Universitat Autònoma de Barcelona

Photometric redshift calibration in DES Y3

Giulia Giannini

Photometric redshift calibration in DES Y3

Author: Giulia Giannini

Institut de Física d'Altes Energies (IFAE)

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

Department de Física
Universitat Autònoma de Barcelona

Thesis Advisor:

Dr. Ramon Miquel Pascual

Tutor:

Dr. Maria Pilar Casado Lechuga

Contents

Introduction and Motivation

In this thesis are presented two powerful methods to calibrate the redshift distributions of galaxy samples down the accuracy required by 3x2pt cosmological analyses. In particular, we are going to focus on the redshift calibration pipeline developed and implemented in DES Y3: it is the combination of a method called *Self-Organizing Maps Photometric Redshifts (SOMPZ)* and the more established *clustering redshifts cross-correlations (WZ)*. The first method relies on galaxy samples for which redshifts are known with high accuracy, and uses machine learning techniques to classify both the high quality photometry of the redshift samples and the wide photometry of the target galaxy sample, in order to transfer the redshift information from one class of galaxy to the other. The second method it exploits spatial overlap of another high quality redshift sample, of which this time the photometry is irrelevant. The redshift information is retrieved my measuring an angular two-point cross-correlation between the two samples. Given the complementarity of the methods, the combination of the information on the estimated redshift distribution is particularly powerful for the analysis we are interested in. Systematic uncertainties related to photometric measurements do not affect the WZ method, and viceversa, clustering-related systematic uncertainties as the unknown relation of the galaxy-matter bias, are irrelevant for the SOMPZ method. The combination is performed by sampling redshift distributions out of a joint likelihood, in which the data in different form is compared to a prediction model.

We present the estimated redshift distributions and inherent uncertainties of the WZ method and its combination with SOMPZ for the DES Y3 weak lensing source galaxy sample, while for the fiducial lens sample MagLim we present the whole methodology, which has proven to be more accurate than the original DES Y3 redshift calibration. We study the impact of using the newly calibrated redshift distribution of the MagLim lens sample on cosmological constraints from galaxy clustering and galaxy-galaxy lensing. Assuming a CDM cosmology, we obtain a 0.4 shift in the matter density and clustering amplitude plane compared to the fiducial DES Y3 results, highlighting

the importance of the redshift calibration of the lens sample in multi-probe cosmological analyses.

Part I

Preliminars

Chapter 1

Cosmological Background

Cosmology is the branch of astrophysics studying the origin, composition and evolution of the Universe. Humanity has wondered about these fundamental topics for millennia, observing the sky, forming queries and speculating on what the correct interpretation would be. Among these, an apparently simple question unveils one of the most important answers: "why is the cosmos dark?". The apparently trivial answer unveils a complex layering of interesting elements. In 1880s the Universe was believed to be static, but there was an ongoing debate on whether it was infinite. The first who gave an answer to the darkness question (called the Olbers' paradox) from the correct point of view was an intellectual most known for his novels: the writer Edgar Allan Poe suggested that if the Universe had been indeed infinite and had not had an origin, in any point of the sky we would see as much light as at the surface of the Sun. The light from any distant galaxy would have had the time to travel across an infinite distance and reach us. Even assuming that light would be partially absorbed by gas clouds across the line of sight, it was clear that we were not living the scenario described by the paradox. The conclusions from this rudimentary experimental observation are two: stars emit light for a limited amount of time, and the Universe must be evolving. Fast forward to the present time, today we know the Universe is expanding with accelerated rate. In order to justify today's observations, the expansion is required to be powered by an enormous amount of energy creating a negative pressure, which would have to be the most abundant among the components of the Universe. This so called "dark energy" remains one of the biggest mysteries on modern cosmology.

1.1 The expanding Universe

Modern cosmology begins in 1916 with the formulation of Einstein's general relativity, which links the content of the Universe with its geometry, describing gravity as a geometrical property of spacetime. One of its pillars is the so called *cosmological principle*: the Universe is homogeneous and isotropic, i.e. there is no preferred position or orientation. Despite this being of course an approximation, it is fairly robust at the large scales (100 Mpc) as supported by observations.

During the first 20 years of the XIX century, the Universe was believed to be static, assumption backed up by many measurements of stars' motions in our galaxy. The original description of the Universe using Einstein's theory implied a dynamic spacetime; this led Einstein to introduce the famous cosmological constant Λ, which appeared as a term in his equations to allow a static solution. We will then see that other solutions to the Einstein equations with Λ exist and also allow for an expanding universe, which constitutes the benchmark model in current modern cosmology.

In cosmology, we are not able to directly measure the distance to the observed objects. This would change if we knew in advance the size of a certain category of objects, since we would be able to determine their distance from their apparent size in the sky. These objects are called "standard rulers". **On the other hand, if we knew their intrinsic brightness, we could indirectly derive the distance by measuring the incoming flux, as the intensity of light received is inversely proportional to the square of the distance. These objects, called "standard candles", are** precisely what Edwin Hubble used for his observations in 1929, thanks to which the static vision of the Universe crumbled definitively. By using Cepheid variable stars he observed that the nearby stars' host galaxies were receding from us (Fig. 1.1) with no preferred direction.

Galaxies appearing moving away from us is not in contradiction with the cosmological principle, and does not turn Earth into a special location in the Universe; any other observer in any other location would see galaxies receding from them, with no preferred direction. This follows from the invariance of physical laws across reference systems: homogeneity translates into invariance under spatial translation, while isotropy translates as invariance under rotations of our reference system. The relation between velocity and position can then be written as the famous Hubble law:

$$v = H_0 r, \tag{1.1}$$

where the Hubble constant H_0 linearly relates the radial velocity of a galaxy

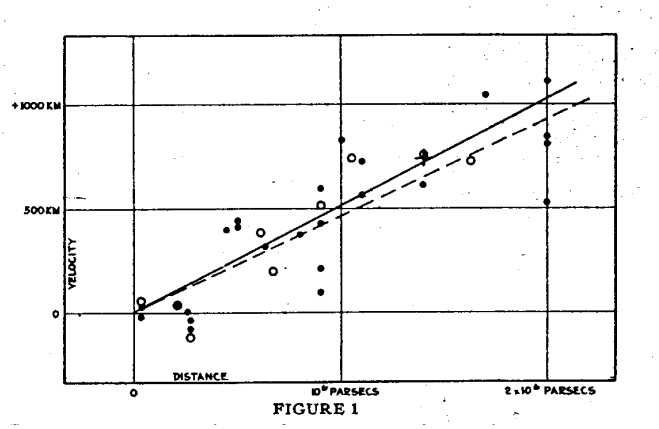

Figure 1.1: Original plot from the 1929 Hubble paper (Hubble, 1929), described as "Radial velocities, corrected for solar motion, are plotted against distances estimated from involved stars and mean luminosities of nebulae in a cluster. The black discs and full line represent the solution for solar motion using the nebulae individually; the circles and broken line represent the solution combining the nebulae into groups; the cross represents the mean velocity corresponding to the mean distance of 22 nebulae whose distances could not be estimated individually". **Note that the y-axis is erroneously labeled as "km", instead of "km/s".**

v and its distance r. Hubble evaluated the relation for the first time in 1929 using nearby galaxies; therefore, he only estimated the value of the Hubble constant at present time t_0: $H(t = t_0) \equiv H_0$. The first estimates of H_0 gave a value of $H_0 = 500$km/s/Mpc. This value is much higher than the one from the most recent measurements, because of unaccounted errors in its calibration. Today, current estimates of H_0 do not converge on a single value, but show large tension (4σ). Local measurements of the current expansion rate, e.g. based on supernovae as standard candles, result in values that gather around 73 km/s/Mpc. Instead, indirect measurements from the early Universe via the cosmic microwave background (CMB) predict the value of H_0 to be around 67 km/s/Mpc. This discrepancy is not understood yet, and it is one of the most intriguing problems in cosmology, the solution of which could possibly lead to the discovery of new physics.

1.1.1 The scale factor

The equivalence principle implies that any possible expansion must be equivalent in any region of the Universe. It is not required, though, to be constant with time. We can then introduce the scale factor a, which describes how distances evolve with cosmic time, i.e. how physical distances are related to

comoving distances:

$$\mathbf{r} = a(t)\chi, \tag{1.2}$$

where $\mathbf{r}$ is the physical distance, whereas χ is called the comoving distance. It is customary to fix $a(t = 0) \equiv a_0 = 1$. Comoving coordinates do not evolve as the Universe expands, while physical distances do. So if we had measured a distance in the past, today it would be larger by a factor $a(t)$. **The relation between the receding velocity and physical distance,** empirically measured by Hubble, can be derived theoretically from Eq. 1.2, by taking the time derivative of the radial part:

$$\frac{dr}{dt} = \frac{da(t)}{dt}d\chi = \frac{\dot{a}}{a}dr \equiv Hdr, \tag{1.3}$$

where we have defined $H \equiv \dot{a}/a$, and the symbol " ˙ " refers to the differentiation with respect to physical time t. The Hubble law is obtained by evaluating Eq. 1.3 at $t = t_0$ (i.e., today).

1.2 The FLRW metric

The metric is used in general relativity to measure distances between vectors: it describes the actual physical distance between two infinitesimally close points in spacetime defined in some arbitrary coordinate system. It is defined as

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu, \tag{1.4}$$

where $g_{\mu\nu}$ is the metric tensor, while ds^2 is the square of the spacetime interval, which is invariant under coordinate transformations (which means that different observers will always measure the same quantity). In the above definition we used the Einstein notation, which states that repeated indices in a single term imply the summation of that term over all the values of the index.

The metric tensor is usually derived by solving the Einstein's equations, but can also be obtained through purely geometrical considerations. We can start from the Minkowski metric which describes a flat spacetime, defined as

$$ds^2 = -c^2dt^2 + r^2. \tag{1.5}$$

We can generalise this metric for an expanding Universe by rewriting the radius in Eq. 1.5 in spherical coordinates:

$$ds^2 = -c^2dt^2 + a^2(t)[d\chi^2 + \chi^2 d\Omega^2], \tag{1.6}$$

where Ω is the solid angle. If we generalise it once more to consider the spacetime curvature, we obtain:

$$ds^2 = -c^2dt^2 + a^2(t)[d\chi^2 + S_k^2(\chi)d\Omega^2]. \tag{1.7}$$

This is known as *Friedmann-Lemaitre-Robertson-Walker* or *FLWR* metric. The term S_k depends on the curvature k of the Universe:

$$S_k = \begin{cases} \sin\chi & k > 0 \ (k = 1) \\ \chi & k = 0 \\ \sinh\chi & k < 0 \ (k = -1) \end{cases} \tag{1.8}$$

The radial part of the metric is independent of curvature, which affects only the measurements of angles.

1.3 Standard Model of Cosmology

1.3.1 Friedmann Equations

The main laws describing the evolution of the content of the Universe in cosmology follow from the application of the theory of general relativity, and in particular, from Einstein's equations:

$$G_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu}. \tag{1.9}$$

This beautiful equation relates the geometry of the Universe with its energy content: on the left hand side we have the Einstein tensor $G_{\mu\nu}$, the cosmological constant Λ and the metric $g_{\mu\nu}$ describing the spacetime geometry; on the right hand side we have the stress-energy tensor $T_{\mu\nu}$, which depends on the energy content of the Universe and describes the source of the gravitational field. The Einstein tensor can be expressed as:

$$G_{\mu\nu} = R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R, \tag{1.10}$$

where $R_{\mu\nu}$ is the Ricci curvature tensor, while R is the Ricci scalar, defined as the contracted form of the Ricci tensor $R_{\mu\nu}$, which is also obtained as a contraction of the Riemann tensor $R_{\alpha\mu\beta\nu}$, i.e. $R \equiv g^{\mu\nu} R_{\mu\nu} = g^{\mu\nu} g^{\alpha\beta} R_{\alpha\mu\beta\nu}$.

The Ricci tensor is linear in the second derivatives of the metric $g_{\mu\nu}$ and non linear in the first derivatives, and vanishes in a flat spacetime. It is generally expressed using the Christoffel symbols $\Gamma^{\mu}_{\alpha\beta} = \frac{1}{2} g^{\mu\nu} \left[g_{\alpha\nu,\beta} + g_{\beta\nu,\alpha} - g_{\alpha\beta,\nu} \right]$[1].

[1]The notation "$_{,\lambda}$" indicates differentiation with respect to the variable x^{λ}.

These are useful when describing an expanding spacetime, as they are essential in describing the *geodesic*, the shortest path across two points, which in an euclidian Universe is a straight line, but it becomes more complex in a curved spacetime. In an expanding Universe it is not possible to find a coordinate system in which the Cristoffel symbols vanish. The Riemann tensor is also referred to as the curvature tensor, as it vanishes in a flat spacetime. Its dependence on the derivatives of the metric can be understood looking at its definition:

$$R_{\alpha\beta\mu\nu} \equiv \Gamma^{\alpha}_{\beta\nu,\mu} - \Gamma^{\alpha}_{\beta\mu,\nu} - \Gamma^{\alpha}_{\kappa\nu}\Gamma^{\kappa}_{\beta\mu} + \Gamma^{\alpha}_{\kappa\mu}\Gamma^{\kappa}_{\beta\nu}. \tag{1.11}$$

Assuming a FLRW metric, and using the Einstein equation, we obtain that all the components of $R_{\mu\nu}$ are proportional to either $\ddot{a}$ or $\dot{a}^2$. It is useful to spell out the non-vanishing different components of the Ricci tensor; the ones with $\mu = \nu = 0$ and $\mu = \nu = i$. A series of paper filling but straight-forward calculations lead to the two components:

$$R_{00} = -3\frac{c^{-2}\ddot{a}(t)}{a(t)}, \tag{1.12}$$

$$R_{ij} = \frac{c^{-2}\ddot{a}(t)a(t) + 2c^{-2}\dot{a}(t)^2 + 2k}{a(t)^2} g_{ij}. \tag{1.13}$$

As for the stress-energy tensor, for a smooth isotropic universe it can be defined as $T^{\nu}_{\mu} = diag(-\rho c^2, p, p, p)$ where ρ and p stand for the energy density and the pressure. We now have all the ingredients to plug into the Einstein equations, and we can now derive the first Friedmann equation:

$$\left(\frac{\dot{a}(t)}{a(t)}\right)^2 = \frac{8\pi G}{3}\rho(t) + \frac{\Lambda c^2}{3} - \frac{kc^2}{a(t)^2}, \tag{1.14}$$

where $\dot{a}/a$ is the Hubble rate $H(t) \equiv \dot{a}/a$ we have already defined in Eq. 1.3. Equation 1.14 emphasizes that the rate of expansion depends on the Universe geometry and energy density.

The second Friedmann equation follows from the spatial components instead:

$$\frac{\ddot{a}(t)}{a(t)} = \frac{-4\pi G}{3}\left(\rho(t) + \frac{3p(t)}{c^2}\right) + \frac{\Lambda c^2}{3}, \tag{1.15}$$

Next, we perform a transformation that allows us to incorporate the cosmological constant in the energy content of the Universe, by considering it to be a fluid with density and pressure. This can be interpreted as "moving" the

Λ term from the left (geometrical) side of the Einstein equation to the right side:

$$\rho(t) \to \rho(t) + \frac{\Lambda}{8\pi G} \tag{1.16}$$

$$p(t) \to p(t) - \frac{\Lambda}{8\pi G} \tag{1.17}$$

We can then rewrite the Friedmann equations as

$$H(t)^2 = \frac{8\pi G}{3}\rho(t) - \frac{Kc^2}{a(t)^2}, \tag{1.18}$$

$$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{3}\left(\rho(t) + \frac{3p(t)}{c^2}\right). \tag{1.19}$$

The second equation highlights that the acceleration depends on the energy density as well as the pressure of the components of the Universe; Equation 1.19 is sometimes referred to as Friedmann acceleration equation. The Friedmann equations relates two quantities that can be measured: the expansion rate H and the density of energy/matter. The Hubble rate depends on the Universe geometry and total energy density, whereas the acceleration depends on the density and on the pressure.

1.3.2 Energy Content of the Universe

We now move on to examine the content of the Universe. From the cosmological principle, we can expect to be able to describe the properties of its components only using the mean density and pressure. In the previous section we have defined the energy-momentum tensor T, and in order to understand how the Universe evolves we have to understand how each component of the tensor evolves with time. Let us start from the law of conservation of the energy-momentum tensor:

$$T^{\mu}_{\nu;\mu} \equiv T^{\mu}_{\nu,\mu} + \Gamma^{\mu}_{\alpha\mu}T^{\alpha}_{\mu} - \Gamma^{\mu}_{\alpha\mu}T^{\alpha}_{\mu} = 0. \tag{1.20}$$

In the above equation, T^i_0 components vanish. The $\nu = 0$ component of the above system of equations is equivalent to the continuity equation, while the spatial part is equivalent to the Euler equation. Let us focus on the $\nu = 0$ component. By substituting the non-vanishing Christoffel's symbols, we derive the continuity equation for a perfect fluid in an expanding Universe:

$$\frac{\partial\rho}{\partial t} + 3\frac{\dot{a}}{a}\left(\rho + \frac{p}{c^2}\right) = 0. \tag{1.21}$$

In cosmology, we can approximate all the components of the Universe as diluted fluids, which obey a simple equation of state:

$$p(t) = w\rho(t)c^2, \tag{1.22}$$

with w a dimensionless constant. Using Equation 1.22, we can now rewrite the continuity equation as:

$$\frac{d\rho}{\rho} = -3(1+w)\frac{da}{a}. \tag{1.23}$$

Assuming that the Universe is constituted by one fluid, from the equation of continuity we can rewrite the equation of state as function of the scale factor a:

$$\rho(a) = \rho_0 a^{-3(1+w)}. \tag{1.24}$$

Using Equation 1.24 we can predict how density and scale factor for the different components in the Universe evolve with time. We are considering the following constituents:

- *Non-relativistic matter:* we can assume it to have zero pressure, as the random motion of the particles is much slower than the speed of light. Therefore this implies that $w = 0$. Non relativistic matter includes both baryonic and dark matter;
- *Relativistic particles:* photons and other massless particles have pressure $p = \rho/3$, leading to $w = 1/3$. We refer to this as radiation;
- Dark energy: the substance required to explain the accelerated expansion of the Universe, is characterised by an equation of state with $w < 1/3$, which would cause the right-hand side of the second Friedmann equation 1.19, to become positive. The particular case of $w = -1$ is what is called cosmological constant Λ.

Having considered the main components of the Universe and their equations of state, we can now compute the evolution in time of each of the components, which we summarise as follows:

$$p = w\rho c^2 \begin{cases} \text{Matter} & w_m = 0 & \rho_m = \rho_{m,0}a^{-3} \\ \text{Radiation} & w_r = \frac{1}{3} & \rho_r = \rho_{r,0}a^{-4} \\ \text{Dark Energy} & w_{\text{de}} = w(a) < -\frac{1}{3} & \rho_{\text{de}} = \rho_{\text{de},0}a^{-3[1+w(a)]} \\ \Lambda & w_\Lambda = -1 & \rho_\Lambda = \rho_{\Lambda,0} \end{cases} \tag{1.25}$$

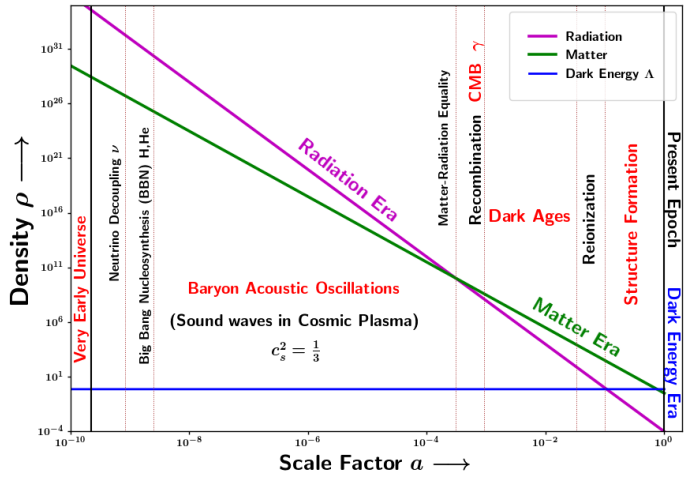

Figure 1.2: Evolution of density of matter, radiation, and cosmological constant Λ. Radiation was dominant in the first portion of the Universe's life, while the period from recombination onwards is referred to as Matter era. Only very recently, dark energy has become the dominant component, an issue referred to as the "coincidence problem". Credits: Swagat Saurav Mishra.

The *standard cosmological model* considers a Universe consisting of Dark Energy (the cosmological constant Λ with $w = -1$), non-relativistic matter (mostly cold dark matter) and radiation. This is referred to as ΛCDM, and it is considered the benchmark in cosmology.

In the wCDM model, a more generalised form of dark energy evolves with time, with equation of state:

$$w(a) = w_0 + (1-a)w_a = O(1-a)^2. \tag{1.26}$$

We define the cosmological parameters as

$$\Omega_s \equiv \frac{\rho_s(t_0)}{\rho_{cr}}, \tag{1.27}$$

where ρ_{cr} is the critical density, the value of density obtained when assuming flat curvature ($k = 0$) in the first Friedmann equation at current time:

$$\rho_{cr} = \frac{3H_0^2}{8\pi G}. \tag{1.28}$$

The sub-index s can represent any of the Universe constituents. It's general use to refer to all the matter in the universe with $\Omega_m = \Omega_b + \Omega_c$ (respectively baryonic matter and cold dark matter).

We can finally rewrite the first Friedmann equation in terms of the cosmological parameters.

$$H^2(a) = H_0^2[\Omega_r a^{-4} + \Omega_m a^{-3} + (1-\Omega_0)\, a^{-2} + \Omega_\Lambda], \tag{1.29}$$

where $Omega_0 = \Omega_r + \Omega_m + \Omega_\Lambda$ is the current total density of the Universe. The current relative densities as measured to the percent level and below by the Planck Collaboration are reported in Table 1.1.

1.4 Distances

As already introduced in the previous section, in an expanding Universe there are two main possible ways to describe distances: a *physical* distance that increases simultaneously with the expansion, and a *comoving* distance that doesn't.

If we consider a far away light source, the total comoving distance travelled by its light is

$$\chi = c \int_{t_e}^{t_o} \frac{dt'}{a(t')}, \tag{1.30}$$

Table 1.1: Cosmological parameters from *Planck* (Planck Collaboration, 2018). The parameter h is the Hubble constant H_0 in units of 100 km/s/Mpc. The total matter $\Omega_{m,0}$ is further divided into a dark-matter-only component and a baryonic $\Omega_{b,0}$ component. The curvature parameter value has been obtained combining *Planck* data with the data from Baryonic Acoustic Oscillations (BAO).

parameter	*Planck* TT,TE,EE+lowE+lensing
h	0.674 ± 0.005
$\Omega_{m,0}$	0.315 ± 0.007
$\Omega_{b,0}h^2$	0.0224 ± 0.0001
$\Omega_{\Lambda,0}$	0.685 ± 0.007
$\Omega_{r,0}$	$4.18343 \cdot 10^{-5}$
Ω_K	0.0007 ± 0.0019 (+BAO)

where t_e is the emitted time and t_o is the observed time. If we integrate the factor in Eq. 1.30 between $t = 0$ and an arbitrary time, we obtain what is called the *conformal time*. When multiplied by c is referred to as *comoving horizon*. This distance defines the limit at which information could have propagated from the Big Bang, also implying that regions separated by distances greater than this are not casually connected together. The comoving horizon is continuously extending at the speed of light, including new portions of the Universe that become in causal contact. For the special case in which we are integrating over the whole age of the Universe, the

Let us consider the angle $\Delta\theta$ subtended in the sky by an object of known size ℓ. We can use the metric to express its size as the distance between the two extremities of the object:

$$\ell = \sqrt{ds^2} = a(t)\chi\Delta\theta. \tag{1.31}$$

We can therefore define the *angular diameter distance* d_A as:

$$d_A \equiv \frac{\ell}{\Delta\theta} = a\chi \tag{1.32}$$

This definition holds only in an flat Universe. If we consider curvature, we have to rewrite Equation 1.32 as:

$$d_A = \frac{a}{H_0\sqrt{|\Omega_k|}} \begin{cases} \sinh(\chi H_0\sqrt{\Omega_k}) & \Omega_k > 0 \\ \sin(\chi H_0\sqrt{-\Omega_k}) & \Omega_k < 0. \end{cases} \tag{1.33}$$

Another way to measure distances to an object is to measure the flux emitted by an object of known luminosity (such as Chepheids stars, or Type Ia

Supernovae). The flux emitted by a point-like source with luminosity L is spread over spherical shells of area $4\pi d^2$. In an expanding Universe, the distance between the emitter and the observer increases over time, therefore it is necessary to write the flux in terms of the comoving distance and scale factor:

$$F = \frac{L(\chi)}{4\pi\chi^2(a)}, \tag{1.34}$$

where $L(\chi)$ is the luminosity for a spherical shell with comoving radius $\chi(a)$. If we assume that all the photons are emitted all at fixed energy, then L is directly proportional to the number of photons going through the spherical shell located at comoving radius χ. Due to expansion, the number of photons that crosses the shell within a specific time interval will decrease with time as function of a. The energy per unit time will then be proportional to a^2. Our equation for the observed flux will then become:

$$F = \frac{L_e(\chi)a^2}{4\pi\chi^2(a)}, \tag{1.35}$$

where L_e is the emitted luminosity. If we define the *luminosity distance* as

$$d_L \equiv \frac{\chi}{a}, \tag{1.36}$$

we can write the observed flux as function of d_L:

$$F = \frac{L_e(\chi)a^2}{4\pi d_L^2(a)}. \tag{1.37}$$

1.4.1 Redshift

The radiation emitted by bodies in movement with respect to the restframe system is subjected to the Doppler Effect: the measured wavelength increases or decreases because of the relative speed between the observer and the source, causing the light to be shifted towards one end of the spectrum (redshifted or blueshifted). In astronomy, because of the Universe expansion causing all galaxies to recede from us, it is convenient to define the relative change in wavelength of light as follows:

$$z \equiv \frac{\lambda_2 - \lambda_1}{\lambda_1}, \tag{1.38}$$

where z is called *redshift*.

We can consider the distance $\lambda_1 = ct_1$ between two wave peaks at the emission time and the reception time. Given the relative velocity of the emitter

with respect to the receiver, in the time interval $\delta t = 1/\nu_1$ the distance between two peaks changes by a factor vt_1. Therefore

$$\lambda_2 = \lambda_1 + \frac{v\lambda_1}{c} = \lambda_1(1 + \frac{v}{c}), \tag{1.39}$$

which using Eq. 1.38 leads to $z = v/c$. Note that for the relativistic case the equation would be

$$z = \sqrt{\frac{1 + v/c}{1 - v/c}} - 1. \tag{1.40}$$

Cosmological Redshift

We can now generalise this to our expanding Universe. The interesting outcome is that the wavelength of light from distant sources in an expanding Universe will be stretched proportionally to the scale factor:

$$\frac{\lambda_o}{\lambda_e} \equiv 1 + z = \frac{a_o}{a_e} = \frac{1}{a_e}. \tag{1.41}$$

Every cosmological object is receding from us (except for the ones in our neighbourhood where the peculiar velocity dominates over the cosmological redshift). For all the objects that are not standard rulers or candles, the redshift is usually the only information about their distance that can be measured. We can now rewrite the comoving distance in terms of redshifts

$$\chi(t) = c \int \frac{dt'}{a(t')} = c \int \frac{dz'}{H(z')} \tag{1.42}$$

At small redshifts we can write the comoving distance in terms of $\chi \sim z/H_0$.

Gravitational Redshift

A consequence of general relativity is that when light is travelling through a divergence of the gravitational potential, e.g., the emitter has a greater gravitational potential ϕ_1 than the receiver ϕ_2, a gravitational redshift occurs due to the loss of energy that the photon invests to "escape" from the gravitational attraction. The gravitational redshift is defined as

$$z_G \equiv \frac{\lambda_2 - \lambda_1}{\lambda_2} = \frac{\phi_2 - \phi_1}{c^2}. \tag{1.43}$$

The gravitational redshift is generally small compared to cosmological or Doppler redshifts. It is easy to show that the combined redshift is:

$$1 + z = (1 + z_C)(1 + z_D)(1 + z_G). \tag{1.44}$$

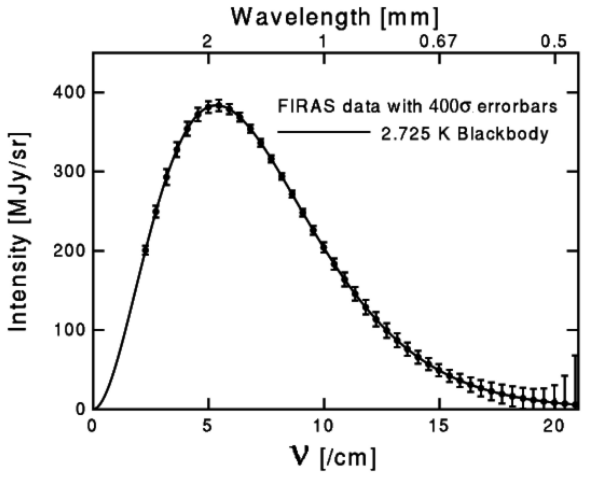

Figure 1.3: Spectrum of the CMB radiation as measured by the COBE FIRAS experiments, compared to the black-body spectrum for a 2.725K source. The errorbars are magnified by 400 times.

1.5 Cosmic Microwave Background

Some of the most outstanding discoveries happen by chance. In 1964 Prof. Penzias and Prof. Wilson were ready to test their new radiometer at the Bell Laboratories, intended for research in radio astronomy. While calibrating their radiometer they detected a steady, constant noise, much larger than they had anticipated. In order to measure any of their planned objects, they had to find the source of the noise and get rid of it. The signal was 100 times stronger than the radio waves they wanted to measure and it was present day and night, independently from the direction. For weeks, they tried to eliminate the noise, with no success. Left with few options, they even considered it could be caused by birds nesting on the antenna. Despite their efforts, which at least left them with a sparkly clean antenna, the noise was still there. They eventually realised they had just measured for the first time the Cosmic Microwave Background, which had been predicted by theory, beating prof. Dicke that had long planned a targeted experiment which he was about to implement.

Why is there such a remnant radiation, and what kind of radiation is it? As we know the Universe is expanding, in earlier times it must have been smaller, and as direct consequence, denser and hotter. In the first 380,000 years of its life, the Universe was a hot plasma of particles, opaque to light, as photons were constantly interacting with electrons via Compton or Coulomb scattering, and ionising hydrogen and helium atoms immediately after they were formed.

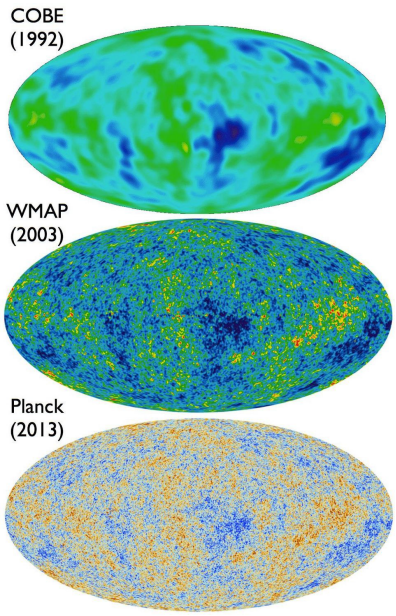

Figure 1.4: Anisotropies in the CMB temperature map as observed by COBE (Smoot, 1992), WMAP (Bennett, 2015) and Planck (Planck Collaboration, 2019). The fluctuations are highlighted with increasing resolution, always using a color scale of the order 10^5.

In these extreme conditions, photons were in thermodynamic equilibrium, such that radiation in these extreme conditions could be completely characterised by a black-body spectrum. Radiation in fact thermalised during the numerous interactions happening in the primitive Universe.

Radiation in thermal equilibrium at a given temperature T is described by a Bose-Einstein distribution; therefore photons permeating the Universe are described by the following relation between their energy density and temperature:

$$\rho_r \propto T_\gamma^4. \tag{1.45}$$

The temperature of the radiation decreases as the Universe expands with a factor of a with respect to the current size, due to the increase in a factor of a to the wavelength associated to this radiation:

$$T_\gamma(a) = \frac{T_{\gamma,0}}{a} \tag{1.46}$$

As the Universe expanded, the temperature of the radiation gradually dropped, and at about 380,000 years after the Big Bang, the temperature was about $T \simeq 3300K$*, equivalent to* $\sim 1eV$*. At this point the number of neutral atoms reached and surpassed the number of ionised atoms, as the radiation was cooling. This epoch is called recombination. The density of free-electrons dropped, which greatly increased the photons' mean free path, turning the Universe transparent. The era of decoupling between radiation and matter occurred abruptly when photons ceased to interact with electrons, and their free path became comparable with the size of the Universe back then. This is equivalent to the rate of Compton scattering of photons being approximately equal to the rate of expansion of the universe* H_0. From this moment on, the thermalised radiation has been too cold to interact with matter and kept cooling to this day, reaching a temperature of $T_o = 2.72548 \pm 0.00057K$ (Fixsen, 2009), equivalent to an energy of $k_B T_0 = 2.3 \times 10^-4eV$, or to a radiation density: $\rho_\gamma \sim 4.65 \times 10^{-34}(T_0/2.726K)^4 gr/cm^3$.

The COsmic Background Explorer (COBE) satellite in 1992 confirmed that the CMB had a black body spectrum. Since then, the CMB has been observed with greater accuracy by WMAP and more recently by *Planck*. The measured blackbody spectrum agrees with the one of radiation in thermal equilibrium with an astonishing precision, making the CMB the most accurate black-body spectrum in nature. Despite being homogeneous and isotropic at the percent level, all three experiments measured small anisotropies. Beyond simpler discrepancies caused by the motion of Earth with respect to the CMB rest frame, differences of the order of 1 part in 10^5 are present. These fluctuations are

believed to be the remnant of the quantum fluctuations present at the time of recombination, which eventually led to the structures (galaxies, clusters of galaxies, filaments, etc.) we see today.

1.6 The Large Scale Structure of the Universe

So far we have treated the Universe as smooth and uniform, assumptions based on the cosmological principle. While this is valid at large scales, when considering scales smaller than 100 Mpc it become obvious that it does not hold anymore. We know in fact that the density across the Universe is far from uniform: we have stars and galaxies, and then immense empty spaces where only few atoms can be found. What caused these structures to form?

1.6.1 The theory of inflation

In a time t after the Big Bang, light travelled a distance d_H given by the comoving horizon (defined in section 1.4). In a Universe dominated by radiation the expansion goes as $a \sim t^{1/2}$, while if it is dominated by matter $a \sim t^{2/3}$, leading to $d_H \sim a^2$ or $d_H \sim a^{3/2}$ respectively. This implies that the comoving horizon grows faster than any other physical distance λ, which grows with the scale factor: $\lambda \sim a$. Therefore, if we looked in the past, we would reach a time when $\lambda > d_H$, meaning that all physical scales would be causally disconnected. Figure 1.5 helps us to identify the moment in which the scale λ entered the causal horizon, $\lambda = d_H$, by plotting the physical distance λ as a function of the scale factor a. The existence of the causal horizon implies two problems for the standard cosmological model:

1. CMB regions separated by more than 1 degree are supposed to never have been in causal contact. The physical distance corresponding to larger angles is larger than the causal horizon at $z = 1100$. Therefore there is apparently no physical process which could explain why the temperature of the CMB map is identical for portions of the sky that are separated by more than 1 degree;

2. Physical scales smaller that 100 Mpc came in causal contact when the Universe was dominated by radiation. The coupling between matter and radiation at that time would have erased any inhomogeneities that could have been present in the initial conditions. This creates a problem in understanding how structure formed.

The most simple inflation model proposes the existence of a scalar field called *inflaton* ϕ, which has energy density $\rho = 1/2\dot{\phi}^2 + V(\phi)$, where $V(\phi)$ is the potential. As this is a homogeneous scalar field, it behaves similarly to a single particle moving in a potential. The scalar field is expected to "slowly roll" toward its ground state, meaning it has little kinetic energy $T(\phi) = 1/2\dot{\phi}^2$. The potential energy is nonzero, however, so the pressure $\rho = T(\phi) - V(\phi)$ is

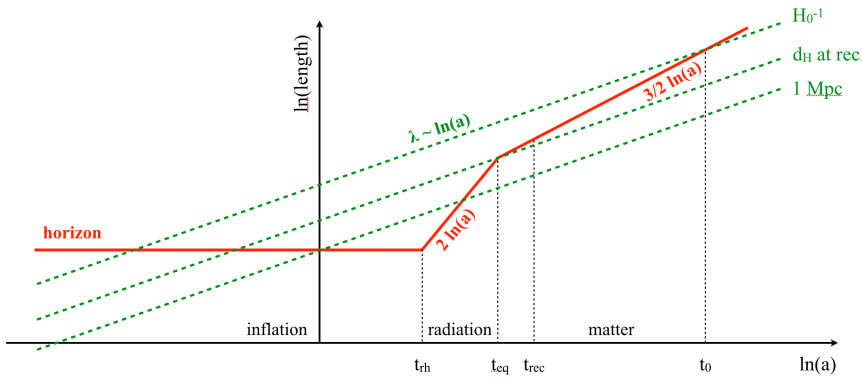

Figure 1.5: Time evolution of the physical scales for three fluctuations of increasing size, during the inflationary phase and the subsequent eras. Without inflation, the fluctuations would initially be outside the horizon, which would have not made possible the existence of those fluctuations in the first place, as they enter the horizon in the radiation and matter dominated eras.

negative, inducing an exponential expansion and greatly diluting the particles in the Universe. After reaching its minimum, inflation ends, as the field will oscillate and decay into lighter particles. If the density is dominated by a constant value, therefore also H will be constant (Figure 1.5). The quantum fluctuations in the inflaton field make the inflation time slightly different for different locations in space, resulting in a non uniform density. These fluctuations are believed to be the initial seeds that then lead to the formation of the small anisotropies imprinted in the CMB measurements, as we have seen in section 1.5, which then lead to the much larger structures we can observe today.

Finally, in order to fully solve the second problem, the existence of cold dark matter is required, as this type of matter does not interact with radiation and its inhomogeneities would have survived the coupling with radiation at the end of the radiation era.

1.6.2 Structure Formation

As introduced in section 1.6.1, the currently accepted explanation on how the massive structures we see today formed is provided by the inflationary theory.

Immediately after inflation, most of the Universe was not in causal contact, as the physical distances increased exponentially, much faster than the Hubble horizon. After inflation stopped, larger and larger scales continued to be included within the horizon, because the latter continued to grow with

cosmic time. Fluctuations whose size is larger than the Hubble horizon are frozen, because those scales are not in causal contact yet; smaller ones can instead interact gravitationally and can therefore grow. Figure 1.5 shows how the physical scale of fluctuations evolves with time, or with the scale factor, with respect to the Hubble horizon.

Linear evolution

In order to describe the evolution of the various components of the Universe, and during the different eras in which radiation, matter or dark energy dominated, we must obtain the three equations describing our cosmological perfect fluids: the continuity equation to ensure the conservation of mass, the Euler equation which represents the conservation of momentum, and the Poisson equation. The linear evolution of the fluctuations can be obtained by introducing perturbations to the density ρ and the pressure p: The perturbed FLRW metric and the perturbed stress energy tensor read:

$$g_{\mu\nu} = \begin{pmatrix} -(1+\frac{2\Phi}{c^2}) & 0 & 0 & 0 \\ 0 & a^2(1-\frac{2\Phi}{c^2}) & 0 & 0 \\ 0 & 0 & a^2 f_K(\chi)^2(1-\frac{2\Phi}{c^2}) & \\ 0 & 0 & 0 & a_K^2(\chi)^2(1-\frac{2\Phi}{c^2}) \end{pmatrix}, \tag{1.47}$$

$$T_{00} = \delta\rho c^2, \quad T_{ii} = \delta p, \quad T_0^i = (\rho c^2 + p)v^i, \tag{1.48}$$

where Φ is the gravitational potential related to the perturbations $\delta\rho$ and δp, and v^i the peculiar velocity with respect to the expansion of the Universe.

We can obtain the continuity equation and the Euler equation respectively from the 00-th and the ii-th components of the stress energy tensor derivative. By keeping the first order only, the two equations become:

$$\delta\rho_{,0} = -\left(\rho + \frac{p}{c^2}\right)\left(\frac{v^i_{,i}}{a}\right), \tag{1.49}$$

$$\frac{\delta p_{,i}}{a} = -\left(\rho + \frac{p}{c^2}\right)\left(Hv^i + v^i_{,0} + \frac{\Phi_{,i}}{a}\right). \tag{1.50}$$

The Poisson equation instead derives from the 00-th component of the perturbed Einstein equation:

$$\Phi_{,i,i} = 4\pi G a^2\left(\delta\rho + 3\frac{\delta p}{c^2}\right). \tag{1.51}$$

The combination of the time derivative of Equation 1.49 with 1.50 and 1.51, for a dark matter dominated Universe, gives us:

$$\ddot{\delta} + 2H\dot{\delta} - 4\pi G\bar{\rho}\delta = 0. \tag{1.52}$$

There are two solutions to this differential equation, one growing with time and one decaying. For the growing mode, the solution can be written as:

$$\delta(\mathbf{x}, t) = \frac{D(t)}{D(t_0)}\delta(\mathbf{x}, t_0), \tag{1.53}$$

with $D(t)$ the "growth factor", which depends on the cosmological parameters. The above solution holds only when the Universe is dominated by dark matter.

