Studying tomographic cross-correlations between CMB gravitational lensing potential and galaxy surveys

Author: mgr. Chandra Shekhar Saraf

Advisors: dr hab. Paweł Bielewicz and dr hab. Michał Chodorowski

CENTRUM ASTRONOMICZNE IM. MIKOŁAJA KOPERNIKA
POLSKIEJ AKADEMII NAUK

A thesis submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in
Astronomy and Astrophysics

August, 2023

Dedicated to
Papa & Maa

Abstract

Gravity has captured the curiosity of mankind for centuries. It is one of the four known fundamental forces of nature and the dominant cause for apples falling to the ground, Earth going around the Sun, and the formation of the large scale structures in the Universe. A revolution in our understanding of gravity was brought by Isaac Newton who explained how planets move around the Sun. Another milestone was placed by Albert Einstein who gave us the General Theory of Relativity (GR). It expresses gravity as a fundamental property of the spacetime fabric of the cosmos. The natural application of GR was to construct a model of the entire Universe to understand its evolution. The most accepted model of the Universe backed by observational confirmations is the Lambda Cold Dark Matter or the ΛCDM model. It explains the expanding Universe as well as the formation and evolution of the large scale structure in the Universe. Although the ΛCDM model is consistent with observations, it is not free from challenges and disparities. Explanation of the nature of dark energy and dark matter, accelerated expansion of the Universe, and tensions in estimations of some cosmological parameters (like σ_8 or Hubble tension) have raised questions on the validity of ΛCDM model and a need of its thorough testing on cosmological scales.

One of the many different ways to test cosmological models is to use the Cosmic Microwave Background (CMB) radiation in synergy with probes of the large scale structure like galaxies, galaxy clusters, quasars, etc. In this thesis, we focus on cross-correlation between CMB gravitational lensing potential, estimated from CMB anisotropy maps, and photometric galaxy surveys. The surveys are prone to several systematic errors which can alter the observed redshift distribution of sources and can cause unphysical variations in their number density in the sky. In this thesis, *we study the impact of various systematic effects on CMB gravitational lensing and photometric galaxy surveys cross-correlation measurements and estimation from them cosmological parameters, in particular galaxy bias, amplitude of cross-correlation or* σ_8 *parameter.*

The photometric redshifts of galaxies are accompanied by errors that generally broadens the shape and changes the median redshift of the galaxy distribution. The redshifts of galaxies are misestimated due to catastrophic errors, which changes the photometric redshift distribution of galaxies. On the other hand systematics like photometric calibration errors, which arise due to fluctuations in the limiting magnitude of surveys, lead to unphysical variations in the number density of galaxies over the survey area. In Chapter 3 we study the impact of these systematic errors on estimation of cross-correlation between CMB lensing potential measured by the *Planck* satellite and photometric galaxy catalogues from the *Herschel* Extragalactic Legacy Project.

Future galaxy surveys will have larger area coverage in the sky and increase in the magnitude depth, thus observing significantly larger number of galaxies than their predecessors. These features of the upcoming surveys will enable us to perform cross-correlation analyses with galaxies divided into redshift bins. These *tomographic* cross-correlation measurements allow us to map the evolution of cosmological parameters and test the validity of cosmological model at different

redshifts. However, tomographic measurements suffer from an additional systematic arising from the redshift bin mismatch of objects due to photometric redshift errors. This issue is addressed in Chapter 4. Using a suite of Monte Carlo simulations of the Rubin Observatory Legacy Survey of Space and Time (LSST) and the *Planck* lensing map we thoroughly study how the scatter of objects between redshift bins biases the inferences based on tomographic cross-correlation analysis and show how to avoid these biases using scattering matrix formalism. We propose and test new, fast method of estimation the scattering matrix which is well-suited for the analysis of upcoming large galaxy surveys. We also demonstrate in Chapter 5 the application of scattering matrix to tomographic analysis of cross-correlation between galaxy catalogue from the Dark Energy Spectroscopic Instrument Legacy Imaging Survey and the *Planck* gravitational lensing map.

The collection of works in this thesis have shown (i) the mitigation strategies of different systematics on cross-correlation measurements of CMB lensing with photometric redshift galaxy catalogues, (ii) that in the case of LSST survey biases in tomographic analysis due to redshift bin mismatch of objects are of order of one standard deviation for the amplitude of cross-correlation and σ_8 parameter, (iii) and that to avoid biases in tomographic analysis one needs to use the scattering matrix approach which properly takes into account redshift bin mismatch of objects.

Streszczenie

(Abstract in Polish)

Grawitacja przyciągała ciekawość ludzkości od wieków. Jest ona jedną z czterech znanych podstawowych sił natury i główną przyczyną powodującą spadanie jabłek, krążenie Ziemi wokół Słońca i formowanie się wielkoskalowej struktury we Wszechświecie. Izaak Newton dokonał rewolucji w naszym rozumieniu grawitacji, co umożliwiło wyjaśnienie ruchu planet wokół Słońca. Kolejny kamień milowy został postawiony przez Alberta Einsteina, który przedstawił Ogólną Teorię Względności (OTW). Wyraża ona grawitację jako podstawową właściwość czasoprzestrzennej struktury kosmosu. Naturalnym zastosowaniem OTW było zbudowanie modelu całego Wszechświata umożliwiającego zrozumienie jego ewolucji. Najczęściej akceptowany model Wszechświata potwierdzony obserwacjami to model ΛCDM (zimnej ciemnej materii ze stałą kosmologiczną, ang. Lambda Cold Dark Matter). Wyjaśnia on rozszerzający się Wszechświat jak również formowanie się i ewolucję wielkoskalowej struktury Wszechświata. Chociaż model ΛCDM jest spójny z obserwacjami, nie jest on wolny od pewnych wyzwań i rozbieżności. Potrzeba wyjaśnienia natury ciemnej energii i ciemnej materii, przyspieszająca ekspansją Wszechświata i rozbieżności w oszacowaniach niektórych parametrów kosmologicznych (jak np. rozbieżności parametrów σ_8 czy stałej Hubble'a) stawiają pytania dotyczące poprawności modelu ΛCDM i wskazują na potrzebę dokładnego przetestowania modelu w skalach kosmologicznych.

Jednym z wielu różnych sposobów testowania modelu kosmologicznego jest wykorzystanie promieniowania mikrofalowego tła w połączeniu z obiektami próbkującymi wielkoskalową strukturę takimi jak galaktyki, gromady galaktyk, kwazary itp. W tej rozprawie skupiam się na badaniu korelacji wzajemnej między potencjałem soczewkującym promieniowanie tła, oszacowanym z map anizotropii promieniowania, a fotometrycznymi przeglądami galaktyk. Przeglądy galaktyk są narażone na kilka błędów systematycznych, które mogą zmienić obserwowany rozkład źródeł ze zmierzonymi przesunięciami ku czerwieni i spowodować niefizyczną zmienność gęstości rozkładu galaktyk na niebie. W rozprawie badam wpływ różnych efektów systematycznych na pomiary korelacji wzajemnej między soczewkowaniem grawitacyjnym promieniowania tła a fotometrycznymi przeglądami galaktyk i na oszacowane z pomiarów parametry kosmologiczne, w szczególności obciążenie rozkładu galaktyk względem ciemnej materii (ang. galaxy bias), amplitudę korelacji wzajemnej lub parametr σ_8.

Fotometryczne przesunięcia ku czerwieni galaktyk są obarczone błędami, które w ogólności poszerzają kształt i zmieniają medianę przesunięć ku czerwieni rozkładu galaktyk. Przesunięcia ku czerwieni galaktyk są również błędnie oszacowane z powodu tzw. katastrofalnych błędów, które zmieniają rozkład galaktyk z fotometrycznymi przesunięciami ku czerwieni. Z drugiej strony, błędy systematyczne takie jak fotometryczne błędy kalibracji, które powstają wskutek fluktuacji limitu jasności pozornej przeglądów, prowadzą do niefizycznej zmienności gęstości galaktyk w obszarze przeglądu. W rozdziale 3 badam wpływ tych błędów systematycznych na oszacowanie korelacji wzajemnej między potencjałem soczewkującym promieniowanie tła zmierzone przez satelitę *Planck* a fotometrycznymi katalogami galaktyk z *Herschel* Extragalactic Legacy Project.

Przyszłe przeglądy galaktyk będą pokrywać większe obszary nieba i będą głębsze w jasności pozornej, zatem będą również obserwować większą liczbę galaktyk niż poprzednie przeglądy. Te własności nadchodzących przeglądów umożliwią nam analizę korelacji wzajemnej dla galaktyk w określonych przedziałach przesunięć ku czerwieni. Takie tomograficzne pomiary korelacji wzajemnej pozwolą nam odwzorować ewolucję parametrów kosmologicznych i przetestować poprawność modelu kosmologicznego przy różnych przesunięciach ku czerwnieni. Jednakże, tomograficzne pomiary są obarczone dodatkowym błędem systematycznym powstającym wskutek błędnego przyporządkowania źródeł do przedziałów przesunięć ku czerwieni z powodu błędów fotometrycznie wyznaczonych przesunięć ku czerwieni. Problem ten jest analizowany w rozdziale 4. Z pomocą pakietu symulacji Monte Carlo przeglądu galaktyk Rubin Observatory Legacy Survey of Space and Time (LSST) i map potencjału soczewkującego dla danych z satelity *Planck* badam w nim dokładnie jak rozproszenie obiektów między przedziałami przesunięć ku czerwieni wpływa na wnioskowanie oparte na analizie tomograficznej korelacji wzajemnej i przedstawiam sposób na uniknięcie tego wpływu za pomocą formalizmu macierzy rozpraszania. Proponuję tam i przedstawiam testy nowej i szybkiej metody oszacowania macierzy rozpraszania, która jest szczególnie dobrze dostosowana do analizy nadchodzących dużych przeglądów galaktyk. W rozdziale 5 przedstawiam również zastosowanie macierzy rozpraszania do tomograficznej analizy korelacji wzajemnej katalogu galaktyk z Dark Energy Spectroscopic Instrument Legacy Imaging Survey i map potencjału soczewkowkującego dla danych z satelity *Planck*.

Zbiór badań przedstawionych w tej rozprawie pokazał (i) sposoby na zmniejszenie wpływu różnych błędów systematycznych na pomiary korelacji wzajemnej między potencjałem soczewkującym promieniowanie mikrofalowe tła a katalogami galaktyk z fotometrycznie zmierzonymi przesunięciami ku czerwieni, (ii) że w przypadku przeglądu LSST obciążenia w analizie tomograficznej, z powodu błędnego przyporządkowania obiektów do przedziałów przesunięć ku czerwieni, są rzędu wielkości jednego standardowego odchylenia dla amplitudy korelacji wzajemnej i parametru σ_8, (iii) że aby uniknąć obciążenia w analizie tomograficznej konieczne jest użycie zaproponowanej metody wykorzystującej macierz rozpraszania, która we właściwy sposób bierze pod uwagę błędne przyporządkowanie obiektów do przedziałów przesunięć ku czerwieni.

Declaration

The work described in this thesis was undertaken between October 2018 and June 2023 while the author was a doctoral student at the Nicolaus Copernicus Astronomical Center, Polish Academy of Sciences (Warsaw, Poland) under the supervision of dr hab. Paweł Bielewicz from the National Centre for Nuclear Research (Warsaw, Poland) dr hab. Michał Chodorowski from Nicolaus Copernicus Astronomical Center, Polish Academy of Sciences. The author has completed his coursework at the Nicolaus Copernicus Astronomical Center, Polish Academy of Sciences. No part of this thesis has been submitted for any other degree at any other university or research institute.

The contents of Chapter 3 of this thesis have appeared in the following paper:

- **Chandra Shekhar Saraf**, P. Bielewicz, M. Chodorowski, 2022, MNRAS, 515, 1993: **Cross-correlation between *Planck* CMB lensing potential and galaxy catalogues from HELP**

Acknowledgements

During my doctoral studies, I have been blessed to receive a great deal of support and assistance from many people. It will be practically impossible to mention each and every person in this acknowledgement, but they all have helped me grow both on a personal and professional level.

I would first like to thank my thesis supervisors, dr hab. Paweł Bielewicz and dr hab. Michał Chodorowski, for their constant support and guidance. This thesis would not have materialised without their counselling and warmth. I thank Paweł for believing in me and help me create independent temperament towards research, and I thank Michał for numerous discussions on science, society and philosophy. I would also like to extend my gratitude to Professor Agnieszka Pollo for being a warm, guiding light throughout my PhD.

I am deeply indebted to my strongest supporters throughout this journey - my *Maa* (Renu Saraf), *Papa* (Anil Kumar Saraf). I owe all my achievements to their uncountable sacrifices. I will also remain in forever debt of my *brother* (Ayush Saraf), for he took care of the family all those times when I was far from home.

I cherish my journey during the PhD that has allowed me to meet, learn and befriend such an amazing group of people, each one talented in their own right. In particular, I'd like to thank dr. Deepika Ananda Bollimpalli and dr. Swayamtrupta Panda for their love, support and guidance during my stay in Poland. I thank Ankan Sur, Rajeev Singh Rathour and Ruchi Mishra for being my support system at the institute, specially during the long state of pandemic. I also thank Abhipsa Panda, Marzena Śniegowska, Suhani Gupta, Abinash Adhikary, Ayush Moharana, Tilak Pawar, Sergen Odzemir, Lorenzo Gavassino, John Martinez, Tathagata Saha, Mohammad Naddaf, and many others who made me feel at home with their cheerfulness and positive attitude towards life. I specially thank Aleksandra Olejak for being the best office mate I could have hoped for. I also thank my flatmates during the course of my stay in Poland Samaresh Mondal, Jose Ortuno Macias, Angelos Karakonstantakis and Biswaraj Palit.

I would like to thank my 'old' friends Vishnu Lohar, Himadri Saini, Puneet Mewara, Kajal Thakar, Vismay Mori, Bhakti Chitroda, Avik Dasgupta, and Binal Patel who were an invisible support for me, despite the vast physical distance. They have been a second family to me, whom I can always rely on.

Finally, I would like to acknowledge the Nicolaus Copernicus Astronomical Center, the professors, the staff for their support and assistance during this journey. I also express my gratitude to the owners of an English book store '*thebooks.pl*' in Warsaw, which helped me keep my reading habit alive.

Contents

Chapter 1

Introduction

1.1 The standard model of cosmology

The standard model of cosmology or the ΛCDM model provides us with a comprehensive framework to understand the structure, origin, and evolution of the Universe. It is a theoretical framework that combines various branches of physics, such as General Relativity and particle physics, to describe the fundamental principles governing the Universe on its largest scales. By studying the properties and behavior of matter, energy, and the fabric of spacetime, the standard model of cosmology aims to explain the vast array of phenomena observed in our Universe. At its core, the ΛCDM model is based on the theory of General Relativity, which was developed by Albert Einstein in the early 20^{th} century. General Relativity describes gravity as the curvature of spacetime caused by the presence of matter and energy. This theory allows us to understand the behavior of objects on both small and large scales, from the motion of planets in our solar system to the expansion of the entire Universe.

In the early 20^{th} century, Edwin Hubble showed that the Universe is not static but actually expanding (Hubble, 1929). This realization paved the way for the development of the Big Bang theory, which forms the foundation of the standard model of cosmology. According to the Big Bang theory, the Universe originated from an incredibly hot and dense state approximately 13.8 billion years ago. It has been expanding ever since, gradually cooling down and giving rise to the diverse structures we observe today. The standard model of cosmology provides a framework for understanding the formation of large scale structure, the distribution of matter and energy, and the overall evolution of the Universe over billions of years. The current state of the expanding Universe is dominated by a yet-unknown 'dark energy' with roughly 68% contribution to the Universe. There is an invisible form of matter called 'dark matter' which makes up about 27% of the observable Universe and rest is in account of the 'visible' baryonic matter. The ΛCDM model accounts for all these properties of the Universe and is very-well established owing to a number of observations taken over past few decades.

The ΛCDM model is built upon a vast array of observations collected by analyzing the emissions from distant objects in the Universe covering the entire electromagnetic spectrum, detecting particles from various sources and gravitational waved reaching us from distant ob-

jects. We list below some of the key evidences and observations that have contributed to the development of the ΛCDM model.

Cosmic Microwave Background (CMB): The discovery of the Cosmic Microwave Background radiation in 1965 (Penzias & Wilson, 1965) stands as one of the most significant achievements in modern cosmology. This faint radiation, permeating the entire Universe, is the remnants of the intense heat and radiation that filled the early Universe, about $380,000$ years after the Big Bang. Precise measurements of the CMB by the Wilkinson Microwave Anisotropy Probe (WMAP; Bennett et al., 2003) and the *Planck* satellite (Planck Collaboration et al., 2016a) have provided critical insights into the age, composition, and evolution of the Universe, confirming the predictions of the Big Bang theory.

Redshift and the Expanding Universe: In the year 1929, astronomer Edwin Hubble made a groundbreaking observation that transformed our understanding of the cosmos. By studying the light emitted by distant galaxies, he noticed a systematic shift towards longer wavelengths, known as redshift. Hubble interpreted this as evidence that the Universe is expanding, with distant galaxies moving away from us (Hubble, 1929). This observation paved the way for the Big Bang theory and became a cornerstone of our current understandings of cosmology.

Large-Scale Structure of the Universe: By mapping the distribution of galaxies on large scales, astronomers have revealed the intricate cosmic web of filaments, clusters, and voids that make up the large-scale structure of the Universe. These observations provide insights into the distribution of matter and energy and have helped to constrain models of cosmic evolution, including the growth of structures through gravitational interactions. For a theoretical introduction of the large scale structure one may refer to Peebles, 1980, and to Coil, 2013; Bernardeau et al., 2002 for current updates on the topic.

Supernovae and Dark Energy: The study of supernovae, the explosive deaths of massive stars, has provided critical evidence for the existence of dark energy. In the late 1990s, observations of Type Ia supernovae (Riess et al. 1998; Perlmutter et al. 1999) revealed that the Universe's expansion is accelerating rather than slowing down, as previously believed. This unexpected finding led to the proposition of dark energy, a mysterious form of energy that permeates the Universe and is driving its accelerated expansion. The nature of dark energy remains one of the most intriguing and challenging puzzles in modern cosmology.

Cosmic Inflation: Cosmic inflation is a theory that explains the uniformity and flatness of the Universe on large scales (Linde 1982; Guth 1981; Starobinsky 1979). This theory suggests that the Universe underwent a period of exponential expansion in its earliest moments, smoothing out irregularities and setting the stage for the structures we observe today. The prediction of cosmic inflation aligns with several observations, such as the uniformity of the CMB and the large-scale structure of the Universe.

Dark Matter and Gravitational Lensing: The existence of an invisible form of matter that does not emit or interact with light, except gravitationally, was inferred through various obser-

vations. This 'dark' matter has only been confirmed through the gravitational effects on the motion of stars and galaxies (Rubin et al. 1978; Zwicky 1937; Zwicky 1933), as well as its influence on the bending of light in a phenomenon known as gravitational lensing (Natarajan et al. 2017; Tyson et al. 1998). These observations provide compelling evidence for the presence of dark matter and have significantly influenced the standard model of cosmology.

We have given a brief account of the most important milestones that have shaped our understanding of the Universe and resulted in the most accepted model of our observable Universe, the ΛCDM model. We now look with some detail into the Cosmic Microwave Background radiation and how it enables us to test the cosmological models when combined with the large scale structure.

1.2 Cosmic Microwave Background

The Cosmic Microwave Background (CMB) radiation stands as one of the most remarkable discoveries in the field of cosmology, providing a window into the early Universe and offering crucial evidence for the Big Bang theory. This faint, pervasive glow of electromagnetic radiation fills the cosmos, bathing the entire observable Universe in a nearly uniform sea of microwave photons. The existence of the Cosmic Microwave Background was theoretically predicted as early as the late 1940s by George Gamow, Ralph Alpher, and Robert Herman (Gamow 1948b; Gamow 1948a; Alpher and Herman 1948a; Alpher and Herman 1948b). In the early Universe, around 13.8 billion years ago, the Universe originated in a hot and dense state, popularly known as the Big Bang. The Universe was filled with fundamental particles including quarks, leptons and bosons, forming a "primordial soup" of particles (Weinberg, 1993). As the Universe expanded and cooled, electrons combined with protons to form neutral hydrogen atoms and the photons gradually lost energy. This process, known as recombination, occurred about $380,000$ years after the Big Bang. At the time of recombination, the Universe had cooled enough for photons to "decouple" from matter. Before recombination, photons were continuously scattered by free electrons, preventing them from traveling freely. However, as electrons combined with protons to form neutral atoms, the Universe became mostly transparent to photons, allowing them to travel freely without scattering. The CMB radiation we observe today is essentially the "last scattering surface" of these primordial photons, carrying crucial information about the conditions at that time (Dodelson, 2003). After nearly 13.8 billion years of travelling through an expanding Universe, the CMB has a significantly redshifted thermal blackbody spectrum at a temperature of 2.73 Kelvin (Fixsen et al. 1996; Fixsen 2009) corresponding to microwave frequencies with anisotropies in the sky at $\mathcal{O}(10^{-5})$ level (Smoot et al. 1992a). The statistics of the CMB temperature anisotropies can be understood as arising from the acoustic oscillations in the primordial soup of photons and baryons, in which the initial perturbations were purely adiabatic, Gaussian, and had a nearly scale invariant power spectrum (Padmanabhan 2006; Spergel et al. 2003; Hu and Dodelson 2002).

The epoch of recombination at which the Universe became neutral, was not instantaneous and there is a finite non-zero width associated with the surface of last scattering. Concerning present observations, this surface is very thin compare to the distance to the last scattering

surface. The process of recombination, however, was slow enough that a photon quadrupole could develop. Thomson scattering of photons with quadrupole distribution will give rise to polarization signal from the last scattering surface (Hu and White 1997; Polnarev 1985; Rees 1968).

The CMB was first detected by two radio astronomers, Arno Penzias and Robert Wilson in 1965 (Penzias & Wilson, 1965) at Bell Laboratories in New Jersey while trying to detect radio signals from space. They identified a persistent background noise that seemed to be coming from all directions, independent of any known celestial sources. Their discovery garnered significant attention and won them the Nobel Prize in Physics in 1978. Since its initial discovery, a myriad of ground-based and space-based instruments have been employed to study the Cosmic Microwave Background in exquisite detail. These observations have provided exceptional insights into the composition, geometry, and age of the Universe. The Cosmic Background Explorer (COBE) satellite (Smoot et al., 1992b) launched in 1989, precisely measured the CMB spectrum, revealing tiny temperature fluctuations that serve as the seeds of the cosmic structure we observe today. Subsequent missions like the Wilkinson Microwave Anisotropy Probe (WMAP; Bennett et al., 2003) and the *Planck* satellite (Planck Collaboration et al., 2016a) further refined these measurements, providing high-resolution maps of the CMB temperature fluctuations and polarization patterns. The CMB has since become an indispensable piece of evidence supporting our understanding of the origins and evolution of the Universe.

The precise measurements of the CMB radiations have shaped our current understanding of the evolution of the Universe. The discovery of the CMB radiation is considered a smoking gun of the Big Bang theory. The CMB anisotropies detected by WMAP and *Planck* satellite missions are crucial in understanding the distribution of matter and energy in the early Universe. They serve as a blueprint for the seeds of cosmic structures, such as galaxies and galaxy clusters, that formed later due to gravitational collapse. The variations in temperature provide insight into the initial conditions that led to the current large-scale structure in the Universe.

The CMB power spectrum, which quantifies the distribution of temperature fluctuations at different angular scales, is a valuable tool for determining cosmological parameters. By analyzing the CMB data values of various parameters such as the total density of matter and energy in the Universe, the current rate of expansion of the Universe (quantified through the Hubble constant), the age of the Universe, and the amount of dark matter and dark energy can be derived with remarkable precision. The matter content of the Universe can be deduced from the acoustic peaks in the CMB power spectrum; the ratio of the second peak's amplitude to the first defines the percentage of baryonic matter, while the third peak determines the dark matter. One of the most significant findings from CMB observations is the measurement of the average density of the Universe. Combined with other cosmological data, such as the large-scale distribution of galaxies, these measurements suggest that the Universe is flat or very close to flat, implying that it contains precisely the critical density required for it to remain nearly flat over cosmic time.

The uniformity and isotropy of the CMB was a puzzling observation until the inflation theory provided an explanation that the Universe underwent a rapid expansion phase after the Big Bang (Linde 1982; Guth 1981; Starobinsky 1979). This inflationary period would have left spe-

cific imprints on the CMB anisotropies. The observations of these patterns, such as the nearly scale-invariant spectrum of fluctuations, has been a vital tool for testing and refining inflationary models. While the CMB has provided an extraordinary amount of valuable information, it also presents some intriguing challenges. For instance, there have been several anomalies observed in the CMB anisotropy, such as the "Cold Spot" (Szapudi et al. 2015; Cruz et al. 2005) and "Axis of Evil" (Land & Magueijo, 2005). These anomalies have led to explore possible explanations, such as statistical fluctuations within cosmic variance, instrumental effects, or more exotic scenarios that could hint at new physics beyond our current understanding.

In conclusion, the Cosmic Microwave Background radiation is a cornerstone of modern cosmology. It confirms the Big Bang theory, offers precise measurements of cosmological parameters, supports the ΛCDM model, helps determine the age of the Universe, sheds light on the origin of cosmic structures, and provides a means to test inflationary cosmology. Its discovery and study have significantly advanced our understanding of the history and evolution of the Universe. Future CMB missions like South Pole Telescope (SPT) - 3G (Benson et al., 2014), Advanced Atacama Cosmology Telescope (adv-ACT) (Henderson et al., 2016), Simons Observatory (SO) (Ade et al., 2019) and CMB-S4 (Abazajian et al. 2019; Abazajian et al. 2016) will give us higher resolution maps of CMB polarisation than achieved by *Planck* mission and will aim towards resolving the anomalies from the CMB measurements.

1.3 Weak Lensing of CMB

The well understood theory of the CMB explains its statistical properties, including the characteristic pattern peaks in the anisotropy power spectra (Scott and Smoot 2006; Challinor 2004). With improvement in the quality of data, observations can probe smaller angular scales where various non-linear signals become important. One of the most important small scale effects is the weak lensing. Weak lensing of the CMB is the deflection of the photons originating from the last scattering surface by the large-scale structures, such as galaxy clusters, filaments, and dark matter halos, that lie between us and the CMB photons. The gravitational potential of the large-scale structures along the line of sight acts as a lens, and it alters the paths of the photons, leading to distortions in the observed CMB temperature and polarization patterns (Metcalf and Silk 1997; Seljak 1996; Linder 1990; Cole and Efstathiou 1989). The strength of the lensing effect depends on the density and distribution of matter along the line of sight.

The typical total deflections suffered by the CMB photons are ~ 2 arcminutes. The weak lensing thus, dominates the CMB anisotropy power spectrum at multipoles $\ell \geq 3000$. Although, the magnitude of the deflections are much smaller than the degree-scale primary CMB acoustic peak, these deflections are correlated over the sky over a comparable scale of $2-3$ degrees (Lewis & Challinor, 2006). Weak lensing broadens the acoustic peak, thus, changing the statistics of the size distribution and diffusing the size of the under-dense or over-dense regions. In addition to the percent-level effect on the CMB temperature power-spectrum, weak lensing introduces non-Gaussian signatures in the CMB and also generates 'B-mode' curl-like polarization signal which may be a source of confusion with any primordial signal from the gravitational waves

(Kamionkowski et al., 1997).

Even though the weak lensing changes the statistical properties of the CMB, these signatures of the lensed CMB field can be exploited to extract information about the lensing deflection field. (Hu and Okamoto 2002; Hu 2001; Zaldarriaga and Seljak 1999). The map of the projected lensing potential can be reconstructed from CMB temperature and polarisation data (Millea et al. 2021; Darwish et al. 2021; Planck Collaboration et al. 2014b; Omori et al. 2017). The integrated information of the matter distribution in the Universe, all the way to the surface of last scattering, is contained in these reconstructed maps of the lensing potential.

Weak lensing of the CMB, both independently and with other cosmological datasets, has led to stringer constraints on cosmological parameters including the densities of baryonic matter and the 'invisible' dark matter content in the Universe, the sum of neutrino masses and the dark energy density component in the evolution of Universe (Madhavacheril et al. 2023; Planck Collaboration et al. 2020b). The combination of CMB weak lensing with various tracers of the large scale structure, which we will discuss in the next section, can provide a vast amount of information on the evolution of the Universe. While weak lensing of the CMB holds great promise as a cosmological probe, several challenges remain. Systematic effects, such as foreground contaminations, instrumental noise, and biases, must be meticulously accounted for in data analysis. Moreover, the ongoing and upcoming CMB experiments, such as the Simons Observatory and CMB-S4, are expected to significantly improve the sensitivity and resolution, offering unprecedented opportunities to study weak lensing in more detail.

1.4 Cross-correlation

The CMB weak lensing is an integrated quantity along the line of sight and holds the surface-projected information of the matter distribution in the Universe. The map of CMB weak lensing by itself does not provide direct information on the evolution of the large-scale gravitational potential. However, this information can be obtained by cross-correlating the lensing map of CMB and tracers of the large scale structure like galaxy density, galaxy clusters, quasars, and radio sources with known redshifts. Since galaxies reside in dark matter halos (Mo et al., 2010) they are good tracers of structures causing gravitational lensing of CMB.

Cross-correlation between the CMB weak lensing and the tracers of the large scale structure can be used to study the distribution and evolution of the dark matter gravitational potential, study the relation of luminous and dark matter (Han et al. 2019; Bianchini and Reichardt 2018), determine the amplitude of structure at different redshifts (Peacock and Bilicki 2018; Doux et al. 2018), and measure galaxy groups and cluster masses (Gupta and Reichardt 2021; Raghunathan et al. 2019; Planck Collaboration et al. 2016b). The cross-correlation approach employed over redshift slices is a powerful tool to study the dynamical evolution of dark energy from the onset of cosmic acceleration and to test the validity of the cosmological models as a function of redshift. The deflections of CMB photons are sensitive to the distribution of matter in the Universe and the related gravitational potential which is governed by the theory of General Relativity.

Cross-correlation measurements are affected by the amplitude and growth of the matter power spectrum and how modifications are made to General Relativity. Any deviations from the underlying theory of gravity, General Relativity, will reflect directly on CMB lensing and hence, on cross-correlations. The importance of cross-correlations between CMB lensing convergence and galaxy positions or galaxy shapes in testing the validity of the ΛCDM model has been firmly established with many studies performed over the past years (Miyatake et al. 2022; Robertson et al. 2021; Abbott et al. 2019; Bianchini and Reichardt 2018; Singh et al. 2017; Giannantonio et al. 2016).

Many cross-correlation studies have been performed over the past with optical catalogues like Sloan Digital Sky Survey (SDSS) (Singh et al., 2020), Dark Energy Survey (DES) (Omori et al. 2019a; Omori et al. 2019b), Wide-Field Infrared Survey Explorer (WISE) (Krolewski et al. 2021; Krolewski et al. 2020; Goto et al. 2012), Two Micron All Sky Survey (2MASS) (Bianchini & Reichardt, 2018) and Subaru Hyper Suprime-Cam (for example, Marques and Bernui 2020; Namikawa et al. 2019) and with radio catalogues from Low-Frequency Array (LOFAR) (Alonso et al., 2021). Cross-correlation studies involving CMB lensing and galaxy density maps have been reported by a number of authors (Darwish et al. 2021; Cao et al. 2020; Aguilar Faúndez et al. 2019; Giusarma et al. 2018; Schmittfull and Seljak 2018; Pullen et al. 2015; Cawthon et al. 2015; Kuntz 2015). Several cross-correlation studies have also been reported between CMB lensing and quasar density maps (Zhang et al. 2021; DiPompeo et al. 2015; Han et al. 2019) as well as high-redshift sub-millimeter sources from the *Herschel* Astrophysical Terahertz Large Area Survey (H-ATLAS) have also been used to perform such studies (Bianchini et al. 2015 and Bianchini et al. 2016).

1.5 Motivation and thesis overview

Although the standard model of cosmology or the ΛCDM model is very well established by cosmological observations, it is not free from challenges and disparities. The nature of dark energy (Λ), accelerated expansion of the Universe, and tensions in cosmological parameters (like $\Omega_m - \sigma_8$ or Hubble tension) among others have raised questions on the validity of ΛCDM model. One of the fundamental assumptions of ΛCDM is that General Relativity (GR) is the correct description of gravity. Even though GR has been successfully tested on many occasions, most of its bounds come from the observations made within our Solar System. Thus, the assumption that this theory is a good description of gravity at cosmological scales needs detailed investigations. Studies in this direction have resulted in several alternatives to the ΛCDM model. These alternatives are divided into two broad categories: Dark Energy models that modify the stress-energy content of the Universe by adding a component to the dark energy equation of state, and Modified Gravity theories, including modifications made to the Einstein-Hilbert action producing deviations from the standard Poisson equation valid in GR. CMB and the large scale structure probe the clustering and distribution of matter in the Universe. They provide observational tests for the theory of gravity at cosmological scales and allow to bring strong constraints on departures from GR. Figure 1.1 presents a schematic overview of testing cosmological models through cross-correlation measurements.

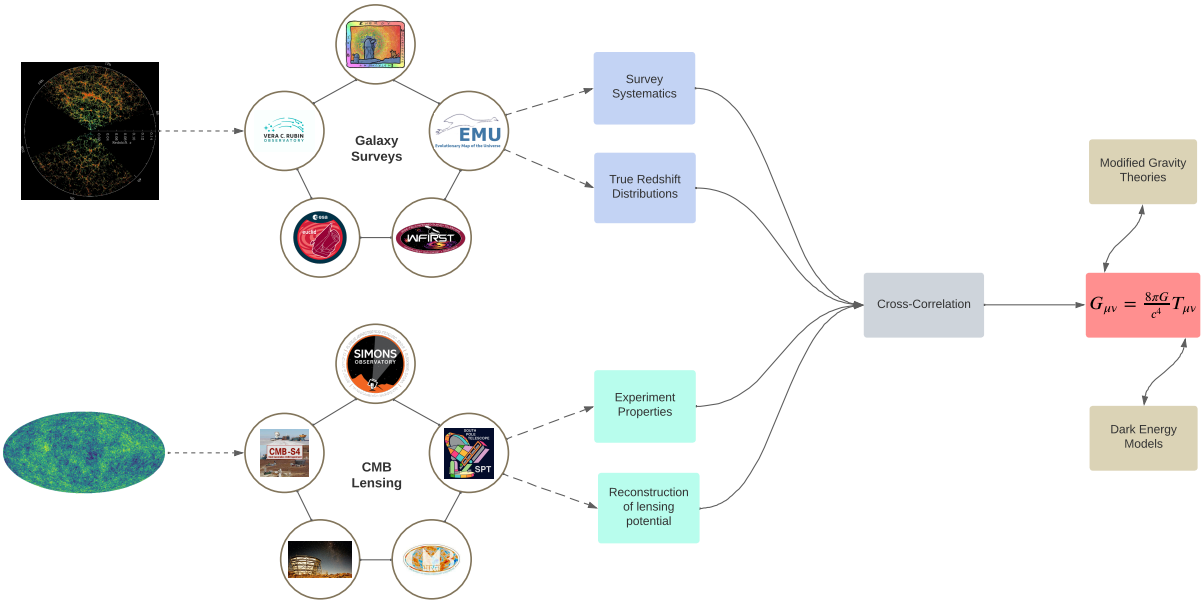


Figure 1.1: A schematic view of the cross-correlation analysis

Observations from the galaxy survey experiments like Sloan Digital Sky Survey (SDSS, Gunn et al. 2006; Strauss et al. 2002), Wide-field Infrared Survey Explorer (WISE, Schlafly et al. 2019; Wright et al. 2010), Kilo-Degree Survey (KiDS, Heymans et al. 2021; de Jong et al. 2015), Hyper Suprime-Cam (HSC, Hikage et al. 2019), Two Micron All Sky Survey (2MASS, Bilicki et al. 2014), and Dark Energy Survey (DES, Abbott et al. 2018) have been the torch-bearer in unveiling the shortcomings of the standard model of cosmology, the ΛCDM model. The baton is now with the upcoming galaxy surveys including the Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST, Ivezić et al. 2019; LSST Science Collaboration et al. 2009), Euclid (Laureijs et al., 2011), Nancy Grace Roman Space Telescope (Spergel et al., 2013), Dark Energy Spectroscopic Instrument (DESI, Dey et al. 2019a), and Spectro-Photometer for the History of the Universe, Epoch of Reionization, and Ices Explorer (SPHEREx, Doré et al. 2014) to provide an in-depth understanding of the workings of our Universe.

The cross-correlation analyses between CMB weak lensing and galaxy samples divided into narrow redshift bins (such as White et al. 2022; Pandey et al. 2022; Chang et al. 2022; Sun et al. 2022; Krolewski et al. 2021; Hang et al. 2021; Peacock and Bilicki 2018) have identified differences in the value of cosmological parameters like the σ_8, Ω_m, or the combined S_8 parameter (defined as $S_8 \equiv \sigma_8 \sqrt{\Omega_m/0.3}$) within the ΛCDM model. These low-redshift cross-correlation probes consistently measure a lower value for S_8 as compared to the high-redshift CMB-only measurements from *Planck* satellite (Planck Collaboration et al., 2020a), resulting in the so-called S_8 tension.

The cross-correlation measurements naturally suffer from similar systematics and errors that affect the number density and redshift distribution of sources in galaxy and quasar surveys. Uncertainties such as photometric redshift errors and catastrophic errors affect the redshift distributions of sources whereas photometric calibration errors bias the number counts of objects

in the sky. One of the primary goals of this thesis is *to robustly test the impact of these error through Monte Carlo (MC) simulations and demonstrate the mitigation of these biases in data analysis.* The cross-correlation analyses over redshift slices additionally suffers from the scatter of objects across redshift bins due to photometric redshift errors, which can lead to a biased estimation of the angular power spectra. *The mitigation of the effects of redshift bin mismatch of objects in a tomographic cross-correlation study* is the second major work of this doctoral thesis. The above-mentioned systematics and errors will bias the cosmological parameters estimated through cross-correlation measurements. This doctoral thesis attempts to present the specifics for a number of systematics in the essence of an unbiased cosmological parameter estimates from the next-generation large scale structure surveys and CMB experiments.

The doctoral thesis is arranged as follows: In Chapter 2, we describe the theoretical angular power spectra for CMB lensing convergence (lensing convergence is the 2-dimensional Laplacian of the lensing potential), galaxy clustering and the cross-correlation between the two. In this chapter, there is also presented the `MASTER` algorithm that we use throughout the body of work to extract full-sky angular power spectrum from observations on partial sky coverage. We also explain briefly the Maximum Likelihood Estimation technique that is used for estimating parameters from power spectra. Chapter 3 presents the first cross-correlation measurements and parameters estimated from the *Planck* CMB lensing potential and photometric redshift galaxy catalogues from the *Herschel* Extragalactic Legacy Project. In this chapter we delve into a detailed study of various systematics that can impact the cross-correlation measurements and subsequently affect the cosmological inferences.

With focus on tomographic cross-correlation measurements with upcoming datasets in Chapter 4, we study in depth through MC numerical simulations how the leakage of objects across redshift bins due to photometric redshift errors results in a biased estimation of angular power spectra in redshift bins. We show that the common methods used in a tomographic analysis to compute the theoretical angular power spectrum with redshift errors are insufficient for correct estimation of parameters. This may lead to apparent tensions on cosmological parameters. Alternatively, we propose using, introduced by Zhang et al., 2010, the scattering matrix formalism with a new, fast method of the matrix estimation to counter the effects of redshift scatter of objects. In Chapter 5, we apply the scattering matrix formalism and re-analyse the cross-correlation study between the *Planck* CMB lensing potential and galaxy catalogue from the Dark Energy Spectroscopic Instrument Legacy Imaging Survey Data Release 8 from Hang et al., 2021. Finally, we summarise all major findings of this doctoral thesis in Chapter 6.

Chapter 2

Theory and Methodology

2.1 Overview

In this chapter, we describe the theoretical and mathematical requisites used throughout this doctoral thesis. The computation of the theoretical angular power spectrum is present in section 2.2. The estimation of full-sky angular power spectra from partial sky coverage maps using the `MASTER` algorithm is detailed in section 2.3. Section 2.4 presents the Maximum Likelihood Estimation technique we use for the estimation of the cosmological parameters.

2.2 Power Spectrum

The gravitational lensing of the CMB photons by the intervening large scale structure acts as a remapping of the unlensed temperature field $\Theta(\hat{\mathbf{n}})$ in the direction $\hat{\mathbf{n}}$ according to (Lewis & Challinor, 2006):

$$\tilde{\Theta}(\hat{\mathbf{n}}) = \Theta(\hat{\mathbf{n}} + \nabla\phi(\hat{\mathbf{n}})) = \Theta(\hat{\mathbf{n}}) + \nabla^a\phi(\hat{\mathbf{n}})\nabla_a\Theta(\hat{\mathbf{n}}) + \mathscr{O}(\phi^2) \tag{2.1}$$

where $\tilde{\Theta}(\hat{\mathbf{n}})$ is the lensed temperature anisotropy field and $\phi(\hat{\mathbf{n}})$ is the CMB lensing potential defined as:

$$\phi(\hat{\mathbf{n}}) = -2\int_0^{\chi_*} d\chi \frac{\chi_* - \chi}{\chi_*\chi}\Psi(\chi\hat{\mathbf{n}}, z(\chi)) \tag{2.2}$$

where χ represents the comoving distance, χ_* is the comoving distance to the surface of last scattering at redshift $z_* \simeq 1100$. $\Psi(\chi\hat{\mathbf{n}}, z(\chi))$ is the three dimensional gravitational potential at position $\chi\hat{\mathbf{n}}$ in photon's path.

The gravitational lensing by foreground matter distribution produces small coherent distortions in the temperature which can be described by the convergence which is related to the lensing potential by a the two-dimensional Laplacian:

$$\kappa(\hat{\mathbf{n}}) = -\frac{1}{2}\nabla^2\phi(\hat{\mathbf{n}}) \tag{2.3}$$

The lensing convergence is related to the line-of-sight matter over-density δ (Bartelmann & Schneider, 2001):

$$\kappa(\hat{\mathbf{n}}) = \int_0^{\chi_*} d\chi \frac{H(\chi)}{c} W^\kappa \delta(\chi\hat{\mathbf{n}}) \tag{2.4}$$

where $H(\chi)$ is the Hubble parameter at comoving distance χ, c is the speed of light in units of km s^{-1} and W^κ is the lensing kernel given by

$$W^\kappa(\chi) = \frac{3\Omega_m}{2c^2} H_0^2 (1+z)\chi \frac{\chi_* - \chi}{\chi_*} \tag{2.5}$$

in which Ω_m and H_0 are the present-day values of the matter density parameter and Hubble constant, respectively. Similarly, one can relate galaxy over-density to the line-of-sight average of the matter over-density as:

$$g(\hat{\mathbf{n}}) = \int_0^{\chi_*} d\chi \frac{H(\chi)}{c} W^g \delta(\chi\hat{\mathbf{n}}) \tag{2.6}$$

with W^g being the galaxy kernel:

$$W^g(\chi) = b(z(\chi))\frac{H(\chi)}{c}\frac{dN}{dz(\chi)} + \frac{3\Omega_m}{2c^2} H_0^2 (1+z)\chi \int_\chi^{\chi_*} d\chi' \frac{H(\chi')}{c}\left(1 - \frac{\chi}{\chi'}\right)(\alpha(\chi') - 1)\frac{dN}{dz(\chi')} \tag{2.7}$$

In the galaxy kernel, $\frac{dN}{dz}$ is the normalized redshift distribution of galaxies, $b(z)$ is the galaxy linear bias that relates the distribution of luminous tracers of the large-scale structure to the underlying total distribution of matter (including dark matter). The second term in Eq. 2.7 accounts for the gravitational magnification of the background objects by foreground sources, and is termed as magnification bias (Turner, 1980). The effect of the magnification bias depends on the slope of the distribution of the integral counts of sources as a function of flux S, i.e. the slope α computed from the $\log N(>S) \propto -\alpha \log S$ distribution.

With the expressions of the lensing kernel and galaxy kernel, the theoretical angular power spectrum is computed using the Limber approximation (Limber, 1953)

$$C_\ell^{xy} = \int_0^{\chi_*} d\chi \frac{W^x(\chi)W^y(\chi)}{\chi^2} P\left(k = \frac{\ell + 1/2}{\chi}, z(\chi)\right) \tag{2.8}$$

where $\{x,y\} = \{\kappa, g\}$, $\kappa \equiv$convergence and $g \equiv$galaxy over-density and $P\left(k = \frac{\ell+1/2}{\chi}, z(\chi)\right)$ is the matter power spectrum which is generated using cosmology codes like `CAMB`[1](Lewis et al., 2000). We compute, using Eq. 2.8, the lensing convergence auto-power spectrum ($C_\ell^{\kappa\kappa}$), galaxy auto-power spectrum (C_ℓ^{gg}), and the cross-power spectrum ($C_\ell^{\kappa g}$) between lensing convergence and galaxy over-density fields.

[1] https://camb.info/

2.3 Extracting full-sky Power Spectra - MASTER

Observations from the extragalactic surveys and CMB experiments are limited to a fraction of the sky. The *Planck* CMB lensing potential map covers $\sim 67\%$ of full sky, masking the regions dominated by emissions from our own galaxy. We need to estimate the full-sky power spectrum for different fields (galaxy density, quasar density, CMB lensing convergence, etc.) from fluctuations observed from a limited area of the sky. We implement our own full-sky estimator based on the `MASTER` algorithm (Hivon et al., 2002) to recover the full-sky power spectrum from the partial sky observations. Figure 2.1 presents a schematic overview of the extraction procedure of the angular power spectrum from maps.

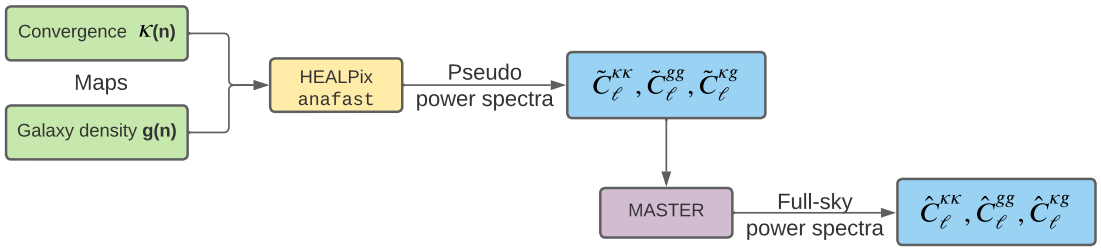


Figure 2.1: Schematic for the extraction of full sky-power spectrum from partial maps

The `MASTER` algorithm is based on the direct spherical harmonic transform of the observed area of the sky, and takes into account mode coupling and pixelization effects to produce an unbiased estimate of the full-sky power spectrum. The incomplete sky coverage leads to a coupling between different spherical harmonic modes and is described by the spherical harmonic transform of the window function. We use the HEALPix `anafast` routine to compute pseudo power spectra $\tilde{C}_\ell$, from the partial sky coverage maps. The full-sky power spectrum $\hat{C}_\ell$ is related to the pseudo power spectrum

$$\tilde{C}_\ell = \sum_{\ell'} M_{\ell\ell'} B_{\ell'}^2 \hat{C}_{\ell'} \tag{2.9}$$

where $M_{\ell\ell'}$ is the mode coupling kernel (see Hivon et al., 2002 for explicit expression and computation strategies for the coupling kernel) and B_ℓ is the pixel window function to account for pixelization effects coming from the finite size of the pixels. The coupling kernel tends to be singular for very small fractions of sky coverage and Eq. 2.9 cannot be inverted directly. To overcome this singularity problem we bin the pseudo power spectrum and the mode coupling kernel in multipoles using binning operators. Eq. 2.9 after binning translates to

$$\tilde{C}_L = \sum_{L'} K_{LL'} \hat{C}_{L'} \tag{2.10}$$

where

$$\tilde{C}_{L'} = \sum_{\ell'} P_{L'\ell'} \tilde{C}_{\ell'} \tag{2.11}$$

and

$$K_{LL'} = \sum_{\ell\ell'} P_{L\ell} M_{\ell\ell'} B^2_{\ell'} Q_{\ell' L'} \tag{2.12}$$

Here, L stands for the multipole bin and $P_{L\ell}$ is the binning operator expressed as:

$$P_{L\ell} = \begin{cases} \frac{1}{2\pi} \frac{\ell(\ell+1)}{\ell^{(L+1)}_{\rm low} - \ell^{(L)}_{\rm low}}, & \text{if} \quad 2 \leq \ell^{(L)}_{\rm low} \leq \ell \leq \ell^{(L+1)}_{\rm low} \\ 0, & \text{otherwise} \end{cases} \tag{2.13}$$

and $Q_{\ell L}$ is the reciprocal binning operator:

$$Q_{\ell L} = \begin{cases} \frac{2\pi}{\ell(\ell+1)}, & \text{if} \quad 2 \leq \ell^{(L)}_{\rm low} \leq \ell \leq \ell^{(L+1)}_{\rm low} \\ 0, & \text{otherwise} \end{cases} \tag{2.14}$$

The final step in the extraction procedure is to subtract the unwanted noise contributions full-sky power spectrum. We perform Monte-Carlo simulations of the observed noise fields, and subtract the ensemble average of these simulations from the recovered full-sky power spectrum. The noise-subtracted full-sky power spectrum then reads

$$\hat{C}_L = \sum_{L'} K^{-1}_{LL'} \tilde{C}_{L'} - \hat{N}_L \tag{2.15}$$

2.4 Maximum Likelihood Estimation

Maximum Likelihood Estimation is a method of estimating parameters given some observed data, my maximizing the likelihood function. The likelihood function represents the joint probability space for the observed data given the free parameters of the model. Maximizing the likelihood function, thus, gives the values of parameters such that the observed data is most probable for the model under consideration.

