

**SCORE-BASED RECONSTRUCTION OF PRIMORDIAL  
GRAVITATIONAL WAVES FROM CMB B-MODE  
POLARIZATION**

**A THESIS**

*submitted in partial fulfillment of the requirements  
for the award of the dual degree of*

**Bachelor of Science - Master of Science**

*in*

**PHYSICS**

*by*

**ANUMANCHI AGASTYA SAI RAM LIKHIT**

**(20048)**

*Under the guidance of*

**DR RAJIB SAHA**



**iiserb**

**DEPARTMENT OF PHYSICS  
INDIAN INSTITUTE OF SCIENCE EDUCATION AND RESEARCH  
Bhopal  
Bhopal - 462066  
April 2025**



भारतीय विज्ञान शिक्षा एवं अनुसंधान संस्थान भोपाल  
Indian Institute of Science Education and Research  
Bhopal  
(Estb. By MHRD, Govt. of India)

---

## CERTIFICATE

This is to certify that **Anumanchi Agastya Sai Ram Likhit**, BS-MS (Physics), has worked on the project entitled '**Score-based reconstruction of primordial gravitational waves from CMB B-mode polarization**' under my supervision and guidance. The content of this report is original and has not been submitted elsewhere for the award of any academic or professional degree.

April 2025  
IISER Bhopal

Dr Rajib Saha  
(IISER Bhopal)

Co-Supervisor Name  
(Co-Supervisor Institute)

Committee Member

Signature

Date

---

---

---

## **ACADEMIC INTEGRITY AND COPYRIGHT DISCLAIMER**

I hereby declare that this project is my own work and, to the best of my knowledge, it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at IISER Bhopal or any other educational institution, except where due acknowledgement is made in the document.

I certify that all copyrighted material incorporated into this document is in compliance with the Indian Copyright Act (1957) and that I have received written permission from the copyright owners for my use of their work, which is beyond the scope of the law. I agree to indemnify and save harmless IISER Bhopal from any and all claims that may be asserted or that may arise from any copyright violation.

## ACKNOWLEDGEMENT

First and foremost, I would like to express my deepest gratitude to my supervisor, Dr. Rajib Saha. His timely support and invaluable insights have guided me in examining every aspect of this work with precision and depth. Above all, his passion for exploring the early universe and his curiosity in problem-solving have deeply inspired and motivated me throughout this journey. Secondly, I am immensely thankful to Dr. Prabu Thiagaraj for welcoming me as a visiting student at the Raman Research Institute during my summer break and providing the opportunity to work at the Gauribidanur Radio Observatory. His tireless commitment, patience, and clear guidance have helped me develop my research skills and grow in confidence in the lab. Working alongside him has been truly inspiring, and I am grateful for his extended support and guidance throughout my academic journey.

Next, I'd like to sincerely thank the faculty members of the Department of Physics and the Department of Data Science and Engineering at IISER Bhopal, whose enriching courses have significantly strengthened my theoretical knowledge and analytical skills. Additionally, I extend my heartfelt gratitude to Dr. Srikanta Pal, PhD scholar of our research group, for his valuable insights and constructive discussions.

I would also like to express my heartfelt appreciation to my friends Katta Naveen, Dr. K. Balaji Naik (aspiring IAS officer and cutting-edge researcher), Kethavath Jagan, Aman, Shalini and Uma Madhuri, whose kindness, humor, and continuous encouragement have made this journey truly memorable. Special thanks to my friends from the Physics department—Divyansh, Sunny, Athul, Snehasis and Debmalya—for the great times we've shared. Additionally, I'm grateful to my lab mates Abhinand, Monish, and PhD scholar Subhankar for the enjoyable and insightful hours spent together in the lab exploring the cosmos.

Lastly, and most importantly, I wish to express my deepest gratitude to my parents, Anumanchi Satyanarayana Rao and Anumanchi Usha Kumari, whose unwavering support, love, and sacrifices have been the cornerstone of my academic journey. Without their constant encouragement and belief in me, none of this would have been possible. I also extend my warmest thanks to my sister, Manogna, and my brother, Vishnu, for their endless patience and love, which have kept my spirits high throughout this journey.

Finally, I would like to thank the late Prof. Stephen W Hawking, whose extraordinary life and contributions to cosmology have continually inspired my curiosity and passion for understanding the universe.

---

# ABSTRACT

The Big Bang theory predicts the universe began in a hot and dense state, rapidly expanding and cooling ever since. The Cosmic Microwave Background (CMB) is one of the strongest pieces of evidence for this theory and now serves as a vital tool to investigate cosmic inflation. In particular, the B-mode polarization signal of the CMB is widely regarded as “smoking-gun” evidence for gravitational waves produced during inflation. Detecting this faint signal provides crucial information about the energy scale of inflation and offers direct confirmation of inflationary theory itself.

In this thesis, we presented a novel approach to isolate and reconstruct the primordial B-mode polarization signal from upcoming CMB observations. We employed score-based diffusion models, formulated through Variance-Exploding Stochastic Differential Equations (VE-SDEs), to reconstruct the B mode signal without explicit prior knowledge of the contaminants. Through end-to-end simulations incorporating realistic instrumental specifications, gravitational lensing, and multiple foreground sources, we demonstrated that Score SDE-based methods accurately recover both B-mode map structures and angular power spectra.

These results indicate that advanced generative models can effectively handle high-dimensional, contaminated CMB data, greatly improving the reliability and sensitivity of B-mode detection. Once trained, these models also enable rapid creation of realistic simulations, making them a valuable tool for next-generation CMB experiments to more precisely measure the tensor-to-scalar ratio  $r$ , validate the inflationary paradigm, and deepen our understanding of the early universe.

# LIST OF FIGURES

1.1	History of Universe (Ref: <a href="http://bicepkeck.org/visuals.html">http://bicepkeck.org/visuals.html</a> ) . . . . .	2
1.2	Top: A simulated full-sky CMB temperature anisotropy map in units of $\mu\text{K}$ . Bottom: The corresponding angular power spectrum $D_\ell^{TT} = \ell(\ell+1)C_\ell^{TT}/2\pi$ , showing how the power in temperature fluctuations varies with the multipole moment $\ell$ . . . . .	4
1.3	<b>Top panels:</b> E-mode (left) and B-mode (right) polarization maps in $\mu\text{K}$ . <b>Bottom panels:</b> The corresponding angular power spectra, $D_\ell^E$ and $D_\ell^B$ , showing how polarization power varies with multipole moment $\ell$ . . . . .	6
1.4	Temperature (left), E-mode (middle), and B-mode (right) power spectra, $D_\ell^{TT}$ , $D_\ell^{EE}$ , and $D_\ell^{BB}$ , showing total (black), scalar (magenta), and tensor (blue) contributions. . . . .	8
1.5	$D_\ell^{BB}$ for different values of $r$ . The dashed line indicates the lensing contribution, which can dominate the primordial B-mode for small $r$ . . . . .	9
1.6	Comparison of total B-mode power (black) with tensor (magenta) and lensed scalar (blue) components at $r = 0.001$ . . . . .	9
1.7	Theoretical CMB angular power spectra of temperature T, E- and B-mode polarization anisotropies as a function of multipole moments for the best-fit CDM parameters derived by Planck (Planck Collaboration VI 2018) with $r = 0.06$ . The instrument noise power spectrum of polarization (dash lines) plotted for different generations of space based CMB surveys, viz. Planck, LiteBIRD and ECHO. . . . .	10
2.1	Visualizing Diffusion Models: Adding noise to data and then denoising to recover the original data. . . . .	12

---

2.2	Gaussian Mixture Model (GMM) fitted to a dataset with two Gaussian components. The model is trained on 300 samples from each distribution with means $\mu_1 = (0, 1)$ , $\mu_2 = (3, -1)$ and covariance matrices $\Sigma_1 = ((1, 0.5), (0.5, 1))$ , $\Sigma_2 = ((2, 0.5), (0.5, 1.0))$ . The left panel shows the probability density function $p(x)$ , while the right panel visualizes the score function $\nabla \log p(x)$ as a vector field overlaid on $\log p(x)$ contours, indicating the direction of steepest ascent. . . . .	13
2.3	Transforming data to a simple noise distribution can be accomplished with an SDE. It can be reversed to generate samples from noise if we know the score of the distribution at each intermediate time step. . . . .	16
3.1	Example illustrating how beam convolution suppresses power at high multipoles. . . . .	25
3.2	Random realizations of simulated B mode noise maps at each ECHO Frequency	26
3.3	B-mode foreground maps at 28 GHz simulated using different sky models in PySM3. . . . .	30
4.1	Overview of the VE-SDE and U-Net denoising approach. . . . .	32
4.2	CMB B-mode maps (top to bottom, three random realizations). <i>Left column</i> : Pure B-mode maps (primordial + lensing). <i>Middle column</i> : Simulated observations obtained by adding detector noise (at 520 GHz) to the pure maps. <i>Right column</i> : Denoised maps output by the VE-SDE score UNet. . . . .	34
4.3	Distribution of correlation coefficients computed between the denoised maps and (1) the ground truth (blue), and (2) the noisy observations (orange). . .	34
4.4	Power spectra of the B-mode maps at 520 GHz. <i>Left</i> : The mean $D_\ell$ over 1,000 simulated realizations. <i>Right</i> : Power spectrum for a single realization, showing the same three categories (pure, observed, and denoised). . . . .	35
4.5	Single realization from each of the four foreground sets (F1–F4). The ground-truth primordial B-mode spectrum (blue), the denoised reconstruction (orange) and the observed (noise + lensing + foreground) spectrum (green) . .	38
4.6	Evolution of a single realization under the reverse VE-SDE at selected times ( $t=1, 0.801, 0.400, 0.200$ ). Each panel shows the progressively denoised power spectrum, culminating in the final $t=0$ result (dashed black) compared to the ground truth (dashed gray). . . . .	39
4.7	Single realization from each of the four foreground sets (F1–F4) comparing the ground-truth (blue) and denoised (orange) primordial B-mode spectra over $\ell = 2 \dots 31$ . . . . .	40

4.8	Mean B-mode power spectra for 5,000 realizations in each foreground set (F1–F4). The ground truth (blue) is closely matched by the denoised solution (orange) across $\ell = 2 \dots 31$ . . . . .	41
4.9	Mean residual ( $\Delta D_\ell$ ) and standard deviation between the denoised and ground-truth spectra across all realizations. The horizontal dashed line at zero indicates perfect agreement. . . . .	41

# CONTENTS

<b>Certificate</b> . . . . .	i
<b>Academic Integrity and Copyright Disclaimer</b> . . . . .	ii
<b>Acknowledgement</b> . . . . .	iii
<b>Abstract</b> . . . . .	iv
<b>List of Figures</b> . . . . .	vii
<b>1. Introduction</b> . . . . .	1
1.1 Inflation . . . . .	1
1.2 Cosmic Microwave Background . . . . .	3
1.3 Gravitational Lensing of the CMB . . . . .	8
1.4 Exploring Cosmic History and Origins . . . . .	9
<b>2. Score Based Generative models</b> . . . . .	12
2.1 Score-Based Generative Modeling through Stochastic Differential Equations . .	15
2.1.1 Variance Exploding (VE) SDE . . . . .	17
2.2 Euler–Maruyama (EM) Discretization . . . . .	19
2.2.1 EM for VE SDE . . . . .	20
<b>3. Cosmological Simulations</b> . . . . .	21
3.1 Theoretical Angular Power Spectrum using CAMB . . . . .	21
3.2 HEALPix . . . . .	22
3.3 Beam Convolution and Instrumental Noise . . . . .	23
3.4 CMB Foregrounds and PySM3 . . . . .	27
3.4.1 Simulations with PySM3 . . . . .	28
<b>4. Reconstructing the Primordial B-mode signal using Score-SDE</b> . . . . .	31
4.1 Score-Based Denoising of Full-Sky CMB B-Mode Maps . . . . .	31
4.1.1 Data Simulations . . . . .	32

4.1.2	Model Development and Training . . . . .	32
4.1.3	Inference and Results . . . . .	33
4.2	Reconstruction of Primordial B mode Angular Power spectrum using Denoising and Delensing . . . . .	35
4.2.1	Data Simulations . . . . .	36
4.2.2	Model Development and Training . . . . .	37
4.2.3	Inference and Results . . . . .	38
<b>5.</b>	<b>Discussion and Conclusions . . . . .</b>	<b>43</b>
	<b>Bibliography . . . . .</b>	<b>47</b>

*If the rate of expansion one second after the Big Bang had been smaller by even one part in a hundred thousand million million, it would have recollapsed before it reached its present size. On the other hand, if it had been greater by a part in a million, the universe would have expanded too rapidly for stars and planets to form.* ~ **Stephen W Hawking.**

# 1. INTRODUCTION

The Universe, as we understand today, began approximately 13.8 billion years ago from an extremely hot and dense state, rapidly expanding and cooling ever since (Lemaître, 1931). Initially, all matter and energy existed as a dense primordial plasma where photons, electrons, protons, and neutrons interacted vigorously. The expanding universe on large scales is mathematically described by the Friedmann–Lemaître–Robertson–Walker (FLRW) metric:

$$ds^2 = -dt^2 + a^2(t) \left[ \frac{dr^2}{1 - kr^2} + r^2 (d\theta^2 + \sin^2 \theta d\phi^2) \right], \quad (1.1)$$

where  $a(t)$  is the cosmic scale factor describing the expansion of the universe, and  $k$  is the spatial curvature parameter taking values  $-1, 0, +1$  corresponding to open, flat, or closed spatial geometries respectively.

The dynamics of this expanding universe are governed by the Friedmann equations (eq 1.2,1.3) derived from Einstein's equations under assumptions of homogeneity and isotropy for a flat space ( $k = 0$ ) :

$$H^2 = \left( \frac{\dot{a}}{a} \right)^2 = \frac{8\pi G}{3}\rho, \quad (1.2)$$

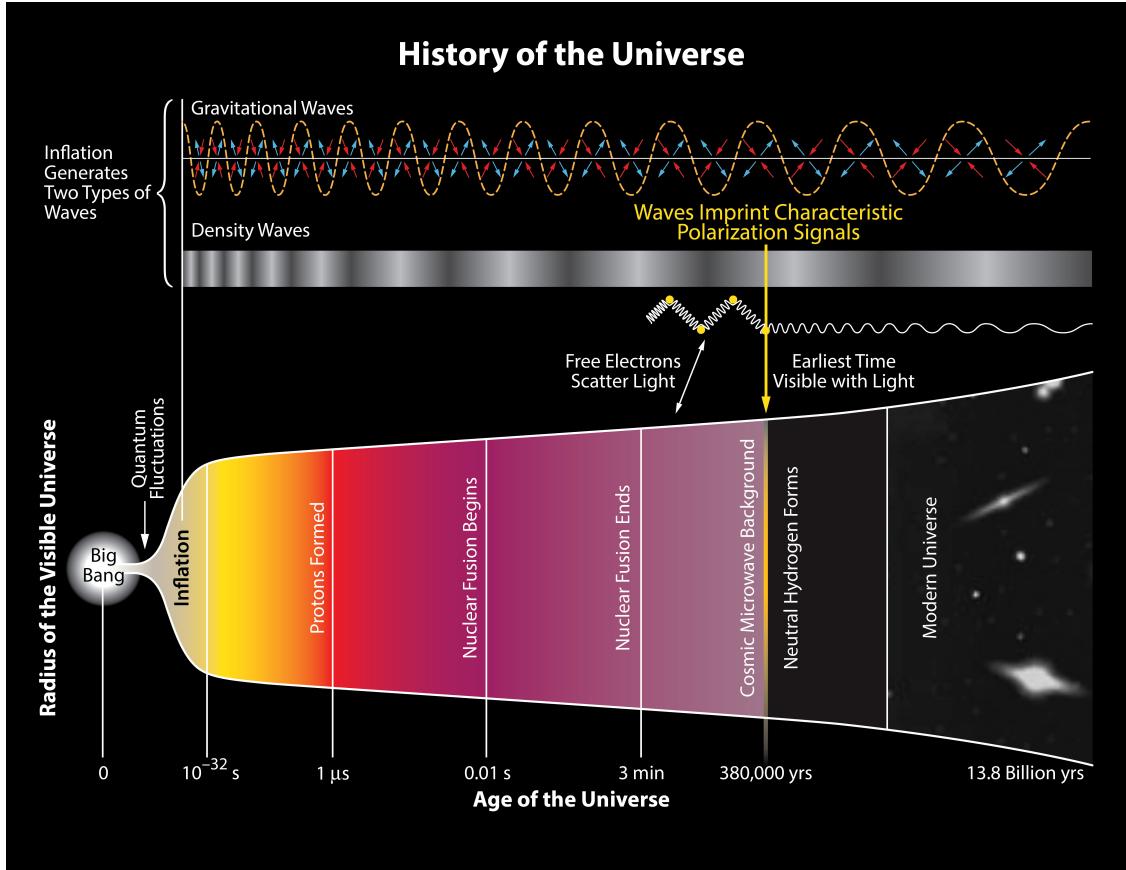
$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p), \quad (1.3)$$

where  $H(t)$  is the Hubble parameter,  $G$  is Newton's gravitational constant, and  $\rho$  and  $p$  represent the energy density and pressure of the cosmic fluid, respectively.

## 1.1 Inflation

However, the standard Big Bang FLRW cosmology faces two major challenges: the horizon problem and the flatness problem. The horizon problem arises from the observation that distant regions of the universe separated by vast distances exceeding their causal horizons exhibit nearly identical temperatures and densities. This uniformity implies that these regions were once in causal contact, contradicting the expectations of standard expansion. Meanwhile, the flatness problem highlights that the universe appears remarkably spatially flat ( $k \sim 0$ ), requiring exceedingly precise initial conditions that are difficult to explain using standard Big Bang model.

