Higher resolution photometric redshifts for cosmological surveys

Alex Alarcon Gonzalez

Higher resolution photometric redshifts for cosmological surveys

Alex Alarcon Gonzalez

Tesi Doctoral
Programa de Doctorat en Física
Director: Dr. Enrique Gaztañaga Balbas
Tutor: Dr. Enrique Fernández Sánchez

Departament de Física
Facultat de Ciències
Universitat Autònoma de Barcelona
2019

Contents

Introduction

In 1915 Albert Einstein published his theory of General Relativity (GR), setting the beginning of a new branch of astronomy, modern cosmology. Under the assumption known as *Cosmological Principle*, where the universe is equivalent in every position, known as *homogeneity*, and the same regardless of the direction, known as *isotropy*, GR provides a framework to explain the origin, evolution and fate of the universe given its energy-matter components are understood. Nowadays, the standard model of cosmology is the so called *Lambda*-CDM (ΛCDM), which describes remarkably well a wide range of galactic and extragalactic observations, but only after including two major components called *dark matter* and *dark energy*. Dark energy is an hypothetical form of energy that tends to accelerate the expansion of the universe, while dark matter is a form of matter different from common (baryonic) observable matter, also of unknown origin. Under this paradigm, the best recent measurements find that the universe energy content is 68% dark energy, 27% dark matter and only 5% luminous ordinary matter and energy, which means the nature of most of the matter-energy content of the universe is currently not well understood.

Galaxy surveys have become an excellent source of information to study the evolution and formation history of the universe. These are maps of the position and redshift of galaxies which measure the large scale structure of the universe and probe its geometry and composition, and ultimately its evolution and fate. Since the first survey of galaxies, namely the CfA redshift survey which dates from 1977 and measured a few thousand galaxies, we have arrived to the current era of massive surveys with over a hundred million galaxies detected, a number that will be further extended by the next generation of future ground and space missions that will start this next decade. Redshift surveys can be broadly classified between spectroscopic and photometric. The former disperse the light of a galaxy to obtain the flux intensity as a function of wavelength, and the redshift is determined by the location of emission and absorption lines while the latter images galaxies by placing different colored

filters at different wavelengths. Spectroscopy is more accurate but has lower efficiency at obtaining the redshift of objects while imaging surveys obtain measurement for many objects at the same time but with lower spectral resolution. The main focus of this thesis is the measurement of photometric redshifts using imaging galaxy surveys and its applications to extract cosmological information. The first part of this thesis sets the cosmological framework and describes the main techniques and galaxy survey data used in this thesis.

Galaxy clustering is one of the most powerful probes of the cosmic expansion and its growth history and it is relatively easy to obtain, since only the positions and redshift of galaxies have to be observed. A 3D distribution map of galaxies at different redshifts (epochs) contains crucial information on the evolution of the universe, but galaxies are not unbiased tracers of the total matter density of the universe, which is mostly composed by dark matter. Redshift space distortions produced by the galaxies peculiar velocities break the degeneracy between this galaxy bias and the growth rate of structures in the universe. However, since we can only observe one single realization of the universe, measurements on large scales are limited by cosmic variance, and measurements of a smaller angular line of sight will be affected by the cosmic large scale structure, known as sample variance. Using multiple tracers over the same area with different galaxy bias allows to cancel this sample variance and increase the amount of information one can extract from these surveys.

In the second part of this thesis we forecast the cosmological and growth history constraints for a redshift survey with high redshift precision, which can come either from spectroscopy or from many narrow band images. Instead of using a 3D clustering measurement, we divide the survey into a series of thin redshift slices where the radial information is measured from the cross correlations between different redshift bins. This allows to naturally combine these measurements with weak lensing analysis. In particular, we explore how splitting the survey into two overlapping subsamples with different bias can cancel the sample variance improve the cosmological information that one could extract.

Physics of the Accelerating Universe survey (PAUS) is a unique imaging redshift survey with a set of 40 narrow band filters of 12.5nm FWHM. Having such high spectral sampling allows to obtain a much better photometric redshift precision than typical imaging broad band surveys. Obtaining redshift measurements for this singular data is the topic of the third chapter. We develop two photometric redshift estimation algorithms that linearly combine templates of the galaxy spectral energy distribution to properly model the continuum and emission line flux of each galaxy. The first photo-z method includes a minimization algorithm that finds the maximum likelihood parameters for

each model, while in the second one we develop an algorithm that finds the Bayesian integral over all parameter combinations. We create a mock galaxy catalog with simulated narrow band fluxes and introduce several effects from real data like noise levels or observed systematic errors in the data. The mock catalog is used to establish the expected photo-z precision under ideal conditions, validate different methods to remove systematic effects and compare the performance of different photo-z algorithms. The measurement of highly precise redshifts enables the science of PAUS and can also be used to calibrate the redshift distribution of lensing surveys.

One of the leading systematic effects in current and future lensing surveys is the uncertainty in the redshift distribution of the source galaxies. Weak gravitational lensing measures the small distortions present in the shapes of millions of background galaxies, which is produced by foreground matter deflecting the light traveling from these sources to us. This observational probe contains significant cosmological information, since it measures the foreground mass distribution and the geometry of the universe, but it is also very sensitive to any biases in the mean redshift of the sample. Photometric redshifts in lensing surveys are typically estimated either from the color information of the galaxies or by performing a cross correlation of the positions of the sample with an overlapping tracer with reliable redshift information. Galaxy colors can be compared to those from models of spectral energy distribution at different redshifts or be related to spectroscopic measurements from calibration samples. Clustering redshifts use a tracer sample with either spectroscopic redshifts or high precision redshifts (like photometric luminous red galaxies), and estimate the global redshift distribution of a sample by measuring the relative correlation signal at different redshift bins of the tracer. These to main sources of redshift information have been used independently to obtain two redshift distributions which are used as a cross check or combined using a parameterization of the redshift distribution. In the fourth part of this thesis we present a methodology that naturally combines the information from colors and clustering in a hierarchical bayesian model. We will test this method in the public MICE N-body simulation, a simulation with over 200 million galaxies over a full octant of the sky that contains realistic clustering and galaxy properties.

We conclude at the end with a summary of the main results of the thesis and its implications for the future.

Part I

Preliminars

This first part contains the material needed in order to set the thesis in context, starting with the cosmological framework and motivation, followed by some of the main techniques used to probe cosmology in imaging galaxy surveys and a presentation of the Physics of the Accelerating Universe Survey.

Chapter 1

Preliminars

1.1 Cosmological background

Newton's *Principia* provided the first theoretical framework of gravity which described what humans had observed for a long time, an attractive force that pulled bodies together and explained a large variety of terrestrial and celestial phenomena. At the beginning of last century, after Einstein's theory of special relativity in 1905, it became clear that the idea that gravity was an instantaneous force that acted between bodies had to be formally incorrect since special relativity showed that simultaneity was different in different inertial frames. A few years later, in 1915, Einstein published the theory of General Relativity (GR), which superseded Newton's universal law of gravity and gave us a new understanding over space and time. Newton's law was replaced by a geometrical interpretation where massive bodies changed the curvature of the spacetime fabric. In other words, gravity was geometry. General Relativity was able to successfully predict the anomalous precession of Mercury's perihelion that Newton's gravity could not explain. A few years later, in 1919, Eddington's experiment famously confirmed the deflection of light from stars surrounding the Sun during a total solar eclipse, as a consequence of the curvature of spacetime around Sun, an effect known as gravitational lens. The era of modern cosmology had started with a testable theory of the universe.

The first cosmological model introduced by Einstein in 1917 included a matter dominated universe. To compensate the effect of gravity and recover a static universe, which was the most accepted model at the time, he introduced an extra constant term in the GR equations which would have a repulsive gravity, known as the cosmological constant, usually denoted as Λ. In the 1920s, Friedmann and Lemaître derived an exact solution to Einstein's equations for either an expanding or contracting universe, assuming the universe was homo-

geneous and isotropic, a simplification which becomes approximately true on large cosmological scales, known as the cosmological principle. In 1929, 5 years after discovering *nebulaes* were not part of the Milky Way but were galaxies on their own, Hubble combined distance observations of nearby galaxies using Cepheid variable stars with Slipher measurements of the redshift velocities of these galaxies and found a direct proportionality: galaxies further apart where moving further away from us. This observation was direct evidence for an expanding universe, which vanished the need for a cosmological constant that made the universe static, since Friedmann and Lemaître solutions to Einstein equations already naturally explained an expanding universe.

The observation that the universe was expanding lead to the idea that, going back in time, the universe would have been very small at some point and possibly had an origin in time. If the universe was static, infinite and homogeneously filled by stars the night sky would be bright, while the most simple astronomical observation shows that the night sky is dark. This is known as Olbers' paradox, which indicates at least one of the statements above must be false. An expanding universe with an origin would mean the universe had a finite age, meaning stars had not always been there, hence why the sky is dark. A dynamical universe with a hot origin, commonly known as Big Bang, was reinforced by the discovery of the cosmic microwave background (CMB, 1965), an isotropic relic radiation consistent with a black body emission that had been redshifted by the expansion of the universe to peak at the microwave region of the spectrum. This radiation is the oldest observable radiation, which happened right at the end of the recombination epoch, an epoch when matter and radiation interacted so heavily that photons could not travel freely, until the universe cooled down with the expansion and became transparent, generating a surface of last scattering, the CMB.

Einstein's gravitational field equations allow to predict the evolution of the universe knowing its matter-energy content. In 1930s, Zwicky found studying the velocities of galaxies of the Coma Cluster that the mass measured from luminous matter was not enough to explain the quick motion of galaxies at the edge of the cluster. He inferred there had to be a hidden *dark matter* which accounted for most of the cluster mass to explain why the galaxies were held together with that speed. Although dark matter has not been directly detected yet, there is overwhelming indirect observational evidence of its existence as a major component of the universe. From the shape of galaxy rotation curves, to the abundance of galaxy clusters of a given mass, the shape of the CMB anisotropies and gravitational lensing; all evidence points towards the presence of a *dark* matter which is several times more abundant than luminous matter but has very weak interaction other than gravitational.

Through the 1980s and during the 1990s evidence begun to accumulate from the CMB anisotropy map and in observations from large scale structure clustering of the flatness of the universe through a process named inflation and of the relatively low matter density present in the universe ($15 - 40\%$). This suggested a new energy density component was missing to reconcile with the observations and the idea of a cosmological constant was recovered. At the end of 1990s two independent groups lead by Riess and Perlmutter carrying observations of type Ia supernovae (SN Ia) arrived to the same extraordinary conclusion that the universe was under a late time accelerated expansion. This Nobel Prize discovery brought definitely back the idea that the cosmological constant was not zero, but indeed had to be positive. Nowadays, we refer as *dark energy* to this component of the universe that is driving the late time accelerated expansion. The most obvious choice for dark energy is a cosmological constant, but a negative pressure fluid that evolved over time is still not ruled out by evidence, among other modifications of gravity. We call a universe with Lambda-Cold Dark Matter (ΛCDM) the standard cosmological model. This model has been so far very successful explaining a large range of observations made over the last decades, including the shape of the CMB anisotropies, the distribution of galaxies in the universe known as large scale structure, the redshift-distance positions of type Ia SN or the abundance of light elements.

The current best measurements of the matter-energy density content of the universe in early epochs from the CMB indicate that about 68% of the total energy density comes from dark energy, 26% from dark matter and only 5% comes from baryonic matter. In order to study the evolution and formation history of the universe a large astronomical effort has undergone in the form of galaxy redshift surveys, which probe the late time universe as opposed to the early CMB universe. If dark energy was measured to have a different value which had evolved over time it would rule the cosmological *constant*. A redshift survey is a portion of the sky where the redshift of galaxies and other objects like clusters or quasars are measured to be able to reconstruct the 3D matter distribution. The first redshift survey dates from 1977, the CfA Redshift survey, which initially measured the redshift for 2200 galaxies. Since then, the number of galaxies measured has risen enormously, in part thanks to recent technical developments in observational astronomy. From the thousands of galaxies from CfA we went to over 200,000 galaxies from the 2dF Galaxy Redshift Survey in 2003, over a million galaxies in the Sloan Digital Sky Survey in 2012 to over a hundred million galaxies in the Dark Energy Survey by 2019, to name a few surveys. Future planned surveys such as Euclid, LSST, DESI or WFIRST and will extend the amount of galaxies even further.

Redshift surveys like 2dF, WiggleZ, BOSS, VVDS, or eBOSS use spectroscopy to obtain very accurate redshifts which allow to map the 3D distribution of galaxies and recover the large scale sctructure of the universe. This enables an accurate reconstruction of the acoustic peak from Baryon Acoustic Oscillations and allows to use information from redshift space distortions produced by galaxies peculiar velocities. However such surveys are shallow or small and suffer from inhomogeneous sampling of the survey area. On the other hand, imaging surveys like SDSS, CFHT, KiDS or DES collect light using a limited amount of filter passbands, measuring a large amount of objects o a greater depth and high completeness. Instead of obtaining the full high resolution spectra of an object one obtains photometry in different wavelength ranges, which translates into lower precision redshift estimation. High quality images allow to measure shapes of millions of objects, which enables the use of weak gravitational lensing as another powerful probe for cosmology. Galaxy shapes are distorted when their light path travels near massive objects and regions in the universe, which allows to map and trace the amount of total matter as a function of line of sight position.

An special type of imaging surveys are those with multiband photometry like ALHAMBRA, COSMOS or PAUS. These include a combination of typical broad bands with several narrower filters. Among them, Physics of the Accelerating Universe survey (PAUS) is a unique example since it has a set of 40 narrow band filters (12.5nm FWHM) which is mapping four contiguous areas of 20 $\deg^2$, each with great photometric redshift precision. PAUS represents a bridge between small and deep pencil beam spectroscopic surveys and shallower wider ones, providing an unprecedented amount of subpercent precision photometric redshifts with imaging techniques that deliver greater depth and area.

1.1.1 Accelerating universe

Einstein's General Relativity field equations read[1] (see for example Weinberg 1972; Tanabashi et al. 2018)

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G_N T_{\mu\nu} + \Lambda g_{\mu\nu}, \tag{1.1}$$

where the left hand side terms describe the geometry of space and time, including the metric $g_{\mu\nu}$ and the Ricci tensor and scalar $R_{\mu\nu}$ and R. The right hand side includes the matter-energy terms, like the stress-energy tensor $T_{\mu\nu}$,

[1] Here and unless otherwise specified we assume natural units, which are defined as $c = \hbar = k_B = 1$ (speed of light, Planck constant and Boltzmann constant).

and the cosmological constant Λ term which is interpreted an effective energy term. The assumption that the universe is homogeneous and isotropic, known as the cosmological principle, allows to find a simple solution to the fields equation (Eq. 1.1) using the Friedmann - Lemaître - Robertson - Walker (FLRW) metric, described by the line element

$$ds^2 = g_{\mu\nu}dx^\mu dx^\nu = dt^2 - a^2(t)\left[d\chi^2 + S_k^2(\chi)d^2\Omega\right], \tag{1.2}$$

which describes an expanding (or contracting) universe with scale factor $a(t)$ which is a function of time, where dt is the time coordinate, $d\chi$ the radial comoving coordinate and $d\Omega$ the solid angle spatial coordinates. The metric also depends on the geometry of the space, which only affects the angles, where the function $S_k^2(\chi) = (\sin\chi, \chi, \sinh\chi)$ for a universe which is (closed, flat, open). The scale factor $a(t)$ relates the physical distance dr with a comoving distance that does not depend on time $d\chi$ as $dr(t) = a(t)d\chi$. Since $a(t)$ does not have units we set its current value to 1, $a_0 = a(t_0) = 1$. Assuming a stress-energy tensor for a perfect fluid, which is a common assumption for describing the universe components,

$$T_{\mu\nu} = -p\ g_{\mu\nu} + (p + \rho)u_\mu u_\nu, \tag{1.3}$$

where $g_{\mu\nu}$ is the spacetime metric from Eq. 1.2, $u_\mu = (1, 0, 0, 0)$, p is the isotropic pressure and ρ the energy density. Using the perfect fluid with FLRW metric and Einstein equations leads to the Friedmann equations, which describe the evolution of the scale factor $a(t)$,

$$H^2(t) \equiv \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G\rho(a)}{3} - \frac{k}{a^2}, \tag{1.4}$$

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p), \tag{1.5}$$

where k is the curvature of the universe, currently measured to be very close to zero. The parameter $H(t)$ is the Hubble parameter and measures the expansion rate of the universe. A third useful equation can be derived from Eqs. 1.4 and 1.5,

$$\dot{\rho} + 3H(\rho + p) = 0, \tag{1.6}$$

which we can use to find the density evolution of each component through his pressure. Matter can be approximated by a pressureless gas $p_m = 0 \rightarrow \rho_m \propto a^{-3}$, radiation by $p_r = \rho_r/3 \rightarrow \rho_r \propto a^{-4}$ the cosmological constant by $\dot{\rho}_\Lambda = 0 \rightarrow p_\Lambda = -\rho_\Lambda$ while a general dark energy fluid with pressure parameter

w has $p_{DE} = w\rho_{DE} \rightarrow \rho_{DE} \propto a^{-3(1+w)}$. Therefore,

$$\rho(a) = \rho_r a^{-4} + \rho_m a^{-3} + \rho_k a^{-2} + \rho_\Lambda \tag{1.7}$$

where here ρ_x refers to the current energy value (since $a_0 = 1$). For a flat FLRW universe (which is consistent with observations) we define the critical density as the present density, $\rho_{\text{crit}} = \frac{3H_0^2}{8\pi G}$. It is common to redefine the current densities relative to the critical density, $\Omega_x = \frac{\rho_x}{\rho_{\text{crit}}}$, which leads to

$$H^2(a) = H_0^2 \left[\Omega_r a^{-4} + \Omega_m a^{-3} + \Omega_k a^{-2} + \Omega_\Lambda\right] = H_0^2 E^2(a) \tag{1.8}$$

These energy values Ω_x represent the content of the universe nowadays and determine the evolution of the expansion rate. The Hubble rate is usually expressed as a function of a reduced hubble constant h, $H_0 = 100\ h\ \frac{\text{km}}{\text{Mpc·s}}$.

1.1.2 REDSHIFT

Redshift is a direct observable property of a photon defined as the wavelength difference between its emission and observation

$$z \equiv \frac{\lambda_{\text{obs}} - \lambda_{\text{em}}}{\lambda_{\text{em}}}. \tag{1.9}$$

There are three main causes of redshift in cosmology, the Doppler effect, gravitational redshift and the cosmological redshift. When objects move apart with a relative velocity v, they create a redshift in the observed photons

$$1 + z = \sqrt{\frac{1 + v/c}{1 - v/c}} \tag{1.10}$$

which for nonrelativistic velocities $v \ll c$ becomes $z \approx v/c$. The second one is a general relativistic effect produced by strong gravitational fields. The third one is the cosmological redshift, produced by the expansion of the universe. Since the comoving wavelength of a traveling light ray must remain constant with the expansion, the physical wavelength between time t, $a(t)$ and now t_0, $a_0 = 1$ changes by the ratio of the scale factors

$$1 + z = \frac{1}{a(t)} \tag{1.11}$$

1.1.3 Distances

A light ray traveling between two points in space satisfies $ds^2 = 0 \rightarrow -c\ dt = a\ d\chi$. Integrating we find the comoving distance between those points

$$\chi(z) = \int_{t_2}^{t_1} \frac{c\ dt}{a(t)} = \int_{a_2}^{a_1} \frac{c\ da}{H\ a^2} = \frac{c}{H_0} \int_0^z \frac{dz}{E(z)} \tag{1.12}$$

A very useful distance is the luminosity distance D_L which is defined as the ratio between the luminosity L at emission and the observed apparent flux F_0,

$$D_L \equiv \sqrt{\frac{L}{4\pi F_0}} = \frac{\chi}{a} = \chi(1+z). \tag{1.13}$$

which relates to the comoving distance. Note how for low redshift, $D_L \approx \frac{zc}{H_0}$, which is the lineal Hubble law.

1.1.4 Cosmic microwave background

One of the fundamental observations that support the standard cosmological model was the discovery of a quasi isotropic and homogeneous radiation, Penzias & Wilson (1965). Since the discovery of the expansion of the universe from Hubble in 1929, if one projected backwards in time arrived to the conclusion that the universe must have been hot, dense and small in the past. In fact, since radiation temperature is inversely proportional to its wavelength, which gets redshifted by the expansion, the evolution of radiation temperature goes as

$$T = \frac{T_0}{a} \tag{1.14}$$

where T_0 is the current temperature. Note how in the past the scale factor is smaller, and the temperature rises. In a hot enough universe, photons, protons and electrons would have enough energy to be in thermal equilibrium and be constantly interacting with each other. With the expansion and cooling of the universe, photons stop having enough energy to ionize hidrogen atoms, which start to form, an epoch known as recombination. The universe became transparent and photons could travel free, generating a very isotropic and homogeneous radiation in all directions that has been traveling (and cooling down) since then. It is known as the cosmic microwave radiation (CMB) due to the redshift of the original wavelength, which now peaks in the wavelength range of radio. The CMB has the nearest black body spectrum measured in nature, since it was generated from a plasma in thermal equilibrium.

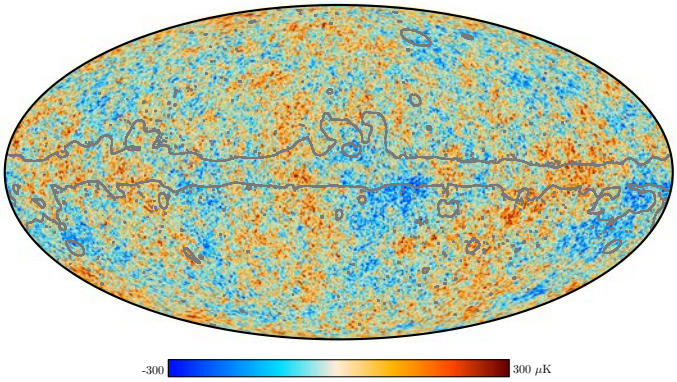

Figure 1.1: Planck 2018 map of the temperature anisotropies of the CMB, extracted using the SMICA method. The gray outline shows the extent of the confidence mask.

The COBE satellite, in 1992, measured the black body spectrum of the CMB and provided definitive evidence for a hot Big Bang universe. Later, two more satellites, WMAP (2003) and Planck (2013), have significantly improved the measurement of the CMB spectrum. Now we know the CMB originated at $z \sim 1100$ and has a temperature of $T = 2.72548 \pm 0.00057$K (Fixsen 2009). While the spectrum is significantly isotropic (once removing a dipole generated by a Doppler shift from our movement in and with the Milky Way), there are small anisotropies at the level of 10^{-5}. Fig. 1.1 shows the anisotropy temperature map measured by Planck (Planck Collaboration et al. 2018). These anisotropies are the result of perturbations from the density in the early universe, which have been projected into the surface of last scattering (or CMB). They contain significant information about the conditions in the early uinverse, and provide the best constraints of the ΛCDM model. Fig. 1.2 shows the angular power spectra of the CMB measured by Planck (2018) and a best fit model to temperature using the ΛCDM model, showing extraordinary agreement.

1.2 Photometry

Photometry is a technique that measures the flux of an object that goes through a passband or filter that blocks light except for a wavelength region. We define the flux of a galaxy measure through a filter with response R in the AB system (Oke & Gunn 1983) as

$$f = \frac{\int d\nu \frac{f_\nu(\nu)}{\nu} R(\nu)}{\int d\nu \frac{g_\nu(\nu)}{\nu} R(\nu)} = \frac{\int d\lambda \lambda f_\lambda(\lambda) R(\lambda)}{\int d\lambda \lambda g_\lambda(\lambda) R(\lambda)} \tag{1.15}$$

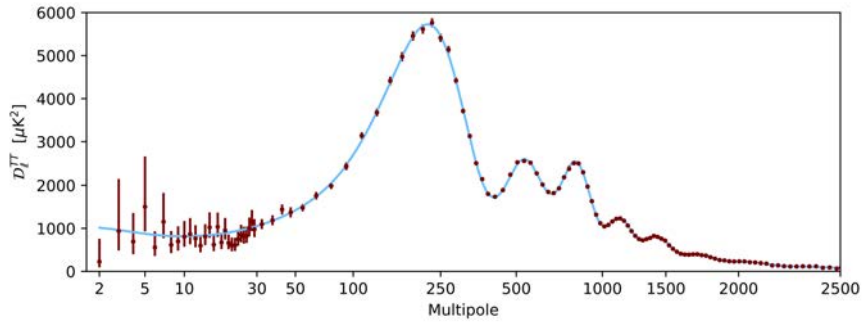

Figure 1.2: The Planck 2018 angular temperature power spectra of the CMB (TT). The blue line is a best-fit model to temperature data using ΛCDM model.

where $f_\nu(\nu)$ is the energy density per unit area, unit time and unit frequency of the object, and $g_\nu(\nu) = 3631\text{Jy}$ is a reference hypothetical source with constant frequency emission that defines the AB system zero point. The response is assumed to be in photon transmission units, which is why the extra ν term is needed to change the units form energy to photon counts. Since sources have very variable amount of energy density, astronomers have usually used a logarithmic scale to label galaxies, calling them magnitudes. The apparent magnitude m of an object is defined as

$$m = -2.5 \log_{10} \left(\frac{\int d\nu \frac{f_\nu(\nu)}{\nu} R(\nu)}{\int d\nu \frac{g_\nu(\nu)}{\nu} R(\nu)} \right) \tag{1.16}$$

The apparent magnitude m is defined as the observed absolute magnitude an object at redshift z would have if it were at a distance of $10pc$ (parsecs),

$$m(z) = M + 5 \log_{10} \left(\frac{D_L(z)}{10pc} \right) \tag{1.17}$$

The difference $m - M$ is the distance modulus,

$$DM = 5 \log_{10} \left(\frac{D_L(z)}{10pc} \right). \tag{1.18}$$

In general, for a filter R and Q the relation includes the K-correction,

$$m_R = M_Q + DM + K_{QR} \tag{1.19}$$

To identify if a galaxy is bluer or redder we define colors as magnitude differences.

1.3 Bayesian Statistics

Bayesian statistics is an interpretation of the concept of probability as the state of knowledge over some logical statement, rather than the frequency of a phenomenon (see Berger (1985); Jaynes & Bretthorst (2003); Ghahramani (2012); Knuth et al. (2014)). We call the current state of knowledge our *prior* information, which can arise from previous experiments, scientific hypothesis or personal belief. Bayesian inference can be used to update the current knowledge in the light of new evidence by writing its conditional implication.

There are two main rules in Bayesian probability, the sum and the product rules. Let x and y be random events. The sum rule states that if the joint probability of x and y is known, then the marginal probability of either of them can be obtained by summing over all values of the other,

$$P(x) = \sum_y P(x, y). \tag{1.20}$$

The product rule states that the joint probability of x and y can be decomposed into the marginal probability of x times the conditional probability of y given x, and viceversa,

$$\begin{aligned} P(x, y) &= P(y|x)P(x) \\ P(x, y) &= P(x|y)P(y) \end{aligned} \tag{1.21}$$

Equating both expressions and rearranging one finds Bayes theorem,

$$P(x|y) = \frac{P(y|x)P(x)}{P(y)} \tag{1.22}$$

We define models as an approximate representation of reality, and a plausible explanation of our data. Let the parameters of the model M be θ (usually a vector). We assume our model can forecast data, $P(D|\theta, M)$, which we call the *likelihood* of the parameters θ. We can obtain the posterior on the parameters given some new data became available from the likelihood using Bayes theorem,

$$P(\theta|D, M) = \frac{P(D|\theta, M)P(\theta|M)}{P(D|M)}, \tag{1.23}$$

where $P(\theta|M)$ is the prior probability of the parameters of the model. It is important to note that the model is not well defined if the prior is not specified. The term dividing is a normalization called the *evidence*, the probabilty that

our data might have come from a given model. Using the sum and product rule,

$$P(D|M) = \int P(D, \theta|M) d\theta = \int P(D|\theta, M) P(\theta|M) d\theta, \tag{1.24}$$

The evidence is usually ignored in parameter estimation, since it is a normalization. However, for model comparison or model selection it becomes very relevant, since the posterior probability of a model given the data is

$$P(M|D) = \frac{P(D|M)P(M)}{P(D)} \tag{1.25}$$

Let M_1 and M_2 belong to the group H of hypothesis or models to be considered. The most likely model between the two will be the one with larger posterior probability, therefore we look at the ratio of the posteriors

$$\frac{P(M_1|D, H)}{P(M_2|D, H)} = \frac{P(D|M_1, H)P(M_1|H)}{P(D|M_2, H)P(M_2|H)}, \tag{1.26}$$

where $P(M|H)$ represents the a priori probability of each model. When these are equal then the posterior ratio equals the ratio of the evidences, also called Bayes factor or odds ratio

$$\text{Bayes factor} = \frac{P(D|M_1, H)}{P(D|M_2, H)}. \tag{1.27}$$

An important property of the Bayesian evidence, Eq. 1.24, is that it carries a preference for simpler models, which is known as the Bayesian Occam's Razor. Fig. 1.3 illustrates this effect, inspired by a figure from Ghahramani (2012). A more complex model, with more parameters, has the ability to forecast more datasets like D, but since the evidence $P(D|M)$ is a probability and must sum to unity, then a more complex model spreads its probability over more datasets. In non Bayesian statistics a more complex model could overfit the data and deceive us into thinking it is the most likely, but in Bayesian statistics there is no fitting and no optimization.

1.4 Physics of the Accelerating Universe Survey (PAUS)

Physics of the Accelerating Universe Survey (PAUS) is an ongoing imaging survey using a unique instrument, PAUCam (Padilla et al. 2019), which

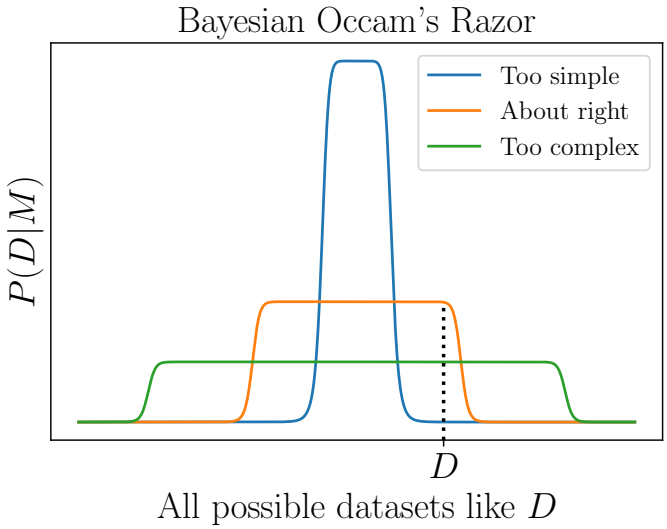

Figure 1.3: Illustration of Bayesian Occam's Razor, inspired by a figure from Ghahramani (2012). The blue line shows that a too simple model can only explain a small subset of possible datasets and will not properly explain D. On the contrary, the green line shows that a too complex model will spread its probability over many more datasets than needed, since probabilities must sum to one, reducing its evidence for a particular dataset D. Both would get rejected against the orange model, which can explain the data without more sophistication. This normalization effect is known as the Bayesian Occam's Razor.

carries a set of 40 narrow band (NB) filters (12.5nm FWHM) that span the wavelength range from 450nm to 850nm (Fig. 1.4) along with the usual ugrizY broad band filters. Such an exquisite wavelength sampling results in subpercent photometric redshift precision down to magnitude $i_{\rm AB} = 22.5$, and as good as $\sigma_z \sim 0.0035(1+z)$ for half of such galaxy sample after selecting with a quality parameter (Martí et al. (2014); Eriksen et al. (2019)).

Wide field galaxy surveys are typically divided into two categories: spectroscopic surveys and imaging surveys. The former obtains a high resolution spectra of the object within some wavelength coverage, which is used to identify sharp features like emission and absorption lines to nail the redshift of the object with very high precision. However, these are expensive to obtain: they require knowing the position of the object beforehand and a large exposure time, which makes it observationally inefficient to observe deep objects over a large area. Such surveys also suffer from incompleteness both because not all objects in the field are targeted and also since a fraction of the measured spectra fail to provide an accurate redshift, for example when only one line is observed or when there is a line confusion.

In contrast, imaging surveys are able to obtain measurements of every object at the same time from a set of bandpass filtered images, which allows to cover large areas faster and to a greater depth. This happens at the expense of getting flux measurements with very poor spectral resolution since the width of typical broad band filters is larger than 100nm, which makes the redshift determination much less precise and easily biased. In this context, PAUS fills a gap between large and wide broad band imaging surveys such as KiDS, DES, Euclid or LSST with limited line of sight resolution, and pencil beam surveys with spectroscopic data but with a very small area covered. Fig. 1.7 shows a comparison between PAUS and different spectroscopic surveys. PAUS will be a few magnitudes deeper (over 10 times denser) than current completed large area flux limited surveys (such as GAMA) and 10 times larger than current completed deep flux limited surveys (such as VVDS/VIPERS/DEEP2). Fig. 1.5 illustrates the effect typical broad band photo-z have on the line of sight resolution compared to a PAUS-like redshift resolution using narrow bands.

This unique combination of area, depth and redshift resolution enables new analysis. Studies on the role of environment in structure formation are limited by poor redshift precision in broad band surveys and by tiny area coverage or very low density in spectroscopic surveys, while PAUS allows to sample with high density several galaxy populations over a much larger area and with sub-percent photo-z accuracy, and in a luminosity-redshift regime unexplored by shallower completed surveys like GAMA, closing up the gap with future surveys like Euclid. It is ideal to study the range of intermediate scales (below

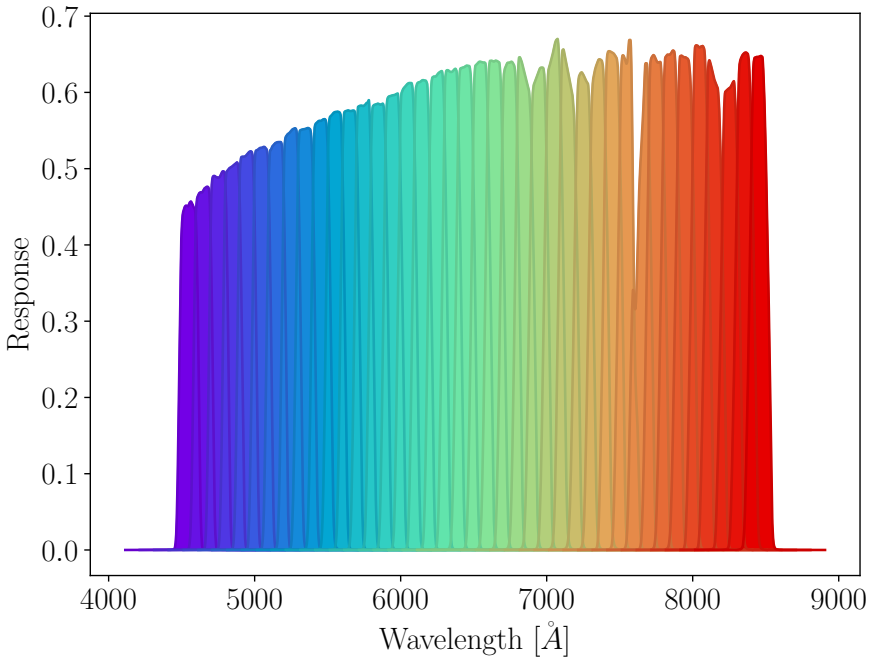

Figure 1.4: PAUS instrument PAUCam carries a set of 40 narrow band filters of 12.5nm FWHM that span the wavelength range of 450nm to 850nm. This figure shows the final filter transmission of each of the narrow band filters taking into account the atmospheric transmission, the quantum efficiency of the CCDs and the telescope throughput.

$10-20$ Mpc), the weakly non-linear regime, where the statistical signal-to-noise ratio of cosmological surveys is largest, a range usually avoided in cosmological analysis as the modeling of nonlinear bias or the impact of baryonic and environment effects in structure formation are not well understood.

PAUS observation fields overlap with deep, high quality shape measurements from broad band surveys as CFHT and KiDS, which will provide substantial constraints on galaxy-shape alignment models, for samples of crucial interest for future surveys like Euclid and LSST. Intrinsic alignments show the interaction between galaxy evolution and the large scale structure of the Universe, but will also become an important astrophysical systematic in cosmology analysis.

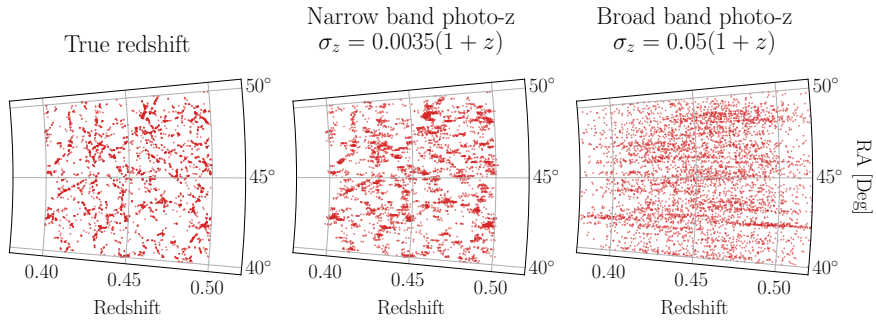

Figure 1.5: Effect of the photo-z in the positions of simulated galaxies from the public MICE2 simulations.

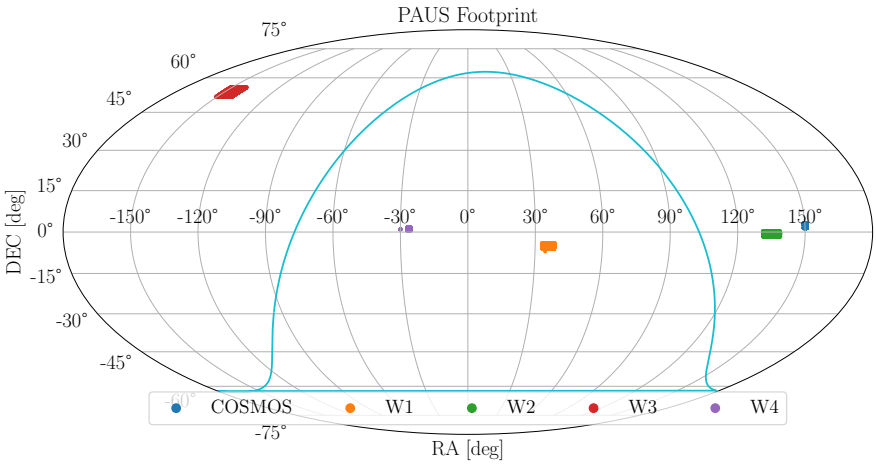

Figure 1.6: PAU Survey fields of observation: COSMOS and CFHTLS W1, W2, W3 and W4. The cyan line represents the Galactic plane.

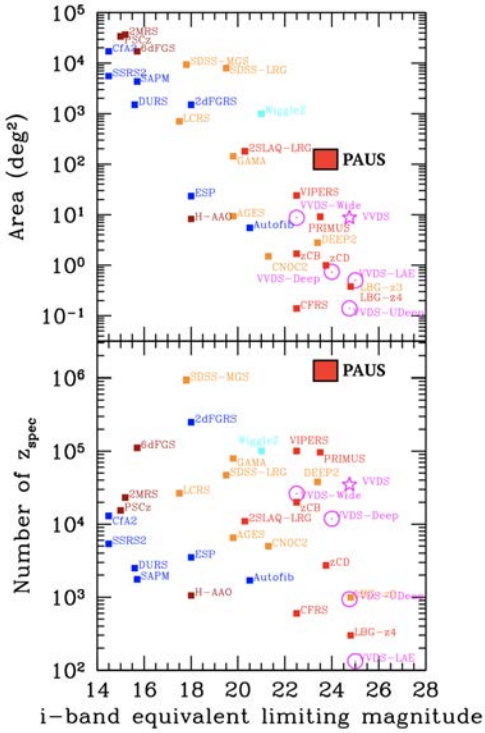

Figure 1.7: Adapted from Le Fèvre et al. (2013). Comparison of the covered area (top) and the number of measured spectroscopic redshifts (bottom) versus depth expressed as the equivalent i-band limiting magnitude, between the PAUS and different spectroscopic redshift surveys. The PAUS numbers are projected quantities expected by the end of the survey.

Part II

Sample variance cancellation forecast with accurate redshifts

In this part we use the Fisher matrix formalism to study the expansion and growth history of the Universe using galaxy clustering with 2D angular cross-correlation tomography in spectroscopic or high resolution photometric redshift surveys. We show how multiple tracers with redshift space distortions cancel sample variance using this formalism where the radial information is contained in the cross-correlations between narrow redshift bins.

Chapter 2

Cosmological constraints from multiple tracers in spectroscopic surveys

2.1 Introduction

One of the most exciting and enigmatic discoveries in the recent years is the late time accelerated expansion of the Universe, confirmed in late 1990s from Type Ia supernovae (Riess et al. 1998, Perlmutter et al. 1999). During the last decade a wide range of observations (see Weinberg et al. 2013) has provided robust evidence for cosmic acceleration, consistent with a ΛCDM model dominated by a new component called *dark energy*, which properties and origin remain unknown.

