Cross-correlations in the Dark Energy Survey: from redshift distribution inference to probes of gravity with the Cosmic Microwave Background

Pauline Eva Vielzeuf

Cross-correlations in the Dark Energy Survey: from redshift distribution inference to probes of gravity with the Cosmic Microwave Background

Pauline Eva Vielzeuf

Doctorado en Fisica

Departament de Fisica

Universitat Autònoma de Barcelona

A thesis submitted for the degree of
Philosophae Doctor (PhD)

Day of defense: October 26th 2018

Director:
Dr. Ramon Miquel Pascual
IFAE ICREA
Edifici Cn, UAB
08193 Bellaterra (Barcelona), Spain
ramon.miquel@ifae.es

Director:
Dr. András Kovács
IFAE
Edifici Cn, UAB
08193 Bellaterra (Barcelona), Spain
akovacs@ifae.es

Declaration of Authorship

I, Pauline Eva Vielzeuf, declare that this thesis titled, 'Cross-correlation in the Dark Energy Survey first year of observation : from redshift distribution inference to cosmological test with the Cosmological Microwave Background' and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Abstract

During the past decades, we have been the witnesses of unprecedented advancements in the understanding of our universe dynamics and evolution. Indeed, together with the theory General Relativity (GR) by Albert Einstein at the beginning of the 1900's, the improvement of observational tools have led to the establishment of the standard model of cosmology the so-called Λ-cold-Dark-Matter model (ΛCDM) that is so far the simplest model that describes best our universe considering observations. However, this standard model suffers from caveats, more specifically the presence in the theory of two dark components, Dark Energy and Dark Matter. Understanding the mystery behind these two components have become the leading objective of observational cosmology today, and of current experiments, as it is the case of the Dark Energy Survey (DES), that will after 5 years of observations image about 300 million galaxies with an unprecedented depth covering one eighth of the sky (5000 sq. deg.).

This thesis is based on the analysis of the first year of observation of DES (DESY1) and more in particular on the use of cross-correlation techniques in cosmological analyses. Here, we expose two of different possible use of cross-correlation. Namely, we first show how cross-correlation techniques has been employed as a tool to infer redshift distribution of objects using the so-called clustering-redshift methodologies. And moreover how for the first time this techniques have entered in the cosmological analyses to correct the mean of the redshift distributions of the DESY1 galaxy catalog with shapes (the weak lensing sample) inferred by typical photometric techniques. We explain how using simulation we have evaluated systematics errors induced by our overall methodology and present the full methodology employed for redshift distribution determination of the DESY1 Weak lensing sample.

In a second time, we expose how we have been using cross-correlations as a probe of ΛCDM cosmology using in one hand cross-correlation between cosmic voids identified in the DESY1 catalogs and the lensing map from the Planck satellite and on the other hand simulated ΛCDM cosmology. More specifically, recent results have suggested an excess signal in the observed void catalogs imprint in the Cosmic Microwave Background temperature maps with respect to simulated ΛCDM cosmology, we reiterate the procedure using this time the lensing maps of the CMB. After optimizing our void catalogs in simulation, looking for population of voids responsible for the more significant detection,we could detect the imprint of DESY1 cosmic voids with a significance of $\sim 7 - 12\sigma$ with a discrepancy of $\sim 3 - 9\sigma$ with respect to ΛCDM simulations.

Pròleg

Durant les últimes dècades, hem estat testimonis d'avenços sense precedents en la comprensió de la dinàmica i evolució de l'univers. Juntament amb la teoria de la Relativitat General (GR) d'Albert Einstein a principis de la dècada de 1900, la millora de les eines d'observació ha conduït a l'establiment del model estàndard de cosmologia anomenat Λ-cold-Dark- Model de matèria (ΛCDM), que és fins ara el model més senzill que descriu millor el nostre univers tenint en compte les observacions. No obstant això, aquest model estàndard pateix advertències, més concretament la presència en la teoria de dos components foscos, Dark Energy and Dark Matter. Entendre el misteri darrere d'aquests dues components s'ha convertit en el principal objectiu de la cosmologia observacional actual, com és el cas del projecte Dark Energy Survey (DES), que després de 5 anys d'observacions tindrà imatges de 300 milions de galàxies amb un profunditat sense precedents, cobrint una vuitena del cel (5000 sq. deg.).

Aquesta tesi es basa en l'anàlisi del primer any d'observació de DES (DESY1) i més en particular en l'ús de tècniques de correlació creuada en anàlisis cosmològiques. S'exposen dos possibles usos de correlació creuada. En primer lloc mostrem com s'han utilitzat les tècniques de correlació creuada com una eina per inferir la distribució de redshift d'objectes mitjançant les anomenades metodologies clustering-redshift. A més, per primera vegada, s'han introduït aquestes tècniques en les anàlisis cosmològiques per corregir la mitjana de les distribucions de redshift del catàleg de galàxies DESY1 amb formes (la mostra de weak lensing) inferida per tècniques fotomètriques típiques. Expliquem com usant la simulació hem avaluat els errors sistemàtics induïts per la nostra metodologia general i presentem la metodologia completa emprada per a la determinació de la distribució del redshift de la mostra de weak lensing DESY1.

En segon lloc, exposem com hem estat utilitzant correlacions creuades per sondejar el model ΛCDM utilitzant per una banda la correlació creuada entre els buits còsmics identificats en els catàlegs DESY1 i el mapa de lents del satèl·lit Planck i d'altra banda cosmologiaλCDM simulada . Més específicament, resultats recents suggereixen un excés de senyal en els catàlegs de buits observats en els mapes de temperatura del fons de microones còsmiques respecte a la cosmologia $LambdaCDM$ simulada, de manera que reiterem el procediment utilitzant aquesta vegada els mapes de lents del CMB. Després d'optimitzar els nostres catàlegs de buits utilitzant la simulació buscant la població de buits responsables de la detecció més significativa, hem pogut detectar l'empremta dels buits còsics DESY1 amb una importància de $\sim 7-12\sigma$ amb una discrepància de $\sim 3-9\sigma$ pel que fa a simulacions ΛCDM.

Contents

List of Figures

List of Tables

Pour pépé Denis

Introduction

Cosmology is define as the science that aims to understand the history and dynamics of our universe at large scales (scales larger than few Mpc). In the past decades, with the formulation of General Relativity theory by Einstein (Einstein (1916)) and the development later on of big experiments focused in observing the largest scales of our universe, our understanding of our universe's dynamic and origin has improved. Indeed, experiments such as Redshift Surveys (see **??**York et al. (2000); The Dark Energy Survey Collaboration (2005); de Jong et al. (2015)) or Cosmic Microwave Background (CMB) experiments (see Boggess et al. (1992); Bennett et al. (2003); Ruhl et al. (2004); Tauber et al. (2010); Fowler et al. (2007)) have seemed to agree in one concordance theory the so-called Λ-*Cold-Dark-Matter* scenario that describes our universe as an universe that started with an explosion (the Big Bang) and is mostly made of two dark components **Dark Matter** and **Dark Energy** whose later is responsible by the observed cosmic acceleration and model by a cosmological constant Λ (Huterer & Turner (1999)). However, even though the standard model of cosmology appears to describe the best our universe, unsatisfactory aspects of this theory remain. Namely, the presence in this model of the two previously mentioned Dark entities in our universe so far not directly observed and that dominates the energy budget of the universe, the main discovery that happened in the last century and lead the community to construct this concordance model will be present in Chapter 1.

The physics behind these two components remains mysterious and the objective of cosmological experiments is to solve this enigma. For this purpose, cosmological efforts are observing to different probes, and one popular cosmological target are the large scales objects, such as galaxies, galaxy clusters, cosmic voids and filaments such are doing the Redshift Surveys that map the large structures of our universe with an increasing accuracy and field of view, various observing procedures, based on specific requirements for the probes they have been dedicated to constrain. The Dark Energy Surveys (DES) has been build in this perspective and will provide catalog of large structures up to redshift $z = 1.2$. DES is currently proceeding it last observations and the observed catalogs from its first year of observation as well as the cosmological analysis made by the collaboration came out recently (see Drlica-Wagner et al. (2018); Elvin-Poole et al. (2017); Troxel et al. (2017); Prat et al. (2017); DES Collaboration et al. (2017)). After only one year over five of observations, DES has shown to already be competitive to infer cosmological

parameters from its galaxy catalogs. In Chapter 2 we will present in details the Dark Energy Survey as well as the main cosmological results obtained with the datasets coming from the first year of observation.

This thesis aims to present the work accomplished in the context of the Dark Energy Survey first year of observation data analysis (DESY1), more in particular developing two different aspects that exists in analyzing galaxy catalogs. The first one, is based on the fact that a challenging aspect in fulfilling the requirements in term of precision in cosmological inference for DES (or other photometric surveys) resides on the accurate estimation of redshift distribution of galaxy catalogs. The bias induced by regular photometric redshift methods could become in the forthcoming cosmological analyses with these type of surveys the dominant term in the error budget and therefore one of the greatest challenge for future surveys such as Euclid (Laureijs et al. (2011)) or the Large Synoptic Survey Telescope (LSST, LSST Science Collaboration et al. (2009)). However, alternative method to photometric redshift have been proposed and in this work we aimed to study in details one of these alternative, the so-called clustering redshift methods (see for example Newman (2008); Ménard et al. (2013); Schmidt et al. (2013); Matthews & Newman (2010)). This method is a relatively new method that allows to infer the redshift distribution of a given sample using the cross-correlation signal it has with a reference sample for which one has redshift information. In chapter 3, we will present the procedure followed in the context of DESY1 to implement clustering-redshift methodology in the cosmological analysis. This methods has enter for the first time in the whole cosmological analysis and this required a detailed study of the systematic budget that the method induced. There we will present in particular the methodology employed and how we caracterized using simulated galaxy catalogs the error we are subject to applying our method as well as the final results obtain for the reconstruction of the redshift distribution of the galaxy catalog with galaxy shape information used in the main DESY1 cosmological analysis (see Hoyle et al. (2018); Gatti, Vielzeuf. et al. (2018); Davis et al. (2017)).

The second main work developed here concern cosmic voids that one can identify in the DES data sample and they imprint that could be detected in the lensing signal from the CMB. Indeed, as it has been mentioned, different cosmological probe exists, and the one with so far the most important constraining power comes from the radiation we observed remnant of the Big Bang explosion. Although this radiation contains usefull information from the early universe, by travelling towards us, the photons from the CMB, similarly to what happen to background galaxy in cosmic shear analyses, will also suffer from deflection effect due to the large structures it crosses, such as cosmic voids. The Planck collaboration have reconstructed a map of the distortions of the CMB photon Planck Collaboration et al. (2016c), and the interest of this work has been to develop an optimized methodology to detect the imprint of cosmic voids identified in the DES catalogs in this map. Namely, using a stacking procedure, we have been comparing the stacked lensing signal at the position of cosmic voids in both ΛCDM simulation and DESY1

void catalogs. This work is particularly insteresting in the sense that previous analysis, have suggest discrepencies between ΛCDM predictions and observations (see Granett et al. (2008); Cai et al. (2014a); Kovács (2018)) applying this void stacking procedure to the CMB temperature anisotropy maps, measuring the so-called integrated Sachs-Wolfe effect (ISW, Sachs & Wolfe (1967)). In chapter 4, after presenting how we have beentesting our methology and optimizing it using simulations, we will expose the results obtain using DESY1 catalogs.

Part I

Concepts

Chapter 1

Cosmological background

Cosmology can be defined as the science of our universe at large scales, which means that it aims to describe our universe as a whole in terms of its history and dynamics. While the whole history of astronomical studies can be dated to antiquity, modern cosmology is only about a hundred years old and for this reason the last century has been the spectator of great discoveries and increasing understanding about our universe at the largest scales. Indeed, starting from Einstein's theory of General relativity (GR hereafter, Einstein (1916)) that allows us to describe the universe as a whole, the community has built, with observations, models of evolution for our universe.

The current preferred theory for our universe is the so-called standard model of cosmology. This is the simplest theory so far in good agreement with observations. It describe a universe that started from an initial explosion (the Big Bang) followed by a brief highly accelerated expansion (inflation)after which an expansion era started and recently entered in an accelerated expansion. Even-though this standard model for our universe seems, at the day of today, consistent with various cosmological probes, it suffers from unanswered questions such as what is the nature of the energy that causes the late cosmic acceleration.

This chapter aims to give a brief overview of, in one hand, the state-of-the-art of present day cosmology by exposing the key discoveries and theoretical developments made in the past century to arrive there, and, on the other hand, the different observational tools and probes currently used in cosmology and useful for the following chapters of this thesis. This chapter will be organized as follows: at first, in section 1.1, the main historical key dates that yielded us the present cosmological scenario will be developed in. Then, section 1.2 will be dedicated to the different probes that are commonly used nowadays cosmological analysis and more specially that will be useful as background in the context of this thesis.

1.1 History and Theory

1.1.1 General relativity (1915) : a theory to describe our universe

At the beginning of the twentieth century, special relativity (SR hereafter) had re-conciliated electromagnetism fundamental laws (the so-called Maxwell laws) with mechanics. Indeed, in Einstein (1905), Albert Einstein extended the principle according to which physical laws should be equivalent from a reference frame to another to the laws of electromagnetism, introducing the Lorentz invariance in the **Minkowski** metric $\eta_{\mu\nu} = diag(1, -1, -1, -1)$, in which the line element ds is expressed as :

$$\begin{aligned} ds^2 = \eta_{\mu\nu} dx^\mu dx^\nu &= c^2 dt^2 - (dx^2 + dy^2 + dz^2) \\ &= c^2 dt^2 - dr^2 - r^2 d\theta^2 - r^2 sin^2(\theta) d\Phi^2 \end{aligned} \tag{1.1}$$

where (x, y, z) (and (r, θ, ϕ)) are the Cartesian (spherical) coordinates.

However, even though on one hand SR theory had shown to be consistent with electrodynamics, on the other hand Newtons theory of gravity was not compatible with it. In this context, Einstein spent the following years attempting to construct a new 'relativistic' theory of gravitation, the so-called General Relativity (GR here after) that he presented ten years latter in Einstein (1916) and for which the first goal was to reconcile Newton gravitational theory to special relativity, which imposes physical laws to be Lorentz invariant.

Einstein's General Relativity is based on the **equivalence principle** that states that a gravitational field is equivalent to the acceleration of an non-inertial referential with respect to the reference one. Consequently, the free-fall of objects in similar initial conditions will follow the same trajectories in a gravitational field, making these trajectory universal and intrinsic properties of space time. From here, it is possible to describe gravitation as a deformation of space-time induced by energy and mass, relating the **energy-momentum** tensor $T_{\mu\nu}$ accounting for the energy density distribution to the **Einstein tensor** $G_{\mu\nu}$ describing the curvature of space time and obtain the so-called Einstein equations:

$$G_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} - \Lambda g_{\mu\nu}, \tag{1.2}$$

where Λ is the cosmological constant[1] and $g_{\mu\nu}$ the metric of our universe.

These equations appeared to be complicated to solve (if not making any particular assumptions) and not having a single solution. However, in Einstein (1917) Einstein exposed a solution to the universe field equation assuming our universe follows a **cosmological principle**. The

[1] Note that Einstein initially introduced the cosmological constant to allow static solutions to equation 1.2. However, even though this constant disappeared after the discovery of the universe's expansion (see 1.1.3), it has been recently reintroduced to explain cosmic acceleration (see 1.1.6).

cosmological principle is based on two hypothesis made on our universe at large scales, namely this principle states that we are in a :

- **Homogeneous universe:** the observed universe at large scales does not depend on the observer position.
- **Isotropic universe:** At large scales, the observed universe does not show any preference from one direction to another.

In the largest scales, these properties have shown to be in agreement with observations. However, the model presented by Einstein is instable, and assume a cosmological constant to ensure a static universe, which was later discarded by observations (see 1.1.3).

1.1.2 The Friedmann equations (1922):

A few years later, A. Friedmann Friedmann (1922) and G. Lemaitre Lemaître (1927), proposed a solution for Einstein's field equations relaxing the assumption of a static universe. Allowing the universe to expand (or contract) with time implies that the Minkowski metric and line element of eq. (1.1) used in special relativity should be modified by introducing the scale factor of the universe $a(t)$ that parametrizes the expansion/contraction. Namely, if an object is at a distance x_0 from an observer today, at an earlier time t, the distance between the observer and the object was $a(t)x_0$. Thus, in an expanding universe, the metric will depend on this scale factor as $g_{\mu\nu} = diag(1, -a^2(t), -a^2(t), -a^2(t))$, having then the line element :

$$ds^2 = -c^2dt^2 + a(t)^2\left(\frac{dr^2}{1-kr^2} + r^2(d\theta^2 + sin^2\theta d\phi^2)\right) \tag{1.3}$$

this metric is known as the Friedmann-Lemaitre-Robertson-Walker (FLRW) metric, and k represents the universe's curvature and can take the values :

$$\begin{cases} k = -1 & \rightarrow \text{open universe} \\ k = 0 & \rightarrow \text{flat universe} \\ k = 1 & \rightarrow \text{close universe} \end{cases} \tag{1.4}$$

If one now inserts this metric in the Einstein field equations (1.2), with the energy-momentum tensor of a perfect isotropic fluid, diagonal and only depending on the density $\rho(t)$ and the

pressure $P(t)$, due to the cosmological principle assumption:

$$T_{\mu\nu} = \begin{pmatrix} \rho & 0 & 0 & 0 \\ 0 & -P & 0 & 0 \\ 0 & 0 & -P & 0 \\ 0 & 0 & 0 & -P \end{pmatrix}, \tag{1.5}$$

one can solve Einstein's equations obtained a series of equations known as the Friedmann equations that describe the evolution of the scale factor of the universe with respect to time as:

$$\frac{\dot{a}^2}{a^2} = \frac{8\pi G}{3c^2}\sum_i \rho_i - \frac{k}{a^2} + \frac{\Lambda}{3} \tag{1.6}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3c^2}\sum \rho_i + 3P_i, \tag{1.7}$$

where the dot stands for the derivative with respect to time, and the sum is done for the different components i of our universe. These two equations can also be combined and give a third one comonly used in cosmology the so-called **continuity equation** :

$$\dot{\rho} + 3\frac{\dot{a}}{a}(\rho + P) \tag{1.8}$$

In general, it is common to parametrize the different components using the so-called equation of state parameter $P = w\rho$. We will then have $w = 0$ for the matter components, $w = 1/3$ for radiation, and finally $w = -1$ for the cosmological constant. Using the equation of state parameter for each component together with its continuity equation (1.8), one can find its density evolution with respect to the universe scale factor $a(t)$ and then using the Friedmann equation (1.6), find the relation between scale factor and time.

- **radiation domination:** $P = \frac{\rho}{3}$, thus from the continuity equation (1.8), $\rho_r \propto a^{-4}$ and then solving the Friedmann equation in an epoch t where the radiation dominates the energy budget of our universe and obtain:

$$a(t) \propto t^{1/2}. \tag{1.9}$$

- **matter domination:** pressure-less fluid ($P = 0$), so proceeding similarly to the previous case but considering now an epoch where the matter is the dominant component of our universe, from eq.(1.8) $\rho_r \propto a^{-3}$ and from eq.(1.6):

$$a(t) \propto t^{2/3}. \tag{1.10}$$

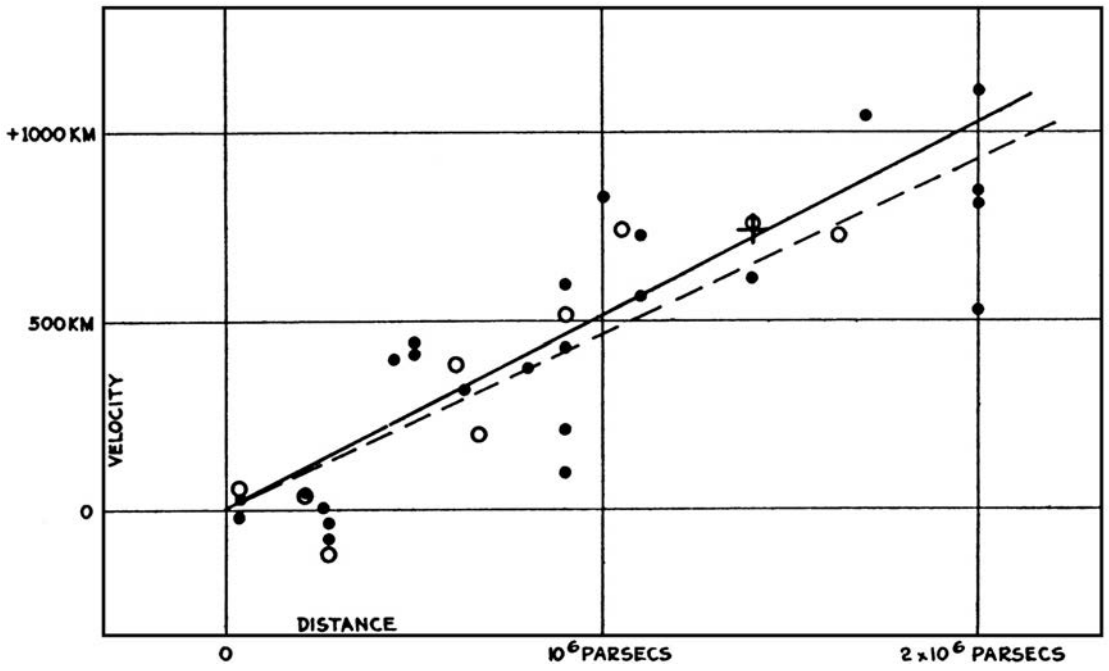

FIG. 1.1.1 – Hubble measurement of the radial velocity of extra-galactic objects with respect to their distance to the observer.

- **The cosmological constant domination:** solving the first Friedmann equation, one has $(\dot{a}/a)^2 = \Lambda/3$ wich thus leads to :

$$a(t) \propto e^{\sqrt{\frac{\Lambda}{3}}t}. \tag{1.11}$$

1.1.3 Hubble and the expanding universe 1929

In parallel to the theoretical advancements presented above, using the relation between Cepheid's[2] period and luminosity discovered by Henrietta Leavitt few years earlier Leavitt & Pickering (1912), Edwin Hubble had been measuring the distance of extra-galactic objects and could observe a proportionality relation between the distances and the radial velocities of the objects (Hubble (1929)). Figure (1.1.1) is the original published result from Hubble showing the measured radial velocities of 'nebulae' with respect to the distance at which we are observing them. As it can be seen in the figure, Hubble could observe in these measurements a linear relation between the two quantities. In other words, on one hand, observed objects tend to recede from us, and, on the other hand the further the object is located, the faster its recession velocity. This has been reported as the first proof of the universe's expansion.

[2]Cepheids are variable stars with a period of pulsation of about a day.

From this work, Hubble estimated the relation between radial velocity and distance, the so-called Hubble law, as:

$$v = H_0 d, \tag{1.12}$$

v being the radial velocity of the observed object, d its distance to the observer and H_0 the proportionality factor now known as the Hubble parameter evaluated at $H_0 = 530 kms^{-1}Mpc^{-1}$ at the time. This parameter has remained a key parameter on the present standard model of cosmology and since Hubble it has been re-estimated by various experiments obtaining values approaching $H_0 \sim 70 kms^{-1}Mpc^{-1}$.

An important aspect to mention here is that in fact Hubble was not directly measuring the radial velocity, he measured the shift towards the red part of the spectra of the different objects. Indeed, due to Doppler effect, if an object is receding (or approaching) to an observer, the emitted light will be observed shifted in the spectra. Thus if one has information on the emitted light, by comparing it to the observed one, one can extract information on the recession (or approaching) velocity. A common indicator of such effect is the so-called **redshift** defined as:

$$1 + z = \frac{\lambda_o}{\lambda_e}, \tag{1.13}$$

where z is the redshift of the observed object and $\lambda_{o,e}$ reffer to an observed and emitted wavelength respectively[3].

Moreover, as developed previously, the scale factor of our universe $a(t)$ is the parameter that rules the change in the universe metric (1.3). Therefore, it can be related to the expansion rate as follows :

$$H(t) = \frac{\dot{a}}{a}. \tag{1.14}$$

On the other hand, it is also possible to relate the observed redshift of an object to the scale factor $a(t)$, namely considering two emitted wave-fronts λ_e, emitted times t_e and $t_e + \delta t_e$ and observed at times t_0 and $t_0 + \delta t_0$. The distance travelled by the two wave-fronts will be similar (in the case where the time between the two emissions is small enough that the scale factor can be considered as constant between them), and one can then relate the two traveling distance as:

$$\int_{t_e}^{t_0} \frac{dt}{a(t)} = \int_{t_e+\lambda_e/c}^{t_0+\lambda_0/c} \frac{dt}{a(t)} \rightarrow \frac{a(t_e)}{\lambda_e} = \frac{a(t_0)}{\lambda_0} \tag{1.15}$$

and from eq.(1.16), one can then infer the relation between redshift and scale factor as:

$$1 + z = \frac{\lambda_o}{\lambda_e} = \frac{a(t_0)}{a(t_e)}. \tag{1.16}$$

[3]Note that the redshift could be related to the radial velocity as $z + 1 = \sqrt{\frac{1+v/c}{1-v/c}}$.

It is then possible to rewrite the Friedmann equation as a function of the Hubble parameter. The first Friedmann equation is often rewritten in terms of the so called energy density parameters Ω_i:

$$\frac{H(a)^2}{H_0^2} = \Omega_{0,r}a^{-4} + \Omega_{0,m}a^{-3} + \Omega_{0,k}a^{-2} + \Omega_{0,\Lambda}, \tag{1.17}$$

where Ω_i is defined for each component of our universe as $\Omega = \rho/\rho_{crit}$ where the critical density ρ_{crit} is the density that make our universe flat today without cosmological constant: $\rho_{crit} = 3H_0^2/8\pi G$ (from eq. (1.6)).

1.1.4 CMB discovery : Big bang theory 1940-1965

1.1.4.1 Prediction and detection

On one hand, from the discovery of the expanding universe, one can easily deduce that in the past the universe was smaller, and on the other hand, in the early 1940s, Gamow (1946), together with his collaborators Alpher et al. (1953), looking for a model able to explain element formation had proposed a theory in which the presence of these elements could have been created in a warm primordial state of an expanding universe, and presented the well known **Big Bang** theory. In this theory, the early universe, dominated by radiation, was in the form of a hot opaque plasma (ionized) and they also predicted that today the universe should be filled with a thermal electromagnetic radiation remnant of this early universe plasma that should have a black-body spectrum, for which they estimate the temperature to be around $\sim 5K$. Indeed, as we go back in time, the universe was warmer. We can then go back to an epoch when the universe energy was such that the photons and electrons were in thermal equilibrium following the photo-ionisation equation :

$$e^- + p \rightleftharpoons H + \gamma. \tag{1.18}$$

As time goes om, the temperature is cooling down (the universe expands), until it reaches a temperature ($T \sim 3700K$) where the thermal-equilibrium breaks and allows the electrons to couple to baryons and form the first neutral atoms: the so-called **recombination**. The photons acquire then a mean free path larger than the universe size and are able to propagate: **the decoupling**. In principle, one should be able to observe this surface at the time where the photons decoupled: **the last scattering surface**.

A few decades later, in 1965, while characterizing different sources of noise of an antenna, Penzias and Wilson noticed an isotropic radio noise that didn't seem to vary with the seasons Penzias & Wilson (1965). This radiation, now known as the Cosmic Microwave Background (CMB,hereafter), was identified as the remnant cosmological radiation predicted by Gamov and Alpher. It appears in the microwave sector of the radio spectra at a redshift $z \sim 1100$ ($\sim 400,000$ years ago). Again, a few decades later, the Cosmic Background Explorerand (COBE) (Boggess

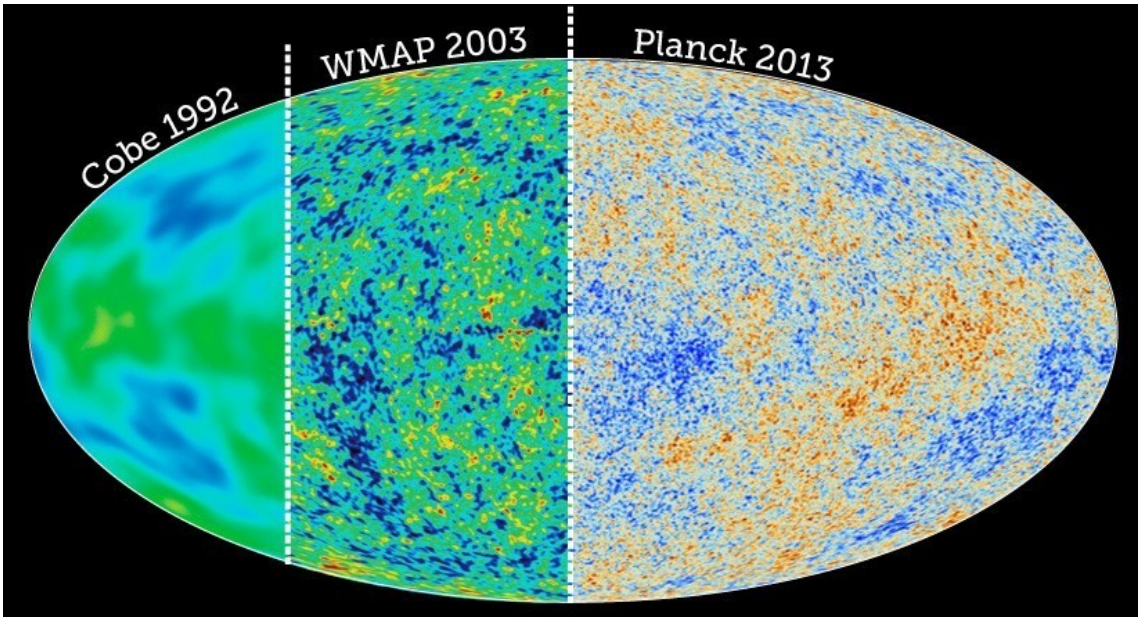

FIG. 1.1.2 – Three different observations of the CMB temperature anisotropies : COBE from 1992 (left), WMAP 2003 (middle) and Planck 2013 (right)

et al. (1992)) mission confirmed the detection as well as its black-body spectrum shape. Moreover, the Far-InfraRed Absolute Spectrophotometer FIRAS instrument measured a temperature of the CMB of $2.72548 \pm 0.00057K$ (Fixsen (2009)).

On the other hand, it is also important to mention here, that beside being a good argument in favor of the big bang theory, the CMB also appeared to be a new probe rich in information about our universe, and it will be develop in section 1.2.4. Indeed, even-though the COBE mission measured a particularly homogeneous radiation, while increasing the precision of the observations we could detect small anisotropies (Smoot et al. (1992)), and since COBE additional mission with better instrumentation have measured these anisotropies with increasing precision. Figure (1.1.2) shows a comparison of the precision on the CMB radiation measurement.

1.1.4.2 Inflation

In the previous section we have developed how the Big-Bang theory had been accepted after the discovery of the cosmic microwave background. However, even-though this theory appears to be convincing after the discovery of the CMB, it is by itself suffering of unexplained problems.

- **The flatness problem :** The fact that observations seem to prefer a flat universe ($\Omega_k = 0$) today appear to be unsatisfactory in the sense that it implies, from the Friedmann equations, that the universe was even flatter before, and thus imposes really strict limits on the initial conditions (a fine-tuning problem).

- **The horizon problem :** The smoothness of the CMB temperature is intriguing in the sense that two points located at large distances (larger than the horizon size[4]) should not in principle show any causality between them.

Nevertheless, in 1981, A. Guth presented a theory of an accelerated period in the early universe that solves these two problems (Guth (1981)). Beside this, the early inflationary scenario also gives a mechanism for the apparition of the structure we are observing today in the universe. Indeed, the quantum fluctuations in the early universe grew exponentially during the inflation era and are the seed of the anisotropies observed today (both in the CMB as developed in 1.2.4 and in the Large Scale Structure of our universe, as it will be developed in 1.2.2).

1.1.5 The presence of Cold Dark Matter: 1937 - today

An additional crucial discovery that occurred in the past century is the existence in our universe of an 'invisible' matter. Indeed, in the early 1930's, Zwicky (Zwicky (1937)) measuring the radial velocities of galaxies in the Coma cluster estimated a total mass $\sim$ 400 times larger than the mass coming from ordinary (visible) matter[5]. After this, several other measurements also showed a missing mass problem in galaxy clusters as well as galaxies themselves (see for example Smith (1936); Babcock (1939); Kahn & Woltjer (1959)). Moreover, in Ostriker & Peebles (1973), using numerical simulations, the stability of galaxy disks was shown to be maintained only in the presence of a surrounding halo of invisible matter. More recently, even if already proposed by Zwicky, gravitational lensing techniques (see 1.2.3) have been used to map the mass distribution and have also found mass excess in galaxies. The presence of this invisible matter was therefore been indirectly observed using different probes and targets. As a consequence, one can separate the matter component in the Friedmann equation ((1.17)) into two distinct terms:

$$\Omega_m = \Omega_{dm} + \Omega_b, \tag{1.19}$$

where Ω_b stands for the ordinary matter (the baryons) and Ω_{dm} accounts for the dark matter.

1.1.6 Accelerating expansion: 1998

Supernovae are events that happen at the death of a star. The explosion occurring liberates a enormous quantity of energy and thus can be observed at high redshift. Moreover, Type 1a supernovae (SN Ia hereafter), have the advantage of being 'standard candle objects', which means that by observing the luminosity function of these events, one can infer the absolute magnitude

[4]The horizon size being the larger distance that light could have travel since the beginning of the universe.

[5] Note that Zwicky had been using the value of the Hubble constant as measured by Hubble. By using the current Hubble constant, the discrepancy between total and visible mass reduces from 400 to $\sim$ 50.

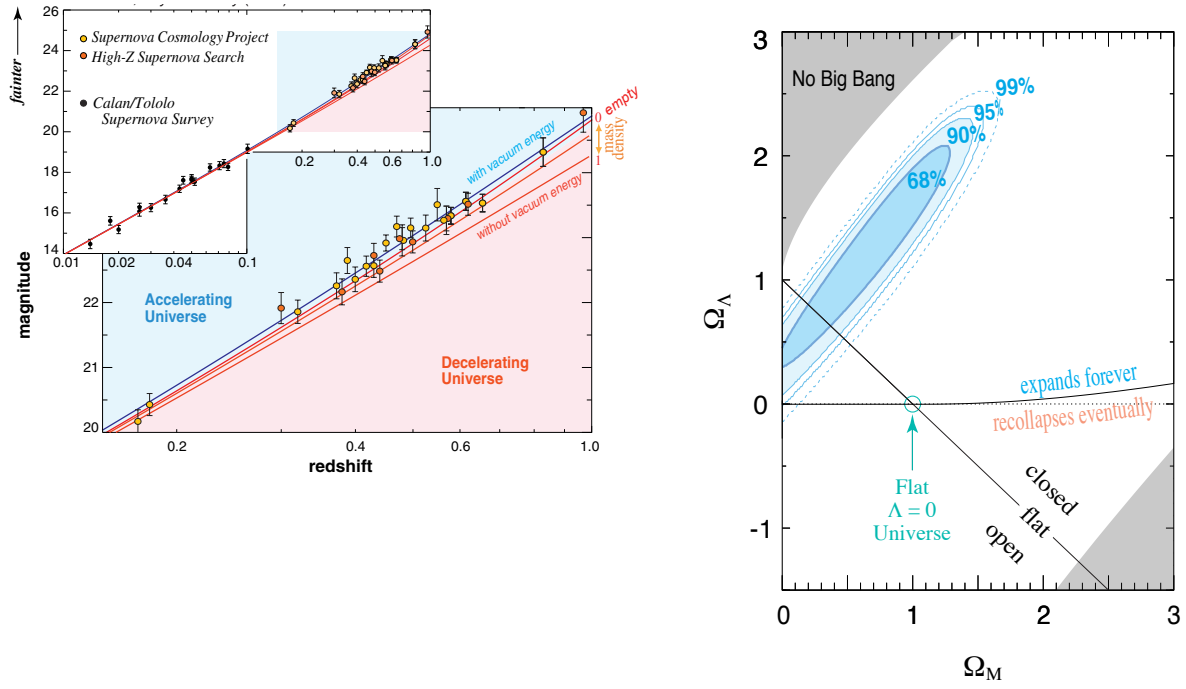

FIG. 1.1.3 – *Left panel:* Results from the two Supernovae experiments showing the magnitude distance relation Perlmutter (2003) *Right panel:* Best fit cosmological parameter estimation results from the The Supernova Cosmology Project Perlmutter et al. (1999).

of the event and therefore its distance, and, as it will be developed more in detail in section 1.2.1, in an expanding universe, the distance evaluation will depends on the expansion rate of our universe and thus on the different terms in the Friedmann equation.

Taking advantage of this property, in the late 1990's, two different teams Riess et al. (1998) and Perlmutter et al. (1999) had been measuring the distance of about $\sim$ 40 SN Ia and found that the SN Ia observed should have been brighter if the universe expansion was slowing down. In other terms, the universe expansion seems to accelerates. The left panel of Fig. (1.1.3) shows the measurements of the magnitude with respect to redshift obtained by the two different projects, as well as the theory prediction on this parameter plane for accelerating and decelerating universes. As it can be seen, at redshift $\gtrsim 0.2$, the measurements tend to be fainter (i.e with higher magnitude) than expected if we were in a decelerating universe.

This tremendous unexpected discovery had as a consequence the return of the cosmological constant in the Friedmann equation (1.6). Indeed this constant, introduced by Einstein to force

a static solution to his equations, had disappeared after the universe expansion's discovery, and came back to parametrize this recent cosmic acceleration. In Huterer & Turner (1999), the concept of **Dark Energy** has been introduced (in analogy to Dark Matter sec.(1.1.5) to name whatever causes this acceleration. On the right panel of Fig.(1.1.3) the cosmological inference for the amount of matter and dark energy today made by the SN Ia observations is shown. As it can be seen, the SN Ia data strongly prefers a cosmological model dominated today by Dark Energy. Namely from these analysis, the present universe seems to be constituted of $\sim 70\%$ of dark energy ($\Omega_\Lambda \sim 0.7$) against $\sim 30\%$ of mostly dark matter ($\Omega_m \sim 0.3$).

1.1.7 Towards a concordance model: today's state of the art

As it has been developed above, General Relativity theory together with observations lead to an evolution scenario for our universe. Namely, our universe was born in an initial explosion **the Big Bang** followed by a short period of accelerated expansion where the energy content was dominated by relativistic particles until the electrons recombined with protons and a residual radiation from this process is now observed in the microwave background signal in the sky: the so-called **cosmic microwave background**. Recombination was followed by a matter dominated era where the expansion got smoother and recently entered an accelerating expansion era caused by an unknown entity that observations tends to prefer as a **cosmological constant** Λ.

Figure(1.1.4) is a representation of this universe scenario. This scenario can be modeled as mentioned above as a Λ-Cold-Dark-Matter standard model with the Friedmann equation (1.6). This concordance model, albeit consistent with current observation, presents two evident intriguing caveats, namely the unexplained nature of Dark Energy and Dark Matter, which together represent about 95% of the present energy content of our universe. The current best fit value from the Planck satellite Planck Collaboration et al. (2018) is:

$$\begin{aligned} \Omega_b &= 0.048 \pm 0.002 \\ \Omega_{dm} &= 0.2647 \pm 0.015 \\ \Omega_\Lambda &= 0.6850^{+0.017}_{-0.016}. \end{aligned} \tag{1.20}$$

1.2 Observables

In the previous section we have presented the main framework that has built the state of the art of the standard model of cosmology today. However, as it has been mentioned, this model is not fully understood and the goal of cosmology today is to test this model with observational tools. In this context, in the last decades a variety of experiments have been imagined using different probes and tools in order to try to unveil the mysteries lying behind dark matter and dark energy.

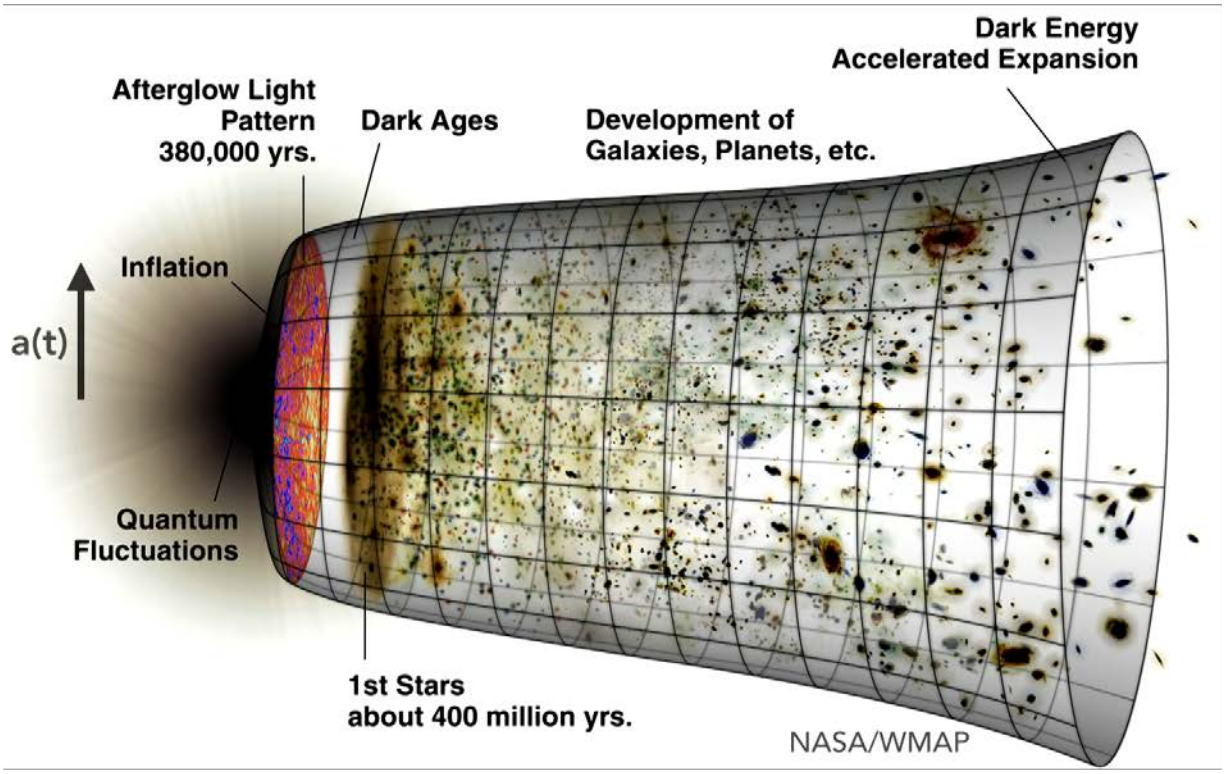

FIG. 1.1.4 – Overview of the different steps on our universe history as predicted by the Big Bang scenario starting from an early-time primordial explosion on the left hand side followed by a short but strong accelerated expansion period (inflation) until photon decoupling and the emission of the cosmic microwave background radiation, a matter dominated era and finally the recent Dark Energy dominated era.

In this section, we aim to present various observational tools that are used in current cosmology and that have been useful in the context of this work. Namely, after exposing some basic definition of distances in cosmology in (1.2.1), we will dedicate a section to the tools developed to study the large structure of our universe and extract cosmological information from it using both clustering information (sec. 1.2.2) and the gravitational lensing effect (sec. 1.2.3), and afterwards, in sec.1.2.4, we will give an overview on how the Cosmic Microwave Background can be used as a cosmological probe.

1.2.1 Cosmological distances

As it has been developed above, one pertinent aspect in cosmological studies resides in the evaluation of distances to the different structures one observes. Indeed, collecting the line of sight information of objects (the so-called redshift (1.16)) provides the cosmological time of the observed objects and mapping these distributions gives us direct information about the time evolution of our universe. However, in an expanding universe, the distance between two fixed

objects is changing through time. Astrophysicists could thus define different estimator of the distances that will present distinct dependencies on the underlying cosmology. Here we will present typical distance definitions commonly used in cosmology.

1.2.1.1 Comoving Distance

The comoving distance reffers to the distance of an object in coordinates of an expanding universe. That is to say, instead of giving the distance of an object at the time it is observed t, it gives the distance of an object as it is now taking into account the universe expansion. In other words, if an object emits a photon at time $t = t(a)$ and we are observing it at time $t = t_0$, the **comoving distance** $\chi(a)$ between the observer and the object is given by :

$$\chi(a) = \int_{t(a)}^{t_0} \frac{cdt'}{a(t')}, \tag{1.21}$$

where c is the speed of light. Using equation (1.16), it is also possible and common to define the comoving distance as a function of the redshift of the object as :

$$\chi(z) = \int_0^z \frac{cdz'}{H(z')}. \tag{1.22}$$

Moreover, if one wishes to measure the distance between two objects at the same redshift with an angular separation $d\theta$, one will have to compute the **transverse comoving distance** $D_M(z)$. This distance will differ from the previous one (1.22) if the curvature density parameter of our universe differs from zero, namely one has :

$$D_M(z) = \begin{cases} \frac{c}{H_0\sqrt{\Omega_k}} sinh\left[\frac{H_0\sqrt{\Omega_k}}{c}\chi(z)\right] & \text{if } \Omega_k > 0 \\ \chi(z) & \text{if } \Omega_k = 0 \\ \frac{c}{H_0\sqrt{-\Omega_k}} sinh\left[\frac{H_0\sqrt{-\Omega_k}}{c}\chi(z)\right] & \text{if } \Omega_k < 0 \end{cases} \tag{1.23}$$

1.2.1.2 Angular Diameter Distance

The angular size of an object depends at which distance we are observing it, there is a linear relation between the observed angular size α , its true size d of an object and the distance at which we observe it :

$$d = \alpha D \tag{1.24}$$

However cosmic expansion make the observed object bigger than for a non-expanding universe, and therefor the angular size of an object that have emitted light at the scale factor a, will evolves as d/a, if one observe the object at the comoving distance χ, then $\alpha = (d/a)/\chi$ inserting this

relation in eq.(1.24), one gets:

$$D_A(z) = \chi(a)a = \frac{\chi(z)}{1+z} \tag{1.25}$$

1.2.1.3 Luminosity Distance

The Luminosity distance D_L correspond to the distance evaluated using the relation that exists between the flux F of light observed to the intrinsic luminosity L and the distance D at which we observe an object :

$$F = \frac{L}{4\pi D^2} \tag{1.26}$$

From here it is then possible if one knows the intrinsic luminosity emitted to evaluate the distance of the object by measuring the incoming flux. However, similarly that for the angular diameter distance, in an expanding universe, the observed flux will appear diluted due to the expansion, in comoving coordinates, eq.(1.26) becomes :

$$F = \frac{L(\chi)}{4\pi\chi^2} \tag{1.27}$$

The luminosity of the source at the comoving distance χ, will decrease as the universe expands by a factor of a^2 giving then the luminosity distance:

$$D_L(z) = (1+z)\chi(z) = (1+z)^2 D_A(z) \tag{1.28}$$

1.2.2 Large Scales Structure og the universe

1.2.2.1 Structure observation and linear evolution

In section 1 we have been deriving the universe's equations under the condition that we are living in a homogeneous and isotropic universe, according to the cosmological principle. However, the first observations of clusters and super-cluster of extra-galactic objects (Abell (1958, 1961)) brought the debate about the limits of the homogeneity of the galaxy distribution. And it was only at the end of the 1970's, by virtue of the development of large scale surveys and the access to the 3D information of objects that allowed us to create 3D maps of galaxies and directly test the departure from homogeneity as a function of scale. Figure 1.2.1 shows the first data from the CfA-II Redshift Survey in 1985 de Lapparent et al. (1986), where we can observe the apparition of structure in the position of galaxies.