Non-linear evolution

If we consider small scales, overdensities become $\delta \gtrsim 1$, therefore the first-order approximation (linear theory) is not sufficient to describe their evolution and the above derivation of the power spectrum becomes inadequate. Gravity in fact becomes stronger, and the increasingly faster downfall of matter is generally described by the "spherical collapse" model. The growth is hierarchical, with fluctuations larger than a critical value δ_c bind forming virialised structures of dark matter called *halos*, which with time grow through accretion of mass or by merging with neighboring structures. As predictions of the non-linear evolution are extremely complex to compute from theory, it is customary to use N-body simulations to inform analytical models.

1.6.3 Statistics of the Matter Density field

In order to describe and quantify how clumpy and clustered matter is, we need to introduce a few statistical tools. It is impossible to predict where a specific galaxy will form. We can only study structure formation as a statistical quantity, comparing the theoretical predictions to the measurements considering a collection of galaxies. The simple average of the density field has no meaning, as we know that at large scales the density is constant. A more useful statistics is the so called *two-point correlation function*, which informs on how clustered matter is and on what is the separation at which matter is typically clustered. Because of the cosmological principle, it depends only on the absolute value of the separation and not on the objects' orientation, and it is defined as

$$\xi(\mathbf{x}, \mathbf{y}) \equiv \langle \delta(\mathbf{x}), \delta(\mathbf{y}) \rangle = \xi(|\mathbf{x} - \mathbf{y}|). \tag{1.54}$$

The Fourier transform of the two-point correlation function is what is called the *power spectrum*.

We start by defining the density field fluctuations in Fourier space. Since the density field is defined in three dimensions, we expand as

$$\delta(\mathbf{x}) = \frac{1}{(2\pi)^3}\int d^3k\tilde{\delta}(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{1.55}$$

By inverting this we obtain the transform of the density contrast:

$$\tilde{\delta}(\mathbf{k}) = \int d^3x\delta(\mathbf{x})e^{-i\mathbf{k}\cdot\mathbf{x}}. \tag{1.56}$$

If we perform a variable change to $\alpha = ||\mathbf{x}-\mathbf{y}|$ therefore can write the Fourier transform of the two-point correlation function as

$$\begin{aligned}\left\langle\tilde{\delta}(\mathbf{k})\tilde{\delta}^*(\mathbf{k})\right\rangle &= \int d^3xe^{-i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}}\int d^3\alpha e^{-i\mathbf{k}'\cdot\alpha}\xi(\alpha) = \\ &= (2\pi)^3\delta_D^3(\mathbf{k}-\mathbf{k}')\int d^3\alpha e^{-i\mathbf{k}'\cdot\alpha}\xi(\alpha) = \\ &= (2\pi)^3\delta_D^3(\mathbf{k}-\mathbf{k}')P(\mathbf{k}),\end{aligned} \tag{1.57}$$

where δ_D^3 is the Dirac delta, and $P(\mathbf{k})$ is the matter power spectrum.

The inflationary theory predicts also the *primordial power spectrum*, which refers to the initial density fluctuations:

$$P(\mathbf{k}) = A_s\mathbf{k}^{n_s}, \tag{1.58}$$

where A_s is the amplitude and n_s is the spectral index, measured by the Planck collaboration to be $n_s = 0.965 \pm 0.004$. The evolution of the density field of the primordial into the linear matter power spectrum today is parameterised by the transfer function as in

$$P(\mathbf{k}, z) = T^2(\mathbf{k}, z)P_0(\mathbf{k}). \tag{1.59}$$

The transfer function takes into account the impact of all physical phenomena that happened throughout the different eras which modified the growth of structure, and it depends on the cosmological model assumed. Therefore the matter power spectrum today also depends on the cosmology.

The amplitude of the power spectrum has to be constrained by observations, **and it is measured by considering the density fluctuations inside a sphere, the radius of which has been conventionally set to 8 Mpc/h, which is why this quantity is called σ_8. Such radius has been chosen because at around that scale the amplitude is close to unity**. The latest *Planck* results have measured $\sigma_8 = 0.811 \pm 0.006$.

1.6.4 Galaxy-matter bias

According to ΛCDM, most of the matter is in form of dark matter, which is driving the structure formation process. Dark matter, however, is not directly accessible; for instance, we do not measure the two-point correlation function of dark matter clumps, but only the two-point correlation function of galaxies. Despite the fact that luminous matter undergoes the same gravitational pull as dark matter, baryons are also subjected to other interactions, which change how the two kinds of matter behave at small scales. This statistical difference is called *galaxy-matter bias*:

$$\delta_g = b\delta. \tag{1.60}$$

We can extend this relation to the two-point correlation functions and the power spectrum:

$$\xi_{gg} = b^2 \xi_{\delta\delta} \tag{1.61}$$

$$P_{gg} = b^2 P_{\delta\delta} \tag{1.62}$$

The bias relates the amplitude of the matter and galaxy fluctuations, as they tend to be different. The latter are generally larger, as galaxies form at the peak of a matter overdensity. Any further difference between the galaxy and matter distribution is parametrised by the cross-correlation coefficient r:

$$r = \frac{\xi_{\delta g}}{\sqrt{\xi_{\delta\delta}\xi_{gg}}}, \tag{1.63}$$

where $\xi_{\delta g}$ is the matter-galaxy cross-correlation function, related to the matter correlation function as

$$\xi_{\delta g} = b \cdot r \; \xi_{\delta\delta}. \tag{1.64}$$

This relation is assumed linear at large scales, but it is known to break down at small scales due to stochastic components, very complex to take into account. On large enough scales though, $\xi_{\delta\delta}$ and $\xi_{\delta g}$ are expected to correctly trace the matter structures, and the r coefficient is expected to be close to unity. It follows that by simultaneously measuring both $\xi_{\delta\delta}$ and $\xi_{\delta g}$ it would be possible to obtain constraints on the bias parameter. The two-point correlation function between galaxies can only be measured from the positions of a galaxy sample. For the matter-galaxy correlation function it is necessary to have an external measurement able to probe the matter distribution directly, which is not possible by measuring positions. Such information can be obtained with the use of weak gravitational lensing, introduced in the following section.

1.7 Weak Gravitational Lensing

One of the Universe's most fascinating phenomenon is gravitational lensing, which refers to the deflection of light trajectories due to the effect of gravitational fields. According to general relativity, the metric of the spacetime is shaped by the energy mass distribution of the Universe. Any massive object, therefore, bends the spacetime. Photons propagate following the shortest path in spacetime, which is defined by $ds^2 = 0$ (also known as geodesic); if the spacetime is bent by the presence of a massive object, the shortest path will not appear as a straight line to the observer anymore. As a consequence, the shape of distant objects can appear magnified, distorted, or even multiple times, due to the gravitational effects of the large scale structure of the Universe. In Figure 1.6, recently captured by the James Webb Space Telescope, the several arches oriented in the same direction are examples of how galaxy images get distorted through gravitational lensing.

Let us start considering the time the photons take to travel from the source to the observer. This can be obtained integrating $ds^2 = 0$ along the line of sight:

$$t = \frac{1}{c} \int \left(1 - \frac{2\Phi}{c^2}\right) dr. \tag{1.65}$$

This equation is similar to the case of geometrical optics, with the light moving through a medium with a refractive index n given by the lensing potential: $n = \left(1 - \frac{2\Phi}{c^2}\right)$.

We are now interested in obtaining the expression for the deflection angle $\hat{\alpha}$, which is the angle between the emitted and received lights paths, as illustrated in Figure 1.7. This can be obtained by solving $\delta t = 0$, whose solutions minimise the light path. This is a variational problem, whose solution is:

$$\hat{\alpha} = \frac{2}{c^2} \int \nabla^p \Phi dr. \tag{1.66}$$

Eq. 1.66 is meant to be integrated along the photons trajectory, **while the gradient of the lensing potential $\nabla^p \Phi$ is evaluated along the perpendicular direction**. Unfortunately, we are not able to integrate over the actual path of the photons, which is very complex and unknown to us. But since the gravitational potential is small ($\Phi/c^2 \ll 1$), the deflection and therefore the angles in play are expected to be small. This allows us to use the Born approximation and integrate over the unperturbed light path connecting the image (not the source) and the observer.

Let us start from a simple case: a point-like body with mass M and gravi-

Figure 1.6: First deep field image from the James Webb Space Telescope. The Einstein rings are very visible. Credits: JWST

tational potential $\Phi = -GM/r$. In this case, the deflection angle becomes

$$\hat{\alpha} = 4\frac{GM}{bc^2} \tag{1.67}$$

where we have introduced the impact parameter b, which is the distance of closest approach of the photons to the body.

A more complex case is when a layer of mass is considered. This is also referred to as "thin lens" approximation. The matter distribution of the lens is described by its surface density:

$$\Sigma(\xi) = \int \rho(\xi, z)dz. \tag{1.68}$$

where ξ is a two-dimensional vector on the lens plane, and z is the radial direction; since we are in the thin lens approximation, we assume most of the

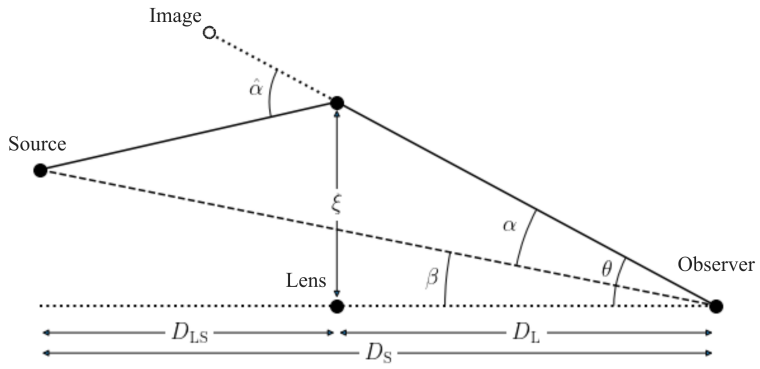

Figure 1.7: idealised version of a gravitational thin lens system. D_S is the distance between the observer and the source; D_L is the distance between the observer and the lens, D_{LS} is the distance between the lens and the source. Having defined the optical axis as the line connecting the observer and the center of the lens, β is the angle between the source and the optical axis; the light coming towards the observer is deflected by α; the angle between the deflected "apparent" image and the optical axis is θ.

density to be localized at $z \approx z_0$. In this case, the deflection angle is

$$\hat{\alpha}(\xi) = \frac{4G}{c^2} \int \frac{(\xi - \xi')\Sigma(\xi')}{|\xi - \xi'|^2} d^2\xi. \tag{1.69}$$

The thin lens approximation, despite being a simplification, is realistic enough that can be applied successfully to describe a number of cases. For instance, in the case of galaxy clusters, the physical size of the clusters is much smaller than the distances between the observer and the cluster, or the cluster and the source of photons. Therefore, assuming a two-dimensional cluster is a fair assumption.

Figure 1.7 represents a thin-lens system; although the figure is lacking one dimension, angles have both an amplitude and a direction. The amplitude of the angles describes the tilt with respect to the z-axis (the horizontal dotted line connecting the observer to the center of the lens), whereas the direction specifies the location in the plane perpendicular to the z-axis.

The key quantities in thin-lens systems are three distances and three angles. In particular: D_S, which is the distance between the observer and the source, D_L, which is the distance between the observer and the lens, and D_{LS}, which is the distance between the lens and the source. Having defined the optical axis as the line connecting the observer and the center of the lens, the three key angles are β, that is the angle between the source and the optical axis; $\hat{\alpha}$, which describes the deflection angle of the light coming towards the observer;

θ, which is the angle between the deflected "apparent" image and the optical axis. Having given these definitions, we can write the lens equation:

$$\theta D_S = \beta D_S + \hat{\alpha} D_{LS}. \tag{1.70}$$

If we introduce the reduce deflection angle as

$$\alpha(\theta) = \frac{D_{LS}}{D_S}\hat{\alpha}(\theta), \tag{1.71}$$

then the lens equation takes the form

$$\beta = \theta - \alpha(\theta). \tag{1.72}$$

This apparently simple equation hides a certain level of complexity: in the strong-lensing regime, the relation between θ and β is non linear and has multiple solutions.

The deflection angle α can be determined knowing the lensing potential associated to the lens. Always assuming the thin lens approximation, we can write the lensing potential as:

$$\psi = \frac{2}{c^2}\frac{D_{\rm LS}}{D_{\rm L} D_{\rm S}} \int \Phi(D_{\rm L}\theta, z) dz \tag{1.73}$$

Note that here z still refers to the coordinate axis, not to redshift. From this equation we can see that the effect will be larger when the angular diameter distance to the lens is comparable to the angular diameter distance between source and lens.

Lensing occurs only when there are spatial deviations $\nabla_\theta \psi$ from the average gravitational potential. These variations can be related to deflection angle; in particular:

$$\alpha(\theta) = \nabla_\theta \psi. \tag{1.74}$$

From Equation 1.74 we can define a very crucial quantity called *convergence* κ, which relates to the laplacian of the potential as follows

$$\nabla^2_\theta \psi = 2\frac{\Sigma(D_L\theta)}{\Sigma_{\rm crit}} \equiv 2\kappa(\theta), \tag{1.75}$$

where we have used the critical surface mass density, which is usefully defined to have inverse dimensions with respect to the surface density, such that the pre-factor of Equation 1.73 cancels out:

$$\Sigma_{\rm crit} = \frac{c^2}{4\pi G}\frac{D_S}{D_L D_{LS}}. \tag{1.76}$$

The convergence κ is a dimensionless quantity which has an important physical meaning: it is the integrated matter density along the line of sight, weighted by a lensing efficiency.

1.7.1 Shear and Magnification

Let's now consider only the weak lensing regime ($\kappa \ll 1$). If the size of the source is much smaller than the scale of variation related to the deflection angle, we can "linearise" the lens equation as follows:

$$\delta\beta \sim A\delta\theta, \tag{1.77}$$

with A being the Jacobian matrix mapping the lensed coordinates θ to the unlensed ones β:

$$A_{ij} = \frac{\partial\beta_i}{\partial\theta_j} = \delta_{ij} - \frac{\partial\alpha_i}{\partial\theta_j} = \delta_{ij} - \partial_i\partial_j\psi. \tag{1.78}$$

In the case of a uniform gravitational potential, the matrix becomes the identity matrix. In the presence of a lensing potential, instead, the mapping between coordinates is determined by the second derivatives of ψ.

It is more convenient to express the matrix A using the convergence field κ and the spin-2 shear field γ, whose definition will follow in a moment. It is standard practice to write γ in terms of its two-components:

$$\gamma \equiv \gamma_1 + i\gamma_2 = |\gamma|e^{2i\varphi}, \tag{1.79}$$

where the amplitude of the shear field describes the degree of the distortion an image goes through, whereas φ is the angle of distortion. We can write the matrix A as follows:

$$A = \begin{pmatrix} 1-k-\gamma_1 & -\gamma_2 \\ -\gamma_2 & 1-\kappa+\gamma_1 \end{pmatrix}. \tag{1.80}$$

With this definition of the matrix A, the convergence and the two components of the shear are related to the second derivatives of the lensing potential:

$$\kappa = \frac{1}{2}(\partial_1\partial_1 + \partial_2\partial_2)\psi = \frac{1}{2}\nabla^2\psi \tag{1.81}$$

$$\gamma_1 = \frac{1}{2}(\partial_1\partial_1 - \partial_2\partial_2)\psi \tag{1.82}$$

$$\gamma_2 = \partial_1\partial_2\psi. \tag{1.83}$$

It is instructive to write A separating the contribution from the convergence and shear components:

$$A = (1-\kappa)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - |\gamma|\begin{pmatrix} cos(2\varphi) & sin(2\varphi) \\ sin(2\varphi) & -cos(2\varphi) \end{pmatrix}. \tag{1.84}$$

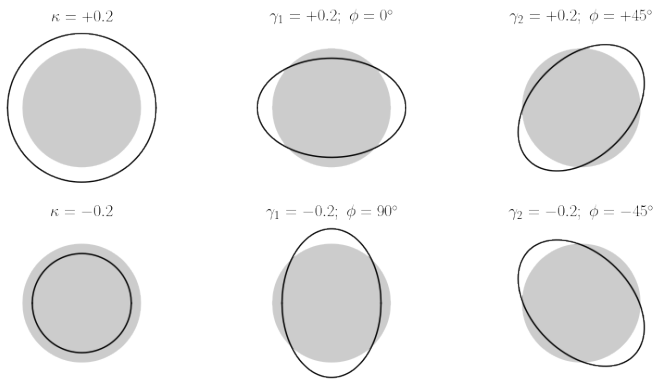

Figure 1.8: Illustration of the effect of the convergence κ and shear field components $\gamma_{1,2}$ on a spherical source. For positive values of convergence, the image is magnified, vice versa for negative values. The shear components produce an anisotropic stretching of the image in different directions.

From this equation we can notice that the shear components are describing the change in shape of the image (e.g., a circle that undergoes a coordinate transformation described by A becomes an ellipse), while the convergence is responsible of the change in size of the image, due to the term $1 - \kappa$. Figure 1.8 illustrates the effect of these distortions.

The amplification matrix A maps the unlensed coordinates β to the lensed coordinates given by θ. What we are actually interested in is the reverse, that is how to obtain the original coordinates giving our observation of the lensed coordinates. We can therefore invert it as

$$d\theta = A^{-1} d\beta. \tag{1.85}$$

The matrix A is guaranteed to be invertible in the weak lensing regime where κ and γ are small, as the determinant is $detA = (1-\kappa)^2 - |\gamma|^2 \sim 1 - 2\kappa$. We can therefore write

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} 1 - \kappa + \gamma_1 & \gamma_2 \\ \gamma_2 & 1 - \kappa - \gamma_1 \end{pmatrix} \tag{1.86}$$

The multiplicative factor $1/\det A$ represents how much the solid angle spanned by the image has changed with respect to the initial unlensed source:

$$\mu \equiv \frac{1}{\det A} = \frac{1}{(1-\kappa)^2 - |\gamma|^2} \sim 1 + 2\kappa \tag{1.87}$$

This is called the *magnification* factor. For the weak lensing regime we can consider at first order the convergence field κ solely responsible for the change in solid angle.

To describe the shape distortions produced by the gravitational field, we can consider a circular source whose image is deformed in such a way to become an ellipse with semi-major axis a and semi-minor axis b. The ellipticity of the distorted image is defined as

$$\epsilon \equiv \frac{a-b}{a+b} = \frac{\gamma}{1-\kappa} \equiv \mathbf{g}, \tag{1.88}$$

where we introduced a new quantity $\mathbf{g}$ called *reduced shear*, which is the quantity that can actually be measured with observations. As we are not able to observe the ellipticity of the source prior to lensing, the relevant quantity becomes the reduced shear itself. A reasonable approximation for weak lensing is $g \sim \gamma$, since $\kappa \ll 1$.

In the real Universe, galaxies do not tend to have a circular shape, but they are characterised by an intrinsic ellipticity ϵ^s. We now have to consider a distortion over an already elliptical shape. We therefore re-define the ellipticity as function of the intrinsic ellipticity and the reduced shear:

$$\epsilon = \frac{\epsilon^s + g}{1 + g^* \epsilon^s}, \tag{1.89}$$

which in the weak lensing regime can be simplified to

$$\epsilon \sim \epsilon^s + \gamma. \tag{1.90}$$

As already pointed out, it is not possible to measure γ for an individual galaxy, as the true ellipticity of the galaxy is not known and completely dominates the shear signal. Even assuming that all the galaxies were perfect disks, their random orientation in the sky would make their ellipticities appear equally randomly oriented. Although the average of the intrinsic ellipticity over a large number of galaxies tends to zero, being a random field, the ellipticity distribution has a finite width. The typical value of shear is around $\gamma \sim 0.01$, which is much smaller than the dispersion of the intrinsic ellipticity $\sigma_\epsilon \sim 0.2$. This is called *shape noise*.

If the mean of the intrinsic ellipticities vanishes, meaning that the galaxies are randomly aligned, then we could approximate $\langle \epsilon \rangle \sim \langle \gamma \rangle$, but this approximation has a flaw: galaxies do not follow a perfect random orientation, they tend to have correlated orientations caused by true gravitational attraction.

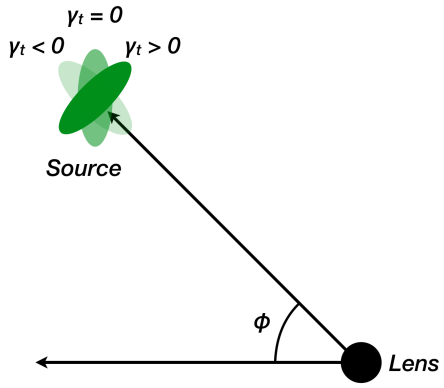

Figure 1.9: Illustration of how the value of the tangential component of the shear of a source galaxy with respect to a lens changes, projected on the sky plane. If $\gamma_t > 0$, it is induced by an overdensity.

This systematic uncertainty called *intrinsic alignment* can pollute the shear measurement signal causing an excess of correlation, and it needs to be estimated and subtracted from the true lensing signal.

Tangential and cross component of the shear field

The two components of the shear $\gamma_{1,2}$ are expressed with respect to a Cartesian frame of reference. However, the sheared image of an object by a spherical distribution of matter is oriented tangentially to the direction towards the center of the sphere. It is therefore more convenient to define another set of shear components, more adequate to describe the distorted images.

For a pair of source-lens galaxies we can then define the *tangential* and *cross components* of the shear field:

$$\gamma_t = -Re\left[\gamma e^{-2i\phi}\right], \tag{1.91}$$

$$\gamma_\times = -Im\left[\gamma e^{-2i\phi}\right], \tag{1.92}$$

where ϕ is the position angle of the source galaxy with respect to the x axis of the Cartesian field, as illustrated in Figure 1.9. By expressing γ_t and $\gamma_\times$ as function of γ_1 and γ_2, it follows:

$$\gamma_t = -\gamma_1 \cos(2\phi) - \gamma_2 \sin(2\phi), \tag{1.93}$$

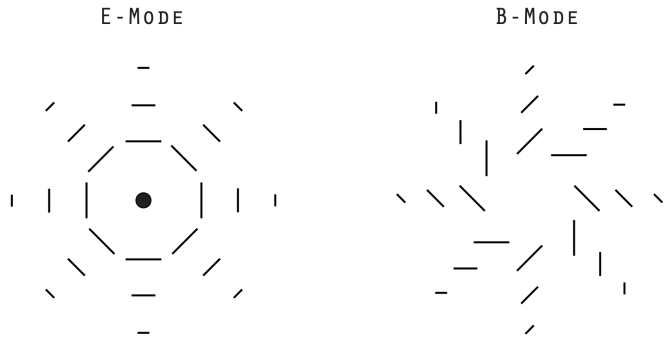

Figure 1.10: The pattern of tangential and cross shear components around a point mass. In the left panel, the images are perpendicular to the line connecting the image itself and the point mass, they are therefore tangential. The lensing signal is all carried by this component, also called E-mode. In the right panel, the cross shear component follows a curl pattern, for which it is called B-mode. For a spherical lens it is expected to be zero, therefore does not contain any cosmological information.

$$\gamma_\times = \gamma_1 \sin(2\phi) - \gamma_2 \cos(2\phi). \tag{1.94}$$

The tangential component is the one carrying all the cosmological signal from lensing if the mass distribution is spherical, while the cross component should vanish. In Figure 1.10 it is showed how the cross component behaves similarly to a curl of the shear field, sometimes referred to as B-mode, as opposed to the E-mode behaviour of the tangential component. At first order, the B-modes are not created by a gravitational lensing signal, therefore they are used as a null test in weak lensing analyses. Once again, it is important to remember that these components have to be observed statistically, averaging over a large amount of galaxies. It would be impossible to measure the shear for single galaxies as the true shape is unknown.

1.7.2 Lensing convergence

As we introduced in the previous section, the lensing convergence field represents the projection of all matter along the line of sight. By using the Poisson equation (Eq. 1.51) and the definition of convergence in Equation 1.81 we can express the relation of the convergence κ with the density contrast δ:

$$\kappa(\theta, \chi) = \frac{3H_0^2\Omega_m}{2c^2} \int_0^\chi d\chi' \frac{\chi'(\chi - \chi')}{\chi} \frac{\delta(\chi'\theta, \chi')}{a(\chi')}, \tag{1.95}$$

where the 2D integral is weighted by a lensing kernel. Assuming a redshift distribution for the sources $n_\kappa^i(z(\chi))$ for each redshift bin i leads to Eq. 1.95

becoming:

$$\kappa(\theta) = \int_0^{\chi_{\rm lim}} d\chi n_\kappa^i(z(\chi))\kappa(\theta,\chi), \tag{1.96}$$

with $\chi_{\rm lim}$ is the limiting comoving distance of the galaxy source sample.

1.8 Angular two-point correlation functions

The power spectrum defined in Equation 1.57 is only useful if an accurate set of redshifts is available for each galaxy sample. For photometric surveys as the Dark Energy Survey, galaxy redshifts are not known with the required accuracy to measure the 3D correlation function. The 2D projected angular correlation function is usually used, since it does not require knowledge of galaxy redshifts to be computed, but only angular positions. The 2D angular correlation function is equivalent to the 3D correlation function integrated over a certain redshift range. It depends only on the absolute value of the separation angle θ between two galaxies in the sky plane, and not on their orientation, because of isotropy. The Fourier transform of the angular correlation function is the angular power spectrum.

Either the angular power spectrum or the two-point angular correlation functions are used in the comparison between data and theoretical predictions, which is the core of any cosmological analysis. In Section 1.6.3 we have described the two-point correlation function for the density field. This is generally referred to as *galaxy clustering*, which as the name suggests is a quantity measuring the amount of clustering of luminous matter. Correlating the positions of lens galaxies provides information on the cosmic expansion history. We have then introduced gravitational lensing, and the shear that can be imprinted on galaxy images by the gravitational field of foreground matter. *Galaxy-galaxy lensing* can be measured from the tangential shear of source galaxies produced by foreground galaxies as function of the projected distance between the two. Its two-point correlation function can therefore be computed from the correlation between the shear of the source galaxies and the position of lens galaxies. It is a powerful way to probe the relation between baryonic and dark matter. Finally, the two-point correlation function between the shear of source galaxies is called *cosmic shear*, which provides us with a measure of the growth of large-scale structure from the pattern of source galaxy shapes, and consequently allows us to infer the properties of dark energy.

The angular power spectrum is usually computed using the *Limber approximation*. For two fields A and B defined on a sphere as projections of the 3D

fields a and b, the angular power spectrum can be written using the window functions $q(\chi)$ as follows:

$$C_{AB}^{ij}(\ell) = \int d\chi \frac{q_A^i(\chi) q_B^j(\chi)}{\chi^2} P_{ab}\left(\frac{\ell+1/2}{\chi}, z(\chi)\right), \tag{1.97}$$

where P_{ab} is the three dimensional power spectrum relative to the two fields a and b, computed for the wave number $k = \ell/\chi$, where ℓ indicates the multipole. This approximation holds as long as the variations in the fields are much smaller than the typical scale of the window functions.

We can now write the angular power spectra for cosmic shear, galaxy-galaxy lensing and galaxy clustering:

$$C_{\kappa\kappa}^{ij}(\ell) = \int d\chi \frac{q_\kappa^i(\chi) q_\kappa^j(\chi)}{\chi^2} P\left(\frac{\ell+1/2}{\chi}, z(\chi)\right), \tag{1.98}$$

$$C_{g\kappa}^{ij}(\ell) = \int d\chi \frac{q_g^i\left(\frac{\ell+0.5}{\chi}, \chi\right) q_\kappa^j(\chi)}{\chi^2} P\left(\frac{\ell+1/2}{\chi}, z(\chi)\right), \tag{1.99}$$

$$C_{gg}^{ij}(\ell) = \int d\chi \frac{q_g^i\left(\frac{\ell+1/2}{\chi}, \chi\right) q_g^j\left(\frac{\ell+1/2}{\chi}, \chi\right)}{\chi^2} P\left(\frac{\ell+1/2}{\chi}, z(\chi)\right), \tag{1.100}$$

where P is the 3D matter power spectrum. The window function for galaxy clustering is denoted by q_g, while the one for shear (often referred to as lensing kernel) is denoted by q_κ; their definitions are:

$$q_g^i(k,\chi) = b^i(k, z(\chi)) \frac{n_g^i(z(\chi))}{\bar{n_g^i}} \frac{dz}{d\chi}, \tag{1.101}$$

$$q_\kappa^i = \frac{3H_0^2\Omega_m}{2c^2} \frac{\chi}{a(\chi)} \int d\chi' \frac{n_\kappa^i(z(\chi'))dz/d\chi'}{\bar{n_\kappa^i}} \frac{\chi'-\chi}{\chi'}, \tag{1.102}$$

with $n_{g/\kappa}^i$ is the redshift distribution of the galaxy sample i, $\bar{n}_{g/\kappa}^i$ the density of galaxies and b^i the linear galaxy-matter bias.

The expected angular correlation functions for galaxy clustering $w(\theta)$, galaxy-galaxy lensing γ_t and cosmic shear $\xi_\pm$ can be computed from the angular power spectra as follows:

$$w(\theta) = \sum_\ell \frac{2\ell+1}{4\pi} P_\ell(\cos\theta)) C_{gg}^{ij}(\ell), \tag{1.103}$$

$$\gamma_t^{ij} = \sum_\ell \frac{2\ell+1}{4\pi\ell(\ell+1)} P_\ell^2(\cos\theta)) C_{g\kappa}^{ij}(\ell), \tag{1.104}$$

$$\xi_{\pm}^{ij} = \sum_{\ell} \frac{2\ell+1}{4\pi\ell^2(\ell+1)^2} [G_{\ell,2}^{+}(\cos\theta) \pm G_{\ell,2}^{-}(\cos\theta)] \times [C_{\kappa\kappa}^{ij}(\ell)], \tag{1.105}$$

with $P_\ell(x)$ the Legendre polynomial of order ℓ, and $G_{+/-}$ the associated Legendre polynomials, given by Equation (4.19) of Stebbins (1996). We note that in the Dark Energy Survey Year 3 analysis, for the galaxy clustering angular correlation function the Limber approximation was not assumed, and we refer to Fang et al. (2020) for a detailed derivation.

The use of several probes is an extremely powerful tool in cosmology, as each measurement is stronger at constraining a particular set of cosmological parameters, and the combination improves the final accuracy of the analysis. The analysis combining galaxy-clustering and galaxy-galaxy lensing is called *2x2pt analysis*, because of the information obtained from two two-point correlation functions. The *3x2pt analysis* also includes cosmic shear, benefiting from the measurement of all three two-point correlation functions, and it has now become the benchmark in cosmological analyses with photometric imaging surveys.

Chapter 2

The Dark Energy Survey

2.1 Project overview

The Dark Energy Survey (DES, Flaugher et al. 2015) is an international collaboration aimed at studying and testing the nature and the properties of dark energy. It aggregates the effort of scientists from 25 institutions in the United States, Spain, the United Kingdom, Brazil, Germany, Switzerland and Australia (see Fig. 2.1). The core of the experiment is the analysis of the images of galaxies taken with the 570-Megapixel digital camera DECam (Flaugher et al. 2015). The camera was mounted at the prime focus of the 4-meter Blanco Telescope, at the Cerro Tololo Inter-American Observatory in Chile and observed for 6 years until the beginning of 2019.

The DES scientific goal is to investigate the cause of the accelerating universe by measuring the histories of cosmic expansion and of structure formation with high precision. According to the Dark Energy Task Force (DETF, Albrecht & Bernstein et al. (2006)), which devised a classification of cosmological experiments, DES is a Stage III survey, as these were defined as near-future, intermediate-scale projects. Stage III ground-based surveys with wide-aperture telescopes have dominated the field up to now, as they are of moderate cost and have managed to hit the marks required by the "precision" cosmology era. DES is the most prominent among these, and it is fulfilling its duty by balancing the intermediate imaging quality (limited by the presence of Earth's atmosphere) with a large area coverage and impressive improvements on the analysis techniques and systematics treatment. These lead to crucial advances in our experience in this kind of analyses and in our knowledge of cosmology, upon which long-term Stage IV experiments will stand.

In order to measure the effects of dark energy on the cosmic expansion and on the growth of structure, DES focuses on four complementary probes:

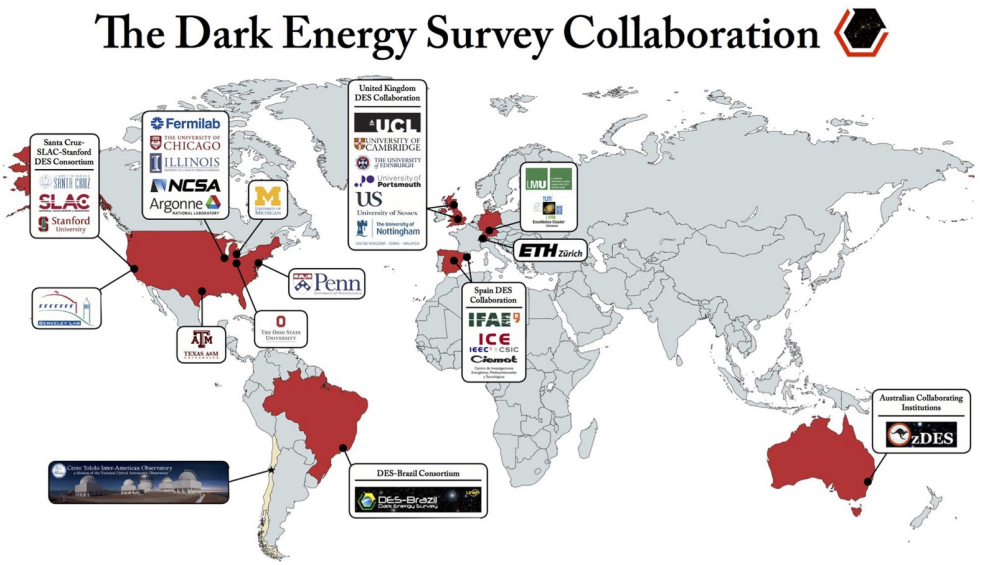

Figure 2.1: Map of the institutions involved in the Dark Energy Survey. Credit: Judit Prat.

- **Supernovae of type Ia** are a crucial tool in order to observe the expansion rate. As already mentioned in Chapter 1, these are used as standard candles as they explode always in similar conditions, therefore releasing similar amounts of energy. More than 3000 supernovae have been observed in the 27 deg^2 chosen fields, more than any other single survey in history. DECam has observed the same patches of the sky every 4-7 nights, therefore allowing a correct monitoring of the supernovae light curves.

- **Baryon Acoustic Oscillations (BAO)** provide a tool to measure the distribution of galaxies across the cosmos. In its early stage, the Universe was a hot plasma of ionised particles. The competition between gravity and pressure filled the fluid with sound waves. During the recombination era, the Universe became transparent to radiation, and the waves were not able to travel anymore as the fluid ceased to exist (pressure dropped to 0). The "frozen" wave fronts remained imprinted in the distribution of matter, and we can now measure their angular separation.

- **Galaxy clusters** provides us with a tool to examine the amount of matter in the Universe. By counting the number of galaxy clusters of a given mass within a given volume of the Universe as a function of redshift we

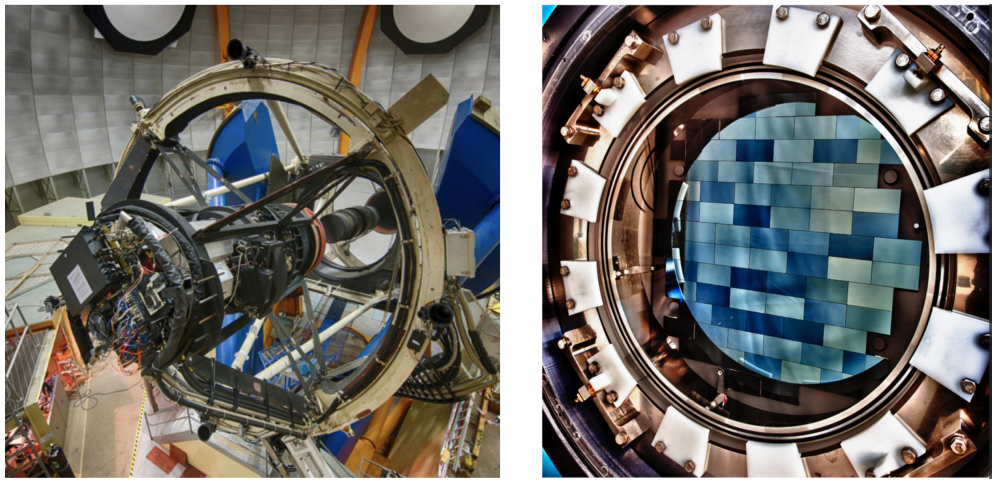

Figure 2.2: Left: the 570-Megapixel digital camera DECam mounted at the prime focus of the Blanco telescope at the Cerro Tololo Inter-American Observatory in Chile. Right: the DECam focal plane, with twelve 2k x 2k pixels CCDs used for guiding and focus, and sixty-two 2k x 4k pixels used for imaging. Credit: Reidar/Hahn/Fermilab.

are able to determine how this quantity has changed over time. Cosmic "clumpiness" is sensitive to dark energy because dark energy influences both the expansion of the universe and the growth of large scale structure.

- **Weak Gravitational Lensing (WL)**, as already described in the previous chapter, is a direct result of general relativity. As radiation travels trough spacetime, its path is perturbed by the gravitational potential of the Large Scale Structure of the Universe. By measuring the distribution of galaxy shapes we are able to indirectly probe the statistical properties of the density field.

The combination of four probes greatly improves the constraints on cosmological parameters. The work presented in this thesis relies on the use of WL observables.

2.2 Instrument

The design of DECam was driven by the need to map a vast portion of the sky while maintaining excellent image quality, high sensitivity in the near infrared and low readout noise. The consensus was reached over a camera with three square degree field of view, a five lens optical corrector and a large focal plane

with 74 fully-depleted, 250-micron thick charged-coupled devices (CCDs). DECam is conceptually similar to a single-lens reflex camera, but of course much larger and heavier, with its 570-Megapixel (see Figure 2.2. The five optical lenses and the five optical filters at the time of construction were the largest and heavier ever built for an astronomical experiment. The camera was aligned and positioned using six pneumatically driven pistols.

The instrumentation was built between 2008 and 2011, assembled at Fermilab using a full scale reproduction of the part of the telescope used to sustain the camera. The front-end electronics was developed as a joint effort of several Spanish and US insitutions: CIEMAT in Madrid, our IFAE and IEEC/CSIC in Barcelona, and Fermilab. The various components were then shipped to the Cerro Tololo Inter-American Observatory over the course of 2011 and 2012, and finally the DECam installation was completed during the fall of 2012.

DECam saw its first light on September 12th 2012. After a period of commissioning, the first scientific observations took place for the Science Verification analysis, covering $\sim 200\text{deg}^2$ at the nominal depth. The full survey began in August 2013 and ended in January 2019, and over the course of 524 nights observed more than 300 million galaxies, 3000 supernovae and 25000 clusters distributed across 1/8 of the sky.

The wide-field survey adopted 10 passes with a $90s$ exposure time for the $griz$ bands and a $45s$ exposure time for the Y band. A $\sim$27 deg^2 deep supernova survey observed in the $griz$ bands.

DES has already performed three out of the four analysis batches planned, analysing half of the total amount of data: the science verification, the analysis of the first year of observation (Year 1) of only 1300 deg^2, published in 2018, and finally the analysis of the first three years of data (Year 3) observed on the full footprint, released to the public in May 2021, ahead of several follow up papers. We are now headed towards the Year 6 analysis, which will represent the legacy of DES to cosmology. This thesis focuses on the results of the Year 3 analysis.

In the following Sections we are going to cover two of the main elements required for the cosmological analysis: the redshift and shear estimations.

2.3 Redshift Estimation

In every weak lensing analysis one of the critical ingredients is the distance to the galaxies in the samples under study. As defined in Section 1.4.1, redshift is the quantity used in astronomy to evaluate the distance to celestial objects, as the two are directly related due to cosmic expansion. As our images can-

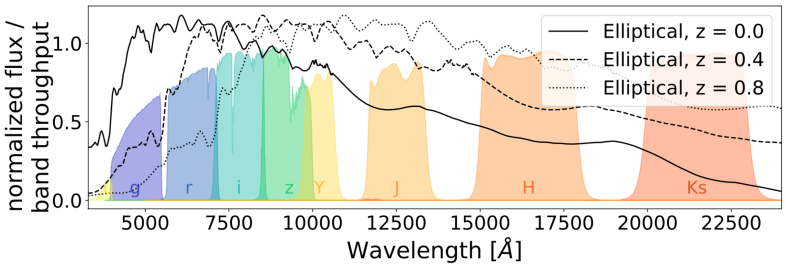

Figure 2.3: Illustration of how redshifts can be estimated from broad band images. Here templates for the same elliptical galaxy plotted at redshifts $z = 0$, $z = 0.4$, and $z = 0.8$, show clearly different colors. Credits: Buchs et al. (2019).

not be anything else than two-dimensional, redshift is crucial to enhance our measurements to three dimensions.

Three main techniques to estimate galaxy redshift will be considered here: spectroscopic, photometric, and clustering redshifts.

2.3.1 Spectroscopic Redshifts

The most accurate way to measure redshifts is to obtain the spectrum of the object, and then compare the wavelengths of its features (absorption lines, line breaks, etc.) with spectra measured in the rest frame. In this way it is possible to identify by how much the spectrum has been shifted in terms of wavelength. Galaxies have numerous clear features in their spectra, due to their gas and stars, thus by observing a galaxy for a sufficient amount of time it is fairly easy to obtain the correct redshift estimate.

The benefits of spectroscopic redshifts are balanced by the large amount of time required to obtain the spectra. It is incredibly expensive to observe a wide portion of the sky with the goal of obtaining spectroscopic measurement of all the galaxies up to a certain depth. In fact, until very recently, all of the so called *spec-z* samples were either confined to a small area, or shallow. Only recently, the DESI instrument has started the largest spectroscopic campaign to date, observing thousands of spectra per day. The first data will be made publicly available during 2023 and will revolutionise the field.