To address these, the theory of inflation (Guth, 1981; Linde, 1982; Hawking, 1982; Guth & Pi, 1985) was developed, describing a phase of extremely rapid accelerated expansion in the early universe (approximately between  $10^{-36}$  and  $10^{-32}$  seconds after the Big Bang). During inflation, any initial fluctuations were stretched to cosmological scales, effectively solving both the horizon and flatness problems by driving the universe toward homogeneity and near-flatness. In most models, this accelerated expansion is driven by a scalar field, commonly known as the inflaton, whose potential  $V(\phi)$  governs the dynamics of the inflationary period.



**Fig. 1.1:** History of Universe (Ref: <http://bicepkeck.org/visuals.html>)

Quantum fluctuations of the inflaton field play a crucial role: as these fluctuations exit the Hubble horizon, they become “frozen in,” later manifesting as classical perturbations. These perturbations include scalar (density) modes, which seed the large-scale structure we observe today, and tensor (gravitational wave) modes, which can imprint themselves on spacetime as primordial gravitational waves.

The tensor-to-scalar ratio, denoted  $r$ , quantifies the relative amplitude of these tensor fluctuations compared to the scalar ones. If  $\nabla_R^2(k)$  and  $\nabla_h^2(k)$  denote the dimensionless power spectra of scalar and tensor perturbations, respectively, the tensor-to-scalar ratio is

defined at a particular pivot scale  $k_*$  as

$$r = \frac{\nabla_h^2(k_*)}{\nabla_R^2(k_*)} \quad (1.4)$$

The amplitude of these fluctuations depends on the shape of the inflaton potential and the rate of the universe's expansion during inflation. In slow-roll models, for instance,  $r$  can be directly related to the potential's slope, making its measurement a powerful probe of high-energy physics. A higher value of  $r$  indicates that inflation produced stronger gravitational waves, whereas a smaller  $r$  suggests weaker tensor modes.

## 1.2 Cosmic Microwave Background

As the universe expanded, it cooled sufficiently to allow electrons and protons to combine into neutral hydrogen atoms, an epoch known as recombination occurring roughly 380,000 years after the Big Bang (redshift  $z \sim 1100$ ). Photons, no longer scattering off charged particles, began to travel freely, preserving a snapshot of the early universe—the Cosmic Microwave Background (CMB) ([Penzias & Wilson, 1965](#)).

The CMB is almost uniform, with an average temperature of 2.7255K, and shows tiny temperature fluctuations (anisotropies) that provide us with some valuable information about primordial perturbations([Coles & Lucchin, 1995](#); [Narlikar, 2002](#); [Hu & Dodelson, 2002](#)). These deviations from the mean temperature are quantified by

$$\Theta(\hat{n}) = \frac{\Delta T(\hat{n})}{T} = \frac{T(\hat{n}) - \bar{T}}{\bar{T}}, \quad (1.5)$$

where  $\bar{T}$  is the average CMB temperature and  $\hat{n}$  denotes a direction on the sky. Since the CMB is observed over the entire celestial sphere, we can expand  $\Theta(\hat{n})$  in spherical harmonics  $Y_{\ell m}(\hat{n})$ :

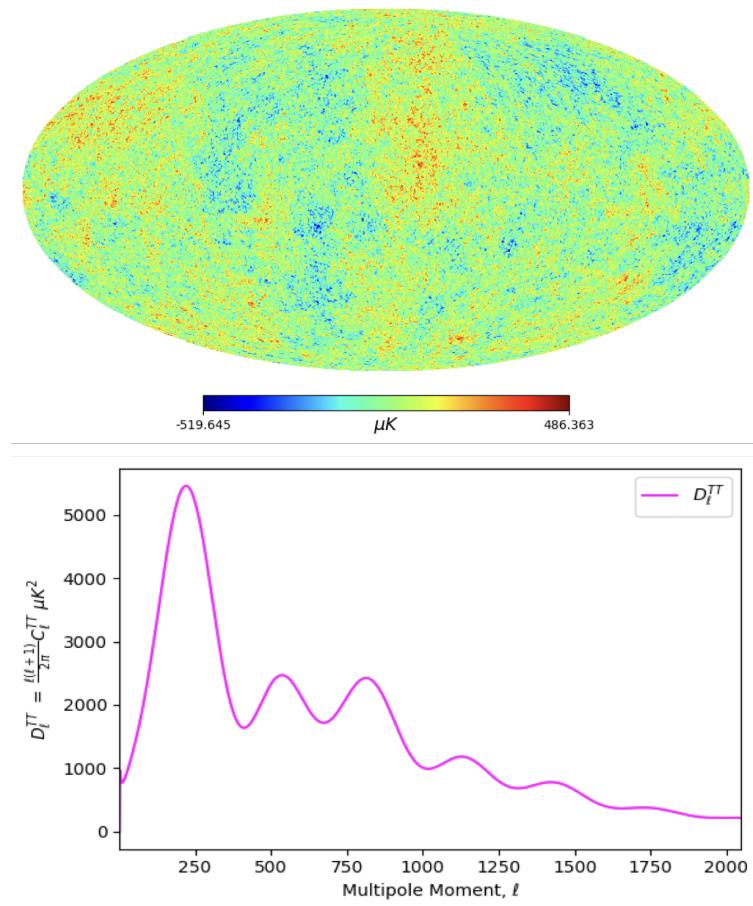
$$\Theta(\hat{n}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m} Y_{\ell m}(\hat{n}), \quad (1.6)$$

with the expansion coefficients given by

$$a_{\ell m} = \int \Theta(\hat{n}) Y_{\ell m}^*(\hat{n}) d\Omega \quad (1.7)$$

where  $d\Omega$  is the differential solid angle element. The theoretical angular power spectrum,  $C_{\ell}^{TT}$ , shows how the temperature fluctuations vary with angular scale and is defined as

$$C_\ell^{TT} = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \langle |a_{\ell m}|^2 \rangle. \quad (1.8)$$



**Fig. 1.2:** Top: A simulated full-sky CMB temperature anisotropy map in units of  $\mu\text{K}$ . Bottom: The corresponding angular power spectrum  $D_\ell^{TT} = \ell(\ell + 1)C_\ell^{TT}/2\pi$ , showing how the power in temperature fluctuations varies with the multipole moment  $\ell$ .

We obtain the theoretical angular power spectrum (Eq. 1.8) from the ensemble of  $a_{\ell m}$  coefficients that characterize the Gaussian random process responsible for the primordial temperature fluctuations. In practice, however, we have access to only one realization of this process—our observed CMB sky. Therefore, the observed angular power spectrum is estimated as

$$\hat{C}_\ell^{TT} = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} |a_{\ell m}|^2 \quad (1.9)$$

The measured angular power spectrum is limited by cosmic variance, which arises because we only have one sky. This uncertainty is quantified by

$$\Delta C_\ell^{TT} = \sqrt{\frac{2}{2\ell+1}} C_\ell^{TT},$$

indicating that fewer modes at low  $\ell$  lead to higher uncertainty.

Polarization in the CMB (Hu & White, 1997) arises primarily from Thomson scattering of photons off free electrons in the presence of a quadrupole temperature anisotropy at recombination. In analogy with electromagnetism, the correlation matrix of the electric field components for a plane wave propagating in the direction  $\hat{z}$  is characterized by the four Stokes parameters:  $I$ ,  $Q$ ,  $U$ , and  $V$ .

$$\begin{pmatrix} \langle E_x E_x^* \rangle & \langle E_x E_y^* \rangle \\ \langle E_y E_x^* \rangle & \langle E_y E_y^* \rangle \end{pmatrix} \equiv \frac{1}{2} \begin{pmatrix} I + Q & U + iV \\ U - iV & I - Q \end{pmatrix} \quad (1.10)$$

we also characterise the polarisation of the CMB using these Stokes parameters. Here, the trace gives the total intensity  $I$ . The parameter  $Q$  represents the difference in intensity between the  $x$  and  $y$  directions, while  $U$  quantifies the difference when the  $x$  and  $y$  axes are rotated by  $45^\circ$ .  $V$  describes circular polarization, which vanishes for Thomson scattering.

$Q$  and  $U$  depend on the chosen coordinate system and transform non-trivially under rotations. In particular, under a rotation by an angle  $\psi$  in the plane perpendicular to the line of sight, the combinations  $Q \pm iU$  transform as spin  $\pm 2$  quantities:

$$(Q \pm iU)' = e^{\mp 2i\psi} (Q \pm iU). \quad (1.11)$$

Because of their spin-2 nature, the linear polarization field is naturally expanded in terms of spin-weighted spherical harmonics,  $\pm_2 Y_{\ell m}(\hat{n})$ :

$$(Q \pm iU)(\hat{n}) = \sum_{\ell, m} a_{\pm 2, \ell m} \pm_2 Y_{\ell m}(\hat{n}). \quad (1.12)$$

To obtain scalar quantities that are invariant under rotations and exhibit definite parity properties, the E-mode (gradient-like) and B-mode (curl-like) components are defined as linear combinations of these coefficients:

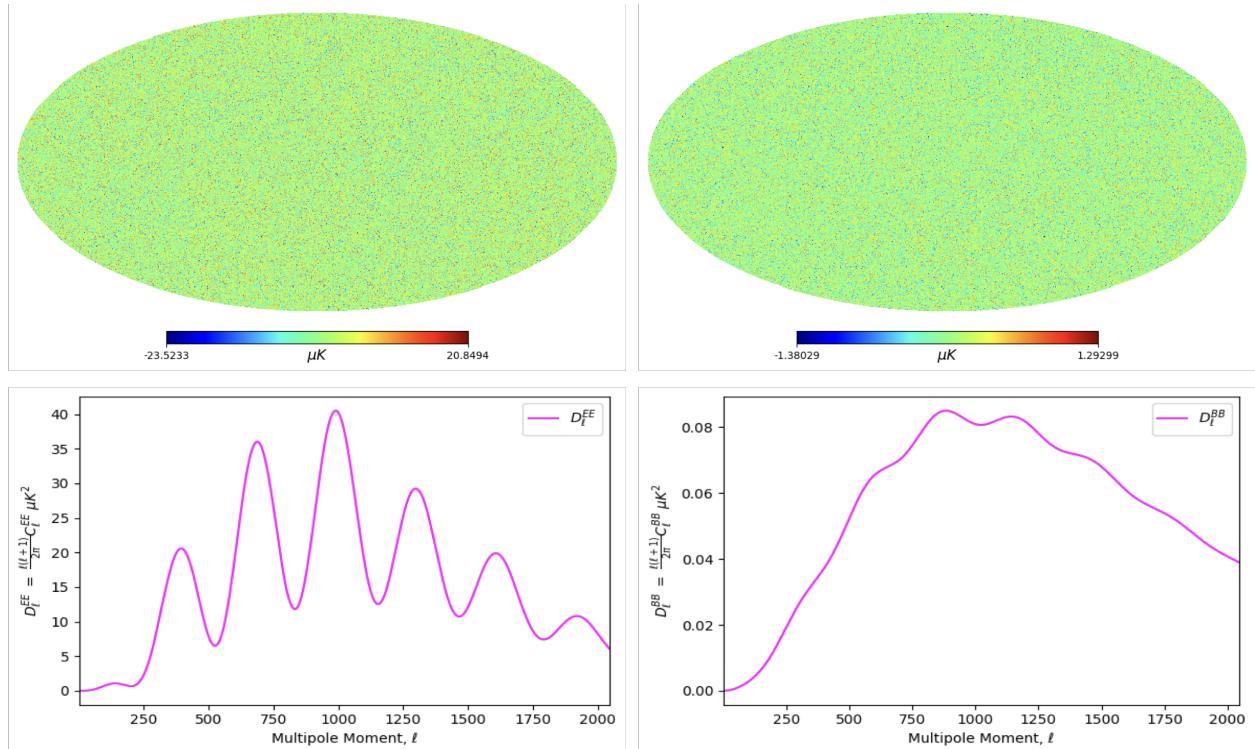
$$a_{\ell m}^E = -\frac{1}{2} (a_{2, \ell m} + a_{-2, \ell m}), \quad (1.13)$$

$$a_{\ell m}^B = \frac{i}{2} (a_{2, \ell m} - a_{-2, \ell m}). \quad (1.14)$$

Analogous to the temperature field, the statistical properties of the polarization fields are characterized by their angular power spectra:

$$C_\ell^{EE} = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \langle |a_{\ell m}^E|^2 \rangle, \quad (1.15)$$

$$C_\ell^{BB} = \frac{1}{2\ell+1} \sum_{m=-\ell}^{\ell} \langle |a_{\ell m}^B|^2 \rangle. \quad (1.16)$$



**Fig. 1.3:** **Top panels:** E-mode (left) and B-mode (right) polarization maps in  $\mu\text{K}$ . **Bottom panels:** The corresponding angular power spectra,  $D_\ell^E$  and  $D_\ell^B$ , showing how polarization power varies with multipole moment  $\ell$ .

In the absence of the leakage between  $E$ - and  $B$ -modes, as in the case for full sky polarization,  $E$ -mode and  $B$ -mode can be written as

$$E(\hat{\mathbf{n}}) = \sum_{\ell m} a_{E,\ell m} Y_{\ell m}(\hat{\mathbf{n}}) \quad (1.17)$$

$$B(\hat{\mathbf{n}}) = \sum_{\ell m} a_{B,\ell m} Y_{\ell m}(\hat{\mathbf{n}}) \quad (1.18)$$

where  $Y_{\ell m}(\hat{\mathbf{n}})$  are the spherical harmonic functions corresponding to a spin-0 field.

The Boltzmann equation in cosmology governs how the photon distribution function evolves in an expanding universe, accounting for both free-streaming and scattering processes (such as Thomson scattering). Solving the Boltzmann equations for linearized polarizations E and B modes (Dodelson, 2003), one obtains:

$$\dot{E}_\ell + k \left[ \sqrt{\frac{(\ell+1)^2 - 4}{2\ell+3}} E_{\ell+1} - \sqrt{\frac{\ell^2 - 4}{2\ell-1}} E_{\ell-1} \right] = \dot{\tau} \left[ E_\ell - \frac{3}{5} \delta_{\ell 2} \left( E_2 - \frac{1}{\sqrt{6}} \Theta_2 \right) \right], \quad (1.19)$$

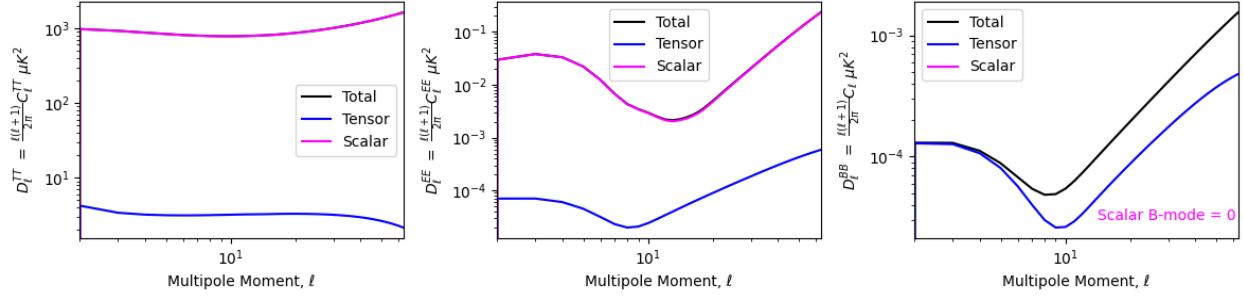
$$\dot{B}_\ell + k \left[ \sqrt{\frac{(\ell+1)^2 - 4}{2\ell+3}} B_{\ell+1} - \sqrt{\frac{\ell^2 - 4}{2\ell-1}} B_{\ell-1} \right] = \dot{\tau} B_\ell. \quad (1.20)$$

- $E_\ell$  and  $B_\ell$  are the multipole moments of the E mode and B mode polarization, respectively.
- $\dot{\tau}$  is the differential optical depth, which quantifies the probability per unit time of Thomson scattering.
- $k$  represents the wavenumber corresponding to the Fourier modes of the perturbations.
- $\Theta_2$  denotes the temperature quadrupole moment.
- $\delta_{\ell 2}$  is the Kronecker delta, ensuring that the source term contributes only for  $\ell = 2$ .

The key outcome is that the B-mode equation (1.20) does not include a source term arising from temperature quadrupoles, indicating that scalar perturbations alone cannot generate B-mode polarization. In the literature (Kamionkowski & Kovetz, 2016; Wang, 2017), it is also shown that tensor perturbations (gravitational waves) contribute to both the temperature anisotropies and the E and B mode polarizations, with the E and B modes generated by gravitational waves possessing roughly equal power on large scales.

From the simulation results presented in Figure 1.4, it is evident that the CMB B-mode polarization is significantly fainter than both the temperature anisotropies and the E-mode polarization. This faintness arises because scalar perturbations, which dominate the temperature (T) and E-mode signals, do not produce any B-mode polarization.

The near absence of scalar contributions in the B-mode makes it a unique and direct probe of tensor modes (primordial gravitational waves). However, the overall B-mode signal remains very weak and is further complicated by contamination from gravitational lensing, astrophysical foregrounds and instrumental noise.



**Fig. 1.4:** Temperature (left), E-mode (middle), and B-mode (right) power spectra,  $D_\ell^{TT}$ ,  $D_\ell^{EE}$ , and  $D_\ell^{BB}$ , showing total (black), scalar (magenta), and tensor (blue) contributions.