Cosmic expansion is parametrized by $\Omega(a)$ and the DE equation of state $\omega(a) = \omega_0 + \omega_a(1-a)$ (Chevallier & Polarski (2001), Linder (2003)) while cosmic growth is parametrized by γ, which gives the growth rate as $f(z) = \Omega(z)^\gamma$. For General Relativity (GR) $\gamma \sim 0.55$, while Modified Gravity models can give different values of γ for the same expansion history (e.g. Gaztañaga & Lobo 2001, Lue, Scoccimarro & Starkman 2004, Huterer et al. 2015). Here we study the dark energy equation of state $\omega(z)$ and growth rate γ constraints using galaxy clustering in spectroscopic surveys. Galaxy clustering is able to probe the expansion and growth history almost independently, unlike weak lensing surveys alone, which are limited to projected, 2D information (see Gaztañaga et al. 2012, Weinberg et al. 2013). Galaxies are easy to observe and by accurately measuring their redshift one can reconstruct the 3D clustering information.

Unfortunately, the relation between galaxy and dark matter is not straightforward, and in the linear regime, for large scales, it can be modeled by a factor called *linear bias* $b(k, z)$, such that $\delta_g(k, z) = b(k, z)\delta_m(k, z)$, where δ_g and δ_m are galaxy and dark matter fluctuations. An independent measurement

is needed to break the degeneracy between bias and γ, as galaxy clustering alone cannot (e.g. see Eq.2.2 below). One can break this degeneracy using cross-correlation with lensing surveys (e.g. Gaztañaga et al. 2012, Weinberg et al. 2013), but in this chapter we will focus on spectroscopic surveys or high resolution photometric surveys (Martí et al. 2014). In this case, to determine bias one can measure the redshift space distortion parameter $\beta \equiv f(z)/b(z)$. Redshift space distortions (RSD) in the linear regime (Kaiser 1987) enhance clustering in the line of sight by a factor $(1 + f)$ due to local infall of bodies as a result of gravity. Measuring with different angles relative to the line of sight one can determine $f(z)$. However, the random nature of fluctuations (sampling variance) limits the accuracy with which one can determine $P(k)$, and with only one tracer that propagates to β and cosmological parameters. McDonald & Seljak (2009) proposed to use multiple tracers of the same underlying distribution to beat this limit measuring along many directions and improve the constraints canceling sampling variance with RSD. Sampling variance cancellation can also be achieved with other observables (e.g. Pen 2004, Seljak 2009). This technique has been explored in recent literature (e.g. White, Song & Percival 2009, Gil-Marín et al. 2010, Bernstein & Cai 2011, Abramo 2012, Abramo, Secco & Loureiro 2016), also for photometric surveys (Asorey, Crocce & Gaztañaga 2014) and combining lensing and spectroscopic surveys (Cai & Bernstein 2012, Eriksen & Gaztañaga 2015a).

We use 2D angular correlations C_ℓ (see §2.2.2) to avoid assuming a cosmology and avoid overcounting overlapping modes without including the full covariance between them (Eriksen & Gaztañaga 2015b). We forecast spectroscopic surveys with narrow redshift bins ($\Delta z = 0.01(1+z)$) such that the radial linear modes will be in the cross correlations between redshift bins. In the fiducial forecast we will compute the correlations using redshift space distortions (RSD) and we include baryon acoustic oscillation measurements (BAO). In this chapter we will study the constraints from single spectroscopic tracers as compared to splitting one population into two tracers. The single tracers are denoted as B1 and B2 and the multiple tracer survey as B1xB2. Note this differs from our series of previous studies (Eriksen & Gaztañaga 2015b and others) where we included one Bright and one Faint population as opposed to two Bright populations. The cosmological parameter error estimation is done using the Fisher matrix formalism described in §2.2.3, and we quantify the relative strength of the surveys through the Figures of Merit (FoMs) defined in §2.2.4, which focus on measuring the expansion and growth history simultaneously. In subsection 2.2.5 we present our fiducial forecast assumptions.

This chapter is organized as follows. In section 2 we present our modeling and fiducial forecast assumptions. Section 3 discusses sample variance

cancellation in surveys with multiple tracers and explores the effect of the relative bias amplitude between two tracers and the dependence on galaxy density. In section 4 we model galaxy bias using a conditional luminosity function (CLF) and halo model to build an apparent limited survey to study the tradeoff between galaxy bias and galaxy density when we split a survey into two subsamples. Section 5 investigates the impact of having partly overlapping redshift bins between two tracers in a multi tracer survey and how this affects the constraints. Moreover, it studies radial resolution by increasing the number of redshift bins. In section 6 we present our conclusions. Appendix A studies the importance of RSD and BAO in the constraints and the degeneracy with cosmological parameters. Appendix 2.B shows the dependence that the constraints have on the bias evolution in redshift.

In this chapter we have produced the results with the forecast framework developed for Gaztañaga et al. (2012), Eriksen & Gaztañaga (2015b), Eriksen & Gaztañaga (2015a), Eriksen & Gaztañaga (2015c) and Eriksen & Gaztañaga (2015d).

2.2 Modeling and forecast assumptions

2.2.1 Galaxy bias

In the local bias model (Fry & Gaztanaga 1993), where fluctuations are small, one can approximate the relation between galaxy overdensities δ_g to matter overdensities δ_m through

$$\delta_g(k,z) = b(k,z)\delta_m(k,z) \tag{2.1}$$

where $b(z,k)$ is the galaxy bias, which can in general depend on the scale and redshift. It also varies between different galaxy populations (galaxies hosted by more massive haloes tend to be more biased, eg. Scoccimarro et al. 2001). Then, for scale independent bias $b(z) = b(k,z)$ the angular correlations $\xi_{gg} \equiv \langle \delta_g \delta_g \rangle$ we have that

$$\xi_{gg}(\theta,z) = b^2(z)\,\xi_{mm}(\theta,z) \propto b^2(z)\; D^2(z) \tag{2.2}$$

Galaxy bias can also include an stochastic component r, see also Eq. 2.27, which is also a common measure of non-linearity

$$r \equiv \frac{\xi_{gm}}{\sqrt{\xi_{gg}\xi_{mm}}}. \tag{2.3}$$

In Gaztañaga et al. (2012) it was shown that it can be treated as a renormalisation of bias in large scales and here it is fixed to $r = 1$. In addition, non local bias can also modify the galaxy correlation function, but this is a smaller effect (Chan, Scoccimarro & Sheth 2012).

We include *redshift space distortions* (RSD, redshift displacement of galaxies due to their peculiar velocities with respect to the comoving expansion) using linear theory, Kaiser (1987), assuming no velocity bias

$$\delta_s(k,\mu) = (b + f\mu^2)\,\delta(k). \tag{2.4}$$

where $\mu \equiv (\hat{z}\cdot\boldsymbol{k})/k = k_\parallel/k$. We define $\beta \equiv f/b$ as the term with specific angular dependence μ in redshift space.

2.2.2 Angular correlation function and Power spectrum

Consider the projection of spatial galaxy fluctuations $\delta^i_g(x,z)$ along a given direction in the sky $\hat{\boldsymbol{r}}$

$$\delta^i_g(\hat{\boldsymbol{r}}) = \int dz\,\phi^i(z)\,\delta^i_g(\hat{\boldsymbol{r}},r,z), \tag{2.5}$$

where $\phi^i(z)$ is the radial selection function for the i-th redshift bin of a given tracer. We define the angular correlation between galaxy density fluctuations as

$$\omega_{ij}(\theta) \equiv \langle \delta^i_g(\boldsymbol{r})\,\delta^j_g(\boldsymbol{r}+\hat{\theta})\rangle. \tag{2.6}$$

Expanding the projected density in terms of spherical harmonics we have

$$\delta^i(\hat{\boldsymbol{r}}) = \sum_{\ell\geq 0}\sum_{m=-\ell}^{\ell} a^i_{\ell m} Y_{\ell m}(\hat{\boldsymbol{r}}) \tag{2.7}$$

where $Y_{\ell m}$ are the spherical harmonics. The coefficients $a^i_{\ell m}$ have zero mean $\langle a^i_{\ell m}\rangle = 0$, as $\langle\delta^i\rangle = 0$ by construction, and their variance form the angular power spectrum

$$\langle a^i_{\ell m} a^j_{\ell' m'}\rangle \equiv \delta_{\ell\ell'}\delta_{mm'}\,C^{ij}_\ell \tag{2.8}$$

which can be related to the angular correlations with

$$\omega_{ij}(\theta) = \sum_{\ell\geq 0}\frac{2\ell+1}{4\pi}\,L_\ell(\cos\theta)\,C^{ij}_\ell \tag{2.9}$$

where $L_\ell(\cos\theta)$ are the Legendre polynomials of order ℓ. The C_ℓ^{ij} can be expressed in Fourier space (Crocce, Cabré & Gaztañaga 2011) as

$$C_\ell^{ij} = \frac{1}{2\pi^2}\int 4\pi k^2 dk\, P(k)\, \psi_\ell^i(k)\psi_\ell^j(k) \tag{2.10}$$

where $P(k)$ is the matter power spectrum and $\psi_\ell^i(k)$ is the kernel for the i-th redshift bin of a given population. For the matter power spectrum $P(k)$ we use the linear power spectrum from Eisenstein & Hu 1998 for linear scales, which accounts for baryon acoustic oscillations (BAO). In real space (no redshift space distortions), taking into account only the intrinsic component of galaxy number counts, this kernel is (Eriksen & Gaztañaga 2015b)

$$\psi_\ell^i(k) = \int dz\, \phi^i(z)\, D(z)\, b(z,k)\, j_\ell(kr(z)) \tag{2.11}$$

where $b(z,k)$ is the galaxy bias, Eq. 2.1. When including RSD, one has to add an extra term that in linear theory is given by (Kaiser 1987, Fisher, Scharf & Lahav 1994, Fisher et al. 1995, Taylor & Heavens 1995)

$$\begin{aligned}\psi_\ell(k) =& \psi_\ell^{Real} + \psi_\ell^{RSD}\\ \psi_\ell^{RSD} =& \int dz\, f(z)\, \phi(z)\, D(z)\, [L_0(\ell)\, j_\ell(kr)\\ &+ L_1(\ell)\, j_{\ell-2}(kr) + L_2(\ell)\, j_{\ell+2}(kr)]\end{aligned} \tag{2.12}$$

where $f(z)$ is the growth rate and

$$\begin{aligned}L_0(\ell) &\equiv \frac{(2\ell^2+2\ell-1)}{(2\ell+3)(2\ell-1)}\\ L_1(\ell) &\equiv -\frac{\ell(\ell-1)}{(2\ell-1)(2\ell+1)}\\ L_2(\ell) &\equiv -\frac{(\ell+1)(\ell+2)}{(2\ell+1)(2\ell+3)}\end{aligned} \tag{2.13}$$

The fiducial modeling includes RSD in the kernel and BAO in the power spectrum, but we will also forecast removing one or both of these effects.

Covariance

Angular cross correlations between a redshift bin i and redshift bin j correspond to the variance of spherical harmonic coefficients $a_{\ell m}$ (Eq. 2.8). Assum-

ing that $a^i_{\ell m}$ are Gaussianly distributed and in a full sky situation, one can then estimate each ℓ angular power spectrum using the $2\ell+1$ available modes,

$$\tilde{C}^{ij}_\ell = \frac{1}{2\ell+1}\sum_{m=-\ell}^{\ell} a^i_{\ell m} a^j_{\ell m}. \tag{2.14}$$

which yields Eq. 2.10. However, in a more realistic situation, we only have partial coverage of the sky so that the different modes ℓ become correlated. Following the approach of Cabré et al. (2007), we bin the $\ell \times \ell$ covariance with a sufficiently large band width $\Delta\ell$ such that it becomes band diagonal, and scale it with $1/f_{Sky}$ (where f_{Sky} is the survey fractional sky). Then, the covariance becomes

$$\mathrm{Cov}\left[\hat{C}^{ij}_\ell, \hat{C}^{kl}_\ell\right] = N^{-1}(\ell)(\hat{C}^{ik}_\ell \hat{C}^{jl}_\ell + \hat{C}^{il}_\ell \hat{C}^{jk}_\ell). \tag{2.15}$$

where $N(\ell) = f_{Sky}(2\ell+1)\Delta\ell$, and the correlation $\hat{C}$ includes observational noise

$$\hat{C}^{ij}_\ell = C^{ij}_\ell + \delta_{ij}\frac{1}{\bar{n}_g} \tag{2.16}$$

where $\bar{n}_g = \frac{N_g}{\Delta\Omega}$ is the galaxy density per solid angle. The first term in Eq. 2.16 is signal and contains sample variance information, while the second is shot-noise. Then, we can define the χ^2 as

$$\begin{aligned}\chi^2 = \sum_{\ell,\ell',i,j} \left(C^{ij}_\ell(\{\lambda_k\}) - \hat{C}^{ij}_\ell\right) \left(\mathrm{Cov}^{-1}\right)_{\ell,\ell',i,j} \\ \times \left(C^{ij}_{\ell'}(\{\lambda_k\}) - \hat{C}^{ij}_{\ell'}\right)\end{aligned} \tag{2.17}$$

where $C^{ij}_\ell(\{\lambda_k\})$ depend on the parameters that we are looking for, $\hat{C}^{ij}_\ell$ are the observed C^{ij}_ℓ's and Cov the covariance matrix between the $\hat{C}^{ij}_\ell$'s. C^{ij}_ℓ include both auto and cross correlations between different redshift bins and also possibly between different tracers if there was more than one.

Nonlinearities

As we are working in the linear regime we have to limit the scales that we include in the forecast. We restrict the forecast to scales between $10 \leq \ell \leq 300$. In addition we apply a further cut in l_{max}

$$\ell_{max} = k_{max}\, r(z_i) - 0.5, \tag{2.18}$$

for which correlations to include, as these are the scales contributing to $C(\ell)$ for a given narrow redshift bin z_i (Eriksen & Gaztañaga (2015b)). In the forecast we use the Eisenstein-Hu power spectrum and the MICE cosmology with a maximum scale k_{max} of (see Eriksen & Gaztañaga 2015c)

$$k_{max}(z) = \exp\left(-2.29 + 0.88z\right). \tag{2.19}$$

2.2.3 Fisher matrix formalism

Even if we don't have any data, we can tell how $\chi^2(\{\lambda_\mu\})$ will vary in the parameters space defined by $\{\lambda_\mu\}$. Expanding χ^2 in the Gaussian approximation around its minimum the Fisher matrix is (Fisher 1935, Dodelson 2003)

$$F_{\mu\nu} = \sum_{l,l'} \sum_{ij,mn} \frac{\partial C_l^{ij}}{\partial \lambda_\mu} \left(\mathrm{Cov}^{-1}\right)_{l,l'} \frac{\partial C_{l'}^{mn}}{\partial \lambda_\nu}, \tag{2.20}$$

and it follows that

$$\mathrm{Cov}\left[\lambda_\mu, \lambda_\nu\right] = \left[F^{-1}\right]_{\mu\nu}. \tag{2.21}$$

The parameters included in the Fisher matrix forecast are (Eriksen & Gaztañaga 2015c)

$$\{\lambda_\mu\} = \omega_0, \omega_a, h, n_s, \Omega_m, \Omega_b, \Omega_{DE}, \sigma_8, \gamma, \text{Galaxy bias}. \tag{2.22}$$

The forecast use one galaxy bias parameter per redshift bin and population, with no scale dependence. Less bias parameters and other bias parameterization give similar results (see Fig. 2.6 or Eriksen & Gaztañaga 2015d). We include all cross-correlations between redshift bins and different populations. We use Planck priors for all parameters except for γ and galaxy bias.

2.2.4 Figure of Merit (FoM)

The Figure of Merit (FoM) for a certain parameter subspace S is defined as

$$\mathrm{FoM}_S \equiv \frac{1}{\sqrt{\det\left[F^{-1}\right]_S}}, \tag{2.23}$$

marginalizing over parameters not in S. This is a good estimator of the error for different dimensional subspaces S. For one parameter, then this is the inverse

error (Eq. 2.21) of the parameter. For two parameters it is proportional to the inverse area included within 1-sigma error ellipse. For three parameters it is the inverse volume within 1-sigma error ellipsoid, and so on. In this work we focus in the figures of merit defined in (Eriksen & Gaztañaga 2015c):

- $\mathrm{FoM}_{\mathrm{DETF}}$. $S = (\omega_0, \omega_a)$. Dark Energy Task Force (DETF) Figure of Merit (Albrecht et al. 2006). Inversely proportional to the error ellipse of (ω_0, ω_a). The growth factor γ is fixed.
- FoM_{ω} : Equivalent to $\mathrm{FoM}_{\mathrm{DETF}}$, but instead of $\gamma = 0.55$ from GR, γ is considered a free parameter and is marginalized over.
- FoM_{γ}. $S = (\gamma)$. Corresponds to the inverse error of the growth parameter γ. Therefore, $\mathrm{FoM}_{\gamma} = 10, 100$ corresponds to 10%, 1% expected error on γ. The dark energy equation of state parameters (ω_0, ω_a) are fixed.
- $\mathrm{FoM}_{\gamma\omega}$. $S = (\omega_0, \omega_a, \gamma)$. Combined figure of merit for ω_0, ω_a and γ.

It is important to note that, when not including priors, the different FoMs scale with area A in the following way

$$\begin{aligned} &\mathrm{FoM}_{\mathrm{DETF}} \propto A, \\ &\mathrm{FoM}_{\omega} \propto A, \\ &\mathrm{FoM}_{\gamma} \propto A^{1/2}, \\ &\mathrm{FoM}_{\gamma\omega} \propto A^{3/2}. \end{aligned} \tag{2.24}$$

Doubling the area would give a factor ~ 2.83 higher $\mathrm{FoM}_{\gamma\omega}$.

2.2.5 Fiducial galaxy sample

We define two galaxy populations based on the following fiducial spectroscopic (Bright[1], B) population. We define a magnitude limited survey, with $i_{AB} < 22.5$ as the fiducial flux limit in the i-band. The fiducial survey area is 14000 deg^2. The fiducial redshift range is $0.1 < z < 1.25$, and the number of redshift bins is 71, with a narrow bin width of $0.01(1 + z)$. Spectroscopic surveys usually have great redshift determination, so we define a Gaussian

[1]This population definition is in correspondence with previous work such as Gaztañaga et al. (2012) and Eriksen & Gaztañaga (2015c)

Area [deg^2]	14,000
Magnitude limit	$i_{AB} < 22.5$
Redshift range	$0.1 < z < 1.25$
Redshift uncertainty	$0.001(1+z)$
zBin width	$0.01(1+z)$
Number of zbins	71
Bias: $b(z)$	2+2(z-0.5)
Density [gal/arcmin^2]	0.4
$n(z)$ - z_0	0.702
$n(z)$ - α	1.083
$n(z)$ - β	2.628

Table 2.1: Parameters that describe our fiducial spectroscopic survey.

spectroscopic redshift uncertainty of $\sigma_{68} = 0.001(1+z)$, much lower than the bin width.

The fiducial bias is interpolated within 4 redshift pivot points, $z = 0.25, 0.43, 0.66, 1.0$, which scale with redshift in the following way,

$$b_B(z) = 2 + 2(z - 0.5). \tag{2.25}$$

Recall that there is one bias parameter per redshift bin and population. The fiducial redshift distribution of galaxies is characterized with the number density of objects per solid angle and redshift as

$$\frac{dN}{d\Omega dz} = N \left(\frac{z}{z_0}\right)^{\alpha} \exp\left(-\left(\frac{z}{z_0}\right)^{\beta}\right), \tag{2.26}$$

and is constructed by fitting a Smail type $n(z)$ (Efstathiou et al. 1991) to the public COSMOs photo-z sample (Ilbert et al. 2010). The values for α, β and z_0 in Eq. 2.26 correspond exactly to the values in Gaztañaga et al. 2012: $z_0 = 0.702$, $\alpha = 1.083$ and $\beta = 2.628$. The normalization N sets the density of galaxies per solid angle, being the fiducial density for this work $n_g = 0.4$ gal/arcmin^2. Table 2.1 summarizes the parameters that characterize our fiducial spectroscopic survey.

2.3 Sample variance cancellation

When two populations in a survey overlap in the same volume (B1xB2) one gets additional cross-correlations and covariance between them. If one is able to split one galaxy sample into two galaxy overdensities in the same area by some observable (i.e. luminosity, color), the resulting subsamples become correlated as they trace the same underlying dark matter fluctuations. As a result, using multiple tracers allow for sampling variance cancellation and can considerably improve the constraints. This multi-tracer technique was first introduced in McDonald & Seljak (2009).

Assume B1 and B2 are two galaxy populations, one with bias b and the other with bias αb. Their density perturbation equations in redshift space (Eq. 2.4) and in the linear regime are

$$\delta_{B1}(k) = (b + f\mu^2)\,\delta(k) + \epsilon_1, \tag{2.27}$$

and

$$\delta_{B2}(k) = (\alpha b + f\mu^2)\,\delta(k) + \epsilon_2, \tag{2.28}$$

where $\mu \equiv k_{\parallel}/k$ is defined to be the cosine of the angle between the line of sight and the wavevector $\hat{k}$, and ϵ_i are stochasticity parameters that can refer to a standard shot-noise or to other random component.

Even when having an infinite galaxy sample, there will be cosmic variance as each mode $\delta(k)$ is a random realization of a Gaussian field. However, if we have two tracers sampling the field we can average over many modes and cancel the sampling variance. To illustrate this we divide Eq.2.28 over Eq.2.27 (with no stochasticity) and obtain

$$\frac{\delta_{B2}}{\delta_{B1}} = \frac{\alpha\beta^{-1} + \mu^2}{\beta^{-1} + \mu^2}, \tag{2.29}$$

where $\beta \equiv f/b$, which has explicit angular dependence, but no dependence on the random field δ, which allows to extract α and β separately, and determine β exactly in the absence of shot-noise. In McDonald & Seljak (2009) the authors compute an analytical example considering a pair of transverse and radial modes ($\mu = 1$ and $\mu = 0$), and already found that can arbitrarily improve the determination of β with respect to the single galaxy in the limit of zero shot-noise.

Splitting one spectroscopic sample into two over the same area increases the number of observables available but also increases the number of nuisance parameters and adds shot-noise to the observables. The cosmological informa-

tion coming from two overlapping tracers is correlated as well as their nuisance parameters, which manages to reduce the error on cosmological parameters (Eriksen & Gaztañaga (2015a)). Decisively, splitting optimizes the constraints by canceling the random nature in the amplitude of the modes (see Eq. 2.29).

The way these mentioned effects propagate into the FoMs is the object of study in Section 3. In the following subsections we show the impact in our forecast of the relative bias amplitude (subsection 2.3.1) and the dependence on galaxy density (subsection 2.3.3) for the single and multi tracer surveys. In subsection 2.3.2 we show the FoMs for $\alpha = 0.5$, which is the fiducial relative bias amplitude value for subsection 2.3.3, section 2.5 and Appendix 2.A.

2.3.1 Relative bias amplitude (α)

In Fig. 2.1 we show $\mathrm{FoM}_{\gamma\omega}$ (§2.2.4) (for other FoM see Fig. 2.11) for the two single tracers (B1 and B2) defined in Eqs. 2.27 and 2.28, without stochasticity, as function of the relative bias amplitude α (Eq. 2.28). They both follow the fiducial configuration from Table 2.1 except for the α parameter. B2 is shown with the fiducial density and with four times less density. Furthermore, we show what happens if we merge both single tracers into one overlapping survey B1xB2, for the two density cases of B2. B1xB2 equals B1 + B2 constraints + extra correlations + extra covariance, and includes all cross redshift and cross population correlations. All lines are normalized to the B1 FoM, which does not depend on α.

In the example considering a pair of transverse and radial modes from McDonald & Seljak (2009), the authors find that the improvement measuring β is proportional to

$$\frac{\sigma_\beta^2(1 \text{ tracer})}{\sigma_\beta^2(2 \text{ tracers})} \propto \frac{(\alpha-1)^2}{\alpha^2}, \tag{2.30}$$

which is minimum at $\alpha = 1$. When doing the full analysis in Fig. 2.1 we take into account the whole range of μ, and our results for 2 tracers (B1xB2) also show a minimum when the bias amplitudes are equal ($\alpha = 1$). Note that we do not expect B1xB2 ratio to B1 to be 1 at $\alpha = 1$ as B1xB2 has twice its total density and extra correlations and covariance between tracers. To remind this and avoid confusion we write the total density in parenthesis next to the population in gal/arcmin^2: B1xB2(0.8), B1(0.4), B2(0.4).

When increasing the bias ratio, $\alpha \neq 1$, we cancel sample variance and we quickly improve our constraints up to a factor 4 from B1(0.4) to B1xB2(0.8). If we reduce four times the density of B2(0.4) the improvement between B1(0.1)

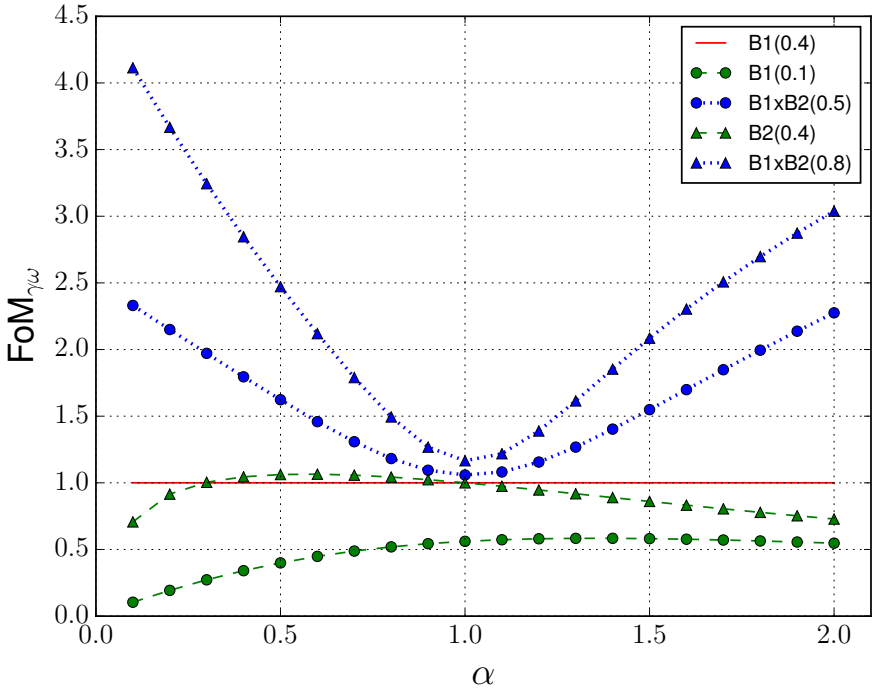

Figure 2.1: FoM$_{\gamma\omega}$ dependence on the relative bias amplitude α (Eq.2.28) for the fiducial density (circles) compared to when B2 density is four times lower, 0.1 gal/arcmin^2, (triangles). The blue dotted lines correspond to B1xB2(0.8 or 0.5), the green dashed to B2(0.4 or 0.1) and the red solid to B1(0.4), where the values inside parenthesis indicate the total density of each population in gal/arcmin^2. All lines are normalized to the B1(0.4) forecast.

and B1xB2(0.5) is a factor ~ 2.3, which is lower because shot-noise is higher. For B2(0.4) the constraints are similar for $\alpha < 1$ (lower bias amplitude), and get worse with $\alpha > 1$ (more bias). Here two effects overlap: RSD effect becomes more important with lower bias which has a great impact in γ constraints, whereas a higher bias increases the amplitude of the correlations, which weakens the impact of shot noise, and in particular improves the ω constraints. For this reason, reducing B2 density has a larger impact on lower bias both in B2 and B1xB2 as it mitigates the benefits from RSD, and also as the signal from correlations is lower then shot noise is more predominant.

For a detailed study of the impact of RSD and BAO with bias and α in all FoMs see §2.A.1 to 2.A.3.

		Fiducial	xBias	No RSD	No BAO	No RSD-xBias	No BAO-xBias
10^{-3} FoM $\gamma\omega$:	B1xB2	13.7	117	1.61	9.25	41.4	64.2
	B2	5.88	36.1	0.62	4.37	16.7	24.7
	B1	5.53	45.4	1.45	3.78	37.9	31.6
FoM γ:	B1xB2	62	190	9.9	58	105	143
	B2	51	152	7.6	49	78	121
	B1	38	147	9.6	38	102	133
FoM ω:	B1xB2	221	615	163	160	395	450
	B2	116	238	82	90	212	204
	B1	147	310	152	100	373	238
FoM DETF:	B1xB2	237	875	209	171	841	787
	B2	129	513	106	104	479	422
	B1	180	801	196	137	797	696

Table 2.2: Sample variance cancellation for multitracing B1xB2(0.8) of two spectroscopic populations, B1(0.4) and B2(0.4), where the values inside parenthesis indicate the total density of each population in gal/arcmin^2. The relative bias amplitude between both populations is set to $\alpha = 0.5$. Each column show the impact of removing different effects, while rows show the single and overlapping population cases.

2.3.2 Fiducial model ($\alpha = 0.5$)

In this subsection we study several effects fixing $\alpha = 0.5$, which will be the fiducial value for the relative bias amplitude in the following subsections, except for section 2.4. Table 2.2 presents four tabulars, one for each FoM, with the two single population cases (B1(0.4), B2(0.4)) and the multitracer case (B1xB2(0.8)) for the rows. In the columns we present the fiducial case (labeled 'Fiducial') and the impact of some physical effects, like fixing bias ('xBias'), computing correlations in real space ('No RSD'), not including BAO wiggles ('No BAO') and combinations of these.

Looking first at the 'Fiducial' column, one sees how the multitracer case has better constraints than the single tracer cases, for all FoMs, due to sample variance cancellations. Comparing to the best single tracer, there is a 133% improvement for FoM$_{\gamma\omega}$, 23% for FoM$_\gamma$, 50% for FoM$_\omega$ and 32% for FoM$_{\rm DETF}$.

Galaxy bias can be fixed from lensing surveys and its cross-correlations with galaxy clustering (see Bernstein & Cai 2011, Cai & Bernstein 2012, Gaztañaga et al. 2012, Eriksen & Gaztañaga 2015c). Fixing bias greatly improves the constraints as it breaks strong degeneracies, but the gains from sample variance cancellations are still present, which shows that they are not caused by measuring bias. RSD allow to measure galaxy bias and the growth separately, but not the random nature of the fluctuations, so fixing bias will not break the degeneracy with the rms amplitude of fluctuations, but multiple tracers will. When removing redshift space distortions ('No RSD'), sample variance cancellations are no longer possible, and the gain for B1xB2(0.8)/B1(0.4) is much lower. Also, without RSD, our ability to measure γ drops, which translates in a much lower FoMs. Not including BAO measurements reduce the FoMs,

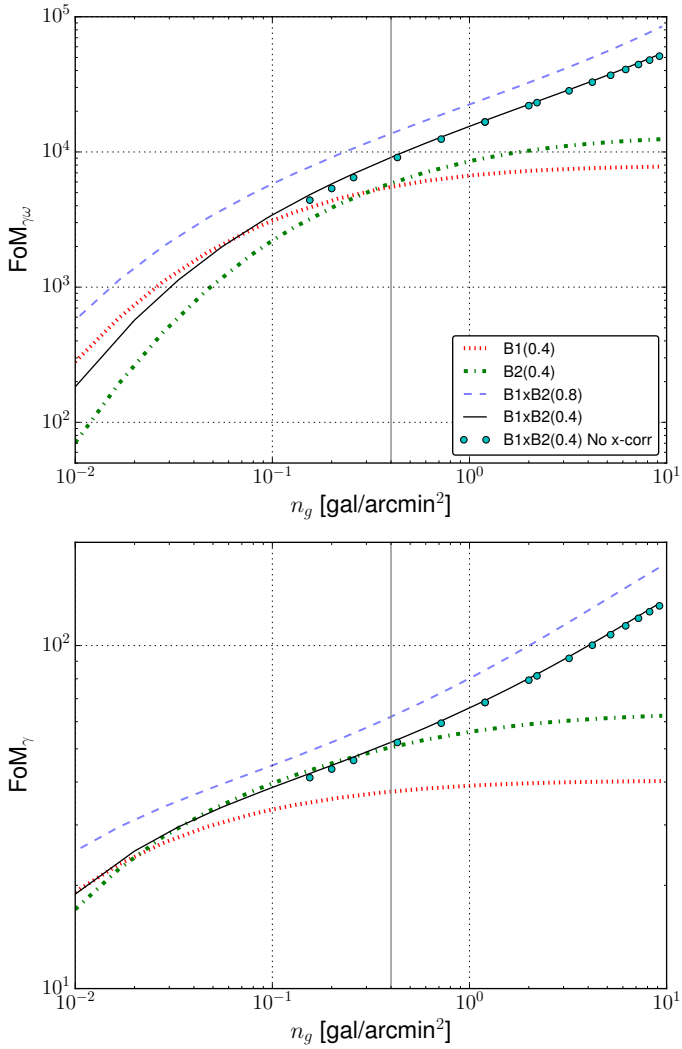

Figure 2.2: Impact of spectroscopic galaxy density on the constraints. The relative bias amplitude is fixed at $\alpha = 0.5$. The red dotted B1(0.4) and green dot-dashed B2(0.4) lines correspond to the single tracers. B1xB2(0.8) (blue dashed line) is the overlapping survey of merging B1(0.4) and B2(0.4), and thus has double the density of each alone. B1xB2(0.4) (black solid) has the same total density as B1(0.4) or B2(0.4) and is identical to B1xB2(0.8) except its tracers have halve the density. The cyan dots show B1xB2(0.4) without cross correlations between B1(0.2) and B2(0.2), which is equivalent to adding the auto correlations from each B1(0.2) and B2(0.2) population alone plus the covariance between them. The vertical line shows the fiducial density, 0.4 gal/arcmin^2.

affecting more the ω constraints while having little impact on FoM_γ (see §2.A for a discussion of the impact of RSD and BAO). We have also checked the effect of weak lensing magnification using the magnification slopes given in Eriksen & Gaztañaga (2015c). We find that they contribute less than 0.5%.

2.3.3 Galaxy density

The auto-correlations for a redshift bin include a shot-noise term (see §2.2.2) due to the discrete nature of the observable (galaxy counts), which depend on the galaxy density. Previously in the introduction of section §2.3 we have discussed that multiple tracers in redshift space can cancel sampling variance, and then our ability to improve our constraints is only limited to the signal-to-noise of the tracers (except when including bias stochasticity). Therefore, if there is no bias stochasticity, by increasing the survey density we can improve our cosmology constraints as much as we want. However, surveys usually have a fixed exposure time, so increasing survey density requires going deeper (longer exposures), which results in a smaller survey area. In this subsection we do not study this trade off between galaxy density and area, but increase galaxy densities for a fixed area. Moreover, spectroscopic surveys are characterized by having very good redshift determination since it has spectra where one can locate the emission lines, but it requires to take longer exposure times which results into lower densities.

Fig. 2.2 shows how $\mathrm{FoM}_{\gamma\omega}$ and FoM_γ depend on galaxy density. B1(0.4) and B2(0.4) correspond to the single tracer surveys. The blue line B1xB2(0.8) is a multiple tracer survey which merges the single tracer surveys B1(0.4) and B2(0.4) over the same area, as the multitracer surveys studied in Fig. 2.1 and Table 2.2. Therefore, it has double of the density of one single tracer alone, and the x-axis refers to the density of one of the subsamples of the survey. On the other hand, the black line B1xB2(0.4) studies the constraints when splitting one single tracer like B1(0.4) into two, keeping the total density, and thus the density of each subsample is reduced by half. Therefore, when comparing to the single tracers the black line addresses the gains from covariance and cross-correlations but adding shot-noise in the subsamples, and the blue line adds the gains from extra density from merging B1(0.4) with B2(0.4). Note that in a real survey we are interested in the gains coming from splitting into two subsamples like B1xB2(0.4), while B1xB2(0.8) shows the combined constraints the way B1(0.4) and B2(0.4) are fiducially defined.

As expected, the single tracer results flatten out at the high density limit and saturate. For multiple tracers there is sample variance cancellation and the constraints improve beyond the single tracer noiseless limit. At lower densities

we observe that B1xB2(0.4) and B1(0.4) lines cross. When shot-noise is already high we do not expect further splitting to improve the constraints. Moreover, B1xB2(0.8) shows better constraints than B1xB2(0.4) as the higher density reduces shot-noise. For the single tracers, in FoM_γ the constraints are similar for B1(0.4) and B2(0.4) at low densities, while we observe a clear difference (due to bias) between them on the noiseless limit. Lower bias populations (B2) get better constraints because in FoM_γ RSD is vital for breaking degeneracies and is enhanced with a lower bias amplitude (see §2.A.1 for details), but when shot noise becomes dominant then this effect disappears. On the other hand, in $\mathrm{FoM}_{\gamma\omega}$ we observe that B1(0.4) and B2(0.4) lines cross, as for the ω constraints a lower bias gives lower constraints in general, but this effect is more noticeable with high noise, as in that case having more signal is more relevant, while in the noiseless limit the constraints are similar for different bias (see, for example, right panel in Fig. 2.15).

The cyan dots correspond to removing cross correlations between B1(0.2) and B2(0.2) in B1xB2(0.4). This is equivalent to adding the correlations (transverse and radial within redshift bins), of each population B1(0.4) and B2(0.4) (B1+B2) plus the same sky covariance. It shows the relative importance of covariance between the tracers and the additional cross correlations in the gains that we are observing. We find that there is only a tiny contribution from cross-correlations ($< 2\%$ at high density), which shows that the multiple tracer improvement comes mainly from sample variance cancellations.

From now on we only show the multitracer B1xB2 with the same total density as the single tracers.

Fixing bias

Fig. 2.3 shows the constraints from B1xB2(0.4), B1(0.4) and B2(0.4) when fixing bias compared to the free bias case (free bias means marginalizing over the bias parameters). When we fix bias we break strong degeneracies and the constraints improve by an order of magnitude. We find that not all improvement comes from measuring bias, as we find similar relative gains with the fixed and free bias cases. The fact that for B1xB2(0.4) the free and fixed bias lines approach comes from the γ constraints, while for the ω constraints the difference is rather flat (not shown). Note that some extreme density values are shown, which are included to study the potential gains from a theoretical point of view.

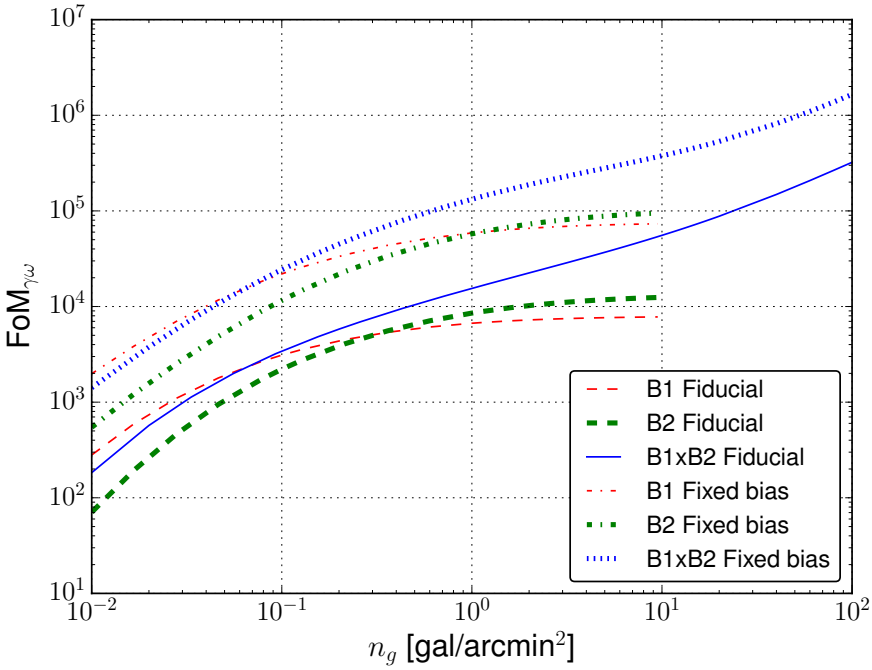

Figure 2.3: Shows the constraints of FoM$_{\gamma\omega}$ for the free and fixed bias case, for the B1(0.4), B2(0.4) and B1xB2(0.4) surveys. Free bias means marginalizing over the bias parameters, and corresponds to the fiducial forecast (Labeled as 'Fiducial').

2.4 Relation between bias and density

In the last subsections we have studied the constraints dependence on galaxy density and relative bias amplitude when splitting one spectroscopic population into two. When splitting by luminosity or absolute magnitude, brighter galaxies tend to live in more massive haloes, which tend to be more biased and less abundant. Therefore, there is a relation between galaxy density and bias. In the other sections we fix $\alpha = 0.5$ and ignore this relation to understand different physical effects from a theoretical point. To account for this effect, in this subsection we model galaxy bias using a conditional luminosity function (CLF) fitted to SDSS data from Cacciato et al. (2013) combined with a halo model (HM). The CLF determines how galaxies with a given luminosity populate dark matter haloes of different mass, $\Phi(L|M)$, while the HM set the abundance of dark matter haloes of a certain mass, $n(M, z)$. Using this modeling we define a magnitude limited survey $18 < r_{AB} < 23$ and we are able to determine the abundance of galaxies and galaxy bias as a function of redshift, halo mass or galaxy luminosity. To define the apparent limited survey we only consider luminosities in redshift such that $r_{AB}(L, z) \in [18, 23]$, since Cacciato et al. (2013) fit the HOD model using the SDSS r-band data.