These structures observed in the sky can be explained from the growth of primordial quantum fluctuation in the early universe (see sec. 1.1.4.1). Indeed the exponential enlargement at early time, made the primordial quantum fluctuations grow and became the observed CMB anisotropies

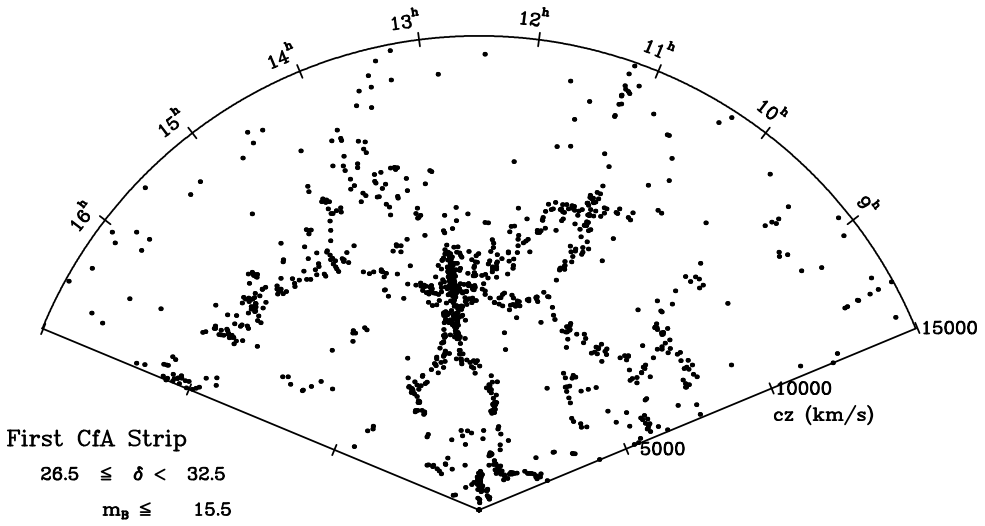

FIG. 1.2.1 – Map of galaxy distribution from the CfA Redshift Survey de Lapparent et al. (1986) showing the hierarchical organization of objects in the sky from galaxy clustering, voids and filaments.

(see section 1.1.4) and later on the Large Scale Structure of the universe as filaments, galaxy clusters, voids,...

The clustering of galaxies and clusters of galaxies can give us information on the initial conditions of our universe and its evolution. Therefore, a variety of surveys are dedicated to probe this structure (see for example The Dark Energy Survey Collaboration (2005); DESI Collaboration et al. (2016); LSST Science Collaboration et al. (2009); York et al. (2000); de Jong et al. (2015)) . The idea here is to compare a linear theory of structure growth to which one has to add non-linear effects to the observed matter distribution. At the end of the 1980s, the increasing amount of observed galaxies by large telescopes opened the path to a statistical analysis of galaxy catalogs.

To model these fluctuations at large enough scales in the matter field, one common way to proceed is to apply linear perturbation theory. To start with, one can express the deviation from homogeneity defining the density contrast δ as:

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x}) - \bar{\rho}}{\bar{\rho}}, \tag{1.29}$$

where $\bar{\rho}$ is the mean universe density and $\rho(\mathbf{x})$ is the density at the 3D location $\mathbf{x}$. If one considers scales smaller than the cosmic horizon, one can use a Newtonian approach to describe the evolution of density perturbations (Peebles (1980)), in which the density contrast for a perfect fluid with no pressure is linked to the gravitational potential through the Poisson equation $\nabla^2\Phi = 4\pi G a^2 \bar{\rho}\delta$. Combining now this equation to the Continuity and Euler equations yields to a second order differential equation for the density contrast evolution :

$$\ddot{\delta} + 2\frac{\dot{a}}{a}\dot{\delta} - \frac{3H_0^2\Omega_m}{2a^3}\delta = 0 \tag{1.30}$$

Two different modes can be identified as solution of this equation, one growing and one decaying mode for the evolution of the perturbation. In general, the growing mode is the one that dominates. We thus write the solution as:

$$\delta(\mathbf{x}, t) = \frac{D(t)}{D(t_0)}\delta(\mathbf{x}, t_0) \tag{1.31}$$

where $D(t)$ represents the growth of the perturbation and is known as the **growth factor** of our universe. This factor will depend on the underlying cosmology.

1.2.2.2 The two-point correlation function

The most common way to evaluate the amount of clustering of a given observable is with the two-point correlation function (2PCF hereafter). The 2PCF is a tool that aims to characterize a distribution is the sense that it quantifies the probability of an object A to be at a distance r from a object B compared to a randomly distributed field. For instance, in the case of a homogeneous distribution of objects, meaning a random distribution of objects without any correlation with a mean density ρ_0, the probability $dN(\vec{r})$ to find an object in the volume dV_2 at the distance r of a volume dV_1 can be expressed as :

$$dN(\vec{r}) = \rho_0^2 dV_1 dV_2. \tag{1.32}$$

If we place us now in the scenario where the distribution of objects differs from a random field, one can rewritte equation (1.32) as :

$$dN(\vec{x}, \vec{r}) = \rho_0^2(1 + \xi(\vec{r}))dV_1 dV_2 \tag{1.33}$$

where $\xi(\vec{r})$ is the two-point correlation function. In the case of a random distribution, this function is zero; otherwise, it can take positive (or negative) values in the case that the objects are more (or less) clustered. In this sense, the 2PCF is an estimation of the amount of clustering of a given sample.

The 2PCF can be expressed as function of the density contrast (eq. 1.29) as :

$$\xi(\vec{r}) =< \delta(\vec{x})\delta(\vec{x}+\vec{r}) > . \tag{1.34}$$

Several measurements of the correlation signal for either galaxy or clusters of galaxies were shown in Totsuji & Kihara (1969),Peebles (1980) to be well approximated by a power-law at small scales ($r \leqslant 150 Mpc/h$) :

$$\xi(\vec{r}) = \left(\frac{r}{r_0}\right)^{-\gamma}, \tag{1.35}$$

where γ is the slope of the power-law and r_0 corresponds to the correlation length. The correlation length corresponds to the scale at which the probability to find an object (i.e. the correlation function) equals unity $\xi(r_0) = 1$. The larger it will be this value, the stronger will be the clustering amplitude of a given sample.

If we are now placing ourselves in a large volume V, it is possible to compute the Fourier transformed of this correlation function, defined as the power spectrum:

$$P(\mathbf{k}) = \int \xi(x) e^{-i\mathbf{x}.\mathbf{k}} d^3x. \tag{1.36}$$

1.2.2.3 The bias

It is important to note that the observable used to estimate the density contrast might be a biased estimator of the underlying matter field. For instance, if one observes galaxies, one observes light, and as it will be explain later, one cannot assume that light is an unbiased tracer of our universe's matter distribution (Kaiser (1984); Fry & Gaztanaga (1993)). Indeed, as it has been seen in section (1.1), most of the matter in our universe is in the form of Dark Matter that we are not able to observe directly. We therefore have to consider the difference between the spatial distributions of luminous and dark matter, and this is done through the so-called tracer bias. In the standard Halo model, galaxies form in collapsed matter over-densities known as Dark matter halos, and there is a correlation between the mass of the host halo and the galaxy properties such as color and luminosity. Therefore, the so called galaxy bias is highly dependent on the considered galaxy sample, the redshift we are observing, as well as the scales we are looking at. We thus define the relation between the tracer density contrast $\delta_i(\vec{k}, z)$ and the underlying matter density contrast as:

$$\delta_i(\vec{k}, z) = b_i(z, \vec{k})\delta(\vec{k}, z) \tag{1.37}$$

where $b_i(z, \vec{k})$ is the bias of the considered tracer i. And consequently, if we aim to relate the correlation function of a given sample i at redshift z to the underlying Dark Matter correlation function one has:

$$\xi_i(\vec{r}, \vec{k}) = b_i^2(z, \vec{k})\xi(\vec{r}, \vec{k}). \tag{1.38}$$

This formalism can also be extrapolated to the case of the cross-correlation signal between two different tracers i and j as:

$$\xi_{i,j}(\vec{r}) = b_i(z)b_j(z)\xi(\vec{r}). \tag{1.39}$$

1.2.2.4 The angular two-point correlation function and the Limber approximation

In practical terms, and as it will be explained later on (see chapters 2 and 3), getting redshift information on every galaxy of our sample (and preserving the statistical power) is quite difficult. A way to avoid this issue is work in two dimensions, using the so called **2-point angular correlation function**, which represents the probability for a galaxy to be positioned at an angle θ in the sky from another object. It can be seen as the projection in 2D of the 3D correlation function (eq. 1.34). Similarly to equation (1.33), the probability to observe two galaxies in the solid angles $d\Omega_1$ and $d\Omega_2$ separated by an angle θ considering the surface galaxy density Σ can be expressed as :

$$dN_{2D} = \Sigma^2 d\Omega_A d\Omega_B (1 + w(\theta)). \tag{1.40}$$

If we consider now the small angle approximation, the angular correlation function $w(\theta)$ can be related to the spatial correlation function $\xi(\vec{r})$ with the Limber equation Limber (1954):

$$w(\theta) = \int_0^\infty d\chi y^4 S^2(y) \int_{-\infty}^\infty dx \xi(\sqrt{x^2 + y^2\theta^2}), \tag{1.41}$$

where $S(y)$ is the so-called survey selection function, defined as the probability for a galaxy at comoving distance χ to be in the sample. And in Fourier space it will take the form:

$$w(\theta) = \int_0^\infty P_g(k) g(k\theta) k dk, \tag{1.42}$$

where P_g is the galaxy power spectrum, and $g(k\theta)$ is the kernel expressed as :

$$g(k\theta) = \frac{1}{2\pi} \int_0^\infty dz J_0(k\theta\chi(z)) \left(\frac{dN}{dz}\right)^2 \frac{dz}{d\chi}, \tag{1.43}$$

where J_0 is the zeroth-order Bessel function and dN/dz is the probility distribution in redshift of the galaxy sample.

Moreover, considering scales small enough, as it is the case for equation (1.35), Peebles (1980) showed that the angular correlation function could also be approximated as a power law :

$$w(\theta) = A\theta^{1-\gamma}, \tag{1.44}$$

where A is the amplitude of the correlation signal.

1.2.2.5 Estimators

In an observational context, the angular correlation signal is measured by counting the excess pairs at a distance r of an object, compared to a random distributed catalog with the same geometry. To do so, different estimators exist. Here we will give an overview of the most popular estimators found in the literature. The simplest estimator known at the 'Natural estimator' was introduced by Peebles and Hauser Peebles & Hauser (1974) and has the form :

$$w(\theta) = \frac{[DD] - [RR]}{[RR]}, \tag{1.45}$$

where $[DD]$ represents the number of pairs of galaxy-galaxy (Data-Data), $[DR]$ are the pairs galaxy-random (Data-Random) and finally $[RR]$ the pairs random-random. However this estimator has been shown to be biased when one gets to large scales (Kerscher et al. (2000)), and thus it is not commonly used in current work. In today analyses, the most popular estimators used are :

- Hewett (Hewett (1982)):
$$w(\theta) = \frac{[DD] - [DR]}{[RR]}. \tag{1.46}$$
- Davis and Peebles (DP hereafter, Davis & Peebles (1983)):
$$w(\theta) = \frac{[DD] - [DR]}{[DR]}. \tag{1.47}$$
- Hamilton (Hamilton (1993)) :
$$w(\theta) = \frac{[DD] \times [RR] - [DR]^2}{[DR]^2} \tag{1.48}$$
- Landy and Szaley (LS hereafter, Landy & Szalay (1993)):
$$w(\theta) = \frac{[DD] - 2[DR] + [RR]}{[RR]} \tag{1.49}$$

At small scales, these different estimators have been shown to perform similarly, as tested in Kerscher et al. (2000). However, in there, the authors have also shown that when one goes to larger scales, the Hamilton and LS estimators were the ones managing to reconstruct better the correlation. Figure (1.2.2) shows a comparison of these different estimators in simulated catalogs from Vargas-Magaña et al. (2013) to recover the underlying correlation signal. In there, the authors used simulated galaxy catalogs with the same geometry as the BOSS survey Dawson et al. (2013) and compared the bias of each 2PCF estimator listed above in recovering the input correlation. As can be seen in the figure, at small scales all estimator seems to work fine. However if one goes to larger scales ($\gtrsim 100 Mpc/h$), all estimators become biased. The DP

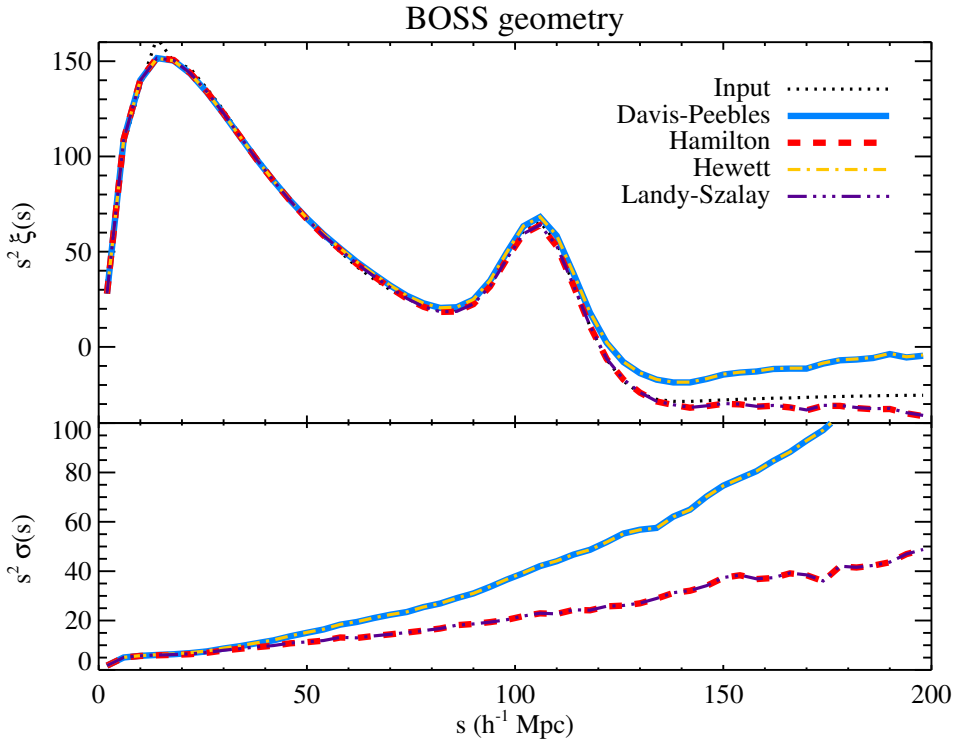

FIG. 1.2.2 – *Top panel:* Comparison of different 2-point correlation functions ξ, as a function of the comoving separation s, evaluated with the David-Peebles (blue solid line), Hamilton (red dashed line), Hewett (yellow dashed line) and Landy-Szalay (purple dashed line) estimators to the input signal in 120 mock catalogs with geometry similar to that of the BOSS survey Vargas-Magaña et al. (2013). *Bottom panel:* root mean square (RMS) of these different recovered 2PCFs, with the same color scheme.

and Hewett estimators behave similarly, as well as the Hamilton and the LS estimators, even if these two seem to be the less biased, being consistent with the results obtained in Kerscher et al. (2000). Nevertheless, in Kerscher et al. (2000), the LS estimator was shown to be less sensitive to the number of randoms in the random catalog. This has made the LS estimator the most popular estimator for correlation analyses in cosmology nowadays.

1.2.3 Gravitational Lensing

As mentioned above, in general relativity the photons follow the geodesics of a curved space-time, where the deformations are made by the gravitational potential. In this context, the theory predicts the deviation of photons travelling towards us when approaching massive objects. This effect is known as **Gravitational Lensing**, and has been detected for the first time observing the deflection of light rays from stars approaching the Sun during a solar eclipse on May, 29, 1919 (Dyson et al. (1920)). In principle, measuring this effect could give us information on the underlying matter field that these photons are crossing. Indeed, studying the distortion of light

coming from background sources when crossing a foreground gravitational field has become an essential tool in modern cosmology, and a variety of experiments are now studying these effects and constraint cosmology with it. (see for example sec.(2) and Prat et al. (2017); Troxel et al. (2017); Hildebrandt et al. (2017); Kilbinger et al. (2013)). In this section, we will introduce the basics of gravitational lensing in cosmological analyses.

1.2.3.1 Propagation and deflection in an inhomogeneous universe

In section 1.1.2, we have presented the General Relativity equations considering an isotropic and homogeneous universe. In the case of gravitational lensing, however, one has to consider scales where homogeneity cannot be assumed anymore. Applying first order perturbation theory, one can reformulate equation (1.3) and obtain the line element equation for light propagation in an inhomogeneous universe as :

$$ds^2 = \left(1 + \frac{2\Psi}{c^2}\right) c^2 dt^2 - a^2(t) \left(1 - \frac{2\Phi}{c^2}\right) dl^2, \tag{1.50}$$

where Φ and Ψ are the two Bardeen potentials that are considered to describe weak fields, in GR, these two fields are equivalent, $\Phi = \Psi$, and the potential of a lens with mass M and radius R will be $\Phi, \Psi \sim GM/R = (c^2/2)(R_S/R)$, G being Newton's gravitational constant and R_S the Schwarzschild radius. One can solve eq.(1.50) in the context of null-geodesics travel ($ds^2 = 0$) and obtain the time of travelling of the photons:

$$t = \frac{1}{c} \int_0^r \left(1 - \frac{2\Phi(r')}{c^2}\right) dr', \tag{1.51}$$

where the integration has to be performed from 0 to the physical coordinate of the source r. In order to derive the equations of light deflection when travelling through an inhomogenous universe, one can use Fermat's principle of minimun travel time. From here, solving the equations for $dt = 0$, one can obtain the **deflection angle** α that corresponds to the difference between the emitted photon and the observed one :

$$\hat{\alpha} = -\frac{2}{c^2} \int \nabla_\perp \Phi dr, \tag{1.52}$$

where the gradient of the potential is perpendicular to the light path.

1.2.3.2 Estimating gravitational lensing effects

If one considers a source **S** that is emitting photons that will deviate an angle $\hat{\boldsymbol{\alpha}}$ in their path when approaching a lens **L**, an observer **O** will observe an image **I** in a different position than the original one. The left panel of Figure(1.2.3) is a schematic representation of light deflection

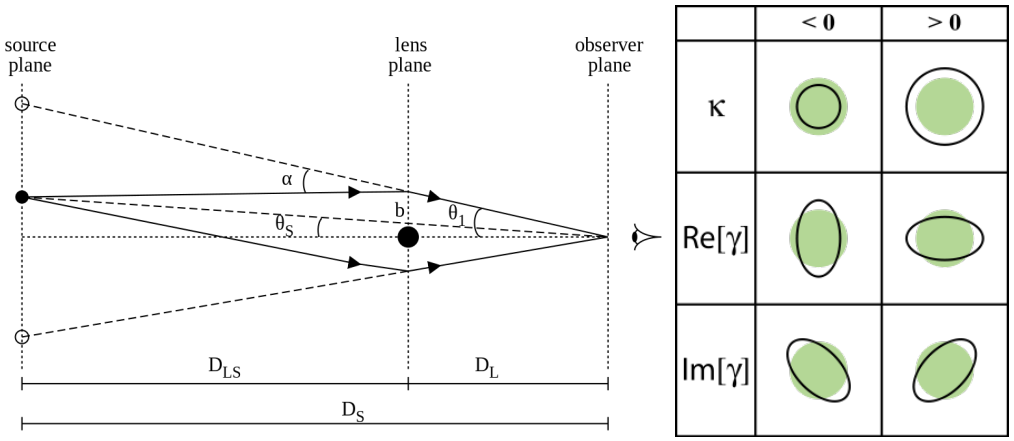

FIG. 1.2.3 – *Left panel:* Diagram representing a typical lens system. *Right panel:* Schematic representation of the convergence (κ) and the two components of the shear (γ) lensing effect on a circular source.

due to gravitational lensing effects. Using a notation similar to that in Figure(1.2.3), in the small angle approximation, one can relate these different quantities as :

$$\theta_I = \theta_S + \frac{D_{LS}}{D_S}\hat{\alpha}. \tag{1.53}$$

To simplify, it's common to rewrite this equation in terms of the angular position difference. In other words, the difference between the observed angular position ($\boldsymbol{\theta_I}$) and the true angular position ($\boldsymbol{\theta_S}$) can be related to the deflection angle induced by graviational lensing as :

$$\theta_I - \theta_S = \alpha, \tag{1.54}$$

where $\alpha = \frac{D_{LS}}{D_S}\hat{\alpha}$ is the reduced deflection angle. Note that for a fixed angular position of the source θ_S, it is possible to obtain various solutions for θ_I. In this case, one will observe various lensed images of the same source in the sky. This effect is called **strong lensing**.

Evaluating the effect of the presence of a gravitational field on the path of light-rays can be done using the so-called **Jacobian matrix**. Indeed, if one now inserts equation (1.52) into equation (1.53) and differentiates with respect to θ_I, one can define the Jacobian matrix as $A = d\theta_S/d\theta_I$:

$$A = \begin{pmatrix} 1 - \frac{\partial^2\Phi}{\partial x_1^2} & -\frac{\partial^2\Phi}{\partial x_1 \partial x_2} \\ -\frac{\partial^2\Phi}{\partial x_1 \partial x_2} & 1 - \frac{\partial^2\Phi}{\partial x_2^2} \end{pmatrix} \tag{1.55}$$

The inverse of the Jacobian matrix A^{-1} describes the local mapping of the source light distribution to the image coordinates.

The consequences of gravitational lensing on an observed object can be described by two distinct effects. The right panel of Figure(1.2.3) show the consequences on the deformation of a circular object induced by these two effects:

- **The convergence $\kappa(\theta)$:** Isotropic deformation induced by the amount of matter in the lens,

$$\kappa(\theta) = \frac{\Sigma(\theta D_L)}{\Sigma_{crit}}. \tag{1.56}$$

  Here, $\Sigma(\theta D_L)$ is the surface mass density :

$$\Sigma(\theta D_L) = \int \rho(\theta D_L, z) dz, \tag{1.57}$$

  $\rho(\theta D_L, z)$ being the mass density, and the critical mass density is defined as :

$$\Sigma_{crit}^{-1}(z_l, z_s) = \frac{4\pi G}{c^2} \frac{D_{ls} D_l}{D_s}. \tag{1.58}$$

  Moreover, it is possible to relate the reduced deflection angle to the convergence field as:

$$\alpha(\theta) = \frac{1}{\pi} \int d^2\theta' \kappa(\theta') \frac{\theta - \theta'}{|\theta - \theta'|^2}. \tag{1.59}$$

- **The shear $\gamma(\theta)$:** Anisotropic deformation due to tidal effects from the graviational field that will distort a circular object into an ellipse. The shear field is a two component field $\gamma = \gamma_1 + i\gamma_2$.

Considering these two effects, it is possible to rewrite the Jacobian matrix of equation (1.55) as :

$$A = \begin{pmatrix} 1 - \kappa - \gamma_1 & -\gamma_2 \\ -\gamma_2 & 1 - \kappa + \gamma_1 \end{pmatrix}. \tag{1.60}$$

defining thus the shear and the convergence as second order derivatives of the gravitational potential. Nevertherless, when one is looking at the lensing effect induced by the large scale structure of our universe, these effects are relatively weak (κ and γ are of order of few percent or less).

The shear is the more important observable in weak gravitational lensing for galaxy cosmological studies (see chapter 2) since it describes the distortion of images, and is by correlating these distortions that we can detect graviatational lensing effects. Since, the effects of the convergence will be only on the size of the observed object and not its shape, using the observed shape of the galaxy one can introduce the reduced shear:

$$g_i = \frac{\gamma_i}{1 - \kappa} \tag{1.61}$$

where $i = 1, 2$[6]. However, in weak lensing analysis we are looking at regimes where both convergence and shear are small compared to unity ($\kappa, \gamma << 1$), which makes the shear itself a

[6]The reduced shear has the same spin-2 properties as the shear.

good approximation to the reduced shear. Moreover, in this regime, the Jacobian matrix can be inverted and each source will be observed as a single lensed image.

1.2.4 Cosmic Microwave Background

As mentioned above, the radiation coming from the early universe has become one of the pillar of the big bang model. Moreover as it will be summarized here, this radiation has also become one of the stronger probes to put constraints on this model. And different observational projects have been carried out to study in detail this radiation, such as the COmic Background Explorer (COBE, Boggess et al. (1992)) , the Wilkinson Microwave Anisotropy Probe (WMAP, Bennett et al. (2003)) or, more recently, the Planck satellite (Tauber et al. (2010),Planck Collaboration et al. (2011)), and the South Pole Telescope (SPT, Ruhl et al. (2004); Padin et al. (2008); Carlstrom et al. (2011)).

The main feature used in the CMB observations is the temperature anisotropy. Indeed even though at large enough scale the CMB temperature look homogeneous, if one looks at smaller scales, one will observe small anisotropies, remnant of primordial quantum fluctuations that grew during inflation (see 1.1.4.2). These fluctuations in the temperature field can be separated in to two main types: the **primordial anisotropies** are temperature variations due to the physics of the universe at recombination, the **secondary anisotropies** are induced by perturbations that photons suffer along the path to us. Additionaly, a temperature **dipole** coming from the Doppler effect on the CMB photons due to the observer movement also has to be taken into account.

1.2.4.1 Temperature anisotropies

As mentioned above, the primordial temperature fluctuations are the ones containing the information of the physics of our universe at the time of recombination. Therefore, it will be possible to infer cosmological information from the temperature anisotropy field :

$$\Theta(\theta, \Phi) = \frac{\Delta T(\theta, \Phi)}{T_0}, \tag{1.62}$$

where θ and Φ are the equatorial coordinates. The amplitude of these fluctuations is of order $\Delta T/T_0 \sim 10^{-5}$, and thus high-precision surveys are required in order to be able to detect them.

Similarly to what is commonly done for Large Scale Structure analysis (1.2.2), it is possible to describe these anisotropies in a statistical way using the angular power spectrum (in temperature) C_l^T. To obtain this power spectrum, one can start by decomposing the temperature field of eq.

(1.62) in spherical harmonics as :

$$\Theta(\theta, \Phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} a_{lm} Y_{lm}(\theta, \Phi) \tag{1.63}$$

The amplitude a_{lm} associated to the spherical harmonic Y_{lm}, quantifies the temperature variations in the angular scales $\Delta\theta$ with :

$$\Delta\theta \sim \frac{\pi}{\theta}. \tag{1.64}$$

It is then common the introduced the C_l coefficients as :

$$C_l = \left\langle |a_{lm}|^2 \right\rangle_m, \tag{1.65}$$

where the mean is evaluated over the $(2l + 1)$ values of a_{lm}.

Figure (1.2.4) shows the Temperature power spectrum as observed by Planck with the best fit ΛCDM model (from Planck Collaboration et al. (2018)). The different features observed in this power spectrum are the reflection of the physical conditions at the time of decoupling. It is possible to identify three main regions (marked in the figure):

- **The Sachs-Wolfe Plateau:** At the largest scales, namely scales larger than the horizon ($l \lesssim 30$). At these scales, the information did not have time to travel from a point to another (no causality), thus the fluctuations come from the initial photon fluctuations and from the Sachs-Wolfe effect[7] Sachs & Wolfe (1967).

- **The acoustic peaks:** For scales smaller than the horizon. Before recombination, fluctuations in the plasma density (small overdense regions) tend to gravitationally attract the matter from less dense region. This effect is opposed by the force induced by radiation pressure that tends to repulse matter. These two opposite effects make oscillate the ionized gas between phases denser and warmer and phases less dense and colder. These oscillations will appear on the temperature power spectrum as peaks (and troughs) known as the acoustic peaks. And the position as well as the amplitude of these peaks will depend on the detailed cosmology.

- **The Silk damping (Silk (1968)):** At the smallest scales. This damping effect is due to the fact that recombination is not an instantaneous process, photons are still interacting from time to time with the baryons, which induces a damping effect at the smallest scales.

[7]The Sachs-Wolfe effect is the fact that photons escaping from potential wells will lose energy while escaping and thus appear colder than the emitted ones.

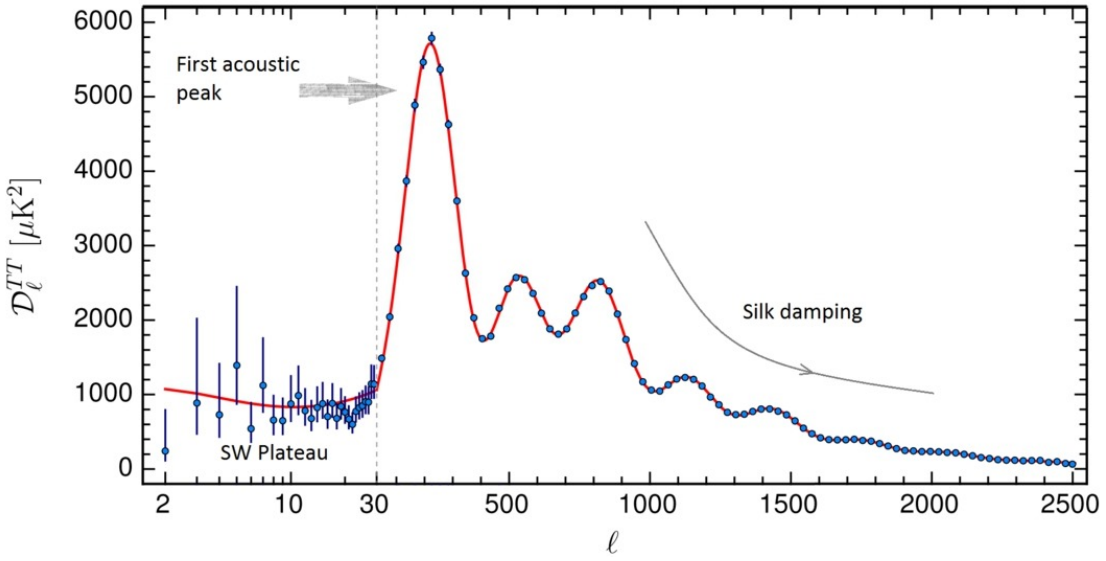

FIG. 1.2.4 – Angular power spectrum ($D_l^{TT} = l(l+1)C_l/2\pi$) from the last Planck analysis Planck Collaboration et al. (2018). The blue error bars represent the measurement by Planck and the red solid line the best fit ΛCDM.

1.2.4.2 Cosmic Microwave Background Lensing

Similarly to the distortions of galaxy shapes due to the underlying matter field used in weak lensing analysis (see sec.1.2.3), the photons that we are observing from the cosmic microwave background are also suffering deflection effects and will be observed lensed. Figure (1.2.5) pictures the deflection of the CMB photons while they are crossing the large structures of our universe (left panel) and the effects on the observed temperature anisotropies due to these lensing effects (right panel).

CMB lensing cosmology is a quite new tool in cosmology that started with the recent publication of reconstructed CMB convergence maps (Planck Collaboration et al. (2016c); Omori et al. (2017)) and that will be used in section 4. The matter density field on the path of the photons coming from the CMB will remap the emitted temperature field of equation (1.62) as:

$$\Theta(\hat{n}) = \tilde{\Theta}(\hat{n} + d(\hat{n})), \tag{1.66}$$

where the tildes represent the unlensed temperature field and $d(\hat{n})$ is the deflection angle of the CMB photons travelling toward us.

As presented in section 1.2.3, from the light deflection angle, it is possible to define a lensing potential. In Lewis & Challinor (2006) the CMB lensing potential in a direction $\hat{\mathbf{n}}$ is defined as[8]:

$$\Phi(\hat{\mathbf{n}}) = -2\int_0^{\chi_{CMB}} d\chi \frac{\chi_{CMB} - \chi}{\chi_{CMB}\chi}\Psi(\chi\hat{\mathbf{n}}; t) \tag{1.67}$$

[8]Assuming a flat universe.

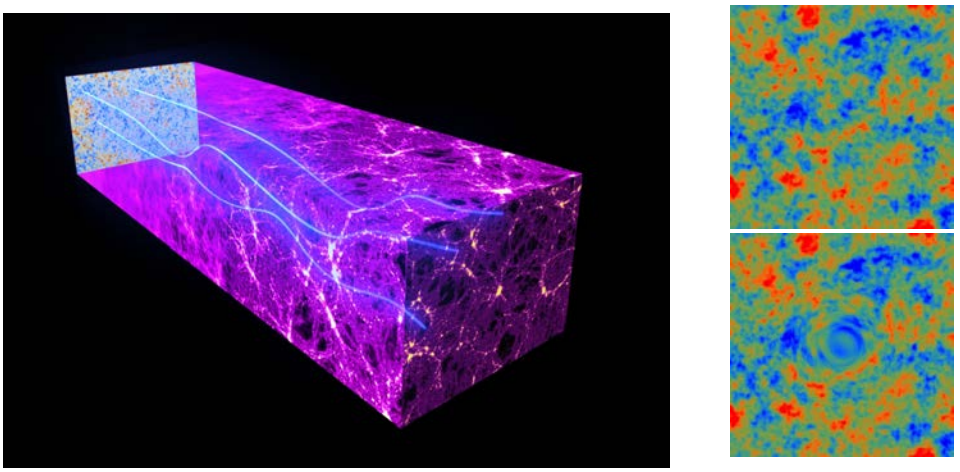

FIG. 1.2.5 – Amplified example of lensing effects on the temperature field of the cosmic microwave background (from Hu & Okamoto (2002)).

where χ is the comoving distance, $\chi_{CMB} \sim 14Gpc$ is the distance to the surface of last scattering, and Ψ it the gravitational potential evaluated at a conformal time χ in the $\hat{\mathbf{n}}$ direction and at the time ($t = \eta_0 - \chi$) where η_0 is the conformal time today [9].

And the gravitational potential can be expressed as a function of the underlying dark matter density ($\delta(\chi\hat{\mathbf{n}}; t)$) field through the Poisson equation :

$$\nabla^2\Psi(\chi\hat{\mathbf{n}}; t) = \frac{3H_0^2\Omega_m}{2a(z)}\delta(\chi\hat{\mathbf{n}}; t), \tag{1.68}$$

with H_0 being the expansion rate today, Ω_m the matter energy density, c the speed of light and $a(\chi)$ the scale factor evaluated at the comoving distance χ.

1.3 Summary and discussion

We have developed in this chapter the historical framework that have yielded to the standard model of cosmology: the Λ-Cold-Dark-Matter model. We have exposed the different key observational discoveries, as well as the associated theoretical framework that have constructed a concordance model starting with a Big Bang followed by an inflation epoch and recently experimenting an accelerating period. This model is mainly composed today of two 'dark' components, the cold Dark Matter ($\sim 25\%$ of the universe energy budget) and the Dark Energy ($\sim 70\%$ of the energy content of the universe today). Even though observations seems to suggest that Dark Energy is another constant of nature, the so-called fundamental constant Λ, the community is not satisfied with this solution and current cosmological experiments aim to unveil the mystery behind these two dark components. This is the case of the Dark Energy Survey, a galaxy survey that will be presented in chapter 2.

[9] The conformal time is defined as $\eta(t) = \int \frac{dt}{a}$.

We have followed the discussion by presenting different tools currently used in cosmology (and particularly in the context of this thesis). Namely, we have started by introducing the tools used in cosmology in the context of Large Scale Structure analyses and, more in particular the statistics of galaxy clustering that will be specially used in chapter 3, where we use clustering information to infer redshift distributions. And afterwards we have presented the main concepts used in gravitational lensing analyses and in the analysis of the cosmic microwave background radiation (and its lensing) that will be used in chapter 4, where we detect the imprints of Large Scale Structures observed by the Dark Energy Survey in the CMB lensing map provided by Planck (Planck Collaboration et al. (2016c)).

Chapter 2

The Dark Energy Survey

2.1 History and goals

In the last decades, as it has been developed in the previous chapter, we have been the witness of tremendous improvement in our understanding of the Universe history and dynamics. Despite the fact that general relativity gave us a standard theory that allows us to describe our universe at large scales, the so called ΛCDM scenario (see sec. 1.1.7), some aspects of the theory remain to be understood, and in particular understanding the true nature of Dark Energy (whether it is a cosmological constant or not) is one of the biggest challenges of current cosmology. In this perspective, a variety of alternative theories have been proposed and one of the main goals is thus in one hand to test these alternative theories and on the other hand to put stronger constraints on the standard cosmological model.

In order to achieve this, it is possible to look at different probes, for example one can analyze the information contained on the cosmic microwave background temperature anisotropies at high redshift, $z \sim 1100$ (see sec. 1.2.4) such it has been done recently with the Planck satellite Tauber et al. (2010); Planck Collaboration et al. (2011), the South Pole Telescope (SPT) Ruhl et al. (2004); Padin et al. (2008); Carlstrom et al. (2011) or the Atacama Cosmology Telescope (ACT) Fowler et al. (2007), indeed looking at the positions and heights of the peaks in the power spectrum of the CMB temperature allows us to estimate cosmological parameters (1.2.4).

On the other hand, at lower redshift ($z \lesssim 2$) it is also possible, as developed in 1.2.2, to test cosmological models using the information one can extract from the large scale structure of the universe, its gravitational history and dynamics (see section 1.2). In order to satisfy the requirements in terms of precision to be competitive with CMB experiments, the community will need to analyze more **accurate** and more **complete** catalogs of this large structure. To collect such samples, two main categories of galaxy surveys have been (and will be) operating:

- **The Spectroscopic Surveys :** They are pointing at specific objects and extracting a precise spectral information for each of them. These types of surveys are providing us the most accurate 3D information of the observed objects, which make them particularly suited to extract the clustering information from galaxy samples. However, pointing to each object is time consuming and this will affect the completeness and size of the observed sample. A non exhaustive list of recent and future spectroscopic surveys is presented in Table 2.1.1.

- **The Photometric Surveys :** They are observing the sky through optical and NIR[1] filters, that is to say they are observing a field at a specific direction using a mechanism (the filters) that only observes the average spectral density within a range of wavelengths and repeat it for a variety of these broad-band filters. These classes of surveys have the advantage of being more complete with respect to the previous one, since they are not requiring to point at each object but will get information on all objects in the observed field (up to a magnitude limit). However, they suffer from the lack of spectral information that will affect their ability to determine the radial position of the objects. Another aspect that is worth to mentioned here is that these surveys can also provide information about galaxy shapes, which make them able to perform weak lensing analyses (see section 1.2.3). Table 2.1.2 is a summary of some recent or future photometric surveys.

This chapter will be dedicated to one of these new-generation galaxy surveys: the Dark Energy Survey (DES, The Dark Energy Survey Collaboration (2005)), which belongs to the category of photometric surveys. The chapter will start in section 2.2 by presenting the DES collaboration and describing the DES camera (DECam) as well as its photometric properties, while 2.3 will present the main galaxy catalogs used in the context of this thesis built from the observed samples after one year of observation of DES (DESY1 hereafter). And then section 2.4 will be dedicated to the main cosmological results recovered by the collaboration using DESY1 catalogs and more in particular the results from the combined probe analysis from DES Collaboration et al. (2017).

2.2 The Dark Energy Camera & the Dark Energy Collaboration

2.2.1 Survey overview

The Dark Energy Survey collaboration is an international community gathering almost 400 scientists spread over 25 institutions in 7 countries (United States, Spain, United Kingdom, Germany, Australia, Brazil and Switzerland). The common effort of this collaboration is to study the nature of dark energy by observing our universe at large scales (The Dark Energy Survey Collaboration (2005)). To do so, the collaboration has built a digital camera, the Dark

[1] NIR stands for Near Infrared.

Survey	Number of galaxies	area (deg^2)	Observation period
Sloan Digital Sky Survey (SDSS Legacy) York et al. (2000)	$\sim 1M$ galaxies ~ 120000 Quasars	7500	observation finished 2000-2008
Wiggle-Z Drinkwater et al. (2010)	239000 Emission Line Galaxies	1000	observation finished 2006-2011
The Baryon Oscillation Spectroscopic Survey (BOSS) Dawson et al. (2013)	1.5M Luminous Red Galaxies 160000 Quasars	10000	observation finished 2008-2014
The Extended Baryon Oscillation Spectroscopic Survey (eBOSS) Dawson et al. (2016)	300000 Luminous Red Galaxies 189000 Emission Lines Galaxies 573000 Quasars	7500 1000 7500	2014-2020
Dark Energy Spectroscopic Instrument (DESI) DESI Collaboration et al. (2016)	$4M$ Luminous Red Galaxies $18M$ Emission Lines Galaxies $2.4M$ Quasars	14000	First light in 2019
Euclid Laureijs et al. (2011)	50M galaxies	15000	Launching in 2021

TABLE 2.1.1 – Summary of recent and future spectroscopic surveys.

photometric Survey	photometric bands	Magnitude limit	Number of galaxies	area deg^2	observation period
Canada-France-Hawaii Telescope Legacy Survey (CFHTLS)	ugriz	$i_{AB} = 24.0$	-	4 fields over $171 deg^2$	observation finished 2004-2009
Kilo Degree Survey (KiDS) de Jong et al. (2015)	ugri	$i_{AB} \lesssim 24.2$	90M	1500	First light in 2011
Dark Energy Survey (DES) The Dark Energy Survey Collaboration (2005)	grizy	$r_{AB} \lesssim 24$	300M	5000	started in 2013 for 525 nights of observation
Hyper Suprime-Cam (HSC) Aihara et al. (2018)	grizy	$r \sim 26$	100M	1400	started in 2014 for 300 nights of observation
Physics of the Accelerating Universe Survey (PAU) Martí et al. (2014)	40 narrow bands	$i_{AB} \sim 23 - 24$	2M	100	started in 2015 123 nights of observation
Large Synoptic Survey Telescope (LSST) LSST Science Collaboration et al. (2009)	ugrizy	$I \lesssim 23.9$	4B	20000	First light 2021
Euclid Laureijs et al. (2011)	R,I,Z YJH	$Y, J, H \lesssim 24$	1.5B	15000	Launching in 2021
COSMOS-30 band) Laigle et al. (2016)	30 UV/Visible/IR bands	-	> 500000	2	Combined sample of various surveys in the COSMOS field

TABLE 2.1.2 – Summary of recent and future photometric surveys.

Energy Camera (DECam, Flaugher et al. (2015)) with a field of view of 2.2 degrees (which makes it suited for wide field observations).

The Dark Energy camera is a high-sensitivity digital camera (570 Megapixels) made of 74 state-of-the-art CCDs about 10 times larger than traditional CCDs. Among these CCDs, 62 are used for science imaging, 4 for guiding and 8 to focus, this allows it to detect long-wavelength photons. Beside the CCDs, the camera has been equipped with additional items, and the left panel of Fig.(2.2.1) is a schematic drawing of the different components of the camera, namely:

- A low noise readout system inside cooled crates, ((a) in the figure)
- 5 optical lenses: the largest optical corrector currently used in astronomy, (C1 to C5 in the figure),
- 5 optical filters: covering wavelenghts from 400 nm to 1080 nm, ((c) in the figure),
- An hexapod : a positioning and alignement system, ((d) in the figure),
- A shutter : which blocks the light from entering the imager between exposures, ((d) in the figure).

The camera has been built between 2008 and 2012 and installed at the prime focus of the Blanco 4-meter telescope at the Cerro Tololo Inter-America Observatory, Chile (CTIO) from February 2012 to September 2012 (see Soares-Santos et al. (2012)).

DECam official first light happened on September 12, 2012; Fig.(2.2.1) shows the 62 CCDs image from DES first light. After this, from November 2012 until February 2013, a first period of observations took place as proof of viability of the survey. The product of this observation period is called the Science Verification (SV) Data, covering a patch of about $150 deg^2$ of the sky (which represent less than 3% of the total expected DES final coverage), which permitted the collaboration to test the well functioning of the camera as well as making the first cosmological inference with it (see Abbott et al. (2016)).

2.2.2 Survey plan

The Dark Energy Survey was originally funded for 5 years of observations with 105 nights of observations per year (summing then 525 nights). As mentioned before, DECam is located in Chile, therefore it is observing at the southern hemisphere, and the expected total area spanned by the camera after the fifth year of observing represents about $5000 deg^2$ of the sky with a limiting depth of about $i_{AB} = 24$ [2]. DES is expected after these five years of observations to

[2]Note that due to some extremely poor weather in Y3, an extra year of observation have been accepted for funding and will start in september 2018

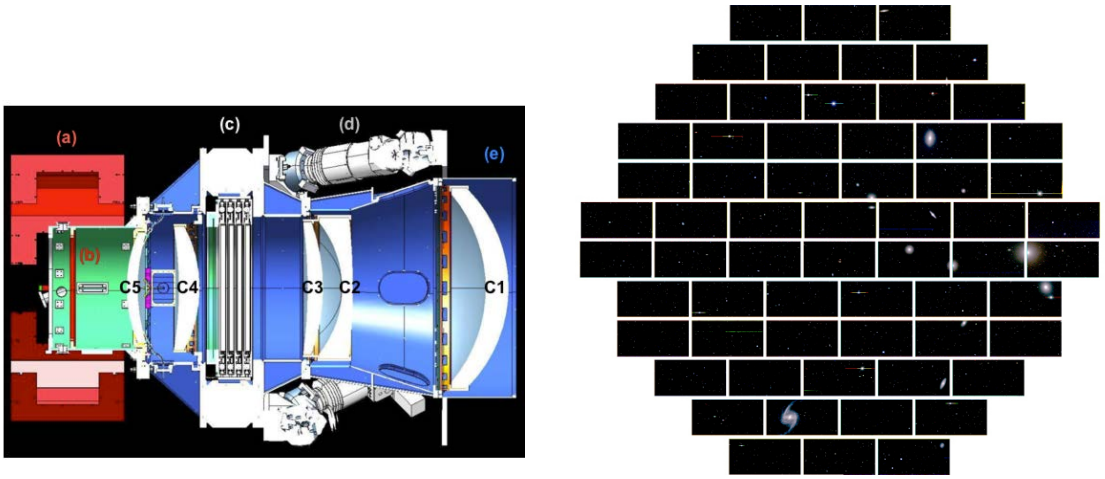

FIG. 2.2.1 – *Left panel:* Schematic image of the Dark Energy Survey camera, where $C1-5$ are the 5 lenses, the readout system (a), the image focal plane (b), the photometric filters (c), the hexapod positioning system (d), and the barrel (e); figure from (Soares-Santos et al., 2012). *Right panel:* Image from the 62 science imaging CCDs from the first light of DECam on September 12th, 2012

provide to the community information on about 300 Million galaxies up to redshift $z = 1.2$, as well as $100,000$ galaxy clusters and thousands of supernovae.

The DES final footprint also presents the advantage of having overlapping regions with other surveys that could be use either as a cross-check (or calibration) sample of our methodology (see for example 3.5.3) or to combine different probes. Fig.(2.2.2) shows the observing area of DES as well as the different overlaps it has with other galaxy surveys (spectroscopic and photometric). In particular, one can see the large overlap with CMB observations by the South Pole Telescope (SPT, with an overlapping area of $\sim 2500 deg^2$) shown in the 'photometric surveys' panel. This has been particularly thought for DES to give SPT complementary information about the redshift of the galaxy clusters SPT finds using the SZ effect.

