**Scuola Internazionale Superiore di Studi Avanzati - Trieste**

**Area of Physics**
**Ph.D. in Astrophysics**

# Cosmic Microwave Background and Large Scale Structure: Cross-Correlation as seen from *Herschel* and *Planck* satellites

**Candidate**
Federico Bianchini

**Supervisor**
Carlo Baccigalupi
Pawel Bielewicz
Andrea Lapi

Thesis submitted in partial fulfillment of the requirements
for the degree of Doctor Philosophiae Academic Year 2015/2016

**SISSA - Via Bonomea 265 - 34136 TRIESTE - ITALY**

# Declaration of Work

The research presented in this thesis was carried out in the Astrophysics Sector of the International School for Advanced Studies (SISSA/ISAS), Trieste, between November 2012 and August 2016. This thesis is the result of my own work except where explicit reference to the results of others has been made.

The content of this thesis is based on the following published or submitted papers:

- **Chapter 3**: Bianchini, F., Bielewicz, P., Lapi, A., et al. 2015, *Cross-correlation between the CMB lensing potential measured by Planck and high-z submillimiter galaxies detected by the HERSCHEL-ATLAS survey*, The Astrophysical Journal, 802, 64

- **Chapter 4**: Bianchini, F., Lapi, A., Calabrese, M., et al. 2016b, *Toward a tomographic analysis of the cross-correlation between Planck CMB lensing and H-ATLAS galaxies*, The Astrophysical Journal, 825, 24

- **Chapter 5**: Bianchini, F., Renzi, A., & Marinucci, D. 2016a, *Needlet estimation of cross-correlation between CMB lensing maps and LSS*, arXiv:1607.05223

- **Chapter 6**: Bianchini, F., & Silvestri, A. 2016, *Kinetic Sunyaev-Zel'dovich effect in modified gravity*, Physical Review D, 93, 064026

# Contents

*Sometimes science is a lot more art, than science. A lot of people don't get that.*

—Rick Sanchez, Rick & Morty

*Alla mia famiglia*

# 1 Chapter 1
# Introduction

> *We know of an ancient radiation*
> *That haunts dismembered constellations,*
> *A faintly glimmering radio station.*

—Frank Sinatra, CAKE

Over the past two decades cosmology has entered the age of *precision cosmology*[1]: the flood of data collected by cosmological experiments have transformed cosmology from a data-starving field into a data-driven field, which allows cutting-edge theories to be tested. A wide set of cosmological observations (Weinberg, 2008) have allowed us to summarize our fair understanding of the basic properties of the Universe in the concordance cosmological model - also known as ΛCDM model - built on the solid theoretical premises of Einstein's General Relativity. Despite the remarkable success in fitting the data, the standard model presents several puzzles on which light needs to be shed. For example, combining different cosmological data sets (Planck Collaboration, 2015c), baryonic matter can account for only 5% of the mass-energy budget of the Universe, while approximately 25% is in the form a *dark* component, namely the Dark Matter (DM) whose properties and features are not known yet. Moreover, the mechanism sourcing the late-time accelerated expansion of the Universe, whether being associated to some other dark component dubbed Dark Energy (DE) or to modifications of the underlying theory of gravity (MG), remains elusive.

Most of the information about the cosmological properties and parameters comes from accurate measurements of Cosmic Microwave Background (CMB) anisotropies, which provide us a snapshot of the early Universe. However, a wealth of information about the later evolution of the Universe can be gained by studying the interaction between CMB photons and its Large Scale Structure (LSS), giving rise to the so-called CMB *secondary anisotropies*. In this framework, two effects have recently

---

[1]As advocated by Peebles in Peebles (2002), the era of precision cosmology is the moment when "*all the important parameters will be established to one significant figure of merit or better, within the cosmological model.*"

emerged as new promising and powerful tools, the *weak gravitational lensing* of CMB photons and the *kinetic Sunyeav-Zel'dovich effect.*

CMB lensing is the deflection of photons trajectories due to the intervening matter distribution between the last scattering surface (the edge of the observable Universe) and us. On one hand it encodes a wealth of information about the geometry and the growth of structure in the Universe, on the other it represents a nuisance since it obscures the primordial inflationary polarization *B*-mode signal. In a period of less than 10 years, from 2007 - the year that I enrolled at the university - up to now, CMB lensing science has turned from the detection regime to being a standard cosmological probe. In fact, it is one of the main driver of the upcoming experimental efforts represented by a plethora of ground-based high-sensitivity CMB experiments such as the Simons Array, the South Pole Telescope (SPT-3G), the Advanced Atacama Cosmology Telescope (AdvACT), the Simons Observatory, and CMB-S4 devoted to an exquisite mapping of the CMB lensing sky.

The kinetic Sunyeav-Zel'dovich (kSZ) effect arises from the inverse Compton scattering of CMB photons off of free electrons, and it is specifically sourced by the bulk motion of these electrons with respect to the CMB rest frame. The large-scale momentum field is a powerful cosmological probe that can place tight constraints on the growth of structure and the epoch of reionization in a complementary way with respect to density fluctuations measurements.

This thesis focuses on how to exploit these two effects to address both cosmological and astrophysical questions, and it features theoretical work, observations, and data analysis. Specifically, we will exploit the synergy between the *Planck* and *Herschel* satellites by developing, validating, and applying cross-correlation algorithms to reconstruct the tomographic signal between the CMB lensing and the positions of the sub-millimeter selected galaxies. As pictorially shown in Fig. 1.1, such a signal is detectable because we are basically probing the same underlying field, i.e. the integrated matter distribution on cosmological scales.[2] Galaxies represent the biased signpost of the DM halos, the most massive of which act as lenses for the CMB photons, providing a clean and independent measure of the relation between luminous and dark matter. Cross-correlation measurements are generically less sensitive to any known (and unknown) systematics, as well as to extract signals hidden in noisy data, even though the full potential of these measurements is still under scrutiny, and it is one of the drivers of the present work. We anticipate that the key aspect of the present work is an accurate interplay between CMB lensing

---

[2]As a little disclaimer we note that the (gradients of) gravitational potentials are technically responsible for the deflection of CMB photons rather than the matter distribution itself, though in the standard scenario (General Relativity) the two can be related through the Poisson equation.
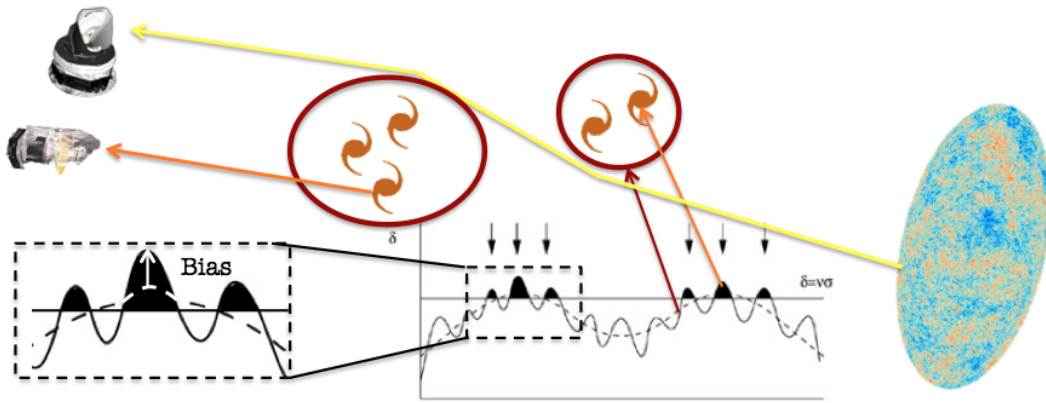
**Figure 1.1:** Pictorial explanation of the origin of the CMB lensing - galaxy density cross-correlation signal. While CMB photons collected by the *Planck* satellite experience several deflections during their cosmic journey because of the matter perturbations (represented by DM halos), the dusty enshrouded galaxies that reside in these halos emit infrared photons captured by *Herschel*.

datasets as seen from *Planck* on one side, and an accurate selection of DM tracers on Herschel data on the other. This results in the most accurate detection of the cross-correlation to date, which allows to exploit it to set cosmological constraints from cross-correlations, which will become tighter and tighter in the incoming years.

The second main topic discussed in this thesis is the impact of modified gravity theories on the kSZ effect of the CMB. Several approaches to detect and study the kSZ have emerged on the market, here we focus on the kSZ signal showing up at small scales in the temperature CMB power spectrum (assuming an efficient foreground removal by means of multi-frequency observations). By being dependent on the details of the growth history, especially at late times, and on the epoch of reionization, the kSZ can be exploited to test gravity on cosmological scales. This field of research is at the early stages of development, though it is very promising both as a cosmological and astrophysical tool.

**Outline of the Thesis**

This thesis is organized as follows. The first two chapters are propaedeutic to follow the main content. In Ch. 2 we review the basics of the current cosmological model, carefully focusing on the physics and applications of two of the main pillars, the CMB and the weak gravitational lensing. The following chapter, Ch. 3, deals with the statistics of random fields, since these are fundamental to model and analyze the data, and provides a thorough description of the exploited *Planck* and *Herschel* datasets. The remaining four chapters represent the core of the work carried out during my Ph.D. and exposed in this thesis. In Ch. 4 we present

the first cross-correlation measurement between the *Planck* CMB lensing and the spatial distribution of the high-$z$ ($z \gtrsim 1.5$) dusty star-forming galaxies detected by *Herschel*-ATLAS, measuring the linear galaxy bias $b$ and reporting a cross-correlation signal somewhat stronger than expected. We improve this analysis with updated datasets in Ch. 5, where we attempt a first tomographic approach by splitting the galaxy sample in two redshift bins and studying the evolution of the cross-correlation amplitude and galaxy bias over cosmic time. Ch. 6 is devoted to the methodology and algorithms to investigate the cross-correlation between CMB and LSS datasets: we discuss the spectral estimation problem in the needlet framework, developing a `MASTER`-like technique to address the bias induced by heavy masking, and compare it with its harmonic counterpart. In Ch. 7 we consider the capability of the kSZ power spectrum measured by the small-scale CMB telescope to constrain models of modified gravity, in particular we focus on a specific $f(R)$ model, the Hu & Sawicki one. Finally, we summarize the main results and future perspectives in the Conclusions.

# 2 | Chapter 2
## Cosmology and gravitational lensing

*Roll up! Roll up for the*
*magical mystery tour!*
*Step right this way!*

—Magical Mystery Tour,
The Beatles

Once you become familiar with the fact that the farther out we look in space, the
further back we see in time, the Universe is put in a new perspective as it unravels
before our eyes. Over the centuries we have come to discover that the Universe is
an incredible *place* and, in my personal view, the most amazing aspect is perhaps
that it can be remarkably well described by mathematical means. In fact, our fair
understanding of the Universe has been summarized in the so-called Standard Hot
Big Bang (HBB) model which rests on the following assumptions:

- On sufficiently large scales the Universe is homogeneous and isotropic (the
  Cosmological Principle (CP)).

- General Relativity (GR) is the theory to describe gravitational interactions
  on cosmological scales.

- The energy budget of the Universe can be modeled in terms of cosmological
  fluids with barotropic equation of state (EoS).

The model states that the Universe evolved from an initial singularity[1] and it has
been expanding ever since. As a result, the primordial Universe was in a hotter
and denser state where all species were in thermal equilibrium. This picture is
corroborated by three fundamental observations: the distance dependent recession
of galaxies from which we can infer the expansion of the Universe (encoded in the
*Hubble law*), the abundances of light elements indicating that the Universe went

---

[1]Curiously, the current paradigm does not explain the *Big Bang* in the HBB (the initial
singularity). In fact, the term HBB was coined by Fred Hoyle, a supporter of the alternative
*steady state* cosmological model.

through an hot and dense phase, and the Cosmic Microwave Background (CMB) which is interpreted as the afterglow of the primeval plasma.

The theoretical limitations of the HBB model, such as the horizon problem or the origin of the structure seeds, can be overcome within the inflationary framework which postulates a period of accelerated expansion in the early Universe, driven by a transient phase of non-zero vacuum energy provided by the potential of a scalar field, the inflaton. As a result, adiabatic and Gaussian perturbations were generated from quantum fluctuations of the inflaton field (as predicted by the simplest inflationary scenario). Then, according to the current paradigm of structure formation, the objects we now observe have originated from the gravitational collapse of such small perturbations in a homogeneous, expanding Universe. The earliest imprint of the primordial inhomogeneities can be observed in the temperature anisotropies of the Cosmic Microwave Background (CMB), a snapshot of the Universe at the time of recombination, allowing for tight constraints on the space of models. This picture of the infant Universe is however distorted by interactions of CMB photons with the evolving medium through which they are propagating. An important effect is the weak gravitational lensing of photons trajectories due to the gradient in the gravitational potentials associated to the Large Scale Structure (LSS) in the Universe. Its characteristic effect on the CMB anisotropies pattern enables the reconstruction of lensing maps which probe the total integrated matter distribution out to the last scattering surface, hence opening a unique window on fundamental physics aspects that affect the structure formation in the Universe such as gravity theories, neutrino sector and Dark Matter (DM).

This Chapter reviews the theoretical framework as well as the key observational tools in cosmology and it is split in three main sections. Sec. (2.1) describes the modeling of the background Universe together with its constituents, and it sketches how structures form and evolve from the initial conditions set by inflation. The physics of the CMB temperature and polarization anisotropies is discussed in Sec. (2.2), while the last part, Sec. (2.3), is dedicated to the weak gravitational lensing as a cosmological probe, focusing on the case of CMB lensing.

## 2.1 FLRW Cosmology

### 2.1.1 Zeroth-order cosmology: the Background Universe

The CP is meant to be statistical, that is the Universe is *on average* homogeneous (invariant under spatial translations, i.e. the cosmological properties are the same

throughout) and isotropic (invariant under spatial rotations, i.e. there are no preferred directions). Homogeneity does not imply isotropy and viceversa; however, if we accept the Copernican Principle, according to which we do not live in a particular position in the Universe, then isotropy guarantees the homogeneity of the whole Universe (Ellis, 1975). While homogeneity can hardly be tested, the isotropy is well established by observations of the CMB (Bennett et al., 1996), the X-ray background (Boughn et al., 2002), and the spatial distribution of galaxies (Marinoni et al., 2012). For a discussion on cosmologies that do not rely on the Copernican Principle, we refer the reader to Clarkson (2012).

The assumption of the CP allows us to define a universal time variable, the *cosmic time*, and to represent the Universe by a time-ordered sequence of three-dimensional spatial slices, each of which is homogeneous and isotropic. If we also consider the Universe not to be static, as Hubble observations suggest, then the Friedmann-Lemaître-Robertson-Walker (FLRW) metric is the most general metric describing an expanding spacetime. When expressed in the *comoving* spherical coordinate system $x^i = (\chi, \theta, \phi)$ it has the following form:

$$\mathrm{d}s^2 = -c^2\,\mathrm{d}t^2 + a^2(t)\gamma_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j \tag{2.1}$$

$$= -c^2\,\mathrm{d}t^2 + a^2(t)\Big[\mathrm{d}\chi^2 + f_K^2(\chi)(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)\Big] \tag{2.2}$$

$$= -c^2\,\mathrm{d}t^2 + a^2(t)\Big[\mathrm{d}\chi^2 + f_K^2(\chi)\,\mathrm{d}\Omega^2\Big], \tag{2.3}$$

where the *scale factor* $a(t)$ (normalized to $a(t_0) = 1$ today) encodes the expansion of the Universe and the *comoving transverse distance*[2] can take - depending on the spacetime curvature - the following forms:

$$f_K(\chi) = \begin{cases} \sin\chi & \text{if } K = 1 \\ \chi & \text{if } K = 0 \\ \sinh\chi & \text{if } K = -1, \end{cases} \tag{2.4}$$

Physical size are related to the comoving ones through $d_\mathrm{p} = a(t)\chi$. The physical velocity of an object is $v = \dot{d}_\mathrm{p} = \dot{a}\chi + a\dot{\chi} \equiv H d_\mathrm{p} + v_\mathrm{pec}$, where we defined the *Hubble parameter* $H \equiv \frac{\dot{a}}{a} = \frac{\mathrm{d}a}{a\,\mathrm{d}t}$ that describes the expansion history of the Universe and the dot denotes proper time derivative, i.e. $\dot{} \equiv \frac{\partial}{\partial t}$. It is customary to normalize its present day value, $H_0$, in terms of the dimensionless parameter $h$ as $H_0 = 100\,h$ km/s/Mpc. The Hubble parameter naturally introduces a characteristic expansion timescale given by $H^{-1}$ which corresponds to the age of the Universe at that time[3]. Similarly to distances we can define the *conformal time* as $\mathrm{d}\eta = \mathrm{d}t/a(t)$: with this

---

[2]Also referred to as the *metric* or *proper distance*.
[3]It is roughly the time in which the scale factor doubles.

definition, the Hubble parameter becomes $\mathcal{H} \equiv \mathrm{d}\ln a / \mathrm{d}\eta = aH$.

An important consequence of the Universe expansion is that light gets stretched. A photon emitted at a time $t_{\mathrm{em}}$ with wavelength $\lambda_{\mathrm{em}}$ will be observed at a time $t_{\mathrm{obs}}$ with a wavelength $\lambda_{\mathrm{obs}}$ given by

$$\frac{a(t_{\mathrm{obs}})}{a(t_{\mathrm{em}})} = \frac{\lambda_{\mathrm{obs}}}{\lambda_{\mathrm{em}}} \equiv 1 + z, \tag{2.5}$$

where the *cosmological redshift* $z$ was defined.

### Distances

Distances in cosmology are not uniquely defined. The *comoving distance* between us and a source at redshift $z$ is given by

$$\chi = \int_t^0 \frac{c\,\mathrm{d}t'}{a(t')} = \int_a^1 \frac{c\,\mathrm{d}a'}{a'^2 H(a')} = \int_0^z \frac{c\,\mathrm{d}z'}{H(z')}. \tag{2.6}$$

In a flat universe ($K = 0$), the comoving transverse distance is simply equal to the comoving distance $\chi$. Note also that both $f_K$ and $\chi$ distances are *not* observable. The causality structure of the Universe is determined by the photons trajectories: in FLRW cosmologies we can define the *particle horizon* $\chi_{\mathrm{ph}}$ as the greatest comoving distance from which light could have reached us by now (assuming the Big Bang at $t_i = 0$)

$$\chi_{\mathrm{ph}} = \int_0^t \frac{c\,\mathrm{d}t'}{a(t')} = \int_0^a \frac{c\,\mathrm{d}a'}{a'^2 H(a')} = \int_z^\infty \frac{c\,\mathrm{d}z'}{H(z')}. \tag{2.7}$$

It is straightforward to see that $\chi_{\mathrm{ph}}(t) = c\eta$. Regions which are separated by a distance larger than the horizon size $d_{\mathrm{ph}}(t) = a(t)\chi_{\mathrm{ph}}(t)$ cannot be in causal contact. It is worth to stress that the particle horizon $\chi_{\mathrm{ph}}$ and Hubble radius $(aH)^{-1}$ are roughly the same for a FLRW cosmology but are generally different.[4] The Hubble radius is the (comoving) distance over which particles can travel in the course of one expansion time. So, if two regions are separated by $\lambda > \chi_{\mathrm{ph}}$, they have *never* communicated, whereas if $\lambda > (aH)^{-1}$, then they are not in causal contact *now*. Cosmologists commonly make use of standard candles, i.e. objects with known luminosity $L$ (like the Type-Ia SN), and standard rulers, i.e. sources with a known physical size $l$ (such as the fluctuations in the CMB). These probes allow us to define the *luminosity distance* $D_L$ as the distance to a standard candle with measured

---

[4]Strictly speaking, this is true if the energy budget of the Universe is dominated by a *pressure-less dust* component.
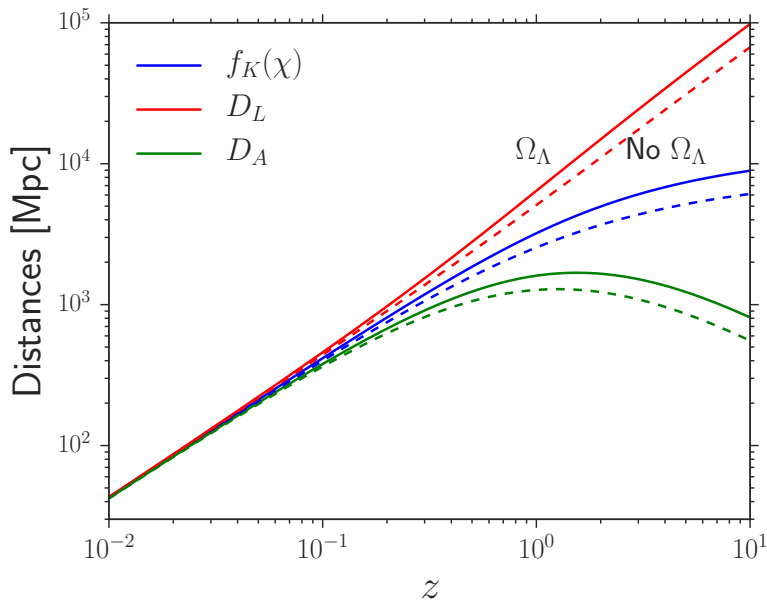
**Figure 2.1:** Different distance measures for a flat cosmology with cosmological constant $\Omega_\Lambda = 0.7$ (solid lines) and matter-only (dashed lines).

flux $F = \frac{L}{4\pi D_L^2(z)}$ as:

$$D_L \equiv \sqrt{\frac{L}{4\pi F}} = \frac{f_K(\chi)}{a} = (1+z)f_K(\chi), \tag{2.8}$$

and the *angular diameter distance* as the distance to an object of dimension $l$ subtending and angle $\delta\theta$:

$$D_A \equiv \frac{l}{\delta\theta} = a(t)f_K(\chi) = \frac{f_K(\chi)}{1+z}. \tag{2.9}$$

From Eq. (2.8) and (2.9) we see that angular diameter and luminosity distances are not independent but they are related through the Etherington's distance-duality relation (Etherington, 1933):

$$D_L = (1+z)^2 D_A. \tag{2.10}$$

The redshift dependence of the three distance measures $f_K(\chi)$, $D_L$, and $D_A$ is shown in (2.1) for a flat cosmology with (solid lines) and without dark energy (dashed lines) in the form of a cosmological constant $\Lambda$, a cosmic component with vacuum energy-like features that we introduce below, hypothesized to source the late-time cosmic acceleration. As can be seen, all the distances are enlarged by the presence of $\Lambda$.

**Friedmann Equations**

So far the discussion focussed on the symmetries of the spacetime and led us to the definition of the FLRW metric. The dynamics of the background Universe, described by the time evolution of $a(t)$, can be worked out by relating the metric to the energy content of the Universe. This is achieved via the Einstein's Field Equations (EFE)

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi G T_{\mu\nu} + \Lambda g_{\mu\nu}, \tag{2.11}$$

where $G_{\mu\nu}$ and $R_{\mu\nu}$ are the Einstein and Ricci tensors computed from the metric, $R = R^\mu_\mu$ is the Ricci scalar, and $T_{\mu\nu}$ is the Stress Energy Tensor (SET) which describes the energy content of the Universe. In order to allow for static solutions ($\dot{a} = 0$) of the EFE, which were considered to describe our Universe before Hubble's discovery of recession of galaxies, Einstein introduced a *cosmological constant* term $\propto \Lambda g_{\mu\nu}$ whose presence does not spoil the Bianchi identity. In principle this term could be equivalently added to the l.h.s., and be interpreted as a geometrical feature of the spacetime, or to the r.h.s. of Eq. (2.11) as a contribution to the SET $T^{\rm vac}_{\mu\nu} = \frac{\Lambda}{8\pi G}g_{\mu\nu}$. The physical interpretation is still debated but a non-zero cosmological constant provides nowadays the simplest explanation for the cosmic acceleration. Homogeneity and isotropy also constrain the form of the SET to be the one of a perfect fluid at rest in comoving coordinates:

$$T^\mu_\nu = (\rho + P)u^\mu u_\nu - P\delta^\mu_\nu, \tag{2.12}$$

where $u^\mu$ is the relative four-velocity between the fluid and the observer, while $\rho = \rho(t)$ and $P = P(t)$ are the *energy density* and *pressure* in the rest-frame of the fluid. In this case, the time-time component of the EFE can be solved to yield the time evolution of the expansion rate $H$

$$H^2 \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho + \frac{\Lambda}{3} - \frac{K}{a^2}, \tag{2.13}$$

while the spatial components of the EFE reduce to a single expression for the acceleration:

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3P) + \frac{\Lambda}{3}, \tag{2.14}$$

where $\rho = \sum_a \rho_a$ and $P = \sum_a P_a$ are the total energy density and pressure of the Universe. Eqs (2.13) and (2.14) are usually called the *Friedmann equations*, from which we see that the total energy density of a flat Universe ($K = 0$) corresponds

to the *critical density* $\rho_{\text{crit}}$ defined as

$$\rho_{\text{crit}} = \frac{3H^2}{8\pi G}.$$  (2.15)

Note that it depends on time and its present value $\rho_{\text{crit}}^0$ is approximately $10^{-27}$ kg/m$^3$. The critical density is useful to define dimensionless density parameters for the different species as $\Omega_i(t) = \frac{\rho_i(t)}{\rho_{\text{crit}}(t)}$. Cosmological fluids are assumed to be *barotropic*, that is, their pressure is given as an explicit function of their energy density so that we can define the fluid EoS $w$ as $P = w(\rho)\rho$ and the *sound speed* as $c_s^2 = \delta P / \delta \rho$. From the conservation of the SET, $\nabla^\mu T_{\mu\nu} = 0$, we find that the time evolution of the energy density obeys

$$\dot{\rho} + 3H(\rho + P) = 0.$$  (2.16)

The continuity equation applies separately to each species as long as their particle number is conserved and their energy exchange can be neglected. Then, for a constant $w$, the energy density scales as $\rho \propto a^{-3(1+w)}$, while for a time-dependent EoS $w(a)$, one has to solve the following integral:

$$\rho \propto \exp\left(-3\int_0^a \frac{\mathrm{d}a'}{a'}[1 + w(a')]\right).$$  (2.17)

With these scalings and the dimensionless densities, we can rewrite Eq. (2.13) as

$$H^2(z) = H_0^2\Big[\Omega_{\text{r}0}(1+z)^4 + \Omega_{\text{m}0}(1+z)^3 + \Omega_{\text{K}0}(1+z)^2 + \Omega_\Lambda\Big].$$  (2.18)

It is useful to classify the different species according to their contribution to pressure as:

- **Matter**: All sources for which the internal pressure is much smaller than the energy density, $|P| \ll \rho$, like the case of a gas of non-relativistic particles (where the mass $m$ is much bigger than the momentum $p$). Baryons (i.e. nuclei and electrons) and DM represent two dust-like species. In this case $w \approx 0$ and the energy density is diluted by cosmic expansion as $\rho \propto a^{-3}$.

- **Radiation**: The term refers to species for which the pressure is about a third of their energy density, $P = \frac{1}{3}\rho$ so that $w = \frac{1}{3}$ and $\rho \propto a^{-4}$. This is the case of a gas of relativistic particles (for which $m \ll p^2$), like photons, neutrinos and massive species while still relativistic.

- **Dark Energy**: Even though the term indicates a plethora of models which source the cosmic acceleration, here we consider a cosmological constant

which exerts a negative pressure $P = -\rho$ so that $w = -1$. In this case the energy density is constant and does not feel the cosmic dilution, $\rho \propto a^0$. The vacuum energy predicted by quantum field theory provides a natural explanation for Dark Energy (DE), however its expected density $\rho_{\mathrm{vac}}$ is off by $\approx 120$ orders of magnitude with respect to the observed one.

**Cosmic Inventory**

What we have sketched above is just a framework that employs few parameters which are not fixed *a priori*. Deciding which values the cosmological parameters should assume for our Universe relies on observational constraints. A combination of cosmological probes, including observations of the CMB[5], the LSS (such as the Baryonic Acoustic Oscillations (BAO) and galaxy clusters), and the Type-Ia SN magnitude-redshift relation has now placed strong constraints on the cosmological parameters. The emerging picture is an almost flat Universe mainly composed of *dark* ingredients we cannot yet theoretically account for. Observations show that the Universe is filled with radiation ($\Omega_{\mathrm{r}}$), matter ($\Omega_{\mathrm{m}}$), and DE ($\Omega_{\Lambda}$) with an EoS that remarkably resembles the one of a cosmological constant $w_{\Lambda} \approx -1$. Ordinary baryonic matter can account for just 5% of the total budget, while most of the matter content is in the form of a clustering dark component which appears to interact only gravitationally or weakly at most. The most popular candidate for DM is a Weakly Interacting Massive Particle (WIMP), an hypothetical elementary particle. WIMPs are generally referred to as Cold Dark Matter (CDM) because their velocity (or thermal pressure) is small so that they behave as a non relativistic fluid. DM properties have a pivotal role in structure formation: if CDM is dominant, then larger structures form by accreting smaller structures. Relativistic species are represented by photons (mainly in the form of CMB) and neutrinos. Summing up all the ingredients we have $\Omega_K = 1 - \Omega_{\mathrm{r}0} - \Omega_{\mathrm{m}0} - \Omega_{\Lambda} \simeq 0$. As can be seen in Fig. (2.2), depending on the species that is the most abundant we can identify three epochs in the cosmic history: the *radiation dominated era* ($a \propto \eta$), the *matter domination era* ($a \propto \eta^2$) and the *DE dominated era* ($a \propto 1/\eta$ ). The transitions between the three eras take place at $a_{\mathrm{eq}} = \frac{\Omega_{\mathrm{r}0}}{\Omega_{\mathrm{m}0}} = 2.93 \times 10^{-4}$ and $a_{\Lambda} = \frac{\Omega_{\mathrm{m}0}}{\Omega_{\Lambda}} = 0.46$ (corresponding to $z_{\mathrm{eq}} = 3410$ and $z_{\Lambda} = 1.18$, respectively). This concordance model goes under the name of $\Lambda$ Cold Dark Matter ($\Lambda$CDM) and the main constituents contribution to the total $\Omega$, as well as other cosmological parameters that will be defined later on, is listed in Table (2.1).

---

[5]From an historical point of view, CMB measurements provide the tightest constraints.

| | Densities | | | |
|---|---|---|---|---|
| **Parameter** | $\Omega_{\mathrm{b}}$ | $\Omega_{cdm}$ | $\Omega_{\Lambda}$ | $\Omega_{\mathrm{r}}$ |
| **Value** | $0.0490^{+0.0005}_{-0.0005}$ | $0.2642^{+0.0005}_{-0.0005}$ | $0.685^{+0.013}_{-0.013}$ | $\simeq 9.2 \times 10^{-5}$ |
| | Others | | | |
| **Parameter** | $H_0$ | $n_s$ | $10^9 A_s$ | $\tau$ |
| **Value** | $67.31^{+0.96}_{-0.96}$ | $0.9655^{+0.0062}_{-0.0062}$ | $2.198^{+0.076}_{-0.085}$ | $0.078^{+0.019}_{-0.019}$ |

**Table 2.1:** Parameters of the vanilla $\Lambda$CDM cosmology computed from the 2015 baseline Planck likelihoods for the CMB temperature power spectra plus low-$\ell$ polarization data (referred to as *Planck* TT+lowP in Planck Collaboration (2015c)).

### Dark Energy or Modified Gravity?

The accelerated expansion of the Universe, first *discovered* in 1998 (Riess et al., 1998; Perlmutter et al., 1999), has raised great interest both in the cosmology and the physics community as a whole: the mechanism behind the acceleration still remains elusive and it represents one of the main scientific drivers of the upcoming experimental efforts. Even though a cosmological constant is consistent with the current observations, it is essential to consider alternatives from a phenomenological perspective. Over the last years a classification scheme has emerged to distinguish between physical scenarios that source the cosmic acceleration: namely the *Dark Energy* and *Modified Gravity* (MG) models. Naively, DE models modify the r.h.s. of EFE by adding an additional component to the SET, while Modified Gravity (MG) models alter the l.h.s. of EFE, hence modifying the Einstein-Hilbert action (i.e. they change GR itself). Giving a thorough overview of all existing theories is far beyond the purposes of this thesis, but we refer the reader to Silvestri & Trodden (2009); Clifton et al. (2012) for complete reviews on the topic. Here we just sketch the main ideas behind the two approaches and, following the prescription of Joyce et al. (2016), we distinguish between DE and MG by means of the *strong equivalence principle* (SEP): any theory which obeys the SEP is classified as DE, while any theory which violates it as MG. Heuristically, the SEP forbids the presence of fifth forces.

The most natural extension to the cosmological constant paradigm is that is a dynamical field (dubbed DE) which relaxes to its present value via some mechanism (Wetterich, 1988; Ratra & Peebles, 1988). This idea is at the base of many different theoretical implementations, that all share the presence of a degree of freedom that drives the background cosmological evolution. Perhaps the most obvious generalization of the canonical scalar field with a potential is the following:

$$S = \int \mathrm{d}^4x \sqrt{-g}\Lambda^4 K(X), \tag{2.19}$$
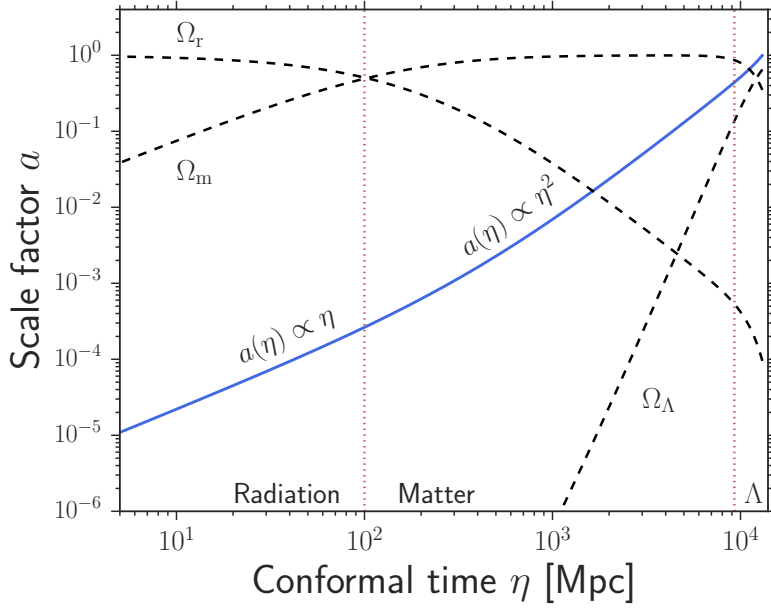
**Figure 2.2:** Cosmic history of the Universe. The blue curve is the scale factor as a function of conformal time, obtained by solving the Friedmann equation, while the dashed lines represent the radiation, matter and cosmological constant densities as function of conformal time. Adapted from Pettinari (2016).

where we defined $X \equiv -\frac{1}{2\Lambda^4}(\partial\phi)^2$ and $K(X)$ is an arbitrary function: these models are usually referred to as $k-$essence (Chiba et al., 2000; Armendariz-Picon et al., 2000). Similar to the canonical case, if the field has an homogeneous profile $\phi = \phi(t)$, these models behave as a (slightly more complicated) perfect fluid with EoS parameter $w$ given by

$$w = \frac{K}{2XK_{,X} - K}. \tag{2.20}$$

The background expansion history can be reproduced by a suitable choice of the functional form for $K$. It is worth to notice that in $k-$essence models the field perturbations do not propagate luminally but with a sound speed given by

$$c_s^2 = \frac{K_{,X}}{K_{,X} + 2XK_{,XX}}, \tag{2.21}$$

allowing for significant differences with $\Lambda$CDM in the structure formation.
On the other hand, modifications to GR lead to new physics on small scales (Lue et al., 2004), introducing new degrees of freedom in the gravity sector. Since these new particles mediate a fifth-force, MG theories must develop some *screening mechanism* in order to evade the very constraining local tests of gravity, like the ones on Solar System scales. In Ch. (7.3) we will discuss a specific model of MG named $f(R)$, where the standard Einstein-Hilbert action gets modified by the presence of a generic function of the Ricci scalar.

## 2.1.2 First-order cosmology: the Perturbed Universe

The assumption of homogeneity and isotropy significantly simplifies the mathematical description of the Universe as a whole, however it just holds when looking at scales larger than approximately 100 Mpc. On smaller scales observations of the galaxies distribution and of the CMB fluctuations anisotropy pattern tell us that the Universe is far from being homogeneous, showing a complex structure and rich hierarchy composed by stars, globular clusters, galaxies, galaxy clusters, filaments and voids. These structures are thought to be originated from primordial fluctuations through gravitational instability. As we mentioned already, the Universe is not homogeneous even on larger scales, as it is seen in the anisotropies of the CMB. Here we sketch the theory of cosmological perturbations, first introduced by Lifshitz (1946); classic literature on the subject is by Kodama & Sasaki (1984); Ma & Bertschinger (1995) and the review by Bernardeau et al. (2002).

**Relativistic perturbations**

The basic idea is to consider small perturbations around the FLRW metric and the perfect fluid SET and couple them through EFE; after subtracting the background, we are left with[6]

$$\delta G_{\mu\nu} = 8\pi G \delta T_{\mu\nu}. \tag{2.22}$$

If the perturbations are small, we can perform a perturbative expansion on both sides and keep the linear terms. Assuming the background metric to be spatially flat, the most general form for the metric fluctuations can be written in conformal coordinates as

$$ds^2 = a^2(\eta)\{-(1+2\phi)\,d\eta^2 + 2B_i\,dx^i\,d\eta + [(1-2\psi)\delta_{ij}^K + E_{ij}]\,dx^i\,dx^j\}. \tag{2.23}$$

The perturbations $\psi$ and $\phi$ are scalars, while $B_i$ and $E_{ij}$ can be split into scalar, vector and tensor parts as:

$$B_i = \underbrace{\partial_i B}_{\text{scalar}} + \underbrace{\hat{B}_i}_{\text{vector}} \quad \text{with} \quad \partial^i \hat{B}_i = 0 \tag{2.24}$$

$$E_{ij} = \underbrace{\partial_{\langle i}\partial_{j\rangle} E}_{\text{scalar}} + \underbrace{\partial_{(i}\hat{E}_{j)}}_{\text{vector}} + \underbrace{\hat{E}_{ij}}_{\text{tensor}} \quad \text{with} \quad \partial^i \hat{E}_i = \partial^i \hat{E}_{ij} = 0, \tag{2.25}$$

---

[6]We remind that here the dot represents proper time derivative $\dot{} \equiv \frac{\partial}{\partial t}$, while the prime denotes conformal time derivative $' \equiv \frac{\partial}{\partial \eta}$.

where

$$\partial_{\langle i}\partial_{j\rangle}E \equiv \left(\partial_i\partial_j - \frac{1}{3}\delta_{ij}\nabla^2\right)E, \tag{2.26}$$

$$\partial_{(i}\hat{E}_{j)} \equiv \frac{1}{2}\left(\partial_i\hat{E}_j + \partial_j\hat{E}_i\right). \tag{2.27}$$

Scalar fluctuations are induced by density inhomogeneities and exhibit gravitational instability leading to structure formation. Vector modes are strongly suppressed by cosmic expansion. Tensor perturbations are an important prediction of the inflationary paradigm and describe gravitational waves. Scalar-Vector-Tensor (SVT) decomposition is powerful because it decouples Einstein equations for scalar, vectors and tensor modes, whose evolutions can be thus treated separately. However, the problem of relativistic perturbation theory is the gauge redundancy in GR: the way to deal with such issues is either to fix a gauge (i.e., a coordinate system) in order to make any two of the four scalar functions $\phi$, $\psi$, $B$ and $E$ vanish or work in terms of *gauge-invariant* perturbations.[7] The most common gauge-invariant linear combinations of these function are the so called *Bardeen's potentials* (Bardeen, 1980; Mukhanov et al., 1992):

$$\Phi \equiv \phi - \frac{1}{a}[a(B-E')]' \quad \text{and} \quad \Psi \equiv \psi + \frac{a'}{a}(B-E'). \tag{2.28}$$

On the r.h.s. of Eq. (2.22) we can perturb the SET to linear order like:

$$\delta T_0^0 = -\delta\rho \tag{2.29}$$

$$\delta T_i^0 = (\bar{\rho} + \bar{P})v_i \tag{2.30}$$

$$\delta T_j^i = \delta P\delta_j^i + \Pi_j^i, \tag{2.31}$$

where bar denotes background quantities and $\Pi_j^i$ is the anisotropic stress tensor. If we substitute the metric from Eq. (2.23) into EFE and keep only the linear order perturbations, we find the *gauge-invariant* perturbation equations (Mukhanov, 2005):

$$\nabla^2\Psi - 3\mathcal{H}(\Psi' + \mathcal{H}\Phi) = 4\pi G a^2 \delta T_0^0 \tag{2.32}$$

$$\partial_i(\Psi' + \mathcal{H}\Phi) = 4\pi G a^2 \delta T_i^0 \tag{2.33}$$

$$\left[\Psi'' + \mathcal{H}(2\Psi + \Phi)' + (2\mathcal{H}' + \mathcal{H}^2)\Phi + \frac{1}{2}\nabla^2(\Phi - \Psi)\right]\delta_{ij} - \frac{1}{2}\partial_i\partial_j(\Phi - \Psi) = -4\pi G a^2 \delta T_j^i. \tag{2.34}$$

---

[7]Here, $B$ and $E$ refer to the scalar components of the metric fluctuations and should not be confused with the $E$- and $B$-modes decomposition of the CMB polarization that we will discuss in Sec. (2.2.1).

These equations can be derived without imposing any gauge conditions and they are valid in an arbitrary coordinate system.

**Conformal Newtonian gauge**

Here we work in the *conformal Newtonian gauge* (or the *longitudinal gauge*), defined by choosing $B = E = 0$ in Eq. (2.23): in this coordinate system the metric takes the form

$$\mathrm{d}s^2 = a^2(\eta)[-(1 + 2\Phi)c^2\,\mathrm{d}\eta^2 + (1 - 2\Psi)\delta_{ij}\,\mathrm{d}x^i\,\mathrm{d}x^j]. \tag{2.35}$$

Note that this metric is similar the usual weak-field limit of GR about Minkowski space: in the Newtonian limit, $\Phi$ plays the role of the gravitational potential, i.e. it governs the dynamics of non-relativistic objects, while the combination (Weyl potential) $\Psi_\gamma \equiv (\Psi + \Phi)/2$ determines the null geodesics, i.e. light propagation. Going to Fourier space, the perturbation equations in Eq. (2.32) expressed in the conformal Newtonian gauge for a mode with comoving wavenumber $k$ become:

$$k^2\Psi + 3\mathcal{H}(\Psi' + \mathcal{H}\Phi) = -4\pi G a^2 \delta\rho \tag{2.36}$$

$$k^2(\Psi' + \mathcal{H}\Phi) = 4\pi G a^2 (\bar{\rho} + \bar{P})\theta \tag{2.37}$$

$$\Psi'' + \mathcal{H}(2\Psi' + \Phi') + \frac{3}{2}\mathcal{H}^2(1 + w)\Phi + \frac{k^2}{3}(\Psi - \Phi) = 4\pi G \delta\rho \tag{2.38}$$

$$k^2(\Psi - \Phi) = 12\pi G a^2 (\bar{\rho} + \bar{P})\sigma, \tag{2.39}$$

where $\theta \equiv \partial_i v^i = \nabla \cdot \mathbf{v}$ is the velocity divergence and $\sigma$ is the anisotropic stress, i.e. the trace-free part of $T_{ij}$. The source terms on the r.h.s. are meant to be the sum over all the relevant matter-energy components: in the absence of anisotropic stress[8] $\Phi = \Psi$. Conservation of the SET, $\nabla_\mu T^{\mu\nu} = 0$, yields the two following equations (for $\mu = 0$ and $\mu = i$) in the conformal Newtonian gauge (Ma & Bertschinger, 1995)

$$\delta' = -(1 + w)(\theta - 3\Psi') - 3\mathcal{H}(c_s^2 - w)\delta \tag{2.40}$$

$$\theta' = -\mathcal{H}(1 - 3w)\theta - \frac{w'}{1 + w}\theta + \frac{c_s^2}{1 + w}k^2\delta + k^2\Phi - k^2\sigma, \tag{2.41}$$

which are the relativistic generalization of the continuity and Euler equation. The equations derived above are fully relativistic and provide an adequate description of perturbations on scales larger than the Hubble radius and for relativistic fluids (like photons and neutrinos).

---

[8]In reality, neutrinos develop anisotropic stress after neutrino decoupling (i.e. they do not behave like a perfect fluid), so that $\Phi$ and $\Psi$ differ from each other by about 10 % between the neutrino decoupling and the matter-radiation equality; during matter domination neutrinos contribution is negligible and the two potentials approach each other.

Looking at the relative size between the perturbation wavenumber and Hubble radius, we can distinguish two regimes: $k \ll \mathcal{H}$ called super-Hubble modes, and $k \gg \mathcal{H}$, referred to as sub-Hubble modes[9]. In the regime when we are deep within the Hubble radius, we can take the non-relativistic limit of the perturbed EFE in order to greatly simplify the description of perturbations. However, these equations are nonlinear and several efforts have been made in the past to find approximate solutions for a wide range of scales and redshifts: the most straightforward approach goes under the name of Standard Perturbation Theory, while other approaches are the so called Renormalized Perturbation Theory and the Effective Field Theory. The efforts aimed at improving the perturbations description towards nonlinear scales are essential for a successful comparison of theory and data in the light of upcoming LSS surveys such as Euclid, Large Synoptic Survey Telescope (LSST), Dark Energy Spectroscopic Instrument (DESI), and Wide Field Infrared Survey Telescope (WFIRST).

**Scales smaller than the Hubble radius**

Assuming a DM dominated Universe, on *sub-Hubble scales* the linearized[10] Poisson, continuity and Euler equations for matter perturbations read:

$$\nabla^2 \Phi = \frac{3}{2}\Omega_{\mathrm{m}}(\eta)\mathcal{H}^2(\eta)\delta \tag{2.42}$$

$$\delta' + \theta = 0 \tag{2.43}$$

$$\mathbf{v}' + \mathcal{H}\mathbf{v} = -\nabla\Phi. \tag{2.44}$$

Combining them, one finds a second-order linear differential equation for the linear density contrast

$$\delta'' + \mathcal{H}\delta' - \frac{3}{2}\mathcal{H}^2\Omega_{\mathrm{m}}\delta = 0. \tag{2.45}$$

In this equation only time derivatives appear, hence solutions can be factorized into a spatial and time-dependent solutions, associated to a growing and decaying mode; in general we can write

$$\delta(\mathbf{x}, t) = D_+(\eta)A(\mathbf{x}) + D_-(\eta)B(\mathbf{x}), \tag{2.46}$$

where $A(\mathbf{x})$ and $B(\mathbf{x})$ are two arbitrary functions describing the initial density field, while $D_\pm$ are the growth and decay functions. In a matter dominated Universe

---

[9]Note that in literature super- and sub-Hubble modes are often (improperly) called super-and sub-horizon modes. See the discussion on the differences in Sec. (2.1.1)

[10]We assume the densities and velocities to be small, i.e. $\delta \ll 1$ and $v \ll c$.

$\mathcal{H} = 2/\eta$, so that

$$D_+(\eta) = \eta^2 \quad \text{and} \quad D_-(\eta) = \eta^{-3}. \tag{2.47}$$

The explicit time-dependence makes clear why the two solutions are known as growing and decaying modes: the relevant mode for structure formation is the former, while the latter quickly fades away. It is customary to normalize the solution by the growing mode today, such that we define the *growth factor* as :

$$D(a) \equiv \frac{D_+(a)}{D_+(a_0 = 1)}. \tag{2.48}$$

### Scales greater than the Hubble radius

What does happen when the perturbation is much larger then the Hubble radius? Assuming that the pressure depends only on the energy density and the EoS $w$ is a constant, then we have $c_s^2 = w$ (which is valid both for matter and radiation). In this case the first two lines of Eq. (2.32) can be combined to yield

$$\Phi'' + 3\mathcal{H}(1 + c_s^2)\Phi' = 0, \tag{2.49}$$

for which $\Phi = $ const is the growing solution. Note that this super-Hubble solution is independent on $w$ as long as it is constant; in particular, the gravitational potential is frozen outside the Hubble radius (during radiation and matter domination). It is also possible to show that a frozen potential at large scales implies that the density contrast $\delta \approx -2\Phi = $ constant as well. Density perturbations are therefore simply proportional to the curvature perturbation: this is a key ingredient for setting the initial conditions and to link inflation with perturbation evolution.

### Transfer function, power spectra and galaxy bias

So far we just described the qualitative behavior of perturbations in an expanding Universe and we have seen that it depends on (i) whether the perturbation mode is larger or smaller than the Hubble radius and (ii) on the species that dominates the cosmic energetic budget. Let's summarize the main results here.

The gravitational potential $\Phi$ is constant when the modes are outside the horizon: if they cross the horizon during the radiation era, they start to oscillate and their amplitude decreases as $a^2$, while during the matter era the potential is constant on all scales. During the radiation domination the radiation fluctuations are constant on super-Hubble scales, while they oscillate if $k \gg \mathcal{H}$; in the matter era, the subhorizon fluctuations in the radiation density oscillate with constant amplitude around a shifted equilibrium point. For CDM, we have that perturbations on super-Hubble scales remain frozen both during radiation and matter domination:

this means that the properties of such modes are preserved from the early Universe. If the perturbations enter the horizon during the radiation-dominated epoch, they grow logarithmically, $\delta_{\mathrm{m}} \propto \ln a$, while the growth is a power-law ($\delta_{\mathrm{m}} \propto a$) if the horizon is crossed during the matter domination. When the Universe becomes dominated by DE, perturbations stop growing. Baryons behavior is somewhat different because they are coupled with photons through Compton scattering prior to the decoupling at $z \sim 1100$, so that we can treat the photons and baryons as a single fluid. On sub-Hubble scales they oscillate with decreasing amplitude together with the photons, while after the decoupling, since the effective baryons pressure drops to approximately zero, their oscillations are not supported by photons, enabling gravitational collapse (if the perturbation scale is below the *Jeans length*). To account for all the interactions between different components and the effects at play on all scales, such as causality, one can define a *transfer function $T(k)$* as (Eisenstein & Hu, 1998)

$$T(k) \equiv \frac{\delta(k, z = 0)}{\delta(k, z = \infty)} \frac{\delta(0, z = \infty)}{\delta(0, z = 0)}, \tag{2.50}$$

which can be defined for each species and, by construction, $T \to 1$ for $k \to 0$. In order to derive $T(k)$ one has to numerically solve the Boltzmann equation (as done by the Code for Anisotropies in the Microwave Background (CAMB) (Lewis et al., 2000) or the Cosmic Linear Anisotropy Solving System (CLASS) (Lesgourgues & Tram, 2011)) but there also can be found fitting formulae in literature, see Bardeen et al. (1986); Eisenstein & Hu (1998). Then the (matter) power spectrum at a given time $\langle|\delta(\boldsymbol{k}, z)|^2\rangle$ is proportional to the square of the transfer function and the growth factor, multiplied by the initial power spectrum $P_{\mathrm{ini}}(k)$ set by inflation (as we will see in Sec. (2.1.3))

$$P_{\delta\delta}(k, z) = D^2(z) T^2(k) P_{\mathrm{ini}}(k). \tag{2.51}$$

The normalization of the matter power spectrum is determined from observations and it is common practice to measure the variance of the smoothed overdensity field over a comoving scale $R$ (usually taken to be 8 Mpc/$h$) at present as

$$\sigma_R^2 = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} P_{\delta\delta}(k, z = 0) |W_R(k)|^2, \tag{2.52}$$

where $|W_R(k)|^2$ is a window function which usually is a top-hat filter in real space.

However, we need to bare in mind that our observations rely on using galaxies to trace the overall matter distribution and as such, they provide a biased picture of

matter clustering: the statistical relation between the dark and luminous matter distribution is described by the *galaxy bias b*. The idea that galaxy formation is enhanced in dense environment was introduced by Kaiser (1984)[11] and generalized by Bardeen et al. (1986); Sheth & Tormen (1999) in the peak-background split framework. Fry & Gaztanaga (1993) then proposed that galaxies $\delta_g$ trace matter distribution $\delta$ through a non-linear transformation, $\delta_g(\boldsymbol{x}) = \mathcal{F}[\delta(\boldsymbol{x})]$, which can be Taylor-expanded to obtain the so-called *local[12] biasing scheme*:

$$\delta_g(\boldsymbol{x}) = \sum_i^{\infty} \frac{b_i}{i!} \delta^i(\boldsymbol{x}). \tag{2.53}$$

The perturbative expansion is usually halted at the linear order and one finds $\delta_g = b_1 \delta$, so that the galaxy power spectrum reads as $P_{gg}(k) = b_1^2 P_{\delta\delta}(k)$. At small scales the linear bias relation breaks down because of nonlinear structure formation and acquires non-local features like the scale dependence, even though for $k \lesssim 0.01 h/\mathrm{Mpc}$ the linear relation should hold. Measuring the galaxy bias of the high-$z$ dusty star forming galaxies (DSFG) will be a central topic of Ch. (4) and (5).

A semi-analytical model that has proven useful to address both nonlinear clustering and galaxy bias is the *halo model*, developed from the pioneering work of Neyman & Scott (1952) and then formalized by Scherrer & Bertschinger (1991); Seljak (2000) (see Cooray & Sheth (2002) for a comprehensive review). The basic idea is that at late times all the matter distribution is contained only in virialized DM objects called *halos*, and the galaxy distribution can be modeled by an appropriate scheme for placing galaxies in these DM halos. The fundamental assumption of the model is that the average halo properties, such as its density profile and its galaxy content, only depend on the halo mass. As a consequence, the total matter power spectrum splits into two terms, $P_{\delta\delta}(k) = P_{\delta\delta}^{1h}(k) + P_{\delta\delta}^{2h}(k)$ where $P_{\delta\delta}^{1h}$ is the 1-halo term describing correlations on small scales between mass within the same halo and $P_{\delta\delta}^{2h}$ is the 2-halo term arising from large-scale correlation of matter in different halos.

---

[11] In order to explain the stronger correlation function of galaxy clusters with respect to the galaxies one.

[12] This is because the galaxy overdensity at some position $\boldsymbol{x}_1$ is not affected by the density displaced in $\boldsymbol{x}_2$, otherwise the bias is said to be non-local and the coefficients $b_i$ would carry spatial dependence.

## 2.1.3 Setting the initial conditions: Inflation

The HBB scenario we depicted so far is a remarkably simple model that allows us to describe the Universe both at the background and at the perturbation level. Nonetheless, the model is unable to explain the following theoretical and observational problems which we summarize briefly here below. It is beyond the scope of this thesis to give a comprehensive description of the physics of the very early Universe, and our discussion is targeted to define the power spectrum of scalar perturbations generated during inflation. For a more complete discussion, we refer the reader to Mukhanov (2005).

**Flatness problem.** The spatial curvature is observationally constrained to be negligible ($|\Omega_K| \ll 1$) (Planck Collaboration, 2015c). Moreover, from the Friedmann equation, one can show that $\Omega_K$ grows with time, so that the spatial curvature becomes vanishingly small as time runs backward. This means that a flat geometry is an unstable solution and if the density today is close to the critical one, as observations suggest, it had to be much more so in the past, requiring unnatural initial conditions (*fine tuning* problem).

**Horizon problem.** Any correlations between regions separated by a distance greater than the particle horizon cannot be naturally explained by the standard model. However, the CMB appears to be isotropic (to $10^{-5}$ precision, see Sec. (2.2)) over much bigger scales[13].

**Monopoles problem.** Most extensions of the standard model of particle physics, like the Grand Unification Theory (GUT), predict the creation of magnetic monopoles which would be produced during transition phases in the early Universe. However, no monopole relics have been observed to date.

**Structure problem.** Despite the solid perturbation theory formalism developed in the sections above, we lack of a mechanism that explains the formation of the primordial seeds that eventually lead to the structure formation via gravitational collapse.

As we already mentioned, these shortcomings are all connected to the Universe initial conditions and can be solved by postulating the existence of a period of accelerating expansion in the early Universe, the so-called *cosmic inflation* (Guth, 1981; Starobinsky, 1982; Linde, 1982). The effect of the cosmic acceleration predicted by the inflationary models[14] is to hugely increase the size of the Universe, washing out any curvature and stretching the geometry of the Universe in a way

---

[13]Barbara Ryden in her cosmology textbook (Ryden, 2003) compares this situation to "*inviting 20.000 people to a potluck dinner and they all bring potato salad*".

[14]The term inflation refers to a *mechanism* rather than a *theory* of the early Universe.

that it becomes spatially flat, hence solving the flatness problem. Moreover cosmic inflation solves the horizon problem by connecting regions that, in the standard HBB model, would be causally disconnected. This can be achieved if the comoving Hubble radius at the beginning of inflation is larger than the largest scale observable today, i.e. the current Hubble radius.[15] Inflation also solves the apparent monopole paradox by completely diluting their density, so that there may be an arbitrarily small number of relics in the observable Universe. Inflationary models also predict an almost scale-invariant power spectrum of primordial density perturbations, which serve as initial conditions to the process of structure formation, as we explain in some more detail below.

For the cosmic expansion to be accelerated, Eq. (2.14) requires the Universe to be dominated by a species of negative pressure which fulfills:

$$\ddot{a} > 0 \Leftrightarrow \rho + 3P < 0 \Leftrightarrow w < -\frac{1}{3}. \tag{2.54}$$

This can be achieved if we consider a Universe in which a cosmological constant is dominant at early times. Then, since $w = -1$, the background expansion is de Sitter and the scale factor increases exponentially in time, $a(t) \propto e^{Ht}$. At the end of inflation, the duration of which can be "tuned" in order to solve the problems discussed above, the cosmic acceleration comes to an end and the cosmological constant decays into a thermal mix of elementary particles, leading to a radiation dominated Universe. This process is the so-called *reheating* and thereafter the Universe can evolve according to the standard HBB model.

Over the years a plethora of inflationary models have emerged on the market (for a comprehensive review of the *zoology* see, for instance, Martin et al. (2014)), and they all share the presence of one or more real scalar fields $\phi_i$, with associated potentials $V_i(\phi_i)$, which drive the accelerated expansion. The particles corresponding to these fields are called *inflatons.*

To sketch how single-field inflation works, consider a Universe dominated by an homogeneous scalar field $\phi$ that behaves as a perfect fluid with

$$\rho_\phi = \frac{1}{2}\dot{\phi}^2 + V(\phi) \tag{2.55}$$

$$p_\phi = \frac{1}{2}\dot{\phi}^2 - V(\phi). \tag{2.56}$$

---

[15]It is worth to stress that it is not the accelerated expansion that solves the horizon problem: the causal connection is established *before* the inflationary epoch and then inflation simply puts these regions out of reach again.

Then, the equation of motion of the scalar field is given by

$$\ddot{\phi} + 3H\dot{\phi} + \frac{\mathrm{d}V}{\mathrm{d}\phi} = 0, \qquad (2.57)$$

which corresponds to an oscillator equation for a scalar field experiencing the damping of the Hubble expansion. Since the inflaton dominates the energy budget, the Friedmann equation becomes

$$H^2 = \frac{8\pi G}{3}\left[\frac{1}{2}\dot{\phi} + V(\phi)\right] - \frac{K}{a^2}, \qquad (2.58)$$

and we immediately notice that if $a$ grows by many orders of magnitude then the term $K/a^2 \to 0$, leading to a spatially flat universe. The sought-after accelerated expansion can be achieved under the *slow-roll condition* that is, $\dot{\phi}^2 \ll V(\phi)$. In this case the $w_\phi \approx -1$ so that the scale factor grows as $a \propto e^{Ht}$, where $H(t) \approx$ const for slowly varying potentials. Several potentials can be tailored in order to satisfy the slow-roll condition and it is common practice to parametrize them with the two *slow-roll parameters* defined as

$$\epsilon \equiv \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{H}\right) = -\frac{\dot{H}}{H^2} \quad \text{and} \quad \eta \equiv \frac{\dot{\epsilon}}{H\epsilon}. \qquad (2.59)$$

If both $\{\epsilon, \eta\} \ll \mathcal{O}(1)$, then slow-roll inflation takes place.

Cosmic inflation also provides a mechanism to generate primordial density fluctuations Starobinsky (1982); Hawking (1982); Bardeen et al. (1983). These primordial fluctuations are the result of the microscopic quantum vacuum fluctuations in the inflation field that, during inflation, get stretched out by the accelerated expansion and imprinted on super-horizon scales. Eventually, the density fluctuations reenter the horizon and serve as initial conditions for the growth of structure in the Universe. The requirement $\epsilon \ll 1$ translates into an approximate time-translation invariance of the background, so that primordial fluctuations generated during inflation are expected to have the same variance on all scales. We can define the power spectrum of the *primordial comoving curvature perturbation*, $\mathcal{R} = \Psi - \frac{1}{3}\frac{\delta\rho_{tot}}{\bar{\rho}_{tot}+\bar{p}_{tot}}$, as

$$\mathcal{P}_\mathcal{R}(k) = \frac{k^3 P_\mathcal{R}(k)}{2\pi^2} = A_s\left(\frac{k}{k_*}\right)^{n_s-1}, \qquad (2.60)$$

where the scale invariance[16] is usually expressed in terms of the *scalar spectral*

---

[16] The exact scale invariance is achieved if $H$ and $\phi$ are constant in time during inflation; however this is only an approximation as we know that the inflaton must have a dynamics otherwise the inflation continues forever.

*index $n_s$*. $A_s$ is the *amplitude of the primordial scalar fluctuations* evaluated at the the pivot scale $k_*$ (commonly taken to be 0.05 $h/$Mpc or 0.02 $h/$Mpc). The scalar spectral index is linked to the slow-roll parameters through $n_s = 1 - 6\epsilon + 2\eta$ and recent measurement from *Planck*, in combination with other experiments, measured a value of $n_s = 0.9655 \pm 0.0062$ (Planck Collaboration, 2015c), excluding $n_s = 1$ at more than $8\sigma$. The primordial power spectrum appearing in Eq. (2.51) can be taken to be $P_{\text{ini}}(k) = 2\pi^2 k^{-3} \mathcal{P}_{\mathcal{R}}(k) = P_{\mathcal{R}}(k)$. Another robust and model-independent prediction from the inflationary mechanism is the production of a stochastic background of primordial gravitational waves that is also expected to be nearly scale invariant. Similarly to the case of scalar perturbations in Eq. (2.60), we define the power spectrum of tensor fluctuations as

$$\mathcal{P}_t(k) = A_t \left(\frac{k}{k_*}\right)^{n_t}. \tag{2.61}$$

The tensor and scalar fluctuations are related through the tensor-to-scalar ratio $r \equiv \frac{\mathcal{P}_t(k_*)}{\mathcal{P}_{\mathcal{R}}(k_*)}$, which represent a valuable probe of the inflation physics since it directly sets the energy scale for inflation as

$$V^{1/4} \approx \left(\frac{r}{0.01}\right)^{1/4} 10^{16} \, \text{GeV}. \tag{2.62}$$

The statistics of primordial perturbations is usually characterized in terms of power spectra, however they are just one of the infinite $n$-point functions that describe a random field (see Sec (3.1)). For a Gaussian random field these moments can be written in terms of the 2-point function, i.e. the power spectrum, while for generic random fields this is not true. Higher-order moments carry information about the nonlinear physics of the early Universe and, in particular, about the interactions of the inflaton. Since the inflationary potential is expected to be very flat, the interactions are expected to be very small as well. The deviations of primordial fields from a perfectly Gaussian distribution are known as *non-Gaussianities* (NG); for a pedagogical review on primordial NG from inflationary models we refer the reader to Chen (2010). In this context, CMB has become one of the best laboratory to study NG: statistics like the bispectrum (3-point function) and the trispectrum (4-point function) of the CMB can be used to assess the level of primordial NG (and possibly its shape) on different cosmological scales and to disentangle it from the one induced by secondary anisotropies and systematic effects (Bartolo et al., 2010). To date, the tightest constraints on NG, and on the inflationary epoch in general, are provided by the latest release of *Planck* collaboration (Planck Collaboration, 2016b).

## 2.2 Cosmic Microwave Background

The CMB was first fortuitously discovered in 1965 by Arno Penzias and Robert Wilson while working on long-distances radio communications at the Bell Laboratories[17] (Penzias & Wilson, 1965). It represents the first compelling evidence of the HBB model proposed by Gamow in the forties (Gamow, 1946) and it is interpreted as being the leftover radiation of the primordial plasma - when the Universe was in a hot and dense state - and as such it is characterized by a blackbody spectrum at a temperature of about 3K . In the early Universe ($10^3 \lesssim z \lesssim 10^8$), protons ($p$), electrons ($e^-$) and photons ($\gamma$) were tightly coupled into the so called photo-baryonic plasma. Protons and electrons interact through Coulomb scattering while photons were coupled to electrons via Compton scattering (Thomson scattering in the low-energy limit), maintaining the three species in thermal equilibrium through:

$$p + e^- \leftrightarrow H + \gamma \tag{2.63}$$

$$e^- + \gamma \leftrightarrow e^- + \gamma. \tag{2.64}$$

The thermal equilibrium allows to define a single temperature $T$ for the photo-baryonic fluid, common to all the species. The cosmic expansion causes the Universe to cool down, hence diluting the density of photons with energy greater than the hydrogen ionization energy, $E_\gamma > 13.6$ eV. The effect is to spoil the balance in the reaction in Eq. (2.63), leading to $p + e^- \to H + \gamma$: this process is dubbed *recombination* and happens at a redshift $z_{\mathrm{rec}} \simeq 1400$ (corresponding to $T_{\mathrm{rec}} \simeq 3900$K), when approximately 50% of the free electrons were captured by protons into hydrogen atoms.[18] The Universe becomes transparent to radiation only when the photon mean free path is larger than the Hubble radius at that time, i.e. when the electron-photon interaction rate $\Gamma(z)$ is similar to the Hubble expansion rate: $\Gamma(z_{\mathrm{dec}}) \simeq H(z_{\mathrm{dec}})$, where we defined the *decoupling* redshift. Observation suggest that the previous condition is satisfied at $z_{\mathrm{dec}} = 1089.90 \pm 0.23$ (Planck Collaboration, 2015c).

Thermal equilibrium ensures that the Spectral Energy Distribution (SED) of the

---

[17]The first observations of the CMB were actually made in the forties by McKellar (McKellar, 1940), who studied the population of excited rotational states of CN molecules in interstellar absorption lines.

[18]In reality the temperature of recombination is much smaller than the binding energy of the hydrogen atom, $T_{\mathrm{rec}} \ll 13.6$ eV. This is mainly because the photon-to-baryon ratio is very large, $\approx 10^9$, and the high-energy tail of the photons distribution is still efficient in breaking hydrogen atom down to lower temperatures.

CMB photons is well described by Planck's statistics

$$B_\nu(T) = \frac{2h v^3}{c^2} \frac{1}{e^{\frac{h\nu}{k_B T}} - 1},$$ (2.65)

where $h$ is the Planck constant, $\nu$ is the frequency and $k_B$ is the Boltzmann constant. The first measurement of the CMB energy spectrum over the frequency range $\nu \in [50, 650]$ GHz has been made in the early nineties with the Far-InfraRed Absolute Spectrophotometer (FIRAS) instrument (Mather et al., 1994) on board of the COsmic Background Explorer (COBE) satellite, showing that it has a Planckian shape with an equivalent brightness temperature of about $T = 2.725 \pm 0.001$ (Fixsen et al., 1996).

Measurements of the CMB blackbody spectrum can tightly constrain energy injections in the primordial plasma. While any energy injection prior to $z_{\text{therm}} \approx 10^7$ would have time to be thermalized[19], a posterior energy release would cause a distortion of the black-body spectrum (Danese & de Zotti, 1982; Burigana et al., 1991). Two main types of *spectral distortions* are usually distinguished, the $y$- and $\mu$-type distortions. The former type arises at $z \lesssim \text{few} \times 10^3 - 10^4$ when the Compton process ceases to be extremely efficient and is characterized by a constant temperature decrement at low and an increment at high frequencies. The latter type of distortion forms at times when the Compton process can still achieve full kinetic equilibrium with the electrons ($z \lesssim \text{few} \times 10^5$), and the CMB spectrum reaches a Bose-Einstein distribution with non-zero chemical potential $\mu$. After the constraining FIRAS measurements on the spectral distortions, $|\mu| < 9 \times 10^{-5}$ and $|y| < 1.5 \times 10^{-5}$ at 95% confidence level, the subject of spectral distortion has recently attracted a renewed theoretical attention (Chluba & Sunyaev, 2012; Chluba, 2014) especially in light of upcoming satellite missions aiming at improving such measurements such as the Primordial Inflation Explorer (PIXIE) (Kogut et al., 2011) and Cosmic Origin Explorer (CORE).

### 2.2.1 Primary anisotropies

**Temperature power spectrum**

The CMB photons carry information about the density, velocity and gravitational potential inhomogeneities imprinted by inflation to the photo-baryonic plasma. These tiny inhomogeneities in the plasma source the tiny departures from the

---

[19]With the term *thermalization* we refer to the process that erases possible spectral distortions of the CMB in the early Universe, hence attempting to restore full thermodynamic equilibrium.

average temperature, the so called *primary* CMB temperature anisotropies[20], which provide a snapshot of the Universe at the time of decoupling.

The CMB temperature sky can be decomposed into

$$T^{obs}(\hat{\mathbf{n}}) = T_0 \left[ 1 + (\boldsymbol{\beta} \cdot \hat{\mathbf{n}}) + \frac{T(\hat{\mathbf{n}})}{T_0} \right], \tag{2.66}$$

where $T_0$ is the mean blackbody temperature, $\boldsymbol{\beta} = \boldsymbol{v}/c$ is the Earth proper velocity vector with respect to the CMB rest frame, and $T(\hat{\mathbf{n}})$ is the CMB temperature fluctuation field. The intrinsic temperature anisotropies are of the order $\Theta \equiv \frac{\Delta T}{T_0} \sim \mathcal{O}(10^{-5})$, approximately two orders of magnitude smaller than the kinematic dipole signal due to the Doppler boosting. Let us stress that Eq. (2.66) is just an approximation as it does not include the contribution from foreground emissions $F_\nu(\hat{\mathbf{n}})$ that has to be taken into account when dealing with real data.

The intrinsic stochastic nature of the quantum fluctuations during inflation prevents us from developing a theory that exactly predicts the pattern of $T(\hat{\mathbf{n}})$, nonetheless the problem can be properly tackled from a statistical point of view. We start by decomposing our scalar CMB temperature field $T(\hat{\mathbf{n}})$ in spherical harmonics $Y_{\ell m}(\hat{\mathbf{n}})$, a frequency-space orthonormal basis for representing functions defined over the sphere (see Sec. (3.1.2)), such that[21]

$$\Theta(\hat{\mathbf{n}}) = \sum_{\ell=2}^{\infty} \sum_{m=-\ell}^{\ell} a_{\ell m}^T Y_{\ell m}(\hat{\mathbf{n}}), \quad \text{where} \quad a_{\ell m}^T = \int d\Omega \, \Theta(\hat{\mathbf{n}}) Y_{\ell m}^*(\hat{\mathbf{n}}), \tag{2.67}$$

and the angular scale $\theta$ and multipole moment $\ell$ are related by $\theta \sim 1/\ell$.