In this thesis, we will estimate two parameters, the galaxy linear bias b and the amplitude of cross-correlation A, from the measurements of the galaxy auto-power spectrum and cross-power spectrum between CMB lensing and galaxy density. The amplitude of cross-correlation, A is a phenomenological parameter that acts as a re-scaling of the cross-power spectrum with respect to the fiducial cross-power spectrum computed for the assumed background cosmology. The amplitude parameter can then be used to test the validity of the cosmological model. The galaxy linear bias, on the other hand, can be used to put constrains on models of structure formation in the Universe.

The galaxy auto-power spectrum scales as b^2, whereas the cross-power spectrum depends on the product of the parameters $b \times A$ and induces a degeneracy in the estimation of parameters. To break this degeneracy, we estimate parameters from a joint likelihood function including both galaxy auto-power spectrum and cross-power spectrum. Thus in total, we perform estimations from three likelihood functions corresponding to C^{gg}_ℓ, $C^{\kappa g}_\ell$ and their joint data vector $C_\ell = (C^{\kappa g}_\ell, C^{gg}_\ell)$. The choice of three likelihood functions helps identify sources of systematics that may affect the sampling of parameter space by the likelihood function. The likelihood function

has the form

$$\mathscr{L}(\hat{C}_\ell|b,A) = \frac{1}{\sqrt{(2\pi)^{N_\ell}\det(\mathrm{Cov}_{\ell\ell'})}}\exp\left\{-\frac{1}{2}[\hat{C}_\ell - C_\ell(b,A)](\mathrm{Cov}_{\ell\ell'})^{-1}[\hat{C}_{\ell'} - C_{\ell'}(b,A)]\right\} \tag{2.16}$$

where N_ℓ is the number of data points, the power spectra with and without hat represents the measured power spectra from data and their corresponding fiducial templates, respectively. The covariance matrix $\mathrm{Cov}_{\ell\ell'}$ for the likelihood function corresponding to C_ℓ^{gg} and $C_\ell^{\kappa g}$ can be computed from the expression of the generalised covariance matrix present in appendix 7.1 by setting $(A,B,C,D) \equiv (g,g,g,g)$ and $(A,B,C,D) \equiv (\kappa,g,\kappa,g)$, respectively. For the joint likelihood function the covariance matrix is given by

$$\mathrm{Cov}_{\ell\ell'} = \begin{bmatrix} \mathrm{Cov}_{\ell\ell'}^{\kappa g,\kappa g} & \mathrm{Cov}_{\ell\ell'}^{\kappa g,gg} \\ \mathrm{Cov}_{\ell\ell'}^{\kappa g,gg} & \mathrm{Cov}_{\ell\ell'}^{gg,gg} \end{bmatrix} \tag{2.17}$$

where the covariance $\mathrm{Cov}_{\ell\ell'}^{\kappa g,\kappa g}$ accounts for the correlation between the CMB convergence and galaxy density fields, can be computed from Eq. 7.13 by setting $(A,B,C,D) \equiv (\kappa,g,g,g)$.

Chapter 3

Cross-correlation between *Planck* CMB lensing potential and galaxy catalogues from HELP

3.1 Overview

In this chapter, we present the first cross-correlation measurements between the photometric redshift galaxy catalogues from the *Herschel* Extragalactic Legacy Project (HELP) and CMB lensing potential measured by the *Planck* satellite. We extract the full-sky power spectra for four HELP fields and estimate two parameters, galaxy linear bias b and amplitude of cross-correlation A using Maximum Likelihood Estimation. This chapter is based on Saraf et al., 2022 published in Monthly Notices of the Royal Astronomical Society.

The chapter is arranged as follows: we describe the CMB lensing and HELP galaxy catalogues in section 3.2. The methodology to extract the full-sky power spectra from the partial sky coverage maps is described in section 3.3. We validate our pipeline for power spectra extraction in section 3.5. The results of cross-correlation measurements from *Planck* CMB lensing potential and HELP galaxy catalogues are present in section 3.6. Finally, we discuss various systematics that may affect the parameters in section 3.7 and summarise our findings in section 3.8.

3.2 Data

3.2.1 CMB Lensing Data

The CMB lensing data used for the cross-correlation measurement presented in this chapter comes the 2018 Public Data Release 3 (PDR3)[1] made by the *Planck* collaboration described in Planck Collaboration et al., 2020b. *Planck* was the third generation satellite mission of the European Space Agency dedicated to precise measurements of the CMB anisotropies, aimed at studying the early Universe and its evolution. Currently the best constraints on cosmological parameters based on CMB measurements come from *Planck* observations. Details about the *Planck* mission and its performance can be found in Planck Collaboration et al., 2016a and Planck Collaboration et al., 2014a.

The CMB lensing potential map we use in our study is based on the *Planck* PDR3 High-Frequency Instrument maps described in Planck Collaboration et al., 2020c. It uses `SMICA` foreground cleaned, DX12 CMB maps to reconstruct the lensing potential employing a quadratic lensing estimator (Hu and Okamoto 2002; Zaldarriaga and Seljak 1999). We use for our baseline analysis, the lensing convergence map derived from a minimum-variance estimate of the CMB temperature and polarization data (hereafter, MV map).

The *Planck* PDR3 provides spherical harmonic coefficients for the lensing convergence map in HEALPix[2] (Górski et al., 2005) format with resolution parameter $N_{\rm side} = 4096$. We build the lensing convergence maps from the spherical harmonic coefficients through the HEALPix `alm2map` routine. The smaller angular scales (or higher multipoles) in the lensing convergence maps are dominated by noise, rendering them impractical for cross-correlation measurements. Thus, for our analysis we downgrade these maps to a lower resolution parameter $N_{\rm side} = 512$, which effectively reduces the processing time for these maps without compromising the quality of our results. The data package also provides the noise power spectrum for MV maps computed from 300 simulations performed by the *Planck* collaboration as well as binary maps to mask out the regions of sky not suitable for analysis.

3.2.2 Galaxy Data

The *Herschel* Extragalactic Legacy Project or HELP[3] is an optical to near-infrared galaxy catalogue with data products derived from 51 premium multiwavelength extragalactic public data sets (Shirley et al. 2021; Shirley et al. 2019). HELP covers 1270 deg^2 in the sky divided into 23 fields defined by the *Herschel* SPIRE extragalactic survey fields, including the *Herschel* Multi-tiered Extragalactic Survey (HerMES; Oliver et al., 2012) and the *Herschel* Astrophysical Terahertz Large Area Survey (H-ATLAS; Eales et al., 2010). The HELP fields have observations spanning the wavelength range $0.36-4.5\,\mu m$ with fluxes in the u,g,r,i,z,y,J,H,K,K_s, and IRAC Ch $1,2,3,4$ photometric filters. The g,r,i,z,y fluxes are available for all HELP fields, the K,K_s

[1] https://pla.esac.esa.int/#cosmology
[2] https://healpix.jpl.nasa.gov/
[3] https://herschel.sussex.ac.uk

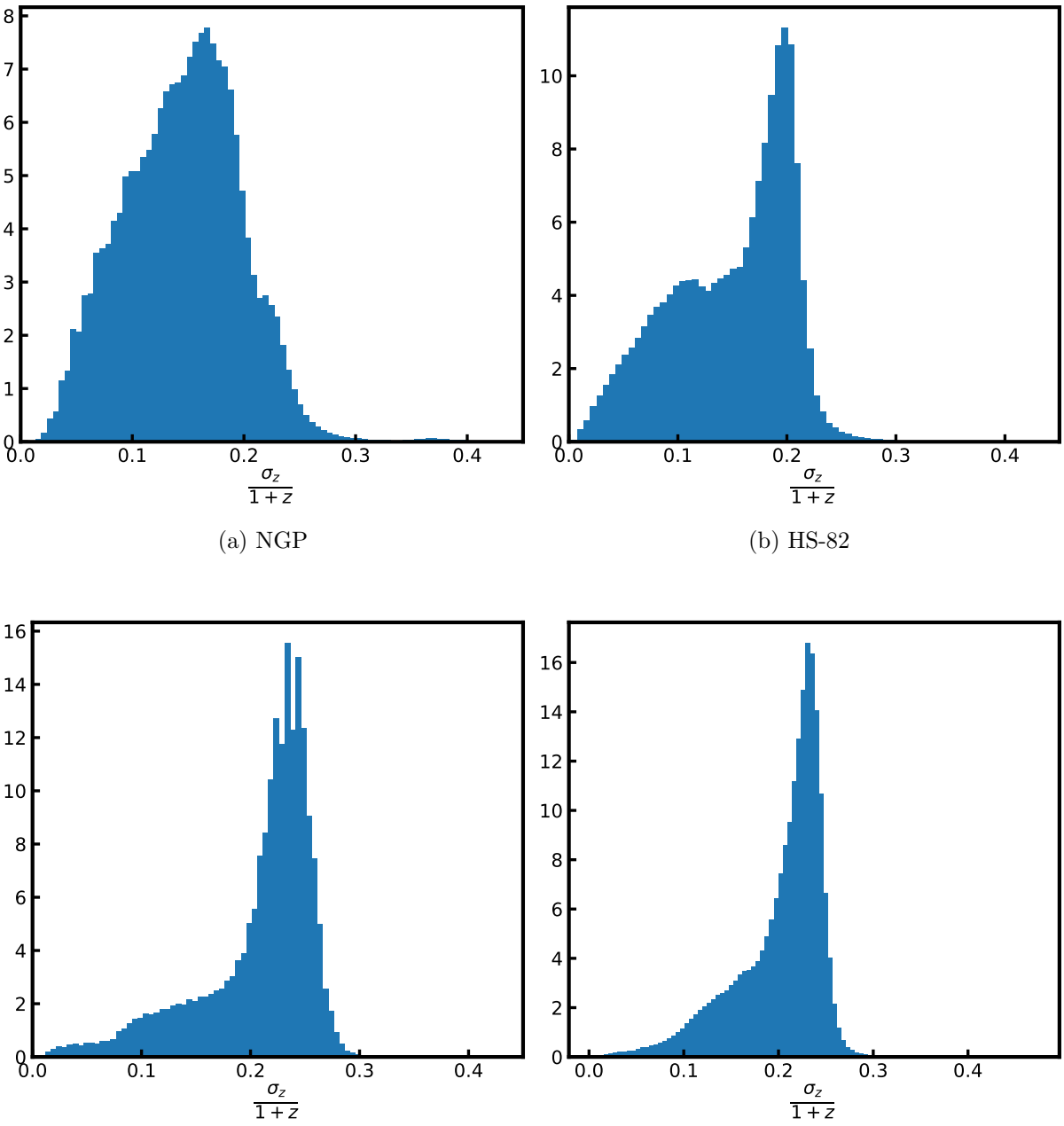
band covers ~ 1130 deg^2 while the IRAC Ch bands are available only on < 300 deg^2. A detailed list of different public surveys contributing to the HELP fields and information on the depth of observations in different photometric bands can be found in appendix A of Shirley et al., 2019.

We filter the available 23 HELP fields based on their physical area and inhomogeneities in the galaxy density due to different depths of the contributing public surveys. We obtain three fields suitable for cross-correlation analysis namely, the North Galactic Pole (NGP), the *Herschel* Stripe-82 (HS-82) and the South Galactic Pole (SGP). We further observe large scale variations in the galaxy density within the SGP field. These variations can be attributed to the fact that the Kilo-Degree Survey (de Jong et al., 2013) only covers approximately one half of the SGP field. while the Dark Energy Survey (Abbott et al., 2018) covers the other half. The depth in the g,r,i bands for the KiDS survey is $25.4, 25.2, 24.2$, respectively, whereas the depth in the same bands for the DES survey is $24.33, 24.08, 23.44$, respectively. The differences in depths contribute directly to the different mean number of galaxies in the two separate halves of the SGP field. Thus, to overcome any systematics that may arise from the galaxy density variation, we divide the SGP field in two parts namely SGP Part-1 and SGP Part-2. We apply additional selection criteria to increase the purity of the galaxy sample. We filter objects with quality control flags `flag-gaia`≤ 2 and `stellarity`< 0.9. The combination of these two parameters removes star-like objects from our sample and increase the probability for the object to be a galaxy.

Table 3.1: Physical properties of HELP patches. f_{sky} is the fraction of sky covered by patches, $[l,b]$ are galactic longitude and latitude respectively for the center of the field, $N_{\rm obj}$ is the number of objects in each field, and $\overline{n}$ is the mean number of objects.

Field	f_{sky}	area [deg^2]	$[l,b]$	$N_{\rm obj}$	$\overline{n}$ [gal pix^{-1}]	$\overline{n}$ [gal str^{-1}]	median z
NGP	0.0043	179.14	[51°,84°]	1311549	96.908	2.426×10^7	0.45
HS-82	0.0062	255.16	[130°,-61°]	6824474	344.862	8.633×10^7	0.60
SGP Part-1	0.0020	85.83	[12°,-68°]	3151922	481.577	1.206×10^8	0.71
SGP Part-2	0.0035	145.32	[-82°,-86°]	6659404	600.975	1.504×10^8	0.71

After removing star-like objects, we constrain our sample based on the photometric redshift errors of the galaxies. The HELP catalogues provide photometric redshifts of objects computed from the aperture fluxes through the Easy and Accurate Z from Yale code (`EAZY`; Brammer et al. 2008). An overview of the HELP photometric redshift estimation pipeline can be found in Duncan et al., 2018a and Duncan et al., 2018b. In Figure 3.1, we show the distributions for relative error on photometric redshifts, $\frac{\sigma_z}{1+z}$, for all the four fields considered in our analysis. We restrict our sample with objects having $\frac{\sigma_z}{1+z} \leq 0.15$ for NGP and HS-82 fields, but larger value of $\frac{\sigma_z}{1+z} \leq 0.25$ for SGP Part-1 and Part-2 fields because the relative error distribution for both parts of the SGP field has a larger median than NGP or HS-82 fields. With these cuts on redshift errors, we create the final catalogue of ~ 18 million objects divided onto NGP, HS-82 and SGP fields. Table 3.1 provides the physical properties for every field including the fraction of sky area covered (f_{sky}), the total area of the field, the galactic longitude and latitude for the center of the field ($[l,b]$), the total number of objects $N_{\rm obj}$, the mean number of objects $\overline{n}$ per pixel and per steradian, and the median redshift of the galaxies in the field.



(a) NGP (b) HS-82

(c) SGP Part-1 (d) SGP Part-2

Figure 3.1: $\frac{\sigma_z}{1+z}$ distributions for HELP patches.

From the final catalogue, we create the galaxy over-density maps at HEALPix resolution parameter $N_{\rm side} = 512$ using the relation

$$g(\hat{\mathbf{n}}) = \frac{n(\hat{\mathbf{n}}) - \overline{n}}{\overline{n}} \tag{3.1}$$

where $n(\hat{\mathbf{n}})$ is the number of objects at angular position $\hat{\mathbf{n}}$ and $\overline{n}$ is the mean number of objects per pixel. In Figure 3.2, we show the convergence and galaxy over-density maps for the part of the sky common to both *Planck* lensing convergence and galaxy density fields. We have filtered out multipoles $\ell \geq 400$ from these maps for illustrative purpose, to show the large-scale fluctuations present in the field. We can evidently conclude from Figure 3.2 that the galaxy-density and convergence fields are homogeneous. We also show in Figure 3.3, the distribution of $g(\hat{\mathbf{n}})$

for the four fields and find that the over-density distribution can be very well approximated by a Gaussian function. This, in turn, confirms that the galaxy density fields have a homogeneous distribution of objects with mean density of objects as quoted in Table 3.1.

Few key points about the cross-correlation analysis performed in this chapter are:

- The number of density of objects in our sample is approximately two orders of magnitude larger than the previous analyses of far-infrared *Herschel* datasets performed, for example in Bianchini et al., 2016 and Bianchini et al., 2015.
- We use the full lensing convergence map (which covers $\sim 67\%$ of the sky) in our analysis, instead of the common area map with the galaxy density fields. In Figure 3.2 we have shown the common area shared between the convergence and galaxy density fields only as an illustration.

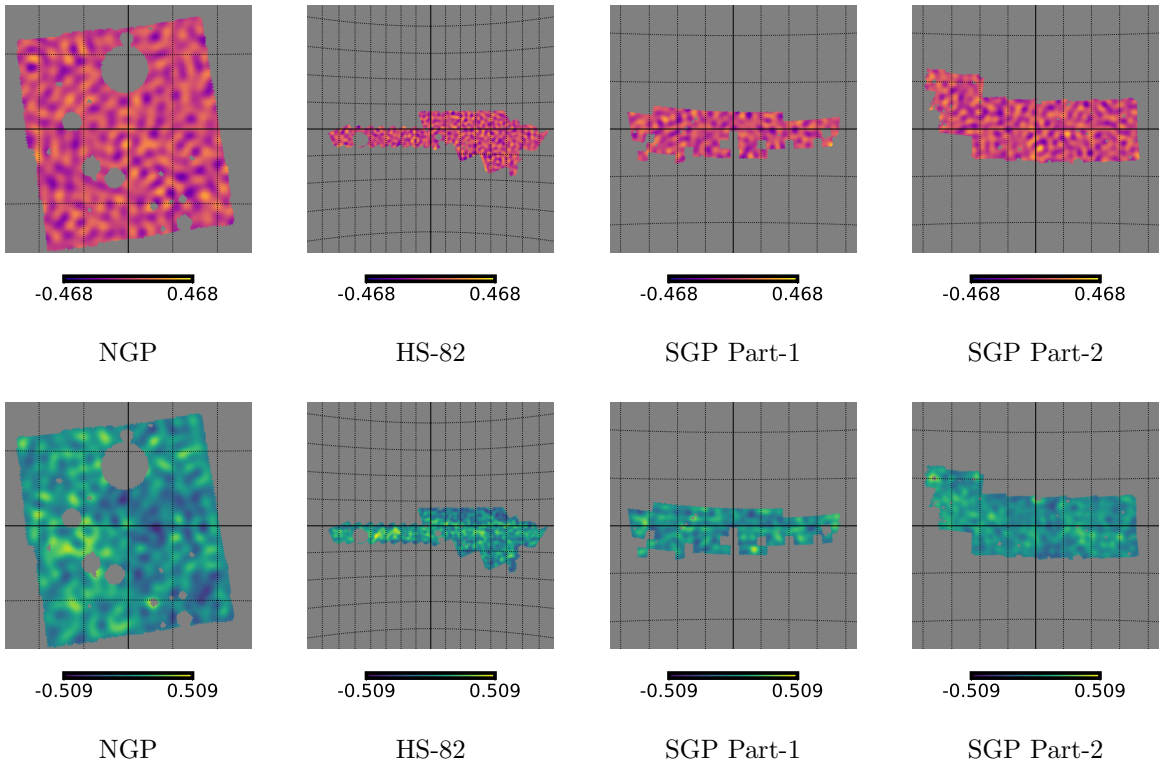
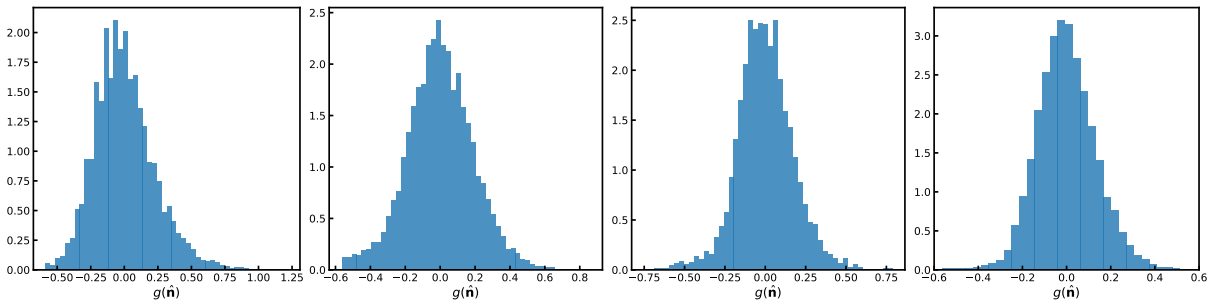


Figure 3.2: Convergence maps (upper row) and galaxy over-density maps (bottom row) of NGP, HS-82, SGP Part-1 and SGP Part-2 fields. Multipoles $\ell \geq 400$ have been filtered out from all maps. The grid spacing is 3° in longitude and 5° in latitude.

The HELP catalogue also provides posteriors of estimated redshifts for every object based on the `EAZY` code. We stack these posteriors together to build the redshift galaxy distributions for every HELP field. The photometric redshift errors translates to the redshift distributions via these posteriors (Budavári et al., 2003). We show the comparison of redshift distributions for HELP fields and CMB lensing kernel in Figure 3.4. The lensing kernel and redshift distributions are normalised to unit maximum for illustrative comparison. The redshift galaxy distribution is then used to compute the theoretical power spectrum using Eq. (2.8). We estimate the slope of



NGP HS-82 SGP Part-1 SGP Part-2

Figure 3.3: The distribution of over-density $g(\hat{\mathbf{n}})$ for the four galaxy fields

the distribution of the integral counts of sources for the four HELP fields to quantify the impact of magnification bias in the galaxy kernel, and find that $\alpha = 1$ for all the four fields.

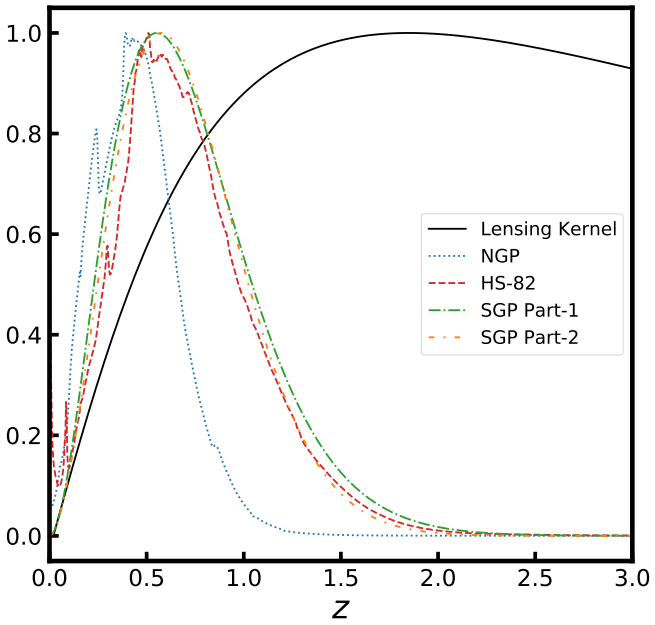


Figure 3.4: CMB lensing kernel W^{κ} compared with redshift distributions for all galaxy patches. Both lensing kernel and redshift distributions are normalised to the unit maximum.

3.3 Full-sky power spectra

From the galaxy density and lensing convergence maps prepared in section 3.2, we compute different angular power spectra for all the four HELP field namely, the lensing convergence auto-power spectrum ($C_\ell^{\kappa\kappa}$), the galaxy auto-power spectrum (C_ℓ^{gg}) and the cross-power spectrum between CMB lensing and galaxy density ($C_\ell^{\kappa g}$). We use our estimator described in section 2.3 to extract the full-sky power spectra $\hat{C}_\ell^{xy}$, where $\{x, y\} = \{\kappa, g\}$. To subtract the noise contribution from power spectra, we perform Monte-Carlo simulations of the lensing convergence and galaxy noise fields, and subtract the ensemble average of these simulations from the recovered full-sky power spectra.. Since the origin of noise in the CMB convergence and galaxy density fields are uncorrelated, we take the noise associated with the cross-power spectrum, $N_\ell^{\kappa g}$, to be zero. We estimate the full-sky power spectra using the `MASTER` algorithm in the multipole range

$0 \leq \ell \leq 1200$, however, the performance of the mode-coupling kernel at the boundaries is noisy and we restrict our analysis to the multipole range $100 \leq \ell \leq 800$ divided into 7 linear multipole bins with binwidth $\Delta\ell = 100$.

3.4 Errors on Power Spectrum

The error on the power spectrum, in analyses based on the common sky area coverage between CMB lensing and galaxy density fields, can be computed from the diagonal of the covariance matrix given by

$$\mathrm{Cov}^{xy}_{LL'} = \frac{1}{(2\ell_{L'}+1)\Delta\ell f_{sky}}\left[C^{xy}_{L}C^{xy}_{L'} + \sqrt{C^{xx}_{L}C^{xx}_{L'}C^{yy}_{L}C^{yy}_{L'}}\right] \tag{3.2}$$

where $\{x,y\} = \{\kappa, g\}$, $\Delta\ell$ is the multipole binwidth, and f_{sky} is the fraction of sky common between x and y.

In our analysis, we use maps with different fraction of sky coverage; $\sim 67\%$ for convergence and $< 1\%$ for HELP fields. Hence, we generalise the expression for covariance matrix from Eq. (3.2) to implement different sky area coverage in the computation of power spectrum errors. The generalised covariance matrix then reads

$$\mathrm{Cov}^{AB,CD}_{LL'} = \frac{1}{(2\ell_{L'}+1)\Delta\ell f^{AB}_{sky} f^{CD}_{sky}}\left[f^{AC,BD}_{sky}\sqrt{C^{AC}_{L}C^{AC}_{L'}C^{BD}_{L}C^{BD}_{L'}} + f^{AD,BC}_{sky}\sqrt{C^{AD}_{L}C^{AD}_{L'}C^{BC}_{L}C^{BC}_{L'}}\right] \tag{3.3}$$

where $\{A,B,C,D\} = \{\kappa, g\}$, f^{AB}_{sky} is the fraction of sky common to fields A and B. $f^{AC,BD}$ is the composite quantity that acts as a weight when taking into account different observed sky areas for different fields. We evaluate $f^{AC,BD}$ from the mode coupling kernel computed during the extraction of power spectra from the `MASTER` algorithm. A detailed derivation of the generalized covariance matrix can be found in Appendix 7.1.

3.5 Pipeline validation

In section 3.3, we described the methodology implemented to compute the full-sky power spectra for the convergence and galaxy density fields using the `MASTER` algorithm. Before proceeding to the estimation of parameters from the power spectra, it is vital to check the performance of our full-sky power spectrum estimator. To validate our numerical algorithm and ensure an unbiased estimation of the full-sky power spectra, we simulate 500 correlated maps of the *Planck* CMB lensing convergence and galaxy over-density fields with statistical properties consistent with HELP catalogues. We follow the procedure of Kamionkowski et al., 1997 and introduce a known degree of correlation using the relation

$$\begin{aligned} \kappa_{\ell m} &= \alpha_1 (C^{\kappa\kappa}_{\ell})^{1/2}; \\ g_{\ell m} &= \alpha_1 \frac{C^{\kappa g}_{\ell}}{(C^{\kappa\kappa}_{\ell})^{1/2}} + \alpha_2 \left[C^{gg}_{\ell} - \frac{(C^{\kappa g}_{\ell})^2}{C^{\kappa\kappa}_{\ell}}\right]^{1/2} \end{aligned} \tag{3.4}$$

where α_1 and α_2 are two complex random numbers drawn from a Gaussian distribution with unit variance for each ℓ and $m > 0$, while for $m = 0$, α_1 and α_2 are real random numbers. C_ℓ's are the fiducial power spectra through which correlations are induced between convergence and galaxy density fields (see chapter 2 for details on the fiducial power spectra). These correlated spherical harmonic coefficients generate maps with only signal and we add noise onto these signal only simulations.

To the simulated convergence maps, we add noise using the minimum variance noise power spectrum $N_\ell^{\kappa\kappa}$, provided in the *Planck* PDR3 data package. We add noise to galaxy over-density maps by generating a number count map in which each pixel has the value drawn from a Poisson distribution with mean

$$\lambda(\hat{\mathbf{n}}) = \overline{n}(1 + g(\hat{\mathbf{n}})) \tag{3.5}$$

where $\overline{n}$ is the observed mean number of sources per pixel for HELP fields and $g(\hat{\mathbf{n}})$ is the corresponding signal only simulated galaxy map generated using Eq. (3.4). The final galaxy over-density map suitable for pipeline validation is build from the number count map using Eq. (3.1). The noise simulated by the Poisson sampling method has statistical properties consistent with observations. The galaxy over-density maps, simulated with the procedure described above, often leads to some pixels where $g < -1$. Then the Poisson sampling for these pixels by Eq. (3.5) is not possible. We reject those simulations in which pixel value $g < -1$ on the area covered by the HELP fields.

With 500 simulated maps of the *Planck* CMB lensing convergence and HELP galaxy fields, we compute the mean power spectrum as

$$\overline{C}_L^{xy} \equiv \langle \hat{C}_L^{xy} \rangle = \frac{1}{N} \sum_{i=1}^{N} \hat{C}_L^{xy,i} \tag{3.6}$$

where $\hat{C}_L^{xy,i}$ represents the i^{th} simulated power spectrum estimate and N is the number of simulations. The error associated with the mean power spectrum is computed from the square root of the diagonal of the covariance matrix given by

$$\mathrm{Cov}_{LL'}^{xy} = \frac{1}{N-1} \sum_{i=1}^{N} (\hat{C}_L^{xy,i} - \overline{C}_L^{xy})(\hat{C}_{L'}^{xy,i} - \overline{C}_{L'}^{xy}) \tag{3.7}$$

In Figure 3.5, we show the mean power spectra, $\overline{C}_L^{\kappa\kappa}$,$\overline{C}_L^{gg}$, and $\overline{C}_L^{\kappa g}$ from 500 simulations of the HS-82 field, computed from our full-sky power spectrum estimator based on the `MASTER` algorithm. The upper panel shows the mean power spectrum estimate in blue data points with red solid curve representing the theoretical power spectrum used for simulations. The bottom panel presents the relative differences between the mean power spectrum and theoretical power spectrum. The power spectra are recovered well within 1σ errors and we do not observe any hints of systematic offsets that can lead to an unbiased estimation of parameters.

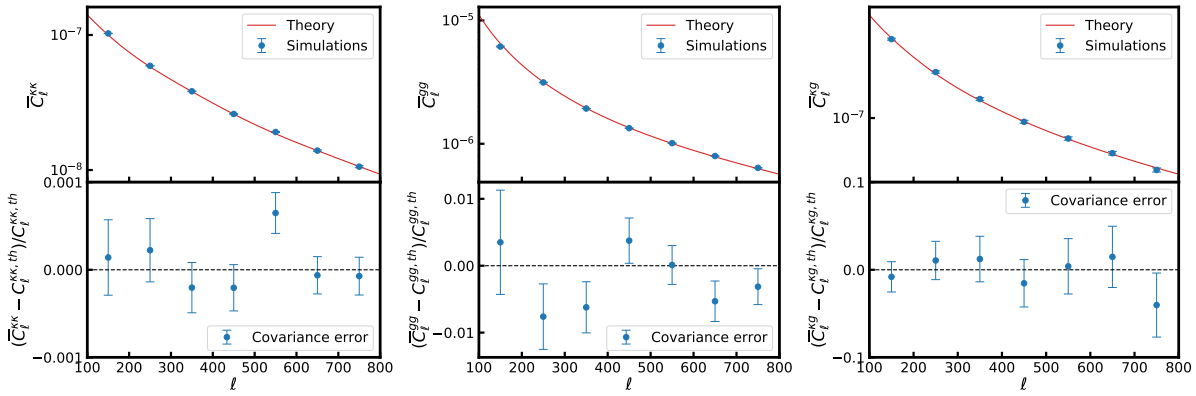


Figure 3.5: *Top* : Average power spectra reconstructed for 500 simulations of HS-82 field. The red line represents the theoretical power spectrum used for simulations. *Bottom*: Error estimated for the reconstructed simulated average power spectra relative to the theoretical power spectrum. *Left to Right*: CMB lensing convergence power spectrum, galaxy auto-power spectrum, and cross-power spectrum

3.6 Results for *Planck* X HELP

We described in section 3.3 the procedure to extract full-sky power spectrum from the convergence and galaxy over-density maps. In section 3.5, we explored the performance of our full-sky estimator through simulations of correlated CMB lensing convergence and galaxy over-density maps. In this section, we present the results for the first cross-correlation measurements between *Planck* CMB lensing convergence and HELP galaxy catalogues.

3.6.1 Noise correlation

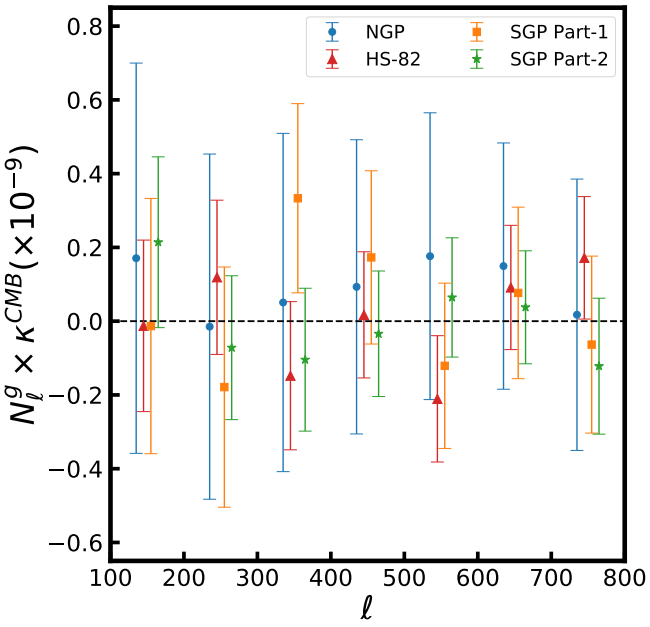


Figure 3.6: Mean correlation between galaxy shot noise obtained from jackknifing approach and CMB lensing convergence map. No significant signal is detected for any patch.

In section 3.3, we argued that the noise associated with the cross-power spectrum, $N_\ell^{\kappa g}$ will

be zero. Here we test this hypothesis with observational data. We estimate the galaxy shot noise by jackknifing approach (Ando et al. 2018; Bianchini and Reichardt 2018) and cross-correlate it with *Planck* CMB convergence map. Under the jackknifing approach, we randomly split the galaxy catalogue into two halves and compute independent over-density maps (δ_g^1 and δ_g^2) from these halves. The combination $(\delta_g^1+\delta_g^2)/2$ then contains both signal and noise, while $(\delta_g^1-\delta_g^2)/2$ gives the noise contribution. We repeat the jackknifing procedure 500 times and show the mean correlation between the difference map $(\delta_g^1-\delta_g^2)/2$ for the four HELP fields and the true *Planck* CMB convergence map in Figure 3.6. The error bars are the square root of the diagonal of the covariance matrix built from 500 independent splits of the galaxy catalogues. The mean noise power spectrum $N_\ell^{\kappa g}$ is consistent with zero correlation for all four HELP fields, having $\chi^2/\nu = 0.083, 0.537, 0.429$ and 0.283 for $\nu = 7$ degrees of freedom, and no correlation probability of $0.99, 0.81, 0.88$ and 0.96 for NGP, HS-82, SGP Part-1 and SGP Part-2, respectively.

3.6.2 Power spectra

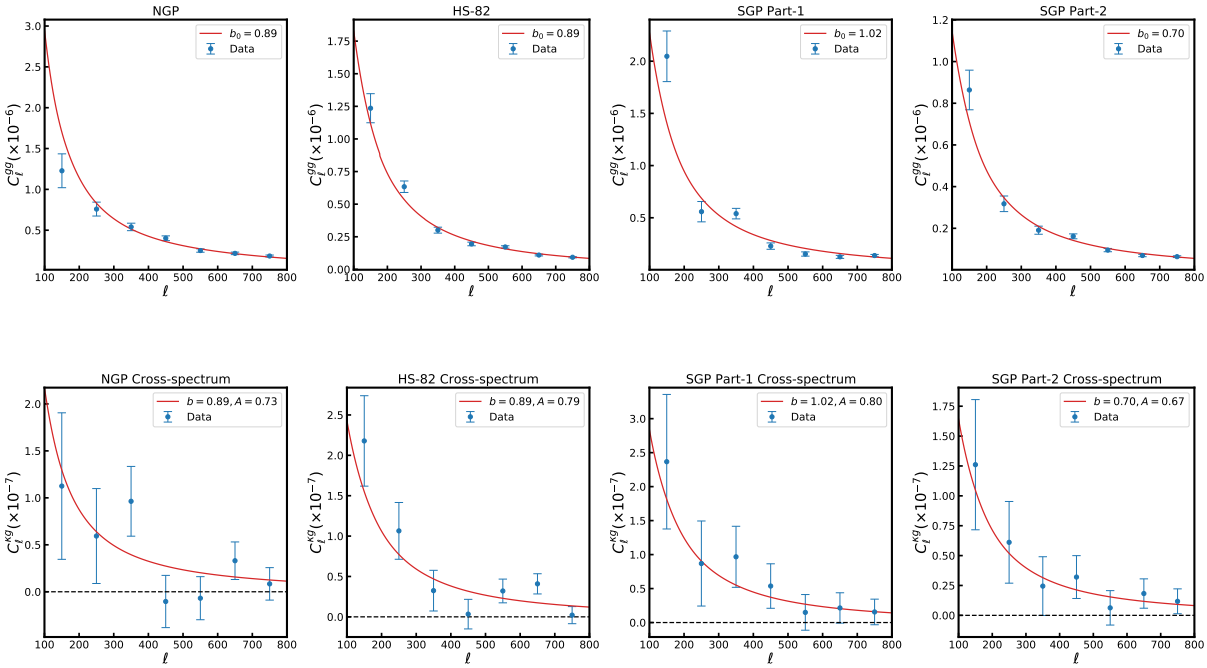


Figure 3.7: Galaxy auto-power spectra (*Top*) and cross-power spectra (*Bottom*) for HELP fields. The blue circles represent the measurements of cross-correlation signal using MV map. The red solid line is the theoretical power spectrum computed using the best-fit values obtained from likelihood analysis using MV map.

The noise subtracted galaxy auto-power spectra and cross-power spectra for all four HELP fields are shown in top and bottom panels of Figure 3.7, respectively. The error bars on the power spectra are computed from the square root of the diagonal of the analytical covariance matrix following Eq. (3.3), using the best fit values of galaxy linear bias b and amplitude of cross-correlation A (the estimation of parameters is discussed in section 3.6.3).

We compute the probability of rejecting the null-hypothesis, that there is no correlation between the HELP galaxy fields and *Planck* CMB lensing convergence, by using the relation

$$\chi^2_{\rm null} = \hat{C}_L^{\kappa g} (\mathrm{Cov}_{LL'}^{\kappa g})^{-1} \hat{C}_{L'}^{\kappa g} \tag{3.8}$$

Table 3.2: Result of no correlation hypothesis rejection for HELP patches using MV convergence map. $\nu = 7$ is the number of degrees of freedom.

Patch	$\chi^2_{\rm null}/\nu$	p-value
NGP	13.4/7	6.32×10^{-2}
HS-82	41.5/7	6.57×10^{-7}
SGP Part 1	16.9/7	1.80×10^{-2}
SGP Part 2	16.4/7	2.16×10^{-2}

The reduced chi square values and associated $p-$values of null hypothesis rejection for correlations between the four HELP fields with MV convergence map are quoted in Table 3.2. Although the $\chi^2_{\rm null}$ value for NGP field is poor, we detect significant cross-correlation signal from all HELP fields.

3.6.3 Parameters

We estimate redshift dependent galaxy linear bias, b and amplitude of cross-correlation A from our analysis using Maximum Likelihood Estimation technique (section 2.4). The amplitude of cross-correlation is expected to obtain a value of unity within the standard ΛCDM model. We account for the redshift dependence in the galaxy linear bias through a simple assumption

$$b(z) = \frac{b_0}{D(z)} \tag{3.9}$$

where b_0 is the galaxy linear bias parameter and $D(z)$ is the growth factor computed through the relation

$$D(z) = \exp\left\{ -\int_0^z \frac{[\Omega_m(z')]^\gamma}{1+z'} \mathrm{d}z' \right\} \tag{3.10}$$

where $\Omega_m(z)$ is the value of matter density parameter at redshift z and $\gamma = 0.55$ is the growth index for general relativity (Linder, 2005).

The form of redshift dependence assumed for the galaxy linear bias (Eq. 3.9) is consistent with other more common relation (Solarz et al. 2015; Moscardini et al. 1998; Fry 1996) given by:

$$b(z) = 1 + \frac{b_0 - 1}{D(z)} \tag{3.11}$$

We employ the publicly available EMCEE (Foreman-Mackey et al., 2013) package to effectively sample the parameter space, and estimate galaxy linear bias parameter b_0 and cross-correlation amplitude A. We use flat priors for the parameters with $b_0 \in [0, 10]$ and $A \in [-5, 5]$. The other cosmological parameters entering the likelihood estimation through theoretical power spectrum templates are set to their constant values for the fiducial background cosmology described in Planck Collaboration et al., 2020a (i.e. the flat ΛCDM cosmology with the best-fit *Planck* + *WP* + highL + lensing parameters, where *WP* stands for the *WMAP* polarisation data at low multipoles, highL is the high resolution CMB data from Atacama Cosmology Telescope (ACT) and South Pole Telescope (SPT), and lensing refers to the inclusion of *Planck* CMB ensing data in the parameter likelihood).

The covariance matrix entering the likelihood function is evaluated using Eq. (3.3). However, from Figure 3.8, we note that the covariance matrix is dominantly diagonal and hence, we use only the diagonal terms in likelihood estimation. We significantly reduce the computation time for sampling the parameter space with minor loss on the accuracy of estimated parameters by implementing the diagonal approximation on the covariance matrix.

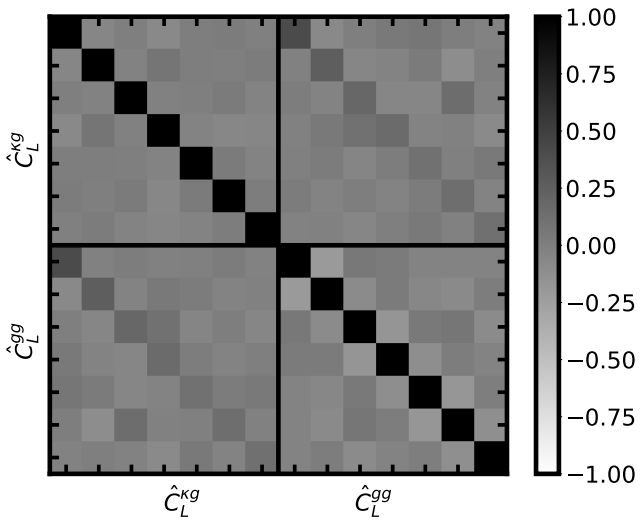
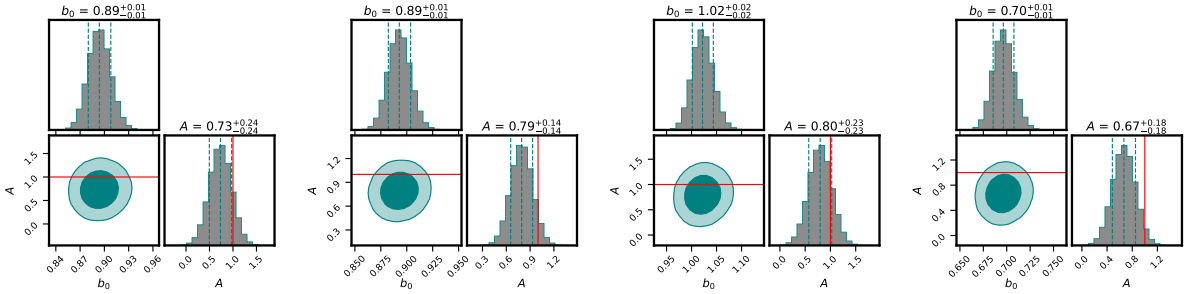


Figure 3.8: Correlation matrices constructed from the covariance matrix mentioned in Eq. (3.7) for HS-82 field. The correlation matrix is dominantly diagonal.

Table 3.3: Galaxy linear bias parameter and cross-correlation amplitude for HELP patches with MV lensing potential map using both separate and joint likelihood functions.

Patch	gg	κg		$\kappa g, gg$		χ^2/ν	p-value
		b_0	A	b_0	A		
NGP	$0.89^{+0.01}_{-0.01}$	$0.97^{+0.42}_{-0.26}$	$0.61^{+0.32}_{-0.30}$	$0.89^{+0.01}_{-0.01}$	$0.73^{+0.24}_{-0.24}$	6.5/5	0.262
HS-82	$0.89^{+0.01}_{-0.01}$	$1.08^{+0.45}_{-0.27}$	$0.71^{+0.28}_{-0.25}$	$0.89^{+0.01}_{-0.01}$	$0.79^{+0.14}_{-0.14}$	8.0/5	0.155
SGP Part-1	$1.02^{+0.02}_{-0.02}$	$0.45^{+0.20}_{-0.13}$	$2.16^{+0.89}_{-0.71}$	$1.02^{+0.02}_{-0.02}$	$0.80^{+0.23}_{-0.23}$	1.4/5	0.919
SGP Part-2	$0.70^{+0.01}_{-0.01}$	$0.26^{+0.14}_{-0.10}$	$1.97^{+1.15}_{-0.72}$	$0.70^{+0.01}_{-0.01}$	$0.67^{+0.18}_{-0.18}$	1.3/5	0.938

The best-fit values of the parameters estimated for our baseline analysis with MV lensing map are presented in Table 3.3. The errors on the parameters are computed as the 16^{th} and 84^{th} percentiles of the marginalised posterior distribution of the parameters. We also report the reduced chi-square values and $p-$values to give an idea for the goodness of fit of the theoretical power spectrum to observed data. We find agreement between the expected and estimated values of the cross-correlation amplitude within $\sim 1\,\sigma$ for NGP and SGP Part-1, within $\sim 1.5\,\sigma$ for HS-82 and within $\sim 2\,\sigma$ for SGP Part-2. In Figure 3.9, we show the marginalised posterior distribution for parameters b_0 and A as well as the $2-$dimensional posterior distribution in the (b_0, A)-plane for cross-correlation of the four HELP fields. The red lines mark the expected value of $A = 1$ within the ΛCDM model.



(a) NGP (b) HS-82 (c) SGP Part-1 (d) SGP Part-2

Figure 3.9: Posteriors of parameter obtained from Maximum Likelihood Estimation for all HELP patches with 68% and 95% confidence contours shown in darker and lighter shades, respectively. The three vertical lines are the median value of posterior and $\pm 1\sigma$ errors. The red line represents the value of $A = 1$ for standard ΛCDM model.

The galaxy linear bias parameter, on the other hand, is constrained entirely by the galaxy auto-power spectrum. This is owing to the fact that b_0 prefers the same value with joint-analysis of power spectra as the one with only galaxy auto-power spectrum. We observe significant variation in the value of b_0 parameter ranging from 0.70 for SGP Part-2 to 1.02 for SGP Part-1 field with less than 2% errors. The variation in the values of galaxy linear bias parameter can be related to the fact that with different HELP fields are composed of observations from different surveys. A detailed list of surveys and the fraction of objects they contribute to the HELP fields is present in Appendix 7.2. As mentioned in section 3.2.2, SGP Part-1 has coverage from KiDS survey while SGP Part-2 is covered by DES survey. This can give rise to the different values of b_0 between the two halves of the SGP field. On the other hand, the NGP and HS-82 fields have a large fraction of objects coming from the PanSTARRS survey, which can result in selection of similar objects and possibly explain the same value of $b_0 = 0.79$ from NGP and HS-82 fields.

3.7 Systematics

We find mild disparities in the amplitude of cross correlation when HS-82 and SGP Part-2 fields are cross-correlated with *Planck* MV lensing potential map. In this section, we explore other systematics that can contribute to a lower than expected value of cross-correlation amplitude.

