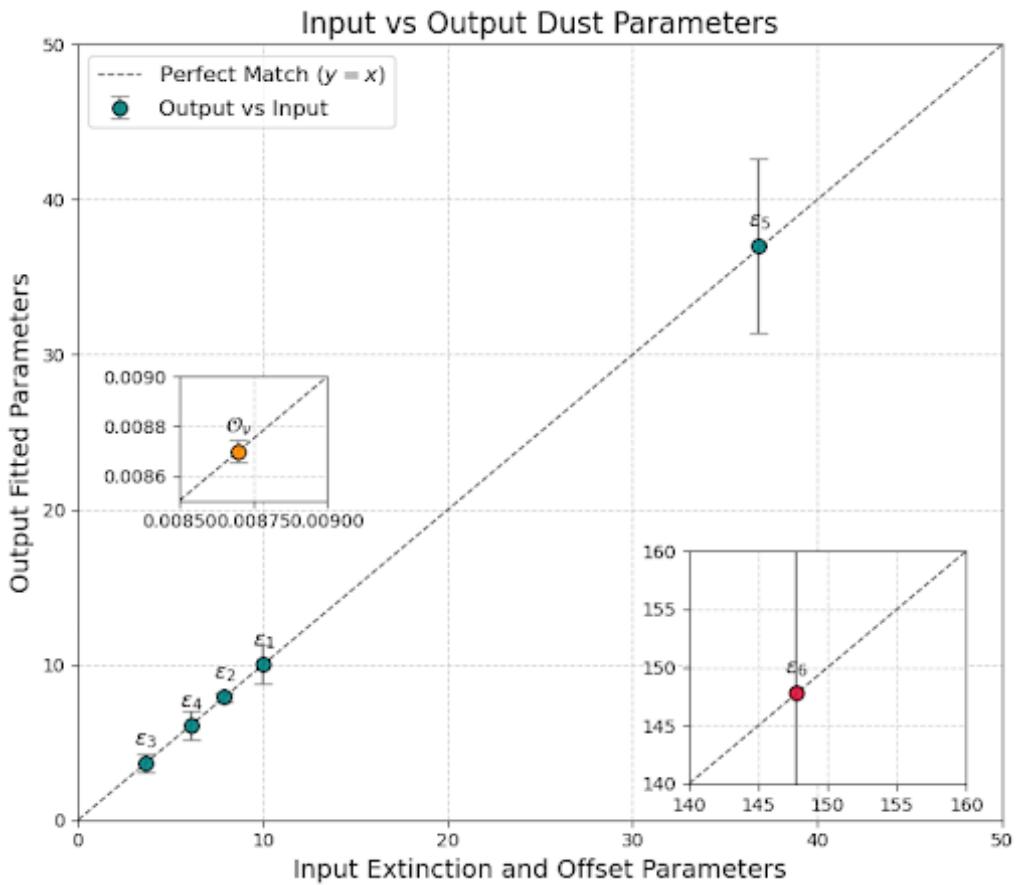


Fig. 3.5: Processing time vs samples at  $nside = 512$

- Fitting a power-law model of the form  $T = a \cdot N^b$  to the observed runtime measurements, we find:

$$T \approx 0.328 \cdot N^{0.98},$$

where  $T$  is the total runtime in minutes and  $N$  is the number of HMC samples.

- The fitted exponent  $b \approx 0.98$  indicates near-linear scaling, consistent with theoretical expectations for iterative MCMC sampling algorithms. This confirms that the computational cost grows predictably with sample size and validates the efficiency of the HMC implementation in this inference framework.

# Chapter 4

## Planck Data Fitting using HMC sampling

### 4.1 Data Fitting using the HMC sampling

In the fitting process, we utilize a total of  $N_{\text{samples}} = 5000$ , which represents the total number of steps in the Markov Chain Monte Carlo (MCMC) process. This ensures robust exploration of the parameter space. The sky is divided into superpixels and subpixels to achieve a balance between resolution and computational efficiency. Specifically, we adopt a superpixel resolution of the  $\text{Nside} = 32$ , with each superpixel containing 256 subpixels at a subpixel resolution of  $\text{Nside} = 512$ . This hierarchical approach allows detailed modeling of variations within the subpixels while maintaining computational feasibility at the superpixel scale, ensuring accurate representation of dust emission and other components in the analysis.

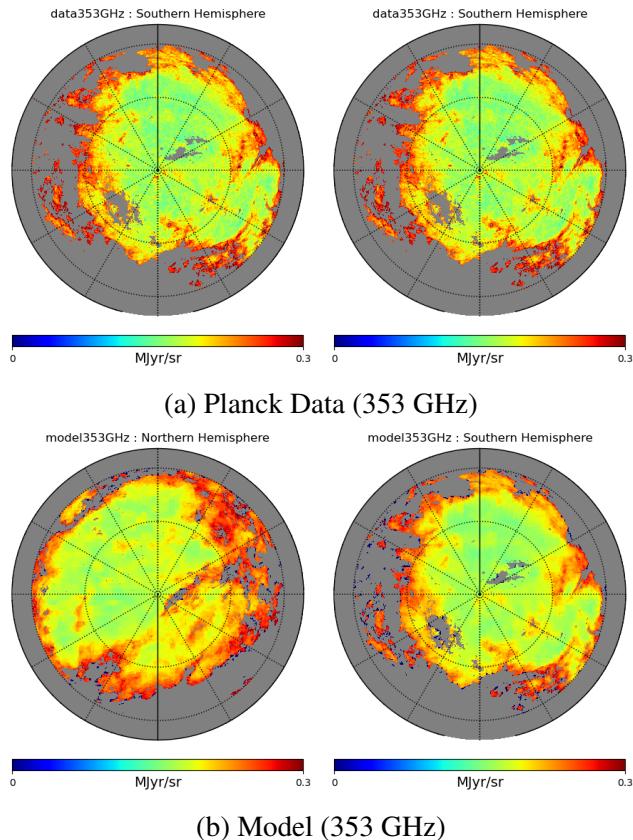


Fig. 4.1: Comparison between Observed Planck data and model outputs at 353 GHz. Each subfigure shows Northern and Southern Hemisphere

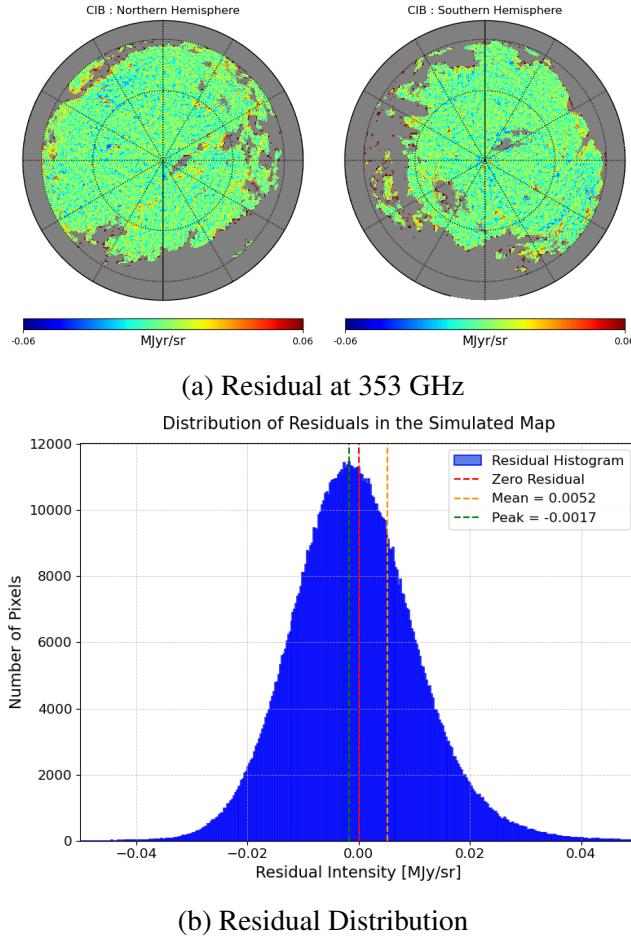


Fig. 4.2: Residual Map and its distribution (Observed - Model mp)

The comparison between the observed data and the model at 353 GHz demonstrates a high level of consistency in reproducing large-scale Galactic dust emission. The observed data reveal brighter regions near the Galactic plane due to the enhanced dust intensity, which is accurately captured by the model. This suggests that the model effectively reconstructs the dominant features of the dust emission. However, some differences are evident, particularly in localized regions, which may indicate the presence of instrumental noise and un-modeled foreground variations. These deviations highlight areas for potential refinement in the modeling approach.

The residual map, along with its corresponding pixel intensity distribution, closely mirrors the behavior observed in the simulated datasets. This similarity indicates that the model accurately captures the large-scale and small-scale structures present in the data. Moreover, the histogram of the residual values demonstrates an approximately Gaussian shape, which aligns well with theoretical expectations for a well-fitted model. Such a distribution suggests that the residuals are primarily dominated by random noise rather than systematic deviations, further supporting the validity and robustness of the simulation and fitting procedures.

## 4.2 Pixel-Dependent Dust-Extinction Coefficient maps

We extracted the dust extinction coefficient maps along with their associated error maps from the posterior distributions obtained through Hamiltonian Monte Carlo (HMC) sampling. The analysis was performed at an angular resolution corresponding to a HEALPix  $N_{\text{side}}$  of 32, which ensures a manageable balance between spatial resolution and statistical robustness in the inference process.

Given that our modeling framework incorporates six distinct dust extinction templates, each representing a specific distance range within the Milky Way, we obtain a total of six sets of extinction coefficient maps. Each of these maps corresponds to a different template and thereby captures the spatial variation in dust extinction across distinct layers along the line of sight. The corresponding uncertainty maps, derived from the posterior standard deviations of the fitted HMC chains, quantify the level of confidence in the inferred extinction structures.