The randomness of the initial conditions does not allow us to predict any particular $a_{\ell m}^T$ coefficient, but different theories tell us about the distribution from which they are drawn. Considering $T(\hat{\mathbf{n}})$ to be a Gaussian real-valued random field, then the harmonic coefficients are complex Gaussian random variables which satisfy the following relations:

$$\langle a_{\ell m}^T \rangle = 0, \, \forall \{\ell, m\}, \tag{2.68}$$

$$\text{Isotropy} \Rightarrow \langle a_{\ell m}^T a_{\ell' m'}^{T*} \rangle = \delta_{\ell \ell'}^K \delta_{mm'}^K C_\ell^{TT}, \tag{2.69}$$

$$T(\hat{\mathbf{n}}) \in \mathbb{R} \Rightarrow a_{\ell-m}^T = (-1)^m a_{\ell m}^{T*}, \tag{2.70}$$

where $\langle \cdot \rangle$ denotes an average over an ensemble of sky realizations and $C_\ell^{TT}$ is the *angular power spectrum*, i.e. the variance of the $a_{\ell m}^T$. The CP guarantees

---

[20]Here, the term *anisotropies* refers to the temperature fluctuations as function of the line-of-sight (LOS) and should not be confused with the notion of *statistical anisotropic* random fields.

[21]The monopole $\ell = 0$ and dipole $\ell = 1$ are left out, see Eq. (2.66).

that the covariance matrix of the harmonic coefficients is diagonal, otherwise the off-diagonal components would introduce a preferred direction in the sky, hence breaking the statistical isotropy of the temperature field. Then, the diagonal part of the covariance matrix, $C_\ell^{TT}$ contains all the statistical information of the CMB field. Even if we are allowed to observe only one realization of the stochastic field, we can construct a power spectrum estimator from a single realization of the sky by replacing the ensemble average in Eq. (2.68) with the spatial (sample) average, i.e. averaging over the $2\ell + 1$ $m$-samples[22]. The commonly used power spectrum estimator reads as (see Sec. (3.2))

$$\hat{C}_\ell^{TT} = \frac{1}{2\ell + 1} \sum_m |a_{\ell m}^T|^2, \tag{2.71}$$

which is unbiased, i.e. $\langle \hat{C}_\ell^{TT} \rangle = C_\ell^{TT}$, and described by a chi-square distribution with $2\ell + 1$ d.o.f., while the diagonal covariance is given by

$$(\Delta \hat{C}_\ell)^2 = \frac{2}{2\ell + 1} C_\ell^2. \tag{2.72}$$

This result is known as *cosmic variance*. The harmonic space formalism can easily be linked to the configuration space 2-point correlation function $C^{TT}(\theta) \equiv \langle T(\hat{\mathbf{n}})T(\hat{\mathbf{n}}') \rangle$ as[23]

$$
\begin{aligned}
C^{TT}(\theta) &= \sum_{\ell m} \sum_{\ell' m'} \underbrace{\langle a_{\ell m}^T a_{\ell' m'}^{T*} \rangle}_{\delta_{\ell\ell'}^K \delta_{mm'}^K C_\ell^{TT}} Y_{\ell m}(\hat{\mathbf{n}}) Y_{\ell' m'}^*(\hat{\mathbf{n}}') \\
&= \sum_\ell C_\ell^{TT} \underbrace{\sum_m Y_{\ell m}(\hat{\mathbf{n}}) Y_{\ell m}^*(\hat{\mathbf{n}})}_{\frac{2\ell+1}{4\pi} P_\ell(\hat{\mathbf{n}} \cdot \hat{\mathbf{n}}')} \\
&= \sum_\ell \frac{2\ell + 1}{4\pi} C_\ell^{TT} P_\ell(\hat{\mathbf{n}} \cdot \hat{\mathbf{n}}'),
\end{aligned}
\tag{2.73}
$$

where $\hat{\mathbf{n}} \cdot \hat{\mathbf{n}}' = \cos\theta$ and we used the addition theorem for spherical harmonics to express the product of $Y_{\ell m}$ in terms of the Legendre polynomials $P_\ell(x)$.

We now seek a theoretical framework that allows us to compute the observables given a set of cosmological parameters. The CMB temperature angular spectrum can be computed via the LOS integration method (see Seljak & Zaldarriaga (1996) for an extensive discussion) as

$$C_\ell^{TT} = 4\pi \int \mathrm{d}\log k \, \mathcal{P}_\mathcal{R}(k) |\Theta_\ell^T(k, \eta_0)|^2. \tag{2.74}$$

---

[22]It is possible to prove that a random field is *ergodic* if it can be described by a Gaussian statistics and if its power spectrum is continuous (Adler & Taylor, 2007).

[23]Sometimes labelled as $w(\theta)$, especially in clustering analysis.

This method elegantly connects the physics at recombination with the projection effects by writing the photon transfer functions $\Theta_\ell^T(k, \eta)$ as a time integral of a set of source functions $S_T(k, \eta)$ multiplied by the spherical Bessel functions $j_\ell(x)$:

$$\Theta_\ell^T(k, \eta_0) = \int_0^{\eta_0} d\eta S_T(k, \eta) j_\ell[k(\eta_0 - \eta)]. \tag{2.75}$$

The temperature source term can be split as[24] ([Lesgourgues & Tram, 2014](#))

$$S_T(k, \eta) = g\left(\frac{1}{4}\delta_\gamma + \Phi\right) + e^{-\tau}(\dot{\Phi} + \dot{\Psi}) + \frac{g}{k}\theta_b, \tag{2.76}$$

where $\tau(\eta)$ is the optical depth $\tau = \int$ for a photon emitted at time $\eta$ and received at $\eta_0$ and $g(\tau) = -\dot{\tau}e^{-\tau}$ is the visibility function which peaks around the last scattering epoch at $\eta = \eta_*$. The first term in Eq. (2.76) contains the intrinsic temperature fluctuations $\frac{1}{4}\delta_\gamma$ and the gravitational redshift term $\Phi$ responsible for the Sachs-Wolfe effect[25]. The second term in $S_T(k, \eta)$ is sourced by the time evolution of the gravitational potentials and gives rise to the so-called integrated Sachs-Wolfe effect (iSW) effect that, in the case of standard GR where the two potentials are equal, reduces to $S_T^{\text{iSW}} \to -2e^{-\tau}\dot{\Phi}$ (see Sec. (2.2.2)). Doppler effects induced by the peculiar motion of the photo-baryonic fluid are given by the $\theta_b$ term.

By looking at the temperature power spectrum (normalized as $\mathcal{D}_\ell^{TT} = \frac{\ell(\ell+1)}{2\pi}C_\ell^{TT}$) shown in Fig. (2.3), we can appreciate a rich structure and distinguish three main features sourced by few physical effects:

1. the *Sachs-Wolfe plateau* at $\ell \lesssim 100$, that almost faithfully traces the initial conditions (which have not had enough time to evolve on such big scales);

2. a prominent peak around $\ell \sim 100$ followed by a series of *acoustic peaks*, produced by modes of the density perturbations which have experienced one or more acoustic oscillations (driven by the competing action of gravity and pressure) before last scattering;

3. an overall damping of the spectrum for $\ell \gtrsim 1000$ caused by photon diffusion at recombination (*Silk damping*) and the fact that these small-scale fluctuations have a size comparable to the width of the last scattering surface (it is rather a *layer* than a surface).

**Polarization power spectra**

CMB photons can be, and in fact are, polarized (at the 10% level). CMB polariza-

---

[24]Neglecting the contribution from polarization.
[25]The Sachs-Wolfe term is sometimes considered to be the sum $\frac{1}{4}\delta_\gamma$ and $\Phi$.

tion represents an invaluable source of information on inflation and on the late-time evolution of the Universe, like reionization and structure formation (through the lensing effects).

Light polarization is commonly described in terms of the *Stokes parameters $I, Q, U$*, and $V$. As we have seen, the CMB total intensity $I$ has a blackbody spectrum along the LOS, while the parameters $Q$ and $U$ describe the linear polarization and $V$ describes the circular polarization (which is expected to vanish for the CMB). Note that the temperature $T \equiv I/4$ is invariant under any rotation in the plane orthogonal to the direction of propagation $\hat{\mathbf{n}}$, while $Q$ and $U$ are not. In particular, we have that under a rotation of an angle $\psi$ they transform as

$$Q' = Q \cos 2\psi + U \sin 2\psi, \tag{2.77}$$

$$U' = Q \cos 2\psi - U \sin 2\psi, \tag{2.78}$$

i.e. they rotate as a spin-2 quantity. The problem with polarization is that $Q$ and $U$ depend on the coordinate system: while such coordinate system is well defined in the flat-sky limit, it becomes ambiguous on the whole sphere (because it is not possible to find a rotationally invariant orthogonal basis on the sphere). Zaldarriaga (1997) and Kamionkowski et al. (1997) have developed an all-sky formalism to describe correlation function of CMB polarization. $Q$ and $U$ Stokes parameters can be combined to yield a pair of spin $s = 2$ quantities as

$$(Q \pm iU)'(\hat{\mathbf{n}}) = e^{\mp 2i\psi}(Q \pm iU)(\hat{\mathbf{n}}), \tag{2.79}$$

which can be decomposed on the spin-weighted spherical harmonics ${}_sY_{\ell m}(\hat{\mathbf{n}})$ as follows

$$(Q \pm iU)(\hat{\mathbf{n}}) = \sum_{\ell m} {}_{\pm 2}a_{\ell m} {}_{\pm 2}Y_{\ell m}(\hat{\mathbf{n}}) = \sum_{\ell m} (a^E_{\ell m} \pm i a^B_{\ell m}) {}_{\pm 2}Y_{\ell m}(\hat{\mathbf{n}}), \tag{2.80}$$

where we have defined the $E$- and $B$-modes as

$$a^E_{\ell m} = -\frac{1}{2}({}_{+2}a_{\ell m} + {}_{-2}a_{\ell m}) \quad \text{and} \quad a^B_{\ell m} = \frac{i}{2}({}_{+2}a_{\ell m} - {}_{-2}a_{\ell m}). \tag{2.81}$$

$E$-modes are invariant with respect to parity transformations ($\hat{\mathbf{n}} \to -\hat{\mathbf{n}}$) while $B$-modes flip sign (similarly to the gradient and curl component of the electromagnetic field). Because of the opposite parity, the cross-correlations between $E$ and $B$ or $T$

and $B$ is vanishing. Then, the non-null power spectra are given by

$$\langle a^E_{\ell m} a^{E*}_{\ell' m'} \rangle = \delta^K_{\ell\ell'} \delta^K_{mm'} C^{EE}_\ell, \tag{2.82}$$

$$\langle a^B_{\ell m} a^{B*}_{\ell' m'} \rangle = \delta^K_{\ell\ell'} \delta^K_{mm'} C^{BB}_\ell, \tag{2.83}$$

$$\langle a^T_{\ell m} a^{E*}_{\ell' m'} \rangle = \delta^K_{\ell\ell'} \delta^K_{mm'} C^{TE}_\ell. \tag{2.84}$$

We now turn to the explanation of the polarizing mechanism for CMB photons. Recall that before recombination the photons interact with free electrons through Compton scattering (Thomson scattering in the low-energy limit) which does not induce polarization *unless* the incident radiation scattering off the electron is anisotropically distributed. The cross-section of the process can be written as

$$\frac{d\sigma}{d\Omega} = \frac{3\sigma_T}{8\pi} |\hat{\boldsymbol{\epsilon}}' \cdot \hat{\boldsymbol{\epsilon}}|. \tag{2.85}$$

where $\hat{\boldsymbol{\epsilon}}' = (\hat{\epsilon}'_x, \hat{\epsilon}'_y)$ and $\hat{\boldsymbol{\epsilon}} = (\hat{\epsilon}_x, \hat{\epsilon}_y)$ are the polarization vectors of the incoming and scattered waves respectively, defined in the plane orthogonal to the direction $\hat{z}$ of the wave propagation. After the scattering, the $\hat{z}$-direction is deflected by an angle $\theta$ defined in the plane that contains both the incident and diffused waves. With this geometrical configuration, it is possible to show that for an incoming unpolarized $(Q, U, V = 0)$ radiation of intensity $I'(\theta, \phi)$, the net polarization produced by the scattering is given by (Kosowsky, 1996):

$$I = \frac{3\sigma_T}{16\pi} \int d\Omega\, I'(\theta, \phi)(1 + \cos^2 \theta) \tag{2.86}$$

$$Q = \frac{3\sigma_T}{16\pi} \int d\Omega\, I'(\theta, \phi) \sin^2 \theta \cos(2\phi) \tag{2.87}$$

$$U = -\frac{3\sigma_T}{16\pi} \int d\Omega\, I'(\theta, \phi) \sin^2 \theta \sin(2\phi). \tag{2.88}$$

Decomposing $I'(\theta, \phi)$ on the spherical harmonics basis as $I'(\theta, \phi) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\theta, \phi)$, one obtains

$$I = \frac{3\sigma_T}{16\pi} \left( \frac{8}{3} \sqrt{\pi} a_{00} + \frac{4}{3} \sqrt{\frac{\pi}{5}} a_{20} \right) \tag{2.89}$$

$$Q - iU = \frac{3\sigma_T}{4\pi} \sqrt{\frac{2\pi}{15}} a_{22}, \tag{2.90}$$

which shows that the presence of a quadrupole term in the distribution of the incoming radiation intensity can induce a linear polarization state of the scattered photon.

We can distinguish two main sources of polarization fluctuations at the time of last-scattering: the density perturbations and the gravitational waves, of scalar
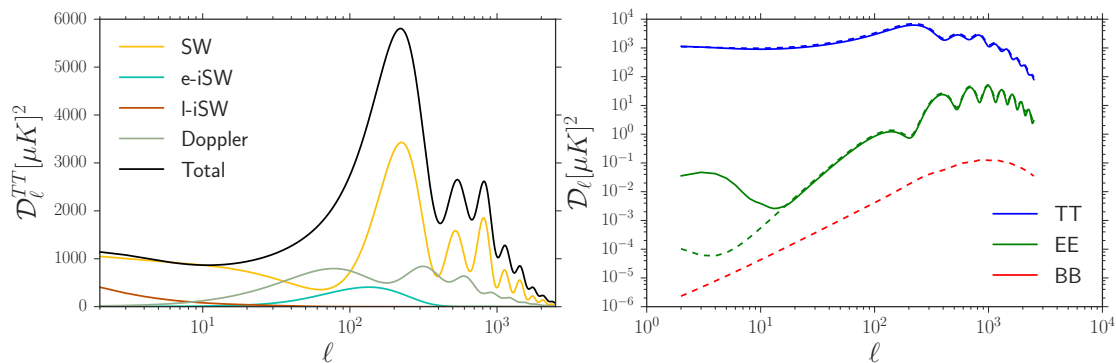
**Figure 2.3:** *Left panel:* CMB temperature angular power decomposed into the scalar contributions. *Right panel:* CMB polarization power spectra with reionization at $z_{\mathrm{re}} = 8$ (solid lines) and without (dashed lines). No primordial tensorial power ($r = 0$) is assumed to predict the theoretical lines.

and tensorial nature respectively. In the former case, a quadrupole is produced as photons start to free-stream so that the electrons see an anisotropic incoming radiation while falling in and out of potential wells; thus, the quadrupole amplitude is proportional to the gradient of the photon fluid. Primordial gravitational waves, introduced in the context of inflation in Sec. (2.1.2), are a prediction og inflation and correspond to traceless and transverse spatial metric perturbations: as the waves propagate through the photo-baryonic fluid, they create ripples in space-time. The associated stretching of photons wavelengths originates a quadrupole which can be converted into polarization via Thomson scattering. The angular power spectra can be evaluated with the same approach of Eq. (2.74) and are shown in the right panel of Fig. (2.3). The acoustic oscillations are activated by the same processes in total intensity through the quadrupole for the *E*-modes. As concerns the *B*-modes, gravitational waves at recombination are able to determine a peak at the degree scale because they rapidly diffuse when entering the horizon since they are not supported by pressure.

## 2.2.2 Secondary anisotropies

After decoupling the mean free path of photons becomes large enough such that they can be considered to freely propagate in the intergalactic medium. In reality, this picture is distorted by interactions of the CMB photons with matter on their way from the last scattering surface to the observer. We refer to these interactions as *secondary anisotropies*, since they are not originated in the primordial Universe, and usually do not include astrophysical foregrounds. For a comprehensive discussion on CMB secondary anisotropies, we refer the reader to the review by Aghanim et al. (2008), while here we introduce the main ones. The weak gravitational lensing

of the CMB is sourced by the intervening LSS and as such can be considered a secondary anisotropy: since this topic is at the core of this thesis, it will be extensively discussed in Sec. (2.3.2).

### Integrated Sachs-Wolfe Effect

This effect is due to CMB photons crossing a time-varying linear gravitational potential. The photons will gain energy if the depth of the potential is shallower by the time they exit, and lose energy otherwise. The change in temperature due to gravitational redshifting is achromatic, i.e. frequency *independent* (contrary to the case of Sunyaev-Zel'dovich effect (SZ) effect), thus it cannot be isolated from primary anisotropies using only spectral information. We can distinguish two types of iSW: the early-iSW, sourced by the damping that potentials feel when entering the horizon during the radiation-matter transition, and the late-iSW, associated to the decay of potentials at the time the cosmic acceleration kicks in. The effect at the power spectrum level is to enhance the fluctuations at low $\ell$, i.e. large angular scales. Assuming a flat geometry, the detection of the iSW effect would hint to the presence of DE (Crittenden & Turok, 1996), thus it represents one of the main probes of DE and in general of the gravity sector, although it is severely affected by cosmic variance. The basic idea of (late-)iSW measurements is to cross-correlate, either reconstructing the 2-point function in harmonic/real domain or with stacking techniques, CMB temperature maps with external LSS tracers of the large-scale potentials.

### Rees-Sciama effect

The iSW effect concerns the linear evolution of potentials. However, CMB photons may happen to fly by structures going through nonlinear evolution: if the crossing time is a non-negligible fraction of the collapsing time-scale, they can experience the gravitational redshifting. This *nonlinear* iSW signal goes by the name of Rees-Sciama effect (RS) (Rees & Sciama, 1968). In the standard structure formation scenario, nonlinearities tend to appear earlier on smaller scales, so this effect should show up at high-$\ell$; nonetheless, the overall amplitude is very small, $\frac{\Delta T}{T} \sim 10^{-6} - 10^{-7}$.

### Sunyaev-Zel'dovich effect

Perhaps the most studied secondary CMB anisotropy, the SZ effect is generally originated by inverse Compton scattering of CMB photons off high-energy free electrons in the hot intracluster gas (Zeldovich & Sunyaev, 1969; Sunyaev & Zeldovich, 1970). This effect can be broadly split into (i) the *thermal* SZ (tSZ),

due to the scattering of CMB photons by the random thermal motion of electrons, and (ii) the *kinetic* SZ (kSZ), sourced by the bulk motion of electrons (that we will thoroughly discuss in Ch. (7)). The CMB temperature change caused by tSZ is proportional to the *gas pressure* $p_e$ in the galaxy cluster integrated over the LOS and can be written as

$$\frac{\Delta T_{\text{tSZ}}}{T_{\text{CMB}}} = f(x)y \quad \text{with} \quad x \equiv h\nu/k_B T_{\text{CMB}} = \nu/(56.78)\,\text{GHz}. \tag{2.91}$$

Here, the frequency independent part is called the *Compton-y parameter*

$$y \equiv \frac{\sigma_T}{m_e c^2} \int p_e \, \mathrm{d}\ell, \tag{2.92}$$

while $f(x)$, in the non-relativistic case, reads as

$$f_{\text{NR}}(x) = x\frac{e^x + 1}{e^x - 1}. \tag{2.93}$$

The tSZ does not induce any CMB temperature change around $\nu \sim 217$ GHz, which is called the SZ null, while below and above this frequency the CMB intensity is decreased and increased respectively. Thus, by looking at CMB temperature maps at different frequencies above and below the SZ null we can detect galaxy clusters up to high-$z$ (Planck Collaboration, 2015e); a key feature of this signal is that is redshift independent. On the other hand, the kSZ induced Doppler boost of CMB photons is sensitive to the *gas momentum* integrated along the LOS and the temperature change, in the non-relativistic regime, reads as

$$\frac{\Delta T_{\text{kSZ}}}{T_{\text{CMB}}} = -\tau_e \frac{v_e}{c}, \tag{2.94}$$

where $\tau_e$ is the optical depth of the cluster. Differently from tSZ, the SED of the kSZ is Planckian, thus it is impossible to disentangle it from CMB with multiband measurements; moreover, the amplitude of this signal is at least an order of magnitude smaller than the tSZ one. kSZ effect has been detected statistically (Hand et al., 2012) and, recently, the detection at the single cluster level has been claimed (Adam et al., 2016).

**Reionization**

After recombination and decoupling the Universe goes from being fully ionized to neutral, entering the so-called *dark ages*. However, the birth of the first astrophysical objects such as stars at later times (around $6 \lesssim z \lesssim 15$, though the details are stil debated), ionizes the surrounding medium (Gunn & Peterson, 1965). The

*reionization* process is not yet fully understood but it is expected to be a *patchy* process: the sources emitting energetic photons form ionized bubbles around them which eventually joint. Since this process happens at times when the free electron density is sufficiently small, the mean free path of CMB photons is slightly affected; in particular, CMB photons are re-scattered by these free electrons and the overall effect is an optical depth dependent suppression of temperature anisotropies, $\frac{\Delta T}{T}(\hat{\mathbf{n}}) \to \frac{\Delta T}{T}(\hat{\mathbf{n}})e^{-\tau(\hat{\mathbf{n}})}$. At the power spectrum level, this translates to a $e^{-2\tau}$ suppression so that the amplitude of the observed spectra is proportional to $A_{\rm s}e^{-2\tau}$, introducing a degeneracy between the primordial power amplitude and the optical depth which can be broken with lensing or polarization information. The effect of reionization on polarization anisotropies is to enhance the power at low-$\ell$, generating a bump in the $C_\ell^{EE}$ and $C_\ell^{TE}$ power spectra.

### 2.2.3 Cross-correlations: combining probes with CMB

As we have seen, by being sensitive to a variety of perturbations over different epochs of the cosmic timeline, the CMB sky contains a wealth of information about fundamental physics, the origin and the evolution of the Universe. While primordial anisotropies give us a snapshot of the physics at recombination, the different secondary effects probe a wide range of environments, from decaying potentials to intracluster medium, along the light-cone. How can we make the best of CMB observations and extract all the relevant information? In the context of precision cosmology, cross-correlation analyses between independent probes have become a useful tool to infer further information about the Universe while providing extra safety checks on known (and unknown) systematics affecting the datasets. This last part can be schematically understood as follows: if we have two independent observables $\{X, Y\}$ their estimates will constitute a signal and a noise term as $\hat{X} = S^X + N^X$ (same for $Y$). Then, the 2-point correlator is given by

$$
\begin{aligned}
\langle \hat{X}\hat{Y} \rangle &= \langle (S^X + N^X)(S^Y + N^Y) \rangle \\
&= \langle S^X S^Y \rangle + \cancel{\langle N^X S^Y \rangle} + \cancel{\langle S^X N^Y \rangle} + \cancel{\langle N^X N^Y \rangle},
\end{aligned}
\tag{2.95}
$$

where we used the fact that, for *uncorrelated* noise, the signal-noise and noise-noise cross terms are vanishing.[26] So, probes combination allows to extract signals hidden in noisy data, providing us a clean and unbiased measurement of the cross-correlation signal which is non-null if the observed fields are correlated, e.g. because

---

[26]Note that the notion of signal and noise is not uniquely defined and depends on the context of the study. As an example, Galactic foregrounds contaminating the CMB are considered to be a nuisance from a cosmological perspective, though they encode a lot of information for a galactic astrophysicist!

of common physical mechanisms that source the signal. Note that, depending on the fields being cross-correlated, the information content can be both cosmological and astrophysical (as the case of the CMB lensing-galaxy density cross-correlation which is the main topic of the present thesis), helping to break the degeneracies affecting the parameters. In particular, CMB secondary anisotropies are expected to be correlated with a variety of external LSS probes like the galaxy distribution, the cosmic shear, the Cosmic Infrared Background (CIB), the $\gamma$-ray background anisotropies and HI intensity data. The reason is that all observables are sensitive in different ways to same underlying fields, i.e. either the distribution of matter or gravitational potentials, and it is exactly the different sensitivity that allows cross-correlation measurements to expand the range of investigation in a safer way. CMB temperature, polarization and lensing data have been cross-correlated with numerous external datasets for different scientific exploitations, here we focus on the CMB lensing, in particular on its cross-correlation with the spatial galaxy distribution.

**CMB lensing vs. LSS**

As we will see in Sec. (2.3.2), the LSS leaves an imprint on CMB anisotropies by gravitationally deflecting CMB photons during their journey from the last-scattering surface to us (Lewis & Challinor, 2006). The net effect is a remapping of the CMB observables, dependent on the gravitational potential integrated along the LOS, and it is sensitive to both the geometry of the Universe and to the growth of the LSS. On the other hand, since CMB lensing is an integrated quantity, it does not provide direct information on the *evolution* of the large scale gravitational potential. However, in the standard structure formation scenario galaxies reside in DM halos, the most massive of which are the signposts of larger scale structures that act as lenses for CMB photons. Then, it is clear that the cross-correlation between CMB lensing maps and tracers of LSS $\delta_g$ enables the reconstruction of the dynamics and of the spatial distribution of the gravitational potential, providing simultaneous constraints on cosmological and astrophysical parameters (Pearson & Zahn, 2014), such as the bias factor $b$ relating fluctuations in luminous and dark matter (Eq. (2.53)), that can be exploited to derive the effective halo masses associated with the tracer populations. This can tighten tests of the time evolution of DM density fluctuations and through that, give constraints on the dynamics of the DE at the onset of cosmic acceleration as well as on neutrino masses. Although the bias factors can also be well determined from the autopower spectra, we must always beware of unaccounted systematic effects. These cross-correlation measurements are not prone to systematics that are not correlated between the

two data sets, thus they are complementary to galaxy clustering measurements and a comparison of the bias estimates from auto- and cross-correlations can uncover unforeseen systematics, both at the data and theory level, on either side. We refer to Sec. (4.2) for the mathematical modeling of the expected cross-correlation signal.

Let us note here that the peculiarity of the sources exploited in the analyses presented in this thesis (see Ch. (4) and (5)) is that their sub-mm flux density remains approximately constant with increasing redshift for $z \gtrsim 1$ (strongly negative K-correction), so that sub-mm surveys have the power of piercing the distant Universe up to $z \gtrsim 3$–4 where the CMB lensing is most sensitive to matter fluctuations. In contrast, the available large–area optical/near-infrared galaxy surveys and radio source surveys reach redshifts only slightly above unity and therefore pick up only a minor fraction of the CMB lensing signal whose contribution peaks at $z > 1$ and is substantial up to much higher redshifts. Quasars allow us to extend investigations much further, but are rare and therefore provide a sparse sampling of the large-scale gravitational field.

CMB lensing itself was first detected by means of cross-correlation analysis (Smith et al., 2007; Hirata et al., 2008). Since the first detections, several analyses have been carried out with the distribution of tracers observed at different wavelengths, including optically-selected sources (Sherwin et al., 2012; Bleem et al., 2012; Planck Collaboration, 2014c; Giannantonio & Percival, 2014; Omori & Holder, 2015; Kuntz, 2015; Giannantonio et al., 2016; Pullen et al., 2016; Baxter et al., 2016; Singh et al., 2016), sub-mm selected sources (Bianchini et al., 2015, 2016b), radio-selected sources (Feng et al., 2012; Planck Collaboration, 2014c; Allison et al., 2015), Infrared (IR)-selected sources (Bleem et al., 2012; Geach et al., 2013; DiPompeo et al., 2014; Giannantonio & Percival, 2014; DiPompeo et al., 2016), X-ray-selected galaxy clusters (Planck Collaboration, 2014c). In Table (2.2) we provide a list of the reported cross-correlations studies between CMB lensing and the galaxy density field, with a summary of the analyses features. Even though the full potential of these measurements is yet to be unravelled, cross-correlation analyses have already been exploited to measure the redshift evolution of the galaxy bias (see Bianchini et al. (2015); Kuntz (2015); Bianchini et al. (2016b); DiPompeo et al. (2016)), to test GR by studying the growth of structure as done by Giannantonio et al. (2016) or measuring the relation between curvature fluctuations and velocity perturbations with the $E_g$ statistics (see Pullen et al. (2016)), to improve the calibration of the cosmic shear data together with CMB lensing - galaxy lensing data (Baxter et al., 2016), and to study primordial-NG (Giannantonio & Percival, 2014). It is worth to stress that *all* the mentioned analyses have reconstructed the 2-point statistics, either in real ($w^{\kappa g}(\theta)$) or harmonic space ($C_\ell^{\kappa g}$): a step towards new statistical

domain for cross-correlation analyses will be taken in Ch. (6), where we develop a needlet-based spectral estimator.

As pointed out by Song et al. (2003), the CMB lensing kernel is well-matched with the one of the unresolved dusty galaxies comprising the CIB since both are tracers of the large-scale density fluctuations in the Universe. In particular, Planck measurements suggest that the correlation between the CMB lensing map and the CIB map at 545 GHz can be as high as 80% (Planck Collaboration, 2014d). Other statistically significant detections have been recently reported by Holder et al. (2013); Hanson et al. (2013); Ade et al. (2014a); van Engelen et al. (2015). Even though there are connections between these studies and those presented here, one needs to bear in mind that, differently from galaxy catalogs, the CIB is an integrated quantity and as such it prevents a detailed investigation of the temporal evolution of the signal. Moreover, the interpretation of the measured cross-correlation is actually more challenging since the precise redshift distribution of the sources contributing to the sub-mm background is still debated. For the sake of completeness, we recall that CIB-galaxy density cross-correlation has been measured to date, see Serra et al. (2014); Schmidt et al. (2014); these studies can be useful to infer the redshift distribution of sources contributing to the CIB and the star formation history.

Recently, CMB lensing has also been combined with galaxy shear data (Hand et al., 2015; Liu & Hill, 2015; Kirk et al., 2016; Harnois-Déraps et al., 2016), demonstrating its complementarity to galaxy shear measurements. They both directly probe the underlying matter field, being insensitive to the galaxy bias and potentially providing constraints on a combination of the mean matter density $\Omega_m$ and the amplitude of matter fluctuations $\sigma_8$. Forecasts analyses have shown (Vallinotto, 2012; Das et al., 2013; Schaan et al., 2016a) that by combining CMB lensing, galaxy clustering and cosmic shear data, it is in principle possible to calibrate shear multiplicative bias, as well as other galaxy related parameters like photo-$z$ errors, while still providing stringent constraints on cosmological parameters. Other cross-correlations involving CMB lensing have been measured in combination with the $\gamma$-ray sky (Fornengo et al., 2015), which represents an high-$z$ leverage to quantify the impact of astrophysical sources to the $\langle \gamma \delta_g \rangle$ correlation for annihilating DM-related studies, and with tSZ data of Compton-$y$ parameter, $\langle \kappa_{CMB} y \rangle$ (Hill & Spergel, 2013; Van Waerbeke et al., 2014).

| Authors | $\kappa_{CMB}$ **data** | $\delta_g$ **data** | **Redshift** | $f_{sky}$ |
|---------|------------------------|---------------------|--------------|-----------|
| Smith et al. (2007) | WMAP3 | NVSS (RGs) | $\sim 1.1$ | 0.58 |
| Hirata et al. (2008) | WMAP3 | SDSS (LRGs) | $0.2 < z < 0.7$ | 0.16 |
| | | SDSS (QSOs) | $z < 2.7$ | 0.15 |
| | | NVSS (RGs) | $\sim 1.0$ | |
| Bleem et al. (2012) | SPT | Spitzer | | |
| | | BCS | $z_{med} \sim 0.5$ | 0.001 |
| | | WISE | | |
| Feng et al. (2012) | WMAP1-7 | NVSS (RGs) | $z_{med} \sim 1.1$ | 0.58 |
| Sherwin et al. (2012) | ACT | SDSS-XDQSO | $z_{med} \sim 1.4$ | 0.008 |
| Planck Team 2013 | Planck 2013 | NVSS (RGs) | $z_{med} \sim 1.1$ | 0.7 |
| | | MaxBCG (GCs) | $0.1 < z < 0.3$ | 0.2 |
| | | SDSS (LRGs) | $z_{med} \sim 0.55$ | 0.25 |
| | | WISE | $z_{med} \sim 0.18$ | 0.6 |
| Geach et al. (2013) | SPT/Planck2013 | WISE (QSOs) | $z_{med} \sim 1.1$ | 0.06 |
| DiPompeo et al. (2014) | Planck 2013 | WISE (QSOs) | $z_{med} \sim 1$ | 0.08 |
| Giannantonio & Percival (2014) | Planck 2013 | NVSS (RGs) | $z_{med} \sim 1.1$ | 0.7 |
| | | 2MASS (IR) | $z_{med} \sim 0.1$ | 0.2 |
| | | SDSS-CMASS | $\bar{z} \sim 0.55$ | 0.25 |
| Bianchini et al. (2015) | Planck 2013 | H-ATLAS (FIR) | $z \geq 1.5$ | 0.01 |
| Omori & Holder (2015) | Planck 2013/2015 | CHFTLens (Optical) | $0.2 < z < 1.2$ | 0.004 |
| Kuntz (2015) | Planck 2013/2015 | CHFTLens (Optical) | $0.2 < z < 1.2$ | 0.004 |
| Allison et al. (2015) | ACT+ACTPol | FIRST (RGs) | $z_{med} \simeq 1.5$ | $\lesssim 0.01$ |
| Giannantonio et al. (2016) | Planck 2015/SPT | DES-SV (Optical) | $0.2 < z < 1.2$ (Tomo) | 0.004 |
| Bianchini et al. (2016b) | Planck 2015 | H-ATLAS (FIR) | $z \geq 1.5$ (Tomo) | 0.01 |
| Pullen et al. (2016) | Planck 2015 | SDSS-CMASS | $z_{med} = 0.57$ | 0.2 |
| Baxter et al. (2016) | SPT-SZ | DES-SV | $0.4 < z < 0.8$ | 0.03 |
| Singh et al. (2016) | Planck 2015 | SDSS-CMASS/LOWZ | $z_{med} = 0.57/0.30$ | 0.2 |

**Table 2.2:** Current observational status of the measurement of the CMB lensing - galaxy density cross-power.

## 2.3 Weak Gravitational Lensing

Gravitational lensing is the deflection of light bundles as they propagate through a gravitational field. This effect is a spectacular consequence of Einstein's GR, although it can also be interpreted within the Newtonian theory of gravity. The basic idea in GR is that photons propagate along geodesics, i.e. the shortest path between two events. While in vacuum the geodesics are simply straight lines, the presence of matter alters the spacetime fabric and, as a consequence, geodesics becomes slightly bent.

Since the Universe today is rather empty, most of the photons only feel weak gravitational potentials and hence, their trajectories experience tiny deflections even over cosmological scales (as the case of CMB photons that we will consider later). However, on small scales, these trajectories happen to pass close to massive structures (like a star, a galaxy or a galaxy cluster) and feel strong gravitational

fields that result in more dramatic deflections. For a comprehensive overview of gravitational lensing, see the review by Bartelmann & Schneider (2001).

### 2.3.1 The lens equation

To illustrate the lensing basics we follow Shirasaki (2016) and consider the trajectory of light from a distant source object in the presence of inhomogeneous matter distribution. The path of the photon is determined by the null geodesic equation,

$$\frac{\mathrm{d}^2 x^\mu}{\mathrm{d}\lambda^2} = -\Gamma^\mu_{\alpha\beta} \frac{\mathrm{d}x^\alpha}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\beta}{\mathrm{d}\lambda}, \tag{2.96}$$

$$g_{\mu\nu} \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda} \frac{\mathrm{d}x^\nu}{\mathrm{d}\lambda} = 0, \tag{2.97}$$

where $\Gamma^\mu_{\alpha\beta}$ denotes the Christoffel symbol evaluated from the metric $g_{\mu\nu}$ as

$$\Gamma^\mu_{\alpha\beta} = \frac{g_{\mu\nu}}{2}(g_{\mu\alpha,\beta} + g_{\mu\beta,\alpha} - g_{\alpha\beta,\mu}), \tag{2.98}$$

with $\partial g_{\mu\alpha}/\partial x^\beta \equiv g_{\mu\alpha,\beta}$. If the gravitational potential $\Phi$ associated to the lens is small enough ($|\Phi| \ll c^2$), the metric that describes an inhomogeneous expanding Universe can be written as follows:

$$\mathrm{d}s^2 = -\Big(1 + \frac{2\Phi}{c^2}\Big)c^2\,\mathrm{d}t^2 + a^2(t)\Big(1 - \frac{2\Phi}{c^2}\Big)\Big[\mathrm{d}\chi^2 + f_K^2(\chi)\,\mathrm{d}\Omega^2\Big]. \tag{2.99}$$

Considering small deflection angles, we can approximate $\mathrm{d}\Omega^2 \approx (\mathrm{d}\theta^1)^2 + (\mathrm{d}\theta^2)^2$ (this is the so called *flat-sky approximation*). Now suppose that the spatial coordinates of the incoming light ray are $x^i = (\theta^1, \theta^2, \chi)$, then the derivative with respect to the affine parameter $\lambda$ reads as

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} = \frac{\mathrm{d}\chi}{\mathrm{d}\lambda} \frac{\mathrm{d}}{\mathrm{d}\chi} \tag{2.100}$$

$$= \frac{\mathrm{d}\chi}{\mathrm{d}x^0} \frac{\mathrm{d}x^0}{\mathrm{d}\lambda} \frac{\mathrm{d}}{\mathrm{d}\chi} \tag{2.101}$$

$$= -\frac{P^0}{a} \frac{\mathrm{d}}{\mathrm{d}\chi}, \tag{2.102}$$

where $P^0 = \mathrm{d}x^0/\mathrm{d}\lambda$. As a result, the transverse components ($i = 1, 2$) of the null geodesic equation can be expressed as a differential equation with respect to the comoving distance $\chi$. By Taylor expanding about $\theta^1$, $\theta^2$, and $\Phi/c^2$ up to first order

in Eq. (2.96), we can obtain the following expression

$$\frac{\mathrm{d}^2(f_K\theta^i)}{\mathrm{d}\chi^2} + K f_K\theta^i = -\frac{2}{c^2}\frac{\partial\Phi}{\partial(f_K\theta^i)}. \tag{2.103}$$

The solution of Eq. (2.103) is given by the so called *lens equation*

$$\beta^i = \theta^i - \frac{2}{c^2}\int_0^\chi \mathrm{d}\chi'\partial_i\Phi(\chi')\frac{f_K(\chi-\chi')}{f_K(\chi)}, \tag{2.104}$$

where $\beta^i$ represents the *unlensed* position of the source on the sky and $\partial_i$ denotes the derivative with respect to $f_K(\chi)\theta^i$. For a fixed source position $\boldsymbol{\beta}$, Eq. (2.104) is a nonlinear equation for the image position $\boldsymbol{\theta}$ that, in certain cases, allows more than one solution, i.e. the source is multiply imaged. The local properties of the lens mapping are given by the Jacobian matrix $\mathbf{A}$

$$A_{ij} \equiv \frac{\partial\beta^i}{\partial\theta^j} = \begin{pmatrix} 1-\kappa-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix} \tag{2.105}$$

$$= \begin{pmatrix} 1-\kappa & 0 \\ 0 & 1-\kappa \end{pmatrix} + \begin{pmatrix} -\gamma_1 & -\gamma_2 \\ -\gamma_2 & \gamma_1 \end{pmatrix} \tag{2.106}$$

where we have introduced the components of the *shear* $\gamma \equiv \gamma_1 + i\gamma_2 = |\gamma|e^{2i\phi}$, and the *convergence* $\kappa$: the former describes an anisotropic stretching of the shape of the image while the latter represents a magnification of its size. In the presence of the both components, a unitary radius circular source is distorted into an elliptical image with major and minor axis given by $(1-\kappa\pm|\gamma|)^{-1}$, oriented along the angle $\phi$, and amplified by a factor $\mu = (\det\mathbf{A})^{-1} = [(1-\kappa)^2 - \gamma^2]^{-1}$. Eq. (2.104) enables us to link the components of $\mathbf{A}$ to the second order derivative of the gravitational potential as follows:

$$A_{ij} = \delta_{ij}^K - \Phi_{ij}, \tag{2.107}$$

$$\Phi_{ij} = \frac{2}{c^2}\int_0^\chi \mathrm{d}\chi' G(\chi,\chi')\partial_i\partial_j\Phi(\chi'), \tag{2.108}$$

$$G(\chi,\chi') = \frac{f_K(\chi-\chi')f_K(\chi')}{f_K(\chi)}, \tag{2.109}$$

from which we find that $2\kappa = \Phi_{11} + \Phi_{22}$. Since the gravitational potential is related to the matter density contrast through Poisson equation (Eq. (2.42)) as

$$\nabla^2\Phi = \frac{3H_0^2}{2}\Omega_{\mathrm{m0}}\frac{\delta}{a}, \tag{2.110}$$

the convergence in a given direction $\hat{\mathbf{n}}$ can be written as a weighted integral of $\delta$ along the LOS:

$$\kappa(\hat{\mathbf{n}}, \chi) = \int_0^\chi \mathrm{d}\chi' \, W^\kappa(\chi')\delta(\chi'). \tag{2.111}$$

The lensing kernel that weights the density contrast reads

$$W^\kappa(\chi) = \frac{3H_0^2 \Omega_{\mathrm{m0}}}{2c^2} g(\chi)(1+z), \tag{2.112}$$

where lensing efficiency $g(\chi)$ is an integral function of geometrical factors and the radial distribution of the sources $\frac{\mathrm{d}N}{\mathrm{d}z}(\chi)$:

$$g(\chi) = f_K(\chi) \int_\chi^\infty \mathrm{d}\chi' \frac{\mathrm{d}N}{\mathrm{d}z} \frac{f_K(\chi' - \chi)}{f_K(\chi')} \xrightarrow{\frac{\mathrm{d}N}{\mathrm{d}z}(\chi') \to \delta^D(\chi' - \chi_s)} \frac{f_K(\chi)f_K(\chi_s - \chi)}{f_K(\chi_s)}. \tag{2.113}$$

Note that the redshift distribution of sources has to be normalized such that $\int \mathrm{d}\chi \frac{\mathrm{d}N}{\mathrm{d}z}(\chi) = 1$ and the equality that follows the arrow in Eq. (2.113) holds only if all sources are at a single redshift $z_s$; as we will see in Sec. (2.3.2), in the case of CMB lensing the last-scattering surface can be considered as a single source plan at $z_s = z_* \simeq 1090$.

## 2.3.2 CMB Lensing

We now move on to tackle the weak gravitational lensing of the CMB temperature and polarization anisotropies. After recombination at $z \sim 1100$, CMB photons can freely propagate through an evolving clumpy medium before being collected by our telescopes. The LSS leaves an imprint on CMB anisotropies by gravitationally deflecting CMB photons during their journey from the last-scattering surface to us. The net effect is a remapping of the CMB observables, dependent on the gravitational potential integrated along the line-of-sight. Thus the effect is sensitive to both the geometry of the Universe and to the growth of the large-scale structure (see Blanchard & Schneider (1987); Cole & Efstathiou (1989) and Lewis & Challinor (2006) for a review); lensing can also be used to break the degeneracy of geometrical parameters that CMB measurement alone cannot (Stompor & Efstathiou, 1999; Sherwin et al., 2011). Moreover, lensing introduces non-Gaussian features in the CMB anisotropy pattern which can be exploited to get information on the intervening mass distribution (Hu & Okamoto, 2002; Hirata & Seijak, 2003), which in turn may give hints on the early stages of cosmic acceleration (Acquaviva & Baccigalupi, 2006; Hu et al., 2006).

CMB lensing is a unique cosmological probe because:

1. CMB, which acts as a backlight, can be considered as a single source plane extended across all sky at a known redshift;

2. it represents the only source plane for lenses at very high redshifts;

3. it is mainly sourced by scales in the linear regime.

However, it is an integrated measure of the matter distribution and it also converts the polarization E-modes into B-modes, hindering the power of primordial gravitational induced B-modes.

**CMB lensing potential**

Let us recall here some useful definitions previously introduced to illustrate the CMB weak gravitational lensing. In FLRW cosmologies a light ray approaching a matter density distribution is deflected by an angle

$$\delta\beta = -2\delta\chi\nabla_\perp\Phi, \tag{2.114}$$

where $\delta\beta$ is the *local* deflection, $\nabla_\perp$ is the spatial gradient on a plane perpendicular to the light propagation direction, and $\delta\chi$ is a small distance along the photon path. The comoving distance that the source appears to have shifted from its actual position because of lensing is $f_K(\chi_* - \chi)\delta\beta = f_K(\chi_*)\delta\theta$, where $\chi_*$ is the comoving distance of the source. Considering that a photon undergoes several deviations during its journey and that we are in the weak lensing regime, we can evaluate the total deflection angle $\boldsymbol{\alpha}$ as

$$\boldsymbol{\alpha} = -2\int_0^{\chi_*} \mathrm{d}\chi \frac{f_K(\chi_* - \chi)}{f_K(\chi_*)}\nabla_\perp\Phi(\chi\hat{\mathbf{n}}, \eta_0 - \chi), \tag{2.115}$$

where $\eta_0 - \chi$ is the conformal time at which the photon was at position $\chi\hat{\mathbf{n}}$.

Eq. (2.115) is cumbersome because it requires the integral to be evaluated along the *perturbed* light path. Nonetheless, since we are working at first order in $\Phi$ (small deflections), we can compute the integral over the unperturbed path of the photon. From the photons geodesic equation (2.96), one has

$$\chi = \eta_0 - \eta - 2\int_{\eta_0}^\eta \mathrm{d}\eta'\Phi(\eta'). \tag{2.116}$$

The Born approximation is equivalent to set $\chi = \eta_0 - \eta$, i.e. calculating the deflection along the unperturbed light path, so that the transverse derivative in Eq. (2.115) becomes the covariant derivative over the LOS $\hat{\mathbf{n}}$, $\nabla_\perp \to \nabla_{\hat{\mathbf{n}}}/f_K(\chi)$. The validity of the Born approximation in the context of CMB lensing has been recently investigated by Calabrese et al. (2015) by means of multiple planes ray-

tracing techniques, who have shown the differences to be small. Recently, CMB weak lensing beyond-Born approximation has been also investigated from a theoretical point of view by Hagstotz et al. (2015); Pratten & Lewis (2016); Marozzi et al. (2016).

Assuming an average potential well to be $|\Phi| \sim 10^{-4}$ and large $\sim 300$ Mpc, a back of the envelope calculation shows that the number of independent deflections is $\sim 14000/300 \sim 50$ (where 14000 Mpc is the comoving distance to the last-scattering surface), meaning that the total r.m.s. deflection should be $\sim \sqrt{50} \times 10^{-4} \sim 7 \times 10^{-4}$, or about 2 arcmin. These deflections are coherent across a scale given approximately by the angular scale subtended by a typical 300 Mpc structure located halfway to the last-scattering surface, $\sim 300/7000 \sim 2\,\mathrm{deg}$. From the deflection angle $\boldsymbol{\alpha}$ we can define the *lensing potential* $\phi(\hat{\mathbf{n}})$ as

$$\phi(\hat{\mathbf{n}}) \equiv -2 \int_0^{\chi_*} \mathrm{d}\chi \, \frac{f_K(\chi_* - \chi)}{f_K(\chi_*) f_K(\chi)} \Phi(\chi\hat{\mathbf{n}}, \eta_0 - \chi), \qquad (2.117)$$

so that we can relate the two as $\boldsymbol{\alpha}(\hat{\mathbf{n}}) = \nabla_{\hat{\mathbf{n}}} \phi(\hat{\mathbf{n}})$ (from now on $\nabla \equiv \nabla_{\hat{\mathbf{n}}}$). The lensing potential can be expanded into multipole moment (all-sky) or Fourier (flat-sky):

$$\phi(\hat{\mathbf{n}}) = \sum_{\ell m} \phi_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}), \qquad (2.118)$$

$$\phi(\hat{\mathbf{n}}) = \int \frac{\mathrm{d}^2 \boldsymbol{\ell}}{(2\pi)^2} \phi(\boldsymbol{\ell}) e^{i\boldsymbol{\ell} \cdot \hat{\mathbf{n}}}, \qquad (2.119)$$

where $(\ell, m)$ and $\boldsymbol{\ell}$ are conjugate to the real space unit vector $\hat{\mathbf{n}}$ in all-sky and flat-sky respectively (see Hu (2000b)). If we consider the primordial perturbations to be Gaussian, then the assumption of linear evolution implies that the lensing potential is also Gaussian, and all the information about the field is carried by its power spectrum:

$$\langle \phi_{\ell m} \phi_{\ell' m'} \rangle = \delta_{\ell\ell'}^K \delta_{mm'}^K C_\ell^{\phi\phi}, \qquad (2.120)$$

$$\langle \phi(\boldsymbol{\ell}) \phi(\boldsymbol{\ell}') \rangle = (2\pi)^2 \delta^D(\boldsymbol{\ell} - \boldsymbol{\ell}') C_\ell^{\phi\phi}. \qquad (2.121)$$

Cosmology predicts the shape and the amplitude of the CMB lensing power spectrum, which depends on geometrical factors and on the metric perturbation evolution (Lewis & Challinor, 2006):

$$C_\ell^{\phi\phi} = 16\pi \int \mathrm{d}\log k \, \mathcal{P}_{\mathcal{R}}(k) \left[ \int_0^{\chi_*} \mathrm{d}\chi \, T_\Phi(k; \eta_0 - \chi) j_\ell(k\chi) \frac{f_K(\chi_* - \chi)}{f_K(\chi_*) f_K(\chi)} \right]^2. \qquad (2.122)$$

Here $j_\ell(k\chi)$ is the spherical Bessel function and $T_\Phi(k;\eta)$ is the linear theory transfer function such that the gravitational potential at later epoch is $\Phi(k;\eta) = T_\Phi(k;\eta)\mathcal{R}(k)$, where $\mathcal{R}(k)$ being the primordial comoving curvature perturbation (set at the inflationary epoch) with power spectrum $\mathcal{P}_\mathcal{R}(k)$. The CMB lensing potential power spectrum, shown in Fig. (2.5), can be related to the deflection angle and to the convergence $\kappa(\hat{\mathbf{n}}) = -\frac{1}{2}\nabla^2\phi(\hat{\mathbf{n}})$ ones through[27]

$$C_\ell^{\alpha\alpha} = \ell(\ell+1)C_\ell^{\phi\phi}, \tag{2.123}$$

$$C_\ell^{\kappa\kappa} = \frac{[\ell(\ell+1)]^2}{4}C_\ell^{\phi\phi}, \tag{2.124}$$

where we used the following relations (Hu, 2000b)

$$\alpha_{\ell m} = -i\sqrt{\ell(\ell+1)}\phi_{\ell m}, \tag{2.125}$$

$$\kappa_{\ell m} = -\frac{\ell(\ell+1)}{2}\phi_{\ell m}, \tag{2.126}$$

$$\tag{2.127}$$

and assuming the curl component of the deflection angle to vanish.

### CMB lensed power spectra

The effect of gravitational lensing on CMB photons can be described as a remapping of the unlensed temperature anisotropies $\Theta(\hat{\mathbf{n}})$ by the deflection field $\boldsymbol{\alpha}(\hat{\mathbf{n}})$. By Taylor expanding in terms of the displacement field we get

$$\begin{aligned}
\tilde{\Theta}(\hat{\mathbf{n}}) &= \Theta(\hat{\mathbf{n}} + \boldsymbol{\alpha}(\hat{\mathbf{n}})) \\
&= \Theta(\hat{\mathbf{n}} + \nabla\phi(\hat{\mathbf{n}})) \\
&= \Theta(\hat{\mathbf{n}}) + \nabla^i\phi(\hat{\mathbf{n}})\nabla_i\Theta(\hat{\mathbf{n}}) + \frac{1}{2}\nabla^i\phi(\hat{\mathbf{n}})\nabla^j\phi(\hat{\mathbf{n}})\nabla_i\nabla_j\Theta(\hat{\mathbf{n}}) + \mathcal{O}(\phi^3),
\end{aligned} \tag{2.128}$$

where the tilde denotes a lensed quantity. In the flat-sky approximation Eq. (2.128) reads as

$$\begin{aligned}
\tilde{\Theta}(\boldsymbol{\ell}) &\approx \Theta(\boldsymbol{\ell}) - \int \frac{\mathrm{d}^2\boldsymbol{\ell}'}{2\pi}\boldsymbol{\ell}' \cdot (\boldsymbol{\ell}' - \boldsymbol{\ell}')\phi(\boldsymbol{\ell} - \boldsymbol{\ell}')\Theta(\boldsymbol{\ell}') \\
&\quad - \frac{1}{2}\int \frac{\mathrm{d}^2\boldsymbol{\ell_1}}{2\pi}\int \frac{\mathrm{d}^2\boldsymbol{\ell_2}}{2\pi}\boldsymbol{\ell_1} \cdot [\boldsymbol{\ell_1} + \boldsymbol{\ell_2} - \boldsymbol{\ell}]\boldsymbol{\ell_1} \cdot \boldsymbol{\ell_2}\Theta(\boldsymbol{\ell_1})\phi(\boldsymbol{\ell_2})\phi^*(\boldsymbol{\ell_1} + \boldsymbol{\ell_2} - \boldsymbol{\ell}),
\end{aligned} \tag{2.129}$$

from which we see that lensing affects the unlensed multipoles by coupling different scales. For a fixed realization of lenses, the mode-coupling effect is to introduce

---

[27]Sometimes the deflection angle $\boldsymbol{\alpha}(\hat{\mathbf{n}})$ is denoted as $\mathbf{d}(\hat{\mathbf{n}})$.
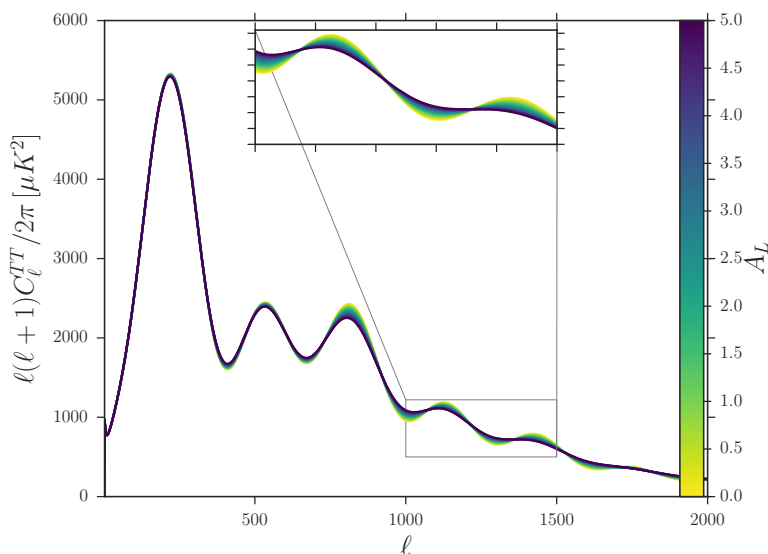
**Figure 2.4:** Effect of the weak gravitational lensing on the temperature angular power spectrum as function of the phenomenological lensing amplitude parameter $A_L$ that rescales the fiducial CMB lensing power spectrum as $C_\ell^{\phi\phi} \to A_L C_\ell^{\phi\phi,\Lambda CDM}$.

off-diagonal components into the covariance matrix of observed temperature, and large-wavelength mode of the lensing potential correlate CMB modes on smaller scales. The Taylor expansion performed above remarkably simplifies the description of the lensing effect on CMB anisotropies, however it is not accurate on all scales, especially when looking at scales comparable to the deflections.

Assuming that both $\tilde{\Theta}(\boldsymbol{\ell})$ and $\phi(\boldsymbol{\ell})$ have the statistics of an isotropic, Gaussian random field and are uncorrelated[28], we find

$$\langle \tilde{\Theta}(\boldsymbol{\ell})\tilde{\Theta}^*(\boldsymbol{\ell'})\rangle = \delta^D(\boldsymbol{\ell} - \boldsymbol{\ell'}) \tag{2.130}$$

$$C_\ell^{\tilde{T}\tilde{T}} \approx (1 - \ell^2 R^\phi)C_\ell^{TT} + \int \frac{d^2\boldsymbol{\ell'}}{2\pi}[\boldsymbol{\ell'} \cdot (\boldsymbol{\ell} - \boldsymbol{\ell'})]^2 C_{|\boldsymbol{\ell}-\boldsymbol{\ell'}|}^{\phi\phi} C_{\ell'}^{TT}, \tag{2.131}$$

where we have defined half the total deflection angle power $R^\phi$

$$R^\phi \equiv \frac{1}{2}\langle|\nabla\phi|^2\rangle = \frac{1}{4\pi}\int \frac{d\ell}{\ell}\ell^4 C_\ell^{\phi\phi} \sim 3 \times 10^{-7}. \tag{2.132}$$

The lensed temperature power spectrum to first-order in $C_\ell^{\phi\phi}$ differs from the unlensed one by a term proportional to $R^\phi$ and by a convolution of the unlensed temperature gradient power spectrum with the deflection power spectrum. As shown in Fig. (2.4), the overall effect is to smooth the CMB acoustic peaks and to

---

[28]We neglect the correlation $\langle T\phi\rangle$ induced by secondary anisotropies such as the iSW and Sunyaev-Zel'dovich effects.

boost the small-scale power in the damping tail. Since weak lensing preserves the brightness but alters the photons directions, the total variance of the temperature field is conserved even though the power is reshuffled among different scales. For an exact calculation of the lensed spectra in the curved sky framework, which relies on real space correlation function methods and is commonly adopted in modern Boltzmann codes, we refer the reader to Challinor & Chon (2002); Challinor & Lewis (2005).

What about the polarization? The lensing effect on the CMB polarization is also described as a remapping of the Stokes parameters $Q \pm iU$ by the deflection field. Assuming no primordial $B$-modes, i.e. $C_\ell^{BB} = 0$, the lensed power spectra to lowest order in $C_\ell^{\phi\phi}$ are (Hu, 2000b)

$$C_\ell^{\tilde{T}\tilde{E}} = (1 - \ell^2 R^\phi)C_\ell^{TE} + \int \frac{\mathrm{d}^2\boldsymbol{\ell}'}{(2\pi)^2}[\boldsymbol{\ell}' \cdot (\boldsymbol{\ell} - \boldsymbol{\ell}')]^2 C_{|\boldsymbol{\ell}-\boldsymbol{\ell}'|}^{\phi\phi}C_{\ell'}^{TE}\cos 2(\varphi_{\ell'} - \varphi_{\ell'}),$$

(2.133)

$$C_\ell^{\tilde{E}\tilde{E}} = (1 - \ell^2 R^\phi)C_\ell^{EE} + \int \frac{\mathrm{d}^2\boldsymbol{\ell}'}{(2\pi)^2}[\boldsymbol{\ell}' \cdot (\boldsymbol{\ell} - \boldsymbol{\ell}')]^2 C_{|\boldsymbol{\ell}-\boldsymbol{\ell}'|}^{\phi\phi}C_{\ell'}^{EE}\cos^2 2(\varphi_{\ell'} - \varphi_{\ell'}),$$

(2.134)

$$C_\ell^{\tilde{B}\tilde{B}} = \int \frac{\mathrm{d}^2\boldsymbol{\ell}'}{(2\pi)^2}[\boldsymbol{\ell}' \cdot (\boldsymbol{\ell} - \boldsymbol{\ell}')]^2 C_{|\boldsymbol{\ell}-\boldsymbol{\ell}'|}^{\phi\phi}C_{\ell'}^{EE}\sin^2 2(\varphi_{\ell'} - \varphi_{\ell'}),$$

(2.135)

where $\varphi_\ell$ and $\varphi_{\ell'}$ are the angles between $\hat{\mathbf{n}}$ and $\boldsymbol{\ell}$ and $\boldsymbol{\ell}'$ respectively. Eq. (2.135) tells us that the lensing of pure $E$-modes generates $B$-mode polarization, even if initially absent. Qualitatively, lensing acts on $C_\ell^{EE}$ and $C_\ell^{TE}$ similarly to what happens to the temperature: the unlensed spectra are convolved with the lensing potential one, with a resulting blurring of spectral features and power shift to the damping tail. Since the acoustic peaks in $C_\ell^{EE}$ are sharper than the temperature ones, the fractional changes in the lensed E-modes are $\mathcal{O}(30\%)$ near the acoustic peaks (Challinor & Lewis, 2005).

Lensing induces an $E$- to $B$-modes power leakage which represents a contaminant in the search for gravitational induced $B$-modes on large angular scales. In order to get a rough estimate of the contamination, consider that the unlensed $C_\ell^{EE}$ peaks around $\ell \sim 1000$, so for $\ell \ll 1000$ we can assume $|\boldsymbol{\ell}| \gg |\boldsymbol{\ell}'|$ for the lensed $B$-mode and find

$$\begin{aligned} C_\ell^{\tilde{B}\tilde{B}} &\approx \int \frac{\mathrm{d}^2\boldsymbol{\ell}'}{(2\pi)^2}|\boldsymbol{\ell}'|^4 C_{\ell'}^{\phi\phi}C_{\ell'}^{EE}\sin^2 2(\varphi_{\ell'} - \varphi_{\ell'}) \\ &= \frac{1}{4\pi}\int \frac{\mathrm{d}\ell'}{\ell'}\ell'^4 C_{\ell'}^{\phi\phi}\ell'^2 C_{\ell'}^{EE}, \end{aligned}$$

(2.136)

which is independent of $\ell$, corresponding to a $B$-modes white noise power spectrum
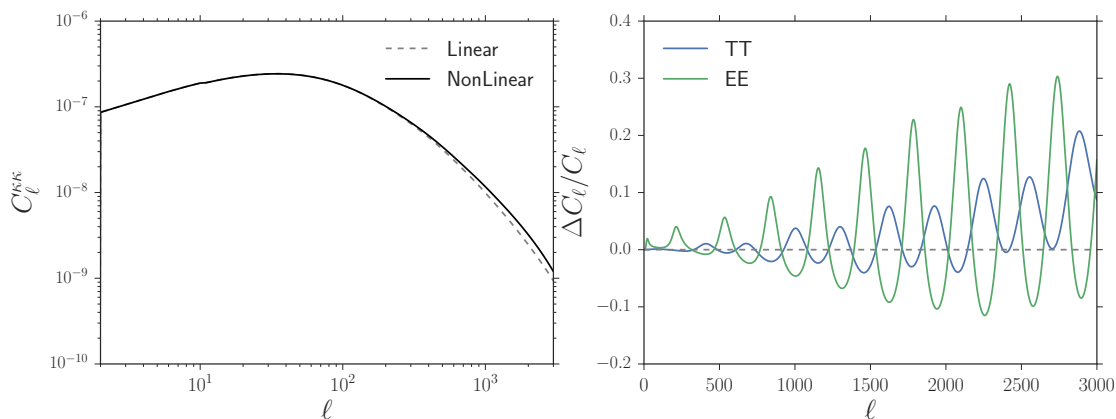
**Figure 2.5:** *Left panel:* Angular power spectra of the CMB lensing convergence evaluated considering linear (dashed line) and nonlinear (solid line) matter evolution. *Right panel:* Comparison between the lensed and unlensed temperature (blue line) and polarization *E*-modes (green line) power spectra, showing the signatures induced by weak lensing.

of amplitude around $C_\ell^{BB} \sim 2 \times 10^{-6}\,\mu\text{K}^2$ for an instrumental white noise level of $\sim 5\mu\text{Karcmin}$.

**CMB lensing reconstruction**

If we assume to know the statistics of the unlensed CMB, it is possible to exploit the information in the 4-point function of the lensed CMB to extract *statistically* the lensing potential. As we have previously seen, the effect of the lensing is to introduce small departures from Gaussianity in the CMB (when marginalized over realizations of the lenses), or statistical anisotropy (for a fixed distribution of the lenses). There are two main methods to detect the CMB lensing: one is to measure the smoothing of the acoustic peaks on small angular scales induced by lensing at the CMB power spectrum level (Reichardt et al., 2009; Dunkley et al., 2010; Keisler et al., 2011; Planck Collaboration, 2014b), while the second method involves measuring the peculiar mode-coupling induced in the CMB by lensing. Methods that directly reconstruct the deflection field can either employ a maximum likelihood approach (Hirata & Seijak, 2003) or optimal quadratic estimators (Hu & Okamoto, 2002). So far, all reconstructed lensing maps have employed the optimal quadratic estimation which is derived under idealized observational conditions, i.e. the CMB is nearly perfect Gaussian field, as well as the lensing field, with negligible mode-coupling induced by instrumental or foregrounds effect. Several papers have investigated the effect of a number of additional sources of mode-coupling, such as finite sky-coverage (Perotto et al., 2010; Namikawa et al., 2013), primordial non-Gaussianity (Lesgourgues et al., 2005), higher-order mode-coupling induced by lensing itself (Kesden et al., 2003; Hanson et al., 2011), and foreground biases

(Fantaye et al., 2012; Osborne et al., 2014; van Engelen et al., 2014).

Let us briefly review the optical quadratic estimator which is based on the first order perturbative expansion, similarly to Eq. (2.129) but truncated at first order. Assuming the different CMB realizations to be lensed by the same lensing potential, the off-diagonal components of the 2-point statistics of the observed fields $\tilde{X} \in \{T, E, B\}$ read as

$$\langle \tilde{X}(\boldsymbol{\ell_1}) \tilde{X}'^*(\boldsymbol{\ell_2}) \rangle_{CMB} = \delta^D(\boldsymbol{L}) + \frac{1}{2\pi} f_{\tilde{X}\tilde{X}'}(\boldsymbol{\ell_1}, \boldsymbol{\ell_1}) \phi(\boldsymbol{L}) + \mathcal{O}(\phi^2), \quad \boldsymbol{L} = \boldsymbol{\ell_1} + \boldsymbol{\ell_2} \tag{2.137}$$

where the average $\langle \cdot \rangle_{CMB}$ represent an average over CMB realizations and $f_{XX}$ denotes the different coupling kernels between the CMB and $\phi$ which depends on the given CMB observable $X$ (see Hu & Okamoto (2002) for their precise expression). By properly filtering the harmonic modes of the observables, we can construct an estimate of the lensing potential field $\hat{\phi}$ as

$$\hat{\phi}(\boldsymbol{L}) = A_\alpha(\boldsymbol{L}) \int \frac{\mathrm{d}\boldsymbol{\ell_1^2}}{2\pi} X(\boldsymbol{\ell_1}) X'(\boldsymbol{\ell_2}) F_\alpha(\boldsymbol{\ell_1}, \boldsymbol{\ell_2}), \quad \alpha \equiv XX' \tag{2.138}$$

$$A_\alpha = \int \frac{\mathrm{d}\boldsymbol{\ell_1^2}}{2\pi} f^\alpha(\boldsymbol{\ell_1}, \boldsymbol{\ell_2}) F_\alpha(\boldsymbol{\ell_1}, \boldsymbol{\ell_2}). \tag{2.139}$$

The function $A_\alpha(\boldsymbol{L})$ is a normalization that guarantees that the estimate is unbiased when averaged over CMB realizations, while $F_\alpha$ is a filter function which can be optimized to yield an optimal estimator and is, in general, a function of the observed lensed and unlensed power spectra comprehensive of instrumental noise. The effect of filtering is to match the observed CMB with its gradient. From the (unbiased) estimated harmonic coefficients of the lensing field one can compute the lensing potential power spectrum using one of the 3! combination of the observable fields $\alpha$ as

$$\langle \hat{\phi}(\boldsymbol{L})_\alpha \hat{\phi}(\boldsymbol{L}')_\beta \rangle = \delta^D(\boldsymbol{L} - \boldsymbol{L}')[\hat{C}_L^{\phi\phi} + N_L^{(0)\alpha\beta}]. \tag{2.140}$$

Here $N_L^{(0)}$ is the lowest order noise on the reconstructed lensing power spectrum (often called Gaussian bias). Each combination of quadratic estimators provides a different reconstruction noise level and it is possible to combine them in a minimum-variance (MV) estimator in order to increase the S/N. The cleanest reconstruction in general depends on the specific observational setup, but for low-noise experiments the combination which has the highest S/N is the one exploiting $E$- and $B$-modes. This is because the observed $B$-modes are a direct result of the lensing effects since the primordial is expected to be small.

The CMB lensing reconstruction scheme outlined above has been derived within the flat-sky framework, for completeness we provide here the full-sky analogue

which has been exploited by the *Planck* team for the internal lensing reconstruction. In this case, if we ensemble average over a fixed realization of lenses, the CMB covariance matrix acquires off-diagonal element given by

$$\text{Cov}[X_{\ell_1 m_1} Z_{\ell_2 m_2}] = \sum_{LM} (-1)^M \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & -M \end{pmatrix} \mathcal{W}^{XZ}_{\ell_1 \ell_2 L} \phi_{LM}, \qquad (2.141)$$

where the fields $X_{\ell m}, Z_{\ell m} \in \{T_{\ell m}, E_{\ell m}, B_{\ell m}\}$ and the precise expression for the covariance response functions $\mathcal{W}^{XZ}$, that tells how different temperature and polarization modes couple to each other, can be found in Okamoto & Hu (2003), while the quantities in parenthesis are the $3j$ Wigner symbols (introduced in Ch. (3.1.2)). From Eq. (2.141) we can se that the off-diagonal information is proportional to $\phi_{LM}$ and can thus be used to estimate the lensing potential. The idea is to inverse-variance filter the observed sky maps $\bar{X}$ and $\bar{Z}$ and construct a general quadratic estimator from the observables combination $XZ$ as

$$\bar{x}_{LM}[\bar{X}, \bar{Z}] = \frac{(-1)^M}{2} \sum_{\ell_1 \ell_2 m_1 m_2} \begin{pmatrix} \ell_1 & \ell_2 & L \\ m_1 & m_2 & -M \end{pmatrix} W^x_{\ell_1 \ell_2 L} \bar{X}_{\ell_1 m_1} \bar{Z}_{\ell_2 m_2}, \qquad (2.142)$$

where $W^x_{\ell_1 \ell_2 L}$ is a set of weight functions that define the estimator $x$. Then, a quadratic estimate of the lensing potential can be written as

$$\hat{\phi}^x_{LM} = \sum_{\ell m} \left[ \mathcal{R}^{x\phi} \right]^{-1}_{LM\ell m} \left[ \bar{x}_{\ell m} - \bar{x}^{\text{MF}}_{\ell m} \right], \qquad (2.143)$$

where $\mathcal{R}^{x\phi}$ is a normalization matrix and the *mean-field* term $\bar{x}^{\text{MF}}_{L'M'}$ accounts for *known* sources of statistical anisotropy in the map (such as masking, beam asymmetry, and inhomogeneous noise), and it is usually estimated by means of Monte Carlo (MC) simulations and subtracted to debias the lensing estimate (Planck Collaboration, 2016a). The minimum-variance estimator can be obtained from the individual estimators as[29]

$$\hat{\phi}^{\text{MV}}_{LM} = \frac{\sum_x \hat{\phi}^x_{LM} \mathcal{R}^{x\phi}_L}{\sum_x \mathcal{R}^{x\phi}_L}, \qquad (2.144)$$

where the sum is taken over the possible combination of estimators $(TT, EE, TE, \ldots)$.