3.7.1 CIB contamination

The Cosmic Infrared Background (CIB) emission is often considered as a potential source of bias when using CMB lensing data products for estimating cosmological parameters (Cao et al. 2020; van Engelen et al. 2014). The residuals of CIB emission can leak into the reconstruction of the CMB lensing potential through CMB temperature maps. The galaxies will, then, correlate with the residuals of CIB emission and may produce smaller values of the amplitude. Figure 3.10 taken from Planck Collaboration et al., 2020b shows the CIB-induced bias in the auto-power spectrum of the MV CMB lensing map (in orange line). As a result of using the `SMICA` CMB maps for lensing potential reconstruction in the *Planck* PDR3, the CIB-induced biases are at sub-percent level for the multipole range used in our analysis. Nevertheless, these CIB biases can

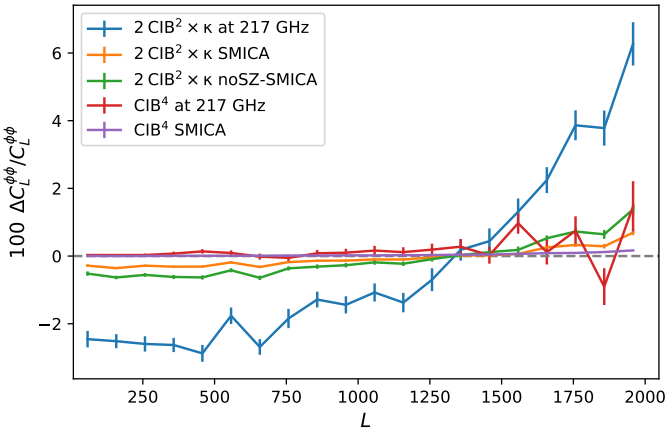


Figure 3.10: Estimates of the expected CIB-induced biases (as a fraction of the lensing potential power spectrum) taken from Planck Collaboration et al., 2020b. The orange line shows the relative CIB contribution for the MV CMB lensing maps.

be larger for the cross-power spectrum. Although, we did not estimate the contribution of CIB to the cross-correlation in this study, it is worthwhile to note that the redshift range of HELP galaxies is low compared to the redshift of CIB sources. Thus, we do not expect significant cross-correlation between HELP galaxies and sources of CIB emission.

3.7.2 Magnification Bias

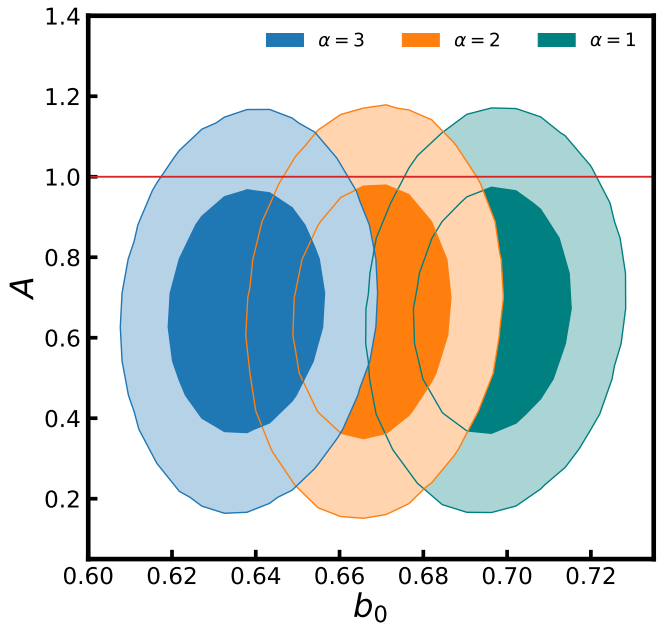


Figure 3.11: Effect of α on the estimated values of cross-correlation amplitude A and galaxy linear bias parameter b_0 for SGP Part-2. There is no significant effect on A.

The second term in Eq. (2.7) accounts for the modification of the observed density of background sources due to weak gravitational lensing by foreground objects. We measured the value of $\alpha = 1$ for all HELP patches used in our study by fitting a straight line to $\log N(>S)$ distribution. In this section we estimate the impact of α on parameters by looking at the change

in b_0 and A for SGP Part-2 field, with $\alpha = 2$ and $\alpha = 3$. The 2-dimensional posteriors in the $b_0 - A$ plane for $\alpha = 1,2,3$ are shown in Figure 3.11. We obtain a value of $b_0 = 0.67 \pm 0.01$ and $A = 0.67 \pm 0.18$ for $\alpha = 2$ and $b_0 = 0.64 \pm 0.01$ and $A = 0.67 \pm 0.18$ for $\alpha = 3$. The value of galaxy linear bias parameter b_0 decreases with higher values of α, but the amplitude of cross-correlation shows no effect. Other fields also show similar trends in the parameters with different α. We do not see any significant change in the cross-correlation amplitude because the cross-correlation with lensing convergence is weak for relatively shallow surveys like HELP. Hence, changes in the cross-power spectrum with different values of α are within errors on the data.

3.7.3 Median redshift

The median redshift of a distribution can be misestimated due to photometric redshift errors. We have imposed restrictions on the distribution of $\frac{\sigma_z}{1+z}$ to be ≤ 0.15 for NGP and HS-82, and ≤ 0.25 for SGP Part-1 and Part-2. These errors can shift the median values of the distributions (present in Table 3.1) from their true values. In this section, we examine the robustness of our results as presented in Table 3.3 against the misestimation of the median redshift of the distributions. For this test, we model the observed redshift distributions from the HELP fields by a function of the form

$$\frac{\mathrm{d}N}{\mathrm{d}z} = a_0 z^{a_1} \exp\left[-\left(\frac{z}{a_2}\right)^{a_3}\right] \tag{3.12}$$

where a_0, a_1, a_2, a_3 are free parameters of the function. The parameter a_2 controls the median redshift of the distribution. The parameter a_2 is not entirely the median redshift but acts as a proxy when shifting the median redshift of the distribution. The best-fit values of parameters a_0, a_1, a_2, a_3 is given in Table 3.4.

Table 3.4: Best fit values of a_0, a_1, a_2, a_3 for the modelling function given by Eq. (3.12).

Patch	a_0	a_1	a_2	a_3
NGP	5.843	1.007	0.602	3.014
HS-82	8.419	1.756	0.579	1.733
SGP Part-1	15.556	1.925	0.415	1.320
SGP Part-2	9.553	1.776	0.552	1.659

Using the functional form of Eq. (3.12), we estimate the minimum amount of shifts required in the median redshift, so that the amplitude of cross-correlation becomes fully consistent with its expectations. We find that for all the HELP fields we should shift the median value of redshift to lower values to account for the observed tension in the amplitude. We need $\sim 10\%$ shift for NGP and SGP Part-1, $\sim 20\%$ for HS-82 and $\sim 25\%$ for SGP Part-2 to alleviate the tension on amplitude of cross-correlation A. Figure 3.12 compares the marginalised and 2-dimensional posterior distributions, with and without shifting the median redshift by $\sim 25\%$, for SGP Part-2 field. The negative sign represents that the shift is towards lower values.

We also perform numerical simulations to get an idea of whether such large shifts in the median redshift is possible with the observed photometric redshift errors of the order of $\frac{\sigma_z}{1+z} \sim 0.2$. We assume true redshift distribution for the HELP fields having the functional form given by Eq. (3.12). For every HELP field, we generate true redshifts for objects drawn from the assumed

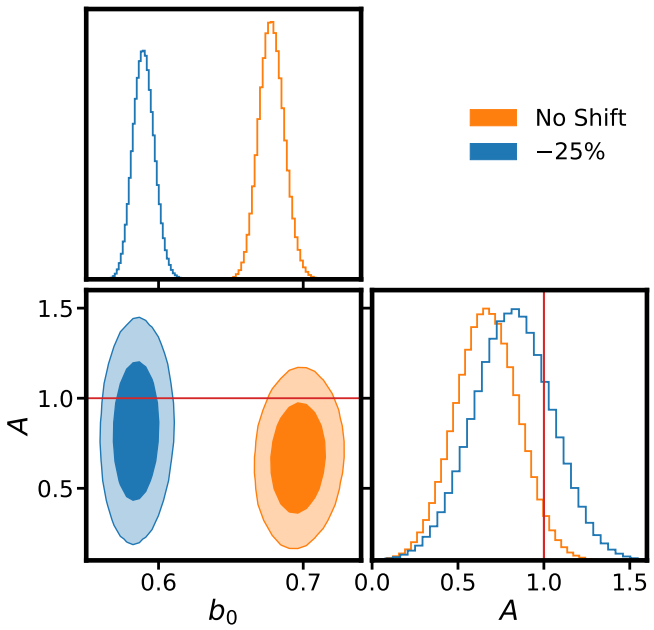


Figure 3.12: The parameters for SGP Part-2 computed with −25% shift in the median redshift of the distribution and without any shift. We show 68% and 95% contours with darker and lighter shaded regions, respectively. The red line represent the expected value $A = 1$.

true redshift distribution. We then redistribute the redshifts, assuming a Gaussian distribution centered at the true redshift and width corresponding to $\frac{\sigma_z}{1+z} \sim 0.2$. We repeat this exercise of redistribution of redshifts and generate a few hundred mock catalogues. Finally, we compare the average median redshift from the mock catalogues with the median redshift of the true redshift distribution. We find that the median redshift is shifted between $5-7\%$ for all HELP fields after taking into account the photometric redshift errors. We conclude that the shifts required to bring the cross-correlation amplitude in unison with unity are not plausible given the photometric redshift uncertainties of the order of $\frac{\sigma_z}{1+z} \sim 0.2$.

3.7.4 Photometric calibration errors

HELP catalogues are made from combination of various public surveys having different limiting magnitudes. The HELP catalogues are, thus, prone to photometric calibration errors which arise due to the variation in the limiting magnitude of surveys across sky. The photometric calibration error can cause the unphysical variations in the number density of objects, which can bias the galaxy auto-power spectrum and affect the estimation of parameters. Although, photometric calibration errors are significant at large angular scales, it may also have non-negligible contribution at small scales when combining different surveys together as in the case of HELP catalogues. Hence, mitigation of calibration errors is one of the most important and challenging aspect for cross-correlation analysis involving combination of different surveys.

To study the impact of the calibration errors on galaxy linear bias parameter and cross-correlation amplitude, we resort to numerical simulations and follow the procedure described in Huterer et al., 2013. The true number counts of objects in a given direction $\hat{\mathbf{n}}$ of the sky is transformed by the calibration field $c(\hat{\mathbf{n}})$ following the relation:

$$N_{\rm obs}(\hat{\mathbf{n}}) = [1 + c(\hat{\mathbf{n}})]\, N_{\rm true}(\hat{\mathbf{n}}) \tag{3.13}$$

in which

$$c(\hat{\mathbf{n}}) = \ln(10)\, s(z)\, \delta m(\hat{\mathbf{n}}) \tag{3.14}$$

where ln() represents natural logarithm, $\delta m(\hat{\mathbf{n}})$ is the variation in the limiting magnitude in a given waveband along the direction $\hat{\mathbf{n}}$, and $s(z)$ is the faint end slope of the luminosity function given by

$$s(z) \equiv \frac{{\rm d}\log_{10} N(z, > m)}{{\rm d}m}\Big|_{m_{\rm max}} \tag{3.15}$$

where $m_{\rm max}$ is the maximal apparent magnitude for a given waveband.

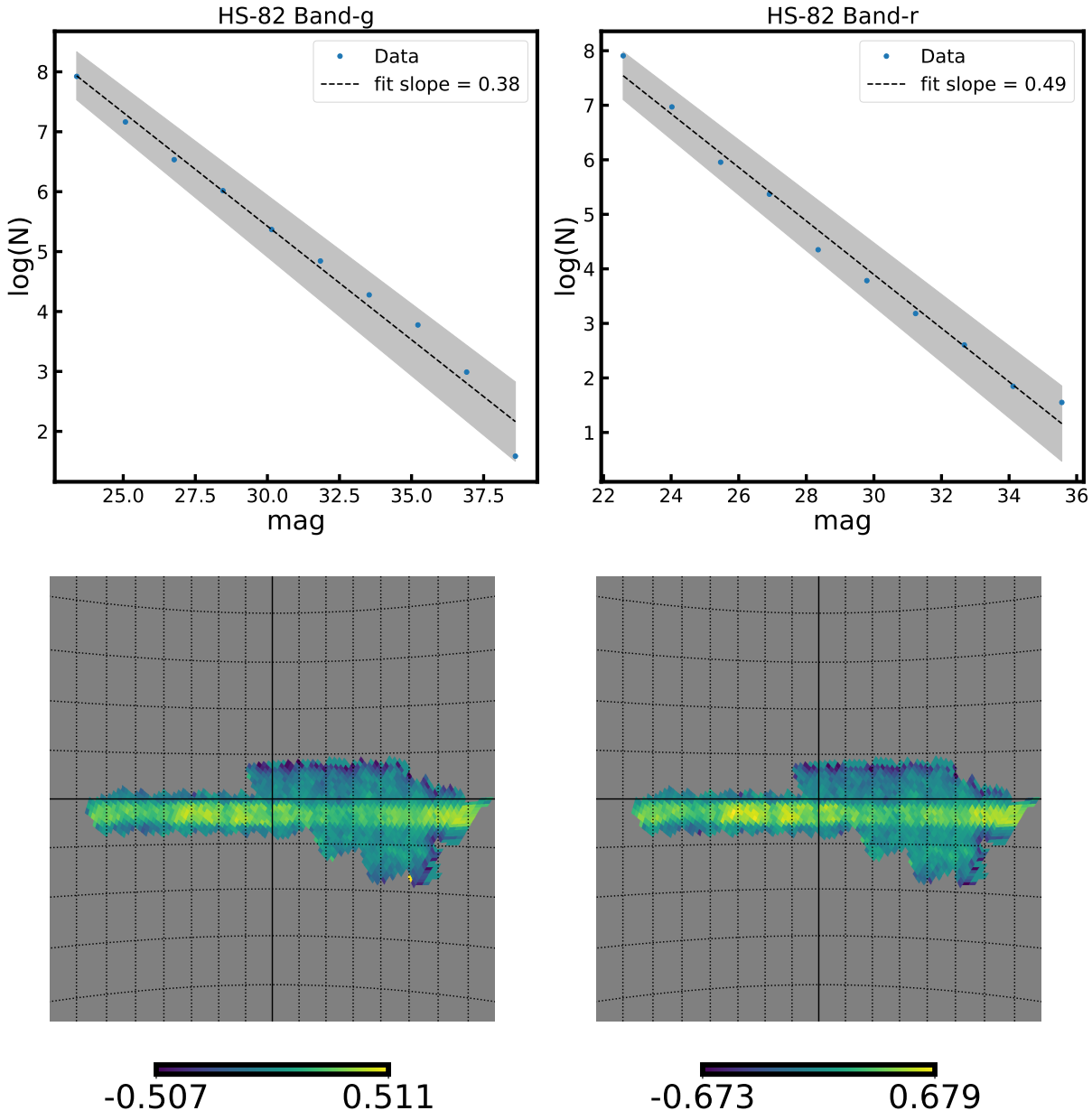


Figure 3.13: *Top*: Faint end slope of the luminosity function and *bottom*: calibration error field $c(\hat{\mathbf{n}})$ for HS-82 in g (*left*) and r (*right*) bands. The grey shaded area in the top panel marks the $1\,\sigma$ interval around the best-fit value of slope.

The correction for photometric calibration errors require a precise estimation of the calibration field $c(\hat{\mathbf{n}})$. We only have a tentative estimate for the $c(\hat{\mathbf{n}})$ field from observational data which may not completely reflect the impact of calibration errors. Thus, we quantify the impact of this systematic using only simulations. We notice from Table 7.1 that g and r bands have the maximum coverage in all HELP fields and hence, we present the influence of calibration errors using simulation of HS-82 field for g and r bands.

The faint end slope of the luminosity function and the corresponding calibration error field maps at HEALPix resolution parameter $N_{\rm side} = 128$ estimated from g and r bands for the HS-82 field are shown in the top and bottom panels of Figure 3.13, respectively. With an estimate of the calibration field, we simulate galaxy number count maps for HS-82 field induced with calibration error by using Eq. (3.13). Finally, we estimate the parameters from these simulations before and after correcting for the calibration errors, and show the comparison of estimated parameters for g and r bands in Figure 3.14. The red lines mark the true values of parameters used during simulations. Although the galaxy linear bias parameter shifts between $1-2\sigma$, the amplitude of cross-correlation remains unchanged. We observe similar behaviour in other HELP fields as well, suggesting that photometric calibration errors may not be significant contributor to the disparity observed in the amplitude.

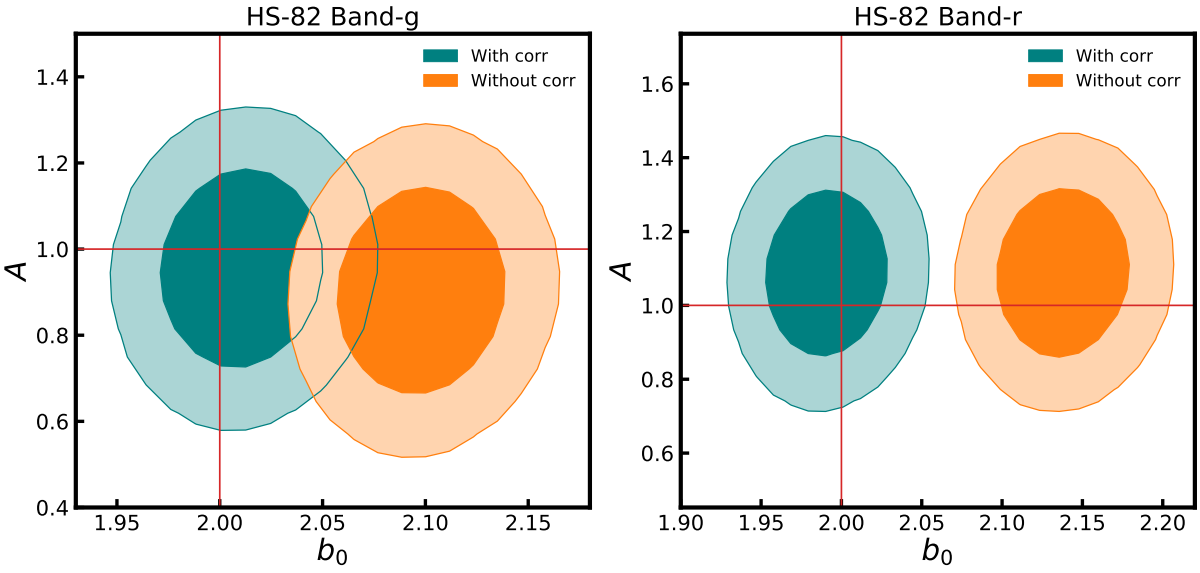


Figure 3.14: Comparison of estimated parameters with and without correction for the calibration error for g and r bands using simulations corresponding to HS-82 field. We show 68 % and 95 % contours with darker and lighter shaded regions, respectively. The red lines represent the true values of b_0 and A parameters used in simulations, i.e. $b_0 = 2$ and $A = 1$.

3.7.5 Catastrophic photometric redshift errors

The catastrophic photometric redshift errors causes the redshifts of galaxies in a photometric survey to be misestimated by a significant amount. The reasons for catastrophic redshift errors are not fully understood, but Muir and Huterer, 2016 relates the outcomes and occurrence of catastrophic errors to the number of photometric filters and their relation to the spectral features that carry principal information about the redshift. The percentage of catastrophic redshift error rates is most likely between $1-10\%$ for current surveys (see Fang et al. 2022; Jouvel et al. 2017;

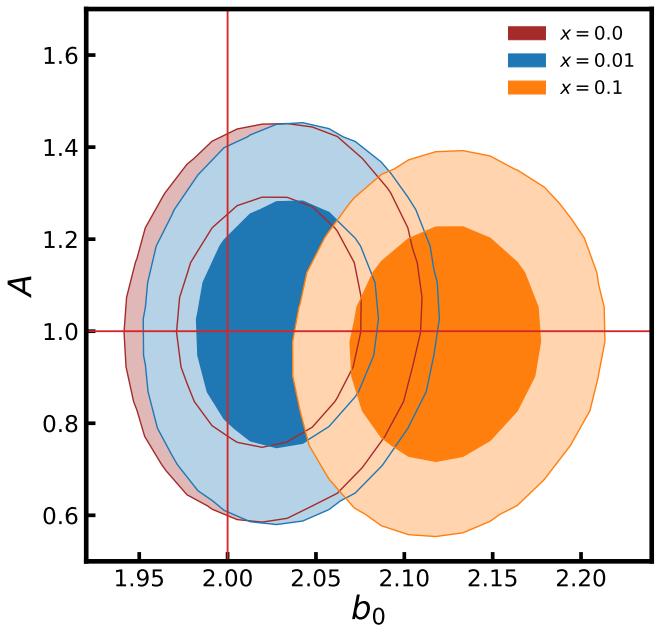


Figure 3.15: Effect of catastrophic errors on the inferred values of cross-correlation amplitude A and galaxy linear bias parameter b_0 for SGP Part-2. We show 68% and 95% contours with darker and lighter shaded regions, respectively, for $x = 0$ (no catastrophic errors), $x = 0.01$ and $x = 0.1$ catastrophic error rate. The red lines represent the true values of b_0 and A used in simulations ($b_0 = 2$ and $A = 1$).

Muir and Huterer 2016). We study the impact of the catastrophic redshift errors on HELP fields by assigning new random redshifts for fraction x of the galaxies from the redshift range $z \in [0.01, 3.0]$. We recompute the redshift distribution with the new assignment of redshifts and simulate the galaxy over-density maps. Figure 3.15 show the 2−dimensional contours in the $b_0 - A$ plane for simulations of SGP Part-2 field with $x = 0.01$ and 0.1. The parameters are estimated from the simulated over-density maps, with the theoretical power spectra computed from redshift distribution without accounting for the catastrophic errors. Figure 3.15 thus represents the parameters contours that one will estimate without correcting for the catastrophic errors. We do not find any systematic effect for the amplitude of cross-correlation for any of the HELP fields.

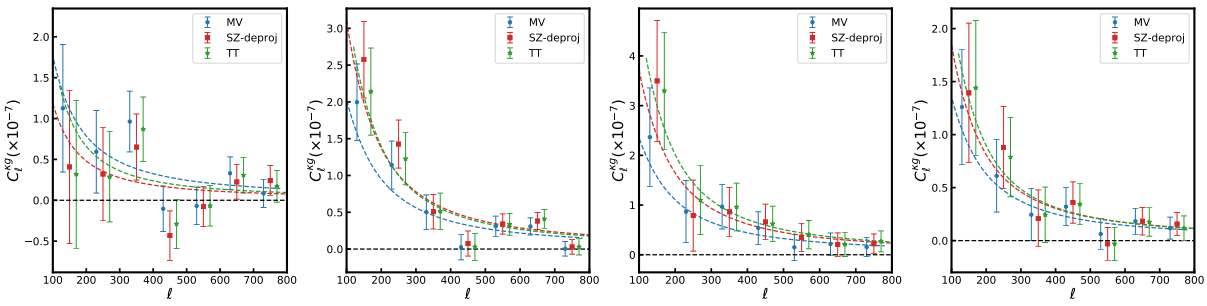
3.7.6 Using different CMB lensing potential maps

In addition to the systematics affecting in the galaxy density maps, we also study the impact of using different CMB lensing convergence maps provided in *Planck* PDR3 datasets (Planck Collaboration et al., 2020b). We use the lensing convergence map derived from the temperature-only Sunyaev-Zeldovich deprojected map (hereafter, SZ-deproj map) and the lensing convergence map reconstructed from only the CMB temperature measurements (TT map). In Table 3.5, we present the parameters b_0 and A, estimated from the MV, SZ-deproj and TT maps. Figure 3.16 show cross-power spectra for cross-correlation of HELP fields with MV, SZ-deproj and TT maps with best-fit theoretical power spectra. A re-estimation of parameters b_0 and A with the SZ-deproj and TT lensing maps produces larger amplitudes in agreement with expectations within $1\,\sigma$ for all HELP fields except NGP. The disparities in the value of the cross-correlation amplitude between MV and SZ-deproj or TT lensing maps can be due to the differences in these

maps. In particular, MV lensing map include CMB polarization measurements along with CMB temperature measurements to reconstruct the lensing potential map, whereas the SZ-deproj and TT maps are reconstructed from the CMB temperature measurements only, with additional subtraction of Sunyaev-Zeldovich signal in the case of SZ-deproj map.

Table 3.5: Galaxy linear bias and cross-correlation amplitude from Maximum Likelihood Estimation from HELP patches using joint likelihood functions for MV, SZ-deproj and TT convergence maps.

Patch	MV			SZ-deproj			TT		
	b	A	χ^2/ν	b	A	χ^2/ν	b	A	χ^2/ν
NGP	$0.89^{+0.01}_{-0.01}$	$0.73^{+0.24}_{-0.24}$	6.5/5	$0.89^{+0.01}_{-0.01}$	$0.39^{+0.27}_{-0.27}$	6.9/5	$0.89^{+0.01}_{-0.01}$	$0.47^{+0.27}_{-0.26}$	8.1/5
HS-82	$0.89^{+0.01}_{-0.01}$	$0.79^{+0.14}_{-0.14}$	8.0/5	$0.89^{+0.01}_{-0.01}$	$1.00^{+0.14}_{-0.14}$	10.5/5	$0.89^{+0.01}_{-0.01}$	$0.89^{+0.15}_{-0.15}$	10.3/5
SGP Part-1	$1.02^{+0.02}_{-0.02}$	$0.80^{+0.23}_{-0.23}$	1.4/5	$1.02^{+0.02}_{-0.02}$	$1.04^{+0.26}_{-0.26}$	1.3/5	$1.02^{+0.02}_{-0.02}$	$1.11^{+0.26}_{-0.26}$	0.9/5
SGP Part-2	$0.70^{+0.01}_{-0.01}$	$0.67^{+0.18}_{-0.18}$	1.3/5	$0.70^{+0.01}_{-0.01}$	$0.75^{+0.21}_{-0.21}$	3.3/5	$0.70^{+0.01}_{-0.01}$	$0.72^{+0.21}_{-0.21}$	3.1/5



NGP HS-82 SGP Part-1 SGP Part-2

Figure 3.16: cross-power spectra (*Bottom*) for cross-correlations of HELP fields with MV (blue cicles), SZ-deproj (red squares) and TT (green stars) maps. The colour dashed lines are their corresponding theoretical power spectra computed using the best-fit values obtained from likelihood analysis quoted in Table 3.5.

3.8 Summary

We have presented measurement of cross-correlation between the minimum-variance CMB lensing convergence map from *Planck* 2018 data release and galaxy catalogues from the *Herschel* Extragalactic Legacy Project. For our analysis, we have selected three of the largest and most uniform fields of the catalogue namely, NGP, HS-82, and SGP divided into two parts. The areas covered by these fields are: ~ 180 deg^2, ~ 255 deg^2, ~ 85 deg^2 and ~ 145 deg^2, respectively.

We have shown that for MV lensing map the no correlation hypothesis can be ruled out with a significance of about $1.7\,\sigma$ for NGP, $\sim 9.2\,\sigma$ for HS-82, $\sim 2.6\,\sigma$ for SGP Part-1 and $\sim 2.5\,\sigma$ for SGP Part-2 field. A joint analysis of galaxy auto-power spectrum and cross-power spectrum using Maximum Likelihood approach gives the galaxy linear bias parameter for different fields ranging from $b_0 = 0.70 \pm 0.01$ for SGP Part-2 to $b_0 = 1.02 \pm 0.02$ for SGP Part-1 field. The

cross-correlation amplitude varies from $A = 0.67 \pm 0.18$ for SGP Part-2 to $A = 0.80 \pm 0.23$ for SGP Part-1 field and a significance of its deviation from one is $1\,\sigma$ for NGP and SGP Part-1, $\sim 1.5\,\sigma$ for HS-82 and $\sim 2\,\sigma$ for SGP Part-2.

Though, a significance of the deviation for the MV lensing map is not very high, especially for NGP and SGP Part-1 fields, in all cases the amplitude is biased towards lower values suggesting that there is some systematic error in the analysis. To check it we have investigated some systematic errors that can account for this deviation, such as the effect of magnification bias caused by weak gravitational lensing and catastrophic photo-z errors which were found to have no notable improvement over the detected tension of cross-correlation amplitude A. We also examined the effect of shifting the estimated median redshift of HELP galaxies and concluded that a lower effective median redshift can increase the estimated value of amplitude A, suggesting that the HELP catalogue may be shallower than expected. However, the amount of shift required to remove the observed tension on the amplitude of cross-correlation, i.e. 20-25 % of the median redshift, is much larger than the potential offset related with photometric redshift errors, i.e. 5 % of the median. We also found out that variations across the fields in magnitude limits of the catalogue caused by photometric calibration errors has no significant effect on the cross-correlation amplitude. The amplitude turned out to be robust with respect to all studied systematic errors that can affect either the redshift distribution or can very the number counts of galaxies in the HELP survey.

Finally, we re-estimate parameters using two different CMB lensing convergence maps, the Sunyaev-Zeldovich deprojected map (SZ-deproj map) and the temperature-only reconstruction (TT) map. For both SZ-deproj and TT CMB lensing maps, the amplitude is higher than MV map by $\sim 1\,\sigma$ and consistent with one for all fields, except NGP field for which its value is lower than for the MV lensing map. These disparities can be explained by differences between MV and SZ-deproj or TT lensing maps, however, it also shows that we need a more robust estimation of the CMB lensing map for cross-correlation studies. We can expect that forthcoming CMB experiments and galaxy surveys will allow us to perform more robust and precise cross-correlation measurements in the future.

Chapter 4

Tomography studies with galaxy clustering and CMB lensing - impact of leakage correction

4.1 Overview

In Chapter 3, we discussed different systematics and their impact on testing cosmological model through cross-correlations between *Planck* CMB lensing potential and HELP galaxy catalogues. Future galaxy surveys like Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST, Ivezić et al. 2019; LSST Science Collaboration et al. 2009), Euclid (Laureijs et al., 2011), and Dark Energy Spectroscopic Instrument (DESI) (Fagrelius 2020; Dey et al. 2019b) will enjoy larger sky coverages (about half of the sky), increased depths and better estimation of photometric redshifts. With these surveys we will be able to perform cross-correlation tomography with galaxy samples divided into narrow redshift bins to map the evolution of the cosmological parameters with redshift. In this chapter, we delve into quantifying the impact of inevitable leakage of objects across redshift bins through Monte Carlo (MC) simulations of LSST galaxy survey and *Planck* CMB lensing convergence.

We describe the simulation setup in section 4.2, the reconstruction of true redshift galaxy distribution from observed photometric redshift galaxy distribution in section 4.3. We present the results before leakage correction in section 4.4 and discuss the correction for leakage across redshift bins through scattering matrix in section 4.5. In section 4.6, we quantify the impact of leakage on cosmological parameters and finally summarise the conclusions from the chapter in section 4.9

4.2 Simulation setup

We use the publicly available code `FLASK` (Xavier et al., 2016) to generate 300 tomographic MC realisations of correlated lognormal galaxy over-density and CMB lensing convergence fields. As mentioned in section 3.5, simulations of Gaussian galaxy over-density fields often lead to pixels with density $g < -1$, from which galaxy number count maps cannot be generated through Poisson sampling. Thus, to avoid negative Poisson sampling, we use `FLASK` code to simulate lognormal over-density fields. The simulated galaxy density follows LSST photometric redshift distribution profile (Ivezić et al. 2019; LSST Science Collaboration et al. 2009) with mean redshift 0.9 and mean surface number density of 40 arcmin^{-2}. The simulated CMB lensing convergence field is consistent with *Planck* observations (Planck Collaboration et al., 2020b). The galaxy density field is induced with Poisson noise and for CMB convergence we use the noise power spectrum provided in the *Planck* 2018 data package[1]. Due to limitations of computational capabilities, the sky area covered in our simulations is 2000 deg^2, however, the results obtained in this study should remain valid for the planned area of the LSST survey, provided the errors are appropriately scaled by the fraction of sky coverage. The mock galaxy samples are divided into 9 disjoint tomographic bins with redshift intervals $(0.0, 0.2, 0.4, 0.6, 0.8, 1.0, 1.4, 1.8, 2.2, 3.0]$, as marked by the dashed vertical lines in Figure 4.1. The fiducial angular power spectra for each redshift bin used to generate correlated maps are computed through Eq. (2.8) following Limber approximation (Limber, 1953). The matter power spectrum $P(k,z)$ is generated using the public software CAMB[2] (Lewis et al., 2000) and we assume $\alpha = 1$ in the galaxy kernel (Eq. 2.7), so that there is no effect of magnification bias on our simulations, leaving study of its impact for future work.

We use a redshift dependent model of galaxy bias (Solarz et al. 2015; Moscardini et al. 1998; Fry 1996):

$$b(z) = 1 + \frac{b_0 - 1}{D(z)} \tag{4.1}$$

where we take $b_0 \equiv b(z=0) = 1.3$ and $D(z)$ is the linear growth function normalised to unity at $z = 0$:

$$D(z) = \exp\left\{ -\int_0^z \frac{[\Omega_m(z')]^\gamma}{1+z'} \mathrm{d}z' \right\} \tag{4.2}$$

where $\gamma = 0.55$ is the growth index for the General Relativity (Linder, 2005).

For every simulation, `FLASK` code produces CMB convergence map and galaxy number count maps for each tomographic bin along with the catalogue of galaxy redshifts. We term these maps as true datasets. We generate photometric redshifts, z_p, for galaxies in every simulation by drawing positive random values from Gaussian distributions with their true redshift z_t as the mean and standard deviation $\sigma_0(1+z_t)$. We adopt two different values of σ_0: 0.02 and 0.05, to study the dependence of our results on the strength of redshift scatters. The galaxies are again divided into 9 tomographic bins based on their photometric redshifts. In Figure 4.1, we show the true LSST redshift distribution (blue solid curve) divided into 9 photometric redshift bins

[1] https://pla.esac.esa.int/#cosmology
[2] https://camb.info/

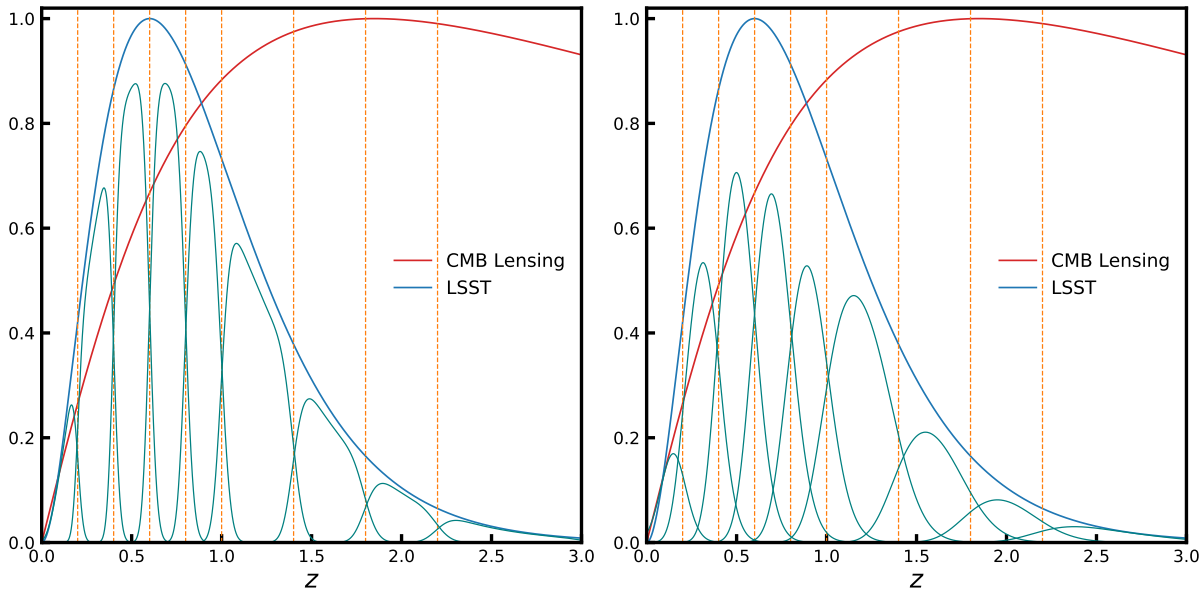


Figure 4.1: True LSST redshift distribution (blue solid curve) divided into 9 photometric redshift bins (shown by solid green curves) for $\sigma_0 = 0.02$ (*left*) and $\sigma_0 = 0.05$ (*right*). The red line denotes the CMB lensing kernel and the orange dashed vertical lines mark the true redshift bins. The full LSST redshift distribution and CMB lensing kernel are normalized to unit maximum.

(solid green lines) and the CMB lensing kernel (red solid curve) in the redshift range $0 \leq z \leq 3$. The green lines represent how the disjoint true redshift bins transform after introducing the photometric redshift errors.

We build galaxy over-density maps from photometric number count maps (hereafter, photometric datasets) with HEALPix[3] (Górski et al., 2005) resolution parameter $N_{\rm side} = 1024$ using Eq. (3.1)

$$g(\hat{\mathbf{n}}) = \frac{n(\hat{\mathbf{n}}) - \overline{n}}{\overline{n}}$$

where $n(\hat{\mathbf{n}})$ is the number of galaxies at angular position $\hat{\mathbf{n}}$ and $\overline{n}$ is the mean number of galaxies per pixel. In Table 4.1, we present from one realisation the comparison of the mean number of objects per pixel (for $N_{\rm side} = 1024$) and median redshift between true and photometric datasets.

For the simulations and analyses presented in this paper, we adopt the flat ΛCDM cosmology with best-fit *Planck* + *WP* + highL + lensing cosmological parameters, as described in Planck Collaboration et al., 2020a. Here, *WP* refers to *WMAP* polarisation data at low multipoles, highL is the high-resolution CMB data from Atacama Cosmology Telescope (ACT), and South Pole Telescope (SPT) and lensing refer to the inclusion of *Planck* CMB lensing data in the parameter likelihood.

4.3 Estimation of true redshift distribution

The true redshift distribution $\frac{{\rm d}N(z_t)}{{\rm d}z_t}$ is estimated from the observed photometric redshift distribution $\frac{{\rm d}N(z_p)}{{\rm d}z_p}$ and some quantification of errors on the photometric redshifts (from cross-validation

[3] https://healpix.jpl.nasa.gov/

Table 4.1: Comparison of physical properties between true and photometric datasets for all tomographic bins. $\bar{n}$ is the mean number of objects per pixel and $\bar{z}$ represents the median redshift of the tomographic bin.

z	$\bar{n}$ (true)	$\bar{n}$ (photo)		$\bar{z}$ (true)	$\bar{z}$ (photo)	
		$\sigma_0 = 0.02$	$\sigma_0 = 0.05$		$\sigma_0 = 0.02$	$\sigma_0 = 0.05$
$[0.0, 0.2)$	14.35	14.77	17.02	0.144	0.144	0.145
$[0.2, 0.4)$	57.02	57.04	57.11	0.311	0.312	0.315
$[0.4, 0.6)$	81.78	81.60	80.68	0.502	0.503	0.506
$[0.6, 0.8)$	82.79	82.58	81.54	0.698	0.699	0.702
$[0.8, 1.0)$	70.58	70.42	69.63	0.896	0.898	0.900
$[1.0, 1.4)$	93.31	93.19	92.60	1.178	1.180	1.183
$[1.4, 1.8)$	44.37	44.40	44.57	1.572	1.575	1.577
$[1.8, 2.2)$	18.44	18.50	18.83	1.968	1.972	1.974
$[2.2, 3.0)$	9.58	9.68	10.22	2.472	2.471	2.471

with some spectroscopic survey or posteriors from machine learning methods). Often these errors are expressed by conditional probabilities $p(z_p - z_t|z_t)$ and $p(z_t - z_p|z_p)$. The true and photometric redshift distributions are then related through (Sheth & Rossi, 2010)

$$\frac{\mathrm{d}N(z_t, z_p)}{\mathrm{d}z_t \mathrm{d}z_p} = \frac{\mathrm{d}N(z_t)}{\mathrm{d}z_t} p(z_p - z_p|z_t) = \frac{\mathrm{d}N(z_p)}{\mathrm{d}z_p} p(z_t - z_p|z_p) \tag{4.3}$$

Thus depending on whether we have estimates of $p(z_p - z_t|z_t)$ or $p(z_t - z_p|z_p)$, the method for estimation of the true redshift distribution is called deconvolution or convolution, respectively.

4.3.1 Convolution method

When $p(z_t - z_p|z_p)$ is known, we can estimate $\frac{\mathrm{d}N(z_t)}{\mathrm{d}z_t}$ for each tomographic bin i using convolution

$$\frac{\mathrm{d}N^i(z_t)}{\mathrm{d}z_t} = \int \mathrm{d}z_p \frac{\mathrm{d}N^i(z_p)}{\mathrm{d}z_p} p^i(z_t - z_p|z_p) \tag{4.4}$$

where $\frac{\mathrm{d}N^i(z_p)}{\mathrm{d}z_p}$ is the photometric redshift distribution of objects for the i^{th} redshift bin. Generally, $p^i(z_t - z_p|z_p)$ is fitted with parametric functions like Gaussian with assumed zero mean (Sun et al. 2022; Marques and Bernui 2020) or modified Lorentzian with zero mean (Hang et al. 2021; Peacock and Bilicki 2018). Sheth and Rossi, 2010 have shown that the quantity $p(z_t - z_p|z_p)$ will be biased and not centered on zero. In our study, we fit the error distribution with a sum of three Gaussians

$$\mathcal{N}(x) = \sum_{i=1}^{3} A_i \exp\left[\frac{(x-\mu_i)^2}{2\sigma_i^2}\right] \tag{4.5}$$

where A, μ, σ control the amplitude, mean and width of the individual Gaussians. The sum of Gaussians can account for the bias in $p(z_t - z_p|z_p)$, as well as other characteristic features of the error distributions like non-Gaussian wings and higher peak in the center. We also check the fit with a higher number of Gaussians but do not find any improvement in the quality of fit beyond three Gaussians. For each of the tomographic bins, we fit for A_i, μ_i, σ_i, and estimate the true redshift distribution using Eq. (4.4).

4.3.2 Deconvolution method

The true redshift distribution $\frac{dN(z_t)}{dz_t}$ can be estimated by a deconvolution method when $p(z_p - z_t|z_t)$ is known:

$$\frac{dN(z_p)}{dz_p} = \int dz_t \frac{dN(z_t)}{dz_t} p(z_p - z_t|z_t) \tag{4.6}$$

We fit $p(z_p - z_t|z_t)$ with a single Gaussian and find the mean to be consistent with zero, in agreement with the unbiased nature of $p(z_p - z_t|z_t)$. We also fit the error distribution with a sum of Gaussians to find no significant improvement in the fit quality. Padmanabhan et al., 2005 proposed a deconvolution method based on Tikhonov regularisation, which lacks a general method to quantify the impact of penalty function on the reconstruction of $\frac{dN(z_t)}{dz_t}$. We use a different approach to deconvolution. It is based on the convolution theorem and kernel-based regularisation (Meister, 2009). The true redshift distribution in our approach is estimated as

$$\frac{dN(z_t)}{dz_t} = \mathscr{F}^{-1}\left[\frac{\mathscr{F}\left[\frac{dN(z_p)}{dz_p}\right]}{\mathscr{F}\left[p(z_p - z_t|z_t)\right]}\right] \tag{4.7}$$

where $\mathscr{F}$ and $\mathscr{F}^{-1}$ represents the Fourier and inverse Fourier transforms, respectively. We estimate the true redshift distribution for the entire redshift range, $0 < z \leq 3$, then the true redshift distribution for each tomographic bin, i, can be expressed as

$$\frac{dN^i}{dz_t} = \frac{dN}{dz_t} W^i(z_t) \tag{4.8}$$

where $W^i(z_t)$ is the Heaviside step function defined as

$$W^i(z_t) = \begin{cases} 1, & \text{if} \quad z^i_{\text{min}} \leq z_t < z^{i+1}_{\text{min}} \\ 0, & \text{otherwise} \end{cases} \tag{4.9}$$

The corresponding photometric redshift distribution for every tomographic bin i is given by

$$\frac{dN^i(z_p)}{dz_p} = \int dz_t \frac{dN^i(z_t)}{dz_t} p^i(z_p - z_t|z_t) \tag{4.10}$$

where $p^i(z_p - z_t|z_t)$ is the error distribution for bin i. Eq. (4.6) will not follow convolution strictly near $z = 0$, since negative redshifts are unphysical. Due to this fact, the reconstructed true redshift distribution will be inaccurate close to redshift $z = 0$, and we expect these inaccuracies to affect to some extent the first two tomographic bins.

4.4 Results

In this section, we present results of estimating true redshift distribution and power spectra for datasets simulated by the FLASK code. We check for any systematics in the estimation for datasets without photometric redshift errors as well as with added errors. The power spectra from the simulated maps are computed from the full-sky estimator based on `MASTER` algorithm described in the section 3.3. We estimate all full-sky power spectra in linearly spaced multipole bins between $50 \leq \ell \leq 1500$ with binwidth $\Delta\ell = 30$.

4.4.1 Estimation of the true redshift distribution

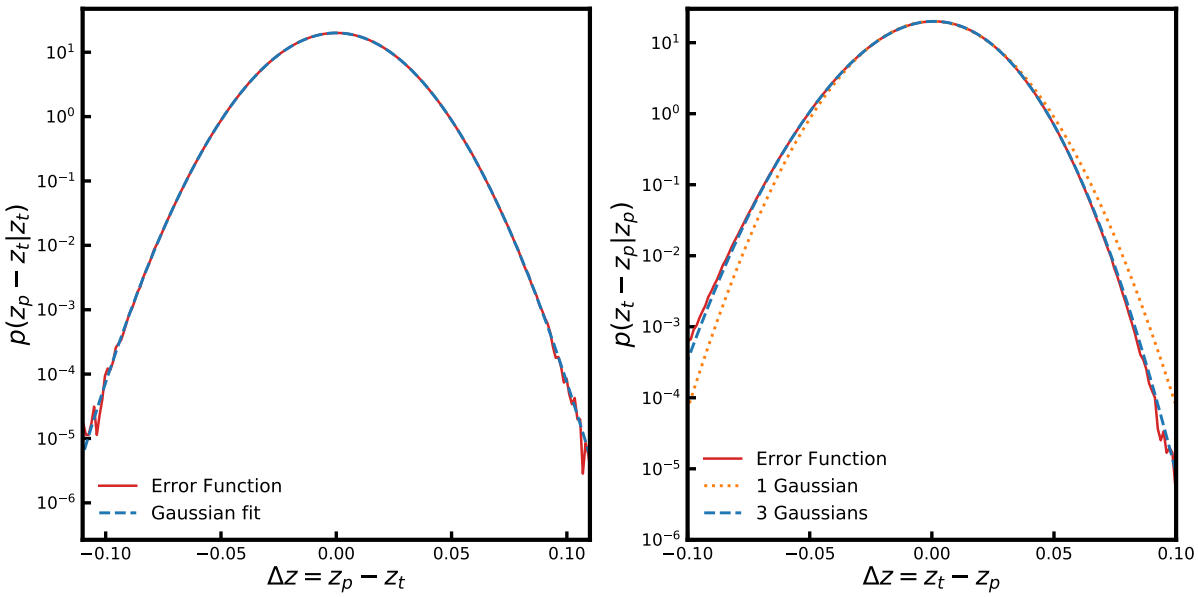


Figure 4.2: *Left*: Fit made to the error function $p(z_p - z_t|z_t)$ (*red solid line*) using single Gaussian (*blue dashed line*). *Right*: Fit made to the error function $p(z_t - z_p|z_p)$ (*red solid line*) using single Gaussian (*orange dotted line*) and sum of three Gaussians (*blue dashed line*).