2.4.1 Conditional Luminosity Function

The conditional luminosity function from Cacciato et al. (2013) has two separate descriptions for the central and satellite galaxies:

$$\begin{aligned}
\Phi(L|M) &= \Phi_c(L|M) + \Phi_s(L|M), \\
\Phi_c(L|M)\,dL &= \frac{\log e}{\sqrt{2\pi}\sigma_c} \exp\left[-\frac{(\log L - \log L_c)^2}{2\sigma_c^2}\right] \frac{dL}{L}, \\
\Phi_s(L|M)\,dL &= \phi_s^* \left(\frac{L}{L_s^*}\right)^{\alpha_s+1} \exp\left[-\left(\frac{L}{L_s^*}\right)^2\right] \frac{dL}{L},
\end{aligned} \tag{2.31}$$

where log is the 10-based logarithm and L_c, σ_c, ϕ_s^*, α_s and L_s^* are all function of halo mass M,

$$\begin{aligned}
L_c(M) &= L_0 \frac{(M/M_1)^{\gamma_1}}{[1 + (M/M_1)]^{\gamma_1 - \gamma_2}}, \\
L_s^*(M) &= 0.562\, L_c(M), \\
\alpha_s(M) &= \alpha_s, \\
\log\left[\phi_s^*(M)\right] &= b_0 + b_1(\log M_{12}) + b_2(\log M_{12})^2.
\end{aligned} \tag{2.32}$$

For the total set of CLF parameters we use the median of the marginalized posterior distribution given in Cacciato et al. (2013) for their fiducial model.

2.4.2 Halo Model

The comoving number density of haloes per unit halo mass can be well described (Press & Schechter 1974, Sheth & Tormen 1999) by

$$\frac{dn_h}{dM} = f(\sigma)\frac{\rho_m}{M^2}\frac{d\ln\sigma^{-1}}{d\ln M}, \tag{2.33}$$

where ρ_m is the mean density of the universe and $\sigma^2(M,z)$ the density variance smoothed in a top hat sphere at some radius $R(M) = (3M/4\pi\rho_m)^{1/3}$,

$$\sigma^2(M,z) = \frac{D^2(z)}{2\pi^2}\int dk\, k^2\, P(k)\,|W(kR)|^2\,, \tag{2.34}$$

where $W(x) = 3j_1(x)/x$. For the differential mass function $f(\sigma,z)$ we use the fit to the MICE simulation from Crocce et al. (2010),

$$f(\sigma,z) = A(z)\left[\sigma^{-a(z)} + b(z)\right]\exp\left[-\frac{c(z)}{\sigma^2}\right] \tag{2.35}$$

with $A(z) = 0.58(1+z)^{-0.13}$, $a(z) = 1.37(1+z)^{-0.15}$, $b(z) = 0.3(1+z)^{-0.084}$, $c(z) = 1.036(1+z)^{-0.024}$. We define the halo mass function in arcmin^2 units as

$$n_h(M,z) \equiv \frac{dN_h/dM}{d\Omega\, dz} = \left(\frac{\pi}{10800}\right)^2 \frac{c\,\chi^2(z)}{H(z)}\frac{dn_h}{dM}(z). \tag{2.36}$$

To model halo bias function we use the fitting function from Tinker et al. (2010),

$$b_h(M,z) = 1 - \frac{A(z)\,\nu^{a(z)}}{\nu^{a(z)} + \delta_c^{a(z)}} + B(z)\nu^{b(z)} + C(z)\nu^{c(z)} \tag{2.37}$$

where $\nu \equiv \delta_c/\sigma(M,z)$, $\delta_c \simeq 1.686$ is the linear density collapse, and where we use the parameter values from Table 2 with $\Delta = 200$ from the same paper (see also Hoffmann, Bel & Gaztañaga 2015 for other values).

2.4.3 Splitting methods

Once the halo mass function and the halo bias function are specified we can determine the galaxy number density and galaxy bias for an apparent limited survey. The average number of galaxies of a given halo mass with $L_1 < L < L_2$ is

$$\Phi(M,z) = \int_{L_1(z)}^{L_2(z)} \Phi(L|M)\, dL, \tag{2.38}$$

and the number density of galaxies per unit redshift is

$$\bar{n}(z) = \int_{M_{min}}^{M_{max}} \Phi(M,z)\, n_h(M,z)\, dM, \tag{2.39}$$

while the corresponding mean galaxy bias is

$$\bar{b}(z) = \int dM\, b_h(M,z)\, \Phi(M,z)\, n_h(M,z)/\bar{n}_g(z). \tag{2.40}$$

Here we define $L_1(z)$ and $L_2(z)$ such that $r_{AB}(L_2(z)) = 18$ and $r_{AB}(L_1(z)) = 23$. We integrate between $M_{min} = 10$ and $M_{max} = 15$ in $\log\left[M/M_{\odot}h^{-1}\right]$ units and consider $\Phi(L|M) = 0$ outside of this boundaries. To split the survey into two subsamples we consider two methods:

- Splitting by halo mass: split the spectroscopic sample introducing a M_{cut} in Eqs. 2.39-2.40 which defines two populations, B1 with $M_{min} < M < M_{cut}$ and B2 with $M_{cut} < M < M_{max}$.
- Splitting by apparent magnitude: split the spectroscopic sample introducing an $L_{cut}(z)$ in Eq.2.38 which defines two populations, B1 with $L_1 < L < L_{cut}$ and B2 with $L_{cut} < L < L_2$. Notice that $r_{AB}(L_{cut}(z)) = r_{cut}$.

Within this two methods we consider two cases, one in which the cutting variable (M_{cut} and r_{cut}) is the same for all redshifts. The another case fix the density ratio (i.e. $\bar{n}_1(z)/\bar{n}_2(z) = const.$) as a function of redshift by fitting the $M_{cut}(z)$ and $r_{cut}(z)$ which produces the corresponding density ratio. This results in a total of four different forecasts. Notice that fixing the density ratio cutting in apparent magnitude $r_{cut}(z)$ or absolute magnitude (luminosity) $L_{cut}(z)$ is the same.

Fig. 2.5 shows the four cases that have just been described for $\mathrm{FoM}_{\gamma\omega}$ in the left panels and FoM_{γ} in the right panels. Two density cases are studied,

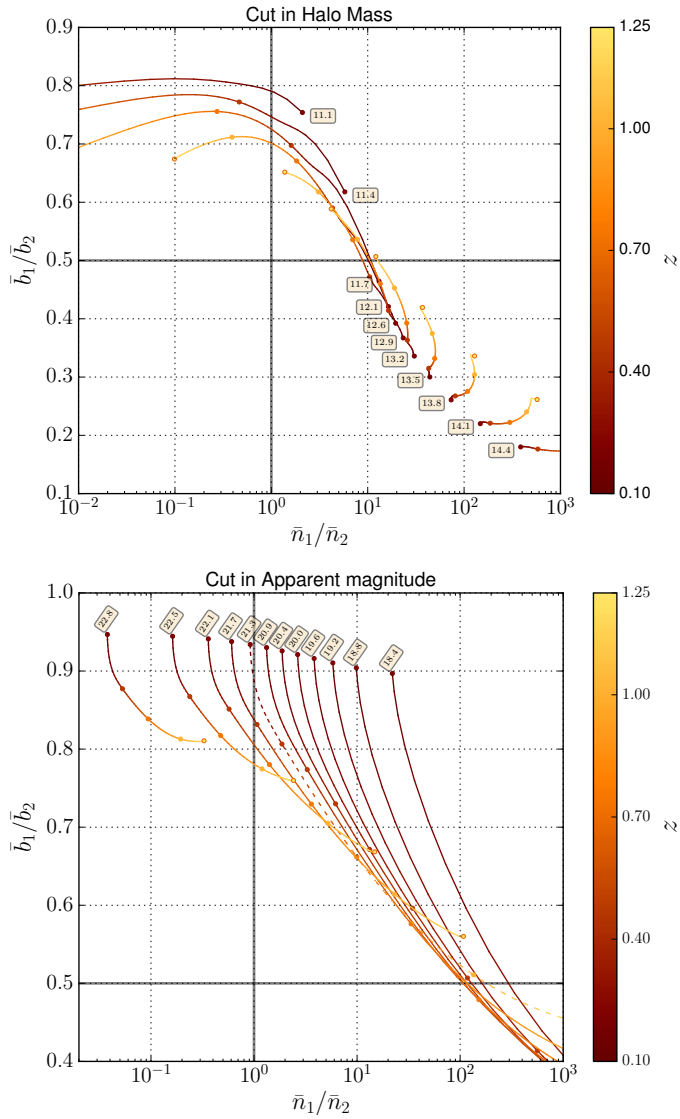

Figure 2.4: Density versus bias ratios between the two subsamples. The top panel shows the split with halo mass, M_{cut}, while the bottom panel shows a split with r-band magnitude, r_{cut}. Each line corresponds to a given M_{cut}/r_{cut}, which value is indicated in a box next to the start of each. The colorbar shows the redshift evolution for each line. There are 5 dots in each line indicating the position of $z = 0.1$, 0.4, 0.7, 1.0, 1.25. The dashed line shows the case where $\mathrm{FoM}_{\gamma\omega}$ is maximum (see the details in the text).

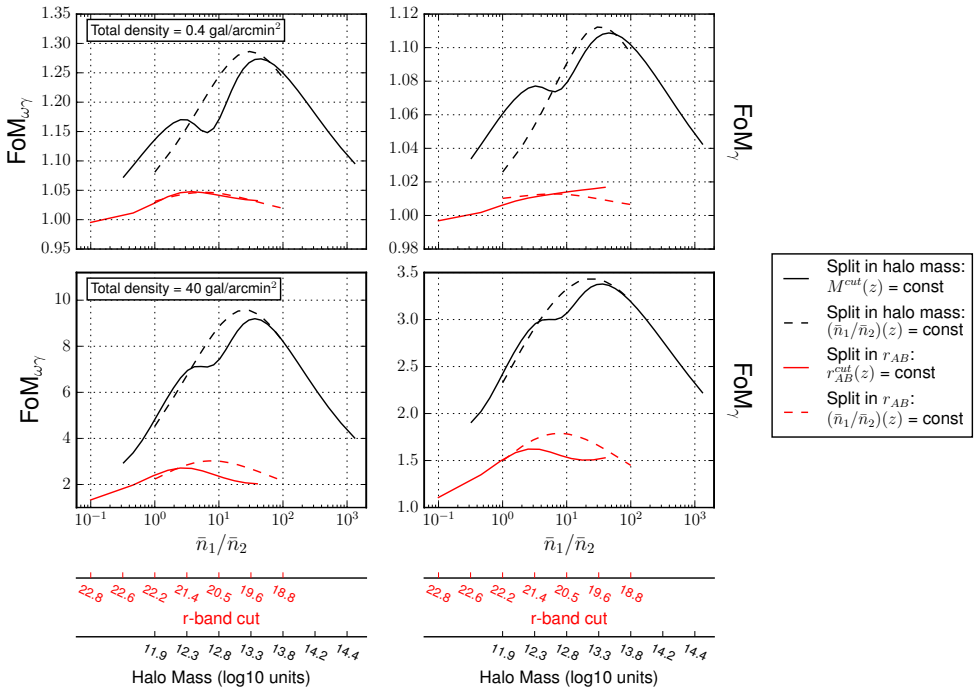

Figure 2.5: FoM$_{\gamma\omega}$ (left panels) and FoM$_{\gamma\omega}$ (right panels) when using a CLF $\Phi(L|M)$ and HM models to build an r-band limited magnitude survey, $r_{AB} = [18, 23]$. Two splitting methods are shown, splitting in halo mass, M_{cut}, and splitting in r-band, r_{cut}. Two cases are studied for each method, splitting with constant M_{cut}/r_{cut} in redshift and splitting with constant density ratio in redshift, $\bar{n}_1(z)/\bar{n}_2(z) \propto const$. The top panels have a total density of 0.4 gal/arcmin^2, while the bottom panels have 40 gal/arcmin^2. The x-axis shows the density ratio between the two subsamples, and the two twin axis show the correspondence of this ratio to a given constant M_{cut} and r_{cut} in redshift. All lines are normalized to the FoM when not splitting the galaxy sample.

0.4 gal/arcmin^2 (top panels) and 40 gal/arcmin^2 (bottom panels). The x-axis shows the density ratio between the two subsamples for each case, while the two twin axis show the correspondence of this density ratio to the cutting variable (halo mass and apparent r-band magnitude) for the two cases in which the cutting variable is constant in redshift. All lines have been normalized to the FoM when not splitting the galaxy sample.

Fig. 2.5 shows that a split of galaxies using the halo mass gives a better improvement in the constraints than splitting with apparent magnitude. Splitting with halo mass improves up to a factor 1.27 in FoM$_{\gamma\omega}$ with low density (top left panel) while splitting with an r-band cut gives a factor 1.05. The peaks are found at halo mass $M_{cut} \simeq 13.5$ ($\log[M/M_\odot h^{-1}]$) and $r_{cut} \simeq 21.3$. Forcing the density ratio between the subsamples to be the same in redshift (labelled as cut with constant density in Fig. 2.5) slightly improves the constraints to a factor 1.29 for a cut in halo mass and leaves it near the same for an r-band split. When using a denser population (bottom left) the improvement raises to a factor 9.2 in FoM$_{\gamma\omega}$ for a halo mass split and a factor 2.7 for r-band split. When fixing the density ratio the factors are 9.6 and 3.0, respectively. The maximum gains are obtained for $\bar{n}_1/\bar{n}_2 \sim 7$ when cutting in r-band and $\bar{n}_1/\bar{n}_2 \sim 30$ when cutting in mass. In practice, one does not need to know the mass or the r-band, but only to have an observational proxy that allows to rank the galaxies to allow the sample split (e.g., richness in the case of halo mass). For FoM$_\gamma$ and low density (top right) the factors are 1.11 and 1.02 for $M_{cut} \simeq 13.5$ and $r_{cut} \simeq 19.4$, although for the r-band cut the maximum would be found at brighter cuts which were numerically unstable. For a denser survey (bottom right), when fixing the density ratio, the constraints improve up to a factor 3.43 for halo mass and 1.79 for r-band.

When splitting a population into two subsamples one want to maximize the bias difference in redshift between them while keeping their densities as similar as possible in order to maximize the FoM. To do so, we would like to have a quantity that increase monotonically with bias with small scatter. Halo mass is such a quantity and so it maximizes the FoM. Splitting in apparent magnitude gives a distribution in halo mass, $\Phi(L|M)$, reducing the bias difference.

Fig. 2.4 shows the density-bias ratio evolution in redshift for different cut values when cutting in halo mass (top panel) and r-band magnitude (bottom panel). The dots in the figure show the position of the 5 ticks from the colorbar ($z = 0.1$, 0.4, 0.7, 1.0, 1.25). For a halo mass cut the bias difference is low when splitting at low halo masses as bias evolves linearly in that regime and the abundance of galaxies overweights that region in front of the high biased one. Cutting at higher masses results into an increasingly greater bias difference,

but also makes a more uneven density split. The maximum in $\mathrm{FoM}_{\gamma\omega}$ is found at $M_{cut} \simeq 13.5$, which has a similar density ratio in redshift $\bar{n}_1/\bar{n}_2 = 40 \sim 50$ and a bias difference of $\alpha = \bar{b}_1/\bar{b}_2 = 0.3 \sim 0.4$.

When splitting with apparent magnitude (Fig. 2.4, lower panel) the density ratios quickly span over large ranges in redshift when the bias difference increases, which limits the amount of improvement. For most magnitude cuts an important part of the distribution is very unevenly splitted, which increase the shot-noise. Furthermore, at a density ratio of $40 \sim 50$, (i.e. the peak with a halo mass cut in Fig.2.5), there is no magnitude cut at any redshift which produces an $\alpha \lesssim 0.55$, which is a factor $1.4 \sim 1.8$ less bias difference than in the halo mass situation. In the high density case we are not shot-noise dominated and thus the improvement goes from a marginal 5% to a 3 times better $\mathrm{FoM}_{\gamma\omega}$.

In addition, Fig. 2.5 shows a relative minimum at $M_{cut} \simeq 12.6$ and a relative maximum at $M_{cut} \simeq 12.1$ for the halo mass cuts at lower density cases, in both $\mathrm{FoM}_{\gamma\omega}$ and FoM_{γ}. Fig. 2.4 shows that although $M_{cut} \simeq 12.6$ has a $10\% \sim 15\%$ greater bias difference depending on redshift it has a more uneven density split. A cut in $M_{cut} \simeq 12.1$ gives a density ratio in redshift which extends over $\bar{n}_1/\bar{n}_2 \sim [0.1, 16]$, with some cuts in redshift being close/equal to a density ratio of unity, which maximally reduces shot noise, whereas a cut in $M_{cut} \simeq 12.6$ results in $\bar{n}_1/\bar{n}_2 \sim [1.3, 20]$. The increment in bias difference does not compensate the induced shot noise. With higher density (Fig. 2.5 lower panels) shot-noise has a lower impact and the relative minimum disappears resulting in a flattened region instead.

Moreover, we have split in absolute magnitude (not shown) by fixing the luminosity cut L_{cut} as a function of redshift. The FoM were worse than with an apparent magnitude cut, and in most cases worse than not splitting the sample at all. Having a magnitude limited survey gives an incompletness of luminosity in redshift, meaning that several redshift ranges have very few galaxies or no galaxies at all, which introduces large amounts of shot-noise.

2.5 Partly overlapping redshift bins

In Fig. 2.6 we show the effect of having partly overlapping redshift bins between two spectroscopic populations (B1xB2) by shifting the beginning of the redshift range z_{min} of one of the populations (B1) while keeping the other fixed. This shifts all the B1 redshift bins with respect to the B2 ones and determines the amount of overlap between them. The total density of both the multitracer and single tracers is equal to 0.4 gal/arcmin^2. In Fig.

2.6, the panels on the left show the FoM normalized to the fully overlapping bins value (i.e. normalized to the B1 $z_{min} = 0.1$ or $0\,\Delta z$ shift value of the FoMs) for $\mathrm{FoM}_{\gamma\omega}$, FoM_{γ} and FoM_{ω}, while the panels on the right show the absolute values. The fiducial forecast line (red solid) shows oscillations that are minimum at the edges of the fiducial binning (marked by vertical grey lines on the plots) and are maximum when the redshift bins half overlap with each other (when B1 bins start in the middle of a B2 bin and viceversa). In the fiducial forecast we parametrize bias with one parameter per redshift bin and tracer. The black dashed lines show an alternative bias parameterization which parametrize the bias with four redshift pivot points $z_i \in [0.25, 0.43, 0.66, 1.0]$ and linearly interpolate between them. We find similar constraints from both bias parameterizations and this shows that the gain does not artificially come from the choice of bias parameterization.

When bins half overlap with each other (when B1 bins start in the middle of a B2 bin and viceversa) the gain is maximum, a factor 1.33 for $\mathrm{FoM}_{\gamma\omega}$, 1.06 for FoM_{γ}, 1.26 for FoM_{ω} and 1.33 for $\mathrm{FoM}_{\mathrm{DETF}}$ (not shown). Having partially overlapping bins induces an effective thinner binning that allows to probe smaller scales which improve constraints. Most of this improvement comes from the cross correlations between both populations, as the smaller scales information comes mostly from cross-correlating with the shifted bins. When removing them (red solid to pink dash-dash-dot line) the gain factors at the peaks reduce to 1.07 for $\mathrm{FoM}_{\gamma\omega}$, 1.00 for FoM_{γ}, 1.07 for FoM_{ω} and 1.12 for $\mathrm{FoM}_{\mathrm{DETF}}$, and for FoM_{γ} (center left panel) shifting bins even leads to worse constraints. When B1 z_{min} starts exactly at the second bin of B2 the constraints drop as all bins perfectly overlap again, but with the forecast having one less bin the FoMs are slightly lower compared to the fiducial forecast. The effect of removing the first bins does not reduce much the FoMs as the first bins are often removed from cutting in k, but the FoMs eventually start to drop when removing more bins.

When fixing bias (blue dotted line) the absolute constraints greatly improve, as expected from breaking degeneracies. FoM_{ω} shows substantially relative lower oscillations (bottom left panel), which means that part of the improvement came form better measuring bias, while FoM_{γ} shows greater oscillations when bias is fixed. $\mathrm{FoM}_{\gamma\omega}$ combines these effects and improves a factor 1.2 at the peak when fixing bias (10% lower than the fiducial). When removing RSD (purple dot-dashed), the constraints for γ reduce considerably and look flat in the absolute values (center right panel). FoM_{γ} shows higher oscillations (center left panel), but now these come from measuring γ directly from the growth rate in front of the power spectrum, and not from RSD. In FoM_{ω} the constraints are worse, but the relative gains are very similar.

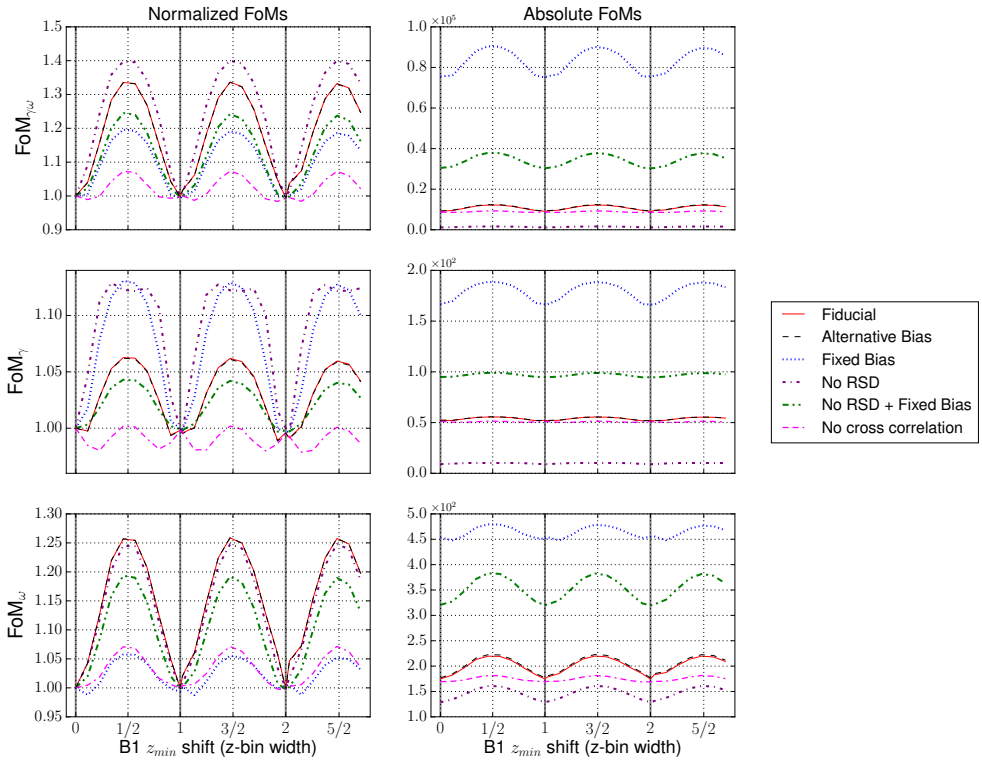

Figure 2.6: Effect of having partly overlapping redshift bins when combining two spectroscopic surveys (B1xB2). The start z_{min} of the B1 redshift range determines the overlap between the redshift bins of both populations. The x-axis shows the B1 z_{min} shift in z-bin width units, Δz. The panels on the left show the FoM normalized to the fully overlapping bins value (i.e. B1 $z_{min} = 0.1$ or shift $= 0$) for FoM$_{\gamma\omega}$, FoM$_{\gamma}$ and FoM$_{\omega}$, while the panels on the right show the absolute values. The black (dashed) line uses 4 redshift pivot points $z_i \in [0.25, 0.43, 0.66, 1.0]$ to parametrize bias instead of the fiducial 1 parameter per redshift bin and population. The pink dash-dash-dot line does not include cross correlations between B1 and B2. The blue dotted line is the fixed bias case, the purple dot-dashed line corresponds to removing RSD, and the green dash-dot-dot line combines fixed bias and no RSD. The grey vertical lines show the fiducial (B2) redshift bin edges.

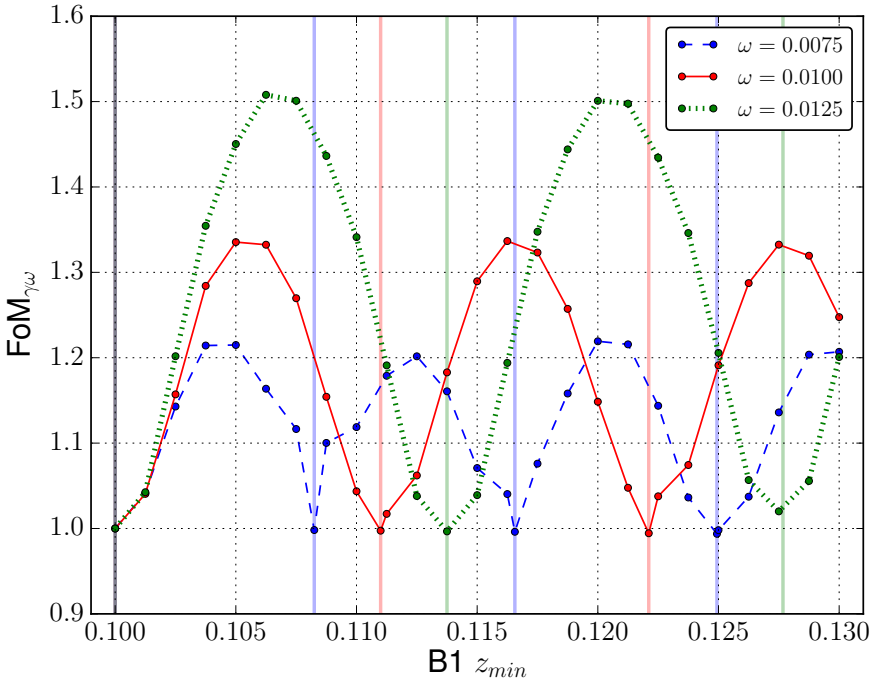

Figure 2.7: Amplitude of the oscillations when having partly overlapping bins between two populations, for different bin widths. For $\Delta z = w(1+z)$, the lines correspond to: red solid, $w = 0.01$ (fiducial value); blue dashed, $w = 0.0075$; green dotted, $w = 0.0125$. All lines are normalized to their respective values at B1 $z_{min} = 0.1$. The vertical lines show the position of the bin edges for each case.

Fig. 2.7 shows the impact of the redshift bin width on the oscillations in FoM$_{\gamma\omega}$. We parametrize the bin width as $\Delta z = w(1+z)$. The lines correspond to: $w = 0.01$ (red solid, fiducial value), $w = 0.0075$ (blue dashed) and $w = 0.0125$ (green dotted). All lines are normalized to their respective values at B1 $z_{min} = 0.1$. It shows that redshift bin width has an important impact on the relative gains. For the thinner binning the relative improvement is only of a factor 1.2, while for the thick binning is $\sim$1.5 (the fiducial is 1.33). This shows that if the binning is narrower the relative gain is lower as the radial resolution is better, but recall that the maximum resolution is limited by only using linear scales (k_{max}).

Fig. 2.8 shows B1xB2 as function of the relative bias amplitude for different B1 z_{min} shift values, which shows the impact of partially overlapping bins from having full overlap (B1 z_{min} shift $= 0.0\,\Delta z$, red solid line) to almost half overlap (B1 z_{min} shift $= 0.45\,\Delta z$, blue dashed). In FoM$_{\gamma}$ (top right panel) increasing the partial overlap has several effects. When bias amplitudes are

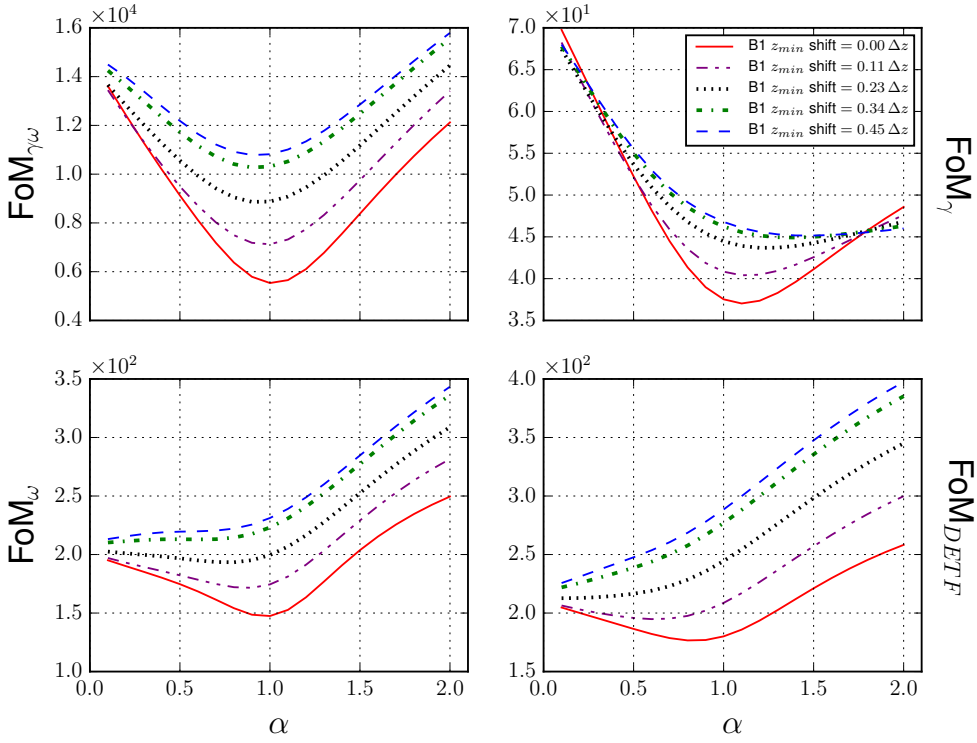

Figure 2.8: Effect of partially overlapping bins in B1xB2 as function of the relative bias amplitude between both populations (α). The lines correspond to B1 $z_{min} = 0.1$ (red solid), 0.10125 (green dash-dot), 0.1025 (black dotted), 0.10375 (purple dash-dot-dot), 0.105 (blue dashed).

similar there is more gain from partial overlap, while when the bias amplitude grows this gain decreases until the point that shifting bins leads to worse constraints. Also, for half overlapping bins FoM$_\gamma$ flattens for $\alpha > 1$. For FoM$_\omega$ and FoM$_{\rm DETF}$ there is always gain from partial overlap. The different lines are closer for lower α where the gain is minimum, which increases until $\alpha = 1$. From that point the lines are quite parallel. When there is full overlap we have the minimum at $\alpha = 1$ (same bias case) and the FoMs increase with the bias difference, but when there is half overlap between the redshift bins of both populations B1xB2 behaves like a single tracer, in the sense that FoM$_\gamma$ decreases with bias while FoM$_\omega$, FoM$_{\rm DETF}$ increase with bias (see Fig. 2.11). On the other hand, FoM$_{\gamma\omega}$ combines the effects from FoM$_\gamma$ and FoM$_\omega$ and keeps the minimum, increasing the FoM for higher partial overlap, meaning the gain is higher when bias is similar.

2.5.1 Radial resolution

In this subsection we study the impact of increasing the number of spectroscopic redshift bins. In the fiducial forecast we use spectroscopic surveys with 71 narrow redshift bins, such that at each bin we mainly account for transverse modes from angular spectra, while the radial information (modes) is contained in the cross correlations between redshift bins. This tomography study can approximately recover the full 3D clustering information when the comoving redshift bin separation, $\Delta r = c\Delta z/H(z)$, corresponds approximately to the minimum linear 3D scale $\lambda^{3D}_{min} = \frac{2\pi}{k_{max}}$, (Asorey et al. 2012). As we are limited by the linear regime, including more bins would eventually lead to include nonlinear modes, which would require modeling the nonlinear angular power spectrum. The bin width $\Delta z = w(1+z)$ is set by the number of redshift bins,

$$w = \sqrt[N_z]{\frac{1+z_{max}}{1+z_0}} - 1, \tag{2.41}$$

which divide the interval $[z_0, z_{max}]$ into N_z redshift bins (see Eriksen & Gaztañaga 2015c).

Fig. 2.9 shows how the constraints improve when increasing the number of redshift bins for $\mathrm{FoM}_{\gamma\omega}$. The lines show B1 (blue dotted), B2 (green dot-dash), B1xB2 increasing both B1 and B2 redshift bins (red solid), and B1xB2 keeping fixed B1 number of redshift bins to 71 (black dashed). Both single populations and combined surveys improve when increasing the number of bins. There are several effects when we increase the number of redshift bins N_z (see Asorey et al. 2012). As the redshift bin width corresponds approximately with the minimum scale, increasing the number of redshift bins in the same redshift range allows for probing smaller scales, which gives more independent modes and improve constraints. The signal to noise at each bin remains nearly constant as the auto power spectrum has higher clustering from having more close pairs while the density per bin is lower, increasing both signal and shot noise. Therefore, increasing N_z increase the number of transverse modes without lowering their signal to noise, which improve constraints. Eventually, the redshift bins will become correlated and this gain will saturate (see Eriksen & Gaztañaga 2015c). In addition, when increasing N_z we also add more cross-correlations between redshift bins, which is very important as RSD depends on the relation between radial and transverse modes ($\mu = k_\parallel/k$).

In Fig. 2.9 we also study the effect of fixing the bins of one of the populations in the overlapping survey. The black dashed line refers to B1xB2 with

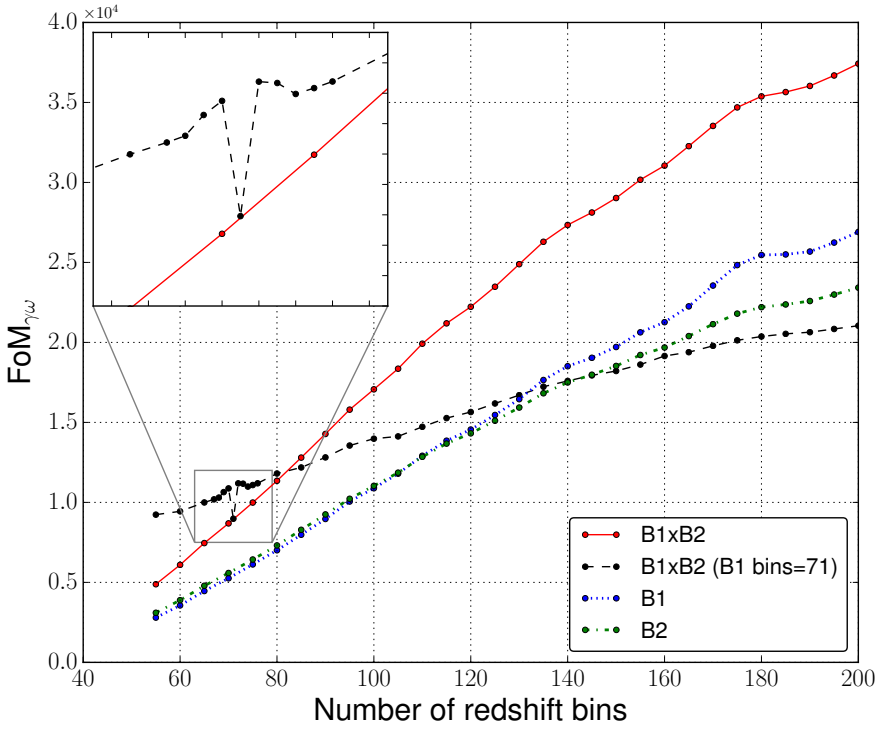

Figure 2.9: Effect of increasing the number of spectroscopic redshift bins in FoM$_{\gamma\omega}$. The lines correspond to: B1xB2 (red solid), B1xB2 with B1 bins fixed to 71 (black dashed), B1 (blue dotted) and B2 (green dot-dashed).

the B1 redshift bins fixed to 71, while the number of B2 bins vary. This results in a flatter improvement than when increasing the bins for both populations as we are adding less redshift bins. An interesting behavior happens when B1 and B2 have similar (but different) number of redshift bins (zoomed region). As we have previously discussed, having two populations with partly overlapping bins improve the constraints. $\mathrm{FoM}_{\gamma\omega}$ improves by a factor 1.2 (1.25) when using one less (more) redshift bins in B2. This gain is equivalent to be using ~ 80 redshift bins for both populations instead of the fiducial 71.

Sample variance cancellation in photometric surveys

Using thicker redshift bins we can model the loss of radial information and study the impact of sample variance cancellation in photometric surveys. Fig. 2.10 shows $\mathrm{FoM}_{\gamma\omega}$ with ($\alpha = 0.5$, 2.0) and without ($\alpha = 1.0$) cancellation for a smaller number of redshift bins (recall that with $\alpha = 1$ cancellation is no longer possible). We show the effect for typical photometric surveys such as DES or EUCLID ($5 \sim 10$ bins), for narrow photometric surveys like PAU (~ 40 bins) and compare it to spectroscopic-like surveys (fiducial 71 bins). The top panel shows that $\mathrm{FoM}_{\gamma\omega}$ improves strongly when increasing the radial resolution, specially in the free bias case where it improves 3 orders of magnitude from 5 to 100 bins as opposed to 2 orders in the fixed bias case. Radial cross-correlations are more important in a free bias forecast, which leads to more radial dependence. Bottom panel shows the sample variance cancellation effect in photometric surveys in the ratio between $\alpha = 2.0$ and $\alpha = 1.0$. We find that the contribution from sample variance cancellation has a larger or at least similar effect in photometric surveys for both free and fixed bias. Note how density (e.g. compare to Fig. 2.2) is as important as redshift accuracy in that it can also change $\mathrm{FoM}_{\gamma\omega}$ by 3 orders of magnitude.

Table 2.3 shows there is a strong density dependence in the ratio, specially in surveys with less number of redshift bins. While for spectroscopic surveys (e.g. 71 bins and 0.4 gal/arcmin^2) the ratios are 2.2 and 2.4 for free and fixed bias, for photo-z surveys (e.g. 5 bins and 8.0 gal/arcmin^2) the ratios become 19.5 and 5.32. Surveys with a low number of redshift bins are more dominated by radial auto-correlations, which are affected by shot-noise, so they become more density dependent, as shown in Table 2.3. For a survey density of 0.4 gal/arcmin^2 auto-correlations contribute up to 72% for a survey with 10 redshift bins with free bias, while only 29% and 10% for 40 and 71 bins. This high density dependence explains the turnover in the free bias ratio for low number of bins in Fig. 2.10, where shot-noise limits the gains from sample variance cancellations, and also produces very large gains for higher densities

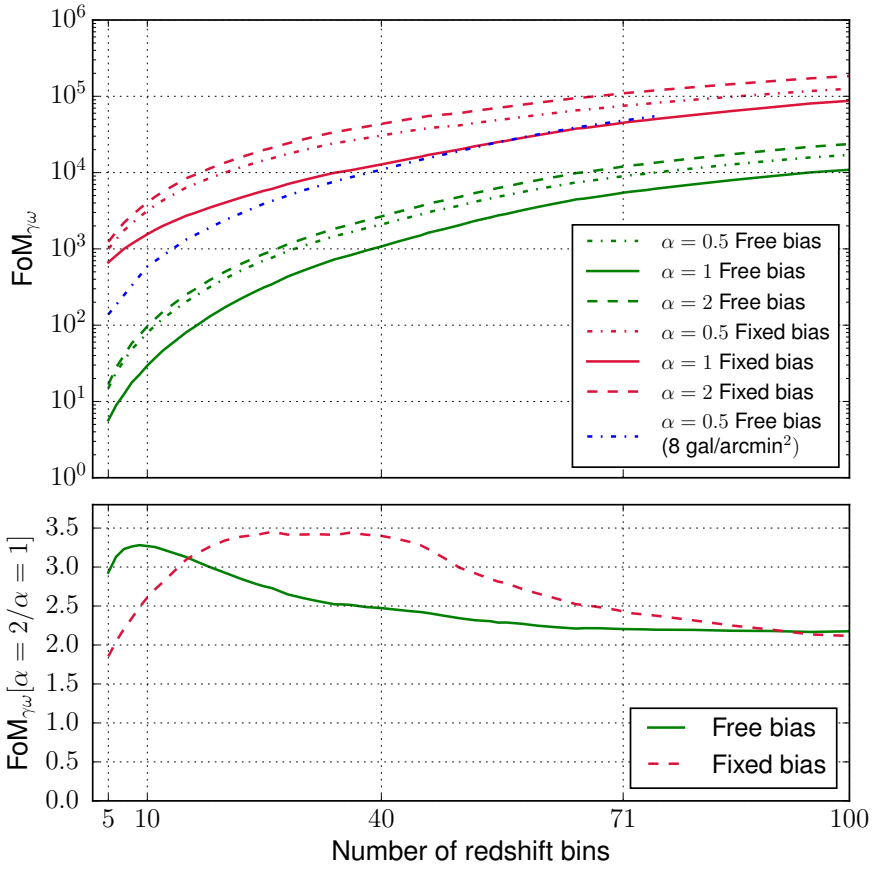

Figure 2.10: Effect of reducing the radial information by using less number of redshift bins, keeping the fiducial density (0.4 gal/arcmin^2). A binning of $5 \sim 10$ bins corresponds to typical photometric surveys, a binning of 40 correspond to narrow photometric surveys and 71 is the fiducial binning for spectroscopic surveys in this work. Top panel shows FoM$_{\gamma\omega}$ for $\alpha = 0.5, 1.0, 2.0$ for free and fixed bias. Bottom panel shows the ratio FoM$_{\gamma\omega}[\alpha = 2/\alpha = 1]$. Free bias with $\alpha = 0.5$ and density 8 gal/arcmin^2 is shown for comparison as a typical photometric survey density.