**Current status of observations**

CMB lensing is a fast evolving field of research. The first observational hints of the lensing effects on CMB were found by Smith et al. (2007) and Hirata

---

[29]Under the assumption that $\mathcal{R}^{x\phi}_{LM\ell m} = \delta^K_{L\ell} \delta^K_{Mm} \mathcal{R}^{x\phi}_L$.
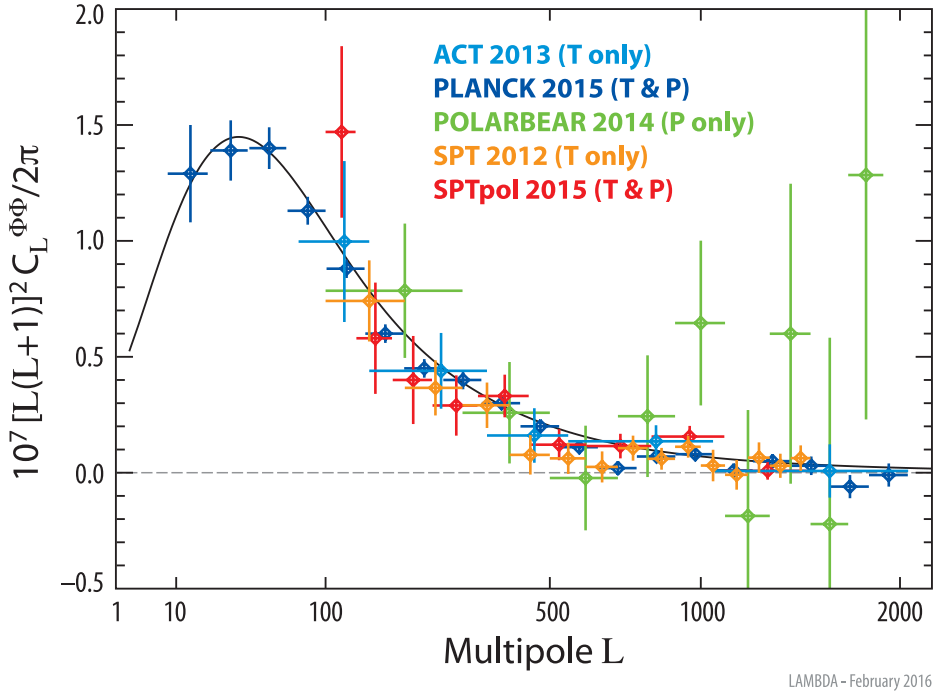
**Figure 2.6:** Status of CMB lensing measurements as of February 2016. Band powers shown in the plot are from ACT, *Planck* full mission, POLARBEAR, SPT, and SPTPol. The black solid line represents the theoretical lensing power spectrum for the best-fit $\Lambda$CDM parameters obtained from the *Planck* 2015 temperature and polarization data. The figure is taken from the Legacy Archive for Microwave Background Data Analysis (LAMBDA) website (`http://lambda.gsfc.nasa.gov`).

et al. (2008), who cross-correlated properly filtered CMB maps obtained from Wilkinson Microwave Anisotropy Probe (WMAP) with external LSS tracers. On the other hand, the first direct evidence of a preference for a lensed CMB is due to Reichardt et al. (2009) who combined Arcminute Cosmology Bolometer Array Receiver (ACBAR) and WMAP data. With the advent of the high sensitivity small scale CMB temperature experiment such as Atacama Cosmology Telescope (ACT) (Das et al., 2011) and South Pole Telescope (SPT) (Keisler et al., 2011), the CMB lensing potential reconstruction has become truly feasible for the first time, with reported detection significance around $\sim 4 - 6\sigma$. The next step forward in CMB lensing science has been made by the Planck team who detected the CMB lensing power spectrum at a significance of about $25\sigma$ and reconstructed an almost full-sky lensing map using only temperature data Planck Collaboration (2014c).[30] The first CMB lensing measurements using polarization data have been recently reported by POLARBEAR (Ade et al., 2014b), SPTPol (Story et al., 2015), and *Planck* (Planck Collaboration, 2016a): recent measurements are reported in Fig. (2.6).

---

[30]Consider that the *Planck* satellite is as sensitive to CMB lensing as COBE was to CMB temperature fluctuations: $T : \kappa = $ COBE : Planck.

## 2.3.3 LSS lensing

CMB photons are not the only ones that experience gravitational deflection by the intervening LSS. Galaxy photons are also affected by weak lensing which causes (i) the distortion the *shape* of galaxy images, and (ii) a change in their *magnitude* and *size* (Bartelmann & Schneider, 2001): these two main effects are at the core of Weak Lensing (WL) measurements. Here, we focus on the relationship between the observables and the cosmological signal in WL.

**Shear measurements**

The most common quantity inferred from observations is the shear of galaxy images. In the weak lensing limit ($|\gamma|, |\kappa| \ll 1$), the shear can be directly estimated from the *observed ellipticity* $\hat{\epsilon}$ as follows

$$\epsilon = \gamma + \epsilon_{\text{int}}, \tag{2.145}$$

where $\epsilon_{\text{int}}$ is the intrinsic ellipticity of source galaxies, for which current observations suggest $\sigma_{\text{int}} = \sqrt{\langle|\epsilon_{int}|^2\rangle} \simeq 0.4$. Measurement of galaxy ellipticities are also complicated by several observational systematics such as the point spread function of the instrument and the blurring caused by atmosphere which have to be finely controlled. Also in this case, what the theory predicts is the statistical correlation of galaxy shears, i.e. the shear power spectrum (or correlation function). The convergence (angular) power spectrum is equivalent to the shear power spectrum in the weak lensing limit, i.e. $C_\ell^{\kappa\kappa} \simeq C_\ell^{\gamma\gamma}$. Moreover, radial information can be used to perform a tomographic analysis of the WL signal by measuring the galaxy shear in redshift slices. Considering two redshift bins $i$ and $j$, with associated radial distribution of sources $\frac{\mathrm{d}N_{i,j}}{\mathrm{d}z}$, we can write the tomographic cross-power spectra in different bins as

$$\langle \kappa_{\ell m}^i \kappa_{\ell'm'}^j \rangle = \delta_{\ell\ell'}^K \delta_{mm'}^K C_\ell^{\kappa^i \kappa^j}, \tag{2.146}$$

which can be related to theory through

$$C_\ell^{\kappa^i \kappa^j} = \begin{cases} \int_0^\infty \mathrm{d}\chi \frac{W_i^\kappa(\chi)W_j^\kappa(\chi)}{f_K^2(\chi)} P_{\delta\delta}\left(\frac{\ell}{f_K(\chi)}, \chi\right) \\ \int_0^\infty \frac{\mathrm{d}z}{H(z)} \frac{W_i^\kappa(z)W_j^\kappa(z)}{f_K^2(z)} P_{\delta\delta}\left(\frac{\ell}{f_K(z)}, z\right). \end{cases} \tag{2.147}$$

We recall that Eq. (2.147) assumes both Limber and Born approximations. The observed auto-power spectra also include a shot-noise term due to random galaxy shape, $C_\ell^{\kappa^i \kappa^j} \to C_\ell^{\kappa^i \kappa^j} + \delta_{ij}\sigma_{int}^2/\bar{n}_i$, where $\bar{n}_i$ is the average number of galaxies per steradian in the $i$-th redshift bin: a large number of sources increases the statistics,

hence lowering the noise level. If we decompose the shear into a tangential and cross component as

$$\gamma_t \equiv -\text{Re}[\gamma e^{-2i\phi}], \tag{2.148}$$

$$\gamma_\times \equiv -\text{Im}[\gamma e^{-2i\phi}], \tag{2.149}$$

we can obtain a rotationally invariant linear combination of the cross-correlation functions of $\gamma_\times$ and $\gamma_t$ as follows:

$$
\begin{aligned}
\xi_\pm^{ij}(\theta) &\equiv \langle \gamma_t(0)\gamma_t(\boldsymbol{\theta}) \rangle \pm \langle \gamma_\times(0)\gamma_\times(\boldsymbol{\theta}) \rangle \\
&= \frac{1}{2\pi} \int_0^\infty \mathrm{d}\ell \, \ell C_\ell^{\kappa^i \kappa^j} J_{0/4}(\ell\theta),
\end{aligned} \tag{2.150}
$$

where the Bessel function $J_0$ ($J_4$) refers to the correlation function $\xi_+$ ($\xi_-$). A common estimator of the shear correlation function is ()

$$\hat{\xi}_\pm(\theta) = \frac{\sum_{ij} w_i w_j [e_t(\theta_i)e_t(\theta_j) \pm e_\times(\theta_i)e_\times(\theta_j)]}{\sum_{ij} w_i w_j}, \tag{2.151}$$

where all galaxy pairs $(i, j)$ separated by an angular distance $|\theta_i - \theta_j| \in \theta$ contribute to the same angular bin with their respective weights $(w_i, w_j)$.

An important systematic effect that has an impact on the above estimator is the intrinsic galaxy alignments due to tidal forces (Hirata & Seljak, 2004). In fact, Eq. (2.151) estimates $\langle \hat{\xi}_\pm \rangle = \xi_\pm + \xi_\pm^{\text{II}} + \xi_\pm^{\text{GI}}$, where $\xi_\pm^{\text{II}}$ measures correlation between the intrinsic ellipticities of neighboring galaxies (known as II) and $\xi_\pm^{\text{II}}$ is sensitive to correlation between foreground Galaxy Intrinsic ellipticity and background galaxy shear (known as GI). The modeling of galaxy alignment is a very active and debated area of research, see Kirk et al. (2015) for a review.

Intrinsic alignments represent a potential contaminant also to CMB lensing-galaxy lensing cross-correlation measurements, entering the 2-point function as a $\langle \kappa_{CMB} \text{I} \rangle$ term. Hand et al. (2015) measured for the first time the cross-correlation between the ACT CMB lensing and the galaxy convergence of Canada-France-Hawai Telescope (CFHT) Stripe 82 (CS82) and reported a lower amplitude with respect to the expect signal. Troxel & Ishak (2014) showed that intrinsic alignment contamination could account for a $\approx 15\%$ reduction of the observed discrepancy; these findings have been later confirmed by Chisari et al. (2015) who modeled the impact of the intrinsic alignment with a model based on early-type[31] galaxy population to CMB lensing measurements, pointing out that at $z \gtrsim 1.2$ alignments remain largely unconstrained. Larsen & Challinor (2015) have considered the impact of intrinsic

---

[31]In this context it means *red* galaxies.

alignments of spiral galaxies on the CMB lensing-galaxy lensing cross-correlation using the quadratic alignment model, finding the signal to be similar in shape with respect to the linear model but with opposite sign, hence it can potentially reduce the overall impact.

**Magnification bias**

The size amplification is called *magnification* and leads to a fluctuation in the size and the flux of individual galaxies. Two competing effects are at play: one one hand WL of photons stretches the observed area and *dilutes* the number density, on the other it changes the *apparent brightness* and allows galaxies in the survey. In a flux-limited survey, the number density of background galaxies is expected to be affected through magnification effects due to the foreground galaxy distribution. Let us see how.

In presence of lensing the galaxy flux is magnified by $f^{\mathrm{obs}} = \mu f^s$, where $f^{\mathrm{obs}}$ and $f^s$ are is the observed and intrinsic galaxy flux respectively. In the WL limit it can be shown that $N(> S)$, the intrinsic counts of sources with observed flux greater than $S$, is related to the cumulative source counts $\tilde{N}(> S)$ observed in a given direction through (Bartelmann & Schneider, 2001)

$$\tilde{N}(> S, \hat{\mathbf{n}}) = \mu^{\alpha-1}(\hat{\mathbf{n}})N(> S). \tag{2.152}$$

Here, $\alpha$ is the logarithmic slope of the cumulative number counts of galaxies at the faint end, $\alpha = \frac{\mathrm{d}\log N(>S)}{\mathrm{d}S}|_{S_{\mathrm{min}}}$. Thus, if $\alpha > 1$ ($\alpha < 1$), the observed number density of objects is enhanced (decreased) by lensing: this effect is the so-called *magnification (anti-)bias*. Since $|\kappa| \ll 1$, the magnification $\mu$ can be approximated as $\mu \approx 1 + 2(\alpha - 1)\kappa$ and the observed number density of galaxies is modulated by the foreground matter distribution as

$$\begin{aligned} \delta_g^{\mathrm{obs}}(\hat{\mathbf{n}}, z) &\simeq \delta_g^{\mathrm{cl}}(\hat{\mathbf{n}}, z) + 2(\alpha(z) - 1)\kappa(\hat{\mathbf{n}}, z) \\ &= \delta_g^{cl}(\hat{\mathbf{n}}, z) + \delta_g^{\mu}(\hat{\mathbf{n}}, z), \end{aligned} \tag{2.153}$$

where $\delta_g^{\mathrm{cl}}(\hat{\mathbf{n}}, z)$ is the intrinsic galaxy clustering term. By substituting in Eq. (2.153) the expression for the convergence $\kappa$ - of the sources in the redshift bin around $z_i$ - we get

$$\delta_{\mathrm{g}}^{\mu}(\hat{\mathbf{n}}, z) = \frac{3\Omega_{\mathrm{m}}}{2c} \frac{H_0^2}{H(z)} (1 + z)\chi(z) \int_z^{\infty} dz' \left(1 - \frac{\chi(z)}{\chi(z')}\right)(\alpha(z') - 1)\frac{dN}{dz'}. \tag{2.154}$$

Cosmic magnification measurements can be carried out by correlating the angular positions of background and foreground galaxy populations (see Scranton et al.

(2005); Gonzalez-Nuevo et al. (2014)) and by counting galaxies (Negrello et al., 2010). Magnification bias can be either a source of information (Hildebrandt et al., 2013) or a nuisance depending on the analysis performed and it can lead to substantial biases if not properly accounted for, as shown for CMB lensing-galaxy density (Bianchini et al., 2015) and iSW (CMB temperature-galaxy density) measurements (LoVerde et al., 2007). The high-$z$ DSFG detected by Herschel have a steep integrated number counts at the limiting flux, hence are potentially sensitive to the magnification bias as we will see in Ch. (4).

# 3 Chapter 3

# CMB-LSS cross-correlation: statistics and datasets

*Statistics: the mathematical theory of ignorance.*

—Morris Kline

Cosmology is intrinsically entwined with statistics. There are threefold reasons for this. Firstly, given the huge size of cosmological structure it is impossible to follow the evolution of all components in a deterministic way. Secondly, we do not observe along a single spatial hyper-surface, rather we have direct observational access to our past light-cone, which means that we see objects at different stages of their evolution. Thirdly the generation of primordial inhomogeneities is an intrinsic stochastic process and as such, we can just predict the statistical properties of cosmological fields (such as the matter density contrast $\delta$ or the intensity of CMB temperature fluctuations). The main implication of these remarks is that we should model the observable Universe as a stochastic *realization* of a stochastic *ensemble*. Moreover, in the vanilla inflationary framework, the initial perturbations are Gaussian and adiabatic meaning that the study of Gaussian random field (GRF), with a specific focus on isotropic and homogeneous ones, is pivotal in cosmology.

This chapter mainly deals with the methodology at the core of the analyses presented in this thesis and with the exploited datasets. We start by describing cosmological random fields with a special focus on the mathematics on the sphere. Then, we discuss the spectral estimation problem in the light of cosmological surveys, while in the last part of the chapter we present the *Planck* and *Herschel* datasets.

## 3.1 Random Fields

### 3.1.1 3D Random Fields

We can describe a perturbation (calculated at a given time $t$) as a random field $f(\boldsymbol{x})$, which basically means that we assign a random number at each point $\boldsymbol{x} \in \mathbb{R}$

according to a probability density function $P[f(\boldsymbol{x})]$.[1] Here, $P[f(\boldsymbol{x})]$ is a distribution that gives the probability of realizing a particular field configuration.[2] Strictly speaking, a random field is a collection of $N$ random variables $f(\boldsymbol{x}_i)$ with $x_i \in \mathbb{R}^n$; the set of functions is called an *ensemble* while each individual function represents a *realization* of the ensemble.

The correlators of the fields are the expectation values of the product of the field at different locations (and times). The 2-point correlation function (2PCF) is defined by

$$\xi(\boldsymbol{x}, \boldsymbol{y}) \equiv \langle f(\boldsymbol{x})f(\boldsymbol{y}) \rangle = \int \mathcal{D}f \, P[f] f(\boldsymbol{x}) f(\boldsymbol{y}), \tag{3.1}$$

where the integral is a functional integral (or path integral) over the field configurations. The requirements of statistical isotropy and homogeneity translate into the following constraints on the correlation function: $\xi(\boldsymbol{x}, \boldsymbol{y}) = \xi(R^{-1}\boldsymbol{x}, R^{-1}\boldsymbol{y})$ for the former and $\xi(\boldsymbol{x}, \boldsymbol{y}) = \xi(\boldsymbol{x} - \boldsymbol{y})$ for the latter, where $R$ is a generic rotation matrix. Combining the two together we find that the 2PCF only depends on the distance between points, i.e. $\xi(\boldsymbol{x}, \boldsymbol{y}) = \xi(|\boldsymbol{x} - \boldsymbol{y}|)$. The same calculations can be performed in the harmonic domain by Fourier transforming the field adopting the following convention:

$$f(\boldsymbol{k}) = \int \mathrm{d}^3 x \, f(\boldsymbol{x}) e^{-i\boldsymbol{k} \cdot \boldsymbol{x}} \quad \text{and} \quad f(\boldsymbol{x}) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} f(\boldsymbol{k}) e^{i\boldsymbol{k} \cdot \boldsymbol{x}} \tag{3.2}$$

Generically the Fourier transform is complex but for real valued fields we have that $f(\boldsymbol{k}) = f^*(-\boldsymbol{k})$. Then, the 2PCF in Fourier space than reads as

$$\begin{aligned}
\langle f(\boldsymbol{k})f^*(\boldsymbol{k}') \rangle &= (2\pi)^3 \delta^D(\boldsymbol{k} - \boldsymbol{k}') \frac{2\pi^2}{k^3} \mathcal{P}_f(k) \\
&= (2\pi)^3 \delta^D(\boldsymbol{k} - \boldsymbol{k}') P_f(k)
\end{aligned} \tag{3.3}$$

where $\mathcal{P}_f(k)$ is the *dimensionless power spectrum* and the homogeneity and isotropy are enforced by the delta function and the dependence only on the module of $k$. The 2PCF and the power spectrum are a Fourier transform pairs:

$$\begin{aligned}
\xi(\boldsymbol{x}, \boldsymbol{y}) = \langle f(\boldsymbol{x})f(\boldsymbol{y}) \rangle &= \int \frac{\mathrm{d}^3 k}{(2\pi)^3} \frac{\mathrm{d}^3 k'}{(2\pi)^3} \langle f(\boldsymbol{k})f^*(\boldsymbol{k}') \rangle e^{i\boldsymbol{k} \cdot \boldsymbol{x}} e^{i\boldsymbol{k}' \cdot \boldsymbol{y}} \\
&= \frac{1}{4\pi} \int \mathrm{d} \log k \, \mathcal{P}_f(k) \int \mathrm{d}\Omega e^{i\boldsymbol{k} \cdot (\boldsymbol{x} - \boldsymbol{y})}.
\end{aligned} \tag{3.4}$$

---

[1] We will consider *centered* field, i.e. $\langle f(\boldsymbol{x}) \rangle = 0$ for the sake of clarity. This is not an issue since the mean can always be subtracted from the field.

[2] The probability that the field $f$ will have a given configuration between $f(\boldsymbol{x})$ and $f(\boldsymbol{x}) + \mathrm{d}f(\boldsymbol{x})$ is calculated as $\mathrm{d}P = P[f(\boldsymbol{x})]\Pi_{\boldsymbol{x}} \, \mathrm{d}f(\boldsymbol{x}) \equiv P[f(\boldsymbol{x})]\mathcal{D}f$.

After the angular integration, Eq. (3.4) reduces to

$$\xi(\boldsymbol{x}, \boldsymbol{y}) = \int \mathrm{d}\log k \, \mathcal{P}_f(k) j_0(k|\boldsymbol{x} - \boldsymbol{y}|), \tag{3.5}$$

where $j_0(x) = \sin x/x$ and the variance is computed as the 2PCF at zero-lag,

$$\sigma^2 = \xi(0) = \int \mathrm{d}\log k \, \mathcal{P}_f(k). \tag{3.6}$$

Higher-order correlations can be evaluated by averaging the product of $n$ fields; some of the most studied polyspectra are the 3PCF or *bispectrum*

$$\langle f(\boldsymbol{k}_1) f(\boldsymbol{k}_2) f(\boldsymbol{k}_3) \rangle = (2\pi)^3 B(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3) \delta^D(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3), \tag{3.7}$$

and the 4PCF or *trispectrum*

$$\langle f(\boldsymbol{k}_1) f(\boldsymbol{k}_2) f(\boldsymbol{k}_3) f(\boldsymbol{k}_4) \rangle = (2\pi)^3 T(\boldsymbol{k}_1, \boldsymbol{k}_2, \boldsymbol{k}_3 \boldsymbol{k}_4) \delta^D(\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{k}_3 + \boldsymbol{k}_4). \tag{3.8}$$

**Gaussian Random Field**

This class of random fields is of particular relevance in cosmology due to the high-level of Gaussianity predicted by inflation for primordial fluctuations. For (homogeneous and isotropic) GRF, $\mathrm{P}[f(\boldsymbol{x})]$ is a Gaussian functional of $f$ fully characterized by its power spectrum[3], or by the 2PCF equivalently. If we discretize the field in $N$ pixels and represent it as a collection of $f_i = f(\boldsymbol{x}_i)$ in a $N$-dimensional vector $\boldsymbol{f} = [f_1, f_2, \ldots, f_N]^T$, the multivariate joint probability distribution of the GRF has the following form

$$\mathrm{P}[\boldsymbol{f}] \propto \frac{e^{-f_i \mathrm{C}_{ij}^{-1} f_j}}{\sqrt{\det(\mathrm{C})}}, \tag{3.9}$$

where $\mathrm{C}_{ij} = \langle f_i f_j \rangle$ is the covariance matrix. Since the Gaussian is even under parity around the mean, any odd expectation value vanishes, e.g. $\langle f(\boldsymbol{x_1}) f(\boldsymbol{x_2}) f(\boldsymbol{x_3}) \rangle = 0$, while any even higher-order correlation function can be written as sum of 2PCF products: this result is known as *Wick's theorem* (or Isserlis theorem) and can be expressed as

$$\langle f(\boldsymbol{x_1}) f(\boldsymbol{x_2}) f(\boldsymbol{x_3}) f(\boldsymbol{x_4}) \ldots \rangle = \sum \text{All possibile 2-point contractions.} \tag{3.10}$$

---

[3]Since the mean is zero.

### 3.1.2 **Fields on the sphere**

Nearly all cosmological observations, from CMB to galaxy surveys, provide us with a sampling of 3D random fields projected on the sphere, mostly in the form of two-dimensional sky-maps.[4] The information content hidden in such maps is usually probed by means of harmonic analysis on the sphere. A popular observable that characterizes the statistical properties of a given cosmic field is the angular power spectrum $C_\ell$ and its reconstruction enables a direct comparison between models and data.

It is common practice to decompose the observed field $X(\hat{\mathbf{n}})$ into spherical harmonics, a frequency-space orthonormal basis for square-integrable, i.e. $\int_{\mathbb{S}^2} d\Omega |X(\hat{\mathbf{n}})|^2 < \infty$, as:

$$X(\hat{\mathbf{n}}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} x_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}), \tag{3.11}$$

where the spherical harmonic coefficients are given by

$$x_{\ell m} = \int_{\mathbb{S}^2} X(\hat{\mathbf{n}}) Y_{\ell m}^*(\hat{\mathbf{n}}) d\Omega. \tag{3.12}$$

The spherical harmonics $Y_{\ell m}$ are the eigenfunctions of the Laplacian on the sphere $\nabla^2_{\mathbb{S}^2} = \frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \left( \sin\theta \frac{\partial}{\partial\theta} \right) + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\varphi^2}$, and are described by two integers, the *multipole* $\ell$ and the azimuthal parameter $m$ that satisfy $\nabla^2_{\mathbb{S}^2} Y_{\ell m} = -\ell(\ell+1) Y_{\ell m}$ and $\partial_\varphi Y_{\ell m} = im Y_{\ell m}$. We recall below some useful properties:

$$\int_{\mathbb{S}^2} d\Omega\, Y_{\ell m}(\hat{\mathbf{n}}) Y_{\ell' m'}^*(\hat{\mathbf{n}}) = \delta_{\ell\ell'}^K \delta_{mm'}^K \qquad \text{Orthonormality} \tag{3.13}$$

$$\sum_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}) Y_{\ell m}^*(\hat{\mathbf{n}}') = \delta^D(\hat{\mathbf{n}} - \hat{\mathbf{n}}') \qquad \text{Completeness} \tag{3.14}$$

$$\sum_{m} Y_{\ell m}(\hat{\mathbf{n}}) Y_{\ell m}^*(\hat{\mathbf{n}}') = \frac{2\ell+1}{4\pi} P_\ell(\hat{\mathbf{n}} \cdot \hat{\mathbf{n}}') \qquad \text{Addition Theorem} \tag{3.15}$$

where $P_\ell(x)$ are the usual Legendre polynomials. The phase of $Y_{\ell m}$ can be chosen such that $Y_{\ell m}^* = (-1)^m Y_{\ell-m}$, so that for a real valued field we have $x_{\ell m}^* = (-1)^m x_{\ell-m}$.

Under the assumption of statistical isotropy, for a finite variance field, the mean of the spherical harmonic coefficients is $\langle x_{\ell m} \rangle = 0$, while their covariance is given by

$$\langle x_{\ell m} x_{\ell' m'}^* \rangle = C_\ell^{XX} \delta_{\ell\ell'}^K \delta_{mm'}^K, \tag{3.16}$$

---

[4]This especially applies when distance information about the sources is unavailable, nevertheless the quantity of interest can always be projected on the sphere, as is the case for tomographic analyses of LSS.

where $C_\ell^{XX}$ is the *angular power spectrum* of $X$ which is related to the angular 2PCF $C^{XX}(\theta)$ through (see Eq. (2.73) for the derivation)

$$C^{XX}(\theta) = \sum_\ell \frac{2\ell + 1}{4\pi} C_\ell^{TT} P_\ell(\hat{\mathbf{n}} \cdot \hat{\mathbf{n}}') \tag{3.17}$$

$$C_\ell^{XX} = 2\pi \int_{-1}^{1} \mathrm{d}\cos\theta \, C^{XX}(\theta) P_\ell(\cos\theta). \tag{3.18}$$

It is also possible to expand the product of two spherical harmonics in terms of spherical harmonics as

$$Y_{\ell m}(\hat{\mathbf{n}}) Y_{\ell'm'}(\hat{\mathbf{n}}) = \sum_{LM} \sqrt{\frac{(2\ell + 1)(2\ell' + 1)(2L + 1)}{4\pi}} \begin{pmatrix} \ell & \ell' & L \\ m & m' & M \end{pmatrix} \begin{pmatrix} \ell & \ell' & L \\ 0 & 0 & 0 \end{pmatrix} Y_{LM}^*(\hat{\mathbf{n}}), \tag{3.19}$$

where the strange quantities in parenthesis that we have introduced are called the Wigner $3j$ symbols (Varshalovich et al., 1988). Integrating the above expression over the whole sphere one finds the *Gaunt relation*:

$$\begin{aligned} \mathcal{G}_{mm'M}^{\ell\ell'L} &\equiv \int_{\mathbb{S}^2} \mathrm{d}\Omega \, Y_{\ell m}(\hat{\mathbf{n}}) Y_{\ell'm'}(\hat{\mathbf{n}}) Y_{LM}(\hat{\mathbf{n}}) \\ &= \sqrt{\frac{(2\ell + 1)(2\ell' + 1)(2L + 1)}{4\pi}} \begin{pmatrix} \ell & \ell' & L \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \ell & \ell' & L \\ m & m' & M \end{pmatrix}. \end{aligned} \tag{3.20}$$

Basically, Wigner $3j$ matrices are a rescaling of the Clebsch-Gordan coefficients and describe the quantum mechanical composition of two angular momentum eigenstates into a third one. The selection rules for Wigner $3j$ symbols, i.e. when they are non-null, are given by

$$\ell + \ell' + L \text{ is even,} \tag{3.21}$$

$$|\ell - \ell'| \leq L \leq \ell + \ell' \tag{3.22}$$

$$m + m' + M = 0. \tag{3.23}$$

These objects are somewhat nasty to calculate and are at the core of the Pseudo-$C_\ell$ method (PCL) methods to calculate the masking induced mode-coupling, as we will see in Sec. (3.2). Harmonic calculations on the sphere can be performed on the computer with standard packages for CMB data analysis such as `HEALPix`[5] (Gorski et al., 2005) and `S`[2]`HAT` (Stompor, 2011). The basic idea is to tesselate the sphere according to some pixelization scheme and to discretize the function into maps. As an example, for `HEALPix`, the resolution of the grid is controlled by the parameter $N_{\mathrm{side}} = 2^k$ (where $k \in \mathbb{N}$) which fixes the total number of pixels $N_{\mathrm{pix}} = 12N_{\mathrm{side}}^2$. In

---

[5] http://healpix.sourceforge.net

particular, spherical harmonic transforms in Eq. (3.12) are operatively estimated as

$$x_{\ell m} \approx \Omega_p \sum_p X(p) Y_{\ell m}^*(p), \qquad (3.24)$$

where $\Omega_p = \frac{4\pi}{N_{\text{pix}}}$ is the solid angle subtended by each pixel $p$. Another complication that arises from pixelization is that we only know the value of the field averaged over the pixel area: the effect is an amount of degradation of the information below a certain scale set by the size of the pixel (basically it acts as a low-pass filter). A way to account for this bias is to deconvolve the reconstructed spectrum for the *pixel window function* $p_\ell$ ($C_\ell^{\text{obs}} = p_\ell^2 C_\ell$).

Before concluding this subsection, let us take a look at the near future: most of the upcoming cosmological surveys will carry out large and deep observations of the sky, providing both angular and redshift information about the cosmic fields. Thus, a formalism for analyzing function on the 3D *ball* - a family of concentric spheres (shells) indexed by a continuos radial parameter, such as the redshift $z$ - will be required. Tomographical analyses (see Ch. (5)) represent a first step in this direction: the idea is to slice the data in redshift bins and then perform the usual 2D spherical analysis. A natural decomposition scheme for full-3D analyses is given by the spherical Fourier-Bessel formalism, while extensions to wavelets (Lanusse et al., 2012; Leistedt & McEwen, 2012) and radial 3D needlets (Durastanti et al., 2014) have also been investigated.

### 3.1.3 Projected fields

Let us now work out the connection between a 3D field $X^{3D}(\boldsymbol{x})$ and its 2D counterpart $X^{2D}(\hat{\mathbf{n}})$, projected on the sphere and integrated along the light-cone according to some arbitrary weight function $W^X(\chi)$:[6]

$$X^{2D}(\hat{\mathbf{n}}) = \int_0^\infty \mathrm{d}\chi\, W^X(\chi) X^{3D}(\chi\hat{\mathbf{n}}, \chi). \qquad (3.25)$$

Using the Rayleigh's expansion

$$e^{i\mathbf{k}\cdot\hat{\mathbf{n}}\chi(z)} = 4\pi \sum_{\ell m} i^\ell j_\ell(k\chi) Y_{\ell m}^*(\hat{\mathbf{k}}) Y_{\ell m}(\hat{\mathbf{n}}), \qquad (3.26)$$

---

[6]To avoid a crowding of equations we restrict to the case of a flat Universe, so that $f_K(\chi) = \chi$.

we can Fourier transform the 3D field $X^{3D}(\chi\hat{\mathbf{n}}, \chi)$ and obtain

$$
\begin{aligned}
X^{3D}(\chi\hat{\mathbf{n}}, \chi) &= \int \frac{\mathrm{d}^3 k}{(2\pi)^3} X^{3D}(\mathbf{k}, \chi) e^{i\mathbf{k}\cdot\hat{\mathbf{n}}\chi} \\
&= 4\pi \sum_{\ell m} i^\ell \int \frac{\mathrm{d}^3 k}{(2\pi)^3} X^{3D}(\mathbf{k}, \chi) j_\ell(k\chi) Y^*_{\ell m}(\hat{\mathbf{k}}) Y_{\ell m}(\hat{\mathbf{n}}).
\end{aligned}
\tag{3.27}
$$

In this way, the projected field becomes

$$
X^{2D}(\hat{\mathbf{n}}) = 4\pi \sum_{\ell m} i^\ell \int \mathrm{d}\chi W^X(\chi) \int \frac{\mathrm{d}^3 k}{(2\pi)^3} X^{3D}(\mathbf{k}, \chi) j_\ell(k\chi) Y^*_{\ell m}(\hat{\mathbf{k}}) Y_{\ell m}(\hat{\mathbf{n}}),
\tag{3.28}
$$

and the spherical coefficients read as

$$
x_{\ell m} = 4\pi i^\ell \int \mathrm{d}\chi W^X(\chi) \int \frac{\mathrm{d}^3 k}{(2\pi)^3} X^{3D}(\mathbf{k}, \chi) j_\ell(k\chi) Y^*_{\ell m}(\hat{\mathbf{k}}).
\tag{3.29}
$$

We calculate the angular cross power spectrum $C_\ell^{XY}$ by taking the ensemble average of $\langle x_{\ell m} y^*_{\ell m} \rangle$ as

$$
\begin{aligned}
C_\ell^{XY} &= (4\pi)^2 \int \mathrm{d}\chi W^X(\chi) \int \mathrm{d}\chi' W^X(\chi') \\
&\quad \times \int \frac{\mathrm{d}^3 k}{(2\pi)^3} \frac{\mathrm{d}^3 k'}{(2\pi)^3} \langle X^{3D}(\mathbf{k}, \chi) Y^{3D*}(\mathbf{k}', \chi') \rangle j_\ell(k\chi) j_{\ell'}(k'\chi') Y^*_{\ell m}(\hat{\mathbf{k}}) Y_{\ell'm'}(\hat{\mathbf{k}}') \\
&= (4\pi)^2 \int \mathrm{d}\chi W^X(\chi) \int \mathrm{d}\chi' W^Y(\chi') \int \frac{\mathrm{d}^3 k}{(2\pi)^3} P_{XY}(k; \chi, \chi') j_\ell(k\chi) j_{\ell'}(k\chi') Y^*_{\ell m}(\hat{\mathbf{k}}) Y_{\ell'm'}(\hat{\mathbf{k}}) \\
&= \frac{2}{\pi} \int \mathrm{d}\chi W^X(\chi) \int \mathrm{d}\chi' W^Y(\chi') \int k^2 \, \mathrm{d}k \, P_{XY}(k; \chi, \chi') j_\ell(k\chi) j_\ell(k\chi')
\end{aligned}
\tag{3.30}
$$

where we used the definition of (isotropic) 3D cross-power spectrum $\langle X(\mathbf{k}, \chi) Y^*(\mathbf{k}', \chi') \rangle = (2\pi)^3 \delta^D(\mathbf{k} - \mathbf{k}') P_{XY}(k; \chi, \chi')$ and the orthonormality condition of spherical harmonics. Summing up, the *exact* relation is

$$
C_\ell^{XY} = \frac{2}{\pi} \int \mathrm{d}\chi W^X(\chi) \int \mathrm{d}\chi' W^Y(\chi') \int k^2 \, \mathrm{d}k \, P_{XY}(k; \chi, \chi') j_\ell(k\chi) j_\ell(k\chi').
\tag{3.31}
$$

**Limber approximation**

The Limber approximation (Limber, 1953) is a handy tool whenever integrals of Bessel functions appear. When deriving it, one usually assumes small angular separations (or large multipoles $\ell$) and that some of the integrand functions vary slower with respect to others. The idea is that if $f(x)$ is a smooth function, then $\int_0^\infty \mathrm{d}x f(x) J_\nu(x) = f(\nu) + \mathcal{O}(1/\nu^2)$, where $\nu \equiv \ell + 1/2$; this can be written in the

following form

$$\frac{2}{\pi} \int k^2 \, \mathrm{d}k \, f(k) j_\ell(k\chi) j_\ell(k\chi') = \frac{\delta^D(\chi - \chi')}{\chi^2} f\left(\frac{\nu}{\chi}\right) \left[1 + \mathcal{O}\left(\frac{1}{\nu^2}\right)\right], \qquad (3.32)$$

where the presence of the $\delta^D$ tells us that the integrand is dominated by the region $\chi \approx \chi'$, and the goodness of the approximation gets better as $\ell \to \infty$, i.e. the corrections are $\mathcal{O}(1/\nu^2)$ (see LoVerde et al. (2007)). Substituting the above expression in Eq. (3.31) (and integrating out the $\delta^D$), one finds the Limber approximated angular power spectrum:

$$C_\ell^{XY} = \int_0^\infty \frac{\mathrm{d}\chi}{\chi^2} W^X(\chi) W^Y(\chi) P_{XY}\left(\frac{\ell + \frac{1}{2}}{\chi}, \chi\right). \qquad (3.33)$$

In literature the power spectrum in Eq. (3.33) is sometimes evaluated at $\nu = \ell$ (as will be the case in Ch. (4) and (5)), which increase the error from $\mathcal{O}(\ell^{-2})$ to $\mathcal{O}(\ell^{-1})$. The approximation is correct to $\mathcal{O}(1\%)$ for $\ell \gtrsim 10$. An extended Limber approximation, in the sense that exploits higher-order terms in the Bessel integral expansion, has been developed by LoVerde et al. (2007). For completeness, we write below the analogous of Eq. (3.33) with the integration in redshift space:[7]

$$C_\ell^{XY} = \int_0^\infty \frac{\mathrm{d}z}{c} \frac{H(z)}{\chi^2(z)} W^X(z) W^Y(z) P_{XY}\left(\frac{\ell + \frac{1}{2}}{\chi(z)}, z\right). \qquad (3.34)$$

Recall that the comoving distances $\chi$ should be replaced by $f_K(\chi)$ for non-flat cosmologies and that Limber approximation is valid as long as (i) the sources stretch in a *wide* range of distances $\Delta\chi$ and (ii) vary smoothly, i.e. $\ell/\chi \times \Delta\chi \gg 1$ (see Simon (2007); Bernardeau et al. (2011)).

## 3.2 Spectral estimation

The scientific exploitation of any cosmological dataset is just the tip of the iceberg: the data analysis that allows us to extract science from observations is a complex and iterative process which requires both computational and physical solid skills. Having in mind the CMB field, the reduction of a dataset usually can be split in the following sequence of steps:

---

[7]In order to switch the Dirac delta from $\chi$ to $z$ recall that $\delta^D[g(x)] = \frac{\delta^D(x-x_0)}{|g'(x_0)|}$ where $x_0$ is the root of $g(x)$. In our case $g(z) = \chi(z) - \chi(z') = \int_0^z \frac{c\,dz''}{H(z'')} - \int_0^{z'} \frac{c\,dz''}{H(z'')}$, so that $g'(z) = \frac{c}{H(z)}$ and $\delta^D[\chi(z) - \chi(z')] = \frac{H}{c}(z')\delta^D(z - z')$

- **Pre-processing**: the goal is to prepare real data so that can be fed to the subsequent pipeline steps. Here, the raw time-ordered data (TOD) collected by the detectors are calibrated and the time-domain systematics are generally flagged and/or removed.

- **Map-making**: after characterizing the instrument's noise specifics, the different frequency maps of the observed temperature and polarization are extracted from the cleaned TOD.

- **Component separation**: the available spectral information is used to isolate the emission from all other (astrophysical) components in the data, hence singling out the CMB signal. One of the outcome of this pipeline step is the production of ancillary data. Though not strictly correct, we can insert in this step the production of additional maps, such as the CMB gravitational lensing convergence, derived from the $T$ and $P$ separated maps.

- **Power spectrum estimation**: CMB temperature $T$ and polarization $E$- and $B$- modes auto- and cross-power spectra are reconstructed from the maps, along with spectra of derived fields such as lensing $\kappa$.

- **Parameter estimation**: the estimated set of CMB power spectra is compared with the theoretical models to study the degeneracies in the parameter space and infer best-fit parameters, usually by means of Bayesian analysis.

### 3.2.1 Power spectrum estimators

Throughout my thesis I have mainly been involved with the last two steps of the CMB pipeline, here I will focus on the power-spectrum estimation techniques: since the statistical properties of GRF are completely described by their power spectra, spectral estimation is a pivotal step.

There exist a variety of well-established methods to recover the underlying angular spectra from observations. In principle, optimal maximum likelihood algorithms can be applied to the spectral estimation problem, assuming that the pixel-pixel covariance matrix is known. The idea behind these methods is the following: considering a data vector $x_i$ (e.g. the CMB temperature in a given pixel) of length $N_d$ with the $x_i$ being Gaussian distributed, then the likelihood function reads as

$$\mathcal{L}(C_\ell|\boldsymbol{x}) = \frac{\exp\left(-\frac{1}{2}x^T C^{-1} x\right)}{\sqrt{(2\pi)^{N_d}\det(C)}}. \tag{3.35}$$

Here $C_\ell$ is the power spectrum to be estimated and the covariance matrix $C$ (in pixel space) can be written as the sum of signal $S_{ij}$ and noise $N_{ij}$ terms as

$C_{ij} = \langle x_i x_j \rangle = S_{ij}(C_\ell) + N_{ij}$. From this, maximum likelihood solution can be found iteratively with Newton-Raphson algorithm (Bond et al., 1998) or with quadratic maximum likelihood estimators (Tegmark, 1997). Despite being optimal, these methods are computational expensive and require $\mathcal{O}(N_{\mathrm{pix}}^3)$ CPU time (as well as a good knowledge of the covariance matrix), and become prohibitively for datasets with $N_{\mathrm{pix}} \gtrsim 10^5$.

As a result, PCL techniques have emerged on the market, see Yu & Peebles (1969); Hauser & Peebles (1973); Wandelt et al. (2001); Hivon et al. (2002); Efstathiou (2004). The idea is to split the estimation of $C_\ell$ in two steps: a biased estimate $\tilde{C}_\ell$ of the power spectrum is first obtained, and then corrected in the harmonic domain for all the different biasing effects such as the beam function, the pixelization, the mapmaking, and mode-mode coupling induced by masking. Schematically, the true power spectrum and the initial estimate are related by the mapping $\tilde{C}_\ell = f(C_\ell)$, where the function $f$ encodes all the biasing effects. If we linearly expand it, we can rewrite the previous relation in matrix notation as $\langle \tilde{C}_{\ell'} \rangle \approx \sum_\ell \alpha_{\ell' \ell} C_\ell + \beta_{\ell'}$, where $\{\alpha_{\ell' \ell}, \beta_{\ell'}\}$ are the coefficients that approximate the effect of $f$, and then invert it to recover the true spectrum as $\hat{C}_{\ell'} = \alpha_{\ell' \ell}^{-1}(\tilde{C}_\ell - \beta_\ell)$. The PCL method is very fast, requiring $\mathcal{O}(N_{\mathrm{pix}}^{3/2})$ CPU time and it is nearly-optimal for temperature (and scalar fields) in practice.[8] Let us see the details of the method (for spin-0 fields).

From a statistical point of view, an unbiased estimator of the (cross-)angular power spectrum is given by (hereafter the hat $\hat{X}$ denotes estimated quantities):

$$\hat{C}_\ell^{XY} = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} x_{\ell m} y_{\ell m}^*. \tag{3.36}$$

Spherical harmonics are particularly appealing because they are statistically orthogonal for full-sky Gaussian-distributed sky-maps, i.e. the covariance is diagonal $\mathrm{Cov}_{\ell \ell'} \propto \delta_{\ell \ell'}^K$, and the power spectrum fully characterizes the behavior of the field. However, real-world observations have to deal with a number of limitations and issues, such as the finite instrumental spatial resolution, the anisotropic noise, and asymmetric beam response. Moreover the incomplete sky coverage, motivated for example by foreground contamination or the instruments scanning strategy, induces a mode-coupling and a power leakage between different multipoles, as well as an overall downward shift of power (Hivon et al., 2002; Efstathiou, 2004). Consider a position dependent weighting scheme $W(\hat{\mathbf{n}})$, i.e. the mask, whose harmonic expansion is $W(\hat{\mathbf{n}}) = \sum_{\ell m} w_{\ell m} Y_{\ell m}^*(\hat{\mathbf{n}})$ and its power spectrum is $\mathcal{W}_\ell$. Then the

---

[8]The problem with the estimation of polarized spectra is the spurious mixing of $E$- and $B$-modes ($E$-to-$B$ leakage) that has to be taken into account. Several methods have been proposed to treat such effect, see Lewis et al. (2002); Bunn et al. (2003); Grain et al. (2009).

pseudo-$x_{\ell m}$ of the masked field $\tilde{X}(\hat{\mathbf{n}}) = X(\hat{\mathbf{n}})W(\hat{\mathbf{n}})$ are given by:

$$
\begin{aligned}
\tilde{x}_{\ell m} &= \int_{\mathbb{S}^2} X(\hat{\mathbf{n}})W(\hat{\mathbf{n}})Y^*_{\ell m}(\hat{\mathbf{n}})d\Omega \\
&= \sum_{\ell'm'} x_{\ell'm'} \int d\Omega \, Y_{\ell'm'}(\hat{\mathbf{n}})W(\hat{\mathbf{n}})Y^*_{\ell m}(\hat{\mathbf{n}}) \\
&= \sum_{\ell'm'} K_{\ell m \ell'm'}[W]x_{\ell'm'},
\end{aligned}
\tag{3.37}
$$

where the kernel $K$ that describes the induced mode-coupling reads as

$$
K_{\ell m \ell'm'}[W] \equiv (-1)^m \sum_{\ell''m''} w_{\ell''m''} \mathcal{G}^{\ell \, \ell' \, \ell''}_{m \, m' \, m''}.
\tag{3.38}
$$

The well known MASTER ([Hivon et al., 2002]) (or XSPECT, the cross-spectra extension by [Tristram et al. (2005)]) approach to obtain unbiased but slightly sub-optimal multipole estimates relates the pseudo-spectrum

$$
\tilde{C}^{XY}_{\ell} = \frac{1}{2\ell + 1} \sum_m \tilde{x}_{\ell m} \tilde{y}^*_{\ell m}
\tag{3.39}
$$

to the underlying power spectrum $C_\ell$ as

$$
\langle \tilde{C}^{XY}_{\ell} \rangle = \sum_{\ell'} M_{\ell\ell'} C^{XY}_{\ell'},
\tag{3.40}
$$

where $M_{\ell\ell'}$ is the *coupling matrix*

$$
M_{\ell\ell'} = \frac{2\ell' + 1}{4\pi} \sum_{\ell''} (2\ell'' + 1)\mathcal{W}_{\ell''} \begin{pmatrix} \ell & \ell' & \ell'' \\ 0 & 0 & 0 \end{pmatrix}^2
\tag{3.41}
$$

The finite size of the observed sky patch (of area $A_{\rm obs}$) approximately determines the smaller multipole that can be recovered, $\ell_{\min} \approx \pi/\sqrt{A_{\rm obs}}$, and the width of $M_{\ell\ell'}$, that characterizes the amount of mode-coupling. The basic idea is to invert eq. (3.40) in order to recover the underlying power spectrum, however for small sky fraction $f_{\rm sky} = \frac{1}{4\pi}\int_{\mathbb{S}^2} W^2(\hat{\mathbf{n}})d\Omega$, one needs to bin the pseudo-power spectrum and the coupling matrix, so that the estimator of the true cross-bandpowers $\hat{C}^{XY}_L$ writes

$$
\hat{C}^{XY}_L = \sum_{L'\ell} K^{-1}_{LL'} P_{L'\ell} \tilde{C}^{XY}_{\ell},
\tag{3.42}
$$

where $L$ is the bandpower index and the binned coupling matrix can be written as

$$
K_{LL'} = \sum_{\ell\ell'} P_{L\ell} M_{\ell\ell'} B^X_{\ell'} B^Y_{\ell'} p^2_{\ell'} F_{\ell'} Q_{\ell'L'}.
\tag{3.43}
$$

Here $P_{L\ell}$ is the binning operator and $Q_{\ell L}$ is its reciprocal. To correct also for angular resolution, finite pixel size and filtering applied to TOD, we added to the binned coupling matrix the beam function (for the observed field $X$) $B_{\ell'}^X$, the pixel window function $p_\ell$, and the effective filtering function $F_{\ell'}$. When dealing with auto-power spectra of maps comprehensive of noise, Eq. (3.42) gets modified as

$$\hat{C}_L^{XX} = \sum_{L'\ell} K_{LL'}^{-1} P_{L'\ell}(\tilde{C}_\ell^{XX} - \langle \tilde{N}_\ell^{XX} \rangle_{\mathrm{MC}}), \qquad (3.44)$$

where $\langle \tilde{N}_\ell^{XX} \rangle_{\mathrm{MC}}$ is the MC estimate of the average noise pseudo-power spectrum. If the true power spectrum varies slowly with respect to the coupling matrix and/or $f_{\mathrm{sky}}$ is large, eq. (3.40) becomes

$$\langle \tilde{C}_\ell^{XY} \rangle \approx C_\ell^{XY} \sum_{\ell'} M_{\ell\ell'} = f_{\mathrm{sky}} C_\ell^{XY}, \qquad (3.45)$$

which is the so-called $f_{\mathrm{sky}}$ approximation (Komatsu et al., 2002). For the analyses discussed in Ch (4) and (5) and performed during my Ph.D., I have developed a code to compute coupling matrices and estimate power spectra of spin-0 masked maps. The computational bottleneck is represented by the calculation of the Wigner 3$j$ symbols: to this end, I have wrapped the FORTRAN rc3jj.f routine from the **SLATEC** library[9] into flexible Python code which calculates the coupling matrix $M_{\ell\ell'}$, as well as other quantities, and can be efficiently interfaced with usual **HEALPix**-based codes (such as the Python implementation called `healpy`[10]). Just as an illustration and example of how a coupling matrix can appear, we anticipate here the mapping in the case of the H-ATLAS dataset considered in this analysis. In the left panel of Fig. (3.1) we show the *unbinned* coupling matrix $M_{\ell\ell'}$ computed for the H-ATLAS mask (shown in the right panel of Fig. (6.2)) that tells us how the different harmonic modes are coupled to each other. As previously mentioned, the width of the diagonal is roughly determined by the sky coverage and mask topology: to highlight this aspect, we plot slices through $M_{\ell\ell'}$ (at fixed $\ell'$) in the right panel of Fig. (3.1).

### 3.2.2 Covariance estimators

An estimate of the power spectrum means almost nothing without the associated covariance matrix. Error bars must be assigned to asses the quality or the significance of measurements and, for example, the full covariance matrix is needed to

---

[9] http://www.netlib.org/slatec/
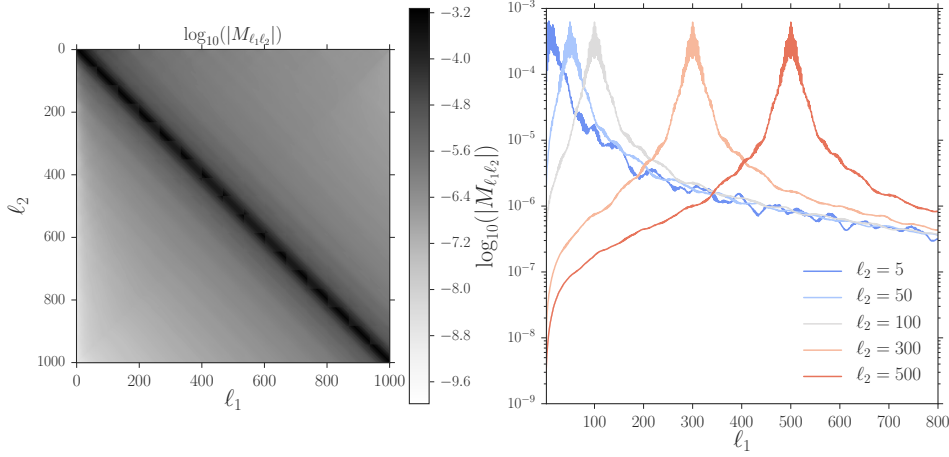[10] https://healpy.readthedocs.io/en/latest/

**Figure 3.1:** *Left panel:* Coupling matrix $M_{\ell\ell'}$ evaluated for the H-ATLAS mask shown in Fig. (6.2). The main diagonal structure, whose width determines the extent of the mode-coupling, is visible. *Right panel:* Slices through the coupling matrix shown in the left panel.

perform the further pipeline steps such as the parameter estimation. The covariance of the power spectrum estimator $\hat{C}_\ell^{XX}$ defined in Eq. (3.36) is:[11]

$$
\begin{aligned}
\mathrm{Cov}(\hat{C}_\ell^{XY}, \hat{C}_{\ell'}^{XY}) &= \langle (\hat{C}_\ell^{XY} - C_\ell^{XY})(\hat{C}_{\ell'}^{XY} - C_{\ell'}^{XY}) \rangle \\
&= \frac{1}{(2\ell+1)(2\ell'+1)} \sum_{mm'} \langle x_{\ell m} y_{\ell m}^* x_{\ell' m'} y_{\ell' m'}^* \rangle - C_\ell^{XY} C_{\ell'}^{XY} - \cancel{C_\ell^{XY} C_{\ell'}^{XY}} \\
&\quad + \cancel{C_\ell^{XY} C_{\ell'}^{XY}} \\
&= \frac{1}{(2\ell+1)(2\ell'+1)} \sum_{mm'} \Big[ \langle x_{\ell m} y_{\ell m}^* \rangle \langle x_{\ell' m'} y_{\ell' m'}^* \rangle + \langle x_{\ell m} x_{\ell' m'} \rangle \langle y_{\ell m}^* y_{\ell' m'}^* \rangle \\
&\qquad\qquad\qquad\qquad + \langle x_{\ell m} y_{\ell' m'}^* \rangle \langle x_{\ell' m'}^* y_{\ell' m'}^* \rangle \Big] - C_\ell^{XY} C_{\ell'}^{XY} \\
&= \cancel{C_\ell^{XY} C_{\ell'}^{XY}} + \frac{C_\ell^{XX} C_{\ell'}^{YY}}{2\ell+1} \delta_{\ell\ell'}^K + \frac{C_\ell^{XY} C_{\ell'}^{XY}}{2\ell+1} \delta_{\ell\ell'}^K - \cancel{C_\ell^{XY} C_{\ell'}^{XY}} \\
&= \frac{\delta_{\ell\ell'}^K}{2\ell+1} \left[ C_\ell^{XY} C_{\ell'}^{XY} + C_\ell^{XX} C_{\ell'}^{YY} \right] .
\end{aligned}
$$
$$(3.46)$$

Then, the cross- and auto-power spectrum variance associated to a given multipole are respectively

$$(\Delta C_\ell^{XY})^2 = \frac{1}{2\ell+1} \left[ (C_\ell^{XY})^2 + C_\ell^{XX} C_\ell^{YY} \right] \tag{3.47}$$

$$(\Delta C_\ell^{XX})^2 = \frac{2}{2\ell+1} (C_\ell^{XX})^2, \tag{3.48}$$

---

[11] Assuming that the fields are Gaussian, hence neglecting the connected part $\langle x_{\ell m} y_{\ell m}^* x_{\ell' m'} y_{\ell' m'}^* \rangle_c$.

with the last one being the same as Eq. (2.72). Let us make a couple of remarks. First of all, if the maps being analyzed are comprehensive of noise (as in the real-world), then the auto-power spectra appearing in the above equations must include a noise term, i.e. $C_\ell^{XX} \to C_\ell^{XX} + N_\ell^{XX}$. Secondly, in the incomplete sky case, things become much more nastier and a *full* analytical treatment is virtually unfeasible. Several authors (Efstathiou, 2004; Tristram et al., 2005; Brown et al., 2005) have proposed analytic approximations if the underlying power spectra meet certain criteria, for example if they vary slow with respect to the coupling matrix. In the *zeroth-order* approximate formula for power spectrum error bars, known as the Knox formula (Knox, 1995), the fractional error on $C_\ell$ typically scales as $1/\sqrt{f_{\mathrm{sky}}}$. Although it is helpful for an assessment of uncertainties, it is known to underestimate errors and does not account for the mode-mode coupling and the $E$-to-$B$ mixing. We will discuss in greater detail these approximations and compare different covariance estimators for cross-power spectra in Ch. (4). The covariance can also be estimated numerically by means of MC simulations with resampling techniques. There exist different ways to calculate the covariance, each with their respective advantages and drawbacks. Here we list the most common ones adopted in literature for cross-correlation analyses, mainly developed for iSW studies, see Cabre et al. (2007); Giannantonio et al. (2008) for thorough discussions.

- **MC method 1**: Perhaps the most used estimator in literature, it consists in measuring the cross-power spectrum (or 2PCF) between $N_{\mathrm{sim}}$ random maps of field $X$, obtained from a fiducial model, and the observed maps of the field $Y$. This method is fast and straightforward to implement. *Issues*: it is model-dependent (like all MC approaches), it does not account for the variance in the $Y$ maps (this is called the *realization bias* since we have only one realization of the $Y$ field), and it assumes no cross-correlation between the fields (the so-called *correlation bias*, but if the expected signal is weak, it should not represent a large bias).

- **MC method 2**: It tries to improve MC1 by measuring the cross-power spectrum between random (MC generated) maps of the $X$ field (from the underlying model) and random $Y$ field maps. *Issues*: it is somewhat more time consuming and still model-dependent but has no dependence on any observed maps unlike the MC1.

- **Jack-knife (JK) method**: The idea is to divide the $Y$ field (e.g., the LSS tracer map) into $M$ patches to create $M$ subsamples (of approximately the same area) by neglecting each patch in turn and evaluating the covariance between them. *Issues*: it underestimates the error, results are dependent on the size and number of discarded patches, it assumes independence of

different patches (not always the case), but it is model-independent.

The mock maps used to estimate covariances can either be obtained through `HEALPix` via the synfast routine, or from N-body simulations. But what are the errors on the errors? The number of simulations exploited is important in determining the covariance matrix estimation convergence: Taylor et al. (2013) have shown that the fractional noise in a covariance matrix of $N_d \times N_d$ elements, obtained from $N_{\rm sim}$ realizations, can be estimated as $\sqrt{2/(N_{\rm sim} - N_d - 4)}$. To numerically estimate the covariance with MC1 and MC2 methods, one resorts to the sample covariance matrix:

$$\mathrm{Cov}_{LL'}^{XY} = \frac{1}{N_{\rm sim} - 1} \sum_{i=1}^{N_{\rm sim}} (\hat{C}_L^{XY,i} - \langle \hat{C}_L^{XY} \rangle_{\rm MC})(\hat{C}_{L'}^{XY,i} - \langle \hat{C}_{L'}^{XY} \rangle_{\rm MC}), \qquad (3.49)$$

while for the JK method the estimator becomes

$$\mathrm{Cov}_{LL'}^{XY} = \frac{M-1}{M} \sum_{i=1}^{M} (\hat{C}_L^{XY,i} - \langle \hat{C}_L^{XY} \rangle)(\hat{C}_{L'}^{XY,i} - \langle \hat{C}_{L'}^{XY} \rangle). \qquad (3.50)$$

## 3.3 Datasets

As anticipated at the beginning of this Chapter, we dedicate the remaining part to provide a brief and introductory description of the most important datasets that we exploit in this thesis. The data, and their combination, will be treated and discussed quantitatively in the following chapters.

### 3.3.1 Planck

The *Planck* satellite is a European Space Agency (ESA) mission devised to the measurement of CMB anisotropies and it represents the third generation of spatial missions devoted to CMB physics, after COBE and WMAP. It was launched on 14th May 2009 (together with its companion ESA's *Herschel* satellite) towards the second Lagrangian point L2 and has collected data until October 2013. The *nominal mission* data, released in 2013, consists in 15 months of temperature-only observations, while the *full mission* data comprehends all the 30 months of the High Frequency Instrument (see below) temperature and polarization data.

The *Planck* telescope is an off-axis tilted Gregorian design, with the optical system composed by a primary mirror of about 1.5 m and the secondary, which focuses the radiation to the detectors, of about 1 m. Their operative temperature is

approximately 45 K. *Planck*'s focal plane contains two separate scientific instruments: the Low Frequency Instrument (LFI) (Bersanelli et al., 2010), an array of radiometers covering three bands centered in 30, 44, and 70 GHz, and the High Frequency Instrument (HFI) (Lamarre et al., 2010), and array of microwave spider-web bolometers operating at higher frequencies from 100 GHz to 857 GHz. Corrugated horns serve as wave-guides to collect the radiation to the instruments detectors. Below we provide a brief description of the two intstruments:

- **HFI**: it observes the sky in six frequency bands centered at 100, 143, 217, 353, 545, and 857 GHz, thus allowing for a characterization of the cosmological CMB signal and to study both Galactic and extragalactic foregrounds. HFI detectors are bolometers that, depending on the placement on the grid wires, can be sensitive only to the intensity of the incoming radiation or to polarization as well. The former are called spider-web bolometers and collect the CMB radiation through a spider-web like grid (to enhance sensitivity, robustness to vibrations and to reduce the cosmic rays cross-section), while the latter are called polarization sensitive bolometers. However, bare in mind that only the four lower frequencies are polarization sensitive. The angular resolution (called *beam*) of the different channels depend on the whole optic chain and can be estimated from observations of planets (like Mars, Jupiter, and Saturn): for HFI the effective beams vary between 10' and 5'.

- **LFI**: it is designed to measure the microwave sky in three bands centered at 30, 44, and 70 GHz. The instrument is composed by differential radiometers in the same fashion of COBE and WMAP, though they represent a major step forward in terms of performances. LFI horns are displaced in the focal plane around the HFI bolometers because since they operate at larger wavelengths, they suffer less from optical aberrations. The number of radiometers is 22, all of them being sensitive to polarization. Their angular resolution is about 33', 28', and 13' at 30, 44, and 70 GHz respectively. Data gathered by LFI are particularly important for monitoring the low frequency Galactic foregrounds, the large scale CMB power spectrum and in particular polarization data can be exploited to study reionization.

To achieve the required sensitivity needed to map the microwave sky at high resolution, the noise level must be suppressed by cooling the instruments down to cryogenic temperatures. In particular, LFI detectors require an operative temperature of about 20 K, while the HFI ones have to be cooled down to 0.1 K: this is achieved with a cooling chain that mixes passive and active cooling system. Schematically, the (active) cryogenic system comprises three different coolers that allow to reach the desired temperature: (i) an hydrogen *sorption cooler* that cools

the LFI focal plane to 20 K and provides a pre-cooling to HFI, (ii) a *Joule-Thomson cooler* based on $^4$He which cools the HFI focal plane down to 4 K, and (iii) a $^3$He/$^4$He *dilution cooler* which is able to bring the temperature of the HFI focal plane down to 100 mK.

*Planck*'s scanning strategy has been designed to optimize the sky coverage, the data redundancy and to perform polarization measurements. *Planck* spins at 1 rpm and its spin axis is approximately oriented with the Sun-L2 axis, so that its solar panels always face the Sun, and to minimize straylight and thermal radiation. The LOS of the telescope is tilted of about 85 deg with respect to the spin axis. What is called *survey* is an almost complete scan of the sky that is performed in 6 months: the full mission data consists of $\sim 5$ surveys.

The broad spectral coverage of *Planck* allows it to characterize and separate the *diffuse* foregrounds (Planck Collaboration, 2015b). In particular, six types have been investigated in temperature: synchrotron emission; free-free emission; Galactic dust thermal emission; dust anomalous emission (from spinning dust); three carbon-monoxide (CO) rotational lines; thermal Sunyaev-Zel'dovich. Instead, the sky components analyzed in polarization, in addition to CMB, are the synchrotron and thermal dust emissions. Foreground maps are obtained using CMB component separation techniques: the four methods exploited in *Planck* are described in Planck Collaboration et al. (2015) and can be divided into two types. The first kind of methods, by only assuming the knowledge of the blackbody spectrum of CMB, remove foregrounds using a combination of multiband data that minimizes the variance of CMB signal. Instead, the other approach exploits an explicit parametrized model of the CMB and foregrounds with their associated likelihoods, and extracts the CMB component by sampling from the posterior distribution of parameters. Note that these two flavors can be implemented both in real and harmonic space. In particular, *Planck* team relied on four algorithms, namely `Commander`, Needlet Internal Linear Combination (`NILC`), Spectral Estimation Via Expectation Maximization (`SEVEM`), and Spectral Matching Independent Component Analysis (`SMICA`), which incorporate the main approaches to component separation. The challenge represented by the treatment of intense, nonlinear and non-Gaussian processes in order to achieve the separation of them from each other, as well as from CMB, motivate a complementarity approach to the problem.

As mentioned earlier, when the CMB map is derived (either directly from the frequency channels or through some component separation method), another branch of the CMB pipeline consists in the extraction of the lensing information, specifically the lensing potnetial. The main *Planck* product exploited in the analyses presented in this thesis is the CMB gravitational lensing potential map,

whose production pipeline is discussed in Planck Collaboration (2014c) and Planck Collaboration (2016a) for the 2013 and 2015 releases respectively. We used both the publicly released *Planck* CMB lensing potential maps for our analyses, here we discuss the main differences between the two data deliveries. The 2013 lensing maps are extracted from the first 15.5 months of observations by applying a (full-sky) quadratic estimator (Hu & Okamoto, 2002) to the 100, 143 and 217 GHz frequency channels[12], which are the most suitable for estimation of the gravitational lensing potential. Nevertheless, the released map is provided as a filtered lensing potential $\phi$ map based on a minimum variance combination of the 143 and 217 GHz temperature anisotropy maps only, because adding the 100 GHz map yields a negligible improvement (Planck Collaboration, 2014c). In addition to the 143 and 217 GHz maps, the 857 GHz *Planck* map is also used as a dust template to project out the diffuse Galactic dust contamination (as well as a part of the CIB fluctuations). The mask associated to this release is obtained by combining three different masks: (i) a galaxy mask based on the temperature analysis one (which is constructed using a combination of 30 and 353 GHz maps) that avoids most of the Galactic foreground power, (ii) a CO and extended-object masks that removes regions believed to be contaminated by CO lines as well as extended nearby objects (such as the two Magellanic clouds and local galaxies), and a (iii) point-source mask that removes the compact objects identified in the *Planck* Early Release Compact Source Catalogue (ERCSC), in the *Planck* SZ clusters (PCC), and in the *Planck* Catalogue of Compact Sources (PCCS).

On the other hand, the publicly released 2015 *Planck* CMB lensing map (Planck Collaboration, 2016a) shown in Fig. (3.2) has been extracted from foreground-cleaned temperature and polarization maps, using the Hu & Okamoto estimator as well. Differently from the 2013 release, these maps have been synthesized from the raw 2015 *Planck* full mission frequency maps using the `SMICA` code (Planck Collaboration et al., 2015). In particular, the released map is based on a minimum-variance combination of *all* five temperature and polarization estimators, and is provided as a mean-field bias subtracted convergence $\kappa$ map rather than the lensing potential $\phi$. The lensing mask associated to the 2015 release covers a slightly larger portion of the sky with respect to the 2013 release: $f_{sky}^{2015}/f_{sky}^{2013} \simeq 0.98$. It is worth to notice that the 2015 mask covers tSZ clusters, whereas the 2013 mask does not. In particular, the 2013 reconstruction masks tSZ clusters in the 143 GHz channel but not in the 217 GHz and since it is a minimum-variance combination, there is still signal in the position of tSZ clusters. By being obtained from `SMICA` maps, in

---

[12]With angular resolution of 10′, 7′, and 5′ and noise levels of 105, 45 and 60 $\mu$K arcmin respectively.
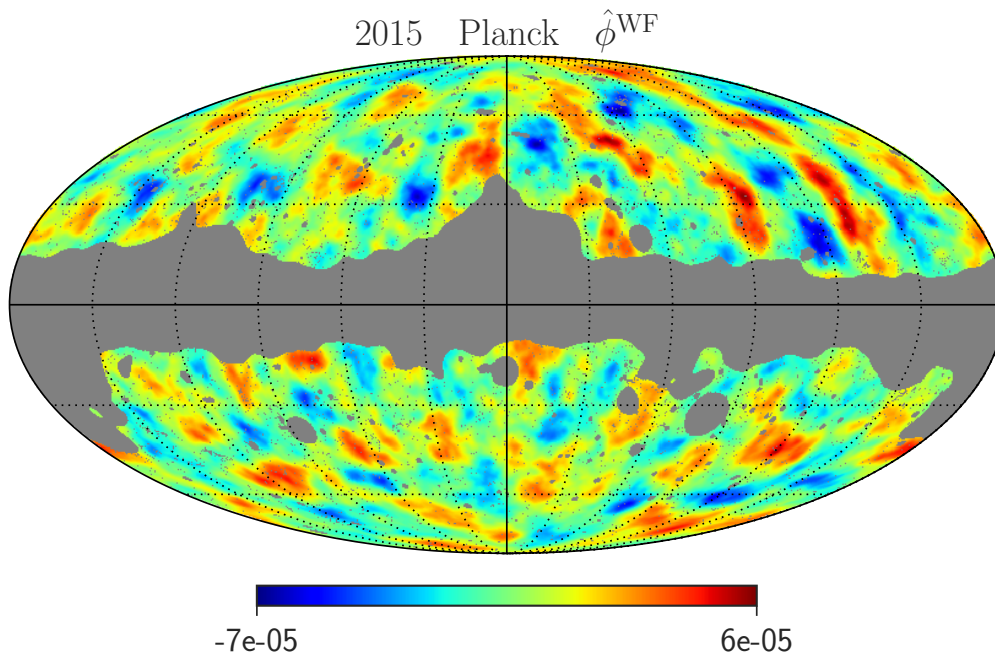
**Figure 3.2:** CMB lensing potential estimated from `SMICA` full-mission $T$ and $P$ maps using the MV estimator. The estimate has been Wiener-filtered as $\hat{\phi}_{\ell m}^{WF} = \frac{C_\ell^{\phi\phi,\text{fid}}}{C_\ell^{\phi\phi,\text{fid}}+N_\ell^{\phi\phi}}\hat{\phi}_{\ell m}$ for a better visual illustration and the map resolution is $N_{\text{side}} = 512$.

the 2015 release the tSZ clusters are masked prior to the reconstruction.

The maps from both data releases are in the `HEALPix` format with a resolution parameter of $N_{\text{side}} = 2048$, corresponding to 50331648 pixels over the sky, with a pixel size of $\sim 1.7'$. In Fig. (3.3) we show the CMB lensing reconstruction noise: the flat part at large scales comes from the contribution of the statistical noise, while the rise at smaller scales represents the limitations due to the instrument angular resolution and noise level. As can be seen, the exploitation of the full-mission temperature and the inclusion of polarization data have the effect of augmenting the *Planck* lensing reconstruction sensitivity by approximately a factor of two. Roughly half of the improvement comes from the reduced noise levels in temperature, while the other half comes from the additional polarization data.