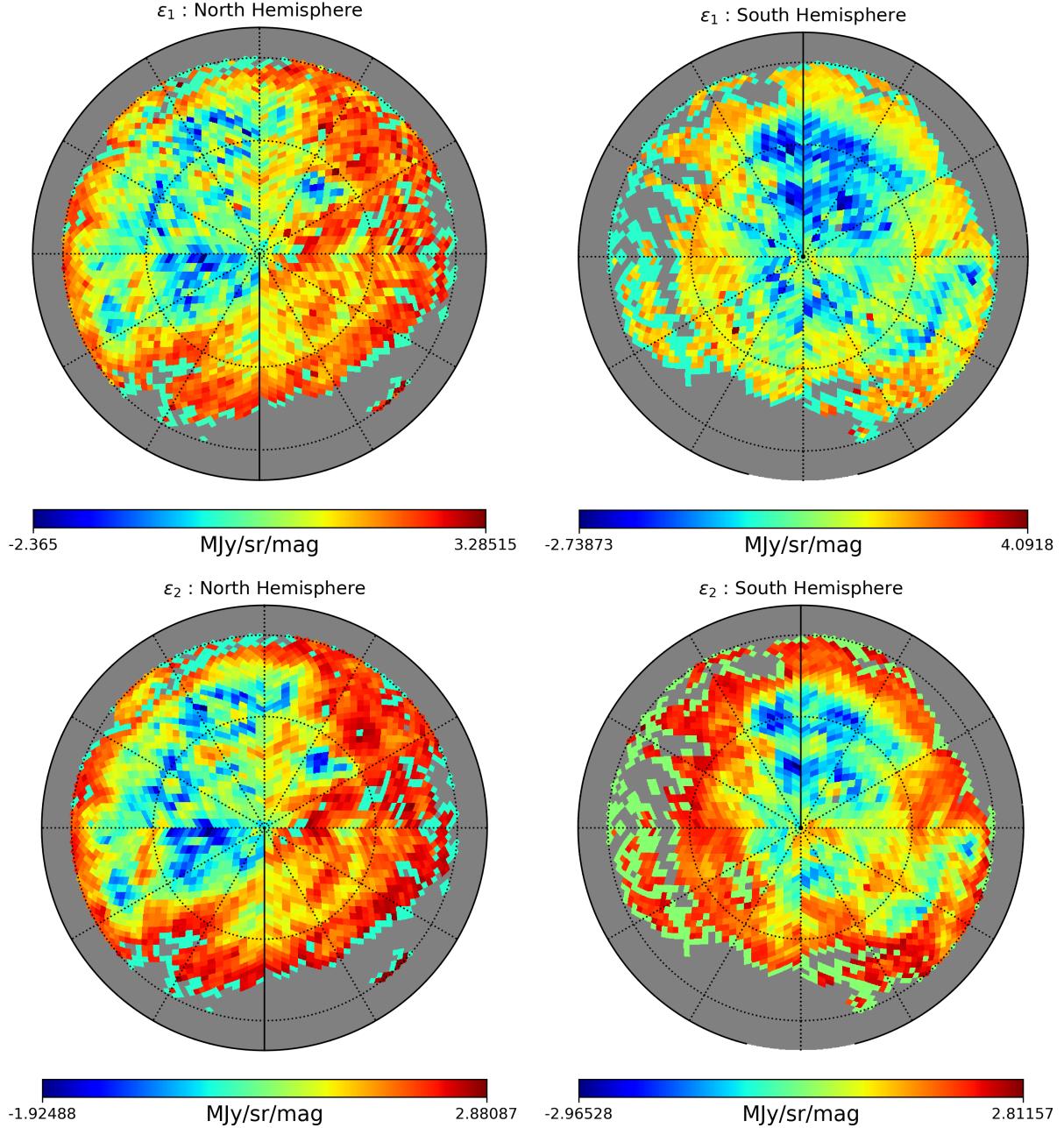


Fig. 4.3: Extinction coefficient maps for distance ranges 10–63 pc (left pair) and 63–251 pc (right pair), showing Northern and Southern hemispheres.

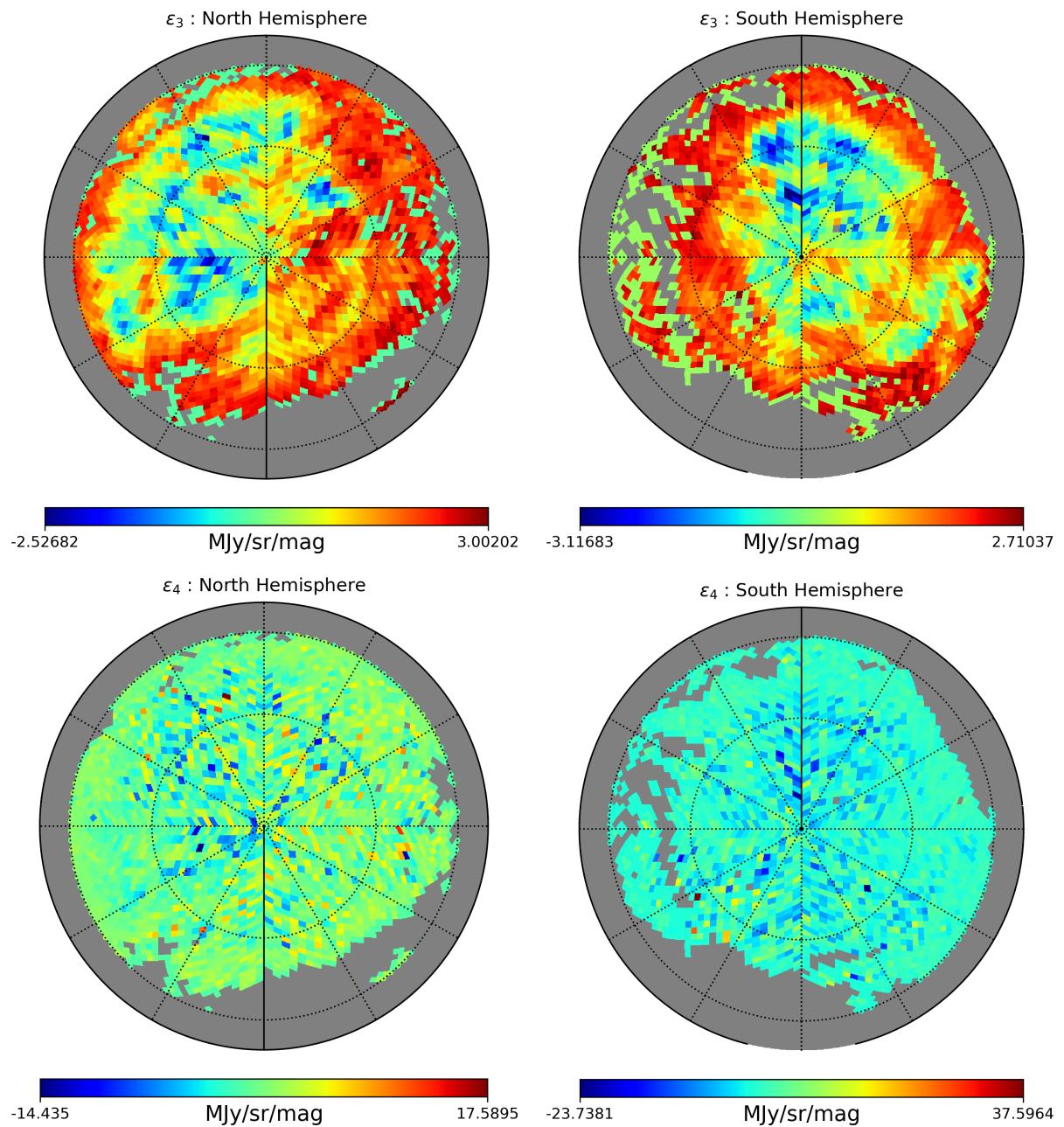


Fig. 4.4: Extinction coefficient maps for distance ranges 251–1000 pc (left pair) and 1000–3881 pc (right pair), showing Northern and Southern hemispheres.

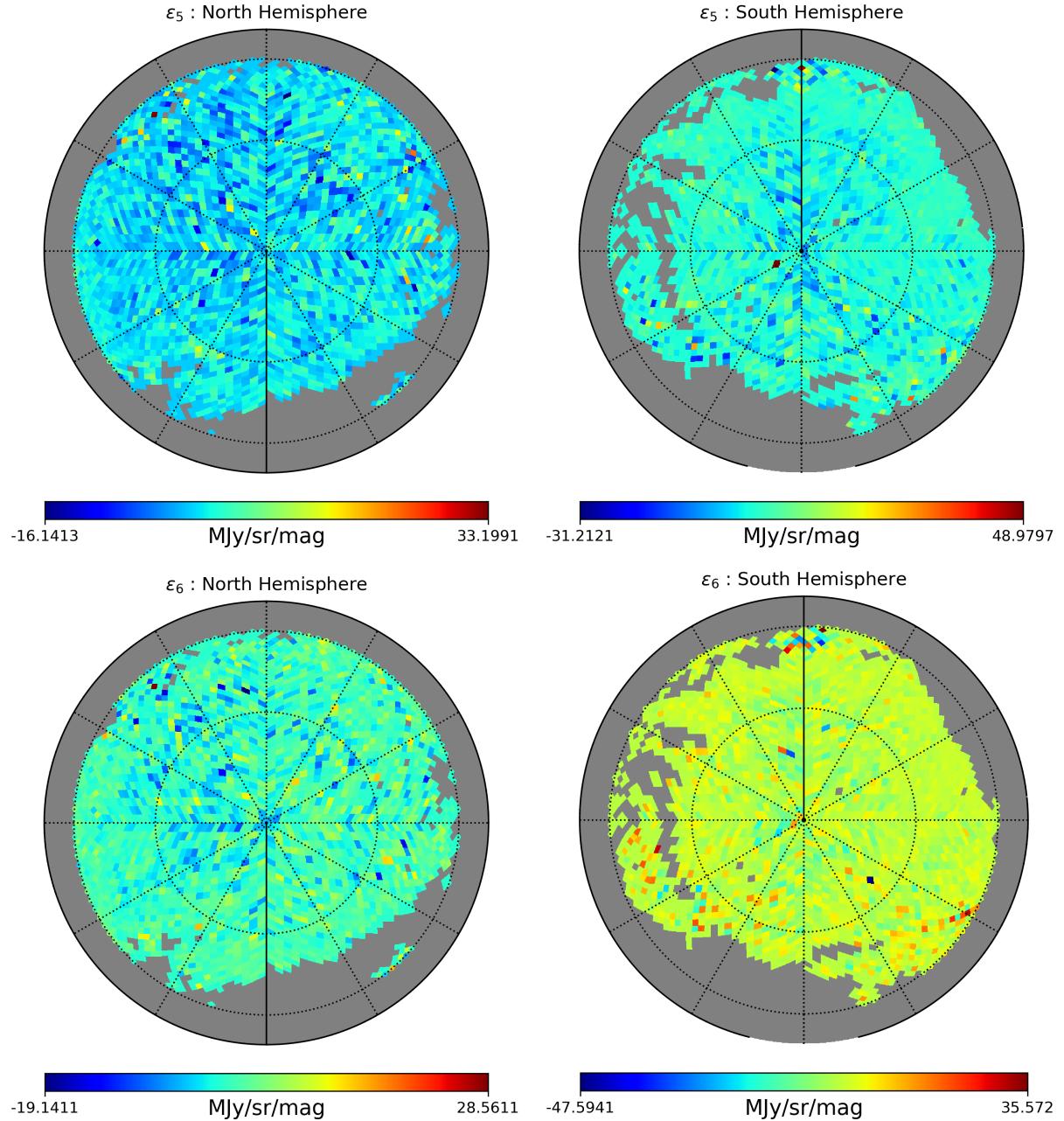


Fig. 4.5: Extinction coefficient maps for distance ranges 3881–15849 pc (left pair) and 15849–63096 pc (right pair), showing Northern and Southern hemispheres.