2.2.3 DES, a photometric survey

2.2.3.1 Galaxy spectral information and spectroscopic redshifts

Before giving more details about the photometric properties of DES, it seems useful to first give a brief overview of galaxy spectral properties that will enter the analysis procedure. Namely, while observing an object, we are able to position it directly on the sky plane (right ascension (RA) and declination (DEC)), in order to get its radial position, one has to measure it indirectly. If one, for instance, had the full chemical composition, in principle one could infer the spectral features that will be absorbed (or emitted) from the object, what is commonly called the spectral energy distribution (SED). And from here it would be straightforward to infer its radial position by estimating the shift between the emitted spectral line and the observed one $1 + z = \frac{\lambda_{obs}}{\lambda_e}$.

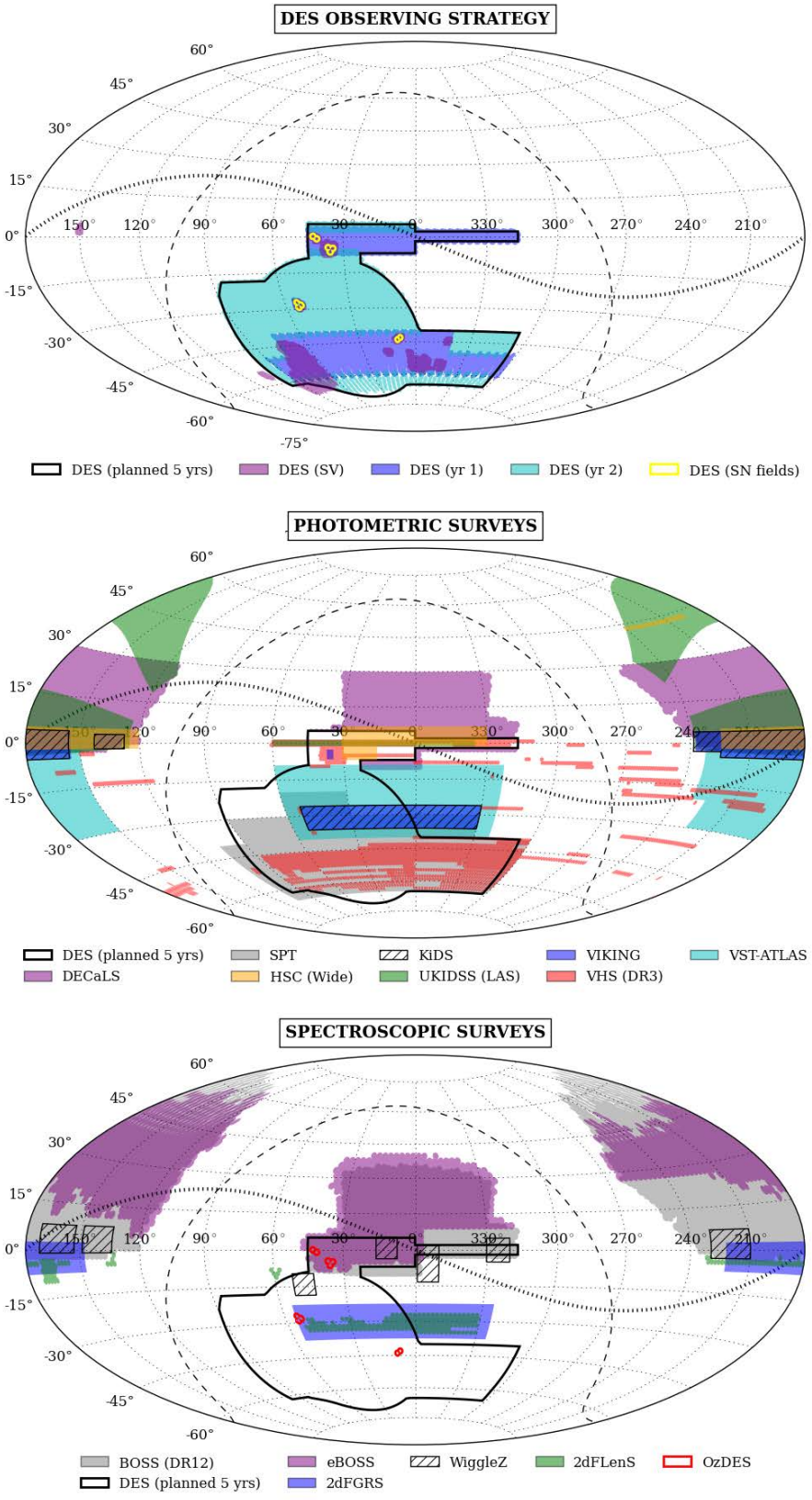

FIG. 2.2.2 – Observation area of DES (top panel), overploted with other recent surveys, photometric (middle panel) and spectroscopic (bottom panel), from Dark Energy Survey Collaboration et al. (2016)

By measuring these shifts on various spectral features, spectroscopic surveys manage to infer precise redshifts ($\frac{\sigma_z}{1+z} \sim 0.1\%$). However, as one can imagine, such measurements are time consuming, and, eventhough there exist techniques to collect spectra of different objects at the same time, the so called MOS (multi-object-spectrographs), some cosmological probes (such as lensing) require redshift information of large number of objects to be competitive in cosmological parameter inference that is not affordable with this type of instrument. However, in photometric data, we are getting these spectral information averaged over the wavelength range covered by each filter. In this sense, the key spectral features for this type of observations are no longer resolved spectral lines but the changes in the overall shape of the SED. For instance, when observing galaxies, two typical 'breaks' in the spectral shape can be identified [3]:

- Lymann break 912 Å: present in all galaxies, caused by hydrogen absorption.
- Balmer break $\sim$ 4000 Å: specially strong in old galaxies, which are typically redder and more luminous.

In the observed data, these two typical breaks will appear more shifted with respect to their initial positions as the redshift of the observed object is increasing. To illustrate this redshift dependence of the spectral feature of a given object, the black line on Fig.(2.2.3) shows a model of galaxy SED at different redshifts. In the figure, one can observe the shifts on the different spectral features with redshift (from bottom to top on the figure). The colored lines in the figure represent different photometric filters through which DES is observing. For this specific model, one can see the Balmer break crossing the filters ($g \longrightarrow r$, $r \longrightarrow i$, and $i \longrightarrow z$) at different redshifts.[4]

2.2.3.2 Photometry

Photometric surveys take observations of specific fields in the sky through different filters, which in a sense would be equivalent to low-resolution spectroscopy. Consequently, this would give the advantage of being much faster in observing than spectroscopic surveys. From these observations, different methodologies exist to extract redshift information from the fluxes observed through the different filters, the so-called photometric redshift (photo-z hereafter). These methods are based on the idea that one can model a relation between the radial position of the observed objects and their observed magnitudes (or colors). The width of the filters in wavelength space will affect the redshift characterization; indeed if one observes through 'narrow-band' filters, one will get more precision on the overall shape of the SED.

[3] Here we have only highlighted as examples two spectral features that are typically entering in the optical filter range.

[4] The galaxy modeled here is a red galaxy for which the Balmer break is pronounced.

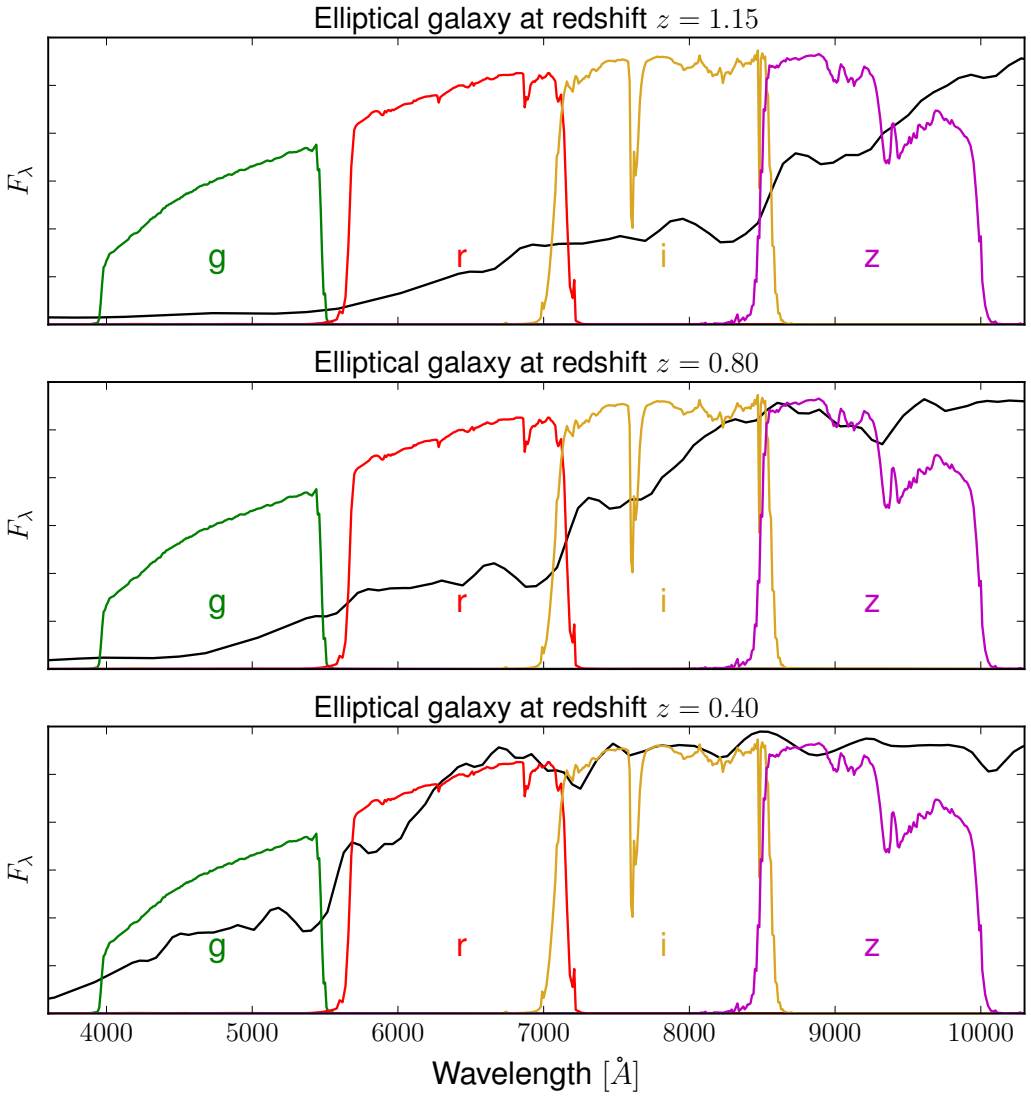

FIG. 2.2.3 – Illustration of the photometric properties of the g, r, i, z filters of the Dark Energy Survey camera DECam overplot with a model of spectral energy distribution of an elliptical red galaxy at redshift z = 0.40 (bottom panel), z = 0.80 (middle panel) and z = 1.15 (top panel); figure from Crocce et al. (2017)

Measuring the overall shape of the spectral energy distribution should allow us to infer in one hand the type of object being observed and on the other the distance at which we are observing it and this is the reason why the two breaks presented in the previous paragraph are playing an important role in photo-z measurements.

In practice, two popular categories of methods to infer single object photometric redshift can be distinguished:

- ***Template Based methods :*** For each galaxy, we observe a articular spectral energy distribution, whose shape depends on the type, age and redshift of the object. The main idea of this method is to use a set of templates of galaxy spectra (simulated or observed), from which we compute the flux distribution at different redshift observed through the same filters, and select the combination of parameters z, T, A that minimize a χ^2:

$$\chi^2(z, T, A) = \sum_{f=1}^{N_f} \left[\frac{f_{obs} - A f_{temp}(z)}{\sigma^f_{obs}} \right]^2 \tag{2.1}$$

  where N_f is the number of photometric filters, $f_{obs,temp}$ are the observed and template fluxes, σ_{obs} are the uncertainties on the observed fluxes, A is a normalization parameter, z is the redshift and T the galaxy type. Selecting the smaller χ^2 will give us the most probable redshift and type of the observed object. In this class of methods, it is also possible to use a Bayesian approach, namely instead of minimizing the χ^2 of eq.(2.1), using it in a probabilistic reasoning, inferring the probability for an object to be at a given redshift considering some measured values, and priors. This Bayesian method is for instance the one used in the context of DESY1 photometric redshift evaluation (see Chapter 3), inferred with the Bayesian Photometric Redshift (BPZ) algorithm (Benítez (2000)).

- ***Training methods :*** Another popular method used to infer photometric redshift is using a machine learning algorithm (ML) such as Neural Network. This is based on the idea that if we have a sample for which we have information in both spectroscopy and photometry, we can train our ML algorithm to find a relation between magnitude (or colors) and redshift. Note that our sample should be first divided into two sub-samples, one training sample (in which the neural network is being trained) and one calibration sample, used to test our training.

We note here that both methods will suffer in case of incompleteness in the template in one side or in the training set on the other. This will have serious consequences on the redshift resolution and bias one can achieve using photometry, namely the typical photo-z resolution of these methods is of the order $\frac{\sigma_z}{1+z} \sim 0.1$ and it is becoming a challenge for the community to

improve these techniques to achieve the requirements of the new-generation surveys in terms of photo-z biases and resolution.

2.3 DES photometry: the samples from the first year of observations

The Dark Energy Survey is providing photometry information observing through 5 different optical filters, g (400 – 550 nm), r (560 – 710 nm), i (700 – 850 nm), z (830 – 1000 nm) and Y (950 – 1065 nm). Fig.2.2.3 shows the efficiency of the different filters from DES used to infer photometric redshift and galaxy shapes[5].

In Year1 (Y1) DES observed through these filters on average three or four times with a typical exposure time of 90 seconds. The DESY1 footprint is about $1800 deg^2$, that is to say about 40% of the final DES footprint. The area covered by this observation is represented by the dark blue shaded region on the top panel of Fig.(2.2.2). The analysis presented here is based on this first year of observations, between August 31st 2013 and February 9th 2014. From this observation, the collaboration has extracted photometry and built a cleaned catalog of galaxies, these gold catalogs of 137 million objects Drlica-Wagner et al. (2018).

In the context of this work, two samples issued from the gold sample will be used, a high-precision photo-z sample used to extract information about the clustering of galaxies and a sample of galaxies with measured shapes, for weak lensing analyses. The following paragraphs will be devoted to describe these two science samples.

2.3.1 A high-precision redshift sample: *redMaGiC*

In order to mitigate the impact of redshift uncertainties in cosmological analyses, it is useful to build a sample by selecting galaxies for which these uncertainties are lower. For instance, when one wishes to extract clustering information, it is possible to define a sample that contains the clustering information and is not as affected by redshift uncertainties as the full gold sample of galaxies. One of the tools that have been used in DESY1 to achieve this is the *redMaGiC* algorithm Rozo et al. (2016). This algorithm permitted to select from the DESY1 gold sample luminous red galaxies (LRG hereafter) that have the advantage of allowing an accurate estimation of their photometric redshift ($\sigma_z/(1+z) \lesssim 0.02$) in the considered redshift ranges). This comes from the fact that this type of objects have a pronounced Balmer break (4000 Å) that allows a precise photometric redshift estimation. Moreover, LRG have shown to be really good candidates to probe large scale structure, in the sense that they can be observed until relatively high redshift with shallow exposures due to the fact that these objects have a high luminosity.

[5]Note that the Y band is not present here because it is mostly used for supernovae observations that will not be treated in this thesis.

The selection methodology followed by the *redMaGiC* algorithm can be described in three main steps :

- Make a fit for all the galaxies of the sample to a red-sequence template and compute from it the best redshift of the fit as well as the template goodness of fit χ^2. The red-sequence templates used in this fitting are typically generated by the *redMaPPer* cluster finder Rykoff et al. (2014).
- Compute the galaxy luminosity using the inferred redshift.
- Select galaxies with a specific luminosity and goodness-of-fit threshold ($L > L_{min}$ and $\chi^2 < \chi^2_{min}$).[6]

The idea being to select similar samples of galaxies at all redshifts, the value of χ^2_{min} will depend on the redshift range. This will produce a galaxy sample with constant comoving density.

In the context of DESY1 analysis, several *redMaGiC* samples have been built, varying the luminosity limits (L_{min}), and more in particular the sample used for cosmological analysis (see 2.4.2,2.4.4), is a combined version of various *redMaGiC* samples with different luminosity cuts depending on the redshift. Moreover, a mask based on the maximal redshift z_{max} to which LRG could be detected for each filter, and weights have also been assigned to each galaxy in the *redMaGiC* science sample to mitigate systematic errors Elvin-Poole et al. (2017).

The final *redMaGiC* sample for Y1 contains about 660000 red galaxies (the different luminosity cuts are in Appendix A) with measured redshift $z_{redMaGiC}$ with a typical error of $\sigma_z/(1+z) = 0.017$. For the cosmological probes that will be presented hereafter, the *redMaGiC* sample has been divided in five redshift bins ([0.15, 0.3], [0.3, 0.45], [0.45, 0.6], [0.6, 0.75], [0.75, 0.9]).

2.3.2 A sample for Weak lensing

The other galaxy catalogs that will be used in this thesis are the catalogs with shapes that entered the cosmological analysis using lensing effects Prat et al. (2017); Troxel et al. (2017); DES Collaboration et al. (2017). In these analysis, two catalogs have been created and used, the `METACALIBRATION` (Huff & Mandelbaum (2017); Sheldon & Huff (2017)) and the `IM3SHAPE` (Zuntz et al. (2013)) catalogs. These catalogs adopt different methodologies to infer the galaxies two component of ellipticities $\mathbf{e} = (\mathbf{e_1}, \mathbf{e_2})$ necessary for weak lensing analysis (see 1.2.3 for terminology) and to calibrate the different shear measurements and estimate the biases we have in these measurements. This bias will be estimated for each redshift bin i and is applied as a

[6]This limit can be tuned to the science application.

multiplicative parameter m_i with :

$$\gamma = (1 + m_i)\gamma_{true} \tag{2.2}$$

where γ and γ_{true} stand for the measured and true shear. This parameter will be considered as a free parameter in the DESY1 analysis and will be marginalized over (see 2.4).

The `METACALIBRATION` **catalog :** This catalog has been produced measuring first the ellipticities of the objects using the `NGMIX` pipeline[7] in which the basic idea it to fit a 2D Gaussian model to the pixel image to each object for all r, i, z band exposures convolved with the point-spread function (PSF)[8]. And in a second phase, to calibrate the measurements of the mean galaxy shear, the `METACALIBRATION` methodology has been applied. `METACALIBRATION` is a self-calibration method that allows to estimate the shear response R_γ and selection response R_S of the estimator, using directly the observed images. R_γ corresponds to the bias of our estimator when infering the correct ellipticity of an object. In order to evaluate this bias, the method applies an artificial additional small (known) shear to each objects and re-evaluate the ellipticity of the artificially sheared objects (evaluating thus the estimator response to shear). By expanding the estimate of the ellipticity in a Taylor series, one gets :

$$\mathbf{e} = \mathbf{e}|_{\gamma=0} + \left.\frac{\partial \mathbf{e}}{\partial \gamma}\right|_{\gamma=0} \gamma + \dots \tag{2.3}$$

where γ stands for the shear of the object, and from here one can then define the shear response as $R_\gamma = \left.\frac{\partial \mathbf{e}}{\partial \gamma}\right|_{\gamma=0}$ and estimate it using the artificially sheared images as :

$$R_{\gamma,i,j} = \frac{e_i^+ - e_i^-}{\Delta\gamma_j}, \tag{2.4}$$

where the $e_i^{+,-}$ stand for the artificially sheared measured ellipticity of the component $i = 1, 2$ of the shear, and $\Delta\gamma$ is the artificial shear applied. In the case of DESY1, $\Delta\gamma = 1\%$ has been chosen.

On the other hand, R_S, the selection response, is the bias induced by the fact that we are selecting galaxies (eg. S/N selection, redshift selection,...) and these different potential selections might have an impact on the measured mean ellipticities, namely if the probability of selecting in our sample a galaxy with an ellipticity $\mathbf{e}$ is higher than for another ellipticity, this will bias our mean shear measurement. In `METACALIBRATION`, the estimation of this bias is done by estimating the shear response when applying the selection to the artificially sheared sample and measuring the

[7] https://github.com/esheldon/ngmix.

[8] The PSF represents the projection of a point source in the observed field into the image.

response to these selection effects in the original catalog :

$$< R_S >= \frac{< e_i >^{S+} - < e_i >^{S-}}{\Delta\gamma_j} \tag{2.5}$$

Where $< e_i >^{S+,-}$ are the mean measured ellipticity on the original images having performed the selection on the artificially sheared sample.

The full response of the METACALIBRATION catalog is then given by the sum of these two responses :

$$S =< R_\gamma > + < R_S > \tag{2.6}$$

Note that this methodology is measuring the weak lensing shear directly on observed images and does not require any priors on galaxy properties from simulations. However, it has been tested in simulated images and has been shown to be able to recover the shear with a precision better than a part in a thousand Sheldon & Huff (2017). The final catalog contains 35 million objects with shapes, and the shear calibration bias, evaluated over different systematic effects has been estimated as $m = 0.012 \pm 0.013$. This systematic is shown to be dominated by blending effects in the catalog [9].

The IM3SHAPE catalog : The IM3SHAPE pipeline is fitting bulges or elliptical models of galaxies to observed objects for all exposures and picking the best fit in the r-band. The code is based on the fitting of six different components, the two ellipticity components $\mathbf{e_{1,2}}$, the half light radius $\mathbf{r}$, the centroid offset $(\mathbf{x_0}, \mathbf{y_0})$ and the amplitude $\mathbf{A}$ of the different observed objects. The calibration of the IM3SHAPE catalog of Y1 has been done using simulated images built to mimic the data sample (Zuntz et al. (2017); Samuroff et al. (2018)). From this study, the shear bias has been evaluated as a two component bias (multiplicative m and additive c_i) that should be applied to each object as:

$$e_i^0 = (1 + m)e_i^t + c_i \tag{2.7}$$

where e_i^0 stands for the observed component i of the ellipticity and e_i^t its true one. These two components have been estimated accounting for selection effects on the catalog, meaning that they depend on any additional cuts that could be made in the catalog. The resulting catalog after applying selection cuts includes 22 million objects, and a multiplicative bias for the shear component of $m = 0.0 \pm 0.025$, which is dominated this time by defects in the simulated images. The final catalog of IM3SHAPE is slightly smaller than the METACALIBRATION one due to the fact the pipeline has been applied only to the r-band images.

Note that these two catalog employ different shape inference methodology and there was no attempt to make them look alike, the objective being to enforce the robustness of our cosmological measurements if the results obtained with both are consistent.

[9] Blending effects effers to the effect of neighbouring objects on the shear measurements.

Moreover, in the context of DESY1 analysis, additional cuts on the catalogs have been made[10]:

- Selecting galaxies having photo-z point estimate within the redshift range $0.2 < z < 1.3$.
- Selecting galaxies that ie in the *redMaGiC* mask: in principle, this would not be necessary for cosmic shear analysis (2.4.3), but it has been done in order to be consistent with the two other 2-point statistics used in the combined analysis.

These additional cuts lead to final catalogs of 26 million galaxies with a galaxy density of 5.5 galaxies $arcmin^{-2}$ for the `METACALIBRATION` catalog and 18 million galaxies with 3.4 galaxies $arcmin^{-2}$ for the `IM3SHAPE` catalog, for a final area for DESY1 analysis of $1321 deg^2$. These catalogs have been divided in four redshift bins ([0.2, 0.43], [0.43, 0.63], [0.63, 0.9], [0.9, 1.3]) and the redshift distribution has been evaluated using a Bayesian template fitting code (BPZ Benítez (2000)) calibrated with two additional techniques, clustering redshift and 30 band photometry. This calibration process will be presented in chapter 3.

2.4 DES first year main results

2.4.1 Cosmological inference methodology

Using the catalogs presented above, various cosmological probes allowed the collaboration to put new constraints on the cosmology of our universe. This section aims to give an overview of the main results that came out from one of these combined probe analysis. More specifically, after an introduction to the overall methodology applied for cosmological inference, we will present the results obtain for the three probes that entered the combined cosmological analysis for DESY1 (Elvin-Poole et al. (2017); Troxel et al. (2017); Prat et al. (2017); DES Collaboration et al. (2017)): the **galaxy clustering** signal (2.4.2), the **cosmic shear** signal (2.4.3) and the **galaxy-galaxy lensing** signal (2.4.4), as well as the results combining these three probes, the **3x2 point** analysis (2.4.5).

These analyses have used the 2-point statistics presented in 1.2.2 computed in all possible pairs of redshift bins of the sample presented in 2.3. The angular correlation functions have been estimated in 20 log-spaced angular bins between $2.5' < \theta < 250'$[11], using the public code `tree-corr`[12] Jarvis et al. (2004) . It is also good to mention that all the analyses presented in the following paragraphs have been at first realized in a "blind" context, meaning that the correlation

[10]Note that the selection response mentioned for `METACALIBRATION` and the two calibration components of the `IM3SHAPE` catalog had to be estimated considering theses additional cuts.

[11]These scales have been selected in order to mitigate contributions due to baryonic interactions at small scales and residual mean shear bias at large scales.

[12]https://github.com/rmjarvis/TreeCorr .

functions have been at first plotted in different axes for the measured and theoretical signals to avoid any confirmation bias. Furthermore, the values of the shear were scaled by an unknown factor. The unblinding occurred in July 2017 and was followed by the DESY1 publications.

In order to estimate cosmological parameters, one should at first consider a model for the measured signal and in then sample the different model parameters on the measurement, obtaining then a set of best fit parameters. For each probe, one would then have to sample in the multi-parameter space the likelihood function $\mathcal{L}$ assumed to be gaussian :

$$log\mathcal{L}(\mathbf{p}) = -\frac{1}{2}\sum_{i,j}\left(D_i - T_i(\mathbf{p})\right)C_{ij}^{-1}\left(D_j - T_j(\mathbf{p})\right) \tag{2.8}$$

where T_i is the theoretical model considered for the redshift bin i, D_i is the measurement done in this bin and C_{ij} is the covariance of the measurement in each angular and redshift bins. This likelihood, is thus sampled over the set of parameters $\mathbf{p}$. These parameters can be divided in two main categories, the **nuisance** parameters, that account for the different uncertainties or artifacts that enter the measurements (systematic or astrophysical), and the **cosmological** parameters that are the parameters we wish to estimate. While the cosmological parameters are the ones ruling the universe evolution and dynamics, the nuisance parameters are observational defects or spurious correlation that will appear in our measurement and have to be estimated or marginalized over. These nuisance parameters depends on the probe used, and in the combined DESY1 analysis a total of 20 nuisance parameters had been considered. Beside these nuisance effects, two different cosmological models have been probed, the standard ΛCDM model which can be describe with six cosmological parameters, the three energy densities parameters (Ω_m,Ω_b,Ω_ν)[13], the value of the Hubble expansion today H_0, the amplitude of the primordial density fluctuations A as well as its spectral index n_S. The amplitude parameter is often convert to the σ_8 parameter that measures the amplitude of the matter power spectrum at $8Mpc/h$. In this specific analysis, we will be more sensitive to the S_8 parameter which can be related to the clustering amplitude and the matter energy density as :

$$S_8 = \sigma_8\left(\frac{\Omega_m}{0.3}\right)^{0.5}. \tag{2.9}$$

The other model sampled is the $wCDM$ model in which the dark energy is allowed to vary with redshift (not a cosmological constant anymore) adding one cosmological parameter to the parameter budget, the equation of state of dark energy w[14]. The cosmological parameters have been sampled in this analysis using flat priors covering ranges much wider that recent results obtained by other observations.

[13]Note that here the sum of neutrino masses has been sampled, even though most other cosmological analyses are commonly fixing it to the limit value of $0.06eV$

[14]The prior used to sample this parameter is a flat prior excluding $w > -1/3$, which would make the model inconsistent with late-time cosmic acceleration.

The theoretical models have been computed using two different algorithms, `CosmoLike` (Eifler et al. (2014)) that uses the `CLASS` (Lesgourgues (2011); Blas et al. (2011)) pipeline to compute the matter power spectrum and `CosmoSIS`[15] (Zuntz et al. (2015)) using `CAMB` (Lewis & Challinor (2011)). And the parameter space has been sampled using both the `MultiNest` (Feroz et al. (2009)) and the `EMCEE` (Foreman-Mackey et al. (2013)) samplers.

2.4.2 Clustering measurement for combined probes

The more straightforward analysis that can be done using these wide galaxy catalogs is to extract information from the clustering of galaxies using the auto (or/and cross) angular correlation signal of the galaxy positions. If one considers that galaxies are tracers of the underlying matter density field, these correlation signals would give us direct information the on matter distribution (baryonic and dark) in the considered bins. However, as it has been mentioned in section 1.2.2.3, galaxies are biased tracers of the underlying matter density field ($\delta_m = b\delta_g$); therefore, the measured correlation signals, besides depending directly on the clustering amplitude (σ_8), will also depend on the galaxy bias of the considered sample b, and the signal will scale as $\propto b^2\sigma_8^2$. Moreover, this bias is highly dependent on the sample considered, for this reason in cosmological analysis, it is often treated as a nuisance parameter and marginalized over. However, the galaxy-galaxy signal (that will be presented in 2.4.4) presents a different dependence on the galaxy bias, so it is possible to break this degeneracy using these two probes together.

The result of the galaxy clustering analysis for DESY1 is presented in Elvin-Poole et al. (2017). In there the galaxy sample considered is the combined *redMaGiC* sample presented in 2.3, divided in the five bins in redshift of width $\Delta z = 0.15$ going from $z = 0.15$ to $z = 0.9$. The 2-point angular correlations have been estimated using the Landy-Szaley estimator (see 1.2.2.5), taking previously into account additional correlations that could appear due to astrophysical and noise effects, see Elvin-Poole et al. (2017).

The theoretical formula used to model this correlation signal is:

$$w^i(\theta) = (b^i)^2 \int \frac{dl}{l} 2\pi J_0(l\theta) \int d\chi \frac{[n^i_{lens}(z(\chi))]^2}{\chi^2 H(z)} P_{NL}\left(\frac{l+1/2}{\chi}, z(\chi)\right), \tag{2.10}$$

where J_0 is 0th order Bessel function, χ stands for the radial comoving distance, n^i_{lens}[16] is the redshift distributions of galaxies in the lens bin i, H(z) is the Hubble expansion rate at redshift z and P_{NL} is the 3D matter power spectrum evaluated at the wave number $k = (l+1/2)/\chi$ and at cosmic time $z(\chi)$.

[15] https://bitbucket.org/joezuntz/cosmosis/wiki/Home .

[16] Here we have called the galaxies lenses to be consistent with the galaxy-galaxy terminology where these same galaxies have been used as lenses.

Although one cannot estimate cosmological parameters without making assumption about the bias of the sample used, it is possible by fixing the cosmological parameters (for instance to the ones inferred in the combined cosmological analysis DES Collaboration et al. (2017)), to infer the value of this bias at each redshift bin. This has been done and the results have shown to be consistent with the biases inferred using galaxy-galaxy lensing (see section 2.4.4 and Fig.(2.4.1)).

2.4.3 Cosmic Shear

The deflection of light that travels from background galaxies to us due to the underlying matter field it crosses makes the observed shapes and sizes different from the original ones due to lensing effects, as presented in 1.2.3. In principle, if one knew the intrinsic shapes and sizes of the background objects, one could have direct information on the matter between us and the background objects by measuring the amount of distortions they suffer. This is, of course, impossible. However, if one assumes the mean ellipticity of galaxies to be randomly oriented ($< e >\sim \gamma$) [17], one can infer the two components of these distortions averaging over a large number of sources, and the alignment of these distortions will be directly dependent on the foreground matter field. It is then possible to use the 2-point statistics auto-correlating the measured distortions to obtain information on the matter distribution. This analysis is called **cosmic shear**. Note that this method has a strong advantage compared to the clustering measurement presented in the last section, in the sense that it does not depend on the galaxy bias. The amplitude of the cosmic shear signal is at first approximation sensitive to the matter fluctuation σ_8 and the matter density Ω_m (at low redshift in the linear regime the scaling is about $\propto \Omega_m^2\sigma_8^2$). The recent cosmological inference from the KiDS survey (Hildebrandt et al. (2017); Köhlinger et al. (2017)) appeared to be in tension with the CMB high redshift results from the Planck satellite (Planck Collaboration et al. (2016b)) at the level of $2-3\sigma$ disagreement.

In this paragraph we will present the results from the DESY1 cosmic shear analysis, with a significance of 25.4σ (signal-to-noise) for the fiducial (`METACALIBRATION`) catalog of galaxy shapes. This is the highest significance signal obtained in today cosmic shear analyses. As matter of robustness, the analysis has been repeated for the two catalogs of galaxy shapes presented in sec. 2.3.2, even-though the `METACALIBRATION` catalog has been taken as fidutial to provide to the community the final DESY1 cosmological results.

As described before, the cosmic shear measurement is done by cross-correlating the two components of the shear γ_t and $\gamma_\times$ (see sec. 1.2.3). These correlation can thus be expressed as:

$$\xi_{+/-} =< \gamma_t\gamma_t > \pm < \gamma_\times\gamma_\times > . \tag{2.11}$$

[17] Ignoring shape correlations from galaxies evolving in the same environment, the so called Intrinsic Alignment effects (IA) that have been treated as a nuisance parameter.

Note that in theory, we should also consider the cross-correlation $< \gamma_t \gamma_\times >$, but in practice this terms vanishes (Schneider et al. (2002)).

The estimator to measure the cosmic shear signal for a pair of galaxy bins i and j is :

$$\hat{\xi}^{ij}_{+/-}(\theta) = \frac{\sum_{ab} W_a W_b \left[\hat{e}^i_{a,t}(\vec{\theta}\hat{e}^j_{b,t}(\vec{\theta}) \pm \hat{e}^i_{a,\times}(\vec{\theta}\hat{e}^j_{b,\times}(\vec{\theta})\right]}{\Sigma_{ab} W_a W_b S_a S_b}, \tag{2.12}$$

where $W_{a,b}$ are weights depending on the shape algorithm[18], S is the shear response correction mentioned in 2.3, defined as $S = R_\gamma + R_S$ for `METACALIBRATION` and $S = 1 + m$ for the `IM3SHAPE` ones. There it has been assumed that the shear response does not depend on the environment and thus the error to apply when one estimates the auto-correlation signal is the square of the mean response S (see Sheldon & Huff (2017)).

Considering a flat universe, the correlation signal between bins i and j can be modeled using the theoretical power spectrum as:

$$\hat{\xi}^{ij}_{\pm}(\theta) = \int \frac{dll}{2\pi} j_{0/4}(\theta l) \int d\chi \frac{q^i_s(\chi) q^j_s(\chi)}{\chi^2} P_{NL}\left(\frac{l+1/2}{\chi}, z(\chi)\right), \tag{2.13}$$

here we kept the same notation than in 2.4.2 and :

$$q^i_s(\chi) = \frac{3\Omega_m H_0^2}{2} \frac{\chi}{a(\chi)} \int_\chi^{\chi_0} d\chi' n^i_s(z(\chi')) \frac{dz}{d\chi'} \frac{\chi' - \chi}{\chi'} \tag{2.14}$$

where Ω_m and H_0 are the mater density and expansion rate (at $z = 0$) cosmological parameters, and n^i_s stands for the redshift distribution of the source sample in the redshift bin i.

Here we also have to take into account the shear calibration systematic presented in 2.3.2:

$$\xi^{ij} = (1 + m^i)(1 + m^j)\xi^{ij}_{true}. \tag{2.15}$$

Sampling the likelihood of eq. (2.8), it is possible, using cosmic shear measurements, without having to make any assumption on the galaxy bias, to estimate the cosmological parameters. Cosmic shear being mostly sensitive to the clustering amplitude (S_8 parameter) and the matter parameter (Ω_m), it is possible to make an estimation of these parameter marginalizing over the other ones. The results obtained in Troxel et al. (2017) are, for ΛCDM, $\sigma_8(\Omega_m/0.3)^{0.5} = 0.782^{+0.027}_{-0.027}$ and $\Omega_m = 0.260^{+0.065}_{-0.037}$. These values have improved by a factor of three the results obtained in the cosmic shear measurement made with the science verification sample (Abbott et al. (2016)). In the top panel of Fig.(2.4.1) is shown the parameter inference in the $(S_8 - \Omega_m)$ plane for the different DES shape catalogs. As it can be seen in the figure, the two shape catalogs

[18]For the `METACALIBRATION` catalog, no galaxy weight has been used W=1, and for the `IM3SHAPE` catalog the weights have been evaluated empirically

give consistent results and the METACALIBRATION catalog obtains stronger constraints on the considered parameters.

As mentioned before, this analysis has also been repeated for a model allowing variation of Dark Energy with redshift (not anymore considering dark energy as a constant as in the standard model), the so called *wCDM* model. Using the methodology and catalogs described above, the constraints on cosmological parameters using cosmic shear are $\sigma_8(\Omega_m/0.3)^{0.5} = 0.777^{+0.036}_{-0.038}$ and the equation of state $w = -0.95^{+0.33}_{-0.39}$. The equation of state parameter remains thus consistent with -1 within errors, not providing thus indication of varying Dark Energy.

2.4.4 Galaxy-galaxy lensing

The last probe used in this analysis is what is known as **galaxy-galaxy lensing**. The idea behind this probe consists in getting information on the matter density field by cross-correlating the shape of a sample of background galaxies (called the sources) with the position of a sample of foreground galaxies (acting as lenses). In principle, the distortion of background galaxies is related to the foreground galaxies mass and therefore one can relate this distortion to the position of a sample of foreground galaxies[19]. The background galaxies will appear orientated perpendicularly to the line connecting the lense to the source. So the *tangential* component of the ellipticity will be sensitive to the foreground mass distribution.

However, as it was the case for the clustering measurement, because the analysis uses galaxies as matter tracer, this signal will also depends on the bias of the lens sample used, and will then scale as $\propto b\sigma_8^2$. But if one uses the same galaxy samples (therefore the same galaxy bias parameter) to measure both signals from clustering and galaxy-galaxy lensing, since both depend on the same bias, it is possible to remove the bias dependency and get an estimation of σ_8.

For this measurement one can model the tangential shear signal measured between the lens bin i and the source bin j as :

$$\gamma_t^{i,j}(\theta) = b^i \int \frac{dll}{2\pi} J_2(l\theta) \int d\chi n_l^i(z(\chi)) \frac{q_s^j(\chi)}{H(z)\chi^2} P_{NL}\left(\frac{l+1/2}{\chi}, z(\chi)\right) \tag{2.16}$$

using the same notation than for the cosmic shear case, and considering the shear calibration systematic, we have now :

$$\gamma_t^{ij} = (1+m^j)\gamma_{true}^{ij}. \tag{2.17}$$

In Prat et al. (2017), the galaxy sample used for this analysis contains the same five redshift bins of the *redMaGiC* sample used in 2.4.2 and presented in 2.3.1 as lenses, and the four redshift bins of the shape catalogs used in 2.4.3 as sources.

[19]Considering scales larger than the size of galaxy haloes.

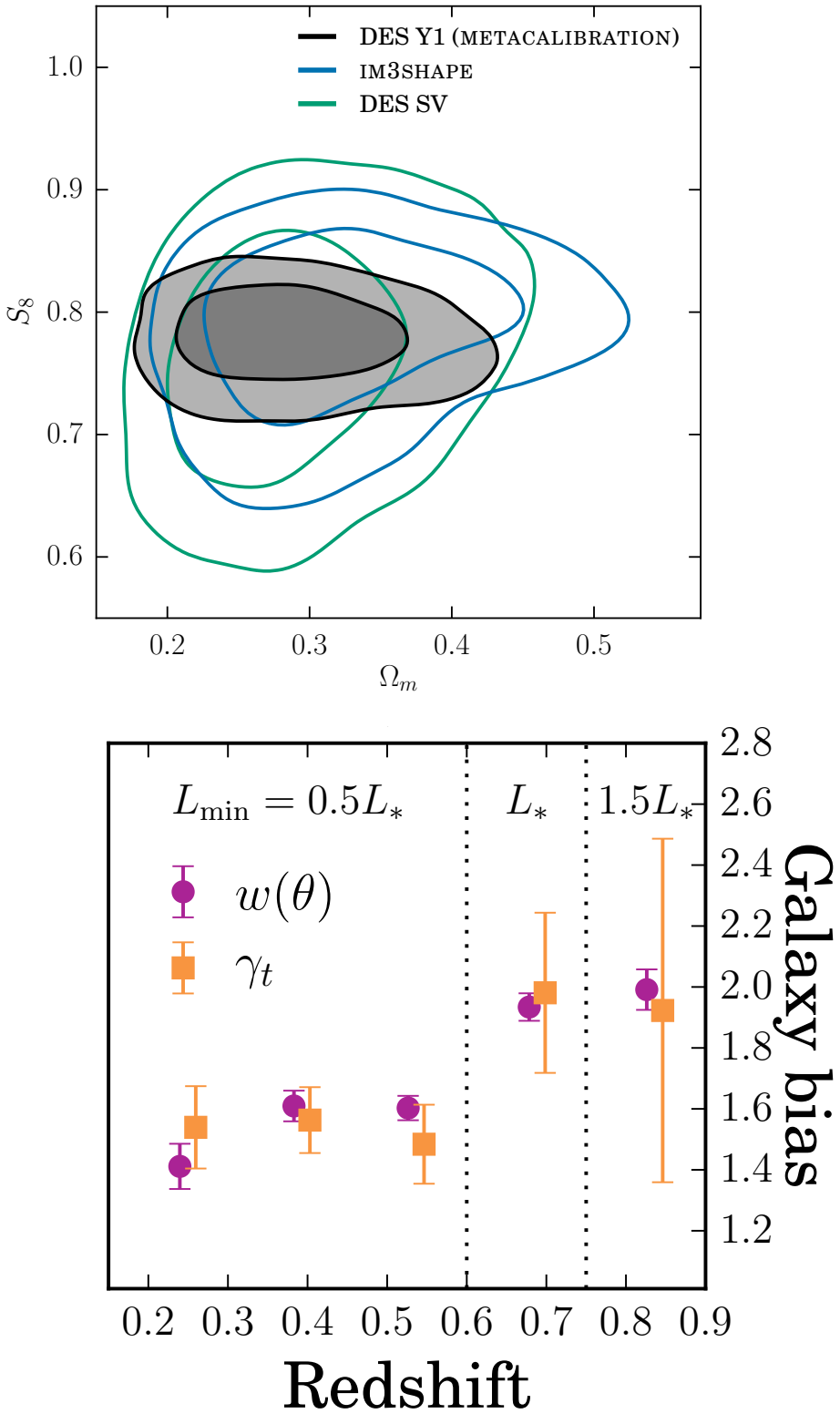

FIG. 2.4.1 – *Top panel:* Cosmological constraints on the S_8 and σ_8 parameters from cosmic shear measurement only comparing the three currently existing galaxy catalogs with shape of DES, the Science Verification catalog, the DESY1 `METACALIBRATION` catalog and the DESY1 `IM3SHAPE` catalog, figure from Troxel et al. (2017). *Bottom panel:* comparison of the galaxy bias estimation for the five *redMaGiC* lens bins using both the clustering (pink) and galaxy-galaxy lensing measurements (orange), figure from Elvin-Poole et al. (2017).

The estimator for tangential shear γ_t depends on the catalog used, since different corrections should be applied (Zuntz et al. (2017)). For the `METACALIBRATION` catalog, the estimators is :

$$< \gamma_t >= \frac{1}{S} \frac{\sum_j w_{l,j} e_{t,j}}{\sum_j w_{l,j}} \tag{2.18}$$

and for the `IM3SHAPE` catalog :

$$< \gamma_t >= \frac{1}{S} \frac{\sum_j w_{l,j} w_{s,j} e_{t,j}}{\sum_j w_{l,j} w_{s,j} (1 + m_j)}, \tag{2.19}$$

where the $w_{l,j}$ and $w_{s,j}$ correspond to weights associated to each pair source-lens.

As mentionned before, using galaxy-galaxy lensing also allows to estimate the bias of the sample in question for each of the redshift bin (fixing cosmological parameters). This bias measurement is shown to be in agreement with galaxy clustering and a comparison of these measurements is shown in the bottom panel of Fig.(2.4.1) for each of the five lens bins. As one can see in the figure, both bias estimations (using clustering and galaxy-galaxy lensing measurements) agree within errors.

2.4.5 Combined cosmological analysis

Now that the three probes have been measured, it is also possible to obtain constraints on cosmology by combining these three two-point statistics (3×2 point). Combining probes has shown in past analysis to improve cosmological parameter estimation, see for instance van Uitert et al. (2018) and Joudaki et al. (2018), where the data from the Kilo Degree Survey has been combined to two different spectroscopic external datasets as lens samples. This has been done in DES Collaboration et al. (2017), where the cosmic shear measurements together with the clustering and galaxy-galaxy lensing measurements have shown to give strong constraints on cosmological parameters. In there, the likelihood of eq. (2.8) has been sampled over twenty nuisance parameters plus six (or seven for $wCDM$ models) cosmological parameters. And the covariance matrix used for this estimation is a halo model theoretical covariance of dimension 457×457 (457 being the size of the final DESY1 data vector) generated by the `CosmoLike` pipeline Krause et al. (2017) [20].

On the top panel of Fig.(2.4.2) we can see the comparison of the results obtained whether one combines the different probes or not. These results have been obtained after marginalizing over the additional parameters. As it can be seen in the figure, the results here are showing a good agreement between clustering statistics and lensing statistics in cosmological parameter inference. Moreover, the combination of the three different probes shows to improve our constraints

[20]The analysis has been reproduced using the `CosmoSIS` pipeline and no discrepancy has been found.

on cosmological parameters. The main constraints obtained by combining the three two-point statistics are for ΛCDM: $\sigma_8(\Omega_m/0.3)^{0.5} = 0.783^{+0.021}_{-0.025}$ and $\Omega_m = 0.264^{+0.032}_{-0.019}$. Moreover, similarly to the cosmic shear analysis, the parameter estimation has also been repeated for the $wCDM$ model, and this time the best fit parameters are $\sigma_8(\Omega_m/0.3)^{0.5} = 0.794^{+0.029}_{-0.027}$, $\Omega_m = 0.279^{+0.043}_{-0.022}$ and the equation of state $w = -0.80^{+0.20}_{-0.22}$. Once again, no evidence for varying dark energy has been found.

Another important results from DES Collaboration et al. (2017) is the comparison of the DESY1 cosmological inference to the one announced by other surveys and probes. In particular, while the KiDS weak lensing analyses (Hildebrandt et al. (2017); Köhlinger et al. (2017)) were showing discrepancy in the S_8 inference with the Planck cosmic microwave background results (Planck Collaboration et al. (2016b)), this discrepancy is not so clear in the DESY1 results. In particular, the bottom panel of Fig.(2.4.2) shows the 68% confidence levels of S_8 and Ω_m in the ΛCDM parametrization for DESY1 alone and combined to external datasets. Thus the blind cosmological analysis made by DESY1 studying large scale structures at low redshift are in fair agreement with the CMB measurements at $z \sim 1100$.

The analysis can also be followed combining the DESY1 3×2 point analysis to the Planck results Planck Collaboration et al. (2016b) as well as the Baryon acoustic oscillation measurements from Beutler et al. (2011); Ross et al. (2015); Alam et al. (2017), and Supernovae luminosity distance from the Joint Lightcurve Analysis (JLA) from Betoule et al. (2014), the different ways to combine these probes and the parameter constraints obtained are represented in Fig.(2.4.2). Namely, for ΛCDM scenario, combining all probes named here, we have obtained $\sigma_8(\Omega_m/0.3)^{0.5} = 0.799^{+0.014}_{-0.009}$ and $\Omega_m = 0.301^{+0.006}_{-0.008}$ which are the tightest constraints made on these parameters up do date. And similarly, the $wCDM$ model results have not shown any indication of varying Dark Energy ($w = -1.00^{+0.04}_{-0.05}$).