	Scalar (density perturbations)	Tensor (gravitational waves)
E-modes	Yes	Yes
B-modes	No	Yes

**Tab. 1.1:** Generation of E and B modes by scalar and tensor perturbations.

### 1.3 Gravitational Lensing of the CMB

Gravitational lensing occurs when the trajectories of CMB photons (LEWIS & CHALLINOR, 2006) are deflected by the gravitational potential of intervening large-scale structures. Although each deflection is typically small (on the order of a few arcminutes), the cumulative effect across the sky systematically remaps the CMB temperature and polarization fields. This remapping can be described by a lensing potential  $\phi(\hat{n})$ , which encodes the projected mass distribution along each line of sight.

In the absence of lensing, the CMB field  $X(\hat{n})$  (where  $X$  can be  $T, Q \pm iU, E$ , or  $B$ ) is observed along the direction  $\hat{n}$ . When lensing is included, photons originate from a slightly different direction on the last-scattering surface due to the gravitational deflection. The observed lensed field  $\tilde{X}(\hat{n})$  is related to the unlensed field  $X(\hat{n})$  by

$$\tilde{X}(\hat{n}) = X(\hat{n} + \nabla\phi(\hat{n})), \quad (1.21)$$

where  $\nabla\phi(\hat{n})$  is the deflection angle. For small deflections, one may expand the above expression in powers of  $\nabla\phi$  and derive a perturbative series for  $\tilde{X}(\hat{n})$ .

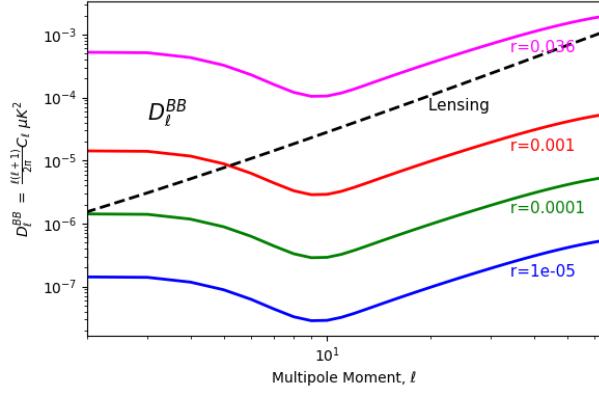
A key consequence of lensing is that it partially converts E-mode polarization into B-mode polarization. Even though scalar (density) perturbations alone do not generate a primordial B-mode signal, lensing of the E-mode field by large-scale structure produces a nonzero lensing B-mode. The lensed Stokes parameters ( $Q \pm iU$ ) acquire additional components from the

remapping:

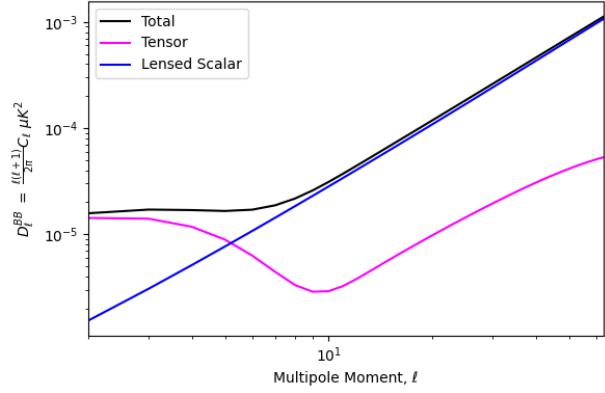
$$(\tilde{Q} \pm i \tilde{U})(\hat{n}) = (Q \pm i U)(\hat{n} + \nabla \phi(\hat{n})). \quad (1.22)$$

When expanded in spin-weighted spherical harmonics, this results in small but significant power in the  $\tilde{B}$ -mode spectrum, denoted  $C_\ell^{BB,\text{lensed}}$ .

$$C_\ell^{BB,\text{Total}} = C_\ell^{BB,\text{tensor}} + C_\ell^{BB,\text{lensed}} \quad (1.23)$$



**Fig. 1.5:**  $D_\ell^{BB}$  for different values of  $r$ . The dashed line indicates the lensing contribution, which can dominate the primordial B-mode for small  $r$ .



**Fig. 1.6:** Comparison of total B-mode power (black) with tensor (magenta) and lensed scalar (blue) components at  $r = 0.001$ .

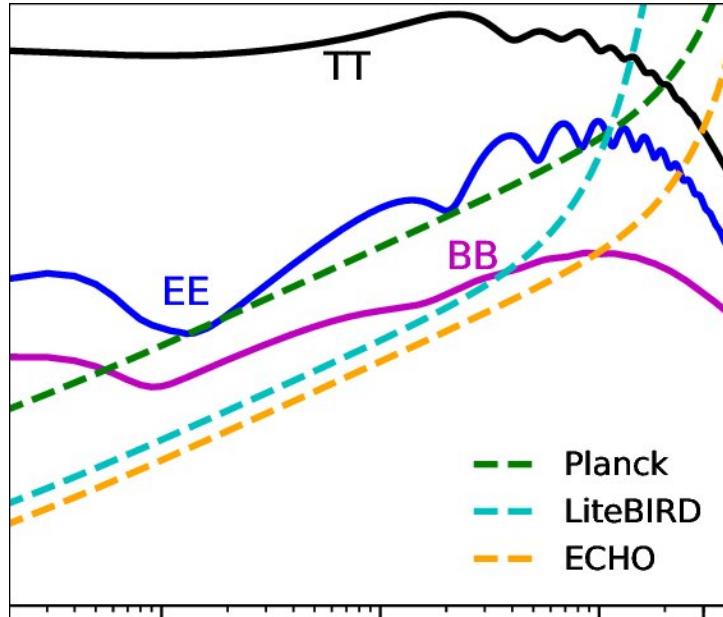
From Figure 1.5, one can see that the primordial B-mode signal is highly sensitive to the value of  $r$ . As  $r$  decreases, the amplitude of the tensor-induced B-modes becomes very small, making it challenging to distinguish from the lensing B-mode signal (dashed line). In Figure 1.6, we focus on the case  $r = 0.001$ , where the lensing contribution (blue) still remains significant across most multipoles.

Because lensing-induced B-modes can dominate over the primordial tensor B-mode for sufficiently small tensor-to-scalar ratios  $r$ , accurate modeling or removal of lensing effects is crucial for detecting primordial gravitational waves.

## 1.4 Exploring Cosmic History and Origins

Exploring Cosmic History and Origins (ECHO), also known as CMB-Bhārat (Adak, Sen, Basak, Delabrouille, Ghosh, Rotti, Martínez-Solaeche & Souradeep, 2022), is a space mission that has been proposed to the Indian Space Research Organisation (ISRO) for the scientific exploitation of the Cosmic Microwave Background (CMB) at the next level of precision and

accuracy. One of the key scientific goals of this mission is to investigate the CMB polarization B-mode signal. The instrumental configuration of ECHO is proposed to detect the target tensor-to-scalar ratio  $r \sim 10^{-3}$  at a  $3\sigma$  significance level, which covers the predictions of a large class of inflationary models.



**Fig. 1.7:** Theoretical CMB angular power spectra of temperature T , E- and B-mode polarization anisotropies as a function of multipole moments for the best-fit CDM parameters derived by Planck (Planck Collaboration VI 2018) with  $r = 0.06$ . The instrument noise power spectrum of polarization (dash lines) plotted for different generations of space based CMB surveys, viz. Planck, LiteBIRD and ECHO.

By improving upon both *angular resolution* and *sensitivity* across multiple frequency bands, ECHO aims to isolate the faint primordial B-mode signal from astrophysical foregrounds, instrumental noise, and gravitational lensing. Table 1.2 presents the proposed instrumental specifications, highlighting the range of frequencies, beam sizes, and expected noise levels.

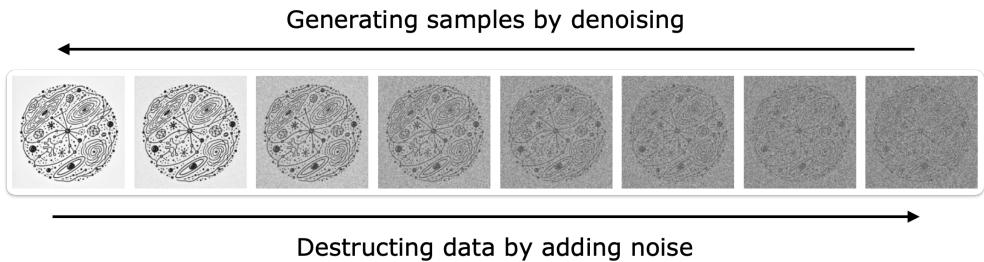
Frequency (GHz)	Beam FWHM (arcmin)	$Q$ & $U$ noise r.m.s ( $\mu\text{K.arcmin}$ )
28	39.9	16.5
35	31.9	13.3
45	24.8	11.9
65	17.1	8.9
75	14.91	5.1
95	11.7	4.6
115	9.72	3.1
130	8.59	3.1
145	7.70	2.4
165	6.77	2.5
190	5.88	2.8
220	5.08	3.3
275	4.06	6.3
340	3.28	11.4
390	2.86	21.9
450	2.48	43.4
520	2.14	102.0
600	1.86	288.0
700	1.59	1122.0
850	1.31	9550.0

**Tab. 1.2:** ECHO instrument specification as proposed in the CMB Bharat proposal.

## 2. SCORE BASED GENERATIVE MODELS

Generative models are machine learning models that aim to generate new samples by understanding the distribution of a dataset. These models have applications in various domains, including physics (Heurtel-Depeiges, Burkhart, Ohana & Blancard, 2023; Kansal, Li, Duarte, Chernyavskaya, Pierini, Orzari & Tomei, 2023; Ullmo, Decelle & Aghanim, 2021; Zhao, Ting, Diao & Mao, 2023; Flöss, Coulton, Duivenvoorden, Villaescusa-Navarro & Wandelt, 2024). Several types of generative models exist in the literature, like auto-regressive models (Chen, Mishra, Rohaninejad & Abbeel, 2018), variational auto-encoders (VAEs) (Kingma & Welling, 2019), generative adversarial networks (GANs) (Goodfellow, Pouget-Abadie, Mirza, Xu, Warde-Farley, Ozair, Courville & Bengio, 2014), normalizing flow models (Rezende & Mohamed, 2016), score models (Song & Ermon, 2020; Song, Sohl-Dickstein, Kingma, Kumar, Ermon & Poole, 2020), diffusion models (Ho, Jain & Abbeel, 2020; Rombach, Blattmann, Lorenz, Esser & Ommer, 2022). Among these, diffusion and score-based models have emerged as the state-of-the-art approaches in image and video synthesis.

Diffusion models, inspired by non-equilibrium thermodynamics, gradually introduce random noise to data through a Markov chain and then learn to reverse this process to generate structured data from noise.



**Fig. 2.1:** Visualizing Diffusion Models: Adding noise to data and then denoising to recover the original data.

Given a probability distribution function  $p(x)$ , its score function  $\nabla \log p(x)$  represents a vector field that points in the direction of the steepest increase in the probability density function. If  $p(x)$  is expressed in an unnormalized form as

$$\tilde{p}(x) = \frac{p(x)}{Z}, \quad \text{with } Z = \int \tilde{p}(x) dx,$$

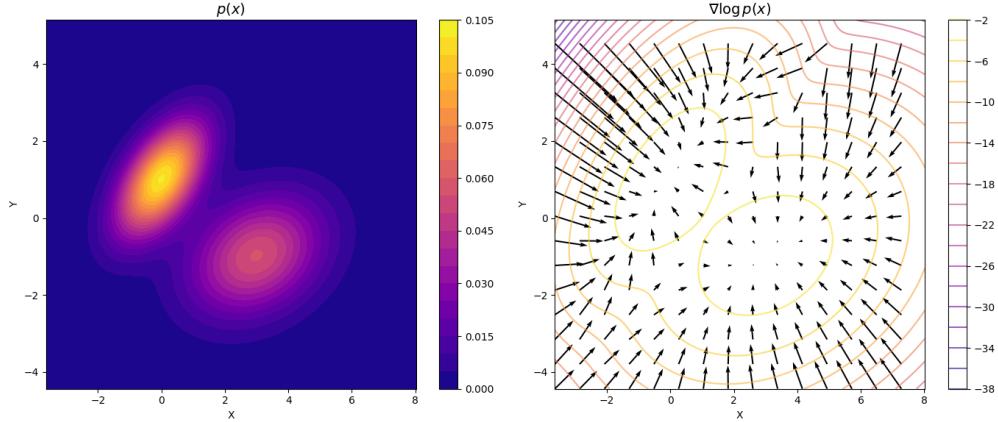
then taking the logarithm yields

$$\log \tilde{p}(x) = \log p(x) - \log Z.$$

Because  $\log Z$  is constant with respect to  $x$ , its gradient vanishes:

$$\nabla \log p(x) = \nabla \log \tilde{p}(x).$$

This cancellation is crucial. In many cases,  $Z$  is computationally intractable, but by working with the score function, we avoid needing to compute or even approximate  $Z$ . In effect, we capture the local geometry of the data distribution through gradients.



**Fig. 2.2:** Gaussian Mixture Model (GMM) fitted to a dataset with two Gaussian components. The model is trained on 300 samples from each distribution with means  $\mu_1 = (0, 1)$ ,  $\mu_2 = (3, -1)$  and covariance matrices  $\Sigma_1 = ((1, 0.5), (0.5, 1))$ ,  $\Sigma_2 = ((2, 0.5), (0.5, 1.0))$ . The left panel shows the probability density function  $p(x)$ , while the right panel visualizes the score function  $\nabla \log p(x)$  as a vector field overlaid on  $\log p(x)$  contours, indicating the direction of steepest ascent.

Hyvärinen's score matching objective (Hyvärinen, 2005) is designed to minimize the discrepancy between estimated score  $s_\theta(x)$  from the model and the true score:

$$J(\theta) = \frac{1}{2} \mathbb{E}_{p(x)} [\|s_\theta(x) - \nabla_x \log p(x)\|^2]. \quad (2.1)$$

However, equation 2.1 can not be used to learn the distribution, as we only have the data samples, not the true score of the distribution  $\nabla_x \log p(x)$ . Hyvärinen (Hyvärinen, 2005) shows an equivalent objective that does not require its direct computation.

$$J(\theta) = \frac{1}{2} \mathbb{E}_{p(x)} [\|s_\theta(x)\|^2] - \mathbb{E}_{p(x)} [s_\theta(x)^\top \nabla \log p(x)] + \frac{1}{2} \mathbb{E}_{p(x)} [\|\nabla \log p(x)\|^2].$$

The last term is independent of  $\theta$  (a constant) and can be dropped. By applying integration by parts we get:

$$\mathbb{E}_{p(x)} [s_\theta(x)^\top \nabla \log p(x)] = -\mathbb{E}_{p(x)} [\nabla \cdot s_\theta(x)],$$

where  $\nabla \cdot s_\theta(x)$  denotes the divergence of the score function estimate. Thus, the objective can be reformulated as:

$$J(\theta) = \mathbb{E}_{p(x)} \left[ \frac{1}{2} \|s_\theta(x)\|^2 + \nabla \cdot s_\theta(x) \right] + \text{const.} \quad (2.2)$$

Although this formulation allows us to optimize  $s_\theta(x)$  without needing the true score explicitly, it introduces a significant computational challenge. In particular, the equivalent formulation requires computing the divergence term, which involves the trace of the Hessian of the log-density function ( $\nabla_x^2 \log p(x)$ ). The computation of this trace would require a number of backward passes proportional to the data dimensionality  $D$  (i.e., one for each diagonal element). In practice,  $D$  can be many thousands, which can render score matching too slow for practical purposes.

To solve this computational problem, [Song & Ermon \(2020\)](#) proposed slice score matching. Instead of computing the full Hessian trace, slice score matching projects the score function onto random directions, significantly reducing the computational cost. Specifically, let  $v$  be a random vector drawn uniformly from the unit sphere  $S^{d-1}$  in  $\mathbb{R}^d$ . The slice score matching objective is then defined as:

$$J_{\text{slice}}(\theta) = \mathbb{E}_{p(x)} \mathbb{E}_{v \sim \mathcal{U}(S^{d-1})} \left[ \frac{1}{2} (v^\top s_\theta(x))^2 + v^\top \nabla_x (v^\top s_\theta(x)) \right]. \quad (2.3)$$

Here,  $v^\top s_\theta(x)$  represents the projection of the score function along  $v$ , and  $v^\top \nabla_x (v^\top s_\theta(x))$  corresponds to the second-order derivative along that direction. By averaging over random directions  $v$ , slice score matching approximates the full divergence (i.e., the trace of the Hessian) without the need for multiple backward passes, making it significantly more efficient for high-dimensional data.

## 2.1 Score-Based Generative Modeling through Stochastic Differential Equations

Stochastic Differential Equations (SDEs) generalize Ordinary Differential Equations (ODEs) by introducing randomness into the system dynamics.

The general form of an SDE is given by:

$$d\mathbf{x} = f(\mathbf{x}, t) dt + g(\mathbf{x}, t) d\mathbf{w}, \quad (2.4)$$

where:

- $f(\mathbf{x}, t)$  is the drift coefficient, governing the deterministic evolution of  $\mathbf{x}(t)$ .
- $g(\mathbf{x}, t)$  is the diffusion coefficient, controlling the intensity of the stochastic component.
- $d\mathbf{w}$  represents the differential of a standard Wiener process (also known as Brownian motion), which introduces randomness into the system.