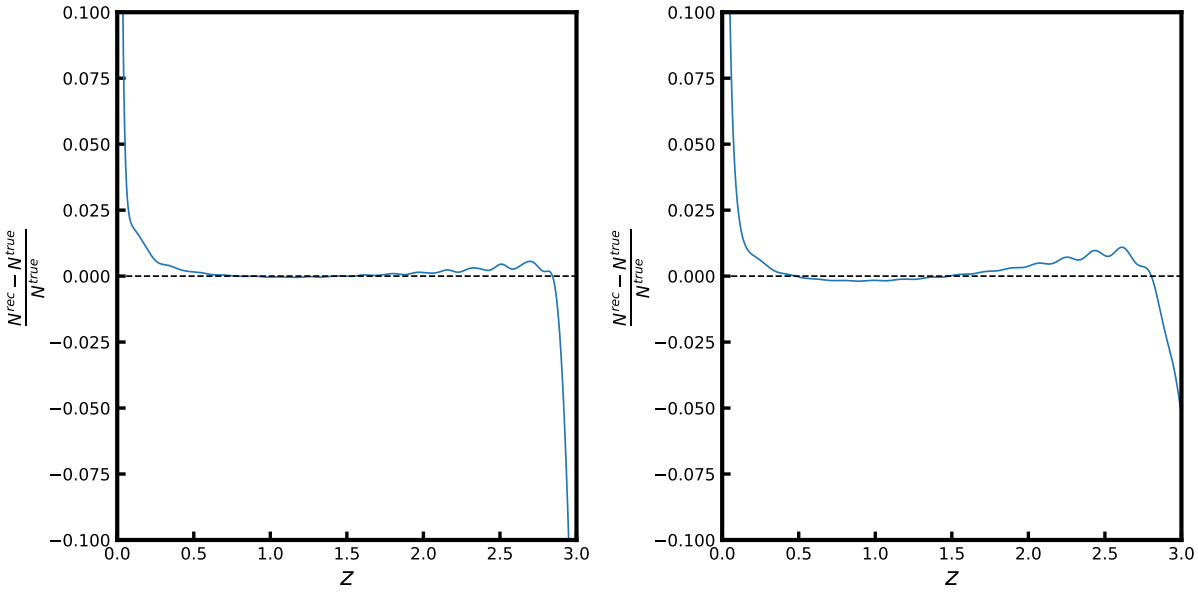


Figure 4.3: Average relative difference between the redshift distribution reconstructed using deconvolution method and the true distribution for 300 simulations with *left*: $\sigma_0 = 0.02$ and *right*: $\sigma_0 = 0.05$. The shaded region represents the fluctuations in reconstruction between 300 simulations.

The estimation of the true redshift distribution requires fitting the error function (i.e. either $p(z_t - z_p|z_p)$ or $p(z_p - z_t|z_t)$) with parametric functions. We have used a single Gaussian for $p(z_p - z_t|z_t)$ and a sum of three Gaussian to fit $p(z_t - z_p|z_p)$. Figure 4.2 shows the quality of fit to the error functions with $\sigma_0 = 0.02$ and compares single Gaussian versus three Gaussians fit of $p(z_t - z_p|z_p)$. The sum of three Gaussians provides a significantly better fit by accurately

capturing the non-Gaussian wings of the error function. The uncertainties in these fits remain within 0.25%, and we do not expect these sub-per cent uncertainties to bias the power spectra or estimation of parameters.

In Figure 4.3, we show the relative error, averaged over 300 simulations, of the redshift distribution reconstructed using the deconvolution method for the two cases of $\sigma_0 = 0.02$ and 0.05. The reconstructed redshift distribution is within $\sim 2\%$ for the entire redshift range with maximum deviations occurring near $z = 0$ possibly due to sharp cuts in the redshift distribution near the boundaries at $z = 0$ and $z = 3$ (section 4.3.2).

4.4.2 Validation from true datasets

Before adding photometric redshift errors in the simulations generated by the `FLASK` code, we verify the validity of this code for simulations without added errors. In the top panel of Figure 4.4 we show, for every tomographic bin, the relative difference between noise-subtracted average galaxy auto-power spectra and their theoretical expectation with error bars computed using

$$\Delta\overline{C}_L^{xy} = \left(\frac{\mathrm{Cov}_{LL}^{xy}}{N_{\mathrm{sim}}}\right)^{1/2} \tag{4.11}$$

where Cov_{LL}^{xy} is the covariance matrix constructed from simulations following Eq. (3.7) and N_{sim} is the number of simulations.

The bottom panel of Figure 4.4 and Figure 4.5 present the relative differences for the cross-power spectrum and CMB convergence auto-power spectrum, respectively. The noise from galaxy density and CMB convergence fields are uncorrelated, and there will be no noise associated with their cross-power spectrum. Their power spectra are consistent with their theoretical expectations and thus, the `FLASK` simulations are free from any internal systematics.

4.4.3 Power spectra from photometric datasets

We compute the power spectra from simulated photometric datasets with the `MASTER` method described in section 2.3. The relative difference between the noise-subtracted average galaxy auto-power spectra ($\overline{C}_\ell^{gg}$) estimated from 300 simulations and the theoretical power spectra is shown in the top panel of Figure 4.6. The blue circles present the offsets for photometric datasets with $\sigma_0 = 0.02$ and the green squares are for $\sigma_0 = 0.05$. The bottom panel of Figure 4.6 show the relative differences in the cross-power spectra between CMB lensing and galaxy over-density ($\overline{C}_\ell^{\kappa g}$). The theoretical galaxy auto-power spectra and cross-power spectra are computed using the redshift distributions estimated by the convolution method (Eq. 4.4) described in section 4.3.1. The estimated galaxy auto-power spectra are smaller than expectations in every tomographic bin, with offsets varying between $2-15\%$ for $\sigma_0 = 0.02$ and between $15-40\%$ for $\sigma_0 = 0.05$. The cross-power spectra comparatively smaller biases, i.e. $< 5\%$ in every tomographic bin for both $\sigma_0 = 0.02$ and $\sigma_0 = 0.05$. We find similar offsets when the true redshift distributions (and hence the theoretical power spectra) are computed using the deconvolution method (following Eq. 4.10). This shows that the offsets in the power spectra are not related with the

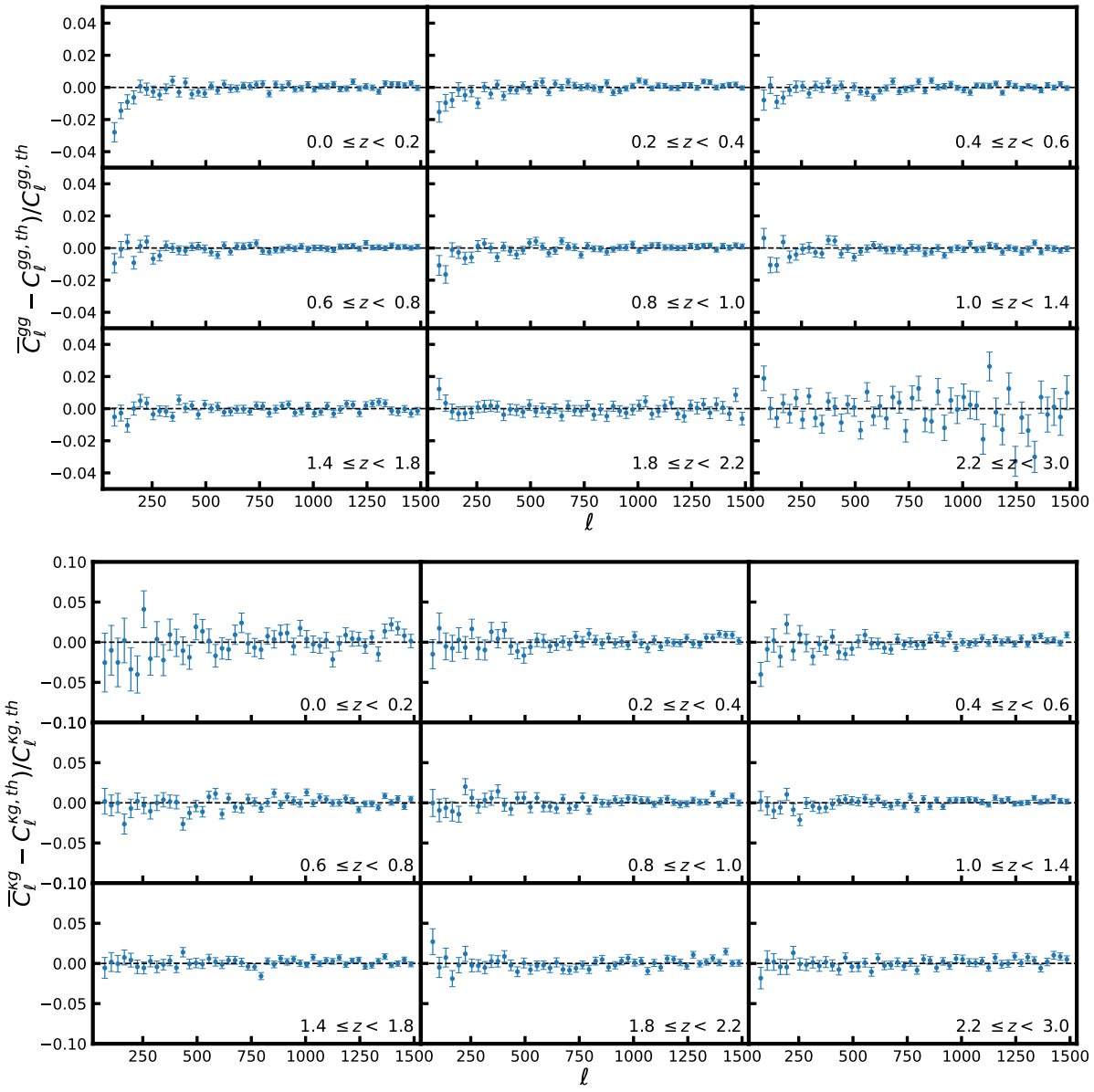


Figure 4.4: Relative errors in (*top:*) the average galaxy auto-power spectrum and (*bottom:*) the cross power spectrum, reconstructed from 300 simulations generated by the `FLASK` code without adding photometric redshift errors. The error bars on the data points are estimated from 300 simulations using Eq. 4.11.

deconvolution method. Furthermore, since larger photometric redshift scatters lead to larger deviations in the power spectra, this confirms that the origin of these offsets is rooted in the leakage of objects from one redshift bin to the other due to photometric redshift errors. These deviations will also impact estimation of cosmological parameters from the power spectra.

4.5 Leakage correction through scattering matrix

Due to errors in photometric redshifts, a fraction of galaxies observed in a photometric redshift bin come from other redshift bins. The leakage of objects across redshift bins changes the strength of correlation in a tomographic analysis as well as result in non-zero correlation between different redshift bins. In this section, we attempt to counter the leakage of objects across redshift

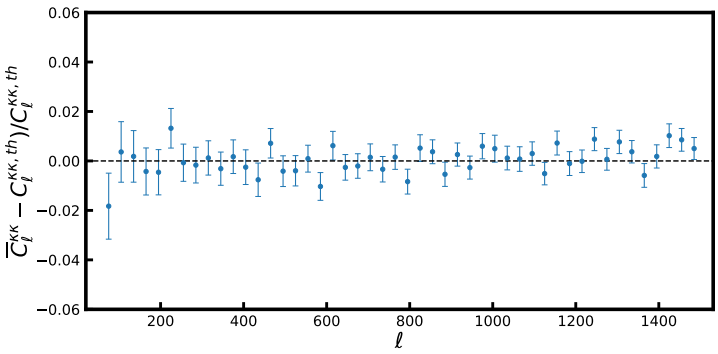


Figure 4.5: Relative errors in the average CMB convergence auto-power spectrum reconstructed from 300 simulations. The error bars on the data points are estimated from 300 simulations using Eq. 4.11.

bins through scattering matrix. If we divide galaxies into n tomographic bins, then Zhang et al., 2010 have shown that galaxy auto-power spectrum between $i^{\rm th}$ and $j^{\rm th}$ photometric bins, $C_{ij}^{gg,{\rm ph}}$, are related to the galaxy auto-power spectra from true redshift bins $C^{gg,{\rm tr}}$ by

$$C_{ij}^{gg,{\rm ph}}(\ell) = \sum_k P_{ki} P_{kj} C_{kk}^{gg,{\rm tr}}(\ell) \tag{4.12}$$

when there are no cross-correlation between true redshift bins. Eq. (4.12) can be generalized for the case with true redshift bins having non-zero correlations, however, using disjoint true redshift bins significantly reduces the complexity. The elements of the scattering matrix P_{ij} are defined as the ratio $N_{i\to j}/N_j^{\rm ph}$, where $N_{i\to j}$ is the number of galaxies moving from $i^{\rm th}$ true redshift bin to $j^{\rm th}$ photometric bin and $N_j^{\rm ph}$ is the total number of galaxies in the $j^{\rm th}$ photometric bin. This definition also produces a natural normalisation $\sum_i P_{ij} = 1$. A similar relation for the cross-power spectra between galaxy over-density in photometric redshift bin i and CMB lensing convergence can be obtained as

$$C_i^{\kappa g,{\rm ph}}(\ell) = \sum_k P_{ki} C_{kk}^{\kappa g,{\rm tr}}(\ell) \tag{4.13}$$

If we collect P_{ij} as elements of the scattering matrix P, then we can compactly write

$$C^{gg,{\rm ph}} = P^{\rm T} C^{gg,{\rm tr}} P \tag{4.14}$$

$$C^{\kappa g,{\rm ph}} = P^{\rm T} C^{\kappa g,{\rm tr}} \tag{4.15}$$

where $P^{\rm T}$ denotes the transpose of matrix P. Eqs. (4.14-4.15) show that the redistribution of galaxies across redshift bins due to photometric redshift errors results is a non-trivial relation between the photometric and true power spectra, weighted by the elements of the scattering matrix. Thus to properly mitigate the effects of leakage on power spectra, a precise estimation of the scattering matrix is necessary.

Zhang et al., 2017 proposed an algorithm to solve problems similar to Eq. (4.14) based on Non-negative Matrix Factorization (NMF) method, that simultaneously approximates the matrices P and $C^{gg,{\rm tr}}$. However, the NMF method proves to be computationally challenging for cases with a large number of tomographic and multipole bins. Here we propose an alternative

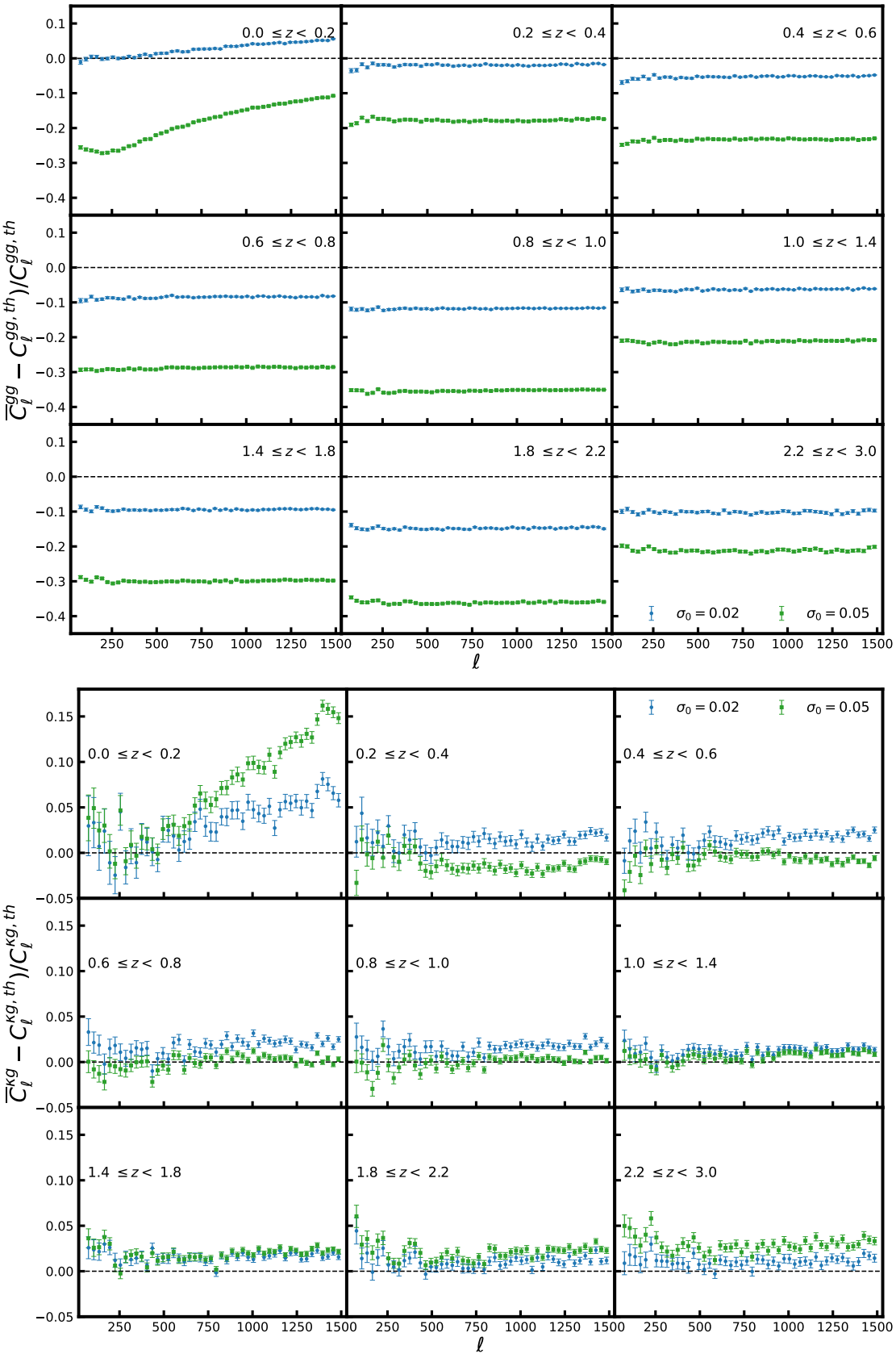


Figure 4.6: Relative errors in (*top:*) the average galaxy auto-power spectrum and (*bottom:*) the cross power spectrum, computed from 300 simulations after adding photometric redshift errors. The error bars on the data points are estimated from 300 simulations using Eq. 4.11.

method for fast and efficient computation of the scattering matrix based on the true and photometric redshift distributions. We first estimate the true redshift distribution, $\frac{dN(z_t)}{dz_t}$, for the entire redshift range following Eq. (4.7) and then we use Eq. (4.10) to compute the redshift distribution $\frac{dN^i(z_p)}{dz_p}$ for every tomographic bin i. The elements of the scattering matrix P_{ij} can then be computed directly by using the relation

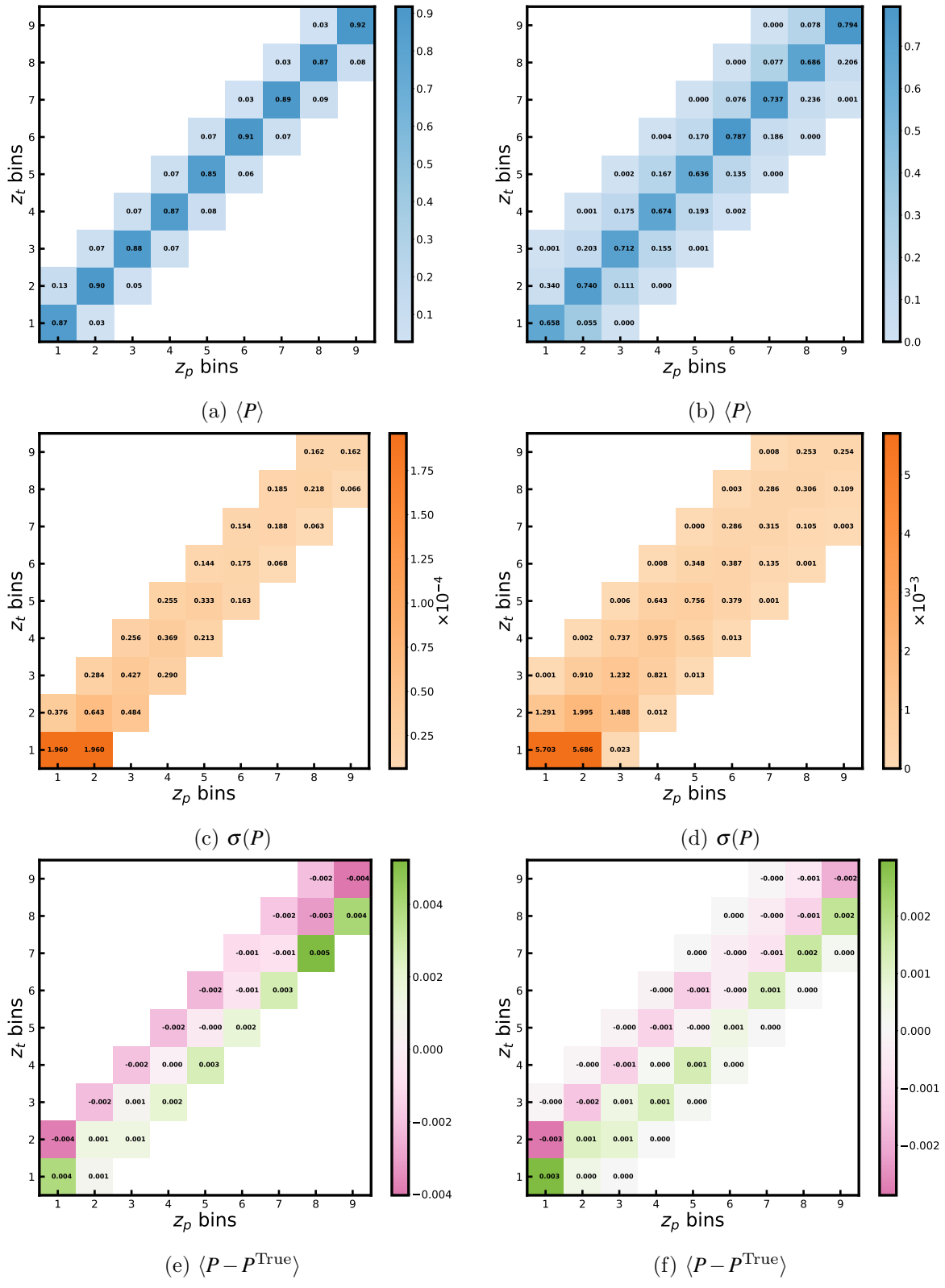
$$P_{ij} = \frac{\int_{z^j_{min}}^{z^{j+1}_{min}} dz_p \frac{dN^i}{dz_p}}{\int_{z^j_{min}}^{z^{j+1}_{min}} dz_p \frac{dN}{dz_p}} \tag{4.16}$$

where $\frac{dN}{dz_p}$ is the observed photometric redshift distribution of galaxies and z^j_{min} is the lower limit of the j^{th} redshift bin. Our method of computing the elements of the scattering matrix is significantly faster than the NMF method, and is only subject to accurate estimation of the true redshift distribution.

Figure 4.7 shows the performance of estimation of the scattering matrix from our proposed method using distributions for the two cases of $\sigma_0 = 0.02$ and 0.05. The average value of the scattering matrix $\langle P \rangle$ and its standard deviation $\sigma(P)$ averaged over 300 simulations are shown in top and middle panels of Figure 4.7, respectively. We note that the scattering matrix elements corresponding to the first true redshift bin have the maximum standard deviation. This behaviour is expected as the objects near $z = 0$ do not strictly follow convolution as discussed in section 4.3.2. The accuracy of the estimation of the scattering matrix can be verified using true scattering matrix P^{True} computed based on exactly counting the number of objects moving from one redshift bin to the other. In the bottom panel of Figure 4.7, we show the difference between the scattering matrix computed from our method and P^{True}, averaged over 300 realisations. We find that $|P - P^{\text{True}}| < 0.006$ for all elements of the scattering matrix, with maximum differences occurring in the first and last tomographic bins, i.e. near the boundaries of the redshift range simulated in our analysis. Hence, the overall precision and accuracy in the estimation of the scattering matrix is found to be good enough and the estimate of the scattering matrix can be used for correcting bin mismatch leakage for the power spectra.

Given an estimate of the true redshift distribution deconvoluted from the observed photometric redshift distribution and scattering matrix, the impact of leakage of the objects can be corrected in two ways: either by transforming true theoretical power spectra $C^{\text{th,tr}}$ to $C^{\text{th,ph}}$ using Eqs. (4.14) and (4.15) and comparing it to the estimated photometric power spectra $\overline{C}^{\text{ph}}$, or by inverting Eqs. (4.14) and (4.15) to transform the estimated photometric power spectra $\overline{C}^{\text{ph}}$ to true power spectra $\overline{C}^{\text{tr}}$ and comparing it to the theoretical true power spectra $C^{\text{th,tr}}$. We use the former approach for figures showing comparison of power spectra while the latter for estimation of cosmological parameters.

In Figure 4.8 we show the noise-subtracted average galaxy auto-power spectra $\overline{C}^{gg}_\ell$ and cross-power spectra $\overline{C}^{\kappa g}_\ell$ for photometric redshift scatter $\sigma_0 = 0.02$. The red solid lines are the theoret-



(a) $\langle P \rangle$ (b) $\langle P \rangle$

(c) $\sigma(P)$ (d) $\sigma(P)$

(e) $\langle P - P^{\mathrm{True}} \rangle$ (f) $\langle P - P^{\mathrm{True}} \rangle$

Figure 4.7: Performance of scattering matrix computed with the method described in section 4.5. *Top*: The average scattering matrix computed from 300 realisations, *middle*: its standard deviation, and *bottom*: the difference between scattering matrix computed from our method and true scattering matrix computed from simulated catalogues. *Left column*: scattering matrix estimated for $\sigma_0 = 0.02$; *Right column*: scattering matrix estimated for $\sigma_0 = 0.05$.

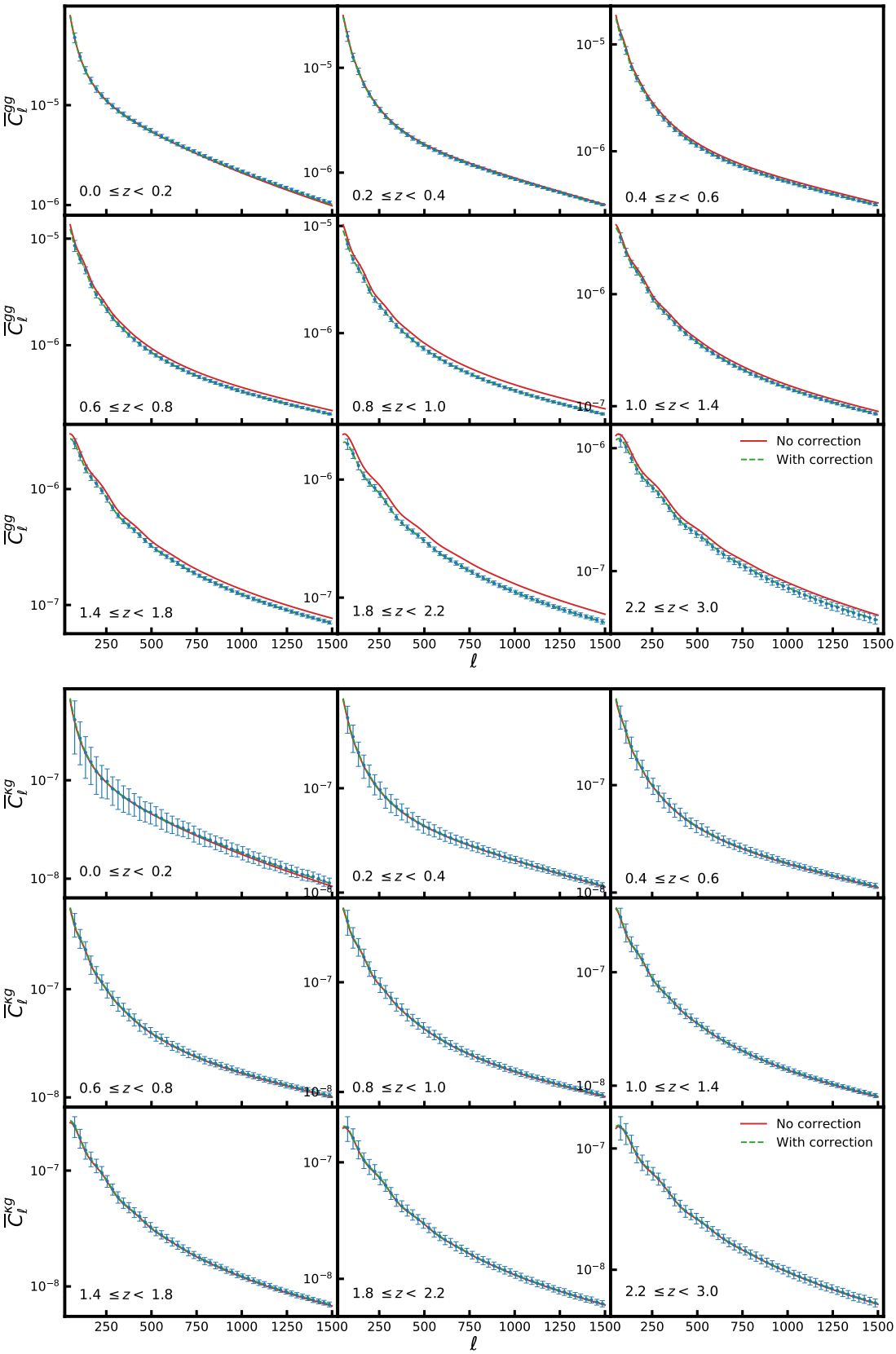


Figure 4.8: (*Top:*) average galaxy auto-power spectrum and (*bottom:*) cross power spectrum estimated from 300 simulations with photometric redshift scatter $\sigma_0 = 0.02$. The red solid line and the green dashed line are the theoretical power spectra before and after leakage correction. The error bars on the data points are estimated from 300 simulations using Eq. 4.11.

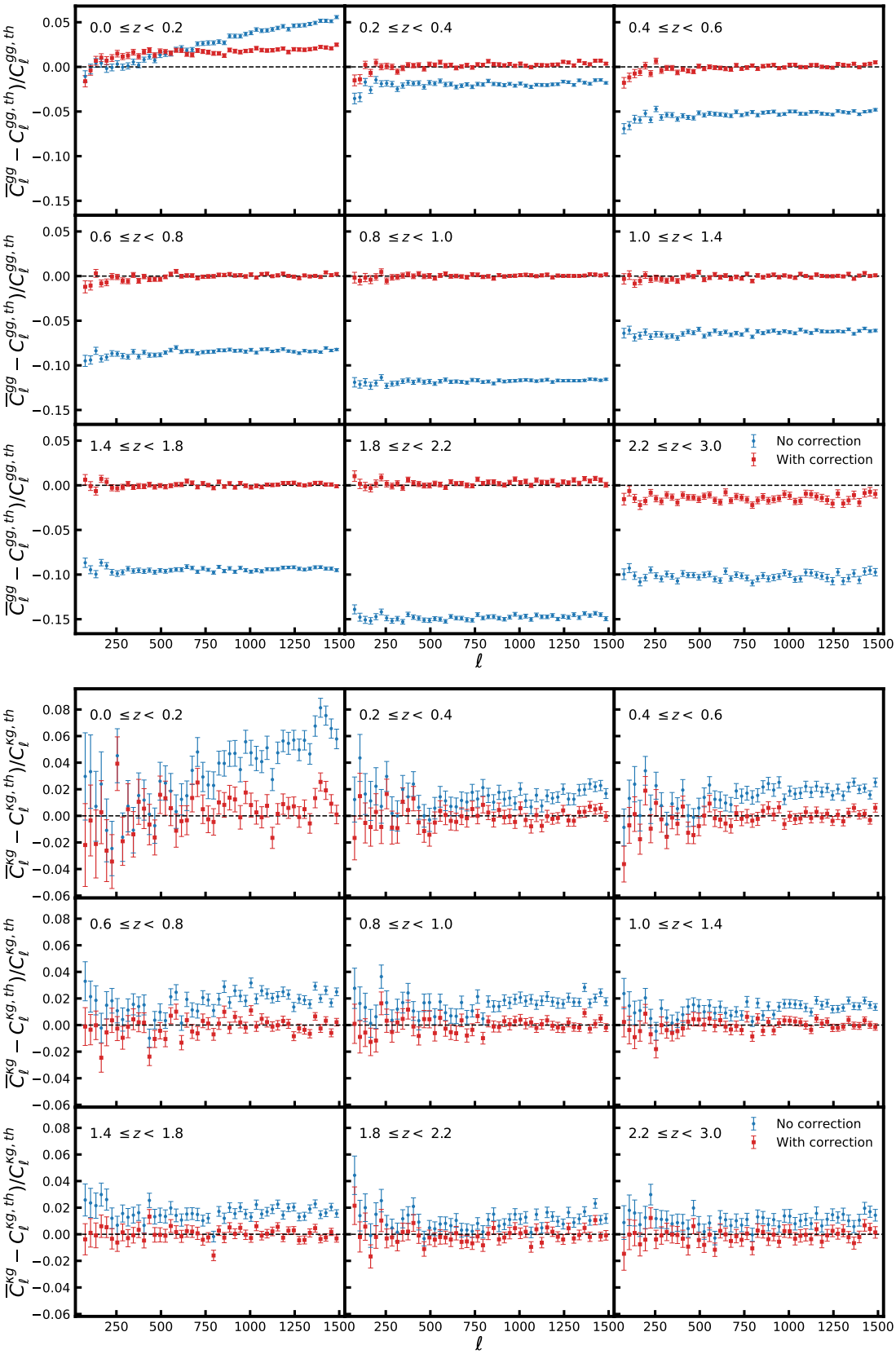


Figure 4.9: (Relative errors on the average *top*) galaxy auto-power spectrum and (*bottom*) cross power spectrum for photometric redshift scatter $\sigma_0 = 0.02$. The blue circles and red squares are offsets before and after correcting the theoretical power spectra for leakage, respectively. The error bars on the data points are estimated from 300 simulations using Eq. 4.11.

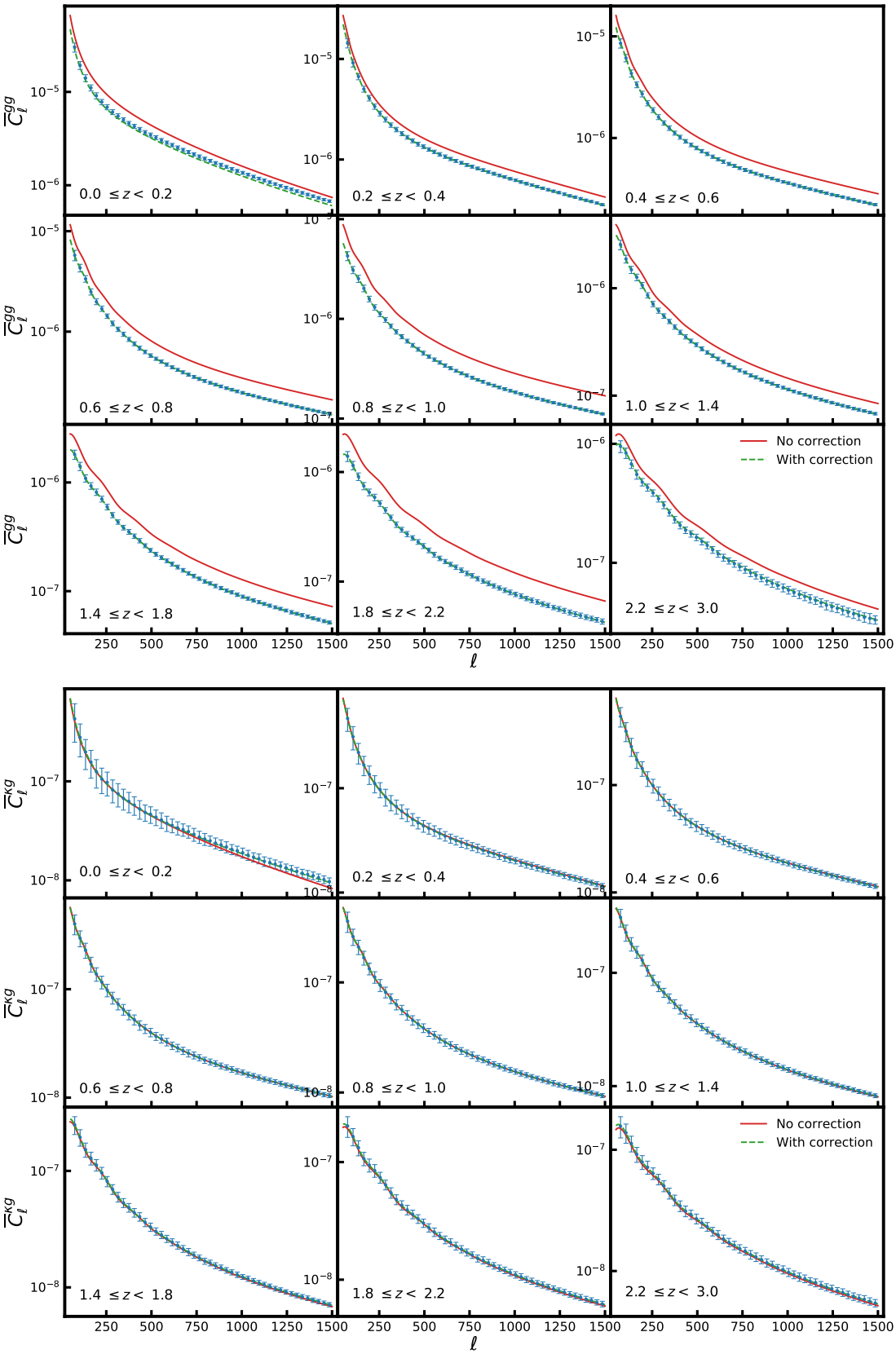


Figure 4.10: (*Top:*) average galaxy auto-power spectrum and (*bottom:*) cross power spectrum estimated from 300 simulations with photometric redshift scatter $\sigma_0 = 0.05$. The red solid line and the green dashed line are the theoretical power spectra before and after leakage correction. The error bars on the data points are estimated from 300 simulations using Eq. 4.11.

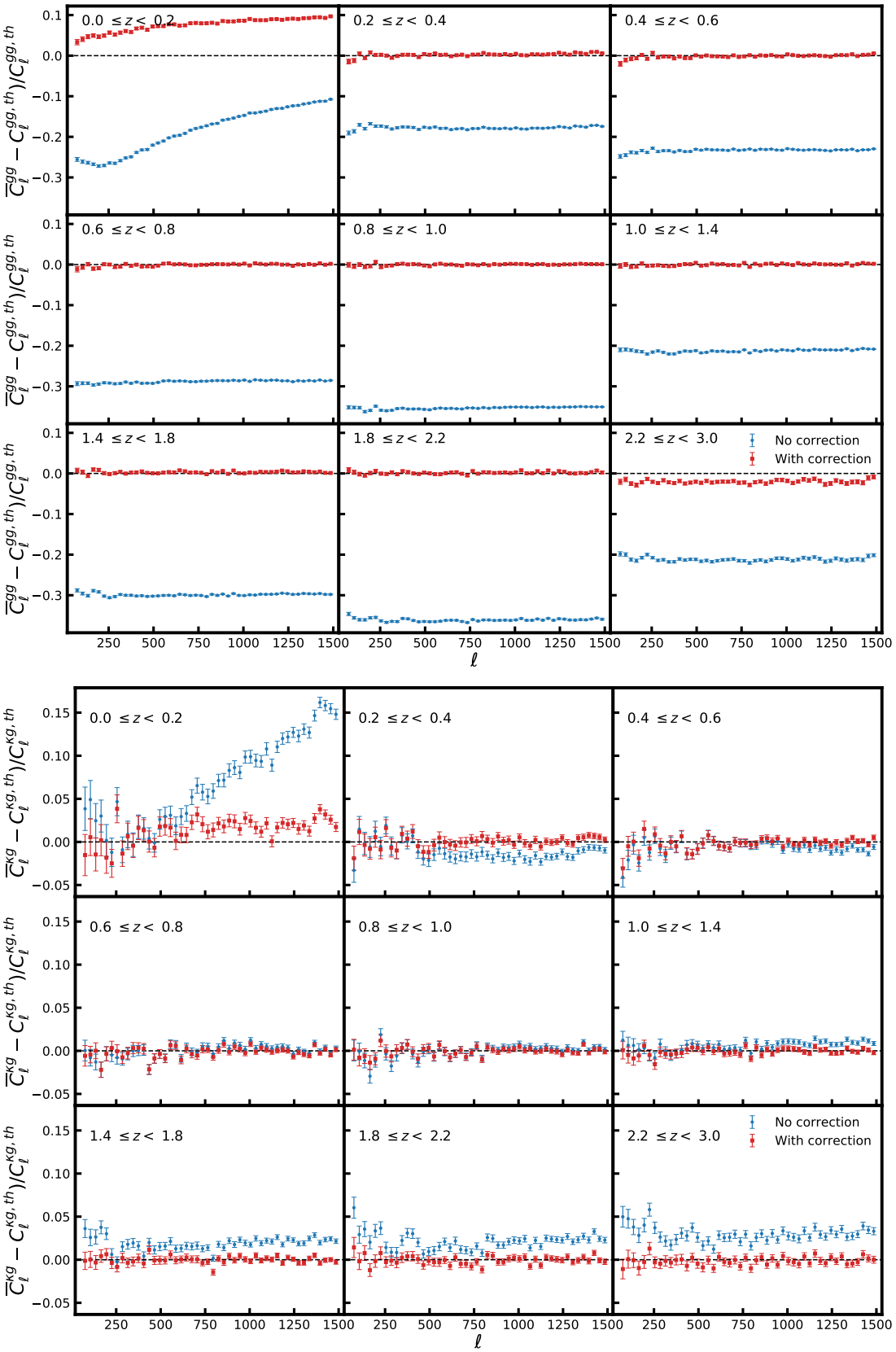


Figure 4.11: Relative errors on the average *top*) galaxy auto-power spectrum and (*bottom*) cross power spectrum for photometric redshift scatter $\sigma_0 = 0.05$. The blue circles and red squares are offsets before and after correcting the theoretical power spectra for leakage, respectively. The error bars on the data points are estimated from 300 simulations using Eq. 4.11.

ical power spectrum without accounting for leakage computed using redshift distribution given by Eq. (4.4) and the green dashed lines are the theoretical power spectra $C^{\rm th,ph}$ estimated using Eqs. (4.14) and (4.15). Figure 4.9 show the relative errors between the estimated average power spectra with $\sigma_0 = 0.02$ and the theoretical power spectra estimated without leakage correction (blue circles) and after accounting for leakage (red squares). Figures 4.10 and 4.11 present similar results for the case of photometric redshift scatter $\sigma_0 = 0.05$.

The theoretical power spectra after leakage correction agree completely with estimated power spectra in all bins, for both cases of $\sigma_0 = 0.02$ and 0.05. The disparity in the first and last bins results directly from the inaccuracy of the deconvolution method near lower and upper bounds of the redshift distribution considered in the analysis. Nevertheless, we notice that even for those tomographic bins the agreement with corresponding theoretical power spectra improves.

4.6 Parameter estimation

In previous sections, we observed that the power spectra in every tomographic bin gets biased due to leakage of objects across redshift bins, which can be corrected by accurate estimation of the scattering matrix. In this section, we study the impact of leakage on the estimation of redshift dependent galaxy linear bias b and amplitude of cross-correlation A from tomographic bins, estimated using Maximum Likelihood Estimation method discussed in section 2.4.

We use flat priors $b \in [0, 10]$ and $A \in [-5, 5]$ for the estimation of parameters while the remaining cosmological parameters are kept constant with values from our fiducial background cosmology described in section 4.2. To effectively sample the parameter space, we use a publicly available software package `EMCEE` (Foreman-Mackey et al., 2013). The best-fit value of the parameters are medians of their posterior distributions, with $\pm 1\,\sigma$ errors being the $16^{\rm th}$ and $84^{\rm th}$ percentile, respectively. We use the average power spectra estimated from 300 simulations for the estimation of parameters, hence, we divide the covariance matrices (see section 2.4 for covariance matrices used in the likelihood function) by $\sqrt{\rm N_{sim}}$, where $\rm N_{sim}$ is the total number of simulations.

Before accounting for leakage, we estimate the galaxy linear bias and cross-correlation amplitude for every tomographic bin using the average galaxy power spectra and the average cross-power spectra estimated from photometric datasets. The theoretical power spectrum templates for tomographic bin i are computed using the relations

$$C_i^{gg,\rm th}(\ell) = \int_0^{\chi_*} \frac{d\chi}{\chi^2} \left(\frac{{\rm d}N^i(z_t)}{{\rm d}z_t} \right)^2 P\left(k = \frac{\ell + 1/2}{\chi}, z(\chi) \right) \tag{4.17}$$

$$C_i^{\kappa g,\rm th}(\ell) = \int_0^{\chi_*} \frac{d\chi}{\chi^2} W^{\kappa}(\chi) \frac{{\rm d}N^i(z_t)}{{\rm d}z_t} P\left(k = \frac{\ell + 1/2}{\chi}, z(\chi) \right) \tag{4.18}$$

where $\frac{{\rm d}N^i(z_t)}{{\rm d}z_t}$ is given by Eq. (4.4). To estimate parameters b and A after accounting for leakage,

we transform the extracted photometric power spectra ($\overline{C}^{gg,\mathrm{ph}}$ and $\overline{C}^{\kappa g,\mathrm{ph}}$) to true power spectra by inverting Eqs. (4.14) and (4.15). The theoretical power spectrum templates for likelihood estimation after leakage correction are computed using Eqs. (4.17) and (4.18) and substituting Eq. (4.8) for $\frac{\mathrm{d}N^i(z_t)}{\mathrm{d}z_t}$. It is important to note that the photometric power spectra in tomographic analysis are a combination of true power spectra as represented in Eqs. (4.14) and (4.15). Thus, the galaxy linear bias in a photometric redshift bin will also be a combination of the galaxy linear bias from the true redshift bins. The estimation of parameters can also be performed directly over the estimated photometric power spectra by properly defining the covariance matrix in the likelihood function. However, transforming the estimated photometric power spectra to true power spectra for parameter estimation saves us from the complexities of defining the covariance matrix as well as reduces the computation time.

The best-fit values of galaxy linear bias and cross-correlation amplitude from all tomographic bins with $\pm 1\,\sigma$ errors, estimated before and after leakage corrections, are quoted in Table 4.2 for $\sigma_0 = 0.02$ and Table 4.3 for $\sigma_0 = 0.05$. The column b_{true} in these tables are the true values of bias for every tomographic bin. The galaxy linear bias for every tomographic bin estimated from photometric datasets are smaller than their expected value for both $\sigma_0 = 0.02$ and 0.05, whereas the amplitude of cross-correlation are consistently higher than the expected value of unity. However, both parameters become consistent with their expected values after correcting for leakage. The galaxy linear bias shows marginal deviations from its true values, the amplitude of cross-correlation is perfectly consistent with its expected value of unity within $1\,\sigma$.