2.5 Discussion

In this chapter, we have been presenting the Dark Energy Survey together with the results it obtained using three different 2-point statistics probes after one year of observations. The Dark Energy survey is entering now in its last observation period and has already shown, after this first observation year, to be competitive in cosmological inference. More in particular, with its photometric camera observing through its five optical filters, the Dark energy survey has been able to constrain cosmological parameters by probing large scale structure at low redshift at similar statistical level (for the first time) to the one made using the high redshift Cosmic Microwave Background information and has obtained constitent results with this last one.

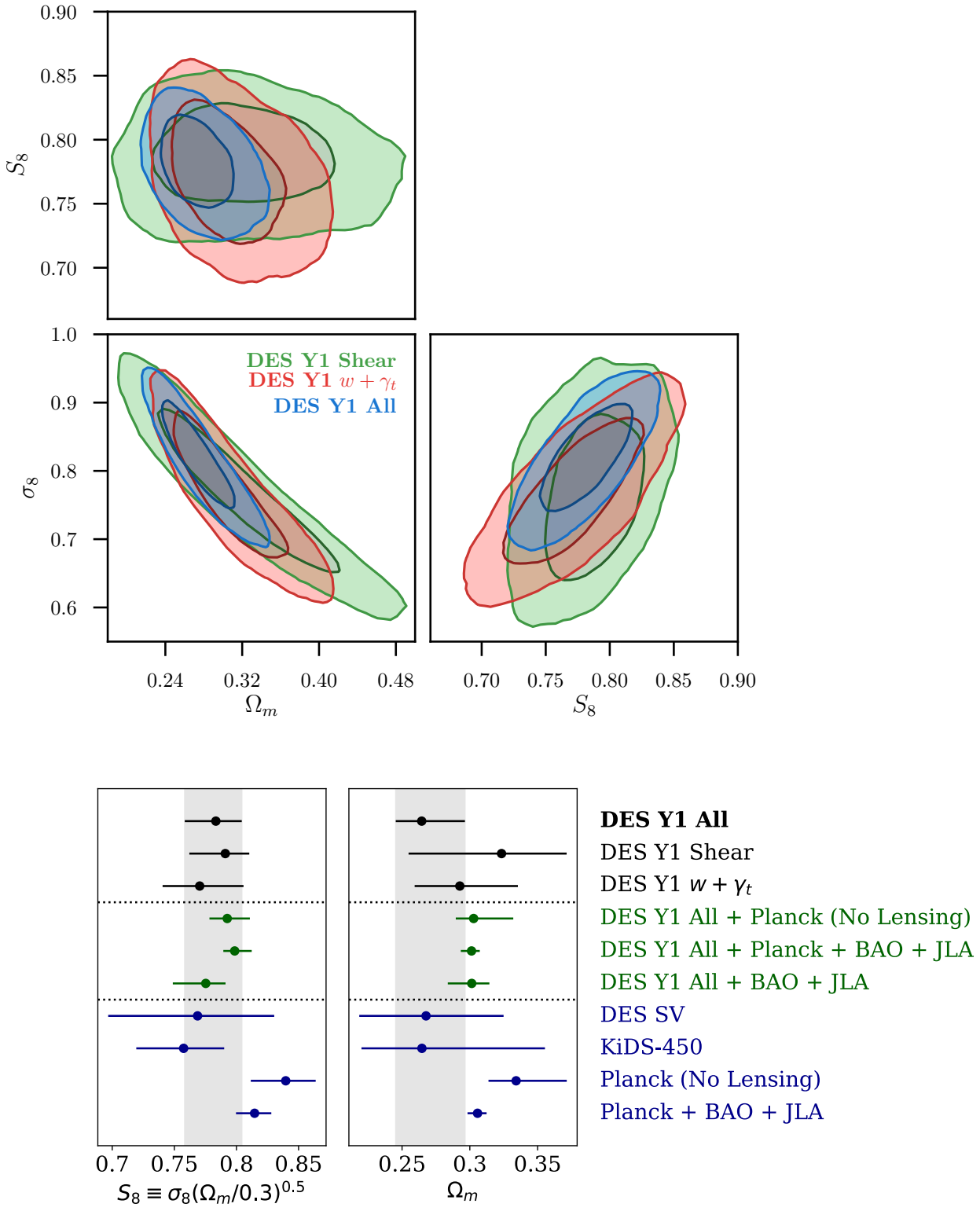

FIG. 2.4.2 – Main results from the DESY1 3×2 point combined constraints on ΛCDM models DES Collaboration et al. (2017), *Top panel :* constraints obtained on the $\sigma_8 - \Omega_m$, $S_8 - \Omega_m$ and $\sigma_8 - S_8$ parameter planes using DESY1 cosmic shear measurements (green), combining clustering and galaxy-galaxy lensing measurements (red) and all three probe combined (blue), the contour representing the 68% and 95% confidence levels. *Bottom panel:* Constraints obtained at the 68% confidence level on the S_8 and Ω_m parameters using various cosmological probes and experiments.

To obtain such results, the collaboration had to make a big effort in processing the data. This effort has been done all in a blind context probing various set of catalogs, sampler algorithms, photometric redshift algorithms. Beside highlighting the robustness of the results showing the consistency of them on the different catalogs or pipelines used, the fidutial `METACALIBRATION` catalog has been able to provide first the bigger detection of cosmic shear signal ever made and combined with other probes and external experiments data the tightest constraints on the $S_8 - \Omega_m$ cosmological parameters.

Moreover, the analysis has also been repeated for alternative model to the ΛCDM namely the $wCDM$ type of model, in which Dark Energy is not a cosmological constant anymore but can vary with redshift, and the results have shown no evidence for such variations (equation of state of Dark energy consistent with -1). This analysis also has open the road to further alternative model testing, indeed considering the large variety of alternative model existing in the literature, it would be also useful to implement such models in the analysis pipeline.

In the future months, the Dark Energy Survey will present the analysis coming from the second and third year of observation, during this period DES has already covered the full expected final area (up to a given magnitude). The analysis will thus be realized in an area more than three times larger than the DESY1.

Part II

Redshift Estimation in DESY1

Chapter 3

Clustering redshift

3.1 Motivations

As mentioned in the previous chapters, one of the most challenging aspects of imaging galaxy surveys, in order to achieve accuracy requirements in cosmological parameter estimation, resides in the determination of redshift distributions of galaxy samples. Indeed, getting spectroscopic redshift for each object restricts us in sample size, and on the other hand getting precise enough photometric redshift becomes challenging as explained in section 2.2.3.2. However, while these two methods make an estimation of the redshift for each objects, it has been mentioned in section 2.4.3 that some cosmological probes, such as weak lensing measurements, do not require individual redshifts but redshift distributions of the sources.

In this context, a new type of methods have emerged, the so called **clustering-redshift methods**, which instead of relying on spectral information of objects, are based on the clustering properties of a given sample. Since it was presented and tested for the first time in Seldner & Peebles (1979), where the authors used the angular cross-correlation of a catalog of galaxies with spectral information with a catalog of quasars to infer its redshift distribution, these techniques seem to offer us today a good alternative to photo-zs.

More recently, in Newman (2008), the authors probed and showed using simulated catalogs that inferring redshift distributions using clustering information was actually feasible. Since then, these new methods has been tested in simulations in various works: Matthews & Newman (2010); Ménard et al. (2013); Schmidt et al. (2013); McQuinn & White (2013),and applied to observed data in Rahman et al. (2015, 2016a,b); Choi et al. (2016); Scottez et al. (2016), and recently Johnson et al. (2017); Hildebrandt et al. (2017); Morrison et al. (2017) used clustering redshift methods in the KIDS (Kilo-Degree Surveys de Jong et al. (2015)) data analysis, as a cross-check of photo-z posterior, and have shown the potential that this new types of methods have for future redshift surveys.

In the case of the Dark Energy Survey first year of observation (DESY1) cosmological analysis, clustering-redshifts have also been used, and for the first time the reconstructed distribution from clustering-redshifts was involved in the analysis not only as a cross-check of posterior distribution given by photometric redshift codes (as it has been done for KIDS) but to directly calibrate the redshift posteriors, obtained by photo-z codes, that were then used for the cosmological analysis (see Hoyle et al. (2018); Elvin-Poole et al. (2017); Troxel et al. (2017)). In Gatti, Vielzeuf. et al. (2018) we used cosmological simulations to determine a fiducial methodology and evaluate the systematic errors that it implies and in Davis et al. (2017) we exposed the results of our methology applied to the weak lensing sample (WL) of DESY1. This chapter aims to explain the methodology followed and present the results obtained in this context.

This chapter is organized as follows: First an introduction to the different clustering-redshift main idea and methods probed in this work will be presented (section 3.2), then a section will be dedicated to present the way these techniques have been included to the DESY1 data analysis, together with a presentation of the different galaxy catalogs involved in this study (section 3.3). Section (3.4) will explain how the systematics of the method were estimated, and finally section 3.5 will discuss the main results obtained for the calibration of the DESY1 WL sample.

3.2 Clustering-redshift methodology

3.2.1 Basic theory

In the literature a variety of methods to recover redshift distributions based on cross-correlation have been discussed (Newman, 2008; Ménard et al., 2013; Schmidt et al., 2013; McQuinn & White, 2013). The underlying idea shared by all these methods is that the spatial cross-correlation between two samples of objects is non-zero only in case of 3D overlap between them. In other terms, if one estimates the cross-correlation signal of two samples in the same footprint, one will get a signal in case of redshift overlap. In principle, the method could be applied to any galaxy sample, nevertheless applying it to a sample of galaxies optimally selected in photometric could also mitigate systematic effects Ménard et al. (2013); Schmidt et al. (2013); Rahman et al. (2016b).

More specifically, one considers two galaxy samples:

1. An *unknown* sample, whose redshift distribution $n_{\mathrm{u}}(z)$ has to be recovered.
2. A *reference* sample, whose redshift distribution $n_{\mathrm{r}}(z)$ is known (either from spectroscopic redshifts or from high-precision photometric redshifts).

In order to recover the redshift distribution of the unknown sample, we have to start by binning the reference sample in sufficiently narrow redshift bins and compute the cross-correlation signal between the 'unknown sample' and each narrow reference bin.

Similarly to section 1.2.2.4, one can express the mean surface density of the unknown sample $\Sigma_u(\theta, z)$ at an angle θ of each object of the reference sample at redshift z as :

$$< \Sigma_u(\theta, z) >_r = \Sigma_u(1 + w_{ur}(\theta, z)) \tag{3.1}$$

and from here, using a specific estimator as explained in 1.2.2.4 evaluate the cross-correlation signal between each reference sample narrow bin and the unknown sample.

In this specific analysis, to evaluate the correlation signal of our two samples, we have made the choice to use the Davis & Peebles estimator Davis & Peebles (1983), which for cross-correlations could be written as:

$$w_{ur}(\theta) = \frac{D_u D_r - D_u R_r}{D_u R_r} \tag{3.2}$$

where $D_u D_r$,and $D_u R_r$ represent the pair count between 'Data-Data' and 'Data-Random' respectively, and 'u' and 'r' represent the unknown and reference samples. This estimator has been selected because it presents the advantage of not requiring a random catalog for the unknown sample, saving us from generating these random catalogs, which, in our case, could have introduced an additional source of errors due to the complicated spatial selection function of the sample (see Zuntz et al. (2017) and section 3.3.2.2). Instead of that, we only need random catalogs for the reference sample for which we have a good knowledge of masking effects and selection function (see Elvin-Poole et al. (2017)).

On the other hand from Limber equation (**??**) and assuming a linear bias model (1.39), one can rewrite the angular cross-correlation of the two considered samples as :

$$w_{\rm ur}(\theta) = \int dz' n_u(\theta, z') n_r(\theta, z') b_u(\theta, z') b_r(\theta, z') w_{DM}(\theta, z'), \tag{3.3}$$

where $_u$ and $_r$ are the unknown and reference redshift distributions (normalized to unity over the full redshift interval), b_u and b_r are the biases of the two samples, and w_{DM} is the dark matter 2-point correlation function. If one assumes that the redshift distribution of the two samples as well as their biases do not depend on the angular scales, this equation becomes :

$$\bar{w}_{\rm ur} = \int dz' n_u(z') n_r(z') \bar{b}_u(z') \bar{b}_r(z') \bar{w}_{DM}(z'), \tag{3.4}$$

where the barred quantities refers to the averaging over angular scales.

Starting from here, various way to estimate this signal have been proposed. In this analysis, we have tried three different clustering-redshift based methods, the 'Schmidt/Ménard' method

from Schmidt et al. (2013); Ménard et al. (2013) that will be explained in 3.2.1.1 , a 'Weighted' method detailed in 3.2.1.2 and the 'Newman's' method from (Matthews & Newman (2010)) presented in 3.2.1.3.

3.2.1.1 Schmidt/Ménard

This method has been introduced in Schmidt et al. (2013). The main underlying idea there (as well as for the method presented in 3.2.1.2) is to assume that the reference sample is divided in bins as narrow as one can, to approximate its redshift distribution in each narrow bin as :

$$n_r(z) \propto N_r \delta_D(z - z') \tag{3.5}$$

where N_r is the number of reference objects in the considered bin and δ_D is Dirac's delta distribution. Hence, eq.(3.4) can be inverted and one can express the redshift distribution of the unknown sample as :

$$n_u(z) \propto \bar{w}_{ur}(z) \frac{1}{\bar{b_u}(z)} \frac{1}{\bar{b_r}(z)} \frac{1}{\bar{w}_{DM}(z)} \tag{3.6}$$

In Schmidt et al. (2013), the authors measured the cross-correlation signal within '1-angular bin'. That is to say, instead of measuring the cross-correlation signal in several angular bins and fitting a power law as it's been usually done for angular correlation analyses (see paragraphs **??** and 1.2.3), one evaluates the mean density of the unknown sample objects around each reference sample object within a single physical anulus going from r_{min} to r_{max} the minimum and maximum comoving distances. In Schmidt et al. (2013), the authors showed that applying a weight proportional to the inverse distance to the reference object improves the signal to noise ratio of the reconstruction (giving more statistical power to closer objects).

In this context, one can estimate the cross-correlation signal between the two samples using the Davis & Peebles estimator as :

$$\bar{w}_{ur} = \frac{N_{R_r}}{N_{D_r}} \frac{\int_{r_{min}}^{r_{max}} dr' W(r')[D_u D_r(r')]}{\int_{r_{min}}^{r_{max}} dr' W(r')[D_u R_r(r')]} - 1 \tag{3.7}$$

where $W(r')$ is a weighting function that is proportional to the inverse radius $(1/r)$.

This approach is one of the most popular in clustering-based redshift analyses and has shown to give consistent results in both observed data and simulations (Schneider et al. (2006); Ménard et al. (2013, 2010); Schmidt et al. (2013); Rahman et al. (2015, 2016a,b); Scottez et al. (2016, 2018)).

3.2.1.2 Weighted method

Another method presented in Ménard et al. (2013) is similar than the previous one (3.2.1.1), in the sense that it only differs in the weighting procedure: whereas before the weights were applied to every pair to compute the cross-correlation signal, here the cross-correlation signal is first computed as a function of angular separation θ and afterwards the weights are applied to the signal. In this sense, eq.(3.7) can be rewritten as:

$$\bar{w}_{ur} = \int_{\theta_{min}}^{\theta_{max}} d\theta W(\theta) w_{ur}(\theta, z) \tag{3.8}$$

where $W(\theta) \propto \theta^{-\gamma}$ is the weight function where we assumed $\gamma = 1$ giving more statistical power for the pairs with small angular separation increasing thus our signal to noise ratio.

3.2.1.3 Newman

The last approach that has been tested is presented in Newman (2008); Matthews & Newman (2010). In there, the authors adopted a different procedure from the two previous ones. Indeed they start from the assumption that the cross-correlation signal could be described by a power law $\xi = (r/r_0)^{-\gamma}$. And then assuming linear biases, it is possible to relate the computed cross-correlation signal to a specific cosmological model as :

$$w_{ur}(\theta, z) = \frac{n_u(z) H(\gamma_{ur}) r_{0,ur}^{\gamma_{ur}} D_A^{1-\gamma_{ur}}}{d\chi/dz}, \tag{3.9}$$

where γ_{ur} is the power law slope of the cross-correlation signal, $H(\gamma_{ur}) = \Gamma(1/2)\Gamma((\gamma_{ur} - 1)/2)/\Gamma(\gamma_{ur})$ (Peebles (1980)), and $D_A(z)$ and (z) are respectively the angular size distance and comoving distance at redshift z. From here, we fitted the angular cross-correlation signal as :

$$w_{ur}(\theta, z) = A_{ur}(z)\theta^{1-\gamma_{ur}} - C_{ur} \tag{3.10}$$

Equalizing now eq.(3.9) and (3.10) we get the redshift distribution for the unknown sample as :

$$n_u(z) = \frac{d\chi/dz}{D_A^{1-\gamma_{ur}} H(\gamma_{ur}) r_{0,ur}^{\gamma_{ur}}} * (A_{ur}(z)\theta^{1-\gamma_{ur}} - C_{ur}) \tag{3.11}$$

Comparing to the two other methods probed in this work, we can see that this approach carries additional degrees of freedom:

- γ_{ur} (as well as the correlation length $r_{0,ur}$) that could be estimated (under the assumption of linear biases) from the auto-correlation signal of both samples $\gamma_{ur} = (\gamma_{uu} + \gamma_{rr})/2$ (and $r_{0,ur}^{\gamma_u r} = \sqrt{(r_{0,uu}^{\gamma_u u} r_{0,rr}^{\gamma_r r})}$)
- C_{ur} (and $A_{ur}(z)$) obtained using chi-square minimization using the covariance of a jack-knife resampling (see section 3.3.1.3)

Here the only parameter that cannot be directly measured is $r_{0,uu}$ but adopting an iterative approach, one can insert a first guess of $r_{0,uu}$ in eq.(3.11), that will then be inserted in the cross-correlation signal of eq.(3.9) to refine the value of $r_{0,uu}$, and we iterate this process until convergence. Following what has been done in Matthews & Newman (2010) where the authors assumed $r_{0,uu} \propto r_{0,rr}$, one can start the iterative procedure with $r_{0,uu} = r_{0,rr}$. However, one can note that this approach could breack down if the bias evolution of the two samples don't follow a proportionality relation; this will be commented in section 3.4.2.5.

3.2.1.4 Bias correction

As mentioned before, the Newman method assumes a linear modeling for the biases, whereas in both Schmidt/Ménard and Weighted methods, the effects of the redshift evolution of the biases was neglected. In principle, one could correct for the redshift evolution of the biases and the dark-matter 2-point correlation function using the 1-bin estimate of the auto-correlation signal of both samples. Indeed, if one divides both the reference and unknown sample in bins narrow enough, the auto-correlations $\bar{w}_{uu}(z)$ or $\bar{w}_{rr}(z)$ could be related to the redshift distributions of the corresponding samples in the considered bin $n_{u,z}(z)$ or $n_{r,z}(z)$ as :

$$\bar{w}_{rr,z} = \int dz' [b_r(z') n_{r,z}(z')]^2 \bar{w}_{DM}(z') \tag{3.12}$$

$$\bar{w}_{uu,z} = \int dz' [b_u(z') n_{u,z}(z')]^2 \bar{w}_{DM}(z') \tag{3.13}$$

and from here correct the reconstructed redshift distribution of eq.(3.6) as :

$$n_u(z) \propto \frac{\bar{w}_{ur}(z)}{\sqrt{\bar{w}_{uu}(z)\bar{w}_{rr}(z)}}. \tag{3.14}$$

The limit of this methodology comes from the fact that, to apply this correction, one needs a good knowledge of the distributions of objects within the narrow bins (which we don't have for the unknown sample), and in the case of the reference sample the auto-correlation estimate could be really noisy depending on the size of the sample. However, in simulated catalogs, one has the

true redshift information, so that would be then feasible in simulations to estimate the impact of neglecting these effects, (see section 3.4.2.1).

3.3 The Dark Energy Survey first year fiducial methodology

As mentioned previously, in the context of the Dark Energy survey first year of observations cosmological analysis, the reconstructed redshift distributions provided by clustering-redshift based methods have not been used directly but as a tool to correct the posterior produced by photo-z code. Indeed, mainly due to the fact that the higher redshift bin of the sample we aimed to reconstruct was not covered by our reference sample (see hereafter, paragraphs 3.3.2.2 and 3.3.2.3), we have been using our reconstructed distribution to correct the bias in the mean of the redshift distribution from photo-z codes. The framework could therefore be separated in two steps :

- *step 1 :* reconstruct the redshift distribution of your unknown sample using clustering based methods;
- *step 2 :* correct the mean of the photo-z posterior of your unknown sample using your reconstructed method

One could imagine different manners to correct the photo-z posterior. In this work, we have been testing two different approaches, and the first part of this section will be dedicated to present these specific methodologies.

Then in a second part, the different samples that have been used for this analysis will be introduced.

3.3.1 Matching procedures

As specified before, the goal was to calibrate the posterior of the redshift distribution obtained using a given photo-z code using the distribution obtained using clustering-redshift based techniques. Two different calibration approaches based on the matching of the two distributions using likelihood functions were used to solve for the shift to be applied to the photo-z posterior coming from the photo-z code.In the context of this work, we have been using the affine-invariant Markov Chain Monte Carlo ensemble sampler `emcee` Foreman-Mackey et al. (2013)[1] to sample the likelihood.

[1] `http://dan.iel.fm/emcee`

3.3.1.1 Shape matching

The goal here would be to correct the distribution obtained with the photo-$z$code ($n_{pz}(z)$) by matching it to the one reconstructed with our methology ($n_{wz}(z)$) within the redshift interval of the considered photometric bin. The corrected distribution is considered to be $n_\Delta = n_{pz}(z - \Delta z)$, allowing then a shift of the uncorrected distribution along the z-direction. We find Δz by minimizing

$$log\mathcal{L} = -\frac{1}{2}\chi^2\left(e^k n_\Delta; n_{WZ}; \hat{\Sigma}^{-1}_{WZ}\right) + Prior(k, \Delta z), \tag{3.15}$$

where e^k is a normalizing factor that allows us to rescale the corrected distribution to unity over the full redshift-bin interval considered. $\hat{\Sigma}^{-1}_{WZ}$ represents the covariance matrix of the clustering redshift proceedure (see 3.3.1.3). In this approach (as well as the one presented in section 3.3.1.2), we have been using flat priors for both k and Δz

3.3.1.2 Mean matching

In a similar way, if we want now to match directly the mean of the photo-z distribution $< z >_{pz}$ to the mean of the clustering redshift distribution $< z >_{WZ}$ and don't take into account the shapes of the distributions anymore, the likelihood function in this specific case can now be define as:

$$log\mathcal{L} = -\frac{1}{2}\chi^2\left(< z >_\Delta; < z >_{WZ}; \hat{\Sigma}^{-1}_{<z>_{WZ}}\right) + Prior(\Delta z), \tag{3.16}$$

where $\hat{\Sigma}^{-1}_{<z>_{WZ}}$ is the covariance matrix of the matching process.

Moreover, for this approach, as it will be detailled in section 3.4.2.6, we saw that matching the means computed in a specific redshift interval around the mean of our reconstructed distribution $n_{wz}(z)$, instead of computing it in the full redshift interval, was helping to remove the noise that can appear in the tails of the distribution (as well as additional systematics that could be dominant in this low density regions, see C). A comparison of different interval size is presented in section 3.4.2.6, and in there we find that the optimal interval to match the means of our distributions is $\pm 2\sigma_{WZ}$ around the mean of the reconstructed clustering-redshift distribution.

3.3.1.3 Error estimation

To estimate the covariance in our procedure (the $\hat{\Sigma}$ terms in the likelihoods mentioned before), we have been using a jackknife re-sampling approach Norberg et al. (2009). To do so, we divided our sample ($\sim 1100 deg^2$, see section 3.3.2.1) in 1000 sub-regions of equal area ($\sim 1 deg^2$ each),

which are larger than the maximum integration scale that we will use in this analysis (fiducial scales used presented in section 3.4.1), and we compute the covariance matrix of these different Jackknife areas as :

$$\hat{\Sigma}(x_i, x_j) = \frac{(N_{JK} - 1)}{N_{JK}} \sum_{k=1}^{N_{JK}} (x_i^k - \bar{x}_i)(x_j^k - \bar{x}_j) \tag{3.17}$$

where N_{JK} is the number of jackknife regions, x_i is the statistic of interest of the i-th bin in the k-th jackknife area. At the moment of inverting the Jackknife covariance matrix, we un-biased it by multiplying it by the Anderson-Hartlap factor $\alpha = (N_{JK} - N_{bins} - 2)/(N_{JK} - 1)$ (Hartlap et al. (2007)), with N_{bins} the number of redshift bins considered in the reconstruction. Moreover, in our procedure, as it will be presented in next sections, we had to reconstruct the redshift distributions of different tomographic bins separately. In principle, we should have considered correlations between different bins and compute the fit for all three bin simultaneously. However, by doing this we have found that systematic errors got propagated from one bin to the other due to off-diagonal terms in the Jackknife covariance having for consequence a wrong estimation of the best fit model for all bins.

Nevertheless, since the effect of correlations between bin is a negligible increase of statistical errors and as it will be shown in the next sections, our methodology is dominated by systematic errors such correlation have been neglected.

3.3.2 The samples used

3.3.2.1 The Buzzard simulations

To test our methodology and evaluate its possible systematic errors, we first used the Buzzard simulations. The Buzzard simulation is a mock DESY1 survey created from a set of dark-matter-only simulations. The simulation and creation of the mock survey data is detailed in DeRose et al. (2017); Wechsler et al. (2017); MacCrann et al. (2018), here we provide only a brief summary. Buzzard-v1.1 is constructed from a set of 3 N-body simulations run using L-GADGET2, a version of GADGET2 (Springel, 2005) modified for memory efficiency. The simulation boxes ranged from 1 to 4 Gpc/h. Light-cones from each box were constructed on the fly. Halos were identified using ROCKSTAR (Behroozi et al., 2013), and galaxies were added to the simulations using the Adding Density Dependent GAlaxies to Light-cone Simulations algorithm (ADDGALS, Wechsler et al. 2017). ADDGALS uses the large scale dark matter density field to place galaxies in the simulation based on the probabilistic relation between density and galaxy magnitude. The latter is calibrated from sub-halo abundance matching in high-resolution N-body simulations. Spectral energy distributions (SEDs) are assigned to the galaxies from a training set of spectroscopic data from SDSS DR7 (Cooper et al., 2011) based on local environmental density. The SEDs are integrated in the DES pass bands to generate *griz*

magnitudes. Galaxy sizes and ellipticities are drawn from distributions fit to deep SuprimeCam $i^{'}$-band data. Galaxies are added to the simulation to the projected apparent magnitude limit of DES Y5 data out to redshift $z = 2$. The galaxy positions, shapes and magnitudes are then lensed using the multiple-plane ray-tracing code Curved-sky grAvitational Lensing for Cosmological Light conE simulatioNS (CALCLENS, Becker 2013). Finally, the catalog is cut to the DES Y1 footprint with $RA > 0$ using the footprint and bad region masking including bright stars, regions of high extinction, etc., used in the actual Y1 data, and photometric errors are added using the DES Y1 depth map (Rykoff et al., 2015). This yields a total masked area of 1108.13 square degrees, 12 million weak lensing source galaxies, and 102120 galaxies in the higher luminosity *redMaGiC* sample used in this analysis, as will be discussed in 3.3.2.2 and 3.3.2.3 respectively.

3.3.2.2 The 'unknown sample'

For DESY1 data analysis, clustering-redshift based methods have been used to calibrate the redshift distributions that entered in the cosmological analysis using cosmic shear (Troxel et al. (2017) and section 2.4.3), hence the sample that had to be corrected using the present methodology is the weak lensing (WL hereafter) source galaxies presented in Zuntz et al. (2017). Due to a lack of reference sample coverage in the higher redshift bin (see 3.3.2.3), we restricted ourselves to the first three bins $z = [(0.2 - 0.43), (0.43 - 0.63), (0.63 - 0.9)]$ used in the cosmological analysis. We also use our method with two different photo-z codes in both observed data and simulations, the Bayesian Photometric Redshift (`BPZ`, (Benítez, 2000; Coe et al., 2006)), and the Directional Neighborhood Fitting (`DNF`, (De Vicente et al., 2016))(as presented in Hoyle et al. (2018)). The following paragraphs aim to give a brief description of these two codes.

- **The Bayesian Photometric Redshift (BPZ)** (Benítez, 2000; Coe et al., 2006) `BPZ` is a template-based method (see section 2.2.3.2). It returns the full probability distribution $p(z)$ for each galaxy given its magnitudes and template libraries. Templates and priors used here are described in more detail in Hoyle et al. (2018), as BPZ has been run on simulations with the same setup used on data. Briefly, templates are generated based on low-redshift models from Coleman et al. (1980) and Kinney et al. (1996), while the redshift evolution and calibration corrections to template fluxes have been computed by matching PRIMUS DR1 Coil et al. (2011) spectroscopic redshifts to DES photometry. The calibration sample comprises 72,176 galaxies. The luminosity prior used in BPZ takes the form of smooth exponential functions, which have been fitted using COSMOS galaxies with accurate photometric redshifts (Laigle et al., 2016).

- **Directional Neighborhood Fitting (DNF)** (De Vicente et al., 2016). DNF is a machine learning algorithm for galaxy photometric redshift estimation. Based on a training sample, DNF constructs the prediction hyperplane that best fits the neighborhood of each target

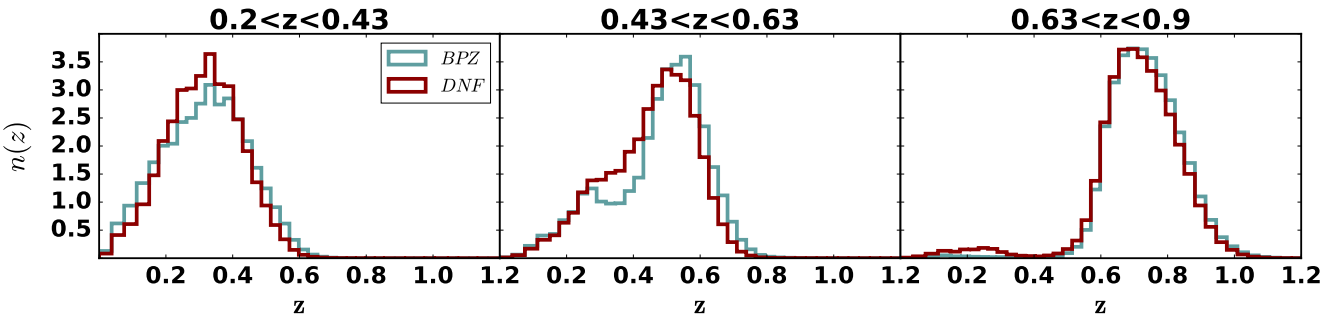

FIG. 3.3.1 – Redshift distribution of the 3 WL tomographic bins from the posteriors obtained running the BPZ (blue) and DNF (dark red) photo-z code on the Buzzard simulation (see 3.3.2.2). All distributions has been normalized over the full redshift interval

galaxy in multiband flux space. It then uses this hyperplane to predict the redshift of the target galaxy. This redshift is used to divide the WL sample into tomographic bins. The key feature of DNF is the definition of a new neighborhood, the Directional Neighborhood. Under this definition two galaxies are neighbors not only when they are close in the Euclidean multiband flux space, but also when they have similar relative flux in different bands, i.e. colors.

Figure 3.3.1 shows the redshift distributions obtained from the two photo-z codes (ran on the Buzzard simulations) for the three tomographic bins of interest. As it can be seen in this figure, the two codes gives similar distributions, except for the second tomographic bin where a secondary 'bump' appear on the BPZ true distribution at $z \sim 0.3$ (BPZ includes there a population of blue galaxies). And Fig.(3.3.2) shows the scatter and bias of BPZ redshift in the simulated Weak Lensing sample.

3.3.2.3 The reference sample

As reference sample, we made the choice to use, instead of a spectroscopic catalog as done in previous works, a high-qualitiy photo-z catalog, the *redMaGiC* higher-luminosity sample (see appendix A for more details on such a choice), created in both simulated and observed data using the red-sequence Matched-filter Galaxy Catalog (*redMaGiC*) algorithm. *redMaGiC* (Rozo et al. (2016)) is an algorithm that aims to select Luminous Red Galaxies (LRG) above a luminosity threshold. *redMaGiC* is doing a fit to red-sequence templates that had been trained using the *redMaPPer* algorithm presented in Rykoff et al. (2014). This selection of objects manage to minimize the redshift uncertainties down to $\sigma_z/(1+z) < 0.02$. The higher-luminosity catalog used in this work is a selection of galaxies with a luminosity cut at $L > 1.5L*$ and span a redshift range of $0.15 < z < 0.85$. The algorithm has been run both in the Buzzard simulated catalog (see sec. 3.3.2.1) and in the DESY1 gold catalog (Drlica-Wagner et al. (2018)). Figure

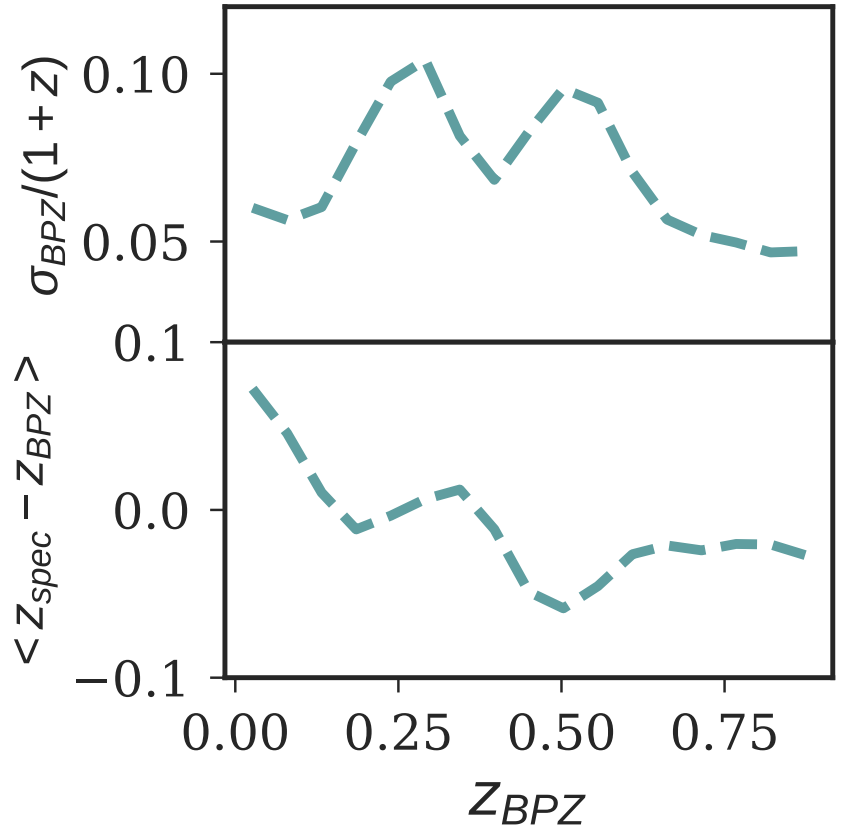

FIG. 3.3.2 – Scatter (top) and bias (bottom) of the WL sample BPZ redshift in Buzzard simulation (cannot be estimate for observed sample due to lack of complete spectroscopic sample)

3.3.3 shows a comparison of the true and photometric redshift distribution of higher-luminosity sample obtained by running the *redMaGiC* algorithm on the Buzzard simulation.

The resulting simulated catalog appears to have $\sim$ 30% less galaxies than the observed catalog. This is mainly due to the fact that the simulated red-sequence selected by *redMaPPer* seems to be redder in simulations, implying that the simulated catalog doesn't reach a maximum redshift as high as the observed *redMaGiC* catalog ($z \leqslant 0.85$ vs. $z \leqslant 0.9$). Nevertheless, the consequences of such a difference will appear as an overestimation of the statistical error, which will be inconsequential as our methodology is systematics dominated (see section 3.4). Additionally, in order to be consistent with the systematic error estimated in simulation, we made the choice when it came to apply our methodology to data to use the *redMaGiC* DESY1 reference sample up to $z = 0.85$ only (see section 3.5) to cover the same redshift range in both data and simulation. Figure 3.3.4 shows the differences between observed and simulated *redMaGiC* catalogs of the scatter (top panel) and redshift bias (bottom panel), and as the figure shows, simulations and observations are consistent. This estimation is straightforward in the simulated catalog, since we have the true redshift for each object. In real data we had to compute this estimation using a

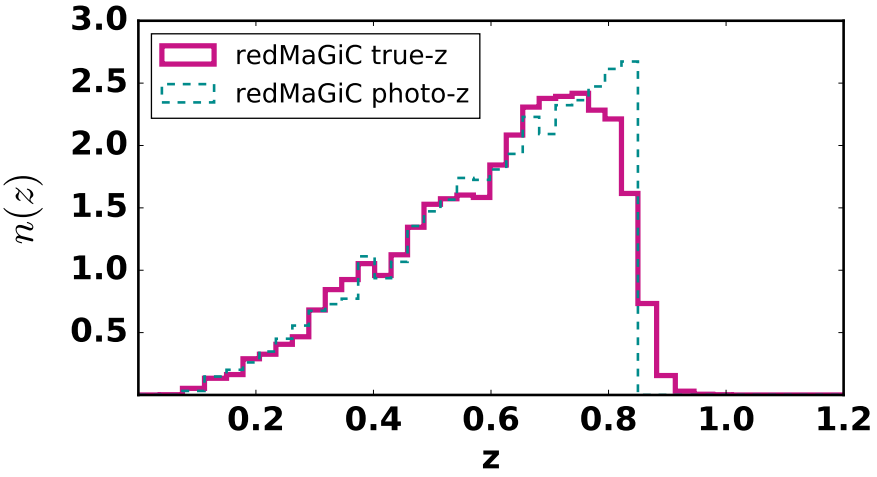

FIG. 3.3.3 – True (continuous pink line) and photometric (dashed cyan line) redshift distribution of the Buzzard Higher-luminosity sample

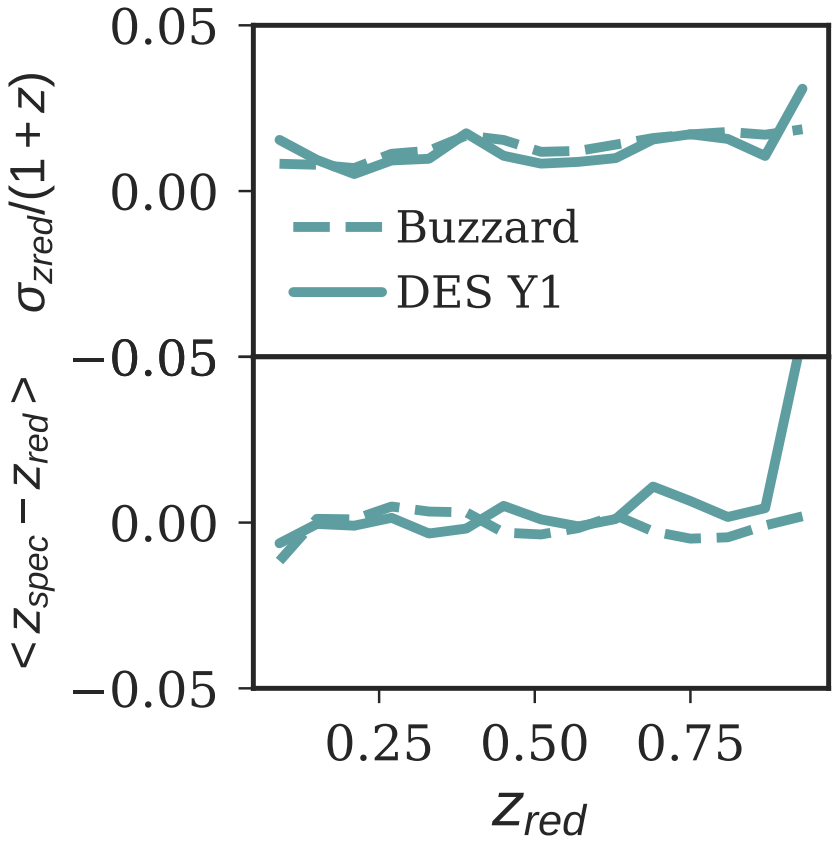

FIG. 3.3.4 – Scatter (top) and bias (bottom) comparison for the simulated (dashed) and data (solid line) *redMaGiC* catalogs

training set of galaxies (and not the full sample). In Cawthon et al. (2017) more details on the robustness of this estimation are given.

It's also interesting to comment that several *redMaGiC* catalogs have been produced in the context of DESY1, we have made the choice of using this specific one because it appears to give us better results in simulations (optimal balance between galaxy density and photometric accuracy), although in appendix A we report the results obtained using as reference sample the 'combined' *redMaGiC* sample presented in Elvin-Poole et al. (2017).

Additionally, to compute the cross-correlation signal using *redMaGiC* as reference, it was necessary to produce a random catalog for this sample. For the simulated sample it has been done uniformly over the survey footprint since in simulated catalog does not contain observational systematics. However, in the observed data analysis, these observational features had to be taken in account by generating the appropriate random catalog (see Elvin-Poole et al. (2017)).

3.4 Systematic evaluation and model parameters optimization

Before applying to our DESY1 WL samples the procedure presented before, we studied the method limits and optimized its underlying parameters using the Buzzard simulation from section 3.3.2.1. In this section, we will present the procedure applied to identify the different systematics that suffers our methodology.

3.4.1 Fiducial methodology

As mentioned before, to proceed with our reconstruction procedure, it is necessary to fix a number of parameters. These parameters that will be presented in this paragraph have been tested in simulations in order to minimize systematic effects. In this paragraph, we set the fidutial methodology that has been employed along our analysis. We found that the optimal combination to obtain the better S/N ratio is the following :

- Clustering redshift method : *Schmidt/Ménard* (see 3.2.1.1)
- Reference sample : *redMaGiC Higher-Luminosity sample binned in 25 uniform redshift bins in the range* $0.15 < z < 0.85$ *, (bin size* $\Delta z_r \sim 0.03$*)*
- Integration scale : $0.5 - 1.5Mpc$
- Matching methodology : *matching the mean within window function* $\pm 2\sigma_{WZ}$ *around the clustering-redshift recovered distribution*
- Bias correction : *no bias correction applied*

For more details on these choices, I refer you to sections 3.4.2.5, 3.4.2.6, and Appendices A, B where comparison of integration scales, methodologies and reference samples are made. The following subsection will present the results obtained and the framework followed to identify systematic effects of our technique using the above fiducial settings (although the results for the shape systematic will also be exposed to justify the choice made on the matching technique).

3.4.2 Characterization of the method systematics

Using the simulated samples presented in the previous section, we aimed to estimate the systematics errors that our methodology suffers correcting the photometric redshift posterior distribution. To do so, we quantify systematic errors using as metric the difference between the mean of the true distribution $< z >_{true}$ to the one recovered after applying our procedure $< z >_{\Delta}$:

$$\Delta z =< z >_{true} - < z >_{\Delta} \tag{3.18}$$

In this work, we could identify three distinct systematics :

- **The Bias systematic:** caused by the evolution of the biases of the two samples and of the dark matter density field that we decided to neglect it and include it in the systematic budget (see eq.(3.4) and section 3.2.1.4)
- **The *redMaGiC* photo-z systematic :** Induced by the photo-z errors of the *redMaGiC* galaxies.
- **The Shape systematic :** depending on the matching procedure, this systematic could come from different aspects. For the 'shape-matching' case, the difference on the shapes of the two distributions has an effect on the recovered photo-z bias. In the 'mean-matching' case, since we are matching within a specific interval of $\pm 2\sigma_{WZ}$ around the clustering-redshift distribution, if the shapes of the two distribution differ outside this interval, it will add an additional systematic.

The estimates of these systematics for each photo-z code and each tomographic bin are shown in Tables 3.4.2 and 3.4.3, and the method used to distinguish and evaluate them is detail hereafter.

3.4.2.1 The bias evolution systematic

In our redshift reconstruction, we chose to neglect the effect of the evolution of the bias of our two samples (and the evolution of the Dark matter density field) which should be sub-dominant

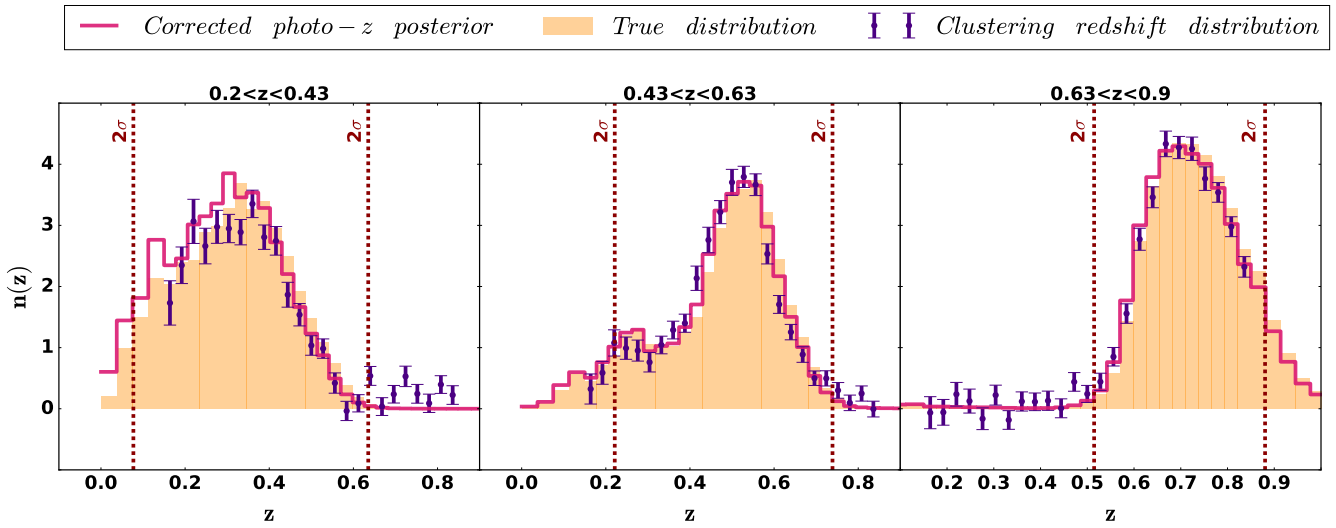

FIG. 3.4.1 – In three different tomographic bins from left to right, True redshift distribution of the buzzard weak lensing sample (orange filled histogram), recovered clustering redshift distribution (purple error bars), photo-z posterior corrected with the shift infered unsing the mean-matching technique within the $\pm 2\sigma_{WZ}$ interval (pink histogram)

effects if these quantities don't evolve too much with redshift (see Ménard et al. (2013)). Also, in principle, if these quantities evolve at scales larger than the ones used in our analysis, the impact of neglecting these evolution should be small. In this sense, binning the sample in narrow bins or selecting the sample in color should decrease the evolution (Ménard et al. (2013); Schmidt et al. (2013); Rahman et al. (2015)).