Bias	Density (gal/arcmin^2)	5 bins	10 bins	40 bins	71 bins
Free	0.4	2.93	3.27	2.47	2.20
	8.0	19.5	16.2	8.81	6.94
	80.0	138	93.4	66.1	38.4
Fixed	0.4	1.85	2.61	3.40	2.43
	8.0	5.32	7.10	8.80	5.29
	80.0	13.4	15.2	29.3	21.3

Table 2.3: Ratio between $\mathrm{FoM}_{\gamma\omega}[\alpha = 2/\alpha = 1]$, which shows the relative contribution of sample variance cancellation (see the text). On the columns there are 4 different binnings (5, 10, 40 and 71 redshift bins), while on the rows there are 3 different total survey densities for free and fixed bias.

as shown in Table 2.3.

Concerning absolute FoMs for photo-z and spec-z surveys, the loss of radial resolution from photometric uncertainties is in some part compensated by a gain in density in photometric surveys. The combined figures of merit $\mathrm{FoM}_{\gamma\omega}$ for a survey without split take the values of 7.19 and 5453 for photo-z (5 bins and 8.0 gal/arcmin^2) and spec-z (71 bins and 0.4 gal/arcmin^2) surveys. Using the fiducial configuration ($\alpha = 0.5$) with sample variance cancellation leads to a $\mathrm{FoM}_{\gamma\omega}$ of 138 and 8979, reducing the ratio between the FoM between both surveys by a factor ~ 12 (compare blue and green lines in Fig. 2.10).

2.6 Discussion

In this chapter we have estimated dark energy (ω_0, ω_a) and growth rate (γ) constraints of multiple tracers in spectroscopic surveys using the Fisher matrix formalism. In the fiducial forecast we use galaxy clustering from 2D angular correlations in 71 narrow redshift bins which include baryon acoustic oscillation (BAO) in the linear Eisenstein-Hu power spectrum (Eisenstein & Hu 1998) and the linear Kaiser effect (Kaiser 1987) to account for redshift space distortions (RSD). To compress the information of our constraining power in one number we define four Figures of Merit: $\mathrm{FoM}_{\gamma\omega}$, FoM_{γ}, FoM_{ω} and $\mathrm{FoM}_{\mathrm{DETF}}$ (§2.2.4). Details of the modelling and fiducial forecast assumptions can be found in section 2.2.

Section 2.3 studied how multiple tracers in the same region of the sky can break degeneracies and improve the constraints on dark energy and growth rate. We split one spectroscopic survey into two populations (named B1 and B2) by some galaxy property such that there is a relative bias amplitude (α)

between both populations. This allows for multi-tracing the same underlying matter distribution, which cancels the random nature (sample variance) of fluctuations in redshift space. Using redshift space distortions and two populations as multiple tracers we quickly improve the constraints when increasing the bias difference between the tracers, with no or very little improvement with no bias difference (Fig. 2.1). Fixing the bias ratio to $\alpha = 0.5$, we increased the galaxy density and showed that B1xB2 FoMs outperform the single tracers by beating the sampling variance limit in the noiseless limit (Fig. 2.2). We showed that this improvement comes from sample variance cancellation and not from additional cross-correlations between B1 and B2, which contribute less than 2%. Also, Fig. 2.3 showed that not all improvement was coming from measuring bias.

In section 2.4 we have set up an r-band limited survey with a CLF and HM to model galaxy bias. We have split the survey into two subsamples cutting in halo mass and r-band magnitude, computing the galaxy bias and galaxy density of each subsample in redshift. In this way, we account for the relation between galaxy bias and galaxy density. We showed that for a cut in halo mass we can improve $\mathrm{FoM}_{\gamma\omega}$ up to a factor 1.3 as compared to doing no split. Splitting in r-band magnitude lead to a factor 1.05 improvement as magnitude scatter halo mass which reduced the bias difference. When increasing the total density of the survey we found huge improvements for both split methods, giving a factor 9.6 and 3.0 in $\mathrm{FoM}_{\gamma\omega}$ for halo mass and r-band cut, respectively.

In section 2.5 we have studied the effect of having partially overlapping redshift bins in a multiple tracer survey. We have shown that as a result of the overlap the FoMs improve (Fig. 2.6), having a peak when the redshift bins are shifted half of the bin width. At the peak the FoMs improve a factor 1.33 for $\mathrm{FoM}_{\gamma\omega}$, 1.06 for FoM_{γ}, 1.26 for FoM_{ω} and 1.33 for $\mathrm{FoM}_{\mathrm{DETF}}$. We have shown that the gain is not artificially produced by the particular bias parameterization, but it is rather coming from the cross-correlations between B1 and B2. We have also shown that the gain is a factor ~ 1.5 in $\mathrm{FoM}_{\gamma\omega}$ when using a 25% thicker binning, while a factor ~ 1.2 when using a 25% thinner binning (Fig. 2.7). Fig. 2.8 showed how FoMs improve from partly overlapping bins for different α values, indicating that there is more improvement when relative bias difference is small, and low or no improvement when there is more bias difference between populations. In Fig. 2.9 we showed how FoMs improve when increasing the number of redshift bins. When using one more or less redshift bin in one of the populations, the improvement in $\mathrm{FoM}_{\gamma\omega}$ is equivalent to having 80 redshift bins in both populations compared to the 71 from the fiducial forecast. In section 2.5.1 we find that the multi-tracer gains are larger for photometric samples, specially when we increase the density. In

fact, having a larger density can compensate the loss in resolution (see Fig. 2.10).

In §2.A.1 to 2.A.3 we have studied the impact of RSD and BAO effects in the FoMs. In Fig. 2.11 we show the improvement from B1xB2 for different bias difference in the four FoMs. The constraints quickly improve in all FoMs, up to a factor 4 in $\mathrm{FoM}_{\gamma\omega}$. In real space (no RSD, Fig. 2.12) sample variance cancellations are no longer possible and the improvements are less than a factor 1.5 in $\mathrm{FoM}_{\gamma\omega}$, having almost no gain in FoM_{γ} (less than 2%). In Fig. 2.13 we have shown that RSD is very important in γ constraints as it breaks degeneracy between galaxy bias and $f(z)$. For FoM_{γ} the B1xB2 ratio RSD/No RSD is $4 \sim 9$, while the ratio BAO/No BAO is lower than 1.08. For dark energy constraints it happens the opposite, in FoM_{ω} the BAO/No BAO rate is $1.4 \sim 1.5$ while RSD/No RSD is ~ 1.0 when bias amplitudes are similar, but up to ~ 1.5 when there is more bias difference. The ω constraints depend more on the shape of the power spectrum and then are enhanced more from including BAO wiggles.

In §2.A.4 we have fixed cosmological parameters to study possible degeneracies. We find that multiple tracers help breaking degeneracies in the dark energy constraints, specially at high densities. In §2.B we studied the impact of a different bias evolution slope in redshift (Fig. 2.15). It shows a dependence for the B1 population, but the B1xB2 forecast is insensitive to the bias slope.

The trends found in Section 3 are in good agreement with a number of studies such as McDonald & Seljak (2009), White, Song & Percival (2009), Gil-Marín et al. (2010), Bernstein & Cai (2011). Direct quantitative comparison is hard to make due to very different survey configurations, along with different modeling and observables. Gil-Marín et al. (2010) suggested that cuts other than halo mass such as peak-height ν might be more competitive for a dark matter haloes split, as they find a $\sim 10\%$ improvement at low redshift with a cut in halo mass. In section 2.4 we have stressed that these gains are highly density dependent and that the split for galaxies can be optimized by looking at the bias-density relation of the tracers.

In this analysis we have assumed a number of idealizations, such as linear theory, deterministic bias, no stochasticity between tracers nor nonlinearities. Several studies indicate that linear theory (Kaiser 1987) start to break down at scales as large as $k_{max} > 0.02\,h\mathrm{Mpc}^{-1}$ (Okumura & Jing 2011, Bianchi et al. 2012), specially at low mass haloes, and that scale dependence in β varies between tracers. As shown in Gil-Marín et al. (2010) even small amounts of nonlinearity can degrade your FoM down to 50%, which emphasize the need for more realistic models for galaxies and nonlinear RSD (e.g. Reid & White 2011,

Okumura et al. 2015). Lately, a number of techniques have been developed to reduce shot-noise and stochasticity under the Poisson level by optimally weighting the tracers (see Seljak, Hamaus & Desjacques 2009, Hamaus et al. 2010, Cai, Bernstein & Sheth 2011, Pearson, Samushia & Gagrani 2016) which become of most interest in combination with multi-tracer surveys and could further improve the FoMs.

In summary, our results suggest that we can improve FoM significantly and break degeneracies in cosmological inference if we split the samples by a density ratio of $\bar{n}_1/\bar{n}_2 \sim 7$ using apparent magnitude as ranking or $\bar{n}_1/\bar{n}_2 \sim 30$ using a mass halo proxy ranking (e.g., richness). Using another proxy for bias (such as local density, see Pujol et al. 2015, or color) to split the sample, will give similar benefits. Splitting volume limited samples does not provide significant improvements. Our analysis also shows that when doing angular clustering tomography is optimal to use overlapping bins for cross-correlation. These finding can be applied to future redshift surveys such as DESI or Euclid and will also work for photometric samples (see Fig. 2.10), such as DES and LSST and its cross-correlations (as shown in Eriksen & Gaztañaga 2015c). We also show large improvement on the FoM with increasing galaxy density. This can be used as a trade-off to compensate a possible loss of radial resolution when using high resolution photometric redshifts (Martí et al. 2014) instead of spectroscopic redshifts.

2.A RSD and BAO effects in sample variance canellation and degeneracies

2.A.1 Redshift space

In Fig. 2.11 we present our results for the four FoMs. The blue (dotted) lines corresponds to the same sky case (B1xB2), the green (dashed) lines to B2 and the red (solid) lines to B1, with all lines being normalized to the B1 lines. Recall from Eqs. 2.27-2.28 that a change in α modifies the B2 bias but leaves the B1 bias equal to the fiducial. Focusing first on the B1xB2 lines, the four FoMs show an improvement compared to the single tracer cases, as expected, due to sample variance cancellations. We can see how the improvement for B1xB2 increases with the bias difference and is minimum when B1 and B2 have the same bias, as then we cannot cancel sample variance. $\mathrm{FoM}_{\gamma\omega}$ shows the biggest gains (factor of 4 and 3 at $\alpha \simeq 0.1$ and $\alpha \simeq 2$) as it combines the gains from FoM_{γ} (factor of 2) and FoM_{ω} (factor of 2). Part of the gain comes from B1xB2 doubling the density of the single tracers, as we have merged both

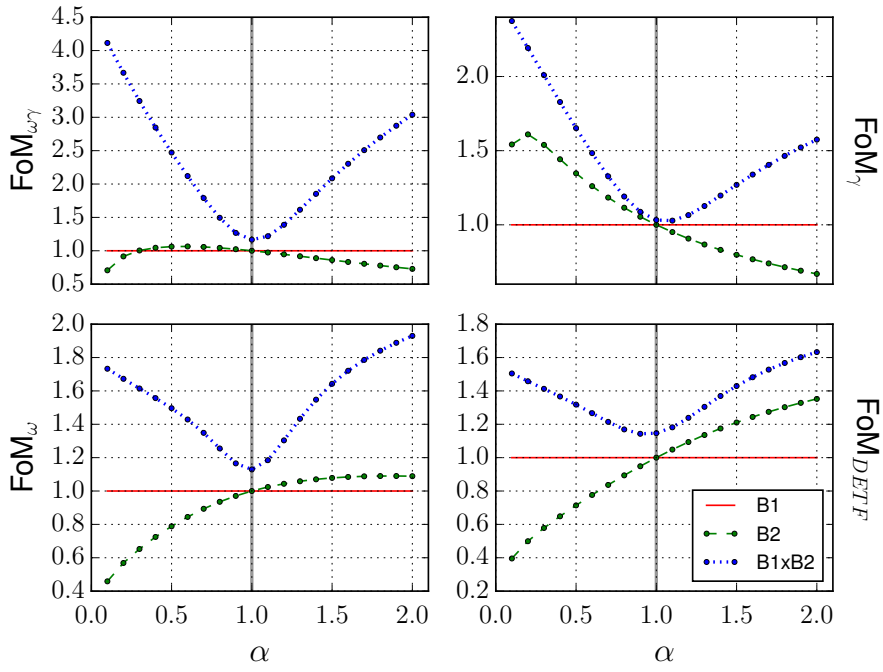

Figure 2.11: FoM dependence on the relative bias amplitude (α) in redshift space. The blue (dotted) line correspond to the same sky case (B1xB2), the green (dashed) line to B2 and the red (solid) line to B1. All lines are normalized to the B1 FoM value (which is constant). The vertical lines indicate the same bias amplitude case ($\alpha = 1$).

single tracers and it reduces shot-noise. To see the effect of splitting B1 into B1xB2 to keep the shot-noise level see Fig. 2.2.

The B2 lines show the impact that bias amplitude has on the single survey case. FoM$_\gamma$ shows that γ constraints improve for lower bias. This is in the line of Asorey, Crocce & Gaztañaga (2014), where the authors find similar results for the dependence of γ constraints on bias for a photometric survey. This is because lower bias gives larger relative importance to RSD, which turns into a better measurement of $f(z)$ and thus γ. On the contrary, ω constraints improve with larger bias, as shown in FoM$_\omega$ and FoM$_{\rm DETF}$. This is due larger bias increasing the signal of correlations and thus reducing the relative impact of shot noise (see Fig. 2.15 bottom right panel for similar trends).

2.A.2 Real space

In Fig. 2.12 we have removed redshift space distortions (RSD). It includes a purple (dashed, flat) line which shows the ratio RSD/No RSD for B1. Looking at the B1xB2 line we see how all FoMs now grow with bias ratio,

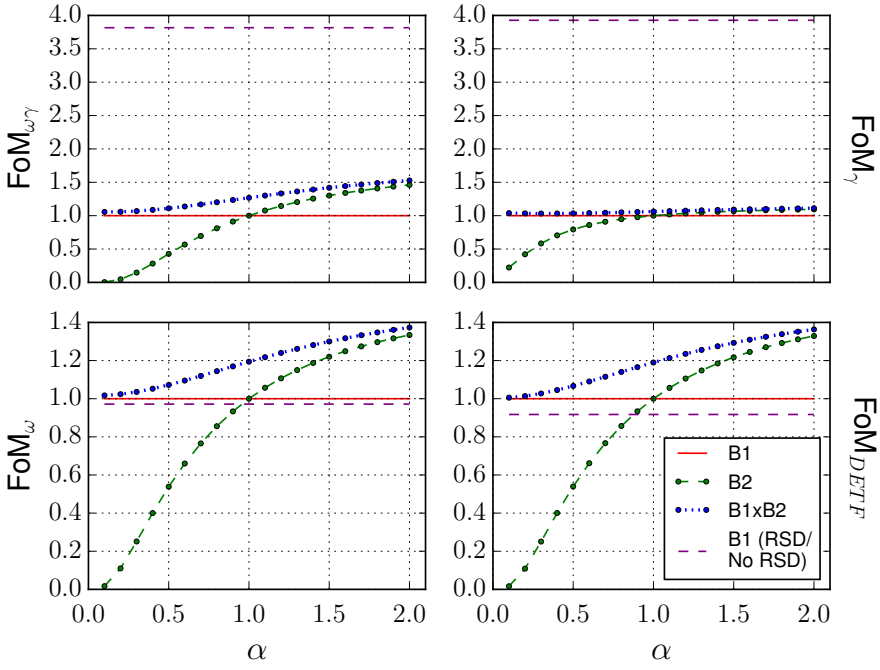

Figure 2.12: FoM dependence on the relative bias amplitude (α) in real space. The blue (dotted) line correspond to the same sky case (B1xB2), the green (dashed with dots) line to B2 and the red (solid) line to B1. All lines are normalized to the B1 FoM value (which is constant). The purple (dashed) line shows the ratio RSD/No RSD for the fiducial bias (B1).

while the characteristic minimum and big improvements from sample variance cancellation have vanished (factor of 1.5 for $\mathrm{FoM}_{\gamma\omega}$ now, as opposed to 4 with RSD).

In real space the density perturbation equation is Eq.2.1, which does not have specific angular dependence and does not allow to cancel sampling variance anymore. The auto correlations for a redshift bin i will then be proportional to b_i^2, while the cross correlations between a bin i and a bin j is proportional to $b_i b_j$. Therefore, signal will increase as bias does. Moreover,

$$\begin{aligned} &\text{when } \alpha \to 0 \text{ then FoM[B1xB2]} \simeq \text{FoM[B1]}, \\ &\text{when } \alpha \gg 1 \text{ then FoM[B1xB2]} \simeq \text{FoM[B2]}. \end{aligned} \tag{2.42}$$

This is a result of B1 correlations dominating over B2 ones at $\alpha \simeq 0$, and viceversa for $\alpha \gg 1$. Also, the B2 FoMs have a steeper slope than B1xB2, as we are only changing the B2 bias, which only affects a subset of correlations on B1xB2. While FoM_ω improves up to a 40% at $\alpha = 2$, the gain in FoM_γ is marginal, and part of it comes from B1xB2 having more density and reducing shot-noise.

The purple (dashed flat) line shows the ratio between the normalizations when computing the correlations in redshift space and in real space. This shows the overall impact of removing RSD in each FoM. As can be seen in the top-right panel, FoM_γ is about 4 times lower, as RSD are crucial for breaking degeneracies between $f(z)$ and $b(z)$, which affect the γ constraints. For the ω constraints, the normalizations are similar. However, as we have discussed, including RSD is still very important as it allows to cancel sampling variance and improve the ω constraints.

2.A.3 Relative impact of RSD and BAO

In Fig. 2.13 we explore the effect that removing RSD and BAO has on the absolute value of the FoMs, for different relative bias amplitudes (α). We show the ratios 'RSD/No RSD' (left panels) and 'BAO/No BAO' (right panels), where 'RSD' and 'BAO' refers to the fiducial forecast that includes both RSD and BAO effects. The left/right panels correspond to removing RSD/BAO for B1xB2 (blue solid), B2 (green dashed) and B1 (red dotted).

The trends in FoM_γ (center panels) show how RSD has a strong relative importance in γ constraints, while BAO in comparison has very little impact. On the other hand, FoM_ω (bottom panels) shows how BAO is more important to measure ω than RSD. This happens because measuring γ depends more on determining the amplitude of the power spectrum $P(k)$, which is enhanced

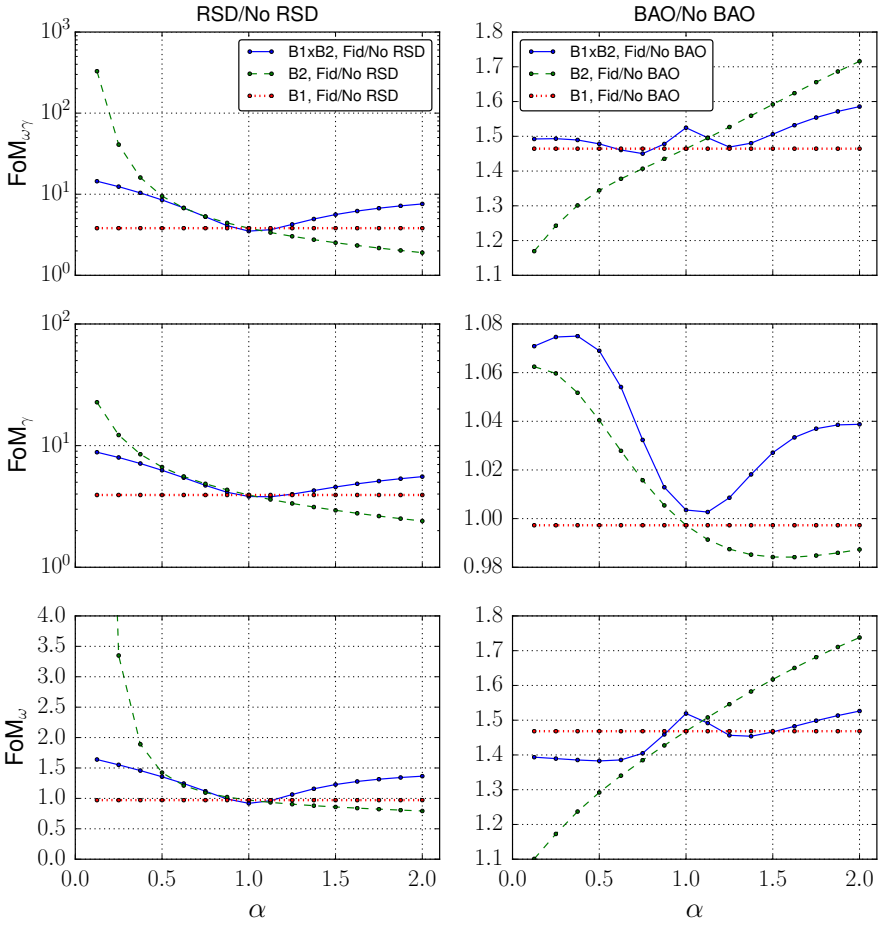

Figure 2.13: Impact of removing BAO and RSD effects on the constraints, for different relative bias amplitudes (α). The left/right panels correspond to removing the RSD/BAO effect for B1xB2 (blue solid), B2 (green dashed) and B1 (red dotted). All the lines show the ratio between the Fiducial forecast where RSD and BAO effects are included and when removing one of the effects. There is no further normalization.

by RSD as it breaks degeneracies between bias and the growth factor, while $\omega(z)$ measurements comes more from measuring its shape and thus including BAO wiggles measurements improves it. In addition, for FoM_ω, B1xB2 (solid) shows a peak at $\alpha = 1$ in BAO/No BAO ratio, because the BAO contribution is relatively higher when RSD is less important. On the other hand, for FoM_γ the BAO contribution in B1xB2 gets enhanced with a better RSD signal. The impact of RSD in B1xB2 is higher for more bias difference in all FoMs, as previously discussed (see §2.A.1), due to sample variance cancellations.

For a single tracer (see B2 in Fig 2.13), the impact of RSD increases very fast at low bias (decreasing α) in all FoMs, because RSD include an additional term such that the signal of the correlations does not drop when bias tend to zero (as happens when we do not have RSD, see B2 in Fig. 2.12). On the contrary, in FoM_ω the impact of BAO improves with more bias because shot-noise has less effect, while in FoM_γ is the opposite and for $\alpha \geq 1$ it has a negative effect, although the effect is tiny. $\text{FoM}_{\gamma\omega}$ (top panels) combines the effects from γ and ω constraints, and shows that RSD has a bigger effect in the constraints than BAO.

2.A.4 Fixing cosmological parameters

In this subsection we investigate how fixing one cosmological parameter affects the forecast by breaking degeneracies. Fig. 2.14 shows the relative gain from fixing each parameter for B1xB2 (left panels) and B1 (right panels), for $\text{FoM}_{\gamma\omega}$ and FoM_γ on the rows, as a function of survey density. For $\text{FoM}_{\gamma\omega}$ in B1xB2 there are strong gains (a factor 2∼3) when fixing Ω_m, Ω_{DE}, Ω_b or h, which constrains the cosmic expansion history, while when fixing σ_8 and n_s $\text{FoM}_{\gamma\omega}$ only improves by a factor 1.2∼1.3. Comparing to the single tracer, in $\text{FoM}_{\gamma\omega}$ multiple tracers help breaking degeneracies for Ω_m, Ω_b, Ω_{DE} and h parameters, specially at the noiseless limit where the relative gains lower from a factor $3.5 \sim 4$ in B1 to a factor 2 in B1xB2. For B1xB2 in FoM_γ the gains are quite low ($< 10\%$) when fixing any parameter, with σ_8 and n_s the most relevant at low density, whereas at high density the gain from σ_8 drops and Ω_m, Ω_{DE} and h become more important. For B1 we find a similar level of degeneracy in FoM_γ.

2.B Bias evolution

In this section we study the impact that different bias evolution has on the constraints. For concreteness, we parametrize the bias evolution in redshift as

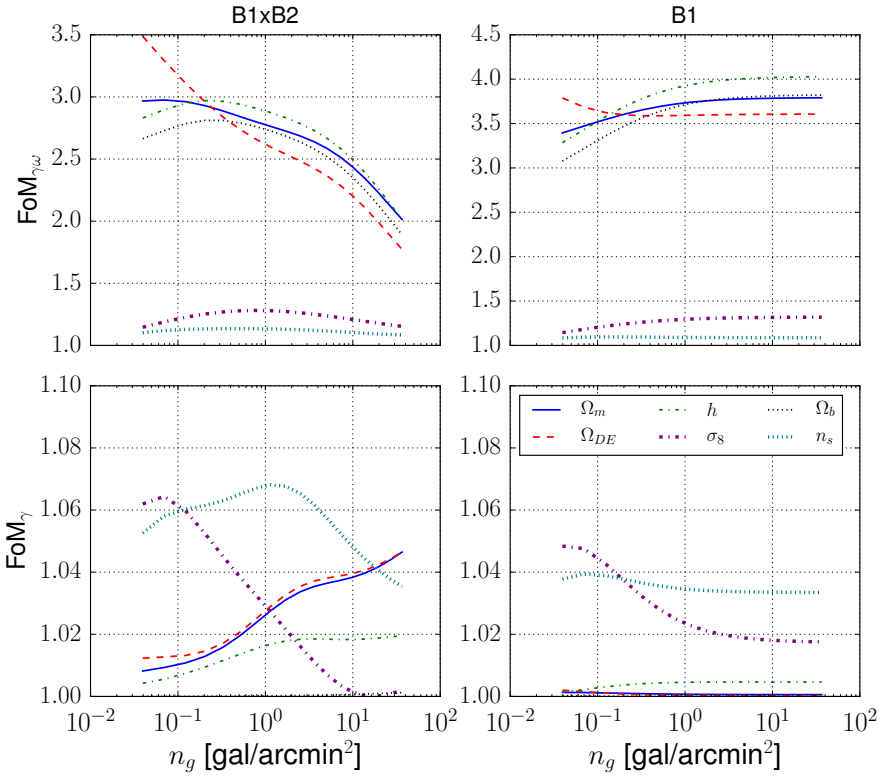

Figure 2.14: Relative gains when fixing a parameter for B1xB2 (left panels) and B1 (right panels), for FoM$_{\gamma\omega}$ and FoM$_{\gamma}$ on the rows, as a function of survey density. All lines are normalized to the respective fiducial forecast values.

$$b_{B1}(z) = 1 + \kappa z, \tag{2.43}$$

where $\kappa = 2$ corresponds to the fiducial forecast value, while we keep $\alpha = 0.5$. Fig. 2.15 shows the FoMs for B1xB2 (left panels) and B1 (right panels) in the columns and $\mathrm{FoM}_{\gamma\omega}$ (top), FoM_{γ} (center) and FoM_{ω} (bottom) for the rows, for different slope values, $\kappa \in [0.0, 0.5, 1.0, 1.5, 2.0]$ (red, green, blue, black, purple). $\mathrm{FoM}_{\mathrm{DETF}}$ is very similar to FoM_{ω} and therefore not shown.

Looking at the multitracer panels (B1xB2) (Fig. 2.15, left panels) it is clear that the constraints are not much affected by the bias evolution history. This shows that for the multi tracer it is much more relevant the relative bias amplitude between the populations rather than the bias evolution in redshift or the bias amplitude itself. Only at low densities we see some gains for higher bias, coming from the ω constraints, while the γ constraints remain very similar for the different bias evolutions at all densities. In other words, the constraining power of B1xB2 does not rely on how biased are the samples themselves, but on the contrast between its subsamples.

For the single tracer (B1) there is instead a clear dependence on the bias evolution in redshift (right panels). In FoM_{γ} the constraints are much more affected by the different bias slopes at the noiseless limit, leading to better constraints for lower slope (and thus lower bias), as then RSD has more relative importance. For higher shot-noise the relative gains from RSD are lower as shot-noise dominates and the lines converge. The opposite happens in FoM_{ω}, the dependence on bias slope is clear when shot-noise is big, while at the noiseless limit the constraints flatten and tend to the same value. As we have previosly remarked, dark energy constraints benefit from higher bias as this increases signal and reduces the relative impact of noise. This is observed clearly in this figures, as the FoMs benefit from higher bias at low densities, while with high densities they tend to the same value as the relative impact of noise is already small. $\mathrm{FoM}_{\gamma\omega}$ combines these effects and we see how all lines cross with each other, leading to better constraints for higher bias at low densities and viceversa at high densities. This results agree with Fig. 2.11 and Fig. 2.2. In addition, a complete unbiased tracer (i.e. $\kappa = 0$) has the best constraints at the noiseless limit for all FoMs.

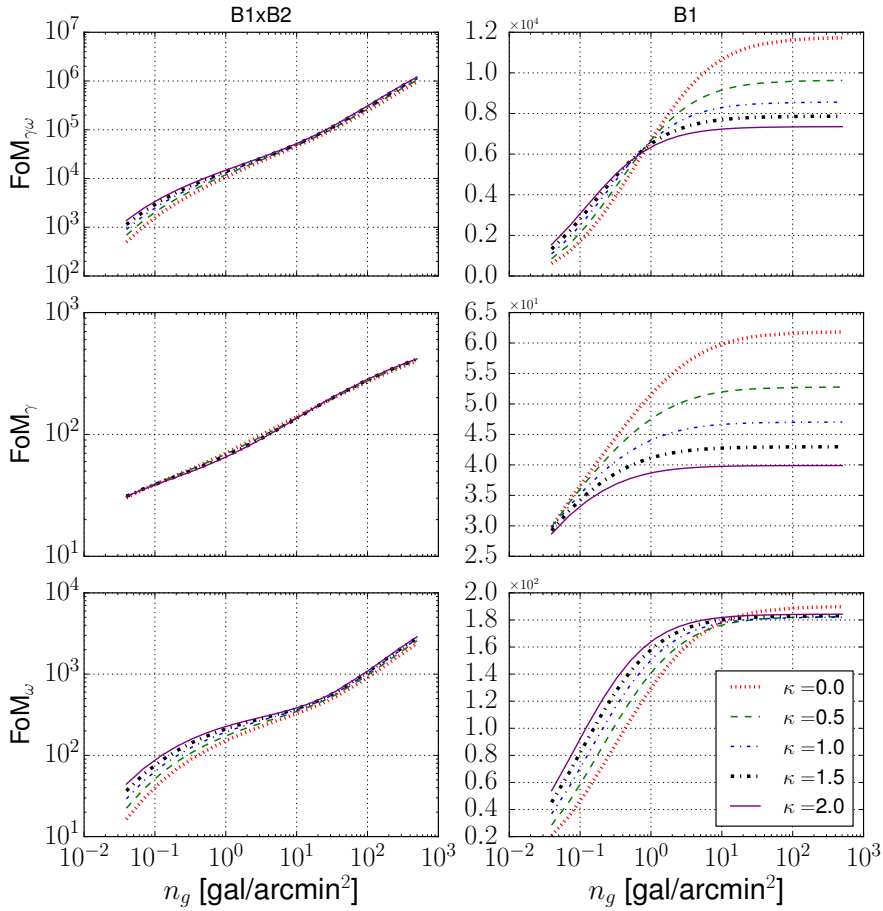

Figure 2.15: FoM dependence on the bias slope κ ($b_{B1}(z) = 1 + \kappa z$), with $\alpha = 0.5$ fixed. The panels on the left (right) shows the FoMs for B1xB2 (B1). On the rows we show FoM$_{\gamma\omega}$ (top), FoM$_{\gamma}$ (center) and FoM$_{\omega}$ (bottom), with FoM$_{\rm DETF}$ very similar to FoM$_{\omega}$ and therefore not included. The lines correspond to $\kappa = 0.0$ (red dotted), $\kappa = 0.5$ (green dashed), $\kappa = 1.0$ (blue dash-dot), $\kappa = 1.5$ (black thicker dash-dot), $\kappa = 2.0$ (purple solid).

Part III

Photometric redshift estimation with narrow bands in PAU Survey

In this part we estimate the redshift of galaxies using narrow band fluxes observed with PAU Survey. We create simulations with narrow band fluxes and add several systematic effects observed in real data. New statistical tools are develop which model the SED as a combination of emission line and continuum fluxes using libraries of synthetic templates. We test the methodology in simulations and finally compare the estimated redshift in real data with spectroscopic measurements.

Chapter 3

Template based photometric redshift estimation with narrow bands from PAU Survey

3.1 Introduction

Imaging galaxy surveys yield filtered images of the sky using a set of band-pass filters. This allows to cover larger areas with greater depth sacrificing spectral resolution, which results in a poorer line of sight, and redshift, resolution. Sharp spectral features accessible in spectroscopic surveys get smoothed out in the filtered bands making their location in the spectra much harder. Instead, photometric redshift focus on the colors of the galaxy, comparing the fluxes of the galaxy accross the bands. PAU Survey lies in between with a set of 40 narrow band (NB) filters in the range $\lambda \in [4500, 8500]$Å in steps of 100Å, which allows for a high resolution photometric spectra measurement that contains both clear information of the sharp features and a calibrated continuum.

Many photometric redshift estimation techniques have been developed over the recent years, but they can be broadly classified between template based and machine learning techniques. Machine learning codes like SKYNET (Bonnett 2015), ANNz2 (Sadeh, Abdalla & Lahav 2016) or DNF (De Vicente, Sánchez & Sevilla-Noarbe 2016) try to learn the relation between the colors and redshift of galaxies through a calibration sample where both colors and spectroscopic redshifts are available. Template based codes like BPZ (Benítez 2000) and LePHARE (Arnouts & Ilbert 2011) predict the flux of each band as a function of redshift by redshifting a library of restframe template SEDs and comparing to the observation.

Machine learning codes suffer from the incompleteness of the calibration

spectroscopic samples, which are shallower and cover small areas. These selection effects and the sample variance from a tiny area can generate biases in the final redshift estimate. On the other hand, template based codes rely on libraries of templates. Templates can be obtained from real galaxies through spectra, but the continuum can be hard to calibrate (Brown et al. 2014) and these also experience the spectroscopic survey selection effects. Many template based codes use synthetic templates instead, created with stellar population synthesis and evolution models, like Bruzual & Charlot (2003); Polletta et al. (2007). Using different stellar parameters such as the initial mass function (IMF), the star formation rate (SFR), metalicity, age or dust one can generate a wide variety of restframe galaxy spectra. However, the physics of the formation and evolution of galaxies are still not perfectly known, and if particular galaxy populations or properties are missing in the library it can lead to a redshift biased estimate. Additionally, template based methods require a zero point offset calibration of each band with respect to the templates, while machine learning self calibrate any possible zero point remaining offset in the training step.

A narrow band survey like PAUS is able to resolve emission lines to a much better accuracy than typical broad band surveys, making the modeling of emission lines a key aspect. These can be already included in the templates like in BPZ or be added following empirical recipes like LePHARE. However, none of these codes present a flexible enough way of incorporating them. In particular, each galaxy has a different amount of flux coming from the emission lines compared to its continuum, and while in a broad band this would represent a small amount of the total flux, in a narrow band this is often the main contribution. For these codes one would need to generate a large grid of templates with a variable relative amount of emission line flux and evaluate for each galaxy and (redshift, template, emision line) step the likelihood, which would largely increase the running time. Instead, we have developed a template based photo-z code, BCNZ2 (Eriksen et al. 2019), which uses a nonnegative linear interpolation of templates to fit the observations. In this way, templates can adapt to each galaxy and find the best combination of continuum and emisson line flux.

Template codes like BPZ and LePHARE have a list of baseline templates, but to improve the sampling of the color space they generate extra linearly interpolated templates between adjacent SEDs in color space (Coe et al. 2006; Ilbert et al. 2009). This gets naturally included in our nonnegative interpolation scheme by including more than one continuum template. Additionally, one can also model the SED of galaxies with very variable bulge to disk colors, for example galaxies that have older stellar populations in the center and

star forming regions on the outskirt disk, or galaxies hosting an active galactic nuclei (AGN) in the center.

The amount of information, i.e. the number of flux observations for each galaxy, is limited. If we feed the linear interpolation fitting scheme with too many models, by for example allowing many continuum SEDs at the same time and modeling each emission line individually, one can easily overfit and degrade the photo-z performance. A Bayesian way of measuring if extra models are required to explain the data is the Bayesian evidence. Computing the Bayes factor (see Eq. 1.27), the ratio between the Bayesian evidence of two sets of models, we can decide, for example, if emission lines or AGN templates need to be added in a particular galaxy SED. It also properly weights each theory or models if we want to marginalize over them to obtain the redshift estimation.

To compare between different SED template model choices and statistical frameworks, we have develop a narrow band galaxy simulated catalog based on simulated galaxies from the public MICE2 simulations (Carretero et al. 2015; Fosalba et al. 2015a; Crocce et al. 2015; Fosalba et al. 2015b). We add several effects from the data, including measured photometric noise and anticipated systematic biases, to try to mimic the data from one of PAUS calibration fields, COSMOS (see Fig. 1.6). We will use the simulations to compare to the expected photo-z performance in an ideal situation where the synthetic models are complete and truly representative of the data, since they have been used to create the observed simulated galaxies.

3.2 Data

In this section we describe the data products we are going to use for the redshift determination and validation.

3.2.1 PAUCam data

PAUS collaboration is building an optical 40 narrow band survey using PAUCam (Padilla et al. 2019) at the William Herschel Telescope (WHT). PAUS has been actively collecting data since 2015 in a set of different runs imaging 5 different fields: COSMOS and the CFHT W1, W2, W3 and W4 fields (see Fig. 1.6). The CFHT fields are larger and represent the main survey, which wishes to obtain a total of 100 $\deg^2$ in four fields, while the COSMOS field ($\sim 1\ \deg^2$) has been targeted as a calibration field since many photometric observations exist, ranging from ultraviolet all the way to far infrared, as well as spectroscopic surveys with relatively high completeness and depth. In this

thesis we will focus on data collected in the COSMOS field between 2015 and 2017.

At the end of each observing night, the data taken at WHT is directly sent to Port d'Informació Científica (PIC) for its storage, orchestration and processing (Tonello et al. 2019). The data reduction process starts with initial de-trending, where a numer of signatures from the instrument are removed from the images, using the NIGHTLY pipeline (see Serrano In prep.; Castander In prep. for details). This includes removing electronic bias with an overscan subtraction, correcting the gain from the different amplifiers and compensation from readout patterns using zero exposure bias frames. A master flat is created from 10s exposures of the dome with homogeneous illumination that are taken every afternoon before the observation, that is then divided on the images to remove the vignetting of the telescope corrector, among other effects such as dead and hot pixels. Each CCD has a different narrow band filter on top, instead of a unique broad band that covers all the focal plane, which produces scattered light at the edges of the images. An intervention of the camera in 2016 significantly reduced this effect, which is partly removed by the pipeline by using a low pass filtering with sigma clipping. Cosmic rays are identified using a Laplacian edge detection (van Dokkum 2001) and masked from the image.

An astrometric solution is added to align the different exposures using SCAMP (Bertin 2011) by comparing to GAIA DR1 (Gaia Collaboration et al. 2016). The point spread function (PSF) is modeled using PSFEX (Bertin 2011) using stars that have been identified with morphology from space observations with the COSMOS Advanced Camera for Surveys (ACS) (Leauthaud et al. 2007; Koekemoer et al. 2007). Photometric calibration follows using Sloan Digital Sky Survey (SDSS) stars that have been previously calibrated (Castander In prep.). Each star brighter than $i_{\rm AB} < 21$ is fitted to obtain a best fit SED from the Pickles stellar library (Pickles 1998) using the u,g,r,i and z broad bands (Smith et al. 2002). We generate synthetic narrow band observations from that best SED and the best fit amplitude and compare to the observations to obtain the zero point. All measurements are combined to obtain one zero point per image and narrow band. This procedure effectively corrects for Milky Way extinction.

Narrow band photometry is obtained using the MEMBA pipeline (Serrano In prep.; Gaztanaga In prep.). We rely on deep overlapping observations from lensing surveys to provide a detection catalog with high quality shape measurements to perform a forced aperture photometry. In COSMOS positions and shape measurements from ACS were used. The half light radius, r_{50}, is used along with ellipticity measurements from Sargent et al. (2007) and the PSF

FWHM to determine the aperture size and shape to target for the 62.5% of the light. A flux measurement is obtained for each individual exposure using this aperture measurement and a background subtraction estimated from a fixed annulus of 30 to 45 pixels around the source, where sources falling in the annulus get sigma clipped. Fluxes measured in different exposures are corrected with the estimated image zeropoints and get combined with a weighted average to produce a narrow band coadded flux measurement. The data reduction pipeline propagates flags for each individual exposure and object, and flagged measurements are not included in the weighted average.