For decades up to the current day though, cosmologist faced the problem of measuring the redshifts of millions of galaxies spanned across the sky, dealt with the use of photometry.

2.3.2 Photometric Redshifts

Wide optical surveys require the redshift information for millions of galaxies across thousands of square degrees of area. It is not feasible yet (too time consuming, too expensive) to obtain a per-galaxy spectrum, therefore the flux measurement is performed using a set of a few broadband filters, which block all radiation apart from a specific wavelength range. For the wide imaging, DES uses the g, r, i, z, and Y filters, corresponding to the range of 400-1065 nm.

The flux going through a given filter i is the integral of all the input light, as follows:

$$F_i \propto \int d\lambda T_i(\lambda) F_\lambda, \tag{2.1}$$

where F_λ is the true flux of a specific object and $T(z)$ is the transmission function, which indicates how much light penetrates the filter. If $T_i(z) = 0$ the filter is opaque while if $T_i(z) = 1$ is transparent.

From the measured fluxes historically there are two broad classes of techniques to retrieve the redshift information: template fitting and training-based methods.

For the template fitting procedure, the flux measured in the broad filters is converted into a low resolution spectrum, and it is compared against a set of well known spectra for which the redshift has been already estimated. Because of the low signal-to-noise, it is much more complex to detect spectral features, so using broad-band filters it is only possible to detect "breaks" (4000 Å break, Balmer break, Lyman decrement), as they are not generally sensitive to the presence of emission lines. The comparison against the template is performed using a standard chi-square minimization procedure:

$$\chi^2(z) = \sum_i^{N_{\text{filters}}} \left[\frac{F_i^{\text{obs}} - cF_i^{\text{tem}}(z)}{\sigma_i} \right]^2, \tag{2.2}$$

where F_i^{obs} and F_i^{tem} are the observed and the template fluxes respectively, σ_i is the uncertainty in the i-th filter and c is a normalisation constant. One of the main shortcomings of this method is the lack of good quality templates at high redshift.

In the second class of methods an empirical relation between magnitudes and redshifts is derived using a training set of objects with measured spectroscopic redshifts. The advantage is that it is not necessary to study the features of the noisy spectra or to make any assumption a priori. The downside is that it

is not a flexible method, it has to be recomputed for each different set of filters and for each survey, as it is necessary that the spectroscopic training sample is suited to each survey. Furthermore, to reach the desired accuracy the training set is generally constituted by bright galaxy, which could potentially introduce a bias when the relation is applied to fainter sources. This set of techniques is increasingly becoming more robust due to the fast advancing of machine learning. We will present one of its most recent application in Chapter 5.

2.3.3 Clustering Redshifts

While for the two previously described techniques the measurement of flux is required, it is not necessary for clustering redshifts (also called WZ). The method exploits the spatial overlap of the galaxy sample with unknown redshifts (called "unknown" sample) with a sample with high quality spectroscopic or photometric redshifts used as reference. The reference sample galaxies can be grouped in redshift bins with small widths because of the high redshift accuracy. A two-point angular correlation function is then measured between the thin redshift bins and the unknown sample, using angular scales that are smaller than the ones used for cosmological measurements. Since galaxies which are closer in the projected sky plane are more likely to also have a similar redshift, measuring this correlation allows us to estimate the redshift distribution. Despite the fact that the requirement of spatial overlap limits the set of adequate samples, this technique has the advantage of not requiring the reference fluxes to be representative of the ones of the unknown sample. The main downside is that it requires knowledge of the redshift evolution of the galaxy-matter bias of the unknown sample, which is the largest systematic uncertainty of the method that needs to be taken care of. This method will be described in detail in Chapter 5.

2.4 Shear estimation

The core of any weak lensing analysis is a correct estimate of the shear field. Correctly measuring the shear of a given sample of galaxies is a very complex task, as the true shape of the galaxies is not accessible. This background contribution is referred to as *shape noise*. Figure 2.4 illustrates how the image of a galaxy is sheared, blurred, pixelised and made noisy. There are several methods that can be used to retrieve the shape measurement, involving algorithms aimed at detecting sources and removing noise due to atmospheric effects, or image simulations aimed at calibrating the measurements. Here we are going to present the pipeline for the DES Y3 analysis:

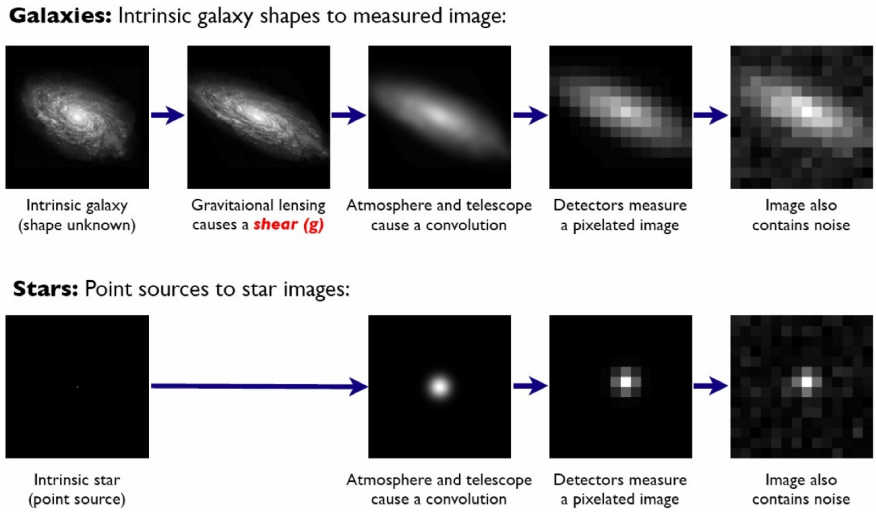

Figure 2.4: Top: different effects applied to a galaxy that needs to be disentangled for the shear measurement. This is a forward process, but experiments only obtain the image on the right, and have to eliminate all sources of noise and pollution of the signal to obtain an unbiased estimate of the shear value. Bottom: similar forward process for a point star, for which there is no gravitational shear. Credits: Bridle et al. (2009)

1. **Object detection.** DECam observed multiple times the same patch in the sky at different epochs, to average out the effect of varying observing conditions. The co-added images improve in terms of signal-to-noise. Objects are usually identified if the co-added flux in that region passes a certain threshold. An algorithm to differentiate galaxies from stars is required, to avoid polluting the sample. Furthermore, there is the issue of "blending": in the case of two or more very close galaxies, or in the case of a faint galaxy nearby a very bright one, the detection could be altered, as the properties could not be identified correctly or even objects could not be detected at all.

2. **PSF estimation.** The Point Spread Function (PSF) describes how a point-like object is measured by an imaging system. Any observed image is the convolution of the incoming light with the PSF. There are two main components to the PSF for a ground-based experiment like DES: the dominant one is caused by air, with the atmospheric turbulence (seeing) distorting the image; the detector also introduces an effect, which can be caused by the telescope's optics, by saturated pixels or by electric fields that slightly deviate charged particles' trajectories in the CCDs. The PSF calibration is performed using moderately bright stars.

3. **Shear measurement.** The galaxy shapes are measured by fitting each galaxy to a set of measured galaxy shapes used as models. A maximum likelihood approach is used to find the best fitting parameters. As already explained in Section 1.7 though, it is impossible to separate the shear from the intrinsic shape of single galaxies. This is the main obstacle in weak lensing. DES developed an algorithm to estimate the response of our shear estimation pipeline called `METACALIBRATION` (Huff & Mandelbaum (2017a); Sheldon & Huff (2017a)). For small shears (assumption valid in the weak lensing regime) we can expand the noisy measurement of the shear field ϵ with respect to the true shear field γ:

$$\begin{aligned} \epsilon &= \epsilon|_{\gamma=0} + \left.\frac{\partial \epsilon}{\partial \gamma}\right|_{\gamma=0} \gamma + \dots \\ &\equiv \epsilon|_{\gamma=0} + R_\gamma \gamma + \dots \end{aligned} \tag{2.3}$$

where R_γ is the shear response matrix, defining how the ellipticities ϵ vary with respect to the true shear γ in absence of lensing $\gamma = 0$. If lensing is absent, then the average over all ellipticities must be equal to zero. We can now create an estimator for an ensemble of shear measurements using the response matrix:

$$\langle \gamma \rangle \sim \langle R_\gamma \rangle^{-1} \langle \epsilon \rangle \,. \tag{2.4}$$

In `METACALIBRATION` the response matrix is measured using finite difference derivatives. These are effectively computed by applying a small, artificial amount of shear $\gamma_\pm = 0.01$ to each image. The measurement of shapes is repeated on the sheared images, and the response for each galaxy is estimated as

$$R_{\gamma_{ij}} = \frac{\epsilon_i^+ - \epsilon_i^-}{\Delta \gamma_j}, \tag{2.5}$$

where $\epsilon_i^\pm$ are the i-th component of the ellipticities measured on images sheared by an artificial shear with j-th component equal to $\pm\gamma$.

4. **Remaining biases in the shear estimator.** Despite the `METACALIBRATION` procedure, small multiplicative and additive biases can still persist, and the two shear components i, j estimated by the algorithm can be parametrised as follows:

$$\gamma_i^{\text{est}} = m_i \gamma_i^{\text{true}} + c_i \tag{2.6}$$

where m_i and c_i are the multiplicative and additive biases respectively. Misestimations of quantities in the previous steps can lead to larger values

of both biases. Additive biases can be detected by performing null tests on data, measuring quantities that are expected to average to zero; multiplicative biases are generally identified using image simulation, where the true shear is known and there is control over the detection systematics.

Part II

DES Year 3 Photometric Redshifts

Chapter 3

Introduction

The Dark Energy Survey (DES, Flaugher et al. 2015, described in Chapter 2) is currently the largest photometric galaxy survey to date, having detected hundreds millions of galaxies over 5000 $\deg^2$ of the southern hemisphere. Together with other ongoing and future galaxy surveys (e.g., Kilo-Degree Survey KIDS, Kuijken et al. 2015; Hyper Suprime-Cam HSC, Aihara et al. 2018; Vera Rubin Observatory Large Survey of Space and Time (LSST), LSST Science Collaboration et al. 2009; Euclid, Laureijs et al. 2011), DES can achieve competitive constraints on cosmological parameters, by studying both the spatial distribution of the detected galaxies and by measuring the tiny distortions in their observed shapes due to gravitational lensing effects induced by the large scale structure of the Universe. The analysis of the first three years (Y3) of DES data (Abbott et al., 2022) placed tight constraints on cosmological parameters, by combining three different measurements of the three two-point (3x2pt) correlation functions that measured galaxy positions and galaxy shapes. The three functions are namely:

1. Cosmic shear, i.e. the 2-point correlation function of galaxy shapes; the DES Y3 measurements (Amon & Gruen et al., 2022; Secco & Samuroff et al., 2022) involve the angular correlation of 100,000,000 galaxy shapes from the weak lensing sample (Gatti & Sheldon et al., 2020b), divided into four tomographic bins. We refer to this as the "source" sample.

2. Galaxy clustering: the 2-point correlation function of the positions of bright galaxies (which we refer to as the "lens" sample) (Rodríguez-Monroy et al., 2022);

3. Galaxy-galaxy lensing: the cross-correlation function of galaxy shapes and the position of the galaxies of the lens sample (Prat et al., 2022).

The modelling of each of these correlation functions requires knowledge of the redshift distributions (from hereafter $n(z)$) of the two samples (lens and source galaxies), which have to be estimated with great accuracy in order to avoid biased cosmological results (Huterer et al., 2006; Cunha et al., 2012; Benjamin et al., 2013; Huterer et al., 2013; Bonnett et al., 2016a; Hildebrandt et al., 2017; Hoyle et al., 2018a; Joudaki et al., 2019; Hildebrandt et al., 2020; Tessore & Harrison, 2020). The optimal solution would be to avail ourselves of spectroscopic observations, providing an accurate redshift measurement of each targeted galaxy. Unfortunately, it is not feasible to obtain said spectra other than for a small fraction of the science sample, due to the required time and cost of the observing campaign. Cosmological surveys like DES therefore have to use for their redshift estimation measurements only a few, noisy, broad-band fluxes, causing cosmologists to apply a great deal of effort and inventiveness to create robust and unbiased redshift calibration pipelines.

Photometric surveys have been relying on different methodologies to derive redshift distributions (Hildebrandt et al., 2010; Sánchez et al., 2014), mostly based on galaxies' multi-band photometry (photo-z methods, or PZ). However, these methods are ultimately limited by the redshift ambiguities in a few band colors, and the limited and incomplete spectroscopic samples available to calibrate the color-redshift relations. This issue can be circumvented by reweighting the redshift distributions of a sample according to the relative abundance of galaxies whose redshifts are known with accuracy (e.g. Lima et al. 2008, Speagle & Eisenstein 2017a, Speagle & Eisenstein 2017b, Hoyle & Gruen et al. 2018a, Tanaka et al. 2018, Wright et al. 2019, Euclid Collaboration et al. 2020, Schmidt et al. 2020). Residual uncertainty coming from the color-redshift degeneracy in this case translates as uncertainty on the measured redshift distribution, often quantified in terms of uncertainties on the moments of the measured $n(z)$.

A substantial improvement on the accuracy of redshift calibration comes from the use of broad-band photometry that covers the full range of optical and near-infrared wavelengths (Masters et al. 2017, Hildebrandt et al. 2017). Unfortunately a complete multi-band photometry is not available over the large areas observed by optical surveys. Nevertheless, optical surveys can choose to observe regions of the sky that overlap with existing spectroscopic and/or near-infrared fields, complementing the optical photometry. Typically, these regions are observed more often to reach a much higher signal-to-noise; these are called "deep fields".

For DES Y3, a new method leveraging the richness of information in the deep fields was developed, using the additional filters and the accurate photom-

etry to break the color/redshift degeneracy, thus overcoming the key limitations of few-bands redshift calibration. The redshift information in the deep fields is validated with the use of spectroscopic or high-quality photometric information (also called photo-z) made available by other surveys. The photometry of all the galaxies "living" in the deep fields is sufficiently good in order to group them in an exhaustive set of phenotypes (Sánchez & Bernstein, 2019a). By also classifying the wide galaxies with noisy photometry, we can build a mapping between the accurate and noisy phenotypes, which allows to transfer the redshift information to the wide field galaxies. This method is named SOMPZ, as the classification is performed by means of Self-Organising Maps (Buchs et al., 2019).

Clustering redshifts methods (Newman 2008; Ménard et al.; Choi et al. 2016; Davis et al. 2017; Morrison et al. 2017; Johnson et al. 2017; Gatti & Vielzeuf et al. 2018; van den Busch et al. 2020) offer an alternative to standard photo-z methods to infer redshift distributions. In short, clustering-z methods exploit the two-point correlation signal between a photometric "unknown" sample and a "reference" sample of high-fidelity redshift galaxies divided into thin bins, to infer the redshift distributions of the photometric sample. One of the biggest advantages of clustering-z methods is that the reference sample does not have to be representative of the photometric sample. Clustering-z methods (or WZ) have been in the past years successfully applied to both data (Rahman et al., 2015, 2016a,b; Scottez et al., 2016; Hildebrandt et al., 2017; Johnson et al., 2017; Davis et al., 2017, 2018; Cawthon et al., 2018; Bates et al., 2019; van den Busch et al., 2020; Hildebrandt et al., 2020) and simulations (Schmidt et al. 2013; McQuinn & White 2013; Scottez et al.; Gatti & Vielzeuf et al. 2018), and they represent one credible supplement to standard photo-z methods for the new, upcoming generation of data sets.

Clustering-z methods have been used both to provide an independent redshift distribution estimate and to calibrate distributions inferred from photo-z methods. In the DES Y1 cosmological analysis we opted for the latter approach (Davis et al. 2017; Hoyle & Gruen et al. 2018b).

Another technique used to calibrate redshift distributions is based on "shear ratios" (Jain & Taylor, 2003; Heymans et al., 2012; Hildebrandt et al., 2020; Prat et al., 2019; Sánchez et al., 2022a), which are ratios of galaxy-galaxy lensing measurements. In particular, the ratio is computed between galaxy-galaxy lensing signals using the same lens galaxy redshift bin but different source bins; this probes the lensing efficiencies of the source galaxies, which are directly related to their redshift distributions. Shear ratios mostly constrain the

redshifts of the sources, but also partially improve the calibration of the lens redshift distributions. Shear ratios constraints are independent from the previous photo-z and WZ methods. The methodology adopted in the DES Y3 analysis is described fully in Sánchez & Prat et al. (2022a).

The redshift calibration strategy for DES Y3 analysis improves in multiple respects on the Y1 strategy. Firs of all, all the three aforementioned methods are used, whereas in DES Y1 we did not incorporate shear ratios in our analysis, as it was used only as an independent validation. Concerning the photo-z estimates, the implementation of the SOMPZ method allowed us to minimise the systematic uncertainties related to the type/redshift degeneracy at given broad-band flux, exploiting small samples of galaxies in the deep fields for which was possible to obtain spectroscopic redshifts. From the clustering-redshift side, we improved over Y1 in the modelling of the clustering signal, accounting for the redshift evolution of the galaxy-matter bias of the reference sample and the clustering of the underlying dark matter density field, which were neglected in the DES Y1 analysis. Moreover, we also marginalised over magnification effects, and we implemented a new method to combine clustering information with redshift distributions from photometry.

The SOMPZ method, the improved instance of the WZ method, and shear ratios have been developed for the calibration of the DES Y3 weak lensing source sample redshift distributions (Gatti & Giannini et al. 2020a, Myles & Alarcon et al. 2020). The same approach has later been applied as an additional calibration to the fiducial DES Y3 lens sample MagLim **cite my paper 7**, with some tweaks due to the different nature of the two samples. We note that the fiducial DES Y3 MagLim $n(z)$ calibration has been performed combining a training-based photometric code (DNF, De Vicente et al. 2016a) and an "Y1-like" instance of the WZ method; the methodology presented here, however, is more robust in terms of systematic uncertainty estimation and control of their propagation throughout the whole pipeline. A thorough comparison between the two calibrations for the lens sample is presented in Chapter 7.

In this part we will focus only on the application of the newly developed clustering redshift method to the DES Y3 weak lensing galaxies, while for the MagLim lens sample we will present the calibration for both methods. In the two cases the combination which results in the final $n(z)$ will be shown.

Chapter 4

Data

In this work we use data from the first three years (out of six) of observations (Y3) of DES, which were taken from August 2013 to February 2016. The DES Data Management (DESDM) team was in charge of processing the raw images (Sevilla et al., 2011; Morganson et al., 2018; Abbott et al., 2018); full details are provided in Sevilla-Noarbe et al. (2020) and Gatti & Sheldon et al. (2020b). The main catalog upon all the DES samples are built is the DES gold catalog, obtained using observations in the *griz* bands. Objects belonging to the gold catalog have passed a number of selections aimed at removing objects in problematic regions of the sky or anomalous detections (e.g., objects with pixels affected by saturation or truncation issues). The gold catalog consists of 388 millions objects (Sevilla-Noarbe et al., 2020). Each object comes with morphological and photometric measurements based on two different pipelines, the Multi-Object Fitting pipeline (MOF) and the Single-Object Fitting pipeline (SOF). The former performs a simultaneous multi-object, multi-epoch, multi-band fit to estimate morphology and photometric information; the latter does not perform the multi-object fit when it comes to crowded objects. The DES Y3 SOF implementation is faster and less prone to fit failures compared to the MOF pipeline, and it does not suffer from any significant loss in terms of accuracy (Sevilla-Noarbe et al., 2020).

We describe in this section the data and simulated products used in this work. We are going to focus on the data used in the clustering-z measurement for the weak lensing sample, and on the data used in both the clustering-z and SOMPZ measurements for the MagLim galaxies. All of the samples in data have also been reproduced in simulation for testing purposes. The samples used are:

- the DES Y3 weak lensing sample (Gatti & Sheldon et al., 2020b), used

as source galaxies in the DES Y3 3x2pt cosmological analysis and which redshift distributions need to be characterised;

- the DES Y3 MagLim sample, used as lens galaxies in the DES Y3 3x2pt cosmological analysis and which redshift distributions need to be characterised;
- the RedMaGiC galaxy catalogue (Porredon et al., 2021a), consisting of luminous red galaxies with excellent photometric redshifts and spanning in the whole DES footprint; it is used as a reference sample for the WZ measurements for the weak lensing source sample;
- BOSS/eBOSS spectroscopic galaxy catalogues, used as reference sample for the WZ measurements for both the lensing source sample and the MagLim sample; these are galaxies with spectroscopic redshift partially overlapping with the DES wide field footprint.
- the DES deep field samples, which are observed in small fields by DES with deeper observations than wide field ones and where information from additional photometric bands are available. Deep fields are a key element of the SOMPZ methodology;
- the DES Balrog sample; this sample consists of software-injected deep field galaxies into DES wide field images and is a key element of the SOMPZ methodology;
- the "redshift" samples, which are a collection of either spectroscopic or multi-band photometric samples collected by other surveys in the DES deep field region. The redshift samples are a key element in the SOMPZ methodology;

4.1 Weak Lensing sample

The DES Y3 weak lensing (WL) sample is created using the metacalibration pipeline (described and tested in Huff & Mandelbaum 2017b and Sheldon & Huff 2017b and applied to the Y3 data in Gatti et al. 2020b) and it is a subset of the gold catalog. The metacalibration pipeline provides a per-galaxy self-calibrated shape measurement, which is free from shear and selection biases. An additional, small calibration based on image simulations (MacCrann et al., 2020) accounts for blending and detection biases. The final catalog consists of ~ 100 million galaxies, spanning the full DES Y3 wide field footprint and with an effective number density of $n_{\rm eff} = 5.59$ gal/arcmin^{-2}. The WL sample is

divided into four tomographic bin through the following procedure, which uses Self-Organising Maps (Section 5.1):

- A first set of edge values are arbitrarily selected.
- Each galaxy of the redshift sample is then assigned to the tomographic bin in which its redshift estimate falls.
- A number of galaxies at this point share the same photometry cell of the wide-field SOM and same tomographic bin, so the cell in its entirety is assigned to the bin to which the majority of its galaxies live.
- The initial bin edges are adjusted to yield approximately the same number of galaxies.
- Finally the whole procedure is repeated with the new bin edges.

After completing this procedure, the final bin edges are [0.0, 0.358, 0.631, 0.872, 2.0] for the Y3 weak lensing source catalogue.

4.2 MagLim sample

The lens galaxy sample considered in this work is the MagLim sample. The MagLim sample is a subset of the DES gold catalog and consists of bright galaxies selected with an *ad-hoc* selection that optimises the number density and the redshift accuracy of the sample (Porredon et al., 2021b). The MagLim sample spans the full DES Y3 wide field footprint, for a total of ~ 4143 deg^2. SOF magnitudes in the riz bands[1] are used for the selection and photometry. The selection is meant to be linear in redshift and magnitude, and reads

$$\begin{aligned} m_i &< 4 * z_{\rm mean} + 18 \\ m_i &> 17.5, \end{aligned} \tag{4.1}$$

where m_i the i-band SOF magnitude and $z_{\rm mean}$ is a per-object redshift estimate from the photo-z code Directional Neighborhood fitting (DNF) De Vicente et al. 2016a; see also next subsection). The sample is then further limited to the redshift range $0.2 < z_{\rm mean} < 1.05$. This leads to a sample that ranges from $18.8 < i_{\rm mag} < 22.2$ The MagLim sample is divided into 6 tomographic bins using DNF $z_{\rm mean}$ and considering the following bin edges:

[1]We exclude the g-band as its photometry is known to be affected by PSF estimation issues (Jarvis et al., 2020).

Bin	z range	N galaxies	n density	C_{flux}
1	$0.20 < z_{\text{ph}} < 0.40$	2 236 473	0.150	0.43
2	$0.40 < z_{\text{ph}} < 0.55$	1 599 500	0.107	0.30
3	$0.55 < z_{\text{ph}} < 0.70$	1 627 413	0.109	1.75
4	$0.70 < z_{\text{ph}} < 0.85$	2 175 184	0.146	1.94
5	$0.85 < z_{\text{ph}} < 0.95$	1 583 686	0.106	1.56
6	$0.95 < z_{\text{ph}} < 1.05$	1 494 250	0.100	2.96

Table 4.1: Summary of the MagLim sample. We have outlined for each tomographic bin the redshift range (selected using DNF z_{mean}), the number of galaxies, the number density, and the magnification coefficient as measured in Elvin-Poole et al. (2021)

$[0.2, 0.4, 0.55, 0.7, 0.85, 0.95, 1.05]$, with a total of a 10,716,506 galaxies, distributed across bins as summarised in Table 4.1. The MagLim sample is used as lens sample in the galaxy-galaxy lensing and galaxy clustering measurements of the DES Y3 2x2 cosmological analysis (Porredon et al., 2021a).

4.2.1 DNF

The photo-z code DNF (Directional Neighborhood fitting is used to define the MagLim selection and to define the MagLim tomographic bins. The DNF algorithm computes a point estimate z_{mean} of redshift of the galaxies by performing a fit to a hyper-plane in color and magnitude space using up to 80 nearest neighbors taken from a reference sample made of spectroscopic galaxies with secure redshift information. For this purpose, a large number of spectroscopic catalogs collected by Gschwend et al. (2018) has been used, including spectra from SDSS DR4 (Abolfathi et al., 2018), OzDES (Lidman et al., 2020), VIPERS (Garilli et al., 2014), and from the PAU spectro-photometric catalog (Eriksen et al., 2019). The total number of spectra used for training is $\sim 10^5$. DNF also provides a redshift estimate z_{DNF} drawn from the redshift PDF for each individual galaxy, although only the quantity z_{mean} is used for the MagLim selection and binning.

4.3 WZ reference sample: RedMaGiC galaxies

One of the reference samples used in the clustering redshifts analysis of the weak lensing sample in DES Y3 consists of RedMaGiC galaxies. The same sample is also used in the DES analysis as lens sample for the galaxy clustering and galaxy-galaxy lensing measurements (Prat et al., 2022; Rodríguez-Monroy et al., 2022). The RedMaGiC algorithm selects red luminous galaxies for which

high-quality photometric estimates exist, through a template fitting scheme. Two samples are defined in order to present different properties: the "high density" sample is selected to have a luminosity larger than $L/L* > 0.5$, which allows to yield a larger number of galaxies, while the "high luminosity" is built to select only the brightest galaxies, with $L/L* > 1$. The combined sample is obtained by simply merging the two samples, using the high density sample for redshift $z < 0.65$, and the high luminosity one otherwise. The sample is binned in 40 bins of $\Delta z = 0.02$, over a $0.014 < z < 0.94$ interval. The choice of the bin width is not expected to impact our measurement, as long as bins are small enough compared to the typical variation scales of the the weak lensing $n(z)$ and the galaxy-matter biases of the two samples. The total number of RedMaGiC galaxies is 3,041,935 in the data.

In order to remove correlations with observational systematics, weights are applied to the RedMaGiC galaxies. In particular, the weights correct any spurious dependence of the number density with respect to any systematic accounted for, to the level needed for two-point correlation functions. Note that due to low-statistics issues, the weights do not resolve fluctuations on scales relevant for this work, but only capture large-scale spurious correlations. The procedure used to compute the weights is described in Rodríguez-Monroy et al. (2022).

4.4 WZ reference sample: BOSS/eBOSS Galaxies

The BOSS/eBOSS galaxy catalog is our main reference for the clustering-z measurement for the case of the lens sample; in the case of the weak lensing source sample, it joins RedMaGiC as second reference sample. It consists of a number of spectroscopic samples from the Sloan Digital Sky Survey (SDSS, Gunn et al. 2006; Eisenstein et al. 2011; Blanton et al. 2017), and combines SDSS galaxies from BOSS (Baryonic Oscillation Spectroscopic Survey, Smee et al. 2013; Dawson et al. 2013) and from eBOSS (extended-Baryon Oscillation Spectroscopic Survey Dawson et al. 2016; Ahumada et al. 2019; Alam et al. 2020). In particular, the BOSS sample includes the LOWZ and CMASS catalogs from the SDSS DR 12 (Reid et al., 2016), while we included the large-scale structure catalogs from emission line galaxies (ELGs Raichoor et al. 2017), luminous red galaxies (LRGs, Prakash et al. 2016) and quasi stellar objects (QSOs) (eBOSS in prep.) from eBOSS. Following Gatti et al. (2020a); Cawthon et al. (2020), we stack together the different samples and use them as a single reference sample. We also create a single random catalog by stacking all the random catalogs of each individual samples. The BOSS/eBOSS sample is divided into 50 bins spanning the $0.1 < z < 1.1$ range of the catalog (width

Spectroscopic Samples			
Name	Redshifts	$N_{\rm gal}$	Area
LOWZ (BOSS)	$z \sim [0.0, 0.5]$	45671	~ 860 deg^2
CMASS (BOSS)	$z \sim [0.35, 0.8]$	74186	~ 860 deg^2
LRG (eBOSS)	$z \in [0.6, 1.0]$	24404	~ 700 deg^2
ELG (eBOSS)	$z \in [0.6, 1.1]$	89967	~ 620 deg^2
QSO (eBOSS)	$z \in [0.8, 1.1]$	7759	~ 700 deg^2

Table 4.2: List of the spectroscopic samples from BOSS/eBOSS overlapping with the DES Y3 footprint used as reference galaxies for clustering redshifts in this work.

$\Delta z \sim 0.02$). The number of galaxies for each sample are listed in Table 4.2, with the final sample consisting of 241,987 objects and covering an area ranging from 14 to 17% of the total DES footprint.

We note that estimates of the magnification coefficients are not available for BOSS/eBOSS galaxies. For our fiducial analysis we assumed magnification values for the BOSS/eBOSS sample to be set to zero. We are confident about this choice for the narrow shape of the MagLim tomographic bins, since the magnification is usually significant in the tails of the distribution when the clustering kernel due to selection effects is larger. We nonetheless verify in this work that our analysis is not very sensitive to the particular choice of the values of the magnification parameters. The area coverage is smaller compared to RedMaGiC galaxies, as shown in Fig. 4.1. Note that some of the galaxies in the BOSS/eBOSS sample are also in the RedMaGiC catalog: $\sim$1 % of the RedMaGiC galaxies are matched to $\sim$10 % of the BOSS/eBOSS galaxies, within 1 arcsec. We did not remove these galaxies from the RedMaGiC sample, as they have a negligible impact both on our constraints and on the covariance between the two samples (as it will be clear in the following sections, the constraints from both samples are systematic-dominated).

4.5 Deep Fields sample

The Deep fields catalog is a key element of the SOMPZ methodology. We provide here a few key details, but we refer the reader to Hartley et al. (2020) for extensive details and the characterisation of the sample.

This work uses four different deep fields, i.e., E2, X3, C3 and COSMOS (COS) covering 3.32, 3.29, 1.94, and 1.38 deg^2, respectively. These are illustrated in Figure 4.1. Each deep field has undergone a scrupulous masking procedure aimed at removing artefacts (e.g., cosmic rays, meteors, saturated

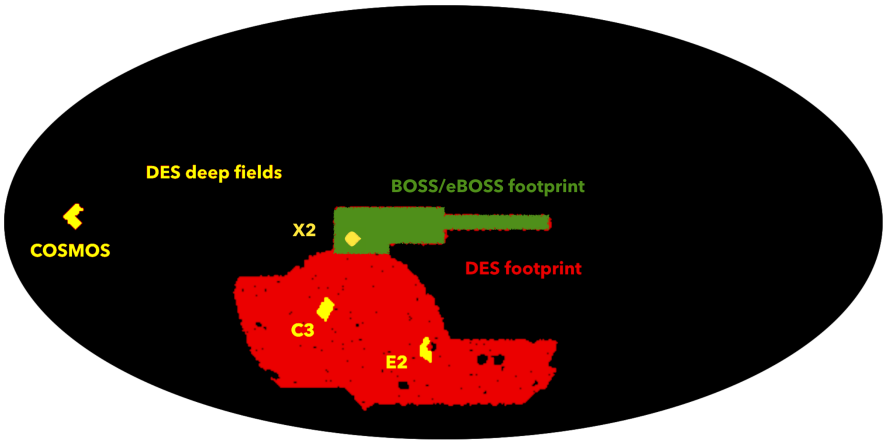

Figure 4.1: Mollweide visualisation of the Dark Energy Survey footprint (in red) and the four deep fields used in the Year 3 analysis: E2, X3, C3 and COSMOS. In green the overlap with the BOSS/eBOSS sample.

pixels, etc.). Considering the final unmasked area overlapping with the UltraVISTA and VIDEO near-infrared (NIR) surveys (McCracken et al., 2012; Jarvis et al., 2013), which is needed to provide photometric information in additional bands, we are left with 5.2 deg^2 of area for a total of 267,229 galaxies with measured u, g, r, i, z, J, H, K_s photometry with limiting magnitudes 24.64, 25.57, 25.28, 24.66, 24.06, 24.02, 23.69, and 23.58. Note that deep field galaxies have deeper photometry and photometry available in more bands compared to the wide field galaxies; this is key for a good performance of the SOMPZ method as it reduces the color-redshift degeneracy.

4.6 Balrog sample

The Balrog sample is another key element of the SOMPZ methodology. It is used to relate galaxies with given deep photometry to observed galaxies with wide field photometry, which are noisier. To this aim we rely on Balrog (Suchyta et al., 2016), a software which injects "fake" galaxies into real images. For this analysis, Balrog was used to inject deep field galaxies into the broader wide field footprint (Everett et al., 2020b). After injecting galaxies into images, the output Balrog images are passed into the DES Y3 photometric pipeline and injected galaxies are detected equivalently to real galaxies, yielding multiple realisations of each injected galaxy. The Balrog sample spans ∼20% of the DES Y3 footprint. We further select injected galaxies using

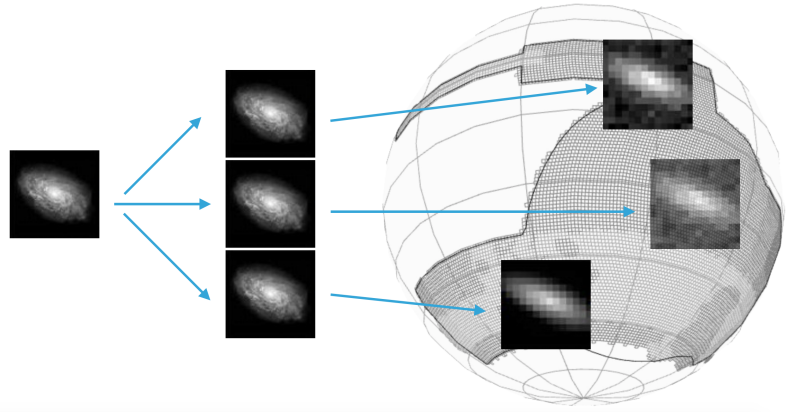

Figure 4.2: Scheme illustrating the operation of BALROG: the practically noiseless deep fields galaxies are injected many times in DES real wide field images; those dichotomous images are then processed through the fiducial DES detection pipeline, to construct a sample containing several noisy representations of the same deep galaxies.

the sample selection. We then construct a matched catalog matching BALROG injected wide field galaxies with their deep field counterparts. The resulting catalog is called the BALROG sample. For MAGLIM, it consists of a total of 351,165 galaxies with both deep and wide photometric information.

4.7 Redshift Samples

The redshift samples used for the SOMPZ section of the analysis consist of galaxies with secure redshift information (either spectroscopic or high quality multi-band photometric) observed in the deep fields. These samples are key to characterise the redshifts of the deep field sample and, in turn, to transfer the redshift information to the wide field sample.

We consider three separate redshift selections, similarly to what has been used in source sample redshift characterisation (Myles & Alarcon et al., 2020):

- a collection of spectra from a number of different public and private spectroscopic samples, from the spectroscopic compilation by Gschwend et al. (2018). We have not restricted ourselves to a few, selected surveys as in the case of the DES Y3 weak lensing sample (Myles & Alarcon et al., 2020), where only zCOSMOS (Lilly et al., 2009a), C3R2 (Masters et al., 2017, 2019), VVDS (Le Fèvre et al., 2013a), and VIPERS (Scodeggio et al., 2018a) were considered. This is because due to the bright nature of the MAGLIM sample we would mostly select high signal-to-noise galaxies,

and using more spectroscopic samples allow us to largely improve the statistic, minimising the impact of possible outliers;

- multi-band photo-z galaxies from the COSMOS field; in particular, we used the `COSMOS2015` 30-band photometric redshift catalog (Laigle et al., 2016), which is equipped with narrow, intermediate and broad bands covering the IR, optical, and UV regions of the electromagnetic spectrum;

- redshifts from the `PAUS+COSMOS` 66-band photometric redshift catalog (Alarcon et al., 2020), which adds 40 narrow band filters from the PAU Survey.

We match these redshift catalogs to our deep field galaxies and keep only those that are selected at least once into our sample selection according to their BALROG injections. Due to the bright nature of the MAGLIM sample, the number of galaxies in our final redshift samples is greatly reduced: for the SPC sample, for example, the unique total number of galaxies passes from 118745 to 17718, a reduction of around 85%.

In some cases, the same galaxy might have redshift information from multiple surveys. Following Myles & Alarcon et al. (2020), we created three slightly different redshift samples, where in case of multiple information from different surveys we use as fiducial the redshift from a specific survey. The samples are:

- 1) *SPC*, where in case of multiple information available we first use the spectroscopic catalog (S), then `PAUS+COSMOS` (P), and finally `COSMOS2015` (C);

- 2) *PC*, where we rank first the `PAUS+COSMOS` catalog before `COSMOS2015`, and we do not include spectroscopic redshifts;

- 3) *SC*: where we first use the spectroscopic catalog before `COSMOS2015`, but we do not include the `PAUS+COSMOS` catalog.

Table 4.3 summarises the number of unique galaxies appearing in each of the three redshift samples, before and after performing the MAGLIM sample selection. The fiducial ensemble of redshift distributions is generated by marginalizing over all three of these redshift samples (SPC, PC, SC) with equal prior, which in practice is achieved by simply concatenating the $n(z)$ samples produced from these three redshift samples. In such a way we marginalise over potential uncertainties and biases in the different redshift catalogs (S, P and C).

	Raw			After MagLim selection		
	SPC	SC	PC	SPC	SC	PC
COSMOS	64139	82856	69686	3299	7231	3721
PAU	18780	-	28780	3950	-	7015
Spec-z	35826	35826	-	10429	10429	-
Total	118745	118682	98466	17678	17660	10736

Table 4.3: Number of unique galaxies belonging to each of the three redshift catalogs (COSMOS, PAU and the spectroscopic collection) for each of the samples SPC (composed by galaxies from Spec-z, PAU, COSMOS in this order), SC (Spec-z, COSMOS), PC (PAU, COSMOS). The sample selection for the MagLim sample applied to the corresponding Balrog injections reduces greatly the size of all samples. For more information, see Section 4.7.

4.8 Simulated Galaxy catalogs

Our methodology is thoroughly validated using simulated catalogs. In particular, we use one realisation of the sets of the Buzzard N-body simulations (DeRose et al., 2022). All the catalogs we used in data have their simulated counterparts, although we adopted some reasonable simplifications, when needed. We give here a brief summary of the Buzzard simulation and the simulated catalog we had to create for this project, i.e., the simulated MagLim sample. The simulated BOSS/eBOSS catalog description is provided in Gatti & Giannini et al. (2020a), whereas the simulated WL sample is described in DeRose et al. (2022).

Buzzard is a synthetic galaxy catalog built starting from N-body lightcones produced by L-GADGET2 (Springel, 2005). Galaxies are incorporated in the dark matter lightcones using the ADDGALS algorithm (DeRose et al., 2019). Buzzard spans 10313 square degrees. The cosmological parameters chosen are $\Omega_{\rm m} = 0.286$, $\sigma_8 = 0.82$, $\Omega_b = 0.047$, $n_s = 0.96$, $h = 0.7$. The simulations are created starting from three lightcones with different resolutions and size (1050^3, 2600^3 and 4000^3 Mpc$^3h^{-3}$ boxes and 1400^3, 2048^3 and 2048^3 particles), to accommodate the need of a larger box at high redshift. Halos are identified using the public code ROCKSTAR (Behroozi et al., 2013) and they are populated with galaxies using ADDGALS (DeRose et al., 2019), which provides positions, velocities, magnitudes, SEDs and ellipticities. Galaxies are assigned their properties based on the relation between redshift, r-band absolute magnitude, and large-scale density from a subhalo abundance matching model (Conroy et al., 2006; Lehmann et al., 2017) in higher resolution N-body simulations. SEDs are assigned to galaxies by imposing the matching with the SED-luminosity-density relationship measured in the SDSS data. SEDs are K-corrected and

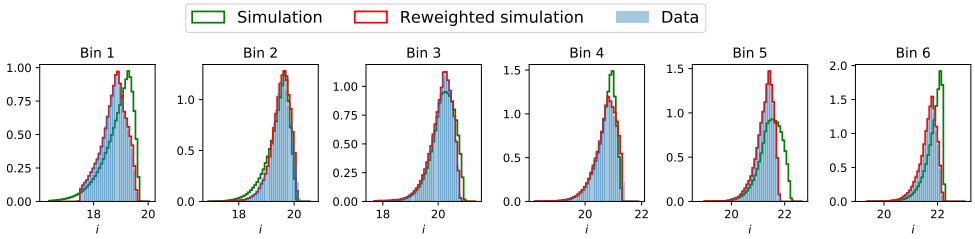

Figure 4.3: Comparison of i-band magnitudes of the six bins of the MagLim sample in data and in simulations. The performed re-weighting is successful at providing us with magnitude distributions very similar to the ones in data.

integrated over the DES filter bands to generate DES *grizY* magnitudes. Ray-tracing is performed through the CALCLENS algorithm (Becker, 2013), to introduce lensing effects, in order to provide weak-lensing shear, magnification and lensed galaxy positions for the lightcone outputs. CALCLENS is run onto the sphere, masked with the DES Y3 footprint, using the HEALPix algorithm (Górski & Hivon, 2011) and is accurate to ~ 6.4 arcseconds.