However, we need to estimate the inaccuracy that this will induce in our fiducial calibration, and to do so we proceed to perform the following test:

1. bin the simulated sample using the mean of the photo-z posterior ;
2. reconstruct the distribution using for each bin the **True redshift** for the *redMaGiC* sample as well as the **True shape** of the posterior for the WL sample.

Figure 3.4.1 shows the different redshift distributions (true, clustering-redshift and corrected photo-z posterior) for this specific scenario. In this framework, we eliminate the two additional systematics. The results are summarized in the first row of Table 3.4.2 for BPZ and 3.4.3 for `DNF` and labeled Δz_A. In the applied procedure, one would expect (without any systematic effect) to recover a shift value consistent with zero, since we are in the case where the photo-z posterior is assumed to be the true one. Instead of that, in Tables 3.4.2 and 3.4.3 we found shifts up to $\Delta z_A \sim 0.02$ in some bins. As it's shown as well in these tables, the level that our methodology is affected by this systematic depends on the considered bin, the matching procedure applied and the photo-z code used.

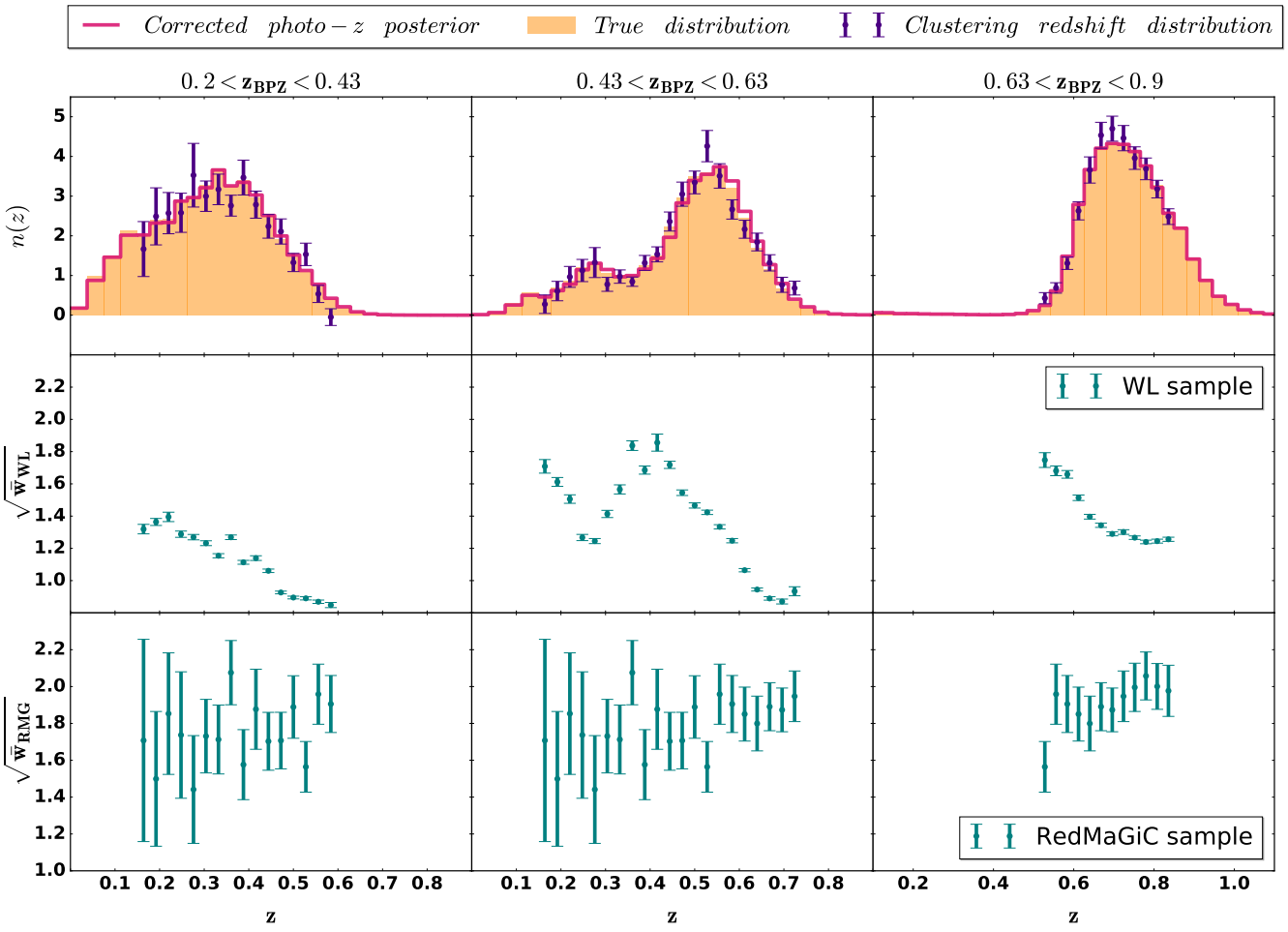

FIG. 3.4.2 – *Top panel:* True redshift distribution (orange), clustering-redshift reconstruction (purple error bars) and corrected photo-z posterior using true redshift for the reference sample and true shape for the photo-z posterior and applying the bias correction presented in 3.2.1.4 (pink histogram). *Middle and bottom panel:* Redshift evolution of the one-bin estimate of the auto-correlation function for the simulated weak-lensing sample (middle) and the simulated higher-luminosity *redMaGiC* sample (bottom) for the three first tomographic bins used in the DESY1 cosmological analysis from left to right.

Additionally, since we are using simulated data, we have the true redshift information for each object, and consequently, this allows us to perform the bias correction presented in section 3.2.1.4 verifying in this sense that the shift observed in Tables 3.4.2 and 3.4.3 is indeed coming from the evolution of the galaxy biases of our two samples and the dark matter 2-point correlation function, neglected in our fiducial methodology. Both samples had been binned in 25 equally-spaced redshift bins, from $z = 0.15$ to $z = 0.85$ and the auto-correlation signal has been computed for each sample in every narrow bin. If the size of the considered bin is narrow enough, then from eq.(3.12) we have $w_{rr} \propto b_r^2 w_{DM}$, and $w_{uu} \propto b_u^2 w_{DM}$ and we can use eq.(3.14) to compute the corrected distribution.

Figure B.0.1 shows the results obtained after bias correction (top panel) using the auto-correlation eq.(3.12) as a function of redshift of the WL sample (middle panel) and *redMaGiC* sample

TABLE 3.4.1 – Recovered shift to apply to the mean of our distribution after applying the bias correction presented in section 3.4.2.1.

	Bin 1		**Bin 2**		**Bin 3**	
	mean match	shape match	mean match	shape match	mean match	shape match
bias evolution systematic (BPZ):	−0.004 ± 0.013	0.001 ± 0.007	−0.007 ± 0.007	0.004 ± 0.003	0.002 ± 0.003	0.001 ± 0.002
bias evolution systematic (DNF):	0.000 ± 0.009	−0.001 ± 0.006	−0.002 ± 0.005	0.002 ± 0.004	0.004 ± 0.003	0.002 ± 0.003

(bottom panel). And Table 3.4.1 shows the recovered shift obtained after correcting for the bias and Dark Matter density field evolutions as described here. As one can see in the table, the recovered shifts are now consistent with zero, hence confirming that the first identified systematic is due to the evolution of the biases and dark matter density field.

It's important to note that this correction is possible only in simulations, where we have the true redshift information of the sample. Even-though a correction for the bias of the reference sample is feasible, as presented in Appendix B, it hasn't been applied since the effect is small and the auto-correlation signal of *redMaGiC* is noisy due to the small number of objects in each narrow bin. The effect of the biases and dark matter density field evolution has been thus considered as a systematic effect in this analysis.

Moreover, as Figure B.0.1 is showing the bias evolution of the WL sample is relatively complex, and therefor correcting from it becomes complicated and hence not appropriate for parametric approaches such it as been done in Matthews & Newman (2010); Schmidt et al. (2013); Davis et al. (2018)

3.4.2.2 The *redMaGiC* photo-z systematic

Previous clustering redshift analysis have been using spectroscopic samples as reference sample. In our case, we made the choice to use a 'high-resolution' photometric sample (the *redMaGiC* higher-luminosity sample presented in section 3.3.2.3), and for a comparison of different *redMaGiC* samples as reference see appendix A. Even though the *redMaGiC* photo-z error is small compared to standard photo-z, we want to estimate the consequence of such a choice. We proceeded as following:

1. bin the simulated sample using the mean of the photo-z posterior;
2. reconstruct the distribution using for each bin the **photo-z redshift** for the *redMaGiC* sample as well as the **True shape** of the posterior for the WL sample.

The results are shown in Fig.(3.4.3), and summarized in Table 3.4.2 for BPZ and 3.4.3 for DNF. Note that Tables 3.4.2 and 3.4.3 present the *redMaGiC* systematic ($\Delta_{z\,redMaGiC}$) as the difference

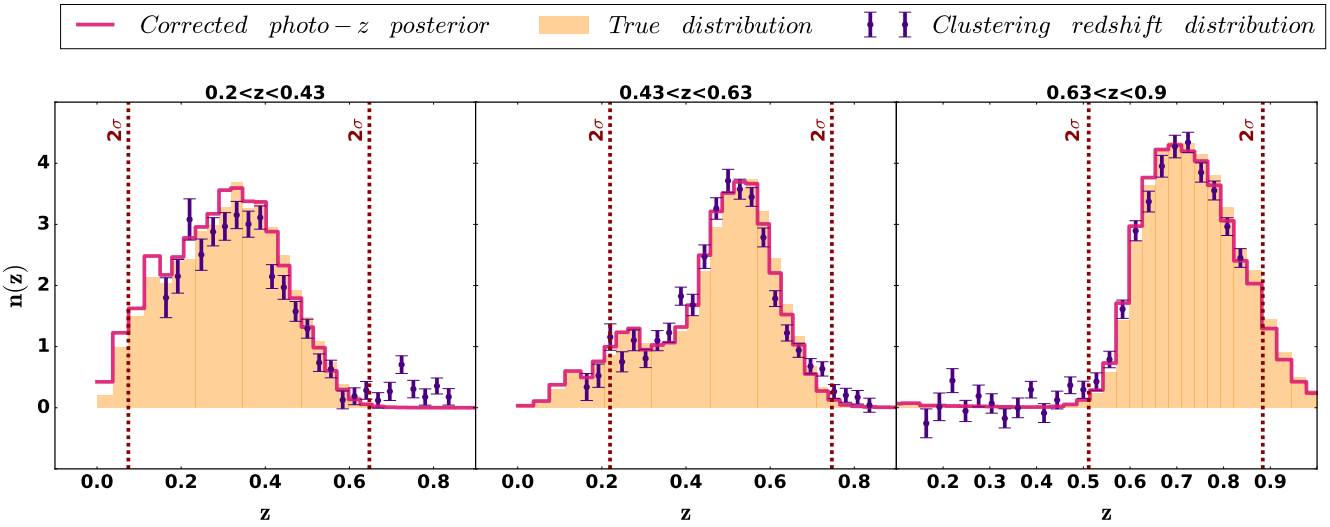

FIG. 3.4.3 – In three different tomographic bins from left to right, True redshift distribution of the buzzard weak lensing sample (orange filled histogram), recovered clustering redshift distribution (purple error bars), photo-z posterior corrected with the shift inferred using the mean-matching technique within the $\pm 2\sigma_{WZ}$ interval (pink histogram)

between the shift obtained in this test (Δz_B) and the one obtained for the **bias evolution** systematic (Δz_A) :

$$\Delta z_{redMaGiC} = \Delta z_B - \Delta z_A \tag{3.19}$$

As one can see, the effect of *redMaGiC* photo-z is sub-dominant compared to the bias systematic ($\sim$ 1 order of magnitude lower). There we are not showing the statistical errors for this systematic, because the statistical error from Δz_B and Δz_A are correlated (we are using the similar data covariances to estimate them) thus the difference between them is close to zero.

3.4.2.3 The shape systematic

Now we want to place ourselves in the more realistic case, by following the steps :

1. bin the simulated sample using the mean of the photo-z posterior;
2. reconstruct the distribution using for each bin the **photo-z redshift** for *redMaGiC* sample as well as the **shape** of the photo-z posterior for the WL sample.

We are thus in a similar case that we will be when we will apply the methodology to the DES data catalogs. By subtracting the effect of the two previous systematics presented from the shift obtained in this case, we can evaluate the effect of the remaining systematic, namely the **shape systematic**, defined as $\Delta z_C - \Delta z_B$.

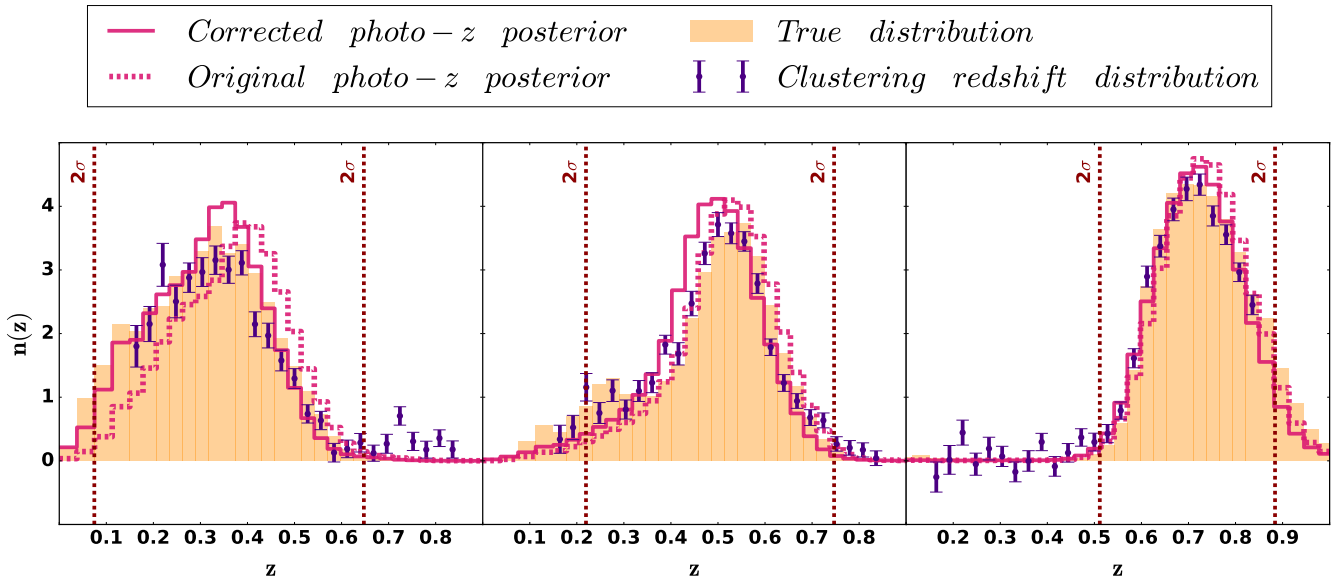

FIG. 3.4.4 – In three different tomographic bins from left to right, True redshift distribution of the buzzard weak lensing sample (orange filled histogram), recovered clustering redshift distribution (purple error bars), original photo-z posterior (dashed pink histogram), photo-z posterior corrected with the shift inferred using the mean-matching technique within the $\pm 2\sigma_{WZ}$ interval (pink histogram)

The results are shown in Fig.(3.4.4), and summarized in Table 3.4.2 for BPZ and Table 3.4.3 for DNF. From these two tables, one can note that this systematic is more important for the shape matching than for the mean matching, in particular when the differences in the shape of the two redshift distributions (true and photo-z) is bigger such as in the second tomographic bin, where a second peak appears in the true distribution.

For the same reason than in 3.4.2.2, we are not showing in Table(3.4.2) and (3.4.3).

3.4.2.4 Total systematic budget

In order to get a estimation of the total systematic effect to take into consideration when we will apply this procedure to data, we decided to add each of the identified systematics in quadrature avoiding therefore any fortunate cancellation between them. This also implied that we ignored any correlation that could exist between different systematics, although one would expect the *redMaGiC* systematic to be uncorrelated to the two others. However, we could expect correlations (or anti-correlations) between the bias systematic and the shape systematic. The total systematic errors provided for the cosmological analysis are shown in the fourth row of Table 3.4.2 for BPZ and Table 3.4.3 for DNF for each tomographic bin and both matching methods applied, although at the end the mean matching procedure was taken as fiducial.

TABLE 3.4.2 – **BPZ systematic errors**. Systematic errors for BPZ, as a function of WL redshift bin and matching procedure. Δz_A, Δz_B and Δz_C refer to the residual shifts in the mean relative to the scenarios A, B and C, as outlined in §3.4.2.1, §3.4.2.2 and §3.4.2.3. For the most realistic scenario (scenario C, §3.4.2.3). When a value in the table is accompanied by an uncertainty, it refers to the statistical uncertainty, estimated from the posterior of the photo-z bias, as explained in §**??**.

	Bin 1		**Bin 2**		**Bin 3**	
	mean match	shape match	mean match	shape match	mean match	shape match
bias evolution systematic: Δz_A	0.020 ± 0.006	0.019 ± 0.005	0.010 ± 0.004	0.020 ± 0.002	0.008 ± 0.003	0.007 ± 0.003
***redMaGiC* photo-z systematic:** $\Delta z_B - \Delta z_A$	-0.009	-0.005	-0.001	-0.006	-0.001	-0.002
shape systematic: $\Delta z_C - \Delta z_B$	-0.011	-0.017	-0.012	-0.032	0.004	0.008
total systematic error	**0.025**	**0.026**	**0.016**	**0.038**	**0.014**	**0.011**

TABLE 3.4.3 – **DNF systematic errors**. Same as table 3.4.2, but for DNF.

	Bin 1		**Bin 2**		**Bin 3**	
	mean match	shape match	mean match	shape match	mean match	shape match
bias evolution systematic: Δz_A	0.007 ± 0.005	0.008 ± 0.004	0.012 ± 0.004	0.010 ± 0.003	0.010 ± 0.003	0.003 ± 0.002
***redMaGiC* photo-z systematic:** $\Delta z_B - \Delta z_A$	−0.005	−0.002	−0.006	−0.001	−0.003	−0.001
shape systematic: $\Delta z_C - \Delta z_B$	−0.007	−0.011	−0.002	−0.032	−0.015	−0.024
total systematic error	**0.012**	**0.013**	**0.013**	**0.034**	**0.019**	**0.025**

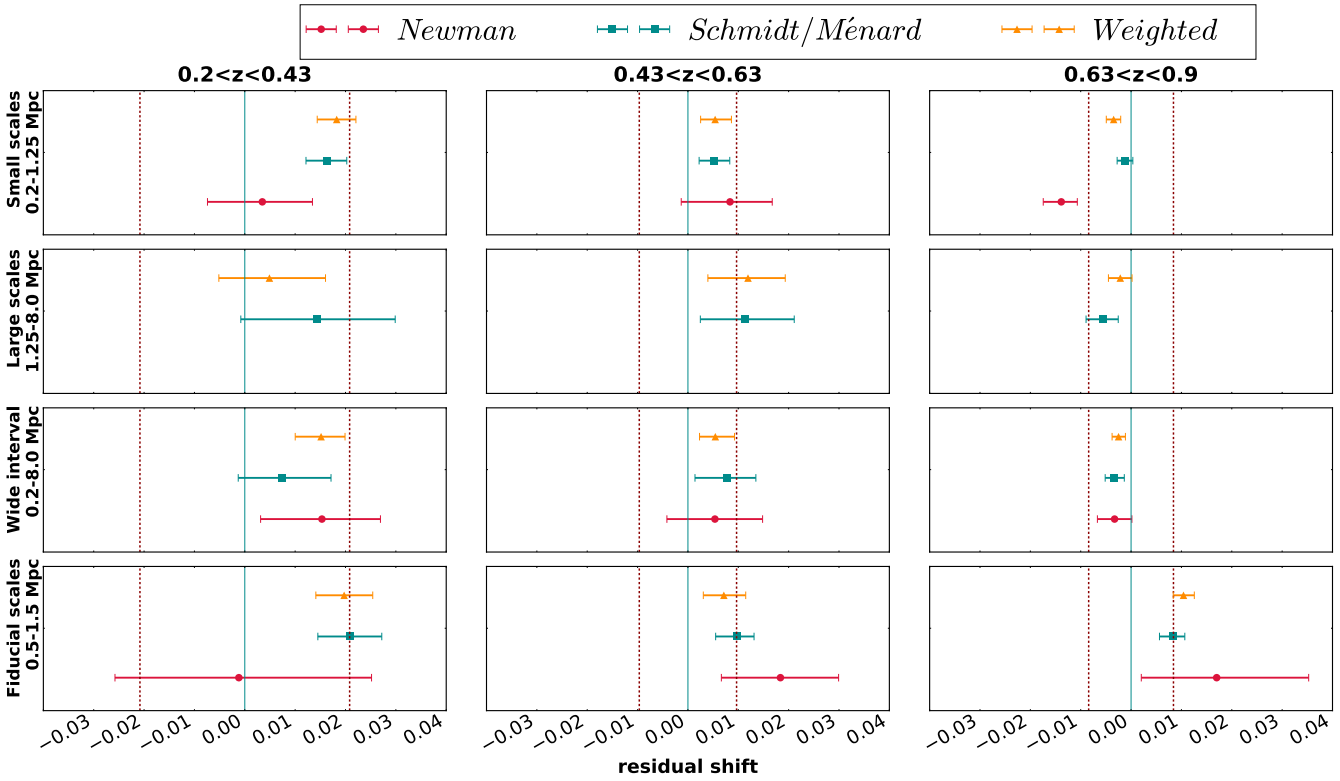

FIG. 3.4.5 – Difference of the value of the **bias evolution** systematic (3.4.2.1) for the different methodologies presented in 3.2, the three different tomographic bins (left to right) and different integration scales (top to bottom)

3.4.2.5 On the choice of the integration scales and the clustering-redshift method

An important parameter that should be set in clustering-redshift methods is the range of scales in which we will integrate our cross-correlation signal. Indeed, intuitively the largest scales will tend to be noisier, and the smaller scale might suffer from non-linear effects (although Schmidt et al. (2013) showed that non-linearities do not have a large impact in the reconstruction). In order to optimize our methodology, before deciding on the scale that will be later on used as fiducial, we have computed the variation in our dominant systematic (the bias evolution systematic) integrating over different ranges of scales for each of the clustering methods presented in section 3.2. In Fig.(3.4.5), the value of this systematic is showed for a small-scale integration range (0.2 – 1.25 Mpc), a large-scale integration range (1.25 – 8.0 Mpc), a wide interval integration range (0.2 – 0.8 Mpc), and finally our fiducial range of integration (0.5 – 1.5 Mpc).

In this figure, one can see that the Schmidt/Ménard method of section 3.2.1.1 and the weighted method of section 3.2.1.2 give similar results, and we observe small differences in the averaging over scales. On the other hand, the Newman method shows bigger differences caused by the fact that in the Newman methodology, one assumes a proportionality relation between the bias evolution of the two samples ($b_u(z) \propto b_r(z)$) which is not a good approximation for the considered samples (see the two bottom panels of Fig. (B.0.1)).

3.4.2.6 On the choice of the window function

In section 3.3.1.2, it has been explained that in order to increase the S/N we were matching the mean within a $\pm 2\sigma_{WZ}$ interval around the clustering-redshift distribution mean (avoiding in this sense the noise that could appear in the tails, or additional systematic effect due to lensing magnification, see appendix C). The choice of this interval had also to be tested, since depending on the shape of the true distribution, we want to cut a maximum of the tails, but a too small interval could be too restrictive and we would lose information specially in broad or bi-modal distributions. To decide on this interval, we re-evaluate each of the systematic errors exposed before using the mean matching criteria with different number N of intervals $\pm N * \sigma_{WZ}$ around the clustering-redshift mean $< z >_{WZ}$. Since the true distribution is differently broad for the three tomographic bins, we probed different values of N for each of them and the results are shown in Fig.(3.4.6). From the figure, one can see that the difference in the systematic values does not vary much with respect to the window size (order of ~ 0.005 variation), yet we can observe differences in the first bin, if the window is too large ($N \sim 3.5$). This is due to the fact that there it includes the positive tail at high-redshift of the reconstructed distribution (see Fig.(3.4.4)) and affect the bias systematic. On the other hand, if the true distribution is broader than the window size, a non-negligible part of the true distribution won't be taken into account and this will increase the shape systematic, as it is the case for the second bin: when $N \sim 1.5$, the second peak of the BPZ true distribution is cut and we observe a change in the shape systematic with N. In order to account for these effects, we decide for each bin to take as systematic the larger value obtained between mean matched in the $\pm 2\sigma_{WZ}$ and $\pm 2.5\sigma_{WZ}$ windows.

3.5 Dark Energy Survey results

In the previous sections, we have presented our fiducial methodology for clustering-redshift calibration and how we have been using simulated galaxy catalogs to evaluate the different systematic errors that we are subject to. Here, we will now expose the results obtained in observed galaxy catalogs using our scheme.

In the context of the cosmological analysis using weak lensing (Troxel et al. (2017); DES Collaboration et al. (2017)), several weak lensing catalogs have been created, and they will be described in section 3.5.1, then the results of the clustering-redshift calibration for these samples will be shown in section 3.5.2, followed by a comparison of results on observed catalogs to those obtained in simulations, and finally we will present two additional tests, performed to validate our methodology using observed data, namely the shear-ratio test from Prat et al. (2017) and the COSMOS 30 band test from Hoyle et al. (2018) in section 3.5.3.

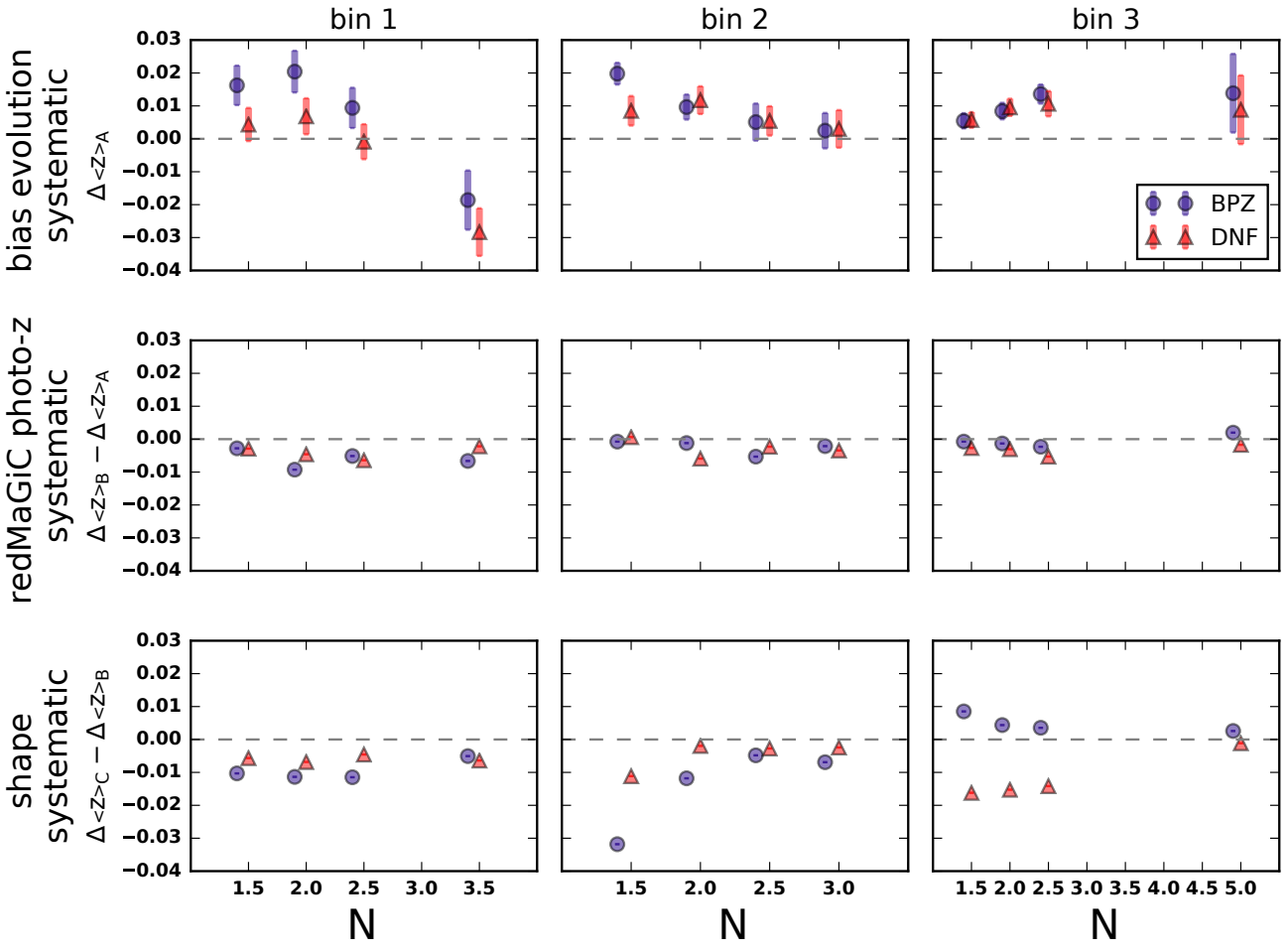

FIG. 3.4.6 – Changes on the three different systematic error value (from top to bottom) when varying the number N of intervals $\pm\sigma_{WZ}$ of the window function ($< z >_{WZ} \pm N\sigma_{WZ}$) for the mean matching case (see 3.3.1.2) and the three WL tomographic bins (from left to right).

3.5.1 Shape Catalogs

In section 2.4.3, we have mentioned that the catalog of galaxy shapes that enters the cosmological analysis is the `METACALIBRATION` (Huff & Mandelbaum (2017); Sheldon & Huff (2017)) catalog. Beside this, in order to test the robustness of the cosmological results, the DES collaboration produced an additional shape catalog, the `IM3SHAPE` (Zuntz et al. (2013)) catalog. To estimate the photometric redshift of each galaxy in these two catalogs, two distinct ways to measure galaxy fluxes have also been probed, the `MOF` (multi-object-fit) fluxes produced by the `NGMIX` [2] algorithm (see Drlica-Wagner et al. (2018)) and the `METACAL` fluxes obtained with the `METACALIBRATION` pipeline.

[2] https://github.com/esheldon/ngmix

From here, it is possible to build different samples varying the measurements and catalogs. In particular, after choosing a catalog of galaxy shapes, the process to build the different catalogs with photometric redshifts can be divided into two steps :

(i) select a flux estimation and a photo-z code to bin the sample in tomographic bins;

(ii) select a flux estimation and a photo-z code to reconstruct the photo-z posterior in each bin.

Once these two steps have been completed, the photometric distributions are inferred by stacking the probability distributions of galaxies as :

$$n^i_{PZ}(z) = \frac{\sum_{j\in i} w_j P_j(z)}{\sum_{j\in i} w_j}, \tag{3.20}$$

where i represents the tomographic bin, j the galaxies assigned to bin i, w_j are weights given by the shape algorithms and $P_j(z)$ is the probability of galaxy j (in i) to have a redshift z. Table 3.5.1 describes the different catalog used in the context of the DESY1 cosmological analyses.

Note that in Hoyle et al. (2018); Troxel et al. (2017); DES Collaboration et al. (2017), the two fiducial catalogs used for cosmological inference or to probe the robustness of parameter estimation are, respectively, the two catalogs presented in the third column of Table 3.5.1, namely one `METACALIBRATION` catalog where `BPZ` was run on `METACAL` fluxes to assign galaxies to tomographic bins and on `MOF` fluxes to infer the redshift distribution of each bin, and an `IM3SHAPE` catalog using `MOF` fluxes for both bin assignment and posterior reconstruction as input for `BPZ`. Moreover, in Hoyle et al. (2018) a sample running `DNF` on `METACAL` fluxes for both bin assignment and posterior estimation on the `METACALIBRATION` shape catalog has been used to verify the impact of the photo-z code in the reconstruction.

By combining the different criteria presented here, it is possible to build a larger variety of catalogs than the ones presented in Table 3.5.1, and these possible catalogs have been in a second step calibrated using the clustering-redshift based methodology. The results of these calibrations are presented in Davis et al. (2018). In the following sections, we will focus on the three samples presented in Table 3.5.1,which are the catalogs entering the cosmological analysis (or used as cross-check to emphasize the robustness of our cosmological inference), and will show the main results applying our calibration as well as the `COSMOS` calibration that will be summarized in section 3.5.3.

Shear catalog	process	BPZ	DNF
METACALIBRATION	bining with	METACAL	METACAL
	stacking with	MOF	METACAL
IM3SHAPE	bining with	MOF	-
	stacking with	MOF	-

TABLE 3.5.1 – Overview of the different processes to estimate the photo-z posterior of the two different shape catalogs (METACALIBRATION and IM3SHAPE) giving a set of three different WL samples that have been calibrated using clustering-redshift methodology

3.5.2 Clustering-redshift methodology applied to DESY1 WL samples

3.5.2.1 Results

In this section, we present the results obtained applying our fiducial scheme to the weak lensing samples described in Table 3.5.1 using as reference sample the observed *redMaGiC* higher-luminosity sample until redshift 0.85 (to be coherent with the simulations, see section 3.3.2.3)[3] . The fifth column of Table 3.5.2 summarizes the different correction shifts recovered for these samples (the errors in the table include both systematic and statistical errors) and Fig. 3.5.1 shows the redshift distributions for the various samples. In the figures, the three tomographic bins are pictured with different colors, the points with error bars correspond to the clustering-redshift reconstruction, where the empty squares represent the points excluded in our matching procedure by the window cut $\pm 2\sigma_{WZ}$ (see section 3.3.1.2), and the lines represent the photo-z posterior (dashed) and the final redshift distribution of our sample after calibration (solid). As it can be seen in the figure as well as in the table, the photo-z codes and clustering-redshift reconstruction agree quite well for the first and third bins (shifts consistent with zero in the table). This is not the case for the second bin, where the shift becomes more important due to the difference in shape of the two distributions (clustering-redshift and photo-z). There we obtained shifts as big as -0.023 ± 0.026, -0.031 ± 0.017 and -0.037 ± 0.014 for the different catalogs.

3.5.2.2 Comparison with results in simulated catalogs

In section 3.2, we have presented how the method systematic errors as well as the method parameters were determined using cosmological simulations. It has also been mentioned in section 3.3.2 that even though both simulated and observed catalogs look alike in most aspects (such as *redMaGiC* redshift bias and scatter), there exist differences between them (eg. density of objects and redshift coverage). However, it is also possible, using the observed sample, to verify

[3] We also have been weighting our observed *redMaGiC* galaxies using the weights taking into account observational systematics and their correlation derived in Elvin-Poole et al. (2017), but it has been verified that more weights have minor consequences in the final result.

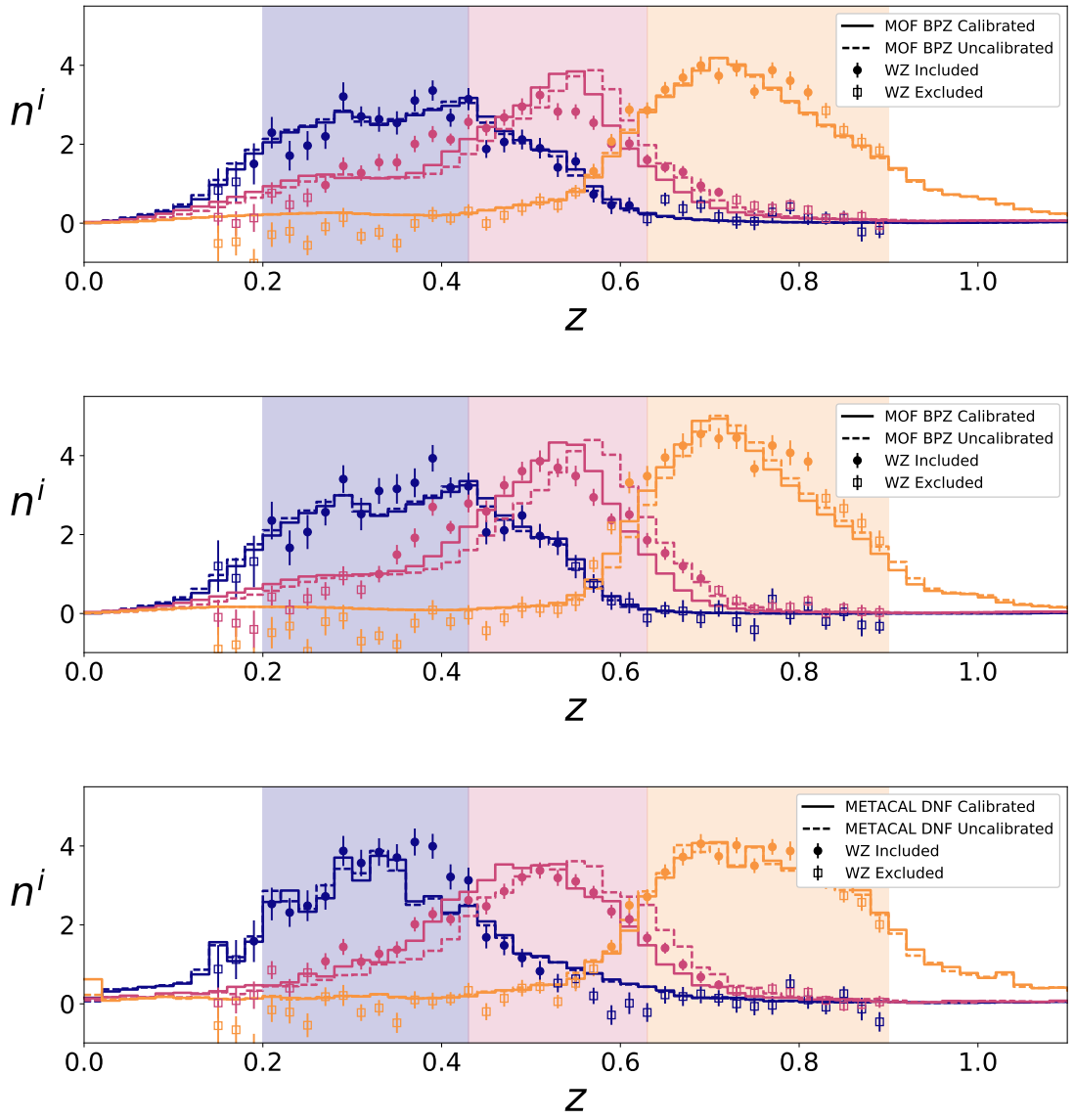

FIG. 3.5.1 – Redshift distributions of the three first tomographic bins (color) of the DESY1 WL samples presented in Table 3.5.1. The error bars correspond to the redshift distribution inferred with the clustering-redshift technique, the empty squares representing the points not taken into account in the mean-matching process (outside the $\pm 2\sigma_{WZ}$ window), the dashed lines are the initial photo-z posteriors and the solid lines the calibrated redshift distribution (figures from Davis et al. (2018))

Shape Catalog	Photo-z	z_{min}	z_{max}	Clustering-redshift Correction	COSMOS Correction	Combined Correction
METACALIBRATION	METACAL/MOFBPZ	0.20	0.43	$+0.007 \pm 0.026$	-0.006 ± 0.020	-0.001 ± 0.016
		0.43	0.63	-0.023 ± 0.017	-0.014 ± 0.021	-0.019 ± 0.013
		0.63	0.90	$+0.003 \pm 0.014$	$+0.018 \pm 0.018$	$+0.009 \pm 0.011$
		0.90	1.30	-	-0.018 ± 0.022	-0.018 ± 0.022
IM3SHAPE	MOFBPZ	0.20	0.43	$+0.008 \pm 0.026$	$+0.001 \pm 0.020$	$+0.004 \pm 0.016$
		0.43	0.63	-0.031 ± 0.017	-0.014 ± 0.021	-0.024 ± 0.013
		0.63	0.90	-0.010 ± 0.014	$+0.008 \pm 0.018$	-0.003 ± 0.011
		0.90	1.30	-	-0.057 ± 0.022	-0.057 ± 0.022
METACALIBRATION	METACALDNF	0.20	0.43	$+0.003 \pm 0.014$	-0.024 ± 0.017	-0.008 ± 0.011
		0.43	0.63	-0.037 ± 0.014	-0.042 ± 0.021	-0.039 ± 0.012
		0.63	0.90	$+0.005 \pm 0.019$	$+0.006 \pm 0.021$	$+0.006 \pm 0.014$
		0.90	1.30	-	$+0.038 \pm 0.020$	$+0.038 \pm 0.020$

TABLE 3.5.2 – Result of the shift to apply to the photo-z posterior of the different samples constructed using the criterion presented in section 3.5.1, the red cells being the final correction applied and entering the cosmological results of Troxel et al. (2017) and DES Collaboration et al. (2017).

the results obtained in simulations and cross-check the choices made on the model parameters. To do so, the idea is to vary the method parameters that have shown to have an impact on the recovered correction shift in simulations, and compare the impact of such variations on the correction shifts in both observed and simulated catalogs.

Firstly, similarly to what has been done in section 3.4.2.5, one can re-compute the correction to apply in the observed sample varying the range of integration scales. Thus, the correction shifts have been re-estimated integrating over different scale ranges (200 – 1250 Mpc, 1250 – 8000 Mpc, 200 – 8000 Mpc). Note that, here it wouldn't be correct to compare the value of the correction shift Δz obtained in simulated and observed samples, because we do not expect them to be the same due to difference in the spectral energy distributions used to generate the simulation and the observed one, but we want to compare the changes in this shift with respect to the integration scales between observed and simulated sample for each bin. The main differences found appeared when we are integrating over large scales ($1250 - 8000 Mpc$) where we observed a variation in the recovered shift of ~ 0.03 and ~ 0.02 for the two first bins. When integrating over this range of scales we obtain a clustering-redshift reconstruction noisier and with more correlation between tomographic bins. This test is emphasizing our fiducial choice of scales ($500 - 1500 Mpc$)

Furthermore, it is also possible to evaluate the influence of the window function size for the mean matching procedure in the observed samples as it has been done for the simulated ones in section 3.4.2.6, there, we shown that the size of the matching window has an impact on the reconstructed mean and we identified this effect as a systematic effect (see section 3.4.2.3). The same comparison could be repeated using observed data and compare the value of the calibration shift for different window sizes. By doing this, we have found that the trend on the change on the shift value recovered is equivalent of the one obtained in simulation, however the difference are stronger, due to the fact that the distribution in the observed sample appear to be much broader than the simulated ones (see figure 3.3.2 and 3.5.1), in particular for the second tomographic bin where the difference on the shift between a mean matching within a $2\sigma_{WZ}$ window and $2.5\sigma_{WZ}$ doubles from simulation to observed catalogs. However, we could note that this variation depends on the shape of the distribution, and the simulated distributions shapes are much narrower than the one recovered in observed catalogs having as consequence a stronger dependence on this cut. Also because of this broad shape, the $2.5\sigma_{WZ}$ window in the second bin is enclosing already all the clustering-redshift reconstructed points (thus the points on the tails of the distribution that suffer more from noise and systematic effect due to lensing magnification see appendix C), we therefor decided to match the two sample means within a window of $2\sigma_{WZ}$.

3.5.3 Additional independant tests

Finally, the methodology presented in this chapter, has also been verified using two other independent tests, the COSMOS-30 bands photometry test Hoyle et al. (2018) and the shear-ratio test Prat et al. (2017). This paragraph will give an overview of these to extra tests as well as a comparison with the results obtained with clustering redshift calibration.

- **The COSMOS redshift distribution calibration :** COSMOS2015 (Laigle et al. (2016)), is a catalog of more than half a million galaxies in a $\sim 2deg^2$ patch of the sky, where the redshift probability distributions were defined using LePhare template-fitting code (Arnouts et al., 1999; Ilbert et al., 2006) on the 30 observed photometric bands with an high accuracy redshift estimation with $\sigma_{/(1+z)\sim 0.01}$[4] for the galaxy that will be use in this test. To calibrate our sample redshift distributions, a sub-sample (of 200, 000 objects) of the DESY1 WL samples galaxies have been matched in color and size (and not in position) to the COSMOS objects. The redshift of these COSMOS matched object has been in a second step re-computed using the DESY1 redshift estimation framework of Hoyle et al. (2018), and compared to the one provide by the 30-band photometry. The difference between the means of the two distributions corresponds to the photo-z bias.

- **The shear-ratio test :** first suggested as a cosmological probe in Jain & Taylor (2003), is based on the fact that the ratio between two tangential shear signals of two different source bins around the same lensed objects only depends on the angular diameter distance and redshift distribution of the concern samples (see section 2.4.4). Nonetheless this ratio appeared to be more sensitive to errors made in estimating the redshift distributions of the concerned samples and in the shear calibration errors than in cosmological parameters. In Prat et al. (2017) it has been shown that the dependence on the redshift distributions could be use as a tool to estimate the photo-z bias of the source sample. Indeed it is possible assuming a cosmology and using the photo-z estimate of the redshift distributions (there the bias in the *redMaGiC* lenses was not considered) to estimate this ratio and match it to one measured allowing a shift on the means distributions of the source samples (marginalizing over shear calibration errors). Using there the different *redMaGiC* lens bins as for the tangential shear measurements (see 2.4.4) one can measure the photo-z bias of the WL sample sources bins.

In principle the recovered shift obtained using these two supplementary tests should be consistent with the one obtained using our clustering-redshift calibration. Figure 3.5.2 resumes the different corrections recovered for each of the tomographic bin using the three calibration tests. Note that in the case of the shear-ratio test, the precision tends to decrease quickly as one goes to higher

[4]this errors concern the galaxies that will be use for this test

bin, this is due to the fact that the relative distance between the lenses and the observer is getting smaller and thus the lensing efficiency is decreasing for these bins, because of this, the higher tomographic bin has not been considered, the only constraint we have on this last bin is then coming from the COSMOS-30 bands calibration.

Nevertheless, in the calibration procedure (Hoyle et al. (2018)), only clustering-based and COSMOS-30 bands corrections have been used, mostly due to the fact that the shear-ratio test depends in a first time in the shear calibration errors (has to be marginalized over) and secondly is covariant with the clustering-redshift techniques since it includes correlation signals of the same galaxy bins and samples.

Assuming the statistical and systematic errors in the COSMOS and clustering-redshift calibration are uncorrelated (Hoyle et al. (2018)), the constraints coming from the two tests have been combined multiplying their Gaussian distributions, and the final result is shown in the last column of table 3.5.2, the red cells in the table correspond to the final correction applied to the fiducial sample entering in the DESY1 cosmological analysis.

3.6 Summary and Conclusions

In this chapter, we started by presenting the overall clustering-redshift based methodology that have been applied in the context of the Dark Energy Survey first year of observation cosmological analysis. In there, clustering-redshift reconstructed distribution has been used to calibrate the bias on the mean of the photo-z posterior of the three lower tomographic bins of the DESY1 WL sample, allowing for a shift of in the z-direction.

We identified the 'Schmidt/Ménard' method integrated over scale in the range of $500-1800 Mpc$ using the *redMaGiC* higher-luminosity sample for reference sample as our fiducial framework to reconstruct the redshift distribution of the DESY1 Weak Lensing sample. Then we opted to calibrate the photo-z posterior distribution matching the means of the two distribution within a window of 2 standard deviations around the clustering-redshift distribution mean ($\pm 2\sigma_{WZ}$).