Table 4.2: Best fit values of galaxy linear bias b and amplitude of cross-correlation A for all tomographic bins with $\sigma_0 = 0.02$, estimated before leakage correction and after correction through the scattering matrix approach. b^{true} is the true value of bias for the tomographic bin.

z	b^{true}	No correction		With correction	
		b	A	b	A
$[0.0, 0.2)$	1.3241	$1.3552^{+0.0002}_{-0.0002}$	$1.0243^{+0.0110}_{-0.0110}$	$1.3357^{+0.0002}_{-0.0002}$	$0.9952^{+0.0124}_{-0.0125}$
$[0.2, 0.4)$	1.3541	$1.3497^{+0.0002}_{-0.0002}$	$1.0269^{+0.0064}_{-0.0064}$	$1.3557^{+0.0002}_{-0.0002}$	$1.0014^{+0.0069}_{-0.0069}$
$[0.4, 0.6)$	1.3909	$1.3623^{+0.0002}_{-0.0002}$	$1.0440^{+0.0053}_{-0.0054}$	$1.3911^{+0.0002}_{-0.0002}$	$0.9990^{+0.0058}_{-0.0057}$
$[0.6, 0.8)$	1.4307	$1.3772^{+0.0002}_{-0.0002}$	$1.0645^{+0.0050}_{-0.0050}$	$1.4308^{+0.0002}_{-0.0002}$	$0.9998^{+0.0054}_{-0.0053}$
$[0.8, 1.0)$	1.4724	$1.3909^{+0.0002}_{-0.0002}$	$1.0816^{+0.0049}_{-0.0049}$	$1.4724^{+0.0002}_{-0.0002}$	$0.9999^{+0.0053}_{-0.0053}$
$[1.0, 1.4)$	1.5337	$1.4907^{+0.0002}_{-0.0002}$	$1.0465^{+0.0037}_{-0.0037}$	$1.5329^{+0.0002}_{-0.0002}$	$1.0003^{+0.0038}_{-0.0038}$
$[1.4, 1.8)$	1.6217	$1.5496^{+0.0002}_{-0.0002}$	$1.0691^{+0.0041}_{-0.0040}$	$1.6223^{+0.0002}_{-0.0002}$	$1.0005^{+0.0042}_{-0.0042}$
$[1.8, 2.2)$	1.7118	$1.5832^{+0.0003}_{-0.0003}$	$1.0963^{+0.0047}_{-0.0047}$	$1.7149^{+0.0003}_{-0.0003}$	$0.9986^{+0.0048}_{-0.0048}$
$[2.2, 3.0)$	1.8277	$1.7318^{+0.0005}_{-0.0005}$	$1.0668^{+0.0047}_{-0.0047}$	$1.8137^{+0.0004}_{-0.0004}$	$1.0050^{+0.0048}_{-0.0047}$

In Figure 4.12 we show the estimates of galaxy linear bias and amplitude of cross-correlation for $\sigma_0 = 0.02$ (left column) and $\sigma_0 = 0.05$ (right column). The blue circles and red squares represent the parameter estimates before and after leakage correction, respectively. The error

Table 4.3: Best fit values of galaxy linear bias b and amplitude of cross-correlation A for all tomographic bins with $\sigma_0 = 0.05$, estimated before leakage correction and after correction through the scattering matrix approach. b^{true} is the true value of bias for the tomographic bin.

z	b^{true}	No correction		With correction	
		b	A	b	A
$[0.0, 0.2)$	1.3241	$1.2172^{+0.0002}_{-0.0002}$	$1.2081^{+0.0100}_{-0.0100}$	$1.3686^{+0.0002}_{-0.0002}$	$0.9872^{+0.0123}_{-0.0123}$
$[0.2, 0.4)$	1.3541	$1.2288^{+0.0002}_{-0.0002}$	$1.0941^{+0.0059}_{-0.0058}$	$1.3555^{+0.0002}_{-0.0002}$	$1.0019^{+0.0069}_{-0.0069}$
$[0.4, 0.6)$	1.3909	$1.2194^{+0.0002}_{-0.0002}$	$1.1354^{+0.0049}_{-0.0049}$	$1.3911^{+0.0002}_{-0.0002}$	$0.9986^{+0.0057}_{-0.0057}$
$[0.6, 0.8)$	1.4307	$1.2088^{+0.0002}_{-0.0002}$	$1.1849^{+0.0045}_{-0.0046}$	$1.4311^{+0.0002}_{-0.0002}$	$0.9998^{+0.0054}_{-0.0054}$
$[0.8, 1.0)$	1.4724	$1.1856^{+0.0002}_{-0.0002}$	$1.2421^{+0.0044}_{-0.0045}$	$1.4729^{+0.0002}_{-0.0002}$	$0.9996^{+0.0052}_{-0.0052}$
$[1.0, 1.4)$	1.5337	$1.3626^{+0.0002}_{-0.0002}$	$1.1333^{+0.0036}_{-0.0036}$	$1.5344^{+0.0002}_{-0.0002}$	$0.9990^{+0.0039}_{-0.0038}$
$[1.4, 1.8)$	1.6217	$1.3589^{+0.0002}_{-0.0002}$	$1.2169^{+0.0040}_{-0.0039}$	$1.6247^{+0.0002}_{-0.0002}$	$0.9988^{+0.0042}_{-0.0042}$
$[1.8, 2.2)$	1.7118	$1.3692^{+0.0003}_{-0.0003}$	$1.2791^{+0.0046}_{-0.0046}$	$1.7138^{+0.0003}_{-0.0003}$	$0.9994^{+0.0048}_{-0.0048}$
$[2.2, 3.0)$	1.8277	$1.6220^{+0.0005}_{-0.0005}$	$1.1614^{+0.0047}_{-0.0047}$	$1.8078^{+0.0004}_{-0.0004}$	$1.0082^{+0.0047}_{-0.0048}$

bars on the data points have been re-scaled by multiplying with $\sqrt{N_{\text{sim}}}$ to match the error budget of a single realization. The top panel of Figure 4.12 presents the evolution of galaxy bias with redshift. The black dashed line marks the fiducial evolution of galaxy bias used in our simulations. The middle panel shows the z-score values for the galaxy linear bias parameter. Without properly accounting for the scatter of objects across redshift bins, the galaxy bias can deviate between $5-30\,\sigma$ when $\sigma_0 = 0.02$, and by $25-110\,\sigma$ when $\sigma_0 = 0.05$. Such large deviations on galaxy linear bias are visible because the errors from likelihood estimation (quoted in Tables 4.2 and 4.3) are between $0.15-0.5\%$ for a single realization. This shows that the estimates for the galaxy bias are very tightly constrained. The bottom panel of Figure 4.12 shows the z-score values for the amplitude of cross-correlation, which can deviate up to $\sim 1.2\,\sigma$ with $\sigma_0 = 0.02$, and up to $\sim 3.5\,\sigma$ when $\sigma_0 = 0.05$. As it is clearly conveyed by Figure 4.12, the parameters galaxy linear bias and cross-correlation amplitude become consistent with their expected values after correcting for the effect of redshift bin mismatch of objects through our scattering matrix formalism.

4.7 A note on σ_8 parameter estimation

We have shown in section 4.6 that the scatter of objects across redshift bins can lead to a biased estimation of parameters, thus altering our inferences about the cosmological model. In this section we will estimate the impact of leakage on the σ_8 parameter. Peacock and Bilicki, 2018 proposed a method to compute the σ_8 parameter from the cross-correlation amplitude using the relation:

$$\sigma_8(z) = A(z)\sigma_{8,0}\,D(z) \tag{4.19}$$

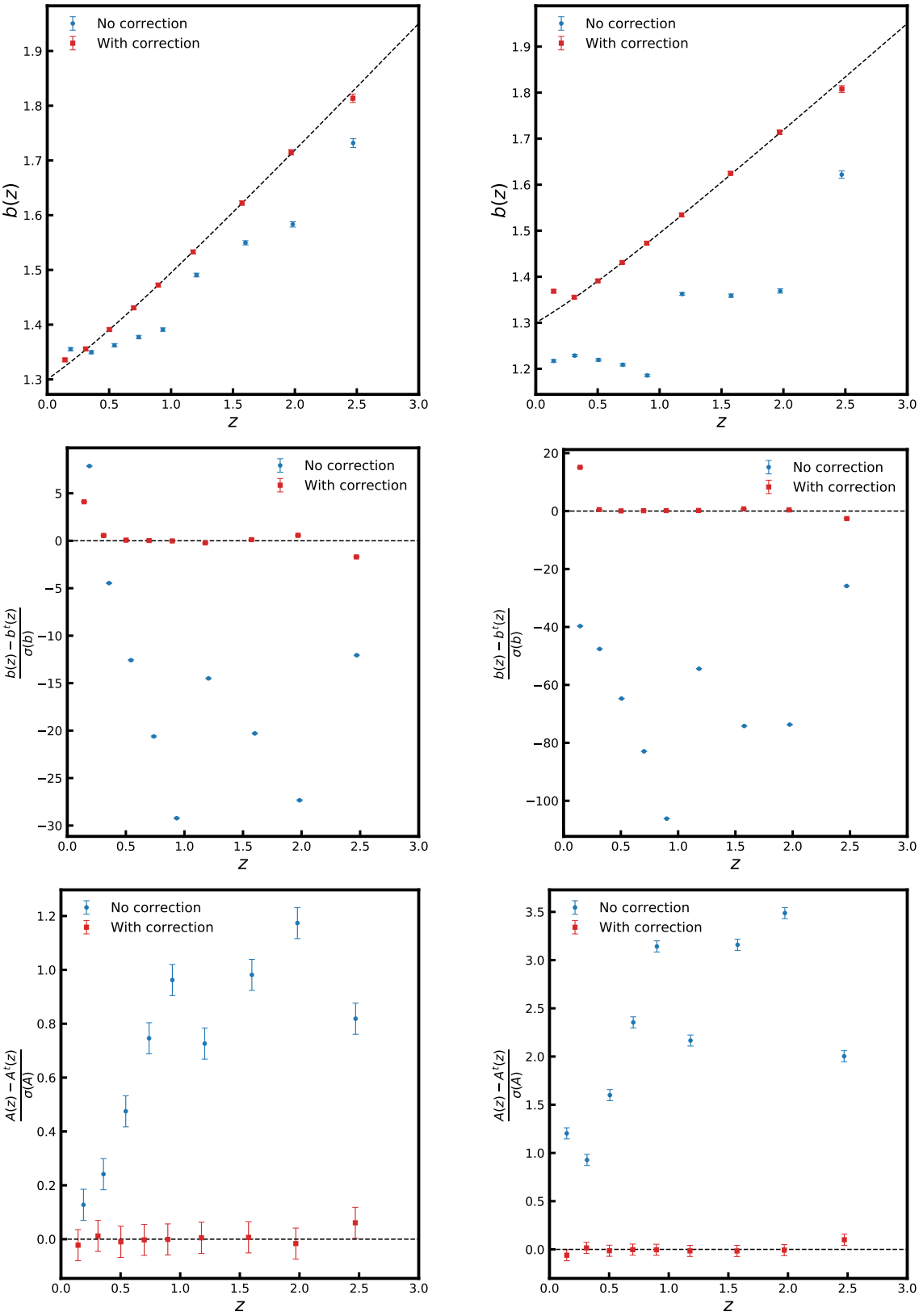


Figure 4.12: Comparison of *top*: galaxy linear bias, *middle*: z-score for galaxy linear bias, and *bottom*: z-score for cross-correlation amplitude before (blue circles) and after (red squares) leakage correction. The *left* column is comparison for $\sigma_0 = 0.02$ and the right column for $\sigma_0 = 0.05$. The error bars on parameters are re-scaled to a single realization.

where $\sigma_{8,0}$ is the value of σ_8 parameter at redshift $z = 0$ and $D(z)$ is the linear growth function given by Eq. (4.2). We compute the value of $\sigma_{8,0}$ for our assumed background cosmology using `CAMB` software. In Figure 4.13, we show the impact of scattering of objects on σ_8 parameter computed using Eq. (4.19), for $\sigma_0 = 0.02$ (in left panel) and 0.05 (on the right). The error bars are multiplied with $\sqrt{N_{sim}}$ (where N_{sim} is the number of simulations) to re-scale it to a single realization. The black dashed lines are the fiducial evolution of σ_8 parameter with redshift. We measure a higher than expected value of σ_8 without taking into account the leakage of objects across tomographic bins. As expected, the biases suffered by the amplitude of cross-correlation are reflected directly in the values of σ_8 parameter. Thus, it becomes crucial to correct for the cross-talk between different redshift bins in a tomographic analysis to get an unbiased estimate of cosmological parameters.

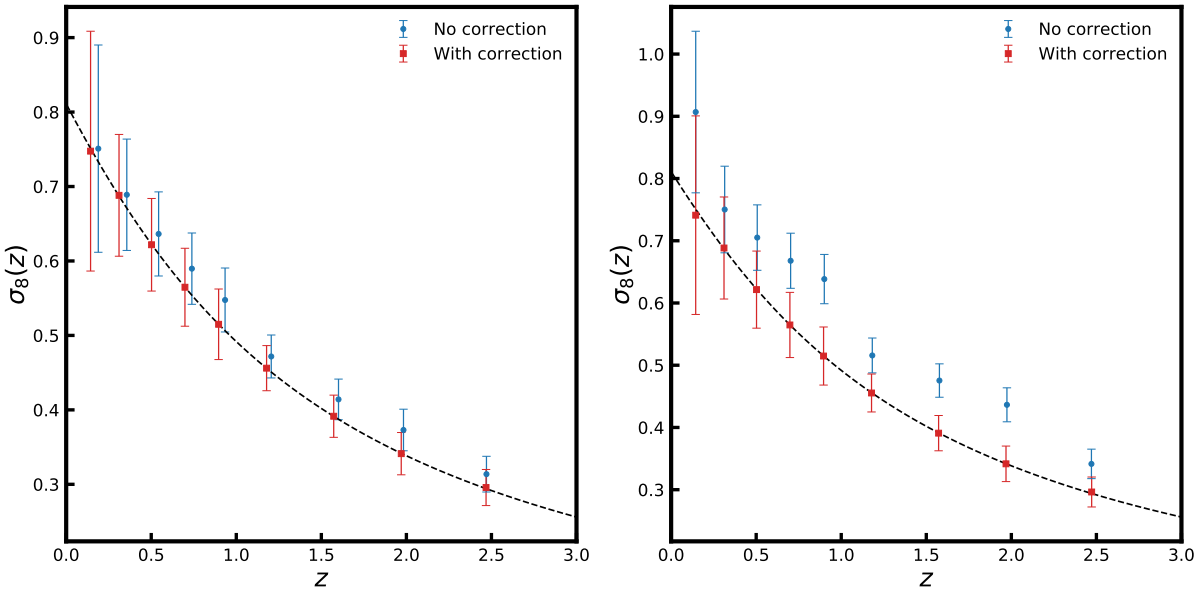


Figure 4.13: Comparison of σ_8 parameter before and after leakage correction for $\sigma_0 = 0.02$ (*left*) and $\sigma_0 = 0.05$ (*right*). The black dashed line is the fiducial evolution of σ_8 parameter used during simulations.

4.8 Impact of redshift binwidth

In section 4.4, we have shown that the photometric angular power spectra measured in a tomographic analysis are biased due to leakage of objects across redshift bins. The biased power spectra can be corrected using scattering matrix as described in section 4.5. We estimate galaxy linear bias and cross-correlation amplitude from 9 redshift bins with different bin-sizes $\Delta z = 0.2, 0.4,$ and 0.8 in the redshift range $z = [0, 1.0), [1.0, 2.2),$ and $[2.2, 3.0]$, respectively. We find that the galaxy linear bias is estimated smaller than their expectations without correcting for scatter of objects. We note that there is an indication from Figure 4.12 that the deviations in the parameters may be dependent on the size of the redshift bin. Thus, in this section, we study the impact of size of redshift bins on the bias incurred by the power spectra and parameters due to photometric redshift scatters.

We create 4 sets of 300 simulated maps using the `FLASK` code with galaxy density following LSST specifications and CMB lensing following *Planck* observations, using the procedure described in section 4.2. We generate photometric redshifts in each simulation from a Gaussian distribution following section 4.2 with $\sigma_0 = 0.02$. Each set of photometric datasets is divided into 3 tomographic bins with redshift intervals $(0.0, z_{\rm min}, z_{\rm max}, 3.0]$, where $z_{\rm min}$ and $z_{\rm max}$ are chosen such that the effect of growth function $D(z)$ on the estimation of parameters from the second redshift bin is marginalized. The values of $z_{\rm min}$, $z_{\rm max}$ and size of the second redshift bin for all 5 sets are present in Table 4.4.

Table 4.4: Properties of the second redshift bin for 4 sets of simulations used to study the impact of size of redshift bins on leakage. $z_{\rm min}$ and $z_{\rm max}$ are the lower and upper redshift bound for the second tomographic bin and Δz is the corresponding binwidth.

	$z_{\rm min}$	$z_{\rm max}$	Δz
Set 1	0.69	0.93	0.25
Set 2	0.63	1.00	0.37
Set 3	0.56	1.08	0.51
Set 4	0.46	1.28	0.82

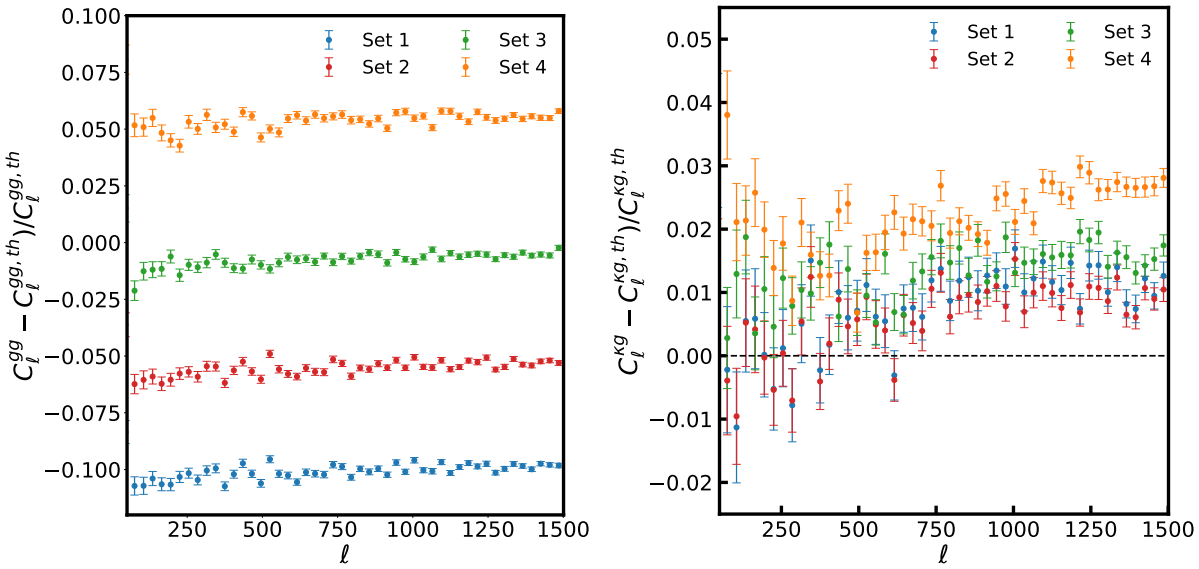


Figure 4.14: Comparison of galaxy auto-power spectra (*left*) and cross-power spectra (*right*) before leakage correction for tomographic bins with sizes 0.25 (Set 1), 0.37 (Set 2), 0.51 (Set 3) and 0.82 (Set 4). The error bars on the data points are estimated from 300 simulations using Eq. 4.11.

We estimate the galaxy auto-power spectra and cross-power spectra from photometric datasets for the second tomographic bin. In Figure 4.14 we compare the extracted power spectra before leakage correction. The offsets in the photometric galaxy auto-power spectra moves from being negatively biased (smaller than expected power spectra) to positively biased (larger than expected) with increasing redshift bin sizes. The cross-power spectra, however, shows only a weak dependence on the redshift bin size. Since the galaxy auto-power spectra moves from negative to

positive biased, it may be possible to find some optimal redshift binwidth for which the the bias incurred by galaxy auto-power spectra will be close to zero. However, it is important to note that offsets in the photometric galaxy auto-power spectra will depend on several factors like photometric redshift errors, magnification bias, systematics such as catastrophic photometric redshift errors and photometric calibration errors. We keep the study of these systematics on the estimated galaxy auto-power spectra for future study.

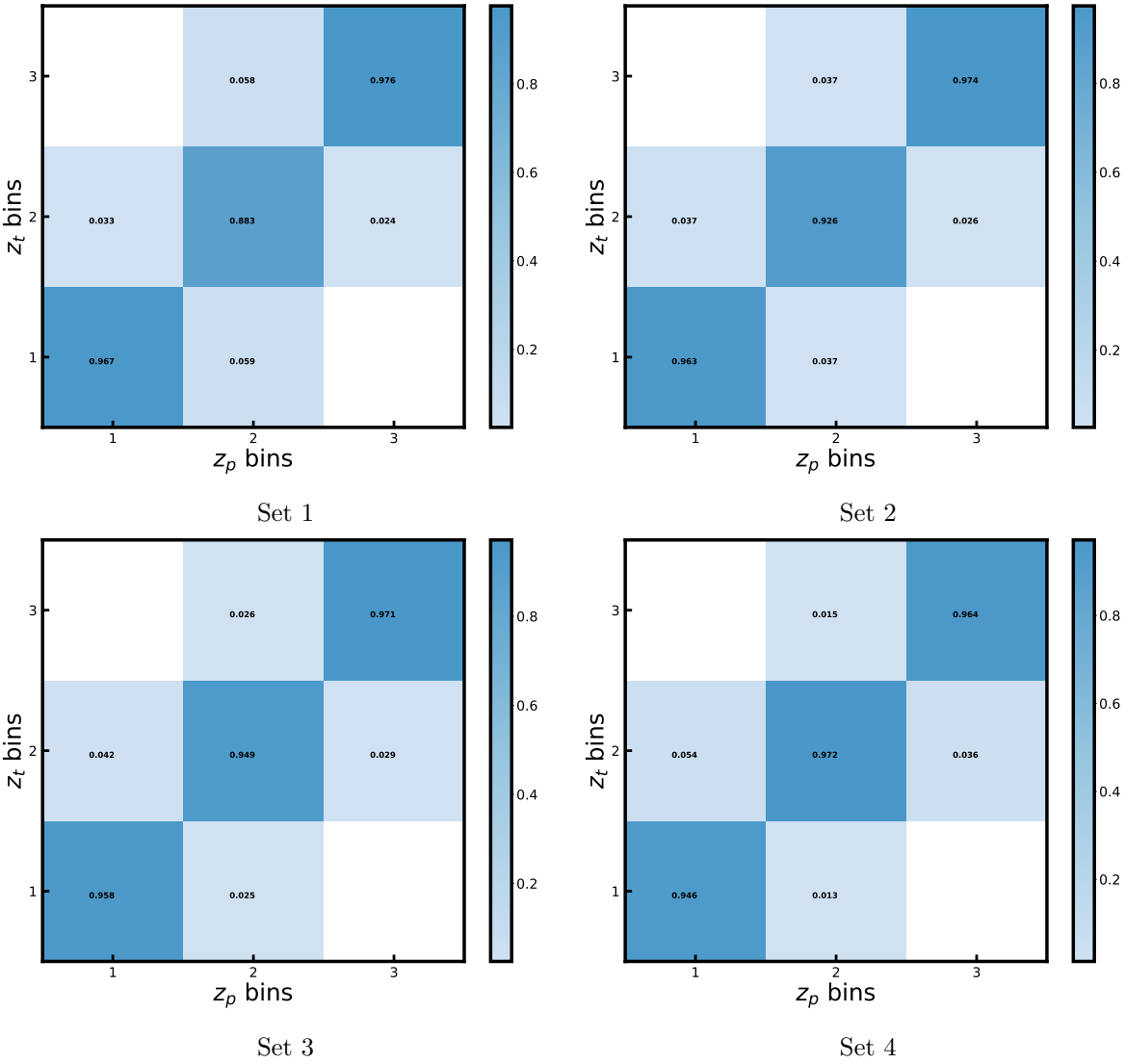


Figure 4.15: Scattering matrices estimated for four different sets of simulations.

In Figure 4.15, we show the scattering matrices computed using the procedure described in section 4.5 for the four set of simulations. We show in Figure 4.16 the impact of different bin sizes on the z-score values of parameters estimated through maximum likelihood estimation before and after leakage correction. The error bars on the parameters are adjusted for a single realization. Due to redshift bin mismatch of objects, the galaxy bias is estimated $\sim 25\,\sigma, \sim 15\,\sigma$ and $\sim 2\,\sigma$ lower than its true value for the redshift bin size $\Delta z = 0.25, 0.37,$ and 0.51, respectively. For binwidth $\Delta z = 0.82$, the galaxy bias is estimated $\sim 12\,\sigma$ higher than its true value. The amplitude of cross-correlation, being anti-correlated to the galaxy linear bias, is estimated larger for $\Delta z = 0.25, 0.37, 0.51$ and marginally smaller for $\Delta z = 0.82$ compared to its true value of $A = 1$.

The parameters, galaxy linear bias and amplitude of cross-correlation, become consistent with their expectations after correcting for leakage. The dependence of the leakage bias on the size of tomographic bins further establishes the importance of using our scattering matrix approach for an unbiased estimation of cosmological parameters.

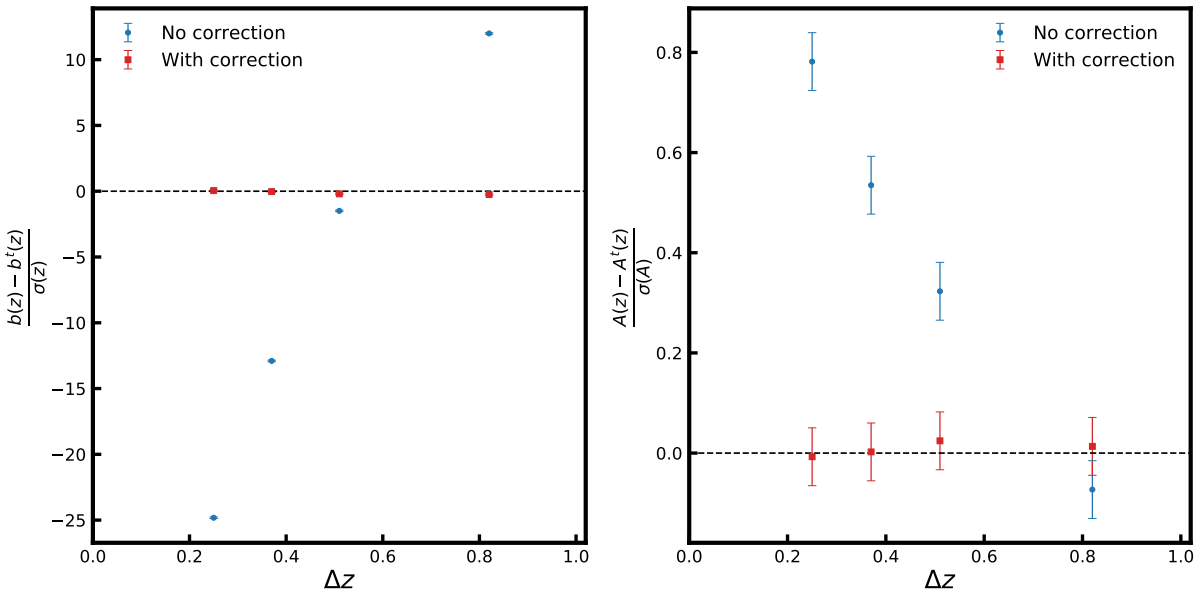


Figure 4.16: Z-score for galaxy linear bias (*left*) and amplitude of cross-correlation (*right*) for tomographic bins with different sizes, before (blue circles) and after (red squares) leakage correction with $\sigma_0 = 0.02$.

4.9 Summary

We present the tomographic study of cross-correlations by performing simulations of *Planck* CMB lensing convergence and galaxy density field mimicking properties of LSST photometric survey. We use the `FLASK` code to simulate log-normal fields and divide the galaxies into 9 redshift bins. We consider photometric redshift errors with standard deviation of $\sigma(z) = 0.02(1+z)$ and $\sigma(z) = 0.05(1+z)$, but do not include catastrophic redshift errors or photometric calibration errors, keeping them for future studies. In this sense, we generate an 'ideal' scenario in our simulations free from other systematics which is crucial to demonstrate the importance of leakage of objects across redshift bins. We We compute galaxy auto-power spectrum and cross-power spectrum between galaxy over-density and CMB convergence fields, and use these power spectra to estimate two parameters, the redshift dependent galaxy linear bias, b, and amplitude of cross-correlation, A employing the maximum likelihood method.

We estimate the true redshift distribution from simulated photometric redshift distribution by the convolution method (section 4.3.1) and compare the estimated photometric power spectra with their theoretical expectations in every tomographic bin. The most important quantity to accurately recover the true redshift distribution from the convolution method is the precise estimation of the error functions $p(z_t - z_p|z_p)$. We estimate the error function with sub-percent accuracy ($< 0.25\%$) by fitting a sum of three Gaussians to $p(z_t - z_p|z_p)$ as shown in Figure

4.2. We find the sum of Gaussians to accurately capture the peculiarities of the error function $p(z_t - z_p|z_p)$ like non-Gaussian tails and higher peak in the center. The galaxy auto-power spectrum measured from photometric datasets are found to be consistently smaller in every bin with respect to their fiducial predictions as shown in Figure 4.6. The offsets vary between $2-15\%$ for simulations with $\sigma(z) = 0.02(1+z)$ and between $15-40\%$ for $\sigma(z) = 0.05(1+z)$. The measured cross-power spectra are also biased with smaller deviations ($< 5\%$) for both cases. The measured power spectra are inconsistent with their expectations due to the scattering of objects from one redshift bin to other due to photometric redshift errors. This conclusion is consistent with the fact that deviations are larger in the case of photometric redshift scatter $\sigma(z) = 0.05(1+z)$.

To alleviate the differences in the power spectra, we implement the scattering matrix approach introduced by Zhang et al., 2010 to correct for the effect of redshift bin mismatch of objects. The scattering matrix describes the fraction of objects in a photometric redshift bin that comes from different true redshift bins. The power spectra in photometric redshift bins then transform as a linear combination of power spectra from different true redshift bins (Eqs. 4.14 and 4.15). Zhang et al., 2017 proposed an algorithm based on Non-negative Matrix Factorisation method to solve for similar numerical problems, however, this method is computationally challenging for large number of data points in the power spectra and number of tomographic bins. To circumvent these challenges, we propose an alternative method for fast and accurate computation of the scattering matrix based on the reconstruction of the true redshift distribution by the deconvolution method (see section 4.5). We show in Figure 4.3 the average fractional errors on reconstruction of the true redshift distribution via the deconvolution method. Figure 4.7 shows that our new method to compute the scattering matrix is robust and only proves inefficient in the first tomographic bin due to cut in the redshift distribution at boundary $z = 0$. With a precise estimation of the scattering matrix, we correct the theoretical power spectra for the tomographic bins to compare with the estimated galaxy power spectra from simulated photometric datasets. Figures 4.8 - 4.11 show that scattering matrix methodology makes the estimated power spectra consistent with the leakage corrected theoretical power spectra.

We quantify the impact of redshift bin mismatch of objects on the estimation of galaxy linear bias and amplitude of cross-correlation. To estimate parameters after leakage correction, we transform the estimated photometric power spectra to estimated true power spectra by inverting the Eqs. 4.14 and 4.15 (as described in section 4.6). Without accounting for the leakage, we estimate smaller values for the galaxy linear bias, while the amplitude of cross-correlation is estimated higher than its fiducial value of unity. It is important to note here that the estimations of lower galaxy bias and higher amplitude are not to be generalized for every tomographic analysis. The offsets suffered by the power spectra and parameters estimated from photometric datasets in a tomographic study will depend strongly on the photometric redshift error distributions as well as the redshift distribution of objects. After correcting for leakage by using the scattering matrix, both parameters are very-well constrained with their expected values. We also study how tomographic bin sizes affect the estimation of galaxy auto- and cross-power spectra. We establish that different bin-widths will result in different magnitudes of errors in the power spectra and cosmological parameters.

The amplitude of cross-correlation is an indicator of the validity of the background cosmological model. Thus, without correcting for the bias resulting from photometric redshift errors, it will become inevitable to make wrong inferences when testing cosmological models with tomographic analyses. Other estimators frequently used to test the cosmological models, like the D_G (Giannantonio et al., 2016) or E_G (Pullen et al. 2016; Zhang et al. 2007) statistics, also employ the ratio of cross-power spectra to galaxy auto-power spectra and are, hence, proportional to the cross-correlation amplitude. We study the relationship between the amplitude of cross-correlation and more familiar σ_8 parameter in section 4.7. We show that the offsets suffered by the amplitude due to scatter of objects are synonymous with the deviations in the σ_8 parameter. With next-generation galaxy surveys like Vera C. Rubin Observatory Legacy Survey of Space and Time (LSST, Ivezić et al. 2019; LSST Science Collaboration et al. 2009), Euclid (Laureijs et al., 2011), and Dark Energy Spectroscopic Instrument (DESI) (Fagrelius 2020; Dey et al. 2019b), the tomographic approach will emerge as a powerful tool to put stringent constraints on the validity of cosmological models. Hence, we propose that the scattering matrix approach developed and presented in this chapter be strictly used for future tomographic studies.

Chapter 5

Impact of leakage correction with DESI Legacy Imaging Survey and *Planck* CMB lensing potential

5.1 Overview

In Chapter 4, we quantified using simulations the impact of leakage of objects across redshift bins on the estimation of power spectra in tomographic cross-correlation measurements. We observed that without correcting for the redshift bin mismatch, the galaxy auto-power spectrum and cross-power spectrum between galaxy density and CMB lensing will be biased in every tomographic bin, depending on the strength of the photometric redshift errors and the size of the redshift bin. The biased estimates for the power spectra will ultimately affect the cosmological parameters, in our case the galaxy linear bias and amplitude of cross-correlation. In this chapter, we apply the leakage correction in the cross-correlation analysis using galaxy catalogues from the Dark Energy Spectroscopic Instrument Legacy Imaging Survey (DESI-LIS; Dey et al. 2019b) prepared by Hang et al., 2021. Hang et al., 2021 studied the cross-correlation between *Planck* CMB lensing and galaxy density field in four tomographic slices covering the redshift range $0 < z < 0.8$ with photometric redshift precision $\frac{\sigma_z}{1+z}$ in the range $0.012 - 0.015$. They found $1.5 - 2.5\,\sigma$ deviations from the expected value of cross-correlation amplitude consistently in all tomographic bins. However, they do not account for the leakage across tomographic bins, and in this chapter we re-analyse the DESI-LIS $\times$ *Planck* cross-correlation measurements and apply the scattering matrix formalism to correct for the tomographic bin mismatch.

The chapter is arranged as follows: we describe the data used in the cross-correlation analysis in section 5.2 and quantify the fits made to the photometric redshift error distributions in section 5.3. Section 5.4 describes the estimation of true redshift distribution using the convolution and deconvolution methods. In section 5.5 we gauge the impact of photometric redshift errors on the estimation of parameters through simulations. We present the impact of leakage correction on measured power spectra and estimates of cosmological parameters in section 5.6. We also study the effects of using different CMB lensing potential maps on the inferred cosmological

parameters as well as translate the influence of leakage correction on to the σ_8 parameter. We finally summarize the findings from this chapter in section 5.7.

5.2 Data

5.2.1 CMB Lensing Data

The CMB lensing convergence map used in our analysis is described Chapter 3. We briefly re-iterate the CMB dataset used for analysis in this section. The minimum-variance CMB convergence map (MV map, hereafter) comes from the 2018 *Planck* data release[1] (Planck Collaboration et al., 2020b). It uses the SMICA DX12 CMB maps to reconstruct the lensing potential from CMB temperature and polarization data covering $\sim 67\%$ of the sky. The *Planck* data package provides spherical harmonics coefficients for the lensing convergence map at HEALPix[2] (Górski et al., 2005) resolution parameter $N_{\rm side} = 4096$, which we downgrade to $N_{\rm side} = 1024$ for our analysis. The data package also provides an estimate of the noise power spectrum $N_\ell^{\kappa\kappa}$ for the minimum variance map along with a binary map masking the regions of sky not suitable for analysis.

5.2.2 Legacy Survey Data

We use the galaxy catalogue and photometric redshifts prepared by Hang et al., 2021 from the Data Release 8[3] of the Legacy Imaging Survey (Dey et al., 2019b). The Legacy Imaging Survey is a combination of observations from the Dark Energy Camera Legacy Survey (DECaLS) observed using the Dark Energy Camera (Flaugher et al., 2015), the Mayall z-band Legacy Survey (MzLS) observed by the Mosaic3 camera (Dey et al., 2016) and the Beijing-Arizona Sky Survey (BASS) observed by the 90Prime camera (Williams et al., 2004). The Legacy Survey covers an area of ~ 17800 deg^2 with sources observed in three optical bands (g, r, z) and three WISE (Wright et al., 2010) bands W_1, W_2 and W_3. The following selection criteria are applied to the data (Hang et al., 2021)

1. PSF-type objects are excluded which eliminates most stars and quasars.
2. Objects are detected in four bands, i.e. `FLUX G|R|Z|W1` > 0.
3. `MW TRANSMISSION G|R|Z|W1` are applied to the fluxes for Galactic extinction correction.
4. Magnitude cuts are applied with $g < 24$, $r < 22$, and $W_1 < 19.5$.

In addition to above selection criteria, Hang et al., 2021 generates completeness map to account for the foreground contamination at the map pixel level. The foreground contaminations include masks for bright stars, globular clusters and incompleteness in optical bands. The galaxies are divided into 4 tomographic bins with redshift intervals $(0.0, 0.3, 0.45, 0.6, 0.8]$. The galaxy count in each pixel is corrected by the completeness in each pixel, we denote by n the completeness corrected counts in each pixel. A summary of four tomographic bins including number of objects and mean density of objects per pixel (corrected for completeness) and per steradian is given in Table 5.1.

[1] https://pla.esac.esa.int/#cosmology
[2] https://healpix.jpl.nasa.gov/
[3] http://legacysurvey.org/dr8/

Table 5.1: Physical properties of DESI-LIS tomographic bins. $N_{\rm obj}$ is the number of objects, $\overline{n}$ is the mean number of objects, and median z is the median redshift of the tomographic bins.

Redshift Bin	$N_{\rm obj}$	$\overline{n}$ [gal pix^{-1}]	$\overline{n}$ [gal str^{-1}]	median z
$0 < z \leq 0.3$	14363105	2.652	2.655×10^6	0.21
$0.3 < z \leq 0.45$	11554242	2.133	2.136×10^6	0.38
$0.45 < z \leq 0.6$	13468310	2.487	2.490×10^6	0.51
$0.6 < z \leq 0.8$	7232579	1.335	1.337×10^6	0.66

The galaxy over-density maps for every tomographic bin is build at the HEALPix resolution parameter $N_{\rm side} = 1024$ using the relation

$$g(\hat{\mathbf{n}}) = \frac{n(\hat{\mathbf{n}}) - \overline{n}}{\overline{n}} \tag{5.1}$$

where $n(\hat{\mathbf{n}})$ is the number of objects in a given pixel and $\overline{n}$ is the mean number of objects per pixel. The galaxy over-density maps smoothed with a Gaussian beam of 60′ FWHM (for illustrative purpose only) are shown in Figure 5.1.

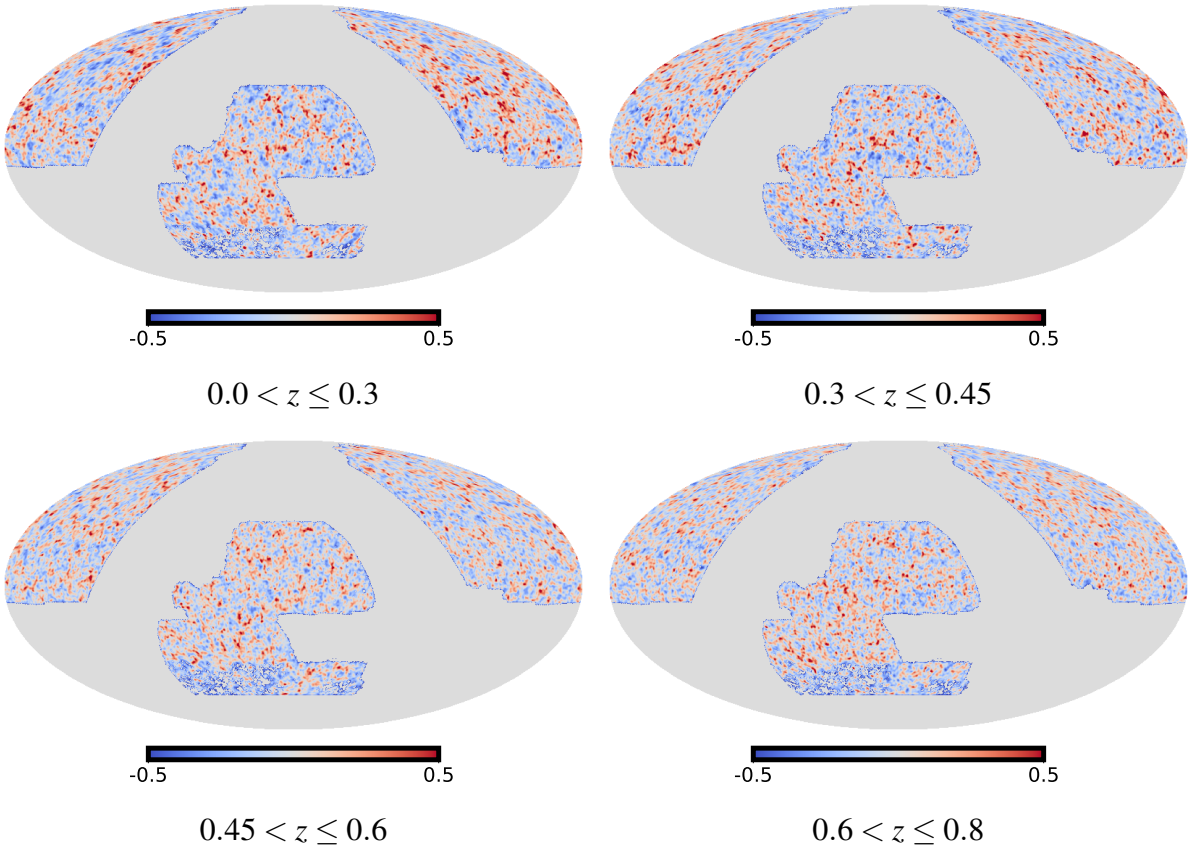


Figure 5.1: Galaxy over-density maps from four DESI-LIS tomographic bins. The maps have been smoothed with a Gaussian beam of 60′ FWHM for illustrative purposes.

Hang et al., 2021 also provides an estimate of the photometric redshift of galaxies based on several spectroscopic observations to assign redshifts in multi-color space. The spectroscopic surveys used to estimate the photometric redshifts of galaxies include GAMA Data Release 2 (Liske

et al., 2015), BOSS Data Release 12 (Alam et al., 2015), eBOSS Data Release 16 (Ahumada et al., 2020), VIPERS Data Release 2 (Scodeggio et al., 2018) and DEEP2 (Newman et al., 2013). Two photometric surveys COSMOS (Ilbert et al., 2009) and DES Y1 redMaGiC (Cawthon et al., 2018) were also included with the spectroscopic surveys for their highly accurate photometric redshifts. The galaxy number count maps and the photometric redshift data used in this study are publicly available at https://gitlab.com/qianjunhang/desi-legacy-survey-cross-correlations.

5.3 Photometric redshift error distribution

Hang et al., 2021 models the redshift error distribution of $\Delta z = z_s - z_p$ ($s \equiv$ spectroscopic; $p \equiv$ photometric) as a function of z_p, $p(z_s - z_p|z_p)$, for every tomographic bin with a modified Lorentzian function given by

$$L(x) = \mathcal{N}\left[1 + \frac{(x-x_0)^2}{2a\sigma^2}\right]^{-a} \tag{5.2}$$

where $\mathcal{N}$ is the normalization and x_0, σ, a are the parameters to be constrained for every tomographic bin. The best fit-values to x_0, σ, a for every tomographic bin taken from Hang et al., 2021 are quoted in Table 5.2.

Table 5.2: The best fit values taken from (Hang et al., 2021) for the parameters defining the modified Lorentzian fit to the photometric redshift error distribution $p(z_s - z_p|z_p)$.

z	x_0	σ	a
(0.0, 0.3]	-0.0010	0.0122	1.257
(0.3, 0.45]	0.0076	0.0151	1.319
(0.45, 0.6]	-0.0024	0.0155	1.476
(0.6, 0.8]	-0.0042	0.0265	2.028

The photometric error distributions compared with the best fit modified Lorentzian function is shown in Figure 5.2. We note that the modified Lorentzian fit proves to be a good estimate only near the peak, but fails to capture the tails of the error distribution. We attempt to model the complete error distribution and find that a sum of 10 Gaussians provide much better fits in every tomographic bin. It is clearly seen in Figure 5.2, where is shown comparison between the fitted sum of Gaussians and the modified Lorentzian of Hang et al., 2021. Because of the striking differences in the fits made to the photometric redshift error distribution, we will study in section 5.6 the changes incurred by the cosmological parameters when using our sum of 10 Gaussians fit over the modified Lorentzian model.

5.4 Estimation of true redshift distribution

In this section, we present two method to estimate the true redshift distribution. The more commonly used convolution method will be used as the simple method of estimating the theoretical power spectra. We will use the true redshift distribution estimated from the deconvolution method for computation of the scattering matrix and correct for the leakage of objects across

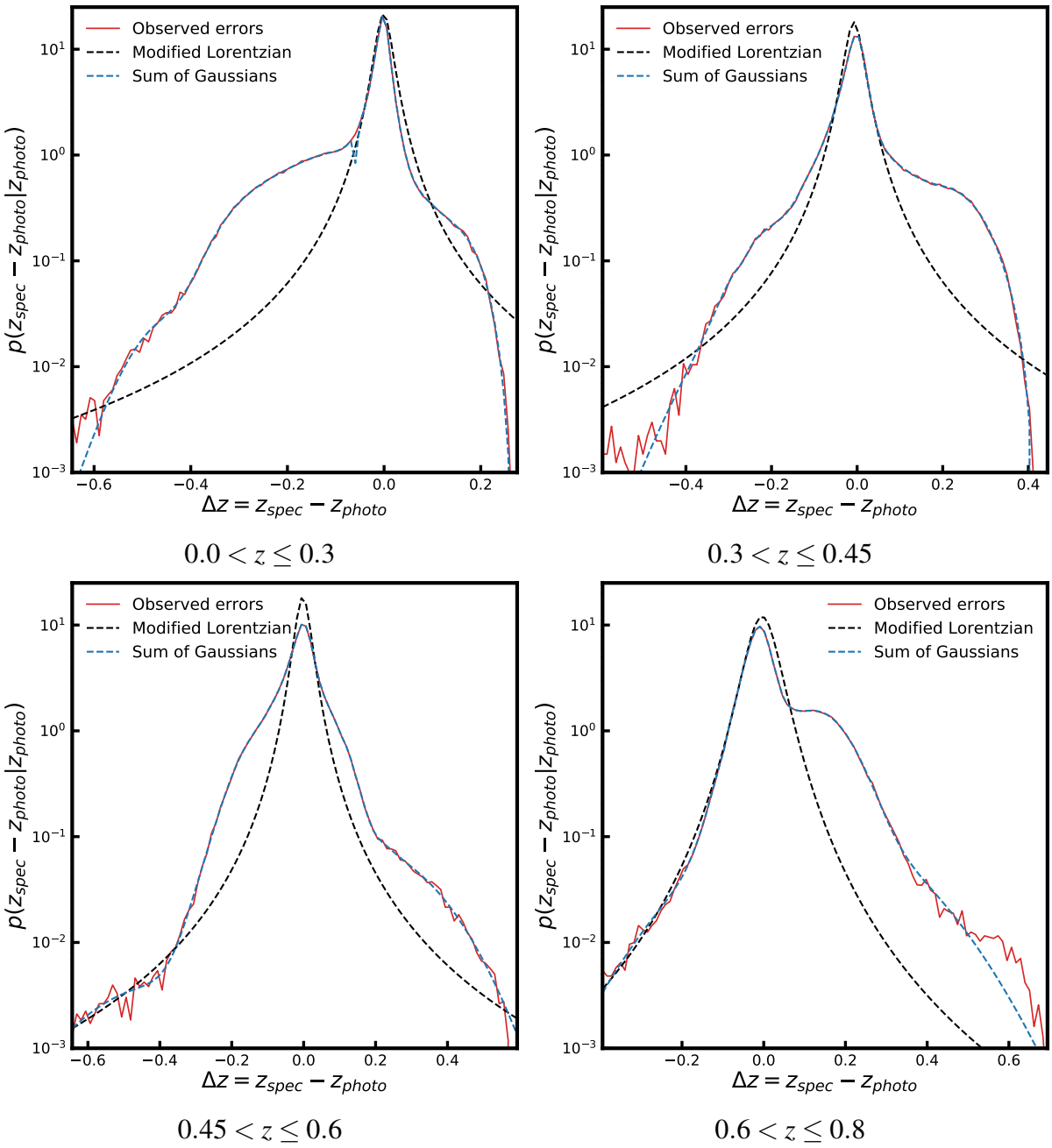


Figure 5.2: Photometric redshift error function (shown in red solid curve) compared with the best-fit modified Lorentzian function obtained by Hang et al., 2021 (black dashed curve) and the sum of 10 Gaussians (blue solid curve).

redshift bins (as described in Chapter 4).

5.4.1 Using convolution method

The true redshift distribution for tomographic bin i can be computed using the relation (section 4.3.1)

$$\frac{\mathrm{d}N^i}{\mathrm{d}z_t} = \int \mathrm{d}z_p \frac{\mathrm{d}N^i(z_p)}{\mathrm{d}z_p} p^i(z_s - z_p|z_p) \tag{5.3}$$

where $\frac{dN^i(z_p)}{dz_p}$ is the observed photometric distribution of galaxies in the i^{th} redshift bin and $p^i(z_s - z_p|z_p)$ is the fit made to the corresponding photometric redshift error distribution (from section 5.3). In Figure 5.3, we show the galaxy redshift distributions for every tomographic bins estimated using Eq. (5.3). The left panel shows the redshift distributions estimated using the modified Lorentzian model and the right panel presents the distributions estimated using our sum of Gaussians approach.