## 3.3.2 Herschel Space Observatory

The *Herschel* Space Observatory (Pilbratt et al., 2010) was launched on the 14th May 2009 together with the *Planck* satellite and represented a huge leap forward in the field of sub-mm/Far-infrared (FIR) astronomy. Before *Herschel*'s arrival, all of the observations were severely limited to small patches, poor angular resolution and a restricted wavelength range was available. However, with its 3.5 m primary
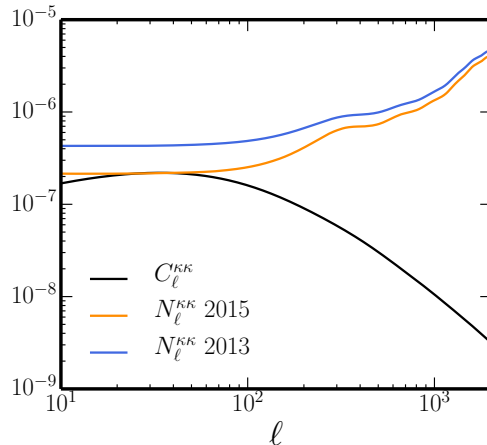
**Figure 3.3:** Lens reconstruction noise levels $N_\ell^{\kappa\kappa}$ for the 2013 (blue solid line) and 2015 (orange solid line) *Planck* releases. The fiducial $\Lambda$CDM CMB convergence power spectrum is plotted as the black solid line.

mirror[13], which is the largest one currently in space, *Herschel* has been able to pierce the deep Universe and to observe distant sources. The wavelength coverage of *Herschel* is from 60 to 680 $\mu$m (roughly corresponding to 440 - 5000 Ghz) at diffraction limited resolution, covering most of the dust emission of a typical galactic SED. Like *Planck*, the satellite has been placed in the second Lagrangian point to ensure that both the Sun and the Earth were always close to each other, thus limiting contaminations and increasing the available field of view. The detectors on board of *Herschel*, PACS and SPIRE, were able to reach high sensitivities by being cooled down to 0.3 K with liquid $^3$He, while most of remaining parts were cooled to 4 - 10 K. The amount of available liquid helium sets the lifetime of the telescope: for *Herschel*, observations have been carried out for approximately 3.5 years. The pointing of the telescope is performed through a combination of gyroscopes and star-tracker cameras: since the launch, the absolute accuracy has been found to be 2" (Pilbratt et al., 2010).

The telescope carried three main instruments, which are a mixture of broadband photometers and spectrometers: below we provide a brief description.

- **SPIRE:** The Spectral and Photometric Imaging Receiver (SPIRE) (Griffin et al., 2010) consisted of an imaging photometer and medium resolution spectrometer. It carried out observation at 250, 350, and 500 $\mu$m band simultaneously (each with a resolution of about $\lambda/\Delta\lambda \sim 3$). The detectors were feedhorn-coupled bolometer arrays with 139, 88, and 43 spider-web bolome-

---

[13]This is the largest size that could fit within the limits of the Arianne 5 rocket used to ship the satellite into space; in comparison the mirror of *Spitzer*, the next largest FIR telescope, is about 0.85m.

ters for the three bands respectively. The high sensitivity of SPIRE meant that observations quickly reached the confusion limit, with $1\sigma$ sensitivities estimated to be around 5.8, 6.3 and 6.8 mJy/beam for the three channels respectively. On the other hand, the SPIRE spectrometer was a medium resolution ($\lambda/\Delta\lambda \sim 1300$ at 200 $\mu$m and $\sim 370$ at 670 $\mu$m) imaging Fourier transform spectrometer covering 194 - 672 $\mu$m. Science-wise, SPIRE was able to shed light on the sub-mm region least explored by previous instruments, thus allowing an investigation of the cold and dusty Universe.

- **PACS**: The Photodetector Array Camera and Spectrometer (PACS) (Poglitsch et al., 2010) was a photometer and medium resolution spectrometer operating in the 60 - 210 $\mu$m regime. The photometer performed observations in three bands centered in 70, 100, 160 $\mu$m (with $\lambda/\Delta\lambda \sim 2$) and could observe either the 70 or 100 $\mu$m channel simultaneously with the 160 $\mu$m band. The mounted spectrometer was a medium resolution spectrometer ($\lambda/\Delta\lambda \sim 1000 - 4000$) covering the wavelengths range between 55 and 210 $\mu$m. PACS provided photometric observations of the peak of galactic dust emission in the nearby Universe at high spatial resolution; its main scientific goals were to map the spectral lines in the FIR, in the Milky Way and in local galaxies.

- **HIFI**: The *Herschel*-Heterodyne Instrument for the Far-Infrared (HIFI) was the high resolution spectrometer operating in the broad wavelengths range between 157 and 625 $\mu$m. Unlike SPIRE and PACS, it only had a single pixel but maps could be made using either a series of pointing or on-the-fly mapping. The main scientific target of the instrument was to investigate the interaction between stars and the interstellar medium in galaxies by searching for molecular rotational lines.

Several surveys have been carried out with the *Herschel* telescope. In this thesis we have made use of data gathered in the context of the H-ATLAS (Eales et al., 2010), here we give a brief overview.

The H-ATLAS is the largest extragalactic key-project carried out in open time with the *Herschel* Space Observatory. It was allocated 600 hours of observing time and covers about $600 \, \text{deg}^2$ of sky in five photometric bands (100, 160, 250, 350 and 500 $\mu$m) with the PACS and SPIRE instruments. The H-ATLAS map-making is described by Pascale et al. (2011) for SPIRE and by Ibar et al. (2010) for PACS. The procedures for source extraction and catalogue generation can be found in Rigby et al. (2011) and Valiante et al. (2016).

Since H-ATLAS is an extragalactic survey, the observed fields were chosen to minimize the Galactic dust contamination. As such, the survey area is divided

into five fields: three equatorial fields centered on 9hr, 12hr, and 14.5hr (Galaxy And Mass Assembly (GAMA) fields, G09, G12, and G15) covering, altogether, $161 \deg^2$; the North Galactic Pole (NGP) block, a rectangular block of $15° \cos(\delta)$ by $10°$ centered on right ascension $\alpha = 199.5°$, declination $\delta = 29°$ and rotated by approximately $8°$ clockwise; and the South Galactic Pole (SGP) block consisting of two concatenated rectangular regions, one of $31.5° \cos(\delta)$ by $6°$ centered on $\alpha = 351.3°$, $\delta = -32.8°$, the other of $20° \cos(\delta)$ by $6°$ centered on $\alpha = 18.1°$, $\delta = -30.7°$. The fields were also chosen depending on the amount of complementarity data available from external surveys operating at different frequencies. For example, spectroscopy is covered by GAMA, Sloan Digital Sky Survey (SDSS), and 2dF Galaxy Redshift Survey (2dFRGS), while much of the area has been covered by Galaxy Evolution Explorer (GALEX) in the ultraviolet. As concerns optical wavelengths, the SDSS has covered both the GAMA and NGP fields in five bands, while GAMA and SGP fields have been observed by the Kilo Degree Survey (KIDS); the SGP will be eventually covered by the Dark Energy Survey (DES).

The scientific targets of H-ATLAS are multiple, concerning both the local and the distant Universe. In particular, a huge number of the discovered objects do not belong to the local Universe: *Herschel* is able to resolve the CIB into discrete sources, thus enabling an investigation of the LSS of the FIR Universe. The $z \lesssim 1$ galaxies detected by the H-ATLAS survey are mostly late-type and starburst galaxies with moderate star-formation rates and relatively weak clustering (Dunne et al., 2011; Guo et al., 2011). High-$z$ galaxies are forming stars at high rates ($\geq$ few hundred $M_\odot \, yr^{-1}$) and are much more strongly clustered (Maddox et al., 2010; Xia et al., 2012), implying that they are tracers of large-scale overdensities. Their properties are consistent with them being the progenitors of local massive elliptical galaxies (Lapi et al., 2011). These DSFG contain a substantial amount of dust and as a consequence, their rest-frame optical/ultraviolet (UV) light can be strongly obscured. It is this high-$z$ H-ATLAS galaxy population that we aim to correlate with the *Planck* CMB lensing map in Ch. (4) and (5). In order to select the these galaxies into a catalogue and, as an example, slice them in bins for a tomographical analysis, the knowledge of their redshift must be at hand. Unfortunately, redshift acquisition is not straightforward for dusty IR selected objects: this is because most of the emission-line redshift indicators lie in the rest-frame optical and UV bands that are severely extinct by dust, and because IR observations are carried out with large beam sizes that make multiband counterpart identification ambiguous. A common method to determine the millimetric photo-$z$, is to perform a SED fitting in the FIR, where redshifts are estimated from the shape of the FIR/sub-mm SED or its colors rather than properties indicated by
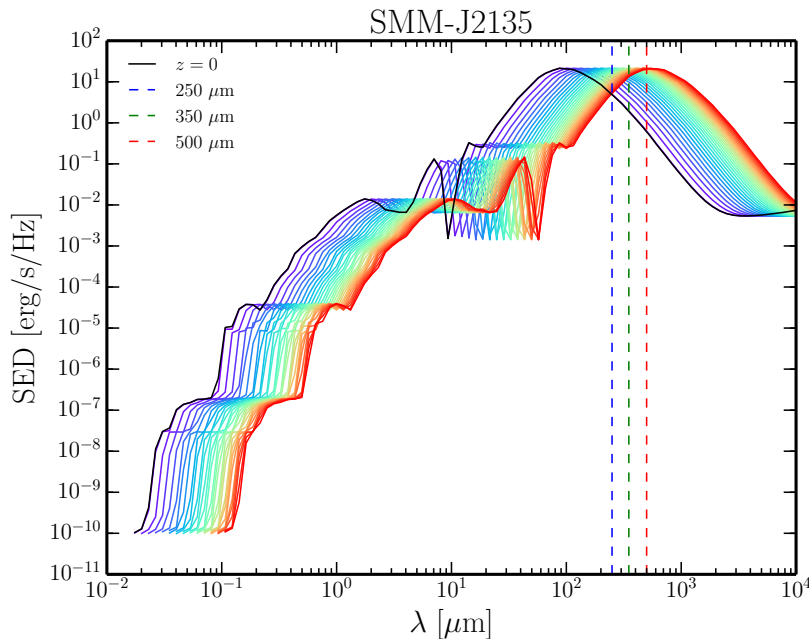
**Figure 3.4:** SMM SED template at $z = 0$ (black solid line): solid colored lines from purpleish to reddish represent rest-frame SED *redshifted* in the range $0 < z \leq 5$. Colored dashed lines mark the wavelengths observed by the SPIRE photometer instrument on board of the Herschel satellite. The prominent *moving* peak that enables the photo-$z$ determination with three photometric points is clearly visible.

its stellar emission features in the optical or emission-lines signatures. Differently from data in the optical or near-IR, FIR data lacks of an extensive photometric wavelength coverage: at most, individual galaxies will have around 10 photometric points in the FIR (and an average of 3-5), while in the optical it is common to have more than 30 bands (Casey et al., 2014). The basic idea assume that the far-IR SED is roughly fixed (e.g. to SMM-J2135 or Arp220) and then to use the FIR colors to infer the galaxy's redshift[14], allowing for a study of the redshift distribution in a statistical sense rather than for individual galaxies. SED fitting process can be divided into (i) methods that compare directly data (i.e. fluxes or colors) with model templates by means of $\chi^2$ or Bayesian techniques, and (ii) methods that fit for simple modified greybody-like functions. The baseline method used here to infer galaxies photo-$z$ is the former and will be discussed in Ch. (4) and (5). The main feature of the FIR SED is the presence of a peak due to dust emission (see Fig. (3.4)); however, note that the dust temperature (which manifests as the SED peak wavelength) correlates with IR luminosity and can be degenerate with redshift. The radial distribution of H-ATLAS galaxies is shown in Fig. (3.5).

---

[14]Of course the accuracy of the method is dependent on the intrinsic variation in SED types within a given population and as a consequence the precision can be poor.
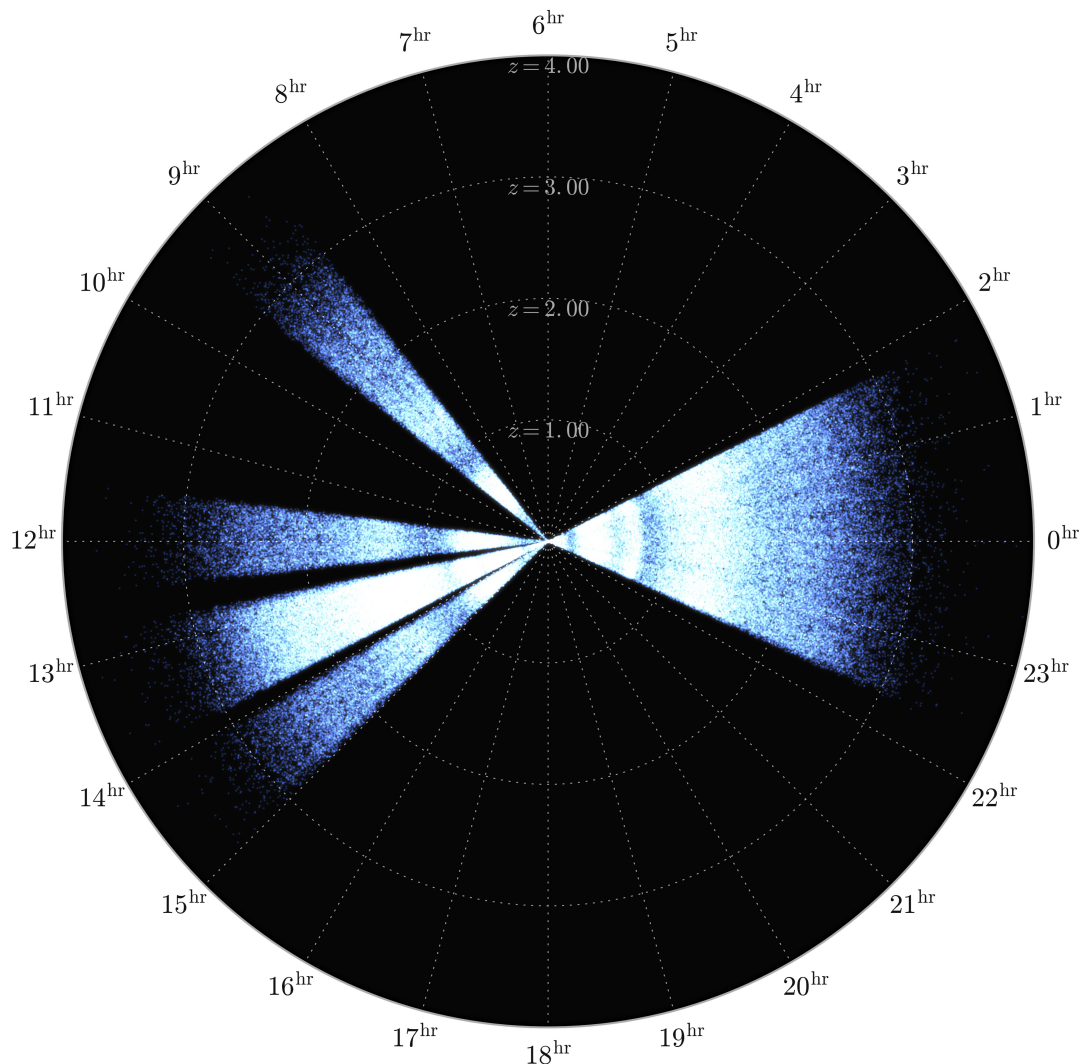
**Figure 3.5:** Angular and radial distribution of sub-mm galaxies in all of the five H-ATLAS patches. Please note that *photometric* redshifts have been used to place galaxies along the redshift axis and even though the spongy structure of the matter distribution is hardly observed - given the photo-$z$ uncertainties that can be important - the radial information can be used in a statistical way. In fact, the two main populations at $z \lesssim 1$ and at $z \gtrsim 1$ can be seen.

# 4

# Cross-correlation in the high-$z$ sky as seen by *Planck* and *Herschel*

In this chapter we present the first measurement of the correlation between the map of the CMB lensing potential derived from the *Planck* nominal mission data and the angular positions of the $z \gtrsim 1.5$ galaxies detected by the *Herschel*-ATLAS (H-ATLAS) survey covering about $600 \, \mathrm{deg}^2$, i.e. about $1.4\%$ of the sky. The hypothesis that there is no correlation between CMB lensing and the galaxy distribution is rejected at a $20\sigma$ significance, and the result is checked by performing a number of null tests. Specifically, the significance of the detection of the theoretically expected cross-correlation is found to be $10\,\sigma$. The galaxy bias parameter, $b$, derived from a joint analysis of the cross-power spectrum and of the auto-power spectrum of the galaxy density contrast is found to be $b = 2.80^{+0.12}_{-0.11}$, consistent with earlier estimates for H-ATLAS galaxies at similar redshifts. On the other hand, the amplitude of the cross-correlation is found to be a factor $A = 1.62 \pm 0.16$ higher than expected from the standard model ($A = 1$). The highly significant detection reported here using a catalog covering only $1.4\%$ of the sky demonstrates the potential of CMB lensing correlations with submillimeter surveys.

## 4.1 Introduction

The scientific context of the present analysis has already been thoroughly discussed in Ch. (2.2.3), here we just recall that the study of cross-correlations between CMB lensing with external tracers of LSS allows us to (i) constrain cosmology by reconstructing the dynamics and spatial distribution of the cosmological gravitational potentials, to (ii) derive the bias factor (as well as the effective halo masses) associated to the tracer populations, and to (iii) uncover known (and unknown) systematics affecting the datasets.

Several catalogs, such as those from the NRAO VLA Sky Survey (NVSS), the

Sloan Digital Sky Survey (SDSS), the Wide Field Survey Infrared Explorer (WISE) have already been cross-correlated with the CMB lensing potential. These surveys cover large areas of the sky but detected sources are mostly at $z \lesssim 1$. The H-ATLAS (Eales et al., 2010) allows us to extend the cross-correlation analysis up to substantially higher redshifts (Lapi et al., 2011; González-Nuevo et al., 2012), representing a novelty with respect to previous studies. In particular, by looking at Table (2.2), one can see that our analysis is the first one involving at the same time

1. a high-noise CMB lensing map;

2. an high-$z$ galaxy sample extremely well characterized in terms of astrophysics;

3. a small sky coverage.

In this chapter we present the first investigation of the cross-correlation between the CMB lensing potential measured by *Planck* and *Herschel*-selected galaxies with estimated redshifts $z \gtrsim 1.5$, i.e. at redshifts higher and closer to the peak of the lensing potential kernel than those of the source samples considered so far (as illustrated by Fig. (4.1)). Our choice of restricting the analysis to $z \gtrsim 1.5$ has a twofold motivation. First, because we aim to reconstruct the evolution of the lensing potential at higher redshifts than done with other galaxy samples, it is expedient to remove the dilution of the signal by low-$z$ sources. Second, as shown by Lapi et al. (2011) and González-Nuevo et al. (2012), the adopted approach for estimating photometric redshifts becomes unreliable at $z \lesssim 1$.

The outline of this chapter is as follows. In Sec. (4.2) we describe the theoretical background and the expected significance of the detection. The datasets are introduced in Sec. (4.3), whereas the analysis pipeline, including the estimator, the simulations used for validation of the algorithm and the error estimation are presented in Sec. (4.4). The measured auto- and cross-power spectra, as well as the null tests, are reported in Sec. (4.5). In Sec. (4.6) we analyze the constraints on the galaxy bias and in Sec. (4.7) we discuss the potential systematic effects that affect the cross-correlation. Finally in Sec (4.8) we summarize our results.

Throughout this chapter we adopt the fiducial flat $\Lambda$CDM cosmology with best-fit *Planck* + WP + highL + lensing cosmological parameters as provided by the *Planck* team in Planck Collaboration (2014b). Here, WP refers to WMAP polarization data at low multipoles, highL refers to the inclusion of high-resolution CMB data of the ACT and SPT experiments, and lensing refers to the inclusion of *Planck* CMB lensing data in the parameter likelihood.
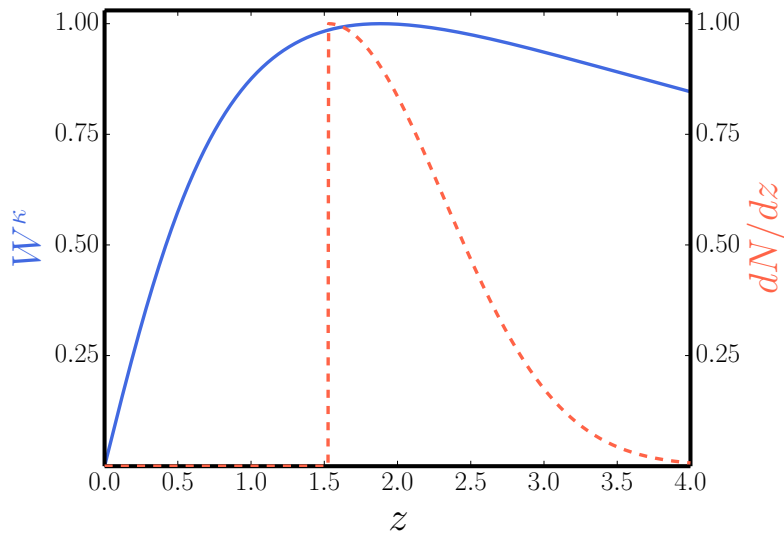
**Figure 4.1:** Estimated redshift distribution of the full sample of H-ATLAS galaxies (dashed red line) compared with the CMB lensing kernel $W^\kappa$ (blue solid line). Both the kernels are normalized to a unit maximum.

## 4.2 Theory and expectations

### 4.2.1 Signal modeling

We outlined the main features of CMB lensing in previous Chapters. Here we briefly recall the quantities which are of relevance for this one. The effect of gravitational lensing on CMB photons can be described as a remapping of the unlensed temperature anisotropies $\Theta(\hat{\mathbf{n}})$ by a two-dimensional vector field in the sky, namely the deflection field $\boldsymbol{d}(\hat{\mathbf{n}})$ (Lewis & Challinor, 2006):[1]

$$\begin{aligned}
\tilde{\Theta}(\hat{\mathbf{n}}) &= \Theta(\hat{\mathbf{n}} + \boldsymbol{d}(\hat{\mathbf{n}})) \\
&= \Theta(\hat{\mathbf{n}} + \nabla\phi(\hat{\mathbf{n}})) \\
&= \Theta(\hat{\mathbf{n}}) + \nabla^i\phi(\hat{\mathbf{n}})\nabla_i\Theta(\hat{\mathbf{n}}) + \mathcal{O}(\phi^2),
\end{aligned} \tag{4.1}$$

where $\tilde{\Theta}(\hat{\mathbf{n}})$ are the lensed temperature anisotropies and $\phi(\hat{\mathbf{n}})$ is the CMB lensing potential:

$$\phi(\hat{\mathbf{n}}) = -2\int_0^{z_*} \frac{c\,dz}{H(z)} \frac{\chi_* - \chi(z)}{\chi_*\chi(z)} \Phi(\chi(z)\hat{\mathbf{n}}, z). \tag{4.2}$$

In this equation, $\chi(z)$ is the comoving distance to redshift $z$, $\chi_*$ is the comoving distance to the last-scattering surface at $z_* \simeq 1090$, $H(z)$ is the Hubble factor

---

[1]To avoid confusion with $\alpha$, the logarithmic slope of the cumulative number counts of galaxies introduced in Ch. (2.3.3), we denote the deflection angle with the letter $\boldsymbol{d}(\hat{\mathbf{n}})$.

at redshift $z$, $c$ is the speed of light, and $\Phi(\chi(z)\hat{\mathbf{n}}, z)$ is the three-dimensional gravitational potential at a point on the photon path given by $\chi(z)\hat{\mathbf{n}}$. Note that the deflection angle is given by $\boldsymbol{d}(\hat{\mathbf{n}}) = \nabla\phi(\hat{\mathbf{n}})$, where $\nabla$ is the the two-dimensional gradient on the sphere. Because the lensing potential is an integrated measure of the projected gravitational potential, taking the two-dimensional Laplacian of the lensing potential we can define the lensing convergence $\kappa(\hat{\mathbf{n}}) = -\frac{1}{2}\nabla^2\phi(\hat{\mathbf{n}})$, which depends on the projected matter overdensity $\delta$ (Bartelmann & Schneider, 2001):

$$\kappa(\hat{\mathbf{n}}) = \int_0^{z*} dz\, W^\kappa(z)\delta(\chi(z)\hat{\mathbf{n}}, z). \tag{4.3}$$

The lensing kernel $W^\kappa$ is

$$W^\kappa(z) = \frac{3\Omega_{\mathrm{m}}}{2c}\frac{H_0^2}{H(z)}(1+z)\chi(z)\frac{\chi_* - \chi(z)}{\chi_*}, \tag{4.4}$$

where $\Omega_{\mathrm{m}}$ and $H_0$ are the present-day values of the Hubble and matter density parameters, respectively.

The galaxy overdensity $g(\hat{\mathbf{n}})$ in a given direction on the sky is also expressed as a LOS integral of the matter overdensity:

$$g(\hat{\mathbf{n}}) = \int_0^{z*} dz\, W^g(z)\delta(\chi(z)\hat{\mathbf{n}}, z), \tag{4.5}$$

where the kernel is

$$\begin{aligned}
W^g(z) &= \frac{b(z)\frac{dN}{dz}}{\left(\int dz'\, \frac{dN}{dz'}\right)} + \mu(z) \\
&= \frac{b(z)\frac{dN}{dz}}{\left(\int dz'\, \frac{dN}{dz'}\right)} + \frac{3\Omega_{\mathrm{m}}}{2c}\frac{H_0^2}{H(z)}(1+z)\chi(z)\int_z^{z*} dz'\left(1 - \frac{\chi(z)}{\chi(z')}\right)(\alpha(z') - 1)\frac{dN}{dz'}.
\end{aligned} \tag{4.6}$$

Assuming that the luminous matter traces the peaks of the dark matter distribution, the galaxy overdensity kernel is given by the sum of two terms. The first one is related to the physical clustering of sources and is given by the product of the linear bias[2] $b(z)$ and the *unit-normalized* redshift distribution $dN/dz$; the second one describes the effect of the *lensing magnification bias* (Ho et al., 2008; Xia et al., 2009) (see also Ch. (2.3.3)). We recall that this effect depends on the slope, $\alpha(z)$, of their integral counts ($N(> S) \propto S^{-\alpha}$) below the adopted flux density limit. Given the sharply peaked redshift distribution of our sources (see Fig. (4.1)) we can safely

---

[2]Throughout the analysis we assume a linear, local, deterministic, redshift- and scale-independent bias factor unless otherwise stated.
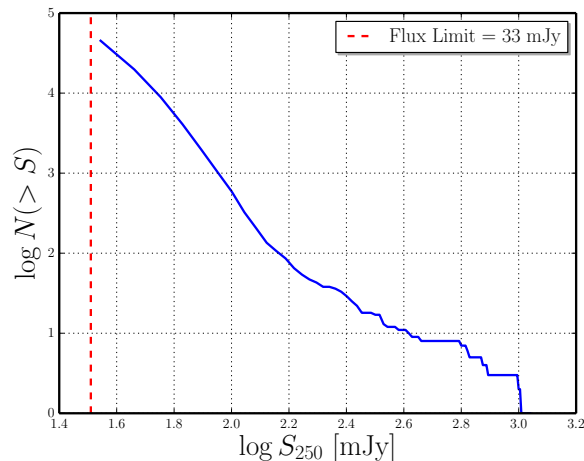
**Figure 4.2:** The cumulative number counts of galaxies as function of the flux at 250 $\mu$m, $S_{250\mu m}$, where the main selection is operated.

assume a redshift- and scale-independent linear bias ($b(z) = $ constant). Previous analyses of the clustering properties of submillimeter galaxies Xia et al. (2012); Cai et al. (2013) indicate $b \simeq 3$ at the redshifts of interest here, and we adopt this as our reference value.

Recent work by Gonzalez-Nuevo et al. (2014) has shown that the magnification bias by weak lensing is substantial for high-$z$ H-ATLAS sources selected with the same criteria as the present sample (see the Sec (4.3.2)). This is because the source counts are steep, although their slope below the adopted flux density limit ($S_{250\mu m} = 35$ mJy) is uncertain. The data (Béthermin et al., 2012) indicate, at this limit, $\alpha \simeq 2$ while for the high-$z$ galaxy subsample considered in this work we find $\alpha \simeq 3$, see Fig. (4.2). In the following we adopt the latter as our fiducial value. The effect of different choices for this parameter value is examined in Sec (4.7). Because the relevant angular scales are much smaller than 1 radian (multipoles $\ell \gtrsim 100$), the theoretical angular cross-correlation can be computed using the Limber approximation (Limber, 1953) as

$$C_\ell^{\kappa g} = \int_0^{z_*} \frac{dz}{c} \frac{H(z)}{\chi^2(z)} W^\kappa(z) W^g(z) P_{\delta\delta}\left(k = \frac{\ell}{\chi(z)}, z\right), \qquad (4.7)$$

where $P_{\delta\delta}(k, z)$ is the matter power spectrum, which we computed using the `CAMB`[3] code (Lewis et al., 2000). The nonlinear evolution of the matter power spectrum was taken into account using the `HALOFIT` prescription (Smith et al., 2003; Takahashi et al., 2012). A more extended discussion on the effect of the nonlinear evolution in CMB lensing maps based on N-body simulations is carried out by Antolini et al.

---

[3]available at `http://camb.info`

(2014). The CMB convergence, $W^\kappa(z)$, and the galaxy redshift distribution $dN/dz$ of the sample analyzed in this chapter (see Sec (4.3.2)) are shown in Fig. (4.1).

Again under the Limber approximation, the CMB convergence, $C_\ell^{\kappa\kappa}$, and the galaxy, $C_\ell^{gg}$, autospectra can be evaluated as

$$
\begin{aligned}
C_\ell^{\kappa\kappa} &= \int_0^{z_*} \frac{dz}{c} \frac{H(z)}{\chi^2(z)} \left[W^\kappa(z)\right]^2 P_{\delta\delta}\left(k = \frac{\ell}{\chi(z)}, z\right); \\
C_\ell^{gg} &= \int_0^{z_*} \frac{dz}{c} \frac{H(z)}{\chi^2(z)} \left[W^g(z)\right]^2 P_{\delta\delta}\left(k = \frac{\ell}{\chi(z)}, z\right).
\end{aligned}
\tag{4.8}
$$

## 4.2.2 Expected S/N

Plugging in the specifics of the H-ATLAS survey (see Ch. (4.3.2) and (3.3.2)), the mean redshift probed by the cross-correlation between CMB lensing and our sample is

$$
\langle z \rangle = \frac{\int_0^{z_*} \frac{dz}{c} z \frac{H(z)}{\chi^2(z)} W^\kappa(z) W^g(z) P_{\delta\delta}\left(k = \frac{\ell}{\chi(z)}, z\right)}{\int_0^{z_*} \frac{dz}{c} \frac{H(z)}{\chi^2(z)} W^\kappa(z) W^g(z) P_{\delta\delta}\left(k = \frac{\ell}{\chi(z)}, z\right)} \simeq 2.
\tag{4.9}
$$

The expected S/N of the convergence-density correlation can be predicted by assuming that both the galaxy overdensity and the lensing fields behave as Gaussian random fields, so that the variance of $C_\ell^{\kappa g}$ is (see Eq. (3.47))

$$
\left(\Delta C_\ell^{\kappa g}\right)^2 = \frac{1}{(2\ell+1)f_{\text{sky}}} \left[(C_\ell^{\kappa g})^2 + (C_\ell^{\kappa\kappa} + N_\ell^{\kappa\kappa})(C_\ell^{gg} + N_\ell^{gg})\right],
\tag{4.10}
$$

where $f_{\text{sky}}$ is the sky fraction covered by both the galaxy and the lensing surveys, $N_\ell^{\kappa\kappa}$ is the noise of the lensing field, and $N_\ell^{gg} = 1/\bar{n}$ is the shot noise associated with the galaxy field. Because our calculations are done in terms of the density contrast, the shot noise is inversely proportional to the mean number of sources per steradian, $\bar{n}$. The signal to noise ratio at multipole $\ell$ is then

$$
\left(\frac{S}{N}\right)_\ell^2 = \frac{\left(C_\ell^{\kappa g}\right)^2}{\left(\Delta C_\ell^{\kappa g}\right)^2} = \frac{(2\ell+1)f_{\text{sky}}\left(C_\ell^{\kappa g}\right)^2}{\left[(C_\ell^{\kappa g})^2 + (C_\ell^{\kappa\kappa} + N_\ell^{\kappa\kappa})(C_\ell^{gg} + N_\ell^{gg})\right]},
\tag{4.11}
$$

and the cumulative S/N for multipoles up to $\ell_{max}$ is

$$
\left(\frac{S}{N}\right)(< \ell_{max}) = \sqrt{\sum_{\ell'=\ell_{min}}^{\ell_{max}} \left(\frac{S}{N}\right)_{\ell'}^2}.
\tag{4.12}
$$

In Fig. (4.3) we show both the S/N per multipole and the cumulative one computed using the specifications for the *Planck* lensing noise (see Sec (4.3.1)) and the mean
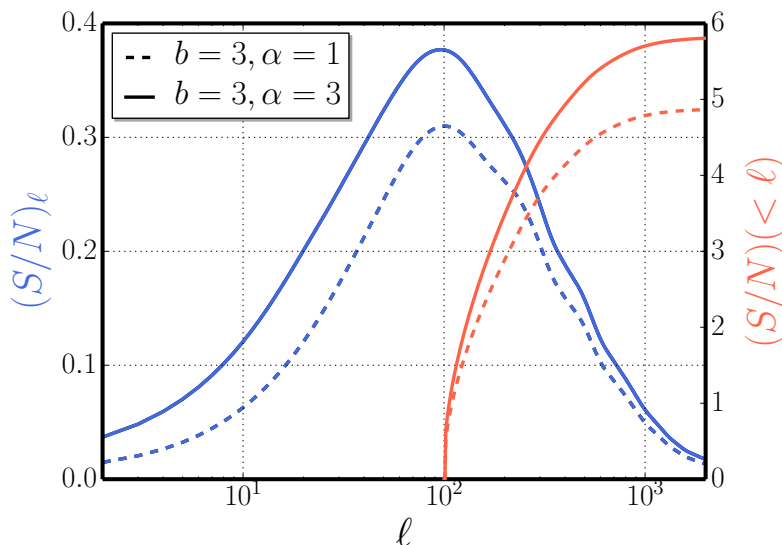
**Figure 4.3:** S/N per multipole (blue lines; left axis) and cumulative S/N (red lines; right axis) evaluated from $\ell_{\min} = 100$ for fiducial models with $b = 3$ and $\alpha = 1$ (no magnification, dashed lines) and $\alpha = 3$ (solid lines).

surface density of our source sample. It must be noted that, because of the limited area covered by the H-ATLAS survey (and split into 5 fields), the cross-correlation is only meaningful on scales below a few degrees. We have therefore limited our analysis to $\ell \geq \ell_{\min} = 100$. This restriction prevents us from exploiting the peak at $\ell \sim 100$ of the signal to noise per multipole. The cumulative S/N saturates at $\ell \sim 1000$. If $b = 3$ and $\alpha = 3$ we expect $S/N \simeq 6$.

## 4.3 Data

Both the *Planck* CMB lensing and *Herschel* submillimeter galaxies datasets used in this analysis have been previously introduced in Ch. (3.3); we recall here the main aspects of interest for this Chapter.

### 4.3.1 *Planck* data

We used the publicly released *Planck* CMB lensing potential map derived from the first 15.5 months of observations (Planck Collaboration, 2014c). The released map is based on a minimum variance combination of the 143 and 217 GHz temperature anisotropy maps only, because adding the 100 GHz map yields a negligible im-
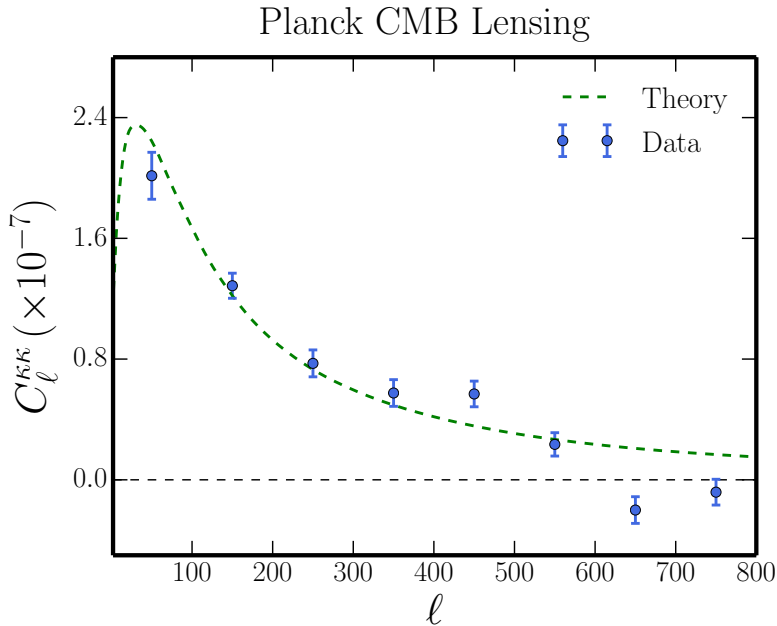
**Figure 4.4:** CMB convergence autopower spectrum as reconstructed from *Planck* data (blue points) on a portion of the sky with $f_{\text{sky}} \simeq 0.6$ compared with the theoretical prediction for our background cosmology (dashed green line).

provement (Planck Collaboration, 2014c).[4] The maps are in the HEALPix format with a resolution parameter of $N_{side} = 2048$, corresponding to 50, 331, and 648 pixels over the sky, with a pixel size of $\sim 1.7'$.

The power spectrum of the lensing potential is very red, and this may introduce a bias when we estimate it within multipole bins. To avoid this problem, we decided to convert the lensing potential map, $\phi$, into the convergence map, $\kappa$, which has a much less red power spectrum. This was done using the relation between the spherical harmonic coefficients of these quantities estimated on the full sky (Hu, 2000b)

$$\kappa_{\ell m} = -\frac{\ell(\ell+1)}{2}\phi_{\ell m} \ . \tag{4.13}$$

The convergence spherical harmonic coefficients were transformed to a map with resolution parameter $N_{side} = 512$ corresponding to a pixel size of $\sim 7'$. This resolution is sufficient for our analysis because the data noise level enables us to detect cross-correlations between the convergence and the galaxy density field only for angular scales larger than $\sim 20'$ ($\ell \lesssim 540$).

The convergence autopower spectrum recovered on approximately 60% of the sky using a modified version of the mask provided by the *Planck* collaboration is shown in Fig. (4.4). The auto-power spectrum has been corrected for the lensing

---

[4]We recall that the angular resolution and the noise level of the 100, 143 and 217 GHz frequency channels are $10'$, $7'$, $5'$ and 105, 45, 60 $\mu$K arcmin respectively.

reconstruction noise power spectrum $N_\ell^{\kappa\kappa}$ which was estimated from the set of 100 simulated lensing maps[5] recently released by the *Planck* team that account for the inhomogeneous noise level. The noise power spectrum was computed by averaging the spectra of the difference maps between the reconstructed and the input lensing map over 100 realizations. The errors on band powers were calculated as the diagonal part of the covariance matrix built from the simulation, as described in Sec. (4.4). The raw auto-power spectrum is not corrected for the bias induced by non-Gaussianity of unresolved point sources and for pseudo-$C_\ell$ leakage effects from masking (we just correct for N0 and N1 bias term adopting the formalism of Planck Collaboration (2014c)). These terms may cause some discrepancy of the power spectrum at high multipoles. Nevertheless, in the range of multipoles relevant for our analysis the power spectrum agrees pretty well with the theoretical one, and proper estimation of the convergence power spectrum is outside the scope of this thesis.

### 4.3.2 *Herschel* fields

We exploited the data collected by the *Herschel* Space Observatory (Pilbratt et al., 2010) in the context of the *Herschel* Astrophysical Terahertz Large Area Survey (H-ATLAS) (Eales et al., 2010), an open-time key program that has surveyed about 600 deg$^2$ with the Photodetector Array Camera and Spectrometer (PACS) (Poglitsch et al., 2010) and the Spectral and Photometric Imaging Receiver (SPIRE) (Griffin et al., 2010) in five bands, from 100 to 500 $\mu$m.

The surveyed area is divided into five fields: three equatorial fields centered on 9hr, 12hr, and 14.5hr (GAMA fields, G09, G12, and G15) covering, altogether, 161 deg$^2$; the NGP block; and the SGP block consisting of two concatenated rectangular regions. The footprint of the survey is shown in the right panel of Fig. (6.2).

The H-ATLAS galaxies have a broad redshift distribution extending from $z = 0$ to $z \simeq 5$ (Pearson et al., 2013). The $z \lesssim 1$ population is mostly made of "normal" late-type and star-burst galaxies with low to moderate star formation rates (SFRs; Dunne et al., 2011; Guo et al., 2011) while the high-$z$ galaxies are forming stars at high rates (SFR $\gtrsim$ few hundred $M_\odot$ yr$^{-1}$) and are much more strongly clustered (Maddox et al., 2010; Xia et al., 2012), implying that they are tracers of large-scale overdensities. Their properties are consistent with them being the progenitors of local massive elliptical galaxies (Lapi et al., 2011). We aim to correlate high-$z$ H-ATLAS galaxies with the *Planck* CMB lensing map.

---

[5]http://irsa.ipac.caltech.edu/data/Planck/release_1/ancillary-data/HFI_Products.html

To select the high-$z$ population, we adopted the criteria developed by González-Nuevo et al. (2012):

1. $S_{250\,\mu m} > 35$ mJy;

2. $S_{350\,\mu m}/S_{250\,\mu m} > 0.6$ and $S_{500\,\mu m}/S_{350\,\mu m} > 0.4$;

3. $3\,\sigma$ detection at $350\,\mu m$;

4. photometric redshift $z_{phot} > 1.5$, estimated following Lapi et al. (2011) and González-Nuevo et al. (2012).

Our final sample comprises a total of 99,823 sources, of which $9,099$ are in G09, $8,751$ in G12, $9,279$ in G15, $28,245$ in NGP and $44,449$ in SGP. The specifics of each patch are summarized in Table (4.1). The redshift distribution of the population is needed in order to predict the amplitude of the cross-correlation. Estimating the uncertainties in the redshift distribution due to photometric redshift errors is not a trivial task.

As stated in González-Nuevo et al. (2012) there is no indication that photometric redshifts are systematically under- or overestimated when the spectral energy distribution of SMM J2135-0102 is used as a template. The median value of $\Delta z/(1+z) \equiv (z_{phot} - z_{spec})/(1 + z_{spec})$ is -0.002 with a dispersion of 0.115. This dispersion corresponds to an rms error on $z$ of $\sigma_{\langle z \rangle} = 0.345$ at the mean redshift $\langle z \rangle \simeq 2$, given by Eq. ((4.9)). To get a rough indication of how many sources were scattered above and below the redshift threshold ($z = 1.5$) by measurement errors we have convolved a gaussian fit to the redshift distribution of sources selected with the first 3 criteria [(1) to (3)] with a gaussian error distribution having zero mean and dispersion $\sigma_{\langle z \rangle}$. The convolved redshift distribution was cut at $z = 1.5$, and the portion at higher $z$ was fitted with a half-normal distribution normalized to unity:

$$\frac{dN}{dz} = \frac{\sqrt{2}}{\sigma\sqrt{\pi}} \exp\left(-\frac{(z-\mu)^2}{2\sigma^2}\right). \tag{4.14}$$

The redshift distributions of the galaxies before and after the convolution are shown in Fig. (4.5).

We built an overdensity map at a resolution $N_{side} = 512$ defined by

$$g(\hat{\mathbf{n}}) = \frac{n(\hat{\mathbf{n}}) - \bar{n}}{\bar{n}}, \tag{4.15}$$

where $n(\hat{\mathbf{n}})$ is the number of objects in a given pixel, and $\bar{n}$ is the mean number of objects per pixel. The CMB convergence and galaxy overdensity maps in the different patches are shown in Fig. (4.6). We filtered out from these fields multipoles $\ell \gtrsim 400$ where $(S/N)_\ell \lesssim 0.3$.
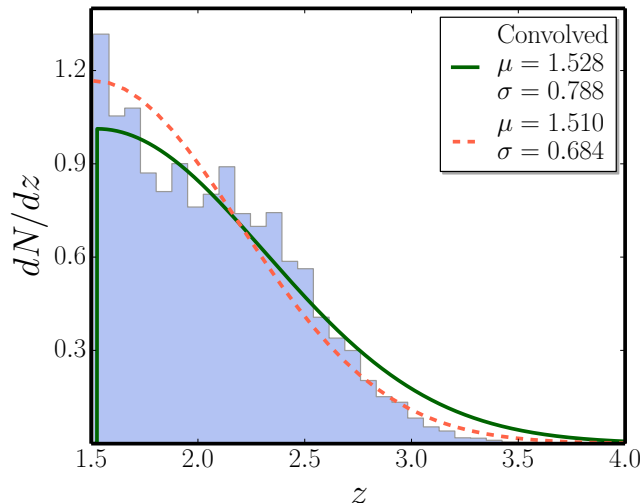
**Figure 4.5:** Redshift distribution of H-ATLAS galaxies for the combined set of patches used in the analysis. The (blue) histogram is the empirical redshift distributions, the dashed (orange) line is the half-normal fit to $dN/dz$ as described in text, while the solid (green) line represents the convolved $dN/dz$ that takes into account errors on photo-z estimation and is used as the fiducial distribution in our analysis. The values of the parameters $\mu$ and $\sigma$ given in the box are the best-fit values and are used in the analytic expression for $dN/dz$ adopted in calculations.

## 4.4 The Cross-Correlation Algorithm

### 4.4.1 Estimator

We computed the angular power spectra within the regions covered by the H-ATLAS survey using a PCL estimator based on the `MASTER` algorithm (Hivon et al., 2002).[6] In Ch. (3.2) we have extensively discussed PCL methods, here we just recall the main aspects and contextualize to the case of the CMB lensing-galaxy cross-correlation. For a survey that covers only a fraction of the sky, different modes of the true cross-power spectrum $C_\ell^{\kappa g}$ are coupled (Hauser & Peebles, 1973). The coupling can be described by the mode-mode coupling matrix $M_{\ell\ell'}$ which relates the pseudo-cross-spectrum $\tilde{C}_\ell^{\kappa g}$ measured from the data

$$\tilde{C}_\ell^{\kappa g} = \frac{1}{2\ell + 1} \sum_{m=-\ell}^{\ell} \tilde{\kappa}_{\ell m} \tilde{g}_{\ell m}^*. \tag{4.16}$$

to the true spectrum

$$\tilde{C}_\ell^{\kappa g} = \sum_{\ell'} M_{\ell\ell'} C_{\ell'}^{\kappa g}. \tag{4.17}$$

---

[6]These regions are inside the area used in the estimation of the CMB lensing map.

**Table 4.1:** H-ATLAS Patches Data

| Patch | $N_{obj}$ | $f_{\rm sky}$ | $\bar{n}\,[\mathrm{gal\,pix^{-1}}]$ | $\bar{n}\,[\mathrm{gal\,sr^{-1}}]$ |
|---|---|---|---|---|
| ALL[a] | 99823 | 0.014 | 2.30 | $5.76 \times 10^5$ |
| NGP | 28245 | 0.004 | 2.25 | $5.64 \times 10^5$ |
| SGP | 44449 | 0.006 | 2.38 | $5.95 \times 10^5$ |
| G09 | 9099 | 0.001 | 2.28 | $5.71 \times 10^5$ |
| G12 | 8751 | 0.001 | 2.13 | $5.35 \times 10^5$ |
| G15 | 9279 | 0.001 | 2.27 | $5.68 \times 10^5$ |

[a] ALL is the combination of all fields together.

However, we cannot directly invert Eq. (4.17) to get the true power spectrum, because for surveys covering only a small fraction of the sky, the coupling matrix $M_{\ell\ell'}$ becomes singular. To reduce the correlations of the $C_\ell$'s it is necessary to bin the power spectrum in $\ell$. We used eight linearly spaced bins of width $\Delta\ell = 100$ in the range $0 \le \ell \le 800$.

Then, the estimator of the true band powers $\hat{C}_L^{\kappa g}$ (hereafter $C_L^{\kappa g}$ denotes the binned power spectrum and $L$ identifies the bin) is given by

$$\hat{C}_L^{\kappa g} = \sum_{L'\ell} K_{LL'}^{-1} P_{L'\ell} \tilde{C}_\ell^{\kappa g}, \tag{4.18}$$

where

$$K_{LL'} = \sum_{\ell\ell'} P_{L\ell} M_{\ell\ell'} p_{\ell'}^2 Q_{\ell'L'}. \tag{4.19}$$

Here $P_{L\ell}$ is the binning operator; $Q_{\ell L}$ and $p_{\ell'}^2$ are, respectively, the reciprocal of the binning operator and the pixel window function that takes into account the finite pixel size. Because of the small size of the sky area covered by the H-ATLAS survey, the power spectrum for $\ell < 100$ is very poorly estimated, and we did not use it in our analysis. However, to avoid the bias coming from the lowest-order multipoles, the first multipole bin is included in the computation of the power spectrum; that is, the inversion of the binned coupling matrix $K_{LL'}$ is performed including the first bin, and the pseudopower spectrum for the first bin is used in the product of Eq. (4.18).

The main assumption in cross-correlation studies is that the noise levels related to the observables being analyzed are uncorrelated, so that we do not need to debias the reconstructed cross-spectrum for any noise term. However, when dealing with autopower spectra, such as $C_\ell^{gg}$ and $C_\ell^{\kappa\kappa}$, we have to correct the estimator given by
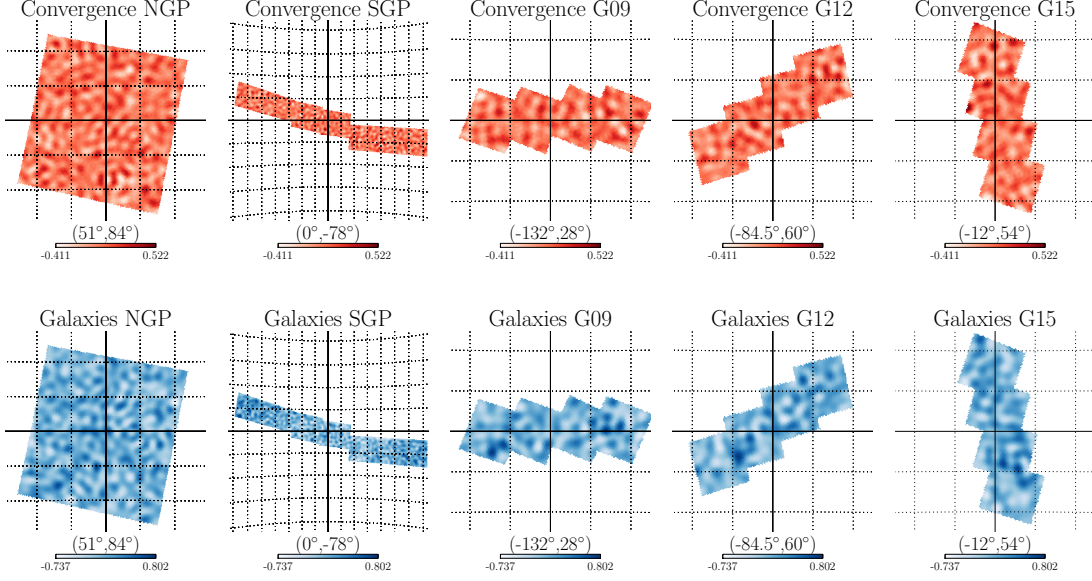
**Figure 4.6:** Convergence maps (upper row) and galaxy overdensity maps (lower row) in the H-ATLAS fields: multipoles $\ell > 400$ for which $(S/N)_\ell \lesssim 0.3$ have been filtered out. Galactic longitude and latitude $(l, b)$ of patch centers are provided in brackets. The grid overlay has spacing of $3°$ in each box.

Eq. (4.18) in order to account for the noise:

$$
\begin{aligned}
\hat{C}_L^{gg} &= \sum_{L'\ell} K_{LL'}^{-1} P_{L'\ell} \Big( \tilde{C}_\ell^{gg} - \langle \tilde{N}_\ell^{gg} \rangle_{\text{MC}} \Big), \\
\hat{C}_L^{\kappa\kappa} &= \sum_{L'\ell} K_{LL'}^{-1} P_{L'\ell} \Big( \tilde{C}_\ell^{\kappa\kappa} - \langle \tilde{N}_\ell^{\kappa\kappa} \rangle_{\text{MC}} \Big),
\end{aligned}
\tag{4.20}
$$

where $\langle \tilde{N}_\ell^{gg} \rangle_{\text{MC}}$ and $\langle \tilde{N}_\ell^{\kappa\kappa} \rangle_{\text{MC}}$ are the average noise pseudospectra estimated from the MC simulations.

## 4.4.2 Covariance matrix

The errors on the cross-power spectrum are described by the covariance matrix (Brown et al., 2005)

$$
\text{Cov}_{LL'}^{\kappa g} = M_{LL_1}^{-1} P_{L_1\ell} \widetilde{\text{Cov}}_{\ell\ell'}^{\kappa g} Q_{\ell' L_2} (M_{L'L_2}^{-1})^T,
\tag{4.21}
$$

where $\widetilde{\text{Cov}}_{\ell\ell'}^{\kappa g}$ is the pseudocovariance matrix given by

$$
\widetilde{\text{Cov}}_{\ell\ell'}^{\kappa g} = \frac{1}{2\ell'+1} M_{\ell\ell'} \Big[ C_\ell^{\kappa g}(b) C_{\ell'}^{\kappa g}(b) + \\
\sqrt{(C_\ell^{\kappa\kappa} + N_\ell^{\kappa\kappa})(C_\ell^{gg}(b) + N_{\ell'}^{gg})(C_{\ell'}^{\kappa\kappa} + N_{\ell'}^{\kappa\kappa})(C_{\ell'}^{gg}(b) + N_{\ell'}^{gg})} \Big].
\tag{4.22}
$$

The corresponding covariance matrix of the galaxy autocorrelation is obtained by replacing in Eq. ((4.21)) the pseudocovariance matrix $\widetilde{\mathrm{Cov}}_{\ell\ell'}^{\kappa g}$ with $\widetilde{\mathrm{Cov}}_{\ell\ell'}^{gg}$ given by

$$\widetilde{\mathrm{Cov}}_{\ell\ell'}^{gg} = \frac{2}{2\ell'+1} M_{\ell\ell'} \Big[ (C_\ell^{gg}(b) + N_\ell^{gg})(C_{\ell'}^{gg}(b) + N_{\ell'}^{gg}) \Big]. \tag{4.23}$$

The analytical expressions for the covariance matrices given above were used in the estimation of the galaxy bias and of the amplitude of the cross-correlation, presented in Sec (4.6).

### 4.4.3 Validation

In order to validate the algorithms used for the computation of the estimators outlined in the previous section and to check that the cross- and autopower spectra estimates are unbiased, we created 500 simulated maps of the CMB convergence field and of the galaxy overdensity field with statistical properties consistent with observations.

Using the theoretical spectra obtained with Eqs. (4.7) and (4.8), we generated full-sky signal maps, injecting a known degree of correlation, so that the simulated CMB convergence and galaxy harmonic modes satisfy both the auto- and the cross-correlations (Kamionkowski et al., 1997):

$$\begin{aligned}
\kappa_{\ell m} &= \zeta_1 \big( C_\ell^{\kappa\kappa} \big)^{1/2}; \\
g_{\ell m} &= \zeta_1 \frac{C_\ell^{\kappa g}}{\big( C_\ell^{\kappa\kappa} \big)^{1/2}} + \zeta_2 \left[ C_\ell^{gg} - \frac{\big( C_\ell^{\kappa g} \big)^2}{C_\ell^{\kappa\kappa}} \right]^{1/2}.
\end{aligned} \tag{4.24}$$

For each value of $\ell$ and $m > 0$, $\zeta_1$ and $\zeta_2$ are two complex numbers drawn from a Gaussian distribution with unit variance, whereas for $m = 0$ they are real and normally distributed.

We also generated 500 noise realizations for both fields. To simulate Gaussian convergence noise maps, we used the convergence noise power spectrum $N_\ell^{\kappa\kappa}$ available at the *Planck Legacy Archive*[7] (PLA). Although this power spectrum is not sufficiently accurate to estimate the convergence power spectrum, it should be sufficiently good for the cross-correlation analysis, which is not biased by the noise term. For the same reason, it is not crucial for our analysis to use the 100 simulations of the estimated lensing maps available at the PLA.

To take into account noise in the simulated galaxy maps, we proceeded in the

---

[7]http://wiki.cosmos.esa.int/planckpla/index.php/Specially_processed_maps

following way. For each signal map containing the galaxy overdensity, we generated a set of simulated galaxy number count maps, where the value in each pixel is drawn from a Poisson distribution with mean

$$\lambda(\hat{\mathbf{n}}) = \bar{n}(1 + g(\hat{\mathbf{n}})), \tag{4.25}$$

where $\bar{n}$ is the mean number of sources per pixel in a given H-ATLAS patch and $g(\hat{\mathbf{n}})$ is the corresponding simulated galaxy map containing only signal. The galaxy number counts map $\lambda(\hat{\mathbf{n}})$ was then converted into a galaxy overdensity map using Eq. (4.15), substituting the real number of objects in a given pixel $n(\hat{\mathbf{n}})$ with the simulated one $\lambda(\hat{\mathbf{n}})$. Note that maps obtained in this way already include Poisson noise with variance $N_\ell^{gg} = 1/\bar{n}$.

We applied the pipeline described above to our set of simulations in order to recover the input cross- and autopower spectra used to generate such simulations. The extracted $\hat{C}_L^{\kappa g}$, $\hat{C}_L^{gg}$, and $\hat{C}_L^{\kappa\kappa}$ spectra averaged over 500 simulations are reported in Fig. (4.7). The mean band power was computed as

$$\langle \hat{C}_L^{XY} \rangle = \frac{1}{N_{\text{sim}}} \sum_{i=1}^{N_{\text{sim}}} \hat{C}_L^{XY,i}, \tag{4.26}$$

where $X, Y = \{\kappa, g\}$, $i$ refers to the $i$-th simulation, and $N_{\text{sim}} = 500$ is the number of simulations. The errors were computed from the covariance matrix as

$$\Delta \hat{C}_L^{XY} = \left( \frac{\text{Cov}_{LL}^{XY}}{N_{\text{sim}}} \right)^{1/2}, \tag{4.27}$$

and the covariance matrix $\text{Cov}_{LL'}^{XY}$ was evaluated from the simulations as

$$\text{Cov}_{LL'}^{XY} = \frac{1}{N_{\text{sim}} - 1} \sum_{i=1}^{N_{\text{sim}}} (\hat{C}_L^{XY,i} - \langle \hat{C}_L^{XY} \rangle_{\text{MC}})(\hat{C}_{L'}^{XY,i} - \langle \hat{C}_{L'}^{XY} \rangle_{\text{MC}}). \tag{4.28}$$

We also show, for comparison, the theoretical error bars obtained from Eq. ((4.10)), modified to take into account the binning. They are in generally good agreement with the MC error estimates, which, however, are slightly larger (by up to $\sim 25\%$).
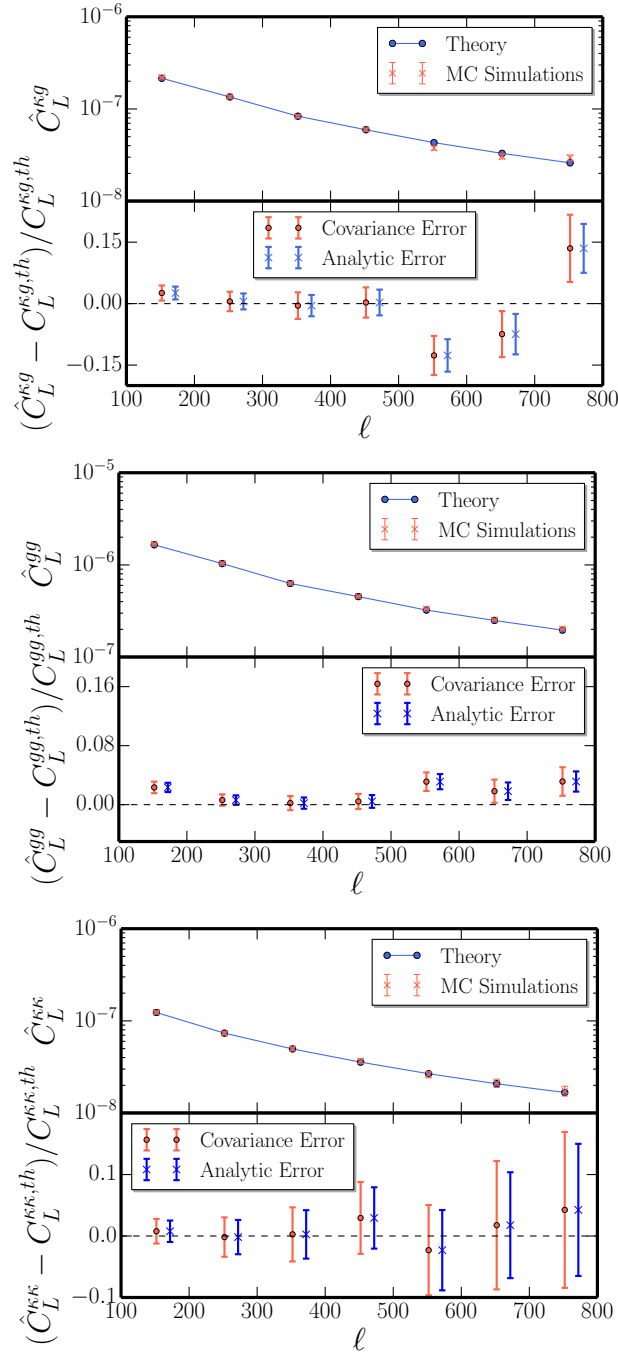
**Figure 4.7:** *Left. Upper panel*: cross-power spectrum of simulated galaxy and lensing maps constructed with $b = 3$. The points connected by the solid blue line represent the binned input cross-spectrum, and the average reconstructed spectrum from 500 simulations is shown by the orange points. *Lower panel*: fractional difference between the input and extracted cross-spectra. Error bars obtained with the simulation covariance matrix (orange points) and with the analytical approximation (blue points) are shown for comparison. *Middle.* As in left plot, but for the galaxy auto-power spectrum. *Right.* As in left plot, but for the CMB convergence autopower spectrum
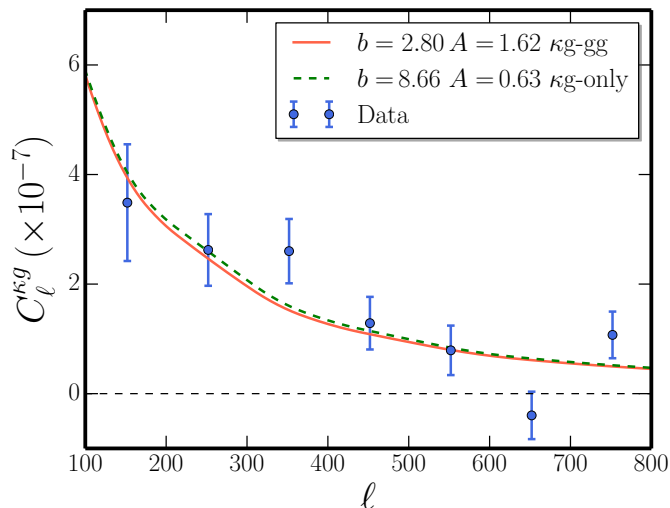
**Figure 4.8:** The CMB convergence-galaxy density cross-spectrum as measured from Planck and *Herschel* data. The data points are shown in blue, with error bars computed using the full covariance matrix obtained from Monte Carlo realizations of convergence maps. The theoretical spectra calculated with the bias values inferred from the likelihood analysis (as described in text) using the cross-correlation data only (solid red line) and the cross-correlation together with the galaxy autocorrelation data (dot-dashed green line) are also shown; we fix $\alpha = 3$ in this analysis. The null (no correlation) hypothesis is rejected at the $20\,\sigma$ level.

## 4.5 Power spectra

### 4.5.1 CMB Convergence-Galaxy Cross-correlation

The recovered cross-spectrum is shown in Fig. (4.8). To compute it we have applied to both maps masks that select the five H-ATLAS patches of interest. The error bars are estimated by cross-correlating 500 MC realizations of simulated CMB convergence maps (consisting of both signal and noise) with the true H-ATLAS galaxy density map, as described in Sec (4.5.3). This method (MC1 in the language of Ch. (3.2.2)) assumes that the two maps are uncorrelated; our error estimates are a good approximation because both maps are very noisy and $C_\ell^{\kappa\kappa,tot} C_\ell^{gg,tot} \gg (C_\ell^{\kappa g})^2$. We have also estimated the errors from cross-correlations of 500 MC realizations of simulated H-ATLAS galaxy density maps with the real *Planck* CMB convergence map. The former approach yields slightly smaller error bars, yet slightly larger than those estimated analytically (see Fig. (4.9)). These error estimates were checked by cross-correlating the publicly available set of 100 simulated lensing maps, which accurately reflect the *Planck* noise properties, with the real H-ATLAS map. The derived error bars are comparable with those found with our baseline approach, and there is no sign of systematic under- or overestimation.
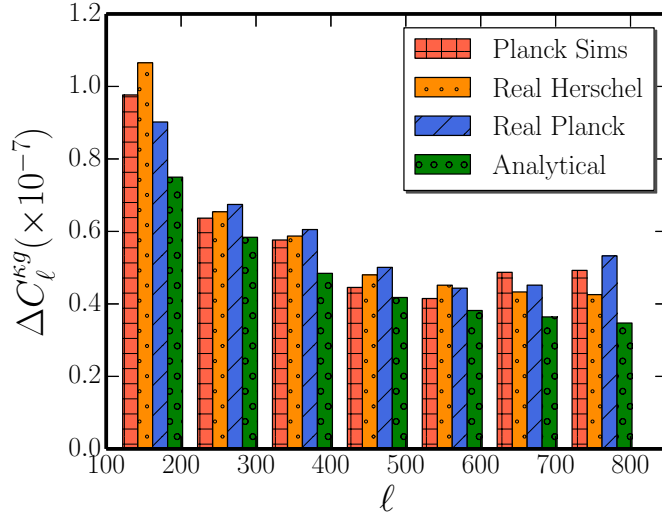
**Figure 4.9:** Error estimates for the cross-power spectrum band powers. The Monte Carlo estimates associated with estimated band powers are shown in orange (500 simulated lensing maps correlated with the real galaxy field). Blue bars represent errors obtained by correlating 500 simulated galaxy maps with the real convergence field, and the green bars represent the analytical approximation to these errors. Error estimates obtained by correlating the real galaxy field with the 100 lensing simulated maps by the Planck collaboration are shown in red.
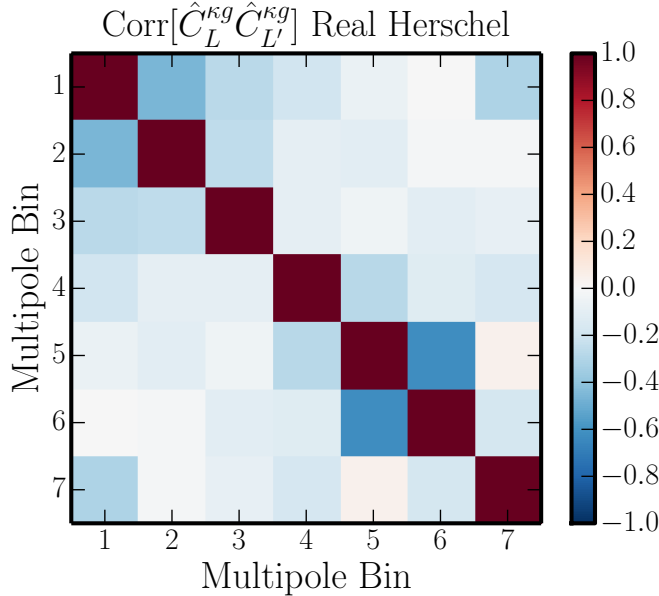


**Figure 4.10:** Correlation matrix $\mathrm{Corr}[\hat{C}_L^{\kappa g}\hat{C}_{L'}^{\kappa g}]$ built from the covariance matrix obtained by correlating 500 simulated lensing maps with the real H-ATLAS galaxy map.

We have exploited the simulations to build the covariance matrix, used to evaluate the probability that the measured signal is consistent with no correlation (our null hypothesis). As can be seen in Fig. (4.10), the covariance matrix is dominated by the diagonal components; however, off-diagonal components are non-negligible and have to be taken into account. The $\chi^2$ was calculated as

$$\chi^2_{\mathrm{null}} = \hat{\mathbf{C}}^{\kappa g}_L \, (\mathrm{Cov}^{\kappa g}_{LL'})^{-1} \, \hat{\mathbf{C}}^{\kappa g}_{L'}. \tag{4.29}$$

For the analysis performed with the whole H-ATLAS sample we obtained $\chi^2_{null} = 83.3$ for $\nu = 7$ degrees of freedom (dof), corresponding to a probability that the null hypothesis holds of $p = 2.89 \times 10^{-15}$. Because the $\chi^2$ distribution has mean $\nu$ and variance $2\nu$, the null hypothesis is rejected with a significance of about $(83.3 - 7)/(14^{1/2}) \simeq 20\,\sigma$. This is the sum in quadrature of the significance of the correlation in each band power, taking into account the correlations between different bins. The results of the $\chi^2$ analysis for each patch are reported in Table (4.2).

### 4.5.2 Galaxy Autocorrelation

We also performed an analysis of the autocorrelation of *Herschel* galaxies on the different patches. The shot noise subtracted autopower spectrum measured for the complete H-ATLAS data set is shown in Fig. (4.11). The error bars on the data points are evaluated from the diagonal part of the covariance matrix built from galaxy simulations with bias $b = 3$. The detected signal is highly significant $(40\,\sigma)$.

### 4.5.3 Null Tests

In order to verify our pipeline and the reconstructed spectra against the possibility of residual systematic errors, we performed a series of null tests, which consist of cross-correlating the real map of one field with simulated maps of the other field. Because there is no common cosmological signal, the mean correlation must be zero.