3.2.2 External COSMOS data

Along with the narrow band data described in the previous section, we add broad band data from the COSMOS2015 catalogue. We includes u^* band data from the Canada-France Hawaii Telescope (MegaCam) and B, V, g, r, i^+, z^{++} broad band data from Subaru, obtained as part of the COSMOS20 survey (Taniguchi et al. 2015). Figure 3.1 shows the broad band transmission curves. We use the 3"" diameter PSF homogenised flux measurements available in the catalogue release and apply several corrections as described and provided in COSMOS2015. The Milky Way interstellar dust reddens the observed spectrum of background galaxies. As described in the previous subsection, PAUS data are corrected for dust extinction in the calibration. Therefore we need to do the same for COSMOS data. Each galaxy has an $E(B-V)$ value from a dust map (Schlegel, Finkbeiner & Davis 1998), and Laigle et al. (2016) provide an effective factor F_x for each filter x according to Allen (1976). For each galaxy the corrected magnitudes are

$$\text{Magcorrected} = \text{Maguncorrected} - E(B-V) * F_x \tag{3.1}$$

3.2.3 Spectroscopic data

To determine the accuracy of the photometric redshift estimation using PAUS, we compare to zCOSMOS DR3 bright spectroscopic data, which has a pure magnitude selection in the range $15 < i_{AB} < 22.5$ (Lilly et al. 2007). This selection yields a sample mainly covering the redshift range $0.1 \lesssim z \lesssim 1.2$ in 1.7 deg^2 of the COSMOS field ($149.47\,\text{deg} \lesssim \alpha \lesssim 150.77\,\text{deg}$,$1.62\,\text{deg} \lesssim \delta \lesssim 2.83\,\text{deg}$, Knobel et al. 2012). This dataset contains 16885 objects of which 10801 remain after removing less reliable redshifts based on a provided confidence class ($3 \leq \text{CLASS} < 5$ Lilly et al. (2007)). This sample covers most of the redshift and magnitude range for PAUS, which makes it especially

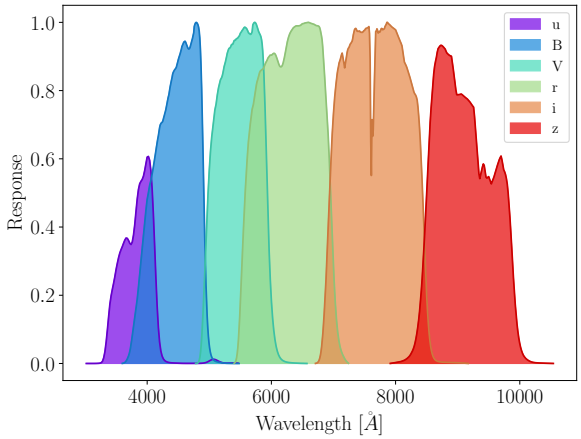

Figure 3.1: Filter curves for u^* band data from the Canada-France Hawaii Telescope (MegaCam) and B, V, g, r, i^+, z^{++} broad band data from Subaru.

interesting for validating the photometric redshift precision. The spectroscopic completeness is shown in Figure 3.2 for the full sample and two subsamples that force a moderately and a highly secure redshift sample by selecting on provided spectroscopic confidence.

3.3 Modeling

3.3.1 Redshift distribution

We want to obtain the redshift probability distribution of a galaxy given some feature vector, typically fluxes, has been observed: $P(z|f)$. We introduce a set of template models that describe the spectral energy distribution (SED) of the galaxy at rest frame. By redshifting the SED and convolving with a filter band we can predict what the observed fluxes would look like at that redshift and ask what is the probability our data could be observed if that was the true SED t and redshift z: $P(f|z,t)$. We can relate both using Bayes rules (see section 1.3),

$$P(z|f) = \sum_t P(z,t|f) = \sum_t P(f|z,t)P(z,t) \tag{3.2}$$

where $P(z,t)$ is the prior probability that SED and redshift could be observed.

As mentioned in section 3.1, emission lines are variable from galaxy to galaxy, both in absolute amplitude with respect to the continuum flux, but also the relative flux of each emission line. Therefore, a flexible enough model

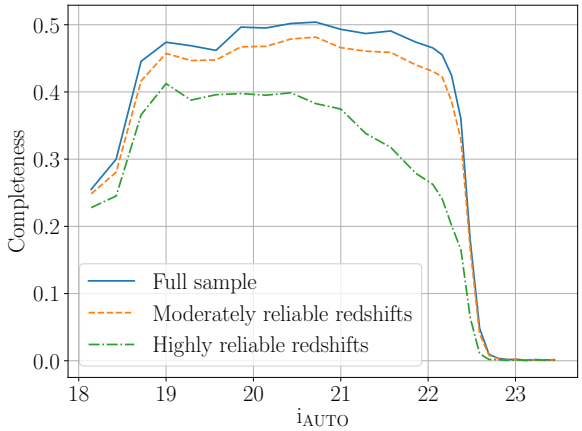

Figure 3.2: Spectroscopic completenes of the zCOSMOS DR3 spectroscopic survey for the full sample (solid line), a moderately secure redshift sample (dashed line) defined by the classes: 3.x, 4.x, 2.5, 2.4, 1.5, 9.5, 9.4, 9.3, 18.5, 18.3 and a highly secure redshift sample (dash-dotted) defined only by the classes 3 and 4.

is required. To this end, we treat the continuum and emission line part of the SED as separate models that can be linearly added. Rather than having one fixed SED we allow for the best combination of continuum and emission line models at each redshift, by defining our model T at a given redshift as

$$T(z) = \sum_{j}^{n} \alpha_j(z)\, t_j(z); \tag{3.3}$$

where α are the parameters that combine each SED t. Then, the redshift distribution reads

$$P(z|f) = \sum_{T \in H} P(z, T|f) = \sum_{T \in H} P(f|z, T) P(z, T) \tag{3.4}$$

where $T \in H$ means the model T defined as in Eq. 3.3 is among those hypothesis H to be considered. A common approach to obtain the likelihood $P(f|z, T)$ is to fit for the maximum likelihood values of the parameters, in this case α_j,

$$P(f|z, T) \propto \exp\left(-0.5\chi^2[z, \alpha_{\min}]\right) \tag{3.5}$$

where we assume Gaussian noise in the fluxes and $\alpha_{\min}$ are the best fit parameters of model T that minimize the χ^2 at redshift z. Negative amplitudes are regarded as non physical and we forbid them, so we restrict to $\alpha_j \geq 0$.

The minimum of Eq. 3.5 would have an analytical solution without the non-negativity constraint, instead we have implemented within BCNZ2 a non-negative quadratic programming algorithm based on Sha et al. (2007) that will iteratively find the minimum. The algorithm is described in section 3.A.1.

A more statistically correct way to estimate the probability from Eq. 3.5 is to compute the evidence, marginalizing over the parameters. Following Eq. 1.24 we have

$$P(f|z,T) = \int P(f,\vec{\alpha}|z,T)d\vec{\alpha} = \int P(f|\vec{\alpha},z,T)P(\vec{\alpha}|z,T)d\vec{\alpha}, \tag{3.6}$$

We can write the Gaussian likelihood $P(f|\vec{\alpha},z,T)$ as

$$P(f|\vec{\alpha},z,T) = \frac{1}{\sqrt{(2\pi)^d}\prod_k \sigma(f_k^{obs})} \exp\left[-\frac{1}{2}\sum_{k=1}^{d}\left(\frac{f_k^{obs}}{\sigma(f_k^{obs})} - \sum_{j=1}^{n_i}\alpha_j \frac{t_{jk}}{\sigma(f_k^{obs})}\right)^2\right] \tag{3.7}$$

where d is the number of bands, n_i is the number of templates in the ith model T from the hypothesis list. Integrating Eq. 3.7 also has an analytic solution but in Eq. 3.6 there is also a prior on the parameters, $P(\vec{\alpha}|z,T)$. This prior will include the non-negativity constraint, and can also include information from the templates, like luminosity functions. We have implemented a code that can efficiently find this integral, based on an Gaussian integral algorithm from Genz (1992). Details of the algorithm are explained in section 3.A.2. Our default choice for the parameters prior will be a top hat function, which can be considered the most 'uninformative',

$$P(\vec{\alpha}|z,T) = \frac{1}{\prod_j(\Delta_j^+ - \Delta_j^-)}\prod_k \Theta(\alpha_k - \Delta_k^-)\Theta(\Delta_k^+ - \alpha_k), \tag{3.8}$$

where $\Theta(x)$ is the Heaviside step function and $(\Delta_k^+ - \Delta_k^-)$ the width of the prior for each parameter α_k. The normalization $\frac{1}{\prod_j(\Delta_j^+ - \Delta_j^-)}$ is responsible for the Bayesian Occam's Razor effect (see section 1.3 and Fig. 1.3).

Fig. 3.3 shows an illustration of the Bayesian evidence for a model T with three templates: a Spiral, a Starburst and an Emission Line template. Each contour shows the 1 and 2 sigma contour of the likelihood (Eq. 3.7) for different redshift (blue: low redshift, red: high redshift) for a simulated galaxy with true redshift around the green area. We want to choose a prior that avoids

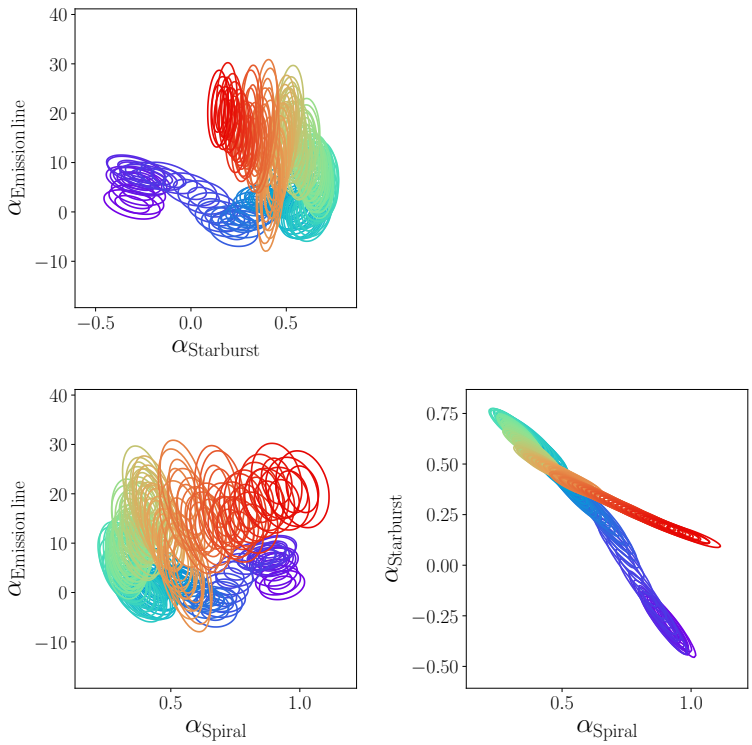

Figure 3.3: Illustration of the Bayesian evidence algorithm. It shows the likelihood (Eq. 3.7) for a simulated galaxy and a model with three templates: a Spiral, a Starburst and an Emission Line template. The galaxy has the true parameter values of $[\alpha_{\rm Sp}, \alpha_{\rm SB}, \alpha_{\rm EmLine}] = 0.5, 0.5, 8$ and true redshift $z = 0.5$ (green-ish area contours). The contours show the 1 and 2 sigma of the likelihood for different redshifts (blue: $z = 0$, red:$z = 1.2$. Note the height of the likelihood (i.e. how well did it fit) is not shown in this plot.

unphysical regions with negative amplitudes, but otherwise we want to include all the likelihood. To achieve that we choose a prior with $\Delta_k^{\pm} = \mu_k \pm 5\sigma_k$, where μ_k and σ_k are the central value and width of the likelihood for parameter α_k, which can be obtained analytically. We clip Δ_k^- to have a minimum value of zero. When Δ_k^+ is found to be zero for some parameter we assign it a value of 10. It is very likely the evidence for that model will be very low.

3.3.2 SED modeling

In this section we describe the pieces that build our SED model. For the continuum flux we choose an SED library that combines synthetic templates from Bruzual & Charlot (2003) and Polletta et al. (2007) libraries similar to Laigle et al. (2016), which include a wide range of elliptical, spiral and starburst

galaxies, with different stellar histories, ages and metallicities. Fig. 3.4 shows a subsample of SEDs from this libraries.

Galaxies have a variable amount internal extinction due to light scattering with the dust inside the galaxy. Since the continuum SED libraries are dust free we model this wavelength dependent effect with an extinction curve $k(\lambda)$ and a color excess $E(B-V)$ that adds this effect to the template,

$$F_{\text{observed}}(\lambda) = F_{\text{no dust}}(\lambda) 10^{-0.4\,E(B-V)\,k(\lambda)}. \tag{3.9}$$

Since this is a multiplicative effect, and the formalism described in the previous section only allows linear combinations of templates, we include this effect by modifying our restframe templates and generating new ones with different amount of dust attenuation. For the starburst galaxies we will use the extinction law from Calzetti et al. (2000), while for redder galaxies like spirals we will model using the Prevot et al. (1984) law. We will not add further extinction to the reddest galaxy templates, like ellipticals. Fig 3.9 shows both extinction curves, with an additional bump around 2175Å with two different amplitudes for the Calzetti law. We generate a grid of templates with different $E(B-V)$ values ranging from 0.05 to 0.5 in steps of 0.05.

Like mentioned previously, having a flexible model that can estimate a different amplitude for the emission line flux in each galaxy is necessary to achieve a good photo-z. We treat the emission line flux as a separate part of the SED, and we compute the flux by modeling the emission lines with a delta function centered around the values from Table 3.1. Since modeling all the emission lines independently would overfit we choose by default to tie each line amplitude to the [OII] line flux with the ratios appearing in Table 3.1, which are obtained from Ilbert et al. (2009). We have found in the data that leaving the [OIII] doublet free as an extra template improves the photo-z in data, so we also consider this setup.

When the likelihood is obtained from the minimization algorithm, we will split our large pool of templates in smaller groups, which will be our models T in Eq. 3.5, to avoid overfitting. This choice has an implicit prior that other choices are not reasonable and we assume they have a zero prior probability (see section 1.3). Also we do not allow in the same model templates with different extinction values or curves. Table 3.2 summarizes the list of models we include in the minimization algorithm. Note we do not allow emission lines in the reddest elliptical templates, since this are not expected a priori.

In the Bayesian evidence algorithm we will let the evidence guide us on which combinations have a larger evidence. Thanks to its preference for simpler models (Bayesian Occam's Razor) models with a lot of templates will

Line	λ[Å]	Ratio
OII	3726.8	1.0
$OIII_1$	4959	0.09
$OIII_2$	5007	0.27
H_α	6562.8	1.77
H_β	4861	0.61
NII_1	6548	0.19
NII_2	6583	0.62
SII_1	6716.44	0.35
SII_2	6730.82	0.35

Table 3.1: List of emission lines included in the SED modeling. The lines are modeled with a delta function centered around the values shown in the second column. In the third column we show the emission ratios of each line relative to the [OII] line emission that are used by default. Leaving the [OIII] doublet free from the other lines is also considered.

Model #	Lines	Ext Law	SED list
1	False	None	Ell1, Ell2, Ell3, Ell4, Ell5, Ell6
2	False	None	Ell6, Ell7, S0, Sa, Sb, Sc
3	True	None	Sc, Sd, Sdm, SB0, SB1, SB2
4	True	None	SB2, SB3, SB4, SB5, SB6, SB7, SB8, SB9,SB10, SB11
5-14	True	Prevot	Sc, Sd, Sdm, SB0, SB1, SB2, SB3
15-14	True	Calzetti	SB4, SB5, SB6, SB7, SB8, SB9, SB10, SB11
25-34	True	Calzetti+Bump 1	SB4, SB5, SB6, SB7, SB8, SB9, SB10, SB11
35-44	True	Calzetti+Bump 2	SB4, SB5, SB6, SB7, SB8, SB9, SB10, SB11

Table 3.2: List of models allowed in the minimization algorithm.

SED list	Lines options	Ext Law allowed
Ell1, Ell4	None	None
Ell4, Ell7	None	None
Ell7, Sc	None	None
Sc, SB0	{None, All, All+[OIII] free}	{None, Prevot}
SB0, SB4	{None, All, All+[OIII] free}	{None, Prevot}
SB4, SB8	{None, All, All+[OIII] free}	{None, Calzetti, Calzetti+Bump1, Calzetti+Bump2}
SB8, SB11	{None, All, All+[OIII] free}	{None, Calzetti, Calzetti+Bump1, Calzetti+Bump2}

Table 3.3: List of models allowed in the Bayesian evidence algorithm.

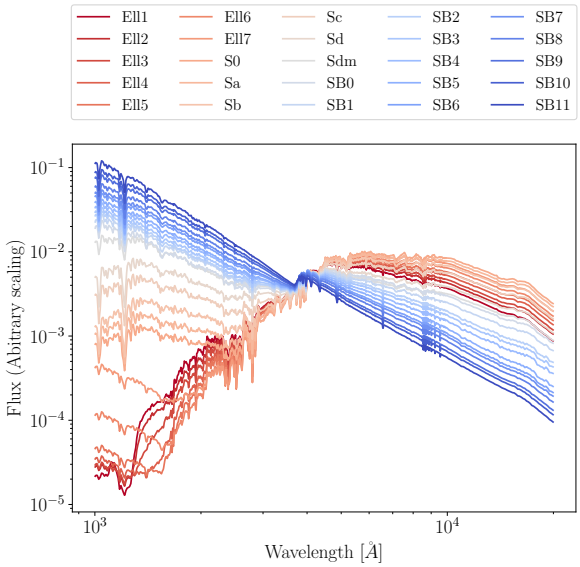

Figure 3.4: Synthetic set of template SEDs including several elliptical (Ell1...Ell7) and spiral (S0...Sdm) templates from Polletta et al. (2007) and starburst (SB0...SB11) from Bruzual & Charlot (2003).

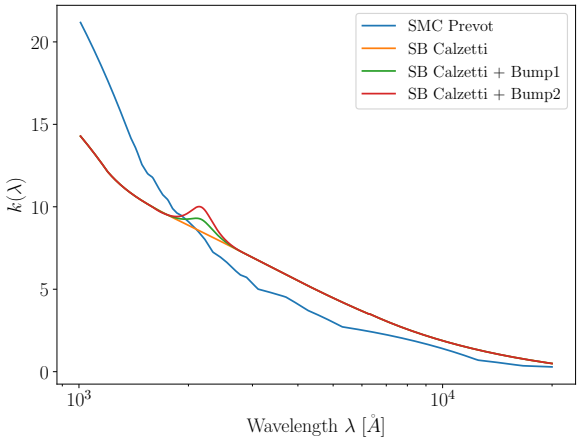

Figure 3.5: Internal extinction curves from Calzetti et al. (2000) and Prevot et al. (1984). The Calzetti are typically applied to starburst galaxies, while the Prevot are applied to spiral galaxies.

be assigned a low evidence, and finally a low probability. We group the templates in groups of 2 such that they keep the interpolation effect across the color-space, and include emission lines and extinction as in the minimization algorithm. Table 3.3 summarizes the list of models included. Note that for the same pair of templates we add a models with different emission line treatment (including no emission lines at all), and we will let the evidence weight them. We find that adding a third continuum template in between the pair of templates leads to very low evidence and contribute very little to the final redshift distribution, so we ignore such combinations.

3.3.3 Population prior

For the model to be complete, we also need to specify the prior on the global redshift distribution and the models T, $P(z, T)$ (see Eq. 3.4). This prior can include information about the population typical redshift and type distributions and include the expected impact of the selection effects that define the galaxy sample of interest (flux limited, color-color selection, etc). In a more general approach, the population prior can also be regarded as part of the redshift problem and be estimated jointly with the individual redshift distributions of each galaxy in a hierarchical bayesian model. We will either use a flat prior or a prior self informed from the data itself.

The population prior is also important to properly balance model choices. For example, in section 3.3.2 we are generating a grid of templates with different extinction $E(B-V)$ values ranging from 0.05 to 0.5 in steps of 0.05. If one increased the number of steps of $E(B-V)$ values within the same range where models are generated and kept using a flat population prior, one would overweight the templates with dust extinction in comparison to the templates without extinction. Another example is when we add a bump to the extinction law (Fig. 3.5), which is not observable at all redshifts in the filter set. To be robust against such model choices, we introduce a prior calibration scheme that compares the observed color space occupation of all models T to the galaxy sample occupation in the same observed color space. Since the color space occupation is fixed by the galaxy sample, if one adds more models with the same observable color space, the prior scheme will appropriately downweight them.

For the prior calibration scheme we will only use broad band colors. The prior calibration works as follows. We introduce a latent variable, the observable colors a galaxy can have $\vec{c}$,

$$P(z, T) = \int d\vec{c}\; P(z, T, \vec{c}) = \int d\vec{c}\; P(z, T|\vec{c})\, P(\vec{c}) \tag{3.10}$$

Instead of integrating over all this space, one can take an approximation by binning the color space and summing over those bins,

$$P(z,T) = \sum_j P(z,T,\vec{c}_j) = \sum_j P(z,T|\vec{c}_j)\, P(\vec{c}_j) \tag{3.11}$$

The term $P(\vec{c}_j)$ is the prior probability of observing the colors defined by $\vec{c}_j$ and can be directly estimated by the number counts of observed galaxies that fall in that bin. The term $P(z,T|\vec{c}_j)$ can be empirically estimated by drawing mock galaxies from each z,T with some rate $\mathcal{P}(z,T)$ and assigning them to a bin defined by $\vec{c}_j$ to estimate $P(\vec{c}_j|z,T)$. Then we can use Bayes theorem,

$$P(z,T|\vec{c}_j) = \frac{P(\vec{c}_j|z,T)\mathcal{P}(z,T)}{\sum_{z,T} P(\vec{c}_j|z,T)\mathcal{P}(z,T)}, \tag{3.12}$$

Note that the rate $\mathcal{P}(z,T)$ is the current knowledge we have about this population prior, which we update through a comparison of the color space occupation of the data and the models. In this work we will assume we know nothing about these models and assign a flat rate $\mathcal{P}(z,T)$. To draw galaxies and estimate $P(\vec{c}_j|z,T)$ we use

$$P(\vec{c}_j|z,T) = \int d\vec{\alpha}\; P(\vec{c}_j,\vec{\alpha}|z,T) = \int d\vec{\alpha}\; P(\vec{c}_j|\vec{\alpha},z,T)P(\vec{\alpha}|z,T) \tag{3.13}$$

which means drawing mock galaxies given model T at redshift z with parameters $\vec{\alpha}$ given some prior probabilities $P(\vec{\alpha}|z,T)$ and assigning them to a bin $\vec{c}_j$ to estimate $P(\vec{c}_j|\vec{\alpha},z,T)$. We will use a flat prior for $P(\vec{\alpha}|z,T)$, like in the data, by drawing numbers between $[0,1]$ (the normalization does not matter since we use colors). However, we discard models where any broad band flux contains more than a 20% of flux from emission lines, since we regard these as not physical.

To bin the color space and obtain $\vec{c}_j$ we will use a self organizing map. A self organizing map (or SOM) is an unsupervised machine learning technique which projects data from a high dimensional space into a lower dimensional grid. The process preserves the topology of the higher dimensional data, which means that objects that were close in the original space will also be closer in the lower dimensional grid. We will use the same method from Masters et al. (2015), where they use a two dimensional grid to generate the SOM.

3.3.4 Combining narrow bands and broad bands

As highlighted in section 3.2, we will add broad bands measurements from deep high quality lensing surveys. While PAUS data obtains the flux measurement from a variable aperture that targets 62.5% of the galaxy light, adapting to the galaxy apparent size and taking into account the PSF of the image, COSMOS broad band data has a fixed aperture of 3" on PSF homogenized images, which means they measure a different fraction of the light, but also that this fraction will be different for galaxies of different size. COSMOS additionally recalibrates their fluxes from aperture to total by averaging the aperture and total fluxes they obtain across several optical and near infrared bands, to try to minimize this effect. Hence, narrow bands and broad bands are in different photometric systems, and this could be variable form galaxy to galaxy. To deal with this effect, we have developed a self calibration algorithm that benefits from the overlap of the Subaru r-band with the narrow bands, with the idea that since they overlap they should yield compatible flux measurements. We introduce a synthetic Subaru r-band flux, defined as a linear combination of narrow band flux measurements.

Let $W_B(\lambda) = \lambda^{-1} T_B(\lambda)$ be a broad band filter that we want to build from a linear combination of 40 narrow band filter $W_N(i,\lambda) = \lambda^{-1} T_N(\lambda)$, where $T(\lambda)$ are the filter responses. We need to solve for coefficients $c(i)$ such that

$$W_B(\lambda) = \sum_{i=1}^{i=40} c(i)\ W_N(i,\lambda) \tag{3.14}$$

where λ is the wavelength which we will bin in integer values. We introduce the notation $< ... >$ which means inner product over wavelength:

$$\begin{aligned} < W_B W_B > \equiv 1 = \int d\lambda\ W_B^2(\lambda) = \sum_\lambda\ W_B^2(\lambda) \\ < W_N W_N > \equiv 1 = \int d\lambda\ W_N^2(\lambda) = \sum_\lambda\ W_N^2(\lambda) \end{aligned} \tag{3.15}$$

which just indicates the transmission curves norm are normalized to unity. We can then define 40 elements of broad-narrow projection vector:

$$BN(j) \equiv < W_B W_N(j) > = \int d\lambda\ W_B(\lambda) W_N(j,\lambda) \tag{3.16}$$

and a 40x40 narrow band overlap matrix

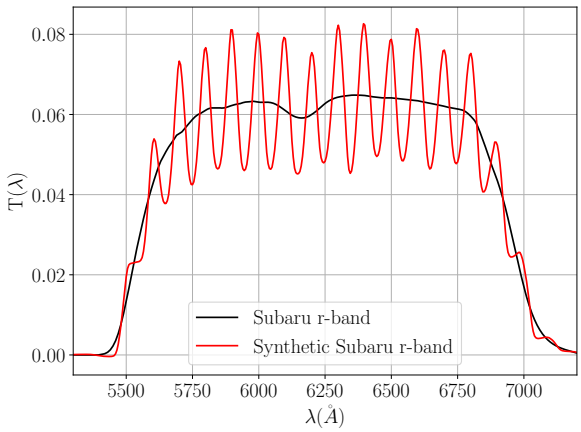

Figure 3.6: Subaru r-band transmission curve compared to a synthetic Subaru r-band built from PAU narrow bands. The overlapping region between narrow bands provides extra resolution (peaks) compared to the center of the filters. The synthetic filter oscillates around the true filter.

$$NN(i,j) \equiv < W_N(i) W_N(j) > = \int d\lambda \; W_N(i,\lambda) W_N(j,\lambda). \tag{3.17}$$

If we multiply Eq.3.14 by $W_N(j,\lambda)$ and then integrate we get

$$BN(j) = \sum_i \; c(i) \; NN(i,j). \tag{3.18}$$

As $NN(i,j)$ is invertible, the unique solution is

$$\begin{aligned} c(i) &= \sum_j NN^{-1}(i,j) BN(j) \\ &= \sum_j \int d\lambda \; W_B(\lambda) W_N(j,\lambda) \; \left[\int d\lambda W_N(i,\lambda) W_N(j,\lambda) \right]^{-1} \end{aligned} \tag{3.19}$$

We can use the $c(i)$ values from Eq. 3.19 to build a synthetic broad band filter from the 40 narrow bands. Fig. 3.6 compares the synthetic filter to the true one. We can also use the $c(i)$ values to compute the synthetic broad band flux for each object given the measured 40 narrow band fluxes.

$$f_{syn}^{BB} = \sum_i c(i) \; f_i^{NB} \tag{3.20}$$

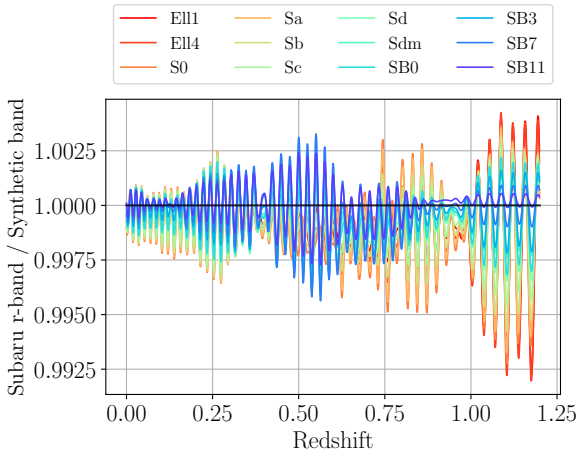

Figure 3.7: Ratio between the Subaru r-band flux and the synthetic Subaru r-band obtained from the narrow band fluxes (using $c(i)$ values from Eq. 3.19) given different SED models (color coded) at different redshifts (x-axis).

Comparing the synthetic broad band with the measured broad band we calibrate the distance between the narrow band and broad band photometric systems. For a given SED the synthetic broad band flux does not completely match the actual broad band flux but it is biased depending on the SED and redshift. This effect can be interpreted as a consequence of the difference between the recovered and true filter from Fig. 3.6. Fig. 3.7 shows this effect with the ratio between the synthetic broad band and the true broad band for different SEDs at different redshifts, showing some oscillations around 1, smaller than 1%. Similarly, Fig. 3.8 shows the flux predicted from the synthetic band and the flux from the Subaru r-band for the emission line model. When the emission line falls where two narrow bands overlap the synthetic r-band flux is larger than the true, and when it falls in the center of a narrow band, the synthetic flux is smaller than the true. Therefore, it is important to add this effect into the modeling by correcting the templates for each SED and redshift by this expected bias.

3.3.5 Zero point systematics

A common approach in the literature is to allow for a recalibration of relative zeropoints between different bands before running the photo-z algorithm (Benítez 2000; Coe et al. 2006; Hildebrandt et al. 2012a; Laigle et al. 2016). A zero point bias doesn't need to come from zero point estimation itself, but can also be caused by incorrect PSF modeling (Hildebrandt et al. 2012a). A way

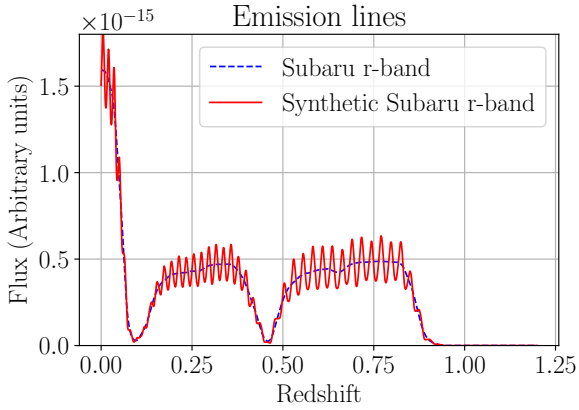

Figure 3.8: Similar to Fig. 3.7 for the Emission line model. It shows the flux contribution in Subaru r-band coming from the emission line model alone, compared to the one found in the synthetic r-band. The flux is in arbitrary units.

to calibrate this zero point shifts is to stack the residual differences between the best fit model and the observations at the true redshift of each galaxy for a subsample of galaxies with spectroscopic redshifts. Once the residual distribution is obtained for each band one can use different statistics to get an estimate of the zero point shift and apply it to the data, and recalibrate again iteratively until the sequence converges. The assumption is that the observations are Gaussian samples of each best model, so the residual distribution should be centered at 0.

One statistic is minimize the sum of the squared differences between the best model and the data divided by the error,

$$\chi^2(\kappa_j) = \sum_i \frac{(\kappa_j \tilde{f}_i - T_i)^2}{(\kappa_j \sigma(\tilde{f}_i))^2}, \qquad \kappa_j = \min\left[\chi^2(\kappa_j), \kappa_j\right] \tag{3.21}$$

where the sum is over the spectroscopic galaxies, T_i is the best model at the true redshift and κ_j is the zero point factor for band j. To avoid outliers biasing the minimum one can clip galaxies with a very bad fit, or sigma clip residuals larger than some value. An alternative statistic more robust to outliers is the median of the ratio between the model and observations

$$\kappa_j = \text{median}\left[\frac{T_i}{\tilde{f}_i}\right] \tag{3.22}$$

or the inverse of the median ratio between observations and model

$$\kappa_j = \left(\text{median} \left[\frac{\tilde{f}_i}{T_i} \right] \right)^{-1} \tag{3.23}$$

which has the advantage of not having a noisy quantity dividing. We will test the three zero point calibration statistics with simulations and then apply them to data.

3.4 Simulated narrow band catalog

We want to test the methodology described in section 3.3 with simulations where the conditions are ideal and the true models and redshifts are known. For this purpose, we have generated a narrow band mock galaxy catalog based on the MICE2 simulated galaxies. The MICE2 simulation is based on a dark matter N-body simulation and contains many galaxy and lensing properties that have been thoroughly validated in Carretero et al. (2015); Fosalba et al. (2015a); Crocce et al. (2015); Fosalba et al. (2015b). The MICE2 cosmology is defined by $\Omega_m = 0.25$, $\Omega_b = 0.044$, $h = 0.7$, $n_s = 0.95$, $\Omega_\Lambda = 0.75$, $\sigma_8 = 0.8$ and $w = -1$. Galaxy SEDs in MICE2 were assigned from the COSMOS galaxy template library (Ilbert et al. 2009) such that they reproduced observed magnitude and color relations as well as observed clustering properties as a function of luminosity and color.

While the MICE2 simulation spans a full octant of the sky up to redshift 1.4, we take a small patch defined by the cuts $3 \leq \text{RA[deg]} \leq 4$, $44 \leq \text{DEC[deg]} \leq 45$, $i_{\text{DES}} < 23.5$ and $z < 1.1$. We take the galaxy positions, RA and DEC, the true redshift and apparent magnitude in the DES i-band, the true COSMOS SED id that defines the continuum of the galaxy, the extinction law id that was used to add the dust attenuation and the amount of dust attenuation parameter $E(B-V)$. The continuum templates and extinction law curves used in MICE are the same shown in Fig 3.4 and Fig. 3.5. Using the continuum template, redshift and extinction we generate narrow band continuum fluxes for every galaxy, as well as continuum broad band fluxes with the filter curves from Fig. 1.4 and Fig. 3.1. We adjust the fluxes such that the Subaru i-band has the same apparent magnitude as the DES i-band from MICE2.

To add emission line fluxes we use the recipe from Ilbert et al. (2009), where the flux from the [OII] line is predicted from the UV luminosity of the NUV broad band from GALEX using the Kennicutt (1998) calibration laws, and the remaining line fluxes are tied to the [OII] flux with a set of ratios. From Kennicutt (1998), the UV luminosity at 2300Å can be related to the star

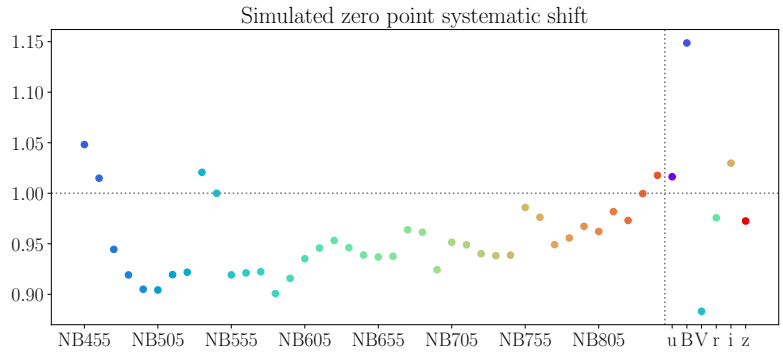

Figure 3.9: Systematic zero point added to the narrow and broad bands of the PAUS mock catalog to mimic the similar effect observed in real data.

formation rate (SFR), $\mathrm{SFR}(M_\odot \mathrm{yr}^{-1}) = 1.4 \times 10^{-28} L_\nu(\mathrm{erg\ s^{-1} Hz^{-1}})$, and the SFR translated to an [OII] emission flux using $\mathrm{SFR}(M_\odot \mathrm{yr}^{-1}) = (1.4 \pm 0.4) \times 10^{-41} L_{[OII]}(\mathrm{erg\ s^{-1}})$, which leads to

$$\log(F_{[OII]})[\mathrm{erg\ s^{-1}\ cm^{-2}}] = -0.4(M_{UV} + \mathrm{DM}(z)) - 6.445 \tag{3.24}$$

where both $\log(F_{[OII]})$ and M_{UV} are dust free, M_{UV} is the absolute UV magnitude and $DM(z)$ the distance modulus. Using Eq. 3.24 and the emission line ratios from Table 3.1, obtained also from Ilbert et al. (2009), we add emission lines to the continuum fluxes. When the galaxy has nonzero extinction we correct the emission line ratios with the extinction law and attenuation value.

We follow by assigning an error to each narrow and broad band by fitting a line to the relation between magnitude and $\log_{10}(S/N)$ obtained in each band from real data in the COSMOS field. We add to the error budget a systematic error of 3% the true flux, which in real data usually appears to account for systematic errors missed by the data reduction pipeline. We make a Gaussian sampling using the true flux and the assigned error to assign a measured simulated flux. We produce a systematic free catalog that only contains pure Gaussian noise.

Based on this clean catalog, we add the systematic effects described in subsections 3.3.4 and 3.3.5. We first add a systematic zeropoint factor to each band, similar to one observed in real data. All galaxies get their fluxes modified by dividing the factors shown in Fig. 3.9. Then we add a systematic shift between the narrow and broad bands by multiplying each galaxy by a value drawn from a Gaussian centered at 1/0.625 and a width of 0.3. We produce two catalogs, one with only the zero point systematic, and a second

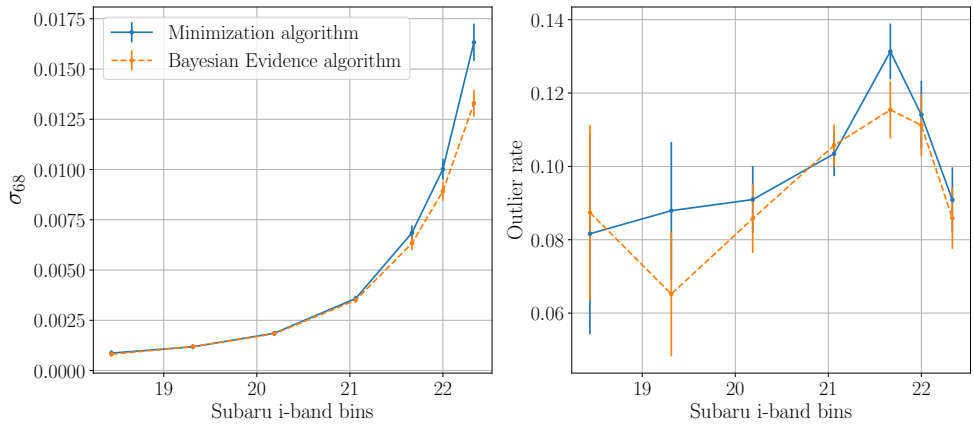

Figure 3.10: Photo-z precision measured with σ_{68} and outlier rate for a clean simulated galaxy catalog with no systematic effects. The errors are measured using 1000 boostrap realizations in each magnitude bin.

adding the narrow and broad band offset on top of the zero point systematic.

3.5 Results

In this section we describe the results found in the simulations from section 3.4 and the results found in the data catalogs described in section 3.2. We define our photo-z point estimate $z_{\rm phot}$ as the mode of the posterior redshift distribution of each galaxy $p(z|f)$ (Eq. 3.4). We will generally use two metrics to asses the photo-z precision, the dispersion σ_{68} of the $\Delta_z = (z_{\rm phot} - z_{\rm true})/(1 + z_{\rm true})$ distribution defined as

$$\sigma_{68} = \frac{P[86] - P[16]}{2} \tag{3.25}$$

where $P[x]$ is the value of the distribution for the percentile x, that is more robust to outliers than the standard deviation. We will also measure the outlier rate as the fraction of outlier galaxies defining as an outlier galaxies that fulfill

$$\text{Outlier} \equiv \frac{|z_{\rm phot} - z_{\rm true}|}{1 + z_{\rm true}} > 3\ \sigma_{68}. \tag{3.26}$$

3.5.1 Simulations

We want to find what is the expected photo-z performance when the conditions are ideal, when the data has been generated from the models and so the models are representative of the data, and the noise is perfectly estimated. We

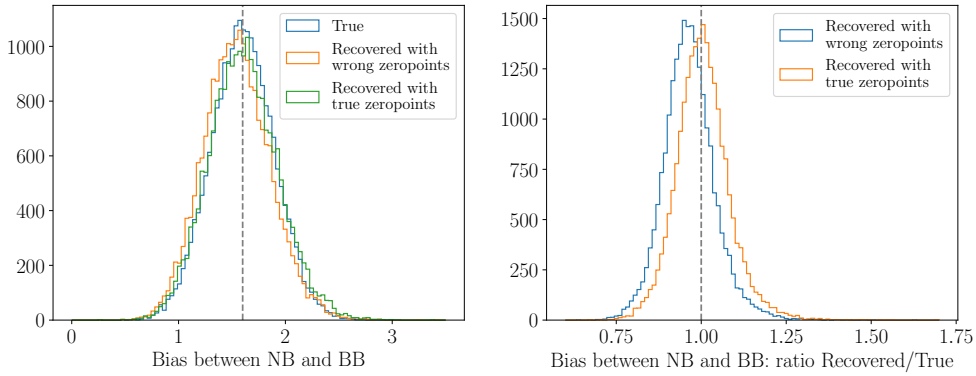

Figure 3.11: Bias between the narrow and broad band photometric systems measured in the simulations. The left panel shows the distribution of biases input to the simulated galaxies (blue) and the recovered distribution in a simulation with band to band zero points shifts (orange) and a simulation without (green). The right panel shows the distribution of ratios between recovered and true bias between NB and BB for each galaxy, showing that when the correction is performed with a catalog that has band to band systematic zero points there is a bias in the recovery (blue distribution compared to orange, that is unbiased). This bias can be corrected afterwards in the band to band zero point calibration.

take the minimization and Bayesian evidence algorithms described in section 3.3 and we run over the clean simulated catalog using a redshift grid between redshift 0 and 1.1 in steps of 0.001. We use a random subsample of 8500 objects with magnitude $i \leq 22.5$. Fig. 3.10 shows the photo-z precision and outlier rate obtained for each algorithm as a function of Subaru i-band bins. The errors are found using 1000 boostrap realizations in each magnitude bin. We use a flat prior for the population prior $p(z, T)$. The σ_{68} grows at fainter magnitudes since the signal to noise decreases. The photometric redshift precision at bright magnitudes is extremely good and quasi spectroscopic. The outlier rate is between $8\% - 13\%$ at all magnitudes, but seems to grow around magnitude 21.5. For a Gaussian distribution only 0.3% of objects would be expected beyond 3σ. While both algorithms produce good results, the Bayesian evidence algorithm recovers a tighter distribution at fainter magnitudes and has a slightly lower outlier rate.