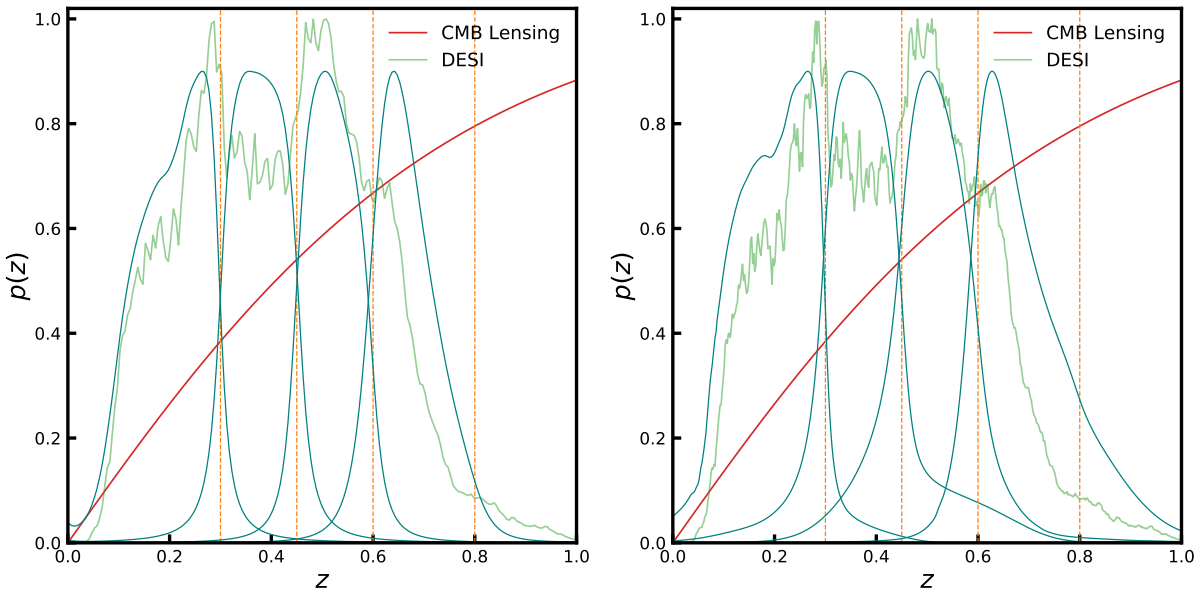


Figure 5.3: The redshift distribution for every tomographic bin shown in blue lines estimated using (*left:*) modified Lorentzian and (*right:*) sum of 10 Gaussians fit to the photometric redshift error distribution. The red line marks the CMB lensing kernel and the green line shows the total redshift distribution of galaxies. The orange vertical dashed lines mark the boundaries for the four tomographic bins. The CMB lensing kernel and redshift distributions are all normalized to unit maximum.

5.4.2 Using deconvolution method

We can also estimate the true redshift distribution by deconvoluting the observed full photometric redshift distribution (see section 4.3.2) with the error distribution $p(z_p - z_s|z_s)$ through the relation

$$\frac{dN}{dz_t} = \mathscr{F}^{-1}\left[\frac{\mathscr{F}[\frac{dN(z_p)}{dz_p}]}{\mathscr{F}[p(z_p - z_s|z_s)]}\right] \tag{5.4}$$

where $\mathscr{F}$ and $\mathscr{F}^{-1}$ represent the Fourier and inverse Fourier transforms, respectively.

We compute the error distribution $p(z_p - z_s|z_s)$ from the spectroscopic sample used to estimate the photometric redshifts of galaxies. As we can see in the left panel of Figure 5.4, fit made with a sum of 9 Gaussians very closely approximates the observed error distribution $p(z_p - z_s|z_s)$. In the right panel of Figure 5.4, we show the true redshift distribution estimated by the deconvolution method (i.e. using Eq. 5.4) compared with the observed photometric redshift distribution.

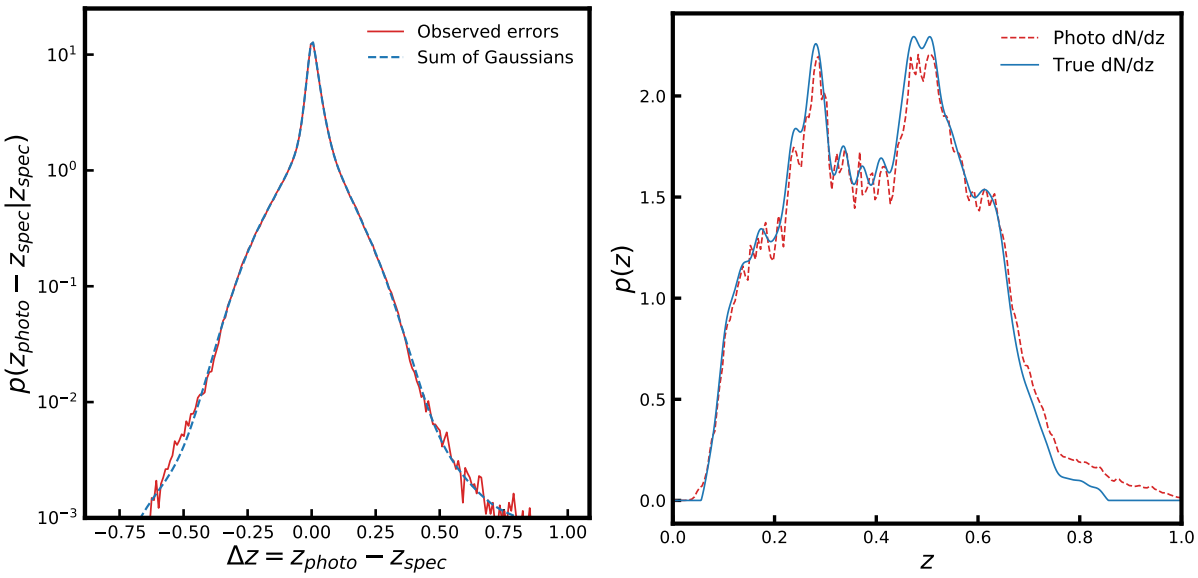


Figure 5.4: *Left*: Error distribution $p(z_p - z_s|z_s)$ fit of a sum of 9 Gaussians. *Right*: Photometric redshift distribution of galaxies (red dashed curve) compared with the true redshift distribution estimated using deconvolution method (blue solid curve).

5.5 Simulations

To get an insight into how the observed redshift errors will impact the parameters: galaxy linear bias and amplitude of cross-correlation, we use the publicly available `FLASK` code (Xavier et al., 2016), similar to section 4.2, and create 300 Monte Carlo simulations of correlated log-normal galaxy density field with DESI-LIS observed physical properties quoted in Table 5.1 and *Planck* CMB lensing convergence field. We assign true redshifts z_t to galaxies in our simulations following the estimate of the true redshift distribution obtained using the deconvolution method (shown in the right panel of Figure 5.4). We generate photometric redshifts using the observed photometric error distribution $p(z_p - z_s|z_s)$ shown in the left panel of Figure 5.4. The redshifts generated in this way closely follows the properties of the real DESI-LIS galaxy catalogue. We divide the simulated galaxy catalogue into 5 redshift bins with intervals $(0.0, 0.3, 0.45, 0.6, 0.8, 1.0]$. We note that due to photometric redshift errors, some objects from the redshift bin $0.6 < z < 0.8$ will also scatter to redshifts $z > 0.8$. Hence, to account for the lost objects outside the redshift boundary at $z = 0.8$, we also include the redshift bin $0.8 < z < 1.0$ when correcting for leakage. The fiducial power spectra necessary for simulations are computed using Eq. (2.8) with the redshift dependent galaxy bias $b(z) = 1 + \frac{b_0 - 1}{D(z)}$, with $b_0 = 1.3$.

From the simulated galaxy density and CMB lensing convergence maps, we estimate their full-sky cross-power spectrum and galaxy auto-power spectrum following the MASTER method described in section 2.3. In the top panel of Figure 5.5, we show the average galaxy auto-power spectra from 300 simulations for four tomographic bins. The shot noise contributions are subtracted from the average galaxy auto-power spectra. The bottom panel of Figure 5.5 presents the average cross-power spectra from 300 simulations. The red solid lines are the theoretical power spectra estimated using Eq. (5.3) without accounting redshift bin mismatch. The green dashed lines are the theoretical power spectra corrected for redshift bin mismatch using Eqs. 4.14

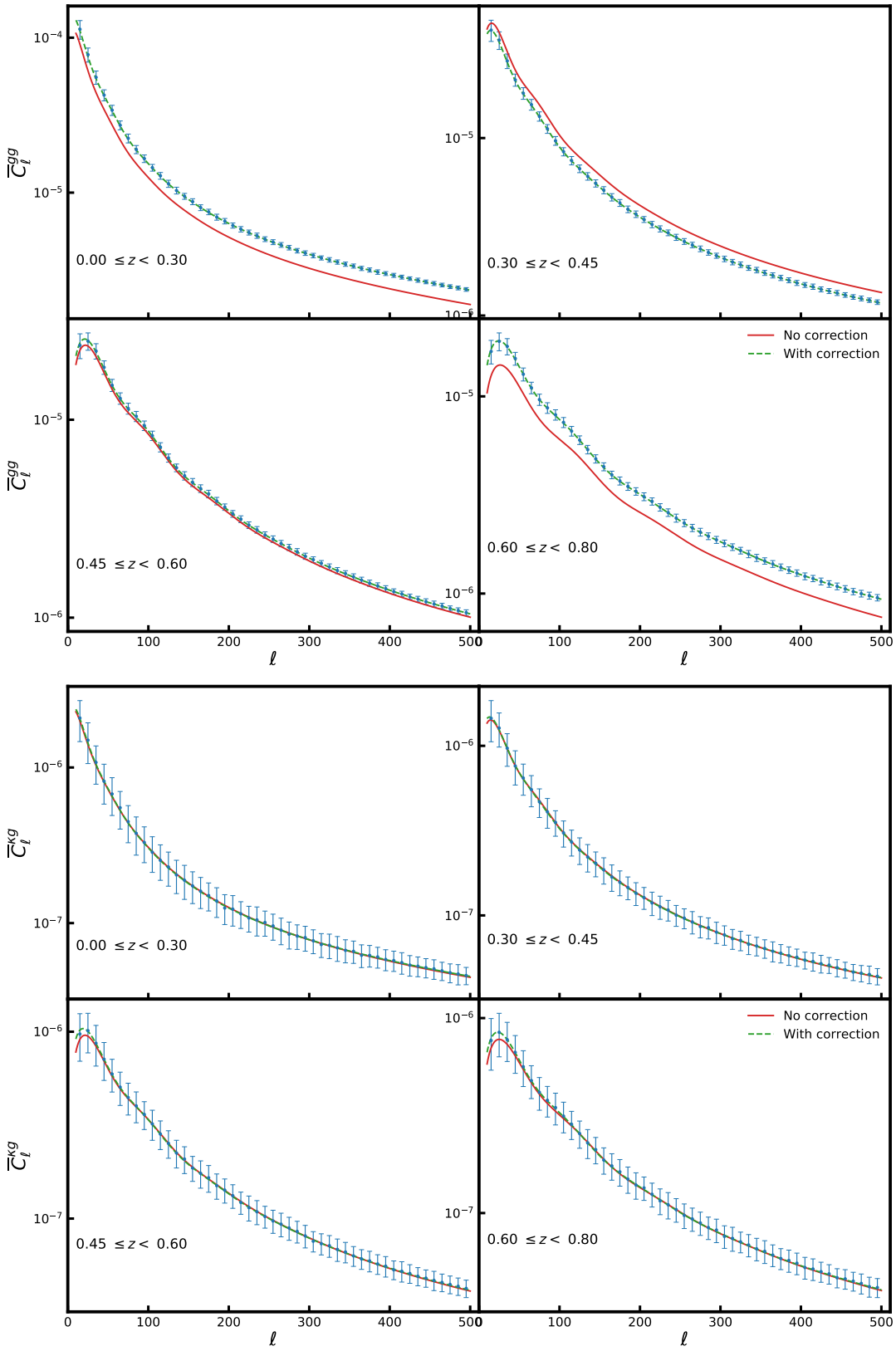


Figure 5.5: *Top:* Galaxy auto-power spectra and *bottom:* cross-power spectra averaged from 300 simulations for four tomographic bins. The red solid lines are the fiducial expectation estimated through Eq. (5.3) and the green dashed lines the theoretical power spectra corrected for the redshift bin mismatch. The error bars on the data points are estimated from 300 simulations using Eq. 4.11.

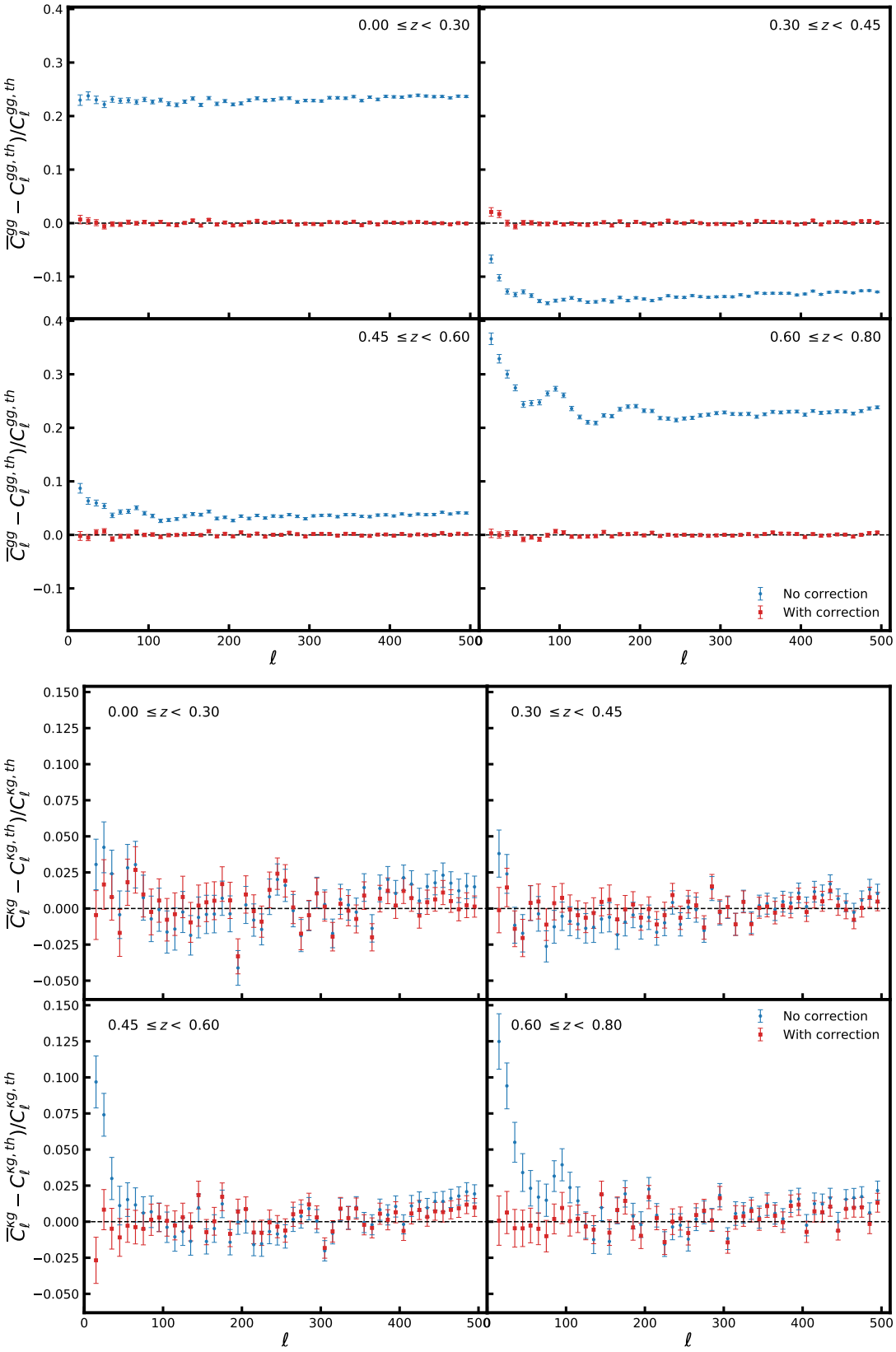


Figure 5.6: Relative errors on the average (*top*) galaxy auto-power spectra and (*bottom*) cross-power spectra without correcting for the redshift bin mismatch (blue circles) and after correcting (red squares). The error bars on the data points are estimated from 300 simulations using Eq. 4.11.

and 4.15.

The top panel of Figure 5.6 shows the relative differences between measured and theoretical auto- and cross-power spectra. There are shown the differences for theoretical spectra without correcting for the redshift bin leakage (i.e. redshift distribution given by Eq. 5.3) as well as for the spectra with the correction for the leakage (using Eqs. 5.6 and 5.7). For the former the galaxy auto-power spectra show deviations between $5-25\%$ due to photometric redshift errors. For each tomographic bin, the average observed galaxy auto-power spectrum is higher than their theoretical expectations, except for the redshift bin $0.35 \leq z < 0.4$ where the average power spectrum is $\sim 15\%$ lower. The leakage corrected theoretical galaxy auto-power spectra for each tomographic bin are completely consistent with measured power spectra. It is important to note the difference in the behaviour of the galaxy auto-power spectra between DESI-LIS simulations and the LSST simulations shown in Chapter 4. With LSST simulations the theoretical galaxy auto-power spectra before leakage correction were all larger than the measured power spectra, in contrast to that shown in Figure 5.5. This difference stems from the different behaviours of the photometric redshift error distributions when correcting for redshift scatters. Hence, it is of utmost importance to accurately quantify the error distributions when extracting the power spectra in cross-correlation analyses. In the bottom panel of Figure 5.6, we show the relative errors acquired by the cross-power spectra. Similar to the LSST simulations in Chapter 4, the cross-power spectra do not show significant variations due to leakage of objects across redshift bins.

5.5.1 Parameters

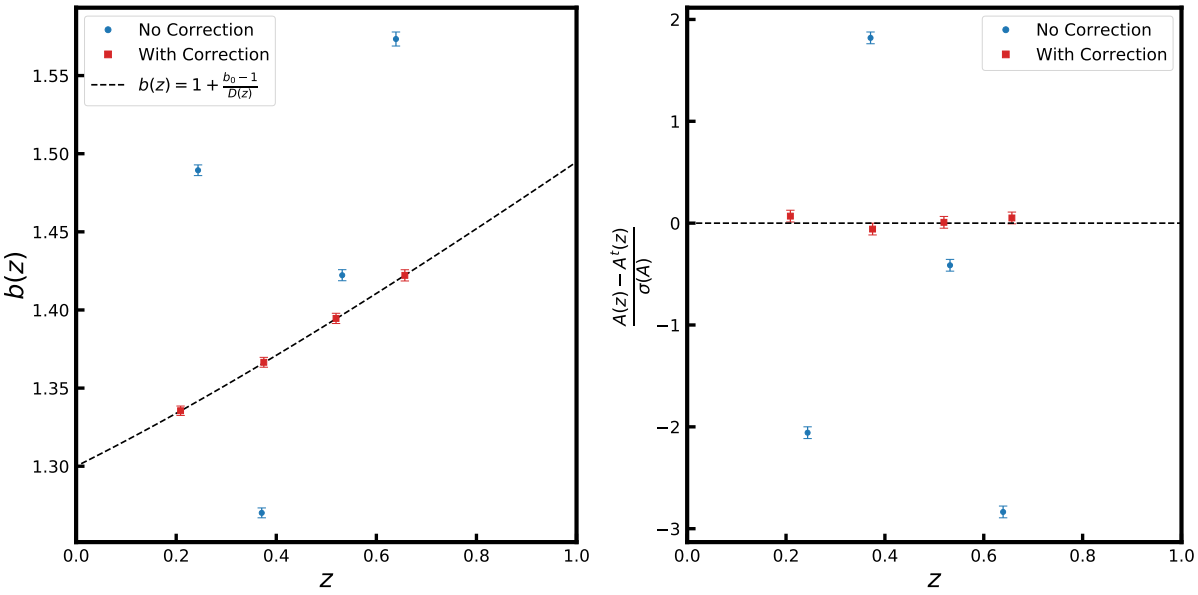


Figure 5.7: *Left*: Estimates for galaxy linear bias $b(z)$ before (blue circles) and after (red squares) correction for leakage. The black dashed lines marks the true evolution of galaxy linear bias (with $b_0 = 1.3$) used for simulations. *Right*: Deviation between the estimated amplitude of cross-correlation from its true value ($A = 1$) in terms of standard deviation of the amplitude.

From the average power spectra measured from simulations, we estimate galaxy linear bias

and amplitude of cross-correlation, using the Maximum Likelihood approach described in section 2.4. In the left panel of Figure 5.7, we compare the galaxy linear bias parameter before and after correcting the measured power spectra. The error bars have been re-scaled to report the errors for a single realisation, thus making it convenient to compare with actual DESI-LIS × *Planck* data analysis in section 5.6. The black dashed line represents the fiducial evolution of galaxy bias used for simulations, i.e. $b(z) = 1 + \frac{b_0 - 1}{D(z)}$ with $b_0 = 1.3$. Before leakage correction, the galaxy bias are estimated higher with $7 - 42\,\sigma$ deviations in three out of four tomographic bins. The second redshift bin gives a $\sim 30\sigma$ lower estimate of the galaxy bias. Such large deviations on the galaxy linear bias are due to very small errors on the estimated bias tightly constrained by the auto-power spectrum. In the right panel of Figure 5.7, we show the deviation between the estimated amplitude of cross-correlation from its true value ($A = 1$) in terms of standard deviation of the amplitude, from the four tomographic bins with error bars re-scaled to single realisation. The amplitude turns out to be lower than expectations with $\sim 0.5\,\sigma$ deviation in the third bin and $\sim 3\,\sigma$ deviation for the fourth bin. The amplitude is $\sim 2\,\sigma$ higher in the second tomographic bin. The amplitude of cross-correlation closely follows its expected value of unity after accounting for leakage across redshift bins. Thus, based on the results from our simulations we can expect the estimated amplitude to be smaller, except second redshift bin, than its true value if power spectra are not corrected for the photometric redshift scatter.

5.6 Results: DESI-LIS × *Planck*

We bin the galaxy auto-power spectra and cross-power spectra between DESI-LIS photometric galaxy catalogues and *Planck* minimum-variance CMB lensing potential map with binwidth $\Delta\ell = 10$ in the multipole range $10 < \ell < 500$, following the analysis choice of Hang et al., 2021. We use the Maximum likelihood Estimation method (described in section 2.4) to estimate two parameters, galaxy linear bias b and amplitude of cross-correlation A from the cross-correlation measurements. We use flat priors $b \in [0, 10]$ and $A \in [-5, 5]$ for the estimation of parameters while the remaining cosmological parameters are kept constant at flat ΛCDM cosmology with best-fit values given in Planck Collaboration et al., 2020a. In this section, we quantify the change in the parameters due to leakage across redshift bins.

5.6.1 Before leakage correction

Hang et al., 2021 estimated the parameters by adopting a two bias parameter model for the linear and non-linear regimes separately:

$$C_\ell^{\mathrm{gg,th}} = b_1^2 C_\ell^{\mathrm{lin,th}} + b_2^2 \Delta C_\ell^{\mathrm{nl,th}} \tag{5.5}$$

where $C_\ell^{\mathrm{lin,th}}$ and the non-linear correction $\Delta C_\ell^{\mathrm{nl,th}}$ are computed using the `CAMB`[4](Lewis et al., 2000) software. In our analysis, we use a redshift dependent bias model through $b(z) = 1 + \frac{b_0 - 1}{D(z)}$, where $D(z)$ is the growth factor computed through Eq. 4.2.

First we estimate parameters assuming theoretical power spectra without correction for the redshift bin mismatch (i.e. with redshift distribution estimated using Eq. 5.3) and the photo-

[4] https://camb.info/

Table 5.3: Best fit values for the galaxy linear bias b and amplitude of cross-correlation A estimated using theoretical power spectra without redshift bin leakage correction.. The first set of parameters b_1 and A are taken from Hang et al., 2021. The second and third sets of b and A are estimated using modified Lorentzian (with parameters from Hang et al., 2021) and sum of Gaussians fit to the error distribution $p(z_s - z_p|z_p)$, respectively. χ^2_r is the reduced chi-square values for the cross-power spectrum with $\nu = 47$ degrees of freedom.

Bin	From Hang et al.		This work					
			Modified Lorentzian			Sum of Gaussians		
	b_1	A	b	A	χ^2_r	b	A	χ^2_r
$(0.0, 0.3]$	$1.25^{+0.01}_{-0.01}$	$0.91^{+0.05}_{-0.05}$	$1.197^{+0.003}_{-0.003}$	$0.891^{+0.054}_{-0.053}$	1.621	$1.437^{+0.004}_{-0.003}$	$0.802^{+0.046}_{-0.046}$	1.177
$(0.3, 0.45]$	$1.56^{+0.02}_{-0.02}$	$0.80^{+0.04}_{-0.04}$	$1.513^{+0.004}_{-0.004}$	$0.850^{+0.044}_{-0.044}$	2.126	$1.731^{+0.004}_{-0.004}$	$0.780^{+0.040}_{-0.041}$	2.150
$(0.45, 0.6]$	$1.53^{+0.01}_{-0.01}$	$0.94^{+0.04}_{-0.04}$	$1.523^{+0.004}_{-0.004}$	$0.954^{+0.042}_{-0.042}$	1.450	$1.790^{+0.005}_{-0.004}$	$0.847^{+0.037}_{-0.036}$	1.440
$(0.6, 0.8]$	$1.83^{+0.02}_{-0.02}$	$0.91^{+0.04}_{-0.04}$	$1.853^{+0.005}_{-0.005}$	$0.901^{+0.038}_{-0.038}$	1.493	$2.220^{+0.006}_{-0.006}$	$0.812^{+0.033}_{-0.033}$	1.509

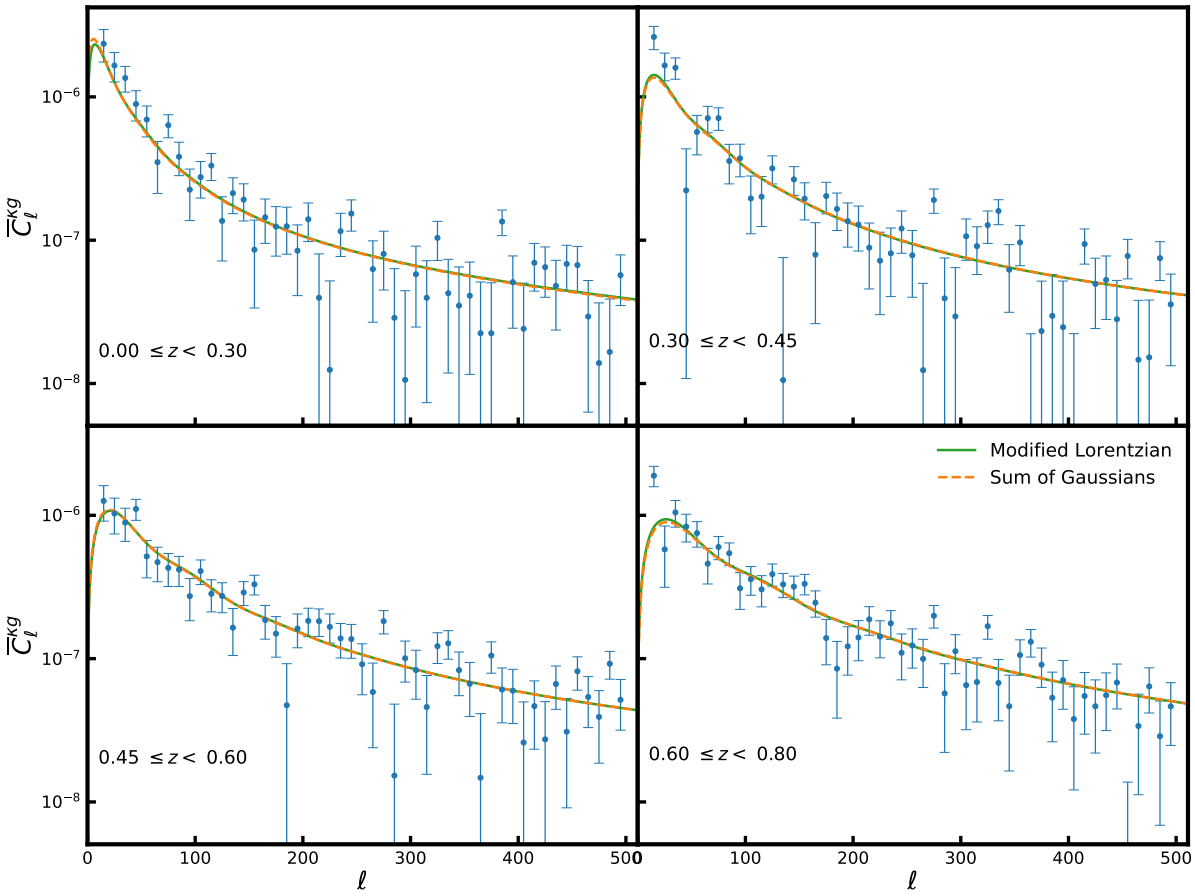


Figure 5.8: Cross-power spectrum measured from four DESI-LIS tomographic bins. The green solid line is the theoretical power spectrum computed following our best-fit estimates of parameters b and A using the modified Lorentzian fit to the error distribution. The orange dashed line is the theoretical power spectrum corresponding to the sum of Gaussians fit to the error distribution. The error bars are computed from the covariance matrix used in the likelihood function (section 2.4 with best-fit values of b and A from Table 5.3.

metric redshift error distribution modelled by modified Lorentzian function of Hang et al., 2021. In this case, we expect the values of the galaxy bias $b(z)$ to be close by the value of b_1 estimated by Hang et al., 2021. Since the sum of 10 Gaussians provides a better fit to $p(z_s - z_p|z_p)$, we also estimate parameters using our sum of Gaussians fit to the error distribution in Eq. 5.3. We present in Table 5.3, the values of parameter b_1 and cross-correlation amplitude A as quoted in table 3 of Hang et al., 2021 and compare it with our best-fit estimates of b and A computed using the modified Lorentzian (with parameters from Hang et al., 2021) and sum of Gaussians fit to the error distribution $p(z_s - z_p|z_p)$, respectively. The amplitude of cross-correlation computed with modified Lorentzian fit are consistent with estimations of Hang et al., 2021 within $1\,\sigma$ errors. The galaxy bias also follows the values of b_1 from Hang et al., 2021 as expected, with maximum difference occurring in the first tomographic bin. The estimated cross-correlation amplitude is in tension at $1.5-2.5\,\sigma$ with respect to its fiducial value of unity.

The values of parameters when computed using the sum of Gaussians approach, differ significantly from the previous estimations, producing consistently smaller values of cross-correlation amplitude and higher values of the galaxy linear bias. The tension in the amplitude increase to $4.3-5.6\,\sigma$ with sum of Gaussians approach. We quote the reduced chi-square values for the measured cross-power spectrum with $\nu = 49-2 = 47$ degrees of freedom in Table 5.3. In Figure 5.8 we show the measured cross-power spectrum and the fiducial power spectrum computed using the values of b and A quoted in Table 5.3 for the four tomographic bins. Both modified Lorentzian and sum of Gaussians approaches yield similar results in their fit to the cross-power spectrum and chi-square values, with striking differences in the values of parameters b and A. The change in parameters are solely due to the differences in modelling the photometric redshift error distributions. Since, the sum of Gaussians model provide a better fit to the photometric redshift error distribution, we treat the parameters computed under this model as our baseline results when correcting for redshift bin mismatch.

5.6.2 With leakage correction

In Chapter 4 and section 5.5, we have shown that the parameters estimated in a tomographic study will be biased unless the leakage of objects across redshift bins is accounted through scattering matrix. We prepare the galaxy over-density map for the new tomographic bin $0.8 < z < 1.0$ in a similar fashion to section 5.2 and measure the galaxy auto-power spectrum and cross-power spectrum with minimum-variance *Planck* CMB lensing convergence map. The addition of this new redshift bin is necessary for correct estimation of the scattering matrix as explained in section 5.5. In the left panel of Figure 5.9, we show the scattering matrix computed for the DESI-LIS photometric catalogue following the methodology described in section 4.5. In the right panel of Figure 5.9 is the mean scattering matrix estimated from 300 Monte Carlo DESI-LIS simulations described in section 5.5. Since we used the DESI-LIS photometric error distribution in our simulations, the similar structure of the scattering matrix from the observed DESI-LIS datasets and from our simulations indicate that the scattering matrix shown in the left panel Figure 5.9 will be robust in correcting the power spectra of the leakage for every tomographic bin.

The measured galaxy auto- and cross-power spectra can finally be corrected for the redshift

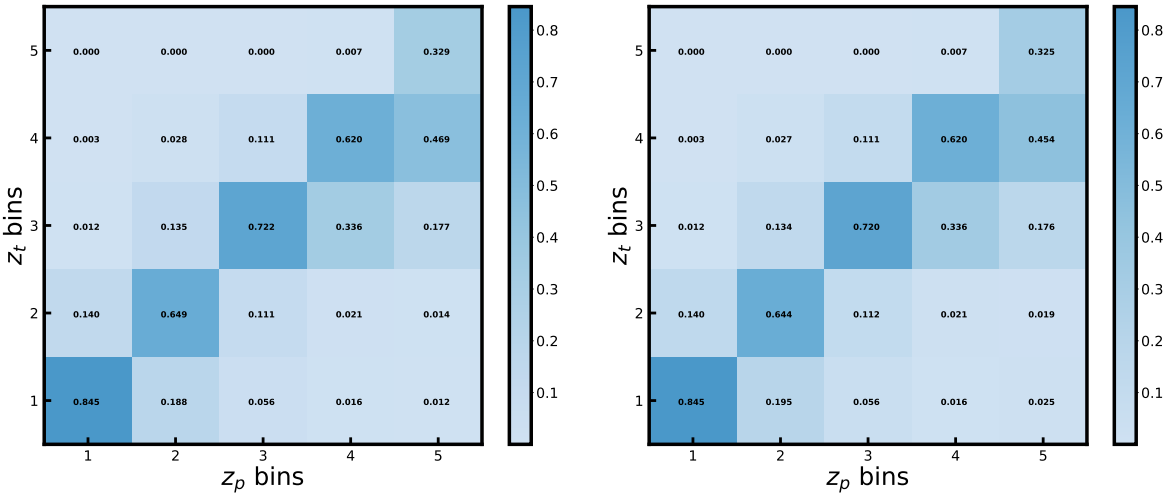


Figure 5.9: *Left*: The scattering matrix estimated for DESI-LIS data. *Right*: The mean scattering matrix estimated from 300 Monte Carlo simulations of DESI-LIS data (described in section 5.5). z_p and z_t bins represent the photometric and true redshift bins, respectively.

bin leakage using the relations

$$\hat{C}^{gg,\mathrm{tr}} = P^{\mathrm{T}^{-1}} \hat{C}^{gg,\mathrm{ph}} P^{-1} \tag{5.6}$$

$$\hat{C}^{\kappa g,\mathrm{tr}} = P^{\mathrm{T}^{-1}} \hat{C}^{\kappa g,\mathrm{ph}} \tag{5.7}$$

where P and P^{T} are the estimated scattering matrix and its transpose. $\hat{C}^{gg,\mathrm{ph}}$ and $\hat{C}^{\kappa g,\mathrm{ph}}$ are the power spectra measured from DESI-LIS catalogue and its cross-correlation with *Planck* CMB lensing map (power spectra presented in section 5.6.1). $\hat{C}^{gg,\mathrm{tr}}$ and $\hat{C}^{\kappa g,\mathrm{tr}}$ represents the estimates of the true power spectra. We use $\hat{C}^{gg,\mathrm{tr}}$ and $\hat{C}^{\kappa g,\mathrm{tr}}$ to estimate parameters after correction for redshift bin mismatch of objects, for reasons described in section 4.6. To compute the parameters from the estimates of the true power spectra using Maximum Likelihood Estimation method, we use the fiducial true power spectrum template for each tomographic bin computed using Eqs. (4.17) and (4.18) with redshift distribution $\frac{\mathrm{d}N^i(z_t)}{\mathrm{d}z_t}$ given by Eqs. (5.4) and (4.8). In Table 5.4, we compare the parameters b and A computed before and after leakage correction for our baseline results.

As expected from our simulations in section 5.5, we observed significant reduction in the galaxy linear bias after leakage correction, expect for the second tomographic bin where the simulations predicted an increase in the galaxy bias after correction. As a consequence of degeneracy between galaxy bias and the amplitude for cross-power spectrum, smaller values of galaxy bias result in increase of the amplitude of cross-correlation by $\sim 2\,\sigma$ for the first and third tomographic bin and by $\sim 3\,\sigma$ for the last tomographic bin making it fully consistent with the ΛCDM predictions. The amplitude in the second redshift bin, however, does not show any substantial variation after correction. We also estimate better χ^2-values for the cross-power spectra after leakage correction. Although, the galaxy linear bias and cross-correlation amplitude estimated taking into account the redshift bin mismatch turn out to be comparable to the estimates by Hang et al., 2021, it is worth to keep in account the differences in modelling of photometric redshift error distributions as depicted in Figure 5.2. In Figure 5.10, we show the leakage corrected

Table 5.4: Best fit values for the galaxy linear bias b and amplitude of cross-correlation A for every tomographic bin estimated with and without taking into account leakage correction in the analysis. χ^2_r is the reduced chi-square values for the cross-power spectrum with $\nu = 47$ degrees of freedom.

Bin	From Hang et al.		This work					
			Before correction			After correction		
	b_1	A	b	A	χ^2_r	b	A	χ^2_r
$(0.0, 0.3]$	$1.25^{+0.01}_{-0.01}$	$0.91^{+0.05}_{-0.05}$	$1.437^{+0.004}_{-0.003}$	$0.802^{+0.046}_{-0.046}$	1.177	$1.189^{+0.004}_{-0.003}$	$0.916^{+0.060}_{-0.060}$	1.004
$(0.3, 0.45]$	$1.56^{+0.02}_{-0.02}$	$0.80^{+0.04}_{-0.04}$	$1.712^{+0.004}_{-0.004}$	$0.786^{+0.041}_{-0.041}$	2.150	$1.594^{+0.004}_{-0.004}$	$0.797^{+0.054}_{-0.054}$	1.367
$(0.45, 0.6]$	$1.53^{+0.01}_{-0.01}$	$0.94^{+0.04}_{-0.04}$	$1.819^{+0.005}_{-0.005}$	$0.838^{+0.036}_{-0.036}$	1.440	$1.581^{+0.004}_{-0.004}$	$0.926^{+0.049}_{-0.049}$	0.805
$(0.6, 0.8]$	$1.83^{+0.02}_{-0.02}$	$0.91^{+0.04}_{-0.04}$	$2.159^{+0.006}_{-0.006}$	$0.815^{+0.033}_{-0.033}$	1.509	$1.768^{+0.004}_{-0.004}$	$0.960^{+0.051}_{-0.051}$	1.213

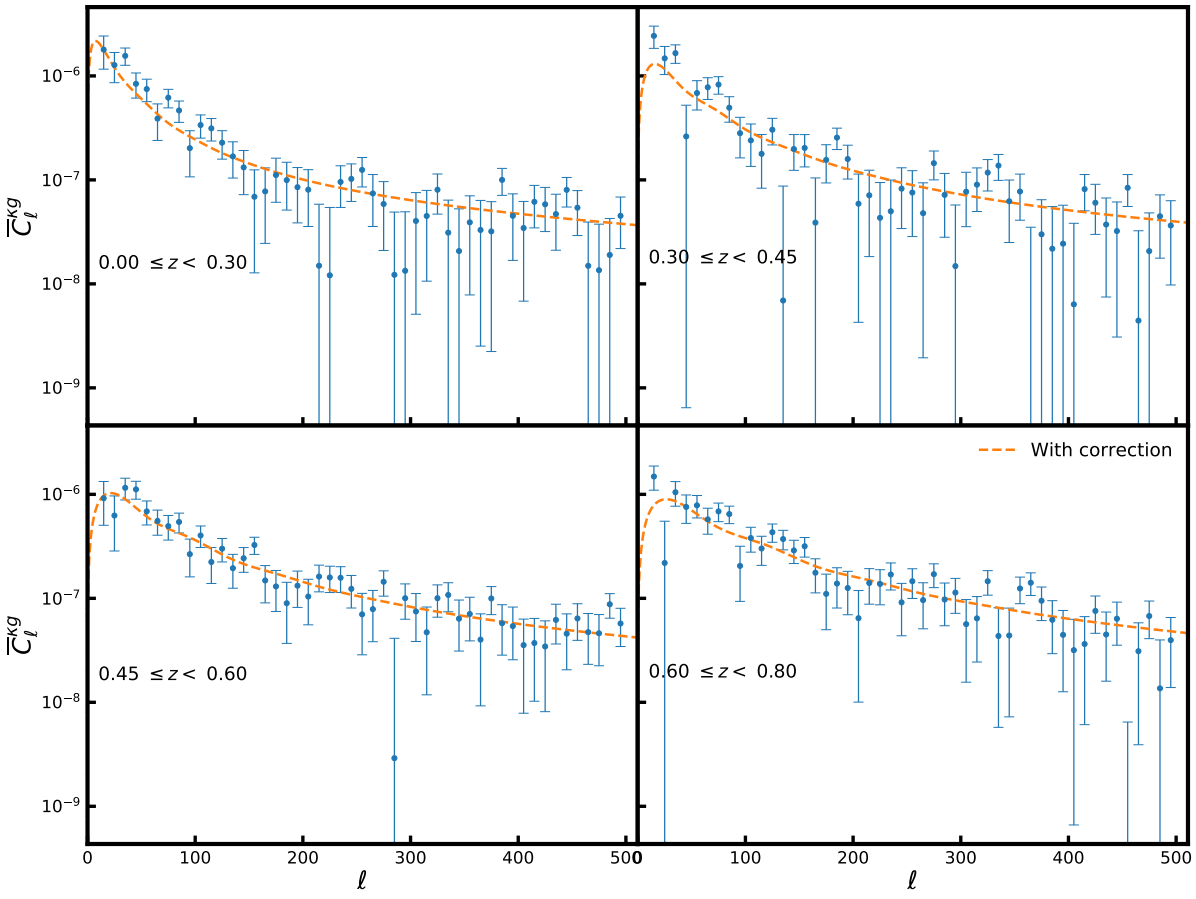


Figure 5.10: Cross-power spectrum measured from four DESI-LIS tomographic bins with MV map after leakage correction. The orange dashed line is the theoretical power spectrum computed with the best-fit values of parameters quoted in Table 5.4. The error bars are computed from the covariance matrix used in the likelihood function (section 2.4 with best-fit values of b and A from Table 5.4.

cross-power spectra for the four tomographic bins and the corresponding theoretical fits using the best-fit parameters mentioned in Table 5.4.

Having estimates of the galaxy bias in different redshift bins we can estimate b_0 parameter,

which can be interpreted as galaxy bias at redshift $z = 0$, in our model of redshift dependence of the galaxy bias $b(z) = 1 + \frac{b_0 - 1}{D(z)}$. We find $b_0 = 1.581 \pm 0.008$ and 1.386 ± 0.008 without and with leakage correction, respectively. In the left panel of Figure 5.11 we show the estimated galaxy linear bias for the four tomographic bins considered in the analysis. The blue circles and red squares represent the galaxy bias estimates without and with correction for leakage of objects across redshift bins. The black dashed and solid lines represent the resulting redshift evolution of galaxy bias with and without leakage correction, respectively. The fit of data points to the best-fitted models is rather poor as the data go up steeper with redshift than lines for our redshift dependent galaxy bias model. Accounting for the scatter of objects between redshift bins leads to significantly lower value of b_0 and will result in notably different inferences about the relation between the dark matter and luminous matter.

In the right panel of Figure 5.11, we present the the deviations between the estimated amplitude of cross-correlation from its true value ($A = 1$) in terms of standard deviation of the amplitude, from the four tomographic bins. By countering the impact of redshift bin mismatch, we reduce the tension in amplitude from $4 - 6\,\sigma$ to $\sim 2\,\sigma$, with complete agreement within errors for the last tomographic bin.

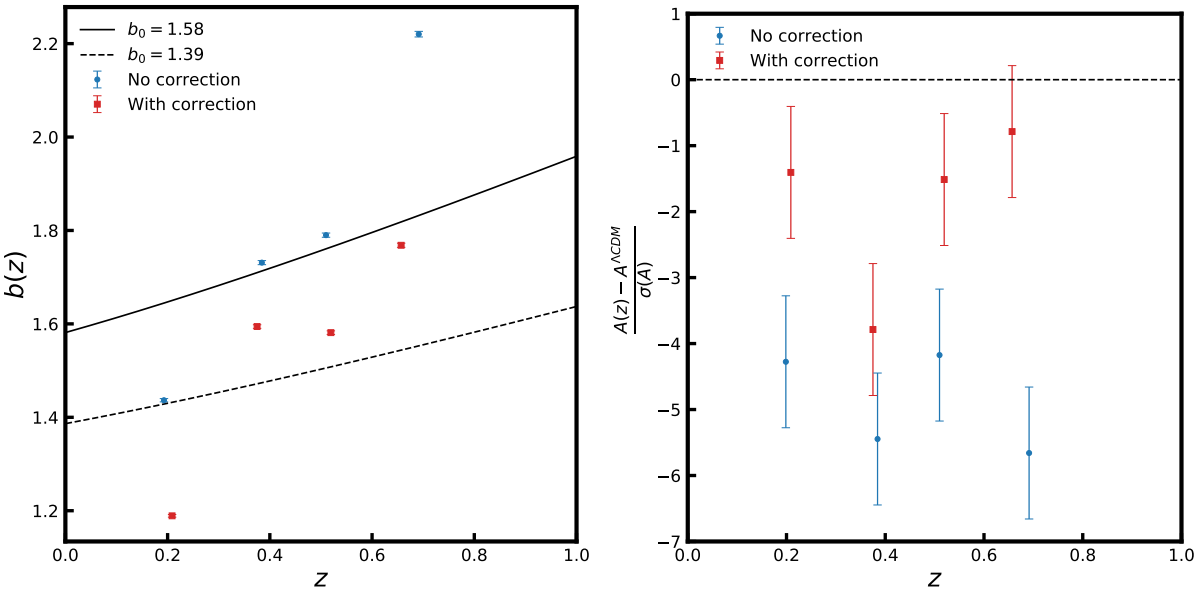


Figure 5.11: *Left*: Estimates for galaxy linear bias $b(z)$ with (red squares) and without (blue circles) taking into account the correction for leakage. The black solid and dashed lines represent the fiducial redshift evolution of galaxy bias for before and after leakage correction scenarios, respectively, with b_0 estimated using Maximum Likelihood Estimation. *Right*: Deviation between the estimated amplitude of cross-correlation from its true value ($A = 1$) in terms of standard deviation of the amplitude.

5.6.3 Using different CMB lensing potential maps

Although accounting for leakage improves the deviations observed on the amplitude of cross-correlation for our baseline analysis, there still remains a $\sim 2\,\sigma$ tension with respect to the prediction of the standard cosmological model. There are a number of systematics that we explored in chapter 3 but have not considered with DESI-LIS datasets, since the goal of this

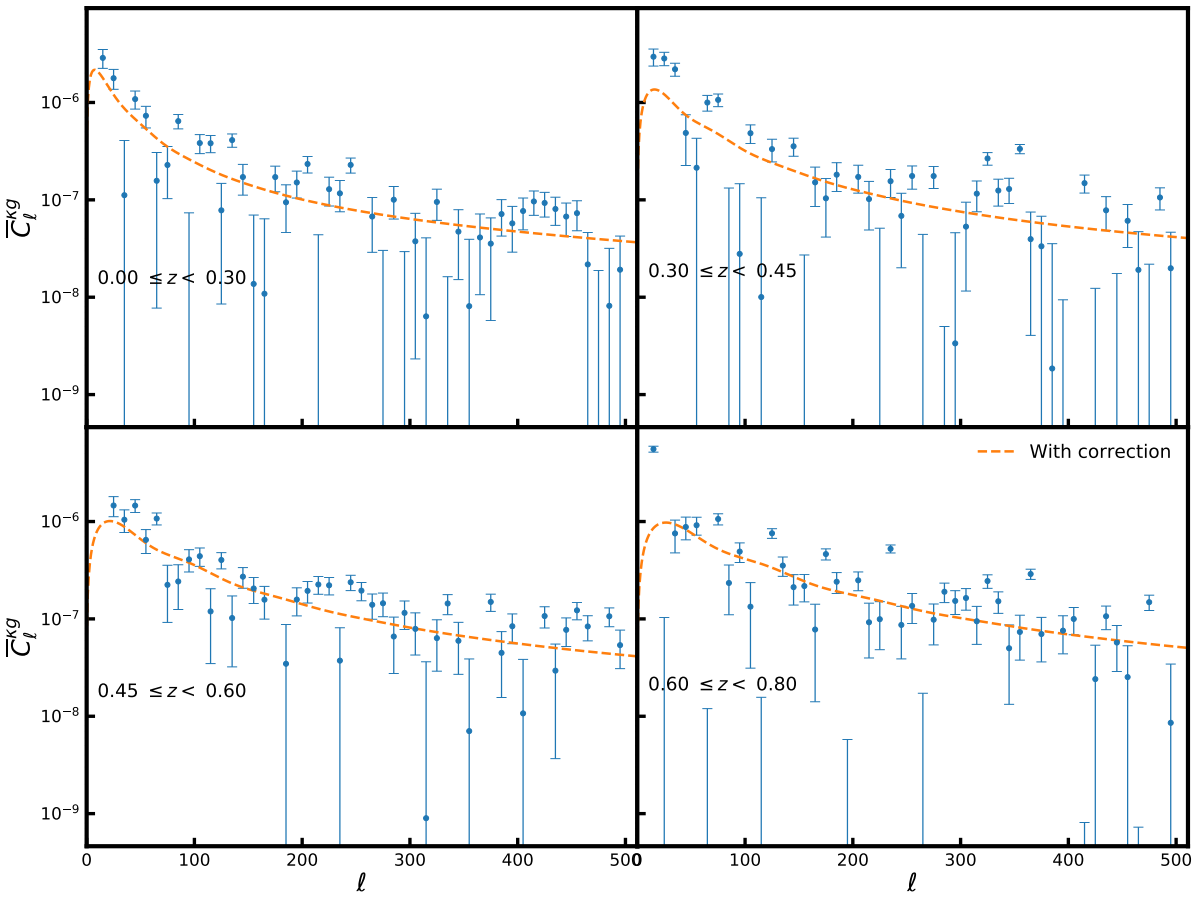


Figure 5.12: Cross-power spectrum measured from four DESI-LIS tomographic bins with SZ-deproj map after leakage correction. The orange dashed line is the theoretical power spectrum computed with the best-fit values of parameters quoted in Table 5.4. The error bars are computed from the covariance matrix used in the likelihood function (section 2.4 with best-fit values of b and A from Table 5.5.

analysis was to convey the importance of leakage correction on the estimation of cosmological parameters. However, one of the important conclusions from chapter 3 is that different CMB lensing convergence maps produce significantly different values of cross-correlation amplitude parameter. Hence, in this section, we compare estimates of the galaxy bias and cross-correlation amplitude obtained using also other provided by the *Planck* team CMB lensing convergence maps i.e. Sunyaev-Zeldovich deprojected (SZ-deproj) and temperature only (TT) convergence maps. The minimum-variance (MV) map used in the baseline analysis combines the *Planck* temperature and polarization measurements through a minimum-variance approach, whereas the TT map is reconstructed from only the *Planck* temperature measurements. The SZ-deproj map further removes the regions of sky with strong Sunyaev-Zeldovich sources from the temperature measurements used for reconstruction of the lensing convergence map.