In score-based generative models through SDEs, the SDE formulation provides a framework for transforming structured data into noise through a stochastic process. More importantly, for any stochastic differential equation of the form given in Eq. 2.4, Anderson (1982) demonstrated that there exists a reverse SDE of the form:

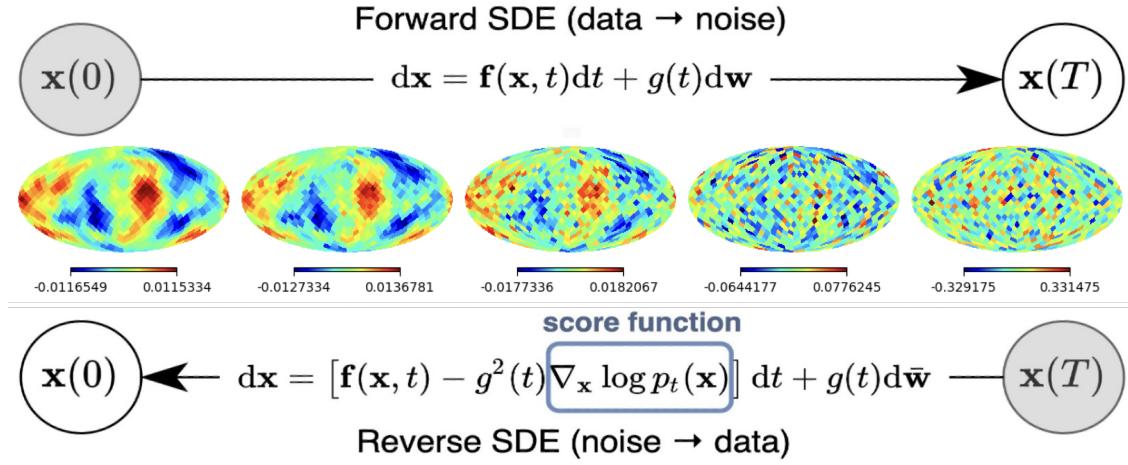
$$d\mathbf{x} = [f(\mathbf{x}, t) - g(\mathbf{x}, t)^2 \nabla_{\mathbf{x}} \log p_t(\mathbf{x})] dt + g(\mathbf{x}, t) d\bar{\mathbf{w}}, \quad (2.5)$$

where:

- $\bar{\mathbf{w}}$  is a standard Wiener process in the reverse-time direction.
- $dt$  represents an infinitesimal negative time step.
- $\nabla_{\mathbf{x}} \log p_t(\mathbf{x})$  is the score function.

This reverse-time SDE enables generative modeling by reconstructing structured data from noise, using score estimates learned through deep neural networks. There exist various choices for the drift  $f(\mathbf{x}, t)$  and diffusion  $g(\mathbf{x}, t)$  functions, which we will explore in detail later in this chapter.

To train a model that predicts the score function, we need to define a suitable objective. In score-based generative models, our aim is to learn a neural network  $s_{\theta}(x_t, t)$  that approximates the true score  $\nabla_{x_t} \log p(x_t)$ . However, since the true score is not directly accessible, we resort to a score matching approach by assuming a Gaussian noising process.



**Fig. 2.3:** Transforming data to a simple noise distribution can be accomplished with an SDE. It can be reversed to generate samples from noise if we know the score of the distribution at each intermediate time step.

Consider the forward noising process:

$$x_t = x_0 + \sigma_t \epsilon, \quad \epsilon \sim \mathcal{N}(0, I). \quad (2.6)$$

This defines the conditional density:

$$p(x_t | x_0) = \mathcal{N}(x_t; x_0, \sigma_t^2 I) = \frac{1}{(2\pi\sigma_t^2)^{D/2}} \exp\left(-\frac{\|x_t - x_0\|^2}{2\sigma_t^2}\right). \quad (2.7)$$

Taking the logarithm gives:

$$\log p(x_t | x_0) = -\frac{D}{2} \log(2\pi\sigma_t^2) - \frac{\|x_t - x_0\|^2}{2\sigma_t^2}.$$

The true score is the gradient with respect to  $x_t$ :

$$\nabla_{x_t} \log p(x_t | x_0) = -\frac{1}{\sigma_t^2} (x_t - x_0).$$

Using the reparameterization  $x_t - x_0 = \sigma_t \epsilon$ , we obtain:

$$\nabla_{x_t} \log p(x_t | x_0) = -\frac{\epsilon}{\sigma_t}.$$

Since we do not have access to the true score during training, we train the neural network

$s_\theta(x_t, t)$  to match the above expression. This leads us to define the following loss function:

$$\mathcal{L}(\theta) = \mathbb{E}_{t, x_0, \epsilon} \left[ \lambda(t) \left\| s_\theta(x_t, t) + \frac{\epsilon}{\sigma_t} \right\|^2 \right], \quad (2.8)$$

where  $\lambda(t)$  is a weighting function that balances the contributions from different noise levels.

This loss function makes the model approximate the score function at each time step  $t$ , which is essential for effectively reversing the diffusion process to recover structured data from noise.

### 2.1.1 Variance Exploding (VE) SDE

$$f(x, t) = 0 \quad \text{and} \quad g(x, t) = g(t) \quad (2.9)$$

The forward SDE is given by:

$$d\mathbf{x} = g(t) d\mathbf{w},$$

where  $d\mathbf{w}$  is the increment of a standard Wiener process (i.e., Brownian motion), and  $g(t)$  controls the noise level at time  $t$ .

For the VE SDE, the diffusion coefficient  $g(t)$  is directly related to the time-dependent noise scale  $\sigma(t)$ . A widely used approach is to define an exponential noise schedule as:

$$\sigma(t) = \sigma_{\min} \left( \frac{\sigma_{\max}}{\sigma_{\min}} \right)^t, \quad t \in [0, 1]. \quad (2.10)$$

At  $t = 0$ , the noise level is minimal, i.e.,  $\sigma(0) = \sigma_{\min}$ , and at  $t = 1$ , the noise reaches its maximum value,  $\sigma(1) = \sigma_{\max}$ .

The forward SDE is written to ensure that the variance of  $\mathbf{x}$  is  $\sigma(t)^2$ . One way to express it is:

$$d\mathbf{x} = \sqrt{\frac{d}{dt} \sigma(t)^2} d\mathbf{w}. \quad (2.11)$$

Since

$$\sigma(t)^2 = \sigma_{\min}^2 \left( \frac{\sigma_{\max}}{\sigma_{\min}} \right)^{2t},$$

its time derivative is:

$$\frac{d}{dt} \sigma(t)^2 = 2 \sigma(t)^2 \ln \left( \frac{\sigma_{\max}}{\sigma_{\min}} \right).$$

Thus, the forward SDE becomes:

$$d\mathbf{x} = \sqrt{2 \ln \left( \frac{\sigma_{\max}}{\sigma_{\min}} \right) \sigma(t)} d\mathbf{w}$$

$$d\mathbf{x} = \sigma_{\min} \left( \frac{\sigma_{\max}}{\sigma_{\min}} \right)^t \sqrt{2 \log \frac{\sigma_{\max}}{\sigma_{\min}}} d\mathbf{w} \quad (2.12)$$

### Marginal Probability Standard Deviation

Since the forward SDE is driftless, the cumulative variance added to the data from time 0 to  $t$  is given by

$$\text{Var}(x(t)|x(0)) = \int_0^t g(s)^2 ds. \quad (2.13)$$

Substitute the expression for  $g(s)^2$ :

$$\text{Var}(x(t)|x(0)) = \int_0^t 2 \ln \left( \frac{\sigma_{\max}}{\sigma_{\min}} \right) \sigma(s)^2 ds.$$

Using  $\sigma(s)^2 = \sigma_{\min}^2 \left( \frac{\sigma_{\max}}{\sigma_{\min}} \right)^{2s}$ , we obtain

$$\text{Var}(x(t)|x(0)) = 2 \ln \left( \frac{\sigma_{\max}}{\sigma_{\min}} \right) \sigma_{\min}^2 \int_0^t \left( \frac{\sigma_{\max}}{\sigma_{\min}} \right)^{2s} ds.$$

Let

$$a = \left( \frac{\sigma_{\max}}{\sigma_{\min}} \right)^2.$$

Then the integral becomes

$$\int_0^t a^s ds = \frac{a^t - 1}{\ln a} = \frac{\left( \frac{\sigma_{\max}}{\sigma_{\min}} \right)^{2t} - 1}{2 \ln \left( \frac{\sigma_{\max}}{\sigma_{\min}} \right)}.$$

Substitute back to get

$$\text{Var}(x(t)|x(0)) = 2 \ln \left( \frac{\sigma_{\max}}{\sigma_{\min}} \right) \sigma_{\min}^2 \cdot \frac{\left( \frac{\sigma_{\max}}{\sigma_{\min}} \right)^{2t} - 1}{2 \ln \left( \frac{\sigma_{\max}}{\sigma_{\min}} \right)} = \sigma_{\min}^2 \left[ \left( \frac{\sigma_{\max}}{\sigma_{\min}} \right)^{2t} - 1 \right]. \quad (2.14)$$

Taking the square root, the marginal probability standard deviation is

$$\text{std}(t) = \sqrt{\text{Var}(x(t)|x(0))} = \sigma_{\min} \sqrt{\left( \frac{\sigma_{\max}}{\sigma_{\min}} \right)^{2t} - 1}. \quad (2.15)$$

This  $\text{std}(t)$  represents the cumulative noise added to the data up to time  $t$  and is used to normalize the score in the denoising score matching objective.

### Reverse-Time VE SDE

For generation, we need to reverse the corruption process. The reverse-time SDE is derived by reversing the forward diffusion and incorporates the *score function*  $\nabla_x \log p_t(\mathbf{x})$ :

$$d\mathbf{x} = \left[ -\frac{d}{dt}\sigma(t)^2 \nabla_x \log p_t(x) \right] dt + \sqrt{\frac{d}{dt}\sigma(t)^2} d\bar{\mathbf{w}},$$

where  $d\bar{\mathbf{w}}$  denotes the reverse-time Wiener process.

$$d\mathbf{x} = \left[ -2\sigma(t)^2 \ln\left(\frac{\sigma_{\max}}{\sigma_{\min}}\right) \nabla_x \log p_t(x) \right] dt + \sqrt{2 \ln\left(\frac{\sigma_{\max}}{\sigma_{\min}}\right)} \sigma(t) d\bar{\mathbf{w}} \quad (2.16)$$

## 2.2 Euler–Maruyama (EM) Discretization

To numerically solve the SDE using the Euler–Maruyama method, we discretize the time interval  $[0, 1]$  into  $N$  steps with  $\Delta t = 1/N$  i.e., we have the time points  $t_n = n\Delta t$  for  $n = 0, 1, 2, 3, \dots, N$ , where  $N\Delta t = 1$  is the final time.

For a general SDE of the form Eq 2.4, the EM method updates the solution  $x(t)$  using the following rule

$$x_{n+1} = x_n + f(x_n, t_n) \Delta t + g(x_n, t_n) \Delta \mathbf{w}_n \quad (2.17)$$

- $x_n$  is the approximation of  $x(t_n)$
- $\Delta \mathbf{w}_n$  is the Brownian increment given by

$$\Delta \mathbf{w}_n = \mathbf{w}_{t_{n+1}} - \mathbf{w}_{t_n} \sim \sqrt{\Delta t} \mathcal{N}(0, 1) \sim \mathcal{N}(0, \sqrt{\Delta t}) \quad (2.18)$$

Similarly, for the Reverse Stochastic Differential Equation 2.5 we discretize as follows.

$$x_{n+1} = x_n + [f(x_n, t_n) - g(x_n, t_n)^2 \nabla_x \log p_{t_n}(x_n)] \Delta t + g(x_n, t_n) \bar{\mathbf{w}}_n. \quad (2.19)$$

- $x_n$  is the approximation of  $x(t_n)$  along the reverse-time trajectory.
- $\Delta \bar{\mathbf{w}}_n$  is the increment of the reverse Brownian motion.

The reverse process is integrated backward in time, so we define  $t_0 = 1$  and  $t_N = 0$  (or equivalently, we can use a negative  $\Delta t$  when moving backward).

### 2.2.1 EM for VE SDE

#### Forward VE SDE

$$x_{k+1} = x_k + g(t_k) \Delta \mathbf{w}_k, \quad (2.20)$$

where  $\Delta \mathbf{w}_k$  is the Brownian increment given by

$$\Delta \mathbf{w}_k = \sqrt{\Delta t} \xi_k, \quad \xi_k \sim \mathcal{N}(0, 1). \quad (2.21)$$

Substituting the expression for  $g(t)$  gives

$$x_{k+1} = x_k + \sqrt{\frac{2}{T} \ln\left(\frac{\sigma_{\max}}{\sigma_{\min}}\right)} \sigma(t_k) \sqrt{\Delta t} \xi_k.$$

with

$$\sigma(t_k) = \sigma_{\min} \left( \frac{\sigma_{\max}}{\sigma_{\min}} \right)^{t_k}.$$

#### Reverse VE SDE

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \frac{2 \sigma(t_k)^2}{T} \ln\left(\frac{\sigma_{\max}}{\sigma_{\min}}\right) \nabla_x \log p_{t_k}(\mathbf{x}_k) \Delta t + \sqrt{\frac{2}{T} \ln\left(\frac{\sigma_{\max}}{\sigma_{\min}}\right)} \sigma(t_k) \Delta \bar{\mathbf{w}}_k, \quad (2.22)$$

where

$$\Delta \bar{\mathbf{w}}_k \sim \sqrt{\Delta t} \xi_k \sim \mathcal{N}(0, \Delta t).$$

This discretization allows us to simulate the continuous-time SDEs step-by-step, providing a numerical solution to both the forward and reverse processes.

## 3. COSMOLOGICAL SIMULATIONS

In this chapter, we outline the end-to-end simulation framework to model the Cosmic Microwave Background (CMB) temperature and polarization signals. We begin by discussing how the theoretical angular power spectra ( $C_\ell$ ) are computed using the *Code for Anisotropies in the Microwave Background* (CAMB).

We then describe the use of **HEALPix** for creating full-sky CMB realizations, focusing on its hierarchical pixelization and associated Python wrapper, **healpy**. Next, we incorporate instrumental effects—specifically beam convolution and noise—to demonstrate their impact on the underlying sky signal in realistic observations. Finally, we introduce the PySM3 software for simulating foreground emission from the Galaxy. Combining these components (theory, instrumentation, and foregrounds) provides a comprehensive approach for producing mock CMB data sets under realistic observing conditions, which is crucial for testing data analysis pipelines and forecasting the detectability of primordial B-mode polarization.

### 3.1 Theoretical Angular Power Spectrum using CAMB

To compute the theoretical angular power spectra ( $C_\ell$ ) for both temperature and polarization, we use the publicly available **Code for Anisotropies in the Microwave Background (CAMB)**<sup>1</sup> (Lewis, Challinor & Lasenby, 2000). CAMB is based on solving the Einstein–Boltzmann equations for linearized perturbations in the early universe, outputting accurate theoretical predictions of the CMB power spectra given a set of cosmological parameters.

#### Cosmological Parameters

Throughout this work, we adopt the best-fit  $\Lambda$ CDM cosmological parameters from Planck 2018 (Aghanim, Akrami, Ashdown & et al., 2020). These parameters, along with the tensor-to-scalar ratio  $r$ , are summarized in Table 3.1. Unless otherwise stated, all theoretical  $C_\ell$  calculations assume these values.

---

<sup>1</sup> <https://camb.readthedocs.io/en/latest/>

Parameter	Value
Baryon density, $\Omega_b h^2$	0.02238
Cold dark matter density, $\Omega_c h^2$	0.1201
Hubble parameter, $H_0$ ( $\text{km s}^{-1} \text{Mpc}^{-1}$ )	67.32
Optical depth, $\tau_d$	0.05431
Scalar spectral index, $n_s$	0.966
Amplitude of perturbations, $A_s$	$2.1 \times 10^{-9}$
Tensor-to-scalar ratio, $r$	0.001

**Tab. 3.1:** Cosmological parameters used to generate theoretical CMB power spectra with CAMB.

With these parameters and a specified tensor-to-scalar ratio  $r = 0.001$ , CAMB provides us with  $C_\ell^{TT}$ ,  $C_\ell^{EE}$ ,  $C_\ell^{BB}$ , and cross-spectra such as  $C_\ell^{TE}$ . Algorithm 1 gives us the computational procedure to generate the theoretical angular power spectra.

---

**Algorithm 1** Generation of Theoretical B-mode Spectra with CAMB

---

**Require:**  $\text{nside} \leftarrow 16$  (*nside defines the CMB map resolution*)

- 1: **Import:** `numpy`, `matplotlib`, `healpy`, `camb`
  - 2: `pars`  $\leftarrow$  `camb.CAMBparams(WantTensors=True)`
  - 3: `pars.set_cosmology(H0 = 67.32, omhb2 = 0.0224, omch2 = 0.1201, Alens = 1)`
  - 4: `pars.InitPower.set_params(ns = 0.966, r = 0.001)`
  - 5: Configure reionization: set  $\tau \leftarrow 0.05431$  and enable optical depth.
  - 6: `results`  $\leftarrow$  `camb.get_results(pars)`
  - 7: `powers`  $\leftarrow$  `results.get_cmb_power_spectra(pars, CMBunit = \mu K, lmax = 2 \times \text{nside} - 1, raw_cl = \text{True})`
  - 8: **Note:** `powers` contains the total, lensing, scalar, and tensor power spectra for TT, EE, BB, and TE modes.
- 

## 3.2 HEALPix

HEALPix<sup>2</sup> (*Hierarchical Equal Area iso-Latitude Pixelization*) is a framework for discretizing and Fast Analysis of Data Distributed on the Sphere (Gorski, Hivon, Banday, Wandelt, Hansen, Reinecke & Bartelmann, 2005). It divides the celestial sphere into pixels of equal area, enabling efficient spherical harmonic transforms and map-based data analysis—both crucial for CMB studies. Each map is determined by an integer parameter `nside`, which sets the total number of pixels  $N_{\text{pix}} = 12 (\text{nside})^2$ .