We test the method to calibrate between the NB and BB systems with a synthetic broad band using the simulated catalogs with systematics, which include a band to band bias in the zero point (Fig. 3.9) and a bias between the NB and BB systems. Fig. 3.11 shows the bias between photometric systems measured as the ratio between the observed Subaru r-band and the estimated synthetic Subaru r-band from the narrow bands described in section 3.3.4. We run this estimate in two catalogs, one which has zero point shifts and a catalog

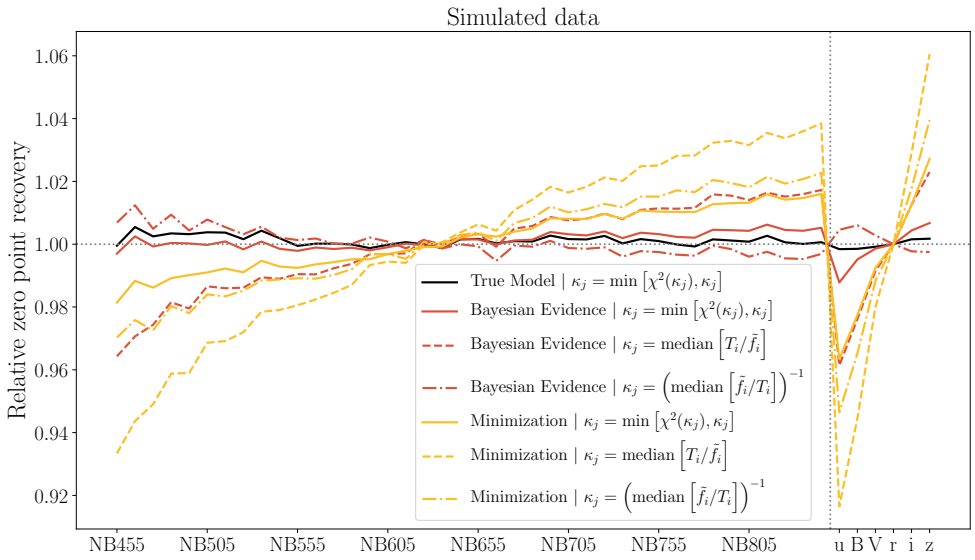

Figure 3.12: Relative zero point recovery measured in a simulation with a systematic zero points between bands and a bias between the narrow and broad band photometric systems. It shows the ratio between the recovered zero point of the bands and the true zero point from Fig. 3.9 including the bias effect measured in Fig. 3.11. We show results for the two photo-z algorithms, Bayesian evidence (red lines) and minimization algorithm (yellow lines), using the three zero point calibration methods (see Eqs. 3.21- 3.23). For reference we add a case where instead of picking the model with a statistical method we use the true model used to generate each galaxy (black line).

that does not. When there are no zero point shifts the estimator recovers an unbiased estimate, while in the presence of zero points biases the estimator gets also a bias. However, this last effect is a global shift that can be corrected later in the zero point calibration between bands. We have found no correlation between the synthetic correction and magnitude nor color.

The last step in the calibration is to measure if there are any residual zero point shifts. We apply the iterative method described in section 3.3.5 on the simulated catalog that has both zero point systematics and a NB to BB bias. The galaxies are previously corrected for the NB to BB bias using the synthetic broad band estimate from Fig. 3.11. Fig. 3.12 shows the recovered zero points relative to the true ones from Fig. 3.9 and the median bias measured from the shift in the right panel of Fig. 3.11. We show results for the two algorithms that are used to select the most likely model for the residual measurement, minimization (yellow lines) and Bayesian evidence (red), for the three zero point methods described in the modeling section, using the χ^2, the median and the inverse median of the residual distribution at each step (Eqs. 3.21-

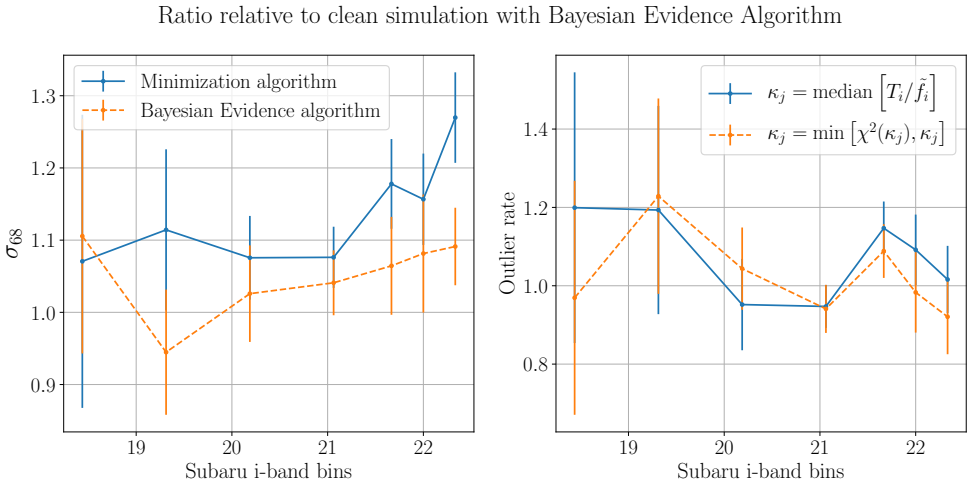

Figure 3.13: Photo-z performance in a simulation with systematic effects which include a band to band zero point shift and a bias between narrow band and broad bands. The systematic calibration procedure has been applied before measuring redshifts. The results are relative to the Bayesian evidence results found with the ideal simulation free of systematics (orange lines in Fig. 3.11), to show the degradation due to systematics. The orange line shows the best case of the zero point recovery algorithm from Fig. 3.12 (Bayesian evidence and χ^2 method), while the blue line shows the opposite end with the worst zero point recovery (Minimization and median method).

3.23). For reference, we include a zero point calibration which fixes the best model to the true model used to generate each galaxy (black line). All lines are normalized with respect to the r-band, since zero point corrections are only sensitive to colors (or flux ratios), not individual fluxes. Using the true model delivers the nearest result to perfect recovery as expected, and the methods that obtain the best recovery are using the Bayesian evidence either with the χ^2 or the inverse median method. We find using the median method estimate, which has the observations dividing (Eq. 3.22), generates a bias towards bluer bands. The reason is that bluer bands, especially the narrow bands, have very small decreasing signal to noise . The minimization algorithm always finds an answer that creates a red color trend.

Fig. 3.13 shows the photo-z impact from applying the systematic calibration. The results are shown relative to the values obtained from the Bayesian evidence algorithm in the clean simulation (orange lines in Fig. 3.11). The orange line shows a run with the best case scenario from the Bayesian evidence algorithm with the χ^2 method for zero point calibration. We only find a $5 \sim 10\%$ degradation from the precision found in the ideal simulations free of systematics in the faint end. In contrast, the blue line shows the worst

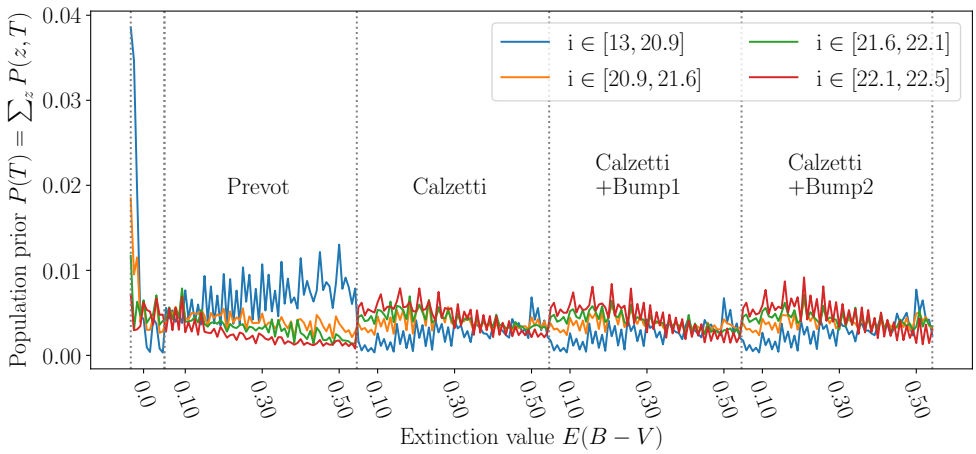

Figure 3.14: Population prior probability of the bayesian evidence models, $p(T|m) = \sum_z p(z, T|m)$, in simulations. The prior is self calibrated comparing the color space occupation of all observed galaxies in the simulation to the occupation of all the models T. Only the broad bands are used. It is presented in four magnitude bins with equal number of galaxies.The five extinction laws (including no extinction) are separated by the vertical lines. Within one extinction law the models are sorted from lower to larger value of extinction. There are three models for the same extinction law and extinciton value which include three different emission line models: no emission line, emission line with fixed ratios and emission line with fixed ratios except for the [OIII] doublet. There are a total of 60 models for each extinction law, except for no extinciton where there are 10.

case from the zero point recovery, using the minimization algorithm with the median zero point calibration. In this case, the photo-z precision becomes a $15 \sim 25\%$ worse at the faint end. Unlike the precision, the outlier rate seems to be similar than in the ideal situation, but note that this is defined with respect to the σ_{68} of the distribution, which has grown, so the distribution tails are larger in absolute amount.

The population prior scheme $p(z, T)$ from section 3.3.3 calibrates the observed color space occupation of the data by comparing it to the occupation from the models, using the broad bands. We only use the marginal information on the models of this prior, by summing over redshift $p(T) = \sum_z p(z, T)$. Fig. 3.14 shows the prior distribution on the models of the bayesian evidence in four magnitude bins which split the galaxies evenly. Note how the red Elliptical galaxies, which do not allow for extra dust reddening, are upweighted with respect to the models with extinction, since there are several models with extinction that have similar observed colors. Also, note how the red models and Prevot extinction law models loose weight for fainter galaxies, as expected. To break color-redshift degeneracies in the prior, we add a population redshift prior $p(z)$ from running the photo-z code first with a flat priors and smoothing

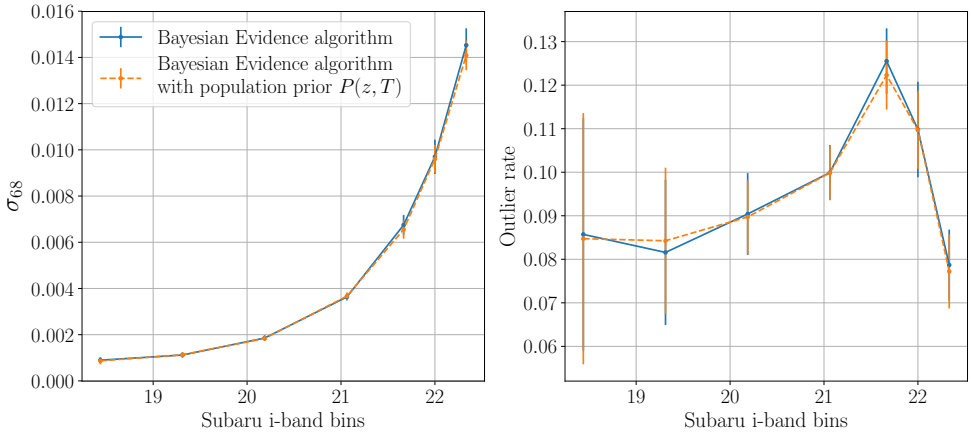

Figure 3.15: Photo-z performance in a simulation with systematic effects, adding a population prior $p(z,T|m)$ self calibrated from the data. We approximate the prior as $p(z,T|m) = p(z|m)p(T|m)$, where the model prior is found from a self calibration with the broad bands, by comparing the color space occupation of the data and all the models (see Fig. 3.14). The redshift prior $p(z|m)$ is found from running the photo-z algorithm with a flat prior $p(z,T|m)$, and smoothing the posterior redshift distribution with a top hat function with a width of 0.1 in redshift.

the posterior redshift distribution of the population with a top-hat filter of size 0.1 in redshift, $p(z,T|m) = p(z|m)p(T|m)$. Fig. 3.15 shows the effect of the population prior in the photo-z. We find the population prior has a very small effect in the photo-z, indicating the data is much more informative than the prior for the redshift determination.

When running the photo-z code we do not only get information on the redshift, but also on the models T. Fig. 3.16 shows the true distribution of extinction values $E(B-V)$ used in the simulation, compared to the distribution found after running the photo-z algorithms. We find all algorithms and zero point calibrations differ slightly from the true extinction distribution. Adding population priors increases the amount of galaxies with no extinction and reduces the amount that have very high extinction, but also changes the shape of the extinction values betwen 0.05 and 0.4, creating a peak around 0.3. We show what is the impact in the extinction of a color trend coming from the zero point recovery by comparing the distribution from Bayesian evidence and χ^2 method with the minimization algorithm using the median method (see Fig. 3.12). We find this creates an excess of galaxies at the largest extinction value, shifts galaxies towards larger extinction values and increases the amount of galaxies with no extinction.

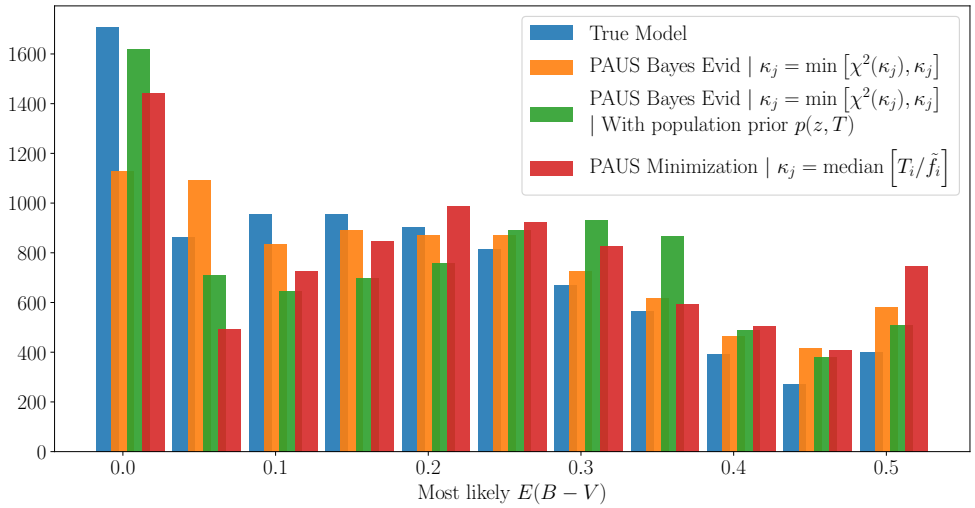

Figure 3.16: Most likely $E(B-V)$ extinction value, where the most likely value for each galaxy is estimated from the mode of the posterior distribution, marginalizing over redshift. We show the true distribution of extinction used to generate the data (blue) compared to the distribution found with the best photo-z configuration (orange, Bayesian evidence and χ^2 method for zero point calibration). The effect of adding a population prior $p(z, T|m)$ is shown (green), which leads to a better recovery of red elliptical galaxies (no extinction), reduces the amount of high extinction galaxies ($E(B-V) > 0.4$) and creates a peak around extinction of 0.3. We compare to the worst photo-z configuration, which introduces a red color trend, finding it shifts the distribution towards larger extinction, and increases both the amount of red and very high extinction galaxies.

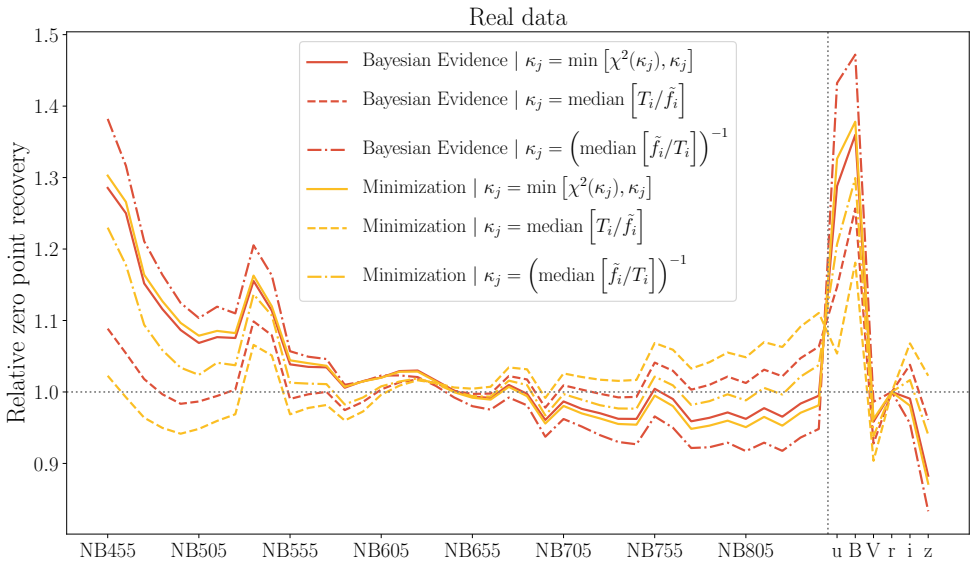

Figure 3.17: Relative zero point recovery measured in real data, using the 40 narrow band data from PAUS in the COSMOS field and the 6 broad band data from COSMOS survey (u,B,V,r,i and z). All lines are shown relative to the r-band, since the zero point calibration can only calibrate colors (or flux ratios). We show results for the two photo-z algorithms, Bayesian evidence (red lines) and minimization algorithm (yellow lines), using the three zero point calibration methods (see Eqs. 3.21- 3.23). Data has been corrected for a bias between narrow band and broad band data using a synthetic broad band.

3.5.2 Data

Now we present the results obtained using real data. We will use the same methods tested in simulations and compare the results between simulated and real data. We use the 40 narrow band data from PAUS and the 6 broad bands from COSMOS survey and compare the photo-z estimates against the zCOSMOS DR3 spectroscopic redshifts. We only use high quality spectroscopic redshifts to minimize spectroscopic failures and remove any masked objects either from PAUS or COSMOS catalogs (see section 3.2 for details about the data reduction and other selections). The three catalogs are matched in sky position by requiring objects to be closer than one arcsecond. To test the photo-z we only use objects with measured flux in all bands, which leads to a total of 7804 objects. Similar to the simulation, we also add a systematic error of 3% the measured flux to account for errors missed in the data reduction pipeline and model uncertainty. We measure the bias between narrow band and broad band data comparing a synthetic broad band with the measured r-band (see section 3.3.4).

Fig. 3.17 shows the zero point recovery in data, showing both photo-

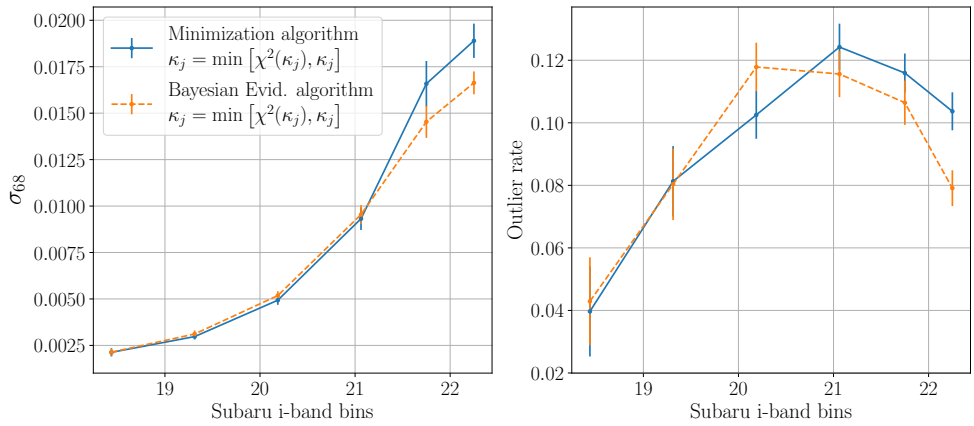

Figure 3.18: Photo-z performance in real data using the 40 narrow band data from PAUS in the COSMOS field and the 6 broad band data from COSMOS survey (u,B,V,r,i and z), compared to the spectroscopic redshift estimate from zCOSMOS DR3 spectroscopic survey. Data has been corrected for a bias between narrow band and broad band data using a synthetic broad band. Zero point corrections are applied using the χ^2 method. The minimization algorithm (solid, blue) and the Bayesian evidence algorithm (dashed, orange) are compared.

z algorithms (Bayesian evidence and minimization) for the three zero point models (see Eqs. 3.21- 3.23). We find a significant blue color trend both in narrow band and broad band data when using the more reliable zero point methods (χ^2 and inverse median). The median zero point method, which we know from simulations is biased and creates a red color trend, compensates the intrinsic blue color trend present in the data. The minimization algorithm delivers a redder zero point than the bayesian evidence for the same zero point method, except for the χ^2 where they look similar. The χ^2 method was the best method in simulations, but could be more sensitive to outliers in the data, therefore the results can differ slightly form the ones found in simulations (see Fig. 3.12). The blue color trend found in the data could be due to some missed effect in the data calibration step and is currently under investigation.

Fig. 3.18 shows a comparison between the photo-z performance of the minimization and bayesian evidence algorithms, using the χ^2 as the zero point correction. We use one less bin above magnitude 21.5 since the amount of spectroscopic data significantly decreases (see Fig. 3.2). Similar to the simulation, we find the Bayesian evidence algorithm performs better, specially at faint magnitudes, finding a better precision and lower outlier rate. We find the minimization algorithm performs a little better at magnitudes brighter than $i < 21$. The σ_{68} is a factor $2 \sim 3$ times larger in the data than in the simulation at magnitudes brighter than $i < 21.5$, and a factor 1.4 larger in the faintest

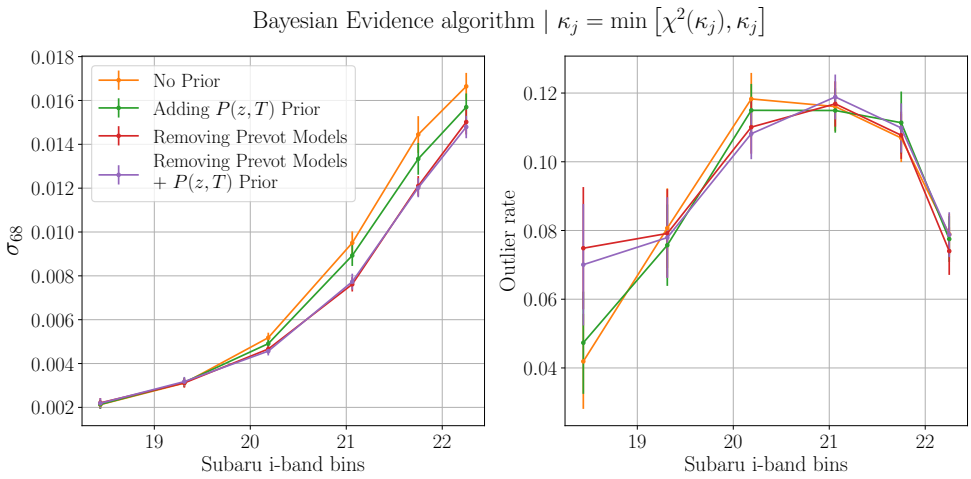

Figure 3.19: Photo-z performance in real data, showing the effect of adding a population prior and removing models with the Prevot extinction law. Removing models with the Prevot extinction law leads to a $10 \sim 25\%$ improvement in σ_{68} at faint magnitudes. The population prior tends to weight the Calzetti models over the Prevot models, which can explain why the prior has a larger impact in data than in simulations. The fact that the prior has a small effect when the Prevot models are discarded reinforces this conclusion.

magnitude bin. The outlier rate is found to be similar between simulatons and data, around $0.08 \sim 0.12$, except for the bright magntiudes where it is smaller in data. We also find a similar trend where the outlier rate grows and peaks around magnitude 21.

The fact the photo-z dispersion is larger hints that several assumptions made in the simulations could be too optimistic. On the one hand, the SED modeling choices could not be completely representative of all the data, since some populations could be missing or the emission line modeling, where one fixes most of the ratios to one value, could be too naive. On the other hand, outliers and spurious measurements are present in the data and could add extra noise to the redshift estimation process. PAUS images have variable amount of scatter light around the edges in some filter bands and uncorrected negative cross talk happening between different CCDs, which can produce such suprious measurements. The data reduction is constantly improving and these effects will have a smaller impact or be removed in future data releases.

We compute a population prior following the same approach showed in simulations, but using real data. Fig. 3.19 presents the effect of the prior using the Bayesian evidence algorithm with the χ^2 zero point method. We find the prior improves the photo-z precision by $5 \sim 10\%$ at magnitudes fainter than $i > 20$, unlike in the simulations where we found very marginal gains. We

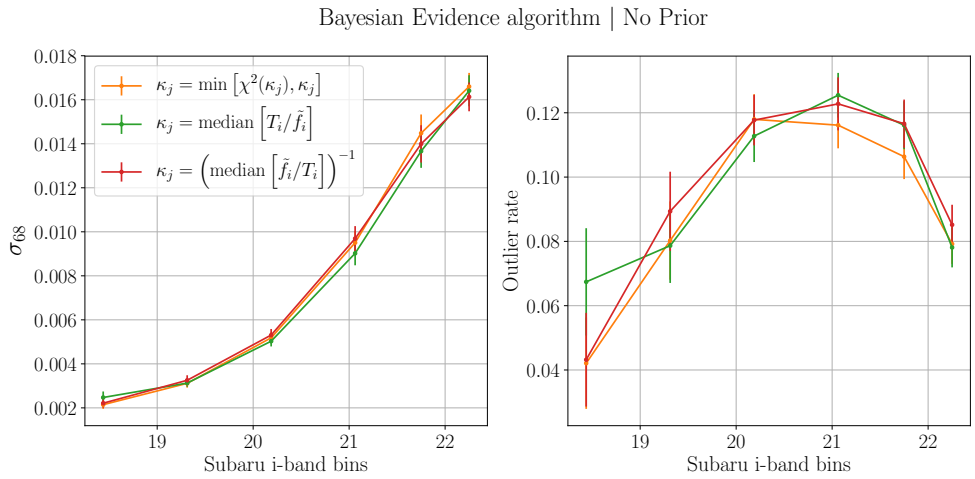

Figure 3.20: Photo-z performance in real data. Comparison without priors between the three zero point recovery methods: the χ^2, the median and inverse median methods (see Eqs. 3.21- 3.23).

have noted that removing models with the Prevot extinction law (see Fig. 3.5) improves the precision by $10 \sim 25\%$ at magnitudes fainter than $i > 20$. We have found the population prior downweights the probability of selecting the Prevot extinction law, both in simulations and data (not shown). We conclude that the population prior is improving the precision in the presence of the Prevot models due to this effect. Adding priors when the Prevot models have been discarded leads to marginal gains again. The outlier rate does not change when adding priors, but it increases for the brightest galaxies when removing the Prevot templates.

We compare the impact in the photo-z precision of using different zero point recovery methods. Fig. 3.20 shows the photo-z precision for the Bayesian evidence algorithm for the χ^2, the median and inverse median methods (see Eqs. 3.21- 3.23), without priors. Fig. 3.21 shows the same but removing models that use the Prevot extinction law and adding population priors. We see the differences between all cases are very small, despite the zero point recoveries being variable. Even using the median zero point method, that we know is biased by construction in our data, and has a very different zero point, leads to a similar (if not slightly better) photo-z result.

Figs. 3.22 and 3.23 show the estimated extinction value $E(B-V)$ distribution without using priors and with a population prior, respectively. In real data we do not have the true extinction value of a galaxy. However, we approximate it with COSMOS 30 band data from Laigle et al. (2016), since

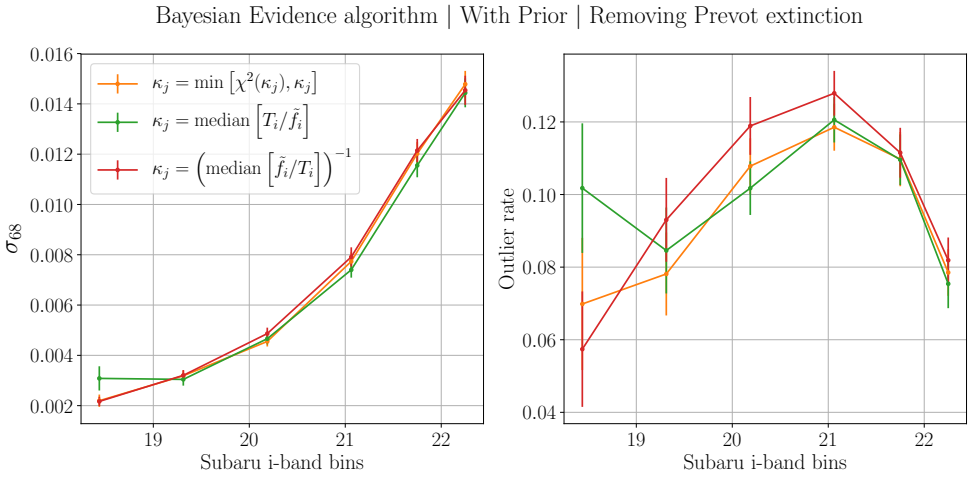

Figure 3.21: Similar to Fig. 3.20 but using population priors and disregarding models with the Prevot extinction law.

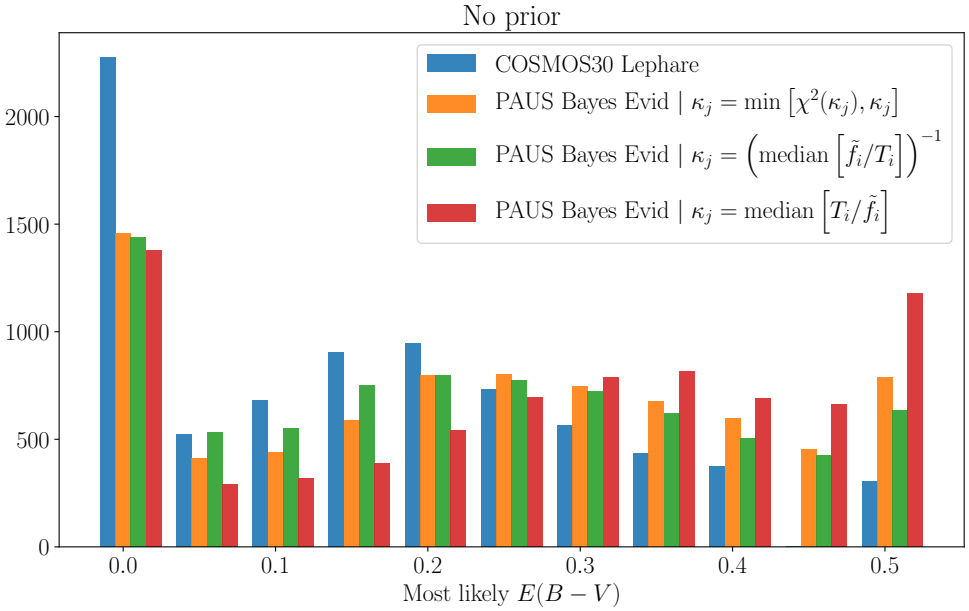

Figure 3.22: Extinction values $E(B-V)$ distribution found in real data. Results from COSMOS 30 band data (Laigle et al. 2016) are used as an approximation to the true values, since this dataset contains data ranging from far ultraviolet until far infrared, providing a more accurate measurement of the galaxy extinction. Distributions are shown for the Bayesian evidence algorithm comparing the three zero point recovery methods. While zero point color trends do not affect the photo-z precision, it can change the recovered extinction values of the galaxy. We find similar results compared to simulations. The median zero point method (which tends to create a red color trend) finds a very large excess of large extinction values ($E(B-V) > 0.4$).

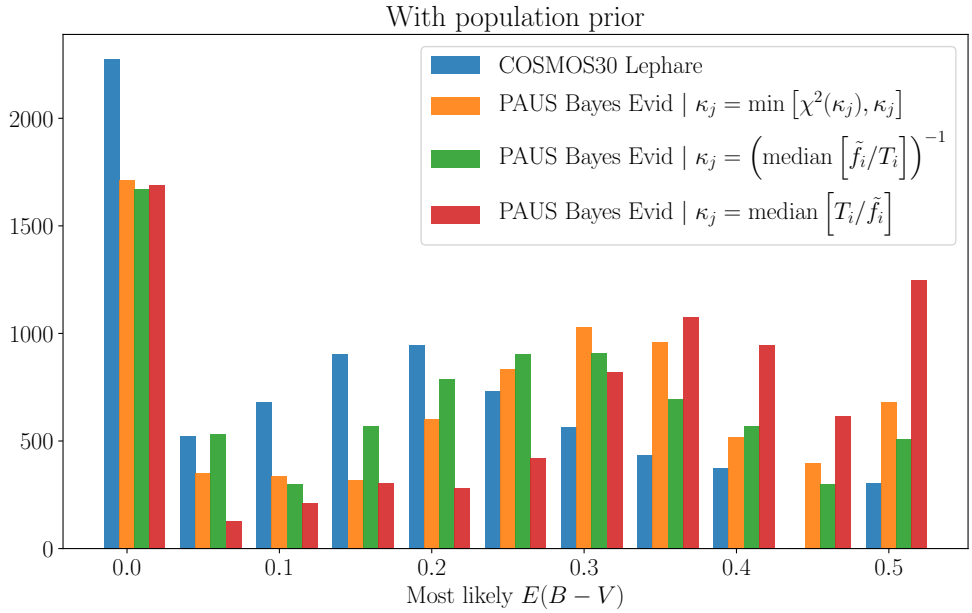

Figure 3.23: Extinction values $E(B-V)$ distribution in real data. Similar to Fig. 3.22 but using population priors. The prior has a similar effect as in the simulations, increasing the amount of galaxies with no extinction and shifting the extinction of galaxies towards larger values.

it contains data spanning from far ultraviolet to far infrared, which allows to constrain extinction models with much better accuracy. Similar to simulations, we find a lack of objects with no extinction and an extinction distribution shifted towards larger extinction values. Adding priors increases the amount of objects without extinction but shifts the remaining of the distribution even further to larger extinction values. We find using the median zero point method delivers an extinction distribution extremely biased towards high extincton ($E(B-V) \geq 0.4$). The inverse median method recovers a distribution closer to COSMOS 30 band compared to the distribution from the χ^2 zero point method. This could be as a result of this method being more robust to model inaccuracies or outliers in the data.

3.6 Discussion

In this chapter we have used data from the Physics of the Accelerating Universe Survey (PAUS) to measure photometric redshifts. PAU Survey is a new survey which features a unique instrument, PAUCam, which contains a novel set of 40 narrow band filters of 12.5nm FWHM, spaced by 10nm in the range between 4500Å$-$8500Å. We have used data taken in the COSMOS field, a small area of $\sim 1.5\text{deg}^2$ which was targeted for calibration purposes between 2015 and 2017. The COSMOS field contain many observations, including a multi band survey, the COSMOS survey, which has photometry in over 30 bands, containing mainly intermediate and broad bands, but also two narrow bands, which span from ultraviolet to infrared wavelength. A spectroscopic survey called zCOSMOS bright overlaps in the same area and provides redshifts for galaxies until $i_{\text{AB}} \leq 22.5$, which we can use to validate our photo-z precision.

One of the main challenges to obtain photometric redshift estimates was to properly calibrate the zero point offsets between bands, by comparing models to data. Especially, the zero point between the narrow bands and the external broad band flux measurements, which used different photometry extraction strategies. We developed mock fluxes based on galaxy properties from the public MICE2 N-body simulations which contained the same 40 narrow bands and 6 broad bands from the COSMOS field. We added the observed signal to noise from the data to generate measured fluxes and artificially included zero point offsets and a bias between the narrow and broad band system, as observed in data.

To correct a bias between the narrow and broad band system we developed a self calibration method building a synthetic broad band from a com-

bination of narrow band measurements. Then, the zero point offsets can be corrected by using an iterative algorithm that compares measurements against a best model prediction using a subset of galaxies with spectroscopy. We have tested in simulations that both methods work and obtain a photo-z precision which is only $5\% - 10\%$ larger than in a ideal setup with no systematics at all.

We have developed two photometric redshift algorithms that rely on a linear combination of galaxy SED models. Then, emission lines can be properly modeled for each galaxy, since these can have a very different amplitude with respect to the galaxy continuum flux. One method finds the best model parameters using a minimization algorithm which restricts to non negative values. This code is called BCNZ2 and has been used in Eriksen et al. (2019) in an early demonstration of PAUS photo-z, and is similar to EAZY (Brammer, van Dokkum & Coppi 2008), although we have a different treatment for galaxy internal extinction and emission lines. We have developed another photo-z code which instead of finding the best parameters integrate over all possible combinations within some prior distribution which includes the non negativity constraint. This integral is the Bayesian evidence, and we found it gives an increasingly better photo-z precision at fainter magnitudes, both in data and simulations. The photo-z precision found in data is exceptional compared to typical photo-z surveys. For high signal to noise galaxies we obtain quasi spectroscopic precision thanks to the fine spectral resolution from the narrow bands, while for the fainter objects with $i_{\rm AB} = 22 - 22.5$ we obtain a 1.4% precision, despite the decreasing signal to noise of the observations.

The photo-z dispersion in data was larger than in simulations. The simulation mimics the same noise levels and some of the known systematic effects from data, but does not contain all the richness and complexity of galaxy SEDs that real data can have. In particular, the results in simulations are found using the same models used to generate the data. In this sense, in the simulations the models are a complete representative set of the observations, while in real data this will likely be different. Some galaxy populations can be missing and the emission line modeling, where the line fluxes were fixed by a ratio, is rather simple. Moreover, data is not perfect and also shows outliers from uncorrected imaging effects such as scatterlight or negative crosstalk, which can severely bias a measurement. However, we are hopeful that with better modeling and improved data reduction we can obtain a photo-z precision in data closer to simulations.

3.A Algorithms

In this section we present details of the two photo-z algorithms that are used in this chapter.

3.A.1 Minimization algorithm

To minimize the likelihood defined in Eq. 3.5 we restrict to non-negative parameters to avoid unphysical solutions. Also a combination of negative emission lines could fit spurious data coming from negative cross-talk between different CCDs. We have implemented a non-negative quadratic programming iterative algorithm based on Sha et al. (2007), which defines

$$A_{xy} \equiv \sum_i \frac{f_i^x f_i^y}{\sigma_i^2}, \quad b_x \equiv \sum_i \frac{f_i^x \tilde{f}_i}{\sigma_i^2}, \tag{3.27}$$

for templates x and y, where the sum is over the bands and $\tilde{f}_i$ and σ_i are the flux and flux error. The amplitudes α_x get updated iteratively using

$$m_x = \frac{b_x}{\sum_y A_{xy}\alpha_y}, \quad \alpha_x \to m_x \alpha_x \tag{3.28}$$

In the implementation the minimum is estimated at the same time for a set of galaxies, for all the different redshift bins.