4.8.1 Simulated MagLim sample

In order to define a simulated MagLim sample, the photo-z code DNF has been run on a subset of the Buzzard simulations, restricted to *i-band* magnitudes $i < 23$, so as to reduce the running time without affecting the final result (note that the MagLim selection presents a cut at $i < 22.2$). The goal is to attain similar number density and color distributions as in data. Due to the small but existing differences in magnitude/color space between the Buzzard simulation and the DES data (DeRose et al., 2019), we expect the simulated sample to not be a perfect copy of the data sample, although we do not expect this to have a sensible impact on any of the conclusions drawn in this work.

The direct application of the fiducial MagLim selection (Eq. 4.1) to the Buzzard catalog leads to slightly different number densities and color distributions with respect to data. We therefore re-define a more adequate MagLim selection for Buzzard, with the goal of achieving the same number density as the data sample. The new Buzzard MagLim selection is a piece-wise linear selection in redshift and magnitude, similar to Eq. 4.1 but with coefficients re-defined by minimising the quadratic sum of the difference in number density with the values in data, for each tomographic bin, in order to avoid discontinuities in the selection. Such a re-defined selection guarantees similar number densities as the data sample. We then ensure similar color distributions by an additional re-weighting procedure of the mock catalog, so as to resemble the

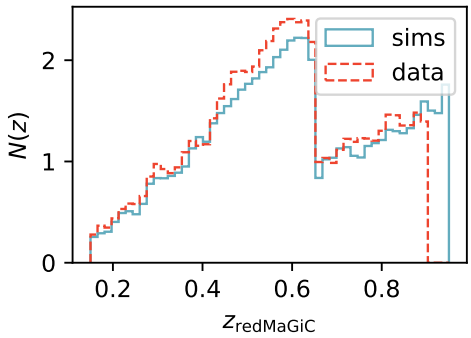

Figure 4.4: Redshift distributions of the RedMaGiC samples, binned using the RedMaGiC photo-z estimates, in data and in simulations.

color distributions of the data sample. In particular, we iteratively re-weight based on i, r magnitudes and i-r colors, with the final distribution matching closely the data one, as shown in Figure 4.3.

4.8.2 Simulated RedMaGiC catalogue

The simulated RedMaGiC sample is selected applying the same algorithm as in data to the Buzzard simulations. We computed the cross-correlations using the same bins as in data: 40 bins of width $\Delta z = 0.02$ spanning the $0.14 < z < 0.94$ range of the RedMaGiC catalog. In Fig. 4.4 are compared the two versions of RedMaGiC , in data and in simulations. Overall the agreement is good, and any discrepancies are caused by the small differences in the evolution of the red-sequence in Buzzard. The number of galaxies in simulations are 2,594,036. The different number density with respect to the data is also caused by the aforementioned discrepancy in the evolution of the red-sequence between data and simulations. This implies that the statistical uncertainties obtained with the simulated sample are slightly larger. In any case, this does not constitute an issue, as the error budget is dominated by systematic uncertainties. The clustering weights are also applied in simulations using a similar recipe, with the only difference that Buzzard only models depth variations across the footprint, while data are of course subject to a larger number of systematics. which are not fully reproduced in simulations.

4.8.3 Simulated BOSS/eBOSS catalogues

To replicate the spectroscopic BOSS/eBOSS sample in simulations, we selected bright galaxies with similar sky coverage and redshift distribution as the ones in data. We did not try to further match other properties of the sample, e.g. the galaxy-matter bias likely differs from that of the real data. We note that the clustering-z methodology corrects for the reference bias, so at no point in the analysis of the real data are we assuming that the simulations have the same bias.

4.8.4 Simulated Deep catalog

The simulated true fluxes from Buzzard are used as the deep measurements, but we further assign a realistic error by using the limiting flux for each mock deep band. We use the same uncertainties as in data, but as the Buzzard simulation has a different zero point, those values have to be converted in magnitude using zero point of 30, and then is converted to a flux uncertainty for a zero point of 22.5, which is the zero point of the Buzzard fluxes. We do not differentiate between fields, as it has been proven in Myles & Alarcon et al. (2020) that this had no impact on the simulated redshift distribution. The size of the sample is 968759 galaxies. We use the true redshift for the redshift sample and to compare our inferred redshift distributions to the true ones.

4.8.5 Simulated Balrog catalog

We mimic the Balrog algorithm by randomly selecting positions over the full Y3 footprint and run the corresponding error model on the galaxies of the simulated deep catalog to obtain noisy versions, according to the exposure times of each location. The deep galaxies can be injected an arbitrary number of times and we set this at 10. Only the wide counterparts of the deep galaxies that respect the MagLim selection defined in the Buzzard simulation are then included in the sample, yielding the final number of 250193 galaxies.

Chapter 5

DES Y3 Photo-z estimation methodology

In this section we are first going to present the Self-Organizing Maps Photo-z (SOMPZ) and the clustering-z (WZ) methods in detail, and then we are going to describe how the are combined. Such a combination is powerful because it exploits the complementarity of the two methods, which are affected by two very different sets of biases and uncertainties. The application to data will follow in the subsequent Chapter. Shear ratios are combined at the level of the cosmological inference, as an additional likelihood. As they are not the main focus of this thesis, we are not going to discuss them further. Nevertheless, our cosmological results in Chapter 8 include the constraints from shear ratios.

5.1 SOMPZ

The SOMPZ methodology estimates wide field redshift distributions by exploiting a mapping between wide field galaxies and deep field galaxies with deeper and more precise photometry. Extracting the redshift information from deep, several band photometry in order to estimate the redshift of an observed wide field galaxy amounts to marginalizing over deep photometric information (Buchs et al., 2019). The classification of galaxies by their properties is essential, and the chosen tool for this analysis is the *Self-Organizing Map*.

Self-Organizing Maps

A Self-Organizing Map (Kohonen, 1982, SOM) is a machine learning based algorithm, specifically an artificial neural network. It produces a mapping of a higher dimensional input space to a lower dimensional one. It is in fact

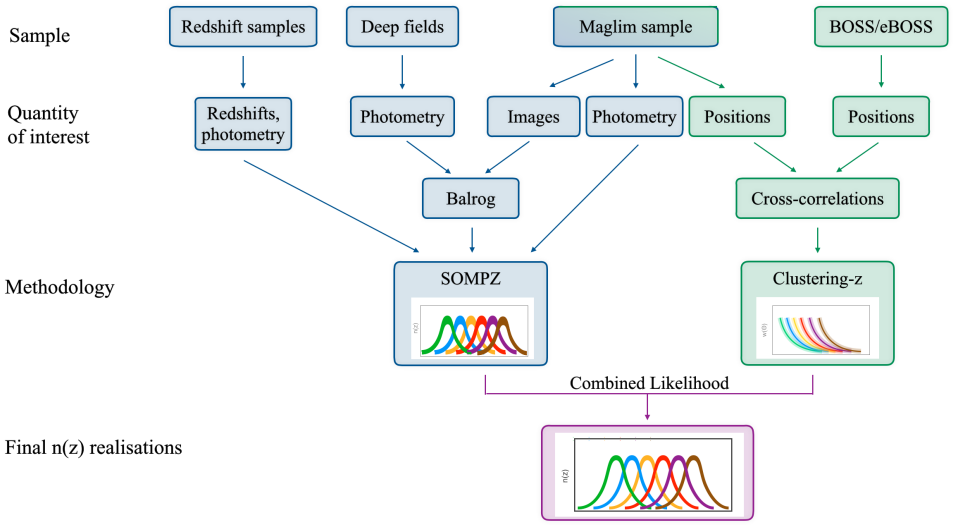

Figure 5.1: Flowchart illustrating the MagLim redshift distributions calibration scheme. The two methodologies included in the analysis are SOMPZ and clustering redshifts. Inspired by the flowchart in Myles & Alarcon et al. (2020).

particularly useful to reduce the dimensionality of a system, while preserving the distances between objects. The algorithm works in two steps. During the training phase, for each item of the training sample the most resembling node of the network (which here we will refer to as cell) is chosen through competitive learning. During the assignation phase, the trained SOM is able to classify different datasets following the recipe obtained with the training. It is an unsupervised method, meaning that the output variable, in our case the redshift, is not used for the training phase.

In a competitive learning neural network, the nodes compete to find the one that most closely resemble each item training example, until the Best Matching Unit (BMU) is found. The winner is then updated to make it more specialized. There are two ways in which this can happen: through a hard or soft competition. The hard competition has a winner-takes-all approach, in which only one output unit (the winner) gets updated; with the soft competition instead, both winner and its neighborhood are adapted. In our SOM algorithm, the soft approach has been chosen.

The characteristics of the SOM (e.g., the resolution and the topology) must be chosen beforehand. Despite being able to represent any complex topology, it is common use to define 2D SOMs because of their easy interpretation. Let us consider an input vector $\mathbf{x}$ (the galaxy properties, e.g. fluxes) and a weight

vector $\mathbf{w}_k$ assigned to each one of the SOM cells, with the cells being denoted by the index k. The basic steps of the algorithm are the following:

- At the first iteration, the normalized weights are generated randomly or from data samples;

- For each item in the training (galaxies in our case), the algorithm calculates which weight vector $\mathbf{w}_k$ is the nearest neighbour of the input vector x. The winning cell is the BMU. The metric used to determine the quality of the matching is the χ^2 distance:

$$d^2(\mathbf{x}, \mathbf{w_k}) = (\mathbf{x} - \mathbf{w_k})^T \mathbf{\Sigma^{-1}} (\mathbf{x} - \mathbf{w_k}) \tag{5.1}$$

  where Σ is the covariance matrix of the input vector x.

- Since we are using the soft competition method, neighbouring cells to the BMU are also updated. The function that determines which cells and how they are updated is the following:

$$\mathbf{w_k}(n_t + 1) = \mathbf{w_k}(n_t) + a(n_t) H_{b,k}(n_t)[\mathbf{x}(n_t) - \mathbf{w_k}(n_t)]. \tag{5.2}$$

  In the above equation, n_t indicates the current step of the procedure, and $n_t + 1$ the "updated" step; $H_{b,k}(n_t)$ is a Gaussian kernel that encodes the distance from the BMU, and it assumes the value of unity in case the index k refers to the BMU cell and decreases to 0 for distant cells. An Euclidian distance on the 2D map is used as a metric to compute $H_{b,k}(n_t)$. The function $a(n_t)$ is called learning rate function, and encodes the responsiveness of the map to the input vector as a function of the training time step. The function is a monotonically decreasing function, built to make the SOM less responsive as the training progresses.

After the training, new sets of data can be assigned to the trained SOM using the same χ^2 distance. The Bayesian probability of a galaxy of belonging to a specific cell k is given by:

$$-2 \log(k|\mathbf{x}, \mathbf{\Sigma}) = (\mathbf{x} - \mathbf{w_k})^T \mathbf{\Sigma^{-1}} (\mathbf{x} - \mathbf{w_k}) + const. \tag{5.3}$$

To simplify handling and reduce the computational load of the methodology, we are going to keep a single integer for each galaxy, representing only the cell maximising this probability.

SOMPZ photo-z inference

Let us consider the probability distribution function for the redshift of a galaxy $p(z)$; let us assume such a probability to be conditioned on the observed wide field color-magnitude $\hat{\mathbf{x}}$, covariance matrix $\hat{\Sigma}$, and selection $\hat{s}$. The probability can be written by marginalizing over deep photometric color $\mathbf{x}$ as follows:

$$p(z|\hat{\mathbf{x}}, \hat{\mathbf{\Sigma}}, \hat{s}) = \int d\mathbf{x}\ p(z|\mathbf{x}, \hat{\mathbf{x}}, \hat{\mathbf{\Sigma}}, \hat{s}) p(\mathbf{x}|\hat{\mathbf{x}}, \hat{\mathbf{\Sigma}}, \hat{s}). \tag{5.4}$$

The large dimensionality of this form prevents us from applying it to realistic data. This problem can be circumvented by discretising the color space $\mathbf{x}$ and ($\hat{\mathbf{x}}$, $\hat{\mathbf{\Sigma}}$) in *cells* c and $\hat{c}$, each spanning a portion of the whole sample and representing a specific galaxy phenotype, respectively of the deep and wide field. The galaxy samples are arranged in cells/phenotypes using SOMs. The choice of the topology of the cells follows Buchs et al. (2019), where a two-dimensional representation of the color space was chosen as it ensures an immediate visualisation of the data not possible otherwise. Once we compressed our data into SOMs, we can write the $p(z)$ for the group of galaxies living in a particular wide cell $\hat{c}$, by marginalizing over deep field phenotypes c:

$$p(z|\hat{c}, \hat{s}) = \sum_c p(z|c, \hat{c}, \hat{s}) p(c|\hat{c}, \hat{s}). \tag{5.5}$$

Source case. For the source galaxies, the selection function $\hat{s}$ is represented by the weak lensing selection, which requires that the galaxies are weighted by the shear response R (introduced in Section 2.4). The tomographic bins are selected such that they have equal number density; the procedure is explained in Section 4.1. At this point we can marginalise over all wide cells $\hat{c}$ belonging to each tomographic bin $\hat{b}$:

$$p(z|\hat{b}, \hat{s}) = \sum_{\hat{c}\in\hat{b}} p(z|\hat{c}, \hat{s}) p(\hat{c}|\hat{s}, \hat{b}) \tag{5.6}$$

$$= \sum_{\hat{c}\in\hat{b}} \sum_c p(z|c, \hat{c}, \hat{s}) p(c|\hat{c}, \hat{s}) p(\hat{c}|\hat{s}, \hat{b}). \tag{5.7}$$

Unfortunately there are very few galaxies for each $(c, \hat{c})$ pair, and in many cases there are none. This makes the term $p(z|c, \hat{c}, \hat{s})$ quite difficult to estimate. However, we can reasonably assume that the $p(z)$ for galaxies assigned to a given deep cell c should not depend on the noisy wide photometry of that

galaxy. Therefore we can relax the selection:

$$p(z|\hat{b},\hat{s}) \approx \sum_{\hat{c}\in\hat{b}} \sum_{c} p(z|c,\hat{b},\hat{s}) p(c|\hat{c},\hat{s}) p(\hat{c}|\hat{s}) \tag{5.8}$$

$$\approx \sum_{\hat{c}\in\hat{b}} \sum_{c} p(z|c,\hat{s}) p(c|\hat{c},\hat{s}) p(\hat{c}|\hat{s}). \tag{5.9}$$

Equation 5.8 is used whenever possible, while in case there is no galaxy satisfying both c and $\hat{b}$, Equation 5.9 is used.

Lens case. Since the MagLim tomographic bins $\hat{b}$ are defined based on DNF, we are going to construct one set of SOMs (one deep and one wide) for each bin $\hat{b}$. Assigning all galaxies belonging to a tomographic bin to a wide SOM is straightforward. Since for the MagLim sample in data there are no weights to be taken into consideration, we will drop the selection function as $\hat{s}_{\rm MagLim} = 1$.

At this point we can marginalise over deep field phenotypes c and over all wide cells $\hat{c}$ in $\hat{b}$. Here, we are computing $p(z|\hat{b})$ for each bin separately from different sets of SOMs, so Equation 5.6 becomes:

$$p(z|\hat{b}) \approx \sum_{\hat{c}} \sum_{c} p(z|c,\hat{c},\hat{b}) p(c|\hat{c},\hat{b}) p(\hat{c},\hat{b}). \tag{5.10}$$

For the simulations case, galaxies need to be weighted in order to match the color distribution in data, so Equation 5.6 is used instead. The same issue regarding the lack of galaxies that are in one (c, $\hat{c}$) pair is also present. Given that each tomographic bin occupies each SOM entirely. Equation 5.8 can be written as:

$$p(z|\hat{b}) \approx \sum_{\hat{c}} \sum_{c} p(z|c,\hat{b}) p(c|\hat{c},\hat{b}) p(\hat{c},\hat{b}). \tag{5.11}$$

In both instances, we obtain each of the terms appearing in Eq. 5.9 and Eq. 5.11, by assigning galaxy samples to the SOM cells, as follows:

- $p(\hat{c})$ is computed collecting *wide* field galaxies from either the weak lensing sample or the MagLim sample into a wide field SOM (for the lens case, one per tomographic bin);
- $p(c)$. In order to construct the deep SOM we have to use our Balrog sample, consisting of all detected and selected Balrog realisations of the galaxies in the wide field, each associated to its own "noiseless" replica

in the deep sample. For the MagLim case, we assign to the deep SOM associated to a tomographic bin, galaxies whose Balrog wide replica is selected in that specific wide bin.

- $p(c|\hat{c})$ is computed from the deep/Balrog sample. It consists of all detected and selected Balrog replicas of the deep galaxies injected in the wide field. We therefore can arrange the deep/Balrog sample simultaneously into a wide and deep SOMs. We call this term the *transfer function*. We weight the deep field galaxies according to their detection rate measured from Balrog. An alternative to Balrog would be using a sub-section of the wide field and deep fields overlap, giving us both deep and wide photometry for a limited number of galaxies. However, the area of overlap is small and the particular observing conditions found in this area will not be representative of the overall observing conditions found in the Y3 footprint as highlighted in Myles & Alarcon et al. (2020).

- $p(z|c)$ is computed from the *redshift* sample, which is a subset of the *deep* sample, for which we have both credible redshifts, 8-band deep photometry, and thanks to Balrog also wide-field realisations.

SOM properties

As in Buchs et al. (2019) and Myles & Alarcon et al. (2020), we use 2D squared SOMs with periodic boundaries, which makes the visualisation easier without compromising the efficiency. We parametrize the SOMs using luptitudes and lupticolors, following Buchs et al. (2019). Luptitudes are defined in Lupton et al. (1999) as inverse hyperbolic sine transformation of fluxes:

$$\mu = \mu_0 - a \sinh^{-1} \frac{f}{2b} \mu_0 = m_0 - 2.5 \log b, \tag{5.12}$$

where m are magnitudes, f are fluxes, $a = 2.5 \log b$ and b is a softening parameter that defines at which scale luptitudes transition between logarithmic and linear behaviour. For the deep SOM we compute 7 lupticolors with respect to the i-band

$$\mu = (\mu_1 - \mu_i, ..., \mu_7 - \mu_i), \tag{5.13}$$

where the index from 1-7 runs over the deep bands *urgzJHK*. We avoid using the g-band for the wide wield galaxies, as any observational systematics and chromatic effects are more evident in the g band. With only two lupticolors available in the wide SOM, we decided to add the i-band luptitude, as

Buchs et al. (2019) find empirically that addition of the luptitude improves the training performance:

$$\mu = (\mu_i, \mu_r - \mu_i, \mu_z - \mu_i). \tag{5.14}$$

The resolution of the SOMs must be chosen balancing three elements: first, the resolution should be sufficiently high to characterise the sample as accurately as possible; on the other hand, a resolution too high may cause computational problems, and cause a fraction of the SOM cells to not have any galaxy assigned to them. To ensure that we get a correctly estimated $n(z)$ is crucial, in fact, that the number of wide field galaxies assigned by the transfer function to a deep cell with no redshift information is kept under 1 %.

The resolution of the wide SOM has been chosen to be 32x32 cells, for both sources and lenses. Regarding the deep SOM, the weak lensing sample uses a resolution of 64x64 cells, while the MagLim sample 12x12. The reason behind the fewer cells in the deep SOM lies in the MagLim selection: the bright magnitude-redshift cuts must be applied also to the wide-component of the deep and redshift samples, and only the deep galaxies whose wide component is selected are included in the sample. This results in smaller deep and redshift samples covering a very small portion of the color space, compared to the weak lensing sample Myles & Alarcon et al. (2020). Besides, reducing the number of deep cells means yielding more galaxies in each one, ensuring that the number of cells with no redshift galaxies is below 1 %. Likewise, shot noise caused by a small number of redshifts in a deep cell can also play a significant role in biasing the estimate. We therefore performed a test to identify the optimal SOM size which would minimise these issues. We first computed several estimates in the Buzzard simulations using different resolutions for the deep SOM. We then evaluated which setting was producing the smallest shift on the mean redshift with respect to the true value.

As mentioned at the beginning of this section, SOMs are neural networks, thus they require to be trained before being able to classify galaxies. For the weak lensing sample, the deep galaxies were used to train the deep SOM. For the lenses, after ensuring that the redshift samples and the MagLim sample span the same luptitude-lupticolor space (achieved using Balrog to obtain the redshift samples wide photometry), we decided to use the redshift sample for the deep SOM training. We instead use the wide samples to train the all wide SOMs.

5.2 Clustering redshifts (WZ)

Clustering redshift (WZ) is a widely used method (Newman 2008; Ménard et al.; Davis et al. 2017; Morrison et al. 2017; Scottez et al.; Johnson et al. 2017; Gatti & Vielzeuf et al. 2018; van den Busch et al. 2020; Hildebrandt et al. 2020; Cawthon et al. 2020) to infer or calibrate redshift distributions of galaxy samples. It relies on the assumption that the cross-correlation between two samples of objects is non-zero only in the case of overlap of the distribution of objects in physical space, due to their mutual gravitational influence.

Various implementations of the clustering redshift methodology differ in their details, but they all agree on one key aspect: the "target" sample (hereafter dubbed "unknown" sample), which has to be calibrated, has to be cross-correlated with a "reference" sample divided into thin redshift bins. The reference sample consists of either high-quality photometric or spectroscopic redshift galaxies, and has to spatially overlap with the unknown sample. In what follows, we will try describe the general methodology implemented in DES Y3, clarifying when needed the differences in the implementation when calibrating either the source or the lens sample.

Assuming linear galaxy-matter bias, we can express the clustering $w_{\rm ur}$ between the unknown sample and each of the reference sample thin bins as function of the separation angle θ between the unknown and reference sample:

$$w_{\rm ur}(\theta) = \int dz' n_{\rm r}(z') n_{\rm u}(z') b_{\rm r}(z') b_{\rm u}(z') w_{\rm DM}(\theta, z') + M(\theta), \tag{5.15}$$

where $n_{\rm r}$ and $n_{\rm u}$ are the redshift distributions of the reference and unknown sample, $b_{\rm r}$ and $b_{\rm u}$ are the galaxy-matter biases of both samples, $w_{\rm DM}$ is the clustering of dark matter and $M(\theta)$ denotes contributions due to magnification. Note that we are assuming Limber approximation (Limber, 1953), but this has been shown to have no impact on the results (McQuinn & White, 2013).

In our methodology, we use a single estimated value from the cross-correlation signal for each thin redshift bin. In practice, we do this by measuring the correlation function as a function of angular separation and then averaging it with a weight function to produce the single estimate:

$$\bar{w}_{\rm ur} = \int_{\theta_{\rm min}}^{\theta_{\rm max}} d\theta \; W(\theta) w_{\rm ur}(\theta), \tag{5.16}$$

where $W(\theta) \propto \theta^{-1}$ is a weighting function. The integration limits in the integral in Eq. 5.16 are set to fixed physical scales (1.5 to 5 Mpc).

We use the Davis & Peebles (1983) estimator for the cross-correlation signal,

$$w_{\rm ur}(\theta) = \frac{N_{\rm Rr}}{N_{\rm Dr}} \frac{D_{\rm u} D_{\rm r}(\theta)}{D_{\rm u} R_{\rm r}(\theta)} - 1, \tag{5.17}$$

where $D_{\rm u}D_{\rm r}(\theta)$ and $D_{\rm u}R_{\rm r}(\theta)$ represent data-data and data-random pairs. The pairs are normalized through $N_{\rm Dr}$ and $N_{\rm Rr}$, which is the total number of galaxies in the reference sample and in the reference random catalog. The choice of the Davis & Peebles estimator over the Landy & Szalay (1993) estimator was driven by the advantage of having to produce a catalogue of random points for only one sample; this avoids creating high-fidelity random catalogues applying the very complex selection function for the DES Y3 source galaxy catalogue. For our analysis, we only rely on random points for the reference sample, whose selection function and mask are well understood. The correlation estimates were computed using `treecorr`[1].

The angular scales considered have been chosen to span the physical interval between 1.5 and 5.0 Mpc. These bounds, applied to data as well as simulations, are selected so that the upper bound is below the range used for the galaxy clustering cosmological analyses, therefore granting the WZ likelihoods to be essentially independent of the assumed cosmology, and allowing us to produce n(z) samples in an MCMC chain that runs independently of the cosmological ones. We perform the cross-correlations of either the source or the lens samples with the reference sample divided into 50 bins of width $\Delta z = 0.02$, spanning the $0.1 < z < 1.1$ interval. In the case of the lens sample, we also weigh each galaxy of the MagLim sample by the clustering weights computed in Rodríguez-Monroy et al. (2022).

Since the $n_{\rm r}$ are binned in narrow bins we can approximate the number density of the sample of reference as a Dirac delta, and the revised expression becomes:

$$\bar{w}_{\rm ur} \approx n_{\rm u} b_{\rm r} b_{\rm u} \bar{w}_{DM} + \bar{M}. \tag{5.18}$$

The above equation relates the redshift distribution of the unknown sample to the measured clustering signal $\bar{w}_{\rm ur}$. The barred quantities indicate they have been "averaged" over angular scales; in what follows, for simplicity, we will drop the barred notation. The galaxy-matter biases of the reference can be estimated from the auto-correlation of the reference sample. Usually the galaxy-matter bias of the unknown sample cannot be inferred and is treated as a nuisance parameter. Later in this section, we introduce a flexible model that captures any reasonable redshift evolution of the galaxy-matter bias of the unknown sample. In the case of the source sample, we marginalise over the parameters of such a model using reasonably wide priors. In the case of the lenses, however, due to the relatively good per-galaxy redshift estimate provided by DNF, we use the autocorrelation of the sample to inform the priors of such a model (see section 7.2.2). We assumed the galaxy-matter bias to be

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

described by a single number at all scales; this is true at large scales in the linear regime, but we do not expect this to necessary hold at the small scales used in this work (1.5 to 5.0 Mpc). In the non-linear regime, even the fact that the terms inside the integral factorizes into $b_{\rm r}(z_i)b_{\rm u}(z_i)w_{\rm DM}(z_i)$ is not guaranteed (Bernardeau et al., 2002; Desjacques et al., 2018). The linear-bias assumption introduces a systematic uncertainty that depends on the scales adopted and the samples under study and that will be quantified in the following sections.

The other terms in the above equation are the clustering of dark matter $\bar{w}_{\rm DM}$, which can be estimated from theory and we found to be not very sensitive to the cosmological parameters, and the magnification term, which is expected to have a little impact and can be estimated if magnification coefficients for the samples are provided.

The evolution of the quantities $b_{\rm r}(z_i)$, $b_{\rm u}(z_i)$, $w_{\rm DM}(z_i)$, $M(z_i)$ need to be correctly characterized to recover the redshift distribution of the unknown sample. We turn now to how to model or estimate these terms in more details.

- **The galaxy-matter bias evolution of the reference sample** $b_{\rm r}(z)$. We can estimate $b_{\rm r}(z)$ by measuring and the auto-correlation function of the reference sample, split into thin redshift bins ($\delta z = 0.02$). This can be done if the redshifts of the reference sample are accurate enough and under the assumption of linear biasing. The angle-averaged auto-correlation function is:

$$w_{\rm rr}(z_i) = \int dz' \left[b_{\rm r}(z')n_{\rm r,i}(z')\right]^2 w_{\rm DM}(z'). \tag{5.19}$$

  If the bins are sufficiently narrow, both the biases and $w_{\rm DM}$ can be assumed as constant; they can therefore be removed from the integrals.

$$w_{\rm rr}(z_i) = b_{\rm r}^2(z_i)w_{\rm DM}(z_i) \int dz' n_{\rm r,i}^2(z'), \tag{5.20}$$

  The redshift distributions of the narrow bins and $w_{\rm DM}(z)$ are required to estimate $b_{\rm r}(z_i)$ from Eq. 5.20.

- **The galaxy-matter bias evolution of the unknown sample** $b_{\rm u}(z)$. This can theoretically be estimated in the same way as $b_{\rm r}(z)$. However in some cases, the redshift quality of the unknown sample is not sufficient to create thin bins; in this case $b_{\rm u}$ likely varies substantially across the sample, so the information on $b_{\rm u}$ from the auto-correlation is weak. The limiting factor of clustering-z methods is precisely this degeneracy between

$b_{\rm u}$ and $n_{\rm u}$. Mitigation schemes exist, based on the use of additional information to constrain the evolution of $b_{\rm u}$: e.g., Matthews & Newman 2010 use the additional constraints coming from the auto-correlation function of the tomographic bins (without dividing the samples into thin bins); or the method implemented in van den Busch et al. 2020, who use the additional constraint coming from the normalisation of the redshift distribution of the full unknown catalog not divided into tomographic bins. However, these methods are not free from shortcomings. In DES Y3 we parametrize $b_{\rm u}$ in a flexible way, effectively treating it as a free function. Depending on the sample, we impose priors on its parameters either based on simulations (source sample) or the autocorrelation of the sample divided into thin bins (lens sample).

- **The dark matter 2-point correlation function** $w_{\rm DM}(z)$. This can be modeled assuming a given cosmology and a non-linear power spectrum:

$$w_{\rm DM}(z_i) = \int d\theta W(\theta) \sum \frac{2\ell+1}{4\pi} P_\ell(\cos\theta) \frac{1}{\chi(z_i)^2 H(z_i)} P_{\rm NL}\left(\frac{l+1/2}{\chi(z_i)}, z_i\right), \quad (5.21)$$

where $P_\ell(x)$ is a Legendre polynomial of order ℓ, χ is the comoving distance, and $H(z_i)$ is the Hubble rate at redshift z_i. $P_{\rm NL}(k,\chi)$ is the 3D non-linear matter power spectrum. Under the Limber approximation, the wavenumber k is set equal to $(l+1/2)/\chi(z_i)$. The redshift evolution of $w_{\rm DM}(z_i)$ depends little on the particular value of cosmological parameters, whereas the dependence of the overall amplitude of $w_{\rm DM}(z_i)$ with respect to cosmology is absorbed by our systematic functions. Based on this, we hold cosmology fixed when computing $w_{\rm DM}(z_i)$, assuming the values in Planck Collaboration 2018). We then verify *a posteriori* that this approximation is valid by repeating our analysis using very different values for the cosmological parameters ($\Omega_{\rm m} = 0.4$, $\sigma_8 = 0.7$), finding that the impact on our conclusions is negligible. Note that some of the mitigation schemes adopted in literature to correct the galaxy-matter bias evolution of the unknown sample also automatically estimate $w_{\rm DM}(z_i)$ from the data (Matthews & Newman, 2010; van den Busch et al., 2020), but they are not adopted in this work.

- **Magnification signal** $M(z_i)$. Weak lensing magnification (Narayan, 1989; Villumsen et al., 1997; Moessner & Jain, 1998) changes the observed spatial density of galaxies: the enhancement in the flux of magnified galaxies can locally increase the number density, as more galaxies

pass the selection cuts/detection threshold of the sample; at the same time, the same volume of space appears to cover a different solid angle on the sky, generally causing the observed number density to decrease. For a flux limited sample, the net effect is driven by the slope of the luminosity function of the sample, here conveniently parametrized through the parameter α, and it has an impact on the measured clustering signal. Formally, the magnification term depends on the galaxy-matter bias and parameter α of the two samples, as well as on the redshift distribution of the unknown sample: $M(z_i; \alpha_\mathrm{r}, \alpha_\mathrm{u}, b_\mathrm{r}, b_\mathrm{u}, n_\mathrm{u})$. More details about our modelling of the magnification effects are given in Appendix A.1, although we anticipate magnification effects have a negligible impact on our analysis, due to our analysis choices. To keep our notation light, when possible, we will simply indicate magnification effects as $M(z_i)$, dropping the dependence on other factors.

As long as the approximation of thin reference bins and linear galaxy-matter bias hold, and using the model for the magnification term described in Appendix A.1, Eq. 5.18 simplifies to a linear system of equations, which can be solved providing an estimate of $n_\mathrm{u}(z_i)$. Several implementations of the WZ method have followed this approach (Newman, 2008; Ménard et al.; Schmidt et al., 2013; McQuinn & White, 2013), including the DES Y1 analysis (Gatti et al., 2018). A different approach, however, can be taken. In a forward modelling fashion, given an estimate of the $n_\mathrm{u}(z_i)$ (provided, for instance, by a photo-z method), Eq. 5.18 can be used to evaluate the expected correlation signal $w_\mathrm{ur}(z_i)$ and compare it to the one measured in data (see, e.g., Choi et al. 2016). This is the approach followed in this thesis.

Last, we note that many of the terms described above that are modeled in this work were not modeled in the DES Y1 analysis. In fact, $b_\mathrm{r}(z_i)$, $b_\mathrm{u}(z_i)$, $w_\mathrm{DM}(z_i)$ were assumed to be constant within each photo-z bin, and simulations were used to estimate the systematic error induced by this assumption. In DES Y1 magnification effects had not been modeled either: the redshift range where magnification effects were expected to have a non-negligible impact (i.e., the tails of the redshift distributions) had been simply excluded from the constraints. On the contrary in this work all these terms are modeled.

5.3 Combination of SOMPZ and WZ - how to sample from the joint likelihood

The redshift information from the SOMPZ and the clustering redshifts methods has to be then combined. Standard methods that combine photo-z estimates

with WZ constraints usually imply shifting around their mean the redshift distributions estimated using some photo-z code until a good match with the $n(z)$ estimated by the WZ constraints. This method can introduce biases (Gatti et al., 2018; van den Busch et al., 2020) if the two redshift distributions estimated from the two methods differ significantly in their shapes. This basically happens if the matching scheme is not flexible enough and does not properly tae into account all the uncertainties of the two methods. Our combination approach is different: we use a joint likelihood framework, by assigning a likelihood of measuring the data, obtained through either SOMPZ or WZ, given some model, which is going to contain different parameters describing the measured quantities. All the uncertainties of the two methods are automatically taken into account in this framework. The combined, final redshift distributions will therefore be sampled from the joint likelihood function of SOMPZ+WZ, and will not suffer from the aforementioned problems.

The SOMPZ likelihood describes the likelihood of obtaining the SOMPZ $n(z)$ measurement given the occupancy of the SOM deep cells, taking into account the fact that the occupancy of the cells is affected by sample variance (which is the largest form of uncertainty in the SOMPZ method). The likelihood is fully analytical and it is described in more detail in section 7.2.1, where we explain how systematic uncertainties are dealt with.

Concerning the WZ likelihood, it basically describes the probability of having a certain clustering amplitude $\bar{w}_{\rm ur}$ given a $n(z)$ (and all the parameters of the model). The clustering amplitude is computed following Eq. 5.18; the likelihood is assumed to be Gaussian, and the covariance is taken from the clustering measurement uncertainties. The detailed description of the WZ likelihood can be found in section 7.2.2, with its application for the DES Y3 source sample, and in section 7.2.2 for its use for the DES Y3 MagLim sample.

In order to combine the SOMPZ likelihood with the WZ likelihood, one could in principle draw $n(z)$ samples from the SOMPZ likelihood and importance-sample them by the value of their WZ likelihood with each $n(z)$ draw. In the case of high dimensional spaces, this can become inefficient. In many dimensions, the regions of phase space where the likelihood is high is a very small fraction. If guesses are isotropic while the "good" probabilities are confined in a specific part and have a direction, the process is bound to be inefficient. The issue is also that if one were to use the samples of one to calculate the likelihood of the other, most of the samples that are considered having high probability for one of the distributions, are likely not to pass the threshold for the other one. This happens because the WZ likelihood naturally accepts "smooth" $n(z)$s, since the WZ measurement is performed over a large area and it is not affected by sample variance; the opposite is true for the SOMPZ

likelihood, where the dominant source of uncertainty is sample variance.

In order to sample from the joint distribution and marginalise over the parameters relative to either one, Metropolis Hastings algorithms can be used. These algorithms start with a guess for the parameters of the two likelihoods, based on the parameter space allowed by the parameters prior. A proposal for a new point is then provided given a proposal distribution; if the probability for the set of parameters is larger than the previous one, then it is accepted, otherwise it is kept only a random fraction of the times. However, because of the reasons mentioned earlier (high dimensionality and different portions of the parameter space probed by the two likelihoods), the Metropolis Hasting algorithms are bound to have extremely low acceptance rates.

One solution would be to modify the proposal distribution to allow the next sampling to happen in a small neighborhood in parameter space, resulting in a higher acceptance probability; however this considerably slows down the process, as it becomes harder to sample the totality of the high likelihood portion of parameter space in a timely manner.

The Hamiltonian Monte-Carlo algorithm (HMC) provides a clever way to deal with this problem. The HMC is a Markov chain method which uses the mathematical properties of Hamiltonian systems to obtain a sequence of random samples which are distributed according to a target probability distribution for which direct sampling is difficult.

The HMC algorithm starts by considering each point of the parameter space to be a point-mass, with potential and kinetic energies U and K. The potential energy is determined by the multidimensional probability distribution $P(q)$, of which the logarithm is taken to avoid dealing with negative values. If we have a probability function $P(q = x_1, x_2, \ldots)$ we can write the potential energy of the system as

$$U = -\log P(q). \tag{5.22}$$

In our case, $P(q)$ is the joint SOMPZ and WZ likelihood, multiplied by the parameters priors (i.e., the posterior), with q being the parameters of the likelihood. The kinetic energy is defined as:

$$K = \frac{1}{2} p^T M^{-1} p \tag{5.23}$$

were p are "momenta" assigned to the point-like particles, and M a mass matrix, which is a free parameter of the HMC method that is chosen to optimize the convergence of the algorithm. The total energy is of course $H = U + K$, and correspond to the joint probability function of P(p,q):

$$P(p, q) = e^{-H} = P(q) e^{p^T M^{-1} p}. \tag{5.24}$$

The idea then is to solve the Hamiltonian equation of motion, obtaining the time evolution of the system given an initial point in the phase space (q_0, p_0); every new point q', indeed, can be considered as a "valid" sample of $P(q)$, which is the joint SOMPZ and WZ posterior. The system of equations is solved numerically; the time-step is chosen to be sufficient to allow the "particle" to both remain in the area and at the same time explore the phase space. For each step the time of integration will be randomly chosen within an interval. To perform the numerical integration of these equations, the derivatives of the Hamiltonian with respect to p and q need to be computed. The considerable advantage of the HMC is that almost every sample drawn can be accepted despite being in very high dimensions, while keeping two next samples uncorrelated from each other, due to the larger step. This is really hard to obtain with other sampling as standard Metropolis Hastings. We will see more details on its punctual application in Sections 7.2.1 and 7.2.2.

Chapter 6

Clustering redshifts: calibration of the DES Y3 Weak Lensing Source Redshift Distributions

6.1 Introduction

The DES Y3 3x2pt analysis features the largest shape catalogue ever built, counting more than 100 million galaxy shapes (Gatti et al., 2020c). In order to achieve an accurate cosmological analysis it is crucial that the redshift distributions are unbiased. In particular, it has been shown that biases on the mean redshift of samples used for weak lensing analysis lead to wrong estimates on the cosmological parameters (Bonnett et al., 2016b; Hoyle et al., 2018b; Tanaka et al., 2018; Troxel et al., 2018; Hildebrandt et al., 2020). Here we discuss the fiducial DES Y3 calibration of the weak lensing sample, applying the methodology described in Chapter 5. In particular, we focus on the part of the calibration concerning the clustering redshift measurement and the combination with the SOMPZ estimates obtained by Myles & Alarcon et al. (2020).

As already seen, the methodology consists of a photometric method and a clustering-based one. The photo-z estimates for the DES Y3 weak lensing sample are provided by a Self-Organizing-Map-based scheme (hereafter SOMPZ, Buchs et al. 2019; Myles & Alarcon et al. 2020). The SOMPZ method provides an ensemble of redshift $n(z)$ for a given galaxy sample, which captures the uncertainties in the redshift distributions at all orders (i.e., not only in the mean or width of the distributions). The clustering constraints (WZ) are then incorporated through a rigorous joint likelihood framework where the cluster-

ing data is forward modelled as a function of the input $n(z)$, and the specific WZ systematics are marginalized over. This scheme allows to draw $n(z)$ samples conditioned on both clustering and photometric measurements, improving the $n(z)$ estimates by correctly taking into account the significance of the information provided by each source of information. This combined approach has proven to be more robust than SOMPZ or WZ applied individually (Gatti & Giannini et al., 2020a), as the combination exploits the complementarity of both methods and reduces the overall $n(z)$ uncertainty. In either route, the DES Y3 cosmological analysis is done by sampling over the finite set of realizations generated by SOMPZ+clustering-z.

For the clustering measurement, we use two different reference samples: luminous red galaxies from the RedMaGiC catalog with accurate photometric redshifts; and a spectroscopic sample from the combined BOSS (Baryonic Oscillation Spectroscopic Survey, Dawson et al. 2013) and eBOSS (extended-Baryon Oscillation Spectroscopic Survey, Dawson et al. 2016; Ahumada et al. 2019; Alam et al. 2020) catalogs. The two samples have properties that complement each other: using photometric redshifts (even though high quality) rather than spectroscopic redshifts is counterbalanced by the higher statistical power of the RedMaGiC sample, which is defined over the full DES footprint and is characterised by a higher number density than BOSS/eBOSS galaxies. The latter also cover only $\approx 17\%$ of the DES Y3 footprint. On the other hand, the BOSS/eBOSS sample spans a wider redshift range and has spectroscopic redshift estimates.