In Table 5.5 we show the best-fit values of parameters b and A estimated using Maximum Likelihood Estimation method for cross-correlation of DESI-LIS tomographic bins with SZ-deproj and TT CMB lensing convergence maps. The parameters estimated with the SZ-deproj map are found to be $\sim 1\,\sigma$ higher in the second and $\sim 2\,\sigma$ for the fourth tomographic bin, in comparison with MV map. With SZ-deproj map, the amplitude in the last tomographic bin is estimated at $A = 1.043 \pm 0.059$, in agreement with the ΛCDM expectations of unity. However, the

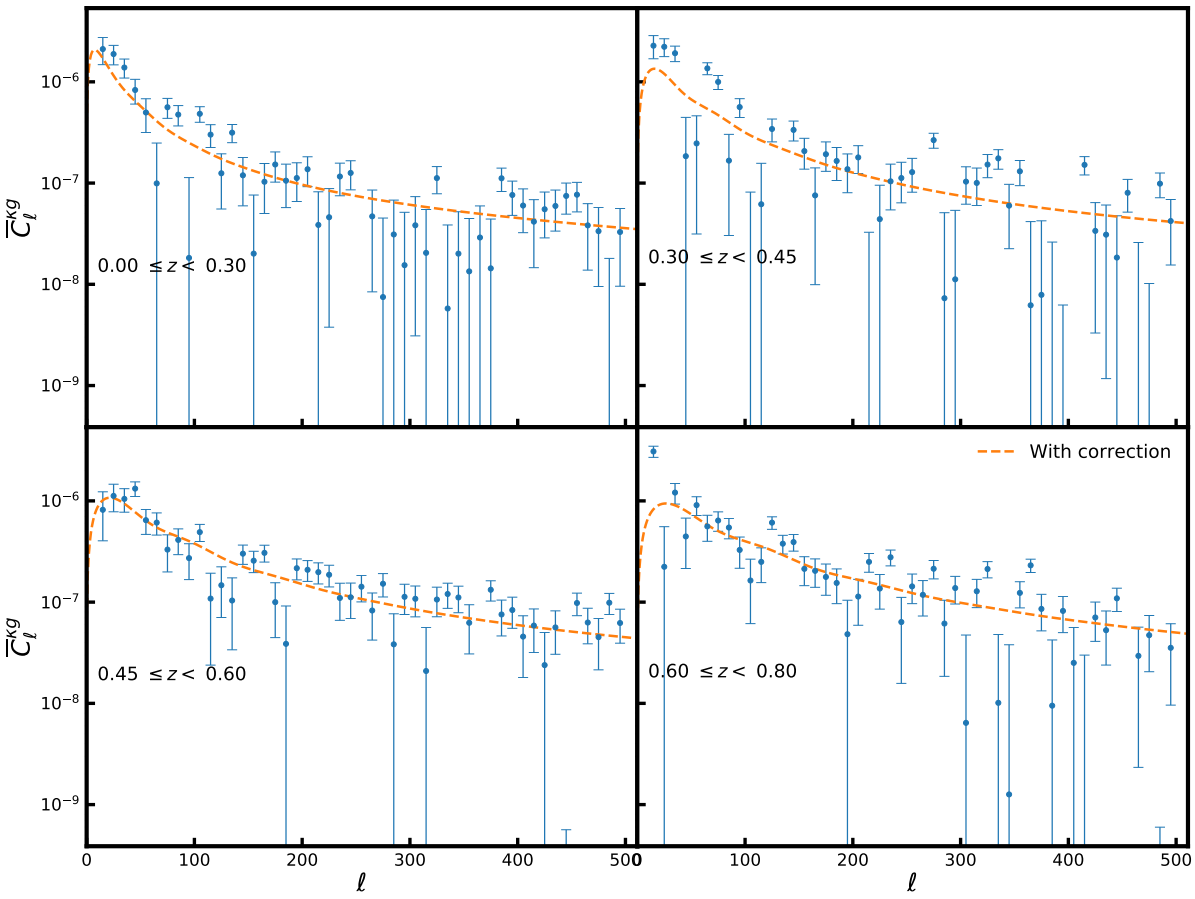


Figure 5.13: Cross-power spectrum measured from four DESI-LIS tomographic bins with TT map after leakage correction. The orange dashed line is the theoretical power spectrum computed with the best-fit values of parameters quoted in Table 5.4. The error bars are computed from the covariance matrix used in the likelihood function (section 2.4 with best-fit values of b and A from Table 5.5.

Table 5.5: Galaxy linear bias and cross-correlation amplitude from Maximum Likelihood Estimation from DESI-LIS tomographic bins (taking into account leakage correction) using joint likelihood functions for SZ-deproj and TT-only convergence maps.

Bin	SZ-deproj			TT		
	b	A	χ^2_r	b	A	χ^2_r
(0.0, 0.3]	$1.190^{+0.003}_{-0.003}$	$0.914^{+0.069}_{-0.069}$	2.720	$1.189^{+0.003}_{-0.003}$	$0.875^{+0.066}_{-0.067}$	1.781
(0.3, 0.45]	$1.596^{+0.004}_{-0.004}$	$0.831^{+0.062}_{-0.062}$	5.095	$1.595^{+0.004}_{-0.004}$	$0.823^{+0.060}_{-0.060}$	3.706
(0.45, 0.6]	$1.583^{+0.004}_{-0.004}$	$0.910^{+0.057}_{-0.056}$	4.266	$1.582^{+0.004}_{-0.004}$	$0.967^{+0.054}_{-0.054}$	1.748
(0.6, 0.8]	$1.772^{+0.004}_{-0.004}$	$1.043^{+0.059}_{-0.059}$	8.815	$1.769^{+0.004}_{-0.004}$	$1.009^{+0.056}_{-0.057}$	3.218

reduced χ^2 values for SZ-deproj map are consistently poorer than the MV map, due to increased fluctuations in the measured cross-power spectra. The TT map, on the other hand, removes the tension on the cross-correlation amplitude for the third and fourth tomographic bins. The amplitude for the second bin is also found to be $\sim 0.6\sigma$ higher than the MV map. Figures 5.12 and 5.13 show the cross-power spectra for the four tomographic bins with SZ-deproj and TT

maps, respectively, after correction for redshift bin mismatch. The orange dashed lined are the theoretical fits using the best-fit parameters mentioned in Table 5.4. The overall estimates of the cross-correlation amplitude for SZ-deproj and TT maps show better agreement with expected value of one than estimate for MV map, however, the reduced χ^2 values from SZ-deproj and TT maps indicate poorer fit of the theoretical power spectra to data than for MV map.

5.6.4 Estimation of σ_8 parameter

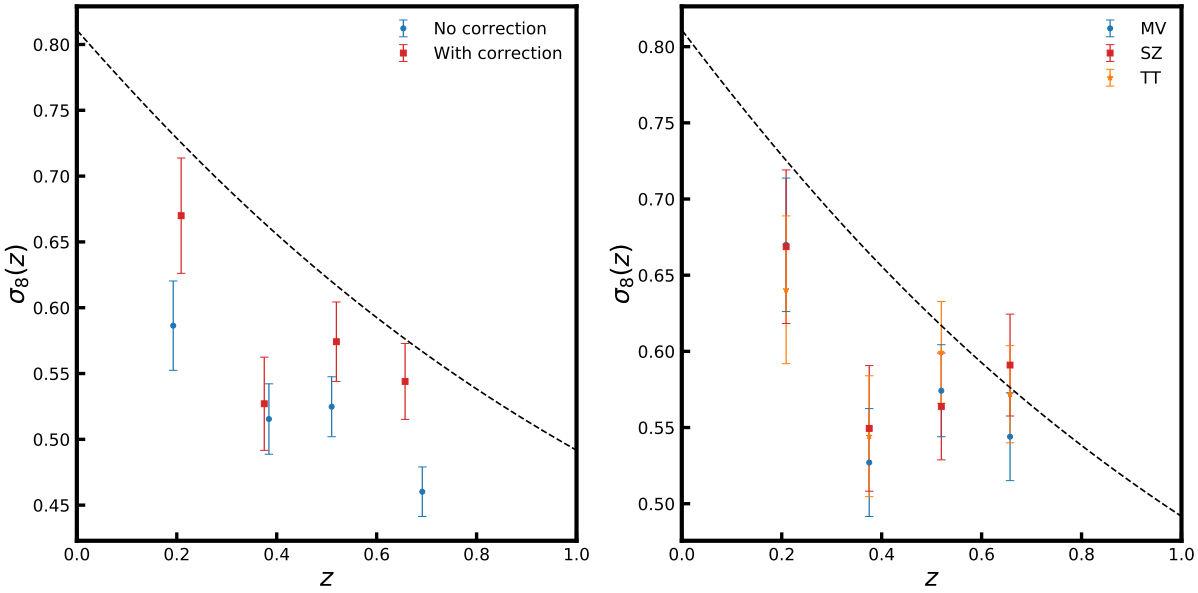


Figure 5.14: *Left*: Comparison of σ_8 parameter computed using cross-correlation measurement with MV map before and after leakage correction *Right*: Comparison of σ_8 parameter after leakage correction computed from cross-correlations with MV, SZ-deproj and TT CMB lensing maps. The dashed line represents the redshift evolution of the σ_8 parameter for our fiducial cosmology.

The amplitude of cross-correlation can be translated to the σ_8 parameter as suggested by Peacock and Bilicki, 2018. We follow the procedure from section 4.7 to estimate the σ_8 parameter from the cross-correlation measurements between DESI-LIS galaxy survey and *Planck* CMB lensing potential using Eq. (4.19). We show in the left panel of Figure 5.14, the impact of leakage correction on σ_8 parameter. The σ_8 parameter before leakage correction is estimated from our baseline analysis using the sum of Gaussians fit to the error distribution. The black dashed line is the redshift evolution of σ_8 for our fiducial cosmology. We observe improvement in the σ_8 values after correcting for leakage, similar to the amplitude of cross-correlation. In the right panel of Figure 5.14, we compare the σ_8 parameter computed from the cross-correlation measurements with MV, SZ-deproj and TT CMB lensing maps. As with the amplitude of cross-correlation, the best-fit values of the σ_8 parameter estimated from cross-correlations with TT and SZ-deproj maps show better agreement with expected values for the fiducial ΛCDM model. The TT map provides completely consistent values of σ_8 with expectations for the last two higher redshift bins.

5.7 Summary

In this chapter, we have performed the cross-correlation analysis between minimum-variance CMB convergence map from Planck Collaboration et al., 2020b and photometric galaxy catalogues from the Data release 8 of the Legacy Imaging Survey prepared by Hang et al., 2021. The galaxy density field is divided in to four redshift slices covering the range $0 < z < 0.8$, with photometric redshift precision $\frac{\sigma_z}{1+z}$ in the range $0.012 - 0.015$. Hang et al., 2021 estimated a $1.5 - 2.5\,\sigma$ tension in the value of cross-correlation amplitude from all tomographic bins with respect to the standard cosmological model.

Hang et al., 2021 approximated the photometric redshift error distribution with a modified Lorentzian distribution. We show that this model does not appropriately capture the tails of the error distribution. In our baseline analysis, we adopt a sum of Gaussians model and make a precise fits to the redshift error distributions. We find that our sum of Gaussians model gives $\sim 2\,\sigma$ smaller values of cross-correlation amplitude, and significantly higher estimates of the galaxy linear bias than Hang et al., 2021. This clearly shows that precise modelling of the redshift error distributions plays a crucial role in the estimation of cosmological parameters.

We have shown in Chapter 4 that tomographic analyses will produce biased estimates of parameters due to redshift bin mismatch of objects. We correct for this scatter of objects across redshift bins using our scattering matrix formalism developed in Chapter 4. For our baseline analysis with sum of Gaussians fit to the redshift error distributions, we find a $\sim 2 - 3\,\sigma$ improvement in the amplitude of cross-correlation after correction for leakage. We show through simulations that the scattering matrix used to correct the power spectra for the redshift bin mismatch also gives a $\sim 2 - 3\,\sigma$ improvement in cross-correlation amplitude. The correction reduces the tension from $4 - 5\,\sigma$ to $\sim 2\sigma$ for three tomographic bins, whereas alleviating the tension for the last bin. We also estimate galaxy linear bias and cross-correlation amplitude using Sunyaev-Zeldovich deprojected (SZ-deproj) and temperature-only (TT) reconstructed CMB lensing convergence maps. The SZ-deproj and TT maps produce $1 - 2\,\sigma$ higher values of amplitude than MV map, in agreement with the conclusion from Chapter 3, but with poorer χ^2 values. The TT map completely resolves the deviation on the amplitude for last two tomographic bins.

Finally, we estimate the impact of leakage correction on the σ_8 parameter and find that the leakage corrected power spectra yields better estimates, reducing the tension from $4 - 5\,\sigma$ to $\sim 1\,\sigma$ for three out of four redshift bins. The $\sigma_8 - \Omega_m$ tension is one of the biggest challenges of the modern cosmology. We conclude that an accurate modelling of the photometric redshift error distribution and accounting for the scatter of objects across redshift bins due to photometric redshift errors are very important for unbiased estimations of the cosmological parameters. In this light, it becomes crucial to include scattering matrices in cross-correlation analyses with future datasets.

Chapter 6

Summary of the thesis

An overview of each of the preceding chapters constituting the backbone of this thesis can be found in Section 1.5. Here we summarize the most important results and conclusions of the studies presented in the thesis.

The main goal of the collection of works in this thesis was to investigate the role of various systematics errors on testing cosmological models through cross-correlation measurements between CMB lensing potential and tracers of the large scale structure, with particular attention paid to tomographic cross-correlation analyses. The important contributions from this thesis are:

1. **First cross-correlation measurements** between *Planck* CMB lensing potential map and photometric redshift galaxy catalogues from the *Herschel* Extragalactic Legacy Project (HELP).

2. **The robust modelling of various systematics that affect the redshift distribution and number density of objects in a galaxy survey** (see Chapter 3). The systematics such as photometric calibration errors, catastrophic errors, magnification bias, etc will bias the cross-correlation measurements. We provide mitigation strategies for these galaxy survey systematics through Chapter 3.

3. **The redshift bin mismatch of objects due to photometric redshift errors can** cause biased estimation of cosmological parameters from cross-correlation analysis. This will **lead to apparent tensions on cosmological parameters like σ_8**. The incurred bias depends on the distribution of photometric redshift errors, the redshift distribution of objects and the size of the redshift bin. We present one of the **first tests of performance of the tomographic cross-correlation analysis based on Monte Carlo simulations**. In the case of LSST survey and photometric redshift error $\sigma_0 = 0.02$, the bias is of order of $1\,\sigma$ for the cross-correlation amplitude or σ_8 parameter. The bias can be corrected using scattering matrix approach developed in Chapter 4. Although the scattering matrix formalism on its own was introduced by Zhang et al., 2010, **we propose in the thesis a new, computationally faster method of estimation of the matrix** well-suited for analysis of large datasets from the upcoming galaxy surveys.

4. **Application of the scattering matrix** approach in the cross-correlation analysis for the

Planck lensing map and DESI-LIS catalogue (Hang et al., 2021) results in **higher values by around $1\,\sigma$ of estimated σ_8 parameter** showing better agreement with expectations from the ΛCDM model (Chapter 5).

5. **Different CMB lensing potential maps** namely, minimum variance reconstruction from the *Planck* CMB temperature and polarization measurements, Sunyaev-Zeldovich deprojected and temperature only reconstruction of the lensing potential, **yields slightly different sets of amplitude of cross-correlation**. The differences are of order of $0.5-1\,\sigma$ for both of the analysed catalogues, i.e. HELP and DESI-LIS. They indicate how big are uncertainties related with possible systematic errors of the CMB lensing maps.

The tomographic cross-correlations have been employed on several occasions to estimate cosmological parameters and they consistently report lower than expected value of σ_8 or S_8 parameter than expected in ΛCDM model (White et al. 2022; Pandey et al. 2022; Chang et al. 2022; Sun et al. 2022; Krolewski et al. 2021; Hang et al. 2021; Peacock and Bilicki 2018). The tensions on cosmological parameters have been a major motivation towards alternative models of our Universe. The results of this thesis indicate that the redshift bin mismatch of objects in tomographic analysis will drive the estimated parameters away from their true values. The shift in the parameters will depend on the photometric redshift error distributions and the redshift distribution of objects. Thus, for an unbiased estimation of cosmological parameters from future tomographic cross-correlation measurements, it is of utmost importance to carry out the analysis taking into account the redshift bin mismatch of objects along with other systematics. This thesis provides a robust, novel and computationally fast scattering matrix approach, thoroughly tested using a suite of realistic Monte Carlo simulations, for unbiased estimation of cosmological parameters. This approach and methods of correction for systematic errors presented in the thesis will be especially well-suited and very valuable for the analysis of large datasets from upcoming generation of multi-wavelength photometric galaxy surveys such as LSST, Euclid (Laureijs et al., 2011), Nancy Grace Roman Space Telescope (Spergel et al., 2013), and Dark Energy Spectroscopic Instrument (DESI, Dey et al. 2019a).

Bibliography

Abazajian, K., Addison, G., Adshead, P., Ahmed, Z., Allen, S. W., Alonso, D., Alvarez, M., Anderson, A., Arnold, K. S., Baccigalupi, C., Bailey, K., Barkats, D., Barron, D., Barry, P. S., Bartlett, J. G., Basu Thakur, R., Battaglia, N., Baxter, E., Bean, R., ... Zonca, A. (2019). CMB-S4 Science Case, Reference Design, and Project Plan. *arXiv e-prints*, Article arXiv:1907.04473, arXiv:1907.04473.

Abazajian, K. N., Adshead, P., Ahmed, Z., Allen, S. W., Alonso, D., Arnold, K. S., Baccigalupi, C., Bartlett, J. G., Battaglia, N., Benson, B. A., Bischoff, C. A., Borrill, J., Buza, V., Calabrese, E., Caldwell, R., Carlstrom, J. E., Chang, C. L., Crawford, T. M., Cyr-Racine, F.-Y., ... Kimmy Wu, W. L. (2016). CMB-S4 Science Book, First Edition. *arXiv e-prints*, Article arXiv:1610.02743, arXiv:1610.02743.

Abbott, T. M. C., Abdalla, F. B., Alarcon, A., Allam, S., Annis, J., Avila, S., Aylor, K., Banerji, M., Banik, N., Baxter, E. J., Bechtol, K., Becker, M. R., Benson, B. A., Bernstein, G. M., Bertin, E., Bianchini, F., Blazek, J., Bleem, L. E., Bridle, S. L., ... SPT Collaborations. (2019). Dark Energy Survey year 1 results: Joint analysis of galaxy clustering, galaxy lensing, and CMB lensing two-point functions. Phys. Rev. D, *100*(2), Article 023541, 023541. https://doi.org/10.1103/PhysRevD.100.023541

Abbott, T. M. C., Abdalla, F. B., Allam, S., Amara, A., Annis, J., Asorey, J., Avila, S., Ballester, O., Banerji, M., Barkhouse, W., Baruah, L., Baumer, M., Bechtol, K., Becker, M. R., Benoit-Lévy, A., Bernstein, G. M., Bertin, E., Blazek, J., Bocquet, S., ... NOAO Data Lab. (2018). The Dark Energy Survey: Data Release 1. ApJS, *239*(2), Article 18, 18. https://doi.org/10.3847/1538-4365/aae9f0

Ade, P., Aguirre, J., Ahmed, Z., Aiola, S., Ali, A., Alonso, D., Alvarez, M. A., Arnold, K., Ashton, P., Austermann, J., Awan, H., Baccigalupi, C., Baildon, T., Barron, D., Battaglia, N., Battye, R., Baxter, E., Bazarko, A., Beall, J. A., ... Simons Observatory Collaboration. (2019). The Simons Observatory: science goals and forecasts. *JCAP*, *2019*(2), Article 056, 056. https://doi.org/10.1088/1475-7516/2019/02/056

Aguilar Faúndez, M., Arnold, K., Baccigalupi, C., Barron, D., Beck, D., Bianchini, F., Boettger, D., Borrill, J., Carron, J., Cheung, K., Chinone, Y., El Bouhargani, H., Elleflot, T., Errard, J., Fabbian, G., Feng, C., Galitzki, N., Goeckner-Wald, N., Hasegawa, M., ... Polarbear Collaboration. (2019). Cross-correlation of CMB Polarization Lensing with High-z Submillimeter Herschel-ATLAS Galaxies. ApJ, *886*(1), Article 38, 38. https://doi.org/10.3847/1538-4357/ab4a78

Ahumada, R., Allende Prieto, C., Almeida, A., Anders, F., Anderson, S. F., Andrews, B. H., Anguiano, B., Arcodia, R., Armengaud, E., Aubert, M., Avila, S., Avila-Reese, V., Badenes, C., Balland, C., Barger, K., Barrera-Ballesteros, J. K., Basu, S., Bautista, J., Beaton,

R. L., ... Zou, H. (2020). The 16th Data Release of the Sloan Digital Sky Surveys: First Release from the APOGEE-2 Southern Survey and Full Release of eBOSS Spectra. ApJS, *249*(1), Article 3, 3. https://doi.org/10.3847/1538-4365/ab929e

Alam, S., Albareti, F. D., Allende Prieto, C., Anders, F., Anderson, S. F., Anderton, T., Andrews, B. H., Armengaud, E., Aubourg, É., Bailey, S., Basu, S., Bautista, J. E., Beaton, R. L., Beers, T. C., Bender, C. F., Berlind, A. A., Beutler, F., Bhardwaj, V., Bird, J. C., ... Zhu, G. (2015). The Eleventh and Twelfth Data Releases of the Sloan Digital Sky Survey: Final Data from SDSS-III. ApJS, *219*(1), Article 12, 12. https://doi.org/10.1088/0067-0049/219/1/12

Alonso, D., Bellini, E., Hale, C., Jarvis, M. J., & Schwarz, D. J. (2021). Cross-correlating radio continuum surveys and CMB lensing: constraining redshift distributions, galaxy bias, and cosmology. MNRAS, *502*(1), 876–887. https://doi.org/10.1093/mnras/stab046

Alpher, R. A., & Herman, R. (1948a). Evolution of the Universe. Nature, *162*(4124), 774–775. https://doi.org/10.1038/162774b0

Alpher, R. A., & Herman, R. C. (1948b). On the Relative Abundance of the Elements. *Physical Review*, *74*(12), 1737–1742. https://doi.org/10.1103/PhysRev.74.1737

Ando, S., Benoit-Lévy, A., & Komatsu, E. (2018). Angular power spectrum of galaxies in the 2MASS Redshift Survey. MNRAS, *473*(4), 4318–4325. https://doi.org/10.1093/mnras/stx2634

Bartelmann, M., & Schneider, P. (2001). Weak gravitational lensing. Phys. Rep., *340*(4-5), 291–472. https://doi.org/10.1016/S0370-1573(00)00082-X

Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Page, L., Spergel, D. N., Tucker, G. S., Wollack, E., Wright, E. L., Barnes, C., Greason, M. R., Hill, R. S., Komatsu, E., Nolta, M. R., Odegard, N., Peiris, H. V., ... Weiland, J. L. (2003). First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Preliminary Maps and Basic Results. ApJS, *148*(1), 1–27. https://doi.org/10.1086/377253

Benson, B. A., Ade, P. A. R., Ahmed, Z., Allen, S. W., Arnold, K., Austermann, J. E., Bender, A. N., Bleem, L. E., Carlstrom, J. E., Chang, C. L., Cho, H. M., Cliche, J. F., Crawford, T. M., Cukierman, A., de Haan, T., Dobbs, M. A., Dutcher, D., Everett, W., Gilbert, A., ... Yoon, K. W. SPT-3G: a next-generation cosmic microwave background polarization experiment on the South Pole telescope (W. S. Holland & J. Zmuidzinas, Eds.). *Millimeter, submillimeter, and far-infrared detectors and instrumentation for astronomy vii* (W. S. Holland & J. Zmuidzinas, Eds.). Ed. by Holland, W. S., & Zmuidzinas, J. *9153*. International Society for Optics and Photonics. SPIE, 2014, 552 –572. https://doi.org/10.1117/12.2057305.

Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. (2002). Large-scale structure of the Universe and cosmological perturbation theory. Phys. Rep., *367*(1-3), 1–248. https://doi.org/10.1016/S0370-1573(02)00135-7

Bianchini, F., Bielewicz, P., Lapi, A., Gonzalez-Nuevo, J., Baccigalupi, C., de Zotti, G., Danese, L., Bourne, N., Cooray, A., Dunne, L., Dye, S., Eales, S., Ivison, R., Maddox, S., Negrello, M., Scott, D., Smith, M. W. L., & Valiante, E. (2015). Cross-correlation between the CMB Lensing Potential Measured by Planck and High-z Submillimeter Galaxies Detected

by the Herschel-Atlas Survey. ApJ, *802*(1), Article 64, 64. https://doi.org/10.1088/0004-637X/802/1/64

Bianchini, F., Lapi, A., Calabrese, M., Bielewicz, P., Gonzalez-Nuevo, J., Baccigalupi, C., Danese, L., de Zotti, G., Bourne, N., Cooray, A., Dunne, L., Eales, S., & Valiante, E. (2016). Toward a Tomographic Analysis of the Cross-Correlation between Planck CMB Lensing and H-ATLAS Galaxies. ApJ, *825*(1), Article 24, 24. https://doi.org/10.3847/0004-637X/825/1/24

Bianchini, F., & Reichardt, C. L. (2018). Constraining Gravity at Large Scales with the 2MASS Photometric Redshift Catalog and Planck Lensing. ApJ, *862*(1), Article 81, 81. https://doi.org/10.3847/1538-4357/aacafd

Bilicki, M., Jarrett, T. H., Peacock, J. A., Cluver, M. E., & Steward, L. (2014). Two Micron All Sky Survey Photometric Redshift Catalog: A Comprehensive Three-dimensional Census of the Whole Sky. ApJS, *210*(1), Article 9, 9. https://doi.org/10.1088/0067-0049/210/1/9

Brammer, G. B., van Dokkum, P. G., & Coppi, P. (2008). EAZY: A Fast, Public Photometric Redshift Code. ApJ, *686*(2), 1503–1513. https://doi.org/10.1086/591786

Brown, M. L., Castro, P. G., & Taylor, A. N. (2005). Cosmic microwave background temperature and polarization pseudo-C_l estimators and covariances. MNRAS, *360*(4), 1262–1280. https://doi.org/10.1111/j.1365-2966.2005.09111.x

Budavári, T., Connolly, A. J., Szalay, A. e. S., Szapudi, I., Csabai, I., Scranton, R., Bahcall, N. A., Brinkmann, J., Eisenstein, D. J., Frieman, J. A., Fukugita, M., Gunn, J. E., Johnston, D., Kent, S., Loveday, J. N., Lupton, R. H., Tegmark, M., Thakar, A. R., Yanny, B., ... Zehavi, I. (2003). Angular Clustering with Photometric Redshifts in the Sloan Digital Sky Survey: Bimodality in the Clustering Properties of Galaxies. ApJ, *595*(1), 59–70. https://doi.org/10.1086/377168

Cao, Y., Gong, Y., Feng, C., Cooray, A., Cheng, G., & Chen, X. (2020). Cross-correlation of Far-infrared Background Anisotropies and CMB Lensing from Herschel and Planck Satellites. ApJ, *901*(1), Article 34, 34. https://doi.org/10.3847/1538-4357/abada1

Cao, Y., Gong, Y., Feng, C., Cooray, A., Cheng, G., & Chen, X. (2020). Cross-correlation of far-infrared background anisotropies and cmb lensing from herschel and planck satellites. *The Astrophysical Journal*, *901*(1), 34. https://doi.org/10.3847/1538-4357/abada1

Cawthon, R., Davis, C., Gatti, M., Vielzeuf, P., Elvin-Poole, J., Rozo, E., Frieman, J., Rykoff, E. S., Alarcon, A., Bernstein, G. M., Bonnett, C., Carnero Rosell, A., Castander, F. J., Chang, C., da Costa, L. N., De Vicente, J., DeRose, J., Drlica-Wagner, A., Gaztanaga, E., ... DES Collaboration. (2018). Dark Energy Survey Year 1 Results: calibration of redMaGiC redshift distributions in DES and SDSS from cross-correlations. MNRAS, *481*(2), 2427–2443. https://doi.org/10.1093/mnras/sty2424

Cawthon, R., Giannantonio, T., Fosalba, P., Elsner, F., Leistedt, B., Benoit-Levy, A., Kirk, D., Crocce, M., Dodelson, S., Holder, G., Omori, Y., Dark Energy Survey Collaboration, & South Pole Telescope Collaboration. CMB Lensing Tomography with the Dark Energy Survey Science Verification galaxies. In *Aps april meeting abstracts. 2015*. APS Meeting Abstracts. 2015, April, Y2.003, Y2.003.

Challinor, A. (2004). Anisotropies in the Cosmic Microwave Background. *arXiv e-prints*, Article astro-ph/0403344, astro–ph/0403344. https://doi.org/10.48550/arXiv.astro-ph/0403344

Chang, C., Omori, Y., Baxter, E. J., Doux, C., Choi, A., Pandey, S., Alarcon, A., Alves, O., Amon, A., Andrade-Oliveira, F., Bechtol, K., Becker, M. R., Bernstein, G. M., Bianchini, F., Blazek, J., Bleem, L. E., Camacho, H., Campos, A., Carnero Rosell, A., ... Williamson, R. (2022). Joint analysis of DES Year 3 data and CMB lensing from SPT and Planck II: Cross-correlation measurements and cosmological constraints. *arXiv e-prints*, Article arXiv:2203.12440, arXiv:2203.12440. https://doi.org/10.48550/arXiv.2203.12440

Coil, A. L. The Large-Scale Structure of the Universe (T. D. Oswalt & W. C. Keel, Eds.). In: *Planets, stars and stellar systems. volume 6: Extragalactic astronomy and cosmology* (T. D. Oswalt & W. C. Keel, Eds.). Ed. by Oswalt, T. D., & Keel, W. C. Vol. 6. 2013, p. 387. https://doi.org/10.1007/978-94-007-5609-0_8.

Cole, S., & Efstathiou, G. (1989). Gravitational lensing of fluctuations in the microwave background radiation. MNRAS, *239*, 195–200. https://doi.org/10.1093/mnras/239.1.195

Cruz, M., Martínez-González, E., Vielva, P., & Cayón, L. (2005). Detection of a non-Gaussian spot in WMAP. MNRAS, *356*(1), 29–40. https://doi.org/10.1111/j.1365-2966.2004.08419.x

Darwish, O., Madhavacheril, M. S., Sherwin, B. D., Aiola, S., Battaglia, N., Beall, J. A., Becker, D. T., Bond, J. R., Calabrese, E., Choi, S. K., Devlin, M. J., Dunkley, J., Dünner, R., Ferraro, S., Fox, A. E., Gallardo, P. A., Guan, Y., Halpern, M., Han, D., ... Wollack, E. J. (2021). The Atacama Cosmology Telescope: a CMB lensing mass map over 2100 square degrees of sky and its cross-correlation with BOSS-CMASS galaxies. MNRAS, *500*(2), 2250–2263. https://doi.org/10.1093/mnras/staa3438

de Jong, J. T. A., Verdoes Kleijn, G. A., Boxhoorn, D. R., Buddelmeijer, H., Capaccioli, M., Getman, F., Grado, A., Helmich, E., Huang, Z., Irisarri, N., Kuijken, K., La Barbera, F., McFarland, J. P., Napolitano, N. R., Radovich, M., Sikkema, G., Valentijn, E. A., Begeman, K. G., Brescia, M., ... Vriend, W.-J. (2015). The first and second data releases of the Kilo-Degree Survey. A&A, *582*, Article A62, A62. https://doi.org/10.1051/0004-6361/201526601

de Jong, J. T. A., Verdoes Kleijn, G. A., Kuijken, K. H., & Valentijn, E. A. (2013). The Kilo-Degree Survey. *Experimental Astronomy*, *35*(1-2), 25–44. https://doi.org/10.1007/s10686-012-9306-1

Dey, A., Rabinowitz, D., Karcher, A., Bebek, C., Baltay, C., Sprayberry, D., Valdes, F., Stupak, B., Donaldson, J., Emmet, W., Hurteau, T., Abareshi, B., Marshall, B., Lang, D., Fitzpatrick, M., Daly, P., Joyce, D., Schlegel, D., Schweiker, H., ... Levi, M. Mosaic3: a red-sensitive upgrade for the prime focus camera at the Mayall 4m telescope (C. J. Evans, L. Simard, & H. Takami, Eds.). *Ground-based and airborne instrumentation for astronomy vi* (C. J. Evans, L. Simard, & H. Takami, Eds.). Ed. by Evans, C. J., Simard, L., & Takami, H. *9908*. International Society for Optics and Photonics. SPIE, 2016, 99082C. https://doi.org/10.1117/12.2231488.

Dey, A., Schlegel, D. J., Lang, D., Blum, R., Burleigh, K., Fan, X., Findlay, J. R., Finkbeiner, D., Herrera, D., Juneau, S., Landriau, M., Levi, M., McGreer, I., Meisner, A., Myers, A. D.,

Moustakas, J., Nugent, P., Patej, A., Schlafly, E. F., ... Zhou, Z. (2019a). Overview of the DESI Legacy Imaging Surveys. AJ, *157*(5), Article 168, 168. https://doi.org/10.3847/1538-3881/ab089d

Dey, A., Schlegel, D. J., Lang, D., Blum, R., Burleigh, K., Fan, X., Findlay, J. R., Finkbeiner, D., Herrera, D., Juneau, S., Landriau, M., Levi, M., McGreer, I., Meisner, A., Myers, A. D., Moustakas, J., Nugent, P., Patej, A., Schlafly, E. F., ... Zhou, Z. (2019b). Overview of the DESI Legacy Imaging Surveys. AJ, *157*(5), Article 168, 168. https://doi.org/10.3847/1538-3881/ab089d

DiPompeo, M. A., Myers, A. D., Hickox, R. C., Geach, J. E., Holder, G., Hainline, K. N., & Hall, S. W. (2015). Weighing obscured and unobscured quasar hosts with the cosmic microwave background. MNRAS, *446*(4), 3492–3501. https://doi.org/10.1093/mnras/stu2341

Dodelson, S. (2003). *Modern Cosmology*.

Doré, O., Bock, J., Ashby, M., Capak, P., Cooray, A., de Putter, R., Eifler, T., Flagey, N., Gong, Y., Habib, S., Heitmann, K., Hirata, C., Jeong, W.-S., Katti, R., Korngut, P., Krause, E., Lee, D.-H., Masters, D., Mauskopf, P., ... Zemcov, M. (2014). Cosmology with the SPHEREX All-Sky Spectral Survey. *arXiv e-prints*, Article arXiv:1412.4872, arXiv:1412.4872. https://doi.org/10.48550/arXiv.1412.4872

Doux, C., Penna-Lima, M., Vitenti, S. D. P., Tréguer, J., Aubourg, E., & Ganga, K. (2018). Cosmological constraints from a joint analysis of cosmic microwave background and spectroscopic tracers of the large-scale structure. MNRAS, *480*(4), 5386–5411. https://doi.org/10.1093/mnras/sty2160

Duncan, K. J., Brown, M. J. I., Williams, W. L., Best, P. N., Buat, V., Burgarella, D., Jarvis, M. J., Małek, K., Oliver, S. J., Röttgering, H. J. A., & Smith, D. J. B. (2018a). Photometric redshifts for the next generation of deep radio continuum surveys - I. Template fitting. MNRAS, *473*(2), 2655–2672. https://doi.org/10.1093/mnras/stx2536

Duncan, K. J., Jarvis, M. J., Brown, M. J. I., & Röttgering, H. J. A. (2018b). Photometric redshifts for the next generation of deep radio continuum surveys - II. Gaussian processes and hybrid estimates. MNRAS, *477*(4), 5177–5190. https://doi.org/10.1093/mnras/sty940

Eales, S., Dunne, L., Clements, D., Cooray, A., De Zotti, G., Dye, S., Ivison, R., Jarvis, M., Lagache, G., Maddox, S., Negrello, M., Serjeant, S., Thompson, M. A., Van Kampen, E., Amblard, A., Andreani, P., Baes, M., Beelen, A., Bendo, G. J., ... White, G. J. (2010). The Herschel ATLAS. PASP, *122*(891), 499. https://doi.org/10.1086/653086

Efstathiou, G. (2004). Myths and truths concerning estimation of power spectra: the case for a hybrid estimator. MNRAS, *349*(2), 603–626. https://doi.org/10.1111/j.1365-2966.2004.07530.x

Fagrelius, P. A. The Dark Energy Spectroscopic Instrument (DESI) Overview. In *American astronomical society meeting abstracts #235. 235*. American Astronomical Society Meeting Abstracts. 2020, January, 446.01, 446.01.

Fang, X., Eifler, T., Schaan, E., Huang, H.-J., Krause, E., & Ferraro, S. (2022). Cosmology from clustering, cosmic shear, CMB lensing, and cross correlations: combining Rubin observatory and Simons Observatory. MNRAS, *509*(4), 5721–5736. https://doi.org/10.1093/mnras/stab3410

Fixsen, D. J. (2009). The Temperature of the Cosmic Microwave Background. ApJ, *707*(2), 916–920. https://doi.org/10.1088/0004-637X/707/2/916

Fixsen, D. J., Cheng, E. S., Gales, J. M., Mather, J. C., Shafer, R. A., & Wright, E. L. (1996). The Cosmic Microwave Background Spectrum from the Full COBE FIRAS Data Set. ApJ, *473*, 576. https://doi.org/10.1086/178173

Flaugher, B., Diehl, H. T., Honscheid, K., Abbott, T. M. C., Alvarez, O., Angstadt, R., Annis, J. T., Antonik, M., Ballester, O., Beaufore, L., Bernstein, G. M., Bernstein, R. A., Bigelow, B., Bonati, M., Boprie, D., Brooks, D., Buckley-Geer, E. J., Campa, J., Cardiel-Sas, L., ... DES Collaboration. (2015). The Dark Energy Camera. AJ, *150*(5), Article 150, 150. https://doi.org/10.1088/0004-6256/150/5/150

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. (2013). emcee: The MCMC Hammer. PASP, *125*(925), 306. https://doi.org/10.1086/670067

Fry, J. N. (1996). The Evolution of Bias. ApJ, *461*, L65. https://doi.org/10.1086/310006

Gamow, G. (1948a). The Evolution of the Universe. Nature, *162*(4122), 680–682. https://doi.org/10.1038/162680a0

Gamow, G. (1948b). The Origin of Elements and the Separation of Galaxies. *Physical Review*, *74*(4), 505–506. https://doi.org/10.1103/PhysRev.74.505.2

Giannantonio, T., Fosalba, P., Cawthon, R., Omori, Y., Crocce, M., Elsner, F., Leistedt, B., Dodelson, S., Benoit-Lévy, A., Gaztañaga, E., Holder, G., Peiris, H. V., Percival, W. J., Kirk, D., Bauer, A. H., Benson, B. A., Bernstein, G. M., Carretero, J., Crawford, T. M., ... Zuntz, J. (2016). CMB lensing tomography with the DES Science Verification galaxies. MNRAS, *456*(3), 3213–3244. https://doi.org/10.1093/mnras/stv2678

Giusarma, E., Vagnozzi, S., Ho, S., Ferraro, S., Freese, K., Kamen-Rubio, R., & Luk, K.-B. (2018). Scale-dependent galaxy bias, CMB lensing-galaxy cross-correlation, and neutrino masses. Phys. Rev. D, *98*(12), Article 123526, 123526. https://doi.org/10.1103/PhysRevD.98.123526

Górski, K. M., Hivon, E., Banday, A. J., Wandelt, B. D., Hansen, F. K., Reinecke, M., & Bartelmann, M. (2005). HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere. ApJ, *622*(2), 759–771. https://doi.org/10.1086/427976

Goto, T., Szapudi, I., & Granett, B. R. (2012). Cross-correlation of WISE galaxies with the cosmic microwave background. MNRAS, *422*(1), L77–L81. https://doi.org/10.1111/j.1745-3933.2012.01240.x

Gunn, J. E., Siegmund, W. A., Mannery, E. J., Owen, R. E., Hull, C. L., Leger, R. F., Carey, L. N., Knapp, G. R., York, D. G., Boroski, W. N., Kent, S. M., Lupton, R. H., Rockosi, C. M., Evans, M. L., Waddell, P., Anderson, J. E., Annis, J., Barentine, J. C., Bartoszek, L. M., ... Wang, S.-i. (2006). The 2.5 m Telescope of the Sloan Digital Sky Survey. AJ, *131*(4), 2332–2359. https://doi.org/10.1086/500975

Gupta, N., & Reichardt, C. L. (2021). Mass Estimation of Galaxy Clusters with Deep Learning II. Cosmic Microwave Background Cluster Lensing. ApJ, *923*(1), Article 96, 96. https://doi.org/10.3847/1538-4357/ac32d0

Guth, A. H. (1981). The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems (L.-Z. Fang & R. Ruffini, Eds.). *Phys. Rev. D*, *23*, 347–356. https://doi.org/10.1103/PhysRevD.23.347

Han, J., Ferraro, S., Giusarma, E., & Ho, S. (2019). Probing gravitational lensing of the CMB with SDSS-IV quasars. MNRAS, *485*(2), 1720–1726. https://doi.org/10.1093/mnras/stz528

Hang, Q., Alam, S., Peacock, J. A., & Cai, Y.-C. (2021). Galaxy clustering in the DESI Legacy Survey and its imprint on the CMB. MNRAS, *501*(1), 1481–1498. https://doi.org/10.1093/mnras/staa3738

Henderson, S. W., Allison, R., Austermann, J., Baildon, T., Battaglia, N., Beall, J. A., Becker, D., De Bernardis, F., Bond, J. R., Calabrese, E., Choi, S. K., Coughlin, K. P., Crowley, K. T., Datta, R., Devlin, M. J., Duff, S. M., Dunkley, J., Dünner, R., van Engelen, A., ... Wollack, E. J. (2016). Advanced ACTPol Cryogenic Detector Arrays and Readout. *Journal of Low Temperature Physics*, *184*(3-4), 772–779. https://doi.org/10.1007/s10909-016-1575-z

Heymans, C., Tröster, T., Asgari, M., Blake, C., Hildebrandt, H., Joachimi, B., Kuijken, K., Lin, C.-A., Sánchez, A. G., van den Busch, J. L., Wright, A. H., Amon, A., Bilicki, M., de Jong, J., Crocce, M., Dvornik, A., Erben, T., Fortuna, M. C., Getman, F., ... Wolf, C. (2021). KiDS-1000 Cosmology: Multi-probe weak gravitational lensing and spectroscopic galaxy clustering constraints. A&A, *646*, Article A140, A140. https://doi.org/10.1051/0004-6361/202039063

Hikage, C., Oguri, M., Hamana, T., More, S., Mandelbaum, R., Takada, M., Köhlinger, F., Miyatake, H., Nishizawa, A. J., Aihara, H., Armstrong, R., Bosch, J., Coupon, J., Ducout, A., Ho, P., Hsieh, B.-C., Komiyama, Y., Lanusse, F., Leauthaud, A., ... Yamada, Y. (2019). Cosmology from cosmic shear power spectra with Subaru Hyper Suprime-Cam first-year data. PASJ, *71*(2), Article 43, 43. https://doi.org/10.1093/pasj/psz010

Hivon, E., Górski, K. M., Netterfield, C. B., Crill, B. P., Prunet, S., & Hansen, F. (2002). MASTER of the Cosmic Microwave Background Anisotropy Power Spectrum: A Fast Method for Statistical Analysis of Large and Complex Cosmic Microwave Background Data Sets. ApJ, *567*(1), 2–17. https://doi.org/10.1086/338126

Hu, W. (2001). Mapping the Dark Matter through the Cosmic Microwave Background Damping Tail. ApJ, *557*(2), L79–L83. https://doi.org/10.1086/323253

Hu, W., & Dodelson, S. (2002). Cosmic Microwave Background Anisotropies. ARA&A, *40*, 171–216. https://doi.org/10.1146/annurev.astro.40.060401.093926

Hu, W., & Okamoto, T. (2002). Mass Reconstruction with Cosmic Microwave Background Polarization. ApJ, *574*(2), 566–574. https://doi.org/10.1086/341110

Hu, W., & White, M. (1997). A CMB polarization primer. *New Astronomy*, *2*(4), 323–344. https://doi.org/10.1016/s1384-1076(97)00022-5

Hubble, E. (1929). A Relation between Distance and Radial Velocity among Extra-Galactic Nebulae. *Proceedings of the National Academy of Science*, *15*(3), 168–173. https://doi.org/10.1073/pnas.15.3.168

Huterer, D., Cunha, C. E., & Fang, W. (2013). Calibration errors unleashed: effects on cosmological parameters and requirements for large-scale structure surveys. MNRAS, *432*(4), 2945–2961. https://doi.org/10.1093/mnras/stt653

Ilbert, O., Capak, P., Salvato, M., Aussel, H., McCracken, H. J., Sanders, D. B., Scoville, N., Kartaltepe, J., Arnouts, S., Le Floc'h, E., Mobasher, B., Taniguchi, Y., Lamareille, F., Leauthaud, A., Sasaki, S., Thompson, D., Zamojski, M., Zamorani, G., Bardelli, S., ... Zucca, E. (2009). Cosmos Photometric Redshifts with 30-Bands for 2-deg^2. ApJ, *690*(2), 1236–1249. https://doi.org/10.1088/0004-637X/690/2/1236

Ivezić, Ž., Kahn, S. M., Tyson, J. A., Abel, B., Acosta, E., Allsman, R., Alonso, D., AlSayyad, Y., Anderson, S. F., Andrew, J., Angel, J. R. P., Angeli, G. Z., Ansari, R., Antilogus, P., Araujo, C., Armstrong, R., Arndt, K. T., Astier, P., Aubourg, É., ... Zhan, H. (2019). LSST: From science drivers to reference design and anticipated data products. *ApJ*, *873*(2), 111. https://doi.org/10.3847/1538-4357/ab042c

Jouvel, S., Delubac, T., Comparat, J., Camacho, H., Carnero, A., Abdalla, F. B., Kneib, J. P., Merson, A., Lima, M., Sobreira, F., da Costa, L., Prada, F., Zhu, G. B., Benoit-Levy, A., De La Macora, A., Kuropatkin, N., Lin, H., Abbott, T. M. C., Allam, S., ... Brownstein, J. (2017). Photometric redshifts and clustering of emission line galaxies selected jointly by DES and eBOSS. MNRAS, *469*(3), 2771–2790. https://doi.org/10.1093/mnras/stx163

Kamionkowski, M., Kosowsky, A., & Stebbins, A. (1997). Statistics of cosmic microwave background polarization. Phys. Rev. D, *55*(12), 7368–7388. https://doi.org/10.1103/PhysRevD.55.7368

Krolewski, A., Ferraro, S., Schlafly, E. F., & White, M. (2020). unWISE tomography of Planck CMB lensing. J. Cosmology Astropart. Phys., *2020*(5), Article 047, 047. https://doi.org/10.1088/1475-7516/2020/05/047

Krolewski, A., Ferraro, S., & White, M. (2021). Cosmological constraints from unWISE and Planck CMB lensing tomography. *arXiv e-prints*, Article arXiv:2105.03421, arXiv:2105.03421.