- **Three-Dimensional Structure and Emissivity Variation of Galactic Dust:** The extinction coefficient maps, obtained for six distinct distance layers, reveal the spatial and depth-dependent structure of dust emission in the Galaxy. Nearby layers (e.g., 10–63 pc) show relatively smooth distributions, while intermediate layers (e.g., 251–3881 pc) display filamentary and clumpy structures indicative of dense dust regions. This progression captures the tomographic complexity of the interstellar medium. The non-uniform coefficient values across the sky also indicate spatial variation in dust emissivity, suggesting localized differences in grain properties, temperature, or radiation field intensity. These maps highlight that dust emission cannot be accurately modeled with a single template and underscore the importance of incorporating multi-layer extinction data in foreground modeling.
- **Model Robustness and Uncertainty Characterization:** Moreover, the smoothness and realism of the coefficient distributions, coupled with reasonable uncertainty magnitudes,

validate the effectiveness of the HMC-based Bayesian inference framework in accurately capturing both the structure and uncertainty of Galactic dust.

From the Hamiltonian Monte Carlo (HMC) inference, we obtain the mean dust extinction coefficients corresponding to the six distance layers as follows:  $\epsilon_{\nu}^1 = 1.20$ ,  $\epsilon_{\nu}^2 = 1.19$ ,  $\epsilon_{\nu}^3 = 1.18$ ,  $\epsilon_{\nu}^4 = 1.16$ ,  $\epsilon_{\nu}^5 = 1.12$ , and  $\epsilon_{\nu}^6 = 1.11$ . All values are in the units of MJy/sr/mag. These values represent the average scaling of each extinction template required to fit the observed Galactic foreground emission at 353 GHz.

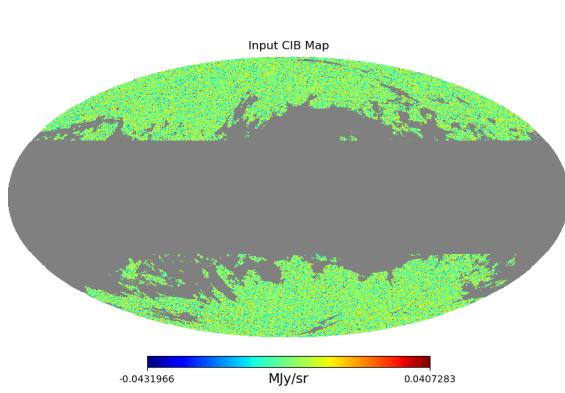
Additionally, the inferred global offset parameter, denoted by  $O_{\nu}$ , is found to be 0.169 MJy/sr. This offset accounts for the monopole component not captured by the extinction templates, including contributions from the Cosmic Infrared Background (CIB), zodiacal light, or instrument calibration uncertainties. The overall decreasing trend in  $\epsilon_{\nu}^i$  with increasing distance supports the expected attenuation of extinction contributions from more distant layers, as well as the cumulative nature of dust absorption along the line of sight.

# Chapter 5

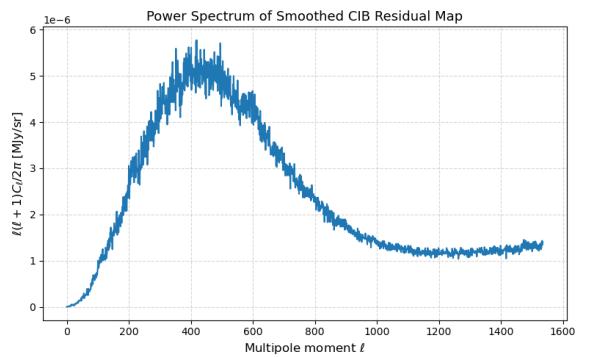
## Conclusions and Future Directions

### 5.1 CIB map and its Power spectrum

To further evaluate the performance of the component separation, we analyze the angular power spectra of the residual maps obtained from our HMC-based fitting pipeline. This includes both the simulated residual map—constructed by fitting synthetic Planck-like data—and the residual map derived from actual Planck observations. These spectra are compared against the input CIB power spectrum to quantify the recovery fidelity. A close match, especially at intermediate multipoles, indicates that the method effectively extracts the CIB signal while mitigating Galactic dust contamination. These power spectrum's are plotted using the valid pixels. The maps are all multiplied with masked regions we are using in this project, and the power spectrum are also calculated using the valid pixels only.



(a) Input CIB map at 353 GHz



(b) Angular power spectrum of the input CIB map

Fig. 5.1: Input CIB map at 353 GHz and its corresponding angular power spectrum.

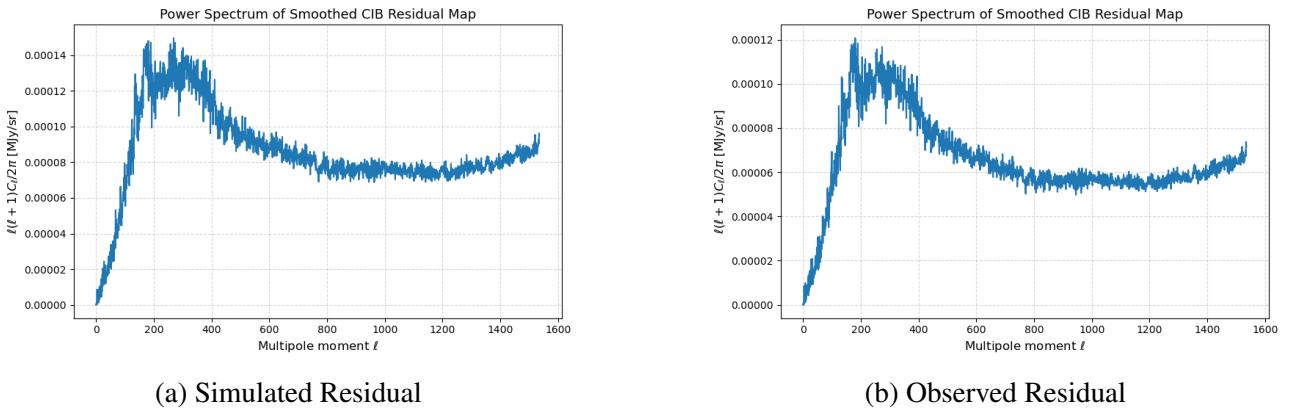


Fig. 5.2: Power spectrum of simulated and observed residual maps

The input power spectrum exhibits the expected rise at intermediate multipoles followed by a gradual decline, characteristic of CIB anisotropies. The residual power spectrum obtained from the simulated data closely follows the input spectrum across a wide multipole range, validating the correctness of the simulation pipeline and the effectiveness of the HMC-based inference in recovering the CIB structure.

The observed residual power spectrum also qualitatively matches the input spectrum, particularly in the range  $100 \lesssim l \lesssim 700$ , though slight deviations appear at higher multipoles. These discrepancies may arise from remaining Galactic foregrounds, instrument noise, or partial masking effects. Nonetheless, the overall consistency across all three spectra supports the robustness of our component separation approach and its ability to isolate the CIB signal from observational data.

## 5.2 Correlation between the Dust extinction coefficient Maps

To investigate the mutual dependencies among the six distance-dependent dust extinction coefficient maps  $\{\epsilon_\nu^i\}_{i=1}^6$ , we performed a pixel-wise correlation analysis across the full sky. These extinction coefficients were inferred from the Hamiltonian Monte Carlo (HMC) fitting procedure using the Planck 353 GHz foreground data and Green et al. (2015) extinction templates.

For each HEALPix pixel not masked in the mask map, we extracted the corresponding set of dust extinction coefficient values from the six coefficient maps. This yielded a valid dataset of shape  $(N_{\text{pix}}, 6)$ , where each row represents the vector of inferred dust extinction coefficients at a given sky location.

We explored the joint statistical distributions of these parameters using two complementary approaches:

- **Corner Plot:** A global corner plot was generated to visualize the marginal distributions and pairwise covariances of the six extinction coefficients. This provided qualitative insight into the degree of correlation and degeneracy between adjacent distance layers, as expected due to the projection effects and line-of-sight overlap inherent in dust extinction modeling.
- **Pearson Correlation Matrix:** We quantified the linear correlation between each pair of extinction maps by computing the Pearson correlation coefficient:

$$r_{ij} = \frac{\text{Cov}(\epsilon_\nu^i, \epsilon_\nu^j)}{\sigma_{\epsilon_\nu^i} \sigma_{\epsilon_\nu^j}},$$

where  $\text{Cov}(\cdot, \cdot)$  denotes the covariance and  $\sigma$  is the standard deviation of each coefficient map over the unmasked pixels. The resulting  $6 \times 6$  matrix was visualized as a heatmap, allowing direct interpretation of the strength and sign of correlations. A high correlation (e.g.,  $r_{ij} \sim 0.8$ ) indicates that two distance layers share significant structural similarity, likely due to their overlapping dust contributions.

This analysis helps to identify degeneracies in the extinction model and informs the interpretation of the posterior parameter uncertainties. It also validates the smooth spatial behavior and mutual consistency of the inferred dust extinction coefficients across Galactic distance.