To select this fiducial framework we developed a methodology in simulated catalogs to identify the different systematic errors that our methodology is subject to, and the underlying framework parameter have been choose in order to minimize these systematic. More specifically, we have been able to distinguish between three different sources of systematic errors: the **bias evolution systematic**, the ***redMaGiC* systematic** and the **shape systematic**. We concluded that the dominant source of error in our methodology came from the evolution in redshift of the bias of our different samples, and more in particular if this evolution has a non-trivial behaviour, as it is the case for our simulated WL sample, this systematic error can be in some tomographic bin

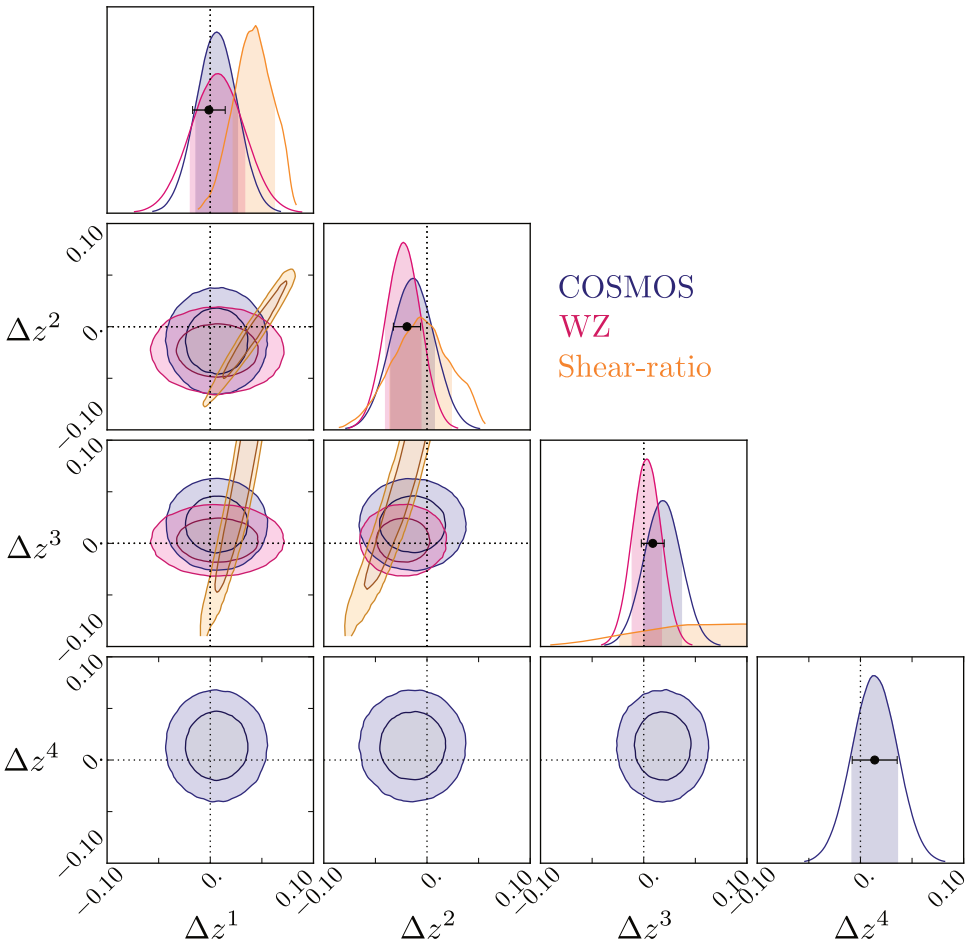

FIG. 3.5.2 – Comparison of the recovered photo-z posterior bias using the three independent tests (different colors) for the different WL sample tomographic bins (figure from Hoyle et al. (2018))

large as ~ 0.02. We have also shown that the systematic errors induced by the *redMaGiC* photo-z was sub-dominant in the total systematic budget, emphasizing here that using good-quality photometric redshift sample as reference could be a good alternative to spectroscopic samples that often give noisy recontruction due to the low number of objects.

After having studying the overall methodology in simulated catalogs and provide its systematic budget, we applied our fiducial scheme to the different observed catalogs created by the DES collaboration. We recover the different correction shifts to apply to the photo-z posteriors for the three tomographic bins considered and combined this correction to an additional independent measurement of these shift using COSMOS-30 bands galaxies. These efforts, lead us to a measurement of the mean redshift with a precision of $\sim \pm 0.015$, which a significant but sub-dominant uncertainties in the total error budget of the DESY1 cosmological analysis.

As exposed here, redshift distributions inferred by clustering-redshift methodology has played for the first time a role in cosmological parameter inference. However future observations of DES (the third year being currently analyzed) as well as future galaxy surveys, the enlargement of the observed areas together with the observed depth improvements will reduce the statistical errors such that the uncertainties in redshift distributions could become the dominant source of errors in the cosmological analysis. Therefore, inferring all more precise redshift distributions will be a challenging aspect of future broad-band photometry surveys, (such as Euclid Laureijs et al. (2011) or LSST Science Collaboration et al. (2009)). And clustering-redshift based methodology could become an important tool for these analysis. In this perspective, improving the methodology to reduce the systematic effects could be useful (if not necessary), and different approaches to do it have been mentioned here (correcting from the bias of the reference sample, taking into account lensing magnification effects,...).

Part III

Cosmic voids vs CMB

Chapter 4

CMB lensing around voids

4.1 Definition and interest

In chapter 1, we have exposed the different tools and probes currently used to attempt to understand what belong behind the mysterious Dark matter and Dark Energy. In there, we have seen that two of the most popular probes for this purpose are in one hand the large structure (see sec. 1.2.2) and on the other hand, the cosmic microwave background (see sec.1.2.4). In this chapter we will present a work where we aim to combine this two probes measuring the lensing imprint of large structures, namely cosmic voids in the cosmic microwave background radiation.

Cosmic voids are define as under-dense extended regions in the cosmic web. These regions, are in fact occupying most of the space of our close universe, they are surrounded by the large structures mentioned before (filaments, galaxy clusters, walls) and have typical size of tens of Mega-parsecs. The fact that these regions are low-density environments, raised a recent cosmological interest in the sense that they are lowly subjects to non-linear effects that high-density regions suffer, and containing low amount of matter, and also represent zones of space where dark energy dominates.

As mentioned above, such as galaxies, galaxy clusters or other large structure observed in our universe, cosmic voids are good candidates for constraining cosmology. However the analysis of these structure had to wait the era of large galaxy surveys to reach the statistical power that allow them to become a cosmological probe.

Similarly to what has been presented in chapter 2, one can also use 2-point correlation statistics to compute the correlation signal between galaxy shape and void position. This is similar to what as been done in 2.4.4 but by substituting the galaxy lens catalog by cosmic void catalog. However, the void-lensing signal are difficult to detect for individual voids (see e.g. Amendola

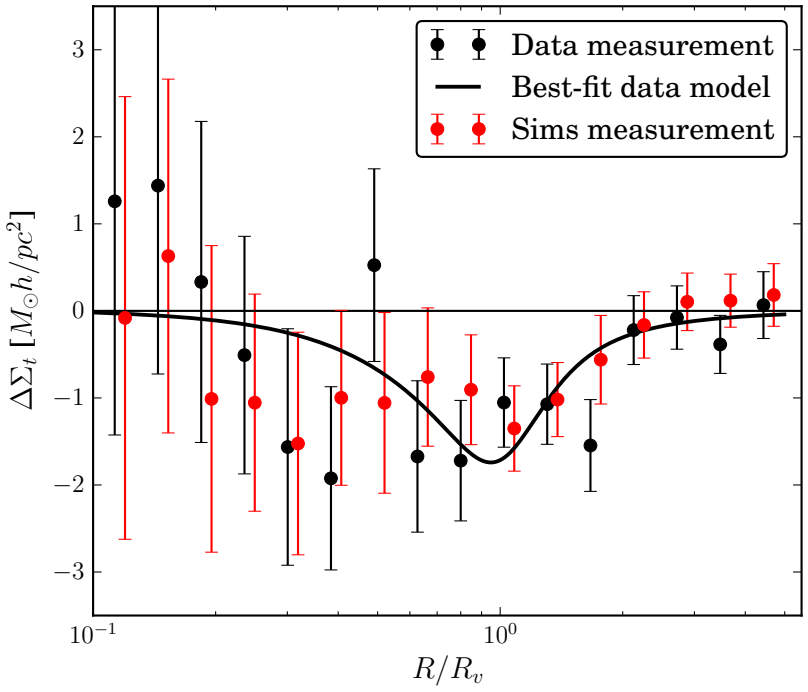

FIG. 4.1.1 – Lensing signal detetected at void location in both simulated catalog (red) and observed DES-SV galaxies (black), from Sánchez et al. (2017)

et al., 1999), but recent work has shown that a stacking methodology could help to increase the signal-to-noise ratio and make the detection possible (Krause et al., 2013; Davies et al., 2018).

The first detected signal of lensing effect due to cosmic voids has been realized Melchior et al. (2014) using ∼ 900 cosmic voids identified in the SDSS photometric catalog with the algorithm. More recently, Sánchez et al. (2017) have used the DES science verification data using a 2.5D void finder. Figure (4.1.1) shows the void-lensing signal measured in both DES science verification data and simulated galaxy catalogs.

It has been exposed in chapter 1, one unsatisfying aspect of the nowadays standard model of cosmology resides on the apparition of Dark Energy and the cosmological constant Λ in the energy budget of our universe after the discovery of cosmic acceleration. Therefore, in the past decades, theorists have imagine alternative cosmological models able to be consistent with cosmic acceleration. One type of these models, is based on the fact that general relativity is not a correct theory in some specific regime the so-called **modified gravity models**. From here, considering that general relativity has been tested with good accuracy in small scales and dense regions such is the solar system, it becomes natural to think as cosmic voids as probe of general relativity since there are define as large-scale under dense regions, and indeed a variety of general relativity models predict different behaviours (than the standard models) in low-density/large-scale regimes.

This chapter will be organized as follows, we will start in section **??** we will describe void properties and introduce void identification giving a spetial attention to the void finding procedure employed in this specific work in 4.2.2.2, then in section 4.3 we will present how combining voids with CMB maps has also been use as an additional probe and the interest of void imprints in the CMB lensing maps. The dataset used in this work as well as the different void catalogs probed will be presented in section 4.4 followed by the methodology applyied in section 4.5. And finally section 4.6 will present the results obtained in this work combining voids detected in the DESY1 catalogs with the CMB convergence map from Planck.

4.2 Void definition

4.2.1 Void properties

As mentioned above, a key aspect in cosmic void analysis resides on the fact that voids are in average spherical, from this, it is possible to look at different features that present these averaged void with respect to redshift or void population. For instance, as exposed in Bertschinger (1985), according to the evolution of perturbations in a Einstein-de Sitter universe, negative perturbations will grow creating a thin overdense region at its surrounding, a **compensation wall**. The density profile of voids of the averaged void can thus be measure, and Hamaus et al. (2014) has proposed an empirical parametrization of this profile as :

$$\frac{\rho_v(r)}{\bar{\rho}} - 1 = \delta_c \frac{1 - (r/r_s)^\alpha}{1 + (r/r_v)^\beta} \tag{4.1}$$

Figure(4.2.1) shows the density profile of different size voids in simulation with respect to the distance from the void center.

4.2.2 Void Finding

4.2.2.1 Finding void in galaxy catalogs

The identification of voids in the cosmic web is an intricate process that is affected by specific survey properties such as tracer quality, tracer density, or masking effects. The void properties also depend meaningfully on the methodology to define the voids (for a review see Nadathur & Hotchkiss, 2015).

In the literature various void-finder have been have been built and run in different void tracers (Dark matter halos, galaxies, cluster of galaxies). In Colberg et al. (2008), the authors are comparing 13 different void finder algorithms using a variety of tracers. In there, the authors had

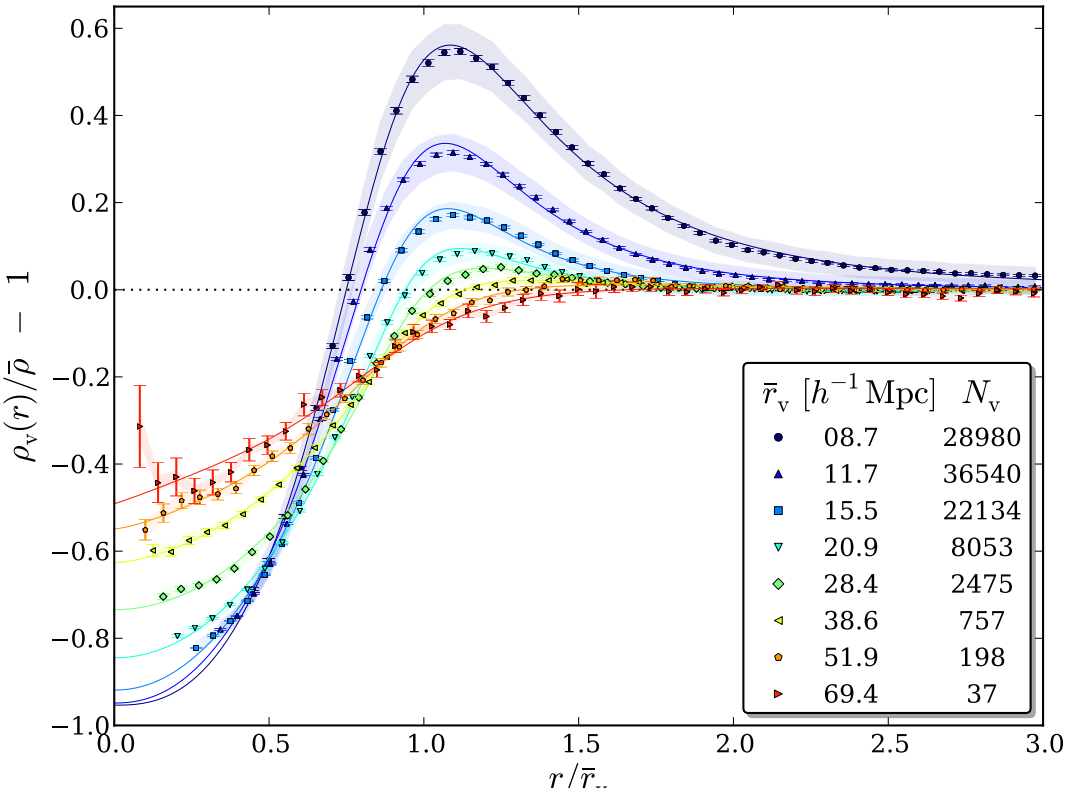

FIG. 4.2.1 – from Hamaus et al. (2014)

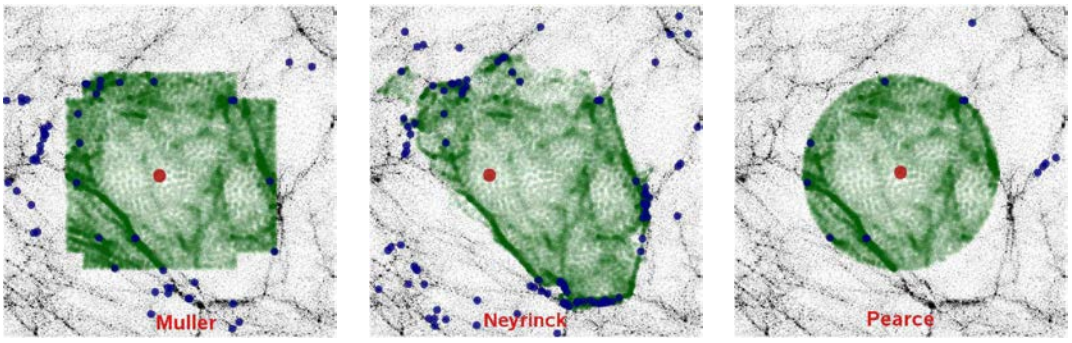

FIG. 4.2.2 – Example of a cosmic void identified by three different void finder algorithms Muller (left panel), Neyrinck (middle panel) and Pearce (right panel) on a $40Mpc/h$ patch of the Millennium simulation, from Colberg et al. (2008)

compared the different void populations characteristics identified running the different algorithms on the Millennium simulation (Springel et al. (2005)). Figure (4.2.2) is an example of a void find by three different algorithms probed by Colberg et al. (2008).

More recently, Cautun et al. (2018) studied how modified gravity models can be tested with different void definitions. They concluded that void lensing observables are better indicators for tests of gravity if defined in 2D projection such as “tunnels" or “troughs". Such scenarios are in

fact the only viable void definitions in photo-z surveys like DES given the significant smearing effect of redshift uncertainties.

4.2.2.2 The 2.5 void finder

In the context of this thesis, similarly to the previous chapter, we have been analysing the Dark Energy Survey observed catalogs and more in particular, we identified voids in the DES first year of observation galaxy catalog, thus treating object with redshift estimate with photometry. Therefore, to create the void catalog, we have been using a modified version of 2.5D void finder presented in Sánchez et al. (2017). This void finder is in a sense a tuned void finder for photometric datas. Indeed, the idea behing the finder is to spot underdensed regions in redshift slices. Due to photometric uncertainties, define the position of cosmic void becomes more complicated. However, finding underdense regions in relatively thin slice of size comparable to photo-z errors allows us to find elongated voids. Namely, Sánchez et al. (2017) showed that significant real underdensities can be identified even using photo-z data in tomographic slices of width roughly twice the typical photo-z uncertainty (see also Pollina et al., 2018, for DES tests and results using other void definitions).

The code works as following :

- divide the sample in redshift slices, $100 Mpc/h$ slices as shown to be a good compromise considering *redMaGiC* redshift accuracy (see Sánchez et al. (2017))
- Compute the density field for each slice by counting the galaxy number in each pixel and smoothing the field with a Gaussian pixel with a predefined smoothing scale.
- select the most underdone pixel and grow around it the void until it reach the mean density
- save the void and erase it from the density map and reiterate the process with the following underdensed pixel

A free parameter in the method is the scale of the initial Gaussian smoothing applied to the projected galaxy density field. For instance, Kovács et al. (2017) found that $\sigma = 20\ h^{-1}$Mpc is a preferable choice for ISW measurements using the whole void sample in the stacking procedure. For weak galaxy lensing measurements with DES voids, however, Sánchez et al. (2017) reported that the smaller $\sigma = 10\ h^{-1}$Mpc smoothing is preferable.

Figure(4.2.3) shows an example of cosmic voids identified using this void finder in on redshift slices of simulated galaxy catalog.

We have also applied the random point methodology of **?** to eliminate voids in the edges with a potential risk of mask effects. This method evaluates the density of random points that have

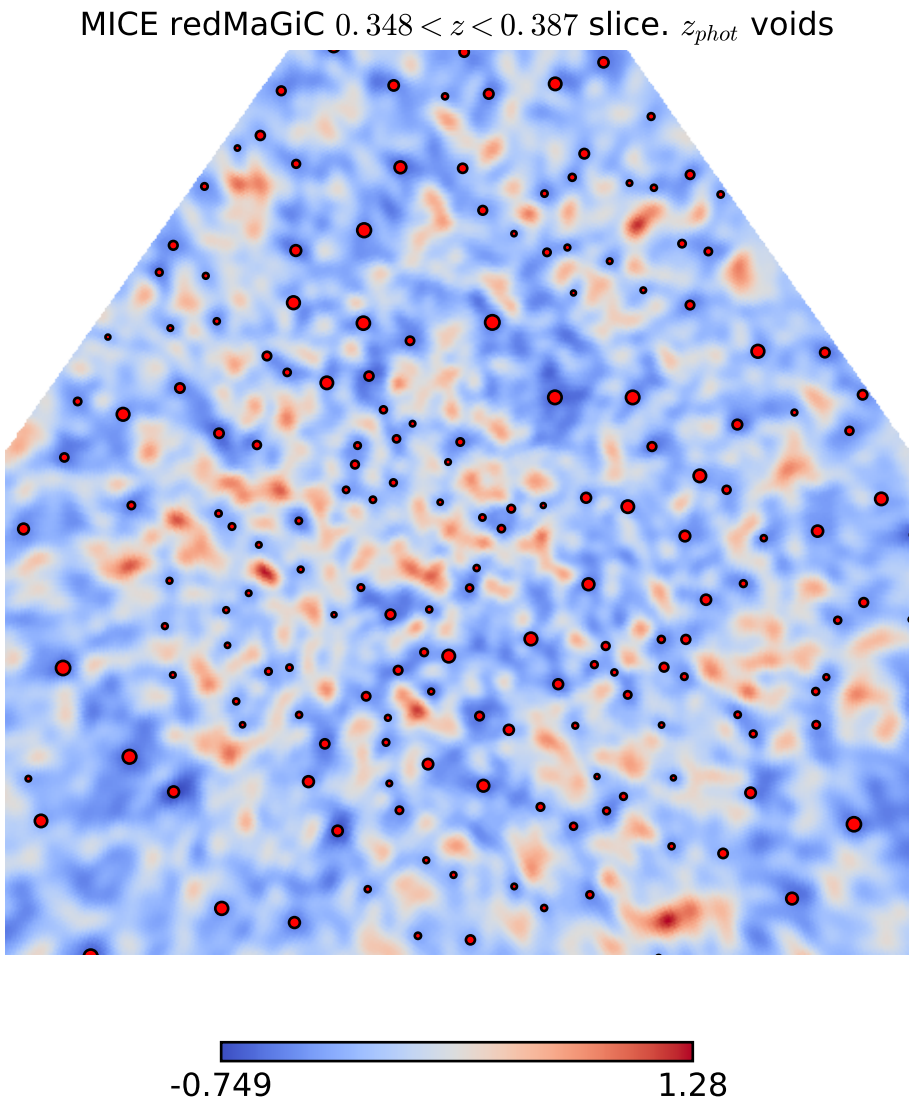

FIG. 4.2.3 – Cosmic voids (red circles) identified in one redshift slice ($0.348 < z < 0.387$) of simulated galaxy catalog, the background color being the smoothed galaxy density field.

been drawn within the mask inside each void. Then, voids with a significant part of their volume laying outside the mask will have a lower random point density and therefore can be identified and excluded. In the final catalog, we have also excluded voids of radius $Rv < 20Mpc/h$ that are expected to be spurious given the photometric redshift uncertainties.

We determine the radii of the voids as well as their redshift (defined as the mean redshift of the slice in which it has been identified). The void finder also provides two additional characteristic quantities related to the under-density of voids:

- *the mean density contrast:* $\bar{\delta} = \rho/\bar{\rho} - 1$ where ρ is the mean density inside the void and $\bar{\rho}$ is the mean density of the corresponding redshift slice;
- *the central density contrast:* The density contrast evaluated at one quarter of the void radius $\delta_{1/4} = \delta(r = 0.25R_v)$.

4.3 Cosmic voids and the CMB

Beside the hereabove exposed different tools to use cosmic voids as cosmological probes, it is also possible to combine the cosmic void found at low redshift to high redshift observation such are the cosmic microwave background radiation observations. In the 80's it has been proposed in Blanchard & Schneider (1987) that similarly to the deflection effects induced by the large structures that we observed on the sky on the photons coming from background galaxies, we should have a lensing effect on the photons that come from the CMB (see section 1.2.4.2). And in principle as is has been presented for cosmic shear cosmology and galaxy-galaxy lensing (see 2.4.3 and 2.4.4), one can infer cosmological informations by measuring the lensing effects of Large Scale Structures on the incoming CMB photons.

Relatedly, the CMB lensing imprint of other elements of the cosmic web have also been measured recently. He et al. (2017) reconstructed the correlation of filamentary structures in the Sloan Digital Sky Survey (SDSS) data and CMB lensing convergence (κ, herafter), as seen by *Planck*. Then, Baxter et al. (2018) stacked the κ maps of the South Pole Telescope (SPT) on locations of galaxy clusters in the Dark Energy Survey first year data set (DES Y1 The Dark Energy Survey Collaboration, 2005), finding good consistency between simulated and observed results.

The prospects of cosmological parameter constraints from CMB lensing probed using cosmic voids are discussed by Chantavat et al. (2016), but, more importantly, the role of void definition, environment, and type have also been studied in simulations. Nadathur et al. (2017) found the parameters of voids, such as their radius (R_v) or their mean under-density ($\bar{\delta}$), are imperfect indicators of the peak positions of the gravitational potential, but their combination is helpful to define meaningful subgroups of voids with different lensing capabilities. Averaging together the contributions from all voids with different type produce an average convergence that is closer to zero and thus potentially harder to measure. This suggests that the sensitivity of detection of void lensing effects could be significantly improved by consideration of sub-populations but the details remain to be explored for such measurements. Figure(4.3.1) shows the CMB lensing profile at void position recovered by Nadathur et al. (2017) for different void populations[1]. As

[1] In Nadathur et al. (2017), the authors identify void populations using both their size and density with the parameter $\lambda_v = \bar{\delta}_g \left(\frac{R_v}{Mpc/h}\right)$

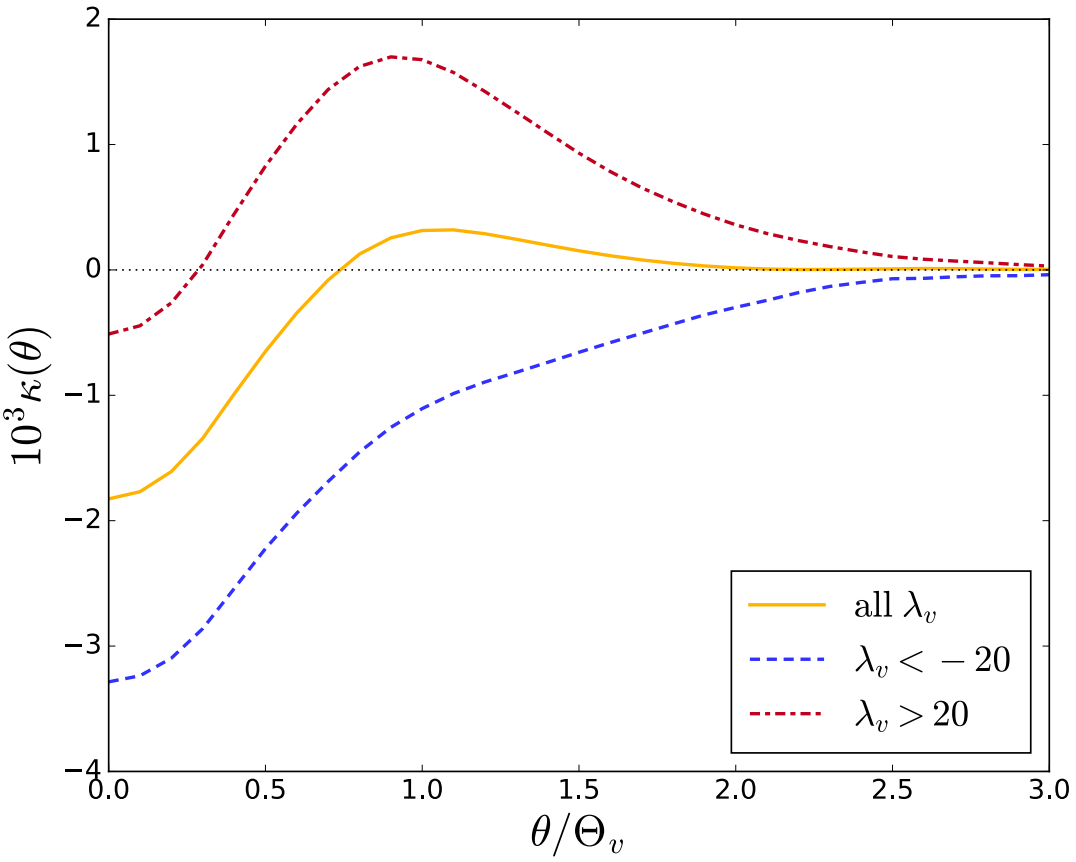

FIG. 4.3.1 – from Nadathur et al. (2017)

it can be see in the figure, smaller λ_v (meaning smaller and deeper voids) induces a stronger negative signal in the CMB lensing profile.

Following their own stacking measurement strategy, Cai et al. (2017) have, for the first time, detected a CMB lensing signal using cosmic voids (catalogue created by Mao et al., 2017) identified in the CMASS galaxy tracer catalogue of the Baryon Oscillation Spectroscopic Survey Data Release 12 (BOSS DR12). They identified voids using the void finder algorithm (Neyrinck, 2008).

Cai et al. (2017) did not optimize the detection of the CMB lensing measurement in their paper because the main point was to complement the stacking measurements of the integrated Sachs-Wolfe effect (ISW) (Sachs & Wolfe, 1967). An evidence for ISW and CMB lensing imprints of the same cosmic voids helps to confirm the reality of each effect. They discuss that an ISW-lensing dual probe is valuable from the point of view of modified gravity, since the two effects are closely related: lensing depends on the sum of metric potentials, whereas ISW depends on their time derivative. Hints of the general coexistence of both the ISW and CMB lensing signatures have been found by the *Planck* team (Planck 2015 results. XXI., 2016).

More importantly, the amplitude of the CMB κ imprint of voids is expected to shed new light on the problem of excess ISW signals coming from the largest voids. In the strength of the signal, typical cosmic voids have shown good consistency with expectations, given the substantial cosmic variance (see Nadathur & Crittenden, 2016, for a comprehensive summary). Discordantly, large-scale hills in the gravitational potential, or *super*voids, have shown excess signals (see e.g. Granett et al., 2008; Cai et al., 2014b; Kovács, 2018). It is typically assumed that neither modifications of the concordance model, given other precise constraints, nor e.g. simple modified gravity scenarios seem to alleviate the ISW tension (e.g. Cai et al., 2014a). If confirmed elsewhere in the Universe, or detected in other void observables, such as CMB lensing imprints, these excess signals pose an important challenge for the ΛCDM model (Nadathur et al., 2012).

Cai et al. (2017) reported a CMB lensing signal of BOSS voids that is compatible with simulated imprints, with somewhat higher-than-expected signal in the center of the voids. They did not perform a comprehensive stacking analysis of void sub-groups. They argued, however, that their sample is dominated by rather small voids surrounded by over-densities (voids-in-clouds) which presumably are suboptimal to detect CMB κ imprints (see again Nadathur et al., 2017). The conclusion was that the puzzling excess ISW signal was seen in the BOSS DR12 data, especially for the most significant big and deep voids, but the lensing counterpart seemed inconclusively noisy with hints of an excess imprint.

In Kovács et al. (2017), we have recently attempted to probe these claims in the the first year footprint of the Dark Energy Survey and identified 52 voids and 102 superclusters at redshifts $0.2 < z < 0.65$ using the void finder tool described in Sánchez et al. (2017). The heart of that method is a restriction to 2D slices of galaxy data, and measurements of the projected density field around centers defined by minima in the corresponding smoothed density field. Similarly to the Granett et al. (2015) analysis of SDSS super-structures, our tests revealed a significant mean line-of-sight elongation for the super-structures that is caused by the photo-z uncertainties. All in all, we found a $\Delta T \approx -10\ \mu K$ cold imprint of voids, formally with $A_{\rm ISW} \approx 8 \pm 6$ ISW amplitude[2], that is 1.2σ higher than the imprint of such super-structures in the simulated ΛCDM universe. We also found $A_{\rm ISW} \approx 8 \pm 5$ for superclusters and in combination we constrained $A_{\rm ISW} \approx 8 \pm 4$ with DES Y1 super-structures, i.e. a 2σ discrepancy. These measurements, although hinting again at a large ISW amplitude, were indecisive because of the significant noise level.

4.4 Data sets

In the context of this thesis, we aimed to measure the lensing signal on the CMB lensing map produced using the Planck satellite observations (Planck Collaboration et al. (2016c)) at the location of voids found in the DESY1 galaxy catalog. And more in particular, since an important

[2]In a ΛCDM cosmology, by definition $A_{\rm ISW} = 1$.

source of error that affects our void finding procedure is photo-z uncertainty; photometric DES data does not provide a precise redshift estimate for the galaxy tracers of voids as spectroscopic survey does, we have chose to use LRG identified by the *redMaGiC* algorithm similarly to what has been presented in 2.3.1. This analyses, as it will be developed bellow, has been at first tested and optimized using the simulated galaxy catalog namely the MICE-GC catalog, on which the *redMaGiC* algorithm has been ran and a CMB lensing map is available.

4.4.1 Simulations - the MICE galaxy mock and κ map

The MICE simulated sample is a N-body light-cone from the *MICE Grand Challenge* (MICE-GC), that contains about 70 billion dark-matter particles in a $(3h^{-1}Gpc)^2$ comoving volume. The details on the creation of this simulation can be found in Fosalba et al. (2015a); Crocce et al. (2015); Fosalba et al. (2015b), here we will give a brief overview. MICE has been developed at the Marenostrum supercomputer at BSC[3] running the GADGET2 (Springel, 2005) code. The simulation assumed a flat standard ΛCDM model with input fiducial parameters $\Omega_m = 0.25$, $\Omega_\Lambda = 0.75$, $\Omega_b = 0.044$, $n_S = 0.95$, $\sigma_8 = 0.8$ and $h = 0.7$ from the Five-Year Wilkinson Microwave Anisotropy Probe (WMAP) best fit results (**?**). Moreover, the catalog has been created and validated to follow local observation constraints such as luminosity functions, galaxy clustering (with respect to different galaxy populations) and color-magnitude diagrams.

The *redMaGiC* algorithm has been run in the MICE mock galaxy catalogue with the same algorithms applied to the observed DES Y1 observed catalog. We utilize this MICE-*redMaGiC* mock galaxy catalogue to trace the large-scale galaxy distribution and to identify cosmic voids.

The lensing maps (Fosalba et al., 2015b) corresponding to the MICE simulation have been produced using the 'Onion Universe' methodology presented in Fosalba et al. (2008). An all-sky map has been constructed by cloning the simulation box from the MICE simulation and translating it around the observer, and afterwards splitting the light cone into concentric shells of size $dz \sim 0.003(1+z)$ and an angular resolution of $\Delta\theta \sim 0.85\ arcmin$. Moreover, the map produced has been validated using different lensing probes such as auto and cross-correlations with foreground MICE galaxy and dark matter particles (see Fosalba et al. (2015b) for details). Figure (4.4.1) is showing the full CMB lensing convergence map from the MICE simulation.

The observed lensing convergence map used in this analysis is the one provided by the Planck collaboration Planck Collaboration et al. (2016c). This map has been built using a quadratic estimator, with the lensing reconstruction methodology described in Okamoto & Hu (2003), to the foreground-cleaned CMB temperature and polarization maps build using the component separation method (Planck Collaboration et al. (2016a)). The map is provided with a pixel resolution of $N_{side} = 2048$ which corresponds to a pixel size of 1.7 $arcmin$.

[3]Barcelona Supercomputing Center, www.bsc.es.

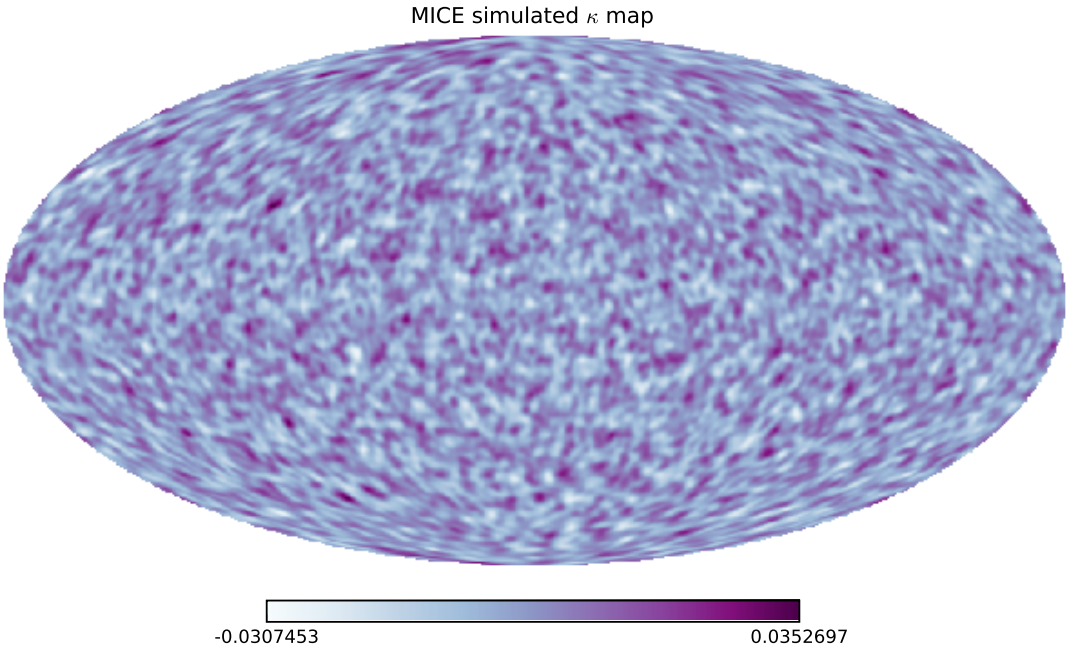

FIG. 4.4.1 – `healpix` map of the CMB lensing signal of the full sky using the MICE simulation.

4.4.2 The different void catalogs

4.4.2.1 Cosmic voids in the MICE galaxy mocks

We examine how potential systematic effects modify the resulting void populations. We compare the void parameter distributions for different tracer densities and various initial Gaussian smoothing applied to the density fields.

Edge/mask effects may lead to different mean void properties as at survey boundaries the full extent of underdense regions around minima may not be captured with good precision. Our initial tests revealed that the indeed the voids identified using *redMaGiC* tracers in the MICE octant have different properties compared to void properties of DESY1-like survey patches inside the octant. We therefore used the official mask used in the DESY1 cosmological analysis (Elvin-Poole et al. (2017)). We note that it is possible to put two independent DESY1-like patches on the MICE octant. Therefore, to gain statistical power and more precision on our expected void lensing signals, we will study two MICE Y1-like void catalogues.

We thus run our void finder using two different *redMaGiC* samples as tracers. The *redMaGiC* high-luminosity sample applies a stronger cut in luminosity ($L > 1.5L_*$) which offers higher precision in photometric redshift. On the other hand, the *redMaGiC* high-density sample has a more relaxed luminosity cut ($L > 0.5L_*$), resulting in an increased galaxy density. We then

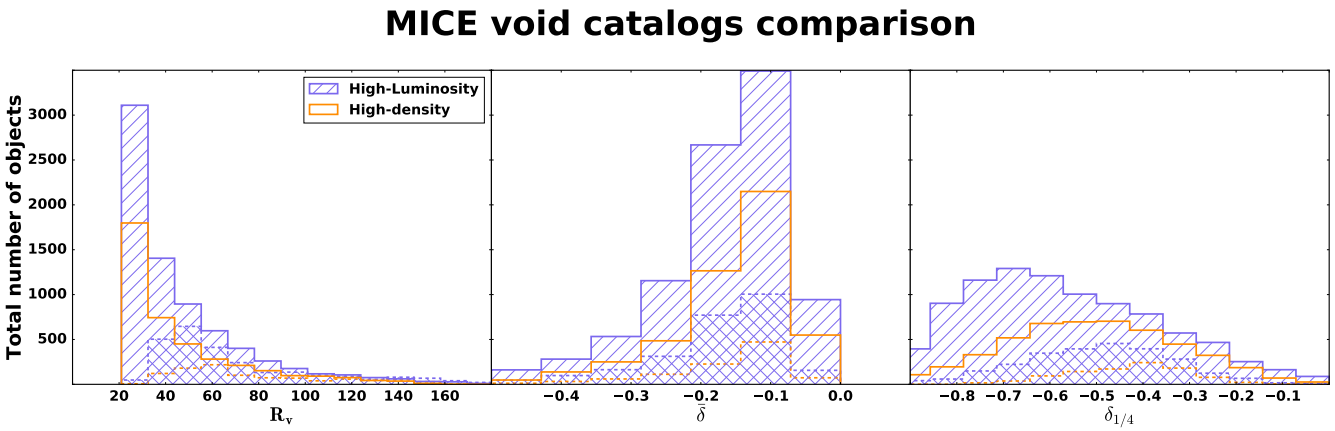

FIG. 4.4.2 – Comparison of the different void catalog obtained using different tracers, the *redMaGiC* High-Luminosity sample (purple) and *redMaGiC* high-density sample (orange) and different smoothing scale for the void finder $10 Mpc/h$ (solid lines) and $20 Mpc/h$ (dashed lines)

further probe systematic effects by running the void finder on these two two samples using different initial Gaussian smoothing scales, namely 10 Mpc/h and 20 Mpc/h. Figure 4.4.2 shows a comparison of the resulting void catalogs for the different tracer density and smoothing scales. We compare the void catalogues in terms of three characteristic parameters of voids; distribution in size (R_v), distribution of mean density ($\bar{\delta}$) and distribution in central void density ($\delta_{1/4}$). We observe the following properties:

- Comparing the different resulting catalogs a higher number of voids is detected when the tracer density is lower (*redMaGiC* high-luminosity sample). Such effects are expected and have already been studied in simulations (see e.g. Sutter et al., 2014).
- A larger smoothing scale decreases the number of voids for both tracer densities.
- The mean void radius is shifted towards larger values for larger smoothings, as potential merging of smaller voids into larger encompassing voids is expected.
- The previous effect is also manifested in the density contrast distributions. Smaller voids tend to be deeper and therefore the excess of small voids in the catalogs constructed using 10 Mpc/h smoothing is observable as a shift towards deeper mean density contrasts in the population.

We aim to test how such competing effects found in alternative void samples influence the signal-to-noise (S/N) of CMB lensing × MICE voids.

4.5 Simulated cross-correlation analyses

4.5.1 Stacking methodology

The CMB lensing imprint of single voids is practically impossible to detect (see e.g. Krause et al., 2013). We therefore apply an averaging method using cutouts of the CMB map at void positions (see e.g. Kovács et al., 2017, and reference therein). This stacking procedure can be described with the following steps:

- we define a catalogue of voids and potentially select subgroups to probe their specific imprint type;
- we cut out patches of the CMB convergence map centered at the void center position using `healpix` tools (Górski et al., 2005);
- we re-scale the patches given the angular size of voids;
- we smooth the CMB convergence map with an arbitrary smoothing scale of 1° to remove noise at scales smaller than the typical void size;
- we stack all patches and measure the average signal in different concentric radius bins around the void center.

More formally, we rely on the following formula in the stacking to express the size of the patch (S_p) in pixels.

$$S_p = 2 * N * \frac{R_v}{n_{\rm pixels}} \tag{4.2}$$

where $n_{\rm pixels}$ is the required resolution in pixel numbers, and N the number of void radii we wish to enclose in each patch. We chose a quite large relative patch size enclosing five void radii. Such choice is helpful to detect the lensing imprint of void surroundings such as the compensation wall that may appear in the outer void density profile. Two examples of our stacked images are shown in Figure 4.5.1. We do not apply any other filtering in the stacking procedure such as exclusion of large-scale modes up to $\ell < 10$ (see Cai et al., 2017, for related results).

4.5.2 Optimization of the measurement

The imprint of voids on the CMB lensing maps depends on their properties. Nadathur et al. (2017) showed that simulated cosmic voids, identified with the methodology, trace the peaks of the underlying gravitational potential differently given different density, size, and environment. They reported that voids can be grouped based on a combined density-radius observable to have distinct lensing profiles. In particular, they found that the more numerous compensated

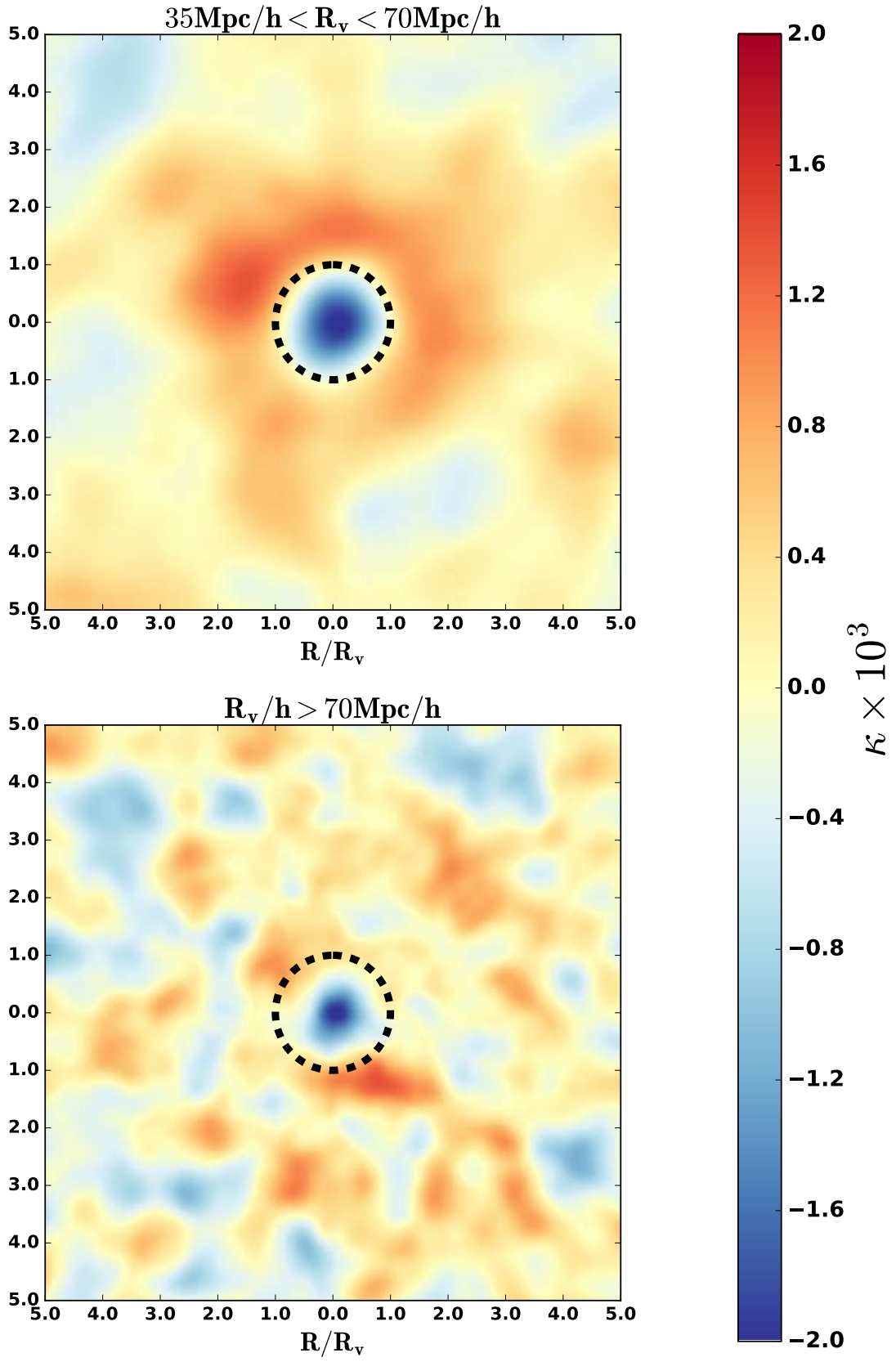

FIG. 4.5.1 – Examples of the stacked lensing convergence signal obtained for medium size voids (top) or large voids (bottom).