We cross-correlated our 500 simulated CMB lensing maps (containing both signal and noise) with the real H-ATLAS galaxy density contrast map and our 500 simulated galaxy maps constructed using $b = 3$ with the true *Planck* CMB convergence map. The error bars on the cross-power spectra were computed using the covariance matrices obtained from these simulations. As illustrated in Fig. (4.12) in both cases no significant signal was detected. In the first test we obtained $\chi^2 = 7.2$
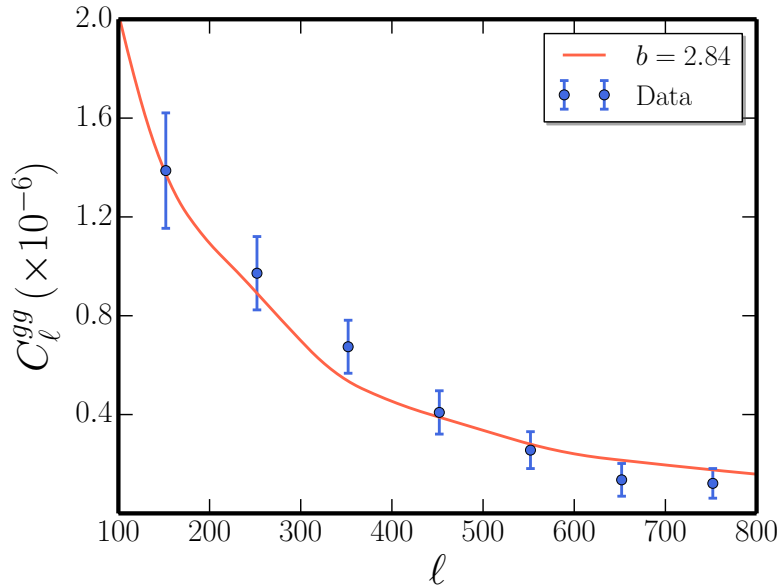
**Figure 4.11:** Galaxy density autopower spectrum for the whole sample of H-ATLAS galaxies. The data points are shown in blue, and the solid (red) line is the theoretical $C_\ell^{gg}$ evaluated for the best-fit value of the bias obtained using a likelihood analysis on the galaxy autospectrum data.

| Patch | $\chi^2_{\rm null}/\nu$ | p-value | Significance |
|-------|-------------------------|---------|--------------|
| ALL | 83.31/7 | $2.89 \times 10^{-15}$ | $20.3\sigma$ |
| NGP | 34.03/7 | $1.70 \times 10^{-5}$ | $7.2\sigma$ |
| SGP | 27.77/7 | 0.002 | $5.6\sigma$ |
| G09 | 22.41/7 | 0.002 | $4.1\sigma$ |
| G12 | 22.26/7 | 0.002 | $4.1\sigma$ |
| G15 | 29.23/7 | $1.0 \times 10^{-4}$ | $5.9\sigma$ |

**Table 4.2:** Significance of No Cross-correlation Hypothesis Rejection

corresponding to a probability of the null hypothesis (no correlation) $p = 0.41$, and in the second one we have $\chi^2 = 5.9$ and $p = 0.55$.

A further test consisted of cross-correlating the galaxy distribution in one patch of the sky with the lensing map in another. We moved in turn the three H-ATLAS GAMA fields and the SGP field to the position of the NGP patch and shifted the NGP galaxies to the SGP area. Then we cross-correlated each shifted galaxy map with the convergence field in the same position. The errors on the cross-correlations were obtained as above. All of the cross-spectra are consistent with no signal.
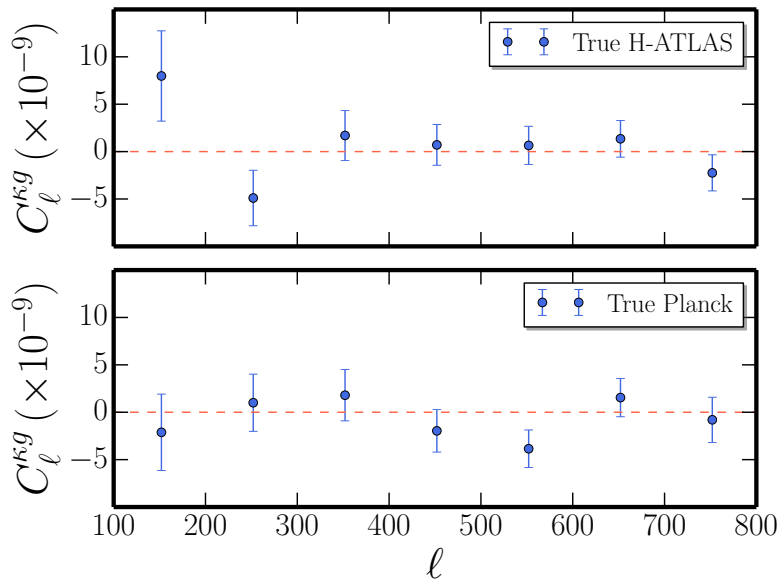
**Figure 4.12:** Results of null tests. *Upper panel*: mean correlation between the true H-ATLAS map including all of the five patches and 500 simulated CMB lensing maps. *Lower panel*: mean cross-spectra between the true *Planck* lensing map and 500 simulated galaxy maps with $b = 3$. No significant signal is detected in either case.

## 4.6 Constraints on bias and amplitude of cross-correlation

We now discuss the cross-correlation signal of cosmological origin. We introduce a phenomenologically motivated parameter, $A$, that scales the expected amplitude of the cross-power spectrum, $C_\ell^{\kappa g}$, of the *Planck* CMB lensing with the H-ATLAS galaxy overdensity map as $A\, C_L^{\kappa g}(b)$. Obviously, its expected value is one. The introduction of such parameter enables consistency checks both at the data and theory level: deviations from unity might point toward an incorrect modeling of the signal (e.g. wrong redshift distribution estimate, scale-dependent bias and magnification bias), new physics (e.g. growth history different from assumed $\Lambda$CDM and primordial NG) or systematics affecting the datasets (e.g. residual correlated contamination in the maps). Because the theoretical cross-spectrum is also basically proportional to the galaxy bias, there is a strong degeneracy between these two parameters. In order to break this degeneracy, we use also the galaxy autopower spectrum which depends only on $b$.

Let us note that, depending on the subject of the study, previous works in literature adopted different strategies to interpret the cross-correlation signal. We can distinguish two main approaches: either the bias is fixed (or a bias template is assumed) from prior knowledge and the observed cross-spectrum $\hat{C}_\ell^{\kappa g}$ is fit for an

amplitude parameter $A$ (Sherwin et al., 2012; Planck Collaboration, 2014c; Allison et al., 2015; Giannantonio et al., 2016), or the cosmology is assumed to be known and the reconstructed cross-spectrum is fit for the bias value (Feng et al., 2012; Bleem et al., 2012; Geach et al., 2013; DiPompeo et al., 2014; Omori & Holder, 2015). Both estimated $A$ or $b$ can then be converted in derived parameters; in particular, it is possible to construct statistics sensitive to deviations from $\Lambda$CDM, such as $E_g$ (Pullen et al., 2016) and $D_g$ (Giannantonio et al., 2016), by combining CMB lensing-galaxy cross-spectra with galaxy auto-spectra. We stress that the analysis presented here is the first one that constrain the cross-correlation amplitude $A$ and linear galaxy bias $b$ by combining the two-point statistics $C_\ell^{\kappa g}$ and $C_\ell^{gg}$. This analysis scheme has been subsequently exploited by Kuntz (2015) to reanalyze the cross-correlation signal between the CHFTLens galaxies with the *Planck* CMB lensing firstly studied by Omori & Holder (2015).

The best-fit values of the amplitude and of the galaxy bias were obtained by means of Bayesian analysis. In the following, we first describe the likelihood functions and present constraints on the redshift-independent galaxy bias and on the cross-correlation amplitude using galaxy autocorrelation data alone, cross-correlation data alone, and combining both data sets. In this analysis, the cosmological parameters and the counts slope $\alpha$ are kept fixed to the fiducial values. In order to efficiently sample the parameter space, we use the Markov chain Monte Carlo (MCMC) method assuming uninformative flat priors. For this purpose we employ `emcee` (Foreman-Mackey et al., 2013), a public implementation of the affine invariant MCMC ensemble sampler (Goodman & Weare, 2010). In this chapter, each quoted parameter estimate is the median of the appropriate posterior distribution after marginalizing over the remaining parameters with uncertainties given by the $16^{th}$ and $84^{th}$ percentiles (indicating the bounds of a 68% credible interval). For a Gaussian distribution, as is the case when combining both data sets, these percentiles correspond approximately to $-1\sigma$ and $+1\sigma$ values, and the median of the posterior is equal to the mean and maximum likelihood value.

We assumed Gaussian likelihood functions for the cross- and autopower spectra. For the galaxy autopower spectrum it takes the form

$$
\begin{aligned}
\mathcal{L}(\hat{C}_L^{gg}|b) = {} & \frac{1}{\sqrt{(2\pi)^{N_L}\det(\mathrm{Cov}_{LL'}^{gg})}} \\
& \times \exp\left\{-\frac{1}{2}[\hat{C}_L^{gg} - C_L^{gg}(b)]\,(\mathrm{Cov}_{LL'}^{gg})^{-1}\,[\hat{C}_{L'}^{gg} - C_{L'}^{gg}(b)]\right\},
\end{aligned}
\tag{4.30}
$$

| | $gg$ | $\kappa g$ | | $\kappa g + gg$ | | | |
|---|---|---|---|---|---|---|---|
| Patch | $b$ | $b$ | $A$ | $b$ | $A$ | $\chi^2_{\mathrm{th}}/\nu$ | p-value |
| ALL | $2.84^{+0.12}_{-0.11}$ | $8.66^{+4.23}_{-4.37}$ | $0.63^{+0.52}_{-0.20}$ | $2.80^{+0.12}_{-0.11}$ | $1.62^{+0.16}_{-0.16}$ | $12.6/5$ | $0.03$ |
| NGP | $2.72^{+0.22}_{-0.21}$ | $7.92^{+5.38}_{-6.38}$ | $0.53^{+1.35}_{-0.26}$ | $2.75^{+0.22}_{-0.21}$ | $1.27^{+0.28}_{-0.29}$ | $23.1/5$ | $3 \times 10^{-4}$ |
| SGP | $2.67^{+0.19}_{-0.19}$ | $0.78^{+1.86}_{-0.61}$ | $3.48^{+2.63}_{-1.95}$ | $2.69^{+0.18}_{-0.18}$ | $1.56^{+0.23}_{-0.23}$ | $5.7/5$ | $0.34$ |
| G09 | $3.79^{+0.35}_{-0.37}$ | $8.99^{+4.02}_{-5.06}$ | $1.11^{+0.96}_{-0.36}$ | $3.72^{+0.35}_{-0.32}$ | $2.11^{+0.41}_{-0.41}$ | $6.9/5$ | $0.22$ |
| G12 | $3.43^{+0.35}_{-0.33}$ | $3.34^{+6.84}_{-2.55}$ | $2.04^{+3.41}_{-1.23}$ | $3.36^{+0.35}_{-0.33}$ | $2.05^{+0.47}_{-0.46}$ | $13.7/5$ | $0.02$ |
| G15 | $3.14^{+0.33}_{-0.35}$ | $8.57^{+4.85}_{-6.54}$ | $0.97^{+1.72}_{-0.38}$ | $3.13^{+0.34}_{-0.34}$ | $2.06^{+0.45}_{-0.47}$ | $18.4/5$ | $2 \times 10^{-3}$ |

**Table 4.3:** H-ATLAS galaxy linear bias and cross-correlation amplitude as determined using both separately and jointly the reconstructed galaxy auto- and cross-spectra in the different patches

where $N_L = 7$ is the number of multipole bins and $\mathrm{Cov}^{gg}_{LL'}$ is the covariance matrix computed as described in Sec (4.4.2).

By sampling this likelihood for the measured H-ATLAS galaxy power spectrum $\hat{C}^{gg}_L$ we obtained constraints on the galaxy bias. Estimated values of the bias for all patches as well as for each of them are presented in Table (4.3). The results for the different patches are consistent with each other within $\lesssim 2\sigma$ and we note the GAMA patches to be appear a bit more biased with respect to NGP and SGP. The global value, $b = 2.84 \pm 0.12$, is consistent with earlier estimates. For example, Xia et al. (2012) found an effective value of the bias factor $b_{eff} \simeq 3$ (no error given) "for the bulk of galaxies at $z \simeq 2$". The Planck Collaboration (2014e) found, from their analysis of the CIB, a slightly lower value ($b_{eff} \simeq 2.6$), as expected because a large contribution to the CIB comes from fainter, presumably less biased, sources.

We used the measured cross-spectra to constrain the $b$ and $A$ parameters in the same fashion. As noted above, the cross-spectra basically measure the product $A \times b$. The likelihood function is given by

$$\mathcal{L}(\hat{C}^{\kappa g}_L | b, A) = \frac{1}{\sqrt{(2\pi)^{N_L} \det(\mathrm{Cov}^{\kappa g}_{LL'})}}$$
$$\times \exp\left\{ -\frac{1}{2} [\hat{C}^{\kappa g}_L - A\, C^{\kappa g}_L(b)]\, (\mathrm{Cov}^{\kappa g}_{LL'})^{-1}\, [\hat{C}^{\kappa g}_{L'} - A\, C^{\kappa g}_{L'}(b)] \right\}, \quad (4.31)$$

where $\mathrm{Cov}^{\kappa g}_{LL'}$ is the covariance matrix (Eq. ((4.21))). The results are shown in Table (4.3).

Finally, we studied the constraints on $b$ and $A$ by combining the cross- and galaxy autospectra. For the joint analysis we used the Gaussian likelihood function that

takes into account correlations between the cross- and the autopower spectra in the covariance matrix. We organized the extracted cross- and autoband powers into a single data vector as

$$\hat{\mathbf{C}}_L = (\hat{\mathbf{C}}_L^{\kappa g}, \hat{\mathbf{C}}_L^{gg}), \tag{4.32}$$

which has 14 elements. The total covariance matrix is then written as the composition of four $7 \times 7$ submatrices

$$\mathrm{Cov}_{LL'} = \begin{bmatrix} \mathrm{Cov}_{LL'}^{\kappa g} & (\mathrm{Cov}_{LL'}^{\kappa g-gg})^{\intercal} \\ \mathrm{Cov}_{LL'}^{\kappa g-gg} & \mathrm{Cov}_{LL'}^{gg} \end{bmatrix} \tag{4.33}$$

where the mixed covariance that takes into account the correlation between the two observables is

$$\mathrm{Cov}_{LL'}^{\kappa g-gg} = M_{LL_1}^{-1} P_{L_1\ell} \widetilde{\mathrm{Cov}}_{\ell\ell'}^{\kappa g-gg} Q_{\ell'L_2}(M_{L'L_2}^{-1})^{\intercal} \tag{4.34}$$

$$\widetilde{\mathrm{Cov}}_{\ell\ell'}^{\kappa g-gg} = \frac{2}{2\ell'+1} M_{\ell\ell'} \left[ (C_\ell^{gg}(b) + N_\ell^{gg})(C_{\ell'}^{gg}(b) + N_{\ell'}^{gg}) C_\ell^{\kappa g}(b) C_{\ell'}^{\kappa g}(b) \right]^{1/2}. \tag{4.35}$$

In the above expressions, $\mathrm{Cov}_{LL'}^{\kappa g}$ and $\mathrm{Cov}_{LL'}^{gg}$ are the covariance matrices evaluated using Eq. ((4.21)).

The full 2-dimensional posterior distributions of the $b$ and $A$ parameters, as well as the marginalized ones obtained from this analysis, are shown in Fig. (4.13). Numerical values of the parameters are presented in Table (4.3), where the best-fit values and the errors are evaluated as the $50^{\text{th}}$, $16^{\text{th}}$, and $84^{\text{th}}$ percentiles, respectively, of the posterior distributions. The $\chi^2$ values are evaluated as $\chi_{th}^2 = [\hat{\mathbf{C}}_L^{\kappa g} - A_{bf}\mathbf{C}_L^{\kappa g}(b_{bf})](\mathrm{Cov}_{LL'}^{\kappa g})^{-1} [\hat{\mathbf{C}}_L^{\kappa g} - A_{bf}\mathbf{C}_L^{\kappa g}(b_{bf})]$, where $b_{bf}$ and $A_{bf}$ are the best-fit values. Note that the posterior distributions of $b$ and $A$ obtained using only cross-correlation data are far from being Gaussian. As a sanity check, we derived a theoretical upper limit on $A$ considering that cross-spectrum cannot be larger than the geometric mean of the two autospectra: $A \leq (C_L^{\kappa g,th}(Cov_{LL'}^{\kappa g})^{-1}\sqrt{\hat{C}_{L'}^{\kappa\kappa}\hat{C}_{L'}^{gg}})/(C_L^{\kappa g,th}(Cov_{LL'}^{\kappa g})^{-1}C_{L'}^{\kappa g,th}) \sim 2.5$.

The $\chi^2$ value of the best-fit theoretical spectrum is $\chi_{th}^2 = 12.6$ for $\nu = 5$ dof ($\chi_{th}^2/\nu = 2.5$). The significance of the detection of the theoretically expected cross-correlation signal was evaluated as the ratio between the estimated amplitude $A$ and its error $\sigma_A$: $A/\sigma_A \simeq 10$, corresponding to a $10\sigma$ significance.

The constraint on the bias factor from the joint fit of the galaxy autocorrelation and of the cross-correlation power spectra, $b = 2.80^{+0.12}_{-0.11}$, is consistent with earlier estimates (Xia et al., 2012). On the other hand, the cross-correlation amplitude is $A = 1.62 \pm 0.16$ times larger than expected for the standard $\Lambda$CDM model for the evolution of large-scale structure. This is at odds with the results of the
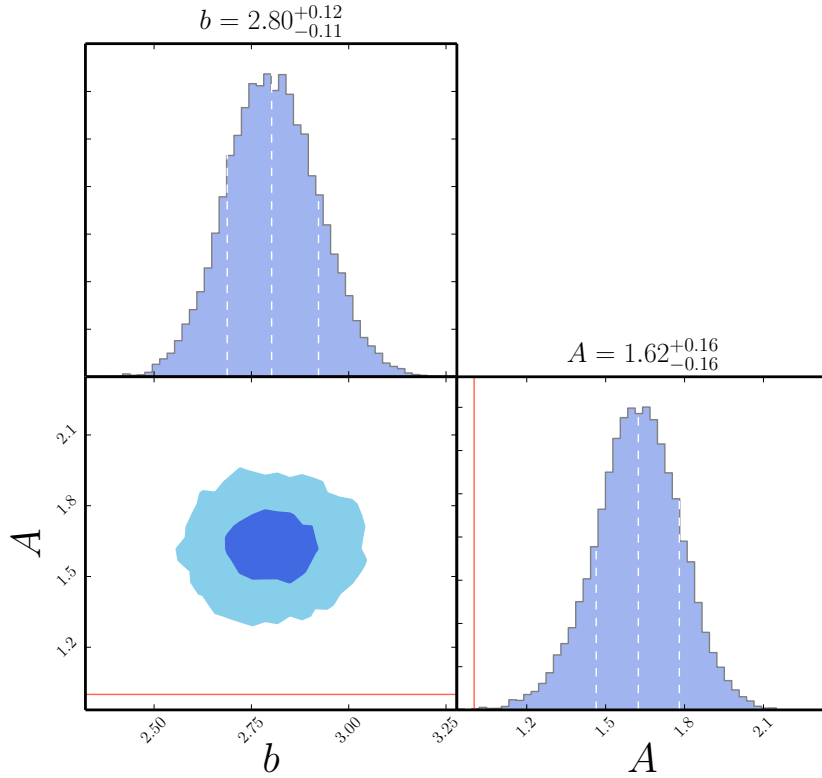
**Figure 4.13:** Posterior distribution in the $b - A$ plane with the 68% and 95% confidence contours (darker and lighter colors, respectively), together with the marginalized distributions of each parameter with $1\sigma$ errors shown by the dashed white lines, obtained by combining the convergence-galaxy cross-correlation and the galaxy autocorrelation data for each patch. The solid red line represents the standard case in which $A = 1$, and $\alpha$ is set to 3 for the analysis.
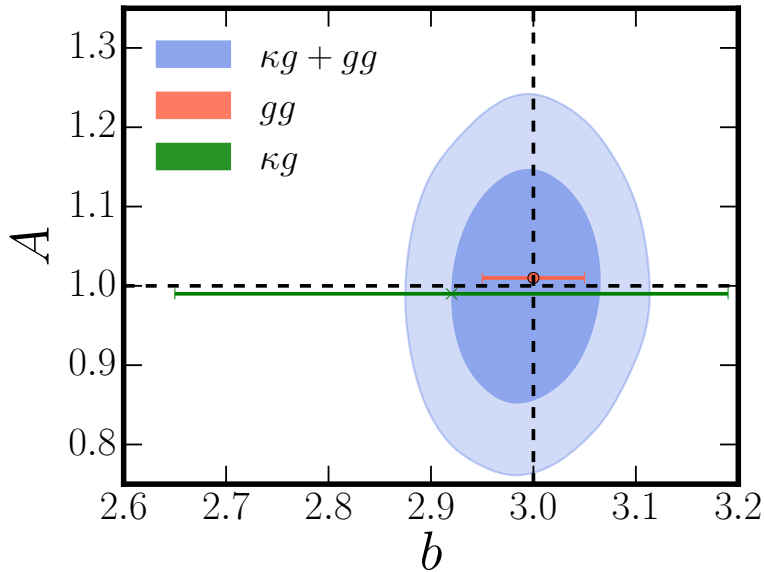
**Figure 4.14:** Validation of the MCMC pipeline and constraining power comparison. We show posterior distribution in the $b - A$ plane with the 68% and 95% confidence contours (darker and lighter colors, respectively), obtained by combining the CMB convergence-galaxy cross-correlation and the galaxy autocorrelation mock data with $b = 3$ and $A = 1$. The $1\sigma$ red and green error bars represent the bias constraints found using only the $C_\ell^{gg}$ and $C_\ell^{\kappa g}$ (fixing $A = 1$), respectively.

cross-correlation analyses presented in the Planck Collaboration (2014c) paper, which are consistent with $A = 1$ except, perhaps, in the case of the MaxBCG cluster catalog. Possible causes of the large value of $A$ are discussed in the following section.

In order to provide further proof of the reliability of the MCMC analysis, as well as to give a comparison of the constraining power of the different two-point statistics exploited here, we conclude this subsection with Fig. (4.14): there, we show the results of the MCMC analysis on mock data using $C_\ell^{gg}$, $C_\ell^{\kappa g}$, and $C_\ell^{gg} + C_\ell^{\kappa g}$. To produce this plot we first output two correlated *signal plus noise* $\kappa$ and $g$ maps with the fiducial model by setting $b = 3$ and $A = 1$. We then extract both the cross- and auto-power spectra and fit separately $\hat{C}_\ell^{gg,\text{fake}}$ and $\hat{C}_\ell^{\kappa g,\text{fake}}$ (assuming $A = 1$) for the bias value, while the combination of $\hat{C}_\ell^{\kappa g,\text{fake}} + \hat{C}_\ell^{gg,\text{fake}}$ is fit for the bias and amplitude parameters. As we can see, in all cases the input model is recovered, i.e. $b = 3$ and $A = 1$ within $1\sigma$, and there are no signs of spurious biases. In particular, the galaxy auto-power spectrum seems to constrain tighter the value of the linear galaxy bias, though it is expected to be much more sensitive to residual systematics.
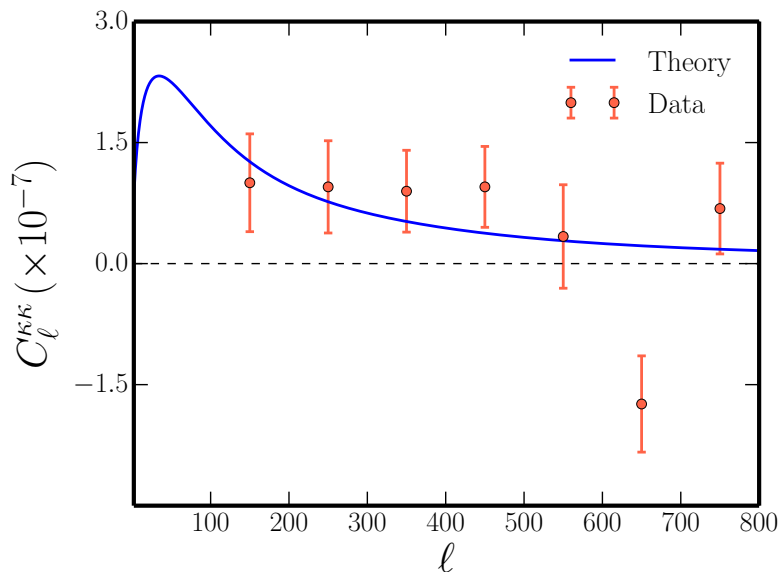
**Figure 4.15:** CMB convergence autopower spectrum recovered using the H-ATLAS mask. Theory line as in Fig. (4.4).

## 4.7 Discussion

The correlation between the CMB lensing potential and the distribution of high-$z$, submillimeter selected galaxies was found to be stronger than expected for the standard cosmological model. We now address on one side the possibility that the tension between the estimated and the expected value of the amplitude $A$ is overrated because of an underestimate of the errors and, on the other side, astrophysical effects that may enhance the measured signal.

### 4.7.1 Noise Levels

Due to the inhomogeneity of the noise level in the *Planck* survey, the H-ATLAS patches used for the cross-correlation may have slightly higher than average effective noise. To check this possibility, we reconstructed the CMB convergence autopower spectrum for each of the H-ATLAS patches. Error bars were derived from 100 simulated *Planck* lensing maps. The results of the analysis performed combining the five patches show some excess power for $\ell \sim 400$–$500$ (Fig. (4.15)). Considering the patches separately we find that the main features of the CMB lensing power spectrum are recovered in the two largest patches, whereas the power spectrum in the three GAMA fields seems to be dominated by noise. Thus, there is an indication of a slight underestimate of the noise bias in the latter fields, but the effect on the combined patches is marginal.
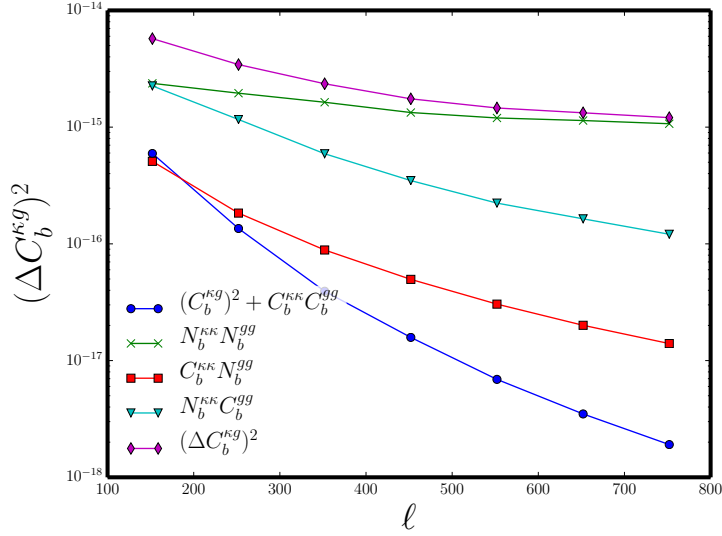
**Figure 4.16:** Contributions to the cross-spectrum variance $(\Delta C_\ell^{\kappa g})^2$ [see Eq. ((4.36))]. Blue line: signal only term. Green line: noise only term. Red and cyan lines: mixed signal and noise terms.

To understand which is the main statistical error source on the cross-power spectrum, we have analyzed the contributions to the error budget. The autospectra contain a signal and a noise term as $\hat{C}_L^{XX} = C_L^{XX} + N_L^{XX}$, so that the errors on the cross-spectra can be written as

$$
\begin{aligned}
f_{\text{sky}}(2L+1)\Delta\ell\big(\Delta\hat{C}_L^{\kappa g}\big)^2 = &\left[C_L^{\kappa\kappa}C_L^{gg} + (C_L^{\kappa g})^2\right] \\
&+ N_L^{\kappa\kappa}N_L^{gg} + C_L^{\kappa\kappa}N_L^{gg} + C_L^{gg}N_L^{\kappa\kappa}.
\end{aligned}
\tag{4.36}
$$

The first term represents the cosmic variance, the second one the pure noise, and the remaining are mixed signal-noise terms. As can be seen from Fig. (4.16), the main contribution to the $C_L^{\kappa g}$ variance is given by the noise-only term. Moreover, the relative amplitude of the mixed terms is telling us that most of the error comes from the lensing noise. In order to reduce the errors of the reconstructed cross-spectrum, it is important to reach high sensitivity in reconstructing the CMB lensing potential. This, of course, does not include the possible systematic errors discussed below.

## 4.7.2 Astrophysical systematics

First we have checked the effect on the auto- and cross-spectra of errors of photometric redshift estimates. To this end we have redone the full analysis using the initial redshift distribution, $dN/dz$, i.e. the one represented by the dashed red line
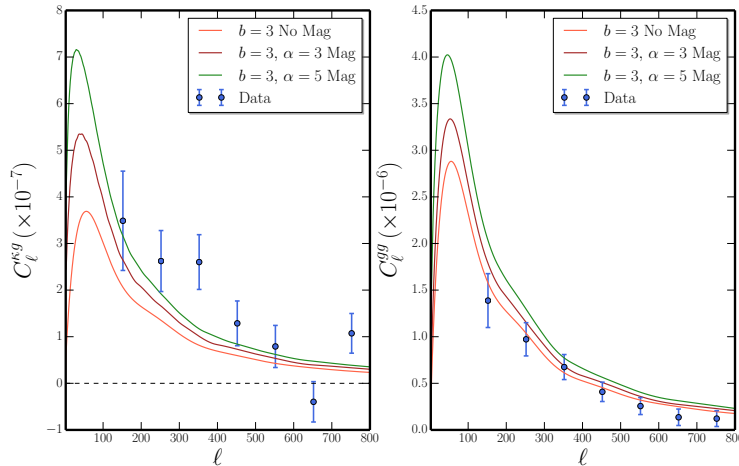
**Figure 4.17:** Effect of lensing magnification bias on the cross-power spectrum (left panel) and on the galaxy autopower spectrum (right panel). In both panels, theory lines are plotted for bias values $b = 3$, while the slope of the galaxy number counts as function of flux is set to $\alpha = 1$ (no magnification) and $\alpha = 3, 5$ as described in the legend.

in Fig. (4.5). As shown in Fig. (4.19), we get a slightly higher value of the cross-spectrum amplitude ($A = 1.70^{+0.16}_{-0.17}$) and a somewhat lower value of the galaxy bias ($b = 2.59^{+0.11}_{-0.11}$). The reason for that is easily understood. As shown by Fig. (4.5), the convolution of the initial $dN/dz$ with the smoothing kernel (representative of the uncertainties on estimated redshifts) results in a broadening of the distribution. This translates into a decrease of the expected amplitude for both the cross- and the autopower spectra. Hence, in order to fit the same data, we need a higher value of the galaxy bias and, consequently, a lower value of the cross-spectrum amplitude $A$. Because the derived value of $b$ is quite sensitive to the adopted redshift distribution, the agreement with other, independent determinations implies that our $dN/dz$ cannot be badly off. Therefore, it looks unlikely that the higher than expected value of $A$ can be ascribed to a wrong estimate of $dN/dz$.

Our choice of a constant $b$ over the redshift range spanned by the H-ATLAS catalog is obviously an approximation, and the effective values of $b$ may be different for the cross- and the galaxy autopower spectra. To check the effect of this approximation on the estimates of $C_\ell^{\kappa g}$ and $C_\ell^{gg}$ we have computed the effective values of the bias for the two cases

$$
\begin{aligned}
b_{eff}^{\kappa g} &= \frac{\int \frac{\mathrm{d}z}{c} b(z) \frac{H(z)}{\chi^2(z)} W^\kappa(z) \frac{dN}{dz} P_{\delta\delta}(k, z)}{\int \frac{\mathrm{d}z}{c} \frac{H(z)}{\chi^2(z)} W^\kappa(z) \frac{dN}{dz} P_{\delta\delta}(k, z)}, \\
(b_{eff}^{gg})^2 &= \frac{\int \frac{\mathrm{d}z}{c} b^2(z) \frac{H(z)}{\chi^2(z)} (\frac{dN}{dz})^2 P_{\delta\delta}(k, z)}{\int \frac{\mathrm{d}z}{c} \frac{H(z)}{\chi^2(z)} (\frac{dN}{dz})^2 P_{\delta\delta}(k, z)},
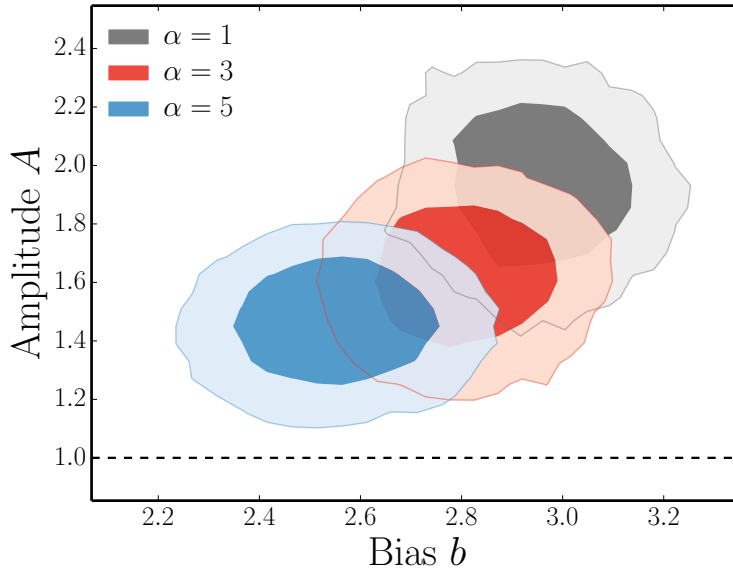\end{aligned}
\tag{4.37}
$$

**Figure 4.18:** Effect of fixed slope of number counts $\alpha$ on the inferred values of cross-correlation amplitude $A$ and bias $b$. We show $1-$ and $2\sigma$ contours (darker and lighter shaded regions, respectively). As the $\alpha$ parameter increases, both $A$ and $b$ shift toward smaller values.

using the bias evolution model $b(z)$ from Sheth & Tormen (1999) for halo masses in the range $10^{12}$–$10^{13}\,\mathrm{M}_\odot$. We find that $b_{eff}^{\kappa g}$ is only slightly larger (by $\simeq 6\%$) than $b_{eff}^{gg}$. Hence, considering a redshift-dependent bias factor would only marginally affect the expected cross-spectrum.

Weak lensing by foreground structures modifies the *observed* density of background sources compared to the real one (Ho et al., 2008; Xia et al., 2009) and is especially important for high-redshift objects. The effect on the galaxy overdensity kernel is described by the second term on the right-hand side of Eq. ((4.6)). The effect of the magnification bias on both $C_\ell^{\kappa g}$ and $C_\ell^{gg}$ is illustrated in Fig. (4.17) where we show the expected power spectra for $A = 1$, $b = 3$, and three values of $\alpha$: 1 (no magnification bias), 3, and 5. The impact of the magnification bias is clearly stronger for $C_\ell^{\kappa g}$.

Fitting the joint data for $\alpha = 1$ we find $b = 2.95^{+0.12}_{-0.11}$ and $A = 1.93^{+0.18}_{-0.19}$ while for $\alpha = 5$, $b = 2.55^{+0.13}_{-0.12}$ and $A = 1.46 \pm 0.14$. The contour plots in the $A - b$ plane are shown in Fig. (4.18). Higher values of $\alpha$ imply lower values of $A$, but even for $\alpha = 5$ the data require $A > 1$.

Another systematic effect that can bias our measurement of the CMB convergence-galaxy cross-correlation is the leakage of CIB emission into the lensing map through the temperature maps used for the lensing estimation, as it correlates strongly with the CMB lensing signal Planck Collaboration (2014d). The 857 GHz *Planck* map

used by Planck Collaboration (2014c) as a Galactic dust template also removes the portion of the CIB fluctuations that have a spectral index similar to that of Galactic dust. However, as noted in that paper, this approach is liable to problems due, for example, to variation of Galactic dust spectral indices across the sky, as well as to the mismatch between the beams at 100/143/217 and 857 GHz.

The H-ATLAS galaxies are well below the *Planck* detection limits (their flux densities at 148 GHz are expected to be in the range 0.1–1 mJy, hence are much fainter than sources masked by Planck Collaboration (2014c)). Thus they are part of the CIB measured by *Planck*. If they are only partially removed by the use of the 857 GHz map, they are potentially an important contaminant of the cross-correlation, resulting in an enhancement of the observed signal. The shot-noise correction applied by the *Planck* analysis removes only partly the contamination by infrared sources because their main contribution to the fluctuation field is due to clustering.

Foreground induced biases to CMB lensing reconstruction have been extensively studied by van Engelen et al. (2014) and Osborne et al. (2014). These authors concluded that the impact of these sources of systematic errors should be small due to *Planck*'s resolution and noise levels. However, a calculation of the bias on the cross-spectrum discussed in this chapter is beyond the scope of the present chapter. Having at hand CMB lensing maps at different frequencies would allow to investigate the CIB leakage issue in more detail.

Clusters of galaxies, which trace the large-scale potential responsible for the CMB lensing, are visible at millimeter and submillimeter wavelengths via the scattering of CMB photons by hot electrons (Sunyaev-Zel'dovich effect) and might therefore contaminate the cross-correlation signal to some extent. However, the redshift range populated by galaxy clusters only marginally overlaps with the redshift distribution of our sources, so that this contamination is negligible.

## 4.8 Conclusions

We have presented the first measurement of the correlation between the lensing potential derived from the *Planck* data and a high-$z$ ($z \geq 1.5$) galaxy catalog from the *Herschel*-ATLAS survey, the highest redshift sample for which the correlation between *Planck* CMB lensing and tracers of large-scale structure has been investigated so far. We have shown that the expected signal is remarkably strong, in spite of the small area covered by the H-ATLAS survey (about 1.3% of the sky), suggesting that cross-correlation measurements between CMB lensing maps
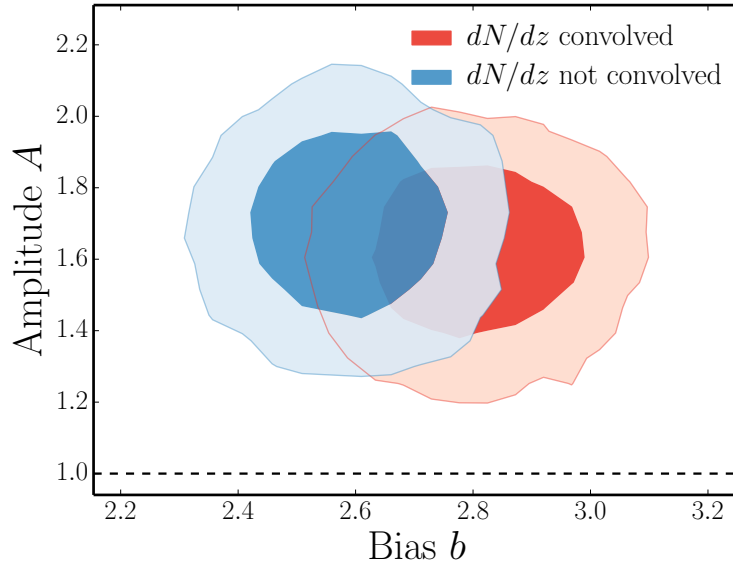
**Figure 4.19:** Posterior distributions for *A* and *b* obtained using the convolved (red contours) and the unconvolved *dN/dz* (blue contours).

and galaxy surveys can provide powerful constraints on the evolution of density fluctuations, on the nature of the DE, and on properties of tracers of the matter distribution, provided that a good control of systematic errors for both data sets can be achieved.

The null hypothesis (no correlation) was rejected with a significance of about $20\,\sigma$ and the significance of the detection of the theoretically expected cross-correlation signal was found to be $10\,\sigma$. The reliability of this result was confirmed by several null tests. A joint analysis of the cross-spectrum and of the autospectrum of the galaxy density contrast yielded a galaxy bias parameter of $b = 2.80^{+0.12}_{-0.11}$, consistent with earlier estimates for H-ATLAS galaxies at similar redshifts. On the other hand, the amplitude of the cross-correlation was found to be a factor $1.62 \pm 0.16$ higher than expected from the standard model.

We have investigated possible reasons for the excess amplitude. Some of them, such as the redshift dependence of the bias parameter or the contamination by the Sunyaev-Zel'dovich effect, were found to be negligible. Others, such as the magnification bias due to weak gravitational lensing or errors in the photometrically estimated redshifts, can contribute significantly to the observed excess but cannot fully account for it. A possible culprit is some residual contamination of convergence maps by unresolved infrared sources (Osborne et al., 2014; van Engelen et al., 2014), adding a substantial contribution to the measured correlation between the lensing convergence and the H-ATLAS high-*z* sources, which are unresolved by *Planck*. However, a detailed calculation of this effect is complicated and beyond the scope

of the present chapter.

We have also investigated the possibility that the tension between the observed and the expected cross-correlation amplitude was overrated because the noise level of the convergence maps in the regions used for the cross correlation is above typical values. This turned out to be the case in the three GAMA fields, but the effect on the combination of fields was found to be marginal.

An exquisite mapping of the CMB lensing pattern is one of the major goals of operating and planned CMB probes because of its relevance in studying cosmological structure formation and the properties of the dark energy. Forthcoming data releases by *Planck*[8] as well as future CMB lensing measurements from suborbital probes will be most relevant to further address the results presented here and improve the constraining power of these studies, both in a cosmological and astrophysical context.

---

[8]The CMB lensing map from 2015 *Planck* release will be used in the next chapter, while the definitive one, within the last data release expected for late 2016 could not be included in the present analysis.

# 5

# Toward a tomographic analysis of the Planck-Herschel cross-correlation

> I'm looking through you, where did you go
> I thought I knew you, what did I know
> You don't look different, but you have changed
> I'm looking through you, you're not the same
>
> —I'm Looking Through You, THE BEATLES

We present an improved and extended analysis of the cross-correlation between the map of the CMB lensing potential derived from the *Planck* mission data and the high-$z$ galaxies detected by H-ATLAS in the photometric redshift range $z_{\rm ph} \geq 1.5$. We compare the results of the previous chapter, based on the 2013 *Planck* dataset, with the 2015 release and investigate the impact of different selections of the H-ATLAS galaxy samples. Significant improvements over our previous analysis have been achieved thanks to the higher signal-to-noise ratio of the 2015 CMB lensing map released by the *Planck* collaboration. The effective galaxy bias parameter, $b$, for the full galaxy sample, derived from a joint analysis of the cross-power spectrum and of the galaxy auto-power spectrum is found to be $b = 3.54^{+0.15}_{-0.14}$. Furthermore, a first tomographic analysis of the cross-correlation signal is implemented, by splitting the galaxy sample into two redshift intervals: $1.5 \leq z_{\rm ph} < 2.1$ and $z_{\rm ph} \geq 2.1$. A statistically significant signal was found for both bins, indicating a substantial increase with redshift of the bias parameter: $b = 2.89 \pm 0.23$ for the lower and $b = 4.75^{+0.24}_{-0.25}$ for the higher redshift bin. Consistently with our previous analysis we find that the amplitude of the cross correlation signal is a factor of $1.45^{+0.14}_{-0.13}$ higher than expected from the standard $\Lambda$CDM model for the assumed redshift distribution. The robustness of our results against possible systematic effects has been extensively discussed although the tension is mitigated by passing from 4 to $3\sigma$.

# 5.1 Improvements over previous analysis

In this chapter we revisit the angular cross-power spectrum $C_\ell^{\kappa g}$ between the CMB convergence derived from *Planck* data and the spatial distribution of high-$z$ sub-mm galaxies detected by H-ATLAS (Eales et al., 2010). The present analysis improves over that presented in our previous chapter and paper (Bianchini et al., 2015, hereafter B15) in several aspects:

- we adopt the new *Planck* CMB lensing map (Planck Collaboration, 2016a, see Sec. (5.2.1));

- we treat more carefully the uncertainty in the photometric redshift estimates of the H-ATLAS galaxy sample (see Sec. (4.3.2));

- we move toward a *tomographic* study of the cross-correlation signal (see Sec. (5.4.2)).

The rest of this chapter is organized as follows: in Sec. (5.2) we describe the novel aspects of the handling of datasets, while the analysis method is presented in Sec. (5.3). In Sec. (5.4) we report and analyze the derived constraints on the galaxy bias parameter, discussing potential systematic effects that can affect the cross-correlation. Finally in Sec. (5.5) we summarize our results.

The theoretical background and cross-correlation formalism has been extensively presented in the previous chapters, in particular we refer the reader to Ch. (4.2) for the modeling of the expected signal.

The expected signal-to-noise (S/N) of the detection for the CMB convergence-density correlation can be estimated assuming that both fields behave as Gaussian random fields, so that the variance of $C_\ell^{\kappa g}$ is

$$\left(\Delta C_\ell^{\kappa g}\right)^2 = \frac{1}{(2\ell+1)f_{\text{sky}}}\Big[(C_\ell^{\kappa g})^2 + (C_\ell^{\kappa\kappa} + N_\ell^{\kappa\kappa})(C_\ell^{gg} + N_\ell^{gg})\Big], \qquad (5.1)$$

where $f_{\text{sky}}$ is the sky fraction observed by both the galaxy and the lensing surveys, $N_\ell^{\kappa\kappa}$ is the CMB lensing reconstruction noise level, $N_\ell^{gg} = 1/\bar{n}$ is the shot noise associated with the galaxy field, and $\bar{n}$ is the mean number of sources per steradian. For the H-ATLAS - *Planck* CMB lensing cross-correlation the sky coverage is approximately 600 deg$^2$ ($f_{\text{sky}} \simeq 0.01$) and we assume a constant bias $b = 3$ and a slope of the galaxy number counts $\alpha = 3$: restricting the analysis between $\ell_{min} = 100$ (as lower multipoles are poorly reconstructed) and $\ell_{max} = 800$, we expect $S/N \simeq 7.5$.

## 5.2 Improved CMB lensing and DM tracers from data

### 5.2.1 *Planck* data

We make use of the publicly released 2015 *Planck* CMB lensing map (Planck Collaboration, 2016a) that has been extracted by applying a quadratic estimator (Okamoto & Hu, 2003) to foreground-cleaned temperature and polarization maps. These maps have been synthesized from the raw 2015 *Planck* full mission frequency maps using the SMICA code (Planck Collaboration, 2016a). In particular, the released map is based on a minimum-variance (MV) combination of all temperature and polarization estimators, and is provided as a mean-field bias subtracted convergence $\kappa$ map.

For a comparison, we also use the earlier CMB lensing data provided within the *Planck* 2013 release (Planck Collaboration, 2014c). Differently from the 2015 case, the previous lensing map is based on a MV combination of only the 2013 *Planck* 143 and 217 GHz foreground-cleaned temperature anisotropy maps. The lensing mask associated to the 2015 release covers a slightly larger portion of the sky with respect to the 2013 release: $f_{\rm sky}^{2015}/f_{\rm sky}^{2013} \simeq 0.98$.

The map is in the HEALPix format with a resolution parameter $N_{side} = 2048$. We downgraded them to a resolution of $N_{side} = 512$ (corresponding to an angular resolution of $\sim 7'.2$).

### 5.2.2 *Herschel* data

Our sample of sub-mm galaxies is extracted from the same internal release of the full H-ATLAS catalogue as in B15. The survey area is divided into five fields: NGP, SGP and the three GAMA fields (G09, G12, G15). As previously mentioned (see Ch. (3.3.2) and (4.3.2)), the H-ATLAS galaxies have a broad redshift distribution extending from $z = 0$ to $z \simeq 5$ (Pearson et al., 2013). We are interested in selecting the high-$z$ ($z \gtrsim 1$) galaxies.

We have selected a sub-sample of H-ATLAS galaxies complying with the following criteria:

1. flux density at $250\,\mu$m, $S_{250} > 35$ mJy;

2. $\geq 3\,\sigma$ detection at $350\,\mu$m

3. photometric redshift greater than a given value, $z_{{\rm ph},min}$, as discussed below.

For our baseline analysis we set $z_{\mathrm{ph},min} = 1.5$. The sample comprises 94,825 sources. It was subdivided into two redshift bins ($1.5 \leq z_{\mathrm{ph}} < 2.1$ and $z_{\mathrm{ph}} \geq 2.1$) containing a similar number of sources (53,071 and 40,945, respectively), thus comparable shot-noise.

The estimate of the *unit-normalized* redshift distributions $dN/dz$ to be plugged into Eq. (4.6), hence into Eq. (4.7), is a very delicate process because of the very limited spectroscopic information. Also, the fraction of H-ATLAS sources having accurate photometric redshift determinations based on multi-wavelength optical/near-infrared photometry is rapidly decreasing with increasing redshift above $z \simeq 0.4$ (Smith et al., 2011; Bourne et al., 2016). However, if a typical rest-frame SED of H-ATLAS galaxies can be identified, it can be used to estimate the redshifts directly from *Herschel* photometric data.

Lapi et al. (2011) and González-Nuevo et al. (2012) showed that the SED of SMM J2135-0102, 'The Cosmic Eyelash' at $z = 2.3$ (Ivison et al., 2010; Swinbank et al., 2010) is a good template for $z \gtrsim 1$. Comparing the photometric redshift obtained with this SED with spectroscopic measurements for the 36 H-ATLAS galaxies at $z \gtrsim 1$ for which spectroscopic redshifts were available González-Nuevo et al. (2012) found a median value of $\Delta z/(1+z) \equiv (z_{\mathrm{phot}} - z_{spec})/(1+z_{spec}) = -0.002$ with a dispersion $\sigma_{\Delta z/(1+z)} = 0.115$. At lower redshifts this template performs much worse. As argued by Lapi et al. (2011) this is because the far-IR/sub-mm SEDs of H-ATLAS galaxies at $z > 1$ are dominated by the warm dust component while the cold dust component becomes increasingly important with decreasing $z$, amplifying the redshift–dust temperature degeneracy. That's why we restrict our analysis to $z_{\mathrm{ph}} \geq 1.5$.

Pearson et al. (2013) generated an average template for $z > 0.4$ H-ATLAS sources using a subset of 53 H-ATLAS sources with measured redshifts in the range $0.4 < z < 4.2$. They found that the redshifts estimated with this template have an average offset from spectroscopic redshift of $\Delta z/(1 + z) = 0.018$ with a dispersion $\sigma_{\Delta z/(1+z)} = 0.26$.

In the following we will use the SED of SMM J2135-0102 as our baseline template; the effect of using the template by Pearson et al. (2013) is presented Sec.. (5.4.4). To allow for the effect on $dN/dz$ of random errors in photometric redshifts we estimated, following Budavari et al. (2003), the redshift distribution, $p(z|W)$, of galaxies selected by our window function $W(z_{\mathrm{ph}})$, as

$$p(z|W) = p(z) \int dz_{\mathrm{ph}} W(z_{\mathrm{ph}}) p(z_{\mathrm{ph}}|z), \qquad (5.2)$$

where $p(z)$ is the fiducial true redshift distribution, $W = 1$ for $z_{\mathrm{ph}}$ in the selected

**Table 5.1:** Statistics of H-ATLAS fields

| | $N_{obj}$ | | | $\bar{n}$ [gal ster$^{-1}$] | | |
|---|---|---|---|---|---|---|
| Patch | $z_{\mathrm{ph}} \geq 1.5$ | $1.5 \leq z_{\mathrm{ph}} < 2.1$ | $z_{\mathrm{ph}} \geq 2.1$ | $z_{\mathrm{ph}} \geq 1.5$ | $1.5 \leq z_{\mathrm{ph}} < 2.1$ | $z_{\mathrm{ph}} \geq 2.1$ |
| ALL[a] | 94825 | 53071 | 40945 | $5.76 \times 10^5$ | $3.22 \times 10^5$ | $2.49 \times 10^5$ |
| NGP | 26303 | 15033 | 11039 | $5.63 \times 10^5$ | $3.22 \times 10^5$ | $2.36 \times 10^5$ |
| SGP | 43518 | 24722 | 18422 | $5.95 \times 10^5$ | $3.38 \times 10^5$ | $2.52 \times 10^5$ |
| G09 | 8578 | 4590 | 3922 | $5.72 \times 10^5$ | $3.02 \times 10^5$ | $2.61 \times 10^5$ |
| G12 | 8577 | 4611 | 3881 | $5.34 \times 10^5$ | $2.87 \times 10^5$ | $2.41 \times 10^5$ |
| G15 | 7849 | 4115 | 3681 | $5.66 \times 10^5$ | $2.97 \times 10^5$ | $2.65 \times 10^5$ |

[a] ALL is the combination of all fields together.

interval and $W = 0$ otherwise, and $p(z_{\mathrm{ph}}|z)$ is probability that a galaxy with a true redshift $z$ has a photometric redshift $z_{\mathrm{ph}}$. The error function $p(z_{\mathrm{ph}}|z)$ is parameterized as a Gaussian distribution with zero mean and dispersion $(1 + z)\,\sigma_{\Delta z/(1+z)}$. For the dispersion we adopt the conservative value $\sigma_{\Delta z/(1+z)} = 0.26$.

A partial estimate of the effect of the dust temperature–redshift degeneracy in contaminating our $z > z_{\mathrm{ph,min}}$ sample derived from *Herschel* colors by cold low-$z$ galaxies is possible thanks to the work by Bourne et al. (2016) who used a likelihood-ratio technique to identify SDSS counterparts at $r < 22.4$ for H-ATLAS sources in the GAMA fields and collected spectroscopic and photometric redshifts from GAMA and other public catalogues. A cross-match with their data set showed that about 7% of sources in GAMA fields with estimated redshifts larger than 1.5 based on *Herschel* colors have a reliable ($R \geq 0.8$) optical/near-IR counterpart with photometric redshift $< 1$. The fiducial redshift distribution for the GAMA fields was corrected by moving these objects and the corrected, unit normalized, redshift distribution was adopted for the full sample. The result is shown in Fig. (5.1) for $z \geq 1.5$ and for the sub-sets at $1.5 \leq z_{\mathrm{ph}} \leq 2.1$ and $z_{\mathrm{ph}} > 2.1$. As mentioned above, photometric redshifts based on *Herschel* colors become increasingly inaccurate below $z \sim 1$. Thus the low-$z$ portions of the $p(z)$'s in Fig. (5.1) are unreliable.

For each H-ATLAS field we created an overdensity map in `HEALPix` format with a resolution parameter $N_{side} = 512$. The overdensity is defined as $g(\hat{\mathbf{n}}) = n(\hat{\mathbf{n}})/\bar{n} - 1$, where $n(\hat{\mathbf{n}})$ is the number of objects in a given pixel, and $\bar{n}$ is the mean number of objects per pixel. As a last step, we combined the *Planck* lensing mask with the H-ATLAS one. The total sky fraction retained for the analysis is $f_{\mathrm{sky}} = 0.013$. The specifics of each patch are summarized in Table (5.1).
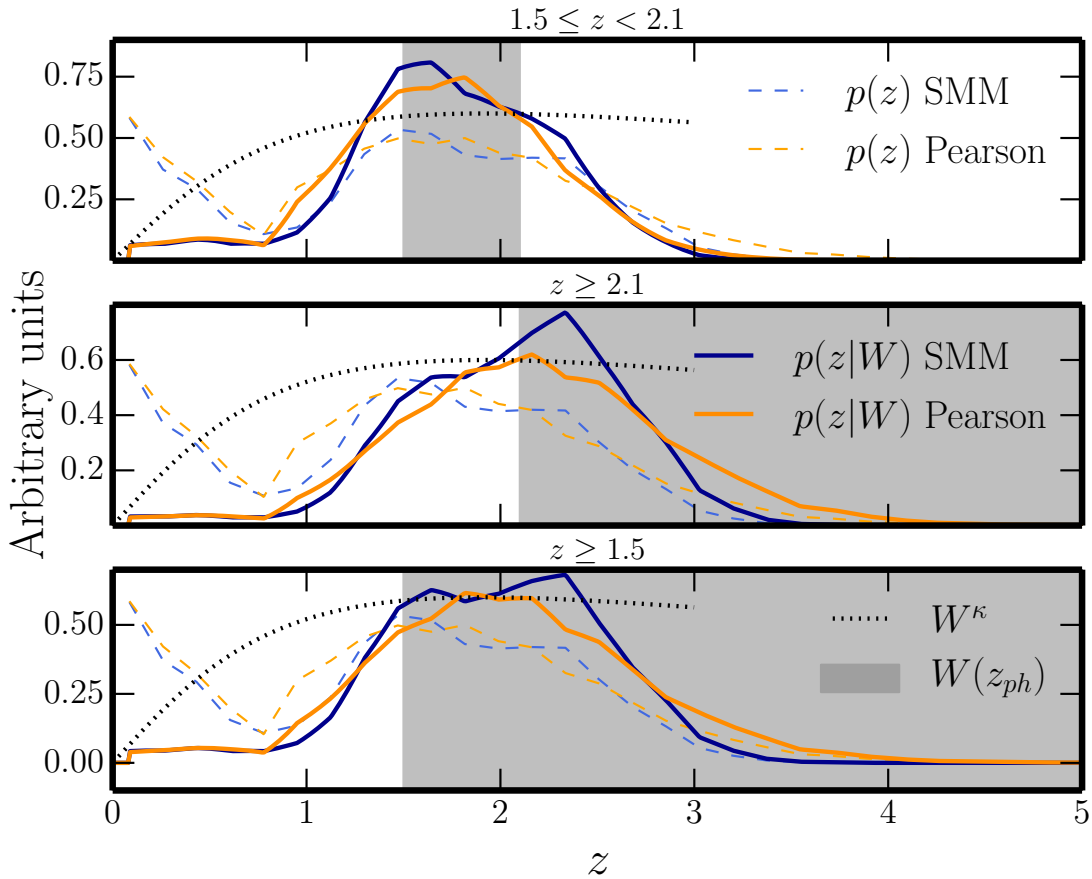
**Figure 5.1:** Redshift distributions of the galaxy samples selected from the full H-ATLAS survey catalogue. The fiducial *true* redshift distribution $p(z)$ for the full sample is shown, in each panel, by the dashed lines, while the solid lines show the redshift distributions $p(z|W)$ obtained implementing the top-hat window functions $W(z_{\mathrm{ph}})$ represented, in each panel, by the shaded area. The blue and the orange lines refer to redshift distributions based on the SED of SMM J2135-0102 and on the Pearson et al. (2013) best fitting template, respectively (see Sec.. (5.4.4)). These redshift distributions were used for the evaluation of the theoretical $C_\ell$'s. The dotted black line in each panel shows the arbitrarily normalized CMB lensing kernel $W^\kappa(z)$.

## 5.3 Analysis method

### 5.3.1 Estimation of the power spectra

Similarly to Ch .(4), we estimate cross- and auto-power spectra using a PCL method based on the `MASTER` algorithm (Hivon et al., 2002) (see also Ch. (3.2) and (4.4)), reconstructing the signal in seven linearly spaced bins of width $\Delta\ell = 100$ spanning the multipole range from 100 to 800.

In order to estimate the full covariance matrix and the error bars we make use of the publicly available set of 100 realistic CMB convergence simulations[1], that accurately reflect the *Planck* 2015 noise properties, and cross-correlate them with the H-ATLAS galaxy density contrast maps. Because there is no correlated cosmological signal between CMB lensing simulations and real galaxy datasets, we also use them to check that our pipeline does not introduce any spurious signal. The mean cross-spectrum between the *Planck* simulations and the H-ATLAS maps is shown in Fig. (5.2) which shows that it is consistent with zero in all redshift bins. For the baseline photo-$z$ bin we obtain $\chi^2 = 9.5$ for $\nu = 7$ degrees-of-freedom (d.o.f.), corresponding to a probability of random deviates with the same covariances to exceed this chi-squared (p-value) of 0.22. In the other two redshift bins we find $\chi^2 = 12.6$ for the low-$z$ one and $\chi^2 = 6.1$ for the high-$z$ one, corresponding to a p-value $p = 0.08$ and $p = 0.53$ respectively.

### 5.3.2 Estimation of the cross-correlation amplitude and of the galaxy bias

Following B15 we introduce a phenomenologically-motivated amplitude parameter $A$ that globally scales the observed cross-power spectrum with respect to the theoretical one as $\hat{C}_L^{\kappa g} = AC_L^{\kappa g}(b)$. We analyze the constraints on the parameters $A$ and $b$ combining the information from the cross-spectrum and the galaxy auto-spectrum. For the joint analysis we exploit Gaussian likelihood functions that take into account correlations between the cross- and the auto-power spectra in the covariance matrix. The extracted cross- and auto-band-powers are organized into a single data vector as

$$\hat{\mathbf{C}}_L = (\hat{\mathbf{C}}_L^{\kappa g}, \hat{\mathbf{C}}_L^{gg}), \tag{5.3}$$

---

[1] http://irsa.ipac.caltech.edu/data/Planck/release_2/all-sky-maps/maps/
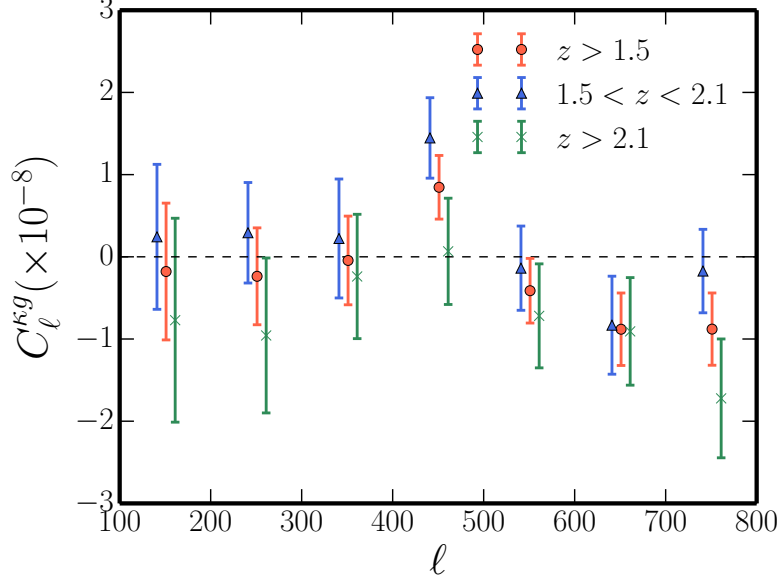component-maps/lensing

**Figure 5.2:** Null test results. Mean cross-spectrum $C_\ell^{\kappa g}$ between $N_{sim} = 100$ simulated *Planck* CMB lensing maps and the H-ATLAS galaxy density maps for the three redshift bins considered: $z \geq 1.5$ (blue circles), $1.5 \leq z < 2.1$ (green crosses), and $z \geq 2.1$ (red triangles). Band-powers are displaced by $\Delta\ell = \pm 10$ with respect to the bin centers for visual clarity. The error bars are calculated as the square root of the covariance matrix diagonal derived from the same set of simulations and divided by $\sqrt{N_{sim}}$.

which has $N_L = 14$ elements. The total covariance matrix is then written as the composition of four $7 \times 7$ submatrices:

$$\mathbf{Cov}_{LL'} = \begin{bmatrix} \mathrm{Cov}_{LL'}^{\kappa g} & (\mathrm{Cov}_{LL'}^{\kappa g - gg})^{\mathsf{T}} \\ \mathrm{Cov}_{LL'}^{\kappa g - gg} & \mathrm{Cov}_{LL'}^{gg} \end{bmatrix}. \tag{5.4}$$

The covariance matrices are approximately given by (see Ch. (3.2.2)):

$$\begin{aligned}
\mathrm{Cov}_{LL'}^{gg} &= \frac{2}{(2L+1)\Delta L f_{\mathrm{sky}}} \left[ C_L^{gg}(\boldsymbol{\theta}) + N_L^{gg} \right]^2 \delta_{LL'}; \\
\mathrm{Cov}_{LL'}^{\kappa g} &= \frac{1}{(2L+1)\Delta L f_{\mathrm{sky}}} \\
&\quad \times \left[ (C_L^{\kappa g}(\boldsymbol{\theta}))^2 + (C_L^{\kappa\kappa} + N_L^{\kappa\kappa})(C_L^{gg}(\boldsymbol{\theta}) + N_L^{gg}) \right] \delta_{LL'}; \\
\mathrm{Cov}_{LL'}^{\kappa g - gg} &= \frac{2}{(2L+1)\Delta L f_{\mathrm{sky}}} \left[ (C_L^{gg}(\boldsymbol{\theta}) + N_L^{gg}) C_L^{\kappa g}(\boldsymbol{\theta}) \right] \delta_{LL'},
\end{aligned} \tag{5.5}$$

where the $\Delta L$ is the bin size, $\boldsymbol{\theta}$ is the parameters vector, and $\delta_{LL'}$ is the Kronecker delta such that they are diagonal. Then, the likelihood function can be written as

$$
\begin{aligned}
\mathcal{L}(\hat{\mathbf{C}}_L|\boldsymbol{\theta}) &= (2\pi)^{-N_L/2}[\det \mathbf{Cov}_{LL'}]^{-1/2} \\
&\times \exp\left\{-\frac{1}{2}\big[\hat{\mathbf{C}}_L - \mathbf{C}_L(\boldsymbol{\theta})\big]\big(\mathbf{Cov}_{LL'}\big)^{-1}\big[\hat{\mathbf{C}}_{L'} - \mathbf{C}_{L'}(\boldsymbol{\theta})\big]\right\}.
\end{aligned}
\tag{5.6}
$$

The parameter space is explored using `emcee`[2], an affine invariant Markov Chain Monte Carlo (MCMC) sampler (Foreman-Mackey et al., 2013), assuming flat priors over the ranges $\{b, A, \mathcal{A}_{\text{bias}}\} = \{[0, 10], [-1, 10], [0, 10]\}$[3] ($\mathcal{A}_{\text{bias}}$ will be defined in Sec.. (5.4.5)). This analysis scheme is applied independently to each redshift bin.

The covariance matrices built with the 100 *Planck* lensing simulations were used to compute the error bars for the cross-power spectra (the ones shown in Figs. (5.2), (5.4), (5.6), (5.11)), to address bin-to-bin correlations and to evaluate the chi-square for the null-hypothesis rejection. On the other hand, we used the diagonal analytical approximation of Eq. (5.5) to evaluate the bias-dependent covariance matrices used to sample the posterior distribution and for error bars on the galaxy power spectra estimation (error bars shown in Figs. (5.5), (5.7), (5.9), (5.12)). As in B15, we decided to rely on an analytical approximation of the covariance matrices that depend on the estimated parameters, i.e. the linear galaxy bias. This simple approximation is able to capture the covariance matrices features as shown in Sec. (5.4.2) where we compare results obtained with (i) the diagonal approximation given by Eq. (5.5); (ii) the non-diagonal (bias-dependent) analytical prescription derived and exploited in B15 that accounts for the mask induced mode-coupling; (iii) and the full covariance matrix evaluated from the set of 100 Planck CMB lensing simulations as described above.

## 5.4 Results and discussion

Before performing a thorough quantitative analysis with the data, we can qualitatively show that the galaxy density field is correlated with the lensing convergence, which means that peaks in the galaxy distribution correspond to regions of enhanced lensing signal: this is shown in the right panel of Fig. (5.3). To output such plot we have smoothed both the CMB convergence and galaxy maps with a 0.5 deg Gaussian filter, binned the density field $\delta$ and found the corresponding pixels in the $\kappa$ maps corresponding to the given density bins. Then, we took the average

---

[2] http://dan.iel.fm/emcee
[3] Note that we constrain separately $(b, A)$ and $(A, \mathcal{A}_{\text{bias}})$.
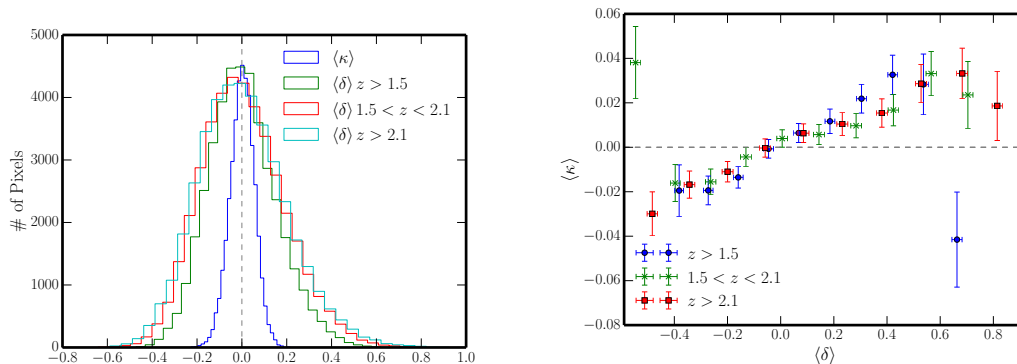
**Figure 5.3:** *Left panel:* Histograms of pixel values in CMB convergence and galaxy density in $z-$bins maps, after applying Gaussian smoothing with $\sigma = 0.5$ deg. *Left panel*: binned values of the H-ATLAS galaxy overdensity samples in the three $z$-bins versus the averages CMB convergence in corresponding pixels. The position of points center along $\delta$ is given by the average density in each bin, while horizontal error bars are given by the scatter in $\delta$ in the same bin. Vertical error bars are obtained by measuring the scatter in the same measurements when using the set of 100 *Planck* lensing simulations in place of the real lensing map.

value of $\delta$ and $\kappa$ in each bin and plot these against each other, showing a strong correlation for all data splits (photo-$z$ bins).

## 5.4.1 Comparison between the 2013 and the 2015 *Planck* results

Figure (5.4) compares the cross-spectra $C_\ell^{\kappa g}$ between the $z_{\mathrm{ph}} \geq 1.5$ H-ATLAS galaxy sample and the 2013/2015 *Planck* CMB lensing maps. For a fuller comparison, the figure also shows the results obtained using the galaxy catalogue built by B15 adding to the requirements (i-iii) mentioned in Sec.. (4.3.2) the color criteria introduced by González-Nuevo et al. (2012) (hereafter GN12): $S_{350\,\mu\mathrm{m}}/S_{250\,\mu\mathrm{m}} > 0.6$ and $S_{500\,\mu\mathrm{m}}/S_{350\,\mu\mathrm{m}} > 0.4$. The error bars were derived by cross-correlating the 100 simulated *Planck* lensing realizations with the sub-mm galaxy map and measuring the variance in $C_\ell^{\kappa g}$.