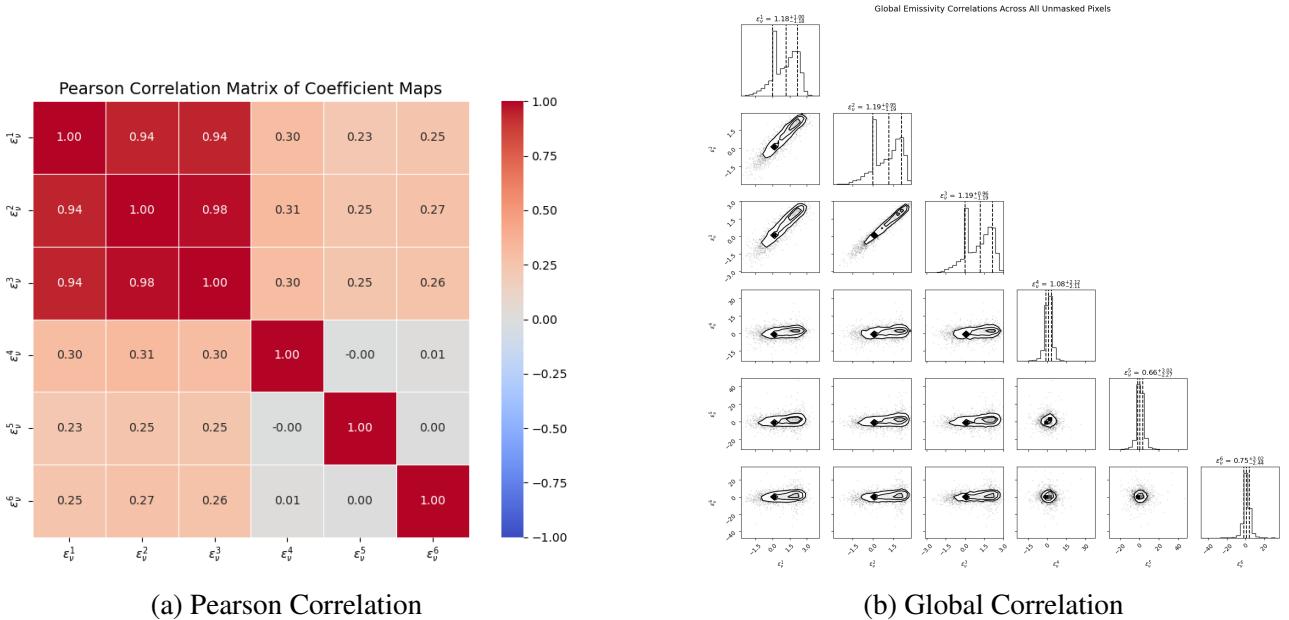


Fig. 5.3: Correlation Plots

- **Correlated behavior among nearby extinction layers:** The corner plot reveals strong positive correlations between the first three extinction coefficients ( $\epsilon_\nu^1$ ,  $\epsilon_\nu^2$ , and  $\epsilon_\nu^3$ ), as shown by the elongated joint distributions and overlapping marginal histograms. This observation is quantitatively supported by the Pearson correlation matrix, where the coefficients among these layers exceed  $r > 0.94$ . These strong correlations are expected, as nearby layers trace overlapping regions of the Galactic plane and share common structural features in the dust emission. The result highlights that line-of-sight degeneracies are more prominent in the local interstellar medium.
- **Statistical independence at larger distances:** In contrast, the extinction coefficients corresponding to the distant Galactic layers ( $\epsilon_\nu^4$ ,  $\epsilon_\nu^5$ , and  $\epsilon_\nu^6$ ) show considerably weaker correlations with one another and with the nearby layers, typically with Pearson coefficients around  $r \sim 0.25$  or less. This behavior reflects the increasing spatial separation of dust structures at high distances and reduced signal-to-noise in the outer Galaxy. It indicates that the HMC-inferred parameters are able to statistically disentangle distant dust layers, thereby validating the tomographic capability of the multi-template modeling approach.

## 5.3 Comparison of the input and observed CIB map

In this study, we use Minkowski Functionals (MFs) to quantify the morphology and statistical properties of the normalized Cosmic Infrared Background (CIB) maps. The MFs serve as

higher-order statistics beyond the power spectrum and are particularly sensitive to non-Gaussian features in the data.

### 5.3.1 Map Normalization and Derivatives

The input map  $u(\hat{n})$ , defined on the sphere, is first normalized within the valid (unmasked) region as:

$$u(\hat{n}) = \frac{T(\hat{n}) - \mu}{\sigma}, \quad (5.1)$$

where  $T(\hat{n})$  is the input intensity map,  $\mu$  and  $\sigma$  are the mean and standard deviation of the unmasked pixel values, and  $\hat{n}$  denotes a direction on the sphere.

From the normalized map, we compute the first derivatives  $\nabla u$  using the spherical harmonic representation:

$$|\nabla u| = \sqrt{\left(\frac{\partial u}{\partial \theta}\right)^2 + \left(\frac{1}{\sin \theta} \frac{\partial u}{\partial \phi}\right)^2}, \quad (5.2)$$

which yields the gradient magnitude map, required for evaluating contour lengths. We also compute the second derivatives using spherical harmonics and use them to evaluate the local curvature  $\kappa$ , defined as:

$$\kappa = \frac{2u_1u_2u_{12} - u_1^2u_{22} - u_2^2u_{11}}{(u_1^2 + u_2^2)} \cdot \frac{1}{2\pi}, \quad (5.3)$$

where  $u_1 \equiv \frac{\partial u}{\partial \theta}$ ,  $u_2 \equiv \frac{1}{\sin \theta} \frac{\partial u}{\partial \phi}$ , and the second derivatives  $u_{11}, u_{12}, u_{22}$  are computed numerically.

### 5.3.2 Definition of Minkowski Functionals

We compute the three primary Minkowski Functionals  $V_0(\nu)$ ,  $V_1(\nu)$ , and  $V_2(\nu)$  as a function of threshold  $\nu$ , where the threshold ranges from  $\nu = -5$  to  $+5$  in steps of  $\delta = 0.5$ . The functionals are defined as:

$$V_0(\nu) = \frac{1}{A} \int_{\Sigma_\nu} dA, \quad (\text{area fraction}) \quad (5.4)$$

$$V_1(\nu) = \frac{1}{4A} \int_{\partial\Sigma_\nu} |\nabla u| dl, \quad (\text{contour length}) \quad (5.5)$$

$$V_2(\nu) = \frac{1}{2\pi A} \int_{\partial\Sigma_\nu} \kappa dl, \quad (\text{Euler characteristic}) \quad (5.6)$$

Here,  $\Sigma_\nu$  is the excursion set defined by  $\{\hat{n} \mid u(\hat{n}) > \nu\}$ ,  $\partial\Sigma_\nu$  is its boundary,  $\kappa$  is the geodesic curvature, and  $A$  is the total area of the valid unmasked sky. In the discrete implementation, these expressions become:

$$V_0(\nu) \approx \frac{1}{N_{\text{pix}}} \sum_{i \in \Sigma_\nu} 1, \quad (5.7)$$

$$V_1(\nu) \approx \frac{1}{\delta N_{\text{pix}}} \sum_{i \in \partial \Sigma_\nu} \frac{1}{4} |\nabla u_i|, \quad (5.8)$$

$$V_2(\nu) \approx \frac{1}{\delta N_{\text{pix}}} \sum_{i \in \partial \Sigma_\nu} \kappa_i. \quad (5.9)$$

### 5.3.3 Analytical Expectation for Gaussian Random Fields

For comparison, we compute the theoretical Minkowski Functionals  $V_0^{(G)}(\nu)$ ,  $V_1^{(G)}(\nu)$ , and  $V_2^{(G)}(\nu)$  expected from a Gaussian random field (GRF) as:

$$V_0^{(G)}(\nu) = \frac{1}{2} \operatorname{erfc} \left( \frac{\nu}{\sqrt{2}} \right), \quad (5.10)$$

$$V_1^{(G)}(\nu) = \frac{1}{8\sqrt{2}} \frac{1}{r_c} \exp \left( -\frac{\nu^2}{2} \right), \quad (5.11)$$

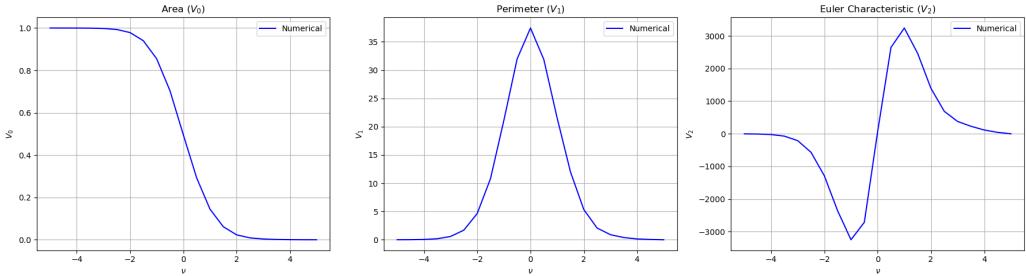
$$V_2^{(G)}(\nu) = \frac{1}{4\sqrt{2}\pi^{3/2}} \frac{1}{r_c^2} \nu \exp \left( -\frac{\nu^2}{2} \right), \quad (5.12)$$

where  $r_c = \sigma_0/\sigma_1$  is the correlation length,  $\sigma_0^2 = \langle u^2 \rangle = 1$ , and  $\sigma_1^2 = \langle |\nabla u|^2 \rangle$ , computed from the gradient map.

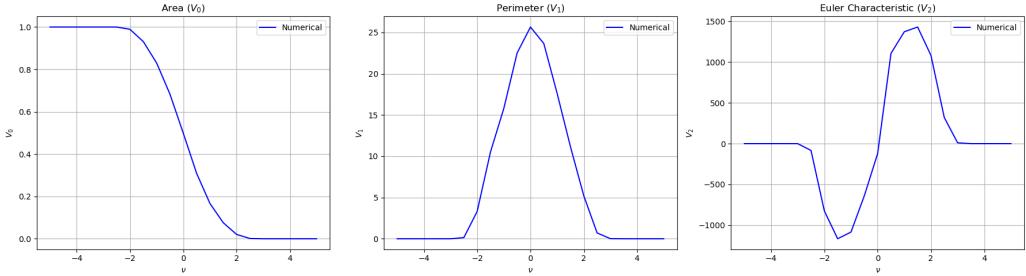
To assess the fidelity and statistical similarity of our recovered CIB maps, we compare the MFs computed from two key maps: (i) the observed CIB + noise map, obtained after dust foreground removal and (ii) the simulated input CIB + noise map, generated under the assumption that the input CIB is Gaussian. Importantly, the Gaussian approximation for the CIB is strictly valid only over the full sky. However, due to masking of contaminated regions, the MF analysis is restricted to a partial sky, and the influence of masking must be carefully considered when interpreting non-Gaussian signatures.