This Chapter is organised as follows: a brief summary of the results from the SOMPZ analysis is given in 6.2; the detailed description of the treatment of systematic uncertainties and how the WZ likelihood is built, is presented in 7.2.2. Finally, the combined results are shown in Section **??**.

6.2 Results from SOMPZ

The WZ measurement is to be combined with the $n(z)$ obtained with the SOMPZ method. The details of the SOMPZ implementation for this method are described in Myles & Alarcon et al. (2020). Here we are only reporting a summary, focusing on the delivered product.

The samples from which the redshift information is acquired are the same as the ones outlines in Section 4.7, with the only difference that only four spectroscopic surveys were used. These are zCOSMOS (Lilly et al., 2009b), C3R2 (Masters et al., 2017), VVDS (Le Fèvre et al., 2013b), and VIPERS (Scodeggio et al., 2018b). The source galaxies are much deeper than MagLim

therefore it is important to rely on surveys that present the least amount of bias possible up to faint magnitudes.

The product of the SOMPZ calibration is a set of redshift distributions that encompass the systematic uncertainties of the method. The relevant systematics are the following:

- Sample Variance: fluctuations in the underlying matter density field determine the abundance of observed deep field galaxies of a given 8-band colour and at a given redshift. This uncertainty is a main contributor to the uncertainty budget in all of our tomographic bins;
- Shot Noise: shot noise in the counts of deep field galaxies of a given 8-band colour and at a given redshift;
- Redshift Sample Uncertainty: biases in the redshifts of the redshift galaxy samples used. This uncertainty is estimated by performing our inference with multiple different underlying redshift samples, and it is non-negligible in the third and fourth tomographic bins and dominant in the third tomographic bin;
- Photometric Calibration Uncertainty: uncertainty in the 8-band colour of deep field galaxies. This uncertainty is non-negligible in the first tomographic bin.

The total amount of uncertainty in the mean redshift of each tomographic bin is (0.015, 0.010, 0.007, 0.012). The dominant uncertainty is sample variance, caused by the small area of the deep fields, from which we are inferring the type/redshift relation. It may be possible that the deep fields are located in special regions of the Universe, thus the properties measured would not correspond to the rest of the weak lensing sample. An analytical model called *3sDir* has been developed to quantify the uncertainty introduced by sample variance, and by the limited number of galaxies with excellent redshifts in the deep fields, which causes shot noise. The 3sDir computes the likelihood of galaxies having a specific redshift and colour phenotype, given that a number of galaxies have been observed to have the same redshift and colour phenotype. The 3sDir likelihood is used in the joint sampling of the final $n(z)$, together with the WZ likelihood.

6.3 Characterization of sources of uncertainty

Following the description of the clustering signal presented in §5.3, we are able to produce a model for the $w(z)$ signal across the full redshift range covered

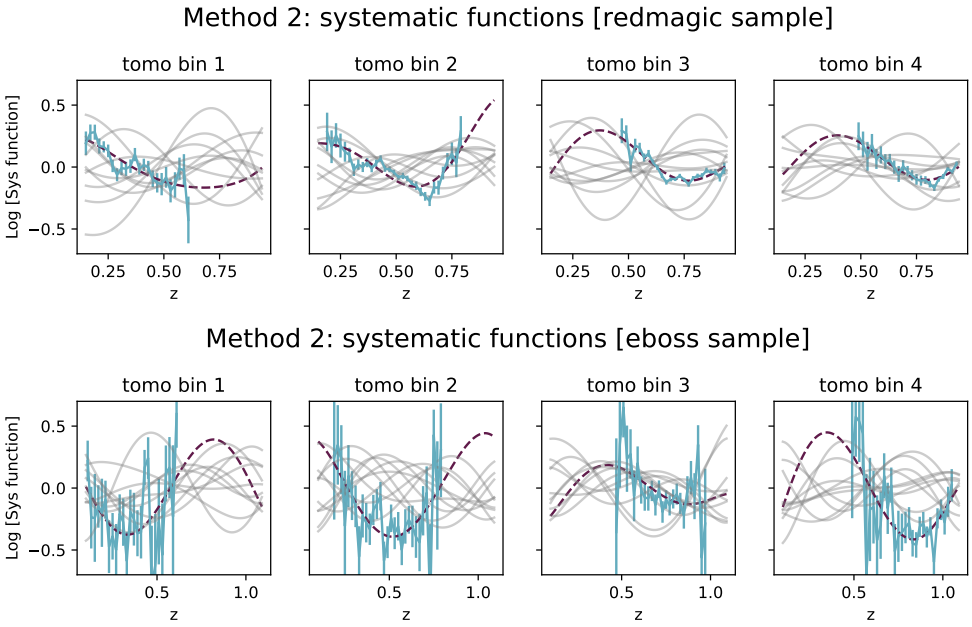

Figure 6.1: Systematic uncertainties of the WZ method as measured in simulations following Eq. 6.1, for the 4 tomographic bins and for the two reference samples considered (RedMaGiC upper panels, BOSS/eBOSS, lower panels). The measured systematic uncertainties are represented by the light blue lines; the purple dashed lines represent the best fitting model. The grey lines represent 10 random realisations of the systematic uncertainty model assumed for the WZ method and described by Eq. 7.2. Note that the rapid upturn of the systematic function in bin 2 in the case of the RedMaGiC sample is due to a rapid evolution of the galaxy-matter bias of the unknown sample, related to a strong evolution of the properties of the galaxy population. Such an evolution is also present in other bins, but it is milder. When the BOSS/eBOSS sample is used, the lower sensitivity does not allow to appreciate this rapid change in the slope of the systematic function.

by the reference samples given an input redshit distribution and produce a likelihood for the observed $w(z)$ data. In practice, this allows us to constrain the full shape of the redshift distributions. Here we use the Buzzard simulations to set the priors for the systematic-error parameters within this model.

Systematic uncertainty determination

Recall that in §5.2, specifically Eq. 7.5, the cross-correlation signal is modelled starting from a proposed value for $n_{\rm u}(z)$ (e.g. provided by SOMPZ), the (measurable) reference-population properties $b_{\rm r}(z)$ and $\alpha_{\rm r}(z)$, and nuisance parameters for the (poorly known) bias and magnification properties of the source population $b_{\rm u}(z)$ and $\alpha_{\rm u}(z)$. We will set these last two as constant over redshift and marginalize over broad priors on these constants, to flexibly model the magnification signal. The underlying function $w_{\rm DM}(z)$ is estimated assuming a cosmological model.

The final component of the $\hat{w}_{\rm ur}$ model is a function $\mathrm{Sys}(z, \mathbf{s})$ that multiplies the true clustering signal and will absorb the systematic errors described in Section 5.2: failures of the linear-bias model itself; the unknown and redshift-dependent $b_{\rm u}(z)$; and possible errors in the $n_{\rm r}(z)$ functions for RedMaGiC bins. The parameters $\mathbf{s}$ of this systematic function will be marginalized as well, as per Eq. 7.6.

Our strategy will be to determine what the $\mathrm{Sys}(z, \mathbf{s})$ function is in the Buzzard simulation, and then produce a prior on the $\mathbf{s}$ parameters which allows marginalization over a broad family of functions with similar form of deviation from unity. The $\mathrm{Sys}(z, \mathbf{s})$ function is given substantial freedom for low-order, smooth variation with z, as we expect from all of the systematic errors, leaving the finer-scale information in $w_{\rm ur}(z)$ to constrain fine-scale behavior in $n_{\rm u}(z)$, i.e. the shape of $n_{\rm u}(z)$.

The blue data points in Fig. 6.1 plot the $\mathrm{Sys}(z, \mathbf{s})$ functions observed in the Buzzard simulations, for both reference samples. Namely they plot:

$$\mathrm{Sys}_{\rm sim}(z_i) = \frac{w_{\rm ur}(z_i)}{\hat{w}_{\rm ur}(z_i))}, \tag{6.1}$$

where the model uses the true $n_{\rm u}(z), b_{\rm r}(z)$, and $n_{\rm r}(z)$ values. We evaluate and plot this ratio only in the z interval where the $w_{\rm ur}$ signal is large enough to have good signal-to-noise and subdominant magnification contribution. The RedMaGiC $w_{\rm ur}(z)$ uses RedMaGiC photo-z's for binning, just as the real data do.

The $\mathrm{Sys}(z, \mathbf{s})_{\rm sim}$ ratio deviates from unity due to systematic effects, as expected. We quantify this by the RMS of $\log[\mathrm{Sys}(z, \mathbf{s})_{\rm sim}]$, which are measured

to be (0.11, 0.07, 0.07, 0.11) for the RedMaGiC tomographic bins and (0.18, 0.15, 0.10, 0.15) for BOSS/eBOSS. From this we conservatively decide that the $\mathrm{Sys}(z, \mathbf{s})$ function needs to have the freedom to have RMS (log) fluctuations of ≈ 0.15 as 1σ deviations under its $p(\mathbf{s})$ function.

We seek a parametric function $\mathrm{Sys}(z, \mathbf{s})$ and a prior $p(\mathbf{s})$ which have these desired properties:

- The function and prior yield a good fit to the $\mathrm{Sys}(z, \mathbf{s})_{\mathrm{sim}}$ measured in Buzzard.
- The prior can be tuned to yield typical RMS variations in $\log[\mathrm{Sys}(z, \mathbf{s})]$ at similar level to that seen in Buzzard.
- The parametric form allows a similar smoothness of variation as seen in Buzzard, i.e. similar number of "wiggles" across the $0 < z < 1.2$ range where the WL source galaxies lie.
- The RMS of $\log[\mathrm{Sys}(z, \mathbf{s})]$ as we vary $\mathbf{s}$ under the prior $p(\mathbf{s})$ is a flat function of z.
- The prior on $\mathbf{s}$ is simple to construct and to use in a Hamiltonian Monte Carlo chain.

We chose the $\mathrm{Sys}\,(\mathrm{z}, \mathbf{s})$ function to be given by:

$$= \sum_{k=0}^{M} \frac{\sqrt{2k+1}}{0.85} s_k P_k(u), \tag{6.2}$$

$$u \equiv 0.85 \frac{z - 0.5(z_{\max} + z_{\min})}{(z_{\max} - z_{\min})/2}. \tag{6.3}$$

with $P_k(z_i)$ being the k-th Legendre polynomial, M is the maximum order, and the second line linearly remaps the z interval $[z_{\min}, z_{\max}]$ to $[-0.85, 0.85]$. The fraction under the summation makes the basis functions close to orthonormal so that the RMS of $\log[\mathrm{Sys}(z, \mathbf{s})]$ is $|\mathbf{s}|^2$. The prior $p(\mathbf{s})$ is chosen to be a simple diagonal normal distribution with standard deviations $\{\sigma_{s0}, \ldots, \sigma_{sM}\}$ and means of zero. Mathematical details of this choice for the systematic function and its prior are given in Appendix A.2.

A distinct set of nuisance parameters $\mathbf{q} = \{\mathbf{p}, \mathbf{s}\}$ (with $\mathbf{p} = \{b_{\mathrm{u}}^{'}, \alpha_{\mathrm{u}}^{'}\}$) are assigned to each combination of tomographic bin and reference sample, and each of these 8 sets of w_{ur} measurements are fit independently. We set $[z_{\min}, z_{\max}]$ to span the full range of the reference catalog, $[0.14, 0.90]$ for RedMaGiC and $[0.10, 1.06]$ for BOSS/eBOSS. We set $M = 5$ and we set the σ_{s_i} to yield an

expectation value of 0.15 for the RMS of $\log[\mathrm{Sys}(z, \mathbf{s})]$. The order M was chosen by finding the value beyond which the RMS residual stopped decreasing for a fit of Eq. 7.2 to the $\mathrm{Sys}(z, \mathbf{s})$ function found in the simulated RedMaGiC $w_{ur}(z)$ data. The σ_{s_i} prior is set to make the simulated $\mathrm{Sys}(z, \mathbf{s})$ functions be $\approx 1\sigma$ fluctuations from a constant. Since e^{s_0} is approximately the mean bias of the unknown sample, and we expect the mean bias b_r to be more uncertain than the variation with redshift, we treat the prior on s_0 somewhat differently, giving it a wide prior $\sigma_{s0} = 0.6$. The RMS of 0.15 is then allocated among the remaining elements $k \geq 1$ of $\mathbf{s}$ which model redshift-dependent systematic errors.

The nuisance parameter $b_{\mathrm{u}}^{'}$ used in magnification estimation is given a Gaussian prior with $(\mu, \sigma) = (1., 1.5)$ (which encompasses the bias of the weak lensing sample as measured in simulation). The other magnification nuisance $\alpha_{\mathrm{u}}^{'}$ is given a mean estimated from image-injection simulations (Appendix A.1) and a conservatively large uncertainty of $\sigma = 1$.

The dashed curves in Fig. 6.1 plot the $\mathrm{Sys}(z, \mathbf{s})$ functions obtained from the maximum-posterior fits to the simulations' $w_{\mathrm{ur}}(z)$ data, combining the priors on the nuisance parameters with the likelihood of Eq. 7.6. In all cases, the best fit models succeed in capturing the slowly varying component of the systematic. In some bins, some of the rapid variations in redshift are not well captured - this is expected, as we truncate the polynomial of the $\mathrm{Sys}(z, \mathbf{s})$ function to order $M = 5$. While this could be improved by increasing the maximum order M, we find in practice that these small discrepancies cause no significant bias in the recovered redshift distributions when the method is applied in simulations (see below). The fitted functions remain well-behaved over the full w_{ur} redshift range even though the fit is done only for redshifts with strong signals. We conclude that this formulation of the systematic errors is sufficient to model the systematic errors in our clustering-z measurement in the Buzzard simulation, and we assume that marginalization over $\mathbf{q}$ will allow us to capture the uncertainties present in the real data as well.

The grey curves in Fig. 6.1 show a few examples of $\mathrm{Sys}(z, \mathbf{s})$ functions obtained by random sampling of the prior $p(\mathbf{s})$. This illustrates the flexibility of our model for the systematic uncertainty, which is able to model a large variety of curves.

Application of the method in simulations

Once our family of systematic functions is determined, we may proceed to validating the performance of the combination of SOMPZ and WZ on the Buzzard simulations. This combination is implemented (both in simulations

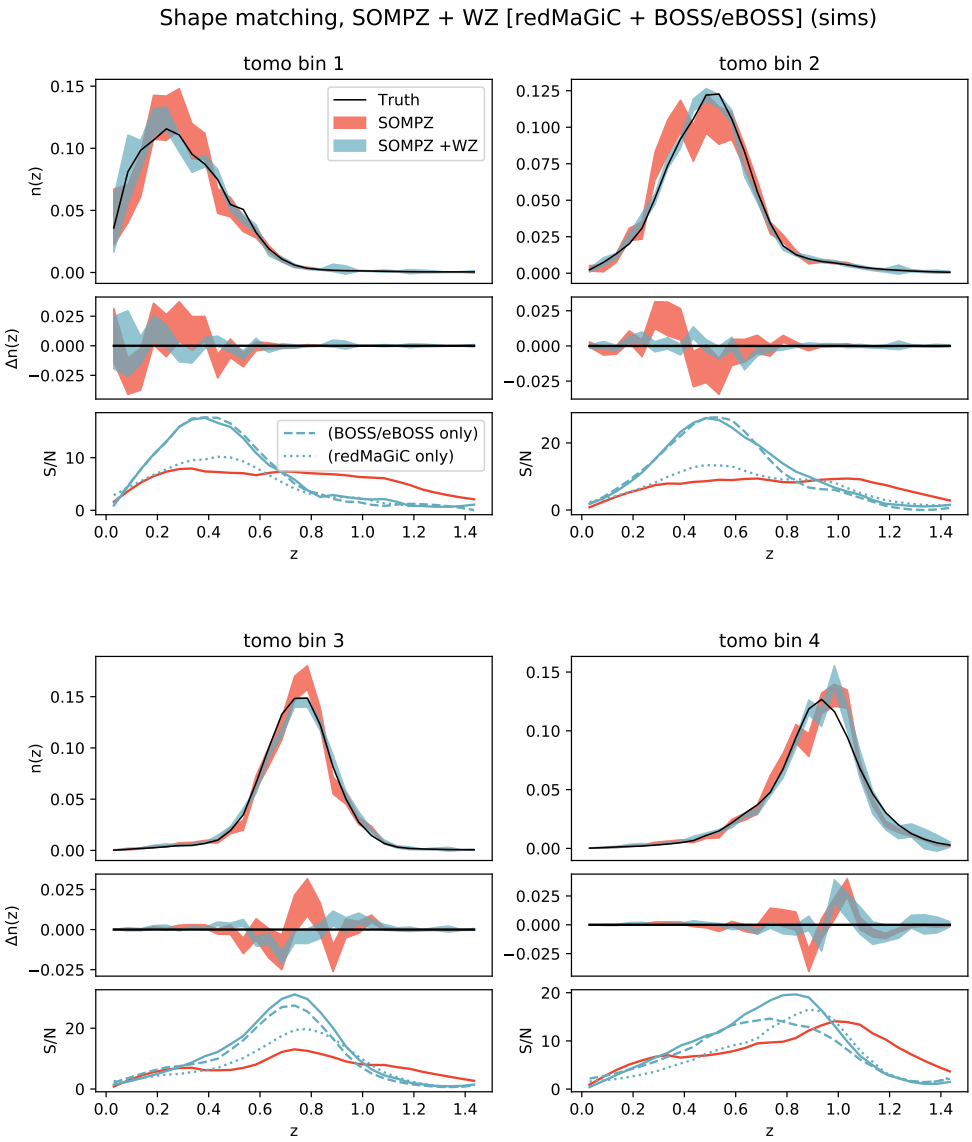

Figure 6.2: For each tomographic bin, three panels are shown. *Upper panels:* SOMPZ redshift distributions, as estimated in simulations, with and without clustering information. The bands encompass 68% confidence interval of the SOMPZ $n(z)$ realisations. *Central panels:* difference between the recovered $n(z)$ and the true $n(z)$ in simulations. *Lower panels:* S/N, defined as the ratio between the SOMPZ $n_{\rm u}(z)$ and its 68% confidence interval of the SOMPZ realisations, with and without clustering information. The dashed (dotted) line has been obtained only using clustering-z constraints from RedMaGiC (BOSS/eBOSS) galaxies.

and in data) by sampling the $n_{\rm u}(z)$ functions for all 4 tomographic bins from a posterior defined by the product of:

- the SOMPZ probability defined by Myles & Alarcon et al. (2020);
- the clustering-z probability defined by Eq. 7.6 for the $w_{\rm ur}(z_i)$ measured against the RedMaGiC sample, marginalized over $\mathbf{q}$ as described in Appendix A.2;
- and likewise, the marginalized clustering-z probability derived for the BOSS/eBOSS sample, marginalized over $\mathbf{q}$ as described in Appendix A.2;

The clustering-z probabilities use $w_{\rm ur}(z)$ over the full redshift range of their respective reference samples. The reference-sample magnification coefficients $\alpha_{\rm r}$ and the cosmology used to derive $w_{\rm DM}(z)$ are held fixed to nominal values. We verify below that the choices of $\alpha_{\rm r}$ and cosmology have insignificant impact on the outcome of the WZ method. We consider the RedMaGiC and BOSS/eBOSS likelihoods independent, i.e. they do not share the WL galaxy bias uncertainty. We did this because we did not split our systematic function into different source of errors, owing to an increasing complexity in the modelling. Given the flexibility of $\mathrm{Sys}(z, \mathbf{s})$ and the conservative choice on the RMS of $\log[\mathrm{Sys}(z, \mathbf{s})]$, considering the RedMaGiC and BOSS/eBOSS likelihoods independent should not be an issue for the methodology.

Fig. 6.2 compares the Buzzard true redshift distribution to the distributions drawn from only the SOMPZ likelihood and the distributions drawn from the joint SOMPZ+WZ posterior. The distributions of the mean redshifts per bin in the lower panels are not shown, but it is reported in Table 6.1. It shows that the WZ likelihood adds little information on these mean z's. The plots in Fig. 6.2, however, shows that the addition of the clustering-z likelihood produces a remarkable improvement in the fidelity of the shape of $n_{\rm u}(z)$ to the truth. To better quantify the improvement, we also show the signal-to-noise (S/N) of the $n_{\rm u}(z)$ estimates, defined as the ratio between the SOMPZ $n_{\rm u}(z)$ and the 68% confidence interval of the SOMPZ realisations. The S/N is generally increased by the inclusion of the clustering-z information; in particular, the S/N is increased up to a factor of 3 in the relevant redshift range where $n(z)$ is substantially different from 0. In the same S/N panels of Fig. 6.2, we also show the contribution to the S/N increment due to RedMaGiC galaxies or BOSS/eBOSS galaxies alone. The latter sample mostly contributes in the redshift range $0.8 < z < 1.0$, whereas most of the clustering-z information at lower redshift comes from RedMaGiC galaxies.

The SOMPZ method has strong fine-scale fluctuations in $n_{\rm u}(z)$ due to sample variance on the small regions of sky used for its deep imaging and spec-

troscopy. The clustering-z correlation functions, on the other hand, are measured over the full DES Y3 footprint and have high S/N level. Although the clustering signal has a strong systematic uncertainty from the unknown WL bias, this systematic is slowly varying as a function of redshift and has less fine-scale fluctuations. The clustering-z likelihood is thus able to drive the $n_{\rm u}(z)$ outputs to a smooth distribution, at least over redshifts where clustering-z reference samples are available.

We remind the reader that the clustering information alone cannot be used to infer the full $n_{\rm u}(z)$, as the reference samples used in this work do not span the whole redshift range relevant for the DES Y3 $n_{\rm u}(z)$. Nonetheless, we can try to understand in simulations if the WZ constraints would be unbiased independently of the SOMPZ information. We did this by importance-sampling realisations of the true $n_{\rm u}(z)$s shifted around their mean redshift, and by assigning to each sample a weight through the likelihood given by Eq. 7.6. This test allowed us to recover the true $n_{\rm u}(z)$ within uncertainties, hence proving the method to be unbiased; for more details, see Appendix A.2.

Finally, we verify that the choices of the the parameters $\alpha_{\rm r}$ or the cosmology assumed to compute $w_{\rm DM}$ do not impact the methodology. We find that assuming different values for the cosmological parameters ($\Omega_{\rm m} = 0.4$, $\sigma_8 = 0.7$) results in a shift in $\Delta\langle z\rangle < 10^{-3}$ on the calibrated SOMPZ redshift distributions. Concerning magnification, in order to roughly asses the impact of the exact values of the magnification coefficients $\alpha_{\rm r}$, we verified that assuming values for $\alpha_{\rm r}$ that are $-1\times$ the fiducial ones resulted in shifts $\Delta\langle z\rangle < 10^{-3}$ (see Appendix A.1 for more details). We conclude that the full-shape likelihoods can be calculated in advance of and independent from the cosmology chains.

6.4 Results in data

We apply the clustering-z methods to DES Y3 data by first measuring the angle-averaged $w_{\rm ur}(z_i)$ (Eq. 5.16) of each WL source tomographic bin sample against the RedMaGiC and BOSS/eBOSS samples described in §4. These cross-correlation data are plotted in Fig. 6.3. Note the exceptionally high S/N level of the RedMaGiC data in particular, even at the rather fine binning of $\Delta z = 0.02$ that we use throughout. Bin-by-bin estimates of the reference bias $b_{\rm r}(z_i)$ are obtained using Eq. 5.20, with a dark-matter $w_{\rm DM}(z_i)$ predicted from theory for nominal cosmological parameters (Planck Collaboration, 2018).

Note that for the RedMaGiC galaxies we calculated $b_{\rm r}(z_i)$ applying the correction to the galaxy-matter bias of the reference sample described by Eq. 5.19, using the fraction of the RedMaGiC galaxies which have a spectroscopic redshift. As RedMaGiC galaxies with spec-z counterparts tend to have

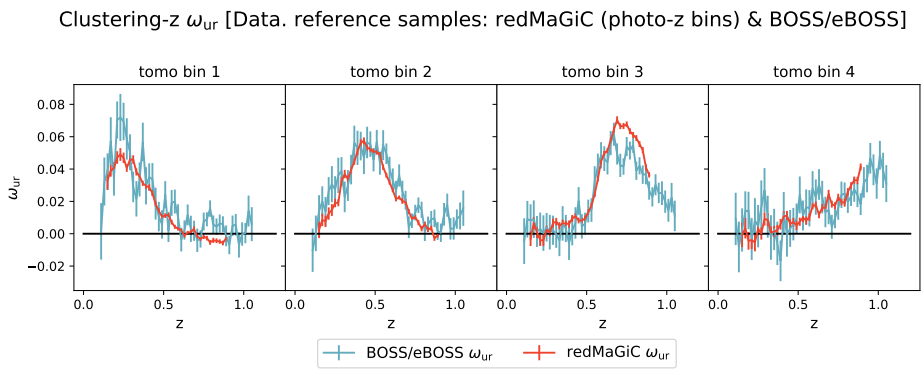

Figure 6.3: The measured $w_{\rm ur}(z)$ for the DES Y3 data are plotted for each of the four tomographic bins, using reference samples from BOSS/eBOSS (blue) and RedMaGiC (red).

brighter magnitudes compared to the full RedMaGiC sample, we have applied a magnitude re-weighting to those galaxies before computing the correction, so as to up-weigh (down-weigh) RedMaGiC galaxies under (over) represented in the spec-z subsample. After the re-weighting, the spec-z sample had the same magnitude distribution of the full RedMaGiC sample. Imperfections in this process should be small based on the tests in previous sections and are included in the systematic uncertainties of the two methods.

Last, we note that the RedMaGiC estimates show a small, negative tail at high redshift, for the first WL tomographic bin. We believe this is due to a systematic effect non corrected by our lens weights, rather than magnification, which should be positive at those redshifts, according to our estimates. The Balrog estimates of the magnification coefficients should also include realistic systematic and observational biases, which might lead to negative magnification; the fact that our estimates are nonetheless positive indicates that this effect is due to some systematic that affects the RedMaGiC number density and that anti-correlates with the WL density distribution. We know, indeed, that the RedMaGiC sample is affected by some residual systematics, which does not affect cosmology (Abbott et al., 2022; Pandey et al., 2021), but manifests as a scale-, redshift- and sky-area-independent phenomenological decorrelation parameter. Given the small amplitude of this effect, the fact that we also have constraints from another independent sample (BOSS/eBOSS), and that our clustering-z constraints are compatible with SOMPZ and shear-ratio (Sánchez et al., 2022b) prior to combination, we believe this should have a negligible impact on our results.

Following the procedure used on the simulations, we define the WZ likeli-

hood using Eqs. 7.5 and 7.6. We assume fiducial values for the magnification parameters for the RedMaGiC sample, as estimated using Balrog (Suchyta et al., 2016; Everett et al., 2020a). We do not have an estimate of the magnification parameters for BOSS/eBOSS galaxies available, so we assumed the same values used for RedMaGiC galaxies. We confirm, however, that assuming values for the magnification parameters that are $-1\times$ the fiducial ones resulted in no relevant effect on the mean of the resultant redshift distributions. The nuisance-parameter priors derived from simulations in §7.2.2 are used, including those specifying the allowed variation with z in $b_{\rm u}(z)$ and other elements of the Sys$(z, \mathbf{s})$ function.

Before applying the method, we checked that the fiducial $\hat{w}_{\rm ur}$ model on data (obtained using SOMPZ $n_{\rm u}(z)$ as baseline) was compatible with the measured $w_{\rm ur}$ marginalised over the systematic function Sys(z). This check has been performed separately for RedMaGiC and BOSS/eBOSS. We then use the Hamiltonian Monte Carlo method to draw samples from the joint posterior distribution of the SOMPZ likelihood and the clustering-z likelihoods for both RedMaGiC and BOSS/eBOSS data. Fig. 6.4 show the 68% confidence interval of the $n_{\rm u}(z)$ samples from the SOMPZ+WZ posterior, as well as those from the pure SOMPZ posterior. At redshifts where clustering-z information is available, it greatly reduces the point-by-point uncertainties in $n_{\rm u}(z)$, just as in the simulations. The WZ method is thus very successful at reducing the impact sample variance on SOMPZ estimators. This combined estimator also shows no sign of negative tail at high redshift in the first tomographic bin (as seen, instead, in the clustering measurement, Fig. 6.3). This stresses the importance of a combined analysis, which is more robust and it is able to remove some of the potential problems or systematics affecting each of the two estimators when used individually.

The averages and standard deviations of the mean-z distributions of the SOMPZ and SOMPZ+WZ posteriors are listed in Table 6.2. As expected from the simulations, the clustering-z information does not substantially alter the bin means derived from photo-z methods. The significant improvement in shape accuracy, as seen in Fig. 6.4, is the principal product of the clustering-z method for DES Y3 analyses.

6.5 Conclusions

In this Chapter we have presented part of the calibration of the DES Y3 weak lensing source galaxies redshift distributions. The overall strategy involves three methods: a photo-z method (SOMPZ), a clustering based method (WZ), and shear ratios, based on ratios of galaxy-galaxy lensing measurements. Here

Table 6.1: **Simulations**. The mean redshift estimates of the SOMPZ distributions with and without clustering-z information, in simulations.

case	**tomo bin 1**	**tomo bin 2**	**tomo bin 3**	**tomo bin 4**
True $\langle z \rangle$:	0.315	0.513	0.743	0.910
SOMPZ $\langle z \rangle$:	0.312 ± 0.008	0.505 ± 0.005	0.746 ± 0.003	0.907 ± 0.005
SOMPZ + WZ:	0.312 ± 0.009	0.507 ± 0.005	0.747 ± 0.004	0.907 ± 0.005

Table 6.2: **Data**. The mean redshift estimates of the SOMPZ distributions with and without clustering-z information.

case	**tomo bin 1**	**tomo bin 2**	**tomo bin 3**	**tomo bin 4**
SOMPZ $\langle z \rangle$:	0.318 ± 0.009	0.513 ± 0.006	0.750 ± 0.005	0.942 ± 0.011
SOMPZ + WZ :	0.321 ± 0.008	0.517 ± 0.006	0.749 ± 0.005	0.940 ± 0.010

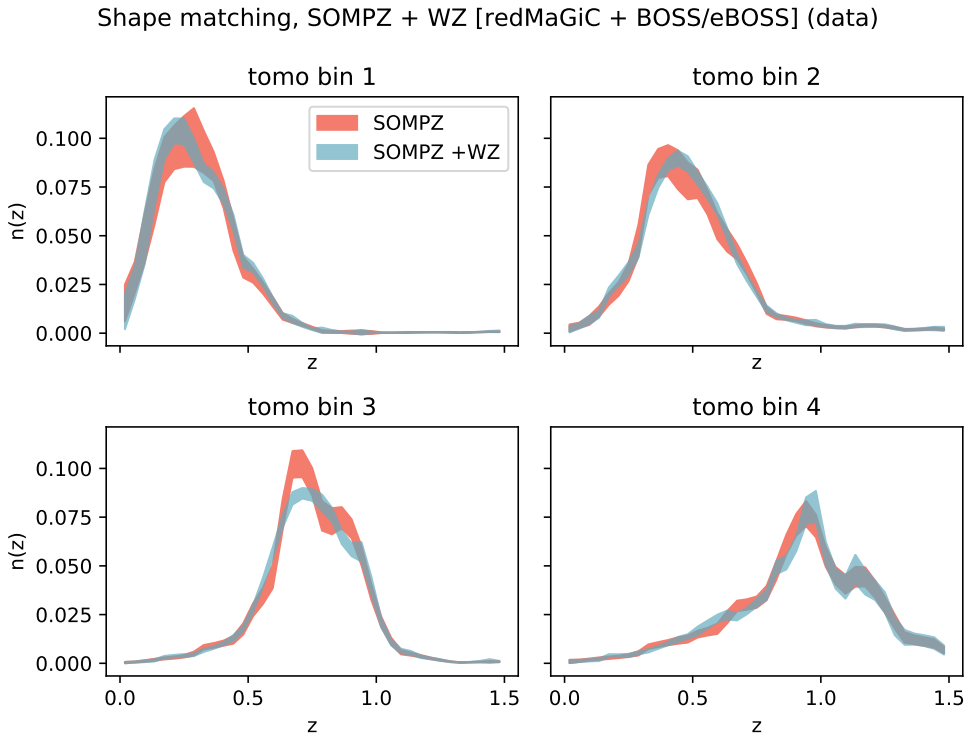

Figure 6.4: SOMPZ redshift distributions, as estimated in data, with and without clustering information. The bands encompass the statistical and systematic uncertainties of the distributions.

we focused on the WZ part of the calibration and on its combination with the SOMPZ method.

The WZ measurements for the calibration have been obtained cross-correlating the weak lensing sample with two redshift catalogs with accurate redshifts: the RedMaGiC galaxies (luminous red galaxies with excellent photometric redshifts) and the BOSS/eBOSS galaxies (with spectroscopic estimates). The cross-correlation have been performed dividing the reference samples into thin redshift bins. We developed a procedure to forward model the clustering signal given a proposal redshift distribution. We characterised the systematic uncertainty of our forward model using simulations, and found that it is limited by the lack of prior knowledge of the redshift evolution of the galaxy-matter bias of the weak lensing sample. We developed a flexible model able to take into account any realistic redshift evolution of the galaxy-matter bias, and used the simulations to inform the priors of such a model.

We then developed a framework which allowed us to sample proposal redshift distributions from the joint likelihood of WZ and SOMPZ conditioned on our measurements, and compared the samples with the ones obtained from the SOMPZ likelihood alone. We found that for the DES Y3 analysis, the clustering information does not help much reducing the scatter on the mean of the redshift distributions compared to SOMPZ, as the latter method performance is superior. Nonetheless, the WZ likelihood does significantly improve the uncertainty on the shape of the redshift distributions, reducing the scatter in the shape compared to the SOMPZ only samples. This is due to the fact that the SOMPZ realisations suffer from cosmic variance, as they are estimated from relatively small deep fields; as a consequence, SOMPZ realisations show the characteristic "peaky" features typical of sample variance, which are reduced by the combination with the clustering measurements. We then presented the WZ+SOMPZ calibration on data, finding consistent results with our simulation tests.

We remind the reader that the final DES Y3 redshift calibration strategy includes an additional step, not addressed here: the samples of the WZ+SOMPZ likelihood, describing the redshift distributions of the WL sample and their uncertainties, are further constrained by the shear-ratio likelihood during the cosmological inference process.

Although current WZ methods and implementations are still affected by systematics and are not simply sample variance limited, they do represent one of the most credible alternative to standard photo-z methods for the next generation of data sets. This is because the new data will be deeper, which will make the calibration of standard photo-z methods significantly harder, whereas WZ methods will be less affected (they only need a proper redshift

coverage by the reference sample). The main WZ systematic that requires additional work is the modelling of the evolution of the galaxy-matter bias of the unknown sample. For future works, we plan to use the auto-correlation of the unknown sample in small deep fields with good photometry and redshift estimates to inform better the parameters of our flexible galaxy-matter bias model. Moreover, for the future analyses, we also plan to combine the WZ and photo-z estimates in a more principled way, using the hierarchical Bayesian methodology developed by Sánchez & Bernstein (2019b) and Alarcon et al. (2019).

Chapter 7

Combination of SOMPZ and WZ: redshift calibration of the DES Y3 MagLim lens galaxies

7.1 Introduction

For the DES Y3 3x2pt analysis, two different lens samples were used. The first sample is defined by selecting luminous red galaxies through the RedMaGiC algorithm (Rozo et al., 2016), which retains galaxies with high quality photometric redshift, by fitting each galaxy to a red-sequence template. The galaxies passing the RedMaGiC selection have, however, a low number density, and the final sample comprises roughly 3,000,000 galaxies. The second sample slightly compromises on the redshift accuracy to the benefit of a larger number density. The MagLim sample (Porredon et al., 2021c) is a magnitude-limited sample with a number density more than 3 times greater than RedMaGiC , comprising roughly 10,000,000 galaxies. In the fiducial DES 3x2pt (Abbott et al., 2022) and 2x2pt analyses Porredon et al. (2021a) that rely on the MagLim sample, the redshift distributions of the sample have been characterised using the machine learning photometric redshift code Directional Neighbourhood Fitting (DNF) (De Vicente et al., 2016b). In its current implementation, the DNF code provides per-galaxy redshift estimates using nearest neighbour techniques. The redshift distributions were then further calibrated using clustering redshift (WZ), which relies on cross-correlation measurements with spectroscopic samples (Cawthon et al., 2020). This calibration step also placed uncertainties on the redshift distribution estimates, which were modelled by "shifting" and "stretching" the redshift distributions.

This work presents an additional and more sophisticated calibration of the redshift distributions of the lens sample, and studies the impact of these new redshift distribution estimates on the cosmological constraints using DES Y3 galaxy clustering and galaxy-galaxy lensing measurements (2x2pt). In particular, we adopt an approach similar to the one adopted to characterise the redshift distributions of the DES Y3 weak lensing (WL) sample, presented in Chapter 6. This methodology also combines photometric and clustering constraints to produce redshift estimates, and it is more powerful than the fiducial redshift calibration adopted for the lenses for a number of reasons. The photometric information is used to produce redshift estimates using a self-organizing-map-based scheme (SOMPZ), which allows a meticulous control over all the (known) potential sources of uncertainties affecting the estimates. The SOMPZ method works by leveraging the DES deep fields, which have deeper observations with additional photometric bands and overlap with many-band redshift surveys available. It is possible to reproduce realistic selection functions in the deep fields from the injection of galaxies into actual DES images using the sophisticated image simulation tool Balrog (Everett et al., 2020b).

The chapter is organised as follows. Section 7.2 is devoted to the characterisation of the method's uncertainties. Section **??** presents the redshift distributions MagLim sample produced using the techniques described in this work. Appendix B.1 complements the results with a validation of the methodology in simulations; Appendix B.3 discusses the impact of different redshift uncertainties marginalisation techniques on the cosmological parameters estimation.

7.2 Characterization of sources of uncertainty

In this section, we present the characterisation of the systematic uncertainties of our methodology. The dominant sources of uncertainties for the SOMPZ method are sample variance and shot noise. In the clustering redshift method, the main uncertainty is caused by the lack of prior knowledge on the redshift evolution of the galaxy-matter bias of the MagLim sample. This is modelled by a flexible systematic function, informed by a measurement of the MagLim auto-correlation function in data. Other, minor sources of uncertainties are related to magnification effects and the approximation of linear bias (Gatti & Giannini et al., 2020a). We provide further details on each source of uncertainty in the following subsections. A full catalog-to-cosmology validation of the method (in simulations) is then presented in Appendix B.1.

7.2.1 SOMPZ uncertainties

For the SOMPZ method we consider the following sources of uncertainty:

- *sample variance of the deep fields*: main uncertainty, caused by the limited area of the deep fields. We model the effect of sample variance by means of the 3sDir analytical model described in §7.2.1;
- *shot noise in the deep and redshift samples*: this is induced by the limited number of galaxies available in the deep and redshift samples. We model the effect of shot noise by means of the 3sDir analytical model described in §7.2.1;
- *SOMPZ method uncertainty*: this uncertainty stems from discretising the color space in the SOMPZ mapping. We do estimate its impact on the SOMPZ estimates by replicating the SOMPZ methods multiple times in simulations, and incorporate its effects by using Probability Integral Transforms (PITs) (§ 7.2.1);
- *photometric calibration*: related to uncertainties in the calibration of the deep fields zeropoint, it is accounted for in the SOMPZ estimates by means of PITs (§ 7.2.1).
- *redshift sample biases*: these biases stem from uncertainties and biases in the redshift estimates of the redshift samples. Their impact is accounted for in our methodology by marginalising over three different combinations of redshift samples (§ 7.2.1);
- *transfer function*: any bias induced by an erroneous estimation of the transfer function due to a size-limited Balrog sample; we anticipate this to be negligible following the results from Myles & Alarcon et al. (2020) (§ 7.2.1).

In the following sections we will proceed to describe in detail how we account for each of the items listed above.

Mean

Uncertainty	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6
	$0.2 < z < 0.4$	$0.4 < z < 0.55$	$0.55 < z < 0.7$	$0.7 < z < 0.85$	$0.85 < z < 0.95$	$0.95 < z < 1.05$
Sample Variance & shot noise	0.015	0.010	0.010	0.008	0.009	0.009
SOMPZ method	0.004	0.003	0.005	0.001	0.007	0.005
Redshift samples	0.009	0.001	0.006	0.003	0.004	0.007
Zeropoint	0.008	0.007	0.004	0.005	0.005	0.005
SOMPZ	0.315 ± 0.015	0.445 ± 0.010	0.630 ± 0.010	0.776 ± 0.008	0.895 ± 0.009	0.983 ± 0.012
SOMPZ+WZ	0.316 ± 0.014	0.456 ± 0.008	0.632 ± 0.008	0.780 ± 0.007	0.893 ± 0.008	0.985 ± 0.010
SOMPZ (with all unc)	0.317 ± 0.020	0.447 ± 0.012	0.634 ± 0.013	0.778 ± 0.010	0.897 ± 0.011	0.988 ± 0.013
SOMPZ+WZ (with all unc)	0.315 ± 0.016	0.463 ± 0.010	0.633 ± 0.009	0.781 ± 0.008	0.893 ± 0.009	0.990 ± 0.012

Width

Sample variance & shot noise	0.007	0.005	0.003	0.003	0.004	0.009
SOMPZ method	0.003	0.003	0.0007	0.0003	0.002	0.0001
Redshift samples	0.001	0.005	0.0007	0.0006	0.0003	0.001
Zeropoint	0.003	0.004	0.001	0.0004	0.001	0.001
SOMPZ	0.077 ± 0.007	0.093 ± 0.007	0.065 ± 0.004	0.081 ± 0.004	0.071 ± 0.004	0.096 ± 0.009
SOMPZ + WZ	0.080 ± 0.004	0.089 ± 0.004	0.060 ± 0.002	0.077 ± 0.003	0.074 ± 0.004	0.105 ± 0.006
SOMPZ (with all unc)	0.081 ± 0.008	0.096 ± 0.007	0.067 ± 0.005	0.081 ± 0.004	0.073 ± 0.005	0.098 ± 0.009
SOMPZ + WZ (with all unc)	0.080 ± 0.005	0.081 ± 0.005	0.060 ± 0.002	0.073 ± 0.003	0.074 ± 0.004	0.101 ± 0.007

Table 7.1: Summary of values for systematic uncertainties and center values for mean (top panel) and width (bottom panel) for the $n(z)$ distributions. The various components are computed as described in section 7.2 and as they are not completely independent it is not expected that they sum up to the total value. The values related to SOMPZ and SOMPZ+WZ refer to Figure 7.3, and include only the 3sDir uncertainty due to sample variance and shot noise (and the redshift samples uncertainty), because it was logistically not possible to add the SOMPZ method and the zeropoint sources of uncertainty before the combination with WZ. As a comparison, the "SOMPZ (with all unc)" includes all uncertainties. The final $n(z)$ which has been used in the cosmological analysis is the bottom line.