The exploitation of the 2015 CMB lensing map has the effect of shrinking the error bars, on average, by approximately 15% with respect to the previous data release due to the augmented *Planck* sensitivity. All shifts in the cross-power spectra based on the 2013 and on the 2015 releases are within $1\,\sigma$. As illustrated by Fig. (5.5), the auto-power spectra, $C_\ell^{gg}$, of H-ATLAS galaxies in the present sample and in the B15 one are consistent with each other: differences are well within $1\,\sigma$. Table (5.2) shows that the various combinations of lensing maps, galaxy catalogues
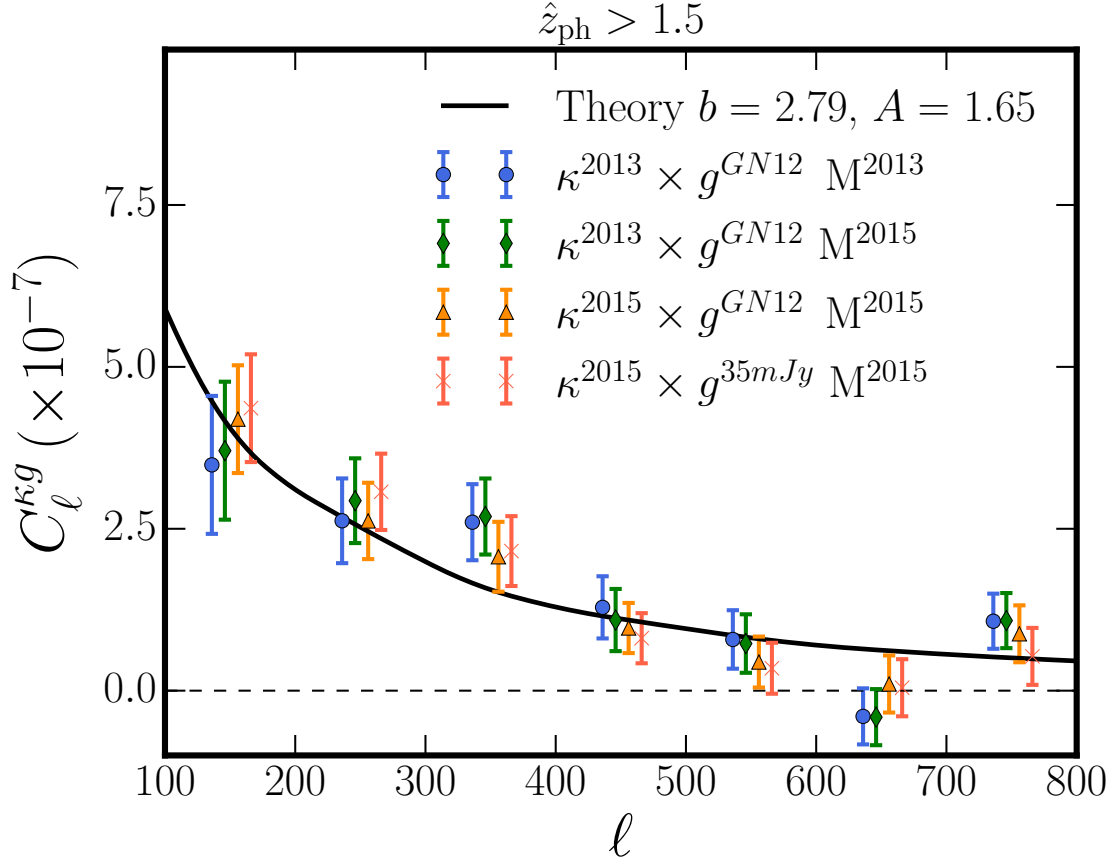
**Figure 5.4:** Comparison of the cross-spectra between the *Planck* CMB lensing maps and the $z_{\rm ph} \geq 1.5$ H-ATLAS galaxy density maps obtained using the 2013 and the 2015 *Planck* results. The results labelled "$[k^{2015} \times g^{35{\rm m}Jy}]M^{2015}$" refer to the analysis done with the present sample of $z_{\rm ph} \geq 1.5$ galaxies. The GN12 superscript refers to the sample of H-ATLAS galaxies used by B15, based on slightly more restrictive selection criteria; results for this sample are shown for both the 2013 or the 2015 *Planck* masks ($M^{2013}$ and $M^{2015}$) and convergence maps ($\kappa^{2013}$ and $\kappa^{2015}$). The solid black line shows the theoretical cross-spectrum for the best-fit values of the bias factor and of the cross-correlation amplitude, $A$, found for the $[\kappa^{2015} \times g^{35{\rm m}Jy}]M^{2015}$ adopting the redshift distribution estimated by B15. The estimated bandpowers are plotted with an offset along the $x$-axis for a better visualization. The error-bars were computed using the full covariance matrix obtained via Monte Carlo simulations as $\Delta C_L^{\kappa g} = \sqrt{{\rm diag}({\rm Cov}^{\kappa g})}$.
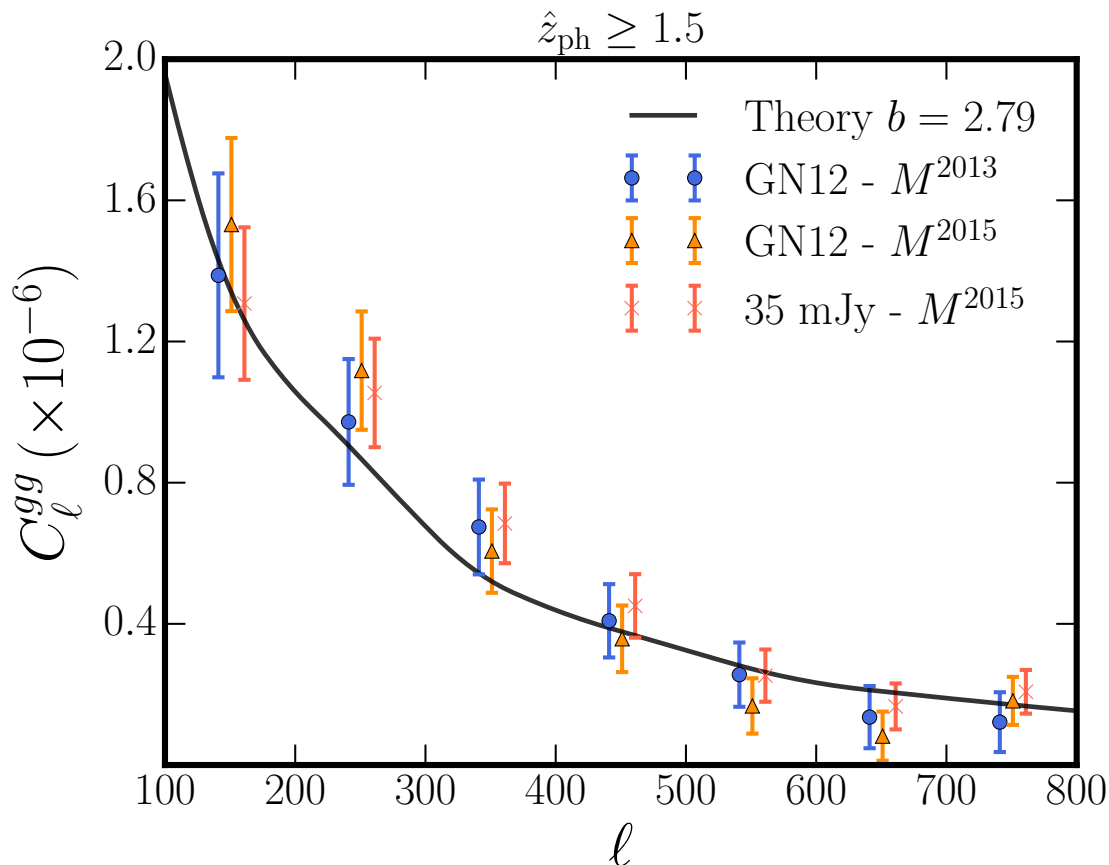
**Figure 5.5:** Comparison of the auto-power spectra of H-ATLAS galaxies with $z_{\text{ph}} \geq 1.5$ in the present sample and in that selected by B15 (GN12) using slightly more restrictive criteria. The effect of using different masks (the 2013 and 2015 *Planck* masks) is also shown. The solid black line represents the auto-power spectra for the best-fit value of the bias parameter given in the inset (see also Table (5.2)) and for the redshift distribution of B15. The estimated bandpowers are plotted with an offset with respect to the bin centers for a better visualization. The error bars are computed using the analytical prescription ($\sqrt{\text{diag}(\text{Cov}^{gg})}$ with $\text{Cov}^{gg}$ given by Eq. (5.5) and evaluated using the estimated bandpowers).

and masks we have considered in Fig. (5.4) lead to very similar values of the *A* and *b* parameters, *if the redshift distribution of B15 is used*. Note that the errors on parameters given in Table (5.2) as well as in the following similar tables, are slightly smaller than those that could be inferred from the corresponding figures. This is because the errors on each parameter given in the tables are obtained marginalizing over the other parameter.

| Datasets | Mask | $b$ | A |
|---|---|---|---|
| $\kappa^{2013} \times g^{GN12}$ | 2013 | $2.80^{+0.12}_{-0.11}$ | $1.62^{+0.16}_{-0.16}$ |
| $\kappa^{2013} \times g^{GN12}$ | 2015 | $2.86^{+0.12}_{-0.12}$ | $1.68^{+0.19}_{-0.19}$ |
| $\kappa^{2015} \times g^{GN12}$ | 2015 | $2.85^{+0.12}_{-0.12}$ | $1.61^{+0.16}_{-0.16}$ |
| $\kappa^{2015} \times g^{35\text{mJy}}$ | 2015 | $2.79^{+0.12}_{-0.12}$ | $1.65^{+0.16}_{-0.16}$ |

[a] The analysis is performed on the $z_{\text{ph}} \geq 1.5$ sample for consistency with B15.

**Table 5.2:** Comparison of the $\{b, A\}$ values obtained from the joint $\kappa g + gg$ analysis for the combinations of maps and masks reported in Fig. (5.4) and adopting the redshift distribution of B15.

## 5.4.2 Tomographic analysis

As discussed in the previous sub-section, if we use the redshift distribution of B15 the impact of the new *Planck* convergence map and mask, and of the new H-ATLAS overdensity map on the $A$ and $b$ parameters is very low. However, significant differences are found using the new redshift distribution for $z_{\text{ph}} \geq 1.5$ built in this chapter and shown by the blue solid line in the bottom panel of Fig. (5.1). Compared to B15, the best fit value of the bias parameter increases from $b = 2.80^{+0.12}_{-0.11}$ to $b = 3.54^{+0.15}_{-0.14}$ and the cross-correlation amplitude decreases from $A = 1.62 \pm 0.16$ to $A = 1.45^{+0.14}_{-0.13}$ (see Tables (5.2) and (5.3)).

As in B15, we get a highly significant detection of the cross-correlation, at $A/\sigma_A \simeq 10.3\,\sigma$ and again a value of $A$ higher than the expected $A = 1$ is indicated by the data.

Figure (5.6) shows the cross-correlation power spectrum for the 3 redshift intervals we have considered. The error bars were estimated with Monte Carlo simulations as described above. Their relative sizes scale, as expected, with the number of objects in each photo-$z$ interval, reported in Table (5.1). In all cases, the detection of the signal is highly significant. The chi-square value for the null hypothesis, i.e. no correlation between the two cosmic fields, computed taking into account bin-to-bin correlations, is $\chi^2_{\text{null}} = \hat{\mathbf{C}}^{\kappa g}_L \, (\text{Cov}^{\kappa g}_{LL'})^{-1} \, \hat{\mathbf{C}}^{\kappa g}_{L'} \simeq 88$ for $\nu = 7$ d.o.f., corresponding to a $\simeq 22\sigma$ rejection for the full sample ($z_{\text{ph}} \geq 1.5$). For the $1.5 \leq z_{\text{ph}} < 2.1$ and $z_{\text{ph}} \geq 2.1$ intervals we found $\chi^2_{null} \simeq 47$ and $\chi^2_{null} \simeq 64$, respectively, corresponding to $10.7\sigma$ and $15\sigma$ rejections.

There is a hint of a stronger cross-correlation signal for the higher redshift interval. The indication is however weak. A much stronger indication of an evolution of the clustering properties (increase with redshift of the bias factor) of galaxies is apparent in Fig. (5.7) and in Table (5.3). However the auto–power spectrum for
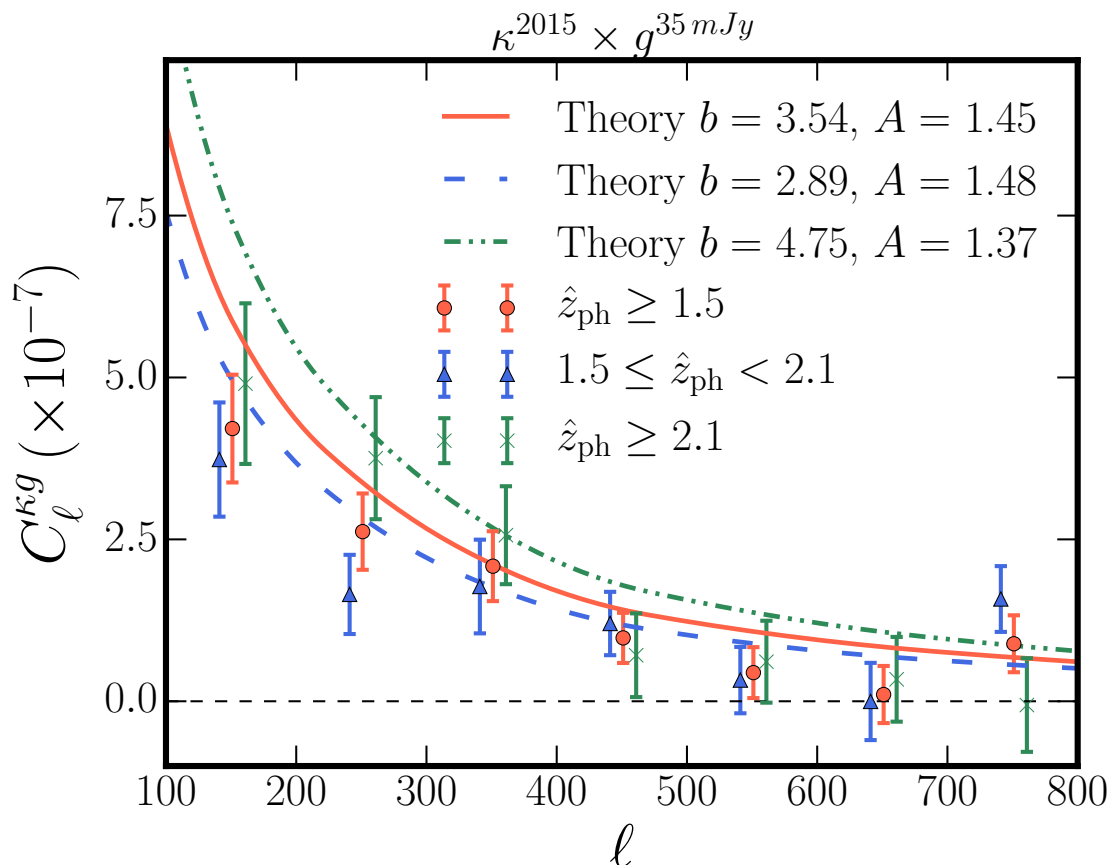
**Figure 5.6:** Cross-power spectra between the 2015 *Planck* CMB lensing map and the H-ATLAS galaxy sample for different redshift intervals: $z_{\rm ph} \geq 1.5$ (red circles), $1.5 \leq z_{\rm ph} < 2.1$ (blue triangles), and $z_{\rm ph} \geq 2.1$ (green crosses). Uncertainties are derived as for bandpowers in Fig. (5.4). The red solid, blue dashed and green dot-dashed lines are the corresponding cross-power spectra for the best-fit bias and amplitude parameters obtained combining the data on the auto- and cross-power spectra (see Table (5.3)). The adopted redshift distributions are shown by the blue lines in Fig. (5.1).
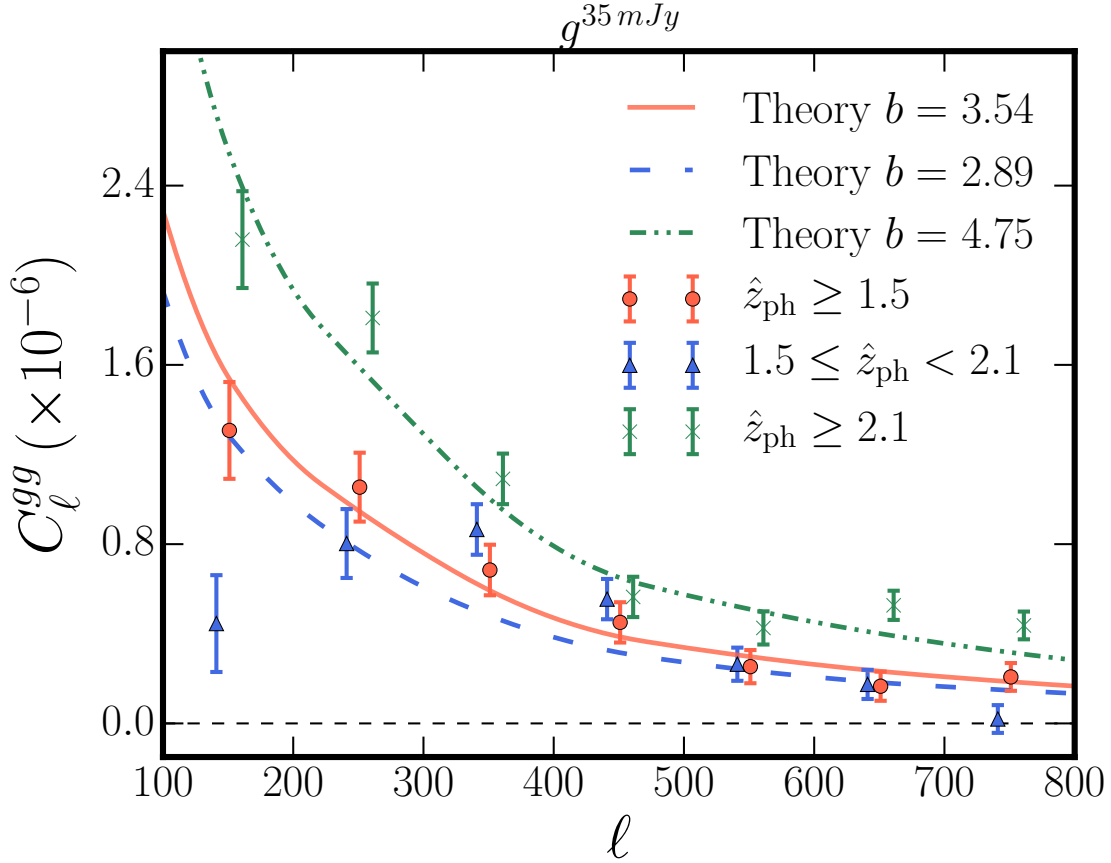
**Figure 5.7:** H-ATLAS galaxy auto-power spectra in different redshift intervals: $z_{\rm ph} \geq 1.5$ (red circles), $1.5 \leq z_{\rm ph} < 2.1$ (blue triangles), and $z_{\rm ph} \geq 2.1$ (green crosses). Uncertainties are derived as for bandpowers in Fig. (5.5). The red solid, blue dashed and green dot-dashed lines are the galaxy auto-power spectra for the *Planck* cosmology and the best-fit bias and amplitude found for the $z \geq 1.5$, $1.5 \leq z < 2.1$, and $z \geq 2.1$ photo-$z$ bins respectively. The theory lines refer to the $dN/dz$ built in this chapter and also used in Fig. (5.6).

**Table 5.3:** Linear bias and cross-correlation amplitude as determined using jointly the reconstructed galaxy auto- and cross-spectra in the different redshift bins.

| Bin | $b$ | $A$ | $\chi^2/\text{d.o.f.}$ | p-value |
|---|---|---|---|---|
| Diagonal covariance matrices approximation (Eq. (5.5)) | | | | |
| $z_{\text{ph}} \geq 1.5$ | $3.54^{+0.15}_{-0.14}$ | $1.45^{+0.14}_{-0.13}$ | $10.6/12$ | $0.56$ |
| $1.5 \leq z_{\text{ph}} < 2.1$ | $2.89^{+0.23}_{-0.23}$ | $1.48^{+0.20}_{-0.19}$ | $29.5/12$ | $0.003$ |
| $z_{\text{ph}} \geq 2.1$ | $4.75^{+0.24}_{-0.25}$ | $1.37^{+0.16}_{-0.16}$ | $9.6/12$ | $0.65$ |
| Non-diagonal covariance matrices approximation (Eq. (4.21)) | | | | |
| $z_{\text{ph}} \geq 1.5$ | $3.53^{+0.15}_{-0.15}$ | $1.45^{+0.14}_{-0.13}$ | $8.75/12$ | $0.72$ |
| $1.5 \leq z_{\text{ph}} < 2.1$ | $2.88^{+0.23}_{-0.25}$ | $1.48^{+0.20}_{-0.19}$ | $23.1/12$ | $0.03$ |
| $z_{\text{ph}} \geq 2.1$ | $4.74^{+0.24}_{-0.24}$ | $1.36^{+0.16}_{-0.16}$ | $8.5/12$ | $0.75$ |
| Covariance matrices from MC simulations | | | | |
| $z_{\text{ph}} \geq 1.5$ | $3.56^{+0.17}_{-0.17}$ | $1.39^{+0.15}_{-0.14}$ | $8.37/12$ | $0.76$ |
| $1.5 \leq z_{\text{ph}} < 2.1$ | $2.91^{+0.24}_{-0.24}$ | $1.48^{+0.22}_{-0.21}$ | $33.7/12$ | $7.5 \times 10^{-4}$ |
| $z_{\text{ph}} \geq 2.1$ | $4.80^{+0.25}_{-0.25}$ | $1.37^{+0.17}_{-0.17}$ | $14.5/12$ | $0.27$ |

[a] The redshift distributions derived in this chapter and shown by the black lines in Fig. (5.1) were adopted. The best-fit values and $1\sigma$ errors are evaluated respectively as the 50th, 16th, and 84th percentiles of the posterior distributions. The $\chi^2$'s are computed at the best-fit values. The results obtained including off-diagonal terms of the covariance matrices and using covariances based on simulations are also shown for comparison.

the $1.5 \leq z_{\text{ph}} < 2.1$ interval shows a puzzling lack of power in the first multipole bin. This feature, not observed in the cross-power spectrum for the same photo-$z$ bin, may be due to systematic errors in the photometric redshift estimate.

The 68% and 95% confidence regions for the amplitude $A$ and the bias $b$, obtained from their posterior distributions combining the data on auto- and cross-spectra, are shown in Fig. (5.8). We have $b = 2.89 \pm 0.23$ and $A = 1.48^{+0.20}_{-0.19}$ for the lower redshift bin and $b = 4.75^{+0.24}_{-0.25}$ and $A = 1.37 \pm 0.16$ for the higher-$z$ one (see Table (4.3)). The reduced $\chi^2$ associated with the best-fit values are close to unity, suggesting the consistency of the results, except for the $1.5 \leq z_{\text{ph}} < 2.1$ interval for which there is a large contribution to the $\chi^2$ from the auto-spectrum for the first multipole bin. In order to test the stability of the results with respect to the chosen covariance matrices estimation method, we redo the analysis with the non-diagonal approximation of B15 and the full covariance matrices from simulations: results
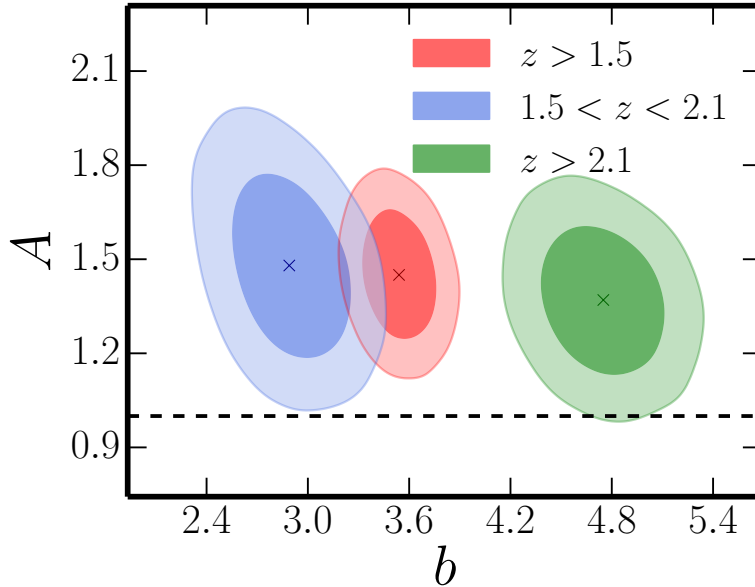
**Figure 5.8:** Posterior distributions in the $(b, A)$ plane with the 68% and 95% confidence regions (darker and lighter colors respectively) for the three redshift intervals: $z_{\rm ph} \geq 1.5$ (red contours), $1.5 \leq z_{\rm ph} < 2.1$ (blue contours), and $z_{\rm ph} \geq 2.1$ (green contours). The dashed line corresponds to the expected amplitude value $A = 1$ (a magnification bias parameter $\alpha = 3$ is assumed). The colored crosses mark the best-fit values reported in Table (4.3) for the three photo-z intervals.

are reported in Table (4.3). As can be seen, in the former case the inclusion of non-diagonal terms induced by mode-coupling results in negligible differences with respect to our baseline analysis scheme. In the latter case we observe a rather small broadening of the credibility contours (dependent on the $z$-bin), from 2% to 17% for $b$ and from 6% to 10% for the cross-correlation amplitude $A$, with the biggest differences reported for the baseline $z \geq 1.5$ bin. The central value of $A$ for the baseline redshift bin is diminished by approximately 4% even though $A > 1$ at $\gtrsim 2\sigma$. However, one might argue that the limited number of the available Planck CMB lensing simulations imposes limitations to the covariance matrices convergence. Given the stability of the results, we therefore adopt the diagonal approximation of Eq. (5.5) as our baseline covariance matrices estimation method.

### 5.4.3 Cross-correlation of galaxies in different redshift intervals

Both the auto- and the cross-power spectra depend on the assumed redshift distribution; hence the inferred values of the (constant) bias and of the amplitude are contingent on it. A test of the reliability of our estimates can be obtained from the cross-correlation $C_\ell^{g_1 g_2}$ between positions of galaxies in the lower redshift interval, $1.5 \leq z_{\rm ph} < 2.1$ (indexed by subscript 1), with those in the higher redshift

interval, $z_{\rm ph} \geq 2.1$ (subscript 2). Assuming, as we did in Eq. (4.6), that the observed galaxy density fluctuations can be written as the sum of a clustering term with a magnification bias one as $\delta^{\rm obs}(\hat{\bf n}) = \delta^{\rm cl}(\hat{\bf n}) + \delta^{\mu}(\hat{\bf n})$, the cross-correlation among galaxies in the two intervals can be decomposed into four terms:

$$C_\ell^{g_1 g_2} = C_\ell^{{\rm cl}_1 cl_2} + C_\ell^{{\rm cl}_1 \mu_2} + C_\ell^{\mu_1 {\rm cl}_2} + C_\ell^{\mu_1 \mu_2}. \tag{5.7}$$

The first term results from the intrinsic correlations of the galaxies of the two samples and it is due to the overlap between the two redshift distributions: if the two galaxy samples are separated in redshift, this term vanishes. The second term results from the lensing of background galaxies due to the matter distribution traced by the low-$z$ sample galaxies, while the third one is related to the opposite scenario: again, it is non-zero only if there is an overlap between the two $dN/dz$. The fourth term results from lensing induced by dark matter in front of both galaxy samples. The relative amplitude of these terms, compared to the observed galaxy cross-power spectrum, can provide useful insights on uncertainties in the redshift distributions.

The measured $\hat{C}_\ell^{g_1 g_2}$ is shown in Fig. (5.9). The expected contributions of the four aforementioned terms are computed using the bias values reported in Table (4.3), and the redshift distributions shown in Fig. (5.1). We remind that the assumed value for the rms uncertainty is $\sigma_{\Delta z/(1+z)} = 0.26$. The figure shows that the expected amplitude of the intrinsic correlation term is dominant with respect to the magnification bias related ones and that the observed signal is in good agreement with expectations. No signs of inconsistencies affecting redshift distributions are apparent.

### 5.4.4 Effect of different choices for the SED

The assumed SED plays a key role in the context of template fitting approaches aimed at photo-$z$ estimation. It is then crucial to test the robustness of the results presented here against different choices for it. To this end we constructed a catalogue with photo-$z$ estimates based on the best fitting SED template of Pearson et al. (2013) and applied the full analysis pipeline described in Sec. (5.3). In Fig. (5.10) we show the comparison between the estimates of photo-$z$ using the SED of SMM and the one of Pearson et al. (2013).

The cross- and auto-power spectra extracted adopting the SED template of Pearson et al. (2013) are shown in Fig. (5.11) and (5.12) respectively. In Fig. (5.13) we compare the credibility regions for the bias $b$ and cross-correlation amplitude $A$
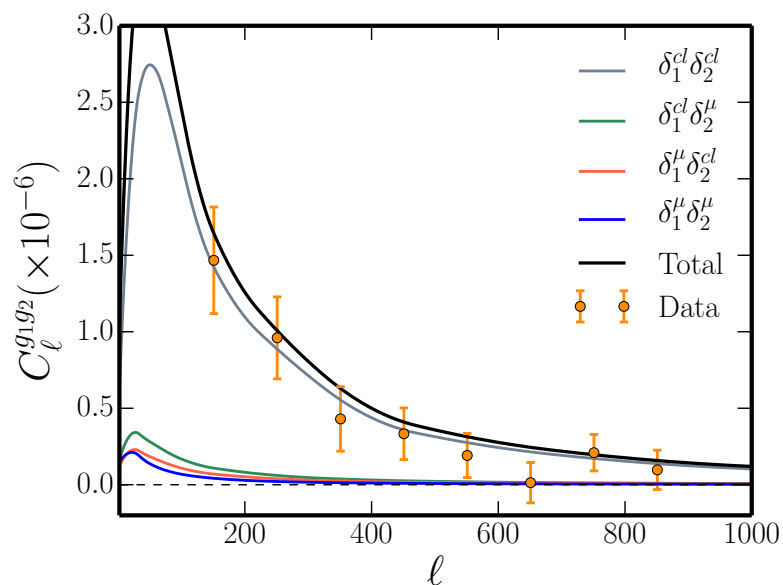
**Figure 5.9:** Cross-correlation of angular positions between galaxies in the low-$z$ and in the high-$z$ redshift interval. The solid lines show the expected contributions from the various terms appearing in Eq. (5.7). Note that the "Total" line is *not* a fit to the data.
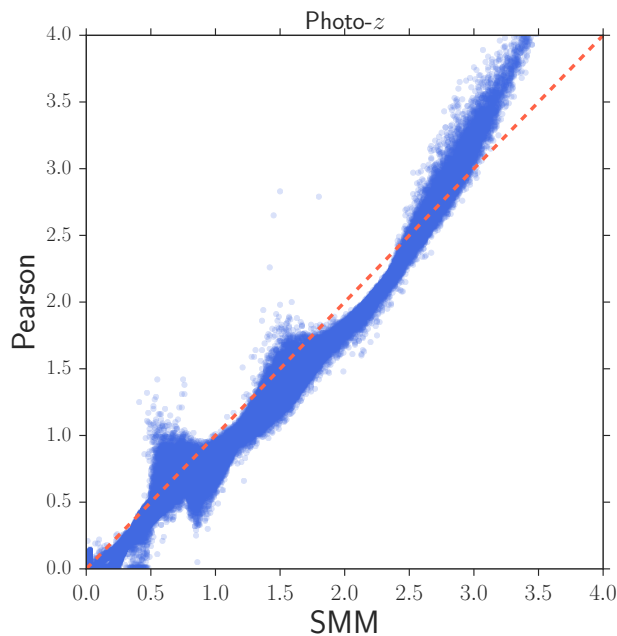


**Figure 5.10:** Scatter plot comparing different photo-$z$ estimates using the fiducial SED of SMM and the one of Pearson et al. (2013) for the H-ATLAS galaxies used in this analysis and selected with the criteria discussed in Sec. (5.2.2).
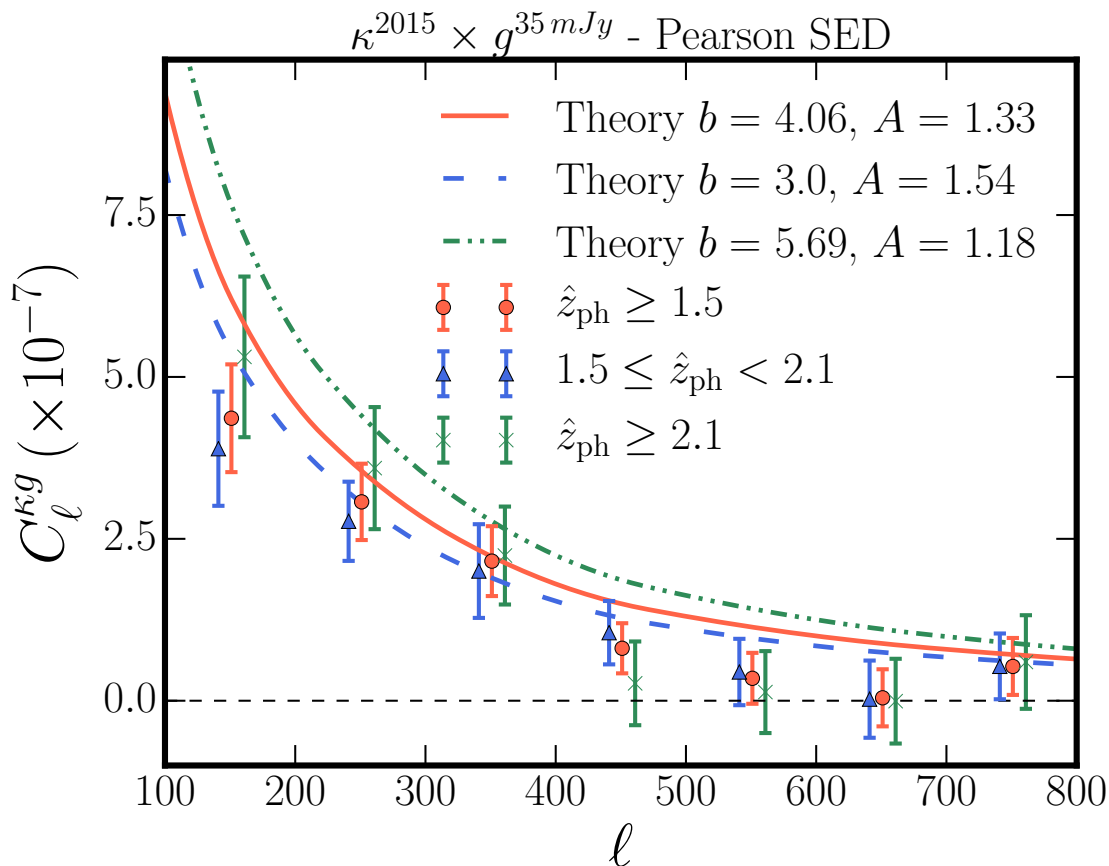
**Figure 5.11:** Cross-power spectra between the 2015 *Planck* CMB lensing map and the H-ATLAS galaxy sample built with the SED of Pearson et al. (2013) for different redshift intervals: $z_{\rm ph} \geq 1.5$ (red circles), $1.5 \leq z_{\rm ph} < 2.1$ (blue triangles), and $z_{\rm ph} \geq 2.1$ (green crosses). Uncertainties are derived as for bandpowers in Fig. (5.4). The red solid, blue dashed and green dot-dashed lines are the corresponding cross-power spectra for the best-fit bias and amplitude parameters obtained combining the data on the auto- and cross-power spectra (see Table (5.3)). The adopted redshift distributions are shown by the orange lines in Fig. (5.1).

obtained with the $dN/dz$ based on the Pearson et al. (2013) best fitting template (filled contours) with that based on the baseline SMM J2135-0102 SED in the three photo-$z$ intervals. The best fit parameter values for the Pearson et al. (2013) SED are reported in Table (5.4).

The Pearson et al. (2013) SED leads to a cross-correlation amplitude consistent with SMM J2135-0102–based results for the $1.5 \leq z_{\rm ph} < 2.1$ interval and for the full $z_{\rm ph} \geq 1.5$ sample, although the deviation from $A = 1$ has a slightly lower significance level: we have $A > 1$ at $\simeq 2.5\,\sigma$ (it was $\simeq 3.5\,\sigma$ in the SMM J2135-0102 case). For the high-$z$ bin we get consistency with $A = 1$ at the $\simeq 1\,\sigma$ level. Also, there no longer a lack of power in the first multipole bin of the galaxy auto-power spectrum for the lower-$z$ interval. The shifts in the $A$ parameter values are associated to
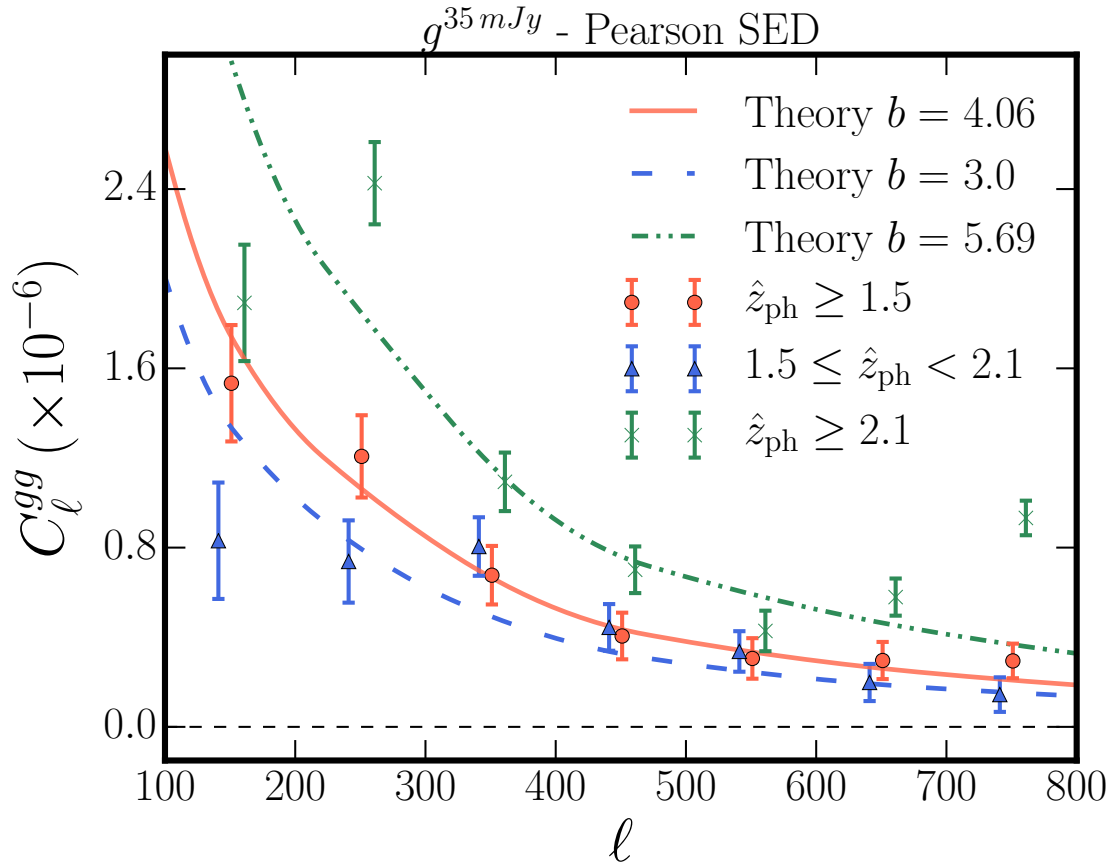
**Figure 5.12:** H-ATLAS galaxy auto-power spectra in different redshift intervals: $z_{\text{ph}} \geq$ 1.5 (red circles), $1.5 \leq z_{\text{ph}} < 2.1$ (blue triangles), and $z_{\text{ph}} \geq 2.1$ (green crosses). The SED template of Pearson et al. (2013) was adopted to estimate photo-$z$. Uncertainties are derived as for bandpowers in Fig. (5.5). The red solid, blue dashed and green dot-dashed lines are the galaxy auto-power spectra for the *Planck* cosmology and the best-fit bias and amplitude found for the $z \geq 1.5$, $1.5 \leq z < 2.1$, and $z \geq 2.1$ photo-$z$ bins respectively. The theory lines refer to the $dN/dz$ built in this chapter and also used in Fig. (5.11).

changes in the bias value: as we move toward higher redshifts, the bias parameter grows increasingly larger compared to that found using the SMM J2135-0102 SED. Adopting an effective redshift $z_{\text{eff}} = 2.15$ for the high-$z$ sample we find that the best fit value $b = 5.69$ corresponds to a characteristic halo mass $\log(M_{\text{H}}/M_\odot) = 13.5$, substantially larger than found by other studies (Xia et al., 2012; Hickox et al., 2012; Cai et al., 2013; Viero et al., 2013; Hildebrandt et al., 2013) to the point of being probably unrealistic.

### 5.4.5 Redshift dependence of the galaxy bias

Using a single, mass independent, bias factor throughout each redshift interval is certainly an approximation, although the derived estimates can be interpreted as
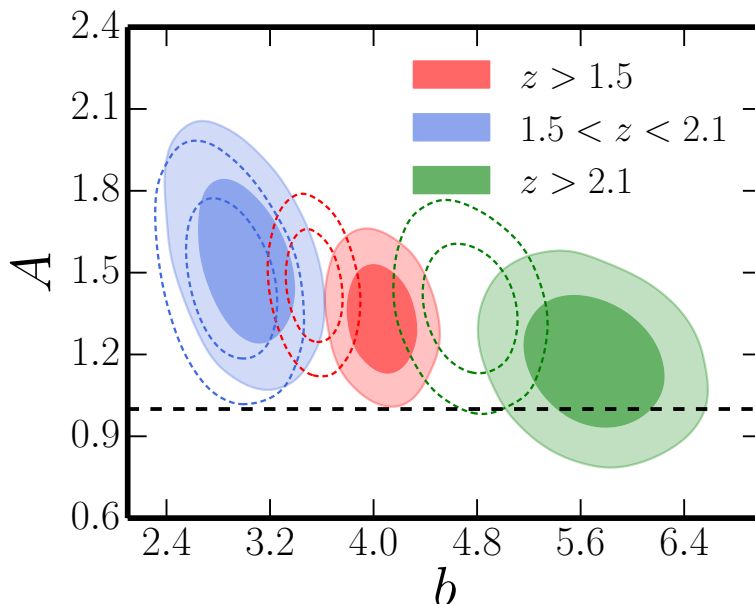
**Figure 5.13:** Posterior distributions in the $(b, A)$ plane with the 68% and 95% confidence regions (darker and lighter colors respectively) plane based on the $dN/dz$ obtained using the Pearson et al. (2013) best fitting template (filled contours) and using the baseline SMM J2135-0102 SED (dashed contours) for the three redshift intervals: $z_{\rm ph} \geq 1.5$ (red contours), $1.5 \leq z_{\rm ph} < 2.1$ (blue contours), and $z_{\rm ph} \geq 2.1$ (green contours).

effective values. In fact it is known (e.g., Sheth et al., 2001; Mo et al., 2010) that the bias function increases rapidly with the halo mass, $M_H$, and with $z$ at fixed $M_H$.

To test the stability of the derived cross-correlation amplitude $A$ against a more refined treatment of the bias parameter we have worked out an estimate of the expected effective bias function, $b_0(z)$, for our galaxy population. We started from the linear halo bias model $b(M_H; z)$ computed via the excursion set approach (Lapi & Danese, 2014). The halo mass distribution was inferred from the observationally determined, redshift dependent, luminosity function, $N(\log L_{SFR}; z)$, where $L_{SFR}$ is the total luminosity produced by newly formed stars, i.e. proportional to the Star Formation Rate (SFR). To this end we exploited the relationship between $L_{SFR}$ and $M_H$ derived by Aversa et al. (2015) by means of the abundance matching technique. Finally, we computed the luminosity-weighted bias factor as a function of redshift

$$b_0(z) = \frac{\int d\log L_{\rm SFR} N(\log L_{\rm SFR}; z) b(L_{\rm SFR}; z)}{\int d\log L_{\rm SFR} N(\log L_{\rm SFR}; z)}, \tag{5.8}$$

where the integration runs above $L_{\mathrm{m}in}(z)$, the luminosity associated to our flux density limit $S_{250} = 35\,\mathrm{mJy}$ at 250 $\mu$m.

To quantify deviations, requested by the data, from the expected effective bias

| Bin | $b$ | $A$ |
|:---:|:---:|:---:|
| $z_{\mathrm{ph}} \geq 1.5$ | $4.06^{+0.18}_{-0.18}$ | $1.33^{+0.13}_{-0.13}$ |
| $1.5 \leq z_{\mathrm{ph}} < 2.1$ | $3.00^{+0.24}_{-0.25}$ | $1.54^{+0.20}_{-0.19}$ |
| $z_{\mathrm{ph}} \geq 2.1$ | $5.69^{+0.36}_{-0.36}$ | $1.18^{+0.16}_{-0.16}$ |

**Table 5.4:** Same as Table (4.3) but for analysis based on SED template of Pearson et al. (2013).

function, $b_0(z)$, we have introduced a scaling parameter $\mathcal{A}_{\mathrm{bias}}$ so that the effective bias function used in the definition of the galaxy kernel $W^g(z)$ [Eq. (4.6)] is $b(z) = \mathcal{A}_{\mathrm{bias}}b_0(z)$.

The 68% and 95% confidence regions in the $(\mathcal{A}_{\mathrm{bias}}, A)$ plane are shown in Fig. (5.14) and the central values of the posterior distributions are reported in Table (5.5), while the corresponding bias evolution is shown in Fig. (5.15). On one side we note that $\mathcal{A}_{\mathrm{bias}}$ is found to be not far from unity, indicating that our approach to estimate the effective bias function is reasonably realistic. The largest deviations of $\mathcal{A}_{\mathrm{bias}}$ from unity happen for the lower redshift interval that may be more liable to errors in photometric redshift estimates. However, the values of the cross-correlation amplitude $A$ are in agreement with the previous results of Table (5.3), showing that our constant bias approach does not significantly undermines the derived value of $A$.

## 5.4.6 Results dependence on flux limit

To check the stability of the results against changes in the selection criterion (ii) formulated in Sec. (4.3.2) we built a new catalogue with objects complying with criteria (i), (iii) and with a (ii.b) $\geq 5\sigma$ detection at $350\,\mu\mathrm{m}$, and applied the pipeline outlined in Sec. (5.3) in the three photo-$z$ intervals. Raising the detection threshold at $350\,\mu\mathrm{m}$ has the effect of decreasing the statistical errors on photometric redshifts because of the higher signal-to-noise photometry and of favoring the selection of redder, higher-$z$ galaxies; the total number of sources decreases by approximately 20%. The credibility regions in the $(b, A)$ plane are presented in Fig. (5.16) while the best fit values of the parameters values are reported in Table (5.6).

The inferred cross-correlation amplitudes are consistent with the previous estimates within the statistical error in all of the three photo-$z$ bins ($A > 1$ at $\sim 2 - 3\,\sigma$). The value of bias parameter for the low-$z$ bin increases (dragging also the value for the full $z_{\mathrm{ph}} \geq 1.5$ sample), while the value for the high-$z$ interval is essentially unchanged. This is likely due to the fact that by requiring at least a $5\sigma$ detection
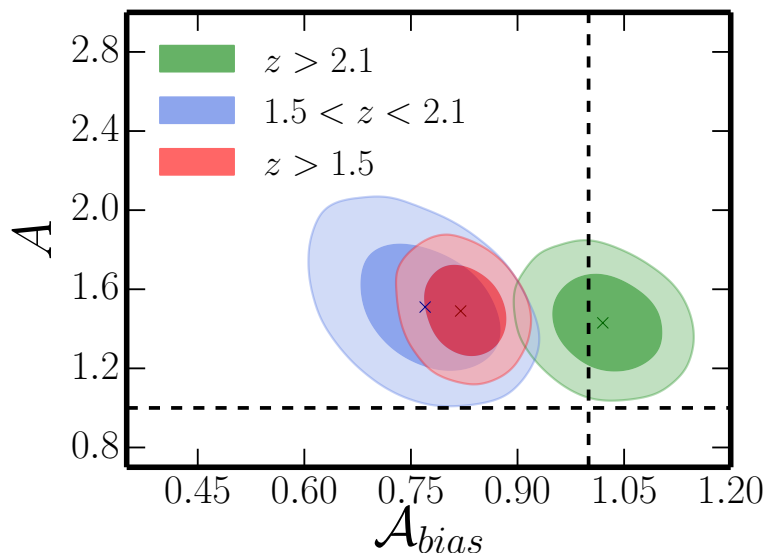
**Figure 5.14:** Posterior distributions in the $(\mathcal{A}_{\text{bias}}, A)$ plane with the 68% and 95% confidence regions (darker and lighter colors respectively) for the three redshift intervals: $z_{\text{ph}} \geq 1.5$ (red contours), $1.5 \leq z_{\text{ph}} < 2.1$ (blue contours), and $z_{\text{ph}} \geq 2.1$ (green contours). The dashed lines correspond to $A = 1$ and $\mathcal{A}_{\text{bias}} = 1$. The colored crosses mark the best-fit values reported in Table (5.5).

| Bin | $\mathcal{A}_{\text{bias}}$ | $A$ | $\chi^2/$ d.o.f. | p-value |
|---|---|---|---|---|
| $z_{\text{ph}} \geq 1.5$ | $0.82^{+0.04}_{-0.04}$ | $1.49^{+0.15}_{-0.15}$ | $9.5/12$ | $0.66$ |
| $1.5 \leq z_{\text{ph}} < 2.1$ | $0.77^{+0.06}_{-0.07}$ | $1.51^{+0.22}_{-0.20}$ | $25.7/12$ | $0.01$ |
| $z_{\text{ph}} \geq 2.1$ | $1.02^{+0.05}_{-0.05}$ | $1.43^{+0.16}_{-0.16}$ | $9.6/12$ | $0.65$ |

[a] The reduced $\chi^2$ are computed at the best-fit values.

**Table 5.5:** Best fit values of the cross-correlation, $A$, and bias, $\mathcal{A}_{\text{bias}}$, amplitudes obtained combining the observed $\kappa g$ and $gg$ spectra.

at 350 $\mu$m, we select objects which are intrinsically more luminous, hence more biased. The high-$z$ sample is not affected by the higher threshold because at such distances we already detect only the most luminous objects (Malmquist bias). At the power spectrum level we find that, for both the total $z_{\text{ph}} \geq 1.5$ sample and the low-$z$ sample, the cross-power spectra are less affected by the modification of the selection criteria, while the galaxy auto-power spectra are systematically above those obtained with the $3\sigma$ selection at $350\,\mu$m. Errors in the photo-$z$ estimates may also have a role, particularly for the low-$z$ sample; a hint in this direction is that the lack of power of $C_\ell^{gg}$ in the lowest multipole bin for the low-$z$ sample is no longer present in the case of the $5\,\sigma$ selection.
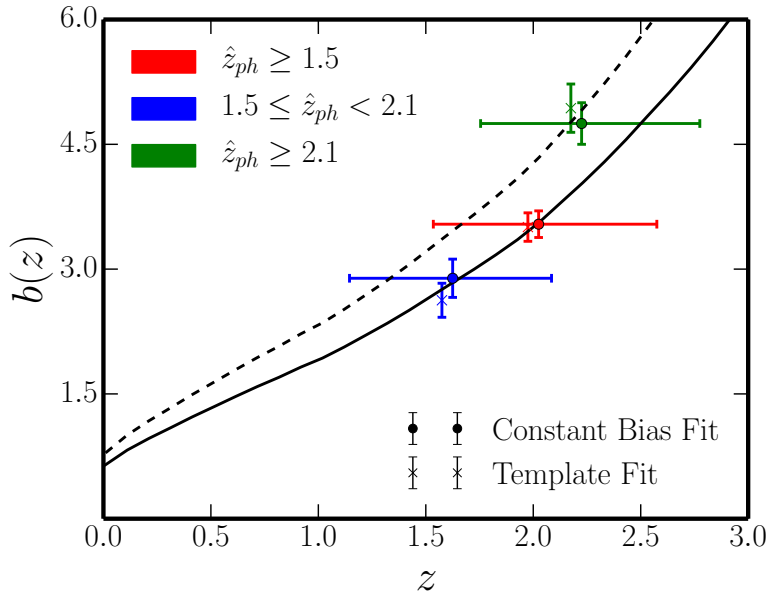
**Figure 5.15:** Effective bias functions. The dashed line corresponds to $b_0(z)$, while the solid line shows $b(z)$ with $\mathcal{A}_{\mathrm{bias}} = 0.82$, the best fit value for $z_{\mathrm{ph}} \geq 1.5$. The data points show the best fit values of the bias parameter at the median redshifts of the distributions for $z_{\mathrm{ph}} \geq 1.5$ (red), $1.5 \leq z_{\mathrm{ph}} < 2.1$ (blue) and $z_{\mathrm{ph}} \geq 2.1$ (green). In the "Template Fit" case $b(z) = \mathcal{A}_{\mathrm{bias}} b_0(z)$. Horizontal error bars indicate the $z$-range that includes 68% of each of the redshift distribution.

## 5.4.7 Other tests

The bias parameter is also influenced by non-linear processes at work on small scales. Thus it can exhibit a scale dependence. At an effective redshift of $z_{\mathrm{eff}} \simeq 2$ the multipole range $100 < \ell < 800$ corresponds to physical scales of $\approx 50 - 6$ Mpc (or $k \approx 0.03 - 0.2 \, h/\mathrm{Mpc}$). Moreover, *Planck* team does not include multipoles $\ell > 400$ in cosmological analysis based on the auto-power spectrum due do to some failed curl-mode tests. We have repeated the MCMC analysis restricting both the cross- and auto-power spectra to $\ell_{max} = 400$ and found $b = 3.58 \pm 0.18$ and $A = 1.47 \pm 0.14$ for the baseline photo-$z$ bin, fully consistent with the numbers shown in Table (5.3). For the low-$z$ bin we obtained $b = 2.76 \pm 0.28$ and $A = 1.46 \pm 0.22$, while for the high-$z$ one we found $b = 4.81 \pm 0.30$ and $A = 1.45 \pm 0.17$. Theeqrefore it looks unlikely that the higher than expected value of $A$ can be ascribed to having neglected non-linear effects, to a scale-dependent bias or to issues associated with the *Planck* lensing map.

To check the effect of our choice of the background cosmological parameters we have repeated the analysis adopting the $\mathrm{WMAP9 + SPT + ACT + BAO} + H_0$ ones (Hinshaw et al., 2013). Both $A$ and $b$ changed by $< 5\%$.

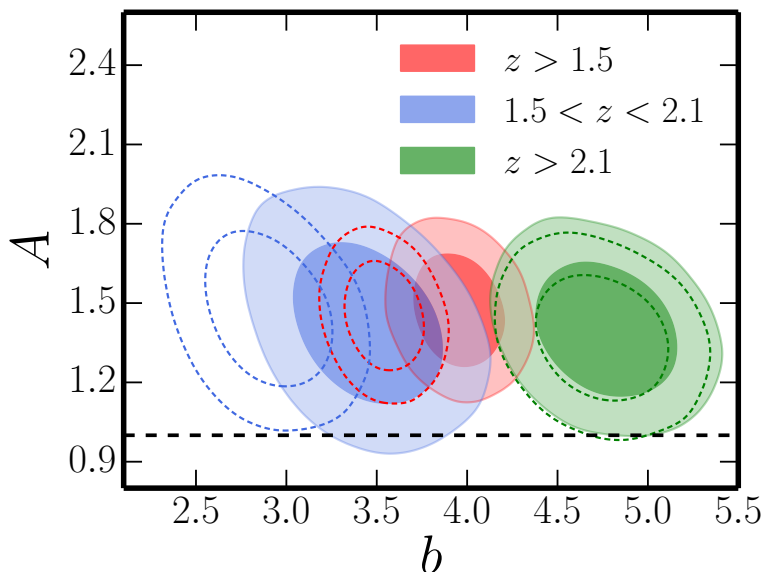The values of the bias parameter are stable and well-constrained in all redshift

**Figure 5.16:** Posterior distributions in the $(b, A)$ plane obtained requiring a $\geq 5\sigma$ detection at $350\,\mu$m (solid contours) compared with distributions obtained with our baseline selection criterion ($\geq 3\sigma$ detection; dashed contours) for the three redshift intervals: $z_{\rm ph} \geq 1.5$ (red contours), $1.5 \leq z_{\rm ph} < 2.1$ (blue contours), and $z_{\rm ph} \geq 2.1$ (green contours).

| Bin | $b$ | $A$ |
|---|---|---|
| $z_{\rm ph} \geq 1.5$ | $3.95^{+0.17}_{-0.17}$ | $1.47^{+0.14}_{-0.14}$ |
| $1.5 \leq z_{\rm ph} < 2.1$ | $3.44^{+0.27}_{-0.27}$ | $1.42^{+0.20}_{-0.20}$ |
| $z_{\rm ph} \geq 2.1$ | $4.77^{+0.26}_{-0.26}$ | $1.40^{+0.17}_{-0.17}$ |

**Table 5.6:** Best fit values of the cross-correlation amplitudes $A$ and galaxy linear bias $b$ obtained requiring a $\geq 5\sigma$ detection at $350\,\mu$m and combining the observed $\kappa g$ and $gg$ spectra.

intervals and can therefore be exploited to gain information on the effective halo masses and SFRs of galaxies. Using the relations obtained by Aversa et al. (2015) one can relate the galaxy luminosities to the SFRs and to the dark matter halo masses, $M_H$. The results are reported in Table (5.7). The SFRs are a factor of several above the average main sequence values (see Rodighiero et al., 2014; Speagle et al., 2014). The host halo masses suggest that these objects constitute the progenitors of local massive spheroidal galaxies (see Lapi et al., 2011, 2014)).

| Bin | $b$ | $\log M_{\mathrm{H}}/M_\odot$ | $\log \mathrm{SFR}\ [M_\odot\ \mathrm{yr}^{-1}]$ |
|---|---|---|---|
| $z_{\mathrm{ph}} \geq 1.5$ | $3.38^{+0.16}_{-0.16}$ | $12.9 \pm 0.1$ | $2.6 \pm 0.2$ |
| $1.5 \leq z_{\mathrm{ph}} < 2.1$ | $2.59^{+0.28}_{-0.29}$ | $12.7 \pm 0.2$ | $2.4 \pm 0.2$ |
| $z_{\mathrm{ph}} \geq 2.1$ | $4.51^{+0.24}_{-0.25}$ | $13.1 \pm 0.1$ | $2.8 \pm 0.2$ |

**Table 5.7:** Effective halo masses, $M_{\mathrm{H}}$, and SFRs derived from the effective linear bias parameters determined using jointly the reconstructed galaxy auto- and cross-spectra in the different redshift intervals. A Chabrier IMF (Chabrier, 2003) was adopted to evaluate the SFR.

## 5.5 Summary and conclusions

We have updated our previous analysis of the cross-correlation between the matter over-densities traced by the H-ATLAS galaxies and the CMB lensing maps reconstructed by the *Planck* collaboration. Using the new *Planck* lensing map we confirm the detection of the cross-correlation with a total significance now increased to $22\,\sigma$, despite of the small area covered by the H-ATLAS survey (about $\sim 1.3\%$ of the sky) and the *Planck* lensing reconstruction noise level. The improvement is due to the higher signal-to-noise ratio of *Planck* maps.

This result was shown to be stable against changes in the mask adopted for the survey and for different galaxy selections. A considerable effort was spent in modeling the redshift distribution, $dN/dz$, of the selected galaxies. This is a highly non-trivial task due to the large uncertainties in the photometric redshift estimates. We have applied a Bayesian approach to derive the redshift distribution given the photo-$z$ cuts, $z_{\mathrm{ph}}$, and the r.m.s. error on $z_{\mathrm{ph}}$.

As a first step towards the investigation of the way the dark matter distribution is traced by galaxies we have divided our galaxy sample ($z_{\mathrm{ph}} \geq 1.5$) into two redshift intervals, $1.5 \leq z_{\mathrm{ph}} < 2.1$ and $z_{\mathrm{ph}} \geq 2.1$, containing similar numbers of sources and thus similar shot-noise levels.

A joint analysis of the cross-spectrum and of the auto-spectrum of the galaxy density contrast yielded, for the full $z_{\mathrm{ph}} \geq 1.5$ sample, a bias parameter of $b = 3.54^{+0.15}_{-0.14}$. This value differs from the one found in B15 ($b = 2.80^{+0.12}_{-0.11}$) because of the different modeling of the redshift distribution, $dN/dz$: when the analysis is performed adopting the same $dN/dz$ as B15 we recover a value of $b$ very close to theirs.

On the other hand, we still find the cross-correlation amplitude to be higher than expected in the standard $\Lambda$CDM model although by a slightly smaller factor: $A = 1.45^{+0.14}_{-0.13}$ against $A = 1.62 \pm 0.16$, for the full galaxy sample ($z_{\mathrm{ph}} \geq 1.5$). A similar excess amplitude is found for both redshift intervals, although it is

slightly larger for the lower-$z$ interval, which may be more liable to the effect of the redshift–dust temperature degeneracy, hence more affected by large failures of photometric redshift estimates. We have $A = 1.48^{+0.20}_{-0.19}$ for the lower $z$ interval against $A = 1.37 \pm 0.16$ for the higher $z$ one. Larger uncertainties in $z_{\mathrm{ph}}$ may be responsible, at least in part, also for the lack of power in the lowest multipole bin of the auto-power spectrum of galaxies in the lower redshift interval. However, reassuringly, the measured cross-correlation of positions of galaxies in the two redshift intervals is in good agreement with the expectations given the overlap of the estimated redshift distributions due to errors in the estimated redshifts. It is thus unlikely that the two redshift distributions are badly off.

We have also tested the dependence of the results on the assumed SED (used to estimate the redshift distribution) by repeating the full analysis using the Pearson et al. (2013) SED. The deviation from the expected value, $A = 1$, of the cross-correlation amplitude recurs, although at a somewhat lower significance level ($\simeq 2.5\,\sigma$ instead of $\simeq 3.5\,\sigma$). However this happens at the cost of increasing the bias parameter for the higher redshift interval to values substantially higher than those given by independent estimates.

The resulting values of $A$ are found to be only marginally affected by having ignored the effect of non-linearities in the galaxy distribution and of variations of the bias parameter within each redshift interval, as well as by different choices of the background cosmological parameters.

The data indicate a significant evolution with redshift of the effective bias parameter: for our baseline redshift distributions we get $b = 2.89 \pm 0.23$ and $b = 4.75^{+0.24}_{-0.25}$ for the lower- and the higher-$z$ interval, respectively. The increase of $b$ reflects a slight increase of the effective halo mass, from $\log(M_H/M_\odot) = 12.7 \pm 0.2$ to $\log(M_H/M_\odot) = 13.1 \pm 0.1$. Interestingly, the evolution of $b$ is consistent with that of the luminosity weighted bias factor yielded, as a function of $M_H$ and $z$, by the standard linear bias model. According to the SFR–$M_H$ relationships derived by Aversa et al. (2015), the typical SFRs associated to these halo masses are $\log(\mathrm{SFR}/M_\odot\,\mathrm{yr}^{-1}) = 2.4 \pm 0.2$ and $2.8 \pm 0.2$, respectively.

If residual systematics in both lensing data and source selection is sub-dominant, then one would conjecture that the selected objects trace more lensing power than the bias would represent, in order to achieve a cross-correlation amplitude closer to 1.

An amplitude of the cross-correlation signal different from unity has been recently reported by the DES collaboration (Giannantonio et al., 2016) who measured the cross-correlation between the galaxy density in their Science Verification data and

the CMB lensing maps provided by the *Planck* satellite and by the SPT. They however found $A < 1$, but for a galaxy sample at lower (photometric) redshifts than our sample. So, their result is not necessarily conflicting with ours, especially taking into account that they found $A$ to be increasing with redshift. Another hint of tension between $\Lambda$CDM predictions and observations has been reported by Pullen et al. (2016), where the authors correlated the *Planck* CMB lensing map with the Sloan Digital Sky Survey III (SDSS-III) CMASS galaxy sample at $z = 0.57$, finding a tension with general relativity predictions at a $2.6\sigma$ level. In another paper, Omori & Holder (2015) compare the linear galaxy bias inferred from measurements of the *Planck* CMB lensing - CFHTLens galaxy density cross-power spectrum and the galaxy auto-power spectrum, reporting significant differences between the values found for 2013 and 2015 *Planck* releases. This case has been further investigated by exploiting the analysis scheme developed in B15 by Kuntz (2015), where the author partially confirms the Omori & Holder (2015) results, finding different cross-correlation amplitude values between the two *Planck* releases.

The CMB lensing tomography is at an early stage of development. Higher signal-to-noise ratios will be reached due to the augmented sensitivity of both galaxy surveys, such as DES, Euclid, LSST, DESI, and of CMB lensing experiments, such as AdvACT (Calabrese et al., 2014) or the new phase of the POLARBEAR experiment, the Simons Array (Ade et al., 2014c). In the near future, the LSS will be mapped at different wavelengths out to high redshifts, enabling the comparison with the results presented in this and other works, the comprehension of the interplay between uncertainties in datasets and astrophysical modeling of sources, as well as the constraining power on both astrophysics and cosmology of cross-correlation studies.

# 6

# Needlet estimation for CMB-LSS cross-correlation

> *Elwood*: What kind of music do you usually have here?
> *Claire*: Oh, we got both kinds. We got country \*and\* western.

> –The Blues Brothers

In this chapter we develop a novel needlet-based estimator to investigate the cross-correlation CMB lensing maps and LSS data. We compare this estimator with its harmonic counterpart and, in particular, we analyze the bias effects of different forms of masking. In order to address this bias, we also implement a `MASTER`-like technique in the needlet case. The resulting estimator turns out to have an extremely good signal-to-noise performance. Our analysis aims at expanding and optimizing the operating domains in CMB-LSS cross-correlation studies, similarly to CMB needlet data analysis. It is motivated especially by next generation experiments (such as Euclid) which will allow us to derive much tighter constraints on cosmological and astrophysical parameters through cross-correlation measurements between CMB and LSS.

## 6.1 The use of needlets for cosmological datasets

As we thoroughly discussed in Ch. (2.2.3), CMB lensing-galaxy cross-correlation measurements can be used to reconstruct the redshift evolution of the galaxy bias (Bianchini et al., 2015; Allison et al., 2015; Bianchini et al., 2016b), the growth of structures (Giannantonio et al., 2016) and to augment the absolute cosmic shear calibration (Baxter et al., 2016). In particular, it is worth to notice that *all* analyses reported to date have reconstructed the 2-point statistics either in harmonic or real space.

In this chapter, we shall use instead a procedure based on needlet ideas, and we shall discuss how to improve its performance for very aggressive masks using a `MASTER`-like correction.

As discussed in many previous references, needlets are a form of spherical wavelets which were introduced in functional analysis and statistics by Narcowich et al. (2006); Baldi et al. (2009) and have then found a number of different applications in the cosmological community over the last decade; we recall for instance Marinucci et al. (2008) for a general description of the methods, Lan & Marinucci (2008); Rudjord et al. (2009); Pietrobon et al. (2010); Donzelli et al. (2012); Regan et al. (2015); Planck Collaboration (2016b) for non-Gaussianity estimation, Delabrouille (2010); Planck Collaboration (2014a, 2015a); Rogers et al. (2016a,b) for foreground component separation, Geller et al. (2008); Leistedt et al. (2015a); Planck Collaboration (2016b) for polarization data analysis, Durastanti et al. (2014); Leistedt et al. (2015b) for extension in 3d framework and Troja et al. (2014); Regan et al. (2015) for trispectrum analysis.

The advantages of needlets, like those of other wavelets system, have been widely discussed in the literature; in short, they are mainly concerned with the possibility to exploit double localization properties, in the real and harmonic domain. Despite this localization in the real domain, we show here that the performance of a needlet cross-correlation estimator deteriorates badly in the presence of very aggressive sky-cuts (i.e., experiments with sky coverage much smaller than 50%). In this chapter, we show how the performance of this estimator can be greatly improved by a `MASTER`-like correction. Thus achieving signal-to-noise figure of merits which are in some aspect superior to the corresponding results for power spectrum methods.

The plan of the chapter is as follows. In section (6.2) we introduce the needlet system along with the spectral estimation in this framework; we then proceed in section (6.3) to introduce the `MASTER`-like algorithm for the needlets cross-correlation estimator. Numerical evidence and some comparison on the performance of these procedures are collected in section (6.4), while final considerations are presented in section (6.5).

## 6.2 Building the cross-correlation estimators

In this section we introduce the spectral estimator in needlet space. We recall that the theoretical background of the CMB lensing-galaxy cross-correlation has been widely discussed in Ch. (2.2.3) and (4.2), while we illustrated how to build an harmonic estimator for cross-correlation analysis in Ch. (3.2.1). Even though the main focus of the chapter is the measurement of the CMB lensing-galaxy cross-correlation, we stress that the estimators presented here can be applied to any *scalar* field on the sphere.

### 6.2.1 Needlet cross-correlation estimator

As mentioned in the introduction, some drawbacks of standard Fourier analysis on the sphere can be mitigated by the exploitation of needlet/wavelet techniques. Related advantages have already been widely discussed in the literature, see again Marinucci et al. (2008); Lan & Marinucci (2008); Donzelli et al. (2012); Troja et al. (2014); Marinucci et al. (2011); Durastanti et al. (2014). Below we provide a brief summary of some of the needlets advantages with respect to other spherical wavelets construction:

1. they have bounded support in the harmonic domain (i.e. localized on a finite number of multipoles);

2. they are quasi-exponentially (i.e. faster than any polynomial) localized in pixel space;

3. they do not rely on any tangent plane approximation;

4. they allow for a simple reconstruction formula (because they represent a tight frame system). Note that this property is *not* shared by other wavelet systems such as the common Spherical Mexican Hat Wavelet (SMHW);

5. they are computationally convenient to implement and natively adapted to standard packages such as `HEALPix`;

6. needlets coefficients can be shown to be asymptotically uncorrelated at any fixed angular distance for growing frequencies.

Here we recall that the spherical needlet system $\psi_{\{jk\}}$ can be obtained by a quadratic combination of spherical harmonics as

$$\psi_{jk}(\hat{\mathbf{n}}) = \sqrt{\lambda_{jk}} \sum_{\ell=[B^{j-1}]}^{[B^{j+1}]} b\left(\frac{\ell}{B^j}\right) \sum_{m=-\ell}^{\ell} Y_{\ell m}^*(\hat{\mathbf{n}}) Y_{\ell m}(\xi_{jk}), \qquad (6.1)$$

where $[\cdot]$ denotes the integer part, $b(\cdot)$ is the filter function in the harmonic domain defined for $x \in [1/B, B]$, and $\{\xi_{jk}\}$ are the cubature points on the sphere corresponding to the frequency $j$ and the location $k$. Since our implementation relies on the `HEALPix` pixelation scheme we can identify the cubature points with the pixel centers, so that the cubature weights $\lambda_{jk}$ can be approximated by $4\pi/N_{\mathrm{p}ix}$, where $N_{\mathrm{p}ix}$ is the number of pixels for the chosen `HEALPix` $N_{side}$ resolution and $k$ represents the pixel number (Pietrobon et al., 2006).

Needlets can be thought of as a convolution of the projection operator $\sum_m Y_{\ell m}^*(\hat{\mathbf{n}}) Y_{\ell m}(\xi_{jk})$ with a filter function $b(\cdot)$ whose width is controlled by the only free parameter $B$: recipes for the construction of the function $b(\cdot)$ can be found in Marinucci et al.