As a result, it is more informative to compare the MFs of the input CIB + noise with the observed CIB + noise, rather than with the pure input CIB. This approach removes the need for unreliable Gaussian predictions on the masked sky. By definition, the input CIB is modeled as a Gaussian random field, and if the MFs from the observed map match those from the input CIB + noise map over the masked sky, we can infer that the output CIB reconstruction is consistent with a Gaussian field. Therefore, we focus on this data-driven comparison. This strategy provides a more robust test of the Gaussianity of the recovered CIB signal under realistic sky coverage and observational noise.

- **Panel (a):** Simulated CIB map obtained by adding Gaussian CIB anisotropies to the multi-template dust model,
- **Panel (b):** Observed CIB map recovered as the residual after dust model subtraction from real Planck 353 GHz data.



(a) Minkowski Functionals for Simulated CIB + Noise map



(b) Minkowski Functionals for Observed CIB + Noise map

Fig. 5.4: Comparison of numerical Minkowski Functionals (blue solid lines) for the Simulated and Observed CIB + noise maps.

The two panels in Figure 5.4 display the morphological statistics across different threshold levels  $\nu$ , specifically:

- $V_0(\nu)$ : Fractional area of excursion sets,
- $V_1(\nu)$ : Total contour length (perimeter),
- $V_2(\nu)$ : Euler characteristic (topological complexity).

**Simulated CIB + Noise Map (Fig. 5.4a):** This map is created by adding Gaussian realizations of CIB anisotropies to instrumental noise and dust model components. The Minkowski Functionals  $V_0(\nu)$ ,  $V_1(\nu)$ , and  $V_2(\nu)$  computed from this simulated map exhibit the expected symmetric and smooth trends across thresholds, characteristic of Gaussian morphology. The area functional  $V_0$  shows a sharp transition near  $\nu = 0$ , while the perimeter  $V_1$  peaks symmetrically, and the Euler characteristic  $V_2$  displays a transition from negative to positive topology. These results confirm that the simulated CIB + noise map behaves as expected for a Gaussian random field under realistic sky coverage.

**Observed CIB + Noise Map (Fig. 5.4b):** The Minkowski Functionals for the observed residual map, obtained after foreground dust subtraction, show deviations in both amplitude and shape relative to the simulated case. Notably, the  $V_2$  functional is more asymmetric and exhibits a less pronounced peak, suggesting possible non-Gaussian features. These discrepancies may arise from residual foreground contamination, imperfect modeling of the dust component, or instrumental systematics. The deviations highlight subtle structural differences between the recovered CIB map and an ideal Gaussian simulation.

The comparison between the Minkowski Functionals of the simulated and observed CIB + noise maps over the masked sky provides a model-independent test of Gaussianity. Although the observed

map generally follows the trends seen in the simulation, the differences — especially in topological measures — point to physical or systematic effects that warrant further investigation. Nonetheless, the similarity in global behavior supports the overall effectiveness of the component separation method.

## 5.4 Conclusion

In this work, we developed a Bayesian framework utilizing Hamiltonian Monte Carlo (HMC) sampling to model and separate the Cosmic Infrared Background (CIB) from Galactic dust emission. By incorporating multiple line-of-sight-dependent dust extinction templates from Green et al. (2015), our method overcomes limitations of single-template HI-based models and allows for a more physically realistic decomposition of the dust emission structure.

Simulated dust maps, constructed using known emissivity values and extinction templates, were shown to closely match the Planck 353 GHz foreground data. This validated our simulation pipeline and justified its use as input for HMC-based inference. Upon applying the method to actual Planck data, the model reproduced large-scale Galactic dust structures with high fidelity, and the resulting residuals displayed approximately Gaussian behavior, suggesting that the model effectively captures the dominant emission components.

The angular power spectrum analysis of both simulated and observed residual maps reveals that our multi-template approach reliably reconstructs the input CIB structure, particularly at intermediate angular scales ( $100 \lesssim \ell \lesssim 700$ ). The residuals from Planck data closely follow the simulated power spectra, with only minor deviations at high multipoles, likely arising from unmodeled noise or residual foregrounds. Additionally, the dust extinction coefficient maps inferred for different line-of-sight distance bins show clear statistical structure: strong correlations among nearby layers and decreased correlations for distant ones.

Furthermore, the Minkowski Functional (MF) analysis provides a complementary morphological validation of the recovered CIB map. By comparing the MFs of the simulated input CIB + noise map with those of the observed residual CIB + noise map over the masked sky, we demonstrate that the reconstructed signal preserves the statistical properties of a Gaussian random field. The close alignment of the MF curves—especially V0 and V1—between these two maps suggests that the component separation is effective not only in a power-spectrum sense but also in higher-order topological terms. Minor deviations observed in the V2 functional (Euler characteristic) highlight regions where subtle non-Gaussian residuals may persist, potentially offering insight into limitations of dust modeling or intrinsic sky features. Overall, the joint spectral and morphological consistency confirms the robustness and reliability of our methodology for full-sky CIB recovery.

## 5.5 Future Directions

We observe a slight positive tail in the residual distribution, indicating potential contributions from unmodeled foreground components or clustered Cosmic Infrared Background (CIB) structures. To address this, we propose refining the analysis by reducing the masking regions, allowing for a broader sky coverage, and increasing the number of MCMC samples ( $N_{\text{samples}} = 10,000$ ) to ensure a more robust exploration of the parameter space. These adjustments aim to minimize the residual asymmetry and achieve a better fit to the observed data. And will try to check the non-gaussianity in the Observed residual maps using Minkowski Functionals in a more refined way.

# Bibliography

- [1] Planck Collaboration. Planck 2013 results. xi. all-sky model of thermal dust emission. *Astronomy & Astrophysics*, 571:A11, 2014. [doi:10.1051/0004-6361/201321525](https://doi.org/10.1051/0004-6361/201321525).
- [2] B. T. Draine and A. Li. Infrared emission from interstellar dust. iv. the silicate-graphite-pah model in the post-spitzer era. *The Astrophysical Journal*, 657(2):810–837, 2007. [doi:10.1086/511055](https://doi.org/10.1086/511055).
- [3] Adak. et al. Bayesian inference methodology to characterize the dust emissivity at far-infrared and submillimeter frequencies. *Monthly Notices of the Royal Astronomical Society*, 531:4876–4890, 2024. [doi:10.1093/mnras/stae1365](https://doi.org/10.1093/mnras/stae1365).
- [4] D. P. Finkbeiner, M. Davis, and D. J. Schlegel. Extrapolation of galactic dust emission at 100 microns to cosmic microwave background radiation frequencies using firas. *The Astrophysical Journal*, 524(2):867–886, 1999. [doi:10.1086/307852](https://doi.org/10.1086/307852).
- [5] Tuhin Ghosh, J. Delabrouille, M. Remazeilles, and F. Boulanger. Bayesian component separation: the commander approach using a hamiltonian monte carlo sampler. *Monthly Notices of the Royal Astronomical Society*, 469(3):3870–3887, 2017. [doi:10.1093/mnras/stx1091](https://doi.org/10.1093/mnras/stx1091).
- [6] R. Gispert, G. Lagache, and J.-L. Puget. Implications of the cosmic infrared background for galaxy evolution. *Astronomy & Astrophysics*, 360:1–9, 2000.
- [7] D. et al. Green. A 3d model of polarized dust emission in the milky way. *The Astrophysical Journal*, 805:188, 2015. [doi:10.1088/0004-637X/805/2/188](https://doi.org/10.1088/0004-637X/805/2/188).
- [8] B. S. Hensley and B. T. Draine. Far-infrared to millimeter emission from interstellar dust: new models and implications. *The Astrophysical Journal*, 906(1):73, 2021. [doi:10.3847/1538-4357/abc59d](https://doi.org/10.3847/1538-4357/abc59d).
- [9] T. R. Jaffe, J. P. Leahy, A. J. Banday, A. W. Strong, W. Reich, P. Reich, and R. Wielebinski. Modeling the galactic magnetic field on the plane in two dimensions. *Monthly Notices of the Royal Astronomical Society*, 431(1):683–694, 2013. [doi:10.1093/mnras/stt189](https://doi.org/10.1093/mnras/stt189).
- [10] Alexander Kashlinsky. Cosmic infrared background and early galaxy evolution. *The Astrophysical Journal*, 601(1):1–12, 2004. [doi:10.1086/379086](https://doi.org/10.1086/379086).
- [11] Alexander Kashlinsky et al. Looking at cosmic near-infrared background radiation anisotropies. *The Astrophysical Journal*, 867(2):55, 2018. URL: <https://arxiv.org/abs/1802.07774>, [doi:10.3847/1538-4357/aae4c8](https://doi.org/10.3847/1538-4357/aae4c8).
- [12] G. Martínez-Solaeche, A. Karakci, and J. Delabrouille. A 3d model of polarized dust emission in the milky way. *Monthly Notices of the Royal Astronomical Society*, 476(1):1310–1330, 2018. URL: <https://hal.archives-ouvertes.fr/hal-01768156>, [doi:10.1093/mnras/sty204](https://doi.org/10.1093/mnras/sty204).