Sample variance and shot noise (*3sDir*)

Sample variance is the dominant uncertainty affecting our SOMPZ estimates, and stems from the limited size and area coverage of the redshift and deep samples, with respect to the whole wide field. The deep fields only cover $\sim 9deg^2$, which means we could be learning the color/redshift relation from a non-representative sample of the sky due to fluctuations in the matter density field; moreover, the finite size of the redshift sample can introduce shot noise effects, preventing a correct sampling of the quantities required for the redshift inference.

Generally the impact of sample variance can be evaluated estimating the redshift distributions in simulations multiple times using different line of sights for the deep fields (e.g. Hildebrandt et al. 2017, Hildebrandt et al. 2020; Hoyle et al. 2018b; Buchs et al. 2019; Wright et al. 2020a). Although we also performed a test where we evaluated the impact of sample variance using the Buzzard simulation, in our standard procedure we use the *three step Dirichlet* (3sDir) approach *3sDir* presented in Sánchez et al. (2020) and applied to the redshift calibration of the DES Year 3 source sample (Myles & Alarcon et al., 2020).

The 3sDir method consists of an analytical sample variance model predicting what the redshift-color distribution would be from the observed individual redshift and galaxy phenotypes (colors) of galaxies coming from smaller deep fields. Using this model we can build an ensemble of redshift distributions (also called $n(z)$) realisations whose fluctuations realistically represent the effect of sample variance. During the cosmological inference, by sampling over these realisations, one can effectively marginalise over the effect of sample variance. Here we provide a short description of the 3sDir method, but we direct the reader to Myles & Alarcon et al. (2020) and Sánchez et al. (2020) for more details. The 3sDir method assumes the probability $p(z, c)$ that galaxies belong to a redshift bin z and color phenotype c to be described by a probability histogram with coefficients f_{zc} (with $\sum f_{zc} = 1$ and $0 \leq f_{zc} \leq 1$). Under this assumption, the expected number counts of galaxies in a deep SOM cell given the coefficients f_{zc} are described by multinomial distribution; if we assume a Dirichlet function for the prior on f_{zc}, the posterior of f_{zc} given the observed number count will also be described by a Dirichlet function. Such a Dirichlet posterior can be used to draw samples and naturally accounts for the effect of shot noise in the data. The effect of sample variance can be introduced by tuning the width of the prior on f_{zc}, which does not change the expected value for f_{zc} in the Dirichlet distribution, but does change its variance to simultaneously account for shot noise and sample variance.

If all the galaxies belonging to the redshift sample were independently drawn, then a Dirichlet distribution parametrized by the redshift sample counts in each couple of redshift bin z and phenotype c, N_{zc}, would fully characterize f_{zc}. However, one subtlety is that sample variance correlates with redshifts; to increase the variance with the correct redshift dependence one can use the fact that two different phenotypes (deep SOM cells) overlapping in redshift are correlated due to the same underlying large-scale structure fluctuations. The 3sDir model assumes that phenotypes at the same redshift share the same sample variance, and therefore groups cells with similar redshifts in *superphenotypes* T. One can then express the f_{zc} as:

$$f_{zc} = \sum f_c^{zT} f_z^T f_T. \tag{7.1}$$

The 3sDir method consists of drawing values of these three sets of coefficients with three Dirichlet functions. In this way, it is possible to include a redshift-dependent variance while conserving the expected value of f_{zc}.

The validation of the *3sDir* method has been carried out in Myles & Alarcon et al. (2020), applied to the weak lensing source sample. The only difference with this work stands in the fact we are performing the 3sDir estimation independently for each tomographic bin, due to their definition.

As reported in Table 7.1, this uncertainty is dominant, both on the mean and width values of the $n(z)$ distributions, computed from the ensemble of realisations provided by the *3sDir* method.

SOMPZ Method Uncertainty

The SOMPZ method relies on the discretisation on the color space spanned by our deep field sample, and this is an approximation that can lead to small biases or additional uncertainties. In order to estimate these, we compute our SOMPZ $n(z)$ a large number of times in the Buzzard simulations. In order to factor out sample variance, each time we randomly select patches of the Buzzard footprint to construct the mock deep fields. In this way, by averaging over all the final $n(z)$ realisations, we can produce an estimate of the $n(z)$ only minimally biased by sample variance, and test the agreement with the true $n(z)$. Due to the computational cost of the SOMPZ pipeline, we decided to produce 300 $n(z)$ replicas. To perform this test, we assumed that the redshift sample would only be limited to one of our four fields, of the size of COSMOS.

We computed the mean redshift offset of the ensemble with respect to the true value, for each tomographic bin. As reported in Table 7.1, these values are smaller than the effect of sample variance. These values are incorporated into our final $n(z)$ ensemble using the PIT method described in the following

section, by additionally shifting each probability integral transform (used to correct for the zeropoint uncertainties) by a value drawn from a Gaussian centered at zero with standard deviation equal to the root-mean-square of the aforementioned mean offset values.

Deep Fields Photometric Calibration Uncertainty

Although the uncertainty in the photometry of each individual galaxy is implicitly accounted for in the SOM training, the uncertainty on the photometric calibrations as a whole must be evaluated by testing how the measured $n(z)$ are affected by changes in the photometric zeropoint in each band. This is relevant for the deep fields, where the relatively precise fluxes are key to constraining reliable $p(z)$ in parts of parameter space that are not subject to selection biases. Ideally, this would be tested by rerunning the full analysis for an ensemble of perturbations of the photometric zeropoint according to the zeropoint uncertainty, but the computational requirements of the Balrog injection procedure make this infeasible. Instead, we produce an analogous ensemble of realizations in simulations, where the Balrog mock photometric survey is reduced to a computationally simpler procedure of adding Gaussian noise to true magnitudes. For each realization of this ensemble, we perturb all deep field magnitudes by a draw from a Gaussian whose width is determined by the photometric zeropoint uncertainty in the Y3 deep fields catalog in a specified band, as computed in Hartley et al. (2020). We then "inject" these perturbed deep field fluxes with a mock Balrog procedure to generate wide field realizations of the galaxies and measure the corresponding $n(z)$. In this way we generate a full ensemble of $n(z)$ realisations reflecting the uncertainty in our redshift calibration due to the photometric calibration. We apply Probability Integral Transforms (PITs) as in Myles & Alarcon et al. (2020) to transfer the variation encoded in the ensemble from simulated $n(z)$ to our fiducial data result.

Redshift Sample uncertainty

As mentioned in § 4.7, we decided to choose three different catalogs to infer our redshift distributions from: a collection of spectroscopic surveys galaxies (Gschwend et al., 2018), PAU+COSMOS redshift as in Alarcon et al. (2020), and COSMOS30 photometric redshifts (Laigle et al., 2016). The reason for availing ourselves of more than one catalog lies in the fact neither of these are exempt from systematic uncertainties: each survey uses different photometry, different model assumptions, and can be affected systematically by selection

effects, incorrect templates, photometric outliers, etc. Since there is a considerable overlap in the number of galaxies belonging to more than one of the redshift catalogs selected for this work, to account for the intrinsic biases we decided to build three samples which are combinations of the aforementioned catalogs. We ranked the redshift catalogs differently for each sample: if a galaxy has information from multiple origins, we assign the redshift from the highest ranked catalog. The three redshift samples *SPC, PC, SC*, are described in Section 4.7.

For each of these, we will perform the complete pipeline, and the final set of realisation will be constructed by an equal fraction $p(R) = 1/3$ from each survey. By placing equal prior probability to each sample, this is equivalent as saying that we do not believe any of the samples is more likely to be correct. But note that for galaxies from which we have information from only one catalog, we are assuming that information to be true, and this is a caveat of this approach.

Transfer function uncertainty

One of the key points in this redshift calibration is the transfer function $p(c|\hat{c})$, the intermediate step necessary to assign redshifts from deep field galaxies to the whole wide field. If the transfer function is inaccurate, regardless of how a precise the color/redshift characterisation is in the deep SOM, it can bias the final $n(z)$ distributions. $p(c|\hat{c})$ depends on the observation conditions in that location, determining if the galaxy is detected or not. Observing conditions vary across the wide field, but for our analysis we are interested in redshift distributions estimated across all the footprint. Balrog injects the same deep galaxies in random wide tiles, but these cover only around $\sim 20\%$ of the DES footprint, but in Myles & Alarcon et al. (2020) they verified that Balrog is adequately sampling the observing conditions in the wide field. They boostrapped the sample by the injected position and recomputed 1000 different transfer functions. They concluded that the dispersion in the final $n(z)$ mean redshift from repeating the analysis using each time a different transfer function was completely negligible. Here we repeated that test, since our deep field sample has less galaxies and might impact differently the transfer function.We found that this is also negligible for our case, with variations on the $n(z)$ mean $< 10^{-3}$ in each tomographic bin, and therefore decided not to propagate this in the final $n(z)$ estimate.

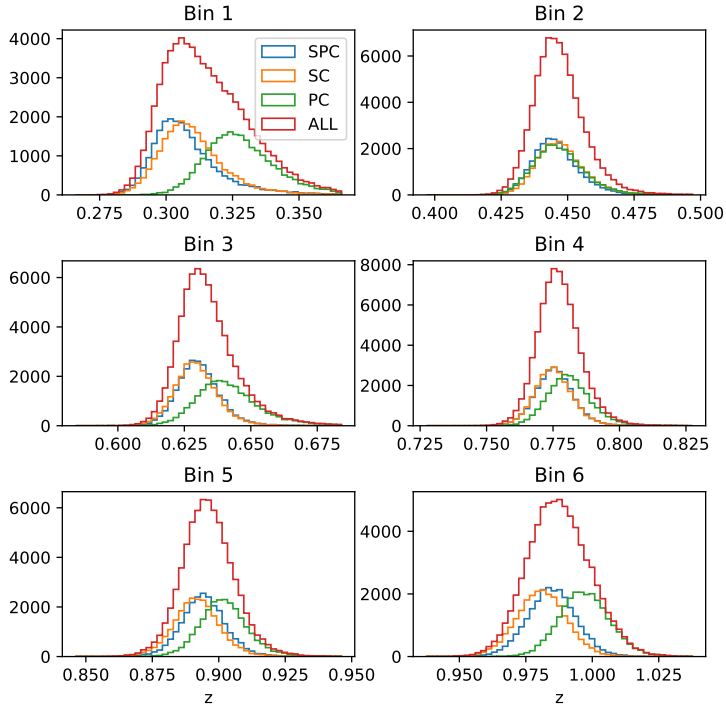

Figure 7.1: Uncertainty on the mean redshift of the three redshift samples: SPC (prioritizes spectra, than PAU photo-z, then COSMOS30), PC (prioritizes PAU photo-z, then COSMOS30) and SC (prioritizes spectra, then COSMOS30). In red the total uncertainty given by their combination.

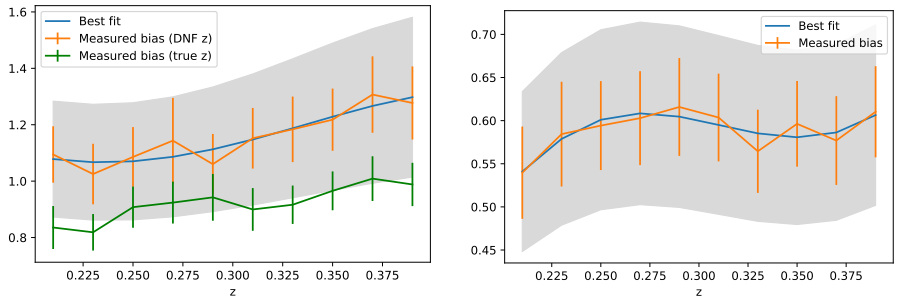

Figure 7.2: *Left panel*: galaxy-matter bias of the MagLim sample as estimated in simulation, dividing the sample into thin bins using the true redshifts (orange points) and divided using the DNF redshift estimates and the methodology outlined in Section 7.2.2. The grey band encompasses the 68% confidence interval of the $\mathrm{Sys}(z, \mathbf{s})$ function. *Right panel*: galaxy-matter bias of the MagLim sample as measured from the data (orange points); the blue line shows the best-fitting $\mathrm{Sys}(z, \mathbf{s})$ function, and the grey band encompasses its 68% confidence interval.

7.2.2 WZ Uncertainties

The WZ systematic uncertainties have been identified and characterised in detail for the WL sample in Gatti & Giannini et al. (2020a). Namely, the systematic budget was found to be dominated by our lack of prior knowledge of the redshift evolution of the galaxy-matter bias of the unknown sample. This is also expected to be the case for the MagLim sample, although the amplitude of the effect might differ from the WL sample (ideally, since the MagLim redshift distributions are narrower, we might expect a smaller impact due to systematics slowly varying with redshift like the galaxy-matter bias).

Similarly to Gatti & Giannini et al. (2020a), we model our systematics by means of a flexible function, $\mathrm{Sys}(z, \mathbf{s})$, which mostly captures the redshift evolution of the galaxy-matter of the unknown sample. The $\mathrm{Sys}(z, \mathbf{s})$ function is parameterized by $\mathbf{s} = \{s_1, s_2, \ldots\}$ that we will marginalize over and is given by:

$$log[Sys\,(z, \mathbf{s})] = \sum_{k=0}^{M} \frac{\sqrt{2k+1}}{0.85} s_k P_k(u), \tag{7.2}$$

$$u \equiv 0.85 \frac{z - 0.5(z_{\max} + z_{\min})}{(z_{\max} - z_{\min})/2}, \tag{7.3}$$

with $P_k(z_i)$ being the k-th Legendre polynomial and $M = 6$ is the maximum order. In this work, we set the prior $p(\mathbf{s})$ to be a simple diagonal normal distribution, with the standard deviations $\{\sigma_{s0}, \ldots, \sigma_{sM}\}$ and means informed by the measured auto-correlation of the MagLim sample.

In Gatti & Giannini et al. (2020a), such a systematic function was let to vary by the typical amplitude of the redshift evolution of the galaxy-matter bias of the WL sample we measured in simulations. In practice, this was achieved by imposing a Gaussian prior with zero mean $p(\mathbf{s})$ on the coefficients $\mathbf{s}$ of the systematic function.

In the case of the MagLim sample, we can use a more informative prior $p(\mathbf{s})$ that uses the information we have from the data about the galaxy-matter bias evolution of the sample. In particular, we rely on the fact that the MagLim sample has good per-galaxy redshift estimates, which allows us to divide the sample in relatively small bins and measure the auto-correlation of such bins. This was not possible for WL sample, due to the poor per-galaxy redshift accuracy.

To this aim, we use DNF 1-point estimates z_{mean} to further divide the MagLim sample in bins of width of $\Delta_z = 0.02$, and we measure the auto-correlation of each bin. We note that the true width of each bin will be much

larger than $\Delta_z = 0.02$, as the DNF photo-z are uncertain. Under the approximation of negligible redshift evolution of the galaxy-matter bias of the MagLim sample over each thin bin, the measured autocorrelaton can be related to the galaxy-matter bias by knowing how broad the true $n(z)$ distribution of each bin is (Gatti et al., 2018; Cawthon et al., 2020):

$$w_{\rm uu}(z_i) = b^2_{\rm u}(z_i) w_{\rm DM}(z_i) \int dz' n^2_{\rm u,i}(z'), \tag{7.4}$$

where $n_{\rm u,i}(z')$ is indeed the true distribution of the thin bin MagLim sample. Such a quantity is estimated using the PDF estimate from DNF $z_{\rm PDF}$.

From this measurement performed in data we can then retrieve the galaxy bias $b_{\rm u}(z)$. We fit the ${\rm Sys}(z, \mathbf{s})$ function presented in Eq. 7.2 to the measured $b_{\rm u}(z)$ and obtain best-fit s values, which we show in Figure 7.2. These best-fit coefficients are then used as the mean value of the Gaussian prior $p(\mathbf{s})$. The best fitting ${\rm Sys}(z, \mathbf{s})$ function to the data is shown in the right panel of Fig. 7.2.

To estimate the width of the prior p($\mathbf{s}$) we took a different approach. First, we estimate the bias evolution in simulations by dividing galaxies into thin redshift bins using: (i) the true redshifts from the simulation; and (ii) the photo-z estimated from the DNF code. When dividing the galaxies with the photo-z from DNF, we further correct the measured auto-correlation using Equation 7.4. These measurements are shown in the left panel of Figure 7.2. The discrepancy between the measured bias evolution from photo-z (equivalent to the application with real data) relative to the measured bias evolution with true redshifts (equivalent to the truth) is a systematic bias. We use the sum in quadrature of this difference with the statistical uncertainty of the bias measurement as the prior width of s_0. For the higher order parameters we estimate the standard deviation of the prior by summing in quadrature the ratio between the two biases and the statistical uncertainty from the bias measurement in data. This allows to best capture the RMS variations of the bias function itself. As can be seen in Fig. 7.2, the 68% confidence interval spanned by the ${\rm Sys}(z, \mathbf{s})$ function both brackets the ideal and real world measurements. The values for the priors are displayed in Table 7.2. Both the prior on the 0-th and higher order coefficients are much tighter than in Myles & Alarcon et al. (2020), were $s_0 = 0.6$ and $s_{1..4} = 0.15$. As already explained, the difference lies in the initial accuracy of the photo-z estimates, that enables the measurement of the auto-correlation of the galaxy sample in thin redshift bins. For the weak lensing source sample such information was not available, and therefore a more conservative prior was deemed appropriate. In the MagLim sample case instead, the greater accuracy on its photo-z allows to extract more information

	Bin 1	Bin 2	Bin 3	Bin 4	Bin 5	Bin 6
s_0	0.1072	0.2155	0.1233	0.0724	0.0636	0.1976
$s_{1..5}$	0.0286	0.0527	0.0407	0.0517	0.0806	0.0436

Table 7.2: Coefficients of the systematic function $\mathrm{Sys}(z, \mathbf{s})$

from the auto-correlation.

Last, we mention that an additional source of uncertainties for the WZ measurement is related to the impact of magnification. We do model magnification effects, but the accuracy of that model is limited by our knowledge of the magnification coefficients for the two sample. In particular, we do not have any prior knowledge of such a coefficients for the BOSS/eBOSS sample. Those coefficients are set to 0 for our fiducial analysis (on the contrary, estimates for the magnification coefficient of the MagLim sample are available). We expect magnification to have a small impact, based on tests performed in Gatti & Giannini et al. (2020a), but we nonetheless test in the following section the impact of having a non null magnification coefficient for the BOSS/eBOSS sample.

7.2.3 Combination of SOMPZ and WZ

In order to combine SOMPZ and WZ constraints, we follow Gatti & Giannini et al. (2020a) and write the clustering likelihood by forward modelling the full clustering signal as function of the SOMPZ redshift distributions estimates $n(z)_{\mathrm{pz}}$. Moreover, we include the systematic function $\mathrm{Sys}(z, \mathbf{s})$ introduced in the previous section, which describes the uncertainties on the WZ measurement, mostly driven by the lack of knowledge of b_{u} and its redshift dependence:

$$\hat{w}_{\mathrm{ur}}(z_i) = n(z)_{\mathrm{pz}}(z_i) b_{\mathrm{r}}(z_i) w_{\mathrm{DM}}(z_i) \times Sys(z_i, \mathbf{s}) + M(\alpha_{\mathrm{u}}, \alpha_{\mathrm{r}}, b_{\mathrm{u}}, n(z)_{\mathrm{pz}}). \quad (7.5)$$

In the above equation, the quantities $\alpha_{\mathrm{u}}(z_i)$ and $\alpha_{\mathrm{r}}(z_i)$ are the magnification coefficients for the unknown and reference samples. See Gatti & Giannini et al. 2020a for full description of the magnification term M. The clustering of dark matter $w_{\mathrm{DM}}(z_i)$ is estimated from theory assuming fixed cosmology. We tested that varying cosmology has a negligible impact on our methodology.

The likelihood of the WZ data conditioned on the target $n(z)$ and all the

systematic parameters reads as:

$$\mathcal{L}\left[\mathrm{WZ}|n_{\mathrm{u}}(z), b_{\mathrm{r}}(z), \alpha_{\mathrm{r}}(z), w_{\mathrm{DM}}(z)\right] \propto \int d\mathbf{s}\, d\mathbf{p}\, \exp\left[-\frac{1}{2}(w_{\mathrm{ur}} - \hat{w}_{\mathrm{ur}})^T \Sigma_w^{-1} (w_{\mathrm{ur}} - \hat{w}_{\mathrm{ur}})\right] p(\mathbf{s}) p(\mathbf{p}), \quad (7.6)$$

were Σ_w is the clustering covariance, estimated through jackknife, and $\mathbf{p} = b_{\mathrm{u}}, \alpha_{\mathrm{u}}$. We implemented a Hamiltonian Monte Carlo sampler (HMC) that simultaneously samples the SOMPZ and WZ likelihood. The HMC does directly take as input the SOMs output of the sample variance estimation (described in 7.2.1), and it perturbs selectively the number counts in the SOMs in such a way to produce realisations that are already more likely to match the clustering redshift data.

7.3 Results in Data

In this section, we present the final redshift distributions for the MagLim sample as obtained in data. We also compare the SOMPZ+WZ redshift distributions with the fiducial DNF+WZ estimates used for the same sample and adopted in the cosmological analysis presented in Porredon et al. (2021a). A complete validation of the method in simulations is presented in Appendix B.1.

We first compare in Figure 7.3 the redshift estimates obtained using the 3sDir method and the estimates obtained including the WZ information as described in section 7.2. Due to logistics, the combination of the two methods was performed before incorporating the SOMPZ and zeropoint errors. As here we are just displaying the effect of the combination, we are showing only how the 3sDir uncertainty from sample variance and shot noise (from the three redshift samples) varies once we add the information from WZ. The combination of the two methods result in stronger constraints on the shape of the $n(z)$, thanks to the complementarity in the information provided by each SOMPZ and WZ. Particularly, the WZ signal strongly correlates across adjacent bins, excluding large portions of possible $n(z)$ shapes allowed by the SOMPZ likelihood alone, which are affected by sample variance fluctuations from the small calibration fields, and resulting in a smoother distribution. The improvement on the uncertainty on the mean is more modest, but not null, as reported in Table 7.1. Usually, WZ data provides limited information on the mean redshift, especially compared to SOMPZ, as the systematic uncertainty on the galaxy bias evolution of the target sample is large and directly degenerate with the mean redshift, as is the case in Gatti & Giannini et al. (2020a). However, in this work we have included a tighter prior on the $\mathrm{Sys}(z, \mathbf{s})$ function describing the

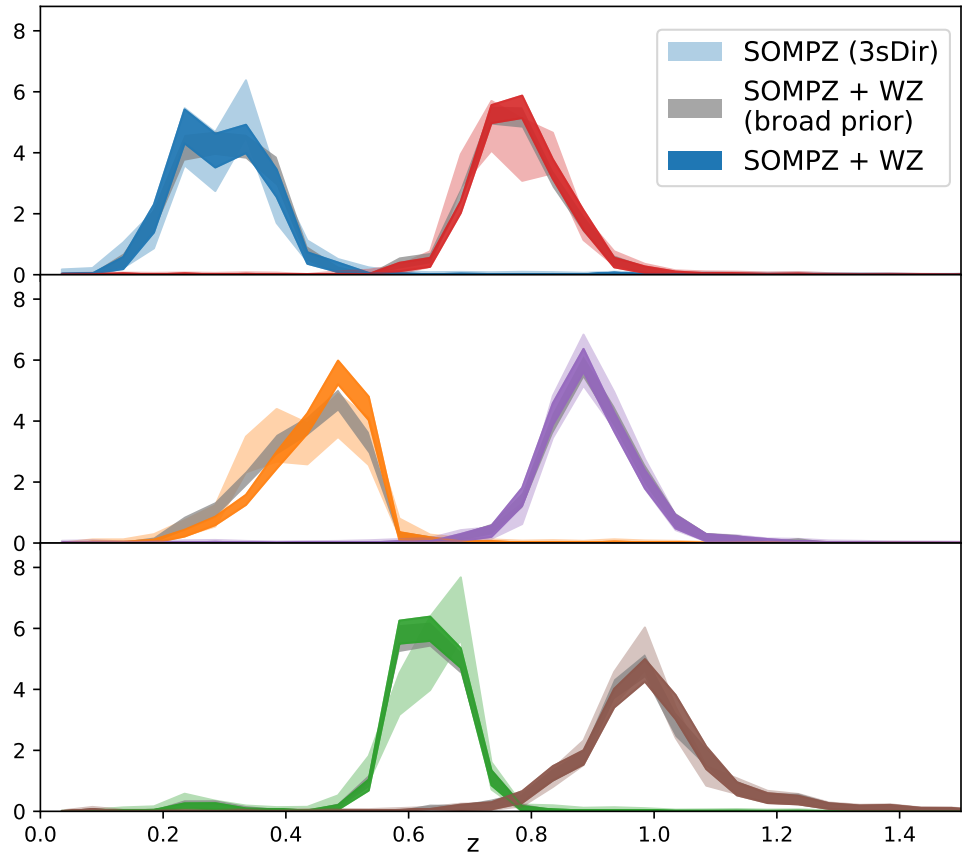

Figure 7.3: 3sDir distributions before (lighter shades) and after the combination with clustering-z (solid shades), and using a broader prior on the galaxy-matter bias function $\mathrm{Sys}(z, \mathbf{s})$ at the combination. In the top row we have bins 1 and 4, in the middle row bins 2 and 5, and in the bottom rows bin 3 and 6. The bands represent the 1σ error from the central value. Note how the combination with WZ tightens the constraint on the shape of the $n(z)$.

galaxy bias evolution uncertainty by measuring it directly from the MagLim auto-correlation function. The addition of the WZ information has a modest impact on the values of the the mean and width of the redshift distributions, at most at the 1σ level; this is somewhat expected, as the WZ and SOMPZ information are independent, but consistent with each other.

7.3.1 Comparison with DNF

We find it interesting to compare the final SOMPZ+WZ redshift distributions with the fiducial ones used for DES Y3, obtained using DNF photometric estimates and clustering constraints (hereafter DNF+WZ). Since the two sets of distributions have been obtained with two different methods, we also briefly discuss the major differences between the two pipelines. The DNF code presented in 4.2.1 produces per-galaxy redshift estimates; these are stacked to produce the redshift distributions for the lens samples. Then, following Cawthon et al. (2020), a clustering redshift measurement is performed, using BOSS/eBOSS galaxies as reference sample, similarly to this work. The DNF $n(z)$are matched to the WZ-estimated $n(z)$through a chi-square fitting; in particular, the DNF $n(z)$are allowed to shift and stretch to improve the χ^2. The maximum-a-posteriori values of the shift and stretch and related uncertainties obtained through this matching procedure are used as a prior for the DNF $n(z)$shift and stretch used in the cosmological inference.

Despite the DNF+WZ and SOMPZ+WZ methods using the same photometric and clustering measurements, the methodologies differ in a number of aspects:

1. **SOMPZ vs DNF uncertainties:** SOMPZ and DNF are both machine learning methods, but they are substantially different in spirit and implementation. DNF is a traditional supervised machine learning code where the likelihood (directional neighborhood) between wide field magnitudes/colors and redshift is learned from training with a subsample of galaxies with both reliable redshift information and measured wide field photometry. On the other hand, in SOMPZ machine learning is only used in an unsupervised fashion (without knowledge of redshift), to group self-similar parts of wide field magnitude/color space together. Then, these groups (wide SOM cells) are probabilistically related using Bayes theorem to the color-redshift relation measured empirically in the calibration deep fields, where much better information is available. The likelihood between each set of wide and deep field photometry is also measured empirically by injecting galaxies of the latter into images of the former. Furthermore, SOMPZ provides a comprehensive list of statistical as well

as systematic uncertainties affecting the calibration samples which are rigorously propagated through the $n(z)$. On the other hand, DNF only describes statistical uncertainties related to the residual differences to the closest training neighbors to the fitted hyperplane of the target galaxies.

2. **Combination:** The clustering information is included and combined with the photometric estimates in a substantially different way. In this work, SOMPZ and WZ are combined by sampling from the joint posterior using the HMC method. No approximation is performed when combining the two likelihoods. On the other hand, matching DNF $n(z)$to the WZ measurements it has been implicitly assumed that the DNF $n(z)$estimates can only be biased at the level of their mean and width, and that inaccuracies in the higher order moments of the $n(z)$can be neglected (or do not affect the matching procedure with the WZ measurements). However, if the DNF and WZ $n(z)$estimates are substantially different beyond their first two moments, the matching might cause biases (Gatti et al., 2018) also in the first and second moments. Furthermore, in the combination of the fiducial method, the DNF shape is only allowed to be modified by shifting and stretching it. Therefore the shift and stretch parameters are centered at the WZ values. This means that the photo-z priors for the cosmological inference only carry uncertainty from the WZ measurement, as this method does not propagate any systematic uncertainties related to uncertainty from the accuracy of DNF or the quality of its training sample photometry. In comparison, SOMPZ+WZ properly combines the statistical significance from SOMPZ and WZ yielding a final uncertainty that truly combines the information from each of them separately. Finally, the SOMPZ+WZ $n(z)$samples also capture the uncertainties in the higher moments of the redshift distributions, whereas the DNF+WZ uncertainties are only relative to the mean and width.

3. **WZ distribution tails:** The WZ measurements used to calibrate the DNF $n(z)$have clipped tails, since the measurements were performed in a restricted redshift window to avoid biases related to un-modelled magnification effects in the tails of the redshift distribution. On the other hand, in this work, when combining the clustering information with SOMPZ estimates, we use the WZ measurements over all the redshift range, since we also marginalise over magnification effects.

4. **WZ galaxy-matter bias:** The WZ measurements used in the DNF+WZ estimates are corrected for the redshift evolution of the galaxy-matter bias of the MagLim sample computed from auto-correlations measure-

		BIN 1	BIN 2	BIN 3	BIN 4	BIN 5	BIN 6
$<z>$	SOMPZ	0.315 ± 0.016	0.463 ± 0.010	0.633 ± 0.009	0.781 ± 0.008	0.893 ± 0.009	0.990 ± 0.012
	DNF	0.292 ± 0.007	0.422 ± 0.011	0.616 ± 0.006	0.762 ± 0.006	0.887 ± 0.007	0.969 ± 0.008
	$\Delta_{<z>}$	1.3	2.7	1.7	1.9	0.5	1.5
σ_z	SOMPZ	0.080 ± 0.005	0.081 ± 0.005	0.060 ± 0.002	0.073 ± 0.003	0.074 ± 0.004	0.102 ± 0.007
	DNF	0.078 ± 0.005	0.094 ± 0.007	0.055 ± 0.003	0.062 ± 0.003	0.075 ± 0.004	0.080 ± 0.007
	$\Delta_{<z>}$	0.2	1.6	1.3	2.2	0.3	2.3

Table 7.3: Values of mean and width of the SOMPZ+WZ final ensemble of distributions and the DNF estimate. The statistical difference $\Delta_{<z>}$ is computed by considering the uncertainties of both methods summed in quadrature, as in $\Delta_{<z>} = (< \mathrm{z} >_{\mathrm{SOMPZ}} - < \mathrm{z} >_{\mathrm{DNF}})/\sqrt{\sigma(< \mathrm{z} >_{\mathrm{SOMPZ}})^2 + \sigma(< \mathrm{z} >_{\mathrm{DNF}})^2}$. We refer to these as are lower limits. Because the WZ measurement is very similar in the two cases, and the uncertainties summed in quadrature are correlated and therefore we are likely underestimating $\Delta_{<z>}$.

ments following Eq. 7.4 (Cawthon et al., 2020). As for this work we use the forward modelling approach described in Section 7.2.2, we instead do not correct directly for the bias, but from the MagLim auto-correlations we determine prior values of the parameters of our $\mathrm{Sys}(z, \mathbf{s})$, and then marginalise over possible bias functions in the sampling from the joint likelihood. We are therefore assuming an uncertainty on the galaxy-matter bias and validating the central value using SOMPZ data.

We must highlight that in Cawthon et al. (2020); Porredon et al. (2021a) several tests were performed to test the robustness of the DNF+WZ method. In particular, Cawthon et al. (2020) tested the performance of the clustering measurements in simulations, whereas Porredon et al. (2021a) tested that matching DNF $n(z)$to the WZ measurements was not introducing biases in the cosmological constraints, and that modelling only the uncertainties in the mean and width of the distributions was sufficient for the DES Y3 cosmological analysis. These tests should cover potential worries raised in points ii), iii) and iv) above for the DNF+WZ method. Having said this, any discrepancy between the SOMPZ+WZ $n(z)$and the DNF+WZ $n(z)$should boil down to the points listed above.

In Figure 7.4, the shapes and uncertainties of the two methodologies are compared, before and after the inclusion of WZ information, respectively in the left and right panel. Visually the DNF+WZ $n(z)$look very similar to the SOMPZ+WZ ones, although some discrepancies can be noticed (e.g., in the second bin). To be more quantitative, we report in Table 7.3 the redshift means and widths of the two sets of distributions, and their agreement. The means and widths are also visually compared in Figure 7.5. The agreement is computed assuming the uncertainties of the two methods to be uncorrelated, which is likely not true; therefore, the reported agreements are optimistic.

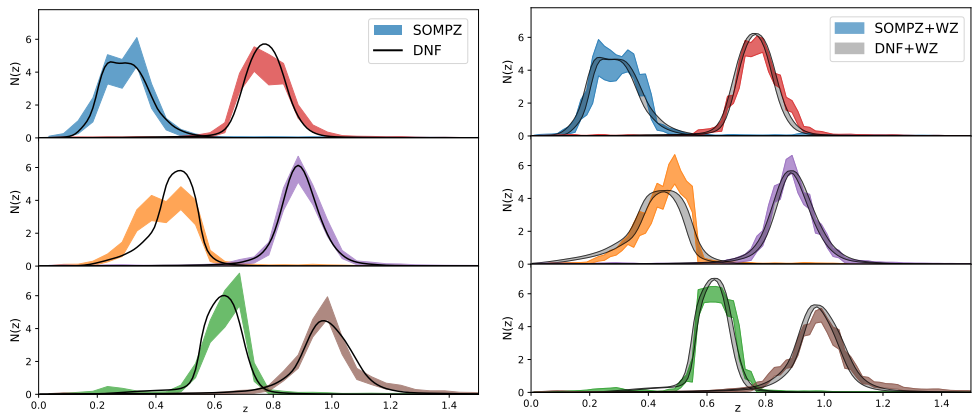

Figure 7.4: Left panel) Final $n(z)$realisations obtained from the SOMPZ methodology alone compared to the fiducial DNF distribution for MagLim (in black). Right panel) Final $n(z)$realisations obtained from both SOMPZ and WZ methodology compared to the fiducial DNF distribution for MagLim (grey bands) after shifting and stretching them to fit WZ measurement. Since in the inference the shift and stretch values are marginalised over, the uncertainties of the gray bands are obtained by sampling over the allowed ranges of shift and stretch defined by the prior, and applied respectively to the DNF estimate. Note that for a fairer comparison of the methods, the two remaining uncertainties were applied to the SOMPZ ensemble (zeropoint and SOMPZ intrinsic), to include all the SOMPZ-related uncertainties. For both plots, in the top row we have bins 1 and 4, in the middle row bins 2 and 5, and in the bottom row bins 3 and 6. .

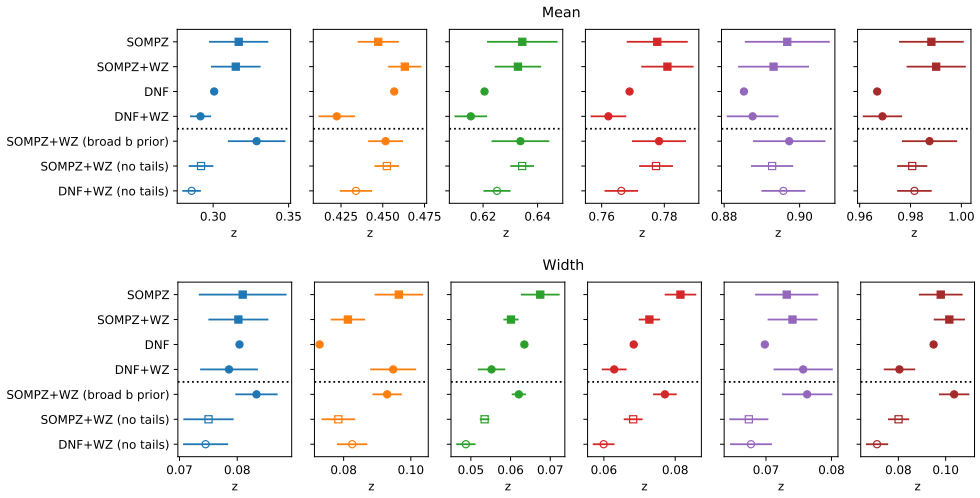

Figure 7.5: Visual representation of the uncertainties on mean (above) and width (below) of the redshift distributions estimated using the SOMPZ (square markers) and DNF (round markers) methods, before and after including the WZ information, for each tomographic bin. Below the dashed line is the comparison of the values computed in the redshift range used for the χ^2 fit of the DNF estimate with the smoothed WZ $n(z)$.

Computing the level of correlation between the two redshift estimates is not trivial. The DNF+WZ estimates and uncertainties are driven only by the WZ measurements in the range where WZ measurements are available and magnification effects are negligible; the tails of the distribution, on the other hand, are described by the DNF estimates. The SOMPZ+WZ estimates receive contributions from both SOMPZ and WZ; if the SOMPZ method was to completely drive our estimates, then the SOMPZ+WZ and the DNF+WZ estimates could be assumed to be independent. This is likely the case for the mean redshift estimates, as we have seen that WZ is not particularly constraining on the mean redshift (see Figure 7.5). The width estimates are inferred more by the WZ measurements, and this might indicate that our tensions are under estimated, because we know that the two calibration methods share part of the WZ information. With this in mind, large tensions between means/widths of the two methods might indicate that either that the DNF+WZ uncertainties are under estimated, or there are some real differences between the two methods (one or both are biased). The reported values in Table 7.3 never exceed 3σ, with some bins showing differences at the 2σ level, which does not point to dramatic differences between the two methods.

From Table 7.3 we note that SOMPZ+WZ uncertainties on the mean are larger than the DNF+WZ ones, while uncertainties on the widths are compa-

rable. This is due to the fact that the uncertainties in the mean redshifts for the SOMPZ estimates are very sensitive to contributions from outliers at high redshift. The DNF+WZ mean redshift estimates (and uncertainties), on the other hand, are driven by the match with the WZ measurements with clipped tails, i.e., they do not take into account uncertainties in the tails, and are therefore smaller. The fact that the modelling of the tails is different between the two methodologies is also responsible for the slightly higher mean redshifts of the SOMPZ+WZ estimates compared to the DNF+WZ estimates. If we restrict the comparison of the aforementioned quantities in redshift intervals that exclude the tails of the distributions, the match between SOMPZ+WZ and DNF+WZ improves (Figure 7.5). We further investigate the importance of the tails on the cosmological constraints in Appendix B.3.1, finding that, despite them being important, they do not drive the main difference between the SOMPZ+WZ and DNF+WZ constraints.

Galaxy-matter bias prior from WZ auto-correlation

We tested the impact on the cosmological parameters of using the same broad prior on the $\mathrm{Sys}(z, \mathbf{s})$ function describing the galaxy-matter bias as was done for the WL sample (Gatti et al., 2020a). In this work we used more informative values computed from the clustering auto-correlation of the MagLim sample, the application of which is explained in more detail in Section 7.2.2. It is particularly interesting to look at the shape of distributions, especially for bin 2. Figure 7.3 shows in grey the 1-sigma bands for the case without using the auto-correlation, and leaving a much broader prior. While in most bins the difference is not appreciable, and the grey bands are very similar to the solid bands, in bin 2 there is an evident difference. It is therefore suggested that this implementation of the auto-correlation information used as priors in the SOMPZ+WZ combination is able to help us constraining the galaxy-matter bias value, in a way that otherwise would not have been possible with traditional methods. In figure 7.5 is shown the comparison over mean redshift and width of the distributions between SOMPZ+WZ with the more informative prior from the auto-correlation, against the broad prior (labelled as "SOMPZ+WZ (broad prior)"). The means and widths are well compatible with the standard SOMPZ+WZ results, and for bins 2 and 3 they are slightly closer to the DNF+WZ results. Even in bin 2, where the shape of the $n(z)$ is substantially different, the values of mean and width do not differ greatly from the standard case, reinforcing the notion that mean and width alone are not sufficient to fully characterise redshift distributions of a lens sample.