3.A.2 Bayesian Evidence integral algorithm

To compute the Bayesian evidence we need to integrate Eq. 3.6. Let us redefine the following terms from Eq. 3.7, $f_k \equiv f_k^{obs}/\sigma(f_k^{obs})$ and $M_{jk} \equiv t_{jk}/\sigma(f_k^{obs})$. Then we can write the exponent in Eq. 3.7 as

$$\begin{aligned}\sum_k^d \left(f_k - \sum_j^{n_i} \alpha_j M_{jk}\right)^2 &= \sum_k^d f_k^2 - 2\sum_k^d \sum_j^{n_i} f_k \alpha_j M_{jk} + \\ &+ \sum_k^d \sum_j^{n_i} \sum_\ell^{n_i} \alpha_j \alpha_\ell M_{jk} M_{\ell k} \\ &= \vec{f}\cdot\vec{f} - 2\vec{B}\cdot\vec{\alpha} + \vec{\alpha}^\intercal \hat{A}\,\vec{\alpha}\end{aligned} \tag{3.29}$$

where in the second step we have defined the matrix $A_{j\ell} \equiv \sum_k^d M_{jk}M_{\ell k}$ and vector $B_j \equiv \sum_k^d f_k M_{jk}$. Then, Eq. 3.7 becomes

$$P(f|\vec{\alpha}, z, T) = \frac{\exp(-\frac{1}{2}\vec{f}\cdot\vec{f})}{\sqrt{(2\pi)^d}\prod_k \sigma(f_k)} \exp\left[-\frac{1}{2}\vec{\alpha}^\intercal \hat{A}\,\vec{\alpha} + \vec{B}\cdot\vec{\alpha}\right] \tag{3.30}$$

and assuming the prior from Eq. 3.8 then Eq. 3.6 becomes

$$P(f|z, T) \propto \frac{1}{\prod_j \Delta_j} \int_0^{\Delta_1} \dots \int_0^{\Delta_{n_i}} d^{n_i}\alpha \, \exp\left[-\frac{1}{2}\vec{\alpha}^\intercal \hat{A}\,\vec{\alpha} + \vec{B}\cdot\vec{\alpha}\right] \tag{3.31}$$

where we have dropped the constants that do not depend on the model. Eq. 3.31 means integrating an uncentered multivariate normal distribution, which we can center with the following change of variable

$$\begin{aligned} P(f|z, T) &\propto \frac{1}{\prod_j \Delta_j} \int_0^{\Delta} \exp\left[-\frac{1}{2}\vec{\alpha}^\intercal \hat{A}\,\vec{\alpha} + \vec{B}\cdot\vec{\alpha}\right] d^{n_i}\alpha \\ &= \frac{\exp(\frac{1}{2}\vec{\mu}\cdot\vec{B})}{\prod_j \Delta_j} \int_0^{\Delta} \exp\left[-\frac{1}{2}(\vec{\alpha}-\vec{\mu})^\intercal \hat{A}\,(\vec{\alpha}-\vec{\mu})\right] d^{n_i}\alpha \\ &= \frac{\exp(\frac{1}{2}\vec{\mu}\cdot\vec{B})}{\prod_j \Delta_j} \int_{-\mu_1}^{\Delta_1-\mu_1} \dots \int_{-\mu_{n_i}}^{\Delta_{n_i}-\mu_{n_i}} \exp\left[-\frac{1}{2}\vec{\theta}^\intercal \hat{A}\,\vec{\theta}\right] d^{n_i}\theta \end{aligned} \tag{3.32}$$

where $\vec{\theta} = (\vec{\alpha} - \vec{\mu})$ and $\vec{\mu} = \hat{A}^{-1}\vec{B}$ corresponds to the peak vector of the distribution (unconstrained maximum likelihood point). The Alan Genz approach (Genz 1992) consists in three transformations that make the numerical integration of Eq. 3.32 more efficient. Following Genz (1992), we define the integral of a Gaussian as function F of its integration limits $\vec{a}$ and $\vec{b}$,

$$F(\vec{a}, \vec{b}) = \frac{1}{\sqrt{|\hat{A}^{-1}|(2\pi)^n}} \int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} e^{-\frac{1}{2}\vec{\theta}^\intercal \hat{A}\,\vec{\theta}} d\vec{\theta}. \tag{3.33}$$

Then our integral becomes

$$P(f|z, T) \propto \frac{\exp(\frac{1}{2}\vec{\mu}\cdot\vec{B})\sqrt{|\hat{A}^{-1}|(2\pi)^{n_i}}}{\prod_j \Delta_j} F(\{-\mu_j\}, \{\Delta_j - \mu_j\}) \tag{3.34}$$

The transformations and code for estimating $F(\vec{a}, \vec{b})$ in Eq. 3.33 are described in the following section 3.A.3. Eq. 3.34 can be interpreted as the total volume

under the multivariate Gaussian (Eq. 3.35) times the fraction of that Gaussian inside the top hat prior limits (Eq. 3.33) divided by the prior hypercube volume ($\prod_j \Delta_j$), since

$$\exp\left(\frac{1}{2}\vec{\mu}\cdot\vec{B}\right)\sqrt{|\hat{A}^{-1}|(2\pi)^{n_i}} = \int_{-\infty}^{\infty}\exp\left[-\frac{1}{2}\vec{\alpha}^{\intercal}\hat{A}\,\vec{\alpha}+\vec{B}\cdot\vec{\alpha}\right]d^{n_i}\alpha \tag{3.35}$$

3.A.3 Alan Genz Gaussian integral

In this subsection we reproduce the algorithm from Genz (1992) used for the integration of Eq. 3.33 using 3 transformations. The first transformation will use the Cholesky decomposition of $\hat{C}\hat{C}^{\intercal} = \hat{A}^{-1}$, $\vec{\theta} = \hat{C}\vec{y}$. In this case $\vec{\theta}^{\intercal}\hat{A}\,\vec{\theta} = \vec{y}^{\intercal}\vec{y}$ and $d\vec{\theta} = |\hat{A}^{-1}|^{\frac{1}{2}}d\vec{y}$. Note also how for this decomposition the new integration limits will be:

$$\begin{aligned}\vec{a}\le\vec{\theta}=\hat{C}\vec{y}\le\vec{b}\\ a_i\le\theta_i=\sum_j^n C_{ij}y_j\le b_i\\ \tilde{a}_i\equiv(a_i-\sum_{j\neq i}C_{ij}y_j)/C_{ii}\le y_i\le(b_i-\sum_{j\neq i}C_{ij}y_j)/C_{ii}\equiv\tilde{b}_i\end{aligned} \tag{3.36}$$

Note that if C is the lower triangular, $C_{ij}=0$ for $i<j$. Hence,

$$F(\vec{a},\,\vec{b}) = \frac{1}{\sqrt{(2\pi)^n}}\int_{\tilde{a}_1}^{\tilde{b}_1}e^{-\frac{y_1^2}{2}}\int_{\tilde{a}_2(y_1)}^{\tilde{b}_2(y_1)}e^{-\frac{y_2^2}{2}}\ldots\int_{\tilde{a}_n(y_1,\ldots,y_{n-1})}^{\tilde{b}_n(y_1,\ldots,y_{n-1})}e^{-\frac{y_n^2}{2}}d\vec{y}. \tag{3.37}$$

Now, the second transformation, $y_i=\Phi^{-1}(z_i)$, where

$$\begin{aligned}\Phi(y) &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y}e^{-\frac{1}{2}\theta^2}d\theta=\frac{1}{2}\left(\mathrm{erf}\left(\frac{y}{\sqrt{2}}\right)+1\right)\\ \Phi^{-1}(z_i) &= \sqrt{2}\,\mathrm{erf}^{-1}(2z_i-1)\end{aligned} \tag{3.38}$$

then $dz_i=\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}y_i^2}dy_i$, so

$$F(\vec{a},\,\vec{b}) = \int_{\Phi(\tilde{a}_1)}^{\Phi(\tilde{b}_1)}\int_{\Phi(\tilde{a}_2(\Phi^{-1}(z_1)))}^{\Phi(\tilde{b}_2(\Phi^{-1}(z_1)))}\ldots\int_{\Phi(\tilde{a}_n(\Phi^{-1}(z_1),\ldots,\Phi^{-1}(z_{n-1})))}^{\Phi(\tilde{b}_n(\Phi^{-1}(z_1),\ldots,\Phi^{-1}(z_{n-1})))}d\vec{z}. \tag{3.39}$$

Finally, the third transformation, $z_i = d_i + w_i(e_i - d_i)$,

$$F(\vec{a}, \vec{b}) = (e_1 - d_1) \int_0^1 (e_2 - d_2) \dots \int_0^1 (e_n - d_n) \int_0^1 d\vec{w} \tag{3.40}$$

where

$$\begin{aligned} d_i &= \Phi\left(\left[a_i - \sum_{j \neq i} C_{ij} \Phi^{-1}(d_j + w_j(e_j - d_j)) \right] / C_{ii} \right) \\ e_i &= \Phi\left(\left[b_i - \sum_{j \neq i} C_{ij} \Phi^{-1}(d_j + w_j(e_j - d_j)) \right] / C_{ii} \right) \end{aligned} \tag{3.41}$$

Note how for some parameters where either $a_i = -\infty$ or $b_i = \infty$ then $d_i = 0$ and $e_i = 1$, respectfully. Now we describe the algorithm to find the integral (Genz 1992).

1. Input $\vec{a}$, $\vec{b}$, $\hat{A}$, N_{max}.
2. Compute Cholesky decomposition $\hat{C}$ for A^{-1}.
3. Initialize $Intsum = 0$, $N = 0$, $d_1 = \Phi(a_1/C_{1,1})$, $e_1 = \Phi(b_1/C_{1,1})$, $f_1 = e_1 - d_1$.
4. Repeat until $N = N_{max}$
   (a) Generate uniform random $w_1, w_2, \dots, w_{m-1} \in [0, 1]$.
   (b) For $i = 2, 3, \dots, m$ set $y_{i-1} = \Phi^{-1}(d_{i-1} + w_{i-1}(e_{i-1} - d_{i-1}))$, $d_i = \Phi((a_i - \sum_{j=1}^{i-1} y_j C_{ij})/C_{ii})$, $e_i = \Phi((b_i - \sum_{j=1}^{i-1} y_j C_{ij})/C_{ii})$, $f_i = (e_i - d_i) f_{i-1}$.
   (c) Set $N = N + 1$, $\delta = (f_m - Intsum)/N$, $Intsum = Intsum + \delta$.
5. Output $= Intsum$

Part IV

Redshift inference from the combination of galaxy colors and clustering in a hierarchical Bayesian model

In this part we estimate the redshift distribution of a galaxy sample for a broad band survey like DES. We implement a hierarchical bayesian model that incorporates the information of colors and clustering and test it in the MICE2 N-body simulations.

Chapter 4

Redshift inference from the combination of galaxy colors and clustering in a hierarchical Bayesian model

4.1 Introduction

Cosmological analysis using imaging surveys need to estimate the redshift distribution of galaxies $n(z) \equiv dn/dzdA$ very accurately. Any unaccounted errors in its estimation can lead to a systematic bias of the cosmological parameters (Hildebrandt et al., 2012b; Bonnett et al., 2016; Hoyle et al., 2018; Hildebrandt et al., 2017). Several redshift estimation techniques have been developed in the last decades, which can be broadly separated between template fitting techniques like BPZ (Benítez 2000), LePHARE (Arnouts & Ilbert 2011) or EAZY (Brammer, van Dokkum & Coppi 2008) which compare theoretical colors assuming a restframe spectral energy distribution for the galaxy and machine learning methods like SKYNET (Bonnett 2015), ANNz2 (Sadeh, Abdalla & Lahav 2016) or DNF (De Vicente, Sánchez & Sevilla-Noarbe 2016) that learn the color redshift relation with a calibration sample. A recent method has been develop that instead of using galaxy colors uses the galaxy positions, by cross correlating them with the positions of a sample with known redshifts (Newman (2008); Ménard et al. (2013); Schmidt et al. (2013)). Recently, lensing surveys like KiDS (Hildebrandt et al. 2017) and DES (Hoyle et al. 2018; Gatti et al. 2018; Davis et al. 2017; Abbott et al. 2018) have used a combination of such methods to either cross check their redshift inferences or combine them under some parametrizations of the redshift distribution. The third source of information is the prior distribution of redshift and type for the population of galaxies, which is usually calibrated using small patches with either spectroscopic or highly precise photometric redshifts.

Sánchez & Bernstein (2019) presented a framework that combined this three pieces of information (photometry, clustering and prior) in a principled way using a hierarchical bayesian model (see also Leistedt, Mortlock & Peiris 2016) which yielded samples of the redshift distribution constrained by all sources of information. In this work we extend the method so that it can be applied to galaxy surveys. We apply the methodology to the public MICE2 N-body simulations and develop a clustering estimator based on a density field produced with a kernel density estimator from a subsample of galaxy positions with known redshifts. We show how to map the density relation from the estimator to the actual density distribution of the unknown galaxies, obtaining priors for a relation and marginalizing over mapping parameters.

This chapter is organized as follows. In section 4.2 we present the methodology and phenotype approach. Section 4.3 describes the simulated galaxy catalog used to test the methodology. We follow with a description of the density estimation used to incorporate the clustering information in section 4.4. We describe the Gibbs sampling steps used to sample the posterior on all the model parameters in section 4.5. Section 4.6 shows the results for a prior which comes from a small patch with sample variance and three cases where this prior is biased inspired in real data situations.

4.2 Framework

In this chapter, we apply the method of SB18 to more realistic simulations. For that, we develop some new key extensions to the methodology, especially in the treatment of galaxy clustering (section 4.4), but we work under the framework presented there. In that approach, galaxy "types" are defined by observed properties rather than rest-frame properties, and we call them *phenotypes*. On one hand, the colors of individual galaxies i are seen as being drawn from a pool of possible phenotypes t_i, redshifts z_i, and angular positions θ_i with some latent intrinsic mean density $n(t, z)$ on the sky. On the other hand, observations yield a noisy set of observable features which will be denoted as F_i. The clustering information is included by considering that the galaxy densities at redshift z are modulated by some factor $1 + \delta_z(\theta)$, where the density estimates $\delta_z(\theta)$ are come from a galaxy population drawn from the same latent density distribution (up to some bias factor). Our notation will be that the vector quantities F, t, z, and $\boldsymbol{\theta}$ denote the full set of properties of all selected galaxies, i.e., $F = \{F_1, F_2, \ldots, F_N\}$ (a summary of all the notation can be found in Table 4.1).

4.2.1 Generative model

As in SB18, the fundamental assumption of the method is that galaxies are drawn from a *Cox process* (Cox, 1955) or *doubly stochastic Poisson process*, i.e., we assume that each galaxy is Poisson sampled from a latent, stochastic density field. The problem simplifies when considering the redshift z as an integer indexing a set of finite-width redshift bins, where each bin has an independent density fluctuation field $\delta_z(\theta)$, i.e., $\langle \delta_{z_i}(\theta)\delta_{z_j}(\theta)\rangle = 0$ for $z_i \neq z_j$. We will also assume that we have a finite set of phenotypes indexed by integer t. Each phenotype has a mean sky density of $n^t = nf_t$, where we place n as the total density of all detectable galaxy phenotypes, and $f_t = p(t)$ being the fraction of the population in each type, with $\sum_t f_t = 1$ as a constraint. Then the redshift distribution of type t will be $p(z|t) = f_z^t$, and we will also denote

$$f_{tz} \equiv p(z,t) = p(z|t)p(t) = f_z^t f_t. \tag{4.1}$$

In this case, we are considering the sky is populated with galaxies of different (but finite) redshifts and phenotypes, where phenotypes specifiy a galaxy's noiseless, observed appearance. We assume there is some selection function s with the probability of a galaxy being selected, possibly depending on sky position, specified as a selection function $p(s|\theta)$. We will always assume that we know nothing about the non-selected galaxies, not even that they exist; the observed data are the positions $\boldsymbol{\theta}$ and features F of the selected galaxies. All galaxies of phenotype t observed under the same conditions are assumed to have the same selection function $p(s|t,\theta)$ and the same probability $p(F,s|t,\theta)$ of being selected and measured to have image features F. Finally, we will allow for some biasing function, $\mathcal{B}_z^t$, depending on both redshift and phenotype, to relate the galaxies' spatial distribution to the underlying density fluctuation δ_z. Now the selected galaxies can be considered as being a Poisson sampling of the following density field:

$$\rho(z,\theta,t|n,f,\mathcal{B},\boldsymbol{\delta}) = nf_{tz}\mathcal{B}_z^t\left[\delta_z(\theta)\right]p(s|t,\theta). \tag{4.2}$$

In that expression, the last two terms describe the spatial variation of the expected detection rate due to density fluctuations and variable observing conditions, respectively.

With knowledge of the survey noise properties and the noiseless appearance of phenotype t, we can determine the likelihood $p(F,s|t,\theta,z)$ of a galaxy of phenotype t at location θ, z being selected and measured with features F. Note this likelihood will not depend on z since the phenotype's observables are independent of z, by construction. Therefore, for each observed galaxy i and

phenotype t, we can assign a feature/selection likelihood

$$\mathcal{L}_{it} \equiv p(F_i, s|t_i, \theta_i). \tag{4.3}$$

This function will depend on the quality of the observations at sky position θ_i and the measurement and selection algorithms. We will assume that this likelihood is known *a priori*, e.g., by the result of analizing the injection of artificial copies of the phenotype into the real survey images (Suchyta et al., 2016).

Then the probability of selecting a set of galaxies $i \in \{1 \dots N\}$ at positions $\boldsymbol{\theta}$ with features F, types $t$and redshifts z takes the standard Poisson form:

$$p\left(F, \boldsymbol{\theta}, t, z|n, f, \mathcal{B}, \boldsymbol{\delta}\right) = \exp\left[-n \sum_t f_t A^t(f^t, \mathcal{B}^t, \boldsymbol{\delta})\right] \times \prod_i \mathcal{L}_{it} n f_{t_i z_i} \mathcal{B}_{z_i}\left[\delta_{z_i}(\theta_i)\right]. \tag{4.4}$$

The exponentiated quantity is, as required for Poisson distributions, the expected number of detections $\langle N \rangle$ for the entire sample, which can be determined from knowledge of the survey properties.

In order to provide the full generative model for the data we must also specify the process $p(\boldsymbol{\delta}|\pi_\delta)$ generating the stochastic density fluctuation fields given some hyperparameters π_δ. For instance, that could be a log-normal distribution where π_δ specifies the power spectrum. We also require priors $p(\mathcal{B})$ and $p(n)$, plus any prior information on $p(f)$ aside from the constraint that $\sum_{tz} f_{tz} = 1$.

4.2.2 Redshift inference

The principal quantity of interest is the underlying redshift distribution

$$n(z) = n \sum_t f_{tz}. \tag{4.5}$$

In most aplications of redshift inference we are only concerned with the shape, not the normalization, of $n(z)$, and therefore we will focus here on the fractions f, rather than n. In addition, in many applications it is also useful to know the individual redshifts of galaxies z, and in order to enable a Gibbs sampling scheme we will need to keep $\mathcal{B}$ and t as conditional variables. We can use Bayes' theorem to write down the posterior joint probability of these variables

Table 4.1: Summary of the notation used throughout this chapter.

F	galaxy set of observed features
t	galaxy phenotype (or simply type)
z	galaxy redshift
θ	galaxy angular position in the sky
s	indicator of successful detection/selection
$\mathcal{L}_{it}$	probability of measuring galaxy i with F_i given t
$F, t, z, \boldsymbol{\theta}$	set of properties for all galaxies in the sample
N	number of galaxies in the sample
N_t	number of types
N_z	number of redshifts
A	effective area of the survey for source detection
n	mean galaxy density per unit solid angle
$n(z)$	mean galaxy density per unit solid angle per z
$\delta_z(\theta)$	density fluctuation at a given z and θ
π_δ	density fluctuation field hyperparameters
$\boldsymbol{\delta}$	set of $\delta_z(\theta)$ for all redshifts and positions
b_z^t	linear galaxy bias for type t at redshift z
$\mathcal{B}$	set of b_z^t for all types and redshifts
f_{zt}	joint type and redshift probability $p(z, t)$
f	set of f_{zt} for all types and redshifts
N_{zt}	number of sources assigned to redshift z and type t
N	set of N_{zt} for all redshifts and types
M_{zt}	number of sources in the prior at redshift z and type t
M	set of M_{zt} for all redshifts and types
Δz	difference between the means of estimated and true $n(z)$'s

of interest:

$$p(f, z, \mathcal{B}, t|F, \boldsymbol{\theta}, \pi_\delta) \propto \int \mathrm{d}n\, \mathrm{d}\boldsymbol{\delta} \; p\left(F, \boldsymbol{\theta}, t, z|n, f, \mathcal{B}, \boldsymbol{\delta}\right) p(\boldsymbol{\delta}|\pi_\delta)\, p(n)\, p(f)\, p(\mathcal{B}). \tag{4.6}$$

We have already derived the first term under the integral in Equation Equation (4.4). In this chapter, as in SB18, we will work with the approximation that we can replace the stochastic density fluctuation $\delta_z(\theta)$ with some determinsitc estimator $\hat{\delta}_z(\theta)$ of the realization of the density fields in the generative probability of Equation (4.4). Under that approximation, and performing the marginalization over n assuming the effective area of the survey is independent of phenotype (see SB18 for the details), the posterior distribution for redshift and phenotype information in Equation (4.6) becomes

$$p(f, z, t, \mathcal{B}|F, \boldsymbol{\theta}) \propto p(f) \prod_i \mathcal{L}_{it_i}\, f_{t_i z_i}\, \mathcal{B}_{z_i} \left(\hat{\delta}_{iz_i}\right), \tag{4.7}$$

$$\hat{\delta}_{iz} \equiv \hat{\delta}_z(\theta_i). \tag{4.8}$$

The roles of the main three sources of information in redshift estimation are clearly present and differentiated in the posterior of Equation (4.7). First, there is a term for the prior probability that any galaxy is of phenotype t and redshift z, f_{tz}. Second, the photometric information for a galaxy is in $\mathcal{L}_{it}$, which is the likelihood of galaxy i resembling phenotype t and passing selection. Third, clustering information enters as the last term, describing the modification of the probability by (our estimator for) the density fluctuation field.

4.2.3 Realistic set up: SOM implementation

For the data implementation of this method to a general galaxy survey, we propose to use a combination of wide and deep survey observations and *self-organizing maps* (SOMs). Deep observations are often available for surveys like the Dark Energy Survey via SNe observations, and these provide essentially noiseless photometric measurements and observations in additional filter bands for galaxies in specific fields (henceforth deep fields, or simply DFs), and provide an empirical sampling of the distribution of galaxies in feature (F) space. In turn, SOMs provide a data-driven way of mapping and discretizing that feature space, so that we can use the cells in a SOM trained in deep photometry as the definition of our galaxy phenotypes t. In that scheme, each cell

c of the so-called *deep* SOM cell constitute a phenotype t.

Another term that we will need in the data application is the noise or measurement likelihood, $\mathcal{L}_{it} \equiv p(F_i, s|t_i, \theta_i)$. We follow the approach of Buchs et al. (2019) and construct the measurement likelihood by training another SOM on wide-field data of the galaxy survey of interest; we will refer to this one as the *wide* SOM and its cells, $\hat{c}$, map the space of wide-field observations in a survey (every galaxy observed will be assigned to a *wide* cell, $\hat{c}$). Crucially, it is possible to inject artificial copies of galaxies with deep photometry, and hence well specified phenotypes, into the real images of the survey, and measure their wide-field properties (Suchyta et al., 2016). Then, for a set of galaxies, we will know both the cells in the deep and wide SOMs ($\hat{c}$ and c), so that we can construct the mapping between deep and wide SOMs which corresponds to our measurement likelihood:

$$\mathcal{L}_{it} \equiv p(F_i, s|t_i, \theta_i) = p(\hat{c}_i, s|c_i, \theta_i) \tag{4.9}$$

One other major part in the application of the method to data is the addition of clustering information, that is, the construction of the density field estimator using a tracer opulation and the treatment of a biasing function relating that estimate to the true underlying density fluctuation field of the selected galaxies. This part will be treated in full detail in Section 4.

4.3 Simulations

We now present a demonstration of the methodology on simulations of galaxy survey data. In the previous work about this method (SB18), a demonstration was performed on a simple toy-model simulation with idealized galaxy properties and noise distributions, and perfect knowledge of the density fluctuation field. Now, instead, we test our methodology on the public MICE2 simulation[1], a mock galaxy catalog created from a lightcone of a dark matter only N-body simulation that contains ∼200 million galaxies over an octant (~ 5000 deg^2) and up to $z = 1.4$. Several important differences with respect to the toy model simulation used in SB18 make this analysis more realistic and allow the method described in here to be applicable to real data analysis.

On one hand, the MICE2 simulation has realistic clustering properties given by a ΛCDM cosmology wth parameters $\Omega_m = 0.25$, $\Omega_b = 0.044$, $h = 0.7$, $n_s = 0.95$, $\Omega_\Lambda = 0.75$, $\sigma_8 = 0.8$ and $w = -1$. In addition, we do not assume true knowledge of the density field but rather infer the clustering information

[1]The data can be downloaded from CosmoHub (Carretero et al., 2017), `https://cosmohub.pic.es/`.

from a set of galaxy tracers, described below. On the other hand, galaxies have realistic spectral energy distributions (SED) assigned from the COSMOS catalog (Ilbert et al. 2009) that preserve the observed color-magnitude distribution as well as clustering observations as a function of colors and luminosity (see Crocce et al. (2015) for more details). Once the galaxy SED is known, mangnitudes are computed based on the luminosity and redshift of the galaxy. The galaxy properties, clustering and lensing in the simulation have been thoroughly validated in Carretero et al. (2015); Fosalba et al. (2015a); Crocce et al. (2015); Fosalba et al. (2015b).

4.3.1 Sample selection

In particular, we select a galaxy sample with a square footprint defined by the cuts $30 \leq \mathrm{RA[deg]} \leq 60$ and $0 \leq \mathrm{Dec[deg]} \leq 35$, representing an area of around 1000 sq. deg., with the redshift range $0.2 \leq z \leq 1.2$ and a magnitude limit $i_{\mathrm{DES}} < 24$. For computing efficiency in this test, we downsample the true galaxy SEDs to simplify the color space and obtain an equally tight color-redshift relation without the need of a higher resolution deep SOM. This procedure removes $\sim 50\%$ of the sample and allows us to significantly reduce the runtime. The downsampling keeps a representative amount of populations (Elliptical, Spiral, Starburst) and dust attenuation laws and values present in the simulation.[2]

The tracer sample is randomly drawn from the full population to have a constant comoving density similar to the RedMagic DES Y1 galaxy sample in the first three lens bins (Elvin-Poole et al. 2018). This is achieved by taking the number of galaxies and redshift range of the third redshift lens bin, assuming it has a top-hat redshift distribution, and converting it into a comoving density assuming MICE cosmology. The target sample is defined as the galaxies that are not selected as part of the tracer sample. The upper panel of Fig. 4.1 shows the redshift distributions of both samples. The redshift binning is chosen to have 20 bins equally spaced in comoving distance between the redshift limits of the catalog, which makes the tracer sample to also have a constant angular density.

4.3.2 The phenotype approach: Deep and wide SOMs

The phenotype method described in section 4.2 is then applied to the simulation. As stated in §4.2.2, the approach can benefit from a sample with

[2]The selection is defined as sed_cos $\equiv c \in 0, 1, 2, 5, 6, 7, 10, 11, 12,$ $15, 16, 17, 21, 22, 23, 24, 25, 29, 30, 35, 36, 37, 38, 39, 41, 42, 43.$

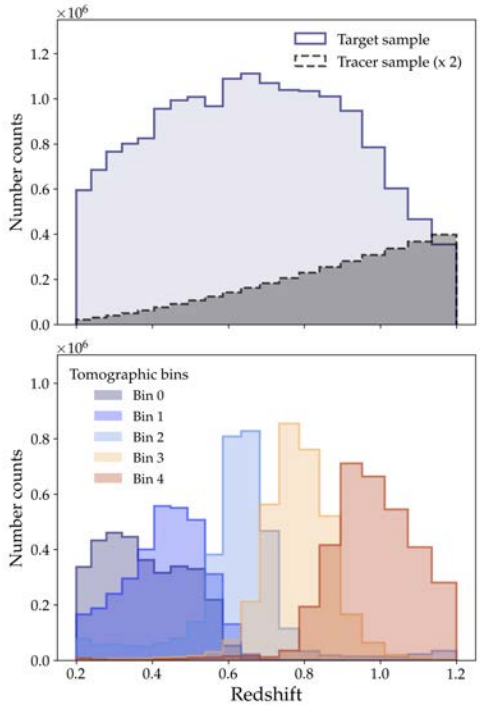

Figure 4.1: *(Upper panel:)* Redshift distribution of the target and tracer samples. The target sample contains the galaxies for which we want to find a redshift distribution. The tracer sample contains galaxies with known redshifts that are used to add the clustering information into the redshift estimation. *(Lower panel:)* Redshift distribution of tomographic bins defined as in §4.3.3.

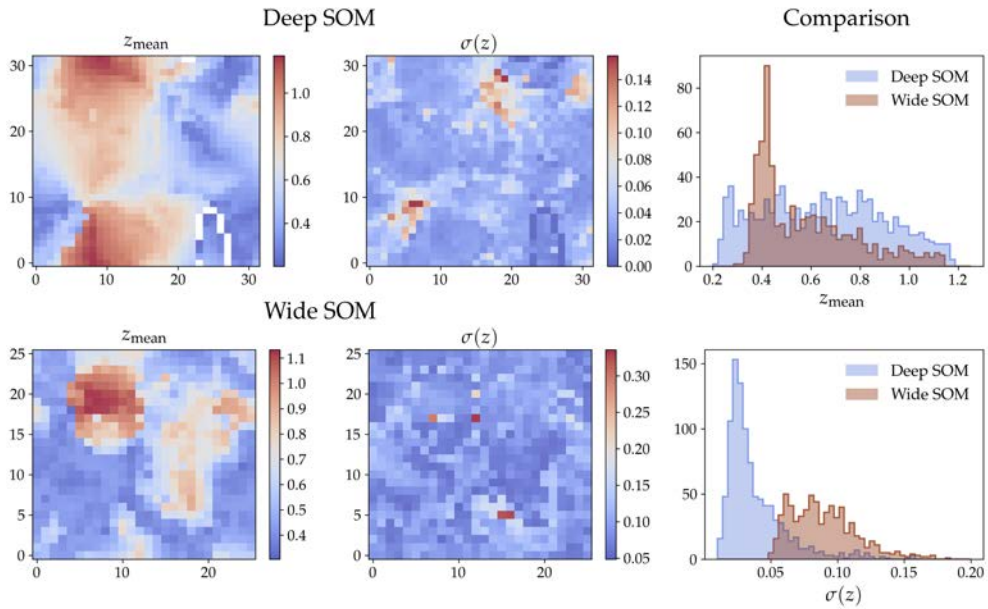

Figure 4.2: Mean redshift and redshift dispersion of cells in deep and wide SOMs described in §4.3.2. The left and central columns show the SOM maps populated with these quantities, while the plots in the right column show the comparison of these distributions. These show how the deep SOM better samples the redshift space of the simulation test, with a lower redshift scatter per cell.

deep photometry and extra bands to define galaxy phenotypes, and then it can have a wide sample with noisier photometry and only a subset of optical bands as observations. We choose among the available bands in MICE2 the DES g,r,i,z bands for both samples and the additional CFHT u, DES Y and VHS J,H,K bands for the deep sample. For the deep sample, we stick to noiseless true fluxes from the simulation, while for the wide sample we add Gaussian noise to the fluxes by fitting a linear relation between magnitude and logarithmic magnitude error for each band using observed noise from the Dark Energy Survey Year 1 public data[3]. We produce deep and wide photometries for all galaxies of the target sample. We finally select only galaxies that have a signal to noise above 5 in each wide band, g, r, i, z. To avoid throwing away too many galaxies, we have shifted by -1.2 magnitudes every galaxy.

Following §4.2.2, we create two self-organizing maps (SOMs) on a squared grid with periodic boundary conditions, similar to the SOM in Masters et al. (2015). The deep SOM is trained with eight colors, defined as $\mathrm{mag} - i$, where $\mathrm{mag} = \{u, g, r, z, Y, J, H, K\}$, in a 32×32 grid. The wide SOM is trained with one magnitude, i, and three colors, $g-i$, $r-i$ and $z-i$, in a 26×26 grid. Each color is renormalized to span the range $[0, 1]$, while the magnitude spans the range $[0, 0.1]$, i.e., , we give colors 10 more weight than magnitudes in creating the wide SOM. Also, to train the wide SOM we only use galaxies with a S/N above 10 in all bands to avoid noise in the map. Once both maps are trained, every galaxy in the target sample gets a best matching unit (bmu) for both the deep and wide SOMs, which is used to accurately calibrate the likelihood probability between the SOMs, $P(c, \hat{c})$, where c and $\hat{c}$ are the labels of the deep and wide SOM cells, respectively.

Figure 4.2 shows the mean redshift and redshift dispersion of the cells in the deep and wide SOMs described above (left and central columns). From the plots, one can note the higher redshift smoothness and the lower redshift dispersion in the deep SOM compared to the wide SOM. This is even more evident from the comparison plots in the right column of Fig. 4.2: the distribution of the mean redshift per cell in the deep SOM is more uniform and samples better the redshift space of the simulation ($0.2 < z < 1.2$), and the redshift dispersion per cell in the deep SOM is significantly lower (median $\sigma(z)$ of 0.030 for the deep SOM vs. 0.086 for the wide SOM).

4.3.3 Tomographic bins

To create tomographic redshift bins we group wide SOM cells together by their mean redshift value. First, we find the redshift probability distribution

[3] `https://des.ncsa.illinois.edu/releases/y1a1/key-catalogs/key-mof`

of each wide SOM cell by summing over the deep SOM cells,

$$P(z|\hat{c}) = \sum_c P(z|c)P(c|\hat{c}), \tag{4.10}$$

where $P(z|c)$ is also estimated using all galaxies in the target sample. We sort the wide SOM cells by their $z_{\rm mean} = \int dz\, z\, P(z|\hat{c})$, and group them creating 5 redshift bins with equal number of galaxies. The true redshift distribution of each tomographic bin is shown in the lower panel of Fig. 4.1. The method presented in this work can be applied to an arbitrary galaxy selection, including any of these tomographic bins. For this work, however, we will focus on Bin 3 and ignore the others. For this reason and to avoid any selection effects in our probabilities, we retrain the wide and deep SOMs using only galaxies from this bin. This choice does not affect the methodology of the work presented here, but makes the results more concise and clear.

4.4 Adding the clustering information

Clustering information can be used in redshift inference when having an overlapping population of galaxies with reliably known redshifts, either from spectroscopic or high-quality photo-z estimates (we call that the tracer population). Section 4.2 describes the formalism used in this work, where clustering information enters in the form of the density field from which galaxies are assumed to be drawn from. We will work under the approximation that we can replace that density field with a deterministic density field estimate discretized in redshift space as a set of redshift bins, $\hat{\delta}_z(\theta)$, obtained from a tracer population that has been drawn from the same generative model up to some biasing relation.

4.4.1 Density estimation

Now we describe the way we estimate the density field in this work. The tracer population, described in 4.3.1, is split in 20 redshift bins equally spaced in comoving distance in the range $z \in [0.2, 1.2]$ using the true redshift from the simulation. The redshift bins are wider than the typical redshift uncertainty of photometric survey tracers, $\sigma_z \sim 0.015(1+z)$ (Rozo et al. 2016; Vakili et al. 2019)), and also wide enough to make them nearly independent from each other. We will defer any attempt to include photo-z errors in the tracer sample to a future study.

Several methods exist to reconstruct the surface density of galaxies centered at the position of those galaxies (e.g. see Cautun & van de Weygaert

(2011); Darvish et al. (2015)). In this work we will use a kernel density estimator (KDE) to estimate the density field at any position of the field, based on counting weighted pairs from a tracer sample and a random sample that describes the selection function of the tracer sample, in the same spirit as conventional clustering$-z$ redshift techniques. We assume that our density field estimate only depends on the distance between a galaxy and tracers (θ_{xT}) and random points (θ_{xR}) and define it as

$$\hat{\delta}_z(x) = \frac{\frac{1}{n_T}\sum_T K(\theta_{xT})}{\frac{1}{n_R}\sum_R K(\theta_{xR})} - 1, \tag{4.11}$$

where n_T is the total number of tracers and n_R is the total number of unclustered random points. We only use tracers from the same redshift bin since we assume each redshift bin is independent. We presume $n_R \gg n_T$ such that the dividing term can be considered a measure of the area surrounding x, taking into account the selection function and mask effects.

Choosing the shape and extent of the KDE is important. Figure 4.3 shows the effect that different KDE shapes has on the field estimate. The top left panel shows a top hat KDE of size $r_{\max} = 30$Mpc. Such a large kernel smooths the density field too much and cannot resolve well large structures, overestimating the density in underdense regions and underestimating the density in crowded places. On the contrary, the top right panel shows a small top hat KDE with $r_{\max} = 3$Mpc. This KDE can resolve better denser structures, although it will still underestimate high density regions, is more affected by shot noise and has no information about the surrounding environment, which can lead to high overdense estimations in underdense regions.

Many cosmological applications of redshift inference will also include the tracer sample as part of their analysis. Allowing large scales in the density estimation can improve its estimation, but will also correlate with the observables of such analyses, and using only very small scales (< 3Mpc) lowers the signal to noise of the estimation and can create noise biases. We set by default $r_{\max} = 15$Mpc, which is a good compromise between both extremes, although we also explore $r_{\max} = 10$Mpc for comparison (cosmological analyses of galaxy surveys such as DES set their minimum scales around 8-12 Mpc). High density structures are smaller than 15Mpc, so a flat kernel of this size would still smooth them, keeping us from accessing most of the information from the tracer sample. The bottom left panel shows a power law KDE $\propto r^{-0.8}$, inspired in the shape of galaxy correlation functions, which obtains information from smaller scales by weighting them more aggressively but still keeps the knowledge of its environment.

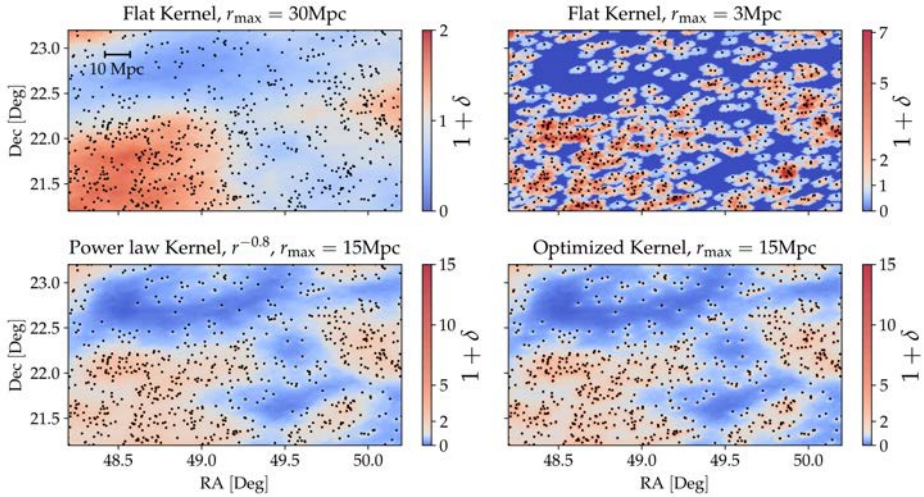

Figure 4.3: Density field estimation using different kernel density estimators from a tracer sample population. Shows the field estimate for a small patch in the highest redshift bin. The black dots show the position of the tracer galaxies, and the background colors show the estimated value of the density field at different positions. The top panels show a flat kernel with a large size ($r_{\rm max} = 30$Mpc, left) and a small size ($r_{\rm max} = 3$Mpc, right). The bottom left panel shows the density with a power law kernel that resolves better the structures. The bottom right panel shows a field estimated with an optimized kernel, which is our default density field estimate.

4.4.2 Biasing relation

Clearly, different KDE shapes deliver different density field estimates. One way to find if a recovered density field is on average an unbiased estimate of a particular target sample is to estimate the relative abundance of target galaxies and randomly distributed points at each value of the field. When we estimate $\hat{\delta} \sim 0$, we expect to find about the same amount of target galaxies and random points (relative to the length of their catalogs), while for $\hat{\delta} \sim 1$ we would expect to find twice more targets than randoms. Figure 4.5 shows, for each redshift bin (color coded), the relation between the average abundance of targets over randoms as a function of a density field estimated with a power law KDE with $r_{\max} = 10$Mpc. If the KDE estimate delivered a perfectly unbiased field estimation this would yield the dashed line. In general, the KDE estimate will not deliver such an estimate, both because the KDE yields a biased estimate and because the tracer will be a biased tracer of the target galaxies. However, we will assume such relation exists, i.e. there is a functional form or mapping $\mathcal{B}$ that relates the estimated KDE field $\hat{\delta}_z$ from the tracers and the true probability $p(\theta|z)$ of a target galaxy being at position θ_i and redshift z,

$$p(\theta_i|z) = \mathcal{B}(\hat{\delta}_z(\theta_i), \{b_i^z\}), \tag{4.12}$$

where $\{b_i^z\}$ are the parameters of $\mathcal{B}$ at redshift z. This is an approximation of a more general approach where the density field is updated locally by the targets as part of the hierarchical model. The parameters $\{b_i^z\}$ are part of the framework parameters (see section 4.2) and they will be sampled along with the other parameters in the HBM (see section 4.5). We choose a polynomial of degree four as our mapping function $\mathcal{B}$, such that

$$\log_{10}(p(\theta|z)) = \sum_{k=0}^{4} b_k^z \ \log_{10}(1.1 + \hat{\delta}_z)^k, \tag{4.13}$$

with the additional constraints that $\int p(\theta|z)d\theta = 1$ and that the derivative must always be positive. Note the arbitrary choice of coordinates in $(1.1 + \hat{\delta}_z)$, instead of the typical $(1 + \hat{\delta}_z)$, to avoid problems with the logarithm.

4.4.3 Optimizing the estimator

We can go one step further and try to optimize the shape of the KDE assuming we have a small calibration patch where the redshift of the target galaxies is known. For this purpose we define a KDE with shape

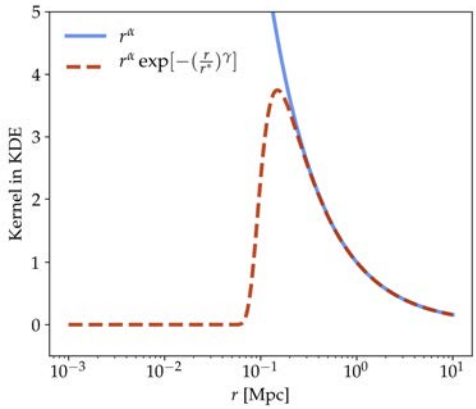

Figure 4.4: Comparison between a power law KDE and a KDE with a power law that truncates at some scale r^*. Such truncation reduces the impact of shot noise in smaller scales and naturally adds a small exclusion region around the positions of tracers.

$$\text{KDE} \propto r^\alpha \exp\left[-\left(\frac{r}{r^*}\right)^\gamma\right], \tag{4.14}$$

which combines a power law with exponent α and an exponential truncation of the power law at scale r^* with width parameter γ. Figure 4.4 compares this KDE shape with a power law. The motivation for allowing a truncation at small scales is to minimize the effect of shot noise, since tracer samples with lower density will sparsely populate smaller scales. Underdense regions will suffer from this effect at larger scales, while really overdense regions would not have this problem until really small scales. However, we will not try to optimize the KDE shape as a function of environment, but rather find an average optimal KDE shape. Moreover, the truncation naturally adds an exclusion region around the position of a known galaxy from the tracer sample.