- [13] A. M. Meisner and D. P. Finkbeiner. Modeling thermal dust emission with planck, iras, and wise maps. *The Astrophysical Journal*, 798(2):88, 2015.  
`doi:10.1088/0004-637X/798/2/88.`
- [14] Planck Collaboration. Planck intermediate results. xxix. all-sky dust modelling with planck, iras, and wise. *Astronomy & Astrophysics*, 586:A132, 2016.  
`doi:10.1051/0004-6361/201424082.`
- [15] Planck Collaboration Int. XVII. Planck intermediate results. xvii. emission of dust in the diffuse interstellar medium from the far-infrared to microwave frequencies. *Astronomy & Astrophysics*, 566:A55, 2014. `doi:10.1051/0004-6361/201323195.`
- [16] Planck Collaboration XVIII. The power spectrum of cosmic infrared background anisotropies. *Astronomy & Astrophysics*, 536:A18, 2011. `doi:10.1051/0004-6361/201116472.`
- [17] E. F. Schlafly and D. P. Finkbeiner. Measuring reddening with sloan digital sky survey stellar spectra and recalibrating sfd. *The Astrophysical Journal*, 737(2):103, 2011.  
`doi:10.1088/0004-637X/737/2/103.`
- [18] Marco Tucci and Carlo Baccigalupi. Cosmic infrared background anisotropies as a window into primordial non-gaussianity. *Monthly Notices of the Royal Astronomical Society*, 460:3678–3690, 2016. URL:  
<https://academic.oup.com/mnras/article/460/3/3678/2589843>,  
`doi:10.1093/mnras/stw1245.`
- [19] Marco Tucci and Carlo Baccigalupi. Exploring the non-gaussianity of the cosmic infrared background and its weak gravitational lensing. *Monthly Notices of the Royal Astronomical Society*, 456:1005–1015, 2016. URL:  
<https://academic.oup.com/mnras/article/456/2/1005/1068928>,  
`doi:10.1093/mnras/stv2753.`
- [20] F. Vansyngel, T. Ghosh, A. J. Banday, A. Rotti, F. Boulanger, and M. Tristram. Statistical simulations of dust polarization maps. *Astronomy & Astrophysics*, 603:A62, 2017.  
`doi:10.1051/0004-6361/201629577.`

## Appendix A : Fitted Error maps

To complement the extinction coefficient maps presented in the main sections of this report, we provide here the corresponding error maps associated with the emissivity estimation at each distance bin. These error maps, computed as the standard deviation of the posterior samples for each pixel, offer insight into the uncertainty in the inferred dust emission profiles across the sky.

Each panel shows the error map split into the Northern and Southern hemispheres, matching the spatial resolution ( $r_{\text{nside}} = 32$ ) of the extinction maps. As expected, the uncertainty is generally higher in regions with lower observational coverage or higher Galactic latitude. The following figures present the error maps for the six distance intervals used in the modeling, ranging from 10 pc to 63 kpc.

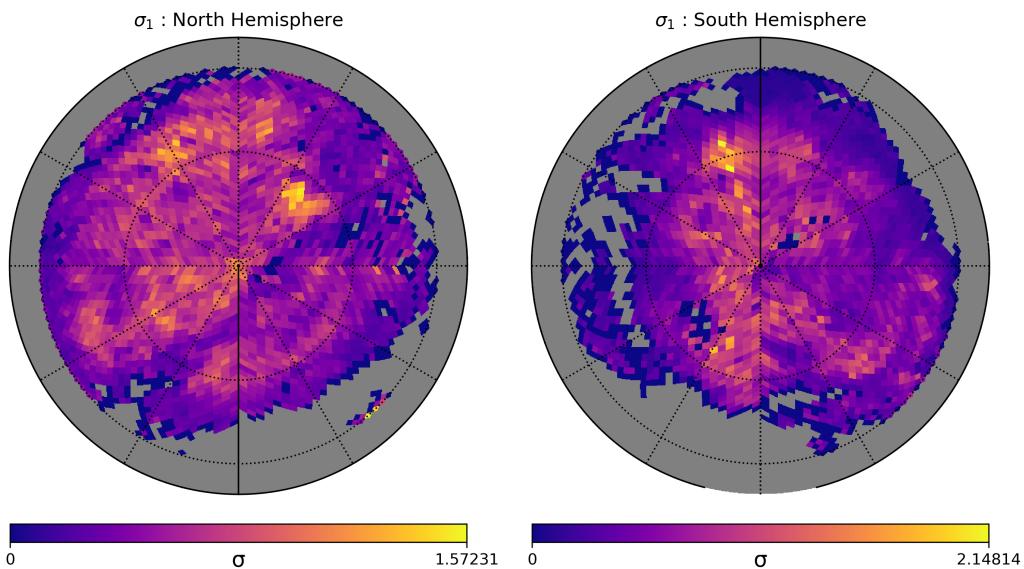


Fig. 5: Error maps (standard deviation of inferred emissivity) for distance range 10–63 pc, shown separately for the Northern and Southern hemispheres.

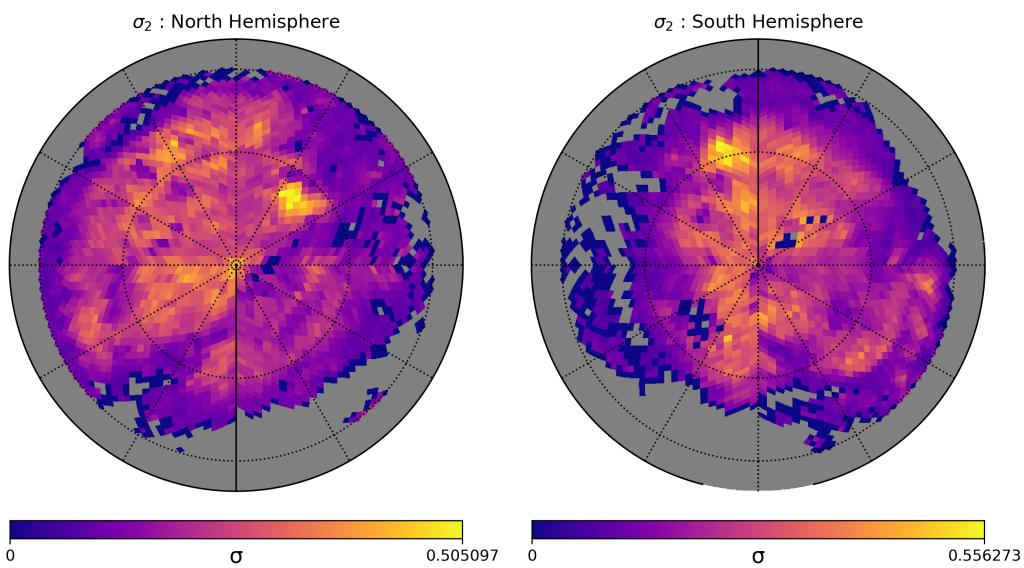


Fig. 6: Error maps for distance range 63–251 pc. These maps reflect the uncertainty in the emissivity fitting across both hemispheres.

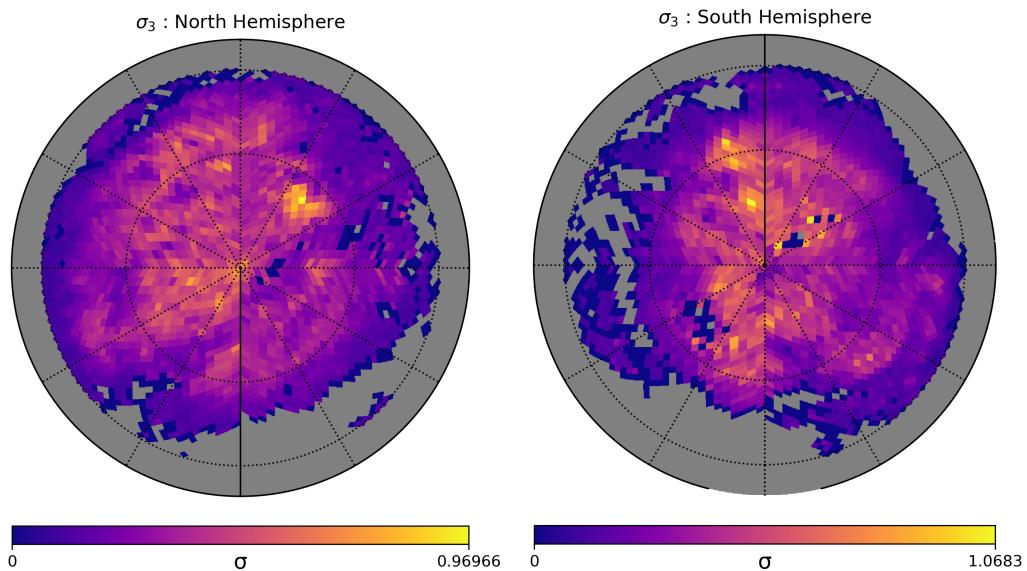


Fig. 7: Error maps for distance range 251–1000 pc.

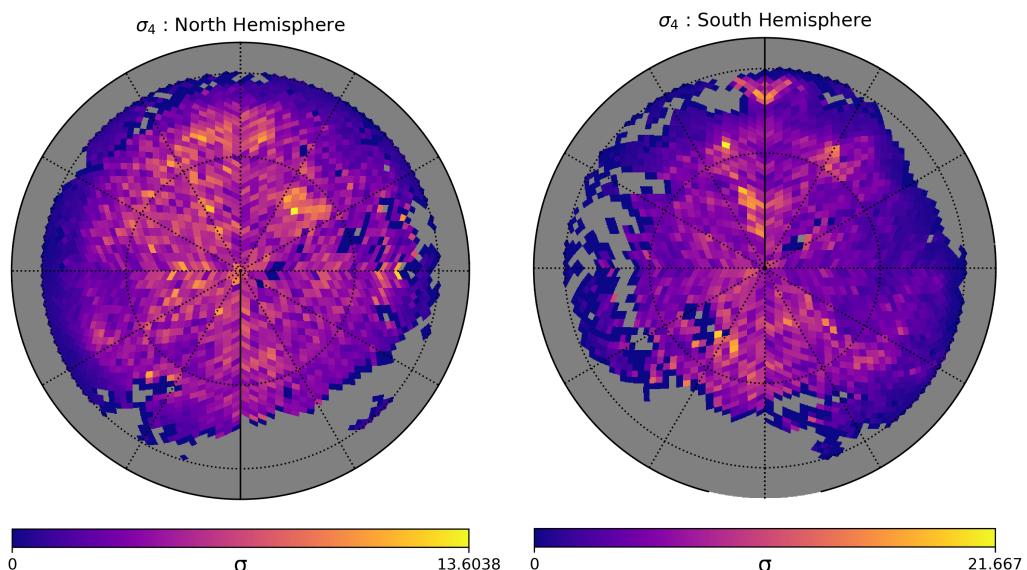


Fig. 8: Error maps for distance range 1000–3881 pc.