(2008); McEwen et al. (2013); Marinucci et al. (2011). For instance, an explicit recipe for the construction of the filter function is as follows (Marinucci et al., 2008):

1. Construct the $C^\infty$-function

$$\phi_1(t) = \begin{cases} \exp\left(-\frac{1}{1-t^2}\right) & t \in [-1,1] \\ 0 & \text{otherwise,} \end{cases} \tag{6.2}$$

   compactly supported in $[-1,1]$;

2. Implement the non-decreasing $C^\infty$-function

$$\phi_2(u) = \frac{\int_{-1}^{u} dt\, \phi_1(t)}{\int_{-1}^{1} dt\, \phi_1(t)}, \tag{6.3}$$

   which is normalized in order to satisfy $\phi_2(-1) = 0$ and $\phi_2(1) = 1$;

3. Construct the function

$$\phi_3(t) = \begin{cases} 1 & t \in [0, 1/B] \\ \phi_2\left(1 - \frac{2B}{B-1}\left(t - \frac{1}{B}\right)\right) & t \in (1/B, 1] \\ 0 & t \in (1, \infty); \end{cases} \tag{6.4}$$

4. Define for $u \in \mathbb{R}$

$$b^2(u) = \phi_3\left(\frac{u}{B}\right) - \phi_3(u). \tag{6.5}$$

A smaller value of $B$ corresponds to a narrower localization in $\ell$-space, while a larger value translates into a more precise localization in real space. Once $B$ is fixed, each needlet can be shown to pick up signal only from a specific range of multipoles determined by the index $j$: the profile of the filter function $b(\cdot)$ is shown in Fig. (6.1) for different frequencies. Needlet coefficients are then evaluated by projecting the centered field $X(\hat{\mathbf{n}})$ on the corresponding needlet $\psi_{jk}(\hat{\mathbf{n}})$ as

$$\begin{aligned}
\beta_{jk} &= \int_{\mathbb{S}^2} X(\hat{\mathbf{n}}) \psi_{jk}(\hat{\mathbf{n}})\, d\Omega \\
&= \sqrt{\lambda_{jk}} \int_{S^2} \sum_{\ell m} x_{\ell m} Y_{\ell m}(\hat{\mathbf{n}}) \sum_{LM} b\left(\frac{L}{B^j}\right) Y_{LM}^*(\hat{\mathbf{n}}) Y_{LM}(\xi_{jk})\, d\Omega \\
&= \sqrt{\lambda_{jk}} \sum_{\ell m LM} x_{\ell m} b\left(\frac{L}{B^j}\right) Y_{LM}(\xi_{jk}) \underbrace{\int_{\mathbb{S}^2} Y_{\ell m}(\hat{\mathbf{n}}) Y_{LM}^*(\hat{\mathbf{n}})\, d\Omega}_{\delta_{\ell L}\delta_{mM}} \\
&= \sqrt{\lambda_{jk}} \sum_{\ell} b\left(\frac{\ell}{B^j}\right) \sum_{m=-\ell}^{\ell} x_{\ell m} Y_{\ell m}(\xi_{jk}).
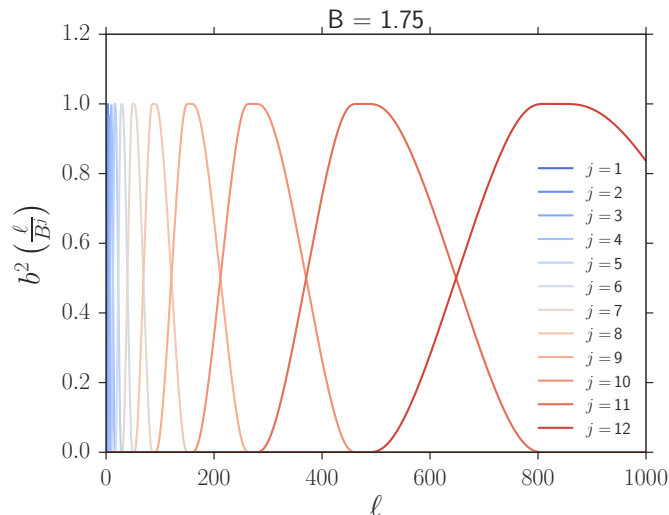\end{aligned} \tag{6.6}$$

**Figure 6.1:** Profile of the filter function in the $\ell$-space for different needlet frequencies $j$. The needlet width parameter is set to $B = 1.75$.

Needlet coefficients corresponding to a given frequency $j$ can themselves be represented as an `HEALPix` map. It is worth to stress that although needlets do not make up an orthonormal basis for square integrable functions on the sphere, they represent a *tight frame* (redundant basis) so that they allow for a simple reconstruction formula, i.e

$$X(\hat{\mathbf{n}}) = \sum_{jk} \beta_{jk} \psi_{jk}(\hat{\mathbf{n}}). \tag{6.7}$$

After computing the needlet coefficients $\beta_{jk}$ from the maps, we can build a spectral estimator as

$$\hat{\beta}_j^{XY} = \frac{1}{N_{\text{pix}}} \sum_k \beta_{jk}^X \beta_{jk}^Y, \tag{6.8}$$

and it is immediate to check that it provides an unbiased estimate of (a binned form of) the angular power spectrum, i.e.

$$\langle \hat{\beta}_j^{XY} \rangle \equiv \beta_j^{XY} = \sum_\ell \frac{2\ell + 1}{4\pi} b^2\left(\frac{\ell}{B^j}\right) C_\ell^{XY}. \tag{6.9}$$

These theoretical predictions can directly be compared to the extracted spectra, allowing for the parameter extraction process. Moreover, as noted in Pietrobon et al. (2006), the analytic relation between $\beta_j$ and $C_\ell$ makes straightforward dealing with beam profiles, pixel window function, and experimental transfer functions. Note that in this chapter we divide the spectral estimator (6.8) and its expected

value (6.9) for a normalizing factor $\mathcal{N}$ given by

$$\mathcal{N} = \sum_\ell \frac{2\ell + 1}{4\pi} b^2 \left( \frac{\ell}{B^j} \right), \tag{6.10}$$

so that in the plots we show $\hat{\beta}_j^{XY} \to \hat{\beta}_j^{XY}/\mathcal{N}$.

The theoretical variance of the cross-correlation power spectrum in needlet space reads

$$(\Delta\beta_j^{XY})^2 \equiv \mathrm{Var}[\hat{\beta}_j^{XY}] = \sum_\ell \frac{2\ell + 1}{16\pi^2} b^4 \left( \frac{\ell}{B^j} \right) \left[ (C_\ell^{XY})^2 + C_\ell^{XX} C_\ell^{YY} \right], \tag{6.11}$$

where the angular auto-spectra can be comprehensive of a noise term, i.e. $C_\ell \to C_\ell + N_\ell$, if present. Moreover, the needlets system is compactly supported in the harmonic domain and as such, for full-sky maps, the random needlets coefficients are uncorrelated by construction for $|j - j'| \geq 2$ (Baldi et al., 2009).

## 6.3 MASTER algorithm for needlets

As mentioned in Sec (6.2.1), one of the main driver behind the development of the needlet spectral estimator is the need to overcome the issues related to Fourier analysis on the sphere in the presence of missing observation. The excellent needlets localization properties in real space represent a key feature for analyzing cosmological data on the partially observed sky, in particular it has been shown that even in the presence of masked regions the random needlet coefficients $\beta_{jk}$ are asymptotically independent (over $k$) as $j \to \infty$ (contrary to the case of random coefficients $x_{\ell m}$) (Marinucci et al., 2008; Baldi et al., 2009). However, as we shall see from simulations in the next section, the estimator defined in Eq. (6.8) becomes biased for aggressive masking: here we formally study the effect of sky-cuts on the needlet power spectrum estimation.

From Eq. (3.37), we find that needlet coefficients computed on a masked sky are given by

$$\tilde{\beta}_{jk} = \sqrt{\lambda_{jk}} \sum_\ell b \left( \frac{\ell}{B^j} \right) \sum_m \tilde{x}_{\ell m} Y_{\ell m}(\xi_{jk}). \tag{6.12}$$

Then, if we consider the statistic

$$\hat{\Gamma}_j^{XY} = \frac{1}{N_{\mathrm{p}ix}} \sum_k \tilde{\beta}_{jk}^X \tilde{\beta}_{jk}^Y, \tag{6.13}$$

it is straightforward to see that its expectation value reads as follows

$$
\begin{aligned}
\langle \hat{\Gamma}_j^{XY} \rangle \equiv \Gamma_j^{XY} &= \sum_{\ell m} b^2 \left( \frac{\ell}{B^j} \right) \langle \tilde{x}_{\ell m} \tilde{y}_{\ell m} \rangle \\
&= \sum_{\ell m} \sum_{\ell' m'} b^2 \left( \frac{\ell}{B^j} \right) K_{\ell m \ell' m'}^2 [W] C_{\ell'} \\
&= \sum_{\ell \ell'} \frac{2\ell + 1}{4\pi} b^2 \left( \frac{\ell}{B^j} \right) M_{\ell \ell'} C_{\ell'},
\end{aligned}
\tag{6.14}
$$

which tells us that $\hat{\Gamma}_j^{XY}$ is an unbiased estimator for a smoothed version of the pseudo angular power spectrum $\tilde{C}_\ell$, similar to the case of $\hat{\beta}_j^{XY}$: in some sense, we can view $\hat{\Gamma}_j^{XY}$ as an estimator of the pseudo-needlet power power spectrum. Using Eq. (3.45), which is valid for slowly varying power spectra and/or large sky fractions, it is possible to relate the two estimators as

$$
\begin{aligned}
\langle \hat{\Gamma}_j^{XY} \rangle &\approx f_{\mathrm{sky}} \sum_\ell \frac{2\ell + 1}{4\pi} b^2 \left( \frac{\ell}{B^j} \right) C_\ell \\
&= f_{\mathrm{sky}} \langle \hat{\beta}_j^{XY} \rangle.
\end{aligned}
\tag{6.15}
$$

Before we conclude this section let us introduce a couple of remarks. We recall first that in the case of a survey with a large sky-cut, inverting the full coupling matrix become unfeasible because of singularities; hence the power spectrum can be estimated only over some subset of multipoles i.e. the power spectrum is recovered only up to some frequencs windows. As discussed earlier in Sec. (3.2.1), the choice of this frequency windows is to a good degree arbitrary; on the other hand, the needlet framework naturally provides a binning scheme which is controlled by a single width parameter $B$ (as well as by the profile of the filter function $b(\cdot)$).

As a second difference, we note that while the PCL approach usually makes use of the *backward modelling*, where measurements are deconvolved for numerical and observational effects to match the theoretical predictions, needlets analysis is oriented towards the *forward modelling*, which turns theoretical (needlet) power spectra into pseudo-spectra that can be directly compared to the raw measurements[1] (see Harnois-Déraps et al. (2016) for a discussion on forward and backward modeling). In particular, in the needlet case it is not feasible to write a closed formula such as Eq. (3.42) to express the original needlet power spectrum as a function of the pseudo one, i.e. $\beta_j = \beta_j(\Gamma_j)$; however, this is not an obstacle for data analysis because the forward estimator can be used just as well to do model checks as

---

[1]Note that pseudo-spectra, either in harmonic or needlet space, depend on the observational setup represented for example by the masking, the smoothing, and the apodization, while this is not the case for theoretical predictions.
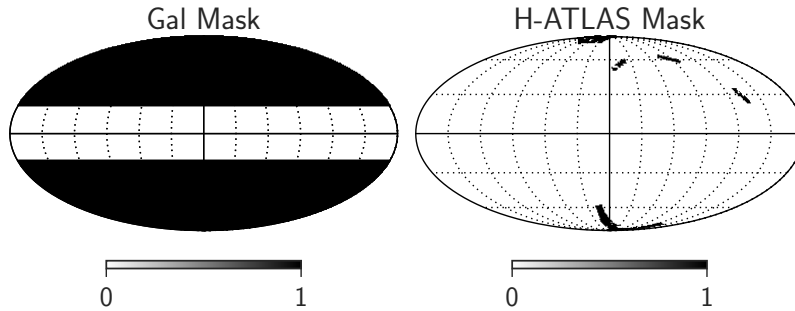
**Figure 6.2:** Masks used for the analysis. The mask with a symmetric Galactic cut at $\pm 20\deg$ ($f_{\text{sky}} = 0.65$) is shown in the left part, while the H-ATLAS mask ($f_{\text{sky}} = 0.013$) is shown on the right one. In both cases the black color denotes observed regions of the sky.

parameter estimation.

## 6.4 Numerical evidence

In this section we describe the simulations setup exploited and the tests performed in order to compare the harmonic and needlet cross-correlation estimators.

### 6.4.1 Simulations

We simulate a set of $N_{sim} = 500$ correlated CMB convergence and galaxy density maps at an `HEALPix` resolution of $N_{side} = 512$ (corresponding to an angular resolution of $\sim 7'.2$). For the galaxies we consider an high-$z$ Herschel-like population with a redshift distribution as described in Ch. (5) and fix $b = 3$ for the present galaxy sample; the precise details of spectra are not fundamental since we are interested in testing the estimators. This simulations set is used in a MC approach (i) to validate the extraction pipelines; (ii) to compute the uncertainty associated with each bin; and (iii) to quantify the degree of correlation among different needlet frequencies. A thorough description of the main steps to obtain correlated CMB lensing and galaxy maps comprising of signal and noise can be found in Ch. (4), here we simply use noise-free maps for validation purposes. Pairs of correlated signal-only Gaussian CMB convergence $\kappa_{\ell m}^{S}$ and galaxy density $g_{\ell m}^{S}$ maps are generated from the three fiducial spectra $C_{\ell}^{\kappa g}$, $C_{\ell}^{\kappa\kappa}$ and $C_{\ell}^{gg}$ (Giannantonio et al., 2008; Bianchini et al., 2015). This is easily implemented using the synfast routine of `HEALPix`. In order to show the effect of masking on the reconstructed statistics, the simulated maps are masked with two different masks: we consider either a Galactic mask that covers the 35% of sky ($f_{\text{sky}} = 0.65$), similar to the one implemented in

*Planck* CMB data, and a much more aggressive H-ATLAS (Eales et al., 2010) mask with sky coverage equal only to 1.3% that comprehends the North Galactic Pole, the South Galactic Pole, and the GAMA fields. The adopted masks are shown in Fig. (6.2).

### 6.4.2 Results

We start by investigating the uncorrelation properties of the needlets coefficients, as a function of the width of the mask. In particular, the covariance matrix of needlet coefficients is computed by means of 500 MC simulations as

$$\mathrm{Cov}_{jj'} \equiv \mathrm{Cov}[\hat{\beta}_j, \hat{\beta}_{j'}] = \langle (\hat{\beta}_j - \langle \hat{\beta}_j \rangle_{\mathrm{MC}})(\hat{\beta}_{j'} - \langle \hat{\beta}_{j'} \rangle_{\mathrm{MC}}) \rangle_{\mathrm{MC}}. \qquad (6.16)$$

The corresponding results are reported in Fig. (6.3) for the full-sky, Galactic and H-ATLAS cases respectively. Numerical evidence is very much consistent with the theoretical expectation: in particular in the full-sky and Galactic mask case the correlation decrease very rapidly outside the main diagonal (where it is trivially unit, which is not reported in the table) in the case of full-sky maps. The decay is still very satisfactory when sky coverage is high, although not complete as for the Galactic mask; on the other hand a very aggressive cut with sky coverage of 1.3% deteriorate enormously the uncorrelation properties (even though $\mathrm{Corr}_{jj'}$ is $\mathcal{O}(0.1)$ and smaller at high frequencies for $|j - j'| \geq 2$), see the bottom panel in Fig. (6.3). The estimated covariances are then used to derive error bars in the cross-correlation estimators reported in Fig. (6.4). Again, the needlet estimator is shown to perform very well in the full-sky and Galactic mask cases whereas Herschel-like framework clearly requires corrections. Error bars decay rapidly for increasing frequencies as expected. For comparison, in Fig. (6.5) `MASTER`-like estimators are reported for the cross-power spectrum, while the equivalent `MASTER` needlet reconstruction discussed in section (6.3) is shown in Fig. (6.6), where we can see that the bias is strongly suppressed.

The most important results are collected in Fig. (6.7), where we report the performance of the `MASTER`-like corrected needlet cross correlation estimator. The figure of merit that we report is the signal-to-noise ratio (S/N), evaluated as the ratio between the analytical expected value of the estimator (numerator) and a measure of variability, which can be either the standard deviation ($\Delta \hat{\beta}_j \equiv \sqrt{\mathrm{Cov}_{jj}}$) or the root mean square error ($\sqrt{\mathrm{MSE}}$). The latter estimator takes in to account also the possible presence of bias, but this is so small that the two measures are largely equivalent. Clearly, an higher value of this figure of merit entails a better

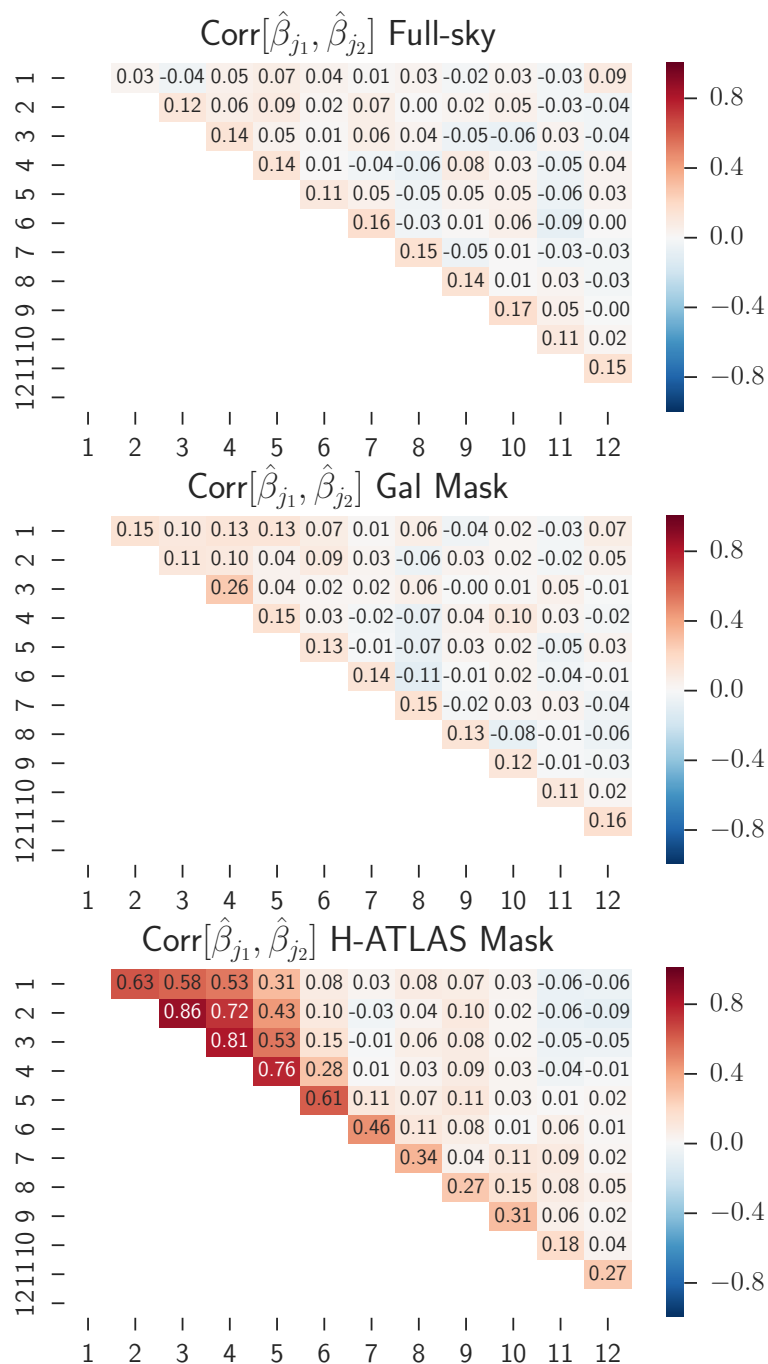$$B = 1.75 \ N_{side} = 512 \ N_{sim} = 500 \text{ - Signal Only}$$



**Figure 6.3:** Cross-correlation coefficient matrices, defined as $\text{Corr}_{ij} \equiv \text{Cov}_{ij}/\sqrt{\text{Cov}_{ii}\text{Cov}_{jj}}$, about the needlet space estimator. From top to bottom we show results for the full-sky, Galactic mask, and H-ATLAS mask cases respectively.
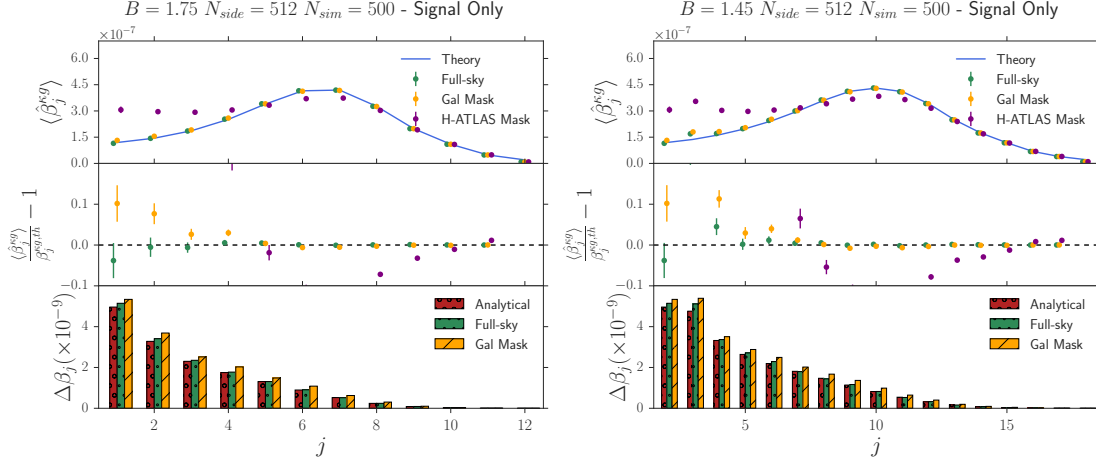
**Figure 6.4:** *Upper panel*: Recovered mean needlet cross-power spectrum between correlated CMB convergence and galaxy density maps for different masks and width parameter ($B = 1.75$ and $1.45$ on the left and right parts respectively). Green, yellow and purple bandpowers represent full-sky, Galactic mask (with $f_{\rm sky} = 0.65$) and H-ATLAS mask (with $f_{sky} = 0.013$) cases respectively. Solid blue line is the generative theoretical input cross-power spectrum. Error bars shown are the diagonal components of the covariance matrices (defined in Eq. 4.1), properly scaled by $\sqrt{N_{\rm sim}}$. *Central panel*: Fractional difference between mean recovered and theoretical needlet cross-spectra for the cases shown in the upper panel. *Lower panel*: Error bars comparison for the cases shown in the upper panel. Note that the lack of power observed for $j = 12$ (or for $j = 18$ if $B = 1.45$) is due to the fact that simulated maps have been generated using spectral information up to $\ell_{max} = 2N_{side} = 1024$, while the needlet frequency $j = 12$ picks up signal in the multipole range of $458 \lesssim \ell \lesssim 1396$ ($551 \lesssim \ell \lesssim 1159$), where the power is partially missing.
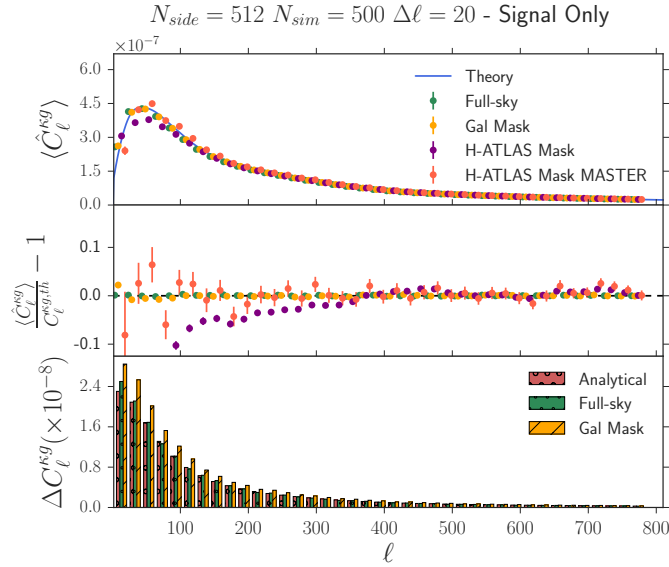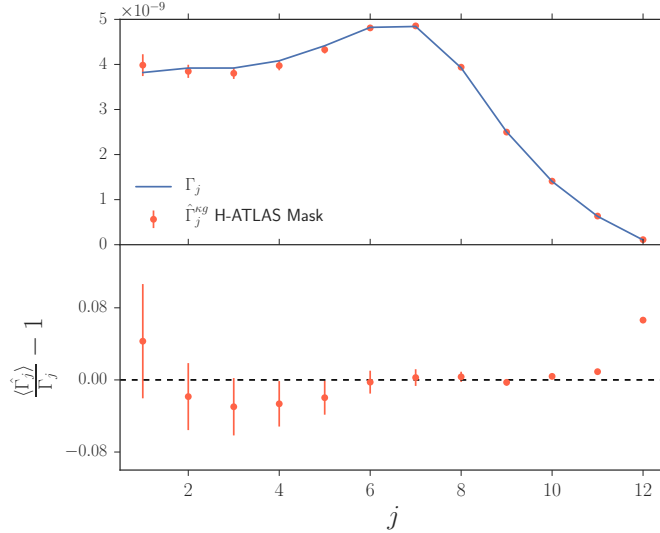


**Figure 6.5:** Same as (6.4) but in harmonic space.

**Figure 6.6:** Mean needlet pseudo power spectrum $\langle \hat{\Gamma}_j^{\kappa g} \rangle$ (orange circles) superimposed to the generative theoretical (pseudo) spectrum (blue line).

performance of the estimator; for comparison, the same statistics are reported also for the MASTER-like cross-correlation estimator in the standard harmonic domain.

The performance of the needlet estimator seems to be extremely satisfactory, with the figure of merit ranging from 1 to 3 for the H-ATLAS case and from 3 to 10 for the Galactic case at the smallest frequencies $j = 3, 4, 5$ (corresponding to multipoles of the order $\ell = 6, 10, 18$ respectively). At higher frequencies, i.e. $j = 10, 11, 12$ (corresponding to central multipoles of the order $\ell = 312, 547, 957$ respectively), the figure of merit is of order 200 (30) when the Galactic (H-ATLAS) mask is applied. To make a rough comparison, the figures of merit for the standard power spectrum cross-correlation estimators are in the order of 9 at $\ell = 10$, and 80 at $\ell = 800$ for the Galactic mask case, while the figure of merit in the H-ATLAS scenario goes from below 1 up to roughly 6 in the same $\ell$-range. To be fair, we stress that these numbers are not strictly comparable, because the bandwidths which are chosen for the standard harmonic domain estimator are constant across the multipoles domain and smaller than the equivalent needlet bandwidths, especially at high frequencies $j$. However, the least one can conclude from these results is that the two procedures have different advantages; in particular, we view as major assets for the needlet based algorithm the very high S/N and the natural choice of bandwidth parameters; while the advantage of the power spectrum based procedure seems the high resolution which can be achieved in multipole spaces.

As a further check we show in Fig. (6.8) the variance of the harmonic and needlet space estimators for the different observational setups as function of multipole $\ell$ and needlet frequency $j$, normalized to the full-sky analytical variance.
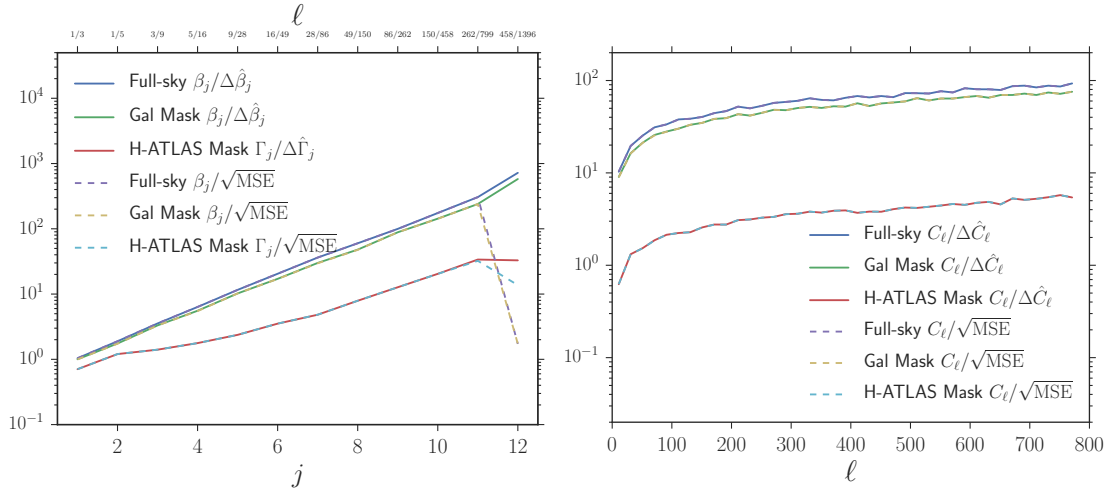
**Figure 6.7:** Different figures of merit discussed in text to assess the goodness of the estimators for the needlet (left panel) and harmonic (right panel) cases respectively. On top of the left plot we quote the multipoles that roughly correspond to a given needlet frequency $j$.
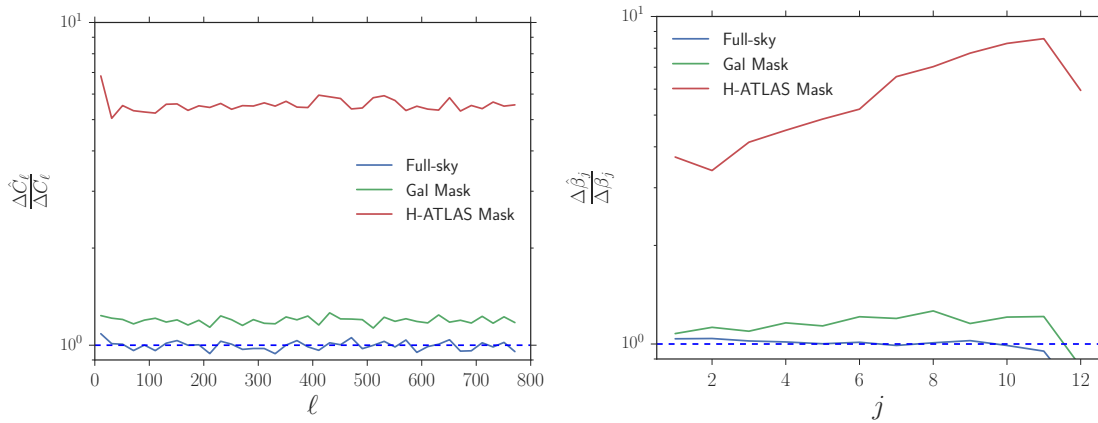


**Figure 6.8:** The variance of cross-power spectrum estimates (left panel) and needlet cross-spectra (right panel) divided by the respective analytical full-sky variance. Note that we calculate the ratio $\hat{\beta}_j/\Delta\beta_j$ for the H-ATLAS case, not the pseudo spectrum $\hat{\Gamma}_j/\Delta\Gamma_j$.

## 6.5 Conclusions

Cross-correlation analyses between independent cosmological datasets have the advantage to be potentially immune to any known (and unknown) systematics, as well as to extract signals hidden in noisy data. In this way, cross-correlation measurements can provide us with a clearer view of the large scale distribution of matter, fundamental to reconstruct the dynamics and the spatial distribution of the gravitational potential that can be then translated into constraints on cosmological parameters, breaking degeneracies with the astrophysical ones.

In this chapter we begin a systematic analysis of the scientific potential associated to the expansion of the analysis domain in CMB-LSS cross-correlation studies to include the localization in the harmonic and spatial domains. In this initial application, by exploiting an ensemble of simulations, we have shown that under the same observational configurations the needlet spectral estimator can outperform the harmonic one thanks to the excellent needlets localization properties in both pixel and frequency space, as well as their optimal window function. Moreover, we have completed an initial needlet based analysis pipeline throughout the implementation of a novel `MASTER`-like approach for needlet spectral reconstruction in the case of aggressive masking ($f_{\text{sky}} \simeq 0.01$), reporting an higher S/N with respect to its harmonic counterpart.

Motivated by these positive indications and results, in future research we plan to explore further the trade-off between S/N and multipole localization, so as to achieve optimal bandwidth selection for a given experimental setting (such as the Euclid coverage mask). We also aim at applying this machinery to accurate CMB maps lensed with ray-tracing techniques (Calabrese et al., 2015) and realistic galaxy mock catalogues based on N-body simulations by adopting, on the CMB side, the projected accuracy and sensitivity of forthcoming polarization oriented CMB probes, targeting the B-modes from cosmological gravitational waves and gravitational lensing. This work is of course preparatory for application to real data, from currently available LSS maps such as Herschel and WISExSCOS Photometric Redshift Catalogue (WISExSCOSPZ) (Bilicki et al., 2016) to upcoming surveys such as Euclid, LSST, DESI, and WFIRST, in order to robustly extract cosmological information from cross-correlation measurements.

# 7 Chapter 7
# Kinetic Sunyaev-Zel'dovich effect in modified gravity

In this Chapter, which concludes the work described in this Thesis, we tackle the interplay between CMB and LSS from a different perspective, focusing not only upon a different effect with respect to CMB lensing, but also investigating its capabilities as a discriminant from cosmologies characterized by modifications of gravity. Here we investigate the impact of modified theories of gravity on the kinetic Sunyaev-Zel'dovich (kSZ) effect of the CMB. We focus on a specific class of $f(R)$ models of gravity and compare their predictions for the kSZ power spectrum to that of the $\Lambda$CDM model.[1] We use a publicly available modified version of `Halofit` to properly include the nonlinear matter power spectrum of $f(R)$ in the modeling of the kSZ signal. We find that the well-known modifications of the growth rate of structure in $f(R)$ can indeed induce sizable changes in the kSZ signal, which are more significant than the changes induced by modifications of the expansion history. We discuss prospects of using the kSZ signal as a complementary probe of modified gravity, giving an overview of assumptions and possible caveats in the modeling.

## 7.1 The kinetic Sunyaev-Zel'dovich effect

As we have already seen in Ch. (2.2.2), the journey of the photons from the last scattering surface to us is not a smooth one. On their way to the observer, they undergo several physical processes which induce additional anisotropies on top of the primordial ones, imprinted at the time of recombination. These can be noticed as secondary effects on the CMB, of which the iSW effect and CMB lensing are common examples. Since they happen at different intermediate redshifts, they carry valuable information on the Universe between the last scattering and today.

---

[1] $f(R)$ gravity is a family of gravity theories in which the Lagrangian density $f$ is an arbitrary function of $R$ and cosmological models are derived from that.

In this chapter we focus on the kSZ effect, which is sourced by the inverse Compton scattering of CMB photons off clouds of moving free electrons (Sunyaev & Zeldovich, 1980; Vishniac, 1987; Jaffe & Kamionkowski, 1998; Shaw et al., 2012). The kSZ signal is sensitive to both the expansion history and the dynamics of cosmological perturbations. Its constraining power as a geometrical probe for DE studies was investigated in Ma & Zhao (2014). We focus on the growth of structure and investigate the prospect of using the kSZ effect to constrain models of modified gravity that address the phenomenon of cosmic acceleration. Indeed we show that the kSZ phenomenon is particularly sensitive to the growth of perturbations at redshifts $z \lesssim 2$ and, as such, we expect it to be a good probe of cosmological models that introduce late times modifications of the growth rate of structure, such as clustering dark energy and modified gravity. In particular, we consider the class of $f(R)$ models introduced by Hu & Sawicki (2007) and we find that the corresponding modifications to the growth rate of structure can indeed induce sizable changes in the kSZ effect. We discuss also the caveats and assumptions associated to the modeling and measurements of the kSZ signal, and conclude giving an outlook of the observational prospective.

A wide set of observational probes has been proposed and exploited to test theories of modified gravity from astrophysical to cosmological scales. In particular, viable $f(R)$ models have been constrained with secondary CMB anisotropies, such as CMB lensing, the iSW effect (Raveri et al., 2014; Planck Collaboration, 2015d) and its cross-correlation with galaxy density (Song et al., 2007b; Giannantonio et al., 2010), galaxy clusters abundances (Jain & Zhang, 2008; Ferraro et al., 2011; Cataneo et al., 2015) and profiles (Lombriser et al., 2012; Wilcox et al., 2015), galaxy power spectrum (Oyaizu et al., 2008; Dossett et al., 2014), redshift-space distortions from spectroscopic surveys (Guzzo et al., 2008; Yamamoto et al., 2010; Alam et al., 2016), weak gravitational lensing (Simpson et al., 2013; Harnois-Déraps et al., 2015), 21-cm intensity mapping (Hall et al., 2013), and dwarf galaxies (Jain et al., 2013; Vikram et al., 2013).

The chapter is organized as follows. In Sec. (7.2) we provide an overview of the kSZ effect and describe our modeling of the kSZ angular power spectrum. The different gravity models considered in this analysis are reviewed in Sec. (7.3). We investigate the effect of the modified cosmic growth history on the kSZ observable and discuss possible caveats in Sec. (7.4). In Sec. (7.5) we discuss the detectability of such an effect with the available and upcoming observations, while the conclusions are presented in Sec. (7.6). Throughout the chapter we assume a fiducial background expansion that closely mimics one of a spatially flat ΛCDM cosmological model consistent with the best-fit cosmological parameters derived using the joint data

sets of 2015 Planck TT + low P (Planck Collaboration, 2015c), namely $\Omega_b = 0.049$, $\Omega_c = 0.264$, $\Omega_\Lambda = 0.687$, $n_s = 0.966$, $H_0 = 67.31$ km s$^{-1}$ Mpc$^{-1}$, and $z_{\rm re} = 9.9$.

## 7.2 The modeling of the kSZ effect

In this section we review the physics and the modeling of the kSZ signal. During their journey from the last scattering surface to the observer, CMB photons experience further Compton scattering off clouds of moving free electrons: the LOS component of the electron momentum induces temperature fluctuations in the CMB sky through Doppler effect. This phenomenon is known as kinetic Sunyaev-Zel'dovich effect and its magnitude is given by (Sunyaev & Zeldovich, 1980)

$$\frac{\Delta T}{T_0}(\hat{\mathbf{n}}) = \sigma_T \int \frac{\mathrm{d}z}{(1+z)H(z)} n_e(z) e^{-\tau(z)} \mathbf{v} \cdot \hat{\mathbf{n}}, \tag{7.1}$$

where $\sigma_T$ is the Thomson scattering cross section, $T_0 \simeq 2.725$ K is the average CMB temperature, $H(z)$, $\tau(z)$, $n_e(z)$, and $\mathbf{v} \cdot \hat{\mathbf{n}}$ are the Hubble parameter, optical depth, free electron number density, and peculiar velocity component along the LOS.

The optical depth out to redshift $z$ is

$$\tau(z) = \sigma_T c \int_0^z \mathrm{d}z \frac{\bar{n}_e(z)}{(1+z)H(z)}, \tag{7.2}$$

where $\bar{n}_e$ is the mean free-electron density,

$$\bar{n}_e = \frac{\chi \rho_g(z)}{\mu_e m_p}, \tag{7.3}$$

with $\rho_g(z) = \rho_{g0}(1+z)^3$ being the mean gas density at redshift $z$ and $\mu_e m_p$ the mean mass per electron. The electron ionization fraction is defined as (Shaw et al., 2012; Ma & Zhao, 2014):[2]

$$\chi = \frac{1 - Y_{\rm p}(1 - N_{\rm He}/4)}{1 - Y_{\rm p}/2}. \tag{7.4}$$

Here, $Y_{\rm p}$ is the primordial helium abundance, fixed to $Y_{\rm p} = 0.24$, and $N_{\rm He}$ is the number of helium electrons ionized.

The kSZ power is expected to be sourced by a contribution from the epoch of

---

[2]Note that in this chapter we switch convention and denote the comoving distance with the letter $x$ to avoid confusion with the electron ionization fraction $\chi$ and stick to the most diffused kSZ literature.

reionization, dubbed the patchy kSZ, and one from the postreionization epoch, namely the homogeneous kSZ.

- **Patchy kSZ**: it is originated by inhomogeneities in the electron density and the ionization fraction at redshift $z \gtrsim z_{\mathrm{re}}$;[3] its power spectrum amplitude and shape depend, at first order, on the time and duration of reionization (Zahn et al., 2005; McQuinn et al., 2005);

- **Homogeneous kSZ:** it is sourced by free-electron density perturbations and peculiar velocities at $z \lesssim z_{\mathrm{re}}$.

Since we focus on the kSZ signal after reionization, we set the upper integration limit in Eq. (7.1) to $z_{\mathrm{re}} = 9.9$ while we assume $\chi = 0.86$, i.e. neutral helium at all redshifts.[4] We discuss more in detail about the possible choices for the reionization redshift and the dependence of kSZ on its value in Sec. (7.4.4).

Writing $n_e = \bar{n}_e(1 + \delta)$, Eq. (7.1) can be recast as

$$\frac{\Delta T}{T_0}(\hat{\mathbf{n}}) = \left(\frac{\sigma_T \rho_{g0}}{\mu_e m_p}\right) \int_0^{z_{re}} \mathrm{d}z \frac{(1+z)^2}{H(z)} \chi e^{-\tau(z)} \mathbf{q} \cdot \hat{\mathbf{n}}, \qquad (7.5)$$

where $\mathbf{q} = \mathbf{v}(1 + \delta)$ is the density weighted peculiar velocity (ionized electron peculiar momentum).

Since the longitudinal Fourier modes of $\mathbf{q}$, i.e those with $\mathbf{k}$ parallel to $\hat{\mathbf{n}}$, experience several cancellations in the LOS integral of Eq. (7.5), only transverse momentum modes contribute to the effect (Vishniac, 1987; Jaffe & Kamionkowski, 1998). Moreover, in the linear regime the velocity field is purely longitudinal, so that only the cross term $\mathbf{v}\delta$ can source the kSZ power spectrum. In the small angle limit, the kSZ angular power spectrum can be written under the Limber approximation (Limber, 1953) as

$$C_\ell = \frac{8\pi^2}{(2\ell+1)^3}\left(\frac{\sigma_T \rho_{g0}}{\mu_e m_p}\right)^2 \int_0^{z_{re}} \frac{\mathrm{d}z}{c}(1+z)^4 \chi^2 \Delta_B^2(\ell/x, z)e^{-2\tau(z)}\frac{x(z)}{H(z)}, \qquad (7.6)$$

where $x(z) = \int_0^z (c dz'/H(z'))$ is the comoving distance at redshift z, $k = \ell/x$ and $\Delta_B(k, z)$ is the power spectrum of the transverse (curl) component of the momentum field. Earlier analytical studies calculated the expression for $\Delta_B^2$ (Vishniac, 1987;

---

[3]We define $z_{\mathrm{re}}$ as the redshift at which hydrogen reionization ends.

[4]This is an assumption: in a more realistic case, helium atoms remain singly ionized until $z \sim 3$ ($\chi = 0.93$) and, below that, they are thought to be doubly ionized ($\chi = 1$). Shaw et al. (2012) calculates for their $\Lambda$CDM case that this would increase the kSZ power by a factor of 1.22 at $\ell = 3000$ relative to their baseline model. Hence, the level of uncertainty on the kSZ power spectrum due to helium reionization is equivalent to that due to the uncertainty on $\sigma_8$.

Jaffe & Kamionkowski, 1998; Dodelson & Jubas, 1995; Ma & Fry, 2002) to be

$$
\Delta_B^2(k, z) = \frac{k^3}{2\pi^2} \int \frac{\mathrm{d}^3 \mathbf{k}'}{(2\pi)^3} (1 - \mu^2) \Big[ P_{\delta\delta}(|\mathbf{k} - \mathbf{k}'|, z) P_{vv}(k', z)
$$
$$
- \frac{k'}{|\mathbf{k} - \mathbf{k}'|} P_{\delta v}(|\mathbf{k} - \mathbf{k}'|, z) P_{\delta v}(k', z) \Big],
$$

$$(7.7)$$

where $P_{\delta\delta}$ and $P_{vv}$ are the linear matter density and velocity power spectra, while $P_{\delta v}$ is the density-velocity cross spectrum and $\mu = \hat{\mathbf{k}} \cdot \hat{\mathbf{k}}'$. We can further simplify Eq. (7.7) by relating the peculiar velocity field to the density perturbations: in the linear regime and on subhorizon scales, under the assumption that $\dot{\Phi} \simeq 0$, we can use the continuity equation for matter to write $\tilde{\mathbf{v}} = i f \dot{a} \tilde{\delta} \mathbf{k} / k^2$, where $f(a, k) = d \log \delta(a, k) / d \log a$ is the linear growth rate. This gives us the following equations

$$
P_{vv}(k) = \left(\frac{f\dot{a}}{k}\right)^2 P_{\delta\delta}(k); \qquad P_{\delta v}(k) = \left(\frac{f\dot{a}}{k}\right) P_{\delta\delta}(k).
$$

$$(7.8)$$

Inserting them into Eq. (7.7), we obtain[5]

$$
\Delta_B^2(k) = \frac{k^3}{2\pi^2} \dot{a}^2 \int \frac{\mathrm{d}^3 \mathbf{k}'}{(2\pi)^3} f^2(k') P_{\delta\delta}(|\mathbf{k} - \mathbf{k}'|) P_{\delta\delta}(k') I(k, k'),
$$

$$(7.9)$$

where

$$
I(k, k') = \frac{k(k - 2k'\mu)(1 - \mu^2)}{k'^2(k^2 + k'^2 - 2kk'\mu)}
$$

$$(7.10)$$

is the kernel that describes how linear density and velocity fields couple to each other. Since we are interested in studying the impact of scale-dependent growth of structure induced by modified gravity scenarios on the kSZ observable, $f$ is kept inside the previous integral. Combining Eq. (7.6) with Eq. (7.9), one obtains the so-called Ostriker-Vishniac (OV) effect (Ostriker & Vishniac, 1986).

When nonlinear structure formation arises, linear perturbation theory does not hold and Eq. (7.9) breaks down: several previous works have investigated the impact of nonlinearities on the kSZ power spectrum (Hu, 2000a; Ma & Fry, 2002; Shaw et al., 2012). In particular, Hu (2000a); Ma & Fry (2002) argue that nonlinearities in the velocity field are suppressed by a factor of $1/k^2$ with respect to those in the density field. Hence it is sufficient to replace the linear $P_{\delta\delta}(k, z)$ in Eq. ((7.9)) with its nonlinear counterpart $P_{\delta\delta}^{\mathrm{NL}}(k, z)$ to capture the effect of nonlinearities, so that

$$
\Delta_B^2(k) = \frac{k^3}{2\pi^2} \dot{a}^2 \int \frac{\mathrm{d}^3 \mathbf{k}'}{(2\pi)^3} f^2(k') P_{\delta\delta}^{\mathrm{NL}}(|\mathbf{k} - \mathbf{k}'|) P_{\delta\delta}(k') I(k, k').
$$

$$(7.11)$$

---

[5]Redshift dependence is suppressed for clarity.

Throughout this chapter, we use Eq. (7.11) and Eq. (7.6) to compute theoretical predictions for the full kSZ angular power spectra. Let us notice that the continuity equation that we have used to go from ((7.7)) to ((7.9)) remains valid in our Jordan frame treatment of $f(R)$ gravity. N-body simulations have shown that $f(R)$ models affect significantly the velocity spectrum, with resulting changes in $P_{vv}$ with respect to the $\Lambda$CDM being more sizable than the corresponding changes in $P_{\delta\delta}$ (Li et al., 2013). Nevertheless, the relative size of these modifications should not hinder the above assumption on negligibility of nonlinearities of the velocity power spectrum in $\Delta_B^2$.

As previously done (Jaffe & Kamionkowski, 1998; Dodelson & Jubas, 1995; Ma & Fry, 2002; Zhang et al., 2004), we assume that gas fluctuations trace dark matter ones at all scales, i.e. there is no (velocity) bias between them. Note that this approximation breaks on small scales, where baryon thermal pressure tends to make gas distribution less clustered than the dark matter one. This suppression effect due to baryon physics can be incorporated into a window function $W(k, z)$, such that (Shaw et al., 2012)

$$P_{gas}^{\mathrm{NL}}(k, z) = W^2(k, z) P_{\mathrm{DM}}^{\mathrm{NL}}(k, z). \tag{7.12}$$

An investigation of the astrophysical processes such as radiative cooling, star formation, and supernova feedback on the kSZ power spectrum achieved by measuring the window function $W^2(k, z)$ in hydrodynamic simulations is provided in Shaw et al. (2012) (the *cooling and star-formation* model, labelled CSF). There, the authors provide an improved fitting formula for $W^2(k, z)$ with respect to the one presented in Gnedin & Hui (1998) and argue that the CSF model is a robust lower bound on the homogeneous kSZ amplitude.

All the previous studies on which we rely to model the kSZ signal, neglect the contribution of the connected term in the transverse component of the momentum power spectrum[6] that arises in the nonlinear regime: a recent investigation of this term is reported in Park et al. (2016).

Since we aim at studying the modified gravity effects on the kSZ power spectrum, in the following analysis we adopt the full nonlinear kSZ modeling without including the thermal gas pressure. However, when comparing with observations, one should take that into account and exploit numerical simulations to calibrate. As discussed in Shaw et al. (2012), the realistic expected homogeneous kSZ power spectrum should lie within the region delimited by the OV signal (lower bound) and the full kSZ (upper bound).

---

[6] As we have seen, $\mathbf{q} \approx \mathbf{v}\delta$, hence $\langle qq \rangle = \langle vv \rangle \langle \delta\delta \rangle + 2\langle v\delta \rangle^2 + \langle v\delta v\delta \rangle_c$

## 7.3 Models

In our analysis we compare the kSZ effect in the standard cosmological model, $\Lambda$CDM to that in the $f(R)$ class of modified gravity. As we discuss, the latter models have an expansion history which is very close to the $\Lambda$CDM one, while they can predict a different cosmic growth history. In our analysis we therefore focus on the effects of the modified growth.

For our standard gravity scenario, we consider a flat $\Lambda$CDM composed of baryonic matter (with critical density $\Omega_b$), pressureless cold DM ($\Omega_c$), and a cosmological constant $\Lambda$ exerting a negative pressure and dominating the current energy budget of the Universe ($\Omega_\Lambda$). The power spectrum of the transverse component of the momentum field $\Delta_B^2(k, z)$ [Eq. (7.11)], which sources the kSZ signal, is sensitive to the growth history through the linear growth rate and to the matter power spectrum, both the linear and nonlinear one. In our analysis, we use $f(z) = \Omega_m(z)^\gamma$ (Linder, 2005), where $\gamma \simeq 0.55$ for general relativity, to evaluate the growth rate, while we make use of CAMB[7] (Lewis et al., 2000) to compute the matter power spectrum. In this way, nonlinearities in $P_{\delta\delta}^{\mathrm{NL}}(k)$ are taken into account via the Halofit prescription of Smith et al. (2003); Takahashi et al. (2012).

For the case of $f(R)$ gravity (Song et al., 2007a; Bean et al., 2007; Pogosian & Silvestri, 2008; De Felice & Tsujikawa, 2010), we consider a specific family of models introduced by Hu and Sawicki in Hu & Sawicki (2007). This represents one of the viable $f(R)$ families, capable of sourcing the late-time accelerated expansion, closely mimicking the $\Lambda$CDM background cosmology on large scales and evading solar system tests because of the built-in chameleon mechanism (Hu & Sawicki, 2007; Khoury & Weltman, 2004). The action of the model in the Jordan frame reads

$$ S = \frac{1}{16\pi G} \int \mathrm{d}^4 x \sqrt{-g}[R + f(R)] + S_m, \tag{7.13} $$

where $f(R)$ is a generic function of the Ricci scalar

$$ f(R) = -m^2 \frac{c_1(-R/m^2)^n}{c_2(-R/m^2)^n + 1}, \tag{7.14} $$

and the matter sector is minimally coupled to gravity. Here, $c_1$, $c_2$, and $n$ are free dimensionless parameters of the theory and $m^2$ is a mass scale. Because of the higher order derivative nature of the theory, there is a dynamical scalar degree of freedom $f_R \equiv df/dR$, dubbed the scalaron, that mediates a long-range fifth-force and is responsible for the modification of structure formation.

---

[7]http://cosmologist.info/camb/

Following Hu & Sawicki (2007), we fix the mass scale to $m^2 = 8\pi G \bar{\rho}_0/3$, where $\bar{\rho}_0$ is the average density of matter today. This effectively corresponds to having $m^2/R \ll 1$ for the entire expansion history, allowing an expansion of ((7.14)) in $m^2/R$, with the scalaron sitting always close to the minimum of the potential, and the model resembling, at linear order, the $\Lambda$CDM one. If we additionally fix $c_1/c_2 \sim 6\Omega_\Lambda/\Omega_m$, then in first approximation the expansion history mimics closely that of a $\Lambda$CDM universe with a cosmological constant $\Omega_\Lambda$ and matter density $\Omega_m$. This leaves us with two free parameters, $c_1/c_2^2$ and $n$; models with larger $n$ mimic $\Lambda$CDM until later in cosmic time, while models with smaller $c_1/c_2^2$ mimic it more closely.

In the following we work in terms of $n$ and the present value of the scalaron $f_R^0$, which can be related to $c_1/c_2^2$ as (Hu & Sawicki, 2007)

$$\frac{c_1}{c_2^2} = -\frac{1}{n}\left[3\left(1 + 4\frac{\Omega_\Lambda}{\Omega_m}\right)\right]^{n+1} f_R^0. \tag{7.15}$$

In order to calculate the kSZ power spectrum, we use the publicly available `MGCAMB` (Hojjati et al., 2011) and `MGHalofit` (Zhao, 2014) to compute the growth rate and the linear and nonlinear matter power spectrum. Following the discussion above, we approximate the expansion history to the $\Lambda$CDM one, and focus on differences due to a different growth rate of structure. `MGCAMB` and `MGHalofit` use the quasistatic approximation for the dynamics of scalar perturbations, which is sufficiently good for the choice of parameters that we have made above and for the range of $f_R^0$ values that we explore (Hojjati et al., 2012). In the following analysis we consider models with $|f_R^0| \in [10^{-6}, 10^{-4}]$ and $n = 1$.

## 7.4 Results and Discussion

We now proceed to calculate the kSZ power spectrum for the above models. Before discussing the final results, let us give an overview of the different quantities that contribute to $\Delta_B^2$ (Eq. (7.11)) and how they differ in these models. In particular, we discuss how the differences in the dynamics of perturbations for $f(R)$ models reflect in a modified kSZ signal. We also give an overview of other physical phenomena to which the kSZ is sensitive regardless of the theory of gravity under consideration, discussing how they can affect our ability to test gravity with kSZ.
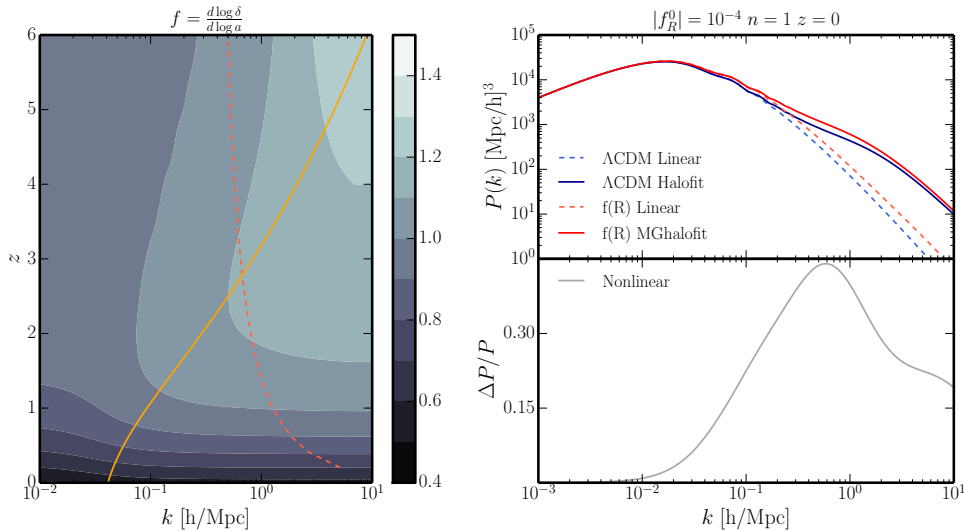
**Figure 7.1:** Right panel: The linear growth rate $f = d \log \delta / d \log a$ as a function of scale $k$ and redshift $z$ for the $f(R)$ model with a $\Lambda$CDM background expansion with $|f_R^0| = 10^{-4}$ and $n = 1$. The characteristic scale dependence of the structure growth is clearly visible: at a given redshift, different scales are characterized by a different growth rate. The red dashed line represents the size-distance relation, $k = \ell / x(z)$, for $\ell = 3000$ while the orange solid line corresponds to the Compton scale $k_C$ associated to the scalaron. Top righ panel: Comparison between matter power spectra in $\Lambda$CDM and $f(R)$ models at redshift $z = 0$: linear predictions are shown as dashed lines, while nonlinear ones are shown as solid lines. Nonlinear matter power spectra in $\Lambda$CDM are evaluated using `Halofit` prescription, while we make use of `MGHalofit` for the Hu and Sawicki model (with $|f_R^0| = 10^{-4}$ and $n = 1$). Bottom right panel: The relative difference between the *nonlinear* matter power spectra in $f(R)$ and $\Lambda$CDM cosmology.

## 7.4.1 Growth history

Being a weighted integral of $\Delta_B^2$ over the redshift, the kSZ power spectrum probes the cosmic growth history across a wide range of scales and redshifts through the evolution of matter and velocity perturbations Shaw et al. (2012). In our analysis of $f(R)$ models, we obtain the linear growth rate $f(z, k)$ using `MGCAMB` and following the procedure outlined in Ma & Zhao (2014): first we output the density contrast as a function of redshift and scale, and then we evaluate its logarithmic derivative with respect to the scale factor. In the left panel of Fig. (7.1) we show the growth rate as a function of scale $k$ and redshift $z$ for the Hu and Sawicki model with $|f_R^0| = 10^{-4}$: it can be seen that $f(R)$ models enhance the growth of matter perturbations on scales smaller than their characteristic length scale, i.e. the scalaron Compton wavelength $\lambda_C^2 \sim f_{RR}/(1 + f_R)$. On the contrary, the growth in $\Lambda$CDM depends only on time.

The differences in the growth history show up in the matter power spectrum. In the

right panel of Fig. (7.1), we plot the present time matter power spectrum, $P_{\delta\delta}(k)$, for $\Lambda$CDM and the Hu-Sawicki $f(R)$ model with $|f_R^0| = 10^{-4}$, including both the linear and nonlinear cases. It can be noticed that the scale-dependent growth of $f(R)$ produces an enhancement of the power spectrum on smaller scales, with the effect kicking in at a $k$ proportional to the Compton wave number. The bigger $f_R^0$, the smaller the $k$ at which the enhancement kicks in.
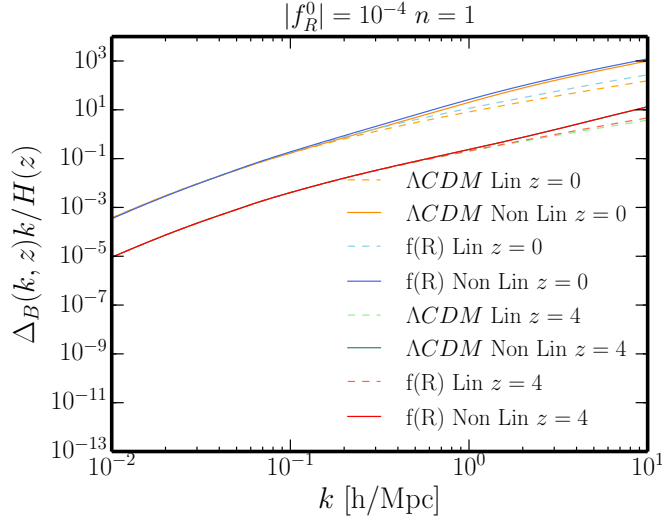


**Figure 7.2:** Power spectrum of the curl component of the momentum field calculated for the different gravity model, regimes, and redshifts.

In Fig. (7.2) we plot the momentum power, $\Delta_B^2$, in terms of the dimensionless quantity $\Delta_B k/H(z)$ introduced in Zhang et al. (2004); Shaw et al. (2012): dashed lines represent linear regime calculations (OV effect), while solid lines include nonlinear corrections to the matter power spectra (full kSZ effect). The plot shows that as cosmic structure evolves over time, the amplitude of the momentum curl component power spectra increases: that is equally true for $f(R)$ and $\Lambda$CDM. However, the modifications of the growth of structure in $f(R)$ imprint a different shape on $\Delta_B^2$ and enhance its power. Nonlinear density fluctuations become relevant when $\Delta_B k/H(z) \approx 1$: as the Universe evolves, density perturbations exceed unity and they increasingly become important at larger scales (Zhang et al., 2004).

## 7.4.2 The kSZ signal

We now move on to understand how the differences at the perturbation level translate into differences in the predicted kSZ angular power spectrum, showing how kSZ measurements can potentially represent a novel test of gravity.
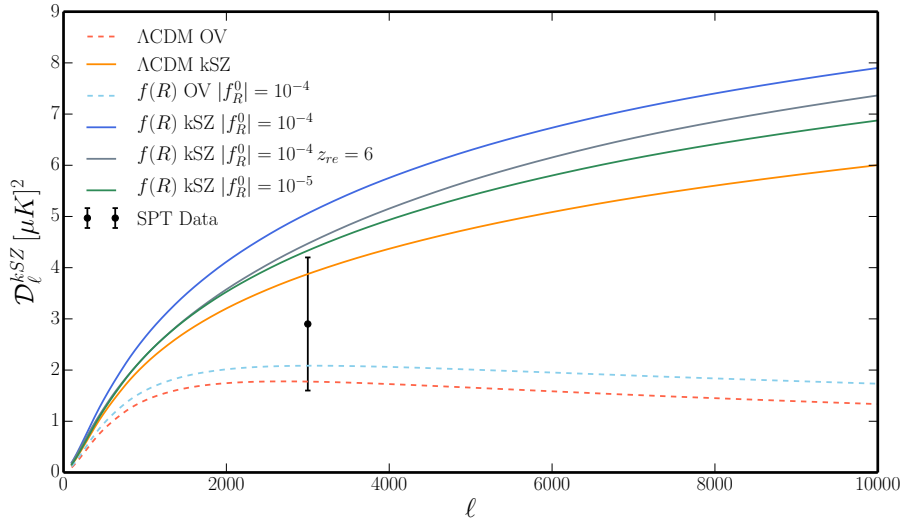
**Figure 7.3:** The homogeneous kinetic Sunyaev-Zel'dovich power spectrum (solid lines) for standard $\Lambda$CDM and Hu and Sawicki models with $|f_R^0| = \{10^{-5}, 10^{-4}\}$ and $n = 1$ as a function of multipole $\ell$. The dashed lines show the linear predictions, i.e. the OV effect. The black data band power $\mathcal{D}_{\ell=3000}^{\mathrm{k}SZ} = 2.9 \pm 1.3\mu\mathrm{K}^2$ ($1\sigma$ confidence level) is taken from the South Pole Telescope (SPT) George et al. (2015)). $z_{\mathrm{re}} = 9.9$ is assumed except where otherwise stated.

In Fig. (7.3) we show the kSZ angular power spectrum as a function of multipole $\ell$ in terms of $\mathcal{D}_\ell \equiv \ell(\ell+1)C_\ell^{\mathrm{k}SZ}/2\pi$. Full kSZ theory spectra are plotted as solid lines while OV calculations are shown as dashed lines. We plot results for the $\Lambda$CDM scenario and for two Hu and Sawicki models with $|f_R^0| = \{10^{-5}, 10^{-4}\}$. The comparison between dashed and solid lines for a fixed model provides an estimate of the impact of nonlinearities in structure formation to the kSZ power spectrum, as also shown in Shaw et al. (2012); Ma & Zhao (2014). Nonlinear structure boosts the $\Lambda$CDM homogeneous kSZ power by a factor $\approx 2.2$ at $\ell = 3000$. On top of nonlinearities, in $f(R)$ cosmology the scale-dependent growth rate enhances the power spectrum by $\approx 2.5$ (2.4) for $|f_R^0| = 10^{-4}$ ($10^{-5}$). The net effect of modifying gravity is to enlarge the expected kSZ signal: the $C_\ell^{\mathrm{k}SZ}$ amplitude increases as $|f_R^0|$ becomes larger (i.e. chameleon mechanism is less efficient). Note that the $\ell$-dependence is not dramatically altered. When comparing full kSZ power for different gravity models, we find $C_\ell^{\mathrm{k}SZ}(|f_R^0| = 10^{-4})$ is approximately 30% larger than the $\Lambda$CDM values, while $C_\ell^{\mathrm{k}SZ}(|f_R^0| = 10^{-5})$ is $\approx 11\%$ larger than the $\Lambda$CDM one.
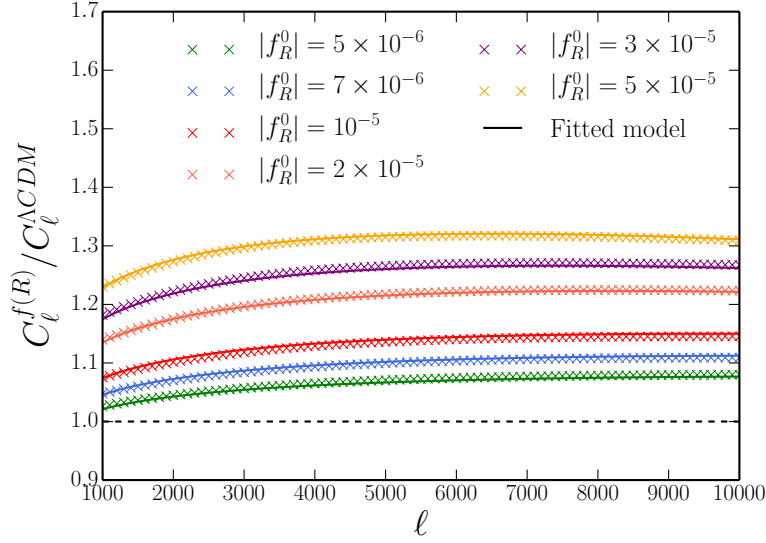
**Figure 7.4:** $C_\ell^{\mathrm{k}SZ}$'s ratio between ΛCDM and Hu & Sawicki model for different values of $|f_R^0|$ with the fitting formula introduced in Eq. (7.16). Notice that the fit has been done using $C_\ell$'s between $\ell \in [10^3, 10^4]$ output for six $|f_R^0|$ values in the range $[5 \times 10^{-6}, 5 \times 10^{-5}]$.

## 7.4.3 Fitting formula

As noted in Ma & Zhao (2014), the numerical evaluation of $C_\ell^{\mathrm{k}SZ}$'s is computationally expensive and makes the exploration of the parameter space through Markov chain Monte Carlo (MCMC) rather unfeasible. To this end, we develop a fitting formula to predict the kSZ angular power spectrum for the Hu-Sawicki $f(R)$ model with $n = 1$. We consider the power spectra ratio $C_\ell^{f(R)}/C_\ell^{\Lambda CDM}$ for several values of $|f_R^0|$ in the representative range $[5 \times 10^{-6}, 5 \times 10^{-5}]$ and perform a fit via the following formula

$$\frac{C_\ell^{f(R)}(|f_R^0|)}{C_\ell^{\Lambda CDM}} = A\left(\frac{\ell}{1000}\right)^B \exp\left(-C\frac{\ell}{1000}\right),\tag{7.16}$$

where $A$ depends on $|f_R^0|$ through $a_0 + a_1 \log_{10}(|f_R^0|) + a_2 \log_{10}^2(|f_R^0|)$ (same dependence for $B$ and $C$). The outcome of this fitting procedure gives:

$$\begin{aligned}
A(|f_R^0|) &= 3.308 + 0.603 \log_{10}(|f_R^0|) + 0.034 \log_{10}^2(|f_R^0|)\\
B(|f_R^0|) &= -0.261 - 0.169 \log_{10}(|f_R^0|) - 0.021 \log_{10}^2(|f_R^0|)\\
C(|f_R^0|) &= 0.182 + 0.044 \log_{10}(|f_R^0|) + 0.002 \log_{10}^2(|f_R^0|).
\end{aligned}\tag{7.17}$$

As can be noticed in Fig. (7.4), the above curves are a very good fit. Hence Eq. ((7.16)) provides an accurate way of modeling the kSZ spectrum in the $f(R)$

models under consideration, allowing one to explore the parameter space in a much faster way. Let us notice that the fitting parameters in ((7.17)), and possibly the optimal fitting curve, may change if a different range of values for $|f_R^0|$ is considered.

### 7.4.4 Caveats *et al.*

The focus of our study is the impact of modifications of gravity on the kSZ signal. However, since we aim at determining whether the kSZ can be a useful probe for modified gravity, in this section we review astrophysical and cosmological effects that could be degenerate with the main signatures we are after, as well as limitations of the theoretical modeling.

**Reionization redshift**

The amplitude of the kSZ power spectrum depends on the redshift $z_{re}$ at which reionization occurs through Eq. (7.6). If reionization ends at higher redshifts, the integral which sources the kSZ power picks up more signal, hence it increases. We find that by fixing $z_{re} = 6$, $C_{\ell=3000}^{kSZ}$ is decreased by approximately 15% and 11% with respect to the baseline $\Lambda$CDM and Hu and Sawicki ($|f_R^0| = 10^{-4}$) results.

**Helium reionization**

So far, we have assumed that helium remains neutral throughout cosmic history ($\chi = 0.86$), although in a more realistic model helium would be singly ionized ($\chi = 0.93$) between $3 < z \leq z_{re}$ and doubly ionized ($\chi = 1$) for $z \leq 3$. The magnitude of $C_\ell^{kSZ}$ scales as the square of the ionization fraction (plus integrated dependence through the optical depth $\tau$), so we expect that helium reionization would boost the power spectrum.

**Applicability of `MGHalofit`**

A fundamental ingredient needed to predict the kSZ power is the matter power spectrum, especially the nonlinear counterpart: as stated in Sec. (7.3), in our analysis we make use of `Halofit` for the $\Lambda$CDM case (as done in Shaw et al. (2012); Ma & Zhao (2014)) and `MGHalofit` for the $f(R)$ model. The accuracy level of the fitting formula to calculate the nonlinear matter power spectrum is crucial if one wants to test gravity with kSZ effect. We recall that `MGHalofit` works for an arbitrary $|f_R^0| \in [10^{-6}, 10^{-4}]$ below redshift $z = 1$ (reaching an accuracy of 6% and 12% at $k \leq 1\,h/\mathrm{Mpc}$ and $k \in (1, 10]\,h/\mathrm{Mpc}$).[8] While it is nontrivial to address

---

[8]Note that the standard `Halofit` accuracy is below 5% ($k \leq 1h/\mathrm{Mpc}$) and 10% ($k \in (1, 10]h/\mathrm{Mpc}$ at $z \leq 3$ (Takahashi et al., 2012), so that previous analytical $\Lambda$CDM cal-

quantitatively and self consistently this issue, we expect nonlinearities to be less important at redshift $z \gtrsim 1$. In the left panel of Fig. (7.5) we show the differential contribution to the kSZ power of redshift slices in the range $0 \leq z \leq z_{re}$ at $\ell = 3000$ for the different cases studied in our analysis. The difference between the OV and kSZ lines (within the same gravity scenario) gives an estimate of the kSZ power enhancement due to nonlinear evolution: moreover, we can see that in the full kSZ case, the bulk of the signal is sourced by structure at $z \lesssim 1$ (especially for the $f(R)$ model). The right panel of Fig. (7.5) shows the differential contribution $d\mathcal{D}_\ell^{kSZ}/dz$ for the $f(R)$ kSZ case with $|f_R^0| = 10^{-4}$ at different $\ell$'s: as the multipole becomes larger, the differential redshift distribution peak shifts toward higher $z$ (but always smaller than 1).