The optimization of the KDE works as follows. We write the probability of the optimized KDE parameters for redshift z as

$$p(\alpha_z, r_z^*, \gamma_z|\theta, z) \propto p(\theta|z, \alpha_z, r_z^*, \gamma_z)\ p(\alpha_z, r_z^*, \gamma_z|z) \tag{4.15}$$

where the last term is the prior on the parameters. Given a sample of targets with known redshifts from a calibration field,

$$p(\theta|z, \alpha_z, r_z^*, \gamma_z) \propto \prod_{i \in z} p(\theta_i|z, \alpha_z, r_z^*, \gamma_z) \tag{4.16}$$

where $p(\theta_i|z, \alpha_z, r_z^*, \gamma_z)$ is the probability of the ith galaxy at redshift z from the calibration sample. We obtain this probability by correcting the KDE

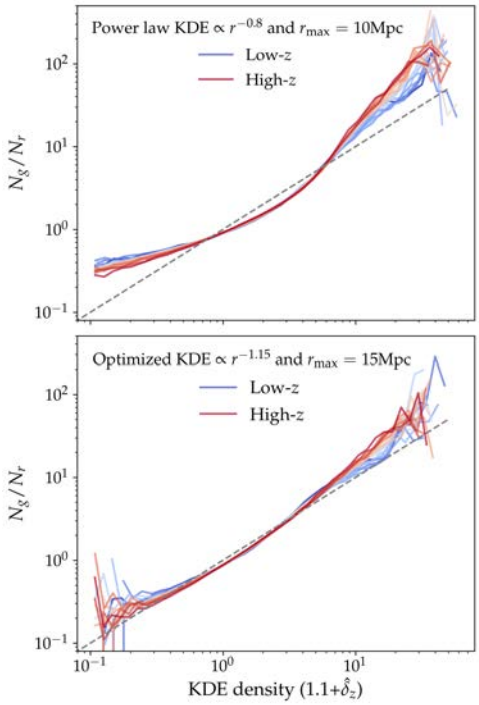

Figure 4.5: *(Upper panel):* Ratio between the abundance of target galaxies and random points as a function of estimated KDE density, for a power law KDE $r \propto r^{-0.8}$ and $r_{\rm max} =$ 10Mpc. The different redshift bins are color coded. If the KDE delivered a perfectly unbiased field estimate of the target galaxies, we would expect to find the dashed line relation. All galaxies have been used without tomographic bin selection to obtain a better estimate. The true redshift of all the target galaxies was used, while in a real data scenario one could only estimate this relation in the smaller calibration fields. *(Lower panel):* Same as upper panel, but using an optimized KDE with $r_{\rm max} = 15$Mpc. The KDE is optimized from a function that combines a power law and an exponential truncation at small scales to deal with shot noise effects (see Fig. 4.4). The optimal parameters are found from a calibration field from $\sim 3.5\text{deg}^2$ where redshifts for the target galaxies are known. It shows a more linear relation, closer to the perfect estimate (dashed line), since the area is larger and the KDE has been optimized.

estimate $\hat{\delta}_z(\theta_i)$ using an abundance vs KDE density relation similar to Fig. 4.5 but estimated using only the galaxies from the calibration sample.

We use a small patch of $\sim 3.5\text{deg}^2$ as the calibration field. We take the average of the parameters across all redshift bins as an estimate for the optimized KDE, since the constraining power in each redshift bin is small. We use top hat priors defined in the range $\alpha_z \in [-2, -0.5]$, $r_z^* \in [0.001, 0.1]$ and $\gamma_z \in [-10, -2]$. We find that we are not very sensitive to the γ parameter so we fix it to its mean value of $\gamma = -4$ and run again. For a KDE limited to $r_{\max} = 15\text{Mpc}$ we find $< \alpha_z >= -1.15$ and $< r_z^* >= 0.018\text{Mpc}$. For a KDE limited to $r_{\max} = 10\text{Mpc}$ we find $< \alpha_z >= -1.0$ and $< r_z^* >= 0.010\text{Mpc}$. Note how a more aggressive power law is preferred when the size of the KDE is larger. Since we have more area we want to avoid increasing the smoothing of structures, and at the same time we are allowed to weight more close by pairs since we have more tracers at reach from which to get the density estimate. Figure 4.5 (lower panel) shows the mapping relation (using again all galaxies) for an optimized kernel with $r_{\max} = 15\text{Mpc}$, which compared to the upper panel is much more close to the ideal relation. This is both consequence of having a larger area (2.25 times more area) and having optimized the KDE. The bottom right panel of Fig. 4.3 shows the density field estimated with an optimized KDE with $r_{\max} = 15\text{Mpc}$, which is our default density field.

The shape of the optimal KDE and the shape of the mapping relation (Fig. 4.5) depend on the tracer sample angular density, among other factors. Here we choose a tracer sample with constant angular density, which minimizes the variation from this effect across redshift. In general, a low density tracer sample will result in an understimation of underdense regions, which by definition are already sparse, and will reduce the smaller scale we can include in the kernel. In the limit where the tracer density was very low there would be a range of low $\hat{\delta}$ with a very flat mapping relation near ~ 1, which is the value the probability $p(\theta|z)$ takes in absence of density information. Such a tracer population would still add information on the higher density regions that it would be able to sample. In our case we assume a density similar to the lower luminosity redshift RedMagic bins (see section 4.3), which is enough to sample relatively well most of the density regions, since we see Fig. 4.5 shows relatively linear relation for most $\hat{\delta}$ values that is not too far from the ideal relation.

4.5 Sampling

Now we turn to the problem of sampling over the redshift and type probability distributions of populations of galaxies and their individual constituents, in the framework of the hierarchical Bayesian model described in the previous sections. It is complicated to simultaneously sample all variables from the joint posterior $p(f, z, t, \mathcal{B}|F, \boldsymbol{\theta})$ in Equation (Equation (4.7)). We will show, however, that it is feasible to draw samples from this posterior using a three-step Gibbs sampler, because the conditional posterior distributions of interest can be easily written and sampled. In the previous work about this method (SB18), the true values of the density field at each position were known, and hence there was no need to sample over the biasing function relating the estimated density field with the truth, making it possible to perform the sampling with a two-step Gibbs sampler. In this work, we will not make that simplification, and we will sample, and marginalize over, the biasing relation between the estimated and true density fields, studied in detail in Section 4.4. This will require the full three-step Gibbs sampler, and it will demonstrate the full implementation of the sampling method, including the development of some key sampling strategies that will enable a future application to real data.

In performing the sampling described above, we will be using information from all galaxies in the sample to constrain the redshift and type probability distributions of the galaxy population. Additionally, the fully Bayesian nature of this scheme allows us to make use of prior information on the relevant quantities, when available. In this work, we will assume that we have access to a subset of galaxies with known z, t, e.g., from a complete spectroscopic survey of a random subsample of targets in a limited, small region of the sky. We will assume such subsample contains a tracer population, with the same selection as the corresponding tracers in the full sample. As we will describe next, this subsample will pose an informative prior on the coefficients f, and it will also be important in sampling over the biasing function parameters described in 4.4.

4.5.1 Three-step Gibbs sampler

Each iteration of the Gibbs sampler comprises three steps which are (i) drawing a sample of f from $p(f|z, t, \mathcal{B}, F, \boldsymbol{\theta})$, (ii) drawing pairs of z_i, t_i for each galaxy i from $p(z_i, t_i|f, \mathcal{B}, F_i, \theta_i)$ using the newly drawn f, and (iii) drawing a sample of the biasing functions $\mathcal{B}$ for each redshift bin given the z_i assignments in step (ii). The conditional distributions can be derived from the joint distribution in Equation (4.7), but next we detail the expressions used in

each of the three steps. The first two steps of the sampler are as in SB18 and hence we will skip the full derivation for brevity (see SB18 for more details), and the third step is new and will be considered in more detail.

(i) The conditional posterior on f depends on the counts of sources of z and t (in the last iteration), $N = \{N_{zt}\}$ where N_{zt} is the number of sources assigned to redshift z and phenotype t, and it also depends on the prior information on f, $p(f)$:

$$p(f|z,t,\mathcal{B},F,\boldsymbol{\theta}) \propto p(f)\prod_{z,t} f_{zt}^{N_{zt}}. \tag{4.17}$$

The prior condition that $\sum f_{zt} = 1$, and $0 \leq f_{zt} \leq 1$, allows us to write the conditional posterior on f as a Dirichlet distribution. Following the derivation in SB18, if $M = \{M_{zt}\}$ are the counts of the prior sample found at each z,t pair, then the prior distribution of f follows a Dirichlet distribution with parameters M, and hence the conditional posterior follows a Dirichlet on the data counts from the last iteration plus the prior counts:

$$p(f|N) \sim \mathrm{Dir}(N+M), \tag{4.18}$$

$$\begin{aligned}\text{with } \mathrm{Dir}(N) \equiv (N+N_zN_t-1)!\,\delta_D\left(1-\sum_{zt} f_{zt}\right) \\ \times \prod_{z=1}^{N_z}\prod_{t=1}^{N_t}\frac{\Theta(f_{zt})f_{zt}^{n_{zt}}}{n_{zt}!}\end{aligned} \tag{4.19}$$

(ii) For each galaxy, the posterior for the z_i, t_i pair conditioned on $f$and $\mathcal{B}$is

$$p(z_i,t_i|f,\mathcal{B},F_i,\theta_i) \propto \mathcal{L}_{it_i}\, f_{t_iz_i}\, \mathcal{B}_z\left(\hat{\delta}_{iz_i}\right) \tag{4.20}$$

where apart from using the f obtained in the first step of the sampler (i), we make use of the measurement likelihood and the clustering terms discussed above. The sampling in this step (ii) will produce pairs of z,t for each galaxy that constitute the next realization of $N = \{N_{zt}\}$, to be used in the step (i) of the next iteration of the Gibbs sampler.

(iii) After we have z assignments for all galaxies in the sample from step (ii), we can now separate galaxies into redshift bins according to those assignments. Then, for each redshift bin, the posterior on the biasing

function of that bin conditioned on all other variables looks like:

$$p(\mathcal{B}_z|f,z,t,F,\boldsymbol{\theta}) = p(\mathcal{B}_z|z,\boldsymbol{\theta}) \propto \prod_{z_g\in z}\left(\mathcal{B}_z(\hat{\delta}_{iz_i}(\theta_g))\right). \tag{4.21}$$

According to this expression, different biasing function parameters, as described in Equation (4.13), will yield different posterior probability for each Gibbs step. In order to explore the biasing functions parameter space and yield samples of this posterior, we use a Markov Chain Monte Carlo (MCMC) sampler for each redshift bin and each step of the Gibbs sampler. In each case, we run 1000 iterations of the MCMC, and the proposal distribution for each case is given by the steps of an MCMC chain run on the prior sample described above in this section. Using the prior chain steps with equal probability on the proposal distribution effectively uses the prior information and enables an informed sampling without the need to blindly tweak the proposal distribution or the parameter limits of the MCMC chains.

4.6 Results

Now we present the results for the third tomographic bin which contains $\sim 3.3\times10^6$ objects. We define a calibration sample where redshift and type are known from one healpy pixel with nside=2^5, which has an area of $\sim 3.5\text{deg}^2$. We apply the same tomographic bin selection to the calibration sample, which leaves a total of 10758 objects. These objects are used to estimate the prior probability $p(z,c)$ and obtain the sampled prior on the mapping function parameters $\mathcal{B}(\{b_i\})$ (see section 4.5 for details about the sampling).

The HBM method yields samples of the individual redshift and type posterior for each galaxy, the redshift and type posterior of the population and samples of the posterior of the mapping function parameters. In this work we focus on the redshift population posterior, marginalizing over all other parameters, since this is what is usually used in photometric cosmological analysis. Current and future weak lensing analysis are very sensitive to small biases in the mean redshift of the distribution, which can become the leading systematic uncertainty. Therefore, we define as a metric the difference between the mean of each sample j of our redshift posterior and the true mean from all the target galaxies,

$$\Delta z_j = \langle z_{\text{est,j}}\rangle - \langle z_{\text{true}}\rangle\,. \tag{4.22}$$

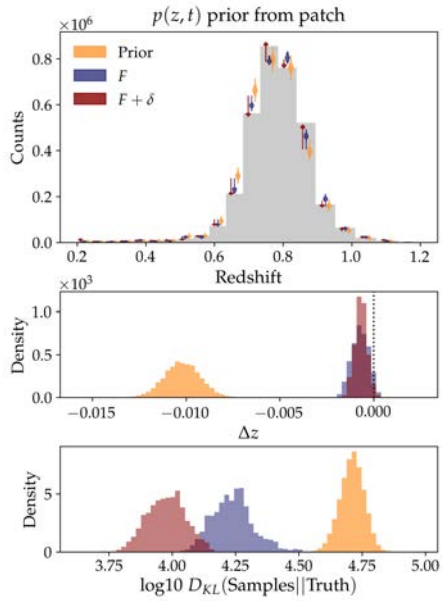

Figure 4.6: Posterior redshift probability distribution, marginalized over type and when including clustering marginalizing over mapping function parameters. The prior is obtained from a small calibration patch with 10758 objects over an area of $\sim 3.5\text{deg}^2$. The distributions are obtained from an HBM with photometry only, F, an HBM with photometry and clustering $F + \delta$ and a distribution from samples drawn from the prior. *Panel:* Shows violin plots for each distribution compared to the true redshift distribution. *Panel:* Shows the posterior distribution on redshift bias Δ_z values. *Panel:* Shows the distribution of Kullback–Leibler divergence ($D_{\rm KL}$) between each sample and the true redshift distribution. The HBM (F) removes most of the redshift bias, since in this case the bias is tightly related to a bias in the type density $p(t)$, due to the sample variance of the calibration patch. The addition of clustering sharpens the distribution and improves the overall shape, reducing the $D_{\rm KL}$ divergence.

Since we draw samples of the full redshift posterior f_z another useful metric that is sensitive to the distribution shape is the Kullback–Leibler divergence ($D_{\rm KL}$) between each sample and the true redshift distribution,

$$D_{\rm KL}(f^{\rm est}_{z,j}||f^{\rm true}_z) = \sum_z f^{\rm est}_{z,j} \log\left(\frac{f^{\rm est}_{z,j}}{f^{\rm true}_z}\right) \tag{4.23}$$

This is a measurement of the relative entropy between the true distribution and the recovered distribution, and can be used to see how much information the photometry and density estimates are adding with respect to the prior knowledge. We will show results for samples drawn from the prior, from an HBM that only includes photometry information and from an HBM that includes both photometry and clustering information, marginalizing over the mapping

relation parameters. We denote the HBM with photometry as F (feature) and the HBM with photometry and clustering as $F + \delta$.

We show one case where the $p(z, t)$ probability comes from the calibration sample, and three cases where the prior is modified and biased. For each case, we show a violin plot of the posterior redshift distribution compared to the true distribution, the distribution of Δz_j differences, the distribution of $D_{\rm KL}$ divergences and a plot showing Δz_j as a function of iteration steps.

4.6.1 Prior $p(z, t)$ with sample variance

Since the prior $p(z, t)$ is estimated from a patch with a small angular size it has a larger sample variance than the full population, which changes the relative density of galaxies at given redshift and type. This effect is larger at lower redshifts, where the volume is smaller. Sample variance mostly affects the density of types $p(t)$, where $p(z, t) = p(z|t)p(t)$, since the same phenotype should yield the same redshift regardless of where it is observed. However, we have seen in Fig. 4.2 that there is a tail of deep cells where the redshift distribution is wider, mostly due to color-redshift degeneracies. As a result, the redshift distribution $p(z|t)$ of these cells is also affected by sample variance. The Dirichlet sampling of the prior neglects sample variance uncertainty, but we expect it to have a small impact on the results since the target population is much larger than the prior sample. Moreover, poor sampling in any of the cells can lead to a noise bias of $p(z|t)$.

Fig. 4.6 compares the posterior from running an HBM with photometry alone (F, yellow), an HBM with photometry and clustering ($F + \delta$, red) and samples drawn from the prior $p(z, t)$. The prior $p(z, t)$ has a mean redshift bias of $\Delta_z = (-1.0 \pm 0.1) \times 10^{-2}$, while when running the HBM we find a bias of $\Delta_z = (-6.1 \pm 4.2) \times 10^{-4}$ with photometry alone and a bias of $\Delta_z = (-6.7 \pm 3.2) \times 10^{-4}$ when adding clustering. An HBM with photometry alone can correct redshift biases that come from a biased type probability $p(t)$. Since sample variance mostly changes $p(t)$, having feature information is enough to remove most of the redshift bias. We find adding the clustering further tightens the Δ_z posterior distribution and also improves the shape of the redshift posterior, leading to a smaller $D_{\rm KL}$ divergence.

4.6.2 Prior $p(z, t)$ with a redshift bias

So far we have assumed our prior is perfectly estimated from the calibration patch. In this section we will add a systematic redshift bias for each type/deep cell $p(z|t)$. We multiply each conditional redshift probability by a

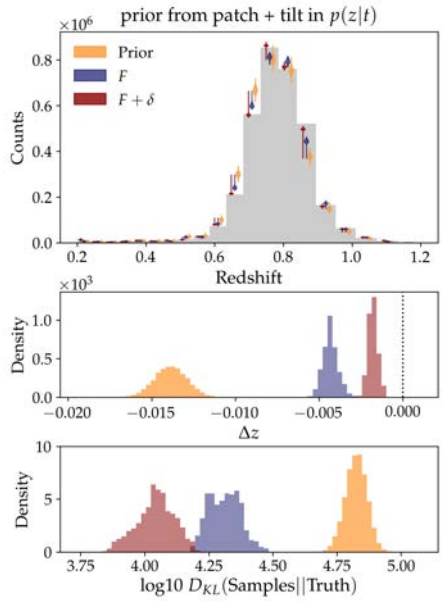

Figure 4.7: Similar to Fig. 4.6. The prior, which is obtained from the same small calibration patch, is systematically biased in the conditional redshift probability of each type $p(z|t)$ towards low redshift. This is achieved by weighting $p(z|t)$ with a straight line that has a value of 20 and 1 in the first and last redshift bin, respectively. The HBM with photometry alone reduces the redshift bias by the same amount as in Fig. 4.6, since it only corrects redshift biases produced by a bias in $p(t)$. The remaining bias can only be corrected with the addition of clustering, which further reduces this bias and improves the redshift posterior shape.

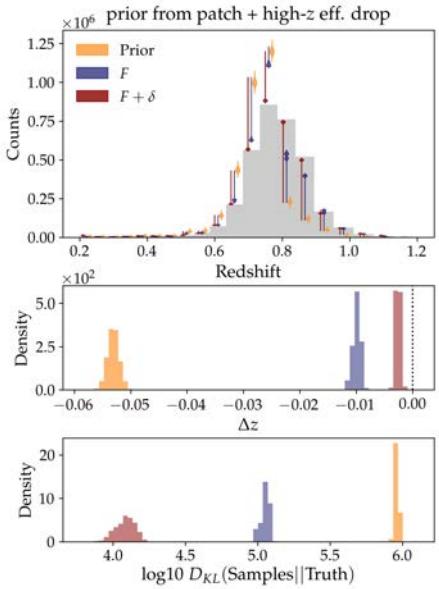

Figure 4.8: Similar to Fig. 4.6. The prior mimics an hypothetical spectroscopic efficiency drop above redshift $z > 0.8$ by reweighting the prior with a factor 0.2 in the last 7 redshift bins. The HBM with photometry is able to correct the redshift posterior in redshift bins far away from $z \sim 0.8$, where the drop happens, by changing the density of deep cells whose redshift probability $p(z|t)$ does not cross $z \sim 0.8$. It increasingly fails to correct the redshift distribution around $z \sim 0.8$ since it cannot modify $p(z|t)$. Adding clustering significantly improves the redshift distribution, removing most of the redshift bias and largely improving the redshift distribution shape.

straight line with a value of 20 and 1 in the first and last redshift bin, respectively. Therefore, the prior $p(z,t) = p(z|t)p(t)$ now has the effect from sample variance and a systematic bias towards low redshift.

Fig. 4.7 shows the HBM results for such prior, which has a mean redshift bias of $\Delta_z = (-1.4 \pm 0.1) \times 10^{-2}$. The HBM with only photometry has a mean posterior redshift bias of $\Delta_z = (-4.4 \pm 0.4) \times 10^{-3}$, while an HBM with photometry and clustering yields $\Delta_z = (-1.8 \pm 0.3) \times 10^{-3}$. Note how the HBM (F) has corrected the same amount of redshift bias as with the sample variance prior (~ 0.01 in redshift), which was related to a bias in the types, but cannot correct any of the systematic bias introduced in $p(z|t)$. The HBM $(F + \delta)$ can use the clustering information to further improve the $p(z|t)$ probability and reduce the total redshift bias. It also reduces the D_{KL} divergence, improving the overall shape.

4.6.3 Prior $p(z,t)$ with a redshift efficiency drop

Spectroscopic surveys usually present sharp selection effects in redshift due to their limited wavelength coverage of the spectra. Using such survey to estimate the prior probability can bias the whole posterior redshift distribution of the weak lensing samples. In this section we use a prior $p(z,t)$ from an hypothetical spectroscopic survey with an efficiency drop above redshift $z > 0.8$ (the last 7 redshift bins). We assume only 20% of the galaxies in the last 7 redshift bins have been measured, which we implement by multiplying by 0.2 the prior $p(z,t)$ in those bins.

Fig. 4.8 shows that this efficiency drop creates a huge redshift bias in the prior of $\Delta_z = (-5.3 \pm 0.1) \times 10^{-2}$. For the HBM ($F$) we find a redshift bias of $\Delta_z = (-9.9 \pm 0.4) \times 10^{-3}$, while for the HBM ($F + \delta$) we find a value of $\Delta_z = (-2.6 \pm 0.4) \times 10^{-3}$. The HBM ($F$) is able to successfully correct redshift bins which are far away from where the efficiency drop happens ($z \sim 0.8$) since there are many deep cells with a very tight redshift-type relation. However, it finds it increasingly hard to recover the redshift distribution closer to the drop, since it cannot update $p(z|t)$. Adding the clustering significantly improves the recovered shape, finding a much better $D_{\rm KL}$ divergence, and solves most of the redshift bias from the prior.

4.6.4 Prior $p(z,t)$ with degraded $z-t$ correlation and biased

So far we have assumed we have a calibration field with spectroscopic data that provides a tight redshift-color relation. Now we explore what happens if we loose this assumption and pretend the redshift information comes from a hypothetical photometric redshift sample. To this effect, we convolve the conditional redshift probability for each type $p(z|t)$ with a top hat function with size 7 redshift bins, which smooths the redshift probability. The median redshift dispersion of the deep cells goes from $\sigma(z) = 0.025$ to $\sigma(z) = 0.1$, significantly reducing the correlation between types and redshift. Furthermore, we add the same systematic redshift bias to each $p(z|t)$ than in section 4.6.2. Note the sample variance in $p(t)$ is left unchanged.

Fig. 4.9 shows the broadening effect in the prior, which now has a redshift bias of $\Delta_z = (-3.9 \pm 0.1) \times 10^{-2}$. The HBM with photometry alone, which can only modify the density of types, is barely able to change the redshift distribution since the correlation between redshift and type has been degraded, finding a redshift bias of $\Delta_z = (-3.15 \pm 0.07) \times 10^{-2}$, and a very similar $D_{\rm KL}$ divergence. In contrast, adding the clustering remarkably improves the redshift bias and shape, leading to a very large decrease in both $D_{\rm KL}$ and Δ_z metrics. In this case, we find a redshift bias of $\Delta_z = (-3.4 \pm 0.3) \times 10^{-3}$. This shows that

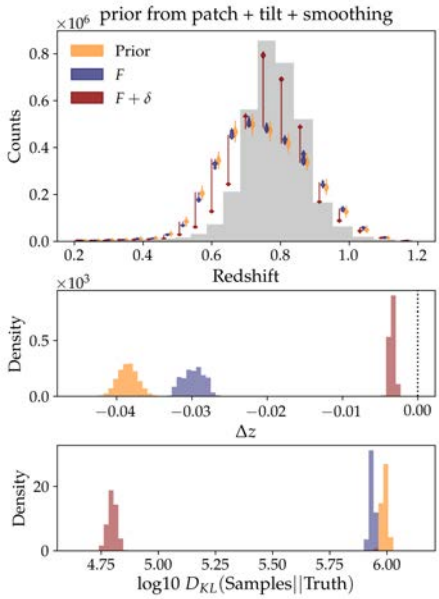

Figure 4.9: Similar to Fig. 4.6. The prior is smoothed by convolving $p(z|t)$ with a top hat function of size 7 redshift bins, increasing the median redshift dispersion of the deep cells goes from $\sigma(z) = 0.025$ to $\sigma(z) = 0.1$, which reduces the correlation between type and redshift for all deep cells. In this case the HBM with photometry alone can barely modify the redshift distribution, since there is little correlation between type and redshift. In contrast, adding the clustering information remarkably improves the redshift distribution recovery and reduces most of the redshift bias. This shows that photometric redshift surveys with wider $p(z|t)$ estimation can be used instead of spectroscopic surveys when clustering is available.

photometric redshift estimates can be more safely used instead of spectroscopic measurements when the clustering is also available, even if such estimates are not so precise and are systematically biased.

4.7 Discussion

Sánchez & Bernstein (2018) presented a hierarchical Bayesian model which can naturally combine the three main sources of information for estimating the redshift probability distributions of galaxies and samples of galaxies in a wide-field survey. These three main sources of information are, namely, prior information, which comes from a subset of galaxies with well measured photometric and (typically) spectroscopic properties, broad-band photometry for the galaxies in the wide-field sample, and the clustering of such against a tracer population with precise and well-characterized redshift estimates. All these sources of information have been used separately in the past, but this is the first method to combine them in a unified and consistent way. In SB18, the method was demonstrated on a simple set of simulations, and was lacking some important pieces that are needed for its application to real data.

In this work, we have expanded the method presented in Sánchez & Bernstein (2018) to include the additional methods needed for its application to a galaxy survey. The method consists of a hierarchical Bayesian model which enables the joint sampling of the redshift distribution of a galaxy sample and its individual constituents. Assuming that the galaxies come from a Poisson sampling of a density field, we propose to characterize that field using a kernel density estimator from the positions of a tracer galaxy population. This approximation enables a sampling process but requires to define a biasing relation, generally redshift dependent, to relate the tracer's spatial distribution to the underlying galaxy density fluctuation. We have detailed here how such biasing function can be constructed, and how we can sample and marginalize over it using prior information.

As opposed to the simple set of simulations used in SB18, we have now tested the methodology on the public MICE2 simulation, a mock galaxy catalog created from a lightcone of a dark matter only N-body simulation with around 200 million galaxies over an octant of the sky. This simulation features realistic galaxy clustering, which we characterize using a limited tracer population, and realistic galaxy properties following observations in the COSMOS catalog. In particular, the realistic set up of the simulations allows us to work in a set up where we can fully employ the phenotype approach proposed in SB18. Under that approach, we assume we have a sample with deep photometry

and extra bands to define galaxy phenotypes, and a wide sample with noisier photometry and only a subset of optical bands as observations. We use two self-organizing maps (SOMs) to characterize the properties of these samples, and we use galaxies with best matching cells in both SOMs to accurately calibrate the likelihood probability that relates observations and phenotypes, as we would do in real data. With this scheme, we construct a set of tomographic bin selections, and choose one of them, at intermediate redshifts, to test the methodology.

In applying the method to the simulation, we always assume there is a small region of the sky (of about 3 sq. deg.) for which the galaxy properties, phenotype and redshift, are well known. We use this set of galaxies as a prior, both for the phenotype and redshift probability distribution and for the biasing function needed for the addition of clustering information from a tracer population. With this setup, we apply the methodology under different cases, comparing the results obtained with and without clustering information in the method and those from just the prior information. As metrics, we use the difference in the mean of the derived and true redshift distributions for the sample, as well as the Kullback-Leibler divergence, which measures the differences in the shapes of the true and recovered redshift distributions.

We perform different tests using different assumptions on the prior sample. When this is just coming from a small patch of the sky but with perfect knowledge, i.e., just including sample variance, the HBM method both with and without clustering information perform similarly well in terms of the mean redshift of the population. This is expected, also consistent with SB18, as sample variance mostly changes the phenotype distribution, and that can be recovered in the HBM without the need of clustering information. The shape of the redshift distribution, however, is better recovered when using clustering information. If a redshift bias is present in the prior, then the HBM method with clustering is able to improve the non-clustering inference in both the mean and the shape of redshift distributions. Additionally, if the prior has some efficiency drop at high redshift, like it may happen in spectroscopic surveys, then this improvement is even more important. Finally, we also test the case where the redshift information in the prior is more uncertain, like in the case of photometric redshift being used, and show the clustering addition to be able to correct for the corresponding broadening in the prior redshift distribution.

The tests performed in this work provide demonstration that the method depicted in SB18 can be used in realistic conditions, and it can still be very powerful at resolving biases that are potentially present in prior samples, even after marginalizing over biasing functions in the addition of clustering information. The method does not guarantee an unbiased posterior, but it uses all

the information at hand to reduce prior biases, and even in all tests performed here, some of which are extreme cases of biased priors, the final biases in the posterior are of order 10^{-3} in the mean of redshift distributions, which is the most important quantity for weak lensing analyses. The success of this method in estimating redshift distributions to the accuracy needed for large cosmological surveys will still depend on the details of the survey, but we now have the full methodology to harness all of the available information and the results in this work show the potential to be a solution to the redshift estimation problem in future photometric galaxy surveys.

Summary and Conclusions

In this thesis we have studied the measurement and applications to cosmology analysis of higher resolution photometric redshifts. In chapter 2 we have presented a forecast of cosmological parameters for a redshift survey with very precise redshift estimates, either from spectroscopy or many narrow band images. We have modeled galaxy clustering in thin redshift slices where the radial information is contained in the cross correlations. This study focused in the use of multiple tracers to cancel sample variance in this 2D projected framework, which improves the survey's constraining power and makes it suitable to combine with other projected cosmological analysis, as weak lensing. We modeled realistic sample density and bias after splitting the survey into two using conditional luminosity functions from SDSS observations. We found a split using a halo mass proxy is more optimal than a luminosity split. This technique can be applied to PAUS data, which can produce overlapping samples with high precision redshift measurements while still delivering a large density. Another application would be the upcoming DESI spectroscopic survey, which will measure luminous red galaxies and emission line galaxies over 14,000 $\deg^2$, which are population with very different galaxy bias.

In chapter 3 we have presented a study of photometric redshifts in PAUS narrow band survey. Using data from a calibration field where multiband photometry and spectroscopic data exists we have measured the photometric redshift precision this unique survey can achieve. We have developed two new photo-z algorithms which linearly combine templates of galaxies to properly model the continuum and emission line flux observable with the narrow band filter set. We developed simulated fluxes from an N-body simulation to validate our methods, which include calibration of several systematic effects of the data prior to the photo-z estimation. The photometric redshift precision found is quasi spectroscopic for the brighter galaxies, and grows with fainter galaxies up to 1.5% dispersion at $i_{\rm AB} = 22 - 22.5$. This quality photo-z estimations enable the science cases of PAUS, which include galaxy evolution studies, multiple tracers analysis, intrinsic alignments and galaxy clustering measurements in

intermediate scales or calibration of redshift distributions of lensing surveys. The photo-z framework developed here is also suitable to properly include more complex modeling of galaxies, like combinations with quasar templates or more sophisticated emission line modeling, which can improve the photo-z estimation in future studies.

Finally, in chapter 4 we have presented a novel method which combines information from colors and clustering in a hierachicale bayesian model to estimate the redshift distribution of a lensing survey. This two main sources of information have historically been used separately and combined later ad hoc, whereas they get naturally combined in this framework. This work shows in complex N-body simulations how clustering information can correct large redshift biases from systematically biased priors or priors affected by sample variance. This method can be applied to current lensing surveys like DES or KiDS and future ones like LSST or Euclid, improving the robustness of the estimated redshift distribution of galaxies. The prior for such surveys can be estimated from photometric redshifts from PAUS data, which provides great photo-z precision for a unique combination of area and depth.

Bibliography

Abbott T. M. C. et al., 2018, Phys. Rev. D, 98, 043526

Abramo L. R., 2012, MNRAS, 420, 2042

Abramo L. R., Secco L. F., Loureiro A., 2016, MNRAS, 455, 3871

Albrecht A. et al., 2006, preprint (arXiv:astro-ph/0609591)

Allen C. W., 1976, Astrophysical Quantities

Arnouts S., Ilbert O., 2011, LePHARE: Photometric Analysis for Redshift Estimate

Asorey J., Crocce M., Gaztañaga E., 2014, MNRAS, 445, 2825

Asorey J., Crocce M., Gaztañaga E., Lewis A., 2012, MNRAS, 427, 1891

Benítez N., 2000, ApJ, 536, 571

Berger J. O., 1985, Statistical decision theory and Bayesian analysis

Bernstein G. M., Cai Y.-C., 2011, MNRAS, 416, 3009

Bertin E., 2011, in Astronomical Society of the Pacific Conference Series, Vol. 442, Astronomical Data Analysis Software and Systems XX, Evans I. N., Accomazzi A., Mink D. J., Rots A. H., eds., p. 435

Bianchi D., Guzzo L., Branchini E., Majerotto E., de la Torre S., Marulli F., Moscardini L., Angulo R. E., 2012, MNRAS, 427, 2420

Bonnett C., 2015, MNRAS, 449, 1043

Bonnett C. et al., 2016, Phys. Rev. D, 94, 042005

Brammer G. B., van Dokkum P. G., Coppi P., 2008, ApJ, 686, 1503

Brown M. J. I. et al., 2014, ApJS, 212, 18

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000

Buchs R. et al., 2019, MNRAS, 2106

Cabré A., Fosalba P., Gaztañaga E., Manera M., 2007, MNRAS, 381, 1347

Cacciato M., van den Bosch F. C., More S., Mo H., Yang X., 2013, MNRAS, 430, 767

Cai Y.-C., Bernstein G., 2012, MNRAS, 422, 1045

Cai Y.-C., Bernstein G., Sheth R. K., 2011, MNRAS, 412, 995

Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682

Carretero J., Castander F. J., Gaztañaga E., Crocce M., Fosalba P., 2015, MNRAS, 447, 646

Carretero J., et al., 2017, PoS, EPS-HEP2017, 488

Castander F., In prep.

Cautun M. C., van de Weygaert R., 2011, arXiv e-prints, arXiv:1105.0370

Chan K. C., Scoccimarro R., Sheth R. K., 2012, Phys. Rev. D, 85, 083509

Chevallier M., Polarski D., 2001, International Journal of Modern Physics D, 10, 213

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

Cox D. R., 1955, Journal of the Royal Statistical Society. Series B (Methodological), 17, 129

Crocce M., Cabré A., Gaztañaga E., 2011, MNRAS, 414, 329

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

Crocce M., Fosalba P., Castander F. J., Gaztañaga E., 2010, MNRAS, 403, 1353

Darvish B., Mobasher B., Sobral D., Scoville N., Aragon-Calvo M., 2015, ApJ, 805, 121

Davis C. et al., 2017, arXiv e-prints, arXiv:1710.02517

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

Dodelson S., 2003, Modern Cosmology

Efstathiou G., Bernstein G., Tyson J. A., Katz N., Guhathakurta P., 1991, ApJ, 380, L47

Eisenstein D. J., Hu W., 1998, ApJ, 496, 605

Elvin-Poole J. et al., 2018, Phys. Rev. D, 98, 042006

Eriksen M. et al., 2019, MNRAS, 484, 4200

Eriksen M., Gaztañaga E., 2015a, MNRAS, 451, 1553

Eriksen M., Gaztañaga E., 2015b, MNRAS, 452, 2149

Eriksen M., Gaztañaga E., 2015c, MNRAS, 452, 2168

Eriksen M., Gaztañaga E., 2015d, preprint (arXiv:1508.00035)

Fisher K. B., Huchra J. P., Strauss M. A., Davis M., Yahil A., Schlegel D., 1995, ApJ, 100, 69

Fisher K. B., Scharf C. A., Lahav O., 1994, MNRAS, 266, 219

Fisher R. A., 1935, Journal of the Royal Statistical Society, 98, 39

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

Fosalba P., Crocce M., Gaztañaga E., Castand er F. J., 2015a, MNRAS, 448, 2987

Fosalba P., Gaztañaga E., Castander F. J., Crocce M., 2015b, MNRAS, 447, 1319

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

Gaia Collaboration et al., 2016, A&A, 595, A2

Gatti M. et al., 2018, MNRAS, 477, 1664

Gaztañaga E., Eriksen M., Crocce M., Castander F. J., Fosalba P., Marti P., Miquel R., Cabré A., 2012, MNRAS, 422, 2904

Gaztañaga E., Lobo J. A., 2001, ApJ, 548, 47

Gaztanaga E., In prep.

Genz A., 1992, Journal of Computational and Graphical Statistics, 1, 141

Ghahramani Z., 2012, Philosophical Transactions of the Royal Society of London Series A, 371, 20110553

Gil-Marín H., Wagner C., Verde L., Jimenez R., Heavens A. F., 2010, MNRAS, 407, 772

Hamaus N., Seljak U., Desjacques V., Smith R. E., Baldauf T., 2010, Phys. Rev. D, 82, 043515

Hildebrandt H. et al., 2012a, MNRAS, 421, 2355

Hildebrandt H. et al., 2012b, MNRAS, 421, 2355

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

Hoffmann K., Bel J., Gaztañaga E., 2015, MNRAS, 450, 1674

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

Huterer D. et al., 2015, Astroparticle Physics, 63, 23

Ilbert O. et al., 2009, ApJ, 690, 1236

Ilbert O. et al., 2010, ApJ, 709, 644

Jaynes E. T., Bretthorst G. L., 2003, Probability Theory

Kaiser N., 1987, MNRAS, 227, 1

Kennicutt, Robert C. J., 1998, ARA&A, 36, 189

Knobel C. et al., 2012, ApJ, 753, 121

Knuth K. H., Habeck M., Malakar N. K., Mubeen A. M., Placek B., 2014, arXiv e-prints, arXiv:1411.3013

Koekemoer A. M. et al., 2007, ApJS, 172, 196

Laigle C. et al., 2016, ApJS, 224, 24

Le Fèvre O. et al., 2013, A&A, 559, A14

Leauthaud A. et al., 2007, ApJS, 172, 219

Leistedt B., Mortlock D. J., Peiris H. V., 2016, MNRAS, 460, 4258

Lilly S. J. et al., 2007, ApJS, 172, 70

Linder E. V., 2003, Physical Review Letters, 90, 091301

Lue A., Scoccimarro R., Starkman G., 2004, Phys. Rev. D, 69, 044005

Martí P., Miquel R., Castander F. J., Gaztañaga E., Eriksen M., Sánchez C., 2014, MNRAS, 442, 92

Martí P., Miquel R., Castander F. J., Gaztañaga E., Eriksen M., Sánchez C., 2014, MNRAS, 442, 92

Masters D. et al., 2015, The Astrophysical Journal, 813, 53

McDonald P., Seljak U., 2009, JCAP, 10, 007

Ménard B., Scranton R., Schmidt S., Morrison C., Jeong D., Budavari T., Rahman M., 2013, arXiv e-prints, arXiv:1303.4722

Newman J. A., 2008, ApJ, 684, 88

Oke J. B., Gunn J. E., 1983, ApJ, 266, 713

Okumura T., Hand N., Seljak U., Vlah Z., Desjacques V., 2015, Phys. Rev. D, 92, 103516

Okumura T., Jing Y. P., 2011, ApJ, 726, 5

Padilla C. et al., 2019, AJ, 157, 246

Pearson D. W., Samushia L., Gagrani P., 2016, MNRAS

Pen U.-L., 2004, MNRAS, 350, 1445

Penzias A. A., Wilson R. W., 1965, ApJ, 142, 419

Perlmutter S. et al., 1999, Apj, 517, 565

Pickles A. J., 1998, PASP, 110, 863

Planck Collaboration et al., 2018, arXiv e-prints, arXiv:1807.06205

Polletta M. et al., 2007, ApJ, 663, 81

Press W. H., Schechter P., 1974, ApJ, 187, 425

Prevot M. L., Lequeux J., Maurice E., Prevot L., Rocca-Volmerange B., 1984, A&A, 132, 389

Pujol A., Hoffmann K., Jiménez N., Gaztañaga E., 2015, preprint (arXiv:1510.01692)

Reid B. A., White M., 2011, MNRAS, 417, 1913

Riess A. G. et al., 1998, AJ, 116, 1009

Rozo E. et al., 2016, MNRAS, 461, 1431

Sadeh I., Abdalla F. B., Lahav O., 2016, PASP, 128, 104502

Sánchez C., Bernstein G. M., 2019, MNRAS, 483, 2801

Sargent M. T. et al., 2007, ApJS, 172, 434

Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525

Schmidt S. J., Ménard B., Scranton R., Morrison C., McBride C. K., 2013, MNRAS, 431, 3307

Scoccimarro R., Sheth R. K., Hui L., Jain B., 2001, ApJ, 546, 20

Seljak U., 2009, Physical Review Letters, 102, 021302

Seljak U., Hamaus N., Desjacques V., 2009, Physical Review Letters, 103, 091303

Serrano S., In prep.

Sha F., Lin Y., Saul L. K., Lee D. D., 2007, Neural Comput., 19, 2004

Sheth R. K., Tormen G., 1999, MNRAS, 308, 119

Smith J. A. et al., 2002, AJ, 123, 2121

Suchyta E. et al., 2016, MNRAS, 457, 786

Tanabashi M. et al., 2018, Phys. Rev. D, 98, 030001

Taniguchi Y. et al., 2015, PASJ, 67, 104

Taylor A. N., Heavens A. F., 1995, in American Institute of Physics Conference Series, Vol. 336, Dark Matter, Holt S. S., Bennett C. L., eds., pp. 381–385

Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G., Gottlöber S., 2010, ApJ, 724, 878

Tonello N. et al., 2019, Astronomy and Computing, 27, 171

Vakili M. et al., 2019, MNRAS, 487, 3715

van Dokkum P. G., 2001, PASP, 113, 1420

Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Rozo E., 2013, Physics Reports, 530, 87

Weinberg S., 1972, Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity

White M., Song Y.-S., Percival W. J., 2009, MNRAS, 397, 1348