<sup>2</sup> <https://healpix.sourceforge.io/>

NSIDE	1	2	4	8	16	32	64	128	256
Npix	12	48	192	768	3072	12288	49152	196608	786432

**Tab. 3.2:** Number of Pixels for corresponding NSIDE value.

`healpy`<sup>3</sup> is a Python wrapper for HEALPix, providing high-level functions to generate and manipulate maps in Python scripts. Some essential functions include:

- `synfast`: Computes random realizations of sky maps given input power spectra ( $C_\ell$ ).
- `anafast`: Computes the power spectrum of a Healpix map, or the cross-spectrum between two maps if `map2` is given.
- `map2alm` and `alm2map`: Computes the  $a_{\ell m}$  of a Healpix map and Computes a Healpix map given the  $a_{\ell m}$  respectively.
- `alm2cl`: Computes (cross-)spectra from  $a_{\ell m}(s)$ .
- `smoothing`: Convolves a sky map with a specified Gaussian beam in real space.
- `mollview`, `cartview`, `gnomview`: Visualization functions for displaying sky maps in different projections, enabling quick inspection of large-scale features or small-scale details of the CMB anisotropies.

By combining `healpy` with the theoretical spectra from CAMB, we can generate mock CMB temperature and polarization maps at a chosen `nside`, and then apply the instrumental and observational effects (beam convolution, noise, etc.) described in Section 3.3. This end-to-end simulation framework allows us to test data analysis pipelines under realistic conditions, assessing how well we can recover the faint B-mode signal.

### 3.3 Beam Convolution and Instrumental Noise

In any real CMB experiment, the true sky signal in harmonic space is represented by the spherical harmonic coefficients  $a_{\ell m}$  of the underlying temperature or polarization fields. However, this true sky signal is inevitably altered by instrumental effects, specifically beam convolution and instrumental noise.

Every telescope possesses a finite angular resolution characterized primarily by its beam pattern, typically approximated as a circular Gaussian with a given full width at half maximum (FWHM),  $\theta_{\text{FWHM}}$ . This finite resolution effectively smooths out small-scale anisotropies,

---

<sup>3</sup> <https://github.com/healpy/healpy>

leading to observed harmonic coefficients being damped versions of the true sky coefficients  $a_{\ell m}$ .

Mathematically, beam convolution in harmonic space can be expressed by multiplying the true spherical harmonic coefficients by a beam window function  $b_\ell$ :

$$\tilde{a}_{\ell m} = b_\ell a_{\ell m}, \quad (3.1)$$

where for a Gaussian beam,  $b_\ell$  is given by:

$$b_\ell = \exp \left[ -\frac{\ell(\ell+1)\sigma_b^2}{2} \right], \quad \text{with} \quad \sigma_b = \frac{\theta_{\text{FWHM}}}{\sqrt{8 \ln 2}}. \quad (3.2)$$

As the multipole moment  $\ell$  increases, corresponding to smaller angular scales, the beam window function  $b_\ell$  rapidly decreases, suppressing the observed power at high  $\ell$ .

---

**Algorithm 2** Simulating Primordial B-mode Realizations with Beam Effects

---

**Require:** B-mode spectrum `cl_BB_tensor` (from Algorithm 1), `nside`, number of realizations `N_sim`, beam FWHM  $\theta_{\text{FWHM}}$

```

1: Set random seed (e.g., np.random.seed(seed))
2: Compute  $\sigma_b \leftarrow \theta_{\text{FWHM}} / \sqrt{8 \ln 2}$ 
3:  $\ell_{\text{max}} \leftarrow 2 \times \text{nside} - 1$ 
4: for  $\ell = 0$  to  $\ell_{\text{max}}$  do
5:    $b_\ell \leftarrow \exp \left[ -\frac{\ell(\ell+1)\sigma_b^2}{2} \right]$  (beam window function)
6: end for
7: Initialize CL_BB  $\leftarrow []$ 
8: for  $i = 1$  to N_sim do
9:   Generate a B-mode map:  $m \leftarrow \text{hp.synfast}(\text{cl_BB_tensor}, \text{nside})$ 
10:  Compute observed spectrum: cl_obs  $\leftarrow \text{hp.anafast}(m)$ 
11:  Apply beam smoothing:  $\text{cl\_obs\_beam}(\ell) \leftarrow \text{cl\_obs}(\ell) b_\ell^2$  for all  $\ell$ 
12:  Append cl_obs_beam to CL_BB
13: end for
14: return CL_BB

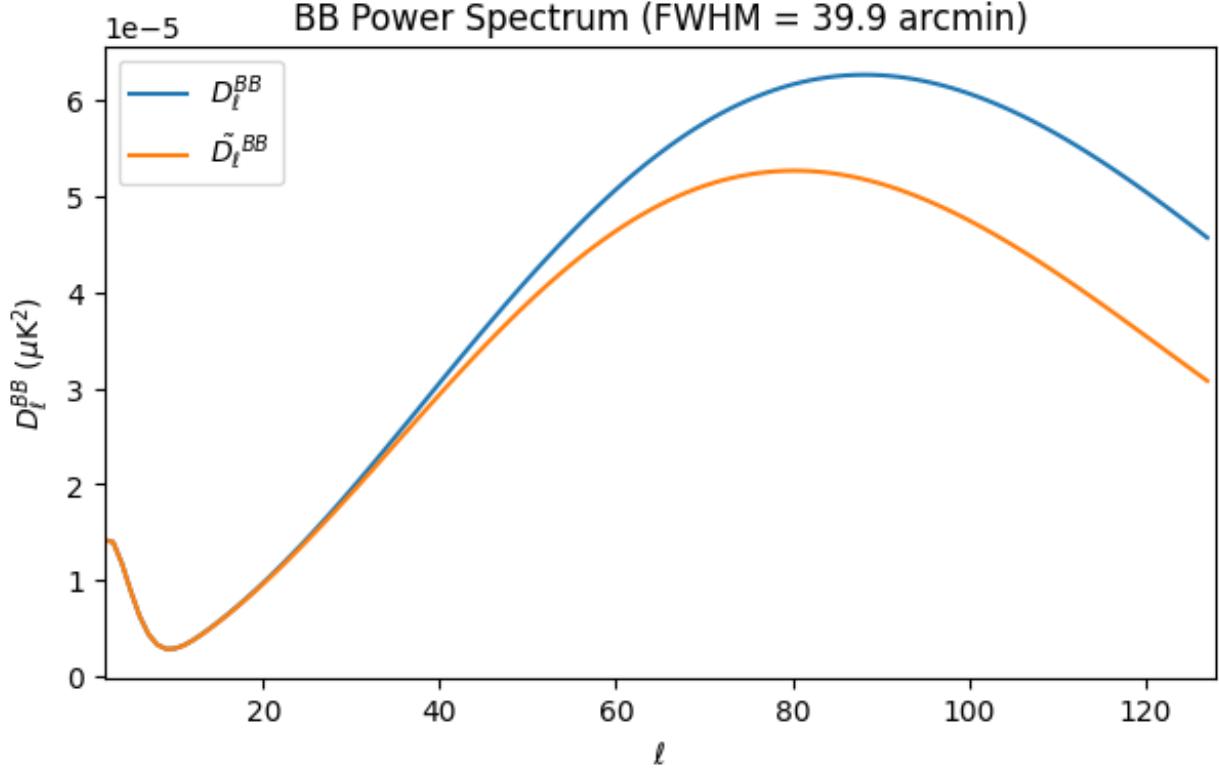
```

---

Additionally, observed data includes instrumental noise, which arises predominantly from thermal (Johnson) noise, photon shot noise, and electronic readout noise. These noise contributions are typically assumed to be Gaussian, white, and uncorrelated, described by noise harmonic coefficients  $n_{\ell m}$  with zero mean:

$$\langle n_{\ell m} \rangle = 0, \quad \langle n_{\ell m} n_{\ell' m'}^* \rangle = N_\ell \delta_{\ell \ell'} \delta_{mm'}. \quad (3.3)$$

Including both beam convolution and instrumental noise, the observed harmonic coeffi-



**Fig. 3.1:** Example illustrating how beam convolution suppresses power at high multipoles.

cients become:

$$\tilde{a}_{\ell m} = b_{\ell} a_{\ell m} + n_{\ell m}, \quad (3.4)$$

and thus the observed power spectrum is:

$$\tilde{C}_{\ell} = b_{\ell}^2 C_{\ell} + N_{\ell}, \quad \text{where } C_{\ell} = \langle |a_{\ell m}|^2 \rangle. \quad (3.5)$$

To simulate the instrumental noise realistically in pixel-space maps, we first calculate the pixel variance ( $\sigma_v^2$ ) using the provided noise r.m.s. ( $\sigma_{rms}$ ) values from Table 1.2. This variance is computed as follows:

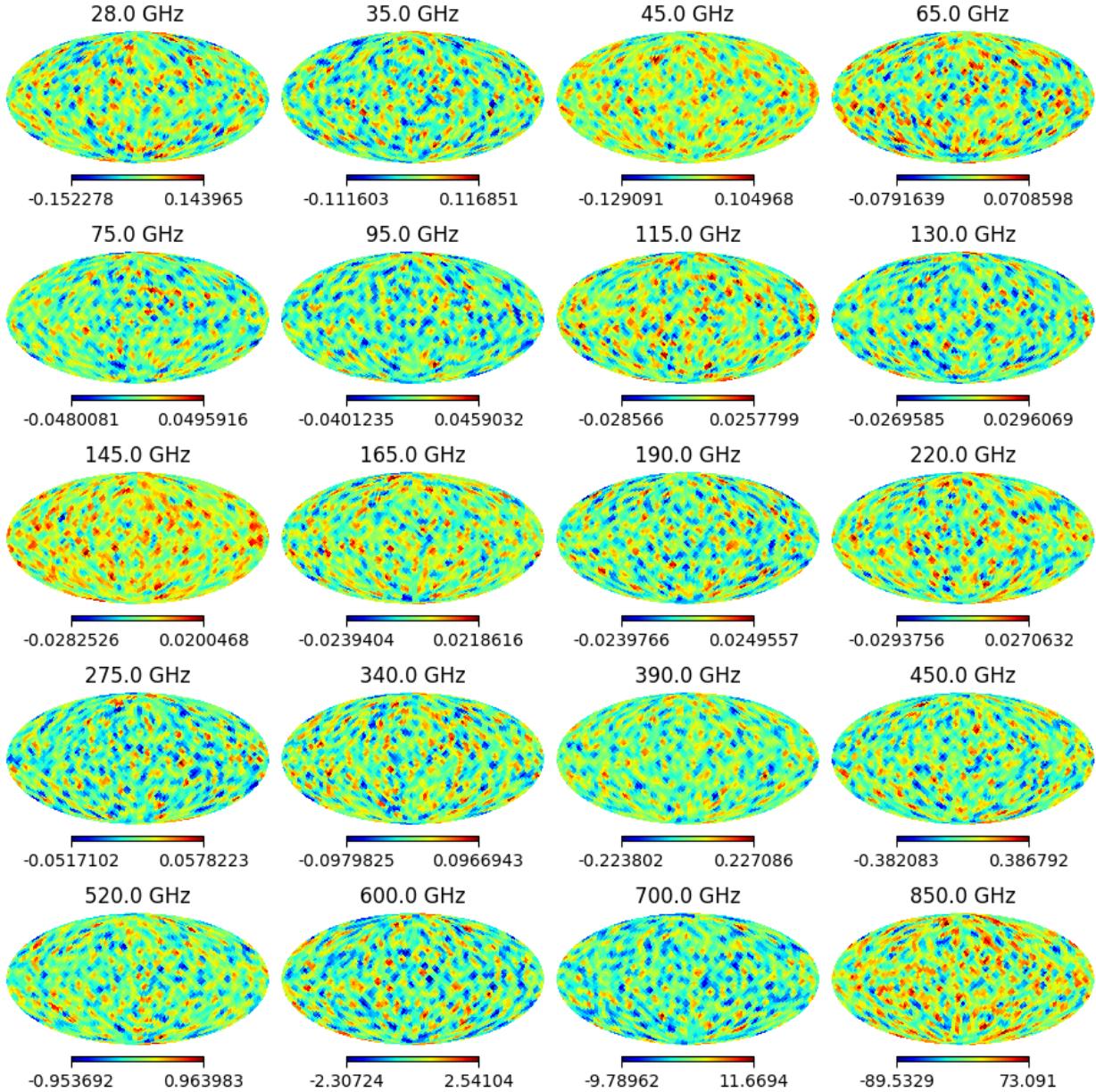
$$\sigma_v^2 = \left( \frac{\sigma_{rms} \times \pi}{60 \times 180} \right)^2 \left( \frac{N_{pix}}{4\pi} \right), \quad (3.6)$$

where  $N_{pix}$  is the total number of pixels used in the map.

Using this pixel variance, we generate Gaussian-distributed noise maps for the Stokes parameters  $Q$  and  $U$  by drawing random realizations:

$$Q_{noise}(\hat{n}) \sim \mathcal{N}(0, \sigma_v) \quad , \quad U_{noise}(\hat{n}) \sim \mathcal{N}(0, \sigma_v).$$

These simulated noise maps are then transformed into harmonic space using spherical harmonic transforms (such as provided by `healpy`), enabling the calculation of noise power spectra ( $N_\ell$ ).



**Fig. 3.2:** Random realizations of simulated B mode noise maps at each ECHO Frequency

Specifically, the noise B-mode power spectrum can be computed from these simulated maps as:

$$N_\ell^{BB} = \langle |n_{\ell m}^B|^2 \rangle, \quad (3.7)$$

where  $n_{\ell m}^B$  are the harmonic coefficients obtained from the noise-only maps. This procedure

**Algorithm 3** Simulating Noise  $C_\ell$  for B-modes

---

**Require:** NSIDE, noise RMS  $rms$  ( $\mu\text{K}$ ) at a frequency ( $v$ ), number of realizations  $N$ , seed

```

1:  $N_{\text{pix}} \leftarrow \text{hp.nside2npix}(\text{NSIDE})$ 
2:  $\sigma_v \leftarrow \sqrt{\left(\frac{rms \cdot \pi}{60 \times 180}\right)^2 \cdot \frac{N_{\text{pix}}}{4\pi}}$ 
3: Initialize noise_cl  $\leftarrow []$ 
4: for  $i = 0$  to  $N - 1$  do
5:   Set np.random.seed(seed+i)
6:   map_q  $\leftarrow \sigma_v \times \mathcal{N}(0, 1, N_{\text{pix}})$ 
7:   Set np.random.seed(24214  $\times$  seed + i)
8:   map_u  $\leftarrow \sigma_v \times \mathcal{N}(0, 1, N_{\text{pix}})$ 
9:   map_t  $\leftarrow \mathbf{0}$  {Zero temperature map}
10:  maps  $\leftarrow [\text{map}_q, \text{map}_u, \text{map}_t]$ 
11:   $(alm\_T, alm\_E, alm\_B) \leftarrow \text{hp.map2alm}(\text{maps}, lmax = 2 \times NSIDE - 1, pol = \text{True})$ 
12:  Append hp.alm2cl( $alm\_B$ ) to noise_cl
13: end for
14: return noise_cl

```

---

allows accurate characterization and modeling of instrumental noise in analyses aimed at extracting faint cosmological signals such as primordial B-mode polarization.

### 3.4 CMB Foregrounds and PySM3

In addition to instrumental effects (Section 3.3), cosmic microwave background (CMB) observations are contaminated by astrophysical foregrounds originating from within our Galaxy. These foregrounds must be accurately modeled and subtracted to isolate the primordial B-mode signal. The dominant Galactic emission processes at microwave frequencies are:

#### 1. Synchrotron Emission

Synchrotron emission is produced by relativistic electrons spiraling around Galactic magnetic field lines. In intensity space, a simple power-law dependence on frequency  $\nu$  is often assumed:

$$I_{\text{syn}}(\nu) \propto \nu^{\beta_{\text{syn}}}, \quad (3.8)$$

where  $\beta_{\text{syn}}$  is the synchrotron spectral index (typically around  $-3$  for the brightness temperature). At lower frequencies ( $\lesssim 60$  GHz), synchrotron usually dominates the polarized foreground signal.

#### 2. Thermal Dust Emission

Dust grains in the interstellar medium absorb stellar/UV radiation and re-emit thermally at microwave to submillimeter wavelengths. This can be approximated as a *modified blackbody* (MBB):

$$I_{\text{dust}}(\nu) = A_0 \left( \frac{\nu}{\nu_0} \right)^{\beta_d} B_\nu(T_d), \quad (3.9)$$

where  $\beta_d$  is the dust emissivity index,  $T_d$  is the dust temperature, and  $B_\nu(T_d)$  is the Planck function. At higher frequencies ( $\gtrsim 100$  GHz), thermal dust dominates the polarized foreground.

### 3. Free–Free Emission

Free–free emission arises when free electrons scatter off ions in ionized gas (H II regions). The brightness temperature can be approximated by:

$$T_{\text{ff}}(\nu) \approx T_e (1 - e^{-\tau_\nu}), \quad \text{where } \tau_\nu \propto \nu^{-2}, \quad (3.10)$$

and  $T_e$  is the electron temperature (a few thousand Kelvin). Free–free emission is typically *weakly* polarized (often assumed nearly zero), but it can significantly contribute to the total intensity over tens to a few hundred GHz.

### 4. Anomalous Microwave Emission (AME)

AME is commonly attributed to *spinning dust* grains, which possess electric dipole moments that radiate when the grains rotate rapidly. The spectral shape often peaks near 20–40 GHz. While detailed models can be complex (e.g. using SpDust2), a simplified parametric approach might look like:

$$I_{\text{AME}}(\nu) \approx A_{\text{ame}} f(\nu; \nu_{\text{peak}}, \dots), \quad (3.11)$$

where  $f(\nu; \nu_{\text{peak}})$  encodes the spinning-dust emission profile and  $A_{\text{ame}}$  is a normalization constant. Although often weaker than synchrotron or dust, AME can still impact polarization channels at these lower frequencies.