7.4 Conclusions

In this Chapter, we presented an alternative calibration of the MagLim lens sample redshift distributions from the Dark Energy Survey (DES) first three years of data (Y3). This new method, which has already been applied to the DES Y3 weak lensing sample (Myles & Alarcon et al., 2020), is based on a combination of a Self-Organising Maps (SOMPZ) based scheme and clustering redshifts (WZ) to estimate redshift distributions and inherent uncertainties. The original redshift calibration of the MagLim has been based on the photo-z code DNF (De Vicente et al., 2016a) and WZ constraints Cawthon et al. (2020). The methodology presented in this paper is meant to be more accurate than the original one. First, the SOMPZ method allows a better control over all the potential sources of uncertainties affecting the estimates compared to DNF; second, the clustering constraints (WZ) are incorporated through a rigorous joint likelihood framework which allows to draw $n(z)$ samples conditioned on both clustering and photometric measurements, improving the $n(z)$ estimates (e.g., the final "SOMPZ+WZ" $n(z)$ have a smaller scatter, or uncertainty, compared to the SOMPZ ones).

We described in detail the methodology followed to produce the alternative MagLim $n(z)$ based on the SOMPZ+WZ approach, together with a detailed report on the main systematics dominating our calibration error budget. Our redshift uncertainties, in particular, are dominated by the impact of sample variance on the SOMPZ estimate (due to the limited area spanned by the deep field sample used in the calibration) and by the effect of the redshift evolution of the galaxy-matter bias of the MagLim sample on the WZ constraints. We then compared our SOMPZ+WZ $n(z)$ with the fiducial DNF+WZ $n(z)$ estimates; in no case did the means and widths of the 6 MagLim tomographic bins show tensions exceeding 3σ, although in some bins we measured differences at the 2σ level. We also found the uncertainties on mean of the redshift distributions of the SOMPZ+WZ method to be slightly larger than the ones of the DNF+WZ method, due to a more conservative calibration of the tails of the redshift distributions. On the other hand, we found the two methods to have a similar constraining power on the widths of the distributions.

Overall, the SOMPZ+WZ method has proven to be robust in minimising residual biases in the $n(z)$ and quite versatile with respect to the systematic uncertainty treatment. We consider this to be an improvement over the fiducial method used for the MagLim redshift calibration in DES Y3. Despite this, there are several improvements that should be taken into consideration for future implementations of the method. Larger deep field area overlapping with additional redshift surveys would greatly benefit the SOMPZ method by

reducing the sample variance and shot noise, the main systematic uncertainty contributors. Also, the photometric uncertainty in the deep field is non negligible at low redshift, where the error budget is dominated by the u-band. Therefore additional data in the u-band in the DES deep fields would reduce the total uncertainty. An ameliorated transfer function could be obtained by increasing the BALROG injection area. Clustering-z could improve its modelling of the evolution of the galaxy-matter bias of the "unknown" sample by obtaining external information from the deep fields. Finally, the implementation of a hierarchical Bayesian methodology would significantly reducing sample variance, by constraining $p(c)$ with the use of the transfer function $p(c|\hat{c})$, and improve the overall efficiency in using clustering to constrain the distribution obtained with photometry (Sánchez & Bernstein, 2019b; Alarcon et al., 2019).

Chapter 8

Impact of the MagLim sample redshift calibration on cosmology

This Chapter describes the impact on the cosmological parameters of using the new redshift calibration presented in Chapter 7, and compares it to the "fiducial" constraints obtained using the DNF+WZ redshift calibration (Porredon et al., 2021a).

As introduced in Chapter 4, the MagLim lens sample has been created with the goal of improving the constraints from the cosmological analyses of galaxy clustering and galaxy-galaxy lensing (Porredon et al., 2021b). The SOMPZ+WZ methodology aims at reduce the systematic uncertainties and, most importantly, to minimise any residual bias left in the MagLim $n(z)$.

8.1 Cosmological Results

In this section, we show the constraints on cosmological and nuisance parameters obtained using the DES Y3 measurements for galaxy-galaxy lensing and galaxy clustering (Prat et al., 2022; Rodríguez-Monroy et al., 2022) (a.k.a. 2x2pt), and the $n(z)$ from the redshift calibration from 7. As in Porredon et al. (2021a), we also include in our analysis an additional likelihood constructed with the Shear Ratios (SR) measurement, introduced in Section 3 (Sánchez et al., 2022a). In brief, this exploits galaxy-galaxy lensing signal at small scales (< 6 Mpc/h) to provide further constraint to the redshift distributions and intrinsic alignment parameters. The ratio of a galaxy-galaxy lensing signal of each lens sample redshift bin computed with respect to two

source sample bins results in a primarily geometric measurement, which has been proven a powerful method for constraining systematics and nuisance parameters. This adds independent information from SOMPZ and WZ to the source redshift calibration.
The posterior distribution obtained follows the Bayes theorem:

$$P(p|D,M) \propto \mathcal{L}(D,p,M)\Pi(p|M), \tag{8.1}$$

where $\Pi(p|M)$ is the prior distribution for all the parameters of the model M. For the cosmological inference we use the `CosmoSIS` pipeline (Zuntz et al., 2015), and we sample the parameter posteriors using the `PolyChord` sampler (Handley et al., 2015a,b).

Our data vector $D = \{w(\theta), \gamma_t(\theta)\}$ is compared to theoretical predictions $T(p) = \{w(\theta,p), \gamma_t(\theta,p)\}$ in a Bayesian fashion, and the posterior of the parameters conditional on the data is evaluated by assuming a Gaussian likelihood for the data:

$$\log \mathcal{L} \propto -\frac{1}{2}(D-T(p))^T C^{-1}(D-T(p)), \tag{8.2}$$

where C is the measurement covariance. In our analysis, we vary 5 (or 6) cosmological parameters assuming a ΛCDM (or wCDM) cosmology: Ω_m, σ_8, n_s, Ω_b, h_{100}, and w for the wCDM case. Moreover, we also marginalise over "astrophysical" nuisance parameters (describing intrinsic alignment effects and the galaxy-matter bias of the lens sample), and calibration parameters (redshift uncertainties, shear measurement uncertainties). In short, our setup (covariance, parameters varied, prior ranges, etc.) is the same as the one adopted in Porredon et al. (2021a), except for the redshift $n(z)$ and uncertainties priors of the lens sample, where the ones obtained in this work have been assumed, and other minor changes that we describe below.

Our analyses were not "blinded", since this work occurred after the "unblinding" of the DES Y3 3x2pt results. We did not perform any cosmological analysis until the redshift distributions were frozen; no changes to the redshift distributions (and uncertainties prior) have been performed after looking at the cosmological constraints. To ensure the robustness of our final estimates, we adopted a p-value criteria on the best-fitting models to our data vector. Following Porredon et al. (2021a), we required the goodness-of-fit $p-$value on unblinded data vectors was larger than 1 per cent. The goodness-of-fit has been computed using the Predictive Posterior Distribution (PPD, Doux et al. 2021) and adopted in the main DES Y3 3x2pt analysis. The PPD methodology derives a calibrated probability-to-exceed p; in the case of goodness-of-fit tests, this is achieved by drawing realisations of the data vector for parameters

drawn from the posterior under study which are then compared to actual observations. The distance metric (χ^2) is computed in data space, which is then used to compute the p-value.

Concerning the redshift uncertainties, as it is the primary goal of this work, we proceeded using the fiducial DES Y3 methodology: we parametrize the redshift uncertainties with two parameters for each tomographic bin, that modify a fiducial $n(z)$distribution with a shift on the mean and a stretch on the width. The fiducial $n(z)$is estimated by averaging the SOMPZ+WZ $n(z)$ realisations. The Gaussian priors on the mean and stretch parameters are centered at the mean and width of the fiducial $n(z)$, while the Gaussian priors width are measured from the variance in the mean and width of the $n(z)$ensemble. This parametrization can be compared directly to the fiducial DES Y3 2x2 analysis Porredon et al. (2021a). In Appendix B.3 we describe an alternative marginalisation of the redshift uncertainties, by marginalising over the full sets of $n(z)$realisations provided by the SOMPZ+WZ method. In principle, this latter method describes better the redshift uncertainties of our method. However, we find that the currently available techniques that marginalise over the full ensemble of realisations during cosmology inference are prohibitively computationally expensive. Therefore we defer its application for future work.

Besides the different $n(z)$, we also ran a few analyses where we marginalised over magnification parameters of the lens samples over wide priors (this is different from Porredon & Crocce et al. (2021a), where magnification parameters have been fixed to some fiducial value). For the fiducial 2x2pt analysis the p-value from the data-model χ^2 using all six bins of MagLim was not sufficient to pass the 1 per cent criteria. After a series of tests the consensus was that the two highest redshift tomographic bins were responsible for worsening the fit. Therefore the analysis in Porredon et al. (2021b) included only the first 4 MagLim bins.

Here, we perform the analyses using all the 6 bins of the MagLim sample, but also using only the first 4 bins, to verify if the same applies also to this work using different redshift distributions. In particular, we consider the following scenarios:

- ΛCDM (wCDM); 4 and 6 lens bins, fixed magnification. This is the fiducial analysis that mirrors the one presented in Porredon et al. (2021a). Five (six) cosmological parameters are varied, including $\Omega_{\rm m}$, σ_8, $n_{\rm s}$, $\Omega_{\rm b}$, h_{100} (and w for the wCDM case). Intrinsic alignment, shear measurement and redshift uncertainties parameters (of both lenses and sources) galaxy-matter linear biases of the lenses also are marginalised over. The magnification coefficients of the lens sample, however, are fixed to the

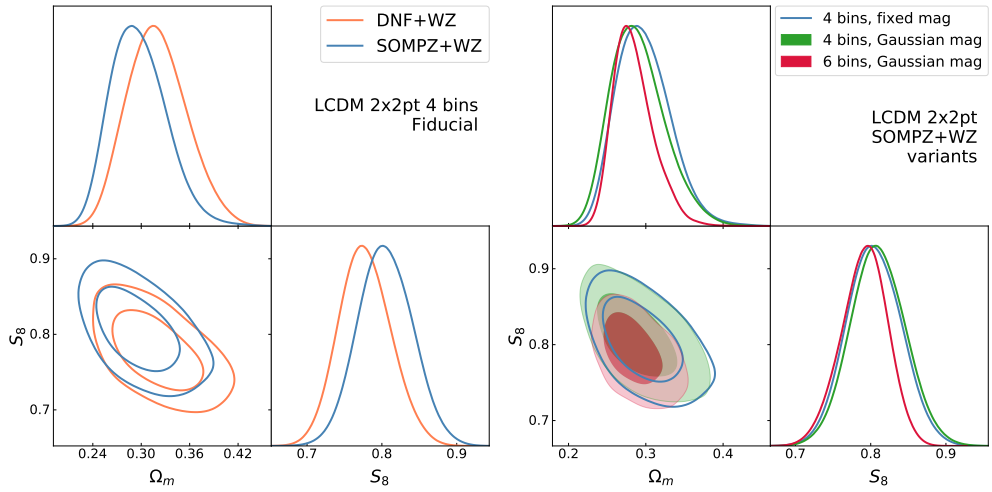

Figure 8.1: *Left panel*: Posterior distributions of the cosmological parameters $\Omega_{\rm m}$, and S_8 for the ΛCDM analysis involving 4 bins and fixed magnification parameters. The "fiducial" posteriors have been obtained using the DNF+WZ redshift distributions, and they are compared to the ones obtained using the SOMPZ+WZ redshift distributions. *Right panel*: Posterior distributions of the cosmological parameters $\Omega_{\rm m}$, and S_8 for the ΛCDM analysis for three different cases: 1) 4 bins and fixed magnification parameters (the blue contours in the two plots share the same analysis choices); 2) 4 bins and marginalised over magnification parameters (in solid green); 3) 6 bins and marginalising over magnification parameters (in solid red). The 2D marginalised contours in both of these figures show the 68 per cent and 95 per cent confidence levels.

values estimated from Balrog Everett et al. (2020a). Uncertainties in the redshift distributions of the lens sample are modelled as a shift and stretch in the distributions.

- ΛCDM (wCDM); 4 and 6 lens bins, free magnification. Same as the ones above, but magnification parameters are marginalised over using Gaussian priors.

In what follows, we will also quote results in terms of the S_8 parameter, defined as $S_8 \equiv \sigma_8(\Omega_{\rm m}/0.3)^{0.5}$. In Table 8.1 we summarise the best fit values of S_8, Ω_m, σ_8, w, and the computed PPD goodness-of-fit p-value for all the different analyses.

Table 8.1: Constraints on the cosmological parameters S_8, Ω_m, and σ_8. For each parameter we report the mean of the posterior and the 68 per cent confidence interval. We also report the PPD goodness-of-fit p-value.

$n(z)$	Model	bins	Magnif.	S_8	Ω_m	σ_8	w	p-value
SOMPZ	ΛCDM	4 bins	Fixed	0.81 ± 0.04	0.30 ± 0.04	0.81 ± 0.07	-	0.029
SOMPZ (broad prior)	ΛCDM	4 bins	Fixed	0.76 ± 0.06	0.31 ± 0.04	0.76 ± 0.09	-	-
SOMPZ	ΛCDM	6 bins	Fixed	-	-	-	-	0.008
SOMPZ	ΛCDM	4 bins	Gaussian	0.79 ± 0.06	0.30 ± 0.05	0.79 ± 0.10	-	0.035
SOMPZ	ΛCDM	6 bins	Gaussian	0.78 ± 0.08	0.31 ± 0.04	0.77 ± 0.09	-	0.065
SOMPZ	wCDM	4 bins	Fixed	0.79 ± 0.06	0.29 ± 0.04	0.81 ± 0.08	-1.2 ± 0.3	0.032
SOMPZ	wCDM	6 bins	Fixed	0.78 ± 0.04	0.30 ± 0.04	0.78 ± 0.06	-0.9 ± 0.3	0.012
SOMPZ	wCDM	4 bins	Gaussian	0.76 ± 0.07	0.30 ± 0.05	0.77 ± 0.10	-1.0 ± 0.3	0.035
SOMPZ	wCDM	6 bins	Gaussian	0.78 ± 0.08	0.31 ± 0.04	0.77 ± 0.09	-0.7 ± 0.2	0.059
DNF	ΛCDM	4 bins	Fixed	0.78 ± 0.04	0.32 ± 0.04	0.76 ± 0.07	-	0.019
DNF	wCDM	4 bins	Fixed	0.78 ± 0.05	0.32 ± 0.05	0.76 ± 0.07	-1.0 ± 0.3	0.024

8.1.1 ΛCDM results

Fiducial results: 4 bins, fixed magnification and comparison with DNF results

The first cosmological constraints we analyse are the ones obtained assuming a ΛCDM cosmology, using 4 lens bins and fixed magnification parameters. This is the "fiducial" setup assumed in the Porredon et al. (2021c) analysis, and we can compare our results directly to the ones obtained using the DNF+WZ $n(z)$. The posterior on the the cosmological parameters Ω_m, and S_8 is shown in the left panel of Fig. 8.1; the marginalised mean values of S_8, Ω_m, and σ_8, along with the 68% confidence intervals, are:

$$\Omega_m = 0.30 \pm 0.04, \tag{8.3}$$

$$\sigma_8 = 0.81 \pm 0.07, \tag{8.4}$$

$$S_8 = 0.81 \pm 0.04. \tag{8.5}$$

The PPD goodness-of-fit test for this analysis results into $p-$value=0.029, well above our threshold (see also Table 8.1). In the left panel of Fig. 8.1 we also compare our results with the constraints obtained using the fiducial DNF+WZ $n(z)$. The size of the posteriors is similar for the two cases, but the two posteriors are slightly shifted; the distance between the posteriors' peaks in the 2D $\Omega_m - S_8$ plane is $d \sim 0.4\sigma$. In DES Y3 we impose a strict 0.3σ threshold for differences in the $\Omega_m - S_8$ plane induced by different analysis choices; these results would apparently violate this criteria. We note, however, that the (arbitrary) 0.3σ threshold adopted by DES refers to differences in the $\Omega_m - S_8$ plane when *noiseless* theory data vectors are assumed. In the presence

of noisy data vectors these differences can become larger, without invalidating our criteria. Having said this, a $d \sim 0.4\sigma$ difference nonetheless imply the large impact that a different redshift calibration of the lens sample can have on the cosmological constraints. This is somewhat different from the results obtained for the source sample $n(z)$ (Amon et al., 2022), where uncertainties in the redshift calibration had a negligible impact on the cosmological constraints.

In Section 7.2.2 we explained how for the combination of the two methods we marginalise over possible functional forms of the unknown galaxy-matter bias of the MagLim sample, by means of the systematic function $\mathrm{Sys}(z, \mathbf{s})$ in our clustering model. The prior on the parameters $\mathbf{s}$ is inferred from the clustering auto-correlation. We tested the impact on the redshift distributions of using a broader prior (the same used in Myles & Alarcon et al. 2020) in Section 7.3. We have tested the impact on using these $n(z)$ for the cosmological inference, and found that there is no change in constraining power and no shift for Ω_{m}, but there is a shift on S_8 such to overlap with the fiducial results from DNF+WZ. Therefore it is clear that the information carried by the auto-correlation is crucial in our cosmological analysis.

Last, we mention that our analysis using 6 lens bins, ΛCDM and fixed magnification coefficients did not pass our $p-$value criteria ($p-$value $= 0.008$), as the fiducial one; hence, we do not show those results here.

4 and 6 bins, free magnification

We then proceed relaxing the fixed priors on the magnification parameters for the lens sample. Instead of fixing them to the values estimated from Elvin-Poole et al. (2021) (as done in the previous section), we leave them as free parameters, using Gaussian priors. In short, Elvin-Poole et al. (2021) estimate the magnification parameters using Balrog, by injecting fake galaxies into the wide field with and without applying a small magnification; the difference between the number of galaxies passing the selection in the two cases is then used to estimate the magnification parameters of the sample. These parameters come with a small uncertainty, which is however ignored in the fiducial analysis, as the magnification parameters are assumed to be fixed to the mean Balrog value. The central values and the uncertainties are reported in table B.1 in B.2. One of the main reasons the DES Y3 fiducial analysis did not vary the magnification parameters was merely computational, as 4 (or 6) additional parameters lengthen the parameter inference process. In principle there is no reason to doubt these estimates. Differences might be caused by the fact that the Balrog injections do not completely sample the full DES Y3 footprint, or in case our injections were not fully representative of the DES sample we are

analysing.

When varying these parameters in our analyses, we find that the $p-$value computed using PPD indicates a good fit of the model to the data not only for the 4 bins case, but also for 6 bins case (see Table 8.1). Adding the last 2 lens bins significantly improves the constraining power on $\Omega_{\rm m}$ by 30% compared to the 4 bins case, whereas the constraints on S_8 are 20% tighter.

8.1.2 wCDM Results

We then proceed analysing the results obtained with wCDM, for all four cases: 4 and 6 bins, fixed and free magnification, as described in the previous section. In general, the 2x2pt constraints on w are loose and affected by the prior ($-2 < w < -0.3$). Parameters posteriors are shown in Fig. 8.2, whereas p-values and parameters constraints are reported in Table 8.1. Including 2 more lens bins increases our constraining power on w by $\sim$20%, although part of the improvement (especially in the case of 6 bins and free magnification) is due to the posterior partially hitting the prior edge.

Freeing the magnification parameters shifts w towards the upper edge of the prior ($w = -0.3$), which evidently results in tighter constraints on S_8; this is caused by projection effects due to the w prior, and it is stronger for the 6 bins case. The constraints on Ω are unaltered in all four scenarios.

8.1.3 Statistical distance to Planck

We compute here the statistical distances between our cosmological constraints and the early Universe ones from the *Planck* satellite (Aghanim et al., 2020). The comparison between the results has been performed considering all the parameters shared by the two analyses; in particular, we followed the approach presented by Lemos et al. (2021) with improved algorithms from Raveri & Doux (2021). We find that for the 4 bins case for ΛCDM (both fixed and free magnification) there is good agreement ($1.15\sigma, 1.11\sigma$), similarly for wCDM with 4 bins we have 0.46σ for both fixed and free magnification. For the 6 bins cases the values are slightly larger than 2σ. For LCDM free magnification, wCDM fixed, and wCDM free magnification we have respectively: 2.4σ, 2.3σ, and 2.2σ.

8.2 Conclusions

In this Chapter we investigated the impact on the cosmological constraints of using an improved methodology for the redshift calibration of the DES Y3

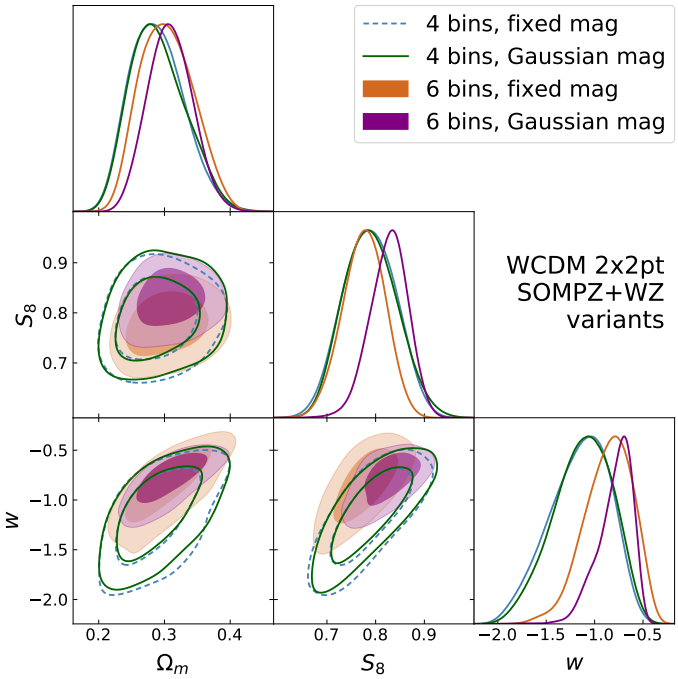

Figure 8.2: Posterior distributions of the cosmological parameters $\Omega_{\rm m}$, and S_8 and w for four different cases: 1) wCDM, 4 bins and fixed magnification parameters; 2) wCDM, 6 bins and fixed magnification parameters, 3) wCDM, 4 bins and free magnification parameters; 4) wCDM, 6 bins and free magnification parameters. The 2D marginalised contours in these figures show the 68 per cent and 95 per cent confidence levels.

MagLim galaxies. In particular, we used the galaxy-galaxy lensing and galaxy clustering measurements (Prat et al., 2022; Rodríguez-Monroy et al., 2022) (a.k.a. 2x2pt), and the $n(z)$ from the combination of SOMPZ and WZ measurements, as described in Chapter 7. The new cosmological constraints have been compared to the results originally presented in Porredon et al. (2021a). In the "fiducial" configuration, which involves using the first 4 lens bins and assuming a ΛCDM cosmology, we obtained as marginalised mean values $\Omega_{\rm m} = 0,30 \pm 0.04$, $\sigma_8 = 0.81 \pm 0.07$ and $S_8 = 0.81 \pm 0.04$. We noted a $\sim 0.4\sigma$ shift in the $\Omega - S_8$ plane compared to the Porredon et al. (2021a) results, but no change in terms of constraining power. The shift indicates that the redshift calibration of the lens sample plays a key role on cosmological constraints from the 2x2pt analysis, contrary to the redshift calibration of the source sample (Amon et al., 2022; Secco et al., 2022). Subsequently, we explored different analysis setups; we tested the case where all the 6 MagLim redshift bins were included, a scenario where the magnification coefficients of the lens sample were marginalised during the inference, and last, we assumed a wCDM cosmology. We found that the inclusion of the last two redshift bins of the MagLim sample help improving the constraints on $\Omega_{\rm m}$ by $\sim 25\%$, and on S_8 by $\sim 20\%$.

We also compared our results to the cosmological constraints from *Planck* (Aghanim et al., 2020), finding a no-tension of $\sim 1\sigma$ between the results when 4 lens bins where considered. We did find a statistical distance of more than 2σ in ΛCDM with free magnification coefficients when including in the analysis the two high redshift bins ($z > 0.85$), which have not been included in the fiducial DES Y3 analysis Porredon et al. 2021a.

As a final comment, despite the SOMPZ+WZ method's ability to produce $n(z)$ samples capturing the redshift uncertainties of our estimates, we could not efficiently marginalise over these realisation during the cosmological inference, due to computational constraints. Our marginalisation strategy followed the one adopted in Porredon et al. (2021a): we adopted the mean of the SOMPZ+WZ samples as our fiducial $n(z)$, and marginalised over a shift in the mean and a stretch of the width of the distribution, using as priors the variances in the mean and widths of the SOMPZ+WZ $n(z)$ samples. While this strategy was deemed sufficient for this current work, we plan to implement the full marginalisation scheme for subsequent analyses of the lens samples with DES Y6 data.

Appendix A

A.1 Magnification effects

We provide in this Appendix more details about the modelling of magnification effects $M(z_i)$ in the cross-correlation signal between the unknown and reference samples. Considering only the dominant terms (which account for the magnification of the unknown sample by the reference sample and the magnification of the reference sample by the unknown sample) and assuming linear bias, this can be written as:

$$M(z_i) = \int d\theta W(\theta) \int \frac{dl\ l}{2\pi} J_0(l\theta) \int \frac{d\chi}{\chi^2} \times \left[b_{\rm r}\alpha_{\rm u} q^{\rm r}_{\delta} q^{\rm u}_{\kappa} + b^{'}_{\rm u}\alpha_{\rm r} q^{\rm u}_{\delta} q^{\rm r}_{\kappa} \right] P_{\rm NL}\left(\frac{l+1/2}{\chi}, z(\chi) \right), \quad \text{(A.1)}$$

where the terms q_δ and q_κ read:

$$q_\delta(\chi) = n[z(\chi')] \frac{dz}{d\chi'}, \quad \text{(A.2)}$$

$$q_\kappa(\chi) = \frac{3H_0^2\Omega_{\rm m}\chi}{c^2 a(\chi)} \int_{\chi}^{\chi(z=\infty)} d\chi' n(z(\chi')) \frac{dz}{d\chi'} \frac{\chi' - \chi}{\chi'}. \quad \text{(A.3)}$$

In the above equations, $n[z(\chi)]$ is either $n_{\rm u}(z)$ or $n_{\rm r,i}(z)$ Under the approximation of thin redshift bins, we can write Eq. A.1 as a discrete summation over redshift bins of width $\Delta\chi$:

$$M(z_i) \ = \ b_{\rm r}(z_i)\alpha_{\rm u}(z_i) \sum_{j>i} \left[D_{ij} n_{\rm u}(z_j) \right] \ + \ b^{'}_{\rm u}(z_i)\alpha_{\rm r}(z_i) \sum_{j>i} \left[D_{ij} n_{\rm u}(z_i) \right], \quad \text{(A.4)}$$

with

$$D_{ij} = \frac{3H_0^2\Omega_{\rm m}}{c^2} w_{\rm DM}(z_i) \frac{\chi(z_i)}{a(z_i)} \frac{\chi(z_j) - \chi(z_i)}{\chi(z_j)} \Delta\chi_j. \quad \text{(A.5)}$$

The magnification coefficient α, for an ideal flux limited sample, can be related to the slope s of the cumulative number counts evaluated at flux limit: $\alpha \equiv 2.5s - 1$, with the slope formally defined as

$$s = \frac{d}{dm} \log_{10} n(< m), \tag{A.6}$$

where $n(< m)$ is the cumulative number count as a function of magnitude m, and s is to be evaluated at the flux limit of the sample. For a sample which is not flux limited, evaluating the coefficient s is more complicated, and Eq. A.6 can not be used. We use two different methods to estimate such coefficients for our samples, depending on whether we estimate them on data or on simulations (see below for further details). Estimates of α for both the reference and unknown samples are needed to properly model magnification effects.

In the Y3 method we model magnification effects according to Eq. A.4. While we absorb the contribution due to $b_{\rm u}$ to the clustering signal into the Sys$(z, \mathbf{s})$ function, we leave $b_{\rm u}^{'}$ as a free parameter in the magnification term. We also leave $\alpha_{\rm u}$ as a free parameter, and marginalise over both parameters analytically when computing the likelihood. By doing so, we absorb uncertainties not only in these values but also in $b_{\rm r}, \alpha_{\rm r}$ and in the linear-bias model adopted for magnification. Hence, formally, the $b_{\rm u}^{'}$ value appearing in the magnification is not assumed to equal the $b_{\rm u}$ that might multiply $w_{\rm DM}$. We do not implement redshift dependence of $\mathbf{p} = \{b_{\rm u}, \alpha_{\rm u}\}$ (although the formalism would allow it) because magnification signals are important only over limited ranges of z (i.e., in the tails, see, e.g., Gatti & Vielzeuf et al. 2018) for a given tomographic bin of the WL sources.

A.1.1 Magnification coefficients estimates

In order to estimate the magnification coefficients of our samples, we adopt two different strategies. For the coefficients in data we use Balrog image simulations (Suchyta et al., 2016; Everett et al., 2020a) in a process briefly described here. Galaxy profiles are drawn from the DES deep fields (Hartley & Choi et al., 2020) and injected into real DES images. The full photometry pipeline (Sevilla-Noarbe et al., 2020), the RedMaGiC , MagLim, and WL sample selection are applied to the new images to produce simulated samples with the same selection effects as the real data. To compute the impact of magnification, the process is repeated, this time applying a constant magnification to each injected galaxy. The magnification coefficients are then derived from the fractional increase in number density when magnification is applied. This method captures

the impact of magnification on both the galaxy magnitudes and the galaxy sizes, including all sample selection effects and potential observational and systematic effects. See Everett et al. (2020a); Elvin-Poole et al. (2021) for further details. The coefficients have been estimated for RedMaGiC in 5 wide redshift bins, centered at $z = (0.25, 0.425, 0.575, 0.75, 0.9)$: yielding the magnification coefficients $\alpha_{\rm r}^{\rm RedMaGiC} = (0.3 \pm 0.7, -1.5 \pm 0.5, -0.7 \pm 0.4, 1.2 \pm 0.5, 1.0 \pm 0.5)$. The coefficients for MagLim are instead computed in 6 bins, centered at $z = (0.3, 0.4075, 0.625, 0.7075, 0.9, 1.0)$ and are $\alpha_{\rm r}^{\rm MagLim} = 0.2 \pm 0.3, 0.1 \pm 0.3, 0.9 \pm 0.2, 1.0 \pm 0.2, 0.8 \pm 0.5, 1.5 \pm 0.7$. The accuracy of these estimates is limited by the number of Balrog injections, which are scarce for a sample as bright as the lens samples. Since the full-shape method formally requires values of the magnification coefficients for each of the bins of the RedMaGiC and MagLim sample, we interpolate these values in z using the `scipy` routine `interp1d`. Although this procedure might not be too accurate given the large uncertainties of the values of $\alpha_{\rm r}$, magnification effects are largely negligible, such that the interpolation details should not impact our main results. For the WL sample, using the same methodology, we infer $\alpha_{\rm u} = (-0.4 \pm 0.2, -0.21 \pm 0.10, 0.00 \pm 0.10, 0.31 \pm 0.07)$, for the 4 tomographic bins, respectively. Note that these values do not have to be interpolated.

For the values estimated for the samples in simulations we adopt a different strategy. In particular, we use the estimated convergence, κ, computed at the location of each galaxy, to apply a small magnification to the galaxy magnitudes (Δm), and then select our samples with and without this Δm applied and compute the fractional change of objects passing the selection $\Delta N/N$ in 10 equally spaced κ bins. The gradient of this relation is then related to the magnification coefficient (Elvin-Poole et al., 2021). This method only captures the effect of magnification on the galaxy fluxes, as it is the only effect expected in simulations. We estimate $\alpha_{\rm r}^{\rm RedMaGiC} = (0.2 \pm 0.4, 0.05 \pm 0.15, 0.00 \pm 0.08, 1.11 \pm 0.12, 1.18 \pm 0.06)$ for the RedMaGiC sample, $\alpha_{\rm r}^{\rm MagLim} = (0.64 \pm 0.04, 1.04 \pm 0.02, 1.08 \pm 0.02, 1.2 \pm 0.01, 1.42 \pm 0.01, 1.52 \pm 0.03)$ for MagLim, and finally $\alpha_{\rm u} = (-0.365 \pm 0.002, -0.655 \pm 0.002, -0.447 \pm 0.002, 0.836 \pm 0.002)$ for the WL sample.

Last, we note that estimates of the magnification coefficients are not available for BOSS/eBOSS galaxies, as we did not try to reproduce the complex BOSS/eBOSS selection function within Balrog image simulations. We also did not estimate these coefficients for the simulated BOSS/eBOSS sample This is not a problem, as we verify below that BOSS/eBOSS does not have the sensitivity to measure magnification effects. When formally needed (for the full-shape method), though, we adopted the same coefficients as the RedMaGiC

sample.

A.1.2 Magnification impact on the clustering measurements

Magnification effects are modelled over the full range of redshift, using as input the estimated magnification coefficients. Nevertheless, their impact is strongly reduced by the combination with the SOMPZ likelihood, which enforces the tails of the redshift distributions to have a small amplitude. To roughly asses the impact of the exact values of the magnification coefficients $\alpha_{\rm r}$ and $\alpha_{\rm u}$, we performed the following test, both in simulations and on data: we verified that assuming values for $\alpha_{\rm r}$ or $\alpha_{\rm u}$ that are $-1\times$ the fiducial ones resulted in shifts $\Delta\langle z\rangle < 10^{-3}$. This highlight the importance of combining SOMPZ and clustering information to achieve a more robust estimator of the redshift distributions.

A.2 Full $\hat{w}_{\rm ur}$ model and analytical marginalisation

We provide here more details about the implementation of the full-shape method. The method assigns a likelihood (Eq. 7.6) of the observed $w_{\rm ur}(z_i)$ given a proposal for the redshift distributions $\{n_{\rm u}(z_i)\}$ along with a set of other relevant parameters. The likelihood uses the model in Eq. 7.5. We will assume that the values of the dark-matter correlation $w_{\rm DM}(z_i)$, the reference-sample properties $b_{\rm r}(z)$ and $\alpha_{\rm r}(z)$, and the magnification coefficients D_{ij} are provided along with $n_{\rm u}(z_i)$. We will consider as nuisance parameters the properties of the unknown population, namely the $\alpha_{\rm u}(z)$ and $b_{\rm u}(z)$ used in magnification terms; plus any parameters $\mathbf{s}$ of the ${\rm Sys}(z,\mathbf{s})$ function that allows for systematic errors.

We will assume here that $\alpha_{\rm u}$ and $b_{\rm u}$ are independent of redshift, though in principle a more general function, linear in some parameters, can be used without altering any of the methods herein. We note that we did not multiply the magnification terms by the systematic function: despite the fact that the magnification terms are not immune to systematic errors, we assumed that it was not necessary to further modelling those, as the $\alpha_{\rm r}$, $\alpha_{\rm u}$ and $\bar{b}_{\rm u}$ parameters provide enough flexibility to the model and the magnification signal is much smaller than clustering to start with. We also note that the $b_{\rm u}$ parameter is used only in the magnification term, and hence can be independent of the bias for clustering that is absorbed into the ${\rm Sys}(z,\mathbf{s})$ function. This allows for the systematic errors in the magnification term to differ from those in the clustering term.

The systematic-error function for clustering is given the exponentiated polynomial form in Eqs. 7.2 and 7.3. Tuning the order M allows us to adjust the

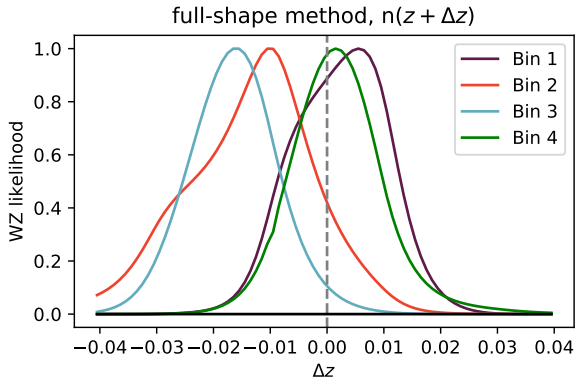

Figure A.1: Likelihood of the shift of the mean of the redshift distributions obtained using the full shape method in simulations, and using true redshift distributions (but shifted around their mean) as proposals distributions.

smoothness of the function, and exponentiation allows us to draw the coefficients $\mathbf{s}$ from 1-d Gaussian priors while maintaining positive $\mathrm{Sys}(z, \mathbf{s})$. Adjusting the σ_s values of these priors tunes the RMS of the systematic variations, in a way made predictable by the orthogonality of the Legendre polynomials. We wish for independent, uniform Gaussian priors on the s_i to propagate into RMS variation of $\log[\mathrm{Sys}(z, \mathbf{s})]$ that is approximately independent of z over $[z_{\rm min}, z_{\rm max}]$. The Legendre polynomials have this property over most of their nominal domain $u \in [-1, 1]$, but not near the edges of this range. For this reason we map $[z_{\rm min}, z_{\rm max}] \to [-0.85, 0.85]$, as indicated by Eq. 7.3.

Eq. 7.6 requires us to marginalize over the nuisance-parameter vector $\mathbf{q} = \{\mathbf{p}\,\mathbf{s}\}$ (with $\mathbf{p} = b_{\rm u}, \alpha_{\rm u}$). Doing so as part of a Markov chain would be unweildly, introducing 8 free parameters for each of the 4 tomographic bins times 2 reference samples. It is far better to execute the marginalization on the fly during sampling if possible. The log-likelihood is not quite quadratic in $\mathbf{q}$—the exponentiation of the polynomial in $\mathrm{Sys}(z, \mathbf{s})$ makes the model $\hat{w}_{\rm ur}$ non-linear in $\mathbf{s}$. We opt to linearize the model about its maximum $\mathbf{s}_0 = \{s_{k,0}\}$:

$$Sys\,(z_i, \mathbf{s}) \approx Sys\,(z_i, \mathbf{s}_0) \times \left[1 + \sum_{k=0}^{M} \frac{\sqrt{2k+1}}{0.85} P_k(u) s_{k,0}\,(s_k - s_{k,0})\right], \quad \text{(A.7)}$$

The deviation of the data from the model can then be rewritten in linear form,

with $\mathbf{w}_{\rm ur}$ being a vector over redshifts, as

$$\mathbf{w}_{\rm ur} - \hat{\mathbf{w}}_{\rm ur} = \mathbf{c}(\mathbf{q}_0) - A\mathbf{q} \tag{A.8}$$

where $\mathbf{c}$ is a vector independent of $\mathbf{q}$ and A is a matrix composed of the linear terms in Eq. A.7 and elements of the magnification terms.

If we assume the nuisance parameters we want to marginalise over to have a Gaussian prior $\mathbf{q} \sim \mathcal{N}(\mu_q, \Sigma_q)$, we can write the full likelihood as follows:

$$\mathcal{L}_{\rm WZ} \approx |2\pi\Sigma_{\rm wz}|^{-1/2}|2\pi\Sigma_p|^{-1/2} \times \int d\mathbf{q}\, \exp\left[-\frac{1}{2}(\mathbf{c}-A\mathbf{q})^T\hat{\Sigma}_{\rm wz}^{-1}(\mathbf{c}-A\mathbf{q})\right] \times \exp\left[-\frac{1}{2}(\mathbf{q}-\mu_q)^T\hat{\Sigma}_q^{-1}(\mathbf{q}-\mu_q)\right], \tag{A.9}$$

This is a Gaussian integral that can be reduced to linear algebra.

In summary, the algorithm for the marginalization in Eq. 7.6 is:

1. Find the values $\mathbf{q}_0$ which maximize the integrand. This is done using Newton iterations.

2. Evaluate the vector $\mathbf{c}$ and matrix A at this value of $\mathbf{q}_0$.

3. Substitute these and the known $\Sigma_{\rm wz}, \mu_q$, and Σ_q into the analytic result for the Gaussian integral above.

Although this marginalization is approximate, it does not actually need to be exact, because the chosen functional form for $\mathrm{Sys}(z, \mathbf{s})$ is somewhat arbitrary. All that is necessary is that the algorithm yields a likelihood $\mathcal{L}$ of the clustering-z data given a proposed $n_{\rm u}(z)$ that decreases in a meaningful way and robust way as the data move away from the naive linear model. We prove that the full shape method recovers the true $n(z)$ within uncertainties in §6.3, assuming the SOMPZ realisations as $n_{\rm u}(z)$ proposals. Here we also show the result of a simpler test, performed in simulations, where the $n_{\rm u}(z)$ proposals are simply taken to be true redshift distributions shifted around their mean. This is a useful test because it is shows that the methodology is unbiased independently of the SOMPZ information. We use Eq. A.9 to assign each true $n_{\rm u}(z)$s (shifted around their mean) a weight, using the clustering measurement and the magnification coefficients from the simulations. The key result is then the likelihood of the shifts Δz, which has to be statistically compatible with 0. This is shown in Fig. A.1; in particular, we obtain

$\Delta z = 0.002 \pm 0.008, -0.013 \pm 0.011, -0.016 \pm 0.008, 0.002 \pm 0.008$ for the 4 tomographic bins, which indicates statistical compatibility with the truth. The models are a good fit to the data, with $\chi^2 = 1.29, 0.67, 0.72, 0.63$ for the RedMaGiC sample, and $\chi^2 = 1.19, 1.20, 0.58, 0.88$ for the BOSS/eBOSS sample.

Appendix B

B.1 Validation in simulations

The validity of our methodology and pipeline has been tested in the Buzzard N-body simulation, introduced in Section 4.8. The measurements of redshift distributions using both phenotypes and clustering were validated in simulations to ensure unbiased estimates with respect to the true redshift distributions. The MagLim sample has been recreated in the Buzzard simulations, as described in Section 4. The sample selection has been altered to reproduce as faithfully as possible the number density and color distributions of the data.