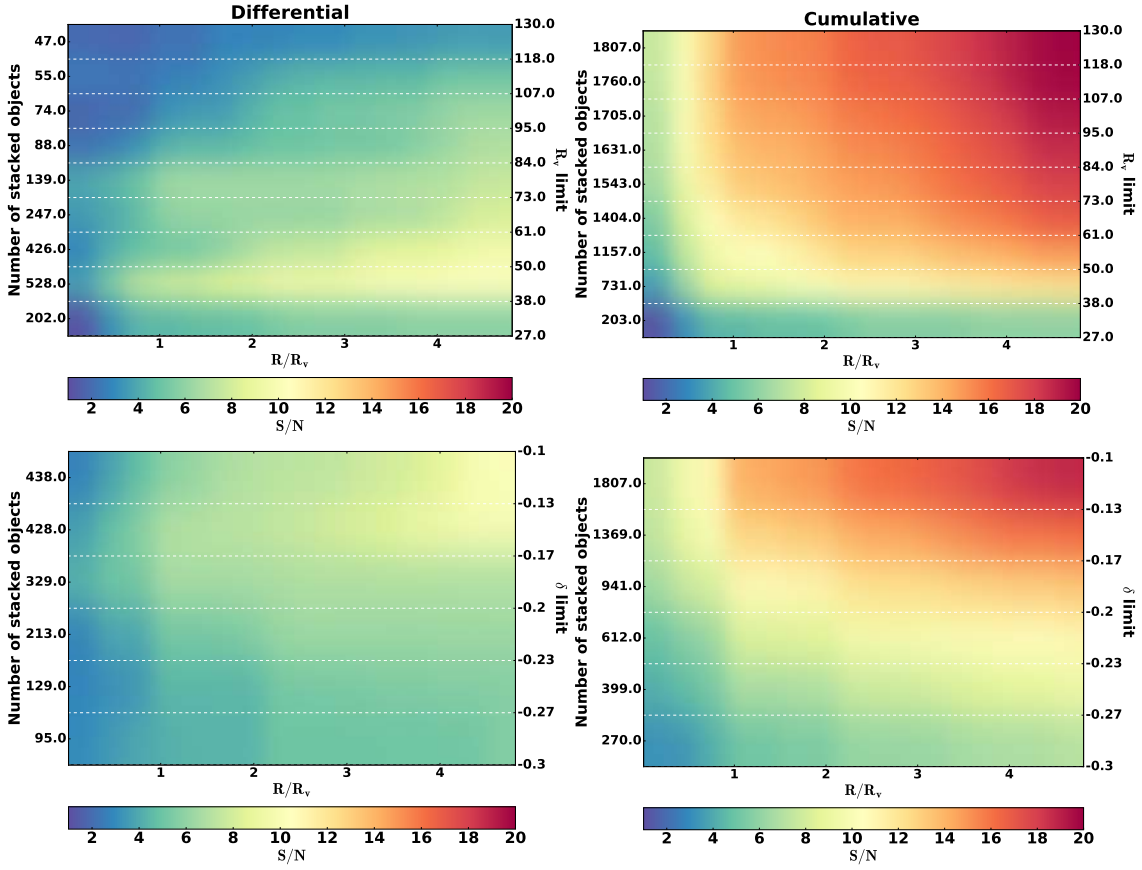

FIG. 4.5.2 – Signal-to-noise measured in a MICE Y1-like mock for different subsets of a void catalogue based on high luminosity *redMaGiC* tracers and 20 Mpc/h smoothing. In each subfigure, we show a cumulative S/N in the radial direction for increasing radius while different images correspond to different ordering schemes (δ or R_v) and averaging technique : differential (left panels) or cumulative (right panels). In the vertical axes on the left, we indicate the number of voids used for the stacking given the corresponding cuts on void parameters on the right side of the figures.

voids (voids-in-clouds) imprint less pronounced lensing signals on the CMB map than voids surrounded by voids (voids-in-voids). These findings appear to be robust against changing the galaxy tracer sample.

While we base our work on a different void definition, we aim to explore the possible optimization of the void catalogue by pruning in a similar manner. We therefore perform the stacking measurement for subsets of our void catalogues for both tracer densities and two different initial smoothing scales. The resulting imprints are characterized by comparing the stacked signal of voids to typical fluctuations of the CMB lensing map at random locations. We measure azimuthally averaged radial profiles from the stacked images. We then estimate a cumulative signal-to-noise (S/N) by gradually adding signals in each radial bins corresponding to larger

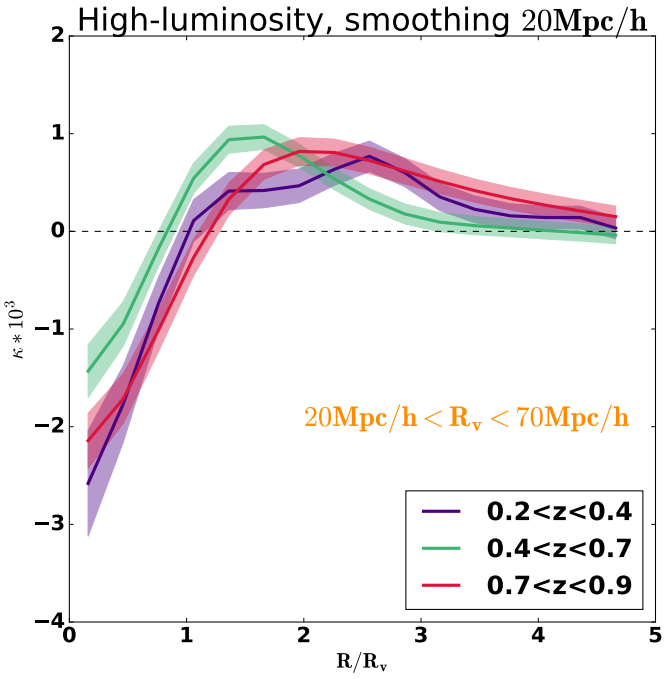

FIG. 4.5.3 – CMB lensing profile for three different redshift bins ([0.2, 0.4], [0.4, 0.7], [0.7, 0.9]) for medium size voids, to test a possible redshift evolution.

radii. This estimator is defined as follows:

$$\left(\frac{S}{N}\right)^2(r) = \sum_{i\leq r, j\leq r}^{r} S(i)cov^{-1}(i,j)S(j) - N_{bins} \tag{4.3}$$

where $S(i)$ is the amplitude of the lensing signal in the radius bin i, cov^{-1} the inverse covariance matrix. We estimate the covariance using 500 random rotations of the void catalogs. We then correct these estimates by multiplying them by the Anderson-Hartlap factor $\alpha = (N_{randoms} - N_{bins} - 2)/(N_{randoms} - 1)$ (**?**), and N_{bins} is the number of radius bin already computed to have an unbiased estimate. We compute the S/N in two different ways:

- *Differential signal-to-noise:* we estimate the signal-to-noise for each radius/density contrast bin.
- *Cumulative signal-to-noise:* we evaluate the signal-to-noise in cumulative way that is to say evaluating it for all catalog below a radius/density contrast threshold.

The S/N measured in differentially stacked images shows how sub-classes of voids contribute to the total detection significance. Similarly, the cumulative stacking, that gradually makes use of all the voids in the sample, indicates which portion of the radius-ordered and density-ordered

data provides the highest detection significance. An example of such tests is shown in figure 4.5.2. Further results for other versions of the void catalogues are presented in figures D.0.1, D.0.2, and D.0.3. We use the mean of the two MICE Y1-like footprints for the signals but consider the noise corresponding to a single Y1 footprint to model the DES expectations.

We make the following observations based on these optimization efforts :

- The most numerous medium size voids contribute most to the total lensing signal, i.e void with radius in the range $40 Mpc/h \lesssim R_v \lesssim 80 Mpc/h$.
- Similarly, the rather shallow but numerous voids show a better chance to measure their CMB lensing imprint, namely voids with mean density in the range $-0.2 \lesssim \bar{\delta} \lesssim -0.1$, again this range of densities are the most frequents in the voids catalogs (see 4.4.2).
- The S/N appears to be high enough to study the sub-classes of voids as suggested by Nadathur et al. (2017) with a caveat that photo-z voids have different properties compared to ZOBOV voids and we can expect less pronounced differences for the subclasses of the former
- While most of the S/N is contained inside the void radius ($R/R_v < 1$) and in the close surroundings ($1 < R/R_v < 2$, measuring the cumulative S/N up to ($R/R_v = 5$) also increases the detectability.[4]
- The highest S/N is achieved by stacking all voids, even if some voids are expected to contribute with less pronounced signal and higher noise at small scales (see Kovács et al., 2017, for a counter-example in the case of ISW imprints)

The best possible configuration is observed when using the high luminosity catalogue with 20 Mpc/h smoothing, reaching a cumulative $S/N \approx 20$ in the full profile up to $R/R_v = 5$. Nevertheless, all measurement configurations show significant CMB lensing signals for voids in a survey like DES Y1 thus we will measure the corresponding observed lensing imprint of all DES void catalogues and smoothing versions.

We note that he main results above are based on the full void sample with a variety of redshifts in $0.2 < z < 0.9$. For completeness, we also performed a simple redshift binning test for voids of size $20 Mpc/h < R_v < 70 Mpc/h$. Figure 4.5.3 shows that MICE voids in the high luminosity sample with 20 Mpc/h smoothing show no clear evidence for redshift evolution in their CMB profiles.

These findings are helpful to identify trends in the observed signal but the realistic uncertainties in the observed *Planck* κ map may exceed the simple variance that we estimated with the Y1-like

[4] In Figure 4.5.2, we can see that a cumulative $S/N \geq 10$ is encolse inside the void surroundings.

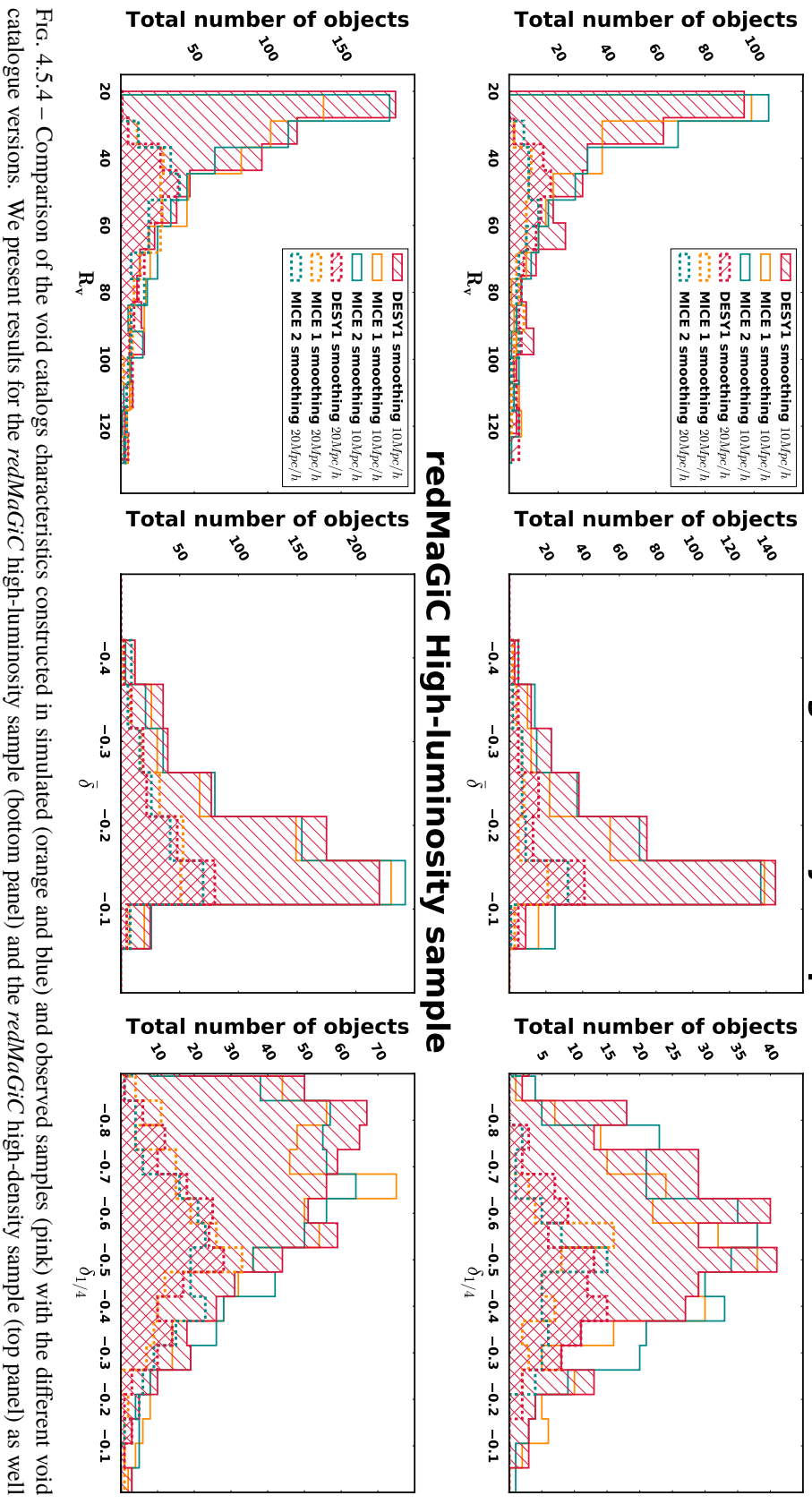

FIG. 4.5.4 – Comparison of the void catalogs characteristics constructed in simulated (orange and blue) and observed samples (pink) with the different void catalogue versions. We present results for the *redMaGiC* high-luminosity sample (bottom panel) and the *redMaGiC* high-density sample (top panel) as well as different smoothing scales for the void finder, 10 Mpc/h (solid lines) and 20 Mpc/h (dashed lines).

High density			
Smoothing	DES Y1	MICE 1	MICE 2
10 Mpc/h	303	262	294
20 Mpc/h	89	64	64
High luminosity			
Smoothing	DES Y1	MICE 1	MICE 2
10 Mpc/h	579	524	564
20 Mpc/h	195	180	181

TABLE 4.6.1 – Number of voids identified in DES Y1 data vs. in two Y1-like MICE patches. High luminosity sample are in better agreement.

patches within the MICE simulation. Therefore, we will determine the DES Y1 measurement error bars by stacking on random locations on the *Planck* map 500 times.

4.6 Results for observations - DES Y1 × *Planck*

4.6.1 Simulated catalogues compared to observed catalogues

In the light of the simulated stacking measurements using the MICE κ map, we aim to measure the DES Y1 tracers × *Planck* CMB κ signal. We thus use the observed *redMaGiC* catalogs from DES Y1, presented in 4.4.2.1, to construct void catalogues with the different tracer densities and initial smoothing scales.

Figure 4.5.4 shows a comparison of the observed and simulated void catalogues. We report a very good agreement in terms of sizes, central density, and mean density for both MICE Y1-like patches when they are compared to DES Y1 data. We find that the simple two-sample Kolmogorov-Smirnov histrogram consistency tests (Kolmogorov (1933); Smirnov (1948)) suggest that in general high luminosity samples are in better match. However, the overall agreement is sufficient, thus we aim to test the consistency of simulations and observations for all void catalogue versions.

4.6.2 Consistency tests of DES Y1 and MICE

We measure the stacked imprint of DES Y1 voids with the same methodology and parameters as in the case of the MICE mock. As an example, we compare the stacked images of the κ imprints of voids in high luminosity catalogues with 20 Mpc/h smoothing in figure 4.5.5.

A visual inspection shows a clear excess negative κ imprint in the center of the voids. The central spots for both MICE and DES Y1 are surrounded by an extended positive ring of overdensities,

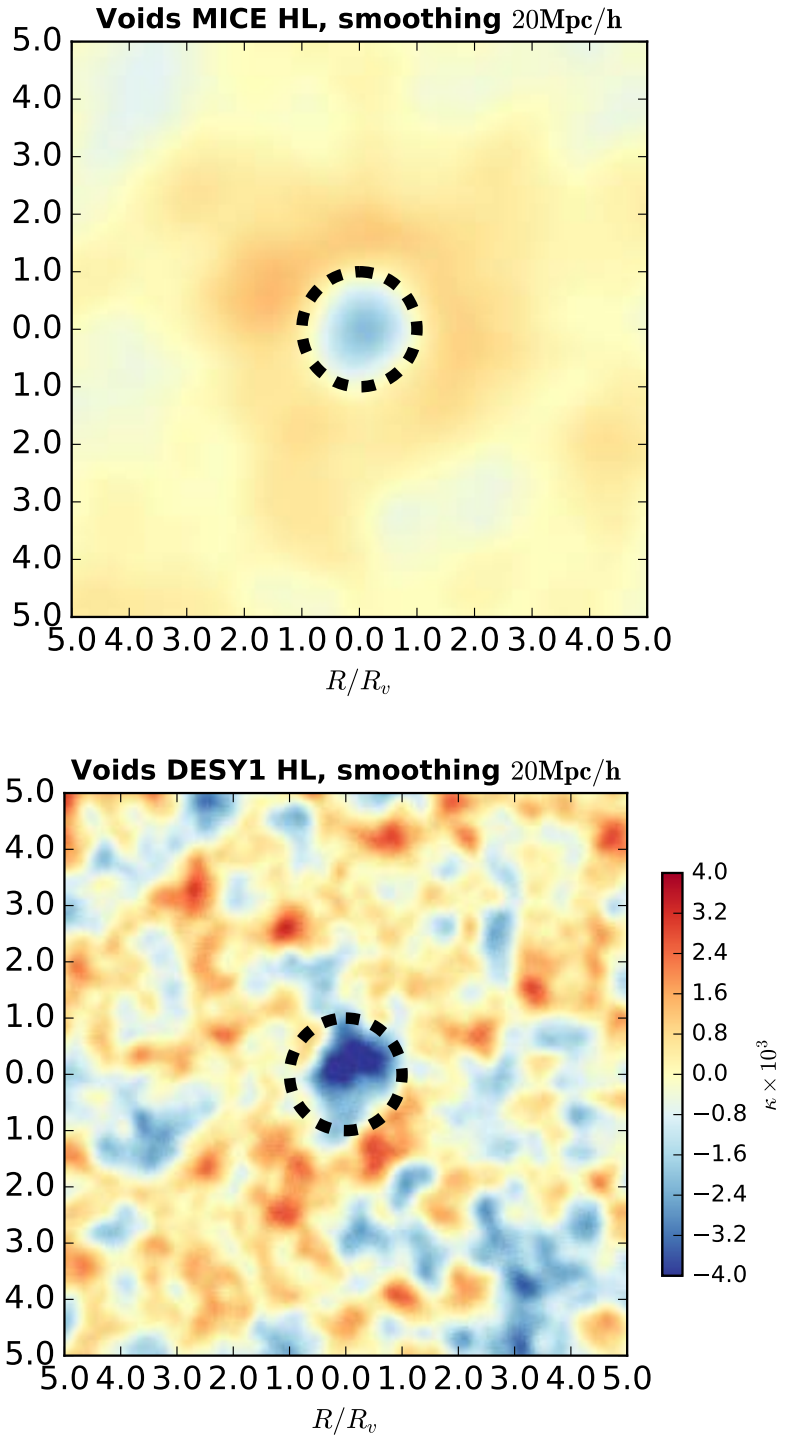

FIG. 4.5.5 – Comparison of simulated (left) and observed (right) stacked lensing signals. Clear excess is observed for DES Y1 data inside the re-scaled void radius $R/R_v = 1$ marked by the dashed circles.

with more noisy reconstruction for DES Y1 due to the presumably higher noise level in the *Planck* map.

We then also measured the azimuthally averaged radial imprint profile in the stacked images to quantify the results. We present the results in figure 4.6.1 for all four void catalogue versions. The shaded region mark 1σ errors computed with 500 random realizations of the stacking measurement on the MICE κ map, while the error bars correspond to 500 random stacking runs on the *Planck* κ map.

We note that our estimates of the stacked CMB κ profile in the MICE mock are in great qualitative and quantitative agreement with the simulated profiles by **?** and Nadathur et al. (2017) even if they used different void definitions and different tracer catalogues.

However, our observed DES Y1 signal shows a clear excess with respect to MICE ΛCDM cosmology, on the other hand , in Pollina et al. (2018) the authors obtained consistent agreement in the void properties between DES voids and the MICE simulation, emphasizing thus the reality of the discrepency observed here.

As a consistency statistic, we first measure the amplitude (A) of the observed and simulated lensing imprints as a function of the distance to the void radius in the usual re-scaled units:

$$A(R/R_v) = \frac{\kappa_{DESY1}(R/R_v)}{\kappa_{MICE}(R/R_v)} \tag{4.4}$$

One may relate this amplitude to the A_L phenomenological lensing amplitude that was estimated by the *Planck* team in their analyses (**?**).

We observe a good general agreement in the sign and the shape of the observed and simulated profiles; negative κ values in the interior of voids plus an extended range of positive convergence in the surroundings. However, we detect a significant *excess* signal in the interior of DES Y1 voids. For 20 Mpc/h smoothing, we obtain $A \approx 2$ for both tracer densities. Furthermore, we find $3 \leq A \leq 4$ when 10 Mpc/h smoothing is applied for both high luminosity and high density tracer catalogues. We note that the convergence of the profiles to zero signal at large distance from the void center is an important null test which proves that our method provides an unbiased estimate.

We aim to characterize further the significance of these excess signals inside the void radius. The significance level of the observed signal compared to zero signal has been evaluated using the cumulative signal-to-noise (4.3) of the first three points of the reconstructed lensing profile; namely the points residing inside the void radius, defined as follows

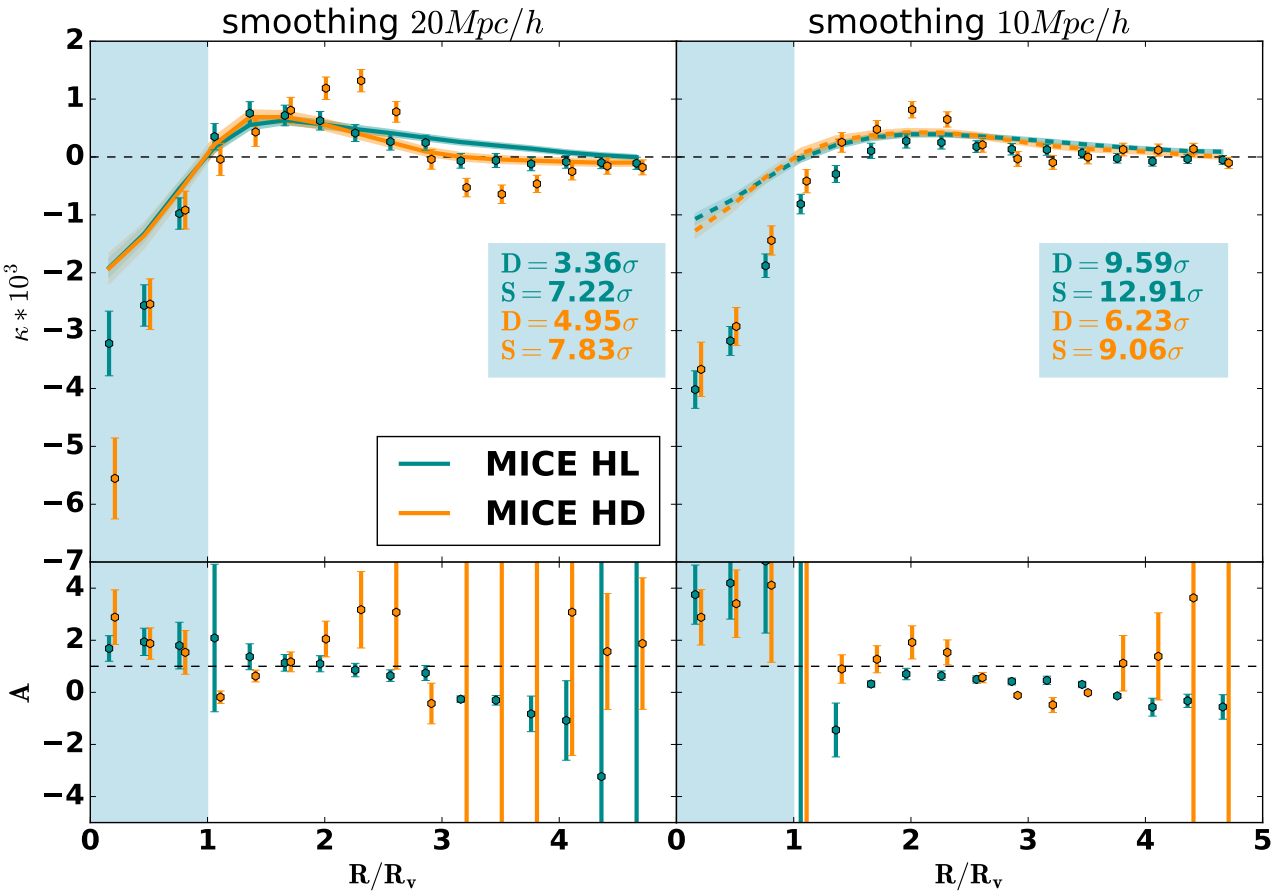

FIG. 4.6.1 – We present our main findings in this figure. Points with 1σ error bars mark the CMB lensing convergence profiles of DES Y1 voids for various configurations indicated by the figure legend and titles. Solid lines with shaded 1σ error ranges show the simulated imprint of MICE voids for the same combinations of smoothing and tracer density. Shaded text boxes indicate the significance (S) and anomaly (A^+) measures of the different configurations. These significance values correspond only to the void interiors that we shaded with the same color. The lower panels show the A amplitude and its 1σ errors of the observed κ signal compared to the MICE estimate in each bin ($\kappa_{DESY1}/\kappa_{MICE}$).

$$S = \sqrt{\sum_{i\leq 3, j\leq 3}^{3} S(i) cov^{-1}(i,j) S(j) - N_{bins}}. \tag{4.5}$$

The measured values of this S detection significance are:

- $S = 7.22\sigma$ (high luminosity, 20 Mpc/h smoothing)
- $S = 7.83\sigma$ (high density, 20 Mpc/h smoothing)
- $S = 12.91\sigma$ (high luminosity, 10 Mpc/h smoothing)
- $S = 9.06\sigma$ (high density, 10 Mpc/h smoothing)

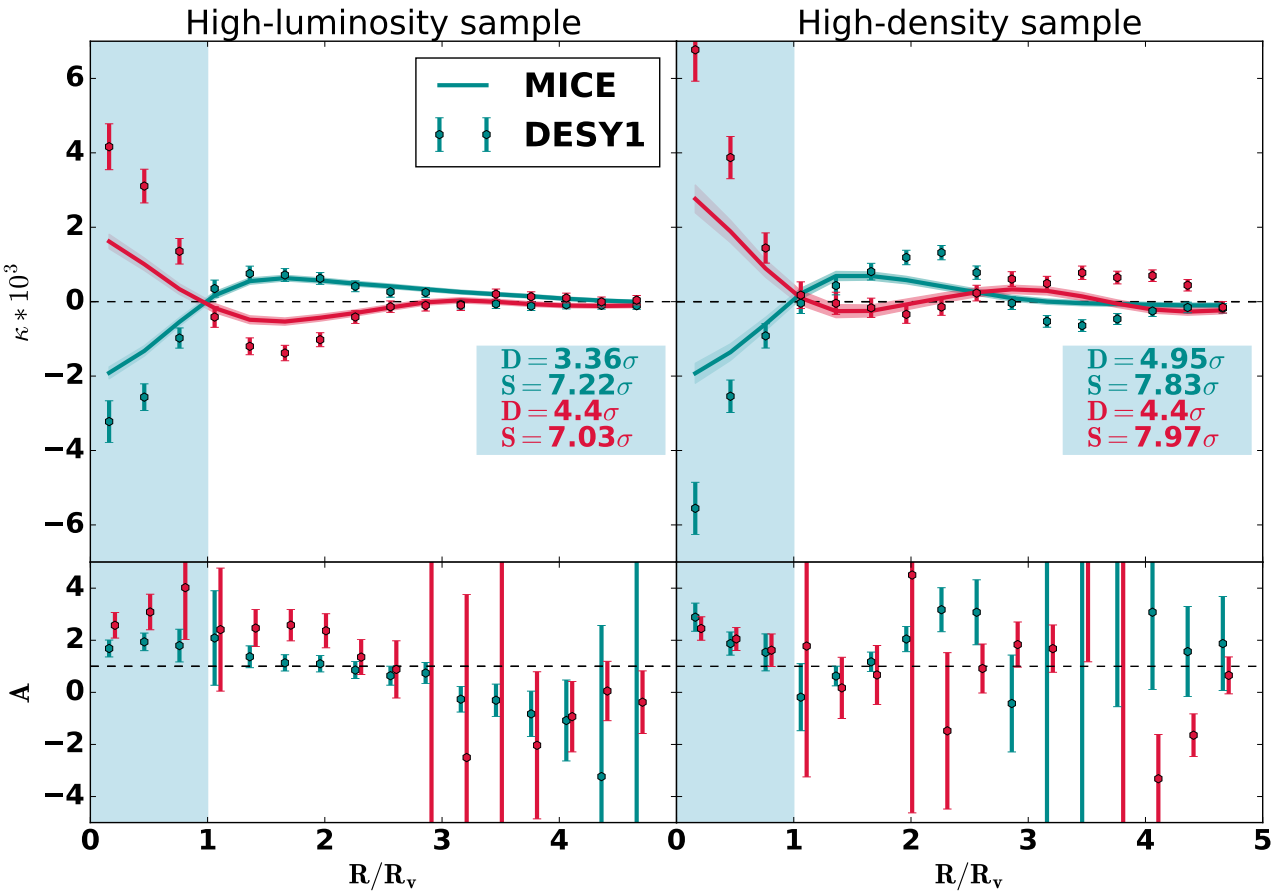

FIG. 4.6.2 – Points with 1σ error bars mark the CMB lensing convergence profiles of DES Y1 voids (blue) and *superclusters* (red) for high luminosity (left) and high density (right) tracer samples, with 20 Mpc/h smoothing in these examples. Solid lines with shaded 1σ error ranges show the simulated imprint of MICE voids (blue) and superclusters (red). Shaded text boxes indicate the significance (S) and anomaly (A^+) measures of the different configurations. These color-coded significance values correspond only to the void and supercluster interiors that we shaded with the same color. The lower panels show the A amplitude and its 1σ errors of the observed κ signal compared to the MICE estimate in each bin ($\kappa_{DESY1}/\kappa_{MICE}$).

Importantly, the detection significance of the DES Y1 clusters × CMB lensing signals by Baxter et al. (2018) is 8.1σ, i.e. comparable to the DES Y1 voids × CMB lensing signal from the very same tracer data set. We note, however, that the overall significance of our void lensing signal is even higher if not only the void interior (first three radial bins) is considered. The outer profile's less significant positive convergence also carries some information, even if not truly helpful to understand the physics of voids itself.

Then, in order to evaluate the significance of the excess part of the signal itself with respect to the simulated template profile, we defined the discrepancy measure D based again on the the

first three bins in the profile in the void interior:

$$D = \frac{\kappa_{DESY1} - \kappa_{MICE}}{N_{DESY1}} \tag{4.6}$$

where N is the noise of the κ signal. We estimated the following values of this D excess discrepancy parameter:

- $D = 3.36\sigma$ (high luminosity, 20 Mpc/h smoothing)
- $D = 4.95\sigma$ (high density, 20 Mpc/h smoothing)
- $D = 9.59\sigma$ (high luminosity, 10 Mpc/h smoothing)
- $D = 6.23\sigma$ (high density, 10 Mpc/h smoothing)

We conclude that the excess signal of DES Y1 voids appears to be robust against changing the tracer density and the void finder configuration. Importantly, the excess is seen when stacking *all* voids in the catalogues. We minimized the role of any *a posteriori* selection bias by analyzing the MICE simulation prior to looking at the DES Y1 data and in fact this pre-analysis showed us that stacking all voids is always beneficial in our measurement configuration. Nevertheless we test how different splits of the void catalogue may affect the results.

4.6.3 Tests of the excess signal - z and R_v dependence

We perform simple tests to verify if the excess signal measured in the observed catalogs is localized in size or redshift space.

Therefore, we first divide our sample into two bins in void radius, $20\ Mpc/h < R_v < 70\ Mpc/h$ and $70\ Mpc/h < R_v < 130\ Mpc/h$. The resulting stacked images are shown in figure E.0.1 for both MICE and DES Y1 to again check their consistency. We find that the DES data shows a similar excess for both small-intermediate and large voids. Then, we split our data into two redshift bins of $0.2 < z < 0.5$ and $0.5 < z < 0.9$. The stacked CMB lensing signals in both simulation and observed catalogs are measured for these sub-classes. We present these results in figure E.0.2. No important difference is seen in the relative strength of the DES Y1 and MICE catalogues at lower and higher redshifts.

We thus conclude that the excess signal is quite robust against subdividing the void catalogues.

4.6.4 Imprint of superclusters - MICE vs. DES Y1

Following Kovács et al. (2017), we inverted our void finder in order to identify superclusters with the same algorithm but instead of looking for minima in the smooth density fields in slices,

it looks for maxima (see again section 4.2.2 for details). We built various supercluster catalogs with the same combinations of tracer density and initial smoothing scale.

Then we applied our stacking methodology to simulated and observed lensing κ maps. An example of the resulting stacked images are shown in figure 4.6.3 for the high luminosity sample with 20 Mpc/h smoothing. We found evidence for an excess CMB lensing signal of superclusters that appears to be comparabe to the excess signal of DES Y1 voids in magnitude.

We again measured the radial κ profile in MICE and DES Y1 for both high luminosity and high density tracers. These imprints are compared to void results in figure 4.6.2.

We verify that the both the simulated and observed imprints are fairly symmetrical around zero. Moreover, the significances of the observed supercluster signals, shown in the lower panels figure 4.6.2, are in good agreement with the DES Y1 void amplitudes. We estimate the following detection and discrepency values for the DES Y1 superclusters:

- $S = 7.03\sigma$ (high luminosity, 20 Mpc/h smoothing)
- $D = 4.4\sigma$ (high luminosity, 20 Mpc/h smoothing)
- $S = 7.97\sigma$ (high density, 20 Mpc/h smoothing)
- $D = 4.4\sigma$ (high density, 20 Mpc/h smoothing)

The overall detection significance of the DES Y1 superclusters $\times$ *Planck* CMB lensing signal is again comparable not just to our DES Y1 void lensing results, but also to the 8.1σ DES Y1 clusters $\times$ CMB κ detection by Baxter et al. (2018) who used the very same DES Y1 data set with a different method testing different scales. We conclude that our robustness test using DES and MICE superclusters clearly verified the existence of a significant excess signal in the observed data with respect to MICE ΛCDM cosmology.

4.7 Discussion & Conclusions

The main objective of this work was to study cosmic voids identified in Dark Energy Survey galaxy samples, culled from the first year of observations. We relied on the *redMaGiC* sample of luminous red galaxies sample of exquisite photometric redshift accuracy to robustly identify cosmic voids in photometric data. We then aimed to cross-correlate these cosmic voids with lensing maps of the Cosmic Microwave Background.

Such a signal has already been detected by Cai et al. (2017) with a significance of 3.2σ. On positions of voids identified in the BOSS footprint, they stacked patches of the publicly available

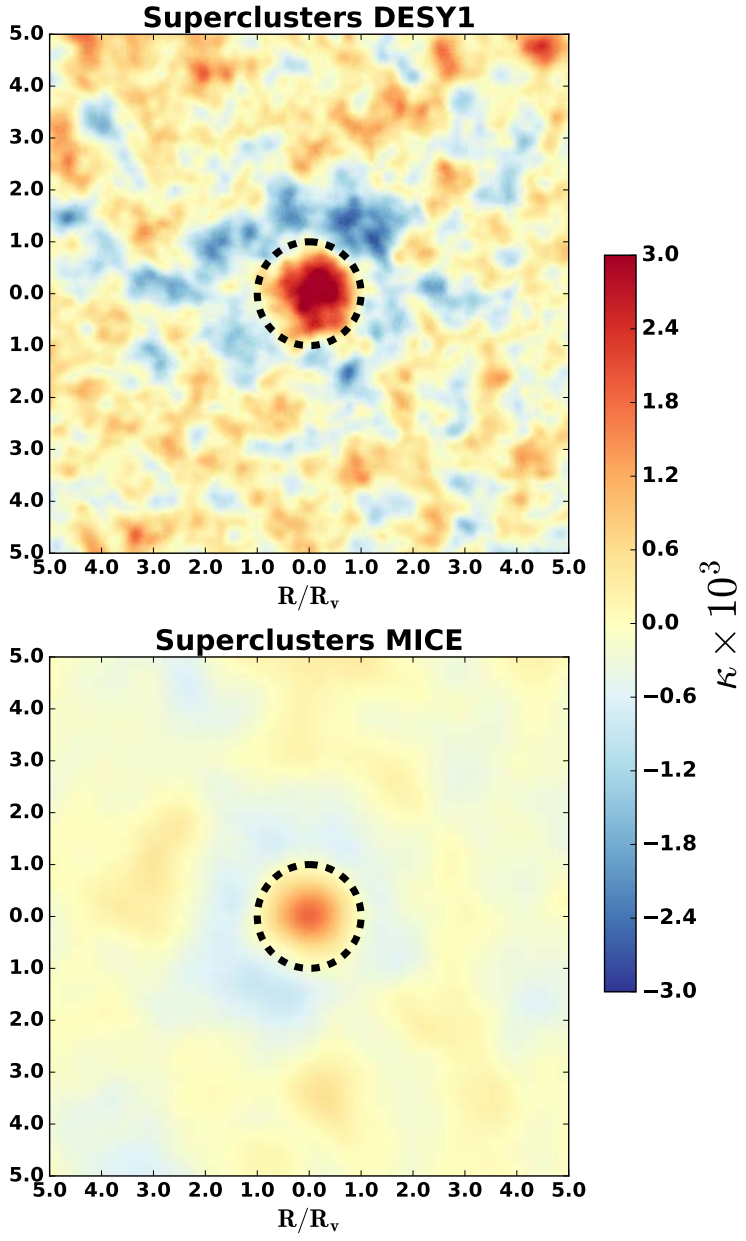

FIG. 4.6.3 – Comparison of simulated (left) and observed (right) stacked lensing signals. Clear excess for DES Y1 data inside the re-scaled supercluster radius $R/R_v = 1$ marked by the dashed circles.

lensing convergence map of the Planck satellite. In general, we followed their methodology but we put more emphasis of simulation analyses to optimally detect a signal with DES data, given different galaxy tracer density, void finder parameters, and sub-classes of voids with different properties (see Nadathur et al., 2017, for details). In particular, we used simulated DES-*like* *redMaGiC* galaxy catalogs together with a simulated lensing convergence map from the MICE *Grand Challenge* N-body simulation to test our ability to detect the CMB lensing imprint of cosmic voids.

We analyzed the signal-to-noise corresponding to the CMB κ profile of MICE *redMaGiC* voids. We considered different void populations. We varied the galaxy density and also the initial smoothing scale applied to density field to find the centres of the voids (see Sánchez et al., 2017, for details). These parameters affect the significance of the measurement as the total number of voids, mean void size, underdensity in void interiors, and their depth in the centres are all affected by these choices and so is the resulting signal and noise.

We first evaluated that the shape and amplitude of the CMB lensing profile in the MICE mock agrees well with previous estimates by Cai et al. (2017), despite using different tracers and void definition. We then comprehensively searched for the best combination of parameters that guarantee the best chance to detect a signal with observed DES data. We concluded that a combination of a 20 Mpc/h initial Gaussian smoothing in the density field and the lower tracer density of the higher luminosity *redMaGiC* galaxy catalogue are preferable to achieve the highest signal-to-noise. Looking into the sub-classes of voids and their contribution to the cumulative signal, we found that the most numerous medium size voids ($40Mpc/h < R_v < 80Mpc/h$)) and rather shallow voids ($-0.2 > \bar{\delta} > -0.1$) add most to the total signal. However, we found that stacking all voids is preferable for the best measurement configuration that one should utilize with observed DES data.

We then identified voids in the observed DES *redMaGiC* catalog and compared their properties with MICE voids. In general, we found a good agreement when comparing observed void catalogues with both DES Y1-like MICE mocks that we used for predictions.

We repeated the simulated stacking analyses now using the observed Planck CMB lensing map. We detected a correlation with $7-12\sigma$ significance depending on the void catalog version considered. However, this detection shows a non-negligible excess with respect to simulation estimates. Firstly, a visual inspection already revealed that most of the discrepancy is originated in the void centres, i.e. where $\sim 60-70\%$ of the S/N is coming from based on our MICE analyses. We thus explored the inner part of voids in greater details using further statistical probes.

In the interior of voids, we found a 3.4σ discrepancy with respect to the MICE ΛCDM cosmology even in the most conservative analysis. We found no evidence for dependence on void radius

and/or redshift by dividing the void catalogs into bins. Furthermore, we inverted the void finder method to find extended underdensities, or *superclusters*, and tested the relation of the CMB κ imprint of these objects in MICE versus in DES Y1. We again detected an excess DES Y1 signal of 4.4σ significance in observed data with respect to MICE estimates.

These excess signals are plausibly related to the *excess ISW* signal identified by Kovács et al. (2017) in the same *redMaGiC* DES Y1 data set. In fact the ISW-lensing dual probe is valuable for probes of modified gravity since lensing depends on the gravitational potential, whereas ISW depends on its time derivative.

The origin of these excess signals of voids and superclusters is of course not necessarily cosmological. However, with more evidence from different observables and different surveys, now involving DES tracers $\times$ *Planck* CMB lensing, at least the excess appears to be robust. Importantly, no CMB color dependence has been seen in the related ISW measurments which may exclude some possibilities. A residual contamination, coming from unresolved extragalactic point sources, may still be blamed though (see e.g. Millea et al., 2012). Dust from galaxies at all redshifts contributes to the CMB temperature fluctuations, which, in turn, would result in a positive correlation between CMB temperatures and galaxy density (see e.g. Ho et al., 2008). However, Hernández-Monteagudo & Smith (2013) reported that realistic contamination of this kind leave a different ISW imprint than what has been observed thus this possibility seems unlikely.

Considering alternative models, neither modifications of the concordance model, given other precise constraints (Nadathur et al., 2012), nor simple modified gravity scenarios seem to resolve the tension in ISW measurements (e.g. Cai et al., 2015). Nevertheless, our detection of excess CMB lensing signals associated with DES Y1 voids and superclusters that also show ISW excess may shed new light in this problem.

In summary, we argue that catalogues of cosmic voids may be more informative about dark energy than presumably optimal techniques. We observe possible problems at the largest scales and averaging in traditional two-point correlation probes may wash out the interesting new features. In the Dark Energy Survey Collaboration, we aim to continue this research along similar lines, including measurements of the imprint of supervoids and superclusters in a bigger catalogue based on three years of observations (Y3). In the near future, beyond a better understanding of the methodologies and possible re-analyses, new cosmic web decomposition data from experiments like the Dark Energy Spectroscopic Instrument (DESI) (DESI Collaboration et al., 2016) and the Euclid mission (Laureijs et al., 2011) will further constrain the lensing and ISW signals of super-structures.

Chapter 5

Conclusions

This thesis have been devoted to the analysis of the first year of observation of the Dark Energy Survey (DES) galaxy catalog. After giving a brief introduction to the state-of-the art of cosmology today as well as the main cosmological probes and tool that have been used for this work in chapter 1, we have presented in chapter 2 DES and the main cosmological results that came out from the first year of observation. In these results we exposed in particularly the methodology followed by the collaboration to extract robust information from galaxy catalogs from DES and how this results, coming from a galaxy survey, are the first to achieve accuracy comparable to Cosmic Microwave Background experiment such as the Planck satellite Tauber et al. (2010); Planck Collaboration et al. (2011).

In the results presented in chapter 2, we have saw that one key element to infer cosmology using cosmic shear or galaxy-galaxy lensing resides in the redshift distribution of the source galaxies.From this point, we have dedicated chapter 3 to the first main project of this thesis, namely redshift distribution inference using rather new methodology the so-called clustering-redshift methods. In a first phase, we have been using simulated galaxy catalog to probe different approaches of these type of methods as well as different intrinsic parameters to define an optimal procedure minimizing the systematic budget induced by the method itself. Whereas in recent experiments clustering-redshift have been used at most to cross-check redshift distibution infered by photometric redshift algorithm, in DESY1, has been used as an additional tool in cosmological analyses. In particular, we have been developing a framework that allows us to use clustering-redshift techniques to correct the distribution inferred by regular photo-z methodologies allowing a shift along the redshift axe, and shown that in one hand, the dominant systematic error in our procedure was due to the fact that we are not able to model the galaxy bias evolution, and on the other hand that the systematic induced by the photo-z error of our high-precision photometric redshift reference sample was sub-dominant.

In Chapter 4, we have been looking for imprints of cosmic voids identified preliminary in the DES galaxy catalogs. Cosmic voids have recently raise a particular in cosmology because of their low-density characteristic that suppose that its form regions where Dark Energy is more dominant than in the rest of the universe, this aspect make them good candidates to test gravity. Beside this, recent work have shown, looking at the cosmic void integrated Sachs-Wolfe (ISW) signal, some discrepancies between the cosmological standard model and observation. This anomaly remains unexplained and gives an additional motivation to look at the imprint of void cosmic void not in the temperature anisotropies map from the CMB, as it is done for the ISW signal, but on the lensing convergence map recently provided by the Planck collaboration Planck Collaboration et al. (2016c). In our analysis, the imprint of voids in the CMB lensing map have been detected both in CDM simulation and DES observation at a significant level and furthermore, the detected signal appear to be similarly to the ISW to be in excess in the observed datasets with respect to the simulated ΛCDM. Both ISW signal and lensing convergence are related to the gravitational potential, and both are showing a significant excess with respect to simulated ΛCDM cosmology, this two excess signal remain mysterious and further investigation will have to be done in order to explain these two excesses and the connections of one with the other. To conclude, in this thesis, we have used two different ways to use cross-correlation in cosmology, in particular we have shown in one hand that cross-correlation techniques could be a tool to infer redshift distribution of galaxies and on the other hand that it is also an essential approach in the cosmological parameter estimation and to probe of the ΛCDM cosmology.

Appendix A

Clustering Redshift : The choice of the reference sample

Along our analysis, we have been using as reference sample the *redMaGiC* higher-luminosity sample. Although in the context of DESY1 cosmological analysis Elvin-Poole et al. (2017); Prat et al. (2017), two additional *redMaGiC* samples were created, the *redMaGiC* High-density ($L > 0.5L_*$) sample and the *redMaGiC* High-luminosity sample ($L > 1.5L_*$) as well as a combined version of these three catalogs has been built as described in Elvin-Poole et al. (2017)), the redshift distribution of the *redMaGiC* combined sample in the buzzard simulation is shown in Fig.(A.0.1). As one can see in the figure the combination of the sample consists in selecting one of the here-above mentioned *redMaGiC* catalogs for a specific redshift range.

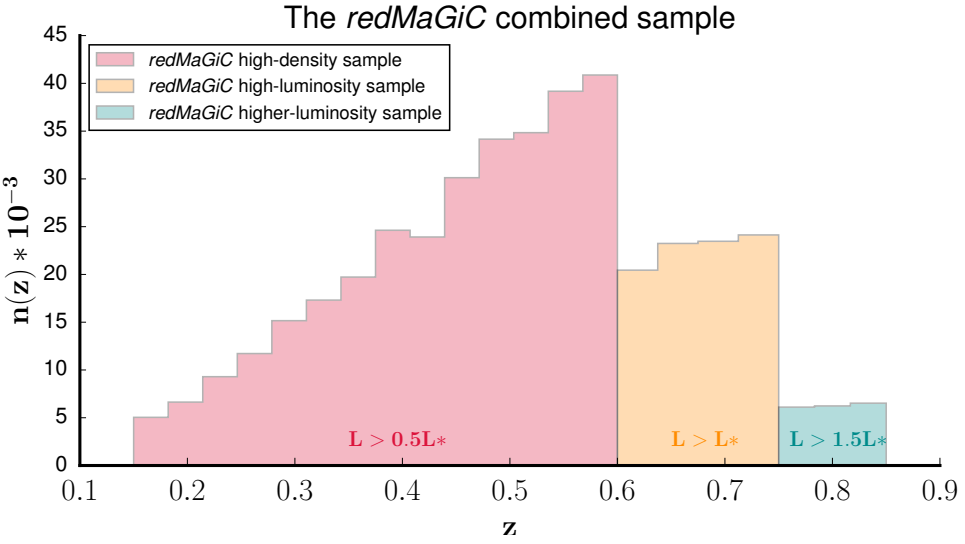

FIG. A.0.1 – Redshift distribution of the combined *redMaGiC* sample in buzzard simulation differentiating with color the various samples that has been used to construct this sample.