#### 3.4.1 Simulations with PySM3

A practical tool for generating foreground sky maps is PySM3 (Python Sky Model) (Zonca, Thorne, Krachmalnicoff & Borrill, 2021). By specifying a set of model presets (e.g., for dust, synchrotron, anomalous microwave emission), as well as the observing frequency and

`nside` (HEALPix resolution), PySM3 produces realistic Stokes  $I, Q, U$  maps covering the main Galactic emission mechanisms.

Model	Type	Key Features
<b>s1</b>	Synchrotron	Power-law spectral scaling with a spatially varying index, based on Haslam 408 MHz (intensity) and WMAP 23 GHz (polarization).
<b>s3</b>	Synchrotron	Extends <b>s1</b> with a curved power-law index (Kogut, 2012), capturing realistic spectral bending around 23 GHz.
<b>d1</b>	Dust	Single-component modified black body (MBB), using Planck Commander maps (Planck Collaboration 2015 (Ade, Aghanim, Arnaud & et al., 2016)). Varying temperature & spectral index for moderate-accuracy simulations.
<b>d4</b>	Dust	Two-component dust model (Finkbeiner, Davis & Schlegel, 1999), allowing multiple dust populations and higher fidelity in complex analyses.
<b>a2</b>	AME	Anomalous Microwave Emission with a 2% polarization fraction. Uses dust polarization angles for $Q/U$ , staying within WMAP/Planck bounds.

**Tab. 3.3:** Selected PySM3 model presets commonly employed in moderate to high-fidelity simulations. PySM3 also includes additional models (e.g. **d2**, **d5**, **s2**, **s5**, **a1**, etc.) for alternative spectral assumptions and small-scale treatments.

In addition to comprehensive families of dust (**d0–d12**), synchrotron (**s1–s7**), and AME (**a1–a2**) models, some presets are particularly popular for moderate to advanced simulations. Table 3.3 summarizes five such models: two for synchrotron, two for dust, and one for AME.

Models can be combined by specifying their labels in a list (`sky_config = [s3, d1, a2]`) when initializing `PySM3.Sky`. One then obtains the total foreground emission from all chosen components at each frequency of interest, typically stored as  $I, Q, U$  sky maps (§3.2). These can be added to synthetic CMB maps and instrumental noise for realistic end-to-end simulations of CMB observations.

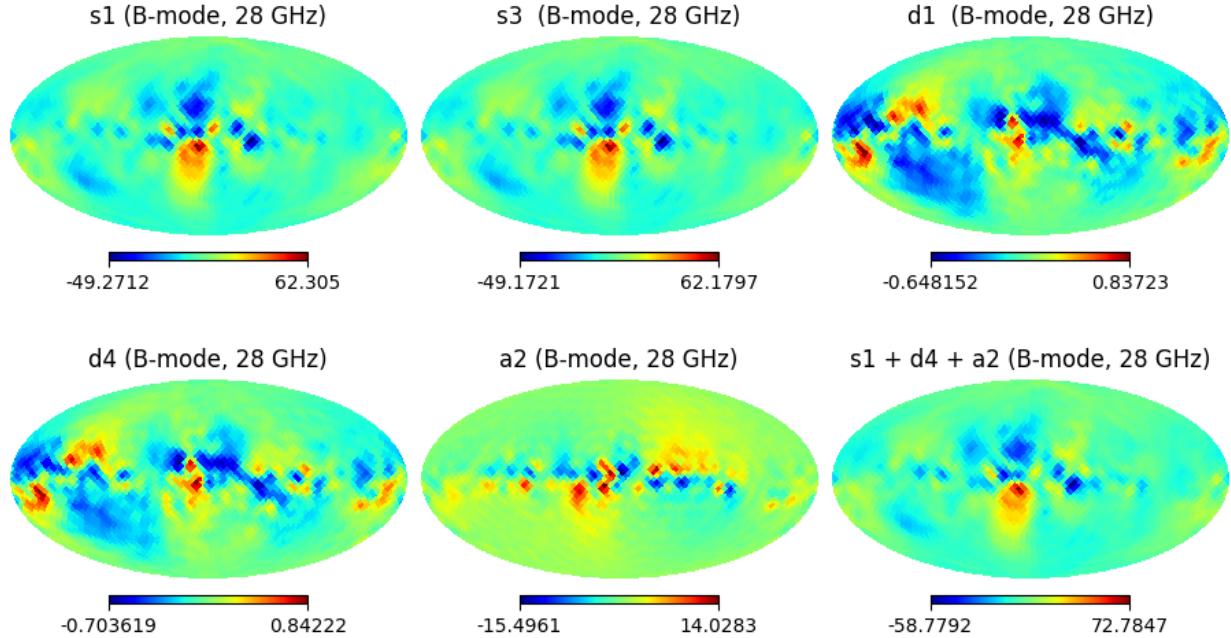
---

**Algorithm 4** Generating a Foreground Map with PySM3

---

**Require:** `sky_config`: List of model presets (e.g. `[s3, d1, a2]`) , `nside`, frequency  $\nu$  (GHz)  
**Ensure:** `maps_IQU(̂n)`: Stokes  $I, Q, U$  maps at `nside`

- 1: `sky ← PySM3.Sky(nsides = nsides, preset_strings = sky_config)`
  - 2: `maps_IQU ← sky.get_emission(ν)`
  - 3: `map_B_foreground ← [almE, almB, almT] ← maps_IQU`
  - 4: `return map_B_foreground`
-



**Fig. 3.3:** B-mode foreground maps at 28 GHz simulated using different sky models in PySM3.

Using this procedure, one can generate full-sky foreground maps at any desired frequency, as shown in Fig. 3.3 ,by combining different dust, synchrotron, and AME models (Algorithm 4).

# 4. RECONSTRUCTING THE PRIMORDIAL B-MODE SIGNAL USING SCORE-SDE

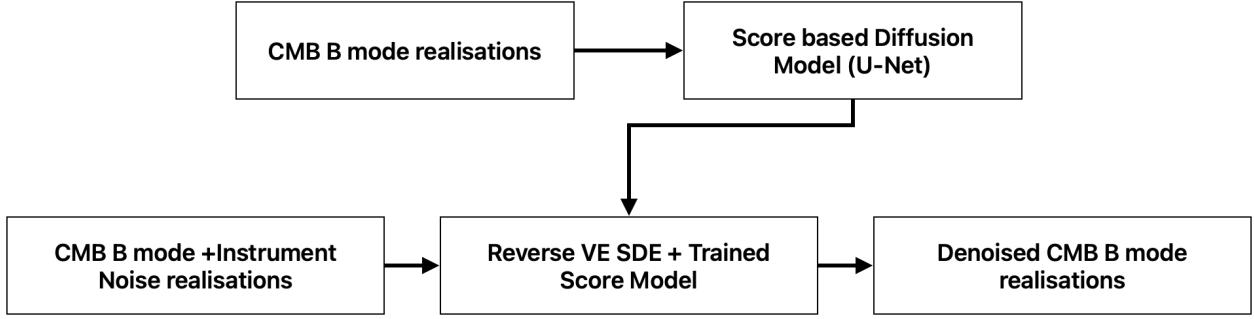
In this chapter, we present developed methodologies to reconstruct the CMB B-mode polarization signal from different contaminations. In the first section, we reconstruct full-sky CMB B-mode maps that combine both primordial and lensing contributions from data contaminated with instrumental noise. Here, we detail the ScoreUNet model, which is trained to estimate the score as a function of time—and present comprehensive results that demonstrate the network’s ability to recover the underlying CMB signal despite the presence of noise.

In the subsequent section, we take this approach a step further by reconstructing the primordial B-mode angular power spectrum directly from power spectra contaminated by gravitational lensing, instrumental noise, and complex foregrounds. Remarkably, our model accurately recovers the primordial signal even without explicit prior knowledge of the foreground components. This result demonstrates the robustness of the score-based denoising framework and its potential for isolating the faint primordial B-mode features in realistic observational scenarios.

## 4.1 Score-Based Denoising of Full-Sky CMB B-Mode Maps

In this section, we present a method to denoise simulated noisy full-sky CMB B-mode polarization maps at 520 GHz, representative of the proposed ECHO mission. Our approach combines a Variance-Exploding Stochastic Differential Equation (VE-SDE; see Section 2.1.1) with a U-Net neural network model (Ronneberger, Fischer & Brox, 2015) to estimate the underlying clean CMB signal from observational noise.

An overview of the methodology is shown in Figure 4.1. First, we train the U-Net to predict the *score function*, using clean simulated CMB maps subjected to the forward VE-SDE noise-adding (diffusion) process. After training, we apply the trained model to noisy



**Fig. 4.1:** Overview of the VE-SDE and U-Net denoising approach.

maps, reversing the diffusion via the reverse VE-SDE, effectively recovering denoised versions of the original B-mode signal.

### 4.1.1 Data Simulations

We simulate  $\sim 10,000$  full-sky CMB B-mode maps using the beam-convolved theoretical power spectrum  $\tilde{C}_\ell = b_\ell^2 C_\ell$ , where  $C_\ell$  includes both primordial ( $r = 0.001$ ) and lensing contributions, following the simulation techniques discussed earlier. For testing purposes, we create a separate set of  $\sim 1,000$  B-mode maps with distinct random seeds. Additionally, we simulate instrumental noise maps at 520 GHz using random realizations with pixel variance  $\sigma_v^2$ , determined from the noise RMS (see Table 1.2) via Algorithm 3. These noise maps are combined with the test CMB maps to produce the “observed” noisy maps used for evaluating denoising performance.

### 4.1.2 Model Development and Training

#### U-Net Architecture

The ScoreUNet model is based on a U-Net-inspired architecture designed to estimate time-dependent score functions. The model begins with a time embedding layer that encodes temporal information, which is added to the feature maps at each stage to make the network aware of the specific perturbations over time.

The encoding path starts with `nside=8` and includes two 1D convolutional layers, each with a kernel size of 4 and stride of 4. While alternative convolution methods for HEALPix data exist, such as CNN on HEALPix sphere (Krachmalnicoff & Tomasi, 2019) and DeepSphere (Perraudin, Defferrard, Kacprzak & Sgier, 2019). We use the defined 1D convolutions for simplicity. The resolution goes from `nside 8` to `nside 4` and then to `nside 2`, while the number of feature channels goes from 0 to 128 and then to 64. Each convolution layer is

followed by group normalization for training stability. In the decoding path, transposed convolutions restore the resolution, first to  $n_{\text{side}} = 4$  and finally back to  $n_{\text{side}} = 8$ , with skip connections retaining finer details from the encoding path. The Swish activation function provides smoother gradients, and the output is normalized by the standard deviation of the perturbation kernel at time  $t$  to ensure consistent scaling across time steps.

### Model Training

We adopt a *denoising score matching* objective at the 1D map level. For a batch of clean maps  $\{\mathbf{x}\} \in \mathbb{R}^{768}$  and random times  $\{t\}$ , we add Gaussian noise  $\sigma(t) \mathbf{z}$  where  $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ . The network predicts the score  $\mathbf{s}_\theta(\mathbf{x} + \sigma(t) \mathbf{z}, t)$ , and we minimize:

$$\mathcal{L}(\theta) = \mathbb{E}_{\mathbf{x}, t, \mathbf{z}} \|\sigma(t) \mathbf{s}_\theta(\mathbf{x} + \sigma(t) \mathbf{z}, t) + \mathbf{z}\|^2.$$

This forces the model to approximate the gradient of the log-likelihood with respect to the noisy data.

We train for 50 epochs with:

- **Learning Rate:** 0.001
- **Batch Size:** 128

By the end of training, the average score-matching loss stabilizes around 200, indicating the network effectively captures the underlying structure of the B-mode maps.

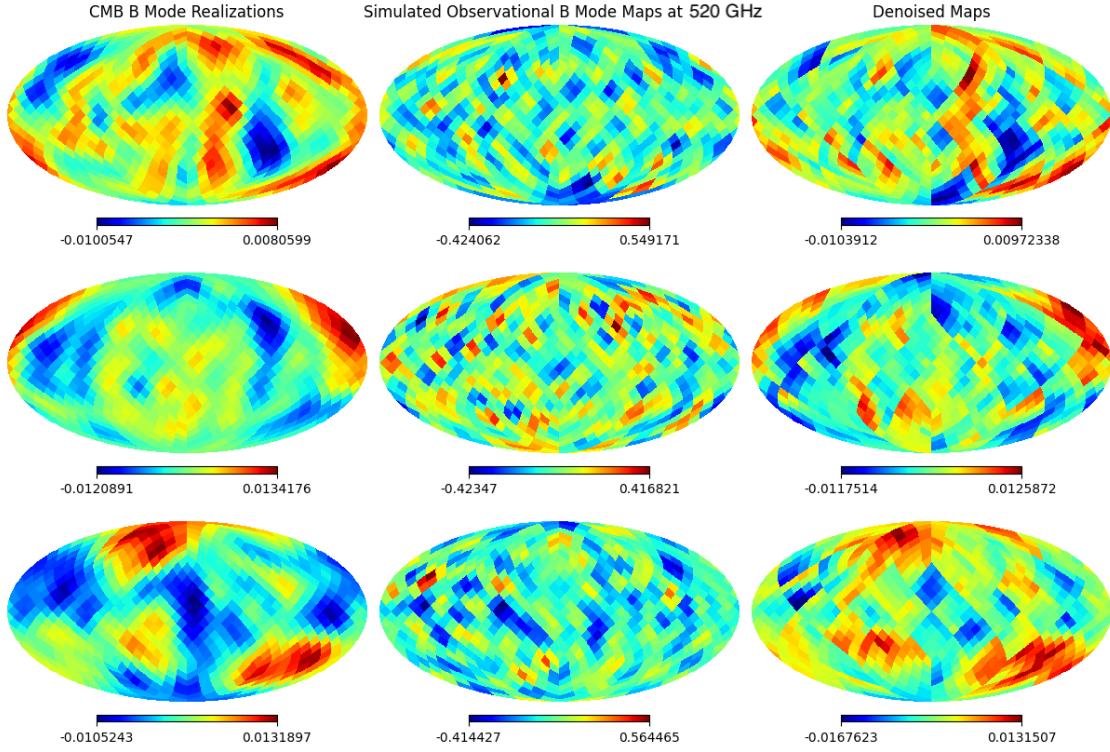
#### 4.1.3 Inference and Results

We compare the ground truth maps (i.e., pure signal without noise), the observed noisy maps, and the denoised maps obtained from reverse SDE using the trained score UNet model.

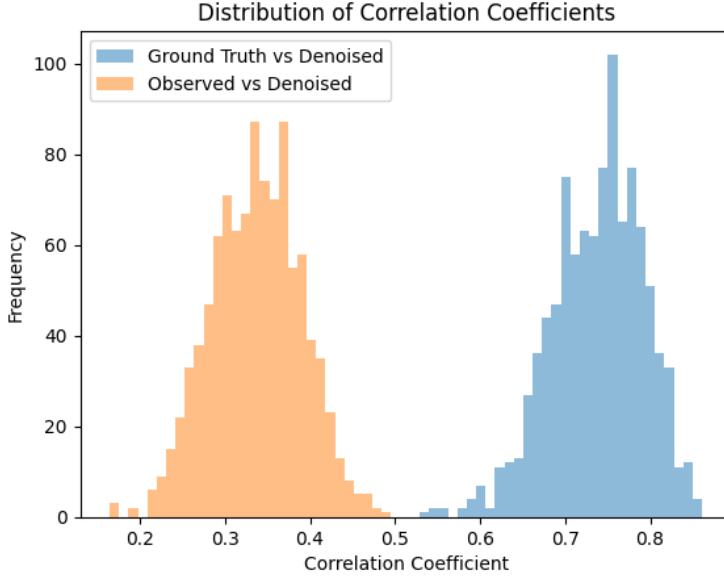
We can see that from figure 4.2, the simulated observational maps exhibit strong noise features that dominate the underlying CMB B mode signal. After denoising, the recovered maps show patterns and the pixel variations that more closely resemble the pure CMB B-mode maps.

We also evaluate the denoised maps by finding the distribution of correlation coefficients between the

1. Denoised maps and the pure CMB B-mode maps.
2. Denoised maps and the simulated observational maps.

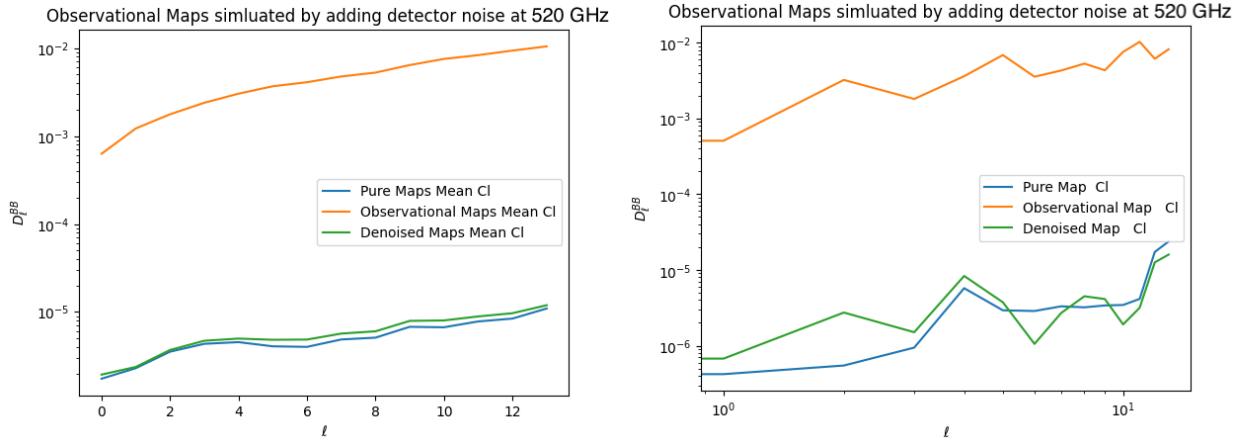


**Fig. 4.2:** CMB B-mode maps (top to bottom, three random realizations). *Left column:* Pure B-mode maps (primordial + lensing). *Middle column:* Simulated observations obtained by adding detector noise (at 520 GHz) to the pure maps. *Right column:* Denoised maps output by the VE-SDE score UNet.