To replicate the SOMPZ methodology, we selected one of the 300 deep fields (and the correspondent Balrog and redshift samples) produced for the SOMPZ method uncertainty characterisation presented in section 7.2.1. We then proceeded to perform the 3sDir analytical sample variance estimation for that one specific realisation. The geometry and resolution of the SOM used in simulations are the same as the ones used in data. The major difference with the methodology applied in data is that we used as redshifts the true values from the Buzzard simulations, and we considered as redshift sample only the galaxies located in deep sample of the size of the COSMOS field (1.38 $\deg^2$), therefore not recreating three different samples. The *3sDir* evaluation of the uncertainty in simulation has been modified accordingly, so as to account for the reduced area of the redshift sample. This means that the shot noise contribution to the total error budget in simulations is slightly larger than the one data. The SOMPZ redshift distributions, and their uncertainties estimated through the *3sDir* method, are in agreement with the true distribution, as shown in Figure B.1.

We also repeated in simulations the same procedure as for data also for the WZ estimates. We created a mock BOSS/eBOSS catalog to use as a reference sample. As in data, also in simulations the BOSS/eBOSS sample is divided into 50 bins spanning the $0.1 < z < 1.1$ range of the catalog (width

$\Delta z \sim 0.02$). Before proceeding with combining the SOMPZ and WZ information through the combined likelihood, the compatibility between SOMPZ and WZ was checked. This was tested by inferring the windowed means and widths of the WZ and SOMPZ redshift estimates, following (Gatti & Giannini et al., 2020a). The window has been determined such that magnification effects related to the WZ measurements can be neglected. As for WZ, we used a "simple " estimator for the redshift distribution, inverting Eq. 7.5 and ignoring magnification effects (this is possible as we are considering only windowed quantities). The means and widths computed in this way for the two methods were compatible within statistical (and systematic) error, hence the SOMPZ and WZ could be combined together.

The posterior obtained in simulations from multiplying the two likelihoods is shown in Figure B.1, in which the effect of the combination immediately stands out: the additional information from clustering redshifts places a tight constraint on the shape of the $n(z)$, while still being in agreement with the true distribution. This larger constraining power derives from the fact in clustering the number density for each redshift bin correlates across neighbouring bins, which restrains the joint likelihood to prefer smoother realisations and reject the ones with more uncorrelated values of clustering.

It is also worth mentioning that a full cosmological analysis was run using the datavector and SOMPZ+WZ $n(z)$in the Buzzard simulations, fully recovering the input value.

B.2 Cosmological and nuisance parameters

In Table B.1 are listed all the parameters varied in our fiducial analysis.

B.3 Redshift uncertainties sampling strategy

How redshift uncertainties are propagated in the cosmological analysis can have an impact on the final result. In this section we discuss different strategies to marginalise over the redshift uncertainties of our sample during the cosmological inference. Because we have can rely on a full ensemble of $n(z)$ shapes capturing our redshift uncertainties, we can compare three different sampling methods:

- **Shift**: we compress the realisations by computing their average, and marginalise over a shift on the mean;

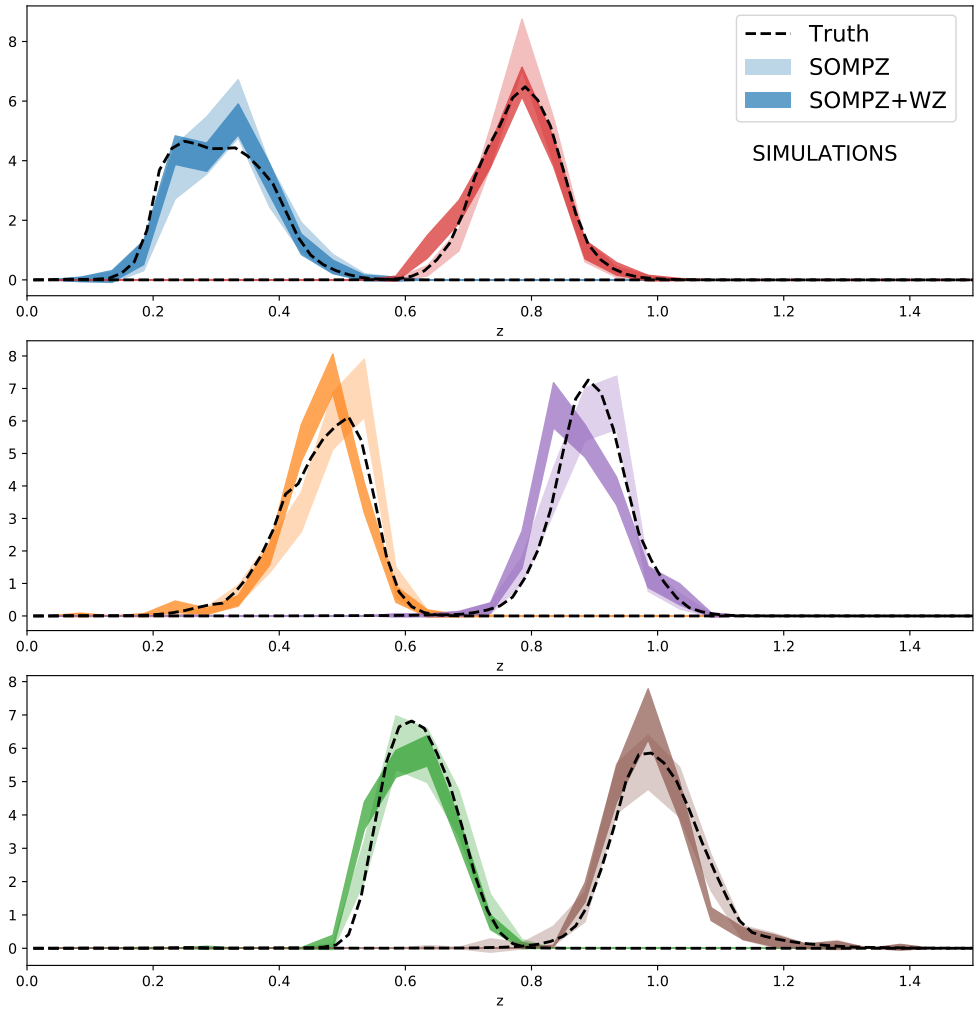

Figure B.1: Estimated $n(z)$ in four tomographic bins using a 12x12 cell deep SOM and 32x32 cell wide SOM trained on Buzzard simulations. In the first row we have bin 1 and 4, in the second bin 2 and 5, and in the third bin 3 and 6. The Redshift sample used here has 100000 galaxies drawn from 1.38 deg^2, the deep sample is drawn from three fields of size 3.32, 3.29, and 1.94 deg^2, respectively from the Buzzard simulated sky catalog. The solid line marks the true value, the outlined violin are the 3sDIR set of $n(z)$ and the solid violins are the realisations once combined with clustering redshifts. We can appreciate the effect of the combined likelihood, resulting in distributions more constrained in terms of shape, and still consistent with the truth.

Table B.1: The parameters and their priors used in the fiducial MagLim ΛCDM and wCDM analyses. The parameter w is fixed to -1 in ΛCDM. Square brackets denote a flat prior, while parentheses denote a Gaussian prior of the form $\mathcal{N}(\mu, \sigma)$.

Parameter	Fiducial	Prior
Cosmology		
$\Omega_{\rm m}$	0.3	[0.1, 0.9]
$A_{\rm s}10^9$	2.19	[0.5, 5.0]
$n_{\rm s}$	0.97	[0.87, 1.07]
w	-1.0	[-2, -0.33]
$\Omega_{\rm b}$	0.048	[0.03, 0.07]
h_0	0.69	[0.55, 0.91]
$\Omega_\nu h^2 10^3$	0.83	[0.6, 6.44]
Linear galaxy bias		
b_i	$1.5, 1.8, 1.8, 1.9, 2.3, 2.3$	[0.8,3.0]
Lens magnification		
C_1	0.43	(0.43, 0.51)
C_2	0.30	(0.30, 0.48)
C_3	1.75	(1.75, 0.39)
C_4	1.94	(1.94, 0.35)
C_5	1.56	(1.56, 0.71)
C_6	2.96	(2.96, 0.95)
Lens photo-z		
$\Delta z_{\rm l}^1$	0.0	$(0.0, 0.0164)$
$\Delta z_{\rm l}^2$	0.0	$(0.0, 0.0100)$
$\Delta z_{\rm l}^3$	0.0	$(0.0, 0.0085)$
$\Delta z_{\rm l}^4$	0.0	$(0.0, 0.0084)$
$\Delta z_{\rm l}^5$	0.0	$(0.0, 0.0094)$
$\Delta z_{\rm l}^6$	0.0	$(0.0, 0.0116)$
$\sigma z_{\rm l}^1$	1.0	$(1.0, 0.0639)$
$\sigma z_{\rm l}^2$	1.0	$(1.0, 0.0624)$
$\sigma z_{\rm l}^3$	1.0	$(1.0, 0.0315)$
$\sigma z_{\rm l}^4$	1.0	$(1.0, 0.0409)$
$\sigma z_{\rm l}^5$	1.0	$(1.0, 0.0515)$
$\sigma z_{\rm l}^6$	1.0	$(1.0, 0.0650)$
Intrinsic alignment		
a_i $(i \in [1, 2])$	0.7, -1.36	$[-5, 5\ ]$
η_i $(i \in [1, 2])$	-1.7, -2.5	$[-5, 5\ ]$
$b_{\rm TA}$	1.0	$[0, 2]$
z_0	0.62	Fixed
Source photo-z		
$\Delta z_{\rm s}^1$	0.0	$(0.0, 0.018)$
$\Delta z_{\rm s}^2$	0.0	$(0.0, 0.013)$
$\Delta z_{\rm s}^3$	0.0	$(0.0, 0.006)$
$\Delta z_{\rm s}^4$	0.0	$(0.0, 0.013)$
Shear calibration		
m^1	0.0	$(-0.006, 0.008)$
m^2	0.0	$(-0.010, 0.013)$
m^3	0.0	$(-0.026, 0.009)$
m^4	0.0	$(-0.032, 0.012)$

- **Shift and stretch**: we compress the realisations by computing their average, and marginalise over both a shift on the mean and on a stretch on the width;
- **Full shape**: we provide as input all the produced realisations and we rank them by one of their properties using the *Hyperrank* method (Cordero et al., 2020), marginalising over the full shape of the distributions.

Using only *shifts* is the methodology usually adopted to model redshift uncertainties in weak lensing sample, as the weak lensing kernel is mostly sensitive to the mean of the redshift distributions. On the other hand, clustering and galaxy-galaxy lensing measurements are also very sensitive to the width of the lens redshift distributions; therefore, the *shift and stretch* approach is preferred. The full shape marginalisation, in theory, is more accurate, because it accounts for the uncertainties in the higher order moments of the distribution; however, depending on the science case, it might not make a huge impact on the final constraints. The full shape marginalisation is implemented via hyperrank (Cordero et al., 2022), which is an algorithm that orders realisations of the ensemble according to a parameter, which facilitates the sampling and marginalization over the $n(z)$ ensemble within the cosmological likelihood Markov chains. Hyperrank was also implemented for the WL sources, although it had a negligible impact on the results. The quantity chosen for the ranking in that case was the mean. We decided for this case it would be more appropriate to perform the optimised ranking of the realisation by the 68% sigma rather than the mean, and we tested it indeed improved the performance of the sampling. To test the different sampling strategies, we built a synthetic noiseless data vector based on theory predictions at fixed cosmology and we used as $n(z)$ the realisations average of the SOMPZ+WZ estimates in data. We then marginalised over redshift uncertainties using the three approaches aforementioned. We performed this test both using 4 or 6 lens bins, although here we are just going to show the posteriors obtained with 4 bins as they are not qualitatively different from the ones with 6 bins. The results of this test are shown in Figure B.2, where we show the posterior of σ_8, Ω_m and for sake of simplicity, two out of the four galaxy-matter linear biases.

Focusing on the shift and shift+stretch contours, one can notice that the width of the contour in the direction perpendicular to the degeneration axis is larger for the shift+stretch. This is related to impact of the additional marginalisation over the width of the distributions. One caveat is that in our marginalisation scheme (as adopted in the main DES Y3 2x2pt analysis), we are implicitly neglecting correlations between the uncertainties in the mean

and widths of the distributions, which usually show a certain degree of correlation (from 10% to 30%, depending from the tomographic bin). These are neglected, which might translate in a slight overestimation of our constraints. When marginalising over the uncertainties using the hyperrank framework, on the other hand, such correlations are implicitly accounted for. Indeed, one can notice that the hyperrank posteriors are slightly tighter than the shift or shift-stretch posteriors.

Unfortunately, we did not manage to successfully apply hyperrank to the data. When performing the cosmological analysis on data using hyperrank, we found significantly less smooth posteriors compared to our tests on simulations. A similar behaviour has also been found when applying hyperrank to the DES Y3 source sample Amon et al. (2022), and it has been interpreted as a consequence of a possible larger degree of complexity of the redshift distributions of our data compared to simulations. We attempted both to artificially smooth our $n(z)$ and to increase the number of samples from the SOMPZ+WZ method, without reaching a satisfactory level. Due to the very high computational cost of running a cosmological chain using hyperrank, we could only test a few different levels of smoothing before deciding to abandon hyperrank for the present work, and choose the shift+stretch as photo-z uncertainty marginalisation methodology. For DES Y6, we plan to apply several tools that will speed up our cosmological inference, enabling more tests on hyperrank, which has great potential and its implementation is a goal for the DES Y6 analysis.

B.3.1 Cosmological constraints with clipped $n(z)$ tails

Here we test whether the difference between DNF+WZ and SOMPZ+WZ constraints (Fig. 8.1) were only due to the different treatment of redshift outliers and of the tails of the redshift distributions. We artificially removed the tails from the DNF+WZ and SOMPZ+WZ $n(z)$ (i.e., we set the distributions to zero), and repeated our cosmological analysis. We used as definition of the tails the same interval used to calibrate the DNF distribution with the WZ constraints adopted in Porredon et al. (2021a). Results for the ΛCDM case, 4 bins and fixed magnification are shown in Fig. B.3. By removing the tails, both posteriors are shifted, which means that the calibration of the tails of the redshift distribution is important for our cosmological analysis. Since the two posteriors are shifted but they still do not overlap, we can assume that the differences in the bulk of the redshift distributions inferred by two methods is also crucially driving the differences at the constraints level seen in Fig. 8.1.

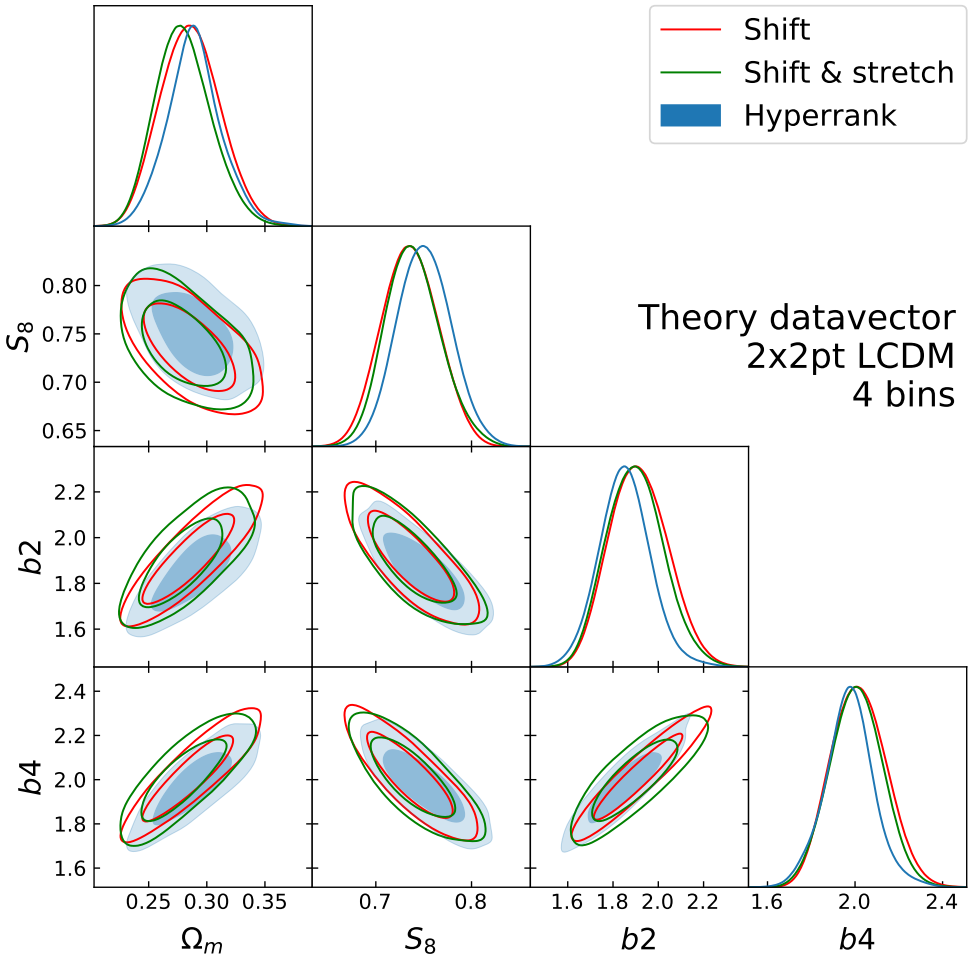

Figure B.2: Posterior distributions of the cosmological parameters $\Omega_{\rm m}$, S_8, and two out of four of the galaxy-matter biases (b_2, b_4) for the ΛCDM analysis involving 4 bins and fixed magnification parameters. These analyses have been obtained assuming a theoretical datavector and adopting different marginalisation schemes on the redshift distribution of the lens sample.

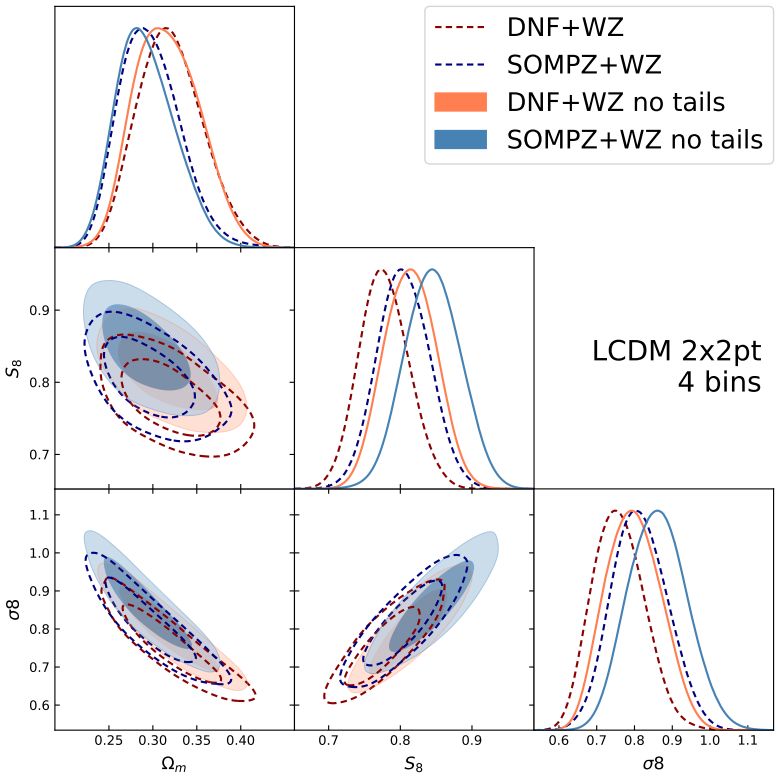

Figure B.3: Same as the left panel of Fig. 8.1, but with two additional posteriors overplotted representing the constraints obtained using the redshift distributions with "clipped" tails.

Model	bins	S_8	Ω_m	σ_8	w	p-value
ΛCDM	4 bins	0.76 ± 0.04	0.31 ± 0.04	0.75 ± 0.07	-	0.052
ΛCDM	6 bins	0.77 ± 0.04	0.28 ± 0.03	0.80 ± 0.07	-	0.033
wCDM	4 bins	0.77 ± 0.05	0.32 ± 0.05	0.76 ± 0.08	-1.0 ± 0.3	0.061
wCDM	6 bins	0.80 ± 0.04	0.33 ± 0.04	0.77 ± 0.06	-0.7 ± 0.2	0.045

Table B.2: Constraints on the cosmological parameters S_8, Ω_m, and σ_8 using flat priors on the magnification parameters. For each parameter we report the mean of the posterior and the 68 per cent confidence interval. We also report the PPD goodness-of-fit p-value.

B.4 Flat prior magnification

We tested the impact on the 2x2pt measurement of marginalising over the magnification parameters using wide flat priors, instead of the BALROG estimated Gaussian priors. Here we report the ΛCDM and wCDM results using free magnification with flat priors. All the other parameters and priors in the analysis are left unchanged.

We show the posteriors on the cosmological parameters Ω_m, and S_8 in Fig. B.4, and we report the values in Table B.2. For ΛCDM, both the 4 and 6 bins case the posteriors prefer lower S_8 values compared to the fixed magnification case. This is a consequence of an interesting degeneracy between S_8 and the magnification parameters, and the fact that some of the magnification posteriors differ from the `Balrog` estimates. This is shown in Fig. B.5. The discrepancy between the two is evident in bin 2 and 3. While the `Balrog` estimates are still in the tails of our data posteriors, indicating that these difference might only be a statistical fluke, the data prefers significantly higher values, which are atypical for a magnitude-limited sample at these redshifts. One explanation might be related to the MAGLIM selection, which is not a pure magnitude-limited selection, but involves DNF estimates; in case of a high rate of photo-z outliers, the selection might significantly differ from the ideal magnitude-limited case, admitting larger values of the magnification coefficient. On the other hand, if the high values preferred by the data are simply a result of a statistical fluctuations, including the cross-correlation measurements between different lens bins in our data vector could help to improve our constraints, as these measurements are mostly sensitive to redshift outliers and magnification parameters. We leave this extension of our data vector to future works (e.g., DES Y6).

Freeing the magnification parameters shifts w towards the upper edge of the prior ($w = -0.3$), and contrary to the ΛCDM case, leaves unaltered the constraints on S_8. This indicates that in the wCDM case, the values of the

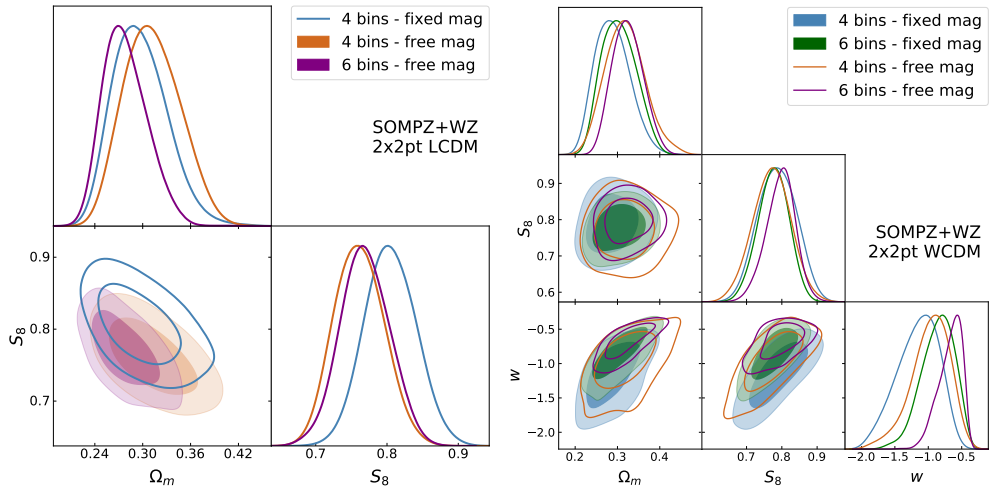

Figure B.4: Posterior distributions of the cosmological parameters $\Omega_{\rm m}$, and S_8 and w. The 2D marginalised contours in these figures show the 68 per cent and 95 per cent confidence levels. Left, ΛCDM: 1) 4 bins and fixed magnification parameters; 2) 4 bins and free magnification parameters with flat priors; 3) 6 bins and free magnification parameters with flat priors. Right, wCDM: 1) 4 bins and fixed magnification parameters; 2) 6 bins and fixed magnification parameters, 3) 4 bins and free magnification parameters with flat priors; 4) 6 bins and free magnification parameters with flat priors.

magnification parameters are more degenerate with w rather than with S_8.

The statistical distance with respect to the Planck measurement has also been computed. Interestingly, for the flat prior case, when including the last two bins the tension rises to 2.96σ for ΛCDM with free magnification. We do not think the increment in the tension for the 6 bins case is related to the particular values of the magnification parameters of the last two bins, as the data prefer values close to the `Balrog` prior (Fig. B.5). For the wCDM case, the tensions are milder (0.46σ and 2.29σ for 4 and 6 bins, respectively, at fixed magnification).

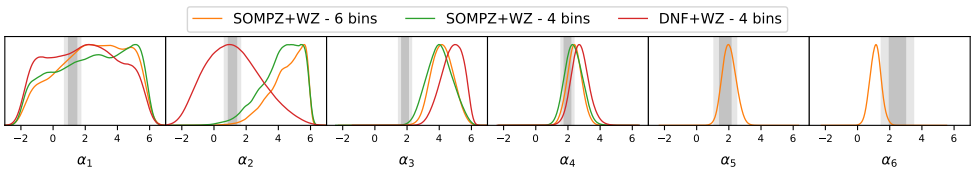

Figure B.5: Posterior distributions of magnification coefficients for 4 and 6 bins. The grey bands represent 68 per cent and 95 per cent confidence interval of the the prior values from `Balrog`; the colored lines represent the posteriors from the ΛCDM analyses, comparing SOMPZ+WZ and DNF+WZ posteriors.

Bibliography

Abbott, T. M. C., Abdalla, F. B., Allam, S., et al., 2018, ApJS, 239, 2, 18

Abbott, T. M. C., Aguena, M., Alarcon, A., et al., 2022, Phys. Rev. D, 105, 2, 023520

Abolfathi, B., Aguado, D. S., Aguilar, G., et al., 2018, ApJS, 235, 2, 42

Aghanim, N., Akrami, Y., Ashdown, M., et al., 2020, Astronomy & Astrophysics, 641, A6

Ahumada, R., Allende Prieto, C., Almeida, A., et al., 2019, arXiv e-prints, arXiv:1912.02905

Aihara, H., Arimoto, N., Armstrong, R., et al., 2018, PASJ, 70, S4

Alam, S., de Mattia, A., Tamone, A., et al., 2020, arXiv e-prints, arXiv:2007.09004

Alarcon, A., Gaztanaga, E., Eriksen, M., et al., 2020, arXiv e-prints, arXiv:2007.11132

Alarcon, A., Sánchez, C., Bernstein, G. M., Gaztañaga, E., 2019, arXiv e-prints, arXiv:1910.07127

Albrecht, A., Bernstein, G., Cahn, R., et al., 2006, arXiv e-prints, astro-ph/0609591

Amon, A., Gruen, D., Troxel, M. A., et al., 2022, Phys. Rev. D, 105, 2, 023514

Bates, D. J., Tojeiro, R., Newman, J. A., et al., 2019, MNRAS, 486, 3059

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

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

Benjamin, J., Van Waerbeke, L., Heymans, C., et al., 2013, MNRAS, 431, 2, 1547

Bennett, C., 2015, Cosmological Constraints Using Planck 2015 and WMAP Data, NASA ADAP Proposal

Bernardeau, F., Colombi, S., Gaztañaga, E., Scoccimarro, R., 2002, Phys. Rep., 367, 1

Blanton, M. R., Bershady, M. A., Abolfathi, B., et al., 2017, AJ, 154, 1, 28

Bonnett, C., Troxel, M. A., Hartley, W., et al., 2016a, Phys. Rev. D, 94, 4, 042005

Bonnett, C., Troxel, M. A., Hartley, W., et al., 2016b, Phys. Rev. D, 94, 4, 042005

Bridle, S., Shawe-Taylor, J., Amara, A., et al., 2009, Annals of Applied Statistics, 3, 6

Buchs, R., Davis, C., Gruen, D., et al., 2019, MNRAS, 489, 1, 820

Cawthon, R., Davis, C., Gatti, M., et al., 2018, MNRAS, 481, 2427

Cawthon, R., Elvin-Poole, J., Porredon, A., et al., 2020, arXiv e-prints, arXiv:2012.12826

Choi, A., Heymans, C., Blake, C., et al., 2016, MNRAS, 463, 3737

Conroy, C., Wechsler, R. H., Kravtsov, A. V., 2006, ApJ, 647, 201

Cordero, J. P., Harrison, I., Rollins, R. P., et al., 2022, MNRAS, 511, 2, 2170

Cordero, J. P., et al., 2020, in preparation

Cunha, C. E., Huterer, D., Busha, M. T., Wechsler, R. H., 2012, MNRAS, 423, 1, 909

Davis, C., Gatti, M., Vielzeuf, P., et al., 2017, arXiv e-prints, arXiv:1710.02517

Davis, C., Rozo, E., Roodman, A., et al., 2018, MNRAS, 477, 2196

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

Dawson, K. S., Kneib, J.-P., Percival, W. J., et al., 2016, AJ, 151, 2, 44

Dawson, K. S., Schlegel, D. J., Ahn, C. P., et al., 2013, AJ, 145, 1, 10

De Vicente, J., Sánchez, E., Sevilla-Noarbe, I., 2016a, MNRAS, 459, 3, 3078

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

DeRose, J., Wechsler, R. H., Becker, M. R., et al., 2019, arXiv e-prints, arXiv:1901.02401

DeRose, J., Wechsler, R. H., Becker, M. R., et al., 2022, Phys. Rev. D, 105, 12, 123520

Desjacques, V., Jeong, D., Schmidt, F., 2018, Phys. Rep., 733, 1

Doux, C., Baxter, E., Lemos, P., et al., 2021, MNRAS, 503, 2, 2688

Eisenstein, D. J., Weinberg, D. H., Agol, E., et al., 2011, AJ, 142, 3, 72

Elvin-Poole, J., et al., 2021, To be submitted to MNRAS

Eriksen, M., Alarcon, A., Gaztanaga, E., et al., 2019, MNRAS, 484, 3, 4200

Euclid Collaboration, Desprez, G., Paltani, S., et al., 2020, A&A, 644, A31

Everett, S., Yanny, B., Kuropatkin, N., et al., 2020a, arXiv e-prints, arXiv:2012.12825

Everett, S., et al., 2020b, in preparation

Fang, X., Krause, E., Eifler, T., MacCrann, N., 2020, J. Cosmology Astropart. Phys., 2020, 5, 010

Fixsen, D. J., 2009, ApJ, 707, 2, 916

Flaugher, B., Diehl, H. T., Honscheid, K., et al., 2015, AJ, 150, 5, 150

Garilli, B., Guzzo, L., Scodeggio, M., et al., 2014, VizieR Online Data Catalog, J/A+A/562/A23

Gatti, M., Giannini, G., Bernstein, G. M., et al., 2020a, arXiv e-prints, arXiv:2012.08569

Gatti, M., Sheldon, E., Amon, A., et al., 2020b, arXiv e-prints, arXiv:2011.03408

Gatti, M., Vielzeuf, P., Davis, C., et al., 2018, MNRAS, 477, 1664

Gatti, M., et al., 2020c, in preparation

Górski, K. M., Hivon, E., 2011, HEALPix: Hierarchical Equal Area isoLatitude Pixelization of a sphere, Astrophysics Source Code Library, record ascl:1107.018

Gschwend, J., Rossel, A. C., Ogando, R. L. C., et al., 2018, Astronomy and Computing, 25, 58

Gunn, J. E., Siegmund, W. A., Mannery, E. J., et al., 2006, AJ, 131, 4, 2332

Handley, W. J., Hobson, M. P., Lasenby, A. N., 2015a, MNRAS, 450, L61

Handley, W. J., Hobson, M. P., Lasenby, A. N., 2015b, MNRAS, 453, 4, 4384

Hartley, W. G., Choi, A., Amon, A., et al., 2020, arXiv e-prints, arXiv:2012.12824

Heymans, C., Van Waerbeke, L., Miller, L., et al., 2012, MNRAS, 427, 1, 146

Hildebrandt, H., Arnouts, S., Capak, P., et al., 2010, A&A, 523, A31

Hildebrandt, H., van den Busch, J. L., Wright, A. H., et al., 2020, arXiv e-prints, arXiv:2007.15635

Hildebrandt, H., Viola, M., Heymans, C., et al., 2017, MNRAS, 465, 1454

Hoyle, B., Gruen, D., Bernstein, G. M., et al., 2018a, MNRAS, 478, 1, 592

Hoyle, B., Gruen, D., Bernstein, G. M., et al., 2018b, MNRAS, 478, 592

Hubble, E., 1929, Proceedings of the National Academy of Science, 15, 3, 168

Huff, E., Mandelbaum, R., 2017a, arXiv e-prints

Huff, E., Mandelbaum, R., 2017b, ArXiv e-prints

Huterer, D., Cunha, C. E., Fang, W., 2013, MNRAS, 432, 4, 2945

Huterer, D., Takada, M., Bernstein, G., Jain, B., 2006, MNRAS, 366, 101

Jain, B., Taylor, A., 2003, Phys. Rev. Lett., 91, 14, 141302

Jarvis, M., Bernstein, G. M., Amon, A., et al., 2020, arXiv e-prints, arXiv:2011.03409

Jarvis, M. J., Bonfield, D. G., Bruce, V. A., et al., 2013, MNRAS, 428, 1281

Johnson, A., Blake, C., Amon, A., et al., 2017, MNRAS, 465, 4118

Joudaki, S., Hildebrandt, H., Traykova, D., et al., 2019, arXiv e-prints, arXiv:1906.09262

Kohonen, T., 1982, Biological Cybernetics, 43, 1, 59, ISSN 0340-1200

Kuijken, K., Heymans, C., Hildebrandt, H., et al., 2015, MNRAS, 454, 4, 3500

Laigle, C., McCracken, H. J., Ilbert, O., et al., 2016, ApJS, 224, 24

Landy, S. D., Szalay, A. S., 1993, ApJ, 412, 64

Laureijs, R., Amiaux, J., Arduini, S., et al., 2011, ArXiv e-prints

Le Fèvre, O., Cassata, P., Cucciati, O., et al., 2013a, A&A, 559, A14

Le Fèvre, O., Cassata, P., Cucciati, O., et al., 2013b, A&A, 559, A14

Lehmann, B. V., Mao, Y.-Y., Becker, M. R., Skillman, S. W., Wechsler, R. H., 2017, ApJ, 834, 37

Lemos, P., Raveri, M., Campos, A., et al., 2021, MNRAS, 505, 4, 6179

Lidman, C., Tucker, B. E., Davis, T. M., et al., 2020, MNRAS, 496, 1, 19

Lilly, S. J., Le Brun, V., Maier, C., et al., 2009a, ApJS, 184, 2, 218

Lilly, S. J., Le Brun, V., Maier, C., et al., 2009b, ApJS, 184, 218

Lima, M., Cunha, C. E., Oyaizu, H., Frieman, J., Lin, H., Sheldon, E. S., 2008, MNRAS, 390, 118

Limber, D. N., 1953, ApJ, 117, 134

LSST Science Collaboration, Abell, P. A., Allison, J., et al., 2009, arXiv e-prints, arXiv:0912.0201

Lupton, R. H., Gunn, J. E., Szalay, A. S., 1999, AJ, 118, 1406

MacCrann, N., Becker, M. R., McCullough, J., et al., 2020, arXiv e-prints, arXiv:2012.08567

Masters, D. C., Stern, D. K., Cohen, J. G., et al., 2017, ApJ, 841, 111

Masters, D. C., Stern, D. K., Cohen, J. G., et al., 2019, ApJ, 877, 2, 81

Matthews, D. J., Newman, J. A., 2010, ApJ, 721, 456

McCracken, H. J., Milvang-Jensen, B., Dunlop, J., et al., 2012, A&A, 544, A156

McQuinn, M., White, M., 2013, MNRAS, 433, 2857

Ménard, B., Scranton, R., Schmidt, S., et al., ????, ArXiv e-prints, arXiv:1303.4722

Moessner, R., Jain, B., 1998, MNRAS, 294, L18

Morganson, E., Gruendl, R. A., Menanteau, F., et al., 2018, PASP, 130, 989, 074501

Morrison, C. B., Hildebrandt, H., Schmidt, S. J., et al., 2017, MNRAS, 467, 3576

Myles, J., Alarcon, A., Amon, A., et al., 2020, arXiv e-prints, arXiv:2012.08566

Narayan, R., 1989, ApJ, 339, L53

Newman, J. A., 2008, ApJ, 684, 88-101

Pandey, S., Krause, E., DeRose, J., et al., 2021, arXiv e-prints, arXiv:2105.13545

Planck Collaboration, 2018, arXiv e-prints, arXiv:1807.06209

Planck Collaboration, 2019, arXiv e-prints, arXiv:1907.12875

Porredon, A., Crocce, M., Elvin-Poole, J., et al., 2021a, arXiv e-prints, arXiv:2105.13546

Porredon, A., Crocce, M., Fosalba, P., et al., 2021b, Phys. Rev. D, 103, 4, 043503

Porredon, A., Crocce, M., Fosalba, P., et al., 2021c, Phys. Rev. D, 103, 4, 043503

Prakash, A., Licquia, T. C., Newman, J. A., et al., 2016, ApJS, 224, 2, 34

Prat, J., Baxter, E., Shin, T., et al., 2019, MNRAS, 487, 1, 1363

Prat, J., Blazek, J., Sánchez, C., et al., 2022, Phys. Rev. D, 105, 8, 083528

Rahman, M., Ménard, B., Scranton, R., Schmidt, S. J., Morrison, C. B., 2015, MNRAS, 447, 4, 3500

Rahman, M., Mendez, A. J., Ménard, B., et al., 2016a, MNRAS, 457, 4, 3912

Rahman, M., Mendez, A. J., Ménard, B., et al., 2016b, MNRAS, 460, 1, 163

Raichoor, A., Comparat, J., Delubac, T., et al., 2017, MNRAS, 471, 4, 3955

Raveri, M., Doux, C., 2021, Phys. Rev. D, 104, 4, 043504

Reid, B., Ho, S., Padmanabhan, N., et al., 2016, MNRAS, 455, 2, 1553

Rodríguez-Monroy, M., Weaverdyck, N., Elvin-Poole, J., et al., 2022, MNRAS, 511, 2, 2665

Rozo, E., Rykoff, E. S., Abate, A., et al., 2016, MNRAS, 461, 1431

Sánchez, C., Bernstein, G. M., 2019a, MNRAS, 483, 2, 2801

Sánchez, C., Bernstein, G. M., 2019b, MNRAS, 483, 2801

Sánchez, C., Carrasco Kind, M., Lin, H., et al., 2014, MNRAS, 445, 1482

Sánchez, C., Prat, J., Zacharegkas, G., et al., 2022a, Phys. Rev. D, 105, 8, 083529

Sánchez, C., Prat, J., Zacharegkas, G., et al., 2022b, Phys. Rev. D, 105, 8, 083529

Sánchez, C., Raveri, M., Alarcon, A., Bernstein, G. M., 2020, arXiv e-prints, arXiv:2004.09542

Schmidt, S. J., Malz, A. I., Soo, J. Y. H., et al., 2020, MNRAS, 499, 2, 1587

Schmidt, S. J., Ménard, B., Scranton, R., Morrison, C., McBride, C. K., 2013, MNRAS, 431, 3307

Scodeggio, M., Guzzo, L., Garilli, B., et al., 2018a, A&A, 609, A84

Scodeggio, M., Guzzo, L., Garilli, B., et al., 2018b, A&A, 609, A84

Scottez, V., Benoit-Lévy, A., Coupon, J., Ilbert, O., Mellier, Y., ????, ArXiv e-prints, arXiv:1705.02629

Scottez, V., Mellier, Y., Granett, B. R., et al., 2016, MNRAS, 462, 1683

Secco, L. F., Samuroff, S., Krause, E., et al., 2022, Phys. Rev. D, 105, 2, 023515

Sevilla, I., Armstrong, R., Bertin, E., et al., 2011, arXiv e-prints, arXiv:1109.6741

Sevilla-Noarbe, I., Bechtol, K., Carrasco Kind, M., et al., 2020, arXiv e-prints, arXiv:2011.03407

Sheldon, E. S., Huff, E. M., 2017a, ApJ, 841, 24

Sheldon, E. S., Huff, E. M., 2017b, ApJ, 841, 24

Smee, S. A., Gunn, J. E., Uomoto, A., et al., 2013, AJ, 146, 2, 32

Smoot, G. F., 1992, Highlights of Astronomy, 9, 281

Speagle, J. S., Eisenstein, D. J., 2017a, MNRAS, 469, 1, 1186

Speagle, J. S., Eisenstein, D. J., 2017b, MNRAS, 469, 1, 1205

Springel, V., 2005, MNRAS, 364, 1105

Stebbins, A., 1996, arXiv e-prints, astro-ph/9609149

Suchyta, E., Huff, E. M., Aleksić, J., et al., 2016, MNRAS, 457, 1, 786

Tanaka, M., Coupon, J., Hsieh, B.-C., et al., 2018, PASJ, 70, S9

Tessore, N., Harrison, I., 2020, The Open Journal of Astrophysics, 3, 1, 6

Troxel, M. A., MacCrann, N., Zuntz, J., et al., 2018, Phys. Rev. D, 98, 4, 043528

van den Busch, J. L., Hildebrandt, H., Wright, A. H., et al., 2020, A&A, 642, A200

Villumsen, J. V., Freudling, W., da Costa, L. N., 1997, ApJ, 481, 578

Wright, A. H., Hildebrandt, H., van den Busch, J. L., Heymans, C., 2019, arXiv e-prints, arXiv:1909.09632

Zuntz, J., Paterno, M., Jennings, E., et al., 2015, Astronomy and Computing, 12, 45