Kuntz, A. (2015). Cross-correlation of CFHTLenS galaxy catalogue and Planck CMB lensing using the halo model prescription. A&A, *584*, Article A53, A53. https://doi.org/10.1051/0004-6361/201526940

Land, K., & Magueijo, J. (2005). Examination of Evidence for a Preferred Axis in the Cosmic Radiation Anisotropy. Phys. Rev. Lett., *95*(7), Article 071301, 071301. https://doi.org/10.1103/PhysRevLett.95.071301

Laureijs, R., Amiaux, J., Arduini, S., Augueres, J. L., Brinchmann, J., Cole, R., Cropper, M., Dabin, C., Duvet, L., Ealet, A., Garilli, B., Gondoin, P., Guzzo, L., Hoar, J., Hoekstra, H., Holmes, R., Kitching, T., Maciaszek, T., Mellier, Y., ... Zucca, E. (2011). Euclid Definition Study Report. *arXiv e-prints*, Article arXiv:1110.3193, arXiv:1110.3193. https://doi.org/10.48550/arXiv.1110.3193

Lewis, A., & Challinor, A. (2006). Weak gravitational lensing of the CMB. Phys. Rep., *429*(1), 1–65. https://doi.org/10.1016/j.physrep.2006.03.002

Lewis, A., Challinor, A., & Lasenby, A. (2000). Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models. ApJ, *538*(2), 473–476. https://doi.org/10.1086/309179

Limber, D. N. (1953). The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field. ApJ, *117*, 134. https://doi.org/10.1086/145672

Linde, A. D. (1982). A New Inflationary Universe Scenario: A Possible Solution of the Horizon, Flatness, Homogeneity, Isotropy and Primordial Monopole Problems (L.-Z. Fang & R. Ruffini, Eds.). *Phys. Lett. B*, *108*, 389–393. https://doi.org/10.1016/0370-2693(82)91219-9

Linder, E. V. (1990). Analysis of gravitationally lensed microwave background anisotropies. MNRAS, *243*, 353–361.

Linder, E. V. (2005). Cosmic growth history and expansion history. Phys. Rev. D, *72*(4), Article 043529, 043529. https://doi.org/10.1103/PhysRevD.72.043529

Liske, J., Baldry, I. K., Driver, S. P., Tuffs, R. J., Alpaslan, M., Andrae, E., Brough, S., Cluver, M. E., Grootes, M. W., Gunawardhana, M. L. P., Kelvin, L. S., Loveday, J., Robotham, A. S. G., Taylor, E. N., Bamford, S. P., Bland-Hawthorn, J., Brown, M. J. I., Drinkwater, M. J., Hopkins, A. M., ... Wright, A. H. (2015). Galaxy And Mass Assembly (GAMA): end of survey report and data release 2. MNRAS, *452*(2), 2087–2126. https://doi.org/10.1093/mnras/stv1436

LSST Science Collaboration, Abell, P. A., Allison, J., Anderson, S. F., Andrew, J. R., Angel, J. R. P., Armus, L., Arnett, D., Asztalos, S. J., Axelrod, T. S., Bailey, S., Ballantyne, D. R., Bankert, J. R., Barkhouse, W. A., Barr, J. D., Barrientos, L. F., Barth, A. J., Bartlett, J. G., Becker, A. C., ... Zhan, H. (2009). LSST Science Book, Version 2.0. *arXiv e-prints*, Article arXiv:0912.0201, arXiv:0912.0201.

Madhavacheril, M. S., Qu, F. J., Sherwin, B. D., MacCrann, N., Li, Y., Abril-Cabezas, I., Ade, P. A. R., Aiola, S., Alford, T., Amiri, M., Amodeo, S., An, R., Atkins, Z., Austermann, J. E., Battaglia, N., Battistelli, E. S., Beall, J. A., Bean, R., Beringue, B., ... Zhang, K. (2023). The Atacama Cosmology Telescope: DR6 Gravitational Lensing Map and Cosmological Parameters. *arXiv e-prints*, Article arXiv:2304.05203, arXiv:2304.05203. https://doi.org/10.48550/arXiv.2304.05203

Marques, G. A., & Bernui, A. (2020). Tomographic analyses of the CMB lensing and galaxy clustering to probe the linear structure growth. J. Cosmology Astropart. Phys., *2020*(5), Article 052, 052. https://doi.org/10.1088/1475-7516/2020/05/052

Meister, A. (2009). Density deconvolution. In *Deconvolution problems in nonparametric statistics* (pp. 5–105). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-87557-4_2

Metcalf, R. B., & Silk, J. (1997). Gravitational Magnification of the Cosmic Microwave Background. ApJ, *489*(1), 1–6. https://doi.org/10.1086/304756

Millea, M., Daley, C. M., Chou, T. L., Anderes, E., Ade, P. A. R., Anderson, A. J., Austermann, J. E., Avva, J. S., Beall, J. A., Bender, A. N., Benson, B. A., Bianchini, F., Bleem, L. E., Carlstrom, J. E., Chang, C. L., Chaubal, P., Chiang, H. C., Citron, R., Moran, C. C., ... Yefremenko, V. (2021). Optimal Cosmic Microwave Background Lensing Reconstruction and Parameter Estimation with SPTpol Data. ApJ, *922*(2), Article 259, 259. https://doi.org/10.3847/1538-4357/ac02bb

Miyatake, H., Harikane, Y., Ouchi, M., Ono, Y., Yamamoto, N., Nishizawa, A. J., Bahcall, N., Miyazaki, S., & Malagón, A. A. P. (2022). First Identification of a CMB Lensing Signal Produced by 1.5 Million Galaxies at z ∼4 : Constraints on Matter Density Fluctuations

at High Redshift. Phys. Rev. Lett., *129*(6), Article 061301, 061301. https://doi.org/10.1103/PhysRevLett.129.061301

Mo, H., van den Bosch, F. C., & White, S. (2010). *Galaxy Formation and Evolution*.

Moscardini, L., Coles, P., Lucchin, F., & Matarrese, S. (1998). Modelling galaxy clustering at high redshift. MNRAS, *299*(1), 95–110. https://doi.org/10.1046/j.1365-8711.1998.01728.x

Muir, J., & Huterer, D. (2016). Reconstructing the integrated Sachs-Wolfe map with galaxy surveys. Phys. Rev. D, *94*(4), Article 043503, 043503. https://doi.org/10.1103/PhysRevD.94.043503

Namikawa, T., Chinone, Y., Miyatake, H., Oguri, M., Takahashi, R., Kusaka, A., Katayama, N., Adachi, S., Aguilar, M., Aihara, H., Ali, A., Armstrong, R., Arnold, K., Baccigalupi, C., Barron, D., Beck, D., Beckman, S., Bianchini, F., Boettger, D., ... SUBARU HSC SSP Collaboration. (2019). Evidence for the Cross-correlation between Cosmic Microwave Background Polarization Lensing from Polarbear and Cosmic Shear from Subaru Hyper Suprime-Cam. ApJ, *882*(1), Article 62, 62. https://doi.org/10.3847/1538-4357/ab3424

Natarajan, P., Chadayammuri, U., Jauzac, M., Richard, J., Kneib, J.-P., Ebeling, H., Jiang, F., van den Bosch, F., Limousin, M., Jullo, E., Atek, H., Pillepich, A., Popa, C., Marinacci, F., Hernquist, L., Meneghetti, M., & Vogelsberger, M. (2017). Mapping substructure in the HST Frontier Fields cluster lenses and in cosmological simulations. MNRAS, *468*(2), 1962–1980. https://doi.org/10.1093/mnras/stw3385

Newman, J. A., Cooper, M. C., Davis, M., Faber, S. M., Coil, A. L., Guhathakurta, P., Koo, D. C., Phillips, A. C., Conroy, C., Dutton, A. A., Finkbeiner, D. P., Gerke, B. F., Rosario, D. J., Weiner, B. J., Willmer, C. N. A., Yan, R., Harker, J. J., Kassin, S. A., Konidaris, N. P., ... Schiavon, R. P. (2013). The DEEP2 Galaxy Redshift Survey: Design, Observations, Data Reduction, and Redshifts. ApJS, *208*(1), Article 5, 5. https://doi.org/10.1088/0067-0049/208/1/5

Oliver, S. J., Bock, J., Altieri, B., Amblard, A., Arumugam, V., Aussel, H., Babbedge, T., Beelen, A., Béthermin, M., Blain, A., Boselli, A., Bridge, C., Brisbin, D., Buat, V., Burgarella, D., Castro-Rodríguez, N., Cava, A., Chanial, P., Cirasuolo, M., ... Zemcov, M. (2012). The Herschel Multi-tiered Extragalactic Survey: HerMES. MNRAS, *424*(3), 1614–1635. https://doi.org/10.1111/j.1365-2966.2012.20912.x

Omori, Y., Baxter, E. J., Chang, C., Kirk, D., Alarcon, A., Bernstein, G. M., Bleem, L. E., Cawthon, R., Choi, A., Chown, R., Crawford, T. M., Davis, C., De Vicente, J., DeRose, J., Dodelson, S., Eifler, T. F., Fosalba, P., Friedrich, O., Gatti, M., ... SPT Collaboration. (2019a). Dark Energy Survey Year 1 Results: Cross-correlation between Dark Energy Survey Y1 galaxy weak lensing and South Pole Telescope+P l a n c k CMB weak lensing. Phys. Rev. D, *100*(4), Article 043517, 043517. https://doi.org/10.1103/PhysRevD.100.043517

Omori, Y., Chown, R., Simard, G., Story, K. T., Aylor, K., Baxter, E. J., Benson, B. A., Bleem, L. E., Carlstrom, J. E., Chang, C. L., Cho, H. M., Crawford, T. M., Crites, A. T., de Haan, T., Dobbs, M. A., Everett, W. B., George, E. M., Halverson, N. W., Harrington, N. L., ... Zahn, O. (2017). A 2500 deg^2 CMB Lensing Map from Combined South Pole Telescope and Planck Data. ApJ, *849*(2), Article 124, 124. https://doi.org/10.3847/1538-4357/aa8d1d

Omori, Y., Giannantonio, T., Porredon, A., Baxter, E. J., Chang, C., Crocce, M., Fosalba, P., Alarcon, A., Banik, N., Blazek, J., Bleem, L. E., Bridle, S. L., Cawthon, R., Choi, A., Chown, R., Crawford, T., Dodelson, S., Drlica-Wagner, A., Eifler, T. F., ... SPT Collaboration. (2019b). Dark Energy Survey Year 1 Results: Tomographic cross-correlations between Dark Energy Survey galaxies and CMB lensing from South Pole Telescope +Planck. Phys. Rev. D, *100*(4), Article 043501, 043501. https://doi.org/10.1103/PhysRevD.100.043501

Padmanabhan, N., Budavári, T., Schlegel, D. J., Bridges, T., Brinkmann, J., Cannon, R., Connolly, A. J., Croom, S. M., Csabai, I., Drinkwater, M., Eisenstein, D. J., Hewett, P. C., Loveday, J., Nichol, R. C., Pimbblet, K. A., De Propris, R., Schneider, D. P., Scranton, R., Seljak, U., ... Wake, D. (2005). Calibrating photometric redshifts of luminous red galaxies. MNRAS, *359*(1), 237–250. https://doi.org/10.1111/j.1365-2966.2005.08915.x

Padmanabhan, T. Advanced Topics in Cosmology: A Pedagogical Introduction (S. Daflon, J. Alcaniz, E. Telles, & R. de la Reza, Eds.). *Graduate school in astronomy: X* (S. Daflon, J. Alcaniz, E. Telles, & R. de la Reza, Eds.). Ed. by Daflon, S., Alcaniz, J., Telles, E., & de la Reza, R. *843*. American Institute of Physics Conference Series. 2006, June, 111–166. https://doi.org/10.1063/1.2219327. arXiv: astro-ph/0602117 [astro-ph].

Pandey, S., Krause, E., DeRose, J., MacCrann, N., Jain, B., Crocce, M., Blazek, J., Choi, A., Huang, H., To, C., Fang, X., Elvin-Poole, J., Prat, J., Porredon, A., Secco, L. F., Rodriguez-Monroy, M., Weaverdyck, N., Park, Y., Raveri, M., ... DES Collaboration. (2022). Dark Energy Survey year 3 results: Constraints on cosmological parameters and galaxy-bias models from galaxy clustering and galaxy-galaxy lensing using the redMaGiC sample. Phys. Rev. D, *106*(4), Article 043520, 043520. https://doi.org/10.1103/PhysRevD.106.043520

Peacock, J. A., & Bilicki, M. (2018). Wide-area tomography of CMB lensing and the growth of cosmological density fluctuations. MNRAS, *481*(1), 1133–1148. https://doi.org/10.1093/mnras/sty2314

Peebles, P. J. E. (1980). *The large-scale structure of the universe*.

Penzias, A. A., & Wilson, R. W. (1965). A Measurement of Excess Antenna Temperature at 4080 Mc/s. ApJ, *142*, 419–421. https://doi.org/10.1086/148307

Perlmutter, S., Aldering, G., Goldhaber, G., Knop, R. A., Nugent, P., Castro, P. G., Deustua, S., Fabbro, S., Goobar, A., Groom, D. E., Hook, I. M., Kim, A. G., Kim, M. Y., Lee, J. C., Nunes, N. J., Pain, R., Pennypacker, C. R., Quimby, R., Lidman, C., ... Project, T. S. C. (1999). Measurements of Ω and Λ from 42 High-Redshift Supernovae. ApJ, *517*(2), 565–586. https://doi.org/10.1086/307221

Planck Collaboration, Adam, R., Ade, P. A. R., Aghanim, N., Akrami, Y., Alves, M. I. R., Argüeso, F., Arnaud, M., Arroja, F., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R. B., Bartlett, J. G., Bartolo, N., Basak, S., Battaglia, P., ... Zonca, A. (2016a). Planck 2015 results. I. Overview of products and scientific results. A&A, *594*, Article A1, A1. https://doi.org/10.1051/0004-6361/201527101

Planck Collaboration, Ade, P. A. R., Aghanim, N., Alves, M. I. R., Armitage-Caplan, C., Arnaud, M., Ashdown, M., Atrio-Barandela, F., Aumont, J., Aussel, H., Baccigalupi, C., Banday, A. J., Barreiro, R. B., Barrena, R., Bartelmann, M., Bartlett, J. G., Bartolo, N., Basak, S., Battaner, E., ... Zonca, A. (2014a). Planck 2013 results. I. Overview of products

and scientific results. A&A, *571*, Article A1, A1. https://doi.org/10.1051/0004-6361/201321529

Planck Collaboration, Ade, P. A. R., Aghanim, N., Armitage-Caplan, C., Arnaud, M., Ashdown, M., Atrio-Barandela, F., Aumont, J., Baccigalupi, C., Banday, A. J., Barreiro, R. B., Bartlett, J. G., Basak, S., Battaner, E., Benabed, K., Benoît, A., Benoit-Lévy, A., Bernard, J. P., Bersanelli, M., ... Zonca, A. (2014b). Planck 2013 results. XVII. Gravitational lensing by large-scale structure. A&A, *571*, Article A17, A17. https://doi.org/10.1051/0004-6361/201321543

Planck Collaboration, Ade, P. A. R., Aghanim, N., Arnaud, M., Ashdown, M., Aumont, J., Baccigalupi, C., Banday, A. J., Barreiro, R. B., Bartlett, J. G., Bartolo, N., Battaner, E., Battye, R., Benabed, K., Benoît, A., Benoit-Lévy, A., Bernard, J. P., Bersanelli, M., Bielewicz, P., ... Zonca, A. (2016b). Planck 2015 results. XXIV. Cosmology from Sunyaev-Zeldovich cluster counts. A&A, *594*, Article A24, A24. https://doi.org/10.1051/0004-6361/201525833

Planck Collaboration, Aghanim, N., Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R. B., Bartolo, N., Basak, S., Battye, R., Benabed, K., Bernard, J. P., Bersanelli, M., Bielewicz, P., Bock, J. J., Bond, J. R., Borrill, J., ... Zonca, A. (2020a). Planck 2018 results. VI. Cosmological parameters. A&A, *641*, Article A6, A6. https://doi.org/10.1051/0004-6361/201833910

Planck Collaboration, Aghanim, N., Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R. B., Bartolo, N., Basak, S., Benabed, K., Bernard, J. P., Bersanelli, M., Bielewicz, P., Bock, J. J., Bond, J. R., Borrill, J., Bouchet, F. R., ... Zonca, A. (2020b). Planck 2018 results. VIII. Gravitational lensing. A&A, *641*, Article A8, A8. https://doi.org/10.1051/0004-6361/201833886

Planck Collaboration, Aghanim, N., Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R. B., Bartolo, N., Basak, S., Benabed, K., Bernard, J. P., Bersanelli, M., Bielewicz, P., Bond, J. R., Borrill, J., Bouchet, F. R., Boulanger, F., ... Zonca, A. (2020c). Planck 2018 results. III. High Frequency Instrument data processing and frequency maps. A&A, *641*, Article A3, A3. https://doi.org/10.1051/0004-6361/201832909

Polnarev, A. G. (1985). Polarization and Anisotropy Induced in the Microwave Background by Cosmological Gravitational Waves. Soviet Ast., *29*, 607–613.

Pullen, A. R., Alam, S., He, S., & Ho, S. (2016). Constraining gravity at the largest scales through CMB lensing and galaxy velocities. *MNRAS*, *460*(4), 4098–4108. https://doi.org/10.1093/mnras/stw1249

Pullen, A. R., Alam, S., & Ho, S. (2015). Probing gravity at large scales through CMB lensing. MNRAS, *449*(4), 4326–4335. https://doi.org/10.1093/mnras/stv554

Raghunathan, S., Patil, S., Baxter, E., Benson, B. A., Bleem, L. E., Chou, T. L., Crawford, T. M., Holder, G. P., McClintock, T., Reichardt, C. L., Rozo, E., Varga, T. N., Abbott, T. M. C., Ade, P. A. R., Allam, S., Anderson, A. J., Annis, J., Austermann, J. E., Avila, S., ... Zhang, Y. (2019). Mass Calibration of Optically Selected DES Clusters Using a Measurement of CMB-cluster Lensing with SPTpol Data. ApJ, *872*(2), Article 170, 170. https://doi.org/10.3847/1538-4357/ab01ca

Rees, M. J. (1968). Polarization and Spectrum of the Primeval Radiation in an Anisotropic Universe. ApJ, *153*, L1. https://doi.org/10.1086/180208

Riess, A. G., Filippenko, A. V., Challis, P., Clocchiatti, A., Diercks, A., Garnavich, P. M., Gilliland, R. L., Hogan, C. J., Jha, S., Kirshner, R. P., Leibundgut, B., Phillips, M. M., Reiss, D., Schmidt, B. P., Schommer, R. A., Smith, R. C., Spyromilio, J., Stubbs, C., Suntzeff, N. B., & Tonry, J. (1998). Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant. AJ, *116*(3), 1009–1038. https://doi.org/10.1086/300499

Robertson, N. C., Alonso, D., Harnois-Déraps, J., Darwish, O., Kannawadi, A., Amon, A., Asgari, M., Bilicki, M., Calabrese, E., Choi, S. K., Devlin, M. J., Dunkley, J., Dvornik, A., Erben, T., Ferraro, S., Fortuna, M. C., Giblin, B., Han, D., Heymans, C., ... Xu, Z. (2021). Strong detection of the CMB lensing and galaxy weak lensing cross-correlation from ACT-DR4, Planck Legacy, and KiDS-1000. A&A, *649*, Article A146, A146. https://doi.org/10.1051/0004-6361/202039975

Rubin, V. C., Ford, J., W. K., & Thonnard, N. (1978). Extended rotation curves of high-luminosity spiral galaxies. IV. Systematic dynamical properties, Sa -> Sc. ApJ, *225*, L107–L111. https://doi.org/10.1086/182804

Saraf, C. S., Bielewicz, P., & Chodorowski, M. (2022). Cross-correlation between Planck CMB lensing potential and galaxy catalogues from HELP. MNRAS, *515*(2), 1993–2007. https://doi.org/10.1093/mnras/stac1876

Schlafly, E. F., Meisner, A. M., & Green, G. M. (2019). The unWISE Catalog: Two Billion Infrared Sources from Five Years of WISE Imaging. ApJS, *240*(2), Article 30, 30. https://doi.org/10.3847/1538-4365/aafbea

Schmittfull, M., & Seljak, U. (2018). Parameter constraints from cross-correlation of CMB lensing with galaxy clustering. Phys. Rev. D, *97*(12), Article 123540, 123540. https://doi.org/10.1103/PhysRevD.97.123540

Scodeggio, M., Guzzo, L., Garilli, B., Granett, B. R., Bolzonella, M., de la Torre, S., Abbas, U., Adami, C., Arnouts, S., Bottini, D., Cappi, A., Coupon, J., Cucciati, O., Davidzon, I., Franzetti, P., Fritz, A., Iovino, A., Krywult, J., Le Brun, V., ... Percival, W. J. (2018). The VIMOS Public Extragalactic Redshift Survey (VIPERS). Full spectroscopic data and auxiliary information release (PDR-2). A&A, *609*, Article A84, A84. https://doi.org/10.1051/0004-6361/201630114

Scott, D., & Smoot, G. (2006). Cosmic Microwave Background Mini-Review. *arXiv e-prints*, Article astro-ph/0601307, astro–ph/0601307. https://doi.org/10.48550/arXiv.astro-ph/0601307

Seljak, U. (1996). Gravitational Lensing Effect on Cosmic Microwave Background Anisotropies: A Power Spectrum Approach. ApJ, *463*, 1. https://doi.org/10.1086/177218

Sheth, R. K., & Rossi, G. (2010). Convolution- and deconvolution-based estimates of galaxy scaling relations from photometric redshift surveys. MNRAS, *403*(4), 2137–2142. https://doi.org/10.1111/j.1365-2966.2010.16258.x

Shirley, R., Duncan, K., Campos Varillas, M. C., Hurley, P. D., Małek, K., Roehlly, Y., Smith, M. W. L., Aussel, H., Bakx, T., Buat, V., Burgarella, D., Christopher, N., Duivenvoorden, S., Eales, S., Efstathiou, A., González Solares, E. A., Griffin, M., Jarvis, M., Faro, B. L.,

... Oliver, S. J. (2021). HELP: the Herschel Extragalactic Legacy Project. MNRAS, *507*(1), 129–155. https://doi.org/10.1093/mnras/stab1526

Shirley, R., Roehlly, Y., Hurley, P. D., Buat, V., Campos Varillas, M. d. C., Duivenvoorden, S., Duncan, K. J., Efstathiou, A., Farrah, D., González Solares, E., Malek, K., Marchetti, L., McCheyne, I., Papadopoulos, A., Pons, E., Scipioni, R., Vaccari, M., & Oliver, S. (2019). HELP: a catalogue of 170 million objects, selected at 0.36-4.5 μm, from 1270 deg^2 of prime extragalactic fields. MNRAS, *490*(1), 634–656. https://doi.org/10.1093/mnras/stz2509

Singh, S., Mandelbaum, R., & Brownstein, J. R. (2017). Cross-correlating Planck CMB lensing with SDSS: lensing-lensing and galaxy-lensing cross-correlations. MNRAS, *464*(2), 2120–2138. https://doi.org/10.1093/mnras/stw2482

Singh, S., Mandelbaum, R., Seljak, U., Rodríguez-Torres, S., & Slosar, A. (2020). Cosmological constraints from galaxy-lensing cross-correlations using BOSS galaxies with SDSS and CMB lensing. MNRAS, *491*(1), 51–68. https://doi.org/10.1093/mnras/stz2922

Smoot, G. F., Bennett, C. L., Kogut, A., Wright, E. L., Aymon, J., Boggess, N. W., Cheng, E. S., de Amici, G., Gulkis, S., Hauser, M. G., Hinshaw, G., Jackson, P. D., Janssen, M., Kaita, E., Kelsall, T., Keegstra, P., Lineweaver, C., Loewenstein, K., Lubin, P., ... Wilkinson, D. T. (1992a). Structure in the COBE Differential Microwave Radiometer First-Year Maps. ApJ, *396*, L1. https://doi.org/10.1086/186504

Smoot, G. F., Bennett, C. L., Kogut, A., Wright, E. L., Aymon, J., Boggess, N. W., Cheng, E. S., de Amici, G., Gulkis, S., Hauser, M. G., Hinshaw, G., Jackson, P. D., Janssen, M., Kaita, E., Kelsall, T., Keegstra, P., Lineweaver, C., Loewenstein, K., Lubin, P., ... Wilkinson, D. T. (1992b). Structure in the COBE Differential Microwave Radiometer First-Year Maps. ApJ, *396*, L1. https://doi.org/10.1086/186504

Solarz, A., Pollo, A., Takeuchi, T. T., Małek, K., Matsuhara, H., White, G. J., Pępiak, A., Goto, T., Wada, T., Oyabu, S., Takagi, T., Ohyama, Y., Pearson, C. P., Hanami, H., Ishigaki, T., & Malkan, M. (2015). Clustering of the AKARI NEP deep field 24 μm selected galaxies. A&A, *582*, Article A58, A58. https://doi.org/10.1051/0004-6361/201423370

Spergel, D., Gehrels, N., Breckinridge, J., Donahue, M., Dressler, A., Gaudi, B. S., Greene, T., Guyon, O., Hirata, C., Kalirai, J., Kasdin, N. J., Moos, W., Perlmutter, S., Postman, M., Rauscher, B., Rhodes, J., Wang, Y., Weinberg, D., Centrella, J., ... Shaklan, S. (2013). Wide-Field InfraRed Survey Telescope-Astrophysics Focused Telescope Assets WFIRST-AFTA Final Report. *arXiv e-prints*, Article arXiv:1305.5422, arXiv:1305.5422. https://doi.org/10.48550/arXiv.1305.5422

Spergel, D. N., Verde, L., Peiris, H. V., Komatsu, E., Nolta, M. R., Bennett, C. L., Halpern, M., Hinshaw, G., Jarosik, N., Kogut, A., Limon, M., Meyer, S. S., Page, L., Tucker, G. S., Weiland, J. L., Wollack, E., & Wright, E. L. (2003). First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters. ApJS, *148*(1), 175–194. https://doi.org/10.1086/377226

Starobinsky, A. A. (1979). Spectrum of relict gravitational radiation and the early state of the universe (I. M. Khalatnikov & V. P. Mineev, Eds.). *JETP Lett.*, *30*, 682–685.

Strauss, M. A., Weinberg, D. H., Lupton, R. H., Narayanan, V. K., Annis, J., Bernardi, M., Blanton, M., Burles, S., Connolly, A. J., Dalcanton, J., Doi, M., Eisenstein, D., Frieman,

J. A., Fukugita, M., Gunn, J. E., Ivezić, Ž., Kent, S., Kim, R. S. J., Knapp, G. R., ... Zehavi, I. (2002). Spectroscopic Target Selection in the Sloan Digital Sky Survey: The Main Galaxy Sample. AJ, *124*(3), 1810–1824. https://doi.org/10.1086/342343

Sun, Z., Yao, J., Dong, F., Yang, X., Zhang, L., & Zhang, P. (2022). Cross-correlation of Planck cosmic microwave background lensing with DESI galaxy groups. MNRAS, *511*(3), 3548–3560. https://doi.org/10.1093/mnras/stac138

Szapudi, I., Kovács, A., Granett, B. R., Frei, Z., Silk, J., Burgett, W., Cole, S., Draper, P. W., Farrow, D. J., Kaiser, N., Magnier, E. A., Metcalfe, N., Morgan, J. S., Price, P., Tonry, J., & Wainscoat, R. (2015). Detection of a supervoid aligned with the cold spot of the cosmic microwave background. MNRAS, *450*(1), 288–294. https://doi.org/10.1093/mnras/stv488

Tristram, M., Macías-Pérez, J. F., Renault, C., & Santos, D. (2005). XSPECT, estimation of the angular power spectrum by computing cross-power spectra with analytical error bars. MNRAS, *358*(3), 833–842. https://doi.org/10.1111/j.1365-2966.2005.08760.x

Turner, E. L. (1980). The effect of undetected gravitational lenses on statistical measures of quasar evolution. ApJ, *242*, L135–L139. https://doi.org/10.1086/183418

Tyson, J. A., Kochanski, G. P., & Dell'Antonio, I. P. (1998). Detailed Mass Map of CL 0024+1654 from Strong Lensing. ApJ, *498*(2), L107–L110. https://doi.org/10.1086/311314

van Engelen, A., Bhattacharya, S., Sehgal, N., Holder, G. P., Zahn, O., & Nagai, D. (2014). Cmb lensing power spectrum biases from galaxies and clusters using high-angular resolution temperature maps. *The Astrophysical Journal*, *786*(1), 13. https://doi.org/10.1088/0004-637X/786/1/13

Weinberg, S. (1993). *The first three minutes : a modern view of the origin of the universe.*

White, M., Zhou, R., DeRose, J., Ferraro, S., Chen, S.-F., Kokron, N., Bailey, S., Brooks, D., García-Bellido, J., Guy, J., Honscheid, K., Kehoe, R., Kremin, A., Levi, M., Palanque-Delabrouille, N., Poppett, C., Schlegel, D., & Tarle, G. (2022). Cosmological constraints from the tomographic cross-correlation of DESI Luminous Red Galaxies and Planck CMB lensing. J. Cosmology Astropart. Phys., *2022*(2), Article 007, 007. https://doi.org/10.1088/1475-7516/2022/02/007

Williams, G. G., Olszewski, E., Lesser, M. P., & Burge, J. H. 90prime: a prime focus imager for the Steward Observatory 90-in. telescope (A. F. M. Moorwood & M. Iye, Eds.). *Ground-based instrumentation for astronomy* (A. F. M. Moorwood & M. Iye, Eds.). Ed. by Moorwood, A. F. M., & Iye, M. *5492*. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series. 2004, September, 787–798. https://doi.org/10.1117/12.552189.

Wright, E. L., Eisenhardt, P. R. M., Mainzer, A. K., Ressler, M. E., Cutri, R. M., Jarrett, T., Kirkpatrick, J. D., Padgett, D., McMillan, R. S., Skrutskie, M., Stanford, S. A., Cohen, M., Walker, R. G., Mather, J. C., Leisawitz, D., Gautier, I., Thomas N., McLean, I., Benford, D., Lonsdale, C. J., ... Tsai, C.-W. (2010). The Wide-field Infrared Survey Explorer (WISE): Mission Description and Initial On-orbit Performance. AJ, *140*(6), 1868–1881. https://doi.org/10.1088/0004-6256/140/6/1868

Xavier, H. S., Abdalla, F. B., & Joachimi, B. (2016). Improving lognormal models for cosmological fields. MNRAS, *459*(4), 3693–3710. https://doi.org/10.1093/mnras/stw874

Zaldarriaga, M., & Seljak, U. (1999). Reconstructing projected matter density power spectrum from cosmic microwave background. Phys. Rev. D, *59*(12), Article 123507, 123507. https://doi.org/10.1103/PhysRevD.59.123507

Zhang, L., Yu, Y., & Zhang, P. (2017). Non-negative Matrix Factorization for Self-calibration of Photometric Redshift Scatter in Weak-lensing Surveys. ApJ, *848*(1), Article 44, 44. https://doi.org/10.3847/1538-4357/aa8c7210.48550/arXiv.1612.04042

Zhang, P., Liguori, M., Bean, R., & Dodelson, S. (2007). Probing gravity at cosmological scales by measurements which test the relationship between gravitational lensing and matter overdensity. *Phys. Rev. Lett.*, *99*, 141302. https://doi.org/10.1103/PhysRevLett.99.141302

Zhang, P., Pen, U.-L., & Bernstein, G. (2010). Self-calibration of photometric redshift scatter in weak-lensing surveys. MNRAS, *405*(1), 359–374. https://doi.org/10.1111/j.1365-2966.2010.16445.x10.48550/arXiv.0910.4181

Zhang, Y., Pullen, A. R., Alam, S., Singh, S., Burtin, E., Chuang, C.-H., Hou, J., Lyke, B. W., Myers, A. D., Neveux, R., Ross, A. J., Rossi, G., & Zhao, C. (2021). Testing general relativity on cosmological scales at redshift z $\sim$ 1.5 with quasar and CMB lensing. MNRAS, *501*(1), 1013–1027. https://doi.org/10.1093/mnras/staa3672

Zwicky, F. (1933). Die Rotverschiebung von extragalaktischen Nebeln. *Helvetica Physica Acta*, *6*, 110–127.

Zwicky, F. (1937). On the Masses of Nebulae and of Clusters of Nebulae. ApJ, *86*, 217. https://doi.org/10.1086/143864

Chapter 7

Appendix

7.1 Generalised Covariance Matrix

We discuss here in detail the expression of covariance presented in Eq. 3.3. We start from the pseudo covariance given by

$$\widetilde{Cov}_{\ell\ell'}^{AB,CD} = \langle(\langle\tilde{C}_\ell^{AB}\rangle - \tilde{C}_\ell^{AB})(\langle\tilde{C}_{\ell'}^{CD}\rangle - \tilde{C}_{\ell'}^{CD})\rangle = \langle\tilde{C}_\ell^{AB}\tilde{C}_{\ell'}^{CD}\rangle - \langle\tilde{C}_\ell^{AB}\rangle\langle\tilde{C}_{\ell'}^{CD}\rangle \tag{7.1}$$

where $\tilde{C}_\ell$ is pseudo power spectrum and A,B,C,D represent scalar fields on sky. Let $\tilde{a}_{\ell m}$ be the spherical harmonic coefficients of $\tilde{C}_\ell$.

$$\begin{aligned}\widetilde{Cov}_{\ell\ell'}^{AB,CD} &= \frac{1}{(2\ell+1)(2\ell'+1)}\sum_{mm'}\left[\langle\tilde{a}^A_{\ell m}\tilde{a}^{B*}_{\ell m}\tilde{a}^C_{\ell' m'}\tilde{a}^{D*}_{\ell' m'}\rangle - \langle\tilde{a}^A_{\ell m}\tilde{a}^{B*}_{\ell m}\rangle\langle\tilde{a}^C_{\ell' m'}\tilde{a}^{D*}_{\ell' m'}\rangle\right] \\ &= \frac{1}{(2\ell+1)(2\ell'+1)}\sum_{mm'}\left[\langle\tilde{a}^A_{\ell m}\tilde{a}^{C*}_{\ell' m'}\rangle\langle\tilde{a}^D_{\ell' m'}\tilde{a}^{B*}_{\ell m}\rangle + \langle\tilde{a}^A_{\ell m}\tilde{a}^{D*}_{\ell' m'}\rangle\langle\tilde{a}^C_{\ell' m'}\tilde{a}^{B*}_{\ell m}\rangle\right]\end{aligned} \tag{7.2}$$

We can express $\tilde{a}_{\ell m}$ in terms of $a_{\ell m}$, the spherical harmonic coefficients of full sky power spectrum C_ℓ, using the mode-mode coupling kernel $K_{\ell m\ell' m'}$ (Hivon et al., 2002) as:

$$\tilde{a}_{\ell m} = \sum_{\ell' m'} a_{\ell' m' K_{\ell m\ell' m'}} \tag{7.3}$$

with which Eq. 7.2 becomes

$$\begin{aligned}\widetilde{Cov}_{\ell\ell'}^{AB,CD} = \frac{1}{(2\ell+1)(2\ell'+1)}\sum_{mm'}\sum_{\substack{\ell_1\ell_2\ell_3\ell_4\\ m_1m_2m_3m_4}}\Big[&\langle a^A_{\ell_1 m_1}a^{C*}_{\ell_2 m_2}\rangle K^A_{\ell m\ell_1 m_1}K^{C*}_{\ell' m'\ell_2 m_2}\langle a^D_{\ell_3 m_3}a^{B*}_{\ell_4 m_4}\rangle K^D_{\ell' m'\ell_3 m_3}K^{B*}_{\ell m\ell_4 m_4} \\ &+ \langle a^A_{\ell_1 m_1}a^{D*}_{\ell_3 m_3}\rangle K^A_{\ell m\ell_1 m_1}K^{D*}_{\ell' m'\ell_3 m_3}\langle a^C_{\ell_2 m_2}a^{B*}_{\ell_4 m_4}\rangle K^C_{\ell' m'\ell_2 m_2}K^{B*}_{\ell m\ell_4 m_4}\Big]\end{aligned} \tag{7.4}$$

Using $\langle a_{\ell m}a_{\ell' m'}\rangle = \delta_{\ell\ell'}\delta_{mm'}\langle C_\ell\rangle$, in Eq. 7.4, we get

$$\widetilde{Cov}_{\ell\ell'}^{AB,CD} = \frac{1}{(2\ell+1)(2\ell'+1)} \sum_{mm'} \sum_{\substack{\ell_1\ell_4 \\ m_1m_4}} \Big[\langle C_{\ell_1}^{AC} \rangle \langle C_{\ell_4}^{DB} \rangle K^{A}_{\ell m \ell_1 m_1} K^{C*}_{\ell' m' \ell_1 m_1} K^{D}_{\ell' m' \ell_4 m_4} K^{B*}_{\ell m \ell_4 m_4} + \langle C_{\ell_1}^{AD} \rangle \langle C_{\ell_4}^{CB} \rangle K^{A}_{\ell m \ell_1 m_1} K^{D*}_{\ell' m' \ell_1 m_1} K^{C}_{\ell' m' \ell_4 m_4} K^{B*}_{\ell m \ell_4 m_4} \Big] \tag{7.5}$$

We develop each term in Eq. 7.5 assuming the large sky coverage (Efstathiou, 2004):

$$\sum_{\substack{\ell_1\ell_4 \\ m_1m_4}} \langle C_{\ell_1}^{AC} \rangle \langle C_{\ell_4}^{DB} \rangle K^{A}_{\ell m \ell_1 m_1} K^{C*}_{\ell' m' \ell_1 m_1} K^{D}_{\ell' m' \ell_4 m_4} K^{B*}_{\ell m \ell_4 m_4} = \sqrt{C_\ell^{AC} C_{\ell'}^{AC} C_\ell^{DB} C_{\ell'}^{DB}} \sum_{\substack{\ell_1\ell_4 \\ m_1m_4}} K^{A}_{\ell m \ell_1 m_1} K^{C*}_{\ell' m' \ell_1 m_1} K^{D}_{\ell' m' \ell_4 m_4} K^{B*}_{\ell m \ell_4 m_4} \tag{7.6}$$

Expanding the mode-mode coupling kernels in terms of sum over pixels and then, applying the completeness relation of spherical harmonics:

$$\begin{aligned} \sum_{\ell_1 m_1} K^{X}_{\ell m \ell_1 m_1} K^{Y*}_{\ell' m' \ell_1 m_1} &= \sum_{\ell_1 m_1} \sum_{pq} w_p^X w_q^{Y*} \Omega_p \Omega_q Y_{\ell m}(\theta_p) Y^*_{\ell_1 m_1}(\theta_p) Y_{\ell_1 m_1}(\theta_q) Y^*_{\ell' m'}(\theta_q) \\ &= \sum_{pq} w_p^X w_q^{Y*} \Omega_p \Omega_q Y_{\ell m}(\theta_p) Y^*_{\ell' m'}(\theta_q) \frac{\delta(\theta_p - \theta_q)}{\Omega_q} \\ &= \sum_{p} w_p^{XY} \Omega_p Y_{\ell m}(\theta_p) Y^*_{\ell' m'}(\theta_p) \\ &= K^{XY}_{\ell m \ell' m'} \end{aligned} \tag{7.7}$$

where w_p is an arbitrary weight function, Ω_p is area of each pixel and we have defined $w_p^{XY} = w_p^X w_p^{Y*}$ as another arbitrary weight function. With Eq. 7.6 and Eq. 7.7, Eq. 7.5 simplifies as:

$$\widetilde{Cov}_{\ell\ell'}^{AB,CD} = \frac{1}{(2\ell+1)(2\ell'+1)} \sum_{mm'} \left[\sqrt{C_\ell^{AC} C_{\ell'}^{AC} C_\ell^{DB} C_{\ell'}^{DB}} K^{AC}_{\ell m \ell' m'} K^{BD*}_{\ell m \ell' m'} + \sqrt{C_\ell^{AD} C_{\ell'}^{AD} C_\ell^{CB} C_{\ell'}^{CB}} K^{AD}_{\ell m \ell' m'} K^{BC*}_{\ell m \ell' m'} \right] \tag{7.8}$$

The product of coupling kernels can be expanded in terms of Wigner-3j symbols as:

$$\begin{aligned} \sum_{mm'} K^{AC}_{\ell m \ell' m'} K^{BD*}_{\ell m \ell' m'} = \sum_{mm'} \sum_{\substack{\ell_1\ell_2 \\ m_1m_2}} w^{AC}_{\ell_1 m_1} w^{BD*}_{\ell_2 m_2} \frac{(2\ell+1)(2\ell'+1)}{4\pi} \sqrt{(2\ell_1+1)(2\ell_2+1)} \\ \times \begin{pmatrix} \ell & \ell' & \ell_1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell_2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell_1 \\ m & -m' & m_1 \end{pmatrix} \begin{pmatrix} \ell & \ell' & \ell_2 \\ m & -m' & m_2 \end{pmatrix} \end{aligned} \tag{7.9}$$

Using the orthogonality relations of Wigner-3j symbols, Eq. 7.9 simplifies as

$$\sum_{mm'} K^{AC}_{\ell m \ell' m'} K^{BD*}_{\ell m \ell' m'} = (2\ell+1) M^{AC,BD}_{\ell\ell'} \tag{7.10}$$

where $M_{\ell\ell'}^{AB,CD}$ is given by (Hivon et al., 2002)

$$M_{\ell\ell'}^{AB,CD} = \frac{2\ell'+1}{4\pi}\sum_{\ell_1}(2\ell_1+1)\left[\frac{1}{2\ell_1+1}\sum_{m_1} w_{\ell_1 m_1}^{AC} w_{\ell_2 m_2}^{BD*}\right]\begin{pmatrix}\ell & \ell' & \ell_1\\ 0 & 0 & 0\end{pmatrix}^2 \tag{7.11}$$

This transforms the expression for pseudo covariance matrix Eq. 7.8 as

$$\widetilde{Cov}_{\ell\ell'}^{AB,CD} = \frac{1}{(2\ell'+1)}\left[M_{\ell\ell'}^{AC,BD}\sqrt{C_\ell^{AC}C_{\ell'}^{AC}C_\ell^{DB}C_{\ell'}^{DB}} + M_{\ell\ell'}^{AD,BC}\sqrt{C_\ell^{AD}C_{\ell'}^{AD}C_\ell^{CB}C_{\ell'}^{CB}}\right] \tag{7.12}$$

The binned covariance matrix for full-sky is given by (Brown et al., 2005):

$$Cov_{LL'}^{AB,CD} = (M_{LL_1}^{AB^{-1}}P_{L_1\ell})\left[\frac{M_{\ell\ell'}^{AC,BD}\sqrt{C_\ell^{AC}C_{\ell'}^{AC}C_\ell^{DB}C_{\ell'}^{DB}} + M_{\ell\ell'}^{AD,BC}\sqrt{C_\ell^{AD}C_{\ell'}^{AD}C_\ell^{CB}C_{\ell'}^{CB}}}{(2\ell'+1)}\right](M_{L'L_2}^{CD^{-1}}P_{L_2\ell'})^T \tag{7.13}$$

Eq. 7.13 is similar to that obtained by Tristram et al., 2005. This expression takes into account different fractions of sky covered by the fields A, B, C, D.

7.2 Extragalactic Surveys in HELP fields

Table 7.1: Contribution of different surveys to HELP fields

Survey	Filter	% of objects			
		NGP	HS-82	SGP Part-1	SGP Part-2
DECaLS	g	15.9	-	-	-
	r	19.4	-	-	-
	z	11.6	-	-	-
DES	g	-	99.3	2.0	98.8
	r	-	99.5	2.2	99.8
	i	-	98.7	2.2	99.7
	z	86.1	99.3	2.0	99.3
	y	-	95.1	2.0	92.8
KiDS	u	-	-	90.2	1.2
	g	-	-	98.7	13.0
	r	-	-	99.8	19.6
	i	-	-	99.3	17.6
	z	-	-	16.6	7.7
PanSTARRS	g	99.3	43.8	0	9.4
	r	100	45.7	0	10.2
	i	100	46.0	0	10.6
	z	100	45.7	0	10.2
	y	99.6	44.2	0	9.6
RCSLenS	g	-	20.6	-	-
	r	-	23.9	-	-
	i	-	3.8	-	-
	z	-	20.9	-	-
	y	-	9.3	-	-
SDSS	u	-	35.8	-	-
	g	-	35.8	-	-
	r	-	35.8	-	-
	i	-	35.8	-	-
	z	-	35.8	-	-
SHELA + SpIES	IRAC12	-	26.1	-	-
UKIDSS-LAS	Y	59.5	10.6	-	-
	J	50.0	8.1	-	-
	H	56.8	10.0	-	-
	K	58.9	13.7	-	-
VISTA	Y	-	2.9	29.9	20.8
	J	-	23.0	34.5	23.0
	H	-	13.9	29.2	16.8
	Ks	-	22.7	31.1	18.3
	Z	-	-	51.3	34.9

7.3 Validation of deconvolution method

In this section, we present the performance of the deconvolution method as described in section 4.3.2 using a toy example.

We generate a fiducial distribution function by random sampling from the LSST photometric redshift distribution profile (Ivezić et al. 2019; LSST Science Collaboration et al. 2009) with mean redshift 0.9, which we term as true distribution. We convolve the true distribution with a Gaussian distribution with $\mu = 0.1, \sigma = 0.02$ and call the resultant as the observed distribution. In the left column of Figure 7.1, we show the true and observed distributions by blue and red solid lines, respectively. We attempt to show the robustness of our deconvolution method by applying it on un-smoothed distributions. In the right column of 7.1, we compare the true distribution (blue solid line) with the distribution recovered using our deconvolution method (red dashed line). The recovered distribution is in good agreement with the true redshift distribution, which validates our deconvolution method to reconstruct the true redshift distribution.

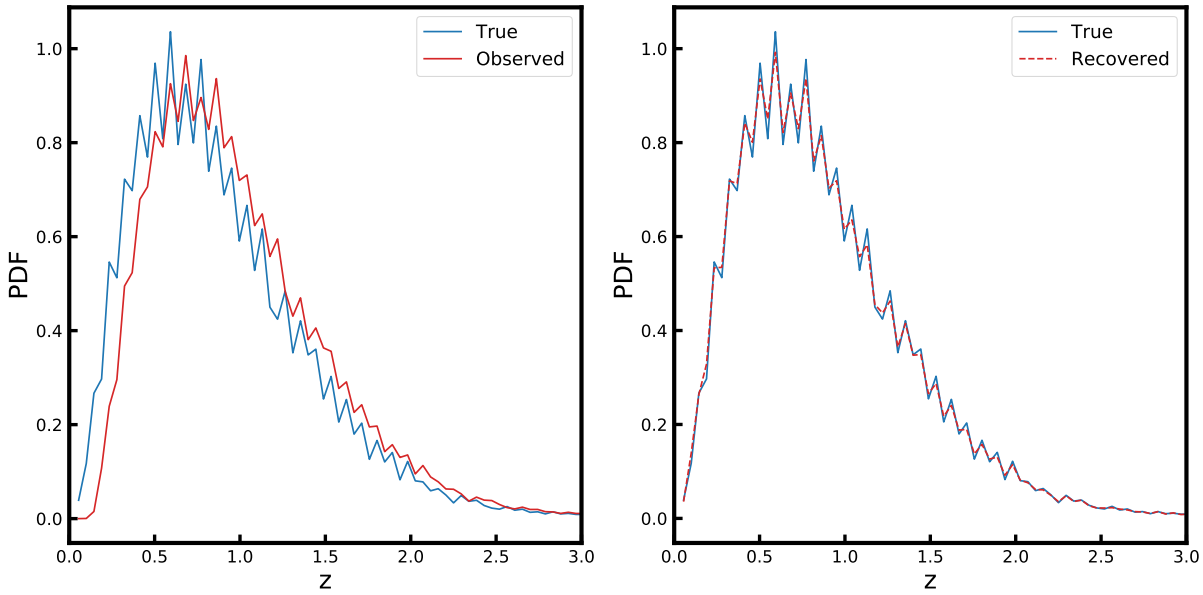


Figure 7.1: *Left*: True (blue solid line) and observed (red solid line) distributions for our toy example. *Right*: Comparison of the true redshift distribution (blue solid line) with the distribution recovered from deconvolution method (red dashed line). The recovered distribution is in excellent agreement with the true redshift distribution.