**Fig. 4.3:** Distribution of correlation coefficients computed between the denoised maps and (1) the ground truth (blue), and (2) the noisy observations (orange).

From figure 4.3, we see that the correlation of denoised maps with pure maps peaks around 0.7–0.8, demonstrating that the denoised maps reliably reconstruct the pure B-mode features. Meanwhile, the observed–denoised correlations center around 0.3–0.4, reflecting how substantially the model reduces the noise component.



**Fig. 4.4:** Power spectra of the B-mode maps at 520 GHz. *Left:* The mean  $D_\ell$  over 1,000 simulated realizations. *Right:* Power spectrum for a single realization, showing the same three categories (pure, observed, and denoised).

Furthermore, reconstructing the angular power spectrum  $D_\ell^{BB}$  from the recovered CMB B-mode maps is very important, and comparing it to the true power spectrum and the observational power spectrum provides a quantitative measure of how well the denoising preserves the underlying statistical properties across different angular scales. Figure 4.4 on the left compares the mean power spectra over 1,000 simulated realizations, and on the right presents a single realization.

Although the methodology presented here successfully reconstructs the total B-mode signal (primordial and lensing combined) without considering foreground contamination, evaluating its performance in more realistic scenarios—particularly with foregrounds included—will be explored in the next section.

## 4.2 Reconstruction of Primordial B mode Angular Power spectrum using Denoising and Delensing

In this section, we focus on reconstructing the primordial B-mode angular power spectrum ( $C_\ell^{BB}$ ) in the multipole range  $2 \leq \ell \leq 31$ . We employ the previously discussed VE-SDE score-based framework, but now directly applied at the power-spectrum level, to separate and recover the primordial gravitational-wave signal from realistic observational scenarios.

Our observational simulations include contamination from gravitational lensing, instrumental noise, and astrophysical foregrounds.

### 4.2.1 Data Simulations

For this study, we generate  $\sim 50,000$  realizations of the beam-convolved primordial (tensor-only) B-mode power spectrum  $C_\ell^{BB}$ . These are produced by sampling the theoretical tensor spectrum (with  $r = 0.001$ ) at different random seeds, then applying the instrument beam window function  $b_\ell^2$ . We use this large set to train our VE-SDE-based score model.

**Test and Validation Sets.** To evaluate reconstruction performance, we simulate an additional  $\sim 20,000$  realizations, ensuring that none share seeds with the training set. Each realization consists of:

- **Primordial (Tensor) B-mode**,  $C_\ell^{(\text{tensor})}$ .
- **Total B-mode**,  $C_\ell^{(\text{total})} = C_\ell^{(\text{tensor})} + C_\ell^{(\text{lensing})}$ ,

where  $C_\ell^{(\text{lensing})}$  is computed from CAMB to account for gravitational lensing. The primordial components ( $C_\ell^{(\text{tensor})}$ ) will serve as the ground truth for validation, while the total B-mode spectra ( $C_\ell^{(\text{total})}$ ) form the basis of our realistic observational scenarios.

**Instrumental Noise.** We incorporate instrumental noise power spectra at 28 GHz, representative of ECHO’s sensitivity (see Table 1.2 for noise details), using Algorithm 3. The noise is added to  $C_\ell^{(\text{total})}$  to reflect the measurement uncertainties typically encountered in real observations.

**Foreground Contamination.** To simulate astrophysical foregrounds, we use four distinct PySM3 sky-model combinations as described in [Pal, Yadav, Saha & Souradeep \(2024\)](#):

$$\begin{aligned} \text{F1: } & (a2, s1, d1), & \text{F2: } & (a2, s1, d4), \\ \text{F3: } & (a2, s3, d1), & \text{F4: } & (a2, s3, d4), \end{aligned}$$

where  $a2$ ,  $s1/s3$ , and  $d1/d4$  denote different models for AME, synchrotron, and dust, respectively. For each foreground configuration (F1 … F4), we generate  $\sim 5,000$  realizations, yielding 20,000 total foreground spectra at 28 GHz. These foreground spectra are added to the lensing + noise-contaminated  $C_\ell^{(\text{total})}$  to create highly realistic mock observations.

**Final Observational Spectra.** Each final observed power spectrum is thus:

$$C_\ell^{(\text{obs})} = C_\ell^{(\text{total})} + N_\ell + F_\ell, \quad (4.1)$$

where  $N_\ell$  represents the instrumental noise contribution at 28 GHz, and  $F_\ell$  denotes one of the four possible foreground contaminations at 28 GHz.

These combined contributions create highly realistic mock observational power spectra from which the primordial B-mode signal must be accurately reconstructed.

### 4.2.2 Model Development and Training

To reconstruct the primordial B-mode power spectrum, we adapt the VE-SDE framework from earlier chapters (see Section 2.1.1) for one-dimensional (1D) data corresponding to  $\ell$ -binned power spectra. Our *ScoreNet1D* model predicts the time-dependent score function  $\nabla_x \log p_t(x)$  across various noise levels. Below is a concise outline of the approach:

**Network Architecture.** We design an MLP-based (Rosenblatt, 1958) score estimator, *ScoreNet1D*, which takes as input:

1. A vector  $\mathbf{x} \in \mathbb{R}^{30}$  representing the B-mode power spectrum in the multipole range  $2 \leq \ell \leq 31$ .
2. A scalar time variable  $t$ , embedded via random Fourier features to capture the noise scale in the VE-SDE.

The output is the score estimate  $\nabla_{\mathbf{x}} \log p_t(\mathbf{x})$ . To ensure correct amplitude scaling, we divide the final layer's output by the noise standard deviation at each  $t$ .

**Loss Function.** We employ a denoising score-matching loss that samples random times  $t$  and adds Gaussian noise  $\propto \sigma(t)$  to each batch. Our goal is to minimize:

$$\mathcal{L}(\theta) = \mathbb{E}_{t, \mathbf{x}, \mathbf{z}} \left\| \sigma(t) \mathbf{s}_\theta(\mathbf{x} + \sigma(t) \mathbf{z}, t) + \mathbf{z} \right\|^2,$$

where  $\mathbf{s}_\theta$  is the network output and  $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ . This formulation ensures that the model approximates the score function at each noise scale.

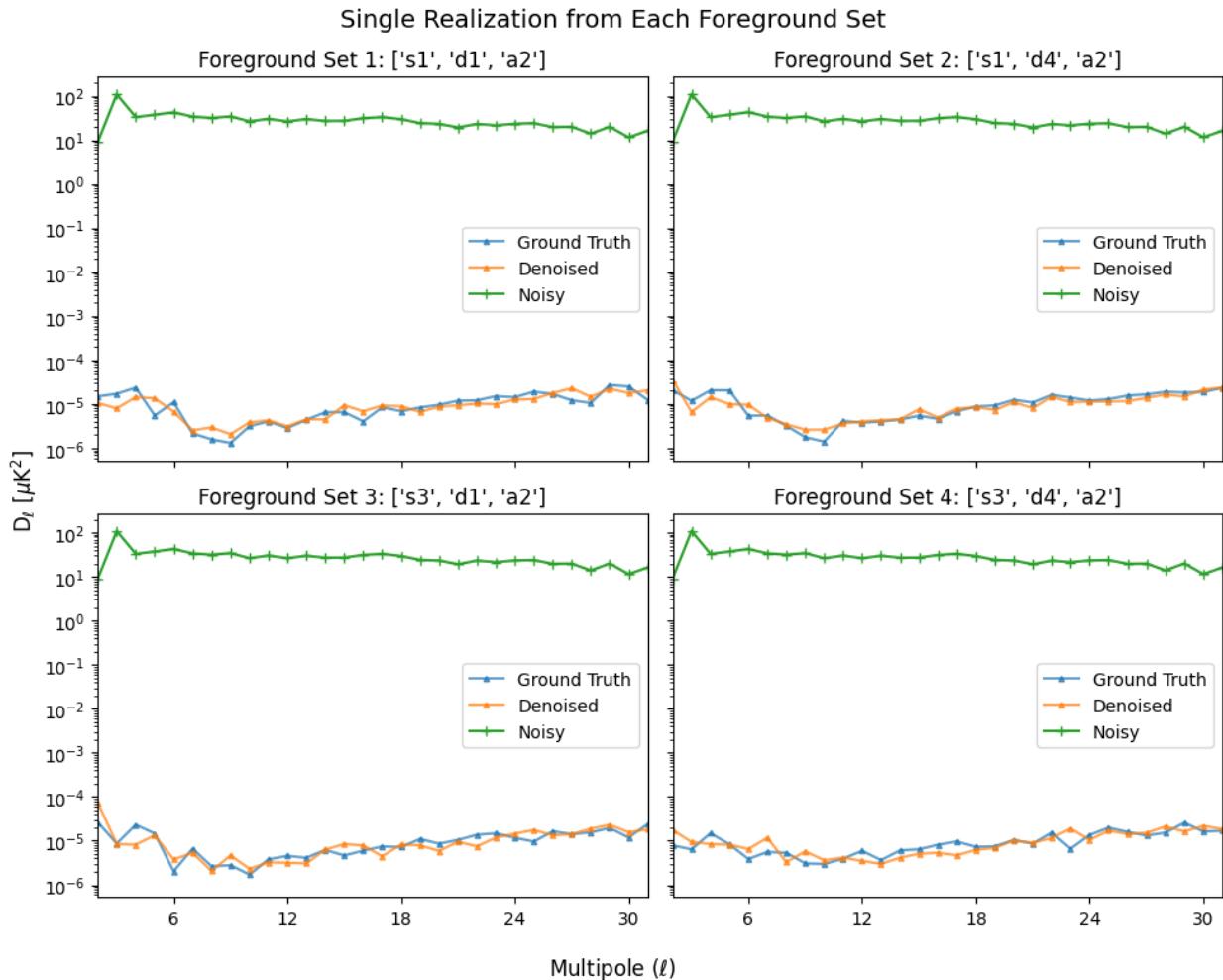
**Training Setup.** We train the model over 150 epochs using the Adam optimizer (Kingma & Ba, 2017) with a learning rate of  $7 \times 10^{-4}$ . Each mini-batch contains random realizations of the primordial B-mode power spectra, and the network updates its parameters to refine

the score estimates at different noise intensities. Convergence is monitored by tracking the average score-matching loss across epochs.

By the end of training, the model effectively learns to denoise power spectra from a wide range of noise amplitudes, providing a foundation for accurately recovering the primordial B-mode signal under realistic observational conditions.

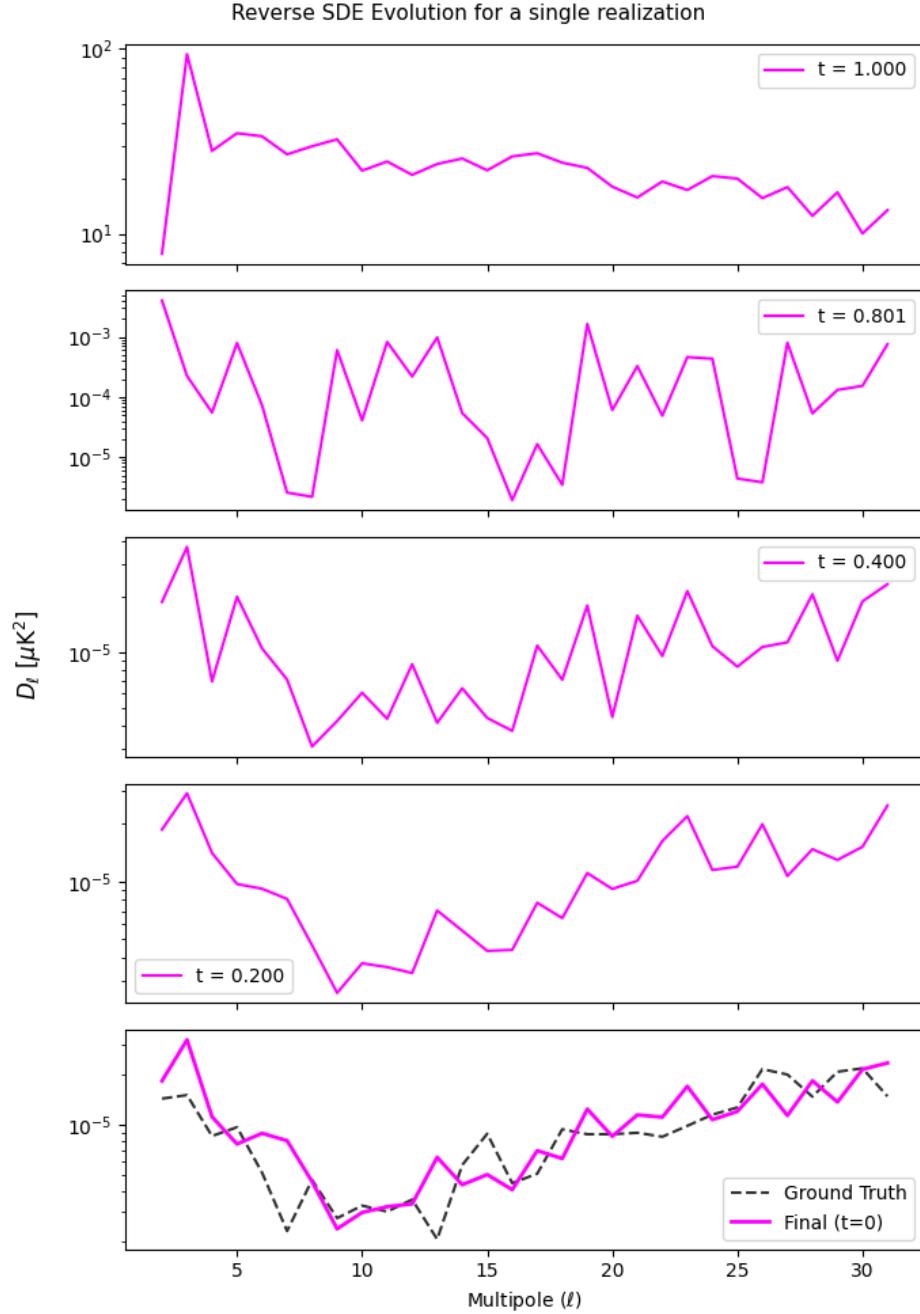
### 4.2.3 Inference and Results

**Inference via Reverse SDE.** After training the ScoreNet1D model, we apply the reverse VE-SDE procedure (§2.1.1) to each observed power spectrum  $C_\ell^{(\text{obs})}$  contaminated by noise, lensing, and foregrounds. Starting from a highly perturbed spectrum at  $t = 1$ , we iteratively

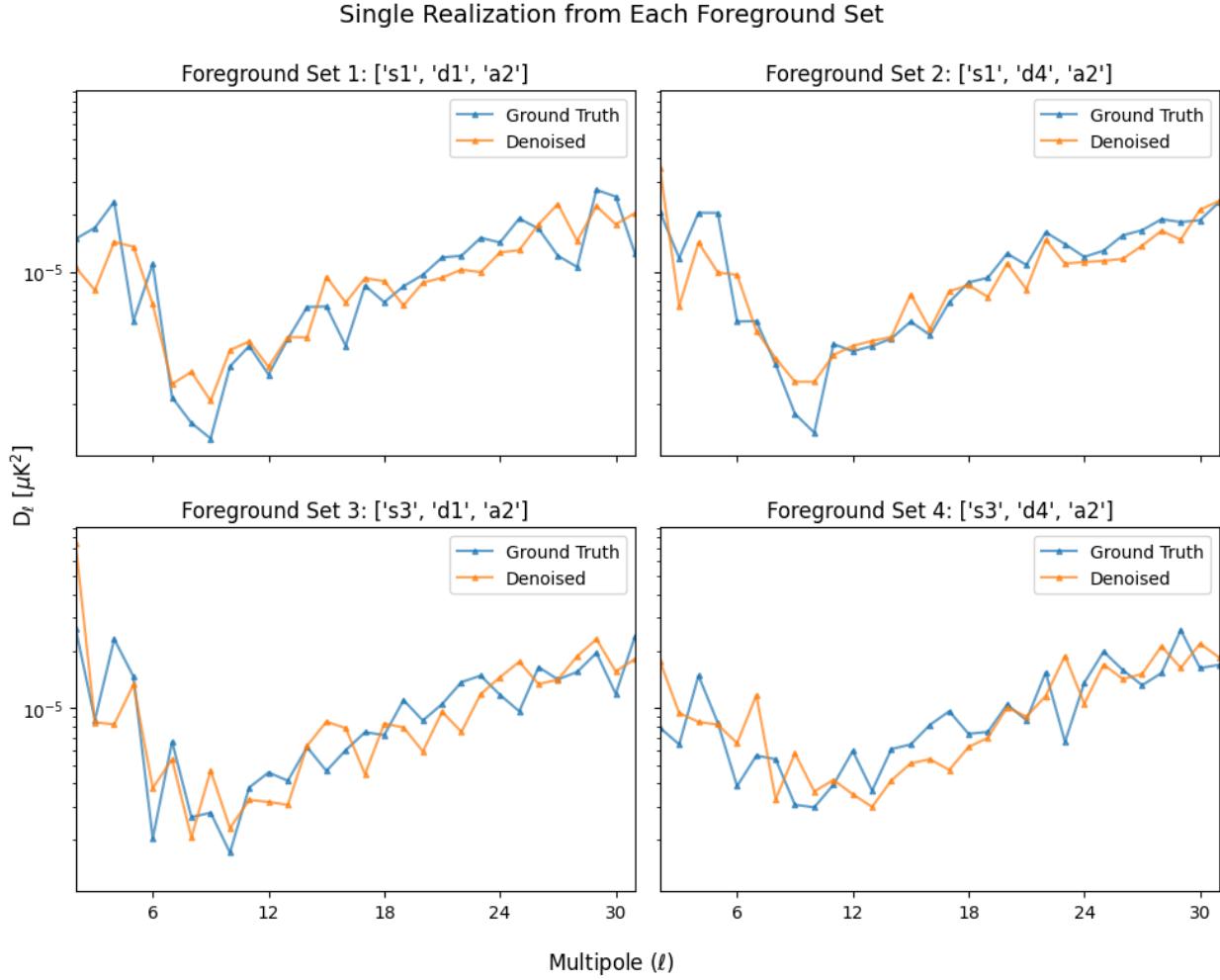


**Fig. 4.5:** Single realization from each of the four foreground sets (F1–F4). The ground-truth primordial B-mode spectrum (blue), the denoised reconstruction (orange) and the observed (noise + lensing + foreground) spectrum (green)

integrate backward in time to  $t = 0$ , guided by the learned score function. This process progressively removes noise and foreground contamination, yielding an estimate of the primordial  $C_\ell^{(\text{tensor})}$  as shown in the Fig 4.6.