**Gastrophysics**

Shaw et al. (2012) have shown that the effect of the baryon physics, i.e. cooling and star formation, is to reduce the gas density in halos, hence hindering the kSZ power boost due to nonlinear density fluctuations. The authors find a reduction of $C_\ell^{kSZ}$ at all angular scales; in particular the power in the CSF model is reduced by $\approx 30\%$ at $\ell = 3000$ with respect to the model without baryon physics. However, they argue that their radiative simulation suffers from the overcooling problem, so that the measured kSZ power is likely to be underestimated.

**Patchy Reionization**

What we have modeled so far is the homogeneous kSZ signal which is sourced in the postreionization epoch: however, the patchy kSZ contribution should be added on top of that. The precise patchy kSZ power spectrum amplitude and $\ell$-shape depend at first order on the details of reionization, i.e. its time and duration (Zahn et al., 2005; McQuinn et al., 2005; Iliev et al., 2006). In particular, the patchy kSZ signal has a different spectral shape with respect to the homogeneous one and peaks on multipoles between $\ell \approx 2000 - 8000$, roughly corresponding to the typical angular size of ionized regions at the reionization redshift (Iliev et al., 2006). Patchy kSZ can allow us to place constraints on the reionization physics, details of which are not yet well understood, assuming a good knowledge and modeling of the homogeneous signal. To this end, 21-cm fluctuation power spectra and its cross-correlation with the CMB anisotropies can shed light on the details of reionization and help disentangle between different kSZ contributions (Jelic et al., 2014). Very recently, Smith & Ferraro (2016) have proposed a new statistic that is capable to separate the kSZ signal from others small-scale CMB secondary

---

culations are similarly affected by the precision of the fitting formula for matter nonlinearities.

anisotropy, and also allow the kSZ signal to be decomposed in redshift bins. The peculiarity of the method is that it does not rely on external information such as galaxy catalogues and it is claimed to be able to cleanly isolate the high-$z$ signal from patchy reionization.
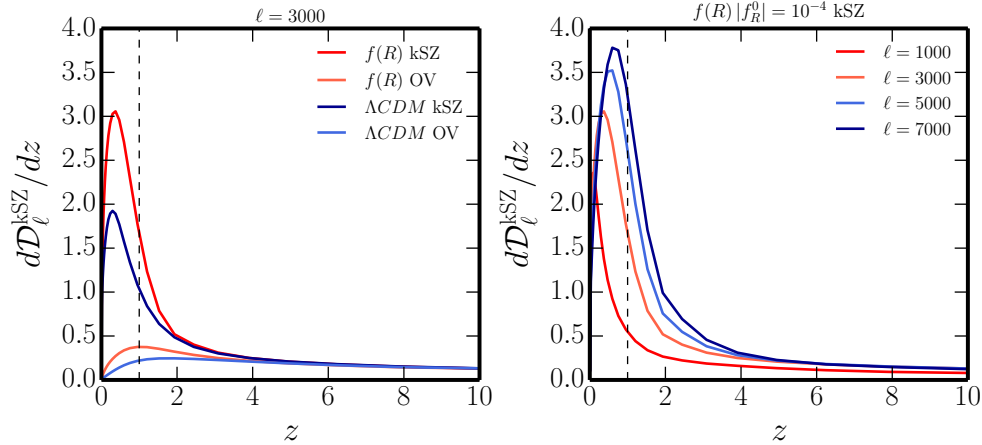


**Figure 7.5:** *Left panel:* the contribution to the kSZ power as function of redshift at $\ell = 3000$ shown for $\Lambda$CDM and Hu & Sawicki model with $|f_R^0| = 10^{-4}$ and $n = 1$. A comparison between the full kSZ and OV effect is also provided. *Right panel:* redshift contribution to kSZ power in the $f(R)$ model for different angular scales. Dashed line marks $z = 1$ in both panels.

## 7.5 Outlook to observations

The kSZ effect was detected observationally for the first time by Hand et al. (2012), applying the pairwise momentum estimator to data from the Atacama Cosmology Telescope (ACT) and the Baryon Oscillation Spectroscopic Survey (BOSS). Recently some authors have performed an analysis of constraints on dark energy and modifications to gravity achievable with the mean pairwise velocity of clusters estimator (Mueller et al., 2015). Other detections using the the pairwise momentum estimator have been reported by the *Planck* team exploiting galaxies from SDSS (Planck Collaboration, 2016c), by the SPT collaboration using galaxies from DES (Soergel et al., 2016), and by the ACTPol team using the BOSS DR11 sources (De Bernardis et al., 2016).

Here, we focus on measurements of the kSZ angular power spectrum obtained from CMB surveys. At high multipoles ($\ell \gtrsim 3000$), there are several sources contributing to the secondary anisotropies of the observed CMB temperature: a number of Galactic and extragalactic astrophysical foregrounds, such as the cosmic infrared

background (CIB), thermal Sunyeav-Zel'dovich (tSZ) and kSZ, radio galaxies, synchrotron and dust emission.[9] Multiband observations are fundamental: the different frequency scalings of these foregrounds can help in separating out the primordial CMB contributions. However, the only contribution that is spectrally degenerate with the primary CMB is the kSZ emission which has a blackbody spectrum.

To the best of our knowledge, the latest observational constraints on the kSZ power spectrum were reported in George et al. (2015) from the combination of the 95, 150, and 220 GHz channel data from the SPT. By jointly fitting the data for the tSZ and kSZ templates (assuming equal power at $\ell = 3000$ from the homogenous (CSF) and patchy contributions), it is found that $\mathcal{D}_{3000}^{\mathrm{k}SZ} < 5.4 \mu\mathrm{K}^2$ at 95% C.L.; when tSZ bispectrum information is added, the derived constraint on the kSZ amplitude is $\mathcal{D}_{3000}^{\mathrm{k}SZ} = 2.9 \pm 1.3 \mu\mathrm{K}^2$ (this data point is shown in Fig. (7.3)). Comparing with our findings, Fig. (7.3), we can see that current kSZ measurements would already have some constraining power. More importantly, we expect kSZ measurements from the ongoing and upcoming CMB surveys to provide a novel complementary probe of gravity on cosmological scales. In particular, it will be essential to rely on arcminute-scale resolution and high sensitivity CMB data since the instrument capability to detect kSZ signal degrades substantially when enlarging the beam size, as noted in Calabrese et al. (2014). In the latter paper, the authors investigate the possibility of using the small-scale polarization information to constrain primordial cosmology in order to remove the primary CMB from temperature measurements, hence isolating the kSZ contribution (assuming an efficient foreground cleaning from multifrequency channels).

Another approach proposed to detect the kSZ signal is to infer the peculiar velocity field $\hat{\mathbf{v}}$ from the observed galaxy number overdensity[10] $\delta_g$ and to stack the CMB temperature maps at the location of each halo, weighted by the reconstructed $\hat{\mathbf{v}}$ field. Recent analysis exploiting this technique has been reported in Planck Collaboration (2016d); Schaan et al. (2016b).

Recently, a different way to detect the kSZ effect based on projected fields have been proposed by Ferraro et al. (2016) and applied in Hill et al. (2016). The basic idea of the method is to (Wiener-)filter the foreground-cleaned CMB temperature maps, square them (to avoid cancellations in the signal), and then cross-correlate them with galaxy position of an external survey. Specificallty, in Hill et al. (2016) the authors have used the (foreground-cleaned) CMB temperature maps built from

---

[9]The effect of weak gravitational lensing is present (but subdominant) at these angular scales and it gently smoothes the CMB peaks.

[10]By solving the linearized continuity equation in redshift space.

multiband *Planck* and WMAP data and the WISE galaxies, yielding a 3.8-4.5 $\sigma$ kSZ detection. It is worth to notice that this method does not require spectroscopic redshift estimate for each cluster but allows to exploit the statistical power of photometric measurements.

## 7.6 Summary

We have revisited the kSZ effect in the context of modified theories of gravity that approach the phenomenon of cosmic acceleration. In particular, we have focused on the class of $f(R)$ models introduced by Hu & Sawicki (2007). We have found that, as expected, the kSZ effect is particularly sensitive to modifications of the growth rate of structure offering, in principle, an interesting complementary probe for modified gravity. Interestingly, we find the kSZ signal to be more sensitive to modifications of the dynamics of cosmological perturbations than to those of the expansion history.

As we have discussed in Sec. (7.4.4), there are several assumptions and caveats in the modeling of the kSZ effect that could hinder its power in constraining modified theories of gravity. We have given a detailed overview of these, elaborating on possible ways of overcoming them, also in light of future experiments. Finally, as this is always the case when the growth rate of structure plays an important role, we expect a degeneracy between the modifications of the kSZ induced by $f(R)$ gravity, that we have discussed, and those that would be induced by massive neutrinos. As part of future work, it would certainly be interesting to explore this latter aspect, as well as to use N-body simulations to get the full nonlinear velocity and density power spectra.

Upcoming high resolution CMB surveys will carry out multifrequency observations in future years (Calabrese et al., 2014), allowing for a reconstruction of the small-scale temperature power spectrum and providing a unique window on the cosmic growth history as well as processes of the epoch of reionization. Given the wide range of modified gravity observational probes and their associated systematics, cross-correlation methods will prove to be robust and complementary tools to determine gravity properties and constrain its behavior over cosmic time.

# Conclusions and future prospects

*There ain't no dark till something shines*
*I'm bound to leave this dark behind*

—Rex's Blues, TOWNES VAN ZANDT

These are exciting times to be a cosmologist and despite the Universe *dark sides*, the future of cosmology looks bright. While *Planck* contributed to consolidate the standard $\Lambda$CDM model, perhaps having the final word on the CMB temperature anisotropies,[11] a lot of work remains to be done in order to address the numerous questions still open. The next cosmological frontiers include, among the others, the measurement of the CMB polarization (especially the *B*-modes) and lensing, as well as the extensive mapping of the LSS at different wavelengths out to high redshifts. The large volume of data gathered by the upcoming CMB and LSS surveys will pose new challenges concerning the best approach to extract cosmological and astrophysical information, how to construct optimal estimators, and how to combine (cross-correlate) different probes. In fact, the main motivations behind this thesis were to move the first steps in this direction.

In this thesis we have studied how to extract cosmological and astrophysical information from the combination of the CMB and LSS datasets. Specifically, we have exploited the synergy between the *Planck* and *Herschel* satellites and measured for the first time the cross-correlation signal between the CMB lensing and the angular positions of the sub-mm galaxies detected in the context of the H-ATLAS survey residing at $z \gtrsim 1.5$ - where the CMB lensing efficiency is at its best - pushing the analysis to a redshift window higher compared to previous studies. Despite the small sky fraction ($f_{\rm sky} \simeq 0.013$) and the high CMB lensing reconstruction noise level, our analysis yielded a detection around $\approx 10\sigma$ (depending

---

[11]Bearing in mind that high-resolution CMB experiments can, and will, say a lot about the small scales temperature anisotropies induced by secondary effects.

on the *Planck* release). This result, validated with a series of null tests and stability checks, confirms that DSFG are excellent tracers of the matter distribution and are found to reside in halos of effective masses around $M_{\mathrm{H}} \simeq 10^{12.9}$ at $z \simeq 2$, supporting our understanding of the structure formation process. Combining the CMB lensing-galaxy density cross-spectrum, $C_\ell^{\kappa g}$, with galaxy clustering measurements, $C_\ell^{gg}$, we have found a cross-correlation amplitude somewhat higher than expected - at the 2.5-3.5$\sigma$ level - for the assumed $\Lambda$CDM cosmology and galaxy redshift distribution. This indicates a possible residual contamination, an incorrect modeling of the signal, or new physics at play. Much effort has been spent also on the astrophysics side for the characterization of the sources in terms of their redshift distribution, a sensitive ingredient of the analysis. As a first step towards a full 3D analysis over the lightcone, we have split the main galaxy samples in two redshift bins and performed a 2D tomographic analysis, reconstructing the galaxy bias evolution over cosmic time.

Motivated by the upcoming experimental efforts, we have also examined the spectral estimation problem in the needlet domain and compared it with the standard harmonic approach, focusing on the masking-induced biasing effects. Even though theoretically equivalent, needlets can benefit from their dual localization in both real and harmonic space and perform better than usual PCL methods. To overcome limitations encountered in the case of heavy masking, as the case of the *Planck-Herschel* cross-correlation, we have developed a novel `MASTER`-like needlet technique.

Finally, we have considered the perspective of theory, by looking at modifications of gravity, but studying manifestations in a different effect, the kSZ. We have focused on a specific class of $f(R)$ models - the Hu & Sawicki one - and identified imprints of the modifications to gravity, through in enhancement of the effect itself with respect to a $\Lambda$CDM case. Nevertheless, we also showed how the same phenomenology could arise fom other effects such as details of reionization, and impact of gastrophysics in modeling of the signal.

**Future prospects**

As mentioned in the introduction, cosmology has entered the *precision era* and the avalanche of data that will be gathered by the community in the upcoming years will propel it to *accuracy era*: this will only be possible by developing suitable statistical techniques and this thesis walks in this direction. At the time of writing this thesis, CMB lensing has been detected at a significance of $\approx 40\sigma$ in the *Planck* data and have become a standard probe in the cosmologist's toolbox, while the cross-correlation science is rapidly moving away from the detection regime. As we

have have learnt, also from our experience, cosmological and astrophysical datasets are now increasingly dominated by systematic, rather that statistical uncertainties.

The work presented here can be extended in different ways. The most puzzling result regarding the cross-correlation analysis is perhaps the observed tension between the cross- and auto-power spectrum, resulting in an amplitude $A > 1$, which should be further investigated. As one of the main limitations of the analysis is the resolution of the CMB lensing maps, a first step would be to exploit low-noise CMB datasets to augment the S/N. Due to the partial overlap with the H-ATLAS survey, the POLARBEAR experiment together the Simons Array, its planned progression, would represent a major step forward in the cross-correlation analysis, enabling a more detailed investigation of the signal.

As we have seen, astrophysical modeling is fundamental in order to interpret the observed signal, meaning that we also need accurate astrophysics if we want correct cosmology measurements. On the galaxies side, the H-ATLAS survey is complete but there is room for improvement on the estimation of the galaxies redshift distribution. In this direction, I started developing a SED template fitting tool based on Bayesian formalism that allows to recover the full photo-$z$ information of the considered galaxy sample.

At the same time, it is essential to complement these findings with appropriate simulations, by exploiting accurate CMB maps lensed with ray-tracing techniques and realistic galaxy mock catalogues based on N-body simulations. The simulative pipeline will help to optimize cross-correlation estimators, to better understand the relations between astrophysical processes and cosmology, and to provide insights on the characterization of future surveys.

Finally, needlet cross-spectrum estimators will be applied to the same set of simulations and above all to real datasets, in preparation of a full-3D analysis.

To conclude, the work presented in this thesis represents an attempt to exploit observations of the CMB secondary anisotropies and LSS to sharpen our understanding of the dark Universe. Much work needs to be done in order to consistently include cross-correlation measurements in the usual cosmological analysis pipelines and to understand how to fully exploit their constraining power.

# Bibliography

Acquaviva, V., & Baccigalupi, C. 2006, *Dark energy records in lensed cosmic microwave background*, Physical Review D, 74, 103510

Adam, R., Bartalucci, I., Pratt, G. W., et al. 2016, *Mapping the kinetic Sunyaev-Zel'dovich effect toward MACS J0717.5+3745 with NIKA*, arXiv:1606.07721

Ade, P. A. R., Akiba, Y., Anthony, A. E., et al. 2014a, *Evidence for Gravitational Lensing of the Cosmic Microwave Background Polarization from Cross-Correlation with the Cosmic Infrared Background*, Physical Review Letters, 112, 131302

—. 2014b, *Measurement of the Cosmic Microwave Background Polarization Lensing Power Spectrum with the POLARBEAR Experiment*, Physical Review Letters, 113, 021301

—. 2014c, *Measurement of the Cosmic Microwave Background Polarization Lensing Power Spectrum with the POLARBEAR Experiment*, Physical Review Letters, 113, 021301

Adler, R. J., & Taylor, J. E. 2007 (New York, NY: Springer New York), 448

Aghanim, N., Majumdar, S., & Silk, J. 2008, *Secondary anisotropies of the CMB*, Reports on Progress in Physics, 71, 066902

Alam, S., Ho, S., & Silvestri, A. 2016, *Testing deviations from $\Lambda CDM$ with growth rate measurements from six large-scale structure surveys at $z = 0.06$-1*, Monthly Notices of the Royal Astronomical Society, 456, 3743

Allison, R., Lindsay, S. N., Sherwin, B. D., et al. 2015, *The Atacama Cosmology Telescope: Measuring radio galaxy bias through cross-correlation with lensing*, Monthly Notices of the Royal Astronomical Society, 451, 849

Antolini, C., Fantaye, Y., Martinelli, M., Carbone, C., & Baccigalupi, C. 2014, *N-body lensed CMB maps: lensing extraction and characterization*, Journal of Cosmology and Astroparticle Physics, 2014, 039

Armendariz-Picon, C., Mukhanov, V., & Steinhardt, P. J. 2000, *Dynamical solution to the problem of a small cosmological constant and late-time cosmic acceleration*, Physical Review Letters, 85, 4438

Aversa, R., Lapi, A., de Zotti, G., Shankar, F., & Danese, L. 2015, *BLACK HOLE AND GALAXY COEVOLUTION FROM CONTINUITY EQUATION AND ABUNDANCE MATCHING*, The Astrophysical Journal, 810, 74

Baldi, P., Kerkyacharian, G., Marinucci, D., & Picard, D. 2009, *Asymptotics for spherical needlets*, The Annals of Statistics, 37, 1150

Bardeen, J. M. 1980, *Gauge-invariant cosmological perturbations*, Physical Review D, 22, 1882

Bardeen, J. M., Bond, J. R., Kaiser, N., & Szalay, a. S. 1986, *The statistics of peaks of Gaussian random fields*, The Astrophysical Journal, 304, 15

Bardeen, J. M., Steinhardt, P. J., & Turner, M. S. 1983, *Spontaneous creation of almost scale-free density perturbations in an inflationary universe*, Physical Review D, 28, 679

Bartelmann, M., & Schneider, P. 2001, *Weak gravitational lensing*, Physics Reports, 340, 291

Bartolo, N., Matarrese, S., & Riotto, A. 2010, *Non-Gaussianity and the Cosmic Microwave Background Anisotropies*, Advances in Astronomy, 2010, 1

Baxter, E. J., Clampitt, J., Giannantonio, T., et al. 2016, *Joint Measurement of Lensing-Galaxy Correlations Using SPT and DES SV Data*, eprint arXiv, 1602, arXiv:1602.07384

Bean, R., Bernat, D., Pogosian, L., Silvestri, A., & Trodden, M. 2007, *Dynamics of linear perturbations in f(R) gravity*, Physical Review D, 75, 064020

Bennett, C. L., Banday, A. J., Górski, K. M., et al. 1996, *Four-Year COBE DMR Cosmic Microwave Background Observations: Maps and Basic Results*, The Astrophysical Journal, 464, L1

Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. 2002, *Large-scale structure of the Universe and cosmological perturbation theory*, Physics Reports, 367, 1

Bernardeau, F., Pitrou, C., & Uzan, J.-P. 2011, *CMB spectra and bispectra calculations: making the flat-sky approximation rigorous*, Journal of Cosmology and Astroparticle Physics, 2011, 015

Bersanelli, M., Mandolesi, N., Butler, R. C., et al. 2010, *Planck pre-launch status: Design and description of the Low Frequency Instrument*, Astronomy and Astrophysics, 520, A4

Béthermin, M., Le Floc'h, E., Ilbert, O., et al. 2012, *HerMES: deep number counts at 250 µ m, 350 µ m and 500 µ m in the COSMOS and GOODS-N fields and the build-up of the cosmic infrared background*, Astronomy & Astrophysics, 542, A58

Bianchini, F., Renzi, A., & Marinucci, D. 2016a, *Needlet estimation of cross-correlation between CMB lensing maps and LSS*, arXiv:1607.05223

Bianchini, F., & Silvestri, A. 2016, *Kinetic Sunyaev-Zel'dovich effect in modified gravity*, Physical Review D, 93, 064026

Bianchini, F., Bielewicz, P., Lapi, A., et al. 2015, *Cross-correlation between the CMB lensing potential measured by Planck and high-z submillimiter galaxies detected by the HERSCHEL-ATLAS survey*, The Astrophysical Journal, 802, 64

Bianchini, F., Lapi, A., Calabrese, M., et al. 2016b, *Toward a tomographic analysis of the cross-correlation between Planck CMB lensing and H-ATLAS galaxies*, The Astrophysical Journal, 825, 24

Bilicki, M., Peacock, J. A., Jarrett, T. H., et al. 2016, *WISE x SuperCOSMOS photometric redshift catalog: 20 million galaxies over 3pi steradians*, arXiv:1607.01182

Blanchard, A., & Schneider, J. 1987, *Gravitational lensing effect on the fluctuations of the cosmic background radiation*, Astronomy and Astrophysics, 184, 1

Bleem, L. E., van Engelen, A., Holder, G. P., et al. 2012, *A measurement of the correlation of galaxy surveys with CMB lensing convergence maps from the South Pole Telescope*, The Astrophysical Journal, 753, L9

Bond, J. R., Jaffe, A. H., & Knox, L. 1998, *Estimating the power spectrum of the cosmic microwave background*, Physical Review D, 57, 2117

Boughn, S. P., Crittenden, R. G., & Koehrsen, G. P. 2002, *The Large Scale Structure of the X-Ray Background and Its Cosmological Implications*, The Astrophysical Journal, 580, 672

Bourne, N., Dunne, L., Maddox, S. J., et al. 2016, *The Herschel -ATLAS Data Release 1 Paper II: Multi-wavelength counterparts to submillimetre sources*, Monthly Notices of the Royal Astronomical Society, stw1654

Brown, M. L., Castro, P. G., & Taylor, A. N. 2005, *Cosmic microwave background temperature and polarization pseudo-Cl estimators and covariances*, Monthly Notices of the Royal Astronomical Society, 360, 1262

Budavari, T., Connolly, A. J., Szalay, A. S., et al. 2003, *Angular Clustering with Photometric Redshifts in the Sloan Digital Sky Survey: Bimodality in the Clustering Properties of Galaxies*, The Astrophysical Journal, 595, 59

Bunn, E. F., Zaldarriaga, M., Tegmark, M., & de Oliveira-Costa, A. 2003, *E/B decomposition of finite pixelized CMB maps*, Physical Review D, 67, 023501

Burigana, C., Danese, L., & de Zotti, G. 1991, *Constraints on the thermal history of the universe from the cosmic microwave background spectrum*, The Astrophysical Journal, 379, 1

Cabre, A., Fosalba, P., Gaztanaga, E., & Manera, M. 2007, *Error analysis in cross-correlation of sky maps: application to the Integrated Sachs-Wolfe detection*, Monthly Notices of the Royal Astronomical Society, 381, 1347

Cai, Z.-Y., Lapi, A., Xia, J.-Q., et al. 2013, *A hybrid model for the evolution of galaxies and active galactic nuclei in the infrared*, The Astrophysical Journal, 768, 21

Calabrese, E., Hložek, R., Battaglia, N., et al. 2014, *Precision epoch of reionization studies with next-generation CMB experiments*, Journal of Cosmology and Astroparticle Physics, 2014, 010

Calabrese, M., Carbone, C., Fabbian, G., Baldi, M., & Baccigalupi, C. 2015, *Multiple lensing of the cosmic microwave background anisotropies*, Journal of Cosmology and Astroparticle Physics, 2015, 049

Casey, C. M., Narayanan, D., & Cooray, A. 2014, *Dusty star-forming galaxies at high redshift*, Physics Reports, 541, 45

Cataneo, M., Rapetti, D., Schmidt, F., et al. 2015, *New constraints on f(R) gravity from clusters of galaxies*, Physical Review D, 92, 044009

Chabrier, G. 2003, *Galactic Stellar and Substellar Initial Mass Function*, Publications of the Astronomical Society of the Pacific, 115, 763

Challinor, A., & Chon, G. 2002, *Geometry of weak lensing of CMB polarization*, Physical Review D, 66, 127301

Challinor, A., & Lewis, A. 2005, *Lensed CMB power spectra from all-sky correlation functions*, Physical Review D, 71, 103010

Chen, X. 2010, *Primordial Non-Gaussianities from Inflation Models*, Advances in Astronomy, 2010, 1

Chiba, T., Okabe, T., & Yamaguchi, M. 2000, *Kinetically driven quintessence*, Physical Review D (Particles, 62, 23511

Chisari, N. E., Dunkley, J., Miller, L., & Allison, R. 2015, *Contamination of early-type galaxy alignments to galaxy lensing-CMB lensing cross-correlation*, Monthly Notices of the Royal Astronomical Society, 453, 682

Chluba, J. 2014, *Science with CMB spectral distortions*, eprint arXiv:1405.6938, arXiv:1405.6938

Chluba, J., & Sunyaev, R. A. 2012, *The evolution of CMB spectral distortions in the early Universe*, Monthly Notices of the Royal Astronomical Society, 419, 1294

Clarkson, C. 2012, *Establishing homogeneity of the universe in the shadow of dark energy*, Comptes Rendus Physique, 13, 682

Clifton, T., Ferreira, P. G., Padilla, A., & Skordis, C. 2012, *Modified gravity and cosmology*

Cole, S., & Efstathiou, G. 1989, *Gravitational lensing of fluctuations in the microwave background radiation*, Monthly Notices of the Royal Astronomical Society, 239, 195

Cooray, A., & Sheth, R. 2002, *Halo models of large scale structure*, Physics Reports, 372, 1

Crittenden, R. G., & Turok, N. 1996, *Looking for a Cosmological Constant with the Rees-Sciama Effect*, Physical Review Letters, 76, 575

Danese, L., & de Zotti, G. 1982, *Double Compton process and the spectrum of the microwave background*, Astronomy and Astrophysics, 107, 39

Das, S., Errard, J., & Spergel, D. 2013, *Can CMB Lensing Help Cosmic Shear Surveys?*, eprint arXiv:1311.2338, 1

Das, S., Sherwin, B. D., Aguirre, P., et al. 2011, *Detection of the power spectrum of cosmic microwave background lensing by the atacama cosmology telescope*, Physical Review Letters, 107, arXiv:1103.2124

De Bernardis, F., Aiola, S., Vavagiakis, E. M., et al. 2016, *Detection of the pairwise kinematic Sunyaev-Zel'dovich effect with BOSS DR11 and the Atacama Cosmology Telescope*, arXiv:1607.02139

De Felice, A., & Tsujikawa, S. 2010, *f(R) Theories*, Living Reviews in Relativity, 13, 156

Delabrouille, J. 2010, *Component separation in CMB observations*, 4

DiPompeo, M. A., Hickox, R. C., & Myers, A. D. 2016, *Updated measurements of the dark matter halo masses of obscured quasars with improved WISE and Planck data*, Monthly Notices of the Royal Astronomical Society, 456, 924

DiPompeo, M. A., Myers, A. D., Hickox, R. C., et al. 2014, *Weighing obscured and unobscured quasar hosts with the cosmic microwave background*, Monthly Notices of the Royal Astronomical Society, 446, 3492

Dodelson, S., & Jubas, J. M. 1995, *Reionization and its imprint of the cosmic microwave background*, The Astrophysical Journal, 439, 503

Donzelli, S., Hansen, F. K., Liguori, M., Marinucci, D., & Matarrese, S. 2012, *On the linear term correction for needlet/wavelet non-Gaussianity estimators*, The Astrophysical Journal, 755, 19

Dossett, J., Hu, B., & Parkinson, D. 2014, *Constraining models of f(R) gravity with Planck and WiggleZ power spectrum data*, Journal of Cosmology and Astroparticle Physics, 2014, 046

Dunkley, J., Hlozek, R., Sievers, J., et al. 2010, *The Atacama Cosmology Telescope: Cosmological Parameters from the 2008 Power Spectra*, The Astrophysical Journal, 739, 20

Dunne, L., Gomez, H. L., da Cunha, E., et al. 2011, *Herschel-ATLAS: rapid evolution of dust in galaxies over the last 5 billion years*, Monthly Notices of the Royal Astronomical Society, 417, 1510

Durastanti, C., Fantaye, Y., Hansen, F., Marinucci, D., & Pesenson, I. Z. 2014, *Simple proposal for radial 3D needlets*, Physical Review D, 90, 103532

Eales, S., Dunne, L., Clements, D., et al. 2010, *The Herschel ATLAS*, Publications of the Astronomical Society of the Pacific, 122, 499

Efstathiou, G. 2004, *Myths and truths concerning estimation of power spectra: the case for a hybrid estimator*, Monthly Notices of the Royal Astronomical Society, 349, 603

Eisenstein, D. J., & Hu, W. 1998, *Baryonic Features in the Matter Transfer Function*, The Astrophysical Journal, 496, 605

Ellis, G. 1975, *Cosmology and verifiability*, Royal Astronomical Society, 16, 245

Etherington, I. 1933, *LX. On the definition of distance in general relativity*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 15, 761

Fantaye, Y., Baccigalupi, C., Leach, S., & Yadav, A. 2012, *CMB lensing reconstruction in the presence of diffuse polarized foregrounds*, Journal of Cosmology and Astroparticle Physics, 2012, 017

Feng, C., Aslanyan, G., Manohar, A. V., et al. 2012, *Measuring gravitational lensing of the cosmic microwave background using cross correlation with large scale structure*, Physical Review D, 86, 063519

Ferraro, S., Hill, J. C., Battaglia, N., Liu, J., & Spergel, D. N. 2016, *The Kinematic Sunyaev-Zel'dovich Effect with Projected Fields II: prospects, challenges, and comparison with simulations*, 14

Ferraro, S., Schmidt, F., & Hu, W. 2011, *Cluster abundance in f(R) gravity models*, Physical Review D, 83, 063503

Fixsen, D. J., Cheng, E. S., Gales, J. M., et al. 1996, *The Cosmic Microwave Background Spectrum from the Full COBE FIRAS Data Set*, The Astrophysical Journal, 473, 576

Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, *emcee : The MCMC Hammer*, Publications of the Astronomical Society of the Pacific, 125, 306

Fornengo, N., Perotto, L., Regis, M., & Camera, S. 2015, *Evidence of Cross-Correlation Between the Cmb Lensing and the Γ-Ray Sky*, The Astrophysical Journal Letters, 802, L1

Fry, J. N., & Gaztanaga, E. 1993, *Biasing and hierarchical statistics in large-scale structure*, The Astrophysical Journal, 413, 447

Gamow, G. 1946, *Expanding Universe and the Origin of Elements*, Physical Review, 70, 572

Geach, J. E., Hickox, R. C., Bleem, L. E., et al. 2013, *A Direct Measurement of the Linear Bias of Mid-infrared-selected Quasars at z approx 1 Using Cosmic Microwave Background Lensing*, The Astrophysical Journal Letters, 776, L41

Geller, D., Hansen, F. K., Marinucci, D., Kerkyacharian, G., & Picard, D. 2008, *Spin needlets for cosmic microwave background polarization data analysis*, Physical Review D, 78, 123533

George, E. M., Reichardt, C. L., Aird, K. A., et al. 2015, *A measurement of secondary cosmic microwave background anisotropies from the 2500 square-degree SPT-SZ survey*, The Astrophysical Journal, 799, 177

Giannantonio, T., Martinelli, M., Silvestri, A., & Melchiorri, A. 2010, *New constraints on parametrised modified gravity from correlations of the CMB with large scale structure*, Journal of Cosmology and Astroparticle Physics, 2010, 030

Giannantonio, T., & Percival, W. J. 2014, *Using correlations between cosmic microwave background lensing and large-scale structure to measure primordial non-Gaussianity*, Monthly Notices of the Royal Astronomical Society: Letters, 441, L16

Giannantonio, T., Scranton, R., Crittenden, R. G., et al. 2008, *Combined analysis of the integrated Sachs-Wolfe effect and cosmological implications*, Physical Review D, 77, 123520

Giannantonio, T., Fosalba, P., Cawthon, R., et al. 2016, *CMB lensing tomography with the DES Science Verification galaxies*, Monthly Notices of the Royal Astronomical Society, 456, 3213

Gnedin, N. Y., & Hui, L. 1998, *Probing the Universe with the Ly forest – I. Hydrodynamics of the low-density intergalactic medium*, Monthly Notices of the Royal Astronomical Society, 296, 44

González-Nuevo, J., Lapi, A., Fleuren, S., et al. 2012, *HERSCHEL -ATLAS: TOWARD A SAMPLE OF \approx 1000 STRONGLY LENSED GALAXIES*, The Astrophysical Journal, 749, 65

Gonzalez-Nuevo, J., Lapi, A., Negrello, M., et al. 2014, *Herschel-ATLAS/GAMA: SDSS cross-correlation induced by weak lensing*, Monthly Notices of the Royal Astronomical Society, 442, 2680

Goodman, J., & Weare, J. 2010, *Ensemble samplers with affine invariance*, Communications in Applied Mathematics and Computational Science, 5, 65

Gorski, K. M., Hivon, E., Banday, A. J., et al. 2005, *HEALPix: A Framework for High Resolution Discretization and Fast Analysis of Data Distributed on the Sphere*, The Astrophysical Journal, 622, 759

Grain, J., Tristram, M., & Stompor, R. 2009, *Polarized CMB power spectrum estimation using the pure pseudo-cross-spectrum approach*, Physical Review D, 79, 123515

Griffin, M. J., Abergel, A., Abreu, A., et al. 2010, *The Herschel -SPIRE instrument and its in-flight performance*, Astronomy and Astrophysics, 518, L3

Gunn, J. E., & Peterson, B. A. 1965, *On the Density of Neutral Hydrogen in Intergalactic Space.*, The Astrophysical Journal, 142, 1633

Guo, Q., Cole, S., Lacey, C. G., et al. 2011, *Which haloes host Herschel-ATLAS galaxies in the local Universe?*, Monthly Notices of the Royal Astronomical Society, 412, 2277

Guth, A. H. 1981, *Inflationary universe: A possible solution to the horizon and flatness problems*, Physical Review D, 23, 347

Guzzo, L., Pierleoni, M., Meneux, B., et al. 2008, *A test of the nature of cosmic acceleration using galaxy redshift distortions*, Nature, 451, 541

Hagstotz, S., Schäfer, B. M., & Merkel, P. M. 2015, *Born-corrections to weak lensing of the cosmic microwave background temperature and polarization anisotropies*, Monthly Notices of the Royal Astronomical Society, 454, 831

Hall, A., Bonvin, C., & Challinor, A. 2013, *Testing general relativity with 21-cm intensity mapping*, Physical Review D, 87, 064026

Hand, N., Addison, G. E., Aubourg, E., et al. 2012, *Evidence of Galaxy Cluster Motions with the Kinematic Sunyaev-Zel'dovich Effect*, Physical Review Letters, 109, 041101

Hand, N., Leauthaud, A., Das, S., et al. 2015, *First measurement of the cross-correlation of CMB lensing and galaxy lensing*, Physical Review D, 91, 062001

Hanson, D., Challinor, A., Efstathiou, G., & Bielewicz, P. 2011, *CMB temperature lensing power reconstruction*, Physical Review D, 83, 043005

Hanson, D., Hoover, S., Crites, A., et al. 2013, *Detection of B-Mode Polarization in the Cosmic Microwave Background with Data from the South Pole Telescope*, Physical Review Letters, 111, 141301

Harnois-Déraps, J., Munshi, D., Valageas, P., et al. 2015, *Testing modified gravity with cosmic shear*, Monthly Notices of the Royal Astronomical Society, 454, 2722

Harnois-Déraps, J., Tröster, T., Hojjati, A., et al. 2016, *CFHTLenS and RCSLenS cross-correlation with Planck lensing detected in fourier and configuration space*, Monthly Notices of the Royal Astronomical Society, 460, 434

Hauser, M. G., & Peebles, P. J. E. 1973, *Statistical Analysis of Catalogs of Extragalactic Objects. 11. the Abell Catalog of Rich Clusters*, The Astrophysical Journal, 185, 757

Hawking, S. W. 1982, *The development of irregularities in a single bubble inflationary universe*, Physics Letters B, 115, 295

Hickox, R. C., Wardlow, J. L., Smail, I., et al. 2012, *The LABOCA survey of the Extended Chandra Deep Field-South: clustering of submillimetre galaxies*, Monthly Notices of the Royal Astronomical Society, 421, no

Hildebrandt, H., van Waerbeke, L., Scott, D., et al. 2013, *Inferring the mass of submillimetre galaxies by exploiting their gravitational magnification of background galaxies*, Monthly Notices of the Royal Astronomical Society, 429, 3230

Hill, J. C., Ferraro, S., Battaglia, N., Liu, J., & Spergel, D. N. 2016, *Kinematic Sunyaev-Zel'dovich Effect with Projected Fields: A Novel Probe of the Baryon Distribution with Planck, WMAP, and WISE Data*, Physical Review Letters, 117, 051301

Hill, J. C., & Spergel, D. N. 2013, *Detection of Thermal SZ – CMB Lensing Cross-Correlation in Planck Nominal Mission Data*, Journal of Cosmology and Astroparticle Physics, 2014, 030

Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, *NINE-YEAR WILKINSON MICROWAVE ANISOTROPY PROBE ( WMAP ) OBSERVATIONS: COSMOLOGICAL PARAMETER RESULTS*, The Astrophysical Journal Supplement Series, 208, 19

Hirata, C. M., Ho, S., Padmanabhan, N., Seljak, U., & Bahcall, N. A. 2008, *Correlation of CMB with large-scale structure. II. Weak lensing*, Physical Review D - Particles, Fields, Gravitation and Cosmology, 78, arXiv:0801.0644

Hirata, C. M., & Seijak, U. 2003, *Analyzing weak lensing of the cosmic microwave background using the likelihood function*, Physical Review D, 67, arXiv:0209489 [astro-ph]

Hirata, C. M., & Seljak, U. 2004, *Intrinsic alignment-lensing interference as a contaminant of cosmic shear*, Physical Review D, 70, 063526

Hivon, E., Gorski, K. M., Netterfield, C. B., et al. 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*, The Astrophysical Journal, 567, 2

Ho, S., Hirata, C., Padmanabhan, N., Seljak, U., & Bahcall, N. 2008, *Correlation of CMB with large-scale structure. I. Integrated Sachs-Wolfe tomography and cosmological implications*, Physical Review D - Particles, Fields, Gravitation and Cosmology, 78, 1

Hojjati, A., Pogosian, L., Silvestri, A., & Talbot, S. 2012, *Practical solutions for perturbed f(R) gravity*, Physical Review D, 86, 123503

Hojjati, A., Pogosian, L., & Zhao, G.-B. 2011, *Testing gravity with CAMB and CosmoMC*, Journal of Cosmology and Astroparticle Physics, 2011, 005

Holder, G. P., Viero, M. P., Zahn, O., et al. 2013, *A Cosmic Microwave Background Lensing Mass Map and Its Correlation With the Cosmic Infrared Background*, The Astrophysical Journal, 771, L16

Hu, W. 2000a, *Reionization Revisited: Secondary Cosmic Microwave Background Anisotropies and Polarization*, The Astrophysical Journal, 529, 12

—. 2000b, *Weak lensing of the CMB: A harmonic approach*, Physical Review D, 62, 043007

Hu, W., Huterer, D., & Smith, K. M. 2006, *Supernovae, the Lensed Cosmic Microwave Background, and Dark Energy*, The Astrophysical Journal, 650, L13

Hu, W., & Okamoto, T. 2002, *Mass Reconstruction with Cosmic Microwave Background Polarization*, The Astrophysical Journal, 574, 566

Hu, W., & Sawicki, I. 2007, *Models of f(R) cosmic acceleration that evade solar system tests*, Physical Review D, 76, 064004

Ibar, E., Ivison, R. J., Cava, A., et al. 2010, *H-ATLAS: PACS imaging for the Science Demonstration Phase*, Monthly Notices of the Royal Astronomical Society, 409, 38

Iliev, I. T., Pen, U.-L., Richard Bond, J., Mellema, G., & Shapiro, P. R. 2006, *kSZ from patchy reionization: The view from the simulations*, New Astronomy Reviews, 50, 909

Ivison, R. J., Swinbank, A. M., Swinyard, B., et al. 2010, *Herschel and SCUBA-2 imaging and spectroscopy of a bright, lensed submillimetre galaxy at z = 2.3*, Astronomy and Astrophysics, 518, L35

Jaffe, A. H., & Kamionkowski, M. 1998, *Calculation of the Ostriker-Vishniac effect in cold dark matter models*, Physical Review D, 58, 043001

Jain, B., Vikram, V., & Sakstein, J. 2013, *Astrophysical tests of modified gravity: constraints from distance indicators in the nearby universe*, The Astrophysical Journal, 779, 39

Jain, B., & Zhang, P. 2008, *Observational tests of modified gravity*, Physical Review D, 78, 063503

Jelic, V., Ciardi, B., Fernandez, E., Tashiro, H., & Vrbanec, D. 2014, *07: SKA - EoR correlations and cross-correlations: kSZ, radio galaxies, and NIR background*, Proceedings of Science, 1

Joyce, A., Lombriser, L., & Schmidt, F. 2016, *Dark Energy vs. Modified Gravity*, 1

Kaiser, N. 1984, *On the spatial correlations of Abell clusters*, The Astrophysical Journal, 284, L9

Kamionkowski, M., Kosowsky, A., & Stebbins, A. 1997, *Statistics of cosmic microwave background polarization*, Physical Review D, 55, 7368

Keisler, R., Reichardt, C. L., Aird, K. A., et al. 2011, *A Measurement of the Damping Tail of the Cosmic Microwave Background Power Spectrum with the South Pole Telescope*, The Astrophysical Journal, 743, 28

Kesden, M., Cooray, A., & Kamionkowski, M. 2003, *Lensing reconstruction with CMB temperature and polarization*, Physical Review D, 67, 123507

Khoury, J., & Weltman, A. 2004, *Chameleon cosmology*, Physical Review D, 69, 044026

Kirk, D., Brown, M. L., Hoekstra, H., et al. 2015, *Galaxy Alignments: Observations and Impact on Cosmology*, Space Science Reviews, 193, 139

Kirk, D., Omori, Y., Benoit-Lévy, A., et al. 2016, *Cross-correlation of gravitational lensing from DES Science Verification data with SPT and Planck lensing*, Monthly Notices of the Royal Astronomical Society, 459, 21

Knox, L. 1995, *Determination of inflationary observables by cosmic microwave background anisotropy experiments*, Physical Review D, 52, 4307

Kodama, H., & Sasaki, M. 1984, *Cosmological Perturbation Theory*, Progress of Theoretical Physics Supplement, 78, 1

Kogut, A., Fixsen, D., Chuss, D., et al. 2011, *The Primordial Inflation Explorer (PIXIE): a nulling polarimeter for cosmic microwave background observations*, Journal of Cosmology and Astroparticle Physics, 2011, 025

Komatsu, E., Wandelt, B. D., Spergel, D. N., Banday, A. J., & Gorski, K. M. 2002, *Measurement of the Cosmic Microwave Background Bispectrum on the COBE DMR Sky Maps*, The Astrophysical Journal, 566, 19

Kosowsky, A. 1996, *Cosmic Microwave Background Polarization*, Annals of Physics, 246, 49

Kuntz, A. 2015, *Cross-correlation of CFHTLenS galaxy catalogue and Planck CMB lensing using the halo model prescription*, Astronomy & Astrophysics, 584, A53

Lamarre, J.-M., Puget, J.-L., Ade, P. A. R., et al. 2010, *Planck pre-launch status: The HFI instrument, from specification to actual performance*, Astronomy and Astrophysics, 520, A9

Lan, X., & Marinucci, D. 2008, *The needlets bispectrum*, Electronic Journal of Statistics, 2, 332

Lanusse, F., Rassat, A., & Starck, J.-L. 2012, *Spherical 3D isotropic wavelets*, Astronomy & Astrophysics, 540, A92

Lapi, A., & Danese, L. 2014, *Statistics of dark matter halos in the excursion set peak framework*, Journal of Cosmology and Astroparticle Physics, 2014, 044

Lapi, A., Raimundo, S., Aversa, R., et al. 2014, *THE COEVOLUTION OF SUPERMASSIVE BLACK HOLES AND MASSIVE GALAXIES AT HIGH REDSHIFT*, The Astrophysical Journal, 782, 69

Lapi, A., González-Nuevo, J., Fan, L., et al. 2011, *Herschel-Atlas Galaxy Counts and High-Redshift Luminosity Functions: the Formation of Massive Early-Type Galaxies*, The Astrophysical Journal, 742, 24

Larsen, P., & Challinor, A. 2015, *Intrinsic alignment contamination to CMB lensing-galaxy weak lensing correlations from tidal torquing*, arXiv:1510.02617

Leistedt, B., & McEwen, J. D. 2012, *Exact Wavelets on the Ball*, IEEE Transactions on Signal Processing, 60, 6257

Leistedt, B., McEwen, J. D., Büttner, M., et al. 2015a, *Analysing the polarisation of the CMB with spin scale-discretised wavelets*, 1

Leistedt, B., McEwen, J. D., Kitching, T. D., & Peiris, H. V. 2015b, *3D weak lensing with spin wavelets on the ball*, Physical Review D, 92, 123010

Lesgourgues, J., Liguori, M., Matarrese, S., & Riotto, A. 2005, *CMB lensing extraction and primordial non-Gaussianity*, Physical Review D, 71, 103514

Lesgourgues, J., & Tram, T. 2011, *The Cosmic Linear Anisotropy Solving System (CLASS) IV: efficient implementation of non-cold relics*, Journal of Cosmology and Astroparticle Physics, 2011, 032

—. 2014, *Fast and accurate CMB computations in non-flat FLRW universes*, Journal of Cosmology and Astroparticle Physics, 2014, 032

Lewis, A., & Challinor, A. 2006, *Weak gravitational lensing of the CMB*

Lewis, A., Challinor, A., & Lasenby, A. 2000, *Efficient Computation of Cosmic Microwave Background Anisotropies in Closed Friedmann-Robertson-Walker Models*, The Astrophysical Journal, 538, 473

Lewis, A., Challinor, A., & Turok, N. 2002, *Analysis of CMB polarization on an incomplete sky*, Physical Review D, 65, arXiv:0108251 [astro-ph]

Li, B., Hellwing, W. A., Koyama, K., et al. 2013, *The non-linear matter and velocity power spectra in f(R) gravity*, Monthly Notices of the Royal Astronomical Society, 428, 743

Lifshitz, E. 1946, *On the gravitational stability of the expanding universe*, J. Phys.(USSR), 10, 116

Limber, D. N. 1953, *The Analysis of Counts of the Extragalactic Nebulae in Terms of a Fluctuating Density Field.*, The Astrophysical Journal, 117, 134

Linde, A. 1982, *A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems*, Physics Letters B, 108, 389

Linder, E. V. 2005, *Cosmic growth history and expansion history*, Physical Review D, 72, 043529

Liu, J., & Hill, J. C. 2015, *Cross-correlation of Planck CMB lensing and CFHTLenS galaxy weak lensing maps*, Physical Review D, 92, 063517

Lombriser, L., Schmidt, F., Baldauf, T., et al. 2012, *Cluster density profiles as a test of modified gravity*, Physical Review D, 85, 102001

LoVerde, M., Hui, L., & Gaztañaga, E. 2007, *Magnification-temperature correlation: The dark side of integrated Sachs-Wolfe measurements*, Physical Review D, 75, 043519

Lue, A., Scoccimarro, R., & Starkman, G. 2004, *Differentiating between modified gravity and dark energy*, Physical Review D, 69, arXiv:0307034 [astro-ph]

Ma, C.-P., & Bertschinger, E. 1995, *Cosmological Perturbation Theory in the Synchronous and Conformal Newtonian Gauges*, The Astrophysical Journal, 455, 7

Ma, C.-P., & Fry, J. N. 2002, *Nonlinear kinetic Sunyaev-Zeldovich effect.*, Physical review letters, 88, 211301

Ma, Y.-Z., & Zhao, G.-B. 2014, *Dark energy imprints on the kinematic Sunyaev-Zel'dovich signal*, Physics Letters B, 735, 402

Maddox, S. J., Dunne, L., Rigby, E., et al. 2010, *Herschel-ATLAS: The angular correlation function of submillimetre galaxies at high and low redshift*, Astronomy and Astrophysics, 518, L11

Marinoni, C., Bel, J., & Buzzi, A. 2012, *The scale of cosmic isotropy*, Journal of Cosmology and Astroparticle Physics, 2012, 036

Marinucci, D., Peccati, G., & Marinucci, Domenico and Peccati, G. 2011, Random fields on the sphere: representation, limit theorems and cosmological applications (Cambridge University Press)

Marinucci, D., Pietrobon, D., Balbi, A., et al. 2008, *Spherical needlets for cosmic microwave background data analysis*, Monthly Notices of the Royal Astronomical Society, 383, 539

Marozzi, G., Fanizza, G., Di Dio, E., & Durrer, R. 2016, *CMB-lensing beyond the Born approximation*, arXiv:1605.08761

Martin, J., Ringeval, C., & Vennin, V. 2014, *Encyclopædia Inflationaris*, Physics of the Dark Universe, 5-6, 75

Mather, J. C., Cheng, E. S., Cottingham, D. A., et al. 1994, *Measurement of the cosmic microwave background spectrum by the COBE FIRAS instrument*, The Astrophysical Journal, 420, 439

McEwen, J. D., Vandergheynst, P., & Wiaux, Y. 2013, in SPIE, ed. D. Van De Ville, V. K. Goyal, & M. Papadakis, 88580I

McKellar, A. 1940, *Evidence for the Molecular Origin of Some Hitherto Unidentified Interstellar Lines*, Publications of the Astronomical Society of the Pacific, 52, 187

McQuinn, M., Furlanetto, S. R., Hernquist, L., Zahn, O., & Zaldarriaga, M. 2005, *The Kinetic Sunyaev-Zel'dovich Effect from Reionization*, The Astrophysical Journal, 630, 643

Mo, H., van den Bosch, F., & White, S. 2010, Galaxy Formation and Evolution (Cambridge: Cambridge University Press), 1

Mueller, E.-M., de Bernardis, F., Bean, R., & Niemack, M. D. 2015, *Constraints on gravity and dark energy from the pairwise kinematic Sunyaev-Zel'dovich effect*, The Astrophysical Journal, 808, 47

Mukhanov, V. 2005, Physical Foundations of Cosmology, Vol. -1 (Cambridge University Press), 442

Mukhanov, V. F., Feldman, H. A., & Brandenberger, R. H. 1992, *Theory of cosmological perturbations*

Namikawa, T., Hanson, D., & Takahashi, R. 2013, *Bias-hardened CMB lensing*, Monthly Notices of the Royal Astronomical Society, 431, 609

Narcowich, F. J., Petrushev, P., & Ward, J. D. 2006, *Localized Tight Frames on Spheres*, SIAM Journal on Mathematical Analysis, 38, 574

Negrello, M., Hopwood, R., De Zotti, G., et al. 2010, *The Detection of a Population of Submillimeter-Bright, Strongly Lensed Galaxies*, Science, 330, 800

Neyman, J., & Scott, E. L. 1952, *A Theory of the Spatial Distribution of Galaxies.*, The Astrophysical Journal, 116, 144

Okamoto, T., & Hu, W. 2003, *Cosmic microwave background lensing reconstruction on the full sky*, Physical Review D, 67, 083002

Omori, Y., & Holder, G. 2015, *Cross-Correlation of CFHTLenS Galaxy Number Density and Planck CMB Lensing*, eprint arXiv, 1502, 3405

Osborne, S. J., Hanson, D., & Doré, O. 2014, *Extragalactic foreground contamination in temperature-based CMB lens reconstruction*, Journal of Cosmology and Astroparticle Physics, 2014, 024

Ostriker, J. P., & Vishniac, E. T. 1986, *Generation of microwave background fluctuations from nonlinear perturbations at the ERA of galaxy formation*, The Astrophysical Journal, 306, L51

Oyaizu, H., Lima, M., & Hu, W. 2008, *Nonlinear evolution of f(R) cosmologies. II. Power spectrum*, Physical Review D, 78, 123524

Park, H., Komatsu, E., Shapiro, P. R., Koda, J., & Mao, Y. 2016, *THE IMPACT OF NONLINEAR STRUCTURE FORMATION ON THE POWER SPECTRUM OF TRANSVERSE MOMENTUM FLUCTUATIONS AND THE KINETIC SUNYAEV-ZEL'DOVICH EFFECT*, The Astrophysical Journal, 818, 37

Pascale, E., Auld, R., Dariush, A., et al. 2011, *The first release of data from the Herschel ATLAS: the SPIRE images*, Monthly Notices of the Royal Astronomical Society, 415, 911

Pearson, E. A., Eales, S., Dunne, L., et al. 2013, *H-ATLAS: estimating redshifts of Herschel sources from sub-mm fluxes*, Monthly Notices of the Royal Astronomical Society, 435, 2753

Pearson, R., & Zahn, O. 2014, *Cosmology from cross correlation of CMB lensing and galaxy surveys*, Physical Review D, 89, 043516

Peebles, P. J. E. 2002, *From Precision Cosmology to Accurate Cosmology*, arXiv preprint astro-ph/0208037, arXiv:0208037 [astro-ph]

Penzias, A. A., & Wilson, R. W. 1965, *A Measurement of Excess Antenna Temperature at 4080 Mc/s.*, The Astrophysical Journal, 142, 419

Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, *Measurements of Omega and Lambda from 42 High-Redshift Supernovae*, The Astronomical Journal, 517, 565

Perotto, L., Bobin, J., Plaszczynski, S., Starck, J.-L., & Lavabre, A. 2010, *Reconstruction of the cosmic microwave background lensing for Planck*, Astronomy and Astrophysics, 519, A4

Pettinari, G. W. 2016, The Intrinsic Bispectrum of the Cosmic Microwave Background, Springer Theses (Cham: Springer International Publishing), 1

Pietrobon, D., Balbi, A., & Marinucci, D. 2006, *Integrated Sachs-Wolfe effect from the cross correlation of WMAP 3 year and the NRAO VLA sky survey data: New results and constraints on dark energy*, Physical Review D, 74, 043524

Pietrobon, D., Cabella, P., Balbi, A., et al. 2010, *Needlet bispectrum asymmetries in the WMAP 5-year data*, Monthly Notices of the Royal Astronomical Society: Letters, 402, L34

Pilbratt, G. L., Riedinger, J. R., Passvogel, T., et al. 2010, *Herschel Space Observatory*, Astronomy and Astrophysics, 518, L1

Planck Collaboration. 2014a, *Planck 2013 results. XII. Diffuse component separation*, Astronomy & Astrophysics, 571, A12

—. 2014b, *Planck 2013 results. XVI. Cosmological parameters*, Astronomy & Astrophysics, 571, A16

—. 2014c, *Planck 2013 results. XVII. Gravitational lensing by large-scale structure*, Astronomy & Astrophysics, 571, A17

—. 2014d, *Planck 2013 results. XVIII. The gravitational lensing-infrared background correlation*, Astronomy & Astrophysics, 571, A18

—. 2014e, *Planck 2013 results. XXX. Cosmic infrared background measurements and implications for star formation*, Astronomy & Astrophysics, 571, A30

—. 2015a, *Planck 2015 results. IX. Diffuse component separation: CMB maps*, Astronomy & Astrophysics, arXiv:1502.05956 [astro-ph.CO]

—. 2015b, *Planck 2015 results. X. Diffuse component separation: Foreground maps*, ArXiv e-prints, 1

—. 2015c, *Planck 2015 results. XIII. Cosmological parameters*, arXiv:1502.01589

—. 2015d, *Planck 2015 results. XIV. Dark energy and modified gravity*, 33

—. 2015e, *Planck 2015 results. XXVII. The Second Planck Catalogue of Sunyaev-Zeldovich Sources*, arXiv:1502.01598

—. 2016a, *Planck 2015 results. XV. Gravitational lensing*, Astronomy & Astrophysics, 10

—. 2016b, *Planck 2015 results. XVII. Constraints on primordial non-Gaussianity*, Astronomy & Astrophysics, 61

—. 2016c, *Planck intermediate results*, Astronomy & Astrophysics, 586, A140

—. 2016d, *Planck intermediate results*, Astronomy & Astrophysics, 586, A133

Planck Collaboration, Adam, R., Ade, P. A. R., et al. 2015, *Planck 2015 results. IX. Diffuse component separation: CMB maps*, astro-ph.CO, 42

Poglitsch, A., Waelkens, C., Geis, N., et al. 2010, *The Photodetector Array Camera and Spectrometer (PACS) on the Herschel Space Observatory*, Astronomy and Astrophysics, 518, L2

Pogosian, L., & Silvestri, A. 2008, *Pattern of growth in viable f(R) cosmologies*, Physical Review D - Particles, Fields, Gravitation and Cosmology, 77, arXiv:0709.0296

Pratten, G., & Lewis, A. 2016, *Impact of post-Born lensing on the CMB*, 25

Pullen, A. R., Alam, S., He, S., & Ho, S. 2016, *Constraining gravity at the largest scales through CMB lensing and galaxy velocities*, Monthly Notices of the Royal Astronomical Society, 460, 4098

Ratra, B., & Peebles, P. J. E. 1988, *Cosmological consequences of a rolling homogeneous scalar field*, Physical Review D, 37, 3406

Raveri, M., Hu, B., Frusciante, N., & Silvestri, A. 2014, *Effective field theory of cosmic acceleration: Constraining dark energy with CMB data*, Physical Review D, 90, 043513

Rees, M. J., & Sciama, D. W. 1968, *Large-scale Density Inhomogeneities in the Universe*, Nature, 217, 511

Regan, D., Gosenca, M., & Seery, D. 2015, *Constraining the WMAP9 bispectrum and trispectrum with needlets*, Journal of Cosmology and Astroparticle Physics, 2015, 013

Reichardt, C. L., Ade, P. A. R., Bock, J. J., et al. 2009, *High-Resolution CMB Power Spectrum from the Complete ACBAR Data Set*, The Astrophysical Journal, 694, 1200

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, *Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant*, The Astronomical Journal, 116, 1009

Rigby, E. E., Maddox, S. J., Dunne, L., et al. 2011, *Herschel-ATLAS: first data release of the Science Demonstration Phase source catalogues*, Monthly Notices of the Royal Astronomical Society, 415, 2336

Rodighiero, G., Renzini, A., Daddi, E., et al. 2014, *A multiwavelength consensus on the main sequence of star-forming galaxies at z 2*, Monthly Notices of the Royal Astronomical Society, 443, 19

Rogers, K. K., Peiris, H. V., Leistedt, B., McEwen, J. D., & Pontzen, A. 2016a, *SILC: a new Planck internal linear combination CMB temperature map using directional wavelets*, Monthly Notices of the Royal Astronomical Society, 460, 3014

—. 2016b, *Spin-SILC: CMB polarisation component separation with spin wavelets*, 13

Rudjord, Ø., Hansen, F. K., Lan, X., et al. 2009, *An Estimate of the Primordial Non-Gaussianity Parameter fNL Using the Needlet Bispectrum from WMAP*, The Astrophysical Journal, 701, 369

Ryden, B. S. 2003, Introduction to cosmology (San Francisco: Addison-Wesley)

Schaan, E., Krause, E., Eifler, T., et al. 2016a, *Looking through the same lens: shear calibration for LSST, Euclid & WFIRST with stage 4 CMB lensing*, arXiv:1607.01761

Schaan, E., Ferraro, S., Vargas-Magaña, M., et al. 2016b, *Evidence for the kinematic Sunyaev-Zel'dovich effect with the Atacama Cosmology Telescope and velocity reconstruction from the Baryon Oscillation Spectroscopic Survey*, Physical Review D, 93, 082002

Scherrer, R. J., & Bertschinger, E. 1991, *Statistics of primordial density perturbations from discrete seed masses*, The Astrophysical Journal, 381, 349

Schmidt, S. J., Menard, B., Scranton, R., et al. 2014, *Inferring the redshift distribution of the cosmic infrared background*, Monthly Notices of the Royal Astronomical Society, 446, 2696

Scranton, R., Menard, B., Richards, G. T., et al. 2005, *Detection of Cosmic Magnification with the Sloan Digital Sky Survey*, The Astrophysical Journal, 633, 589

Seljak, U. 2000, *Analytic model for galaxy and dark matter clustering*, Monthly Notices of the Royal Astronomical Society, 318, 203

Seljak, U., & Zaldarriaga, M. 1996, *A Line-of-Sight Integration Approach to Cosmic Microwave Background Anisotropies*, The Astrophysical Journal, 469, 437

Serra, P., Lagache, G., Doré, O., Pullen, A., & White, M. 2014, *Cross-correlation of cosmic far-infrared background anisotropies with large scale structures*, Astronomy & Astrophysics, 570, A98

Shaw, L. D., Rudd, D. H., & Nagai, D. 2012, *Deconstructing the kinetic SZ power spectrum*, The Astrophysical Journal, 756, 15

Sherwin, B. D., Dunkley, J., Das, S., et al. 2011, *Evidence for dark energy from the cosmic microwave background alone using the atacama cosmology telescope lensing measurements*, Physical Review Letters, 107, arXiv:1105.0419

Sherwin, B. D., Das, S., Hajian, A., et al. 2012, *The Atacama cosmology telescope: Cross-correlation of cosmic microwave background lensing and quasars*, Physical Review D - Particles, Fields, Gravitation and Cosmology, 86, 1

Sheth, R. K., Mo, H. J., & Tormen, G. 2001, *Ellipsoidal collapse and an improved model for the number and spatial distribution of dark matter haloes*, Monthly Notices of the Royal Astronomical Society, 323, 1

Sheth, R. K., & Tormen, G. 1999, *Large-scale bias and the peak background split*, Monthly Notices of the Royal Astronomical Society, 308, 119

Shirasaki, M. 2016, Probing Cosmic Dark Matter and Dark Energy with Weak Gravitational Lensing Statistics, Springer Theses (Singapore: Springer Singapore)

Silvestri, A., & Trodden, M. 2009, *Approaches to understanding cosmic acceleration*, Reports on Progress in Physics, 72, 96901

Simon, P. 2007, *How accurate is Limber's equation?*, Astronomy and Astrophysics, 473, 711

Simpson, F., Heymans, C., Parkinson, D., et al. 2013, *CFHTLenS: testing the laws of gravity with tomographic weak lensing and redshift-space distortions*, Monthly Notices of the Royal Astronomical Society, 429, 2249

Singh, S., Mandelbaum, R., & Brownstein, J. R. 2016, *Cross-correlating Planck CMB lensing with SDSS: Lensing-lensing and galaxy-lensing cross-correlations*, eprint arXiv:1606.08841, arXiv:1606.08841

Smith, D. J. B., Dunne, L., Maddox, S. J., et al. 2011, *Herschel-ATLAS: counterparts from the ultraviolet-near-infrared in the science demonstration phase catalogue*, Monthly Notices of the Royal Astronomical Society, 416, 857

Smith, K. M., & Ferraro, S. 2016, *Detecting patchy reionization in the CMB*, arXiv:1607.01769

Smith, K. M., Zahn, O., & Doré, O. 2007, *Detection of gravitational lensing in the cosmic microwave background*, Physical Review D, 76, 043510

Smith, R. E., Peacock, J. A., Jenkins, A., et al. 2003, *Stable clustering, the halo model and non-linear cosmological power spectra*, Monthly Notices of the Royal Astronomical Society, 341, 1311

Soergel, B., Flender, S., Story, K. T., et al. 2016, *Detection of the kinematic Sunyaev-Zel'dovich effect with DES Year 1 and SPT*, Monthly Notices of the Royal Astronomical Society, 461, 3172

Song, Y.-S., Cooray, A., Knox, L., & Zaldarriaga, M. 2003, *The Far-Infrared Background Correlation with Cosmic Microwave Background Lensing*, The Astrophysical Journal, 590, 664

Song, Y.-S., Hu, W., & Sawicki, I. 2007a, *Large scale structure of f(R) gravity*, Physical Review D, 75, 044004

Song, Y.-S., Peiris, H., & Hu, W. 2007b, *Cosmological constraints on f (R) acceleration models*, Physical Review D, 76, 063517

Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, *A HIGHLY CONSISTENT FRAMEWORK FOR THE EVOLUTION OF THE STAR-FORMING "MAIN SEQUENCE" FROM z 0-6*, The Astrophysical Journal Supplement Series, 214, 15

Starobinsky, A. 1982, *Dynamics of phase transition in the new inflationary universe scenario and generation of perturbations*, Physics Letters B, 117, 175

Stompor, R. 2011, *S2HAT: Scalable Spherical Harmonic Transform Library*, Astrophysics Source Code Library, record ascl:1110.013

Stompor, R., & Efstathiou, G. 1999, *Gravitational lensing of cosmic microwave background anisotropies and cosmological parameter estimation*, Monthly Notices of the Royal Astronomical Society, 302, 735

Story, K. T., Hanson, D., Ade, P. A. R., et al. 2015, *A MEASUREMENT OF THE COSMIC MICROWAVE BACKGROUND GRAVITATIONAL LENSING POTENTIAL FROM 100 SQUARE DEGREES OF SPTPOL DATA*, The Astrophysical Journal, 810, 50

Sunyaev, R. A., & Zeldovich, Y. B. 1970, *The interaction of matter and radiation in the hot model of the Universe, II*, Astrophysics and Space Science, Volume 7, Issue 1, pp.20-30, 7, 20

—. 1980, *The velocity of clusters of galaxies relative to the microwave background. The possibility of its measurement*, Monthly Notices of the Royal Astronomical Society, 190, 413

Swinbank, A. M., Smail, I., Longmore, S., et al. 2010, *Intense star formation within resolved compact regions in a galaxy at z = 2.3*, Nature, 464, 733

Takahashi, R., Sato, M., Nishimichi, T., Taruya, A., & Oguri, M. 2012, *REVISING THE HALOFIT MODEL FOR THE NONLINEAR MATTER POWER SPECTRUM*, The Astrophysical Journal, 761, 152

Taylor, A., Joachimi, B., & Kitching, T. 2013, *Putting the precision in precision cosmology: How accurate should your data covariance matrix be?*, Monthly Notices of the Royal Astronomical Society, 432, 1928

Tegmark, M. 1997, *How to measure CMB power spectra without losing information*, Physical Review D, 55, 5895

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*, Monthly Notices of the Royal Astronomical Society, 358, 833

Troja, A., Donzelli, S., Maino, D., & Marinucci, D. 2014, *The Needlet CMB Trispectrum*, Proceedings of the International Astronomical Union, 10, 48

Troxel, M. A., & Ishak, M. 2014, *Cross-correlation between cosmic microwave background lensing and galaxy intrinsic alignment as a contaminant to gravitational lensing cross-correlated probes of the Universe*, Physical Review D, 89, 063528

Valiante, E., Smith, M., Eales, S., et al. 2016, *The Herschel -ATLAS Data Release 1 Paper I: Maps, Catalogues and Number Counts*, Monthly Notices of the Royal Astronomical Society, stw1806

Vallinotto, A. 2012, *USING COSMIC MICROWAVE BACKGROUND LENSING TO CONSTRAIN THE MULTIPLICATIVE BIAS OF COSMIC SHEAR*, The Astrophysical Journal, 759, 32

van Engelen, A., Bhattacharya, S., Sehgal, N., et al. 2014, *CMB LENSING POWER SPECTRUM BIASES FROM GALAXIES AND CLUSTERS USING HIGH-ANGULAR RESOLUTION TEMPERATURE MAPS*, The Astrophysical Journal, 786, 13

van Engelen, A., Sherwin, B. D., Sehgal, N., et al. 2015, *THE ATACAMA COSMOLOGY TELESCOPE: LENSING OF CMB TEMPERATURE AND POLARIZATION DERIVED FROM COSMIC INFRARED BACKGROUND CROSS-CORRELATION*, The Astrophysical Journal, 808, 7

Van Waerbeke, L., Hinshaw, G., & Murray, N. 2014, *Detection of warm and diffuse baryons in large scale structure from the cross correlation of gravitational lensing and the thermal Sunyaev-Zeldovich effect*, Physical Review D, 89, 023508

Varshalovich, D. A., Moskalev, A. N., & Khersonskii, V. K. 1988, Quantum Theory of Angular Momentum (WORLD SCIENTIFIC)

Viero, M. P., Wang, L., Zemcov, M., et al. 2013, *HerMES: COSMIC INFRARED BACKGROUND ANISOTROPIES AND THE CLUSTERING OF DUSTY STAR-FORMING GALAXIES*, The Astrophysical Journal, 772, 77

Vikram, V., Cabré, A., Jain, B., & VanderPlas, J. 2013, *Astrophysical tests of modified gravity: the morphology and kinematics of dwarf galaxies*, Journal of Cosmology and Astroparticle Physics, 2013, 020

Vishniac, E. T. 1987, *Reionization and small-scale fluctuations in the microwave background*, The Astrophysical Journal, 322, 597

Wandelt, B. D., Hivon, E., & Górski, K. M. 2001, *Cosmic microwave background anisotropy power spectrum statistics for high precision cosmology*, Physical Review D, 64, 083003

Weinberg, S. 2008, Cosmology (Oxford University Press), 593

Wetterich, C. 1988, *Cosmology and the fate of dilatation symmetry*, Nuclear Physics, Section B, 302, 668

Wilcox, H., Bacon, D., Nichol, R. C., et al. 2015, *The XMM Cluster Survey: testing chameleon gravity using the profiles of clusters*, Monthly Notices of the Royal Astronomical Society, 452, 1171

Xia, J. Q., Negrello, M., Lapi, A., et al. 2012, *Clustering of submillimetre galaxies in a self-regulated baryon collapse model*, Monthly Notices of the Royal Astronomical Society, 422, 1324

Xia, J.-Q., Viel, M., Baccigalupi, C., & Matarrese, S. 2009, *The high redshift Integrated Sachs-Wolfe effect*, Journal of Cosmology and Astroparticle Physics, 2009, 003

Yamamoto, K., Nakamura, G., Hütsi, G., Narikawa, T., & Sato, T. 2010, *Constraint on the cosmological model from the multipole power spectrum of the SDSS luminous red galaxy sample and prospects for a future redshift*, Physical Review D, 81, 103517

Yu, J. T., & Peebles, P. J. E. 1969, *Superclusters of Galaxies?*, The Astrophysical Journal, 158, 103

Zahn, O., Zaldarriaga, M., Hernquist, L., & McQuinn, M. 2005, *The Influence of Nonuniform Reionization on the CMB*, The Astrophysical Journal, 630, 657

Zaldarriaga, M. 1997, *All-sky analysis of polarization in the microwave background*, Physical Review D, 55, 1830

Zeldovich, Y. B., & Sunyaev, R. A. 1969, *The interaction of matter and radiation in a hot-model universe*, Astrophysics and Space Science, 4, 301

Zhang, P., Pen, U.-L., & Trac, H. 2004, *Precision era of the kinetic Sunyaev-Zel'dovich effect: simulations, analytical models and observations and the power to constrain reionization*, Monthly Notices of the Royal Astronomical Society, 347, 1224

Zhao, G.-B. 2014, *MODELING THE NONLINEAR CLUSTERING IN MODIFIED GRAVITY MODELS. I. A FITTING FORMULA FOR THE MATTER POWER SPECTRUM OF f(R) GRAVITY*, The Astrophysical Journal Supplement Series, 211, 23