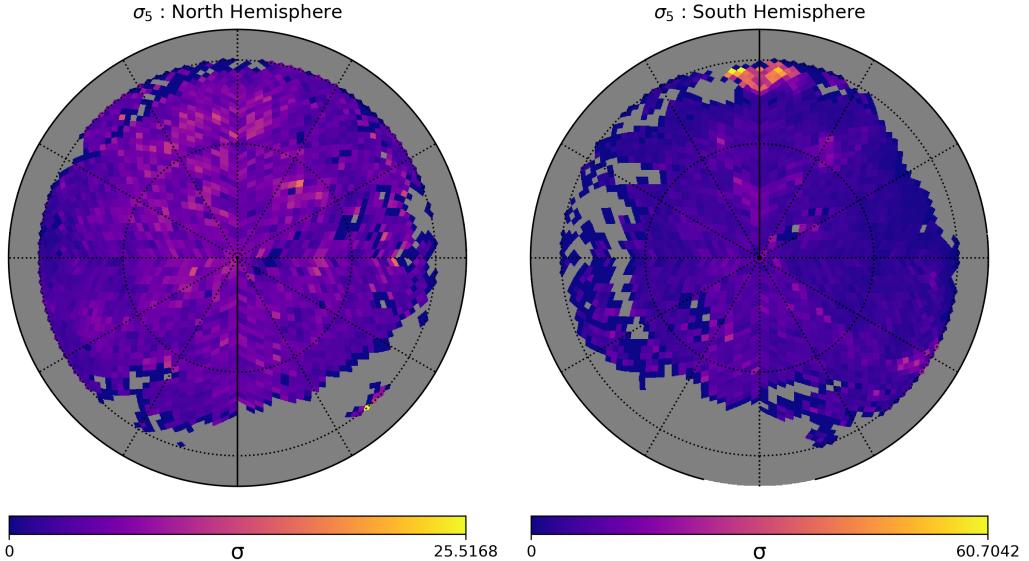


Fig. 9: Error maps for distance range 3881–15849 pc.

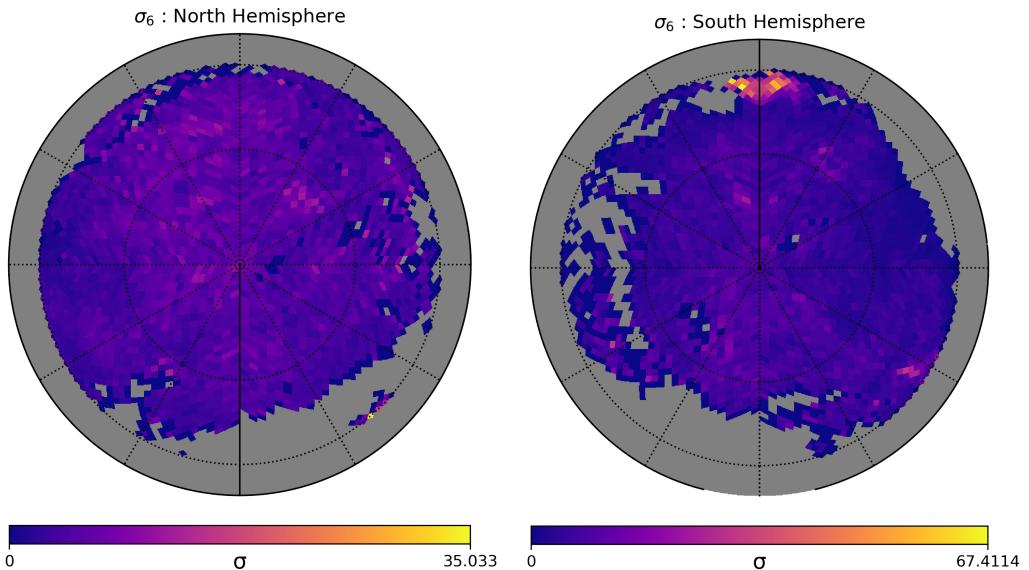


Fig. 10: Error maps for distance range 15849–63096 pc.

The corresponding error maps—representing posterior standard deviations from the HMC chains—quantify the confidence in the inferred extinction coefficients. Low-error regions typically coincide with high signal-to-noise areas, confirming model reliability there. In contrast, high-error zones point to either sparse data, weak signal, or potential degeneracies among templates. This uncertainty quantification is essential for identifying regions of weak constraint and for interpreting residual maps.

## Appendix B : Dust Extinction Coefficients with HI column density

We have examined the correlation between the dust extinction coefficients and the neutral hydrogen (HI) column density to understand the spatial variation and physical dependence of dust emissivity on the interstellar medium's gas content. The HI column density ( $N_{\text{HI}}$ ) map used in this analysis is sourced from the GASS (Galactic All Sky Survey), which provides high-resolution measurements of HI gas across the sky. The original  $N_{\text{HI}}$  map is provided at a resolution of  $\text{Nside} = 512$ .

To facilitate a consistent comparison with the pixel-dependent dust extinction coefficients extracted from our Hamiltonian Monte Carlo (HMC) analysis, which operates at a coarser angular resolution, we compute the mean HI column density at  $\text{Nside} = 32$ . This is achieved by averaging the values of all unmasked high-resolution pixels (at  $\text{Nside} = 512$ ) that fall within each low-resolution superpixel. By excluding masked or unreliable pixels in the averaging process, we ensure that the computed  $N_{\text{HI}}$  values accurately represent the true gas distribution in the corresponding sky regions.

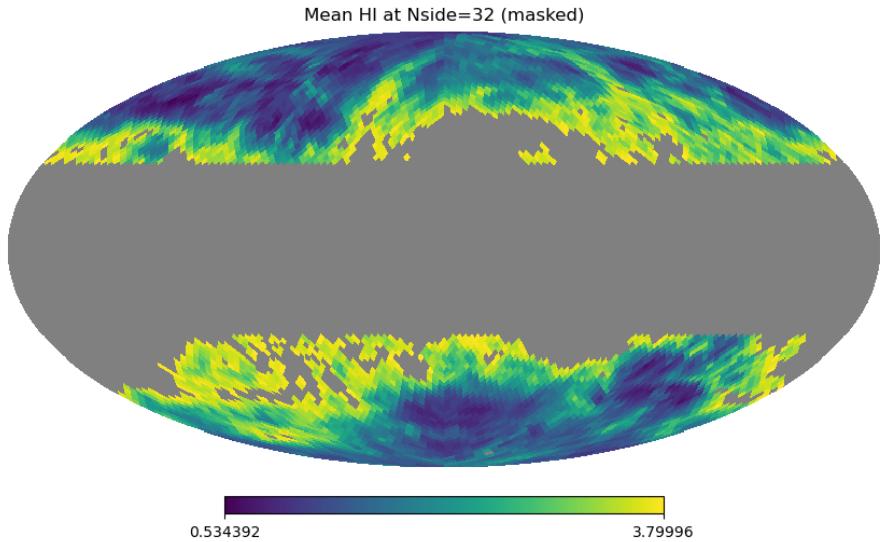


Fig. 11: HI column density map ( $N_{\text{HI}}$ ) from the GASS survey at  $\text{Nside} = 512$ , shown in Galactic coordinates. This map is used to investigate the spatial correlation between neutral hydrogen and dust extinction.

To further investigate the physical relationship between dust and gas in the interstellar medium, we analyze how the pixel-dependent dust extinction coefficients vary as a function of the local neutral hydrogen column density. By averaging the HI column density values from the GASS survey over unmasked pixels at  $\text{Nside} = 512$  and downgrading them to  $\text{Nside} = 32$ , we obtain a consistent spatial scale for comparison with the extinction coefficients derived from our HMC fitting. The following figures illustrates the correlation between the average HI column density and the corresponding dust extinction coefficients across the sky.

### HI vs Dust Extinction coefficients

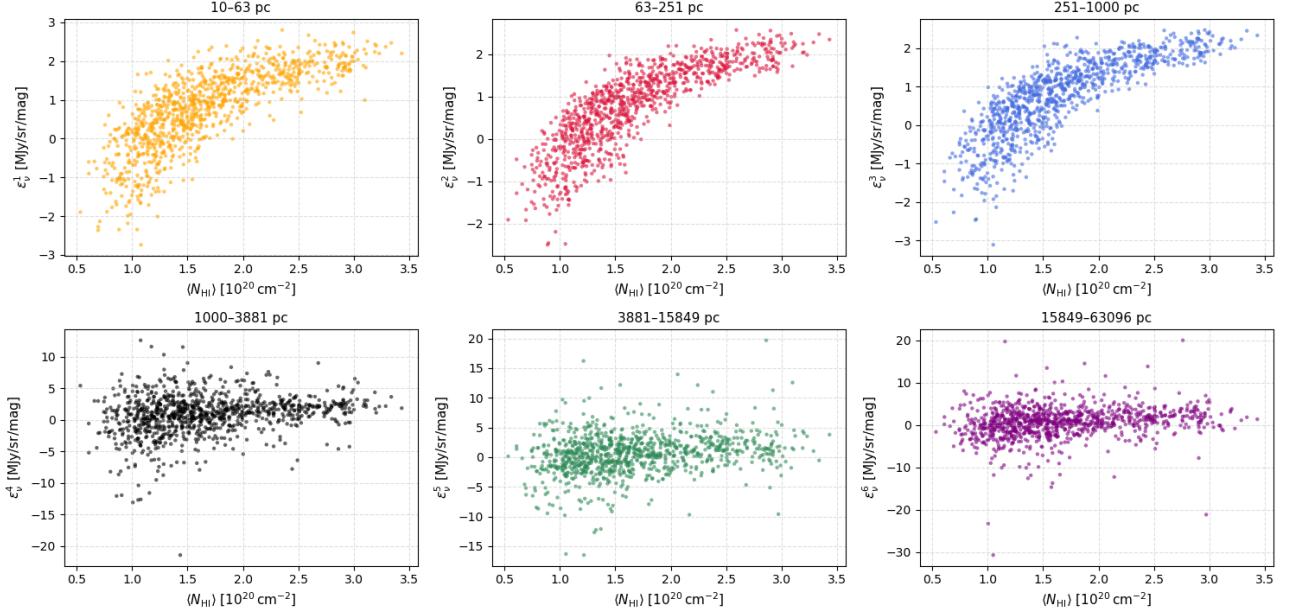


Fig. 12: Variation of dust extinction coefficients with the average HI column density ( $\langle N_{\text{HI}} \rangle$ ) at  $N_{\text{side}} = 32$ .