**Fig. 4.6:** Evolution of a single realization under the reverse VE-SDE at selected times ( $t=1, 0.801, 0.400, 0.200$ ). Each panel shows the progressively denoised power spectrum, culminating in the final  $t=0$  result (dashed black) compared to the ground truth (dashed gray).



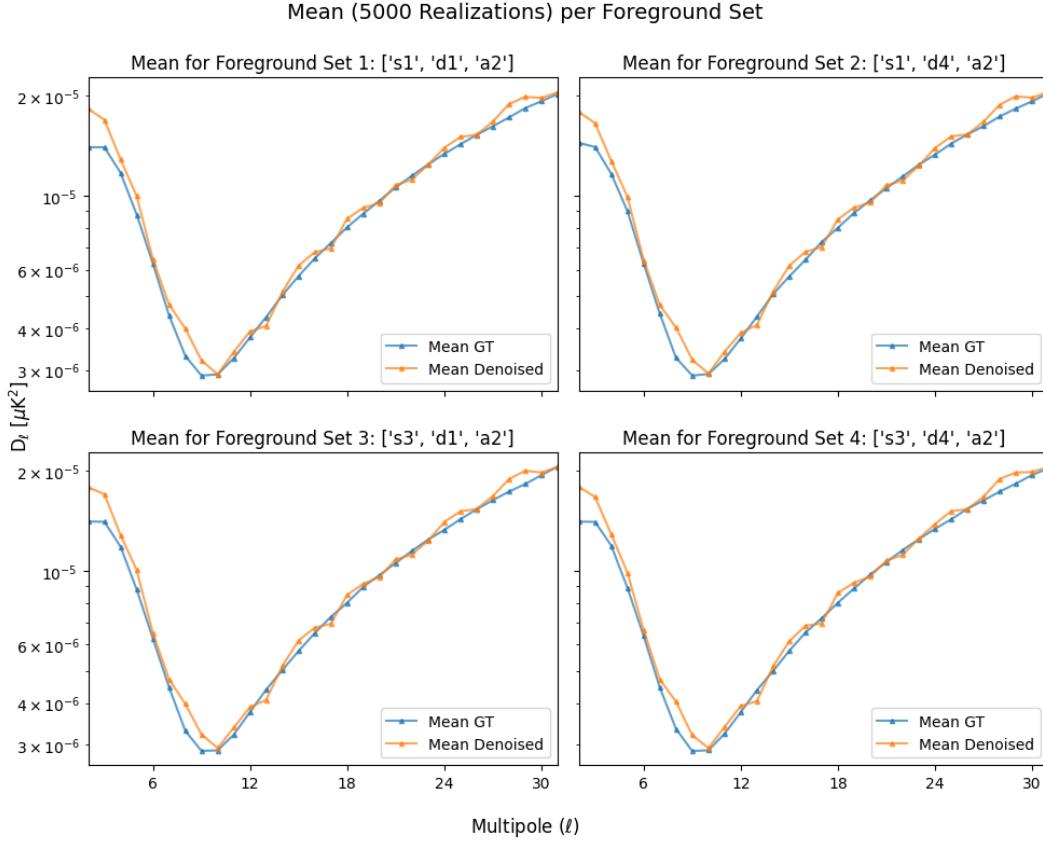
**Fig. 4.7:** Single realization from each of the four foreground sets (F1–F4) comparing the ground-truth (blue) and denoised (orange) primordial B-mode spectra over  $\ell = 2 \dots 31$ .

**Foreground-Specific Comparisons.** Figure 4.5 and Figure 4.7 shows the reconstruction results for each of the four foreground models (F1–F4), comparing:

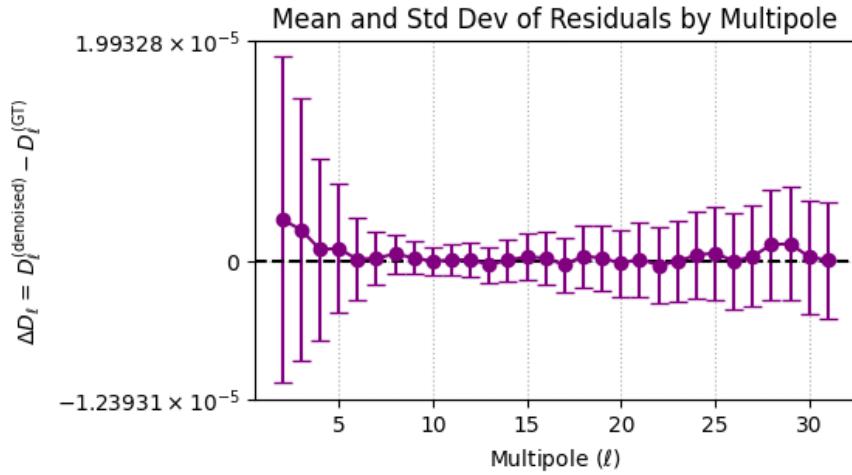
1. The *pure* (ground-truth) primordial B-mode spectrum,
2. The *observed* (noise + lensing + foreground) spectrum,
3. The *reconstructed* spectrum from the reverse SDE.

In each case, the reconstructed curves recover the overall shape and amplitude of the primordial signal across the multipole range  $2 \leq \ell \leq 31$ .

In summary, our VE-SDE-based denoising approach effectively recovers the primordial B-mode power spectrum from realistic observations, even when contaminated by lensing,



**Fig. 4.8:** Mean B-mode power spectra for 5,000 realizations in each foreground set (F1–F4). The ground truth (blue) is closely matched by the denoised solution (orange) across  $\ell = 2 \dots 31$ .



**Fig. 4.9:** Mean residual ( $\Delta D_\ell$ ) and standard deviation between the denoised and ground-truth spectra across all realizations. The horizontal dashed line at zero indicates perfect agreement.

instrumental noise, and multiple foregrounds (F1–F4). As illustrated by Figures 4.5, 4.7, and 4.8, the denoised spectra closely match the ground-truth signal across the multipole range  $2 \leq \ell \leq 31$ . Furthermore, a detailed residual analysis (Figure 4.9) confirms that the denoising procedure remains largely unbiased on average while achieving low scatter in residuals. Overall, these results demonstrate that our score-based framework can reliably isolate primordial B-mode signals and mitigate diverse contamination sources, paving the way for more precise constraints on the tensor-to-scalar ratio  $r$ .

## 5. DISCUSSION AND CONCLUSIONS

In this thesis, we explored the science case for detecting and characterizing primordial B-mode polarization in the Cosmic Microwave Background (CMB)—a key observational window into the physics of inflation. After reviewing the fundamental cosmological framework (including standard Big Bang theory, inflation, and the generation of scalar and tensor perturbations), we discussed how CMB temperature and polarization anisotropies encode information on both early-universe dynamics and late-time effects such as gravitational lensing. We then examined the challenges of measuring the faint primordial B-mode signal, highlighting both astrophysical foregrounds (dust, synchrotron, anomalous microwave emission, etc.) and instrumental systematics (finite beam resolution, detector noise).

Next, we introduced the proposed ECHO (CMB-Bhārat) mission, which aims to achieve unprecedented sensitivity in multiple frequency bands. The mission’s instrumental specifications are designed to isolate a primordial B-mode signal corresponding to a tensor-to-scalar ratio  $r \sim 10^{-3}$  at a  $3\sigma$  level, covering a wide range of inflationary models to test.

On the data-analysis side, we presented modern generative-modeling techniques—specifically, diffusion-based or score-based methods—that can be used for denoising CMB observations. By formulating the problem via a variance-exploding stochastic differential equation (VE-SDE), we demonstrated how to learn the “score” (gradient of the log-density) of CMB maps (or power spectra) and then run a reverse SDE to remove noise and foreground contamination. Our simulations showed that this approach:

- Accurately recovers B-mode maps for high-frequency bands (e.g., 520GHz), significantly reducing detector noise and reconstructing underlying B-mode structures in pixel space.
- Reconstructs the primordial tensor-only  $C_\ell^{BB}$  in the presence of gravitational lensing, foregrounds, and noise. Even with diverse foreground models, our trained score networks successfully recovered the primordial signal with out any prior knowledge about the foregrounds in the training data.

In summary, these results confirm the viability of score-based generative models for future CMB B-mode experiments. Such methods offer a complementary path to traditional foreground-removal pipelines and likelihood-based analyses, particularly in regimes where

the data are high-dimensional, non-Gaussian, or heavily contaminated. As CMB data volume and complexity continue to grow, the development of robust, scalable, and statistically principled tools—like the diffusion-based approach demonstrated here—will be critical for extracting maximal cosmological information, including a definitive measurement of the primordial gravitational-wave background.

In addition, a natural extension of this work involves training the score-based model on a range of tensor-to-scalar ratios  $r$ —for example, spanning from  $r \sim 0$  (almost purely scalar) to an upper bound of  $r = 0.036$  while employing a sufficiently flexible neural network to learn an unbiased score as a function of time over this parameter space. Under such an approach, the same model that denoises CMB maps would also be capable of handling data generated by different inflationary scenarios. In other words, given a map (or power spectrum) of unknown  $r$ , the model could not only remove noise and foregrounds effectively but also provide valuable information about the likely value of  $r$ . Such an outcome significantly enhances the versatility and scientific utility of the score-based framework, especially for upcoming B-mode experiments aimed at probing a wide spectrum of inflationary energy scales.

Furthermore, once the score-based model has been successfully trained across an extensive range of tensor-to-scalar ratios  $r$ , it may also serve as a fast and reliable emulator for generating realistic simulations of CMB B-mode maps or power spectra, effectively replacing or complementing computationally intensive traditional tools such as CAMB. By accurately learning the score over various inflationary scenarios, this generative model would quickly synthesize data corresponding to any desired value of  $r$ , drastically reducing computational resources and accelerating simulation timelines. Such a capability would be particularly beneficial for comprehensive parameter-space explorations, likelihood-free inference methods, or rigorous statistical analyses necessary for next-generation precision CMB polarization studies.

# BIBLIOGRAPHY

- [1] Adak D., Sen A., Basak S., Delabrouille J., Ghosh T., Rotti A., Martínez-Solaech G., Souradeep T., 2022, *Monthly Notices of the Royal Astronomical Society*, 514, 3002–3016
- [2] Ade P. A. R., Aghanim N., Arnaud M., et al. 2016, *Astronomy & Astrophysics*, 594, A13
- [3] Aghanim N., Akrami Y., Ashdown M., et al. 2020, *Astronomy & Astrophysics*, 641, A6
- [4] Anderson B. D. O., 1982, Stochastic Processes and their Applications, 12, 313
- [5] Chen X., Mishra N., Rohaninejad M., Abbeel P., 2018, in Dy J., Krause A., eds, Proceedings of Machine Learning Research Vol. 80, Proceedings of the 35th International Conference on Machine Learning. PMLR, pp 864–872, <https://proceedings.mlr.press/v80/chen18h.html>
- [6] Coles P., Lucchin F., 1995, Cosmology: The Origin and evolution of cosmic structure
- [7] Dodelson S., 2003, Modern Cosmology
- [8] Finkbeiner D. P., Davis M., Schlegel D. J., 1999, , 524, 867
- [9] Flöss T., Coulton W. R., Duivenvoorden A. J., Villaescusa-Navarro F., Wandelt B. D., 2024, Denoising Diffusion Delensing Delight: Reconstructing the Non-Gaussian CMB Lensing Potential with Diffusion Models ([arXiv:2405.05598](https://arxiv.org/abs/2405.05598)), <https://arxiv.org/abs/2405.05598>
- [10] Goodfellow I. J., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y., 2014, Generative Adversarial Networks ([arXiv:1406.2661](https://arxiv.org/abs/1406.2661)), <https://arxiv.org/abs/1406.2661>
- [11] Gorski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, *The Astrophysical Journal*, 622, 759–771
- [12] Guth A. H., 1981, *Phys. Rev. D*, 23, 347

- [13] Guth A. H., Pi S.-Y., 1985, *Phys. Rev. D*, 32, 1899
- [14] Hawking S. W., 1982, *Physics Letters B*, 115, 295
- [15] Heurtel-Depeiges D., Burkhart B., Ohana R., Blancard B. R.-S., 2023, Removing Dust from CMB Observations with Diffusion Models ([arXiv:2310.16285](https://arxiv.org/abs/2310.16285)), <https://arxiv.org/abs/2310.16285>
- [16] Ho J., Jain A., Abbeel P., 2020, CoRR, abs/2006.11239
- [17] Hu W., Dodelson S., 2002, *Annual Review of Astronomy and Astrophysics*, 40, 171–216
- [18] Hu W., White M. J., 1997, *New Astron.*, 2, 323
- [19] Hyvärinen A., 2005, *J. Mach. Learn. Res.*, 6, 695
- [20] Kamionkowski M., Kovetz E. D., 2016, *Annual Review of Astronomy and Astrophysics*, 54, 227–269
- [21] Kansal R., Li A., Duarte J., Chernyavskaya N., Pierini M., Orzari B., Tomei T., 2023, *Physical Review D*, 107
- [22] Kingma D. P., Ba J., 2017, Adam: A Method for Stochastic Optimization ([arXiv:1412.6980](https://arxiv.org/abs/1412.6980)), <https://arxiv.org/abs/1412.6980>
- [23] Kingma D. P., Welling M., 2019, *Foundations and Trends® in Machine Learning*, 12, 307–392
- [24] Kogut A., 2012, *The Astrophysical Journal*, 753, 110
- [25] Krachmalnicoff N., Tomasi M., 2019, *Astronomy & Astrophysics*, 628, A129
- [26] LEWIS A., CHALLINOR A., 2006, *Physics Reports*, 429, 1–65
- [27] Lemaître G., 1931, , 91, 490
- [28] Lewis A., Challinor A., Lasenby A., 2000, *The Astrophysical Journal*, 538, 473–476
- [29] Linde A. D., 1982, *Phys. Lett. B*, 108, 389
- [30] Narlikar J. V., 2002, Introduction to Cosmology (3rd ed.). Cambridge University Press, Cambridge

- [31] Pal S., Yadav S. K., Saha R., Souradeep T., 2024, Accurate and Unbiased Reconstruction of CMB B Mode using Deep Learning ([arXiv:2404.18100](https://arxiv.org/abs/2404.18100)), <https://arxiv.org/abs/2404.18100>
- [32] Penzias A. A., Wilson R. W., 1965, , [142, 419](#)
- [33] Perraудин N., Defferrard M., Kacprzak T., Sgier R., 2019, *Astronomy and Computing*, 27, 130–146
- [34] Rezende D. J., Mohamed S., 2016, Variational Inference with Normalizing Flows ([arXiv:1505.05770](https://arxiv.org/abs/1505.05770)), <https://arxiv.org/abs/1505.05770>
- [35] Rombach R., Blattmann A., Lorenz D., Esser P., Ommer B., 2022, High-Resolution Image Synthesis with Latent Diffusion Models ([arXiv:2112.10752](https://arxiv.org/abs/2112.10752)), <https://arxiv.org/abs/2112.10752>
- [36] Ronneberger O., Fischer P., Brox T., 2015, U-Net: Convolutional Networks for Biomedical Image Segmentation ([arXiv:1505.04597](https://arxiv.org/abs/1505.04597)), <https://arxiv.org/abs/1505.04597>
- [37] Rosenblatt F., 1958, Psychological review, 65, 386
- [38] Song Y., Ermon S., 2020, Generative Modeling by Estimating Gradients of the Data Distribution ([arXiv:1907.05600](https://arxiv.org/abs/1907.05600)), <https://arxiv.org/abs/1907.05600>
- [39] Song Y., Sohl-Dickstein J., Kingma D. P., Kumar A., Ermon S., Poole B., 2020, CoRR, abs/2011.13456
- [40] Ullmo M., Decelle A., Aghanim N., 2021, , [651, A46](#)
- [41] Wang M. S., 2017. <https://api.semanticscholar.org/CorpusID:54028419>
- [42] Zhao X., Ting Y.-S., Diao K., Mao Y., 2023, Can Diffusion Model Conditionally Generate Astrophysical Images? ([arXiv:2307.09568](https://arxiv.org/abs/2307.09568)), <https://arxiv.org/abs/2307.09568>
- [43] Zonca A., Thorne B., Krachmalnicoff N., Borrill J., 2021, *Journal of Open Source Software*, 6, 3783