TABLE A.0.1 – **Total systematic error with *redMaGiC* combined sample as a reference**. The table shows the total systematic error for the mean matching procedure, for the three WL source redshift bins and photo-z codes.

	Bin 1	**Bin 2**	**Bin 3**
BPZ	0.037	0.016	0.007
DNF	0.021	0.015	0.016

Here we re-iterate the analyses presented in 3.4 and characterize the total systematic budget (with the simulated buzzard catalogs) but using this time the combined *redMaGiC* sample as reference sample for the two different photo-z code. Table A.0.1 expose the results obtained in this case, and comparing it to the different corrections obtained in Table 3.4.2 and 3.4.3 one observes enlargement of the total systematic in the first tomographic bin and reduction of it on the third one. The first tomographic bin, as one can see in Fig.(A.0.1), will have a larger overlap with galaxies from *redMaGiC* high-density sample this sample includes more fainter galaxies ($L > 0.5L*$) increasing at the same time the number density of galaxy and the photo-z bias of our references. This last aspect, will thus increase the *redMaGiC* photo-z systematic (section 3.4.2.2) and will be observed mostly in the first bin.

On the other hand, we have now, a reference sample made of three different samples of galaxies, this will have as consequence to make the redshift evolution of the combined *redMaGiC* bias to vary more than the higher-luminosity *redMaGiC* one. Especially on the third bin where the 'change' in sample occurs in the middle of the bin ($z = 0.75$), such a variation will cause an increase in the systematic due to the reference sample bias evolution although what we are observing here is a cancellation of this systematic due to the opposite effects it induces with respect to the unknown sample bias evolution and give us a smaller value as it can be seen in the table.

In our analysis, given that both the bias evolution systematic and *redMaGiC* photo-z systematic are stronger using the combined *redMaGiC* sample as reference, we decided to use as fiducial reference sample the higher-luminosity *redMaGiC* sample.

Appendix B

Clustering-Redshift : galaxy bias correction

As it has been mentioned in section 3.4.2.1, in principle if we could make a good estimation of the auto-correlation signal of both (unknown and reference) samples, it would be possible to correct the **bias evolution systematic**. In 3.4.2.1 to estimate the auto-correlation signal, we have been using the true redshift of our galaxies (since we were using simulated catalogs) binned in narrow redshift bins. Here the goal is to explore the possibility to correct these evolutions in observed catalogs (using now photometric redshift). Fig. B.0.1 compares the evolution in redshift of the auto-correlation signals (in the Buzzard simulation) using the true redshift and the one recovered using photo-z, for the WL sample, the *redMaGiC* higher-luminosity and combined samples. From the figure, one can see that using photometric redshifts does not allow us to recover the true evolution of the auto-correlation signal, specially for the WL sample where the reconstructed signal behaves differently: due to the low-quality of the photometric redshifts, it is not possible to recover the true signal bining the sample in such narrow bins.

However, in the case of the *redMaGiC* samples, one can observed similarity between the two redshift evolutions. Here we will show that it is feasible to correct the small discrepancy induced by the *redMaGiC* photo-z errors, using a sub-sample of *redMaGiC* galaxies for which we have spectroscopic redshifts.

Indeed if one considers Eq.(3.12), and assumes that neither the dark matter density field nor the galaxy bias of our sample evolve considerably with redshift in the scales considered here and the narrow bins treated here, Eq.(3.12) can be rewritten as:

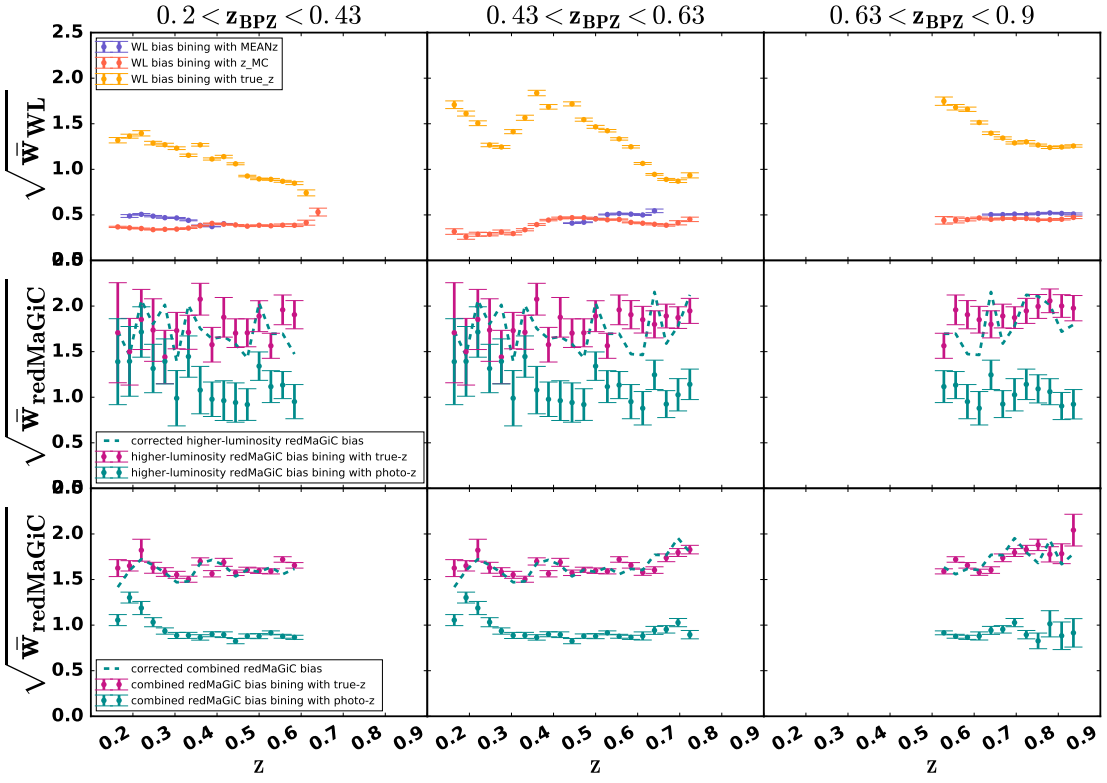

FIG. B.0.1 – *Top panel:* 1-bin estimate of the square root of the auto-correlation functions for the WL sample using alternatively the galaxy true redshift (yellow), photo-z mean redshift MEANz (purple) and a random draw from the photo-z posterior z_{MC} (orange). *Second and third panel:* 1-bin estimate of the square root of the auto-correlation functions for the two different *redMaGiC* samples (combined and higher luminosity) using alternatively the *redMaGiC* true redshift (pink) and photo-z (blue), the dashed blue line being the signal infered after applying the correction presented here.

$$\bar{w}_{rr}^{true}(< z >) = \bar{w}_{DM}(< z >) b_r^2(< z >) \int dz' n_{r,true}^2(z') \tag{B.1}$$

$$\bar{w}_{rr}^{photo-z}(< z >) = \bar{w}_{DM}(< z >) b_r^2(< z >) \int dz' n_{r,photo-z}^2(z'), \tag{B.2}$$

where the quantities outside the integrals are now evaluated at the mean redshift $< z >$ of the reference bins. If one now divides these two equations, one will end with a correction to apply to the photometric auto-correlation signal to recover the true one, namely :

$$\bar{w}_{rr}^{true}(< z >) = \bar{w}_{rr}^{photo-z}(< z >) \frac{\int dz' n_{r,true}^2(z')}{\int dz' n_{r,photo-z}^2(z')} \tag{B.3}$$

In principle, to compute this correction one would need to know the right hand side of the equation (the true redshift distributions, binning both with photometric and spectroscopic redshifts). Even

though we don't have access to this information in the full *redMaGiC* sample, we can use a sub-sample of *redMaGiC* for which we have spectral information. In Fig.(B.0.1), the dashed blue lines represent the signal obtained after applying this correction for both the combined and Higher-luminosity samples. And as it can be seen there, the correction applied allows us to recover the true signal. Note that for the WL sample case, this would not be feasible, due to the fact that the redshift evolution of the bias could not be considered constant within the bin considered, and smaller bins are not possible due to the poor photo-z resolution.

Appendix C

Clustering redshift : Lensing magnification effect

As mentioned in section 3.2 and can be explicitly observed on the reconstructed distribution given by clustering-redshift based methodology (see for example figure 3.4.1), the tails of our reconstructed distributions are noisy and the origin of it could come from lensing magnification effects.

Lensing magnification effects presented first in Narayan (1989) and estimate later in various works Bartelmann & Schneider (2001); Scranton et al. (2005); Ménard et al. (2010); Morrison et al. (2012), causes extra-correlations signal between two samples with large redshift separation. This extra-signal is due to magnification effects of background galaxy by matter distribution along the line of sight that bias the observed spacial density of galaxies.

In the literature, lensing magnification effects in clustering-redshift distribution has been neglected, because expected to be small compare to the real clustering signal [1], although in the redshift ranges where we don't have overlap (in the tails of the distributions) between 'unknown' and 'reference' samples this effects can become dominant and thus influence our estimation of the mean of the distribution.

In principle, one could theoretically estimate the impact of lensing magnification in our reconstruction. Starting with the observed over-density can be then formulate as :

$$\delta_{obs} = \delta_g + \delta_\mu \tag{C.1}$$

[1] in Ménard et al. (2010) and Scranton et al. (2005) this effects is at the order of $\sim 1\%$ at arcminute scale for high redshift source lensed by a lens at redshift $z = 0.5$

where g stands for the galaxy over-density component and μ for the magnified one. In this perspective, one can rewritte the cross-correlation signal of two sample of galaxy as :

$$w_{ur}(\theta) =< \delta_{g,u}\delta_{g,r} > + < \delta_{g,u}\delta_{\mu,r} > + < \delta_{\mu,u}\delta_{g,r} > + < \delta_{\mu,u}\delta_{\mu,r} > \tag{C.2}$$

In chapter 3, only the first term (namely the intrinsic clustering of the two galaxy samples) in the right hand side of the equation has been considered, we are now interested in evaluating how much the additional second and third terms could bias our redshift reconstruction (here we will not be considering the forth term in the equation which considers correlation between magnified objects of the two sample and should be sub-dominant compare to the two other ones, see Heavens & Joachimi (2011); Duncan et al. (2014)). In order to evaluate the impact of these two components, assuming linear bias model, this terms can be expressed in fourier space using Limber equation (see 1.2.2.4) in the flat-sky approximation as :

$$< \delta_{g,u}\delta_{g,r} >= b_u b_r \int_0^{\chi_H} d\chi \eta_u(\chi)\eta_r(\chi) \int_0^{\infty} \frac{kdk}{2\pi} P(k,\chi)J_0(\chi k\theta) \tag{C.3}$$

where $P(k,\chi)$ is the 3D matter power spectrum at wavenumber k and comoving distance χ, $b_{u,r}$ are the biases of our two samples, $\eta_{u,r}(\chi)$ are their radial distribution in co-moving distances, and $J_0(\chi k\theta)$ is the zeroth order Bessel function. And the terms due to lensing magnification could also be evaluate in a fourier frame as :

$$< \delta_{g,1}\delta_{\mu,2} >= b_1(2.5s-1) \int_0^{\chi_H} d\chi \eta_1(\chi)K(\chi) \int_0^{\infty} \frac{kdk}{2\pi} P(k,\chi)J_0(k\chi\theta) \tag{C.4}$$

the subscripts 1 and 2 correspond either to the unknown reference sample, s being the slope of the magnitude number counts which can be expressed, for a flux limited sample, as function of the galaxy number density n and the galaxy magnintude m :

$$s = \frac{dlog_{10}n(>m)}{dm} \tag{C.5}$$

and $K(\chi)$ being the lensing kernel of the sample in the background define as :

$$K(\chi) = \frac{3H_0^2\Omega_m\chi}{c^2 a} \int_{\chi}^{\chi_H} d\chi' \eta_2(\chi')\frac{\chi'-\chi}{\chi'} \tag{C.6}$$

with H_0 and a, the Hubble parameter today and the scale factor.

Using the equations presented here, it is then possible to theoretically estimate the clustering signal of equation C.2, the difficulty here come when one has to estimate the parameter s for our samples. Indeed, in one side the *redMaGiC* sample is not a flux-limited function.

However in the case of small redshift bins one can approximate the *redMaGiC* sample to be flux limited with $m > M_{max} + 5log_{10}(d(z_{bin}))$, and thus s can be estimate evaluating the slope

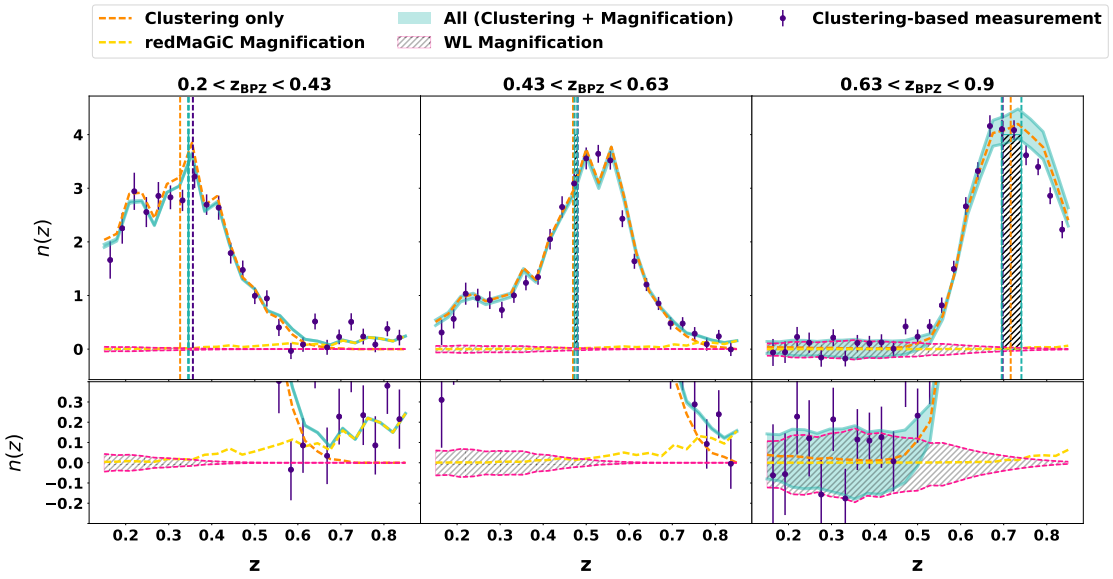

FIG. C.0.1 – *Top panel :* Different effect on the reconstructed redshift distribution due to lensing magnification, the shaded area being the impact of the different magnification components. The points are the clustering redshift measurement done in simulation and the vertical bars are the mean computed for including or not magnification effects. *Bottom panel :* zoom of the tails of the distribution where the magnification effects are stronger.

where the number counts drops binned in magnitude bins. On the other hand, in the case the unknown sample, the weak lensing sample is not flux limited, one should used a more general expression than equation C.5 that imply to characterize well the selection function of the sample. The goal here being to give an estimate of the lensing magnification impact on our reconstruction we choose instead to use characteristic values, specifically $2.5s - 1 = \pm 1.5$ for the slope of the magnitude number count. Computing in this way the different components of equation C.2, one can evaluate in one side the clustering-redshift theoretical signal including the different magnification effects (equation C.4), as well as the effects of each of the magnification terms and compare it to the measured clustering redshift distribution, this is shown in figure C.0.1.

As it can be seen in Fig.(C.0.1), magnification lensing is stronger in the tails of the distributions, and in particular for the first and the third bin, where we were observing positive tails. In the first bin, this effect is mostly drive by the *redMaGiC* magnification whereas in the last bin the WL sample would be the one magnified. The shade region representing the range of cumulative number count slope for the WL sample. And the different vertical lines been the different mean of the distributions taking into account (or not) the different lensing magnification effects, as one can see the differences on this means differ more in the first and last bins, where the magnification effects are dominant on the tails. In particular, as expected in the first bin the correlation signal in the tail is dominated by magnification effects on *redMaGiC* galaxies at high-redshift and this excess signal is also observed in the clustering-redshift reconstruction. For the last bin we can also observed magnification effect on the weak lensing sample, however this effects are highly

dependent on the parameter s and further test should be done to conclude that the signal in the tail comes from this effect.

We conclude here that as it can been on Figure C.0.1, magnification effects could have a non-neglectable effect especially for redshift distribution with long tails, however in our framework (section 3.2) we have been evaluating the mean of our reconstruction within a $\pm 2\sigma_{WZ}$ window, which excludes the redshift ranges where these effects are strong. However, resolution requirement in redshift distribution for future surveys are increasing and correcting from these effect might become essential.

Appendix D

Signal-to-noise optimization

In section 4.5.2, we presented the methodology we have been adopted in order to optimized the void imprints in the CMB lensing maps. We have presented there the cumulative signal-to-noise obtained for one of the various void catalogs we have been testing. Figure D.0.1,D.0.2 and D.0.3 are the cumulative signal-to-noise signals obtained for these additional catalogs.

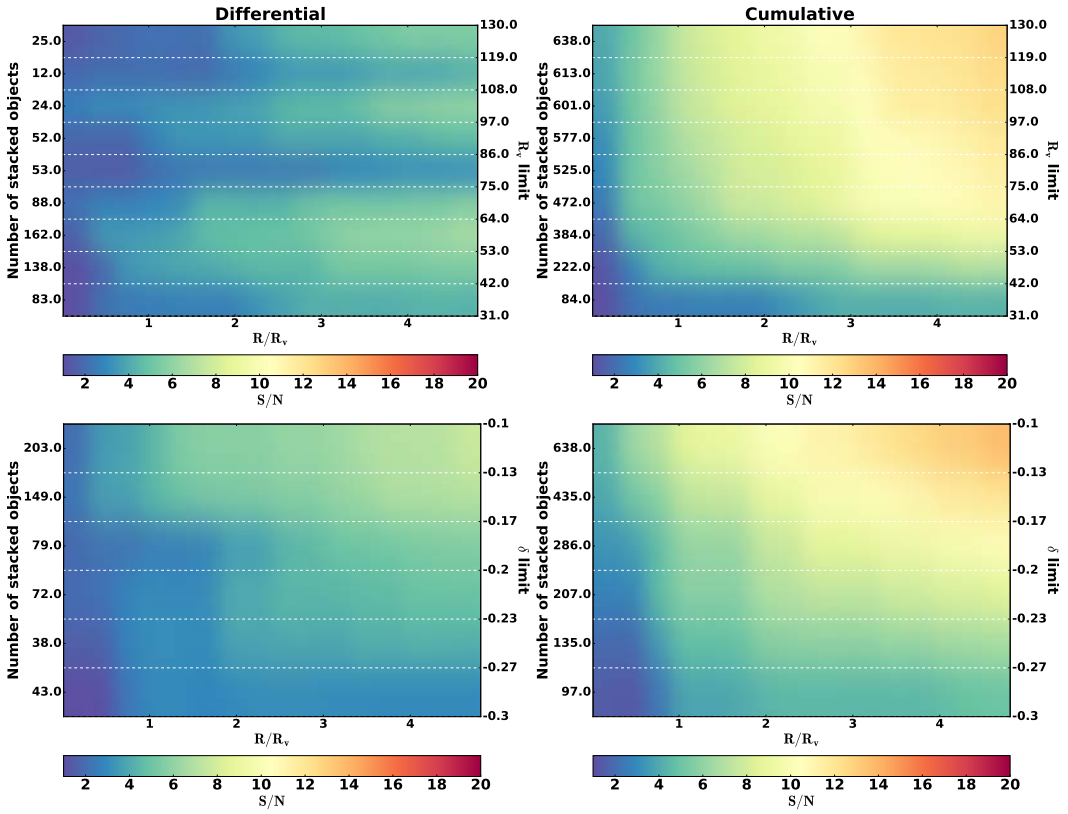

FIG. D.0.1 – Signal-to-noise measured in a MICE Y1-like mock for different subsets of a void catalogue based on high density *redMaGiC* tracers and 20 Mpc/h smoothing. In each subfigure, we show a cumulative S/N in the radial direction for increasing radius while different images correspond to different ordering schemes (δ or R_v) and averaging technique (differential or cumulative). In the vertical axes on the left, we indicate the number of voids used for the stacking given the corresponding cuts on void parameters on the right side of the figures.

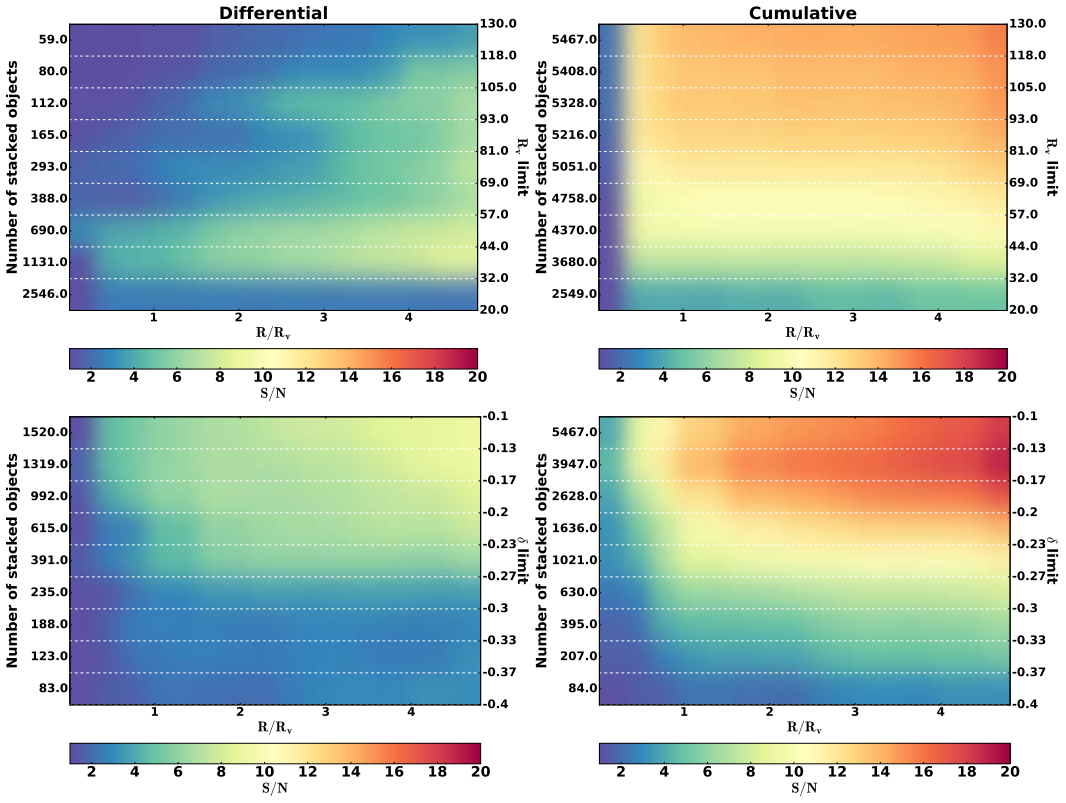

FIG. D.0.2 – Signal-to-noise measured in a MICE Y1-like mock for different subsets of a void catalogue based on high luminosity *redMaGiC* tracers and 10 Mpc/h smoothing. In each subfigure, we show a cumulative S/N in the radial direction for increasing radius while different images correspond to different ordering schemes (δ or R_v) and averaging technique (differential or cumulative). In the vertical axes on the left, we indicate the number of voids used for the stacking given the corresponding cuts on void parameters on the right side of the figures.

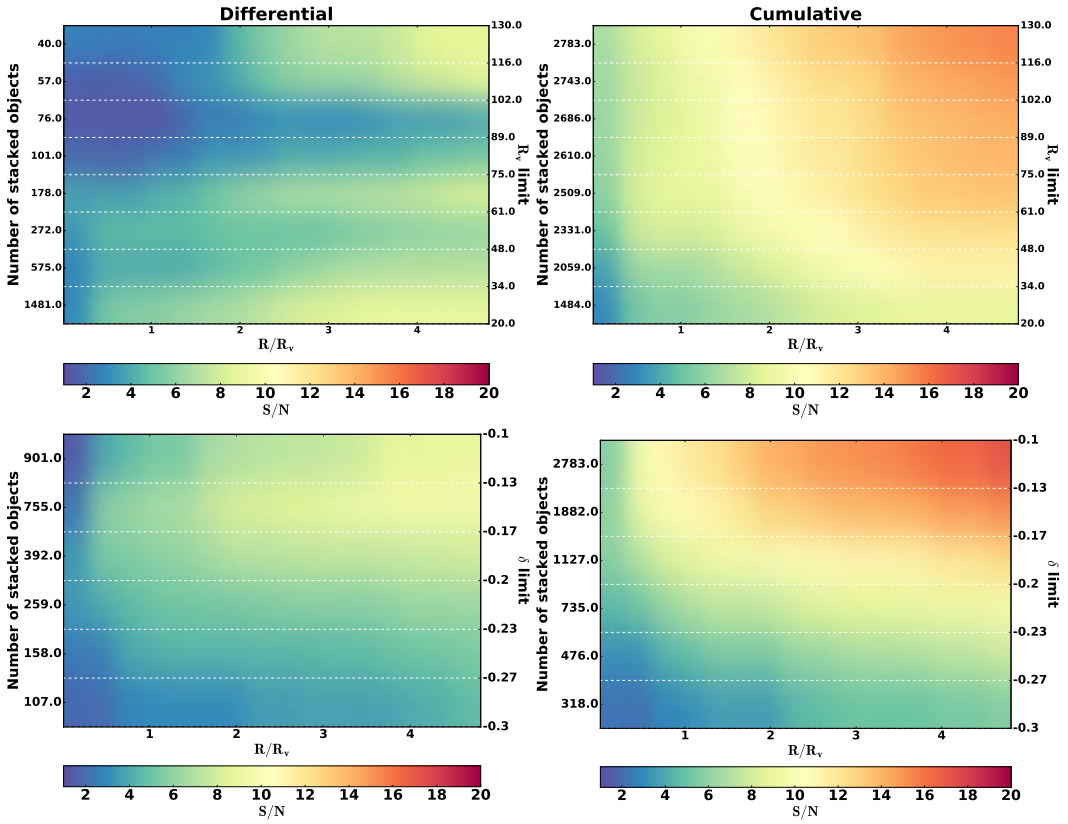

FIG. D.0.3 – Signal-to-noise measured in a MICE Y1-like mock for different subsets of a void catalogue based on high density *redMaGiC* tracers and 10 Mpc/h smoothing. In each subfigure, we show a cumulative S/N in the radial direction for increasing radius while different images correspond to different ordering schemes (δ or R_v) and averaging technique (differential or cumulative). In the vertical axes on the left, we indicate the number of voids used for the stacking given the corresponding cuts on void parameters on the right side of the figures.

Appendix E

Tests of the excess signal

Additionally to the signal measurements of the voids imprint in the CMB maps of chapter 4, we have performed several consitency tests in order to find out weather or not we could identify specific trends in the excess signal detected for the observed catalogs. Namely in a first time we re-iterate our stacking analysis dividing our void samples (both simulated and observed) in a first time into two void radius bins ($20Mpc/h < R_v < 70Mpc/h$ and $70Mpc/h < R_v < 130Mpc/h$) and in a second time into two redshift bins ($0.2 < z_v < 0.5$ and $0.5 < z_v < 0.9$). Figure E.0.1 and E.0.2 are showing the stacked image of these two tests for both MICE and DES.

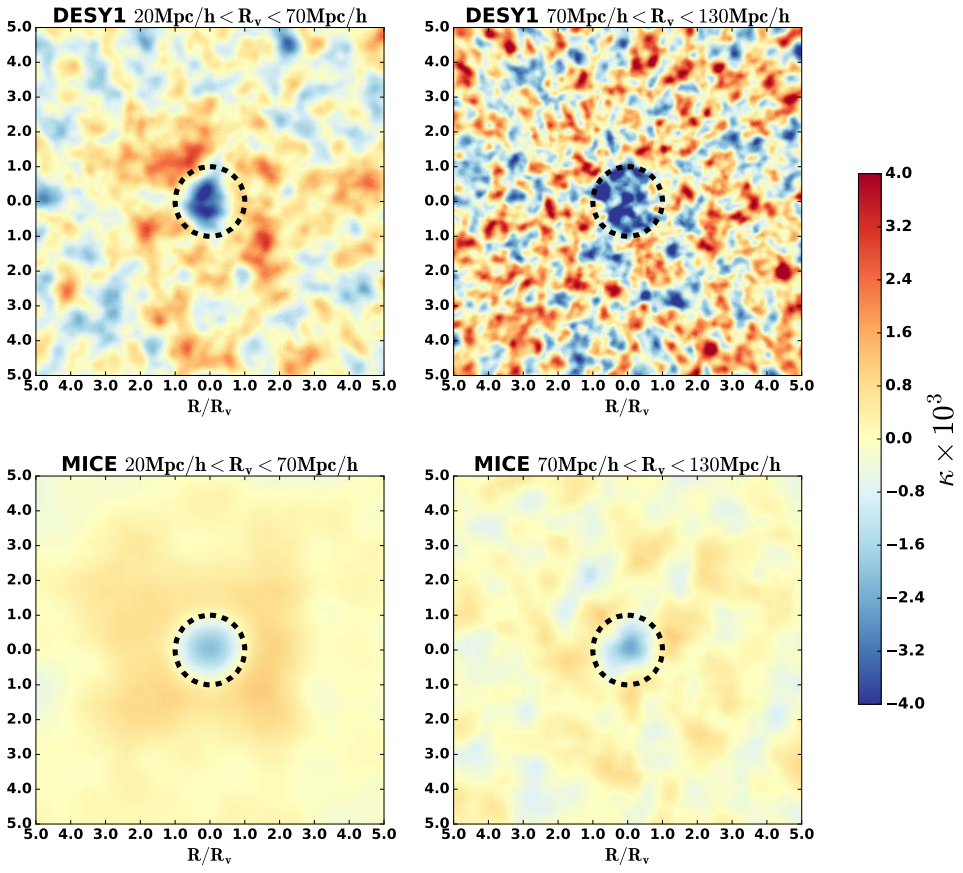

FIG. E.0.1 – Consistency tests of small vs. large voids in observations and simulation. No trend is seen as DES Y1 data shows stronger imprints in both bins.

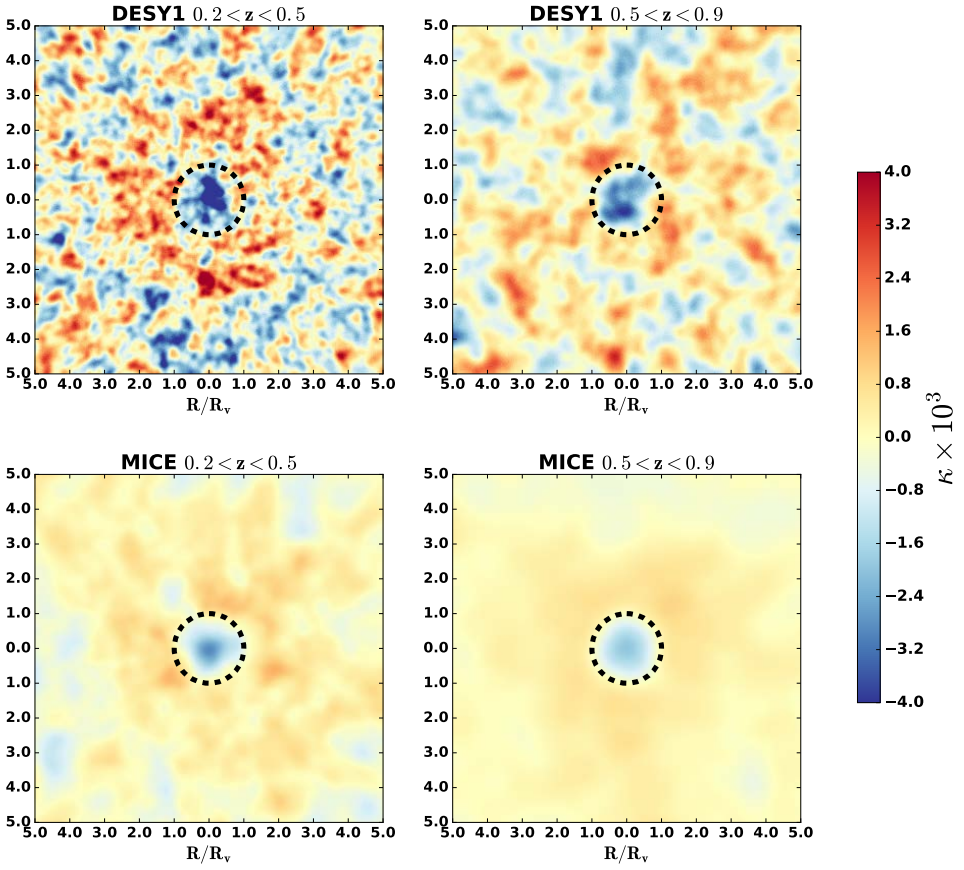

FIG. E.0.2 – Consistency tests of low-z vs high-z voids in observations and simulation. No trend is seen as DES Y1 data shows stronger imprints in both bins.

Bibliography

Abbott T., et al., 2016, , 94, 022001

Abell G. O., 1958, , 3, 211

Abell G. O., 1961, , 66, 607

Aihara H., et al., 2018, , 70, S4

Alam S., et al., 2017, , 470, 2617

Alpher R. A., Follin J. W., Herman R. C., 1953, Physical Review, 92, 1347

Amendola L., Frieman J. A., Waga I., 1999, , 309, 465

Arnouts S., Cristiani S., Moscardini L., Matarrese S., Lucchin F., Fontana A., Giallongo E., 1999, , 310, 540

Babcock H. W., 1939, Lick Observatory Bulletin, 19, 41

Bartelmann M., Schneider P., 2001, , 340, 291

Baxter E. J., et al., 2018, , 476, 2674

Becker M. R., 2013, , 435, 115

Behroozi P. S., Wechsler R. H., Wu H.-Y., 2013, , 762, 109

Benítez N., 2000, , 536, 571

Bennett C. L., et al., 2003, , 583, 1

Bertschinger E., 1985, , 58, 1

Betoule M., et al., 2014, , 568, A22

Beutler F., et al., 2011, , 416, 3017

Blanchard A., Schneider J., 1987, , 184, 1

Blas D., Lesgourgues J., Tram T., 2011, , 7, 034

Boggess N. W., et al., 1992, , 397, 420

Cai Y.-C., Li B., Cole S., Frenk C. S., Neyrinck M., 2014a, , 439, 2978

Cai Y.-C., Neyrinck M. C., Szapudi I., Cole S., Frenk C. S., 2014b, , 786, 110

Cai Y.-C., Padilla N., Li B., 2015, , 451, 1036

Cai Y.-C., Neyrinck M., Mao Q., Peacock J. A., Szapudi I., Berlind A. A., 2017, , 466, 3364

Carlstrom J. E., et al., 2011, , 123, 568

Cautun M., Paillas E., Cai Y.-C., Bose S., Armijo J., Li B., Padilla N., 2018, , 476, 3195

Cawthon R., et al., 2017, preprint, (arXiv:1712.07298)

Chantavat T., Sawangwit U., Sutter P. M., Wandelt B. D., 2016, , 93, 043523

Choi A., et al., 2016, , 463, 3737

Coe D., Benítez N., Sánchez S. F., Jee M., Bouwens R., Ford H., 2006, , 132, 926

Coil A. L., et al., 2011, , 741, 8

Colberg J. M., et al., 2008, , 387, 933

Coleman G. D., Wu C.-C., Weedman D. W., 1980, , 43, 393

Cooper M. C., et al., 2011, , 193, 14

Crocce M., Castander F. J., Gaztañaga E., Fosalba P., Carretero J., 2015, , 453, 1513

Crocce M., et al., 2017, preprint, (arXiv:1712.06211)

DES Collaboration et al., 2017, preprint, (arXiv:1708.01530)

DESI Collaboration et al., 2016, preprint, (arXiv:1611.00036)

Dark Energy Survey Collaboration et al., 2016, , 460, 1270

Davies C. T., Cautun M., Li B., 2018, preprint, (arXiv:1803.08717)

Davis M., Peebles P. J. E., 1983, , 267, 465

Davis C., et al., 2017, preprint, (arXiv:1710.02517)

Davis C., et al., 2018, ,

Dawson K. S., et al., 2013, , 145, 10

Dawson K. S., et al., 2016, , 151, 44

De Vicente J., Sánchez E., Sevilla-Noarbe I., 2016, , 459, 3078

DeRose J., Wechsler R., Rykoff E., et al., 2017, in prep.

Drinkwater M. J., et al., 2010, , 401, 1429

Drlica-Wagner A., et al., 2018, , 235, 33

Duncan C. A. J., Joachimi B., Heavens A. F., Heymans C., Hildebrandt H., 2014, , 437, 2471

Dyson F. W., Eddington A. S., Davidson C., 1920, Philosophical Transactions of the Royal Society of London Series A, 220, 291

Eifler T., Krause E., Schneider P., Honscheid K., 2014, , 440, 1379

Einstein A., 1905, Annalen der Physik, 322, 891

Einstein A., 1916, Annalen der Physik, 354, 769

Einstein A., 1917, Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 142-152.,

Elvin-Poole J., et al., 2017, preprint, (arXiv:1708.01536)

Feroz F., Hobson M. P., Bridges M., 2009, , 398, 1601

Fixsen D. J., 2009, , 707, 916

Flaugher B., et al., 2015, , 150, 150

Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, , 125, 306

Fosalba P., Gaztañaga E., Castander F. J., Manera M., 2008, , 391, 435

Fosalba P., Crocce M., Gaztañaga E., Castander F. J., 2015a, , 448, 2987

Fosalba P., Crocce M., Gaztañaga E., Castander F. J., 2015b, , 448, 2987

Fowler J. W., et al., 2007, , 46, 3444

Friedmann A., 1922, Zeitschrift fur Physik, 10, 377

Fry J. N., Gaztanaga E., 1993, , 413, 447

Gamow G., 1946, Physical Review, 70, 572

Gatti, Vielzeuf. et al., 2018, , 477, 1664

Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M., Bartelmann M., 2005, , 622, 759

Granett B. R., Neyrinck M. C., Szapudi I., 2008, , 683, L99

Granett B. R., Kovács A., Hawken A. J., 2015, , 454, 2804

Guth A. H., 1981,

Hamaus N., Sutter P. M., Wandelt B. D., 2014, Physical Review Letters, 112, 251302

Hamilton A. J. S., 1993, , 417, 19

Hartlap J., Simon P., Schneider P., 2007, , 464, 399

He S., Alam S., Ferraro S., Chen Y.-C., Ho S., 2017, preprint, (arXiv:1709.02543)

Heavens A. F., Joachimi B., 2011, , 415, 1681

Hernández-Monteagudo C., Smith R. E., 2013, , 435, 1094

Hewett P. C., 1982, , 201, 867

Hildebrandt H., et al., 2017, , 465, 1454

Ho S., Hirata C., Padmanabhan N., Seljak U., Bahcall N., 2008, Physical Review D, 78, 043519

Hoyle B., et al., 2018, , 478, 592

Hu W., Okamoto T., 2002, , 574, 566

Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168

Huff E., Mandelbaum R., 2017, preprint, (arXiv:1702.02600)

Huterer D., Turner M. S., 1999, , 60, 081301

Ilbert O., et al., 2006, , 457, 841

Jain B., Taylor A., 2003, Physical Review Letters, 91, 141302

Jarvis M., Bernstein G., Jain B., 2004, , 352, 338

Johnson A., et al., 2017, , 465, 4118

Joudaki S., et al., 2018, , 474, 4894

Kahn F. D., Woltjer L., 1959, , 130, 705

Kaiser N., 1984, , 284, L9

Kerscher M., Szapudi I., Szalay A. S., 2000, , 535, L13

Kilbinger M., et al., 2013, , 430, 2200

Kinney A. L., Calzetti D., Bohlin R. C., McQuade K., Storchi-Bergmann T., Schmitt H. R., 1996, , 467, 38

Köhlinger F., et al., 2017, , 471, 4412

Kolmogorov A., 1933, Giornale dell' Istituto Italiano degli Attuari, 4, 83

Kovács A., 2018, , 475, 1777

Kovács A., Sánchez C., García-Bellido J., the DES collaboration 2017, , 465, 4166

Krause E., Chang T.-C., Doré O., Umetsu K., 2013, , 762, L20

Krause E., et al., 2017, Submitted to: Phys. Rev. D

LSST Science Collaboration et al., 2009, preprint, (arXiv:0912.0201)

Laigle C., et al., 2016, , 224, 24

Landy S. D., Szalay A. S., 1993, , 412, 64

Laureijs R., et al., 2011, preprint, (arXiv:1110.3193)

Leavitt H. S., Pickering E. C., 1912, Harvard College Observatory Circular, 173, 1

Lemaître G., 1927, Annales de la Société Scientifique de Bruxelles, 47, 49

Lesgourgues J., 2011, preprint, (arXiv:1104.2932)

Lewis A., Challinor A., 2006, , 429, 1

Lewis A., Challinor A., 2011, CAMB: Code for Anisotropies in the Microwave Background, Astrophysics Source Code Library (ascl:1102.026)

Limber D. N., 1954, , 119, 655

MacCrann N., et al., 2018, preprint, (arXiv:1803.09795)

Mao Q., et al., 2017, , 835, 161

Martí P., Miquel R., Castander F. J., Gaztañaga E., Eriksen M., Sánchez C., 2014, , 442, 92

Matthews D. J., Newman J. A., 2010, , 721, 456

McQuinn M., White M., 2013, , 433, 2857

Melchior P., Sutter P. M., Sheldon E. S., Krause E., Wandelt B. D., 2014, , 440, 2922

Ménard B., Scranton R., Fukugita M., Richards G., 2010, , 405, 1025

Ménard B., Scranton R., Schmidt S., Morrison C., Jeong D., Budavari T., Rahman M., 2013, preprint, (arXiv:1303.4722)

Millea M., Doré O., Dudley J., Holder G., Knox L., Shaw L., Song Y.-S., Zahn O., 2012, , 746, 4

Morrison C. B., Scranton R., Ménard B., Schmidt S. J., Tyson J. A., Ryan R., Choi A., Wittman D. M., 2012, , 426, 2489

Morrison C. B., Hildebrandt H., Schmidt S. J., Baldry I. K., Bilicki M., Choi A., Erben T., Schneider P., 2017, , 467, 3576

Nadathur S., Crittenden R., 2016, , 830, L19

Nadathur S., Hotchkiss S., 2015, , 454, 889

Nadathur S., Hotchkiss S., Sarkar S., 2012, JCAP, 6, 42

Nadathur S., Hotchkiss S., Crittenden R., 2017, , 467, 4067

Narayan R., 1989, , 339, L53

Newman J. A., 2008, , 684, 88

Neyrinck M. C., 2008, , 386, 2101

Norberg P., Baugh C. M., Gaztañaga E., Croton D. J., 2009, , 396, 19

Okamoto T., Hu W., 2003, , 67, 083002

Omori Y., et al., 2017, , 849, 124

Ostriker J. P., Peebles P. J. E., 1973, , 186, 467

Padin S., et al., 2008, , 47, 4418

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

Peebles P. J. E., Hauser M. G., 1974, , 28, 19

Penzias A. A., Wilson R. W., 1965, , 142, 419

Perlmutter S., 2003, Physics Today, 56, 53

Perlmutter S., et al., 1999, , 517, 565

Planck 2015 results. XXI. 2016, AAP, 594, A21

Planck Collaboration et al., 2011, , 536, A1

Planck Collaboration et al., 2016a, , 594, A9

Planck Collaboration et al., 2016b, , 594, A13

Planck Collaboration et al., 2016c, , 594, A15

Planck Collaboration et al., 2018, preprint, (arXiv:1807.06209)

Pollina G., et al., 2018, preprint, (arXiv:1806.06860)

Prat J., et al., 2017, preprint, (arXiv:1708.01537)

Rahman M., Ménard B., Scranton R., Schmidt S. J., Morrison C. B., 2015, , 447, 3500

Rahman M., Ménard B., Scranton R., 2016a, , 457, 3912

Rahman M., Mendez A. J., Ménard B., Scranton R., Schmidt S. J., Morrison C. B., Budavári T., 2016b, , 460, 163

Riess A. G., et al., 1998, , 116, 1009

Ross A. J., Samushia L., Howlett C., Percival W. J., Burden A., Manera M., 2015, , 449, 835

Rozo E., et al., 2016, , 461, 1431

Ruhl J., et al., 2004, in Bradford C. M., et al., eds, Vol. 5498, Z-Spec: a broadband millimeter-wave grating spectrometer: design, construction, and first cryogenic measurements. pp 11–29 (arXiv:astro-ph/0411122), doi:10.1117/12.552473

Rykoff E. S., et al., 2014, , 785, 104

Rykoff E. S., Rozo E., Keisler R., 2015, ArXiv: 1509.00870,

Sachs R. K., Wolfe A. M., 1967, , 147, 73

Samuroff S., et al., 2018, , 475, 4524

Sánchez C., et al., 2017, , 465, 746

Schmidt S. J., Ménard B., Scranton R., Morrison C., McBride C. K., 2013, , 431, 3307

Schneider P., van Waerbeke L., Mellier Y., 2002, , 389, 729

Schneider M., Knox L., Zhan H., Connolly A., 2006, , 651, 14

Scottez V., et al., 2016, , 462, 1683

Scottez V., Benoit-Lévy A., Coupon J., Ilbert O., Mellier Y., 2018, , 474, 3921

Scranton R., et al., 2005, , 633, 589

Seldner M., Peebles P. J. E., 1979, , 227, 30

Sheldon E. S., Huff E. M., 2017, , 841, 24

Silk J., 1968, , 151, 459

Smirnov N., 1948, doi:https://doi.org/10.1214/aoms/1177730256, 19, 279

Smith S., 1936, , 83, 23

Smoot G. F., et al., 1992, , 396, L1

Soares-Santos M., et al., 2012, Physics Procedia, 37, 1445

Springel V., 2005, , 364, 1105

Springel V., et al., 2005, , 435, 629

Sutter P. M., Lavaux G., Hamaus N., Wandelt B. D., Weinberg D. H., Warren M. S., 2014, , 442, 462

Tauber J. A., et al., 2010, , 520, A1

The Dark Energy Survey Collaboration 2005, ArXiv Astrophysics e-prints,

Totsuji H., Kihara T., 1969, , 21, 221

Troxel M. A., et al., 2017, preprint, (arXiv:1708.01538)

Vargas-Magaña M., et al., 2013, , 554, A131

Wechsler R., DeRose J., Busha 2017, in prep.

York D. G., et al., 2000, , 120, 1579

Zuntz J., Kacprzak T., Voigt L., Hirsch M., Rowe B., Bridle S., 2013, , 434, 1604

Zuntz J., et al., 2015, Astronomy and Computing, 12, 45

Zuntz J., et al., 2017, preprint, (arXiv:1708.01533)

Zwicky F., 1937, , 86, 217

de Jong J. T. A., et al., 2015, , 582, A62

de Lapparent V., Geller M. J., Huchra J. P., 1986, , 302, L1

van Uitert E., et al., 2018, , 476, 4662