### HI vs Extinction Uncertainty

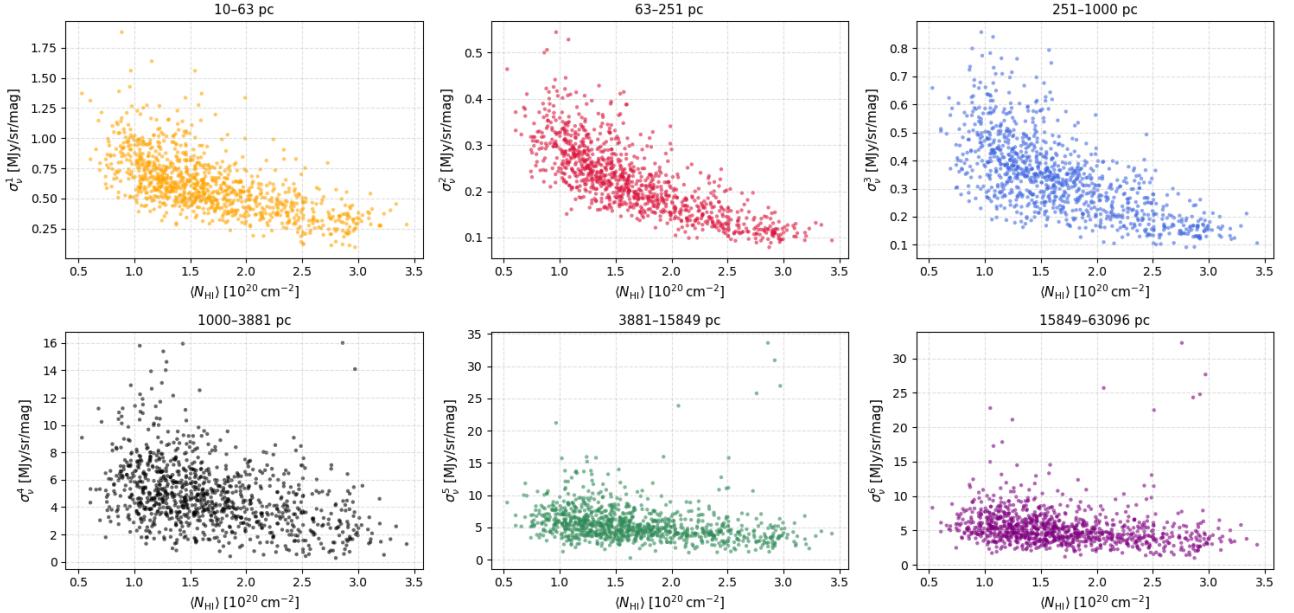


Fig. 13: Variation of dust extinction coefficients errors with the average HI column density ( $\langle N_{\text{HI}} \rangle$ ) at  $N_{\text{side}} = 32$ .

- **Dust Extinction Coefficients vs. HI Column Density:** A strong and consistent positive correlation is found between the dust extinction coefficients  $\epsilon_v^i$  and the average HI column density  $\langle N_{\text{HI}} \rangle$  for the first few distance layers (spanning approximately 10–1000 pc). This trend suggests that in the local interstellar medium (ISM), dust grains are well mixed with neutral atomic hydrogen, and the far-infrared dust emission scales proportionally with the gas content. The tight coupling likely arises due to uniform radiation fields and minimal

contamination from other gas phases in these nearby regions. However, for more distant layers, the correlation progressively weakens and becomes increasingly scattered. This degradation in correlation is attributed to multiple astrophysical factors including: projection effects along the line of sight, the superposition of multiple dust components at varying distances, lower signal-to-noise ratios, and the increasing influence of gas phases such as ionized hydrogen (HII) or molecular gas ( $H_2$ ) which are not effectively traced by HI observations alone.

- **Uncertainty in Extinction Coefficients vs. HI Column Density:** The inferred uncertainties  $\sigma_\nu^i$  associated with the dust extinction coefficients exhibit a declining trend with increasing  $\langle N_{HI} \rangle$  in the nearby layers. This behavior reflects the statistical robustness of parameter estimation in regions with high dust and gas content, where the emission signal is stronger and less prone to noise. In contrast, for distant layers (e.g., beyond 1 kpc), the uncertainties remain significantly larger and tend to be relatively flat as a function of column density. This indicates that even in areas with substantial HI presence, the long-distance extinction coefficients are weakly constrained. This can again be attributed to the lower spatial resolution of distant templates, weaker emission signal, and enhanced degeneracies due to overlapping structures along the line of sight. These findings collectively underscore the reduced diagnostic power of dust-HI correlations in the outer Galaxy and emphasize the need for complementary gas tracers or higher-resolution extinction data for accurate modeling at large distances.

## Appendix C : Validation of Posterior Coefficients

To validate the accuracy of the inferred dust coefficient values, we computed the normalized difference between the input and output emissivities for each superpixel. This difference is defined as:

$$\delta_j = \frac{\epsilon_j^{\text{inp}} - \epsilon_j^{\text{out}}}{\sigma_{\epsilon_j^{\text{out}}}}$$

where  $\epsilon_j^{\text{inp}}$  is the input emissivity value used in the simulation,  $\epsilon_j^{\text{out}}$  is the posterior mean from Hamiltonian Monte Carlo (HMC) sampling, and  $\sigma_{\epsilon_j^{\text{out}}}$  is the standard deviation of the posterior. The histograms of these  $\delta_j$  values were then plotted for HFI frequency at 353 GHz).

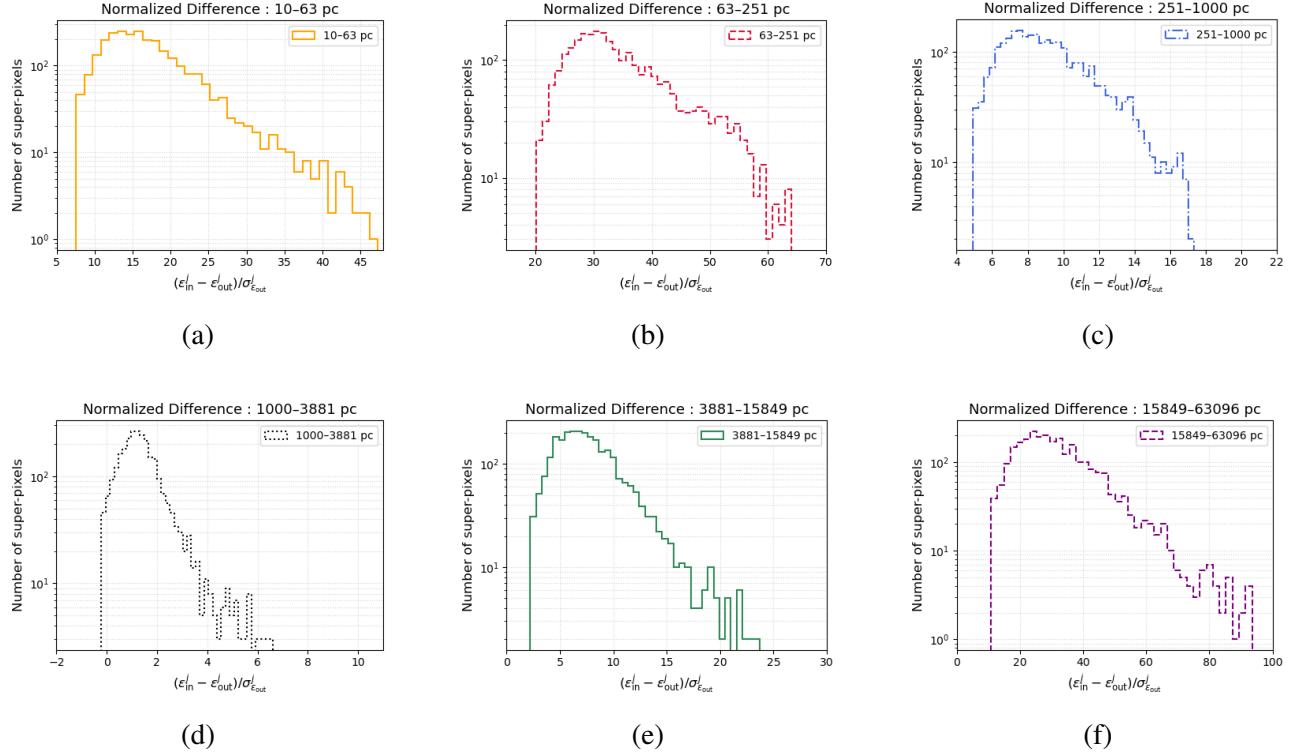


Fig. 14: Histograms of normalized differences between the input emissivity  $\epsilon^{\text{in}}$  and the HMC-derived posterior mean emissivity  $\epsilon^{\text{out}}$  for each distance layer. The values are scaled by the standard deviation  $\sigma_{\epsilon^{\text{out}}}$  of the posterior distribution for each superpixel. These plots validate the accuracy and stability of the inferred dust extinction coefficients across six distance bins.

- The histograms illustrate the normalized residuals of the form  $(\epsilon_{\text{in}}^j - \epsilon_{\text{out}}^j)/\sigma_{\epsilon_{\text{out}}^j}$  for each of the six extinction layers. This metric effectively quantifies how well the posterior dust extinction coefficients  $\epsilon_{\text{out}}^j$  recovered via Hamiltonian Monte Carlo (HMC) match the known input values  $\epsilon_{\text{in}}^j$ , normalized by the posterior uncertainties. For the nearby layers (spanning approximately 10–1000 pc), the distributions are sharply peaked, centered near zero, and symmetric. This indicates high accuracy in parameter recovery, well-behaved posterior distributions, and strong constraints enabled by higher signal-to-noise ratios and minimal line-of-sight confusion.
- For the more distant layers (particularly the bins 3881–15849 pc and 15849–63096 pc), the histograms appear broader and display slight asymmetries. These features reflect the expected increase in uncertainties due to reduced data fidelity, increased template degeneracy, and weaker emission signals at large distances. Nevertheless, the distributions remain unimodal and within reasonable statistical consistency, suggesting that the inference remains stable even

in challenging regimes. This validates the robustness of the multi-template model in extracting physically meaningful parameters across the full range of Galactic distances.

**Note:** Both the results presented in Appendix B and Appendix C require further refinement and validation. Additional tests—such as mock recovery under varied noise realizations or cross-validation with independent extinction datasets—are recommended to fully assess the stability and reliability of